© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1717169

Machine Learning and Deep Learning for Plant Disease Classification and Detection

N. SRI JANAKI RAM[1], N. VENKATA SAI GOPI CHANDU[2], S. AJAY[3], REDDY VEERAMOHANARAO[4]
[1, 2, 3]*Department of Computer Science & Engineering (AI & ML), R.V.R & J.C College of Engineering, (Autonomous), Affiliated to Acharya Nagarjuna University, Guntur, Andhra Pradesh, India*
[4]*Assistant Professor, CSE (AI & ML), (Autonomous), Affiliated to Acharya Nagarjuna University, Guntur, Andhra Pradesh, India*

Abstract- Plant diseases cause significant losses in global agricultural output, and early, accurate identification remains a major challenge. This paper presents a unified deep learning pipeline that simultaneously localises and classifies plant diseases from leaf images. A YOLOv5s model performs region-level detection, while a transfer-learning-based ResNet50 classifier identifies the specific disease category. The system is trained on the PlantVillage dataset, which contains 54,306 images spanning 38 disease classes across 14 crop species. Preprocessing includes image resizing, normalisation, and a combination of standard and advanced augmentation strategies, namely CutMix and MixUp. ResNet50 is fine-tuned in two phases with Cosine Learning Rate decay, and Test Time Augmentation (TTA) is applied at inference for additional accuracy gains. The ResNet50 classifier achieves 99.1% validation accuracy with TTA, while the YOLOv5s detection model reaches a mAP@0.50 of 0.856. The macro-averaged precision, recall, and F1-score stand at 0.974, 0.968, and 0.971, respectively. The integrated pipeline demonstrates reliable end-to-end performance with 97.6% accuracy on 200 held-out images, offering a practical and scalable solution for precision agriculture.

Index Terms—Plant Disease Detection, Deep Learning, ResNet50, YOLOv5, Transfer Learning, Precision Agriculture, CutMix, MixUp, Test Time Augmentation, PlantVillage.

I. INTRODUCTION

Agriculture underpins food security and economic stability for much of the world, yet plant diseases continue to erode crop yields and quality, inflicting severe financial losses on farmers and national economies alike. Conventional disease identification relies on trained agronomists who physically inspect leaf samples—a process that is slow, expensive, and entirely unsuitable for the vast scale of modern commercial farms.

Recent advances in Artificial Intelligence, particularly Convolutional Neural Networks (CNNs), have opened a practical route to automated, real-time plant health monitoring. CNNs excel at extracting subtle visual features from images, and two complementary paradigms have emerged: image classification, which determines whether a leaf is healthy or diseased, and object detection, which additionally localises the affected region within the image. Most published systems address only one of these tasks; real-world deployment demands both.

This work bridges that gap by integrating YOLOv5s-based detection with ResNet50-based classification into a single forward-pass pipeline. The system is trained and validated on the publicly available PlantVillage dataset and is designed with practical deployment in mind, supporting use cases from handheld smartphone capture to web-based diagnostic portals.

II. RELATED WORK

A. Traditional Classification Approaches

Early automated plant disease systems relied on handcrafted features—colour histograms, texture descriptors, and shape statistics—fed into classifiers such as Support Vector Machines (SVM), k-Nearest Neighbours (KNN), and Random Forests. While these methods achieved moderate accuracy on well-curated datasets, they generalised poorly to the photometric variability encountered in the field.

B. Deep Learning-Based Classification

Mohanty et al. [2] demonstrated that a deep CNN trained on PlantVillage could recognise 26 diseases across 14 crops with over 99% accuracy under laboratory conditions. Subsequent work adopted pretrained models such as VGG16, InceptionV3, MobileNet, and ResNet50 via transfer learning, substantially reducing the training data requirement while maintaining high accuracy. ResNet50's residual learning mechanism is particularly well suited to this domain because skip connections allow gradients to flow through very deep networks without degradation.

C. Object Detection for Disease Localisation

Detection frameworks have evolved from computationally expensive sliding-window approaches to end-to-end deep networks. Faster R-CNN [18] introduced region proposal networks, and SSD [19] further improved speed. The YOLO family [4] treats detection as a single regression problem, enabling real-time throughput. Fuentes et al. [22] applied Faster R-CNN and SSD to tomato diseases under field conditions, while Balafas et al. [1] benchmarked multiple classifiers and detectors on the PlantDoc dataset, achieving a maximum classification accuracy of 61% and mAP of 39.3% with YOLOv5, highlighting the difficulty of generalising from laboratory to real-world imagery.

D. Limitations in Existing Work

Existing systems predominantly treat classification and detection as independent pipelines. Class imbalance, reliance on controlled datasets, and the absence of a unified detection-plus-classification output in a single inference pass remain open challenges that this work directly addresses.

III. METHODOLOGY

A. Dataset

The PlantVillage dataset [6] provides 54,306 RGB leaf images covering 38 disease and healthy-plant classes across 14 crop species including tomato, potato, corn, apple, and grape. The dataset was split 80:20 into training and validation subsets using stratified sampling with a fixed random seed of 42, yielding 43,456 training and 10,850 validation images. Class weights were computed from the training distribution and baked into the data generator to mitigate the moderate class imbalance present in the dataset.

B. Preprocessing and Augmentation

All images are resized to 224×224 pixels for ResNet50 and 640×640 pixels for YOLOv5s. Pixel values are normalised to [0, 1] by dividing by 255. During training, standard augmentations—random horizontal and vertical flips, rotation up to 40°, zoom (0.25), brightness variation ([0.6, 1.4]), channel shift (±20), and shear (0.15)—are applied. CutMix and MixUp augmentation are additionally incorporated at a batch level, replacing standard batches with probability 0.4 and 0.3, respectively. These strategies help the model learn to classify partially occluded or mixed-appearance leaves, significantly improving generalisation.

C. ResNet50 Architecture and Training

The ResNet50 backbone, pretrained on ImageNet [11], is used as a feature extractor. A custom classifier head is appended: Global Average Pooling, Batch Normalisation, a 512-neuron Dense layer, Dropout (0.4), a 256-neuron Dense layer (implicit in the 512→1 pipeline), Dropout (0.3), and a 38-class Softmax output. Training proceeds in two phases. Phase 1 freezes all backbone layers and trains only the classifier head for 12 epochs with Adam at a learning rate of 1e-3, reaching a validation accuracy of 87.1%. Phase 2 unfreezes all layers and fine-tunes for up to 30 additional epochs using Cosine Learning Rate decay starting at 1e-5, raising validation accuracy to 98.7%. The total training spans 22 epochs due to early stopping. Test Time Augmentation (five augmented variants per image, softmax scores averaged) pushes final validation accuracy to 99.1%.

D. YOLOv5s Architecture and Training

YOLOv5s is initialised from COCO-pretrained weights. Because PlantVillage contains only classification labels, full-image bounding boxes are automatically generated and stored in YOLO format. A data.yaml file enumerates all 38 classes. The model is trained for a maximum of 50 epochs (early stopping patience of 15) with mosaic augmentation, MixUp (0.15), geometric jitter, and a batch size of 16 at 640×640 resolution. Training converges at epoch 41.

E. Unified Inference Pipeline

At inference, YOLOv5s first processes the input image and returns bounding boxes for detected disease regions with a confidence threshold of 0.5. Each cropped region is resized to 224×224 and classified by ResNet50 with TTA. If no region is detected above the threshold, the full image is passed directly to ResNet50 as a fallback. The output is an annotated image with bounding boxes, disease labels, and confidence scores.

IV. RESULTS AND DISCUSSION

A. ResNet50 Classification Performance

After full fine-tuning and TTA, the ResNet50 model achieves 99.1% validation accuracy on the PlantVillage test set. The macro-averaged precision across all 38 classes is 0.974, macro recall is 0.968, and the macro F1-score is 0.971, confirming consistently strong performance across both majority and minority classes. The class-wise analysis shows that Blueberry Healthy, Tomato Leaf Mold, and Corn Common Rust achieve near-perfect F1-scores of 1.00, 0.99, and 0.99, respectively, owing to their visually distinctive symptoms. The most challenging classes are Corn Gray Leaf Spot (F1: 0.91) and Tomato Spider Mites (F1: 0.91), which share subtle visual features with other early-stage disease conditions.

B. YOLOv5s Detection Performance

The YOLOv5s model achieves a mAP@0.50 of 0.856 and mAP@0.50:0.95 of 0.612 on the validation set. Overall detection precision is 0.874 and recall is 0.831. The best-performing detection classes are Tomato Leaf Mold (AP: 0.93), Corn Common Rust (AP: 0.91), and Tomato Bacterial Spot (AP: 0.90). Corn Gray Leaf Spot records the lowest AP at 0.69, consistent with its classification difficulty.

C. Unified Pipeline Evaluation

The end-to-end pipeline is evaluated on 200 randomly sampled validation images. The combined YOLOv5s + ResNet50 system achieves an overall accuracy of 97.6%. In the 12 cases where YOLOv5s fails to detect a region above the confidence threshold, the full-image ResNet50 fallback correctly classifies the leaf, ensuring no prediction is withheld. This graceful degradation mechanism is critical for deployment robustness.

D. Comparison with Baseline

Balafas et al. [1], the primary benchmark for this work, report a peak classification accuracy of 61% on PlantDoc and a YOLOv5 mAP of 39.3% on PlantDoc. The higher metrics reported here (99.1% accuracy, 85.6% mAP@0.50) reflect both the more uniform nature of PlantVillage images and the comprehensive training strategy employed, including two-phase fine-tuning, CutMix/MixUp augmentation, class weighting, and TTA. Crucially, the system presented here goes beyond prior work by producing simultaneous localisation and classification in a single pass, a capability absent from the baseline.

V. CONCLUSION

This paper presents a unified deep learning framework for plant disease detection and classification that combines YOLOv5s object detection with ResNet50 transfer learning. The system achieves 99.1% classification accuracy with TTA and 85.6% mAP@0.50 for detection on the PlantVillage dataset, and 97.6% end-to-end accuracy on held-out images. The integration of detection and classification into a single inference pipeline, supported by advanced augmentation and two-phase training, represents a practical step toward scalable, field-deployable agricultural AI.

Future work will focus on adapting the system to real-world field datasets such as PlantDoc through domain adaptation, obtaining precise lesion-level bounding box annotations, supporting multi-label inference, deploying lightweight variants on mobile and edge devices, and incorporating hyperspectral or IoT sensor data for pre-symptomatic disease detection.

VI. ACKNOWLEDGMENT

The authors sincerely thank Mr. Reddy Veeramohanarao for his continuous guidance and technical insights throughout this project. Gratitude is also extended to Dr. G. Rama Mohan Babu, Head of the Department of CSE (AI & ML), and Dr. K. Srinivas, Principal of R.V.R & J.C College of Engineering, Guntur, for providing a supportive research environment and necessary computational resources.

REFERENCES

[1] V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, "Machine Learning and Deep Learning for Plant Disease Classification and Detection," IEEE Access, vol. 11, pp. 114352–114377, 2023.

[2] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," Front. Plant Sci., vol. 7, Art. 1419, 2016.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1717169

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE CVPR, 2016, pp. 770–778.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE CVPR, 2016, pp. 779–788.

[5] G. Jocher, YOLOv5 by Ultralytics, v3.1, 2022. [Online]. Available: https://github.com/ultralytics/yolov5

[6] D. P. Hughes and M. Salathé, "An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics," arXiv:1511.08060, 2015.

[7] D. Singh et al., "PlantDoc: A Dataset for Visual Plant Disease Detection," in Proc. ACM IKDD CoDS-COMAD, 2020, pp. 249–253.

[8] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556, 2014.

[9] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," arXiv:1412.6980, 2014.

[10] S. Yun et al., "CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features," in Proc. IEEE ICCV, 2019, pp. 6023–6032.

[11] H. Zhang et al., "MixUp: Beyond Empirical Risk Minimization," in ICLR, 2018.

[12] A. Fuentes, S. Yoon, S. Kim, and D. Park, "A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition," Sensors, vol. 17, no. 9, Art. 2022, 2017.

[13] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in NeurIPS, 2015, pp. 91–99.

[14] W. Liu et al., "SSD: Single Shot MultiBox Detector," in Proc. ECCV, 2016, pp. 21–37.

[15] K. P. Ferentinos, "Deep Learning Models for Plant Disease Detection and Diagnosis," Comput. Electron. Agric., vol. 145, pp. 311–318, 2018.