International Journal of
Computational Learning & Intelligence
Vol. 04 | Issue 02 | 2025

RESEARCH ARTICLE OPEN ACCESS

# Smart Agriculture: Machine Learning And Deep Learning For Advanced Weed Detection And Crop Quality Enhancement

**Y Bhagya Lakshmi . Y Rakesh . S Subramanyam . T Hari Priya . G Anusha Reddy**

Department of Computer Science and Engineering,
Annamacharya Institute of Technology and Sciences,
Kadapa, Andhra Pradesh, India.

DOI: **10.5281/zenodo.15210829**
Received: 27 January 2025 / Revised: 21 February 2025 / Accepted: 27 March 2025


**Abstract** – The evolution of agricultural systems is ongoing, driven by new technologies that seek to refine traditional farming. The primary goal is to not only boost agricultural output per hectare but also to improve crop quality, all while maintaining the natural environment. Notably, weeds present a considerable threat to crops, as they deplete essential nutrients, water, and light, thus diminishing crop productivity. The uniform spraying of entire fields for weed control results in high costs and detrimental environmental impacts. To mitigate the limitations of standard weed control methods, this research proposes the use of Machine Learning (ML) and Deep Learning (DL) methods to identify and classify weeds within crops. For ML-based approaches, various statistical and texture-based features are extracted, including image moments, mean absolute deviation, and gray level co-occurrence matrix (GLCM). The YOLOv8m algorithm is utilized for weed identification, and for weed classification, features from the CottonWeedID15 and Earlycrop-weed datasets are used to train SVM, Random Forest, and ANN models, using SMOTE to balance the classes. Deep learning models, such as VGG and ConvNeXtBase, are also trained on balanced data for automated feature extraction and classification. The best ML result was a 99.5% accuracy with SVM, and the best DL result was 98% accuracy with ConvNeXt and Random Forest. This demonstrates the potential of these methods for efficient agricultural solutions.

**Index Terms** – Agricultural automation, weed detection, weed classification, machine learning, deep learning, computer vision, YOLOv8, CNN, image processing, precision agriculture, sustainable farming, CottonWeedID15 dataset, Early-Crop-Weed dataset, transfer learning, feature extraction, SMOTE, object detection.

## I. INTRODUCTION

Global population growth is a pressing concern, with projections indicating a rise to 9 billion by 2050. This surge necessitates a 70% increase in agricultural output to meet escalating food demands [1]. The agricultural sector, however, faces numerous challenges including limited arable land, soil salinity, climate change, water scarcity, and weed infestation. Artificial intelligence (AI), particularly through computer vision, machine learning, and deep learning, offers promising solutions to these issues [2]. Weeds, categorized by leaf shape into broadleaf, grasses, and sedges, are detrimental to crops. They compete for resources and harbor pests, leading to significant yield losses, estimated at 40% by the European Crop Protection (ECPA). Traditional weed control methods, such as manual removal [3] or mechanical devices, are labor-intensive or limited in application. Chemical spraying, while common, is costly and environmentally harmful. AI-driven systems provide an alternative by accurately detecting and classifying weeds [4].

Weed detection and classification pose challenges due to the visual similarities between crops and weeds, as well as varying lighting conditions. A typical AI-based weed detection system involves image capture, preprocessing, feature extraction, and classification [5]. Recent advancements in AI, driven by GPU capabilities, have introduced sophisticated algorithms. Deep learning models, however, require substantial computational resources. Transfer learning, which leverages pre-trained models, can mitigate this. Studies have shown that fine-tuning deep learning models on agricultural datasets improves accuracy and reduces training time [6, 7]. Major Research Contributions: This research focuses on developing effective machine learning and deep learning models for weed detection and classification. Key contributions include:

- Annotating the CottonWeedID15 dataset, providing a valuable resource for weed identification research [8].
- Evaluating the effectiveness of statistical and texture features, alongside deep learning features, for weed detection.
- Utilizing U2Net for background removal to enhance image analysis.
- Highlighting the potential of deep learning in agriculture to improve crop yields and promote sustainable practices.

## II. LITERATURE SURVEY

The application of machine learning (ML) and deep learning (DL) techniques has become increasingly prevalent in addressing various challenges within agriculture, and weed detection is a significant area of focus. Numerous studies have explored automated methods to identify and classify weeds in agricultural fields, aiming to reduce reliance on manual and indiscriminate chemical control approaches. Early research in this domain often utilized traditional machine learning algorithms coupled with handcrafted feature extraction techniques. For instance, Support Vector Machines (SVMs) have been extensively investigated for weed density classification. These studies frequently relied on texture-based features extracted from images, such as those derived from the Gray-Level Co-occurrence Matrix (GLCM). By analyzing the spatial relationships between pixel intensities, GLCM features capture information about image texture, which can help differentiate between crops and various weed species. Some of these early works reported promising accuracies, reaching up to 84.85% in specific scenarios.

However, the performance of these methods was often heavily dependent on the quality and robustness of the manually engineered features.

Recognizing the limitations of handcrafted features, researchers have increasingly turned to deep learning techniques, particularly Convolutional Neural Networks (CNNs). CNNs have the inherent capability to automatically learn hierarchical representations of image data, eliminating the need for explicit feature engineering. Various CNN architectures, ranging from relatively shallow networks to deeper models like VGGNet, have been explored for weed detection and classification tasks. These models have demonstrated significant potential in handling the complex visual variations present in agricultural environments, including differences in lighting, growth stages, and weed density. To address the challenges associated with limited and often imbalanced agricultural datasets, several strategies have been employed. Generative Adversarial Networks (GANs) have been explored as a means of data augmentation, allowing for the generation of synthetic weed and crop images to increase the size and diversity of training sets. Transfer learning, a technique that leverages knowledge learned from large, general-purpose image datasets (such as ImageNet), has also proven highly effective. By fine-tuning pre-trained CNN models on smaller agricultural datasets, researchers have been able to achieve high accuracy with reduced training data and computational resources.

Furthermore, the field has seen the application of object detection models like YOLO (You Only Look Once) for localizing and classifying individual weeds within an image. These models offer the advantage of performing both detection and classification in a single forward pass, making them suitable for real-time applications. Different versions and adaptations of YOLO have been investigated for specific agricultural contexts, including weed detection in turfgrass and the classification of diverse crop and weed species. Ongoing research efforts continue to focus on several key areas. These include the development of more robust and efficient feature extraction techniques, the optimization of deep learning model architectures for agricultural tasks, and the exploration of advanced data augmentation strategies to improve model generalization. Addressing the challenges posed by complex visual data variations, such as overlapping plants and varying environmental conditions, remains a critical focus. Moreover, the need for large, accurately annotated datasets to effectively train and evaluate these sophisticated models is a persistent challenge that researchers are actively working to overcome.

In summary, the literature reveals a significant shift from traditional machine learning approaches relying on handcrafted features towards deep learning methodologies that automatically learn relevant representations. While substantial progress has been made, ongoing research continues to explore novel techniques and address the remaining challenges to develop highly accurate and practical weed detection systems for real-world agricultural applications.

## III. METHODOLOGY

This study explored both traditional machine learning (ML) and modern deep learning (DL) techniques for the task of weed detection and classification in agricultural imagery. The methodology involved several key phases: dataset acquisition and preparation, image preprocessing to enhance data quality and reduce complexity, extraction of relevant features using both manual and automated approaches, training and evaluation of various classification models, and the application of a state-of-the-art object detection algorithm for weed localization.

*1. Datasets*

Two publicly available datasets were utilized: Early-Crop-Weed [6] and CottonWeedID15 [17]. The Early-Crop-Weed dataset focused on early-stage weeds (velvetleaf, black nightshade) and crops (cotton, tomato), while CottonWeedID15 provided a more diverse set of fifteen weed species common in cotton fields. Both datasets exhibited class imbalance, as visualized in Figures 3 and 4, which necessitated the use of balancing techniques.

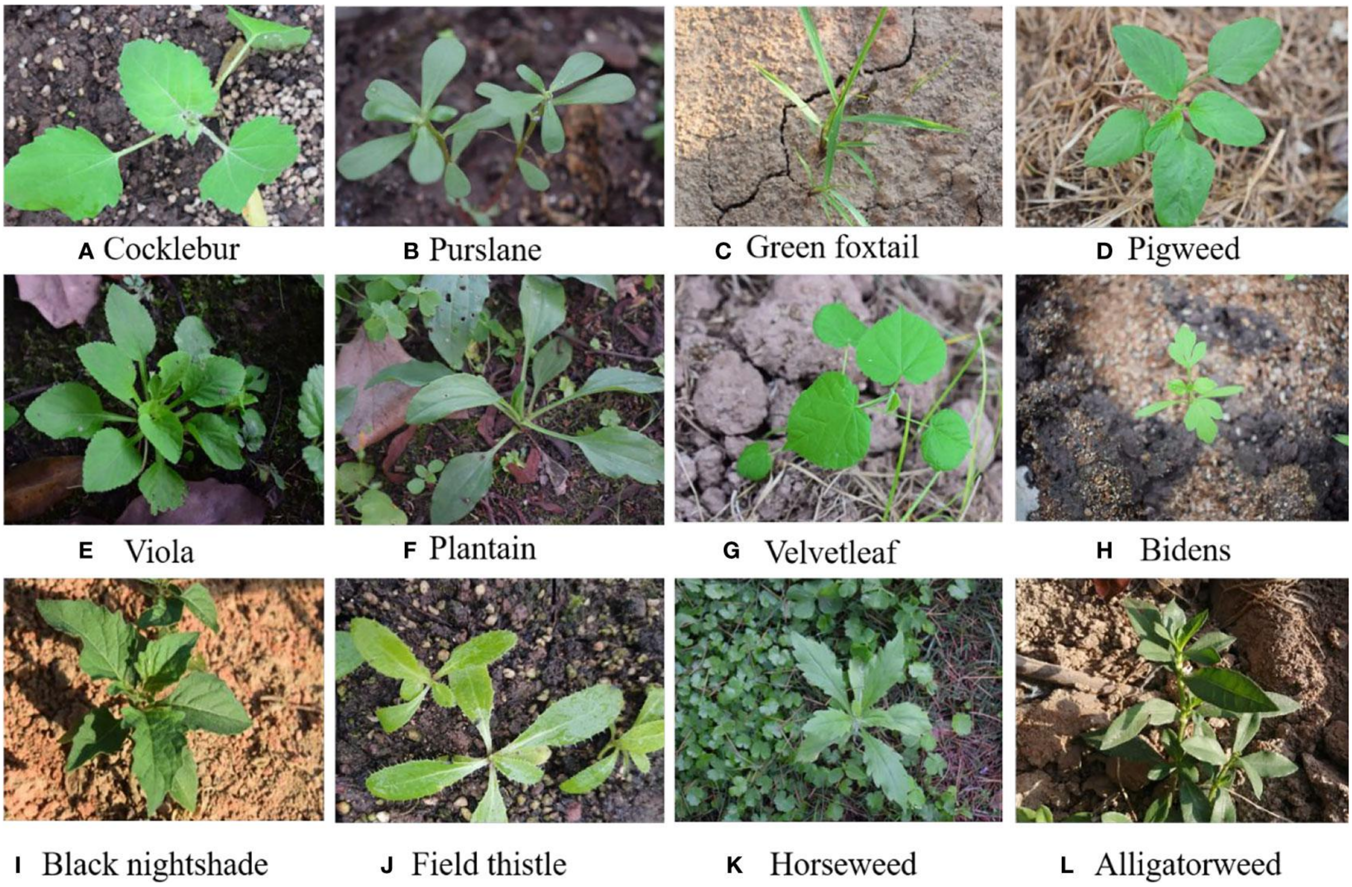


**Fig 1:** Collection of weed datasets

*2. Preprocessing*

Several preprocessing steps were applied to the images to improve model performance and reduce computational cost.

- **Annotation:** For YOLOv8 training, weed leaves in both datasets were annotated with bounding boxes using LabelImg, minimizing soil inclusion to focus on the target (Figure 5).
- **Grayscale and Resizing:** Color images were converted to grayscale and resized to 224x224 pixels to reduce computational load and standardize input size.
- **Background Removal:** The U2-Net deep learning model was employed to remove background noise and focus on the plant regions using generated saliency maps (Figure 6).
- **SMOTE:** The Synthetic Minority Over-sampling Technique (SMOTE) was used to address the class imbalance present in both datasets by generating synthetic minority class samples.

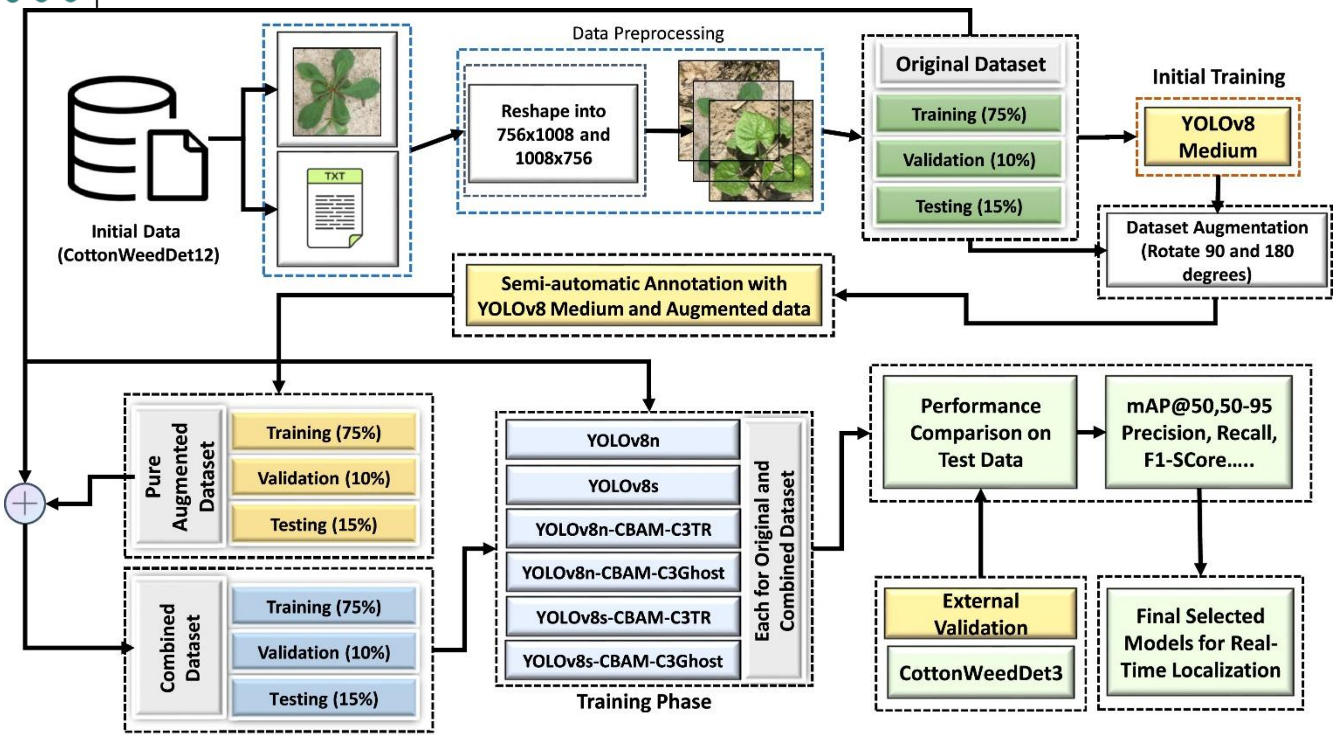


**Fig 2:** Smart Agriculture Model architecture

*3. Manual Features*

Following preprocessing, manual features were extracted, including texture features using the Gray-Level Co-occurrence Matrix (GLCM), statistical features (mean, standard deviation, etc.), and shape-based Hu moments. These features were normalized and used to train traditional ML classifiers: Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks (ANN). SMOTE was applied to the training data before feature extraction to handle class imbalance.

*4. Deep Learning Features*

Deep learning features were automatically extracted using several pre-trained Convolutional Neural Network (CNN) architectures (e.g., VGG16, ConvNeXt) initialized with weights from the ImageNet dataset, employing transfer learning. SMOTE was applied to the training data prior to feature extraction to ensure balanced class representation (Figures 7, 8). The extracted features were then used to train a Random Forest classifier for evaluation.

*5. Experimental Setup and Results*

The datasets were split into training (70%), validation (15%), and testing (15%) sets. For manual features, ANN achieved 89.26% accuracy on CottonWeedID15, and SVM (polynomial kernel) reached 99% on Early-Crop-Weed. Deep learning features with Random Forest yielded 98% accuracy on Early-Crop-Weed and 89% on CottonWeedID15. The YOLOv8-M object detection model achieved a mean average precision (mAP) of 89% after 100 training epochs. The detailed results are presented in tables and figures as referenced in the original text.

## IV. RESULTS AND DISCUSSIONS

Manual Features and Classifiers: After extracting manual features, the datasets were split into training (65%), validation (20%), and testing (15%) sets. The ANN achieved 89.26% testing accuracy on the CottonWeedID15 dataset, while SVM with a polynomial kernel achieved 99% accuracy on the Early-Crop-Weed dataset. Deep Learning Features and Classifiers: Deep learning features were extracted using pre-trained CNN architectures. A Random Forest classifier was used. Features from ConvNeXt yielded the best performance, achieving 98% accuracy on the Early-Crop-Weed dataset and 89% on the CottonWeedID15 dataset. YOLOv8 Object Detection: YOLOv8-M was implemented. The model achieved a mean average precision of 89%. Discussion: Deep Learning vs. Manual Features: Deep learning features, particularly ConvNeXt, achieved higher accuracy. SMOTE Effectiveness: SMOTE effectively addressed class imbalance. YOLOv8 Performance: YOLOv8 demonstrated strong performance in agricultural applications. Related Advancements: IoT devices and optimization algorithms show promise. Agricultural Impact: Deep learning technologies have the potential to improve agricultural productivity. Future Directions: Future research should focus on improving model robustness and exploring multi-sensor data.

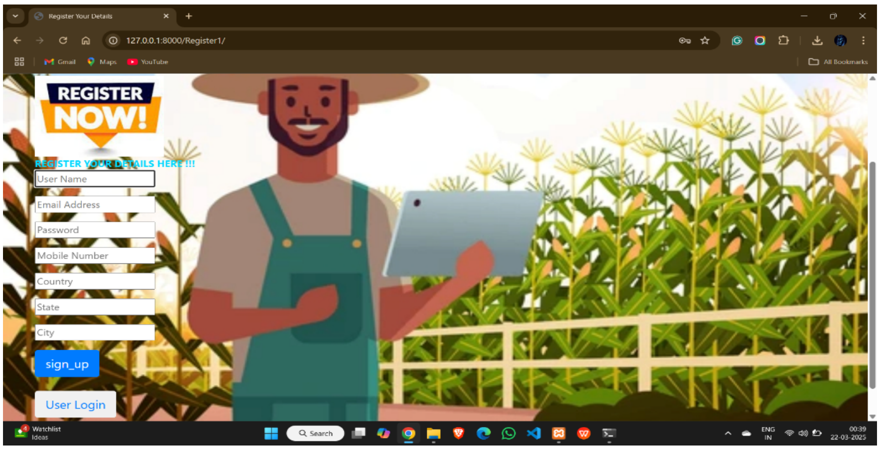


**Fig 3:** Registration page

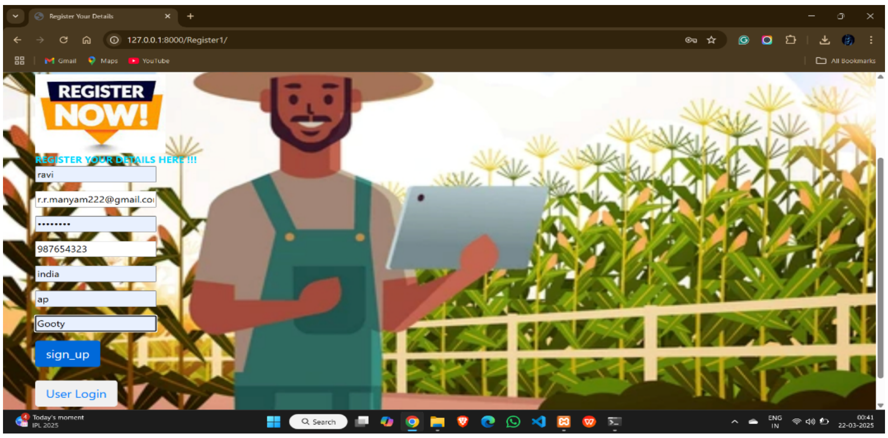


**Fig 4:** Registration details

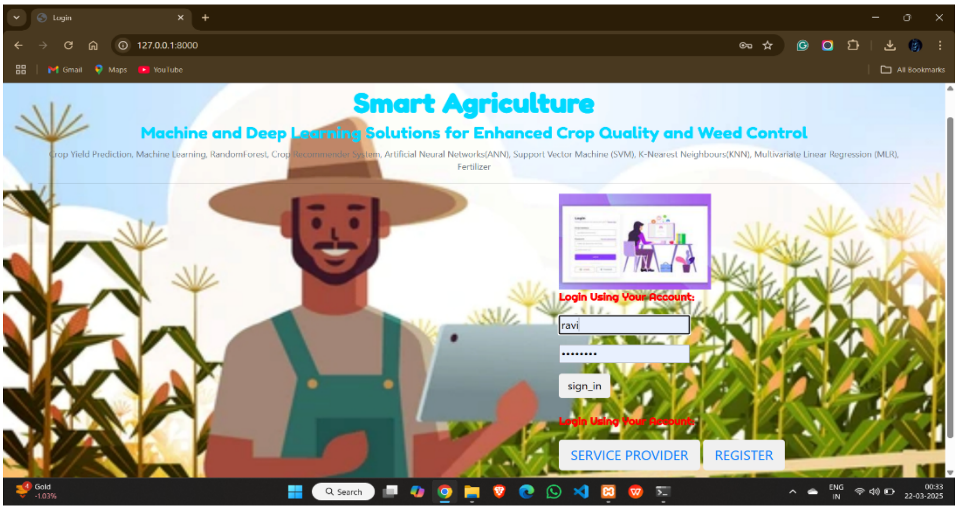


**Fig 5:** Login page

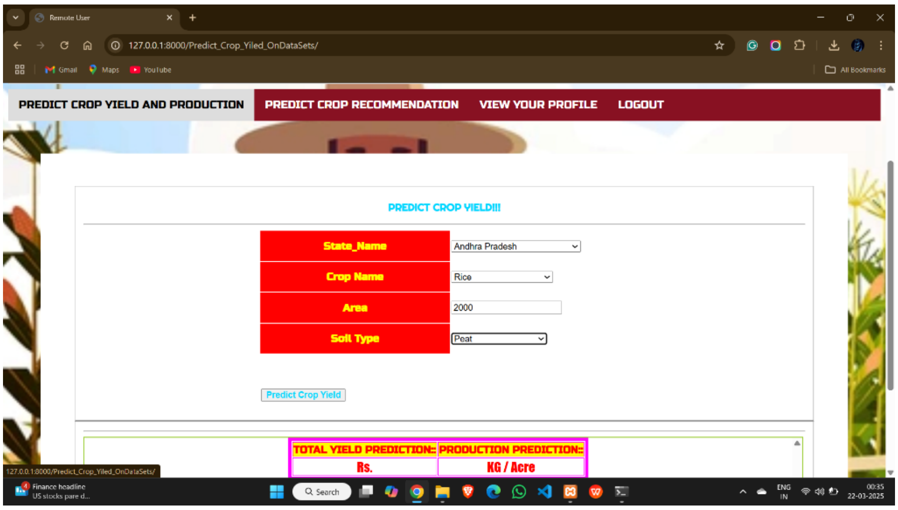


**Fig 6:** Precict Crop Yield and production page

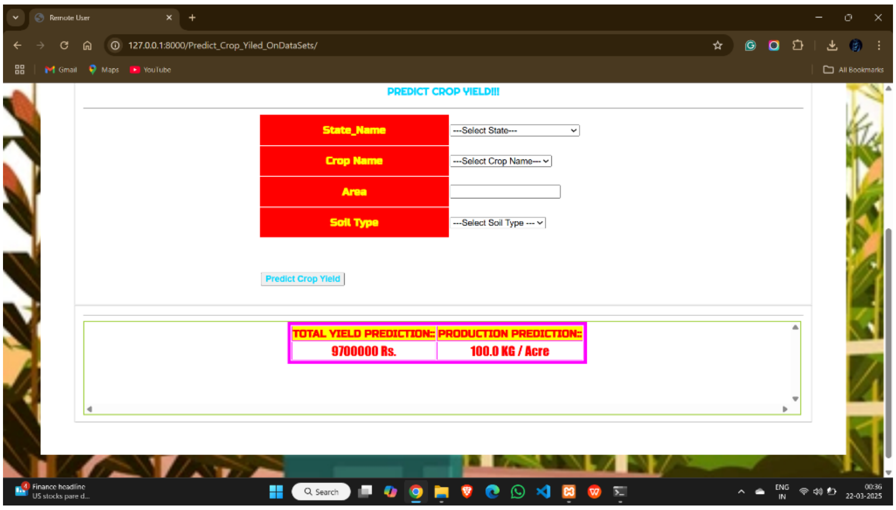


**Fig 7:** Precict Crop Yield and production result

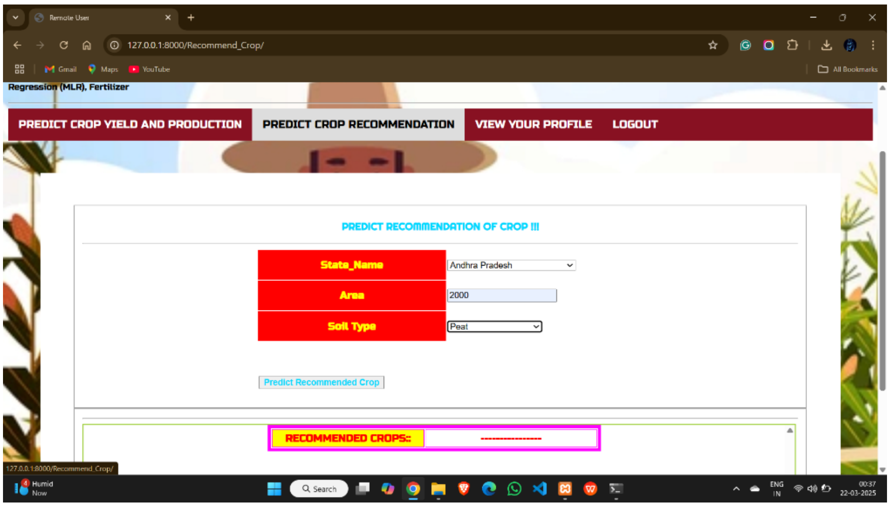


**Fig 8:** Crop Recommendation page

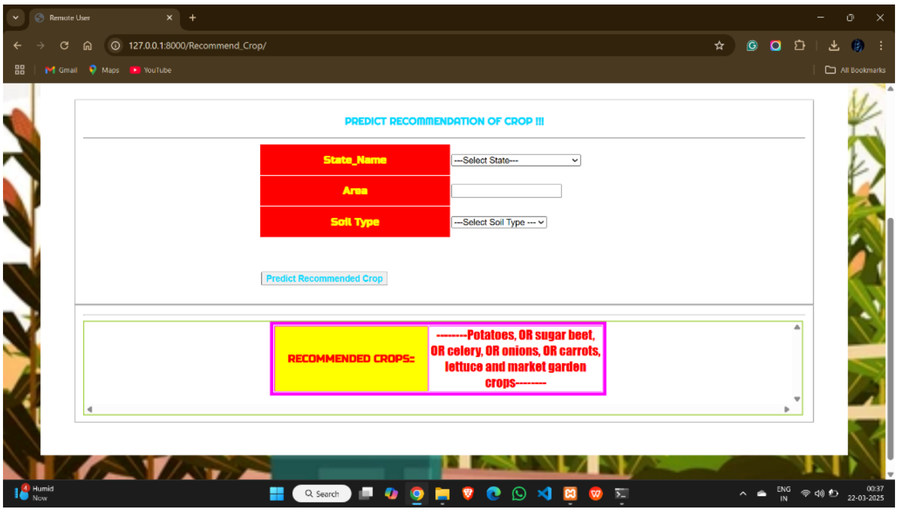


**Fig 9:** Crop Recommendation result

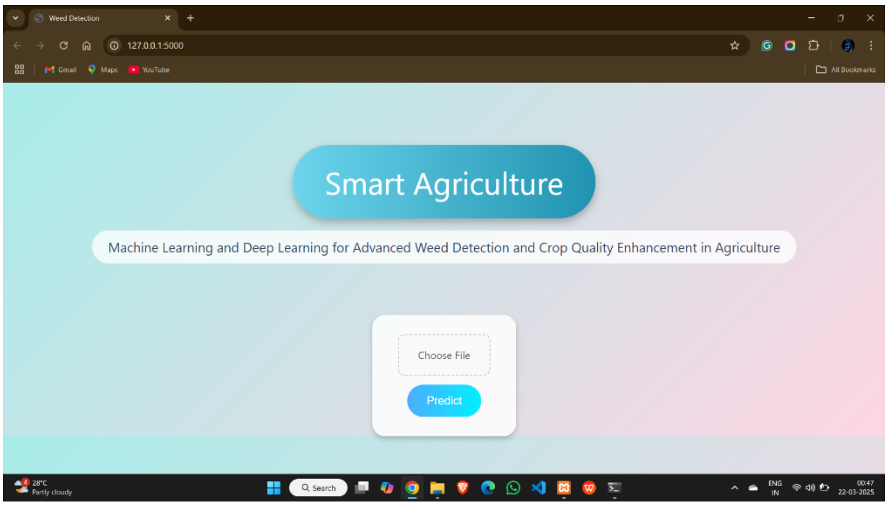


**Fig 10:** Weed classification page

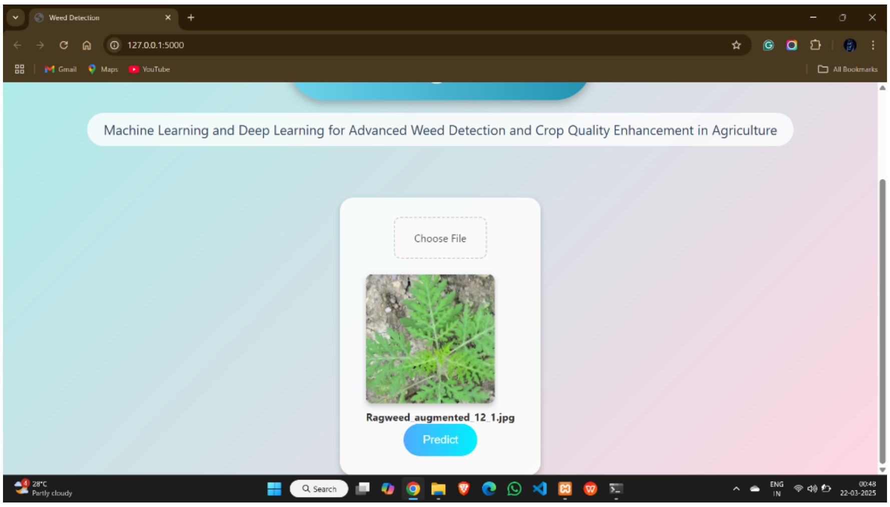


**Fig 11:** Weed image selection

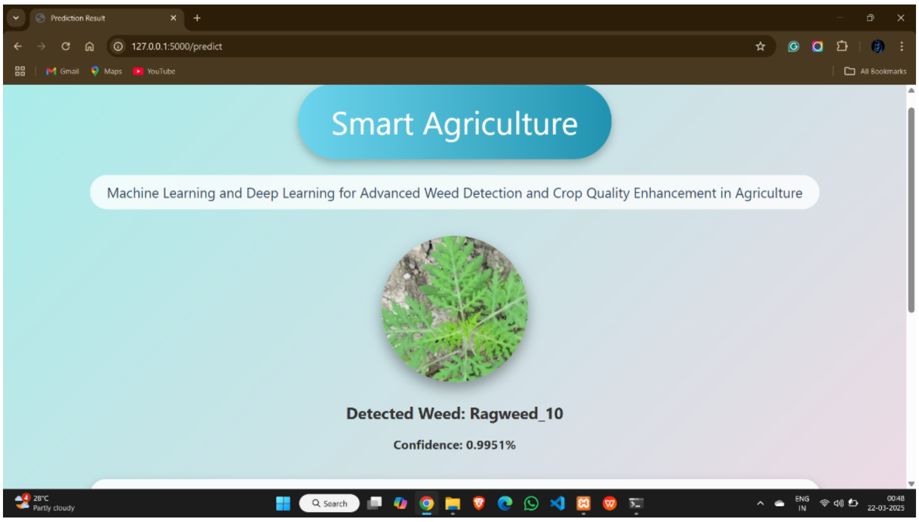


**Fig 12:** Detected weed result

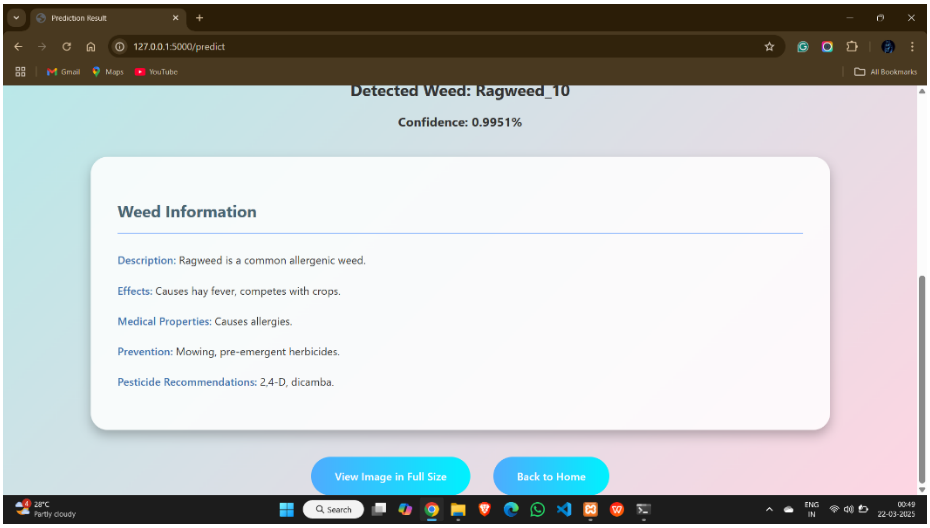


**Fig 13:** Weed information page

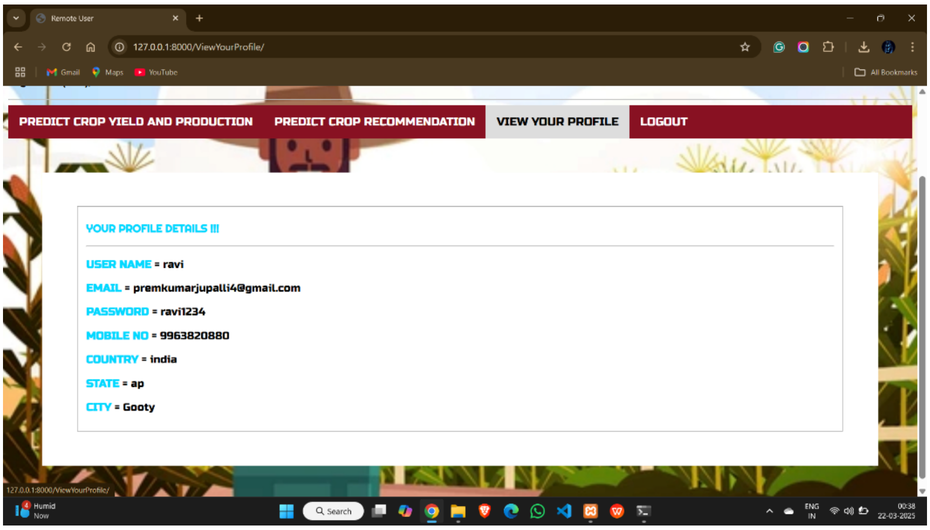


**Fig 14:** User profile page

## V. CONCLUSION AND FUTURE WORK

Automated weed control systems, powered by advanced technologies, are crucial for addressing agricultural weed challenges, improving crop production, and reducing costs. Future research should prioritize building larger, more diverse image datasets to enhance real-time performance in agricultural settings. Deep Convolutional Generative Adversarial Networks (DCGANs) can be utilized for effective crop and weed data augmentation. While deep learning has shown promising results, continuous improvement is necessary. A significant barrier to adopting modern agricultural technologies, particularly deep learning, is the lack of awareness in agricultural countries. Educating farmers about these technologies is essential to improve agricultural practices and crop yields. This study advocates for the integration of advanced technologies, especially deep learning, to transform traditional agriculture. By promoting the adoption of state-of-the-art techniques for sustainable and efficient weed management, this research aims to positively impact the agricultural community and industry. Future research should focus on refining the model's robustness to environmental factors and expanding the dataset to encompass a broader range of agricultural scenarios. Investigating transfer learning techniques and exploring the use of multi-sensor data for more accurate weed identification are potential avenues for improvement. Current research predominantly relies on RGB images for weed detection. Exploring the integration of multispectral data, such as infrared or hyperspectral imagery, could provide additional insights into weed characteristics and improve the model's accuracy, particularly in scenarios where visual cues alone may be insufficient.

## REFERENCES

1. Radoglou-Grammatikis, P., Sarigiannidis, P., Lagkas, T., & Moscholios, I. (2020). A compilation of UAV applications for precision agriculture. *Computer Networks, 168*, 107037.
2. Ashraf, M., & Khan, S. (2020). Weed density classification in rice crop using computer vision. *Computers and Electronics in Agriculture, 179*, 105776.
3. Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In *European Conference on Computer Vision* (pp. 740–755). Springer.
4. Seelan, S. K., Laguette, S., Casady, G. M., & Seielstad, G. A. (2003). Remote sensing applications for precision agriculture. *Remote Sensing of Environment, 88*(1–2), 157–169.

5. Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., & Vellidis, G. (2020). Towards weeds identification assistance through transfer learning. *Computers and Electronics in Agriculture, 174*, 105451.
6. Giselsson, T. M., Jørgensen, R. N., Dyrmann, M., & Midtiby, H. S. (2017). A public image database for benchmark of plant seedling classification algorithms. *arXiv preprint* arXiv:1711.05458.
7. Mujtaba. (2023). Annotated Weed Images. *Kaggle*. https://www.kaggle.com/datasets/mujtabad/annotated-weed-images
8. Espejo-Garcia, B., Mylonas, N., Fountas, S., & Vellidis, G. (2021). Combining generative adversarial networks and agricultural transfer learning for weeds identification. *Biosystems Engineering, 204*, 57–66.
9. Wang, A., Zhang, Y., & Zhang, G. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture, 158*, 226–240.
10. Yu, J., Wu, Y., Chen, P., Liu, Z., & Shen, L. (2019). Deep learning for image-based weed detection in turfgrass. *European Journal of Agronomy, 104*, 78–84.
11. Hasan, M., Mouazen, A. M., & Hossain, M. A. (2021). A survey of deep learning techniques for weed detection from images. *Computers and Electronics in Agriculture, 184*, 106067.
12. Bakhshipour, A., & Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Computers and Electronics in Agriculture, 145*, 153–160.
13. Kaya, A., Fidan, M., & Deveci, M. (2019). Analysis of transfer learning for deep neural network based plant classification models. *Computers and Electronics in Agriculture, 158*, 20–29.
14. Zhuang, X., Wu, J., Du, Z., Zhang, J., & Wang, Z. (2022). Evaluation of different deep convolutional neural networks for detection of broadleaf weed seedlings in wheat. *Pest Management Science, 78*(2), 553–561.
15. Ahmad, T., Iqbal, M., & Irfan, M. (2021). Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. *Computers and Electronics in Agriculture, 187*, 106276.
16. Chen, Z., Li, Y., Xu, M., Wang, T., & Yang, J. (2022). Performance evaluation of deep transfer learning on multi-class identification of common weed species in cotton production systems. *Computers and Electronics in Agriculture, 198*, 107056.
17. Wang, X., Wang, J., & Yang, G. (2022). A deep learning approach incorporating YOLO v5 and attention mechanisms for field real-time detection of the invasive weed *Solanum rostratum Dunal* seedlings. *Computers and Electronics in Agriculture, 198*, 107001.
18. Subeesh, A., John, J. J., & Ramesh, M. V. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. *Artificial Intelligence in Agriculture, 6*, 42–51.
19. Razfar, M. R., Valipour, M., & Karimi, M. (2022). Weed detection in soybean crops using custom lightweight deep learning models. *Journal of Agriculture and Food Research, 10*, 100424.
20. Yang, L., Zhang, Q., & Li, M. (2022). A comparative evaluation of convolutional neural networks, training image sizes, and deep learning optimizers for weed detection in alfalfa. *Weed Technology, 36*(1), 91–100.
21. Jin, T., Luo, J., Chen, Q., & Yu, Y. (2022). Deep learning for detecting herbicide weed control spectrum in turfgrass. *Plant Methods, 18*(1), 112.
22. Zhang, M., Li, S., Wang, J., & Gao, Y. (2022). SE-YOLOv5x: An optimized model based on transfer learning and visual attention mechanism for identifying and localizing weeds and vegetables. *Agronomy, 12*(3), 635.
23. Reedha, S., Murugesan, K., & Vimal, S. (2022). Transformer neural network for weed and crop classification of high-resolution UAV images. *Remote Sensing, 14*(24), 6241.
24. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (Vol. 27).
25. Lu, H., Huang, Y., Zhang, J., & Zhao, Y. (2022). Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. *Computers and Electronics in Agriculture, 196*, 106898.
26. Madapuri, R. K., & Senthil Mahesh, P. C. (2017). HBS-CRA: Scaling impact of change request towards fault proneness: Defining a heuristic and biases scale (HBS) of change request artifacts (CRA). *Cluster Computing, 22*(S5), 11591–11599. https://doi.org/10.1007/s10586-017-1424-0
27. Dwaram, J. R., & Madapuri, R. K. (2022). Crop yield forecasting by long short-term memory network with Adam optimizer and Huber loss function in Andhra Pradesh, India. *Concurrency and Computation: Practice and Experience, 34*(27), e7310. https://doi.org/10.1002/cpe.7310
28. Reddy, B. S. H. (2025). Deep learning-based detection of hair and scalp diseases using CNN and image processing. *Milestone Transactions on Medical Technometrics, 3*(1), 145–155. https://doi.org/10.5281/zenodo.14965660

29. Reddy, B. S. H., Venkatramana, R., & Jayasree, L. (2025). Enhancing apple fruit quality detection with augmented YOLOv3 deep learning algorithm. *International Journal of Human Computations & Intelligence, 4*(1), 386–396. https://doi.org/10.5281/zenodo.14998944
30. Ahmed, S. T., Kaladevi, A. C., Shankar, A., & Alqahtani, F. (2025). Privacy Enhanced Edge-AI Healthcare Devices Authentication: A Federated Learning Approach. *IEEE Transactions on Consumer Electronics*.
31. Singh, K. D., & Ahmed, S. T. (2020, July). Systematic linear word string recognition and evaluation technique. In *2020 international conference on communication and signal processing (ICCSP)* (pp. 0545-0548). IEEE.
32. Syed Thouheed Ahmed, S., Sandhya, M., & Shankar, S. (2018, August). ICT's role in building and understanding indian telemedicine environment: A study. In *Information and Communication Technology for Competitive Strategies: Proceedings of Third International Conference on ICTCS 2017* (pp. 391-397). Singapore: Springer Singapore.
33. Sreedhar Kumar, S., Ahmed, S. T., & NishaBhai, V. B. (2019). Type of supervised text classification system for unstructured text comments using probability theory technique. *International Journal of Recent Technology and Engineering (IJRTE)*, *8*(10).
34. Ahmed, S. T., Basha, S. M., Arumugam, S. R., & Kodabagi, M. M. (2021). *Pattern Recognition: An Introduction*. MileStone Research Publications.