

Desenvolvimento de um Modelo Inteligente para Identificar e Contar Insetos em Armadilhas Adesivas

Artur Mariano da Silva

Instituto Federal de Educação, Ciência e
Tecnologia Sul-rio-grandense (IFSul)
Passo Fundo, RS, Brasil
artursilva.pf041@academico.ifsul.edu.br

Álisson Mello da Rosa

Instituto Federal de Educação, Ciência e
Tecnologia Sul-rio-grandense (IFSul)
Passo Fundo, RS, Brasil
alissonrosa.pf020@academico.ifsul.edu.br

Nicolas Welfer Kirinus

Universidade de Passo Fundo (UPF)
Passo Fundo, RS, Brasil
183948@upf.br

Douglas Lau

Empresa Brasileira de Pesquisa
Agropecuária (Embrapa Trigo)
Passo Fundo, RS, Brasil
douglas.lau@embrapa.br

Alexsander Furtado Carneiro

Instituto Federal de Educação, Ciência e
Tecnologia Sul-rio-grandense (IFSul)
Passo Fundo, RS, Brasil
alexsandercarneiro@ifsul.edu.br

Telmo De Cesaro Júnior

Instituto Federal de Educação, Ciência e
Tecnologia Sul-rio-grandense (IFSul)
Passo Fundo, RS, Brasil
telmojunior@ifsul.edu.br

Resumo—Este estudo propõe um modelo baseado na rede neural YOLOv7 para a identificação e contagem de afídeos alados em imagens de armadilhas adesivas. Comparou-se dois modelos: o primeiro utilizando transferência de aprendizado (*transfer learning*) e o segundo treinado a partir de pesos aleatórios. Embora o modelo treinado com pesos aleatórios tenha demonstrado melhor desempenho geral, as métricas de *recall*, *F1 score* e *mAP@.5:95* revelaram dificuldades na localização precisa dos afídeos, indicando a necessidade de aprimorar o conjunto de imagens e o mecanismo de aquisição. Este estudo ilustra como a inteligência artificial pode ser aplicada ao monitoramento de pragas, contribuindo para práticas agrícolas mais sustentáveis.

Abstract—This study proposes a model based on the YOLOv7 neural network for identifying and counting winged aphids in images from adhesive traps. Two models were compared: the first using transfer learning and the second trained from random weights. Although the model trained with random weights demonstrated better overall performance, the recall, F1 score, and mAP@.5:95 metrics revealed challenges in accurately locating the aphids, indicating the need to improve the image dataset and acquisition mechanism. This study illustrates how artificial intelligence can be applied to pest monitoring, contributing to more sustainable agricultural practices.

I. INTRODUÇÃO

Os afídeos (*Hemiptera: Aphididae*) causam grandes perdas na produção de cereais de inverno, como trigo e aveia, e são vetores de vírus como o do nanismo amarelo, que pode reduzir a produção em até 20% [1], [2]. Para mitigar esses impactos, é crucial monitorar não apenas os afídeos, mas também seus inimigos naturais, os parasitoides (*Hymenoptera: Aphelinidae* e *Braconidae, Aphidiinae*), para orientar o manejo integrado de pragas.

A Embrapa Trigo coordena uma Rede de Monitoramento de Pragas e Cereais de Inverno, em colaboração com instituições de pesquisa, cooperativas e empresas privadas. O objetivo é mensurar os níveis de afídeos e parasitoides em culturas de inverno para apoiar estudos e modelos de simulação. O programa utiliza armadilhas Moericke e adesivas para a captura de insetos [3].

Conforme detalhado em De Cesaro Júnior [4] e Di Domênico [5], periodicamente é realizado um processo manual de coleta, triagem, digitalização, identificação e contagem automática dos insetos pelo software InsectCV [6] [4]. Os dados gerados são armazenados no banco de dados AgroDB [7]. Esses dados são posteriormente usados por modelos de simulação como o ABISM, que simula o ciclo de vida de várias espécies de insetos pragas em diferentes culturas agrícolas [8].

De Cesaro Júnior [4] aponta desafios na automatização do processo manual com armadilhas Moericke, devido ao volume e ao peso do líquido. A literatura sugere armadilhas adesivas como uma solução viável, com trabalhos de Huang et al. [9], Li et al. [10], e Bockmann et al. [11] propondo armadilhas eletrônicas com superfícies adesivas e a utilização de redes convolucionais para a identificação e contagem automática dos insetos.

Portanto, o objetivo deste estudo é comparar duas abordagens de treinamento para modelos baseados no YOLOv7: o *transfer learning* e o treinamento a partir de pesos aleatórios, aplicados à tarefa de detecção de afídeos alados em imagens digitais de armadilhas adesivas. Para otimizar o processo de coleta de imagens destinadas ao treinamento, validação e teste dos modelos, propõe-se o uso de um mecanismo automatizado com base em um robô cartesiano. Espera-se que, com este novo modelo, o desempenho na detecção e contagem de afídeos supere os resultados obtidos no trabalho de Di Domênico [5].

II. REFERENCIAL TEÓRICO

Ferramentas de visão computacional, baseadas em aprendizado profundo, têm demonstrado avanços significativos na detecção de objetos em imagens digitais. À vista disso, a agricultura está adotando essas tecnologias para automatizar tarefas repetitivas, como a contagem de insetos retidos em

armadilhas, que geralmente são realizadas manualmente por especialistas com base nas características morfológicas [12].

Na revisão sistemática realizada por De Cesaro Júnior e Rieder [4] foram selecionados 33 trabalhos, entre 2015 e 2019, que indicaram a tendência pela utilização de DCNNs em detrimento as técnicas de extração manual de características. Essa pesquisa destacou a necessidade de grande volume de dados para treinamento de redes profundas e a aprimoramento de técnicas para a detecção de objetos conectados ou parcialmente sobrepostos e com variação de pose e tamanho.

A. Redes Neurais Convolucionais

Uma Rede Neural Convolutiva, também conhecida como CNN (*Convolutional Neural Network*), é uma rede de aprendizado profundo de visão computacional capaz de reconhecer e classificar características de imagens. A arquitetura da CNN foi influenciada pela organização e funções do córtex visual, sendo projetada para se assemelhar às conexões entre os neurônios no cérebro humano. Nesse sentido, a arquitetura profunda das CNNs é que, com mapeamentos adicionais (não lineares) e hierarquias de características mais avançadas, a rede possa aproximar a função objetivo de maneira eficaz. A profundidade da rede é um parâmetro crucial para o treinamento supervisionado, pois redes profundas representam classes de funções específicas de maneira mais eficaz do que sistemas rasos [13].

B. You Only Look Once (YOLO)

O algoritmo *You Only Look Once* (YOLO) é amplamente reconhecido por sua eficiência em detecção de objetos, demandando menos recursos computacionais em comparação com outros métodos. Introduzido por Redmon et al. em 2015 [14], o YOLO se destaca pela capacidade de realizar detecção de objetos em tempo real com alta precisão. Baseado na rede neural Darknet [15], o YOLO aborda a detecção como um problema de regressão único, permitindo que a rede neural processe a imagem inteira em uma única passagem, melhorando tanto a velocidade quanto a precisão. Pelo fato de realizar a detecção em apenas uma etapa, sem *Region Proposal Network (RPN)*, alcançam precisão menor em comparação com os classificadores de dois estágios (*Faster R-CNN*, o *R-FCN* e o *Mask R-CNN*), porém necessitam de menor consumo computacional, possibilitando ser empregados em computadores de placa única e sistema embarcados [16].

Atualmente, o YOLO está na versão 10, refletindo a rápida evolução e os constantes aprimoramentos desse modelo. Nesse sentido, a tecnologia apresenta constantes melhorias na arquitetura e no desempenho em comparação com versões anteriores. Vale ressaltar que o YOLOv4 [17], usado por Di Domênico em comparação com o *Mask R-CNN*, apresentou similar desempenho às Redes Neurais Convolucionais Profundas (DCNNs), porém com menor demanda computacional.

III. MATERIAIS E MÉTODOS

O conjunto de dados foi criado a partir de um mecanismo de aquisição de imagens digitais (figura 1), utilizando um

Raspberry Pi 4 para controlar um robô cartesiano com motores de passo, baseado no projeto de Imai [18] e fabricado em 3D. O sistema movimenta um sensor de imagem (Raspberry Cam Model 2) com precisão para capturar imagens de armadilhas adesivas.

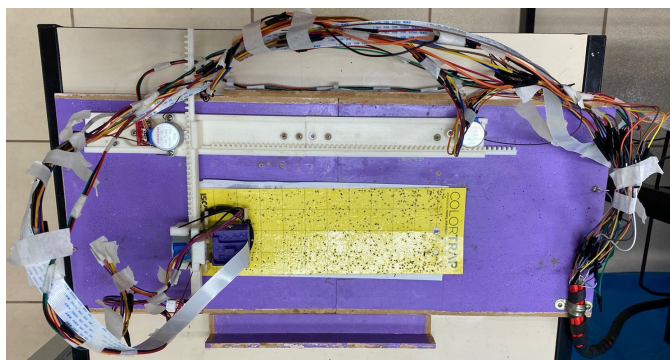


Figura 1. Mecanismo de aquisição de imagens

Uma rotina em Python foi desenvolvida para controlar o robô, capturar imagens e enviá-las para o serviço de inferência do InsectCV. Utilizou-se a biblioteca OpenCV [19] para tratamento e recorte das imagens. O envio automatizado de imagens para o Google Drive gerou 3141 capturas para o dataset.

O YOLOv7 foi o modelo selecionado para o treinamento, fundamentado em estudos anteriores, como o de Di Domênico [5], que mostraram que versões anteriores, como o YOLOv4, apresentaram excelente desempenho com menor demanda computacional em comparação com outras abordagens, incluindo Redes Neurais Convolucionais Profundas (DCNNs). Além disso, a versão 7 possuía a documentação mais abrangente disponível no momento da execução deste trabalho, tornando-a uma escolha adequada. Por esses motivos, optou-se por utilizar o YOLOv7.

A rotulação do dataset, que incluiu um total de 5.231 insetos, foi realizada utilizando as ferramentas Roboflow e LabelImg¹ [20]. Dessa forma, o conjunto de dados consistiu em 1841 imagens destinadas ao treinamento, 650 para validação e 650 para testes. Para aumentar a diversidade do conjunto de treinamento e melhorar a generalização do modelo, foram aplicadas técnicas de *data augmentation* no subconjunto de treinamento, resultando em 5523 imagens, além das 1300 destinadas para validação e testes (6823 imagens no total).

Nesse sentido, dois modelos foram treinados: um utilizando *Transfer Learning* (modelo COCO) por 800 épocas e outro com pesos aleatórios por 1000 épocas, ambos no YOLOv7 p5 com um limiar de IoU de 0,65. O treinamento foi realizado em um servidor com GPU RTX 4070 Ti. O modelo com *Transfer Learning* teve uma convergência mais rápida devido ao pré-conhecimento, enquanto o modelo do zero demandou mais tempo para aprender características a partir de pesos aleatórios.

Por conseguinte, os modelos foram testados na plataforma Trap System para validação em condições reais, utilizando 14

¹Ferramenta recomendada na documentação do YOLO

imagens capturadas pelo robô, com resultados discutidos nas seções seguintes.

IV. RESULTADOS

A. Treinamento dos Modelos

O primeiro experimento utilizou o modelo COCO para detecção de objetos (tabela I). Pode-se perceber que as métricas precisão, *recall* e $mAP@.5$ foram significativamente superiores à $mAP@.5:.95$, indicando dificuldade do modelo em localizar e marcar precisamente o inseto.

A métrica $mAP@.5$ (*Mean Average Precision at IoU = 0.5*) representa a média da precisão para todas as classes, considerando uma detecção correta quando a Interseção sobre a União (IoU) entre a predição e o rótulo verdadeiro é de pelo menos 50%. Já a $mAP@.5:.95$ é uma métrica mais rigorosa, calculando a média das precisões em diferentes limiares de IoU, de 0.5 a 0.95 com um passo de 0.05. Isso avalia a capacidade do modelo em realizar detecções precisas e localizações refinadas. A $mAP@.5:.95$ tende a ser mais baixa, especialmente em situações desafiadoras como a identificação de afídeos, devido à sua forma irregular e características variadas.

Tabela I
RESULTADOS DE TREINAMENTO POR EXPERIMENTO

Métricas/Experimento	Experimento 1	Experimento 2
Precision	0.886	0.881
Recall	0.769	0.807
mAP@.5	0.87	0.914
mAP@.5:.95	0.538	0.603

No segundo experimento, o modelo treinado do zero com pesos aleatórios mostrou melhora em relação ao anterior, como indicado na Tabela I. Dessa forma, para validar completamente os modelos em uma situação que simula a realidade, tornou-se necessária a captura de novas imagens de armadilhas para realizar nova inferência.

B. Validação dos modelos

Para validar os modelos, foram processadas 14 imagens. As inferências foram revisadas manualmente e classificadas como verdadeiros positivos, falsos positivos e falsos negativos. As métricas de desempenho, apresentadas na Tabela II, foram calculadas em comparação com as contagens feitas por especialistas em laboratório.

Tabela II
RESULTADOS DE VALIDAÇÃO DOS EXPERIMENTOS

Métricas/Experimento	Experimento 1	Experimento 2
Precision	0.759	0.954
Recall	0.389	0.402
F1	0.514	0.566

Comparando os dois experimentos, o segundo, iniciado com pesos aleatórios, obteve métricas superiores ao primeiro. Esse resultado desafia a expectativa de que modelos pré-treinados

teriam melhor desempenho, especialmente com dados limitados.

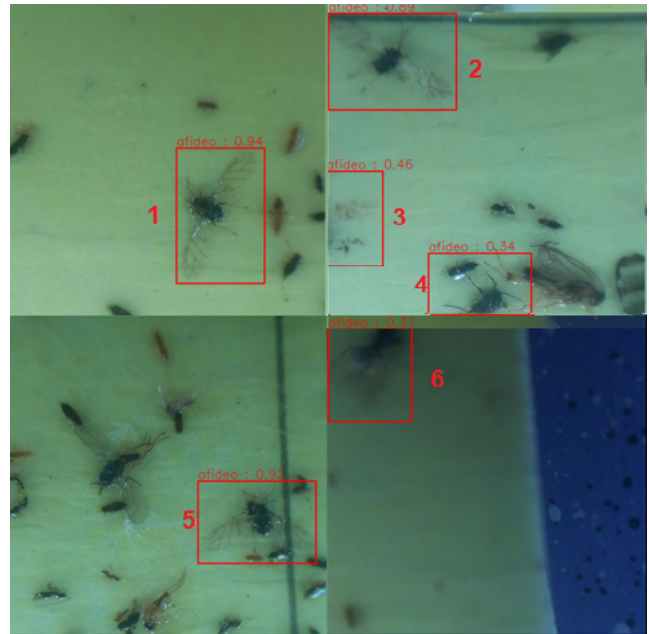


Figura 2. Resultado Inferência

A extração manual dos dados após a validação do modelo está ilustrada na Figura 2, que mostra amostras de imagens da plataforma TrapSystem com inferências do modelo (experimento 2).

A Figura 2 mostra que, apesar da maior assertividade do experimento 2, ainda houve inconsistências. As caixas de delimitação (*bounding boxes*) destacam os insetos identificados, com percentuais de confiança numerados de 1 a 6. As marcações 1, 2, 4 e 5 representam verdadeiros positivos com afídeos visíveis, enquanto as caixas 3 e 6 indicam problemas: a número 3 consiste em falso positivo contendo apenas rastros de insetos, e a caixa 6 não está nítida devido ao mau foco e à baixa iluminação (fato recorrente nas imagens adquiridas pelo mecanismo).

Além disso, a imprecisão na movimentação do sistema de captura é evidente; a imagem no canto inferior direito do mosaico mostra apenas metade da armadilha, o que gera dúvidas sobre a captura completa dos insetos.

Apesar dos desafios nas métricas e na captura de imagens, os resultados são promissores, com a alta precisão, especialmente no segundo experimento. O principal desafio é reduzir os falsos negativos. Melhorar o sistema de captura e ajustar o modelo para lidar com condições adversas são etapas essenciais para otimizar a detecção de afídeos.

C. Discussões

Embora o modelo tenha alcançado uma precisão relativamente alta, as métricas de detecção total de afídeos foram insatisfatórias. A análise das imagens revelou imprecisões na movimentação e na captura, que resultaram na ocultação de alguns insetos. Essa problemática evidenciou que o uso de

materiais plásticos comprometeu a qualidade das imagens, gerando atrito e imperfeições.

O desempenho notável na métrica de precisão do experimento 2 destaca a capacidade do modelo para identificar corretamente a maioria dos afídeos com poucos falsos positivos. No entanto, as métricas de *recall* e F1 foram insatisfatórias, em parte devido a problemas na movimentação do mecanismo, que ocultou ou desfocou alguns insetos (como o inseto 6 na Figura 2).

Ademais, a variação no posicionamento e a sobreposição dos insetos, como destacado por De Cesaro e Rieder [21], podem ter dificultado o reconhecimento pelo YOLO. A disposição não padronizada dos insetos, que às vezes têm asas e outras vezes não, compromete a precisão do modelo. A métrica mAP@.5:.95 reforça essa dificuldade, indicando a necessidade de melhorias na detecção.

O desempenho superior do experimento 2, apesar de utilizar pesos aleatórios, pode ser atribuído à especificidade dos afídeos. O modelo COCO, treinado para uma ampla gama de objetos, pode não ser otimizado para a detecção de insetos, resultando em desempenho inferior.

Os resultados indicam que o uso de um modelo pré-treinado não foi ideal para este estudo. Iniciar o treinamento a partir do zero, como no experimento 2, mostrou-se mais eficaz para a detecção de afídeos. Além disso, o mecanismo de captura de imagens deve ser aprimorado com motores de passo mais precisos, como os NEMA, sistemas de roldanas e materiais metálicos para garantir maior precisão.

V. CONSIDERAÇÕES FINAIS

O controle de afídeos é essencial para a produtividade das culturas tritícolas, e a integração de visão computacional e inteligência artificial oferece uma solução inovadora para automatizar a contagem e o monitoramento desses insetos, aumentando a precisão e reduzindo o trabalho manual. Este estudo visou automatizar a contagem de afídeos e comparar diferentes estratégias de detecção. Os resultados mostraram que o modelo treinado do zero, com pesos aleatórios, superou o modelo pré-treinado COCO em todas as métricas, sugerindo que o treinamento personalizado pode ser mais eficaz do que utilizar modelos pré-existentes para a detecção de afídeos.

Embora a precisão tenha sido satisfatória, as métricas de *recall*, F1 *score* e mAP@.5:.95 indicam desafios na identificação completa dos afídeos, com muitos falsos negativos e dificuldades na localização exata. Recomenda-se expandir o dataset, aplicar *data augmentation* e ajustar parâmetros como o tamanho das imagens e as condições de captura. Um trabalho futuro é desenvolver um dataset e um modelo específicos para parasitoides, pois esses insetos desempenham um papel crucial no controle dos afídeos.

O avanço das redes neurais e técnicas de aprendizado profundo oferece oportunidades para otimizar os resultados, mas é essencial continuar abordando limitações. Este estudo evidencia a importância de tecnologias avançadas na agricultura, destacando a necessidade de investimentos em pesquisa e desenvolvimento para aprimorar o controle de pragas agrícolas

e expandir as aplicações de inteligência artificial e visão computacional.

O presente trabalho foi financiado pelo CNPq (processo 403878/2021-7) e pela Pró-Reitoria de Pesquisa, Inovação e Pós-Graduação do Instituto Federal Sul-Rio-Grandense (IF-Sul), registro PE08230723/043.

REFERÊNCIAS

- [1] S. A. Hogenhout, E.-D. Ammar, A. E. Whitfield, and M. G. Redinbaugh, "Plant virus transmission by insects," *Journal of virology*, vol. 82, no. 9, pp. 3661–3675, 2008.
- [2] P. R. V. d. S. Pereira, M. D. Michelotto, and O. C. Bortolotto, "Dispersão de afídeos e dinâmica de transmissão do vírus do nanismo-amarelo em trigo e cevada," *Boletim de Pesquisa e Desenvolvimento*, vol. 16, no. 7, pp. 12–22, 2016.
- [3] E. Trigo, "Treinamento para aprimorar técnicas de identificação de insetos em cereais de inverno," 2019.
- [4] T. De Cesaro Júnior, R. Rieder, J. R. Di Domênico, and D. Lau, "Insectcv: A system for insect detection in the lab from trap images," *Ecological Informatics*, vol. 67, p. 101516, 2022.
- [5] G. A. Didomenico, "Evaluation of the yolo deep learning model for identification and classification of aphids (hemiptera: Aphididae) and parasitoids (hymenoptera: Braconidae) on winter cereals in southern brazil," *Revista Brasileira de Entomologia*, vol. 64, no. 3, pp. 1–9, 2020.
- [6] T. De Cesaro Jr., R. Rieder, D. Lau, and J. R. D. Domenico, "InsectCV," Programa de Computador. Número do registro: BR512021000542-2, 2021, INPI - Instituto Nacional da Propriedade Industrial.
- [7] A. T. Lazzaretti, J. M. C. Fernandes, W. Pavan, J. Toebe, and R. Wiest, "Agrodb—integration of database management systems with crop models," in *Proceedings of the 8th International Congress on Environmental Modelling and Software*. iEMSs Society, 2016, pp. 194–201.
- [8] R. Wiest, "Crescimento populacional de *Rhopalosiphum padi* (Linnaeus, 1758) (Hemiptera: Aphididae) em trigo: uma aproximação por simulação," Master's thesis, Universidade de Passo Fundo, Faculdade de Agronomia e Medicina Veterinária, Programa de Pós-Graduação em Agronomia, Passo Fundo, 2019.
- [9] R. Huang, T. Yao, C. Zhan, G. Zhang, and Y. Zheng, "A motor-driven and computer vision-based intelligent e-trap for monitoring citrus flies," *Agriculture*, vol. 11, no. 460, 2021.
- [10] W. Li, D. Wang, M. Li, Y. Gao, J. Wu, and X. Yang, "Field detection of tiny pests from sticky trap images using deep learning in agricultural greenhouse," *Elsevier Computers and Electronics in Agriculture*, vol. 183, no. 106048, 2021.
- [11] E. Bockmann, A. Pfaff, M. Schirrmann, and M. Pflanz, "Rapid and low-cost insect detection for analysing species trapped on yellow sticky traps," *Scientific Reports*, vol. 11, no. 10419, 2021.
- [12] P. R. V. da S. Pereira, J. R. Salvadori, and D. Lau, "Identificação de adultos ápteros e alados das principais espécies de afídeos (hemiptera: Aphididae) associadas a cereais de inverno no brasil," Embrapa Trigo, Passo Fundo, Tech. Rep., 2009.
- [13] D. Bhatt, C. Patel, H. Talsania, J. Patel, R. Vaghela, S. Pandya, K. Modi, and H. Ghayvat, "Cnn variants for computer vision: History, architecture, application, challenges and future scope," *Electronics*, vol. 10, no. 20, p. 2470, 2021.
- [14] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.
- [15] J. Redmon, "Darknet: Open source neural networks in c," <https://pjreddie.com/darknet/>, 2013.
- [16] P. Soviany and R. T. Ionescu, "Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction," in *20th SYNASC*, 09 2018, pp. 209–214.
- [17] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.
- [18] Makoto Imai, "Arduino mini cnc plotter," <https://www.thingiverse.com/thing:4608599>, 2020, acesso em: September 27, 2020.
- [19] O. team, "Opencv," <https://opencv.org/>, 2023.
- [20] Tzutalin, "Labelimg," <https://github.com/tzutalin/labelImg>, 2015.
- [21] T. De Cesaro Jr. and R. Rieder, "Automatic identification of insects from digital images: A survey," *Computers and Electronics in Agriculture*, vol. 178, p. 105784, 2020.