Ann. Agric. Res. New Series Vol. 46 (4) : 373-380 (2025)

LeafNet-CBAM: a lightweight attention-enhanced network for weed image classification

Ritika[1&3], Savita Kumari Sheoran[2*] and Arjun Singh[3]

Indira Gandhi University, Meerpur, Rewari, India -123401
**Corresponding Author's Email: savita.sheoran@igu.ac.in*

Received: September 2025; Revised Accepted: November 2025

Abstract

Accurate weed identification is essential for precision cotton farming, as early-stage weed infestation significantly affects crop yield and resource efficiency. This study proposes LeafNet (L-Net), a lightweight and stable convolutional neural network (CNN) designed for cotton-weed image classification under real-field conditions. L-Net integrates depth wise separable convolutions with Convolutional Block Attention Module (CBAM) and residual connections to enhance discriminative feature learning while maintaining low computational complexity. A cotton field image dataset was collected under varying illumination and background conditions and augmented to improve robustness. Experimental results demonstrate that L-Net outperforms EfficientNet-B0, ResNet-50, and MobileNet, achieving an accuracy of 98.99%, a recall of 99.17%, and an F1-score of 99.00%, with minimal false classifications. Ablation and confusion matrix analyses confirm the effectiveness of attention and residual learning in improving model stability and generalization. The proposed L-Net architecture offers a reliable and efficient solution for real-time and edge-based weed detection in precision agriculture.

Keywords: Leafnet, cotton–weed classification, deep learning, convolutional block attention module, precision agriculture

Introduction

Weed infestation severely limits cotton productivity, especially during early growth stages, where competition for nutrients, water, and light leads to yield loss and higher production costs (Adhinata *et al.,* 2024; Bharathi *et al.,* 2024). Manual weed control and herbicide application methods are less environmentally friendly and more expensive, besides causing the emergence of herbicide resistance, soil degradation, and air pollution. Visually similar weeds add more difficulty to the already challenging problem of weed control. Cotton weed management through traditional methods was largely dependent on either manual weeding or the use of broad-spectrum herbicides that both adversely affected the environment. The overuse of pesticides causes damage to the soil, diminishes the variety of species, and creates new weed species resistant to herbicides (Abdelhak, 2024). The vast advances in deep learning (DL) and computer vision (CV) technologies have made the use of automated weed detection in precision agriculture possible(Ashraf and Khan, 2020; Adhinata *et al*., 2024). However, precision agriculture can make the process of weed control easier and more accurate due to the crop-weed identification from field images

[1] Ph.D. Scholar, Indira Gandhi University, Meerpur, Rewari, India -123401
[2]Professor, Indira Gandhi University, Meerpur, Rewari, India-123401
[3]ICAR-Indian Agricultural Research Institute, New Delhi, India-110012

(Visentin *et al.*, 2023; Chouhan *et al.*, 2024). DL and CNN made the detection of weeds by images more precise; however, many approaches use large and complex architectures. These large models usually lead to over fitting when dealing with small datasets of cotton weeds and perform inconsistently (Mesías-Ruiz *et al.*, 2024).

The variability in cotton field images is high due to changes in the lighting, the clutter of the background, and the overlapping of the plant structures, while the visual similarity between the weeds and the cotton plants adds to the problem of classification (Gate *et al.*, 2024). These challenges require light weight, stable, and well-generalizing models. To address this, L-Net, a lightweight CNN designed for cotton weed classification, is proposed. L-Net utilizes depth wise separable convolutions, a mixed activation function, and intense regularization to provide an efficient, stable, and accurate classification on small datasets.

The main contributions of this research are outlined below:

1. A CNN architecture, L-Net, which is less heavy and stable, is presented for the classification of cotton weed images in real-field conditions.
2. An advanced activation mechanism CBAM, which combines GELU and Leaky ReLU is used to enhance the convergence and robustness of the model(Santeramo and Jelliffe, 2024).
3. The exhaustive evaluations done on the cotton weed dataset reveal that the model yields high accuracy, precision, recall, and F1-score with a reduced number of parameters.
4. The architecture proposed is applicable for real-time and edge-based deployment in precision cotton agriculture considering its low memory requirement.

Materials and Methods

The experiment was conduced during the year 2025 of cotton and weed data set collected from the agricultural farm located in Berli Khurd village, Rewari, Haryana (longitude – 76°32′34.8″E and latitude – 28°10′35.4″N). A commonly cultivated cotton variety was grown following the standard package of agricultural practices. Images were captured using an iPhone 15 mobile device equipped with a rear dual-camera system along with a 12 MP ultra-wide camera (f/2.4, 13 mm, 120° field of view), maintaining a 4:3 aspect ratio. The field collection included an aggregate of 1,872 images of the cotton crop and weed images. Representative samples of cotton and weed images are illustrated in Figure 1.

Fig. 1. Cotton and weed images

The dataset is divided into two part soriginal dataset and augmented dataset. The original dataset consisted of 1,872 high-resolution images of cotton and weeds. Among these, 931 images capture cotton plants, while the remaining 941 images represent various weed species present in the field. Each image was uniformly maintained at a resolution of 800 × 600 px. The primary objective of the cottonweed dataset is to facilitate real-time weed detection by effectively capturing discriminative visual attributes such as shape, size, texture, and fine structural patterns. The augmented dataset is indeed a larger one, containing around 4885 images exhibiting cotton and varieties of weeds.

Conventional CNN architectures such as EfficientNet-B0, ResNet-50, and MobileNet are widely used for image classification due to their strong hierarchical feature learning capabilities (Bhardwaj and Gupta, 2025). EfficientNet-B0 is a compound model that scales depth, width, and resolution through the use of MB Conv blocks with the squeeze-and-excitation approach, thereby reaching great accuracy with fewer parameters (Adhinata *et al.*, 2024). ResNet-50 improves learning of deep spatial and semantic features and thus is effective for the separation of crops and weeds that look alike, adopting residual skip connections to overcome the problem of vanishing gradients (Ahmed *et al.*, 2023). MobileNet uses depth wise separable convolutions as a means of making its computational load very light, and consequently allows for deployment in real-time and on edge-device environments.

The prposed L-Net model is designed for efficient multi-scale feature extraction with low parameter complexity. Its encoder-inspired architecture with skip connections allows the effective capture of low-level as well as high-level features, which provides the improvement in the discrimination of crop-weeds classification tasks. The compactness of L-Net design gives more strength against the changing of light conditions, the mixing up of the background and the different growth stages. The main architectural differences among L-Net, EfficientNet-B0, ResNet-50, and Mobile Net are summarised in Table 1.

L-Net presents an innovative DL architecture that leverages attention mechanisms for cotton–weed discrimination with a convolutional backbone of CBAM. While traditional CNN depend on the hierarchical convolutional features, L-Net applies efficient multi-scale feature fusion and the attention-driven refinement to accentuate the crop-weed regions of interest. The integration of the channel and spatial attention mechanisms allows the network to get both high-resolution local details and wider contextual cues,required for successfully separating cotton seedlings and weeds in difficult field conditions.

L-Net was trained simultaneously with EfficientNet-B0, ResNet-50, and MobileNet, all using the same 224×224 normalized input images, and the overall experiment setup is shown in Fig. 2.

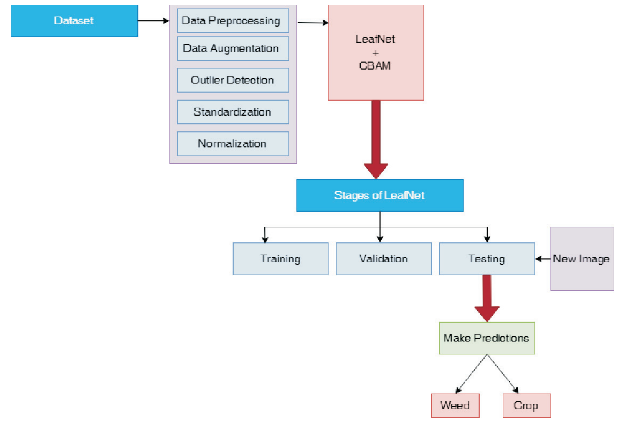

Fig. 2. Methodology for cotton-weed detection

The model was loaded with pre-trained weights, initially restricting some of the early layers of the neural network to only a bit of freezing and retaining the generic features, while allowing the rest of the layers and the CBAM modules to communicate the task-specific patterns to the model by learning from the data. The size of the input images was changed to 224 x 224, with the Adam optimiser, at a learning rate of 1×10{t as the starting point and gradually down via a learning rate scheduler, categorical cross-entropy was used as the loss function for the training process. To target over fitting, batch normalisation and dropout were applied. A batch size of 16, early stopping based on validation loss, and a maximum of 50 epochs were used for training to en-

Table 1. Architectural differences between models

Feature	L-Net	EfficientNet-B0	ResNet-50	MobileNet
Layers	~15–20 (lightweight CNN)	~82	50	53
Key Concept	Leaf-oriented feature extraction using compact convolutional blocks	Compound scaling of depth, width, and resolution	Deep residual learning with identity skip connections	Depthwise separable convolutions with inverted residuals
Computational Efficiency	Very High (low parameters and FLOPs)	High	Moderate–Low	Very High
Accuracy in Weed Detection	Moderate–High(effective for leaf-based patterns)	High	Very High	Moderate–High
Robustness to Variability	Moderate (sensitive to scale and illumination changes)	Strong	Very Strong	Moderate
Transfer Learning	Limited (often trained from scratch or shallow pre-training)	Strong	Strong	Strong

sure stable convergence. The fine-tuned L-Net managed to be computationally efficient while generalising well across different growth stages, soil textures, and lighting conditions.

The dataset was split in a stratified manner to preserve the class distribution into training (70%), validation (15%), and testing (15%) sets. Data augmentation was performed only on the training set that was intended to make the model more capable of generalising in various field situations. The performance on the unseen test set was finally evaluated using accuracy, precision, recall, F1-scoreand confusion matrices. Such a scenario allowed for an unbiased comparison between the performance of L-Net and that of the baseline CNN models while still being able to claim very strong generalization performance.

$$\text{Precision} = \frac{TP}{TP+FP} \quad .. (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad .. (2)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad .. (3)$$

$$\text{F1 score} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \quad .. (4)$$

Results and Discussion

Overall Performance of Models

According to the experimental results, L-Net in cotton-weed classification shows an accuracy of 98.99%, a precision of 98.82%, a recall of 99.17%, and an F1-score of 99.00%. The main reason for this superiority is the model's lightweight structure in combination with multi-scale feature fusion and attention-based refinement, which makes it possible to differentiate visually similar plant and weed species effectively. Among the baseline models, EfficientNet-B0 demonstrated strong performance with an accuracy of 97.68%, followed

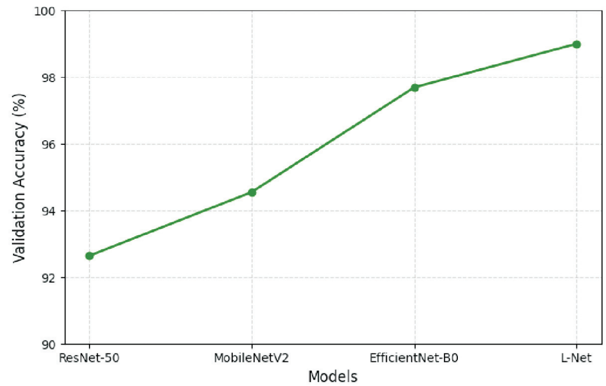

Fig. 3. Accuracy achieved by models

by MobileNet at 94.54%, while ResNet-50 achieved a lower validation accuracy of 92.63%, suggesting over fitting due to the lack of attention mechanisms. Overall, L-Net outperformed all comparative models across evaluation metrics, as illustrated in Figure 3.

The performance results from the models are clearly demonstrated by the summary in Table 2.

The training and validation performance metrics (Table 3) were used to provide a comparative assessment of the proposed and baseline architectures. L-Net was the model with the highest accuracy and the lowest loss. The Efficient Net-B0 model also showed strong performance, with its benefits attributed to compound scaling and effective feature extraction. The MobileNet gave moderate accuracy along with higher loss caused by the small model capacity, which made it perfect for real-time use but less resistant in hard field conditions. The ResNet-50 model reached a flawless training accuracy, with amigo validation lossfalling, and consequently, the validation accuracy dropped, caused by over fitting, indicating the limitations of attention-free deep architectures in the agricultural image classification task.

The comparative performance of MobileNet, ResNet-50, EfficientNet-B0, and L-Net across vali-

Table 2. Performance of models for cotton–weed classification

Model	Validation Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNet	94.54	95.14	94.54	94.56
ResNet-50	92.63	93.45	92.63	92.65
EfficientNet-B0	97.68	97.72	97.68	97.68
L-Net	98.99	98.82	99.17	99.00

Table 3. Comparative performance of architectures

Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
MobileNet	0.9621	0.1847	0.9454	0.2316
VGG16	0.9863	0.0289	0.9263	0.4128
EfficientNet-B0	0.9856	0.0674	0.9768	0.1043
L-Net	0.9938	0.0412	0.9899	0.0198

dation accuracy, precision, recall, and F1-score is illustrated in Figure 4. L-Net consistently achieves the highest values across all evaluation metrics, demonstrating superior classification accuracy, robustness, and generalisation capability compared to both lightweight and deep CNN baseline models.

The analysis of the effectiveness of various architectural on the robustness of models when applying L-Net variants to the cotton-weed dataset outlined numerous significant points (Table 4). The comparison of the entire L-Net architecture and its reduced forms showed that the addition of attention mechanisms and residual connections had a crucial role in the improvement of classification accuracy. The L-Net configuration, which integrates CBAM with residual pathways, consistently outperformed the simpler versions in terms of accuracy and recall, indicating better feature separation and training reinforcement. On the other side, the removal of attention modules, residual connections, or network depth caused a significant drop in performance especially during difficult background and occlusion tests. Thus, the ablation results of L-Net suggest that the use of architectural improvements such as the combination of attention mechanisms and residual learning is still necessary for reliable and accurate cotton seedling classification even with low resource consuming network frameworks operating in challenging agricultural environments.

The complete L-Net achieve the performance of 98.99% accuracy and 99.17% recall, due to the powerful combination of the CBAM attention mechanism and residual connections that allowed for the extraction of both fine-grained and contextual features at the same time. The removal of CBAM, however, drew the accuracy down to 97.35%, thus proving the necessity of the attention mechanisms, while shallow and lower-resolution variants showed even less performance, highlighting the importance of network depth and input resolution. The abolishment of residual connections resulted in an additional drop to 96.60%, and the baseline L-Net without attention or residual learning reached the lowest accuracy of 95.90% confirming the impact of the proposed architectural improvements. The results are presented in Figure 5.

The confusion matrix analysis (Figure 6) re-

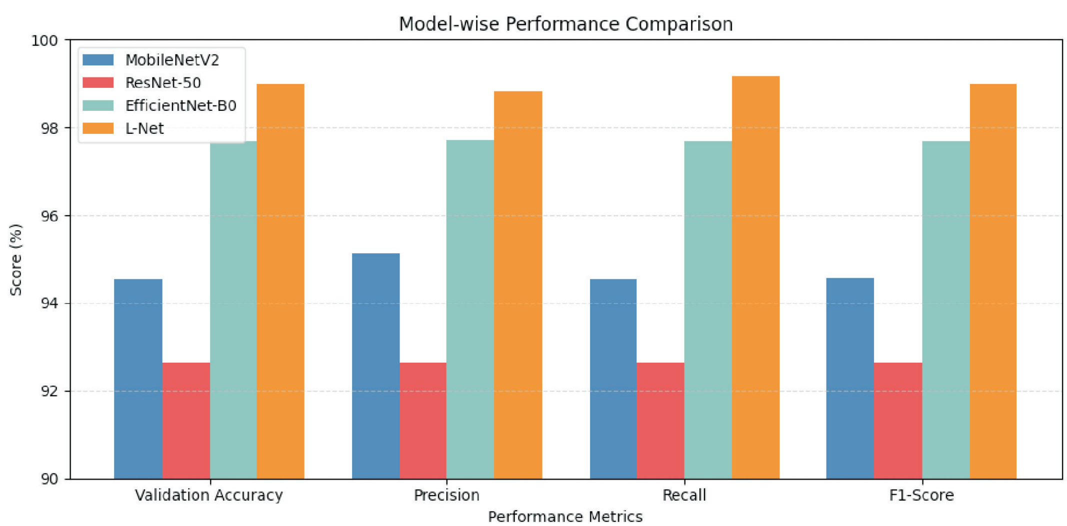

Fig. 4. The comparative performance of MobileNetV2, ResNet-50, EfficientNet-B0, and L-Net

Table 4. Ablation study on L-Net

Model	Attention Type	Accuracy (%)	Recall (%)	Observations
L-Net	CBAM + Residual Connections	98.99	99.17	Full L-Net architecture, best balance between lightweight design and feature enhancement
L-Net w/o CBAM	Residual Connections only	97.35	97.30	Removal of attention slightly degrades feature refinement
Shallow L-Net	CBAM + Fewer Layers	96.80	96.75	Reduced depth limits hierarchical feature extraction
L-Net small input	CBAM + Residual, smaller input	97.05	97.00	Lower input resolution causes minor performance drop
L-Net w/o Residual	CBAM only	96.60	96.55	Absence of residual paths reduces training stability
L-Net baseline	No Attention, No Residual	95.90	95.85	Simplest configuration, fastest but weakest feature learning

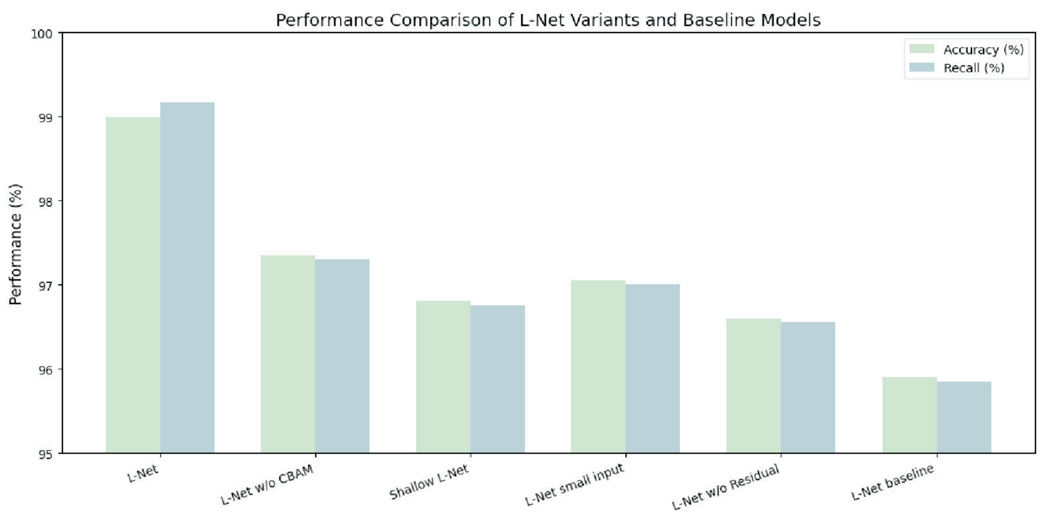

Fig. 5. Accuracy and recall performance of l-net variants under different architectural configurations.

veals the great performance of the proposed L-Net integrated CBAM method in classification as compared to Efficient Net, ResNet-50, and Mobile Net. L-Net achieves uniform results, identifying correctly 404 cotton and 322 weed samples with slight misclassification, which eventually resulted in the highest accuracy of 98.99%. Efficient Net was quite close in performance with further false negatives, while ResNet-50 faced huge misclassification of cotton samples which ultimately led to lower accuracy. MobileNet was able to perform at a moderate level. However, L-Net still had the lowest rates of false positives and false negatives, verifying its superior reliability for cotton-weed classification in difficult field environments.

Effectiveness of L-Net Architecture and Attention Mechanisms

The prposed L-Net model exhibits remarkable efficiency in cotton-weed classification due to the combination of a compact structure with CBAM attention and residual connections. Thus, accurate extraction of local as well as contextual features is achieved, which leads to the impressive performance in terms of accuracy (98.99%) and recall (99.17%). The ablation study verifies that the attention mechanisms play a decisive role in feature discrimination, whereas the residual learning provides the stability and generalization under complex field conditions that help in training.

A comparative study with conventional CNN

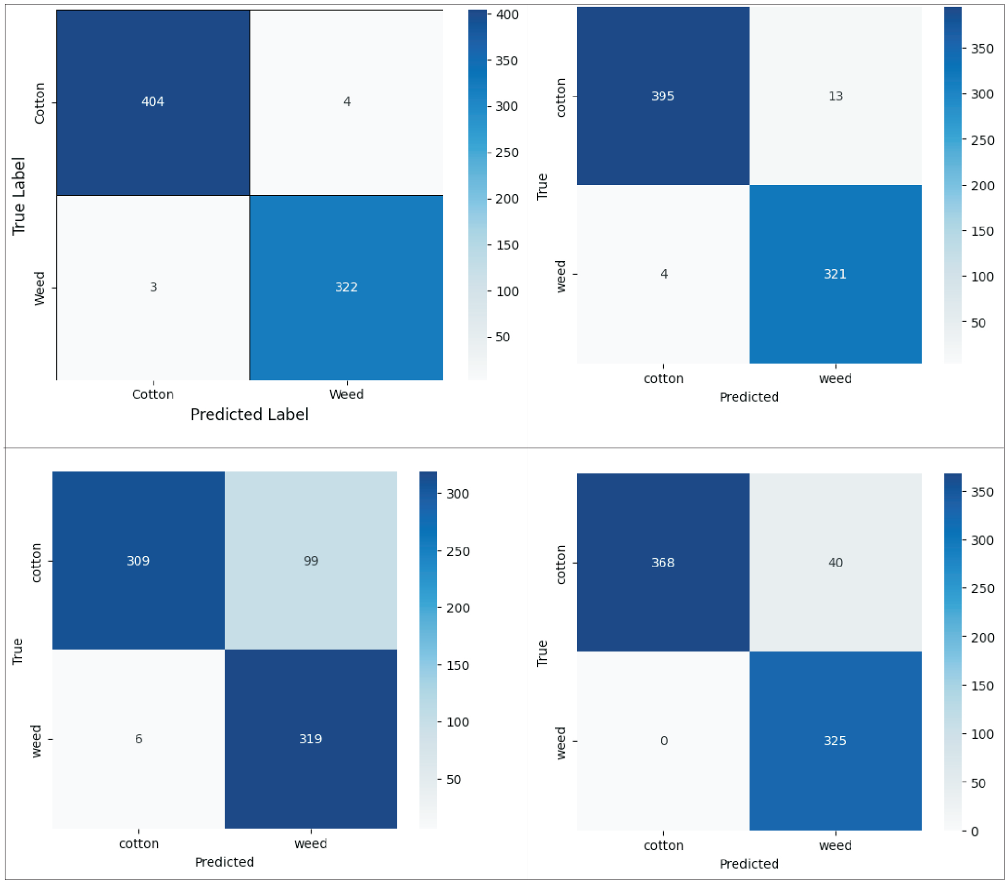

Fig. 6. Confusion matrices of L-Net+CBAM, EfficientNet, ResNet-50, and Mobile Net

networks indicates that L-Net is superior to both heavy and lightweight baseline models in a significant, impressive way. In comparison to baseline models, L-Net is ahead of the mentioned architectures all the time, that is, EfficientNet-B0, ResNet-50, and MobileNet. In spite of the fact that EfficientNet-B0 gets appealing accuracy, it is still a little bit under L-Net. ResNet-50 suffers from over fitting and a decline in generalization, while MobileNet experiences more misclassification as a result of its small feature capacity. These findings emphasize the merit of attention-augmented lightweight architectures over the deeper or iba purely efficiency-driven CNNs in the context of agricultural image classification.

The confusion matrix analysis reveals that L-Net exhibits the minimum number of false positives and false negatives, which is synonymous with maximum crop safety in computer-assisted weed management systems. Moreover, the lightweight structure of the system facilitates its easy movement to edge devices like drones and agricultural robots, thus making it suitable for real-time precision farming. Overall, L-Net is a feasible option for cotton-weed discrimination because of its excellent mix of accuracy, stability, and computational efficiency.

Conclusion

This study presented Leaf Net (L-Net), a lightweight and stable deep learning architecture designed for accurate cotton–weed image classification under real-field conditions. By integrating depth wise separable convolutions, the combination of CBAM attention mechanisms and residual connections in L-Net facilitated the extraction of both detailed local features and global context at the same time while keeping the computational complexity low. A comprehensive performance of L-Net was superior to EfficientNet-B0, ResNet-50, and MobileNet, reaching 98.99% accuracy, 99.17%

recall, and 99.00% F1-score, with very few false positives and negatives.

The role of attention and residual learning in feature discrimination enhancement, training stability, and generalisation capability was confirmed by the ablation studies to be critical ones, especially when working with limited and highly variable agricultural datasets. The confusion matrix analysis further corroborated the reliability of L-Net, emphasising its potency in crop safety by minimising misclassification in the tasks of cotton-weed discrimination. Overall, the proposed L-Net architecture provides a robust and computationally efficient solution for precision agriculture applications. Its lightweight design and high accuracy make it suited for deployment on edge devices, unmanned aerial vehicles, and automated weed management systems. Future work will focus on extending L-Net to multi-class weed recognition, integrating temporal and spectral information, and validating its performance across diverse crops and field conditions to further enhance its applicability in real-world smart farming environments.

References

Abdelhak, M. 2024. Innovative Techniques for Soil and Water Conservation. *Ecosystem Management: Climate Change and Sustainability*, 291–326.

Adhinata, F. D., Wahyono, and Sumiharto, R. 2024. A comprehensive survey on weed and crop classification using machine learning and deep learning. *Artificial Intelligence in Agriculture,* **13**, 45–63. doi: 10.1016/j.aiia.2024.06.005

Ahmed, S. F., Alam, Md. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T. and Rafa, N. 2023. Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review,* **56**: 13521–13617. doi: 10.1007/s10462-023-10466-8

Ashraf, T. and Khan, Y. N. 2020. Weed density classification in rice crop using computer vision. *Computers and Electronics in Agriculture.* **175**: 105590. doi: 10.1016/j.compag.2020.105590

Bharathi, P., Karthickraja, A. and Saravanane, P. 2024. Advanced Precision Agriculture Techniques for Effective Weed Management in Cotton. *Recent Trends in Agriculture and Allied Sciences*, 162–171.

Bhardwaj, A. and Gupta, S. P. 2025. Empowering Sustainable Agriculture Harnessing AI for Enhanced Yields, Resource Efficiency, and Eco-Friendly Farming Practices, in *AI and Ecological Change for Sustainable Development*, (IGI Global Scientific Publishing), 183–216.

Chouhan, S. S., Singh, U. P., Saxena, A., and Jain, S. 2024. Assessing the importance and need of artificial intelligence for precision agriculture. in *Artificial intelligence techniques in smart agriculture*, (Springer), 1–6.

Gate, A., Abate, A. and Semahagn, B. K. 2024. "Cotton Genetics and Breeding," in *Cotton Sector Development in Ethiopia: Challenges and Opportunities*, (Springer), 19–38.

Mesías-Ruiz, G. A., Borra-Serrano, I., Peña, J. M., de Castro, A. I., Fernández-Quintanilla, C. and Dorado, J. 2024. Weed species classification with UAV imagery and standard CNN models: Assessing the frontiers of training and inference phases. *Crop Protection,* **182**, 106721. doi: 10.1016/j.cropro.2024.106721

Santeramo, F. G. and Jelliffe, J. 2024. CBAM and Agriculture: Opportunities, Challenges, and Perspectives.

Visentin, F., Cremasco, S., Sozzi, M., Signorini, L., Signorini, M. and Marinello, F. 2023. A mixed-autonomous robotic platform for intra-row and inter-row weed removal for precision agriculture. *Computers and Electronics in Agriculture,* **214**, 108270. doi: 10.1016/j.compag.2023.108270