Amrita Kumari, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

# Lightweight MobileNetV2–Attention Framework for Maize Leaf Disease Classification under Semi-Realistic Conditions

**Amrita Kumari, Ankush Kumar, and Omvir Singh**
Department of Computer Science and Engineering, Institute of Technology Roorkee

**Abstract- Leaf diseases significantly affect the Maize productivity. Early and timely detection of these diseases help in reducing crop loss. It also ensures sustainable agricultural practices. Traditional examination of leaves by means of manual methods are protracted, subjective, and difficult to scale, especially when working on large agricultural fields. In the current work, a lightweight hybrid deep learning (DL) framework for maize leaf disease detection has been developed. The proposed framework is based on MobileNetV2 integrated with a spatial attention mechanism. The MobileNetV2 backbone allows effective feature extraction and the attention module helps in improving the model's ability to focus on the regions affected by disease. To enhance generalization and reduce overfitting, several training strategies are employed. These include MixUp augmentation, exponential moving average (EMA) and label smoothing. The evaluation of the model is performed on datasets with semi-realistic conditions with moderate background and illumination variability. A peak accuracy of approximately 97.45% was achieved by the proposed approach as confirmed by the evaluation results and exhibits stable convergence. These findings indicate that lightweight architectures, when synergistically combined with attention mechanisms and robust training strategies, can attain high accuracy. The architecture also preserves computational efficiency. This makes them suitable for real-world agricultural deployment.**

**Keywords: Maize leaf disease detection, Deep learning, MobileNetV2, Spatial attention.**

## I. INTRODUCTION

Agriculture is a highly demanding industry that makes a very crucial contribution towards sustaining the ever-increasing global population [1, 2]. The increasing population requires large crop production. Maize is a common staple crop because of its wide range of applications in food, industrial products, and livestock food. However, it was discovered that the leaf diseases are significant in hindering the production of maize. These diseases are Bacterial Leaf Streak, Northern Leaf Blight, Common Rust and Gray Leaf Spot, which decrease the photosynthetic efficiency and eventually lead to yield losses [3, 4]. This is why proper disease control is crucial to maintain the quality of crop and its productivity such that the demand could be fulfilled. Conventionally, the process of detecting the plant diseases is done by manual inspection by experts in the agriculture field. This method can be efficient, however, it is extremely consuming in terms of time, labor-intensive and in most cases subjective. Such means to scale up to large agricultural fields might not be possible, particularly in areas with limited supply of experts. To solve these problems, automated, efficient and scalable systems of disease detection are needed.

Deep learning (DL) methods have been very successful over the years in plant disease classification [3, 5]. In particular, the Convolutional Neural Networks (CNNs) have demonstrated strong performance and this could be attributed to the ability to learn hierarchical representations of image data [3, 6, 7]. However, typical CNN models have been identified to focus mostly on local feature extraction, which makes them inadequate to focus on important context of disease patterns. With the introduction of attention mechanisms, the model has been able to work around the shortcomings by focusing on the

relevant parts of the input image, thus improving its performance and interpretability. Despite many developments, there are still many daunting challenges. Most of the available studies are mostly based on controlled datasets, e.g. the PlantVillage which do not effectively represent field conditions in most of the situations, and perform more poorly when deployed. This is primarily because of differences in lighting, complexity of background, and occlusions [8]. In addition to these, most of the current methods focus on the classification accuracy without sufficient attention to computational efficiency or practical deployment.

To address these issues, in this work we propose that to ensure good generalization, a lightweight architecture with MobileNetV2 and spatial attention mechanism in combination with MixUp augmentation, label smoothing, and EMA be implemented. The proposed method attempts to find a trade-off between the classification accuracy and the cost of computation hence, rendering it suitable in resource constrained implementation. Testing is done under semi-realistic scenarios for the evaluation to be as close as possible to reality. This paper emphasis on the importance of a dependable performance in real scenarios, other than focusing on accuracy in isolated environments.

## II. LITERATURE REVIEW

DL has become a leading method of detecting maize leaf disease because of its capability to automatically learn discriminative features based on image data. The trend in research in this area has changed over the years from the traditional CNN-based models to more advanced hybrid architectures (including attention mechanisms and lightweight models).

The initial studies of maize leaf disease detection were associated primarily with CNNs to classify images. These models showed how practical it was to recognize diseases automatically in real-field environments and how they could be useful not only in the controlled experiments of the laboratory but also under real field conditions [5]. In subsequent research, transfer learning using pre-trained architectures was used to improve performance. ResNet, VGG, Inception, DenseNet, and EfficientNet models were extensively used due to their capability to capture complex hierarchical features. These methods were able to classify with high accuracy. The obtained accuracy was commonly over 95% and CNNs was considered as the credible way to detect the plant diseases [6, 7]. Even though CNN-based models have been known to be highly accurate, they often require considerable computational resources. In order to overcome the limitations of conventional CNNs approaches, recent research started utilizing hybrid architectures [9] and attention mechanisms that helps in improving the feature representation. As an example, the models containing the blocks of Squeeze-and-Excitation (SE) blocks have been reported to have improved classification due to feature-importance per channel [10]. In the same way, both fusion-based architectures that incorporate models such as Xception and DenseNet have been proposed to enhance feature extraction [11]. The combination of segmentation and detection methods has also been used in some studies to enhance the localization of diseases. Other systems, such as PSPNet as well as CNN classifiers and object detection systems (such as Faster R-CNN) have demonstrated promising results in identifying and localizing diseased regions [12, 13].

A number of lightweight models have been demonstrated to perform competitively with lower memory and computational needs [14]. In such a way, small architectures that adopt pruning, quantization, and recalibration of the features give high accuracy with minimum size. This renders them appropriate when it comes to edge devices [15, 16]. Compression methods such as knowledge distillation and parameter pruning have been examined to minimize the complexity of the model without affecting the performance [17, 18].

Lightweight models can greatly reduce the cost of computation, but they may not be able to achieve high classification accuracy in complex environmental conditions. Therefore, the lightweight architectures require to be enhanced with more mechanisms towards enhancing feature representation.

Though improvement has been made, there are still some limitations in the detection of maize leaf disease. First, there are numerous studies that are based on a very tightly controlled dataset, which restricts the generalization to real-world conditions. Second, deep models and hybrid models are highly accurate, but expensive to compute. This renders them unsuitable to be deployed to edges. Third, lightweight models are more efficient, but often have worse feature representations and classification accuracy. Moreover, to gain high accuracy, efficiency, and robustness, little research has been conducted on the synergistic integration of lightweight architectures, attention mechanisms as well as advanced training strategies.

In order to fill these specified gaps, as well as to obtain improved performance under semi-realistic conditions, the current study suggests a new hybrid framework, which will combine MobileNetV2 with a spatial attention mechanism and effective training strategies.

## III. METHODOLOGY

### 3.1 Overview of Proposed Framework

The proposed system is a lightweight hybrid DL model to detect maize leaf disease. The framework is combined with an attention mechanism to improve the feature representation. This has been done in order to get a balance between classification and computational efficiency that is needed to deploy the algorithm in the real world.

The framework consists of the following stages:

- Image acquisition (dataset images with changing conditions)
- Preprocessing and data augmentation
- Feature extraction using a lightweight CNN
- Spatial attention-based feature refinement
- Classification using a softmax-based output layer
- Model optimization for efficient deployment

### 3.2 Data Acquisition and Preprocessing

#### 3.2.1 Data Sources

The dataset consists of the popular PlantVillage [19] and PlantDoc datasets [20]. Fig. 1 shows the maize leaf images with illness along with the healthy leaf.

3.2.2 Preprocessing Steps

To make all input images compatible with the model architecture, all input images are resized to 224 x 224 pixels. The pixel values were normalized so that convergence was enhanced during training.

(a) (b) (c) (d)

Fig. 1. Images of Maize leaf used in this work. (a) Healthy (b) Common Rust (c) Gray Leaf Spot (d) Blight.

#### 3.2.3 Data Augmentation

Data augmentation methods are used in training to enhance generalization. These include random rotation (±15°30°), horizontal and vertical flipping, changing brightness, random cropping and color jittering. In training, MixUp augmentation is applied to produce interpolated samples. This assists to enhance robustness and curbs overfitting.

### 3.3 Feature Extraction Using MobileNetV2

MobileNet V2 is used as the base to extract features due to its lightweight structure and performance.

### 3.4 Spatial Attention Module

To further improve the model in the aspect of focusing on regions that are relevant to the disease, the spatial attention module is added after the feature extraction step.

### 3.5 Classification Head

The refined feature is sent through a classification head comprising of global average pooling (GAP), which is followed by fully connected layer and a softmax activation function.

One of the following classes is predicted by the model:

- Healthy
- Northern Leaf Blight
- Common Rust
- Gray Leaf Spot
- Bacterial Leaf Streak

This approach helps in reducing overconfidence in predictions and improves the model ability to generalize across unseen data.

### 3.6 Training and Optimization Strategy

To achieve stable training and improve model performance and generalization, several optimization strategies are utilized, as mentioned below

- MixUp augmentation
- Label smoothing
- Exponential Moving Average (EMA)
- Two-stage training (freeze → fine-tune)

### 3.7 Lightweight Optimization for Deployment

MobileNetV2 is designed to keep the proposed model lightweight and computationally efficient, making it suitable to be deployed to resource-constrained systems, including mobile devices and embedded systems.

Training Configuration: Optimizer: AdamW Learning Rate: $1\times10^{-3}$
Batch Size: 32
Epochs: 30
Train/Test Split: 80/20

## IV. RESULTS AND DISCUSSION

### 4.1 Experimental Setup and Performance Metrics

The proposed model was evaluated using maize leaf images from both a controlled dataset (PlantVillage) and real-field images. This allowed assessment under both laboratory and practical conditions. The dataset was divided into training (80%) and testing (20%) subsets. Model performance was measured using standard classification metrics that includes accuracy, precision, recall, and F1-score. These metrics provide a balanced view of overall correctness and class-wise prediction behavior.
The evaluation metrics are defined as:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ score = 2\,\frac{Precision \cdot Recall}{Precision+Recall} \quad (4)$$

where TP, FP, FN, and TN denote true positives, false positives, false negatives, and true negatives, respectively.

### 4.2 Quantitative Results

The proposed model achieves 96.03% accuracy and high precision, recall, and F1-scores (as depicted in Table 1). Peak accuracy of 97.45% was achieved. As classification accuracy saturates, distinguishing visually similar disease classes becomes challenging, even small gains reflect meaningful enhancements in feature representation and learning capability. Additionally, a properly optimized lightweight architecture can compete with deeper, more computationally demanding models.

### 4.3 Analysis and Discussion

The model has a high classification accuracy and a few misclassifications. Fig. 2 shows the confusion matrix. Mistakes include mostly similar classes of diseases, including Common Rust and Gray Leaf

Spot. This confusion can be expected because of the similarity of textual and colour characteristics, particularly at the nascent stages.

A spatial attention module is used to improve the performance by ensuring that the model focuses on relevant parts of the image. This is especially beneficial in situations where additional background or lighting differences may hinder feature extraction. The model has a higher discriminatory capacity against subtle visual patterns, by focusing on the disease-affected areas.

Table 1 Classification Report

| | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| Blight | 0.94 | 0.93 | 0.93 | 172 |
| Grey Leaf Spot | 0.86 | 0.89 | 0.87 | 87 |
| Common Rust | 0.99 | 0.98 | 0.99 | 196 |
| Healthy | 1.00 | 1.00 | 1.00 | 175 |
| | | | | |
| Accuracy | | | 0.96 | 630 |
| Macro Avg | 0.95 | 0.95 | 0.95 | 630 |
| Weighted Avg | 0.96 | 0.96 | 0.96 | 630 |

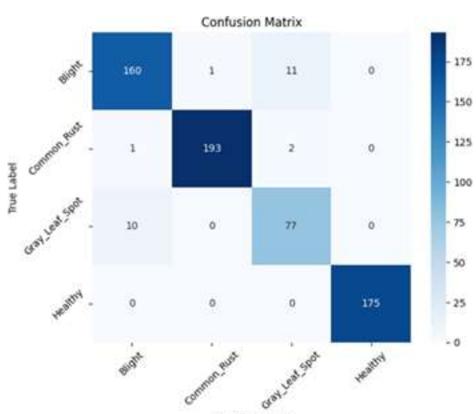


Fig. 2. Confusion Matrix

## V. CONCLUSION

This paper presents a lightweight deep learning model to classify maize leaf diseases, which involves a MobileNetV2 backbone and a spatial attention mechanism. The method provides a tradeoff between classification performance and computational efficiency to implement it in practice. Experimental evidence indicates that the model attains a reasonably good mean accuracy of about 96.03% in the semi-realistic conditions. The attention mechanisms allow the model to concentrate on disease-relevant regions, whereas training strategies such as MixUp augmentation, label smoothing, and exponential moving average can help the model to focus on disease-relevant regions and enhance generalization and stability. This is essential to the applications of agriculture that have resource limitations and real-world variability. Although these are promising findings, a small performance degradation when using field-like conditions suggests that additional testing on more varied data needs to be done. Future research will apply the model to bigger and more diverse real-world data and investigate its implementation on mobile and edge devices to detect diseases in real time.

## REFERENCES

1. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.
2. Kumari, A., Saini, A., & Kumar, A. (2024). A review on Agricultural Practices for Long-term Sustainability. In book: Engineering a Sustainable Future Role of Science and Technology for Achieving SDGs, Walnut Publication, Volume-I, 86-106.
3. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and electronics in agriculture, 145, 311-318.
4. Kumari, A., Anand, N., Saini, A., Kumar, A., & Sharma, S (2026). Predictive Modeling of Maize Leaf Diseases Using Machine Learning Techniques, 21(S2), 233-239.

5. DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E. L., Yosinski, J., Gore, M. A., ... & Lipson, H. (2017). Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. Phytopathology, 107(11), 1426-1432.
6. Fraiwan, M., Faouri, E., & Khasawneh, N. (2022). Classification of corn diseases from leaf images using deep transfer learning. Plants, 11(20), 2668.
7. Ma, Z., Wang, Y., Zhang, T., Wang, H., Jia, Y., Gao, R., & Su, Z. (2022). Maize leaf disease identification using deep transfer convolutional neural networks. International Journal of Agricultural and Biological Engineering, 15(5), 187-195.
8. O'Halloran, T., Obaido, G., Otegbade, B., & Mienye, I. D. (2024). A deep learning approach for Maize Lethal Necrosis and Maize Streak Virus disease detection. Machine Learning with Applications, 16, 100556.
9. Marcelino, J., Sybingco, E., Concepcion, R., Bandala, A., & Demdam, R. (2024, November). Classification of Corn Diseases From Leaf Images Using Deep Convolutional Neural Networks. In 2024 IEEE 16th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM) (pp. 1-6). IEEE.
10. Siam, A. S. M., Hossain, A., Hossain, R. B., & Rahman, M. M. (2024, March). SE-VGG16 maizenet: Maize disease classification using deep learning and squeeze and excitation attention networks. In 2024 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 1-6). IEEE.
11. Mishra, V., Naik, N. S., Kumar, S., Alsamhi, S. H., Saif, A., & Curry, E. (2023, September). Maize plant disease prediction of UAV images for precision agriculture using fusion of multimodal. In 2023 3rd International Conference on Computing and Information Technology (ICCIT) (pp. 353-358). IEEE.
12. Bachhal, P., Kukreja, V., & Ahuja, S. (2023). Real-time disease detection system for maize plants using deep convolutional neural networks. International Journal of Computing and Digital Systems, 14(1), 10263-10275.
13. Reddy, B. R. S., Madhavi, G. B., Lakshmi, C. S., Nagendra, K. V., & Sridevi, R. (2021). Detection of disease in maize plant using deep learning. CABI Digital Library.
14. Shinde, T. (2025). An efficient and scalable framework for lightweight crop disease recognition in low-resource settings. In Proceedings of the Computer Vision and Pattern Recognition Conference (pp. 5534-5541)
15. Alam, M. N. A., Fahim, R. H., Sayed, M. K. B., & Shorna, S. A. (2025, December). CornNetLite: An Ultralight CNN for Corn Leaf Disease Classification in Low-Resource Agricultural Environments. In 2025 IEEE 7th International Conference on Sustainable Technologies for Industry 5.0 (STI) (pp. 1-6). IEEE.
16. Weloday, F., & Su, J. (2026). LWMSCNN-SE: A Lightweight Multi-Scale Network for Efficient Maize Disease Classification on Edge Devices. arXiv preprint arXiv:2601.07957.
17. Abubakar, M., Ibrahim, Y., Ajayi, O. O., & Saminu, S. S. (2026). A Lightweight Maize Leaf Disease Recognition Using PCA-Compressed MobileNetV2 Features and RBF-SVM. Journal of Computing Theories and Applications, 3(3), 334-348.
18. Wang, R., Zhang, W., Ding, J., Xia, M., Wang, M., Rao, Y., & Jiang, Z. (2021). Deep neural network compression for plant disease recognition. Symmetry, 13(10), 1769.
19. https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset
20. https://github.com/pratikkayal/PlantDoc-Dataset