Intelligent Drones with Vision Technology for Sustainable Weed Mapping in Precision Agriculture

Pasquale De Marinis*, Gennaro Vessio and Giovanna Castellano

Department of Computer Science, University of Bari Aldo Moro, Bari, Italy

Abstract

This contribution outlines ongoing research in drone vision conducted at the Computational Intelligence Laboratory of the Department of Computer Science, University of Bari Aldo Moro, for drone-based weed mapping in precision agriculture. The research integrates multispectral imaging, deep learning, and model compression to develop efficient and deployable solutions. A key focus is on transfer learning strategies that enable pre-trained deep learning models on RGB datasets to operate effectively on multispectral imagery, addressing the scarcity of annotated agricultural data. Further work explores knowledge distillation techniques to compress complex segmentation networks into lightweight models suitable for real-time inference on drones and low-power edge devices, preserving high accuracy with minimal computational cost. A fully unsupervised framework—RoWeeder—is also introduced, which exploits crop-row detection to generate pseudo-ground truth for training segmentation models without manual annotations. Across these lines of research, experiments on the WeedMap dataset demonstrate robust weed detection under varying field conditions, advancing the applicability of drone-based systems in sustainable agriculture.

Keywords

precision agriculture, multispectral imaging, transfer learning, knowledge distillation, semantic segmentation, unsupervised learning, UAVs, lightweight models, weed detection, crop-row detection

1. Introduction

In intelligent robotics, drones can be seen as advanced autonomous agents equipped with sensors and intelligent algorithms that enable them to perceive and interact with their environment. Within this framework, computer vision algorithms can enhance drones' perceptual capabilities to support sustainable interventions, such as targeted tasks in precision agriculture (e.g., herbicide application, reduction of chemical usage, optimized crop management, and improved environmental monitoring), advancing both agricultural efficiency and ecological sustainability. In particular, integrating drone vision solutions into precision agriculture is reshaping how we detect, map, and manage weed infestations in modern crop fields.

At the Computational Intelligence Laboratory (CILab) of the Department of Computer Science, University of Bari Aldo Moro, our research is dedicated to designing AI-driven drone systems that enable high-resolution, automated weed mapping. By leveraging multispectral imaging, real-time aerial data acquisition, and advanced deep learning algorithms, we aim to deliver scalable, practical tools for site-specific weed control—ultimately reducing herbicide dependency and enhancing crop productivity.

This paper comprehensively summarizes our ongoing research efforts and experimental findings in this domain. By consolidating our methodologies and results, we seek to demonstrate the real-world applicability of AI-based weed mapping solutions, foster critical engagement, and support further innovations in sustainable agriculture.

The global need for more efficient and ecologically responsible food production is urgent. According to the Food and Agriculture Organization (FAO), the world population is expected to reach nine billion by 2050, doubling current food demand [1]. Simultaneously, agricultural resources are increasingly threatened by climate change, degradation, and overuse. In response, the sector embraces technologies

Ital-IA 2025: 5th National Conference on Artificial Intelligence, organized by CINI, June 23-24, 2025, Trieste, Italy

pasquale.demarinis@uniba.it (P. D. Marinis); gennaro.vessio@uniba.it (G. Vessio); giovanna.castellano@uniba.it (G. Castellano)

0000-0001-8935-9156 (P. D. Marinis); 0000-0002-0883-2691 (G. Vessio); 0000-0002-6489-8628 (G. Castellano)

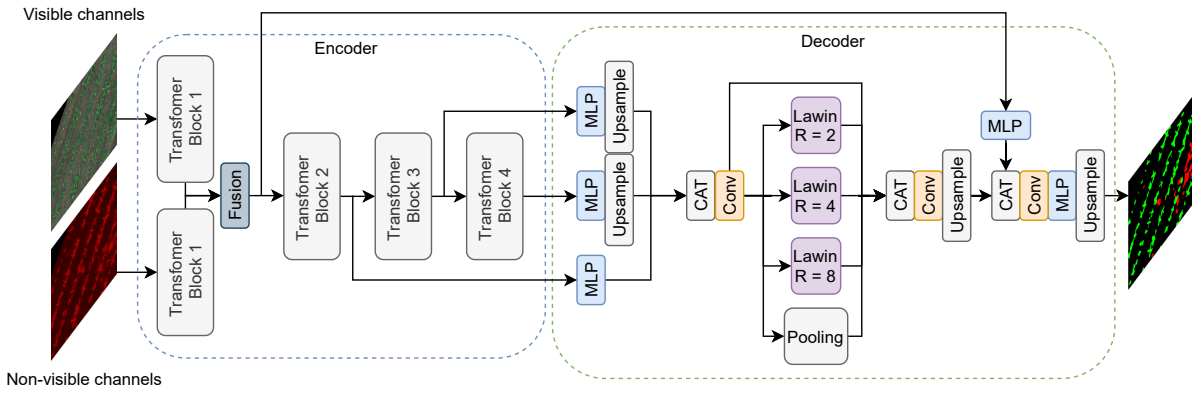

Figure 1: SplitLawin architecture. The model is based on a Vision Transformer backbone but duplicates the first Transformer block so that the first block processes the RGB channels and the second block processes the multispectral channels. The two branches are then fused and passed through the rest of the original model.

such as remote sensing, autonomous robotics, and unmanned aerial vehicles (UAVs) [2, 3]. Among these, UAVs are gaining prominence for their cost-effectiveness, real-time data acquisition, and ability to operate in heterogeneous field conditions.

One of the most pressing challenges in this context is weed mapping—an essential component of Site-Specific Weed Management (SSWM). Weeds compete with crops for nutrients, water, and light, diminishing yields and profitability. Traditional blanket herbicide applications incur high economic costs, pose ecological risks, and foster herbicide-resistant weed species. Weed mapping enables the precise localization of unwanted vegetation, facilitating targeted intervention. However, deploying deep learning models for weed segmentation on UAVs faces critical bottlenecks, including annotation cost, computational complexity, and real-time processing requirements. In this work, we report on three complementary research directions that address these issues:

- *Transfer learning for multispectral data:* Leveraging pre-trained models to adapt RGB-based deep learning architectures for multispectral weed segmentation, enabling efficient feature extraction and improved performance.
- *Knowledge distillation:* Compressing large, complex models into lightweight alternatives suitable for real-time inference on UAVs without significant loss of accuracy.
- *Unsupervised learning via crop-row detection:* Developing a framework that generates pseudo-ground truth from crop-row patterns coupled with an efficient segmentation learning model, eliminating the need for manual annotations while maintaining robust segmentation performance.

2. Transfer Learning for Multispectral Data in Weed Mapping

Multispectral imagery enriches standard RGB data by capturing reflectance patterns across multiple wavelengths, which can benefit the differentiation of crops from morphologically similar weeds. Traditional approaches based on Convolutional Neural Networks (CNNs) have shown promise, but their computational demands limit deployment on UAVs. Furthermore, the usage of multispectral data in precision agriculture limits the application of existing trainings. To overcome this constraint, we proposed a novel weed segmentation framework leveraging lightweight Vision Transformers, explicitly optimized for performance and efficiency in drone-based agricultural systems, coupled with a multispectral adaptation techinque [4].

Our architecture extends Vision Transformer models to handle multispectral UAV imagery by introducing a spectral adaptation strategy. Specifically, starting from the Lawin segmentation model [5], we adapted it to fine-tune RGB-pretrained weights for multispectral inputs, enabling knowledge transfer across both domains and spectral channels. Furthermore, we introduced two architectural variants: *DoubleLawin* and *SplitLawin*, which incorporate a dual-branch encoder. The *DoubleLawin* model

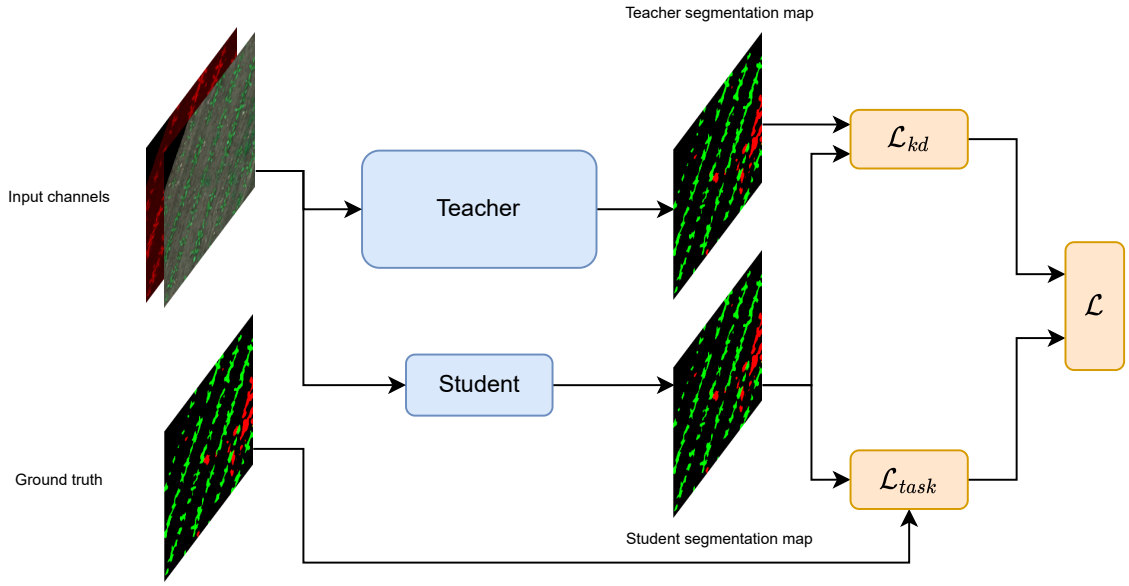

Figure 2: An overview of our knowledge distillation process. The teacher model is a large, complex model that has been pre-trained on a large dataset and fine-tuned on the WeedMap dataset. The student model is a smaller, more efficient model that is trained to mimic the teacher's predictions. The distillation process involves training the student model to minimize the difference between its predictions and the teacher's predictions, using a combination of soft and hard labels.

processes RGB and multispectral channels in parallel using two encoders. In contrast, the *SplitLawin* model splits only the first Transformer block to handle each channel separately before merging them for subsequent processing. This design choice allows for more efficient feature extraction and improved performance on multispectral data. SplitLawin is depicted in Fig. 1. The experimental evaluation on the WeedMap dataset [6] demonstrated that our SplitLawin variant achieved state-of-the-art F1 scores while maintaining a negligible 1ms increase in inference time compared to the baseline. This performance-cost balance makes it viable for real-time onboard execution.

These results indicate two separate findings: first, Transformer-based architectures, when properly adapted, can outperform CNNs in predictive accuracy and computational efficiency; second, pre-training on RGB datasets can be successfully leveraged in multispectral applications.

3. Knowledge Distillation for Lightweight Weed Mapping on Drones

Deploying deep learning models on drones for weed mapping presents a computational bottleneck, primarily when inference must be performed in real time with strict constraints on power and processing capacity. To address this challenge, our research explores using knowledge distillation (KD) to train ultra-lightweight models capable of accurate semantic segmentation in the field [7]. Our KD process is illustrated in Fig. 2.

We investigated several KD techniques using the WeedMap dataset, where a large teacher model (43.5 GMacs) guides the training of a smaller student model to approximate its performance. Our most compact model required only 0.5 GMacs while achieving strong performance across both the Rheinbach and Eschikon subsets. Notably, while the gains on Rheinbach were modest (+2% F1), the Eschikon subset showed a substantial 7% improvement with distillation, underscoring the value of KD in more complex or less homogeneous field conditions.

This performance-efficiency trade-off makes such distilled models ideal for integration in UAV systems and handheld platforms, such as smartphones used in mobile ground control stations or agricultural land robots. The result is a practical, scalable solution for SSWM, capable of delivering real-time insights without the overhead of high-end GPUs.

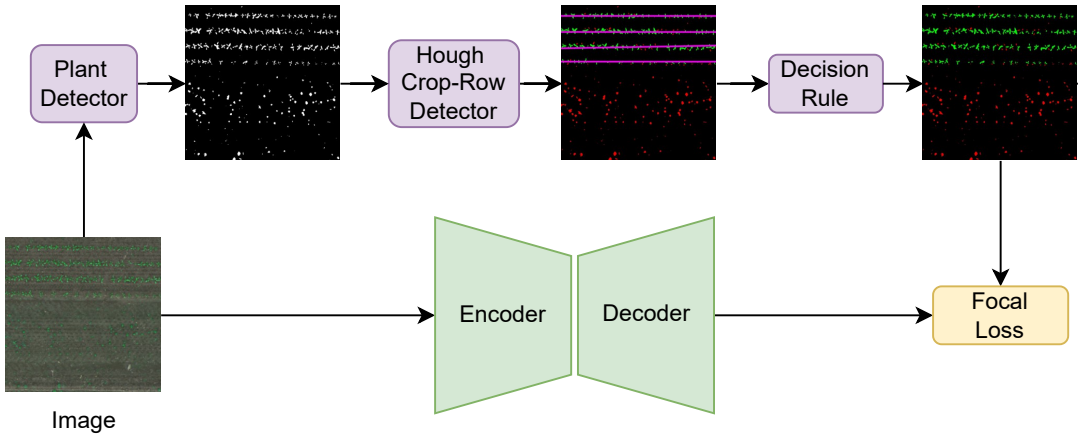

Figure 3: Overview of the RoWeeder framework, illustrating the training process of the semantic segmentation model. In the top branch, the image is fed into a plant detection algorithm, and the mask produced is used to detect crop rows. Every plant on a crop row is classified as a crop; otherwise, it is classified as a weed. This pseudo-ground truth is used to train a deep learning model.

This work highlights how KD can unlock new levels of deployability for AI models in agriculture, transforming drone platforms from data collectors into intelligent agents for sustainable and responsive farm management.

4. Unsupervised Weed Mapping via Crop-Row Detection

We recently proposed RoWeeder, a lightweight, fully automatic framework for unsupervised weed mapping that combines crop-row detection with a noise-resilient deep learning model. The key insight is to exploit the spatial regularity of crop rows—where crops are aligned and weeds typically grow between rows—to generate a pseudo-ground truth for training without manual annotation [8]. The overall framework is illustrated in Fig. 3.

The method first detects crop rows to distinguish intra-row crops from inter-row vegetation. Using this spatial layout, pseudo-labels are generated and used to supervise a segmentation model based on a SegFormer encoder and a custom decoder designed for feature fusion under label noise. This approach bypasses the need for hand-labeled data, which remains a significant bottleneck in deploying deep learning models in agricultural settings.

RoWeeder was evaluated on the WeedMap dataset, achieving an F1 score of 75.3% and outperforming prior weakly supervised methods. Ablation studies on decoder architectures showed that a flat decoder with summation-based fusion yielded the highest robustness to label noise. The system effectively differentiates crops from weeds even in challenging intra-row scenarios, where conventional crop-row methods often misclassify vegetation as crops.

5. Conclusion

This work comprehensively overviewed our ongoing research in drone-based weed mapping for precision agriculture. By integrating multispectral imaging, deep learning, and model compression techniques, we aim to develop efficient and deployable solutions that can operate in real-time on UAV platforms. Our findings highlight the potential of transfer learning and knowledge distillation to enhance the performance of lightweight models. At the same time, our unsupervised framework—RoWeeder—demonstrated the feasibility of generating pseudo-ground truth for training segmentation models without manual annotations.

Looking forward, future research will focus on improving these models by incorporating few-shot semantic segmentation techniques. This approach will allow models to generalize better with minimal labeled data, reducing reliance on large annotated datasets. In parallel, we aim to explore the integration of real-time data collection and model adaptation, enabling drones to make immediate adjustments based on the observed environment. Further work will also investigate the application of these models to a broader range of crops and environmental conditions, expanding the versatility of UAV-based solutions in precision agriculture. Moreover, standard evaluation metrics like accuracy or F1 score may fail to capture the agronomic utility of predictions. Future directions include the development of region-aware metrics that reflect spraying efficiency and crop safety, as well as dynamic temporal modeling to accommodate evolving plant morphology across growth stages.

Acknowledgments

A Ph.D. fellowship funds the research of Pasquale De Marinis within the framework of the Italian "D.M. n. 352, April 9, 2022" – under the National Recovery and Resilience Plan, Mission 4, Component 2, Investment 3.3 – Ph.D. Project "Computer Vision techniques for sustainable AI applications using drones", co-supported by "Exprivia S.p.A." (CUP H91I22000410007).

Declaration on Generative AI

We used GPT-4o and Grammarly to enhance grammar and spelling accuracy in preparing this work. After using these tools, we reviewed and edited the content as necessary and took full responsibility for the publication's content.

References

[1] FAO, How to Feed the World in 2050. Insights from an Expert Meet, FAO (2009).

[2] S. G. Vougioukas, Agricultural Robotics, Annual Review of Control, Robotics, and Autonomous Systems 2 (2019) 365–392.

[3] M. Burke, A. Driscoll, D. B. Lobell, S. Ermon, Using Satellite Imagery to Understand and Promote Sustainable Development, CoRR abs/2010.06988 (2020). arXiv:2010.06988.

[4] G. Castellano, P. De Marinis, G. Vessio, Weed mapping in multispectral drone imagery using lightweight vision transformers, Neurocomputing 562 (2023) 126914. URL: https://www.sciencedirect.com/science/article/pii/S0925231223010378. doi:10.1016/j.neucom.2023.126914.

[5] H. Yan, C. Zhang, M. Wu, Lawin Transformer: Improving Semantic Segmentation Transformer with Multi-Scale Representations via Large Window Attention, 2022. URL: http://arxiv.org/abs/2201.01615. doi:10.48550/arXiv.2201.01615, arXiv:2201.01615 [cs].

[6] I. Sa, M. Popović, R. Khanna, Z. Chen, P. Lottes, F. Liebisch, J. Nieto, C. Stachniss, A. Walter, R. Siegwart, WeedMap: A Large-Scale Semantic Weed Mapping Framework Using Aerial Multispectral Imaging and Deep Neural Network for Precision Farming, Remote Sensing 10 (2018) 1423.

[7] G. Castellano, P. De Marinis, G. Vessio, Applying Knowledge Distillation to Improve Weed Mapping with Drones, in: 2023 18th Conference on Computer Science and Intelligence Systems (FedCSIS), 2023, pp. 393–400. URL: https://ieeexplore.ieee.org/abstract/document/10305897. doi:10.15439/2023F960.

[8] P. De Marinis, G. Vessio, G. Castellano, RoWeeder: Unsupervised Weed Mapping Through Crop-Row Detection, in: A. Del Bue, C. Canton, J. Pont-Tuset, T. Tommasi (Eds.), Computer Vision – ECCV 2024 Workshops, Springer Nature Switzerland, Cham, 2025, pp. 132–145. doi:10.1007/978-3-031-91835-3_9.