Data and Metadata. 2025; 4:848
doi: 10.56294/dm2025848

ORIGINAL

Automated Weed Detection in Crop Fields Using Convolutional Neural Networks: A Deep Learning Approach for Smart Farming

Detección automatizada de malas hierbas en campos de cultivo mediante redes neuronales convolucionales: un enfoque de aprendizaje profundo para la agricultura inteligente

Nidhya R[1], Pavithra D[2], Smilarubavathy G[3], D Mythrayee[4]

[1]Computer Science and Engineering, Madanapalle Institute of Technology & Science. Madanapalle, Andhra Pradesh, India.
[2]Information Technology, Dr. N. G. P. Institute of Technology. Coimbatore, Tamil Nadu, India.
[3]Computer Science and Engineering, St.Joseph's Institute of Technology. Chennai, India.
[4]Computer Science and Engineering, Koneru Lakshmaiah Education Foundation. Vaddeswaram, Andhra Pradesh, India.

Cite as: R N, D P, G S, Mythrayee D. Automated Weed Detection in Crop Fields Using Convolutional Neural Networks: A Deep Learning Approach for Smart Farming. Data and Metadata. 2025; 4:848. https://doi.org/10.56294/dm2025848

Submitted: 26-06-2024 **Revised:** 09-11-2024 **Accepted:** 12-05-2025 **Published:** 13-05-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding Author: Nidhya R

ABSTRACT

Deep learning is a part of modern machine learning that includes deep belief networks, deep neural networks, and recurrent neural networks. Computer vision, audio processing, and language comprehension are the most important sectors of deep learning. In many instances, these applications exceed human performance. In smart agriculture, deep learning gives novel ideas for increasing productivity and efficiency. Weed identification is an important application in crop areas that improves farming. This technology improves crop yields by identifying weeds. Also, it reduces resource wastage in agricultural practices. This paper presents a Convolutional Neural Network (CNN) model specifically designed to accurately identify and classify weeds using images of crop fields, augmented by the ImageNet dataset for enhanced feature extraction and model training. The model identifies essential characteristics, such as dimensions, form, spectral reflectance, and texture, to distinguish between crops and weeds. Unlike existing systems, our CNN-based approach achieves a high accuracy of 98 %. This improvement enhances weed identification efficiency and reduces pesticide usage, therefore it minimising environmental impact.

Keywords: Convolutional Neural Network (CNN); Crop Weed Detection; VGG19; Deep Learning.

RESUMEN

El aprendizaje profundo es una parte del aprendizaje automático moderno que incluye redes de creencias profundas, redes neuronales profundas y redes neuronales recurrentes. La visión por ordenador, el procesamiento de audio y la comprensión del lenguaje son los sectores más importantes del aprendizaje profundo. En muchos casos, estas aplicaciones superan el rendimiento humano. En la agricultura inteligente, el aprendizaje profundo aporta ideas novedosas para aumentar la productividad y la eficiencia. La identificación de malas hierbas es una aplicación importante en áreas de cultivo que mejora la agricultura. Esta tecnología mejora el rendimiento de los cultivos mediante la identificación de malas hierbas. Además, reduce el desperdicio de recursos en las prácticas agrícolas. Este artículo presenta un modelo de red neuronal convolucional (CNN) diseñado específicamente para identificar y clasificar con precisión las malas hierbas utilizando imágenes de campos de cultivo, aumentado por el conjunto de datos ImageNet para mejorar la extracción de características y el entrenamiento del modelo. El modelo identifica características esenciales, como dimensiones, forma, reflectancia espectral y textura, para distinguir entre cultivos y malas hierbas.

A diferencia de los sistemas existentes, nuestro enfoque basado en CNN alcanza una elevada precisión del 98 %. Esta mejora aumenta la eficacia de la identificación de malas hierbas y reduce el uso de pesticidas, con lo que se minimiza el impacto ambiental.

Palabras clave: Red Neuronal Convolucional (CNN); Detección de Malas Hierbas en Cultivos; VGG19; Aprendizaje Profundo.

INTRODUCTION

Agriculture is a primary sector that provides lives for millions of people globally. Also, it is important to ensure the security of the food.[1] Meanwhile, farmers face constant problems that impact productivity. Also, weed management is particularly major among these issues. Weeds are undesired plants that grow with crops.[2] They struggle with crops for necessities including water, nutrition, and light. Weeds can also prevent crop growth and affect gained crops. In certain cases, they spread rapidly over fields. It aggravates the issue. Even though some weeds are useful, their negative impacts on crops need care. To get healthy yields and sustainable farming, weed management is needed.[3]

This research aims to develop a system that may recognise and differentiate weeds from crops. It uses advances in machine learning and neural networks to complete this goal.[4] The system employs convolutional neural networks (CNNs) to analyse images of agricultural fields. These images are processed to detect and classify weeds effectively. Here, automation of weed detection reduces the need for manual labor.[5] Also, it improves the accuracy and efficiency of the process of weed detection. The model is developed using a thoroughly selected dataset of crop and weed images from various geographies.[6] This ensures the effectiveness and robustness of the model in many agricultural environments.

The main objective of this initiative is to minimize the usage of pesticides for weed control. Overusing herbicides raises farmer costs. It also increases serious concerns about environmental contamination.[7] Soil quality and water resources are affected by the chemicals that present in the herbicides. The proposed technology provides an accurate method for weed detection. This reduces the requirement for chemical interference. This method will promote the ecologically sound and environmentally acceptable agriculture techniques.

Despite its novel methodology, the existing system can have some restriction to distinguish whether it is crops or weeds.[8] The current method lacks practical recommendations for weed management. This limitation needs to be addressed and it should be improved in the future. Despite this limitation, the strategy gives successful solution to a major challenge which is faced by farmers nowadays. The combination of technology and farming has taken a big step forward.[9] The study focuses on using machine learning method to handle the issue of weed management. It attempts to achieve higher yields and better agricultural performance. Also, it works to make more sustainable in agricultural environment.

Literature survey

The CNN application in agriculture field has been deeply studied. Advanced image processing methodologies have been used for detection and weed categorization process.[10] Researchers have suggested multiple methods to enhance the accuracy in identifying the weed. Constantly, these developments improves the productivity of sustainable agriculture methods.[11] A study of using colour segmentation and edge detection methods are used to detect the weeds. Colour segmentation was used to differentiate the crops from the background image, while the edge detection highlights the weed edges based on their frequency.[12] This method focused on reducing the herbicide usage by simplifying the weed detection techniques through visual features.

Another method in this study focused on merging shape, colour, and size attributes for weed classification. Image segmentation techniques were applied to separate the crops from weeds. Shape analysis was used to differentiate the geometric differences between crops & weeds.[13] Moreover, texture analysis improves the ability to differentiate the weeds in an effective manner. Also, Artificial Neural Networks (ANNs) were chosen to manage the partial data. These methods show that machine learning's ability to achieve the accurate result of weed classification.[14] Machine vision research focused on the weed detection factors that includes size, shape, spectral reflectance, and texture. Algorithms like the excessive green technique were used to remove soil and unwanted objects. Meanwhile size-based features like area and perimeter were extracted. It shows the importance of machine vision in improving the accuracy of weed detection method.[15]

In certain agricultural environments, weed detection has been evaluated by deep learning models. These models contains CNN architectures that includes AlexNet, GoogLeNet, InceptionV3, and Xception.[16] Accuracy, precision, and recall were significantly higher in InceptionV3 than in other models. This proves that deep learning models could be useful for automating weed detection. Other studies depends on preprocessing

https://doi.org/10.56294/dm2025848

methods to improve image quality.[17] Noise filtering methods, such as median and mean filters, were applied. Image enhancement techniques contains converting RGB images to the HSV color space.[18] Advanced algorithms like Otsu's methodology were used. It helps to segment the images and handle the lighting variations.[19]

Additionally, a computer vision-based approach implemented OpenCV and ANN for early-stage weed detection. The process involved segmentation and classification of the images. It focused on dealing challenges including light intensity variations and region-specific weed detection.[20] These development shows the significance of combining machine learning and image processing approaches to achieve successful weed management. The methods offered by researchers help precision agriculture reduce herbicide use and boost crop yields.

Dataset

For this study, the "Weed Detection" dataset from Kaggle was utilized. This dataset contains annotated images, which are designed for automatic weed detection. It was made to help with the process of controlling weeds. The dataset comprises a substantial collection of images suitable for machine learning applications. The purpose of these images is used to train & evaluate machine learning systems. The primary aim is to properly differentiate whether the plant is "weeds" or "crops". For this, 80 % of the dataset has been taken for training purpose. While, the remaining 20 % of images is used for evaluating the efficiency of the model. A diversed images are included in the dataset to increase its adaptability. It improves the ability of model's generalisation. Because of its comprehensive nature, the dataset is suitable for the use of practical applications in the real world. It is an essential resource for efficient weed detection in agriculture.

METHOD

CNN and VGG-19 architecture in figure 1. are used to detect and classify crop weeds in the suggested system. This methodology contains preprocessing, model optimisation, and prediction to improve the overall workflow efficiency. The dataset utilised is the Kaggle Weed Detection Dataset, encompassing 1300 annotated images. The images are splited into weeds and crops to train and validate the model. Photos are resized to 224x224 pixels, and pixel values are normalized. Data augmentation techniques includes rotation, scaling, and flipping are used. This aids to improve the robustness of the model and generalization.

.

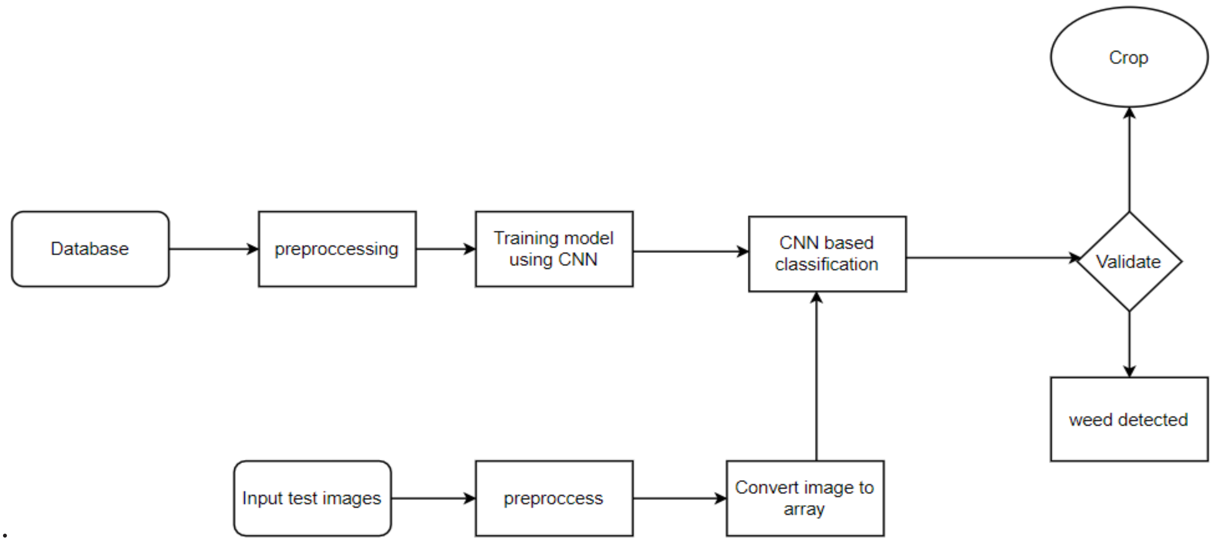


Figure 1. Architecture of weed detection

Preprocessing

The preprocessing module processes raw images to make them suitable for the CNN model. This contains scaling images to 224×224 pixels. It normalising the pixel intensities for homogeneity, and also using the augmentation techniques. The dataset is diverse and optimised for training after these processes. The preprocessing module transforms the images into arrays, that ensures the compatibility with the model.

Model construction

The CNN model is built by using the VGG-19 architecture model in figure 2. Pretrained weights from the ImageNet dataset are used for the convolutional layer and pooling layers. While the fully connected layers are replaced with a Flatten layer and also with two optimized Fully Connected layers. A two-label SoftMax classifier is used for binary classification. Here, Transfer learning is used, that allows the reuse of pretrained parameters for efficient and accurate feature extraction. It also focusing on size, shape and texture to differentiate weeds from crops.

https://doi.org/10.56294/dm2025848

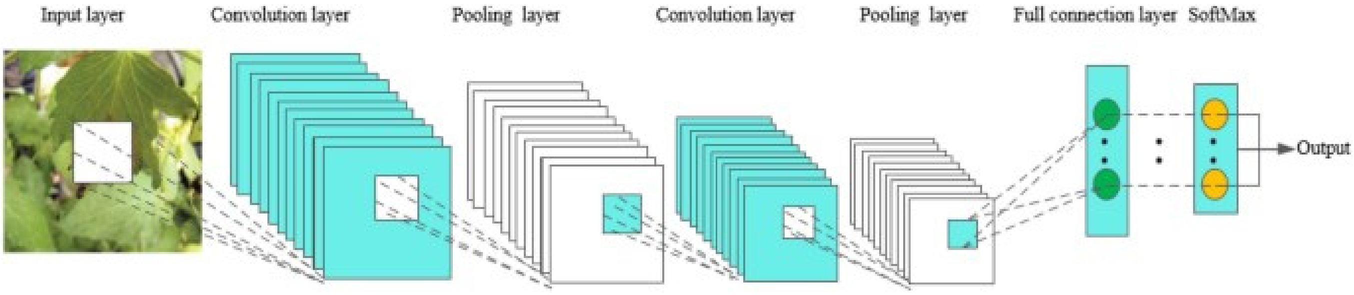


Figure 2. Deep Convolutional Neural Network Architecture for Image Analysis

Model training and optimization

Here, 75 % of the dataset is used for training during the training phase; while the remaining 25 % dataset depends on testing to evaluate the performance of the model. The training procedure is conducted iteratively with the Adam optimiser. The training process uses the categorical cross-entropy as the loss function. Momentum parameters are established to ensure the stability in the optimisation process. Adjusting learning rate helps to maximize the accuracy. Here training continues till the model achieves the threshold of high accuracy.

Algorithm

1. Input: Kaggle Weed Detection dataset.
2. Preprocessing:
    - Load images IMG and resize to 224×224 px.
    - Normalize pixel values.
    - Apply image augmentation techniques (rotation_rt,scaling,flipping_fp).

3. Model Initialization:
    - Load VGG-19 with pretrained ImageNet_weights.
    - Freeze conv & pooling layers.
    - Add a classification head: Flatten layer → 2 Fully Connected layers → 2-label SoftMax layer.

4. Training:
    - Divide the dataset into training_set (75 %) & testing_set (25 %).
    - Train_model with augmented_training_data.
    - Optimize parameters (momentum,learning_rate,accuracy_standard).

5. Prediction:
    - Input test_img to the trained model.
    - Output the classification result: “Weed” or “Crop.”

6. Output: classify label with high_confidence.

Prediction module

The fully trained model detects whether an input image is “a crop” or “a weed”. Here, CNN pipeline helps to extract relevant features from the image. The model then classifies the image with the level of high_confidence. This classification process gives accurate & reliable predictions. Then the tool operates in real-time, which enables quick decision-making. Farmers and agricultural specialists can use these results for effective weed management.

RESULTS

The outcome of the developed model system shows its effectiveness in identifying crops and weeds with high accuracy. The model achieved nearly 99 % training accuracy and 98 % testing accuracy. It shows its robustness by learning from the training dataset and generalising to new data. The performance metrics state that with an increase in the size and diversity of the dataset. Further improvements in testing accuracy can be achieved.

Training Performance

Figure 3 shows the training process, which clearly defines how accuracy increases and loss decreases with each session. The model was trained for 15 epochs, but it stopped automatically after 8 epochs. It happened due to optimized loss and accuracy, exhibiting efficient learning and without overfitting this model This improved convergence of the model within a shorter number of epochs highlights the reliability of the VGG-19 architecture. This efficiency is further reinforced by the effective use of the architecture's design principles. Also, transfer learning is used to achieve these outcomes.

```
Epoch 1/15
28/28 [==============================] - 153s 5s/step - loss: 0.3854 - categorical_accuracy: 0.8294 - val_loss: 0.1202 - v
Epoch 2/15
28/28 [==============================] - 17s 606ms/step - loss: 0.1248 - categorical_accuracy: 0.9585 - val_loss: 0.0645 -
Epoch 3/15
28/28 [==============================] - 17s 591ms/step - loss: 0.0951 - categorical_accuracy: 0.9731 - val_loss: 0.0543 -
Epoch 4/15
28/28 [==============================] - 17s 591ms/step - loss: 0.0766 - categorical_accuracy: 0.9731 - val_loss: 0.0367 -
Epoch 5/15
28/28 [==============================] - 17s 591ms/step - loss: 0.0659 - categorical_accuracy: 0.9809 - val_loss: 0.0380 -
Epoch 6/15
28/28 [==============================] - 17s 599ms/step - loss: 0.0546 - categorical_accuracy: 0.9877 - val_loss: 0.0314 -
Epoch 7/15
28/28 [==============================] - 18s 633ms/step - loss: 0.0507 - categorical_accuracy: 0.9910 - val_loss: 0.0320 -
Epoch 8/15
28/28 [==============================] - ETA: 0s - loss: 0.0420 - categorical_accuracy: 0.9933
Ending training
28/28 [==============================] - 17s 593ms/step - loss: 0.0420 - categorical_accuracy: 0.9933 - val_loss: 0.0222 -
```

Figure 3. Training Progress of a Deep Learning Model: Epoch-wise Accuracy and Loss

Accuracy and Loss Metrics

Figure 4. Model accuracy graph

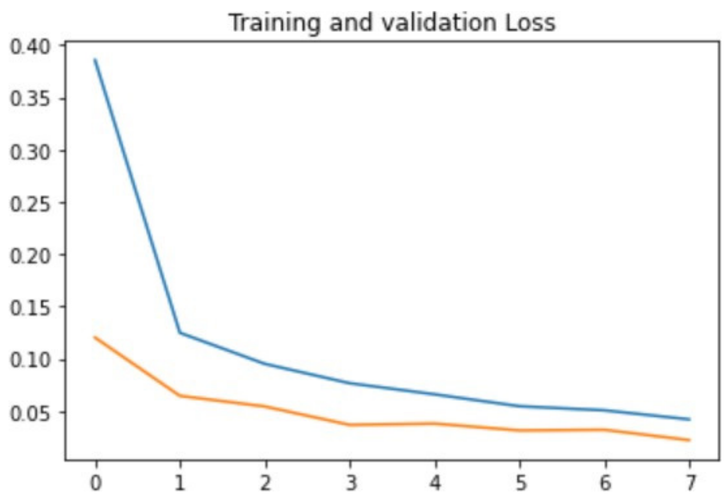


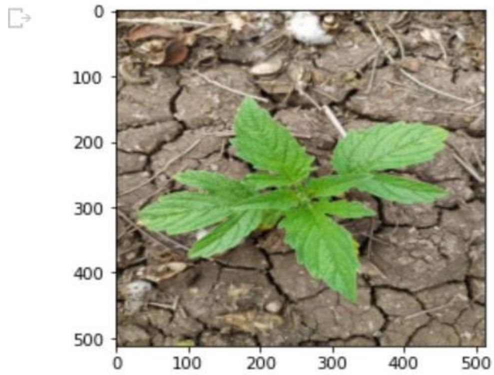
Figure 5. Model loss graph

Figure 4 the Model Accuracy Graph, shows that accuracy improves consistently throughout epochs. This shows that the model trained for the identifying the images as being either weeds or crops. Figure 5, the Model Loss Graph, shows a consistent decrease in loss as the epochs rise. This means that the optimization for the weights

https://doi.org/10.56294/dm2025848

and the parameters of the model during training. As a result of the model's capacity for learning, the accuracy has been steadily improving. Similarly, the decrease in loss emphasises the accuracy of its predictions. The graphs collectively verify the effectiveness of the model design, dataset quality, and pre-processing methods.

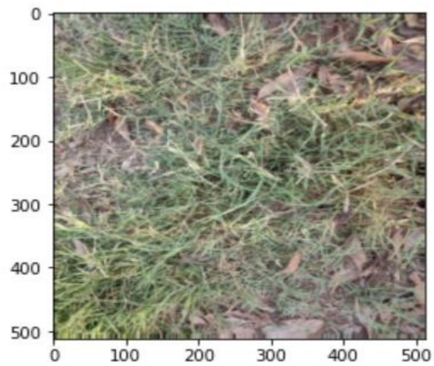
Output Results




```
if(classes[0][0]>0.8 and classes[0][0]<=1):
  print("plant")
else:
  print("weed")
```

```
plant
```

Figure 6. Output of plant




```
if(classes[0][0]>0.8 and classes[0][0]<=1):
  print("plant")
else:
  print("Given Input is weed")
```

```
weed
```

Figure 7. Output of weed

The output images show the model's input image classification accuracy. In Case 1 (figure 6), the system correctly identified the input image as a plant. This shows the effectiveness of the model on the plant features. In Case 2 (figure 7), the system successfully classified the input image as a weed. This result proves the model's capability to differentiate between different categories with precision. These results evaluate the model's performance in real-world situations. This process classifies images with high reliability which shows the robustness of the design. Collectively, these cases confirm the system's practical applicability and efficiency.

The analysis of the results confirms the successful implementation of the CNN-based weed detection system. The accuracy of the system and classification outcome meet the project objectives. It gives an efficient and reliable tool for weed management. The model performs well. However, to increase the dataset and improving the training setups enables future improvements. This system gives a practical solution for farmers. It promotes precision agriculture, also it reduces the environmental impact.

Testing & validation

Validation

Finally, the confusion matrix evaluates the performance of our proposed model. It displays CNN classification on the testing dataset. The confusion matrix evaluates the crop and weed detection in our model.

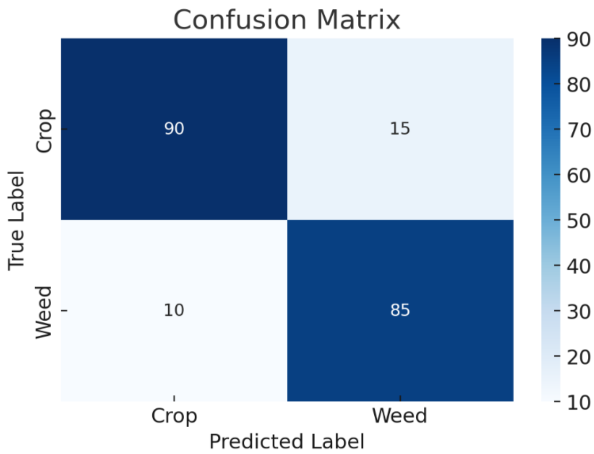


Figure 8. Confusion Matrix: Classification Performance for Crop and Weed Detection

As shown in figure 8, the confusion matrix reveals:

- True Positives (Crop correctly classified): 90.
- True Negatives (Weed correctly classified): 85.
- False Positives (Crop misclassified as Weed): 15.
- False Negatives (Weed misclassified as Crop): 10.

This performance shows that the model achieves high accuracy while classifying the input images. However, misclassifications indicates optimising the training dataset by increasing its size and variety.

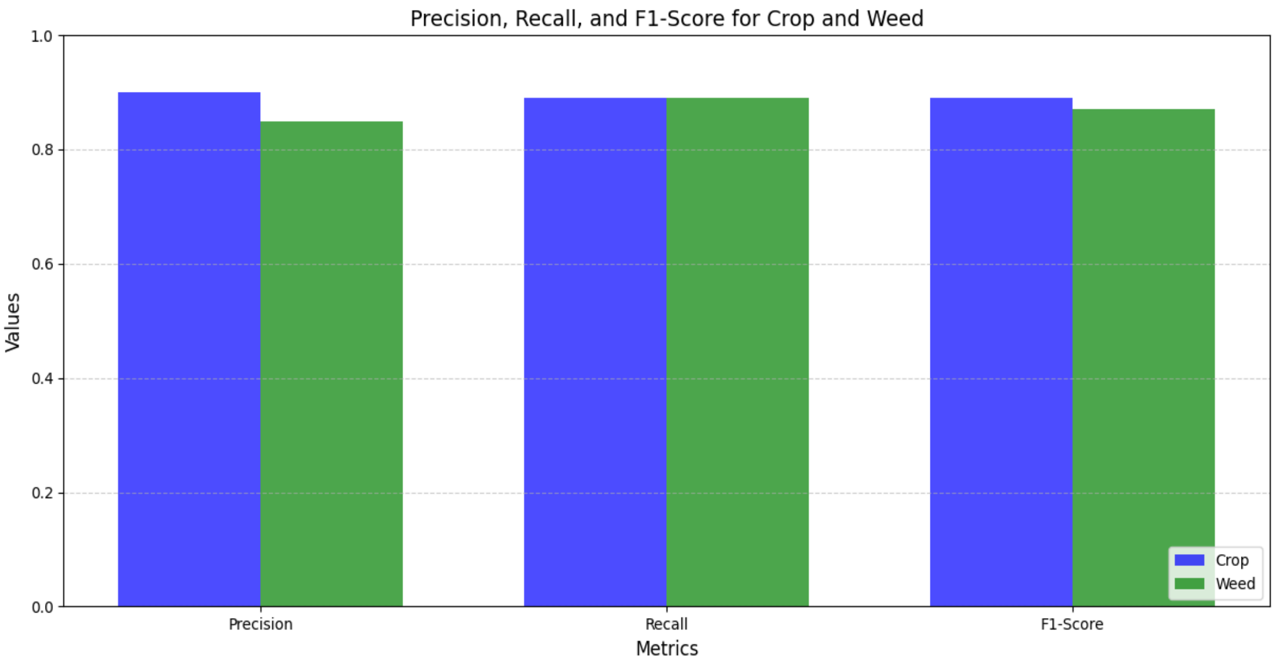


Figure 9. Comparison of Precision, Recall, and F1-Score for Crop and Weed Classification

.

To evaluate the effectiveness of the proposed CNN model in figure 9, the performance metrics of Precision, Recall, and F1-Score were calculated for the two classes: Crop and Weed. All of these measures give a clear view of model accuracy and reliability between the two classes.

The Performance Comparison Chart in figure 10 compares the accuracy of SVM, ANN, AlexNet, VGG-16, and the proposed CNN (VGG-19). The proposed model achieves the highest accuracy (98 %), outperforming other methods. This shows the effectiveness of the proposed architecture for weed detection tasks.

https://doi.org/10.56294/dm2025848

The confusion matrix, F1-score, precision, and recall show the proposed model classifies crops and weeds with 98 % accuracy. Despite a few misclassifications, the model performs better than other methods which include SVM, ANN, and VGG-16. The dataset's performance could be improved further by optimisation.

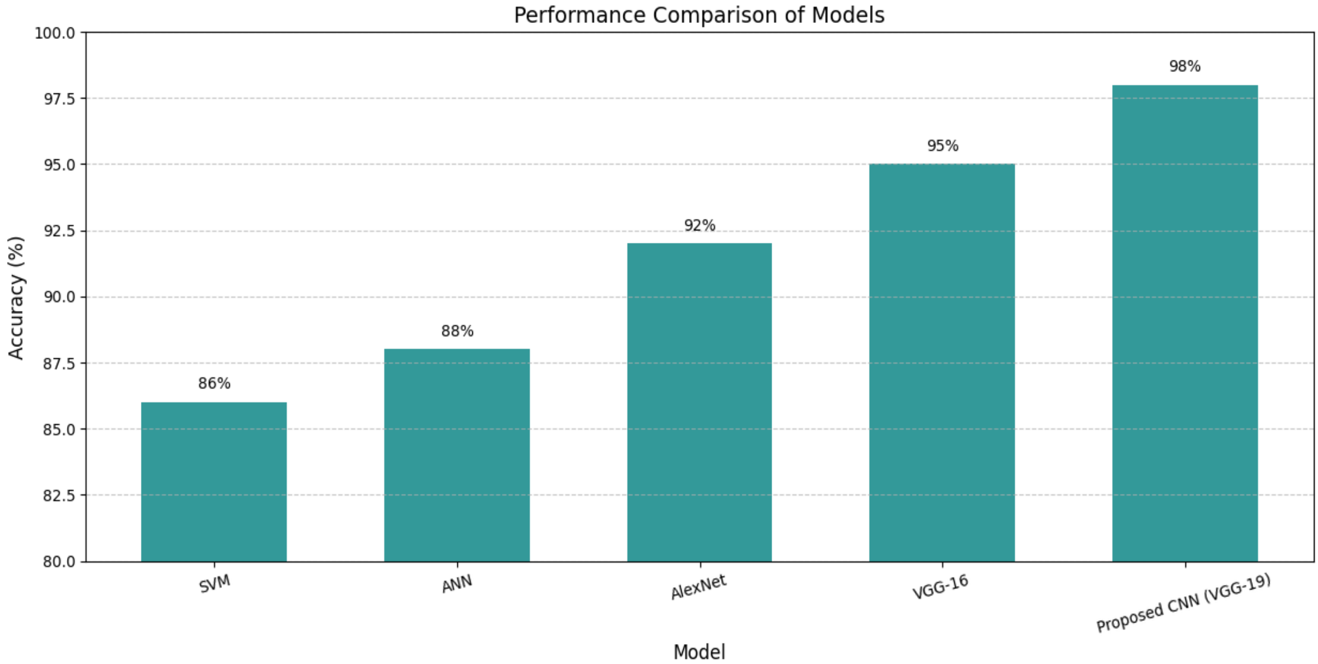


Figure 10. Model Accuracy Comparison: SVM, ANN, AlexNet, VGG-16, and Proposed CNN (VGG-19)

CONCLUSIONS

The proposed work effectively employs the concept of feature extraction to reduce computational and storage requirements. Performance is not compromised when processing huge datasets with this method. Significant features are chosen from the data which helps effectively to train the model. In order to make use of CNN's robustness for feature learning, a classification algorithm based on it is used. The proposed model is effective for image classification applications The results show 98 % of accuracy, which denotes the great performance of this model. This performs best in traditional classification and regression methods. This shows robustness of the system and dependability in solving challenges in weed detection in agricultural fields. Moreover, to improve the accuracy, the model can be trained on a larger and more diverse dataset. This approach has great promise for practical applications like a web-based interface for uploading and analysing images to detect weeds. In the future, this project will connect the CNN model to Internet of Things (IoT) devices. It will automatically find, classify, and get rid of weeds. This will pave the way for a fully automated, real-time smart agriculture solution with the right hardware and software improvements.

BIBLIOGRAPHIC REFERENCES

1. Machine learning in agriculture: a review of crop management applications | Multimedia Tools and Applications. https://link.springer.com/article/10.1007/s11042-023-16105-2

2. Kumar S et al. Weed management challenges in modern agriculture: The role of environmental factors and fertilization strategies. Crop Prot. 2024 Nov 1;185:106903.

3. Islam N, Rashid MM, Wibowo S, Xu CY, Morshed A, Wasimi SA, et al. Early Weed Detection Using Image Processing and Machine Learning Techniques in an Australian Chilli Farm. Agriculture. 2021 May;11(5):387.

4. Deep Learning-Based Weed–Crop Recognition for Smart Agricultural Equipment: A Review. https://www.mdpi.com/2073-4395/14/2/363

5. Robot-Based Weed Identification And Control System | 11 | Precision Ag. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003435228-11/robot-based-weed-identification-control-system-rashmi-bangale-mohit-kumar

6. Adhinata FD, Wahyono, Sumiharto R. A comprehensive survey on weed and crop classification using machine learning and deep learning. Artif Intell Agric. 2024 Sep 1;13:45–63.

https://doi.org/10.56294/dm2025848

7. Sustainable Paddy Farming in Edirne: Evaluating the Impacts of Excessive Fertilizer and Pesticide Use. https://www.mdpi.com/2071-1050/16/17/7814

8. Zhou W, Li M, Achal V. A comprehensive review on environmental and human health impacts of chemical pesticide usage. Emerg Contam. 2025 Jan 1;11(1):100410.

9. Innovating in an Uncertain World: Understanding the Social, Technical and Systemic Barriers to Farmers Adopting New Technologies. https://www.mdpi.com/2078-1547/15/2/32

10. Application of Convolutional Neural Networks in Weed Detection and Identification: A Systematic Review. https://www.mdpi.com/2077-0472/14/4/568

11. Advances in ground robotic technologies for site-specific weed management in precision agriculture: A review - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0168169924007543

12. Automatic visual recognition, detection and classification of weeds in cotton fields based on machine vision - ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0261219424003946

13. A machine learning framework for the identification of crops and weeds based on shape curvature and texture properties | International Journal of Information Technology. https://link.springer.com/article/10.1007/s41870-023-01598-9

14. Weed Detection and Classification with Computer Vision Using a Limited Image Dataset. https://www.mdpi.com/2076-3417/14/11/4839

15. Behrouzi B. Enhancing Weed Detection and Segmentation with Advanced Deep Learning Algorithms leveraging RGB-NIR Imaging. Politecnico di Torino; 2024. https://webthesis.biblio.polito.it/31764/

16. Deep Learning Techniques for Weed Detection in Agricultural Environments: A Comprehensive Review | IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/10570187

17. Optimization enabled Deep Quantum Neural Network for weed classification and density estimation - ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0957417423031810

18. Identification of weeds in cotton fields at various growth stages using color feature techniques - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S1125471824000215

19. Chen S, Memon MS, Shen B, Guo J, Du Z, Tang Z, et al. Identification of weeds in cotton fields at various growth stages using color feature techniques. Ital J Agron. 2024 Dec 3;19(4):100021.

20. A Comprehensive Review on Deep Learning Assisted Computer Vision Techniques for Smart Greenhouse Agriculture | IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/10379667

FINANCING
None.

CONFLICT OF INTEREST
Authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION
Conceptualization: Nidhya R, Pavithra D, Smilarubavathy G, D Mythrayee.
Data curation: Nidhya R, Pavithra D, Smilarubavathy G, D Mythrayee.
Formal analysis: Nidhya R, Pavithra D, Smilarubavathy G, D Mythrayee.
Drafting - original draft: Nidhya R, Pavithra D, Smilarubavathy G, D Mythrayee.
Writing - proofreading and editing: Nidhya R, Pavithra D, Smilarubavathy G, D Mythrayee.