International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XXI): April - June 2025
ISSN 2394 - 7780

GREEN SCAN: AUTOMATED LEAF DISEASE DETECTION USING IMAGE PROCESSING AND MACHINE LEARNING

[1]Md Sahil Zahir, [2]Mohd Ahad Khan, [3]Saurav Kumar, [4]Shahzeb Khan and [5]Dr.Pooja Kapoor
[1, 2, 3, 4]Department of (CSE), Mangalmay Institute of Engineering and Technology, Greater Noida, Uttar Pradesh, India
[5]Research Coordinator, MIET, Greater Noida, Uttar Pradesh, India

ABSTRACT

Leaf disease detection is critical for agricultural productivity and food security. This paper presents Green Scan, an automated system for diagnosing plant leaf diseases using image processing and convolutional neural networks CNNs. The approach involves preprocessing high-resolution leaf images, extracting features via a sequential CNN architecture, and classifying diseases with 88.25% accuracy. Evaluated on a dataset of 20,638 images spanning 15 disease classes (e.g., Bacterial Spot, Early Blight), the system demonstrates robustness and scalability. A user-friendly Gradio web interface enables real-time diagnosis, offering farmers a cost-effective alternative to manual inspection. Limitations and future improvements, including data augmentation and IoT integration, are discussed.

Keywords* — *Leaf disease detection, image processing, convolutional neural network CNN, agriculture, precision farming

I. INTRODUCTION

Plant diseases cause significant crop yield losses globally, with manual diagnosis being time consuming and error-prone. Automated systems leveraging computer vision and machine learning offer scalable solutions for early detection. Green Scan addresses this need by combining image preprocessing, feature extraction, and classification into an end-to-end pipeline

KEY CONTRIBUTIONS

1. A CNN-based model achieving 88.25% accuracy on a diverse leaf image dataset.
2. Integration with a Gradio web interface for real-time farmer use.
3. Analysis of limitations and proposed enhancements for future work.

II. LITERATURE SURVEY

Prior work in leaf disease detection has explored various methodologies:

- **Prakash et al.** used K-means clustering and SVM for segmentation and classification.
- **Indhumathi et al.** employed KMedoid clustering with Random Forests, achieving 92% accuracy on limited datasets.
- **Sardogan et al.** applied CNNs with Learning Vector Quantization LVQ) for tomato disease detection.

While these methods show promise, challenges persist in generalizability and real-time deployment. Green Scan improves upon existing approaches by leveraging transfer learning and a modular architecture optimized for scalability.

III. METHODOLOGY

A. Dataset and Preprocessing

- **Dataset:** 20,638 images from PlantVillage, covering 15 disease classes (e.g., Pepper Bacterial Spot, Potato Early Blight).
- **Splitting**: 80% training 16,511 images), 20% validation 4,127 images).
- **Preprocessing:** Resizing to 180 180 pixels, normalization RGB values scaled to RGF values scaled to)

B. Model Architecture

- A sequential CNN with the following layers:
- Conv2D 16 filters, 3 3 kernel, ReLU activation).
- MaxPooling2D 2 2 pool size).

- Repeated Conv2DMaxPooling blocks 32 and 64 filters).
- Flatten and Dense layers 128 neurons, ReLU) for classification.

```
model = Sequential([
layers.Rescaling(1./255, input_shape=(180, 180, 3)),
layers.Conv2D(16, 3, padding='same', activation='relu'), layers.MaxPooling2D(),
# ... additional layers
layers.Dense(15, activation='softmax')
])
```

C. Training

- **Optimizer:** Adam**.**
- **Epochs:** 5.
- **Batch Size:** 32.

IV. RESULTS AND DISCUSSION

A. Performance Metrics

- **Accuracy:** 88.25% on validation data.
- **Confusion Matrix:** Misclassification observed between visually similar diseases (e.g., Tomato Leaf Mold vs. Late Blight, Fig. 16.

B. Case Studies

- **Correct Predictions**: Pepper Bacterial Spot Fig. 10, Potato Early Blight Fig. 11 .
- **Incorrect Prediction**: Tomato Leaf Mold misclassified as Late Blight Fig. 16.

C. Limitations and Improvements

- **Data Augmentation**: Rotations, flips, and brightness adjustments to enhance generalization.
- **Transfer Learning**: Pretrained models (e.g., ResNet50) to improve feature extraction.
- **IoT Integration**: Deploying models on edge devices for real-time field monitoring.

IV. CONCLUSION

Green Scan demonstrates the feasibility of automated leaf disease detection using CNNs, achieving 88.25% accuracy on a 15-class dataset. Future work will focus on expanding the dataset, refining hyperparameters, and integrating drone-based imaging for large-scale agricultural monitoring. This system has the potential to reduce crop losses, minimize pesticide use, and enhance global food security.

REFERENCES

[1] Prakash, Saraswathy, Ramalakshmi, Mangaleshwari, Kaviya (2017). Detection of Leaf Diseases Classification doi:10.1109/ICIIECS.2017.8275915, IEEE. using Digital Image Processing,

[2] Indumathi, Saagri, Thejuswini, Swarnareka (2019). Leaf Disease Detection and Fertilisation Suggestion, doi:10.1109/ICSCAN.2019.8878781, IEEE.

[3] Anand, Aravinth, Veni (2016). An Application of image processing techniques for Detection of Diseases on Brinjal Leaves Using K-Means Clustering Method, doi:10.1109/ICRTIT.2016.7569531,IEEE.

[4] Kumar, Prasad, Srilekha, Suman, Rao, Krishna (2020). Leaf Disease Detection and Classification based on Machine Learning, doi: 10.1109/ICSTCEE49637.2020.9277379, IEEE.

[5] Kumari, Prasad, Mounika (2019). Leaf Disease Detection: Feature Extraction with K=Means clustering and Classification with ANN, doi: 10.1109/ICCMC.2019.8819750, IEEE.

[6] Singh, Varsha, Misra (2015). Detection of unhealthy region of plant leaves using Image Processing and Genetic Algorithm, doi: 10.1109/ICACEA.2015.7164858, IEEE.

ISSN 2394 - 7780

[7] Sardogan, Tuncer, Ozen (2020). Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm, doi: 10.1109/UBMK.2018.8566635, IEEE.

[8] Jiang, Chen, He, Liang (2019). Real-Time Detection of Apple Leaf diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks, doi: 10.1109/ACCESS.2019.2914929, IEEE.

[9] Zhang, Song, Zhang (2020). Deep Learning Based Object Detection Improvement for Tomato Disease, doi: 10.1109/ACCESS.2020.2982456, IEEE.

[10] Sun, Yang, He, Wu (2020). Northern Maize Leaf Blight Detection Under Complex Field Environment Based on Deep Learning, doi: 10.1109/ACCESS.2020.2973658, IEEE. Online. Available: https://plantvillage.org

[11] Kaur chitranjanjit, kapoor pooja, kaur Gurjeet(2023), “image recognition(soil feature extraction)using Metaheuristic technique and artificial neural network to find optimal output.Eur. Chem. Bull.2023(special issue 6).

[12] Maheshwari Chanana shalu, Kapoor pooja,kaur chitranjanjit(2023),”Data mining techniques adopted by google: A study.: Empirical Economics Letters,22(special issue 2).