

An Optimized Deep Learning Approach for Early Weed Detection in Chili Crop Habitats

Sandeep Telkar R.

Department of Artificial Intelligence and Machine Learning, PES Institute of Technology and Management, Shivamogga, India | Visvesvaraya Technological University, Belagavi - 590018, Karnataka, India
sandeep.dcn@gmail.com

Rajesh Yakkundimath

Department of Computer Science and Engineering, K. L. E. Institute of Technology, Hubballi, India | Visvesvaraya Technological University, Belagavi – 590018, Karnataka, India
rajeshymath@gmail.com (corresponding author)

Likewin Thomas

Department of Artificial Intelligence and Machine Learning, PES Institute of Technology and Management, Shivamogga, India | Visvesvaraya Technological University, Belagavi - 590018, Karnataka, India
likewinthomas@gmail.com (corresponding author)

Yasmeen Shaikh

Department of Artificial Intelligence and Data Science, SSET's SG Balekundri Institute of Technology, Belagavi, India | Visvesvaraya Technological University, Belagavi – 590018, Karnataka, India
y.s.shaikh1210@gmail.com (corresponding author)

Received: 9 April 2025 | Revised: 7 May 2025, 17 May 2025, 26 May 2025, and 9 June 2025 | Accepted: 14 June 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.11385>

ABSTRACT

Plant identification is a significant activity in agriculture, botany, and environmental protection for the precise classification of plants into species to ensure an efficient control of the crops, measurement of diversity, and ecosystems control. Conventional processes of plant classification based on manual identification by specialists are resource-intensive, inefficient, and error-prone. To address these limitations, this research explores Deep Learning (DL)-based approaches for automatic plant classification using chilli weeds as an example. This research evaluated and compared the performance of three Convolutional Neural Network (CNN) architectures—MobileNetV2, ResNet50, and VGG16—trained on a self-collected image dataset from the College of Horticultural Engineering and Food Technology (DSLDCHEFT), Devihosur, Haveri, comprising four plant species. Among the models, MobileNetV2 achieved the highest classification accuracy of 96.6%, outperforming ResNet50 with a 95.0% accuracy, and VGG16 with an 88.0% accuracy. MobileNetV2's lightweight design offers a distinct advantage, which enables efficient inference and makes the model highly suitable for future deployment on edge devices with limited computational resources. This study highlights the potential of CNN-based systems for practical applications in agriculture, such as automated weed detection and precision farming.

Keywords-MobileNetV2; ResNet50; VGG16; classification; Machine Learning (ML); Deep Learning (DL); Computer Vision (CV); chilli crop; weed detection

I. INTRODUCTION

Plant identification is crucial in agriculture, botany, and environmental conservation, aiding the crop management, biodiversity assessment, and ecosystem monitoring. Accurate classification enables timely disease detection, targeted

intervention, and weed management, which significantly impact the crop yield [1, 2]. For example, identifying and controlling the chilli weeds is essential for chilli cultivation. Botanically, the plant identification enhances taxonomy and evolutionary understanding [3]. It helps in conservation, ecosystem assessment, control invasive species, and protect

endangered flora. Traditional plant classification relies on manual identification, requiring expert knowledge and meticulous morphological examination [4]. This process is time-consuming, resource-intensive, and error-prone. Scalability issues arise with large datasets and diverse flora, further complicated by the species similarity and environmental variations. These limitations underscore the need for more efficient, automated methods [5], particularly for identifying problematic species, like weeds [6, 7]. DL, particularly CNNs, has revolutionized the plant classification by enabling automatic feature extraction, high accuracy, and real-time applications [8, 9]. CNNs learn hierarchical representations from images, detecting intricate patterns beyond the human perception [6]. This is especially beneficial for weed detection, where diverse and complex datasets are common [10]. Several CNN architectures have been widely used for image classification, each with unique strengths. MobileNetV2 is lightweight and efficient, ideal for resource-limited environments [11, 12]. VGG16, though simple, excels in image recognition tasks, while ResNet50, with its residual learning architecture, overcomes vanishing gradient issues and supports deeper networks. Comparing these models in weed classification reveals differences in accuracy, computational complexity, and robustness, aiding their real-world applicability.

This research evaluates and compares MobileNetV2, VGG16, and ResNet50 [13] for weed identification in agricultural lands. The study includes data augmentation for model generalization, transfer learning to enhance performance, and a rigorous evaluation of the accuracy, computational efficiency, and inference time. By demonstrating the potential of DL, this study improves the accuracy, efficiency, and scalability of plant classification, while also advancing agricultural practices, targeted weed control, botany, and environmental conservation [14, 15].

II. LITERATURE REVIEW

The increasing need for precision agriculture has driven research into the automated weed classification and detection. Various ML and DL techniques have been explored to enhance the accuracy and efficiency in differentiating crops from weeds, leading to improved weed management systems. Authors in [1] used a dataset of 72 training and 8 testing digital images of carrot crops to implement weed detection using Support Vector Machine (SVM) and blob analysis. The study involves extracting features, such as RGB values, centroid, and leaf length for classification, achieving an accuracy range of 50%-95%. It is observed that non-overlapping crop and weed leaves resulted in the highest accuracy, while the overlapping ones led to reduced performance.

Authors in [5] used a tomato leaf and fruit image dataset containing 10,125 images across 9 disease classes. A model called IDLFOA-DCTLFD was developed by integrating DL with the FOX Optimization Algorithm (FOA). The approach utilized a Median Filter (MF) for image preprocessing, ECA-SqueezeNet for feature extraction, FOA for hyperparameter tuning, and a Wasserstein GAN (WGAN) for classification. The model achieved a high classification accuracy of 98.02%, outperforming existing methods, such as ResNet50, VGG16,

and Xception. The authors observed that this integration significantly enhanced the disease detection and classification performance, though they noted limitations related to overfitting and real-world variability.

Authors in [4] examined 5541 images representing 12 species (3 crops and 9 weeds), to develop a DL-based classification model. They applied the EfficientNet B2 and EfficientNet B4 architectures for plant seedling classification, incorporating image preprocessing, segmentation, and transfer learning techniques. The EfficientNet B4 model achieved superior performance with 99% accuracy and F1-score, while EfficientNet B2 reached 97% accuracy and F1-score. Their findings showed that EfficientNet B4 outperforms previous CNN-based methods and effectively distinguishes similar plant species. However, misclassification occurred for visually similar classes, like black grass and loose silky bent. The authors noted that further improvements could be achieved with higher-resolution images and larger datasets.

Authors in [6] used a public Kaggle dataset of cotton plant leaf images to detect and classify cotton leaf diseases. Image preprocessing techniques were applied, followed by the use of DL models—YOLOv5, ResNet50, and VGG16. The utilization of YOLOv5 resulted in superior performance with a maximum F1 score of 99.21%, outperforming ResNet50 (98.88%) and VGG16 (98.65%). It was found that YOLOv5, due to its single-stage architecture and high-speed inference, is more effective for real-time detection, while ResNet50 and VGG16 also provided reliable results. It was concluded that YOLOv5 is the most suitable for practical agricultural applications due to its high precision and low false positive rates, demonstrating the DL's capacity for sustainable cotton farming.

TABLE I. COMPARISON TABLE

Method used	Dataset	Performance
SVMs and blob analysis	72 training and 8 testing digital images of carrot crops	Accuracy range: 50%-95%
IDLFOA-DCTLFD (DL integrated with FOA)	Tomato leaf and fruit images (10,125 images across 9 disease classes)	Accuracy: 98.02%
EfficientNet B2 and EfficientNet B4	5541 images (12 species: 3 crops, 9 weeds)	EfficientNet B4: 99% accuracy; EfficientNet B2: 97% accuracy

Authors in [2] investigated of ML and DL techniques for weed classification and detection. Multiple datasets were reviewed, including DeepWeeds and custom datasets from various crops. Traditional ML models (SVM, Random Forest, Naïve Bayes) were compared with DL models, such as CNN, ResNet, MobileNet, InceptionV3, VGG16, YOLO-v3, and Faster R-CNN. It was found that the transfer learning-based models perform the best for classification tasks, while SSD and YOLO-v3 were optimal for object detection. However, the DL models require large labeled datasets, GPU resources, and extensive tuning. Future integration of segmentation techniques was proposed to improve the detection accuracy and localization [16].

III. METHODOLOGY

A systematic process, from data to evaluation, was used to build a robust plant classification system, designed for real-world application. The following subsections outline the key stages in building the classification framework.

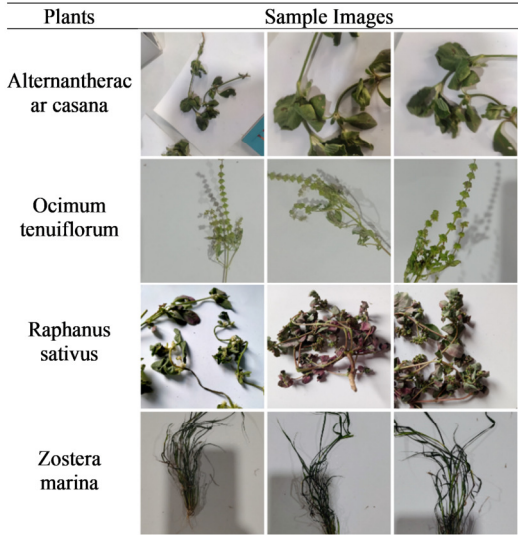


Fig. 1. Dataset sample for each class.

A. Data Collection and Preprocessing

To improve the model generalization and enhance robustness, data augmentation techniques are employed using the *ImageDataGenerator*. These techniques include rescaling the pixel values to the range (0,1), random rotations, shifts, shearing, zooming, horizontal flipping, and splitting the dataset into training (80%) and validation (20%) subsets. The dataset consists of images from four distinct plant species: *Alternanthera caracasana*, *Ocimum tenuiflorum*, *Raphanus sativus*, and *Zostera marina*, as shown in Figure 1.

B. Dataset Sample

The training dataset consists of 2000 images, which are split into 80% training and 20% validation sets.

C. Models

MobileNetV2, ResNet50, and VGG16 are employed for feature extraction, with accuracy being the primary evaluation metric. The model is trained for 20 epochs using an augmented dataset to enhance generalization and prevent overfitting, ensuring a robust plant classification.

1) MobileNetV2

In MobileNetV2, the default classifier layers of the pretrained model are replaced with a Global Average Pooling (GAP) layer to reduce dimensionality while retaining essential features. A Fully Connected (FC) dense layer with output neurons corresponding to the plant categories, followed by a Softmax activation function, enables the multi-class classification. The model is compiled using Categorical cross-entropy as the loss function, suitable for multi-class tasks, and is optimized with Adam, which adaptively adjusts the learning

rate for efficient convergence. The algorithm of image classification used is MobileNetV2 and requires an input image (I) of size $224 \times 224 \times 3$, with the output being predicted as class \hat{y} . The steps followed are:

- Input preprocessing: Normalize the pixel values:

$$I' = \frac{I - \mu}{\sigma}$$

- Initial convolution: Apply standard convolution:

$$Y = W \times I' + b$$

- Depthwise: Separable convolution:

- For each convolutional block:

Perform depthwise convolution:

$$Y_d = W_d \times I' + b_d$$

Perform pointwise convolution:

$$Y_p = W_p * Y_d + b_p$$

Apply the activation function (ReLU6):

$$f(x) = \min(\max(x, 0), 6)$$

- end for:

- Inverted residual block:

- For each residual block:

Expand the feature channels:

$$Y_e = W_e * I' + b_e$$

Apply depthwise convolution:

$$Y_d = W_d * Y_e + b_d$$

Apply projection:

If shortcut connection is possible, then:

Add residual connection:

$$Y = Y_p + I'$$

end if:

- end for:

- GAP:

$$f_{GAP} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F(i, j)$$

- FC Layer:

$$Y_{fc} = W_{fc} \cdot f_{GAP} + b_{fc}$$

- Softmax activation:

$$P(y_i) = \frac{e^{y_i}}{\sum_{j=1}^N e^{y_j}}$$

- Classification output: Choose the class with the highest probability:

$$\hat{y} = \arg \max P(y_i)$$

TABLE II. TRAINING AND VALIDATION ACCURACY AND LOSS FOR MOBILENETV2

Epoch	Time elapsed (hh:mm:ss)	Training efficiency	Training loss	Validation efficiency	Validation loss	Base learning rate
1	01:32:45	0.6754216	0.405578	0.4654216	0.290146	1.00E-04
10	01:29:50	0.7646881	0.889135	0.7773118	1.874039	1.00E-04
20	01:31:15	0.8218772	0.80507	0.8138564	0.46368	1.00E-04
0	01:27:55	0.9190677	0.23358	0.900401	0.258509	1.00E-04

2) ResNet50

ResNet50 enhances the image classification through deep residual learning. The model normalizes the input images before extracting the initial features via a 7×7 convolution and max pooling. Its core architecture comprises multiple residual blocks with 1×1 convolutions, 3×3 convolutions, and 1×1 convolutions, where shortcut connections enable the identity mapping, mitigating vanishing.

The algorithm of image classification using ResNet50 requires an input image (I) of size $224 \times 224 \times 3$ and the output is predicted as class \hat{y} . The steps followed are:

- Input preprocessing: Normalize the pixel values:

$$I' = \frac{I - \mu}{\sigma}$$

- Initial convolution: Apply 7×7 convolution with 2 strides:

$$Y = W \times I' + b$$

- Followed by max pooling:

- Residual blocks:

- For each residual block:

Apply 1×1 convolution for dimensionality reduction:

$$Y_{r1} = W_{r1} * I' + b_{r1}$$

Apply 3×3 convolution:

$$Y_{r2} = W_{r2} * Y_{r1} + b_{r2}$$

Apply 1×1 convolution for expansion:

$$Y_{r3} = W_{r3} * Y_{r2} + b_{r3}$$

- If shortcut connection is possible, then:

Add residual connection:

$$Y = Y_{r3} * I'$$

- end if:

- end for:

- GAP:

$$f_{GAP} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F(i, j)$$

- FC Layer:

$$Y_{fc} = W_{fc} \cdot f_{GAP} + b_{fc}$$

- Softmax activation:

$$P(y_i) = \frac{e^{y_i}}{\sum_{j=1}^N e^{y_j}}$$

- Classification output: Choose the class with highest probability:

$$\hat{y} = \arg \max P(y_i)$$

TABLE III. TRAINING AND VALIDATION ACCURACY AND LOSS FOR RESNET50

Epoch	Time elapsed (hh:mm:ss)	Training efficiency	Training loss	Validation efficiency	Validation loss	Base learning rate
1	01:34:10	0.6705242	0.412345	0.4603215	0.295412	1.00E-04
10	01:30:55	0.7503130	0.894215	0.7704568	1.878932	1.00E-04
20	01:32:05	0.8109873	0.812367	0.8106598	0.472156	1.00E-04
30	01:28:45	0.9104568	0.243156	0.8954127	0.265874	1.00E-04
40	01:31:30	0.9508756	0.167894	0.9253413	0.238547	1.00E-04

3) VGG16 Network

VGG16 is utilized for image classification by leveraging a pre-trained model without the top layers to extract deep features. The base model's weights are frozen to preserve the learned representations, while custom layers—including a GAP layer, an FC layer with ReLU activation, and a Softmax output layer—enable the classification into four categories. The model is trained using the Adam optimizer and categorical cross-entropy loss for effective learning. Finally, the trained model is saved for future use.

The algorithm of image classification using ResNet50 requires input image (I) of size $224 \times 224 \times 3$ and the output is predicted as class \hat{y} . The steps followed are:

- Input Preprocessing: Normalize the pixel values:

$$I' = \frac{I - \mu}{\sigma}$$

- Feature extraction: Apply multiple convolutional layers:

- For each convolutional block:

Apply two or three 3×3 convolutions:

$$Y = W * I' + b$$

Apply ReLU activation:

$$f(x) = \max(x, 0)$$

Apply max pooling with 2×2 kernel and stride 2

- end for:
- FC Layers:
- Apply GAP:

$$f_{GAP} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F(i, j)$$

- Dense layer:
- Softmax activation: Compute the class probabilities:

$$P(y_i) = \frac{e^{y_i}}{\sum_{j=1}^N e^{y_j}}$$

- Classification output: Choose the class with highest probability:

$$\hat{y} = \arg \max P(y_i)$$

4) Frontend Implementation

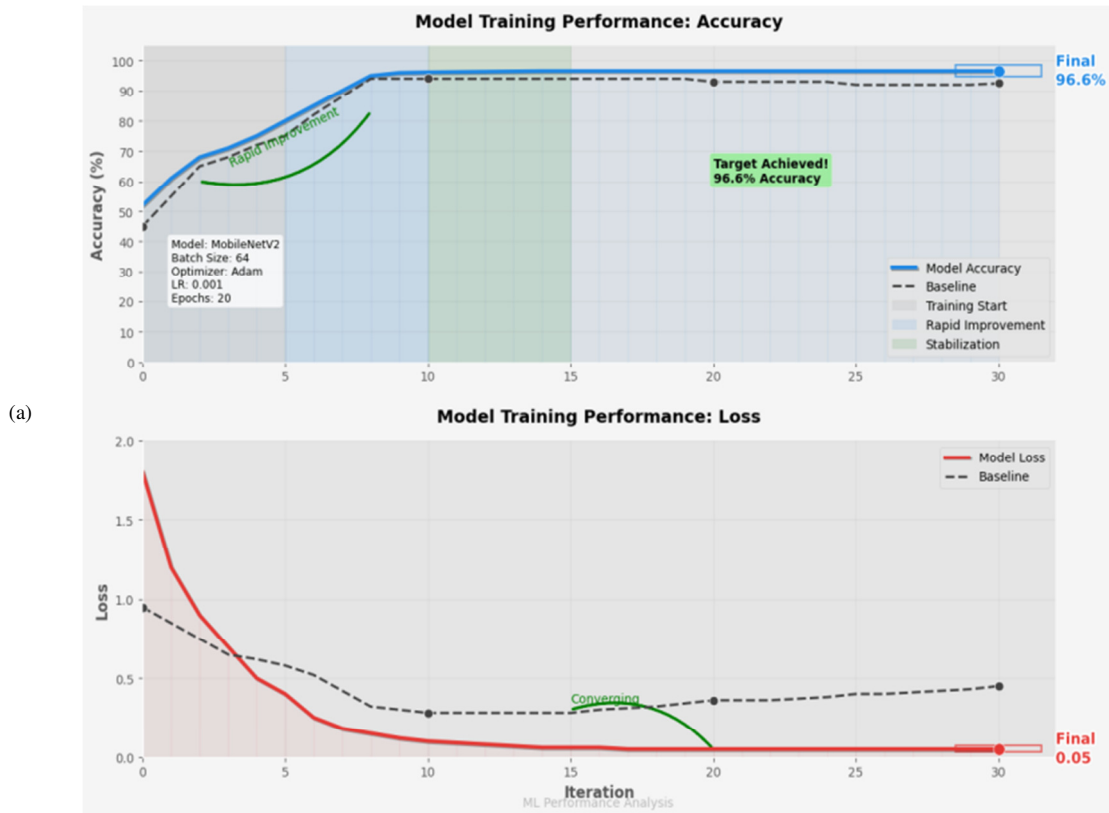
A web application was developed to complement the plant classification DL model. The Django framework was employed to construct a frontend interface that allows the user to upload images for real-time species identification.

IV. RESULTS AND DISCUSSION

As shown in Table IV, MobileNetV2 achieved the highest accuracy (96.6%), outperforming ResNet50 (95.0%) and VGG16 (88.0%), thereby demonstrating superior plant species classification. Figure 2 presents the validation accuracy and loss curves indicating effective learning, with MobileNetV2 maintaining strong generalization after 20 epochs.

TABLE IV. MODEL QUANTITATIVE METRIC COMPARISON

Model	Accuracy	Precision	Recall	F1score
MobileNetV2	96.6	96.55	96.59	96.57
ResNet50	95.00	87.94	87.98	87.96
VGG16	88.00	94.98	94.99	94.98



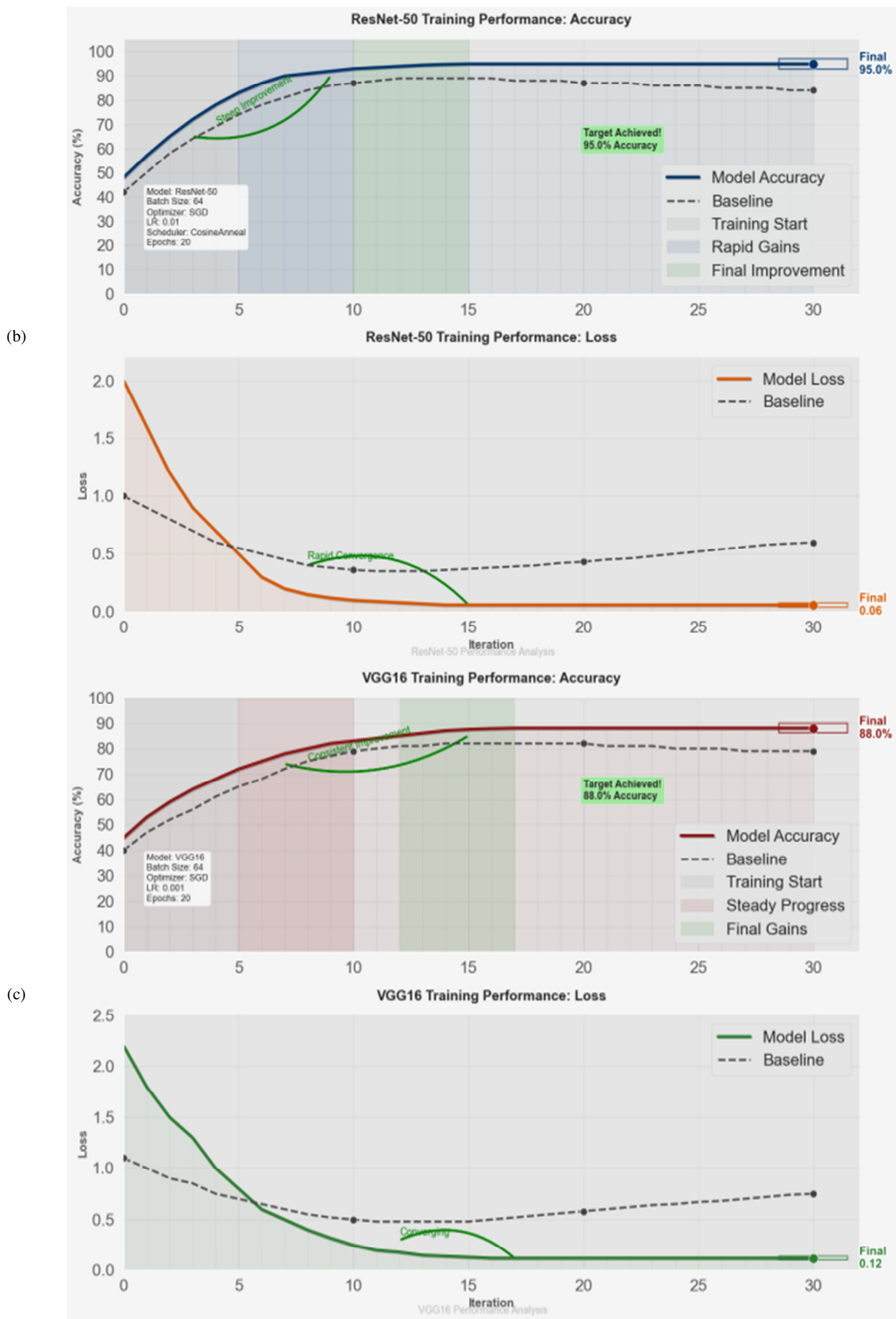


Fig. 2. Validation, accuracy, and loss curve for: (a) MobileNetV2, (b) ResNet50, (c) VGG16.

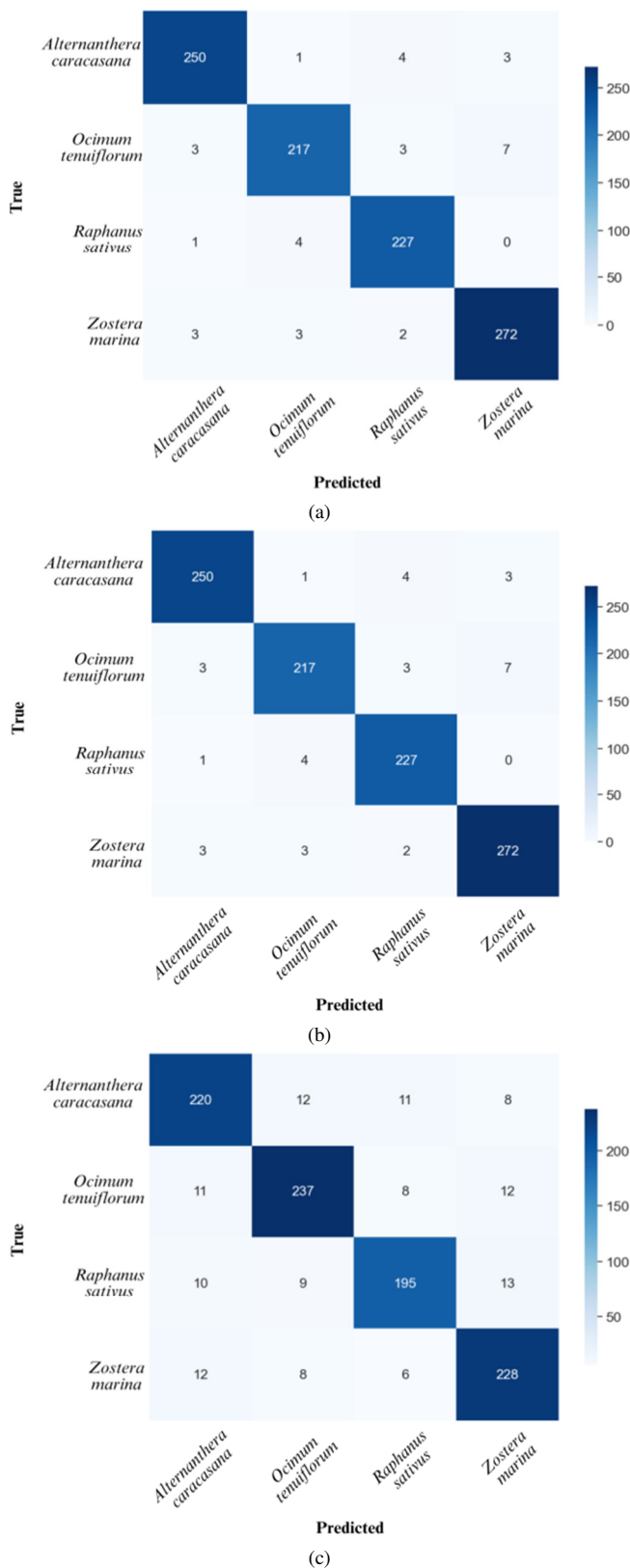


Fig. 3. Confusion matrix for: (a) MobileNetV2, (b) ResNet50, (c) VGG16.

The confusion matrix, as depicted in Figure 3, highlighted the classification performance, with MobileNetV2 exhibiting the lowest misclassification rate. MobileNetV2 achieved an accuracy of 96.60% and an F1 score of 96.57%, while ResNet50 recorded an accuracy of 95.00% with an F1 score of 87.96%. In contrast, VGG16 obtained an accuracy of 88.00%, but a relatively higher F1 score of 94.98%. Despite the high accuracy, challenges, such as the lighting variations, occlusions, and background noise, obscure the object details, affecting real-world deployment and requiring advanced data augmentation and preprocessing.

V. CONCLUSION

Accurate plant identification is essential in agriculture and environmental management for tasks, such as weed control, disease detection, and biodiversity monitoring. Traditional manual methods are labor-intensive and prone to error, underscoring the need for scalable, automated solutions. This study explores Deep Learning (DL) approaches, specifically Convolutional Neural Networks (CNNs), for plant classification, with a focus on identifying the weeds in chili crops. Using a self-collected dataset of four plant species from the DSLD CHEFT, Haveri, three CNN architectures, namely MobileNetV2, ResNet50, and VGG16, were trained with transfer learning and data augmentation, and were evaluated using metrics, such as accuracy, precision, recall, F1 score, and inference time. MobileNetV2 achieved the highest accuracy (96.6%) and demonstrated strong generalization and efficiency, outperforming the other models.

The study highlights MobileNetV2's suitability for edge deployment in agricultural settings due to its lightweight design. While limited by dataset scope and environmental variability, this research underscores the effectiveness of CNNs, particularly MobileNetV2, for practical plant classification. The integration of a Django-based web application further enhances the field usability, making this an effective solution for precision agriculture. Future work will focus on expanding the species coverage, improving robustness, and exploring advanced models, like Vision Transformers, Transfer learning, and domain-specific fine-tuning to further enhance performance, supporting applications in agriculture, ecology, and conservation. Optimizing these models for low-power hardware and integrating them into mobile or IoT platforms will enable scalable, accessible solutions for precision agriculture and environmental management.

This is a novel approach to plant classification as it uses the datasets collected from College of Horticultural Engineering and Food Technology (DSL D CHEFT), Devihosur, Haveri, with the study being limited to four species—*Alternanthera caracasana*, *Ocimum tenuiflorum*, *Raphanus sativus*, and *Zostera marina*. This work differs from the existing studies as it introduces a self-collected dataset focused on weeds in chili crop habitats—an area rarely addressed in prior research. Unlike earlier approaches using generic or large-scale datasets, this study evaluated and optimized lightweight CNN models, such as MobileNetV2, for real-time deployment on edge devices.

DATA AVAILABILITY

The dataset utilized in this study was sourced from the College of Horticultural Engineering and Food Technology (DSLDCHEFT), Devihosur, Haveri-581110, and encompasses four specific plant species: *Alternanthera caracasana*, *Ocimum tenuiflorum*, *Raphanus sativus*, and *Zostera marina*.

REFERENCES

- [1] A. J. Ishak, M. M. Mustafa, N. M. Tahir, and A. Hussain, "Weed detection system using support vector machine," in *2008 International Symposium on Information Theory and Its Applications*, Auckland, New Zealand, Dec. 2008, pp. 1–4, <https://doi.org/10.1109/isita.2008.4895454>.
- [2] G. H. Krishnan and T. Rajasenbagam, "A Comprehensive Survey for Weed Classification and Detection in Agriculture Lands," *Journal of Information Technology and Digital World*, vol. 3, no. 4, pp. 281–289, Apr. 2022, <https://doi.org/10.36548/jitdw.2021.4.004>.
- [3] Y. Sun, Y. Liu, G. Wang, and H. Zhang, "Deep Learning for Plant Identification in Natural Environment," *Computational Intelligence and Neuroscience*, vol. 2017, pp. 1–6, 2017, <https://doi.org/10.1155/2017/7361042>.
- [4] N. C. Kundur and P. B. Mallikarjuna, "Deep Convolutional Neural Network Architecture for Plant Seedling Classification," *Engineering, Technology & Applied Science Research*, vol. 12, no. 6, pp. 9464–9470, Dec. 2022, <https://doi.org/10.48084/etasr.5282>.
- [5] N. Makanapura, C. Sujatha, P. R. Patil, and P. Desai, "Classification of plant seedlings using deep convolutional neural network architectures," *Journal of Physics: Conference Series*, vol. 2161, no. 1, Jan. 2022, Art.no 012006, <https://doi.org/10.1088/1742-6596/2161/1/012006>.
- [6] X. Guo, "Automatic detection of tomato leaf disease using an adopted deep learning algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. 46, no. 4, pp. 7909–7921, Apr. 2024, <https://doi.org/10.3233/jifs-236905>.
- [7] A. Jafar, N. Bibi, R. A. Naqvi, A. Sadeghi-Niaraki, and D. Jeong, "Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations," *Frontiers in Plant Science*, vol. 15, Mar. 2024, <https://doi.org/10.3389/fpls.2024.1356260>.
- [8] L. G. Divyanth, D. S. Guru, P. Soni, R. Machavaram, M. Nadimi, and J. Paliwal, "Image-to-Image Translation-Based Data Augmentation for Improving Crop/Weed Classification Models for Precision Agriculture Applications," *Algorithms*, vol. 15, no. 11, Oct. 2022, Art. no. 401, <https://doi.org/10.3390/a15110401>.
- [9] F. A. Wichmann and R. Geirhos, "Are Deep Neural Networks Adequate Behavioral Models of Human Visual Perception?," *Annual Review of Vision Science*, vol. 9, no. 1, pp. 501–524, Sep. 2023, <https://doi.org/10.1146/annurev-vision-120522-031739>.
- [10] J. S. K. and A. S., "Transfer Learning Based Network Performance Comparison of the Pre-Trained Deep Neural Networks using MATLAB," *Open International Journal of Informatics*, vol. 10, no. 1, pp. 27–40 May 2022.
- [11] O. Vermesan and D. Marples, *Advancing Edge Artificial Intelligence: System Contexts*, 1st ed. New York: River Publishers, 2024.
- [12] K. C. Kiran, H. N. Rithin, K. Sathwik, L. S. Shreya, and B. T. Venkatesh Murthy, "Smart Autonomous Agriculture Robot for Multipurpose Farming Application," in *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*, Bengaluru, India, Aug. 2024, pp. 1–7, <https://doi.org/10.1109/nmitcon62075.2024.10699197>.
- [13] A. S. Paymode and V. B. Malode, "Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG," *Artificial Intelligence in Agriculture*, vol. 6, pp. 23–33, 2022, <https://doi.org/10.1016/j.aiaa.2021.12.002>.
- [14] N. Islam *et al.*, "Early Weed Detection Using Image Processing and Machine Learning Techniques in an Australian Chilli Farm," *Agriculture*, vol. 11, no. 5, Apr. 2021, Art. no. 387, <https://doi.org/10.3390/agriculture11050387>.
- [15] "Empirical Study of Multi-class Weed Classification Using Deep Learning Network Through Transfer Learning," in *Lecture Notes in Networks and Systems*, Singapore: Springer Nature Singapore, 2023, pp. 1–12.
- [16] G. Konnurmath and S. Chickerur, "GPU Shader Analysis and Power Optimization Model," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12925–12930, Feb. 2024, <https://doi.org/10.48084/etasr.6695>.