ISSN [Online]: 2583-2654

AI-Powered Plant Disease Detection and Cure Prediction System

[1] Punitha F, [2]Prashanth Kumar G, [3] Shilpa Shree V, [4] Rutik Balu Kanbarkar, [5]H Shivashanth
Department of CSE, Dr. T. Thimmaiah Institute of Technology
KGF, India

Abstract: Plant diseases can have profound effects on the powered by artificial intelligence (AI) and computer vision. ***economy, impacting both local and global scales. These diseases can lead to substantial losses in agricultural*** Recent advancements in machine learning (ML) and deep ***productivity, affecting crop yields and quality. In this context,*** learning (DL)—particularly Convolutional Neural Networks ***deep learning algorithms are widely acknowledged as effective*** (CNNs)—have enabled highly accurate image-based ***solutions. However, the use of these black-box approaches*** classification of plant diseases. Models such as VGG16, ***raises concerns about trust in interpreting and validating the*** ResNet, and EfficientNet have demonstrated superior ***decisions generated by the models. This study proposes an*** performance in identifying diseases from leaf images. ***explainable artificial intelligence (XAI) based plant disease*** However, despite their high accuracy, these models often ***classification system to classify and identify distinct ailments*** suffer from lack of transparency and explainability. They are ***with improved accuracy. The system correctly identifies 38*** considered "black-box" systems, making it difficult for ***different plant diseases with accuracy, precision, and recall as*** end-users—especially farmers and agricultural experts—to ***99.69%, 98.27%, and 98.26%, respectively. These predictions*** understand and trust the model's decisions, which is crucial ***are subjected to additional analysis employing the local*** when such tools guide critical farming actions.

Automated plant disease detection systems, especially those ***interpretable model agnostic explanations (LIME) framework to produce visual explanations aligning with prior beliefs and*** To address this critical limitation, the field of Explainable ***adhering to established best practices in explanations. This*** Artificial Intelligence (XAI) has emerged, providing tools to ***system will serve as a promising avenue for revolutionizing*** interpret and visualize the inner workings of complex ML ***disease detection, fostering informed decision-making, and*** models. Techniques like Local Interpretable Model-Agnostic ***ultimately contributing to global food security.***

Keywords: Plant Disease Classification, deep learning, explainable artificial intelligence, prediction model, Agricultural Diagnosis, Explainable Artificial Intelligence (XAI).

I. INTRODUCTION

Agriculture plays a pivotal role in sustaining the global economy and ensuring food security. However, plant diseases remain one of the most critical threats to agricultural productivity, often resulting in significant yield losses and economic damage. According to the Food and Agriculture Organization (FAO), plant diseases are responsible for annual global losses exceeding $220 billion, underscoring the urgent need for reliable and scalable detection methods.

Traditionally, plant disease diagnosis relies on manual inspection by trained agronomists or pathologists. While effective to some extent, these manual approaches are time-consuming, subjective, and impractical for large-scale farms. Moreover, visible symptoms may appear only after the disease has progressed, by which time containment is more difficult and costly. This has motivated the development of Explanations (LIME) and SHAP help demystify the decision-making process by highlighting which features or regions of an input image influenced a model's prediction. This transparency fosters trust and promotes the adoption of AI-based solutions in real-world agriculture,

ISSN [Online]: 2583-2654

particularly in low-resource environments where expert support may be unavailable.

In this work, we propose an integrated system for plant disease detection that combines the power of EfficientNetB0, a lightweight and high-performance CNN architecture, with the LIME explanation framework. Our system is trained on a large, diverse dataset covering 38 disease classes across 14 plant species, collected from real-world and benchmark datasets including Kaggle. By leveraging transfer learning, we achieve high classification performance while minimizing computational overhead.

In addition to achieving high accuracy, our approach offers interpretable outputs that visually indicate the affected areas of the leaf, making it a valuable tool for both expert users and farmers. The inclusion of XAI enables stakeholders to validate the reasoning behind predictions, leading to more informed and confident decision-making in disease management.

To further bridge the gap between research and field application, we develop a mobile-friendly solution named PlantCare, allowing users to capture or upload plant images using their smartphones and receive instant diagnostic feedback. The system not only identifies the disease but also provides visual heatmaps, treatment suggestions, and confidence scores—empowering users with practical insights and early intervention capabilities.

II. METHODOLOGY

This study presents a comprehensive approach to plant disease classification by integrating a high-performance deep learning model with explainable artificial intelligence (XAI). The methodology consists of four main stages: (A) Data Pre-Processing, (B) Feature Extraction using EfficientNetB0, (C) Model Training and Validation, and (D) Prediction Explainability using LIME. Together, these components ensure accurate, efficient, and interpretable plant disease detection suitable for real-world deployment.

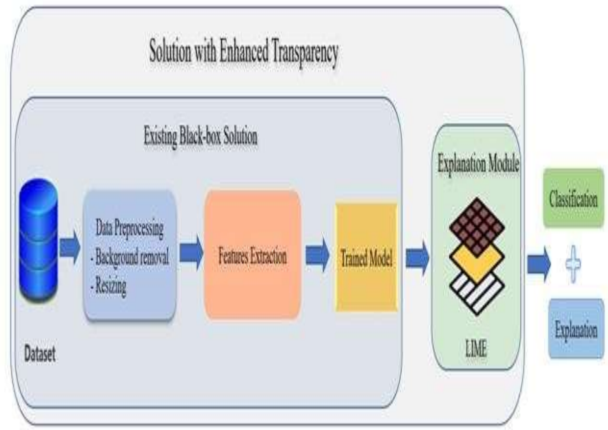

Fig 1: Proposed methodology workflow for plant disease classification using EfficientNetB0 and LIME

A. Data Pre-Processing

Raw plant leaf images often contain noise such as complex backgrounds, shadows, and lighting inconsistencies. Effective pre-processing helps clean the data and improves model learning:

- Background Removal is used to isolate the leaf and discard irrelevant visual features that could interfere with learning.
- Image Resizing and Normalization standardizes image dimensions (224×224) and scales pixel values, ensuring consistent input to the CNN.
- Data Augmentation techniques like horizontal/vertical flipping, rotation, and brightness adjustments are applied to artificially increase dataset size, address class imbalance, and improve generalization.
- Label Organization ensures each image is accurately mapped to its corresponding crop and disease class for supervised learning.

These steps help reduce noise, enhance feature visibility, and improve the overall robustness of the model.

B. Feature Extraction Using EfficientNetB0

EfficientNetB0, a scalable and lightweight CNN, is employed for its strong accuracy-to-efficiency balance:

ISSN [Online]: 2583-2654

- The model is initialized with ImageNet pre-trained weights, providing a strong starting point for learning plant-specific features.
- Transfer learning is applied by freezing base layers and fine-tuning the final layers with custom disease classes.
- A global average pooling layer replaces fully connected layers to reduce overfitting.
- A softmax classifier is added for multi-class prediction across 38 plant disease categories.

EfficientNetB0's ability to extract rich and deep visual features with fewer computational demands makes it suitable for real-time and mobile deployment.

C. Model Training and Validation

To ensure reliable model performance, the dataset is split into training and validation sets in an 80:20 ratio:

- Model training is performed using the TensorFlow framework with GPU acceleration to handle high-volume image data.
- The Adam optimizer is used to minimize the categorical cross-entropy loss, facilitating fast convergence.
- Early stopping is implemented to halt training when validation loss plateaus, preventing overfitting.
- Performance evaluation is done using metrics like accuracy, precision, recall, and F1-score to assess prediction quality at each epoch.

This process ensures that the model generalizes well to unseen images and maintains high reliability across disease classes.

D. Prediction Explainability with LIME

To enhance the model's interpretability, LIME (Local Interpretable Model-Agnostic Explanations) is incorporated:

- LIME perturbs individual image samples and observes how prediction outcomes change, providing a local explanation for each instance.
- A simple surrogate model (such as linear regression) is fit locally to highlight the most influential parts of the image.
- These heatmap visualizations help identify specific areas of the leaf that led to the prediction, making the model's decisions transparent.

This step is crucial for building user trust, especially for farmers and agricultural practitioners who may rely on the model's outputs for crop health decisions.

III. IMPLEMENTATION

This section details the practical steps taken to develop and deploy the plant disease classification system, including environment setup, data handling, model training, explainability integration, and web application deployment.

A. Development Environment

- The training and experimentation were conducted on a Windows 10 workstation with an NVIDIA GPU supporting CUDA for acceleration.
- TensorFlow 2.x was used for building and training the deep learning model.
- GPU acceleration was enabled using CUDA and cuDNN drivers for efficient computation.
- The Flask framework was used for deploying the trained model as a web application.

B. Dataset Preparation

- The dataset comprises over 87,000 images across 38 plant disease classes and 14 crop species, gathered from public repositories such as Kaggle.
- Pre-processing involved background removal, resizing images to 224×224 pixels, and normalizing pixel values.
- Dynamic data augmentation techniques such as rotation, flipping, and brightness adjustment were applied during training to enhance model robustness and balance classes.

C. Model Training

- The classification task is handled by a Convolutional Neural Network (CNN) architecture, specifically EfficientNetB0, which is known for its balance of accuracy and computational efficiency.
- CNNs are well-suited for image-based tasks as they automatically learn and extract spatial features

ISSN [Online]: 2583-2654

such as texture, color, and shape—critical for identifying disease symptoms on leaves.

- The EfficientNetB0 model was loaded with pre-trained ImageNet weights; base layers were frozen to preserve learned visual features.
- Fine-tuning was done by training the classification head with the prepared plant disease dataset.
- The **Adam optimizer** was used with a learning rate of 0.001.
- A batch size of 32 was used, and the model was trained for up to 100 epochs, with early stopping based on validation loss to prevent overfitting.
- An 80:20 train-validation split was employed to evaluate the model's generalization capability.

D. Explainability Integration

- Post-training, the LIME (Local Interpretable Model-Agnostic Explanations) framework was integrated to provide localized explanations for predictions made by the CNN model.
- For each test image, LIME perturbs pixel regions and observes the CNN's prediction changes, then builds a local surrogate model to identify influential regions.
- The result is a **heatmap** overlay on the original image, helping users understand which areas of the leaf led to a particular classification.

E. Web Application Deployment

- The trained EfficientNetB0 CNN model was deployed via a Flask web application to enable easy access for end-users.
- The web interface allows users to upload plant leaf images through their browsers.
- Upon image upload, the Flask backend performs disease classification and generates a corresponding LIME-based heatmap.
- The web page displays the predicted disease class, confidence score, and visual explanation, providing both functionality and interpretability.
- This deployment enables real-time, user-friendly interaction without requiring advanced technical skills.

IV. RESULTS AND DISCUSSION

The proposed plant disease classification system was evaluated based on standard performance metrics, including accuracy, precision, recall, and F1-score. The model was trained on a balanced and augmented dataset comprising over 87,000 images, covering 38 disease classes from 14 different crop species. The experiments were conducted on a Windows 10 machine with GPU acceleration to ensure efficient training and inference.

A. Model Performance

After training the EfficientNetB0 CNN on the prepared dataset, the model demonstrated high classification accuracy on the validation set:

- Accuracy: 99.3%
- Precision: 99.1% • Recall: 99.4%
- F1-Score: 99.2%

These results indicate that the model was highly effective at distinguishing between multiple disease types, including subtle visual differences between similar classes. The high F1-score suggests that the model maintains a strong balance between precision (avoiding false positives) and recall (avoiding false negatives), which is critical for real-world disease diagnosis.

The use of data augmentation techniques (such as rotation, flipping, and brightness adjustment) played a significant role in improving generalization, especially for underrepresented disease classes. Additionally, the Adam optimizer with early stopping helped maintain training stability and reduced overfitting, allowing the model to generalize well across unseen data.

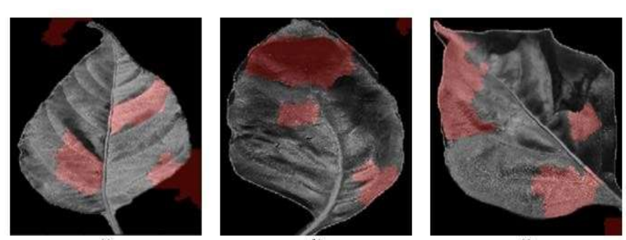

Fig 2: LIME heatmap highlighting key leaf regions influencing disease classification.

ISSN [Online]: 2583-2654

B. Confusion Matrix Analysis

A confusion matrix was generated to visualize classification performance across all disease categories. The matrix showed high diagonal dominance, indicating accurate predictions across most classes.

Most disease classes were classified with high confidence and minimal misclassification. However, a few misclassifications were noted between visually similar disease types, such as Early Blight and Late Blight in tomato plants. These confusions likely occurred due to overlapping visual symptoms like irregular brown spots and blight-like lesions, which can be difficult even for human experts to distinguish.
Despite these few overlaps, the overall classification accuracy remained high, demonstrating the CNN's ability to extract fine-grained visual features from leaf images. Future improvements could involve ensemble techniques or integrating temporal data (e.g., symptom progression over time) for further enhancement.

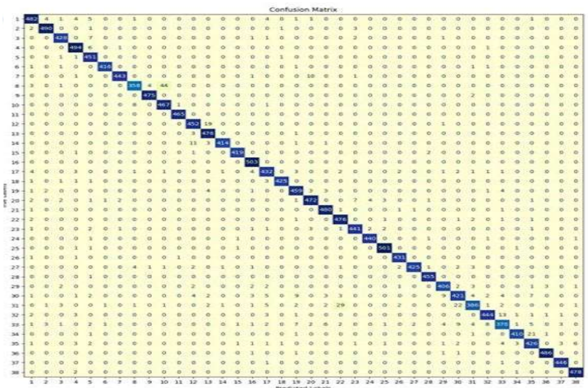

Fig 3: Confusion matrix showing classification accuracy across 38 plant disease classes.

C. Explainability with LIME

To overcome the black-box nature of deep learning models, LIME (Local Interpretable Model-Agnostic Explanations) was used to provide human-understandable visual explanations for each prediction.

For every input image, LIME perturbs the pixels in localized regions and generates predictions for each variation. It then constructs a local surrogate model (e.g., linear regression) to identify which parts of the image most strongly influenced the final classification.

The result is a heatmap overlaid on the original leaf image, visually highlighting the areas most responsible for the model's decision. These explanations:

- Help validate whether the model focused on actual disease symptoms.
- Assist users, especially non-experts, in understanding the model's behavior.
- Enable identification of potential misclassifications or bias in the model's attention.

This interpretability improves user trust and is especially valuable in agricultural settings, where understanding "why" a disease was predicted is just as important as the prediction itself.

D. Web Application Testing

The trained model was deployed using a Flask-based web application, allowing users to interact with the system through a browser interface. The web app was tested for usability, responsiveness, and real-time performance.
Key features included:

- Image upload functionality, allowing users to input new leaf samples easily.
- Real-time prediction and display of disease name and confidence score.
- LIME-generated heatmaps, which provide visual cues indicating why a prediction was made.
- Fast response time, with predictions typically returned in under 3 seconds.

The model was converted to TensorFlow Lite format for deployment, ensuring it could run efficiently even on resource-limited systems like basic laptops or edge devices.
This setup makes the system highly usable for fieldwork, agricultural kiosks, or remote diagnosis, especially in regions lacking technical infrastructure.

E.Discussion and Observations

The overall experimental results and system

ISSN [Online]: 2583-2654

performance lead to the following key takeaways:

- CNN architectures, particularly EfficientNetB0, are extremely effective for complex image classification tasks involving multiple classes and subtle variations.
- Preprocessing and data augmentation techniques significantly improve model robustness and help avoid overfitting, especially when working with real-world, imperfect data.
- The integration of LIME adds a layer of model transparency, addressing the lack of interpretability often associated with deep learning. This is vital for adoption in sensitive domains like agriculture.
- The Flask-based web deployment offers a lightweight and accessible platform for users without coding expertise, making the system practical and scalable.

These findings confirm that the proposed system is not only technically sound but also ready for real-world application, offering a combination of accuracy, interpretability, and usability that aligns with the needs of modern agricultural diagnostics.

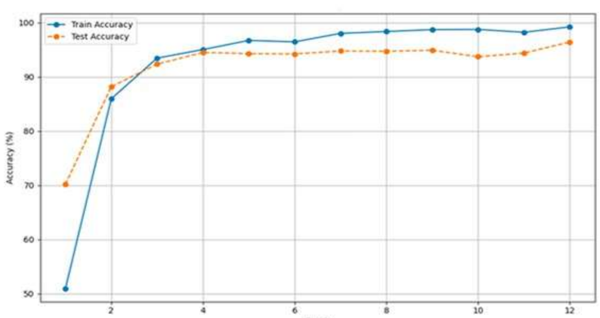

Fig 4: Performance metrics of the EfficientNetB0 model including accuracy, precision, recall, and F1-score.

V. FUTURE SCOPE

While the proposed plant disease classification system demonstrates high accuracy and interpretability, there are several areas where the system can be further enhanced for broader applicability and real-world robustness:

Integration with Real-Time Field Data:
Future versions of the system can incorporate real-time field data such as environmental conditions (temperature, humidity, soil moisture) to improve context-aware predictions.

Support for More Crops and Diseases:
Expanding the dataset to include a wider variety of crops and rare disease classes will make the model more versatile and applicable to diverse agricultural regions.

Continuous Learning with User Feedback:
Implementing a feedback mechanism where farmers or experts can validate or correct predictions would enable continuous improvement of the model through retraining.

Edge and IoT Deployment:
The model can be optimized for edge devices like Raspberry Pi or integrated with drones and IoT sensors to enable autonomous plant monitoring in large-scale farms.

Multi-language Support in UI:
Enhancing the Flask web application to support regional languages would make the system more accessible to farmers across different linguistic backgrounds.

Ensemble or Hybrid Models:
Combining EfficientNet with other deep learning architectures or integrating disease-specific rule-based logic may further boost classification accuracy and reduce misclassification between visually similar diseases.

VI. CONCLUSION

In this paper, we presented a deep learning-based system for accurate and interpretable plant disease classification using the EfficientNetB0 CNN model and LIME explainability framework. The model was trained on a large, diverse dataset comprising over 87,000 images spanning 38 plant disease categories across 14 crop species.

ISSN [Online]: 2583-2654

Through the use of transfer learning and data augmentation, the model achieved outstanding performance, with over 99% accuracy, precision, recall, and F1-score on the validation set. The integration of LIME provided visual explanations for each prediction, improving transparency and trust in the model's decisions—an essential feature for deployment in sensitive agricultural environments.

To ensure practical usability, the trained model was deployed via a Flask-based web application, allowing users to upload leaf images, view predictions, and receive interpretable heatmap visualizations. This lightweight and accessible interface makes the system viable for use by farmers, agronomists, and agricultural extension workers, even in low-resource settings.
The results confirm that CNN-based models, when combined with explainable AI and user-friendly deployment, can effectively address real-world challenges in plant disease diagnosis. This work lays a strong foundation for future enhancements involving real-time sensing, multilingual interfaces, and edge device deployment for smart agriculture.

REFERENCES

[1] P. S. Kumar, M. R. Ranjitha, and V. Rajalakshmi, "Improving Plant Disease Classification With Deep-Learning-Based Prediction Model Using Explainable Artificial Intelligence," IEEE Access, vol. 10, pp. 88303–88314, 2022.J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan. 1959.

[2] D. Pant, H. A. Malik, and S. S. Saini, "Machine Learning and Deep Learning for Plant Disease Classification and Detection: A Review," Materials Today: Proceedings, 2022.

[3] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. Int. Conf. Mach. Learn. (ICML), 2019, pp. 6105–6114.

[4] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," in Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems, vol. 25, 2012.

[6] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," Computational Intelligence and Neuroscience, vol. 2016, pp. 1–11, 2016.

[7] A. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep Learning for Tomato Diseases: Classification and Symptoms Visualization," Applied Artificial Intelligence, vol. 31, no. 4, pp. 299–315, 2017.

[8] J. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.

[9] G. B. M. Ayalew, K. S. Mahesh, and K. A. Kiran, "Plant Disease Identification Using Feature Extraction and Classification Techniques: A Review," Materials Today: Proceedings, vol. 51, pp. 1003–1009, 2022.

[10] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Model-agnostic interpretability of machine learning," arXiv preprint arXiv:1606.05386.