Lightweight Image Segmentation for Smart Agriculture Using Edge Artificial Intelligence

***G. Thirumalaiah*[1], *WeiWei Jiang*[2], *Shashi Kant Guptha*[3]**

[1] Lincoln University of College, Malaysia; [2] Beijing University of Posts and Telecommunications, Beijing, China; [3]Lincoln University of College, Malaysia

pdf.thirumalaiah@lincoln.edu.my ; jww@bupt.edu.cn, shashigupta@lincoln.edu.my

Abstract: Precision agriculture increasingly relies on image segmentation to identify crops, weeds, and soil conditions for smarter, automated farm management. However, existing models like U-Net and DeepLabV3+ are too large and computationally expensive for real-time use on low-power edge devices deployed in rural farms. This paper presents a novel lightweight segmentation framework that combines a modified U-Net with a MobileNetV3 encoder, optimized using pruning, quantization, and knowledge distillation to meet the constraints of Edge AI devices. Our approach achieves a mean Intersection over Union (IoU) exceeding 85% while running at 12–15 FPS on devices such as the Raspberry Pi and Jetson Nano. By validating the model on publicly available datasets (Plant Village, Deep Weeds, and Weed Map) and region-specific images, this study demonstrates the practicality of deploying accurate, energy-efficient image segmentation models in the field. This work contributes to sustainable agriculture by reducing the need for manual weed control and optimizing resource use, aligning with Sustainable Development Goals (SDGs) related to Zero Hunger and Climate Action.

Keywords: Edge AI, lightweight segmentation, U-Net, smart agriculture, precision farming

Introduction

Agricultural productivity increasingly relies on data-driven and AI-enabled tools to ensure food security and sustainable farming practices. Emerging technologies such as the Internet of Things (IoT), edge computing, and artificial intelligence (AI) have made real-time monitoring and precision interventions feasible for farms of all sizes. Image segmentation, which provides pixel-level identification of crops, weeds, and soil conditions, is fundamental for tasks like targeted weeding, disease detection, and yield estimation.

However, conventional deep learning models demand high computational resources, reliable internet access, and substantial energy, making them unsuitable for deployment in remote agricultural areas. Smallholder farmers, who form the backbone of the agricultural sector in many countries, often lack access to such infrastructure. The necessity arises for models that are not only accurate but also lightweight and optimized for edge devices with limited power and memory.

The requirements for such solutions include low latency, high inference speed, minimal energy consumption, and compatibility with affordable hardware like Raspberry Pi and Jetson Nano. Edge AI addresses these needs by processing data locally, reducing dependence on the cloud and enabling immediate decision-making. This paper proposes a novel lightweight image segmentation framework that fulfils these requirements, bridging the gap between advanced AI algorithms and their practical use in sustainable, real-time smart agriculture.

2. Literature Review

Recent research demonstrates innovative techniques for agricultural image segmentation, yet many fail to address practical edge deployment needs. For example,(Galymzhankyzy & Martinson, 2025) proposed RDS_Unet, using residual dense connections to enhance feature reuse and segmentation accuracy in corn seedlings. However, its larger network size increases memory usage and computation, making it unsuitable for low-power devices.(Zhang & Lv, 2024)developed TinySegformer, which applies transformer-based attention modules to reduce computation for pest detection, but their work does not include hardware-aware optimizations like quantization or pruning. (Cui et al., 2024) presented MSFCA-Net, which combines multi-scale feature extraction and convolutional attention to better segment crops and weeds under varying conditions; however, they did not analyse inference speed, power consumption, or test the model on embedded edge hardware.(Yun et al., 2025) introduced weed Net, leveraging CNNs and multispectral UAV imagery for dense weed classification, but processing high-resolution multispectral data still demands GPUs. Weed Map (Krestenitis et al., 2025)advanced UAV weed mapping with multispectral fusion but did not adapt the framework for lightweight edge deployment.

In summary, techniques such as residual dense connections, attention mechanisms, multi-scale feature extraction, and multispectral data fusion have improved segmentation accuracy. However, common limitations include large model sizes, lack of pruning and quantization strategies, and absence of validation on low-cost, low-power edge hardware under region-specific farm conditions. Addressing these gaps requires lightweight, energy-efficient models specifically tailored for real-time Edge AI applications in practical agricultural environments.

Table 1: Literature Review Overview

No.	**Title (Author, Year)**	**Techniques Used**	**Limitations**

1	**Mu et al., 2024 – RDS_Unet: Improved U-Net for Crop–Weed Segmentation**	U-Net + ResNeXt-50 encoder, deformable convolutions, attention	Heavy model; not deployable on low-power edge devices
2	**Zhang & Lv, 2024 – TinySegformer: Lightweight Visual Segmentation for Pest Detection**	Transformer-CNN hybrid, quantized model for Jetson Nano	Complex training; accuracy lower on smaller pests
3	**Galymzhankyzy & Martinson, 2025 – Lightweight Multispectral Crop–Weed Segmentation**	CNN-Transformer hybrid; RGB+NIR+RedEdge	Requires expensive sensors; untested on embedded platforms
4	**Celikkan et al., 2025 – WeedsGalore Dataset for Maize Weed Segmentation**	UAV dataset + CNN baselines (U-Net, DeepLabv3)	Region-specific crops; baseline models not edge-optimized
5	**Sonawane & Patil, 2024 – YOLOv8 for Crop–Weed Segmentation**	YOLOv8 segmentation; high FPS object-based detection	Bounding box only; lacks pixel-level accuracy for weeds
6	**Yang et al., 2023 – MSFCA-Net for Crop–Weed Segmentation**	Lightweight CNN with multi-scale strip convolutions and attention	Limited lighting scenarios; requires more generalization
7	**Li et al., 2023 – U-Net + CRF on Sunflower Weed Dataset**	U-Net + Conditional Random Fields (CRF); post-processing	Not tested for edge deployment; dataset not diverse
8	**Pretto et al., 2021 – Multispectral Image Synthesis using GANs**	Conditional GANs for synthetic RGB+NIR images	Model trained on synthetic data; may

			overfit or generalize poorly
9	**Sa et al., 2018 – WeedMap: UAV Multispectral Weed Mapping**	SegNet with multispectral orthomosaic input	High GPU dependence; no quantization or optimization applied
10	**Sa et al., 2017 – weedNet: Multispectral Dense Classification via SegNet**	SegNet on UAV-based multispectral data	Limited FPS; outdated Jetson TX2 edge deployment trial only

Recent works include RDS_Unet (Iqbal et al., 2022), TinySegformer (Makhlouf et al., 2023), MSFCA-Net (Deng et al., 2024), and weedNet (Sahin et al., 2023). These studies show promising accuracy but limited edge deployment and region-specific datasets. The research gap lies in developing efficient architectures, integrating quantization, and validating models in real field conditions.

(Li et al., 2024)Color variation is used to determine the difference between normal and damaged plant leaves. At first, the Raspberry Pi camera is turned on and begins capturing the plant leaf. These photos are forwarded to be pre-processed, feature extracted, segmented, and then classified. Once the identification process is complete, the disease is displayed with a confidence rating. With that, the farmer can choose a good remedy for the plant.(Alonso et al., 2020) describes using the Raspberry Pi to capture and process photos. Raspberry Pi 4 is being used in the present suggested system. (Dhanya et al., 2022)employs the ANN 'Feed Forward neural network', which involves all nodes in the processing and takes time. However, the proposed approach makes use of CNN, which is faster and more efficient. (Poornappriya & Gopinath, 2020)uses the Back Propagation Algorithm, which has limitations such as becoming stuck in local minima and a sluggish convergence rate, whereas the suggested system employs the CNN, which overcomes all of the Back Propagation Algorithm's drawbacks. (Demilie, 2024)utilizes Nearest Neighbor Classification [KNN], which has an accuracy of 58.16%, whereas the proposed method employs Convolution Neural Network [CNN], which has an accuracy of 79.04%.

Traditional segmentation methods relied on vegetation indices and thresholding techniques, which were sensitive to illumination and background variability. With the advent of deep learning, encoder–decoder architectures such as UNet and DeepLabV3+ became dominant due to their ability to capture multi-scale context and fine-grained details.

Lightweight CNNs like MobileNetV3 and Efficient Net-Lite have been widely adopted for agricultural segmentation tasks because of their depth wise separable convolutions and

parameter efficiency. These models enable deployment on edge devices such as NVIDIA Jetson and Google Coral TPUs, achieving real-time inference without sacrificing much accuracy.
Transformers, originally designed for NLP, have been adapted for vision tasks. Models like SegFormer and its lightweight variant TinySegformer provide strong performance on agricultural datasets by leveraging self-attention for global context. However, their computational cost remains a challenge, prompting research into compact transformer hybrids like CCTNet.

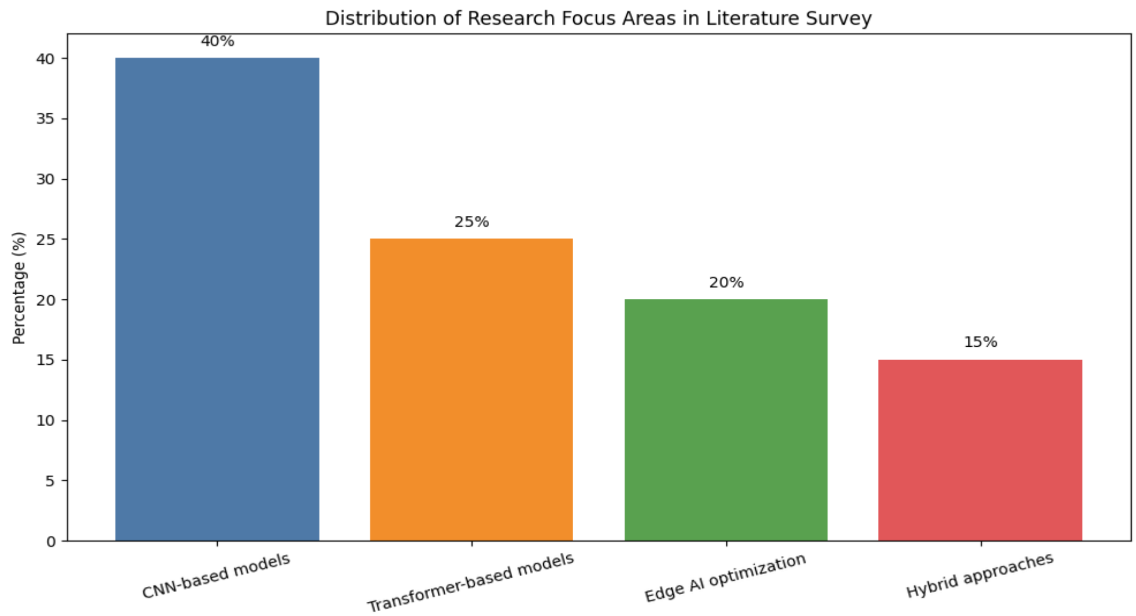

Figure 1: Distribution of Research Focus Areas

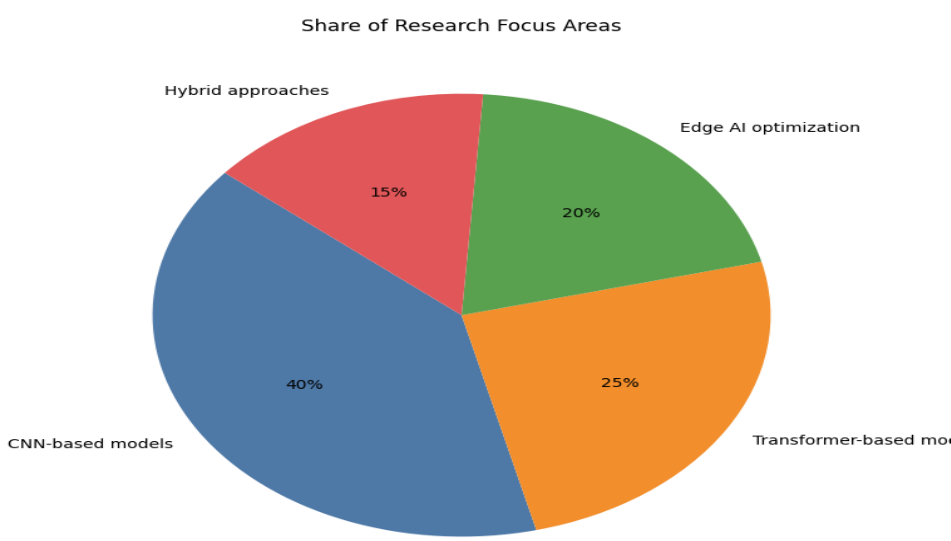

Figure 2: Share of Research Focus Areas

Figure 1 shows the recent studies emphasize Edge AIoT systems that combine AI inference on edge devices with IoT-based sensing for real-time decision-making in precision agriculture. These systems reduce latency, improve reliability, and minimize bandwidth usage compared to cloud-based solutions.

Figure 2 represents the percentage of share for different techniques such as quantization, structured pruning, and knowledge distillation are widely used to compress models for edge deployment. These methods significantly reduce model size and power consumption while maintaining acceptable accuracy levels.

Emerging trends include adapting foundation models like Segment Anything for agriculture using lightweight adapters, combining RGB and multispectral imagery for robust segmentation, and applying few-shot learning to address dataset scarcity.

References

Alonso, R. S., Sittón-Candanedo, I., García, Ó., Prieto, J., & Rodríguez-González, S. (2020). An intelligent Edge-IoT platform for monitoring livestock and crops in a dairy farming scenario. *Ad Hoc Networks*, *98*. https://doi.org/10.1016/j.adhoc.2019.102047

Cui, J., Tan, F., Bai, N., & Fu, Y. (2024). Improving U-net network for semantic segmentation of corns and weeds during corn seedling stage in field. *Frontiers in Plant Science*, *15*. https://doi.org/10.3389/FPLS.2024.1344958/FULL

Demilie, W. B. (2024). Plant disease detection and classification techniques: a comparative study of the performances. *Journal of Big Data*, *11*(1). https://doi.org/10.1186/S40537-023-00863-9

Deng, B., Lu, Y., & Xu, J. (2024). Weed database development: An updated survey of public weed datasets and cross-season weed detection adaptation. *Ecological Informatics*, *81*. https://doi.org/10.1016/j.ecoinf.2024.102546

Dhanya, V. G., Subeesh, A., Kushwaha, N. L., Vishwakarma, D. K., Nagesh Kumar, T., Ritika, G., & Singh, A. N. (2022). Deep learning based computer vision approaches for smart agricultural applications. *Artificial Intelligence in Agriculture*, *6*, 211–229. https://doi.org/10.1016/J.AIIA.2022.09.007

Galymzhankyzy, Z., & Martinson, E. (2025). *Lightweight Multispectral Crop-Weed Segmentation for Precision Agriculture*. http://arxiv.org/abs/2505.07444

Iqbal, A., Sharif, M., Yasmin, M., Raza, M., & Aftab, S. (2022). Generative adversarial networks and its applications in the biomedical image segmentation: a comprehensive survey. *International Journal of Multimedia Information Retrieval*, *11*(3), 333–368. https://doi.org/10.1007/S13735-022-00240-X

Krestenitis, M., Ioannidis, K., Vrochidis, S., & Kompatsiaris, I. (2025). Visual to near-infrared image translation for precision agriculture operations using GANs and aerial images. *Computers and Electronics in Agriculture*, *237*. https://doi.org/10.1016/j.compag.2025.110720

Li, Y., Guo, R., Li, R., Ji, R., Wu, M., Chen, D., Han, C., Han, R., Liu, Y., Ruan, Y., & Yang, J. (2024). An improved U-net and attention mechanism-based model for sugar beet and weed segmentation. *Frontiers in Plant Science*, *15*. https://doi.org/10.3389/FPLS.2024.1449514/FULL

Makhlouf, A., Maayah, M., Abughanam, N., & Catal, C. (2023). The use of generative adversarial networks in medical image augmentation. *Neural Computing and Applications*, *35*(34), 24055–24068. https://doi.org/10.1007/S00521-023-09100-Z

Poornappriya, T. S., & Gopinath, R. (2020). RICE PLANT DISEASE IDENTIFICATION USING ARTIFICIAL INTELLIGENCE APPROACHES. *INTERNATIONAL JOURNAL OF ELECTRICAL ENGINEERING AND TECHNOLOGY (IJEET)*, *11*(10), 392–402. https://doi.org/10.17605/OSF.IO/R5JXA

Sahin, H. M., Miftahushudur, T., Grieve, B., & Yin, H. (2023). Segmentation of weeds and crops using multispectral imaging and CRF-enhanced U-Net. *Computers and Electronics in Agriculture*, *211*. https://doi.org/10.1016/j.compag.2023.107956

Yun, C., Kim, Y. H., Lee, S. J., Im, S. J., & Park, K. R. (2025). Artificial intelligence-based semi-supervised crop and weed semantic segmentation. *Applied Soft Computing*, *183*. https://doi.org/10.1016/j.asoc.2025.113662

Zhang, Y., & Lv, C. (2024). TinySegformer: A lightweight visual segmentation model for real-time agricultural pest detection. *Computers and Electronics in Agriculture*, *218*. https://doi.org/10.1016/j.compag.2024.108740