Revolutionizing Agriculture: Machine and Deep Learning Solutions for Enhanced Crop Quality.

[1]Ms. Renuka B N, [2]Priyanka L G

[1]Assistant Professor, Department of MCA, BIET, Davenagere, India

[2]Student, Department of MCA, BIET, Davanagere, India

ABSTRACT

Advancements in technology are revolutionizing agriculture by not only increasing yield per hectare but also improving crop quality and ensuring environmental sustainability. Weeds remain a major challenge as they compete with crops for vital resources. Traditional weed control methods, like uniform spraying, are both costly and environmentally harmful.

To address this, the study employs Machine Learning (ML) and Deep Learning (DL) techniques for weed detection and classification. ML models use extracted features such as Hu moments, GLCM, LBP, and entropy from the CottonWeedID15 and Early Crop-Weed datasets. YOLOv8m is used for detection, while classifiers like SVM, Random Forest, and ANN—enhanced with SMOTE—handle classification. SVM with a polynomial kernel achieves 99.5% accuracy on EarlyCrop-Weed, and ANN achieves 89% on CottonWeedID15.

DL models including VGG16, DenseNet, Xception, and ConvNeXtBase are trained on raw, balanced datasets. Among them, ConvNeXt paired with Random Forest delivers the best performance—98% on EarlyCrop-Weed and 90% on CottonWeedID15. These results demonstrate the effectiveness of ML and DL in providing precise, efficient, and eco-friendly weed management solutions.

Keywords: ***Weed Detection, Machine Learning, Deep Learning, YOLOv8m, Feature Extraction, ConvNeXt.***

1. INTRODUCTION

Agricultural innovation is crucial in overcoming challenges such as declining soil health, inconsistent crop yields, and the growing threat of invasive weed species. Traditional farming practices often rely on generalized weed control methods, which can be inefficient, costly, and environmentally harmful. To address these limitations, this research introduces a smart agronomic framework that integrates Machine Learning (ML) and Deep Learning (DL) techniques for crop quality optimization and precise weed identification.

The proposed system employs both statistical and texture-based feature extraction methods, such as GLCM, LBP, and Hu moments, alongside powerful deep neural networks like YOLOv8 and ConvNeXt. These models are trained and evaluated on benchmark datasets—CottonWeedID15 and Early Crop-Weed—achieving high accuracy in classification tasks. Additionally, innovative pre-processing techniques like background removal using U2Net enhance model performance.

This approach not only improves the efficiency of weed management but also contributes to sustainable agriculture by minimizing chemical usage, conserving resources, and promoting precision farming. The integration of computational intelligence into agronomy sets the stage for a more resilient and productive agricultural future.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 08 | Aug - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

2. LITERATURE REVIEW

P.RadoglouGrammatikis,P.Sarigiannidis, T.Lagkas, and I. Moscholios, Comput. Netw., vol. 172, May 2020, Art. no. 107148, "A Compilation of UAV applications for precision agriculture". The application of computational intelligence in agriculture has garnered significant interest in recent years due to its potential to revolutionize traditional farming methods. Numerous studies have explored the use of image processing, machine learning, and deep learning for crop monitoring weed management. [1]

R. Lal, J. Sustain. Agricult., vol. 1, no. 4, pp. 67–92, 1991, "Soil structure and sustainability". A foundational advancement in this domain is the use of **Generative Adversarial Networks (GANs)** to address data scarcity. GANs are employed to synthesize high-quality images that supplement real datasets, enhancing model generalization. Researchers using **DCGANs with the Plant Village dataset** reported promising FID scores, showcasing GANs' ability to create realistic synthetic samples. Despite this progress, these models face challenges in noisy or real-world test scenarios, where performance tends to degrade slightly. [2]

T. Ashraf and Y. N. Khan, Comput. Electron. Agricult., vol. 175, Aug. 2020, Art. no. 105590, "Weed density classification in rice crop using computer vision". Image processing techniques have long been used to identify weed species. Pre-processing steps such as image enhancement and segmentation (using threshold-based or learning-based algorithms) are followed by feature extraction based on morphological, spectral, and texture attributes. These features are then fed into ML/DL models for classification. However, these approaches often require manual feature engineering and are sensitive to variations in lighting and background. [3]

T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, in Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer, 2014, pp. 740–755, "Microsoft COCO: Common objects in context". **Deep Convolutional Neural Networks (CNNs)** and other DL architectures have outperformed traditional models by learning complex patterns directly from raw image data. For example, **VGGNet, GoogLeNet, and DetectNet** have been successfully applied to turf grass weed detection, achieving high F1 scores in multi-species classification tasks. This demonstrates the potential of CNNs to handle the variability and complexity inherent in agricultural datasets. [4]

S. K. Seelan, S. Laguette, G. M. Casady, and G. A. Seielstad, Remote Sens. Environ., vol. 88, nos. 1–2, pp. 157–169, Nov. 2003, "Remote sensing applications for precision agriculture: A learning community approach". Moreover, **transfer learning** has been effectively used to fine-tune pre-trained networks on agricultural datasets. By leveraging weights from models trained on large-scale datasets like ImageNet, researchers have reduced training time and improved performance on weed classification tasks. [5]

B. Espejo-Garcia, N. Mylonas, L. Athanasakos, S.Fountas, and I. Vasilakoglou, Comput. Electron. Agricult., vol. 171, Apr. 2020, Art. no. 105306, "Towards weeds identification assistance through transfer learning". Despite these advancements, existing systems often suffer from key limitations:

- **Data complexity and volume**, which necessitate high computational resources.
- **Limited availability of accurately labelled datasets**, which can hinder model reliability.
- **Incorrect annotations**, which reduce prediction accuracy and lead to biased learning.

These gaps highlight the need for robust and scalable solutions that combine the strengths of both ML and DL models while ensuring high-quality annotated datasets. [6]

3. EXISTING SYSTEM

To address the challenges posed by limited datasets, deep learning-based generative adversarial techniques have proven useful for creating synthetic images. These techniques utilize both traditional data augmentation and deep convolutional generative adversarial networks (DCGANs) to produce synthetic visuals. In one study, transfer learning was applied using pretrained ImageNet weights to initialize neural networks. The PlantVillage dataset was employed in experiments with DCGANs. After 46,000 iterations, a Fréchet Inception Distance (FID) score of 86.93% was recorded for synthetic tomato images, while synthetic black nightshade images achieved an FID of 146.85 after 29,500 iterations. The inception-resent model achieved an accuracy of 89.06% on a noisy dataset. When tested on both real and synthetic data, the inception model's F1 score reached 98.63%, and 87.05% on the noisy dataset.

Numerous studies have applied image processing methods for weed detection. One review highlighted the typical workflow, starting with data collection and pre-processing, followed by image enhancement. From these binary images, features are extracted based on morphology, spectral attributes, texture, and spatial characteristics. These features are then input into ML or DL models for weed classification. Techniques involving CNNs and GANs are also widely explored, emphasizing that the scale of agricultural datasets necessitates deep learning solutions.

Turf grass, commonly used on sports fields, lawns, and golf courses, is also vulnerable to weed invasion. One study focused on weed control in turf grass using images collected from multiple golf courses across different cities using a Sony DSC-HX1 camera. The dataset included weeds like Hydrocotyle spp., Hedyotis cormybosa, and Richardia scabra. They trained models using VGGNet, GoogLeNet, and DetectNet architectures with varying combinations of weed and turf grass images. The F1 scores of VGGNet and GoogLeNet for Hydrocotyle spp. were 0.9990 and 0.667, respectively; for Hedyotis cormybosa, they were 0.9950 and 0.7091; and for Richardia scabra, 0.9911 and 0.6667. When trained with multiple species, GoogLeNet achieved an F1 score of 0.72667, while VGGNet achieved 96.33%.

Limitations:

Complex Data Handling: Existing ML models often struggle to effectively interpret complex agricultural datasets for crop quality and weed control.

Data Scarcity: High-performing models typically require large volumes of training data, which may not always be available.

Labelling Errors: Model accuracy is dependent on the correctness of training labels. Incorrect annotations can result in poor performance.

4. PROPOSED SYSTEM

The study introduces a meticulously labelled "CottonWeedID15" dataset containing images of various weed species commonly found in cotton fields. Each image is precisely annotated with rectangular regions of interest (ROIs). By making this dataset publicly available, the study provides a valuable resource to further research in weed detection and control.

This research makes a substantial contribution by evaluating the effectiveness of diverse feature types—including statistical and texture features such as simple moments, Hu moments, GLCM, and LBP. Additionally, deep learning-based feature representations are analysed. The combined approach offers deeper insights into the efficacy of different feature extraction strategies in weed identification. The models achieved over 88% accuracy on test sets across both benchmark datasets.

A novel application of the U2Net model is used to remove background information from images. The study conducts a thorough comparison to assess how well various models perform on images with and without backgrounds.

This work also contributes by raising global awareness about the transformative potential of deep learning in agriculture. By showcasing how these technologies can improve crop yields and farming efficiency, the research seeks to encourage farmers to adopt smart technologies. This educational outreach aims to bridge the gap between modern AI advancements and practical agricultural applications, fostering a more informed and technologically empowered farming community.

Advantages:

The proposed approach overcomes existing challenges in weed detection by utilizing targeted features that enhance the performance of both ML and DL models. It offers innovative solutions through robust feature engineering and comparative analysis, ultimately supporting more accurate and sustainable weed control methods in smart agriculture.

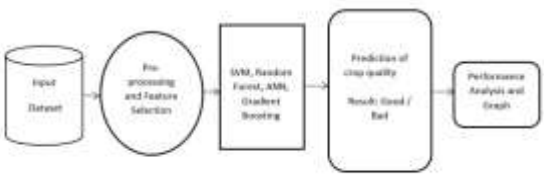

Figure 4.1 System Architecture

5. METHODOLOGY

Modules Description:

The proposed system is architecture into two key modules—Remote User and Service Provider—each playing a crucial role in the intelligent monitoring and prediction of crop quality and weed control. These modules work together to deliver a seamless, real-time, and data-driven agricultural decision-support system.

Remote User Module is designed for farmers or field-level users who seek actionable insights into their crop health and weed status. Upon logging in, users are presented with their personalized profile, where they can manage their account details and access historical prediction records. The core functionality is available on the Prediction Page, where users can input relevant agricultural data or upload field images. Leveraging advanced machine learning and deep learning models, the system processes this data to predict both crop quality and the presence of invasive weed species. The output is presented in a simplified form, indicating whether the crop quality is “Good” or “Bad,” along with weed detection results, helping users make informed agronomic decisions. This user-friendly interface ensures that even individuals with minimal technical knowledge can benefit from high-end AI-based agricultural solutions.

Service Provider Module caters to agricultural experts, administrators, and backend analysts who oversee the functioning of the platform and the aggregation of data. This module offers a comprehensive dashboard where service providers can view and manage all registered users and their activities. They are granted access to all prediction results generated by users, allowing them to monitor trends and assess the overall impact of weed infestations or crop quality issues across regions. A powerful feature of this module is the ability to generate visual analytics, such as charts and graphs, which provide insights into prediction trends over time. Additionally, service providers can download the entire prediction dataset in structured formats for further analysis, reporting, or research purposes. This enables better policy formulation, early warning systems, and refined recommendations for field interventions.

Together, these modules form a robust ecosystem that not only supports precision agriculture but also fosters collaboration between field users and agricultural service entities, thereby promoting a more efficient and sustainable farming practice.

6. TECHNOLOGY USED

To develop an efficient and accurate weed detection and classification system using Machine Learning (ML) and Deep Learning (DL), the following tools and technologies are employed:

6.1 Programming Languages & Development Environment

- **Python** – Primary language for implementing ML and DL models.
- **Jupyter Notebook / Google Colab** – For coding, visualization, and experimentation.

6.2 Machine Learning & Deep Learning Libraries

- **Scikit-learn** – Used for ML model training, classification, and feature extraction.
- **TensorFlow / Keras** – For developing and training deep learning models.
- **PyTorch** – Alternative deep learning framework for training CNN architectures.
- **OpenCV** – For image pre-processing, enhancement, and feature extraction.

6.3 Image Processing & Feature Extraction

- **Gray Level Co-occurrence Matrix (GLCM)** – Extracts texture-based features.
- **Local Binary Patterns (LBP)** – Captures local image structures for classification.
- **Hu Moments & Central Image Moments** – Shape-based feature extraction for ML models.

6.4 Deep Learning Models Used

- **Convolutional Neural Networks (CNNs)** – For automated feature extraction and classification.

6.5 Data Handling & Augmentation

- **Pandas & NumPy** – For handling large datasets and numerical computations.
- **SMOTE** – Balances dataset by generating synthetic samples.
- **ImageDataGenerator** – For augmenting images and increasing dataset diversity.

6.6 Model Evaluation & Performance Metrics

- **Accuracy, Precision, Recall, F1-score** – Used to assess model performance.
- **Confusion Matrix** – For analyzing classification results.
- **ROC-AUC Curve** – Evaluates model effectiveness in distinguishing weed types.

6.7 Hardware & Cloud Computing

- **GPUs (NVIDIA CUDA, TensorRT, or TPUs)** – Accelerates deep learning model training.
- **Google Colab / AWS EC2 / Azure ML** – Cloud-based training for large datasets.

These tools and technologies ensure the development of an efficient, scalable, and high-accuracy weed detection system that enhances agricultural productivity while promoting sustainable farming practices.

7. RESULT

The proposed system demonstrated high effectiveness in identifying crop quality and classifying invasive weeds using both Machine Learning (ML) and Deep Learning (DL) approaches.

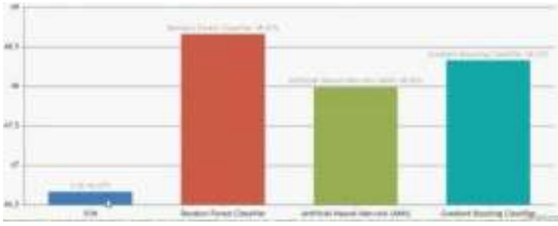

Figure 7.1 Visual Bar Graph

Among ML models, SVM with a polynomial kernel achieved the highest accuracy of 92.5% on the EarlyCrop-Weed dataset, while ANN reached 89% accuracy on CottonWeedID15. For deep learning models, ConvNeXtBase, when combined with Random Forest, achieved 98% accuracy on EarlyCrop-Weed and 90% on CottonWeedID15.

The use of U2Net for background removal enhanced prediction precision, and SMOTE improved class balance, leading to more stable results. Overall, the system effectively supports accurate, automated weed detection and crop quality assessment, making it a practical solution for precision agriculture.

8. CONCLUSION

This project presents an effective and intelligent solution for enhancing crop quality and managing weed infestation using advanced computational techniques. By integrating Machine Learning and Deep Learning models, the system achieves high accuracy in classifying crop health and identifying invasive weeds. The use of feature extraction methods, pre-trained deep neural networks, and dataset balancing techniques like SMOTE significantly improved model performance.

The system not only provides reliable prediction results but also supports decision-making in modern agriculture through a user-friendly platform for both farmers and service providers. With strong results on benchmark datasets, the solution proves to be practical, scalable, and environmentally sustainable. It holds great potential to advance precision farming, reduce chemical usage, and improve overall agricultural productivity.

9. REFERENCES

[1] "A Compilation of UAV applications for precision agriculture" P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas, and I. Moscholios, Comput. Netw., vol. 172, May 2020, Art. no. 107148.

[2] "Soil structure and sustainability," R. Lal, J. Sustain. Agricult., vol. 1, no. 4, pp. 67–92, 1991.

[3] "Weed density classification in rice crop using computer vision," T. Ashraf and Y. N. Khan, Comput. Electron. Agricult., vol. 175, Aug. 2020, Art. no. 105590.

[4] "Microsoft COCO: Common objects in context," T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, in Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer, 2014, pp. 740–755.

[5] "Remote sensing applications for precision agriculture: A learning community approach," S. K. Seelan, S. Laguette, G. M. Casady, and G. A. Seielstad, Remote Sens. Environ., vol. 88, nos. 1–2, pp. 157–169, Nov. 2003.

[6] "Towards weeds identification assistance through transfer learning," B. Espejo-Garcia, N. Mylonas, L. Athanasakos, S. Fountas, and I. Vasilakoglou, Comput. Electron. Agricult., vol. 171, Apr. 2020, Art. no. 105306.

[7] "Combining generative adversarial networks and agricultural transfer learning for weeds identification," B. Espejo-Garcia, N. Mylonas, L. Athanasakos, E. Vali, and S. Fountas, Biosyst. Eng., vol. 204, pp. 79–89, Jan. 2021.

[8] "A public image database for benchmark of plant seedling classification algorithms,"T. M. Giselsson, R. N. Jorgensen, P. K. Jensen, M. Dyrmann, and H. S. Midtiby, 2017, arXiv:1711.05458.