



COMPARAÇÃO ENTRE CLASSIFICADORES PARA IDENTIFICAÇÃO DE PLANTAS DANINHAS UTILIZANDO TÉCNICAS DE PROCESSAMENTO DE IMAGENS NA CULTURA DO FEIJÃO

João G. M. SILVA¹; João M. RIBEIRO²; Allan A. PEREIRA³; Guilherme B. SILVA⁴; Eduarda da S. TRINDADE⁵; Maria V. A. B. F. da CRUZ⁶

RESUMO

Este artigo explora a aplicação de algoritmos de Aprendizado de Máquina na classificação de plantas daninhas em imagens de um talhão de cultivo de feijão. Os resultados indicam que o algoritmo Random Forest supera os demais na segmentação das ervas daninhas, apresentando maior acurácia. A utilização de técnicas como validação cruzada e otimização de hiperparâmetros resultou em melhorias na acurácia dos modelos. A complexidade da classificação de plantas daninhas por meio do processamento de imagens é ressaltada, apesar dos avanços obtidos. Este estudo contribui para alcançar maior precisão na segmentação das plantas daninhas na agricultura do feijão.

Palavras-chave:

RGB; Tecnologias agrícolas; Inteligência Artificial; Aprendizado de Máquina; Agricultura de Precisão;

1. INTRODUÇÃO

As plantas daninhas invasoras são consideradas um dos problemas mais importantes que afetam a produção agrícola em escala global, resultando na redução da capacidade de produção de alimentos, fibras e combustíveis. Elas causam perda de rendimento das safras, dificultam a colheita, infestam os campos durante as estações de cultivo e aumentam o surgimento de doenças e pragas.

Para lidar com o problema, surgem novas tecnologias desenvolvidas baseadas na agricultura de precisão, que, em conjunto com as técnicas de controle tradicionais de ervas daninhas, têm mostrado resultados promissores na produtividade e na redução dos impactos ambientais.

O mapeamento da infestação de plantas daninhas torna-se uma ferramenta útil para a aplicação localizada de herbicidas, fornecendo informações sobre a localização e a distribuição das ervas daninhas em toda a área. Isso permite a prescrição precisa da quantidade e do volume de defensivos necessários, resultando na redução do risco de poluição ambiental e dos custos dos herbicidas. Diante dessa demanda, pesquisas que aplicam técnicas de processamento de imagens e aprendizado de máquina na agricultura têm recebido destaque, pois representam ferramentas para o desenvolvimento de estratégias ou sistemas inteligentes de aplicação de defensivos.

¹ Discente do Bacharelado em Ciência da Computação, IFSULDEMINAS – Campus Muzambinho. E-mail: jgmsgabriel@gmail.com

² Orientador, IFSULDEMINAS – Campus Muzambinho. E-mail: joao.ribeiro@muz.ifsuldeminas.edu.br

³ Coorientador, IFSULDEMINAS – Campus Muzambinho. E-mail: allan.pereira@ifsuldeminas.edu.br

⁴ Coautor, IFSULDEMINAS – Campus Muzambinho. E-mail: bueno1.guilhermesilva@gmail.com

⁵ Coautor, IFSULDEMINAS – Campus Muzambinho. E-mail: eduardasilvatrindade@gmail.com

⁶ Coautor, IFSULDEMINAS – Campus Muzambinho. E-mail: mariavictoriaalvesbueno@gmail.com

Diante do exposto acima, o objetivo deste trabalho é comparar algoritmos de aprendizado de máquina capazes de identificar plantas daninhas. Visto que o algoritmo de maior capacidades resultaria numa importante ferramenta tecnológica de auxílio à tomada de decisão para o manejo de plantas daninhas, colaborando com redução de custos e de impactos ao meio ambiente.

2. FUNDAMENTAÇÃO TEÓRICA

As plantas invasoras se desenvolvem em áreas de cultivo e competem pelos recursos essenciais com as culturas desejadas pelos agricultores. Elas podem ser encontradas em diversos tipos de plantações. As daninhas representam um dos principais desafios enfrentados pelos agricultores, pois resultam em reduções significativas na produtividade e qualidade das safras.

Em um cenário em que a agricultura enfrenta diversos desafios devido aos altos custos de produção, observa-se uma relação direta entre a rentabilidade e a produtividade. Para superar esses desafios, surgiram tecnologias que auxiliam no aumento do potencial produtivo, especialmente com o uso de sensoriamento remoto. A agricultura de precisão é uma abordagem que utiliza tecnologias avançadas com o objetivo de maximizar a eficiência no uso de recursos, como fertilizantes, defensivos agrícolas e água, ao mesmo tempo em que minimiza os impactos ambientais. A agricultura de precisão se baseia na coleta e análise de dados detalhados.

Uma técnica importante para análise de dados, é o uso do Aprendizado de Máquina (ML), também conhecido como *Machine Learning*, é um ramo da Inteligência Artificial (IA) que tem a capacidade de automatizar processos e aprender a identificar padrões.

Dentre os algoritmos, alguns dos utilizados na classificação de plantas daninhas, são: SVM⁷ (*Support Vector Machines*). O objetivo dele é encontrar o hiperplano de separação que maximize a margem entre as classes. Outro algoritmo usado é o *Random Forest*⁸, essa abordagem é baseada no uso de múltiplas árvores de decisão em um processo chamado *bootstrap*. *Naive Bayes*⁹ também é amplamente usado, ele se baseia no teorema de Bayes. Durante o treinamento, o algoritmo calcula probabilidades de classes e de ocorrência de atributos para cada uma delas.

3. MATERIAL E MÉTODOS

Para realizar as análises neste estudo, foi utilizado o ambiente de programação Python juntamente com o serviço do Google Colab¹⁰, uma plataforma baseada em nuvem que permite a

⁷ Disponível em: <<https://shre.ink/al2m>>. Último acesso em 12/08/2023

⁸ Disponível em: <<https://shre.ink/al2X>>. Último acesso em 12/08/2023

⁹ Disponível em: <<https://shre.ink/al2v>>. Último acesso em 12/08/2023

¹⁰ Disponível em: <<https://shre.ink/al23>>. Último acesso em 12/08/2023

execução de código Python, com o uso das bibliotecas, numpy¹¹, OpenCV¹² e pandas¹³. As imagens foram adquiridas utilizando o drone DJI Mavic, no contexto de um talhão situado na cidade de Muzambinho. Estas imagens estão disponíveis nos canais de cores R (vermelho), G (verde) e B (azul). Na área em estudo, verifica-se a presença de cultivo de feijão, com as imagens tendo sido obtidas dois meses subsequentes ao período de plantio. Para a construção da base de dados, empregada no treinamento dos algoritmos, tornou-se necessário redimensionar as imagens originalmente com dimensões de 5472 x 3648 pixels para um formato de 800 x 800 pixels. Este redimensionamento objetivou otimizar os recursos de processamento. Em sequência, procedeu-se ao cálculo das médias dos canais RGB, da escala de cinza, bem como dos índices vegetativos, tais como o VARI (*Índice Visível Atmosfericamente Resistente*), ExG (*Índice de Excesso de Verde*) e GLI (*Índice de Folhas Verdes*). A inclusão de parâmetros adicionais, para além dos canais RGB, foi importante na etapa de segmentação e classificação. A demarcação dos pixels correspondentes a plantas daninhas foi estabelecida por meio de uma classificação manual.

Após a preparação da base de dados, procedeu-se com a implementação dos algoritmos de aprendizado de máquina. Foram empregados três modelos de classificação distintos, o *Random Forest*, *Support Vector Machines* (SVM) e *Naive Bayes* (NB). Com o propósito de mitigar a ocorrência de *overfitting*, a técnica de validação cruzada¹⁴ foi empregada para uma avaliação mais sólida e robusta do desempenho dos modelos. Concomitantemente, visando aprimorar os algoritmos e identificar os melhores hiperparâmetros, utilizando a acurácia como critério de referência, foi adotada a abordagem denominada *Randomized Search CV*¹⁵. Essa abordagem realiza uma exploração das combinações viáveis dentro de um espaço de hiperparâmetros predefinido.

4. RESULTADOS E DISCUSSÃO

Com base nos conjuntos de dados, os algoritmos foram treinados. O algoritmo *Random Forest* demonstrou uma média de acurácia de 84.71%, enquanto o SVM alcançou uma acurácia de 84.11% e o *Naive Bayes* de 80.42%. Mediante a aplicação da técnica de validação cruzada e a otimização dos hiperparâmetros, tornou-se evidente uma elevação na acurácia dos modelos, alcançando 86.93% para o *Random Forest*, 84.65% para o SVM e 81.14% para o *Naive Bayes*.

Para avaliar a eficácia dos algoritmos, empregou-se uma nova imagem de dimensões 250 x 250 pixels, na qual os pixels classificados como plantas daninhas pelos algoritmos foram destacados em branco. Tal procedimento gerou três imagens, ilustradas na Figura 1. Por meio dos

¹¹ Disponível em: <<https://numpy.org/pt/>>. Último acesso em 12/08/2023

¹² Disponível em: <<https://opencv.org/>>. Último acesso em 12/08/2023

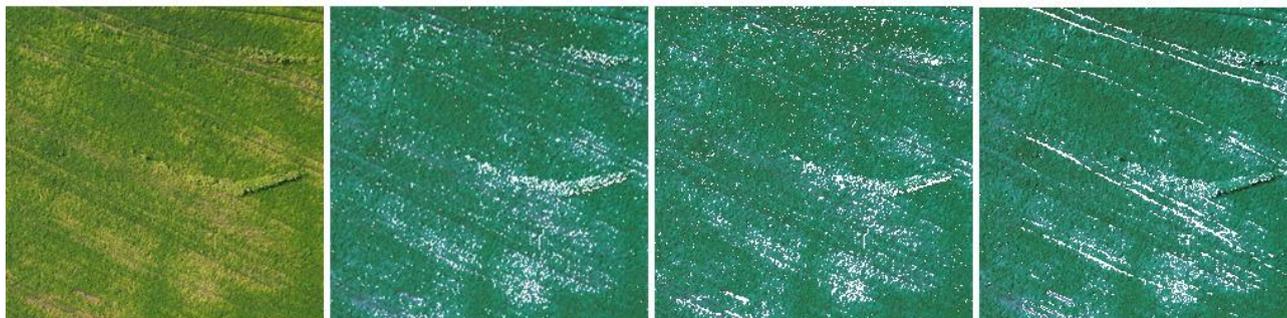
¹³ Disponível em: <<https://pandas.pydata.org/>>. Último acesso em 12/08/2023

¹⁴ Disponível em: <https://scikit-learn.org/stable/modules/cross_validation.html>. Último acesso em 12/08/2023

¹⁵ Disponível em: <<https://shre.ink/al2B>>. Último acesso em 12/08/2023

resultados obtidos, foi possível notar uma concentração de pontos brancos nas regiões correspondentes às áreas com efetiva presença de plantas daninhas. No entanto, também foi observada a presença de diversos pontos brancos em áreas onde não existem plantas daninhas.

Figura 1 - Imagem Original, Random Forest, Naive Bayes e SVM



5. CONCLUSÃO

Neste artigo, foi realizada a aplicação de algoritmos de Aprendizado de Máquina na tarefa de classificação de plantas daninhas presentes em imagens capturadas de um talhão de cultivo de feijão. Os resultados obtidos revelaram que o algoritmo *Random Forest* demonstrou um desempenho superior na segmentação das ervas daninhas nas imagens. Tal superioridade é constatada tanto pela avaliação da acurácia alcançada quanto pela distribuição dos pixels brancos, conforme ilustrado na Figura 1. Este achado, sinaliza a viabilidade da aplicação desta técnica como uma abordagem promissora para o controle e monitoramento eficaz dessas questões agrícolas.

Essa pesquisa evidencia a complexidade associada à classificação de plantas daninhas com base em técnicas de processamento de imagens. Os avanços na acurácia após a aplicação de validação cruzada e otimização de hiperparâmetros destacam a importância dessas estratégias para o aprimoramento dos modelos de aprendizado de máquina. No entanto, os desafios relacionados à segmentação precisa das plantas daninhas sugerem a necessidade contínua de refinamento dessas abordagens, buscando maior precisão e robustez em futuras aplicações na agricultura do feijão.

REFERÊNCIAS

OWEN, M. D. K. Diversas abordagens para o manejo de plantas daninhas resistentes a herbicidas. *Weed Science*, v. 64, n. sp1, p. 570-584, 2016.

ROBERT, P.C. Precision agriculture: a challenge for crop nutrition management. *Plant and Soil*, v. 247, p. 143-149, 2002.

SCHUSTER, I.; NORDMEYER, H.; RATH, T. Comparison of vision-based and manual weed mapping in sugar beet. *Biosystems Engineering*, v. 98, n. 1, p. 17-25, 2007.