INTERPRETABLE DEEP LEARNING MODEL USING CNN AND TRANSFORMER FOR BEAN PLANT DISEASE DIAGNOSIS

P.Geethanjali[1], N.Deekshitha[2], K.Adithya Reddy[3], J.Pranay[4], N.Vasanth Reddy[5]

[1] Assistant Professor, *Department of CSE (ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING) TKR COLLEGE OF ENGINEERING & TECHNOLOGY*

[2,3,4,5] *UG Scholars in Department of CSE (ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING) TKR COLLEGE OF ENGINEERING & TECHNOLOGY*

Abstract: Bean leaf diseases like rust and angular leaf spot lead to significant yield losses in agriculture. Early identification and proper severity estimation are crucial for controlling the spread of infection. The proposed system uses a deep learning model to automatically detect the type of disease and assess its severity from leaf images. The model processes the input image, extracts important features using convolutional and transformer techniques, and classifies the disease severity into categories like mild, moderate, and severe. This approach allows for precise monitoring of crop health and helps farmers make informed decisions about disease management. Grad-CAM visualization is included to highlight the affected areas, ensuring the model's predictions are understandable and trustworthy. The proposed method aims to provide high accuracy and practical usability for real-time agricultural applications.

Keywords: Bean Leaf Disease Detection, Disease Severity Estimation, PVT, GCADSN Attention, Grad-CAM Visualization, Precision Agriculture.

1. INTRODUCTION

Bean plants are commonly grown by many farmers, and they are an important source of food and income. But these plants often get infected by diseases like rust and angular leaf spot. These diseases usually start with small spots, and most farmers don't notice them until they spread across many leaves. Because of busy schedules and lack of expert knowledge, early detection becomes difficult, and by the time the problem is identified, the crop may already be affected badly.

Right now, farmers mainly depend on their own eyes to check for diseases. This method is not always reliable because leaf symptoms can look similar, lighting may not be clear, and sometimes the infection is too mild to notice. Even if a farmer identifies the disease, estimating how serious it is becomes another challenge. This leads to delayed treatment, wrong usage of pesticides, and unnecessary loss of both money and crop quality.

To help farmers with this, our project introduces a simple system that can detect diseases directly from a leaf photo. By using modern deep-learning techniques, the system identifies the disease and also tells how severe the infection is. It even highlights the exact affected areas so that farmers can clearly see what part of the leaf is damaged. This makes it easier for farmers to take quick decisions, control the spread, and protect their crop more effectively.

1.1 Motivation

Farmers depend heavily on bean crops for their livelihood, but these plants are easily attacked by diseases like rust and angular leaf spot, which quietly spread and cause serious damage before anyone notices. Most farmers try to check the leaves on their own to figure out what's wrong, but this often leads to confusion because many disease symptoms look alike, and small infections are not easy to notice during regular field work. Early signs usually go unnoticed, especially when the leaf is viewed under poor lighting or dusty conditions. When diseases are detected late, farmers end up losing a large part of their crop and also spend extra money on treatments that may not even work effectively. Because of this, there is a real need for a simple and quick method that can help farmers spot diseases early and clearly understand the condition of the plant. This situation shows a strong need for a simple, dependable, and fast method that can help farmers catch diseases early and understand how serious the infection is. Building such an automated system can make it easier for farmers to take action on time and protect their crops without depending too much on guesswork. This motivation drives the development of an automated disease detection solution that supports farmers in maintaining healthier crops and achieving better yields.

1.2 Proposed System

The proposed system is designed to offer farmers an efficient way to detect diseases in bean leaves, right from an image of the leaf. Without necessarily having to check physically, this system analyzes the photo and detects the disorders that might be affecting a leaf, whether rust, angular leaf spot, or any other disease covered under the model. This process is faster and easier, and highly more accurate compared to traditional methods of visual inspection.

The system begins with the preprocessing of the leaf image to remove irrelevant background, resize, and enhance the clarity. This helps in making the model focus only on the leaf and its disease patterns rather than getting misled by the soil, shadows, or uneven lighting of the leaf. This stage cleans the image for further proper analysis.

Then, the cleaned image is passed through a deep-learning model with Pyramid Vision Transformer, as it captures not only the big patterns but also the small fine details on the leaf. Finally, a GCADSN attention mechanism enhances the ability of the model to pay more attention to the disease-affected regions, therefore improving accuracy. These joint techniques empower the system to detect even early-stage infections that usually go unnoticed by the human eye.

Once the model identifies the disease, the system also measures the severity level- mild, moderate, or severe. This is an important part of the project because knowing the disease itself is not enough, but the farmers have to know how serious it is to decide the right treatment. The estimation of the severity provides farmers with a clearer understanding of their crop's condition and helps them in taking timely actions.

Finally, the system provides the visual explanation using Grad-CAM. This feature highlights the exact areas on the leaf that the model considered in predicting the disease. This makes the system transparent and easy to trust, as farmers will clearly see which parts of the leaf are infected. With all these features combined, the proposed system becomes practical and user-friendly to support real-time disease monitoring that helps farmers protect their crops more effectively.

2. SEQUENCE DIAGRAM

A sequence diagram is a dynamic UML diagram that shows how objects interact over time. It illustrates the sequence of messages or method calls exchanged between objects to accomplish a specific functionality, highlighting the order of operations

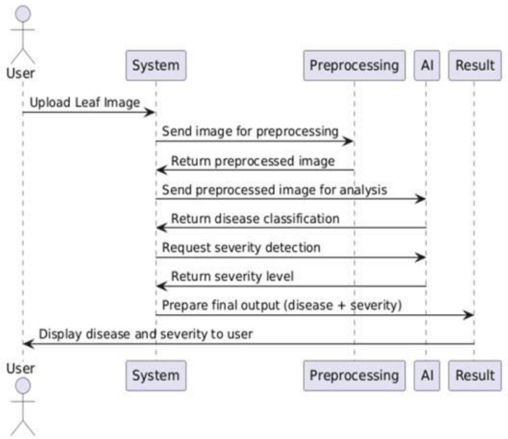


Fig: 1 Sequence diagram of bean plant disease diagnosis

2.1 USECASE DIAGRAM

A Use Case Diagram is a UML diagram that shows how users (actors) interact with a system. It represents the system's functional requirements through use cases, which are the actions or services the system provides. Actors can be people, devices, or other systems that interact with the system. Relationships connect actors to use cases, showing how they use the system. It helps visualize system functionality and ensures clear understanding of requirements.

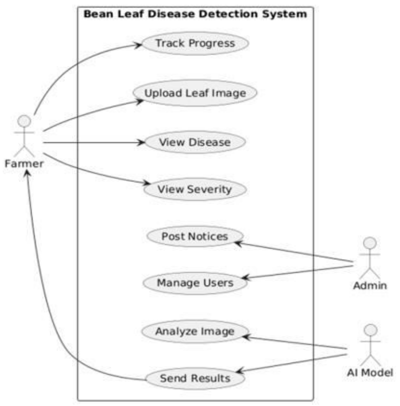


Fig: 2 Use case diagram of bean plant disease diagnosis

3. ALGORITHMS

A. Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a deep learning algorithm widely used for image processing and analysis. In the proposed bean leaf disease detection system, CNN is responsible for extracting important visual features from the input leaf images. It consists of multiple layers such as convolutional layers, activation functions, and pooling layers. The convolutional layers apply filters to the image to detect patterns like edges, textures, and color variations. As the data passes through deeper layers, the network learns more complex features such as disease spots, lesions, and discoloration present on the leaf surface.

The activation function, usually ReLU, introduces non-linearity, allowing the model to learn complex relationships in the data. Pooling layers reduce the size of the feature maps while retaining essential information, which helps in reducing computational cost and preventing overfitting. In this system, CNN focuses on identifying local features that are crucial for distinguishing between healthy and diseased leaves. It automatically learns relevant features without manual intervention, making it highly efficient. Overall, CNN acts as the backbone of the model, providing detailed feature representations that are later used for classification and severity prediction.

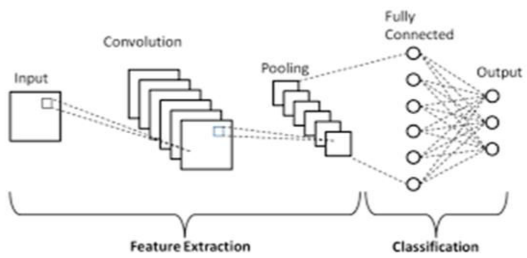


Fig:1. CNN

B. Transformer Model

The Transformer model is an advanced deep learning architecture that uses an attention mechanism to process data and capture relationships within it. In the proposed system, the transformer is used to analyze the overall structure of the leaf image and understand global patterns. Unlike CNN, which focuses on local features, the transformer considers the entire image at once. It divides the image into smaller patches and processes them as a sequence, allowing the model to learn how different regions of the image are related.

The key component of the transformer is the self-attention mechanism, which helps the model focus on the most important parts of the image. This is particularly useful in detecting plant diseases, as it allows the model to identify how infections are distributed across the leaf. The transformer also uses positional encoding to maintain spatial information about the patches. In this system, it complements CNN by providing a global understanding of disease spread and severity. This combination improves the overall accuracy and robustness of the model, especially in complex cases where disease patterns are not localized.

C. Hybrid Model (CNN + Transformer)

The hybrid model combines the strengths of both CNN and transformer architectures to improve the performance of the disease detection system. CNN is effective at capturing local features such as textures, edges, and small disease spots, while the transformer excels at understanding global relationships across the entire image. By integrating both approaches, the system can analyze leaf images more comprehensively.

In this model, the input image is first processed by the CNN to extract detailed feature maps. These features are then passed to the transformer, which analyzes the relationships between different regions of the image. This combination allows the model to detect both small-scale and large-scale disease patterns. The hybrid approach helps in accurately identifying the type of disease as well as estimating its severity.

This method is particularly useful in agricultural applications where diseases may vary in size, shape, and distribution. By combining local and global feature extraction, the hybrid model provides more reliable and precise predictions. It enhances the model’s ability to generalize across different conditions, making it suitable for real-time crop monitoring and disease management.

D. Softmax Classification Layer

The Softmax function is used in the final layer of the model to perform classification. It converts the output values of the neural network into probabilities, which represent the likelihood of each class. In the proposed system, the classes correspond to different levels of disease severity, such as mild, moderate, and severe.

The Softmax function ensures that the sum of all output probabilities is equal to one, making it easier to interpret the results. For example, if the model outputs probabilities of 0.2, 0.5, and 0.3 for the three classes, the system will classify the leaf as having moderate severity since it has the highest probability. This probabilistic approach allows for more informed decision-making.

In addition, Softmax helps in handling multi-class classification problems effectively. It provides a clear and interpretable output, which is essential in practical applications like agriculture. Farmers and experts can use these predictions to assess the condition of crops and take appropriate actions. Overall, the Softmax layer plays a crucial role in converting learned features into meaningful predictions.

F. Grad-CAM (Gradient-weighted Class Activation Mapping)

Grad-CAM is a visualization technique used to make deep learning models more interpretable. In the proposed system, it helps in understanding how the model makes its predictions by highlighting the important regions in the leaf image. Grad-CAM works by using the gradients of the target class flowing into the final convolutional layer to produce a heatmap.

This heatmap indicates which parts of the image contributed most to the model's decision. For example, in a diseased leaf, the highlighted regions may correspond to infected areas such as spots or lesions. This provides visual confirmation that the model is focusing on relevant features rather than irrelevant background details.

In agricultural applications, interpretability is very important because users need to trust the model's predictions. Grad-CAM enhances transparency by showing the reasoning behind the classification. It also helps in debugging and improving the model by identifying potential errors. Overall, Grad-CAM adds an extra layer of reliability and usability to the system, making it more suitable for real-world deployment.

3.1 **SAMPLE DATA**

Sample data is a small subset of the original dataset used to demonstrate the structure and format of data in the bean leaf disease detection system. It contains a few example leaf image records along with their corresponding features and labels. Each record represents a bean leaf and includes attributes such as image ID, leaf condition, disease type (e.g., rust, angular leaf spot, or healthy), and severity level.

The dataset primarily consists of leaf images, from which features are automatically extracted using deep learning techniques such as Convolutional Neural Networks (CNN) and Transformer models. These features include visual patterns like color variations, texture, lesion size, and distribution of infected areas on the leaf surface. Such features help the model analyze the condition of the leaf and identify the presence and severity of disease.

The final label in the sample data indicates the classification result, such as healthy, rust disease, or angular leaf spot, along with severity categories like mild, moderate, or severe. This structured data helps in training and evaluating the model effectively.

This sample data is useful for understanding how the dataset is organized, how image data is processed, and how the model generates predictions based on extracted features. It is mainly used for demonstration, testing, and documentation purposes.

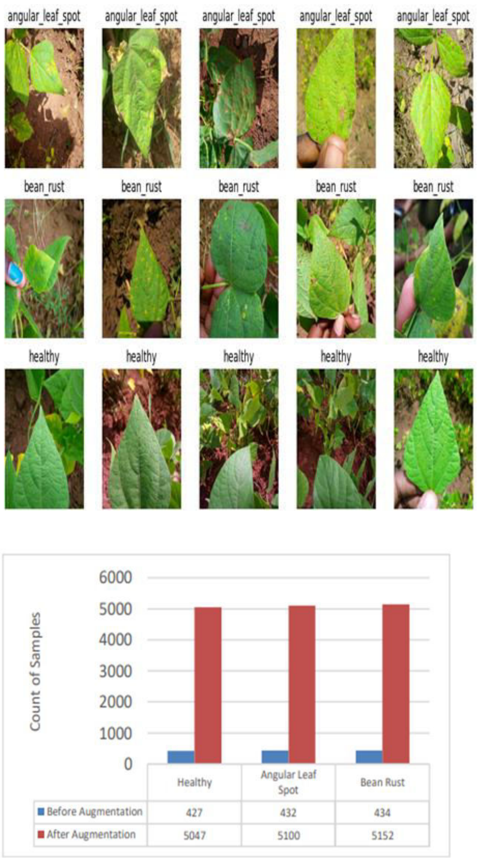


Fig Sample Data

4. Explanation Of Key Functionalities

FEATURE EXTRACTION

Feature extraction is the first step in the implementation process. In this function, the system extracts important visual information from the bean leaf images. These features include color variations, texture patterns, shape of lesions, size of infected regions, and distribution of disease spots on the leaf surface. Deep learning models such as Convolutional Neural Networks (CNN) automatically extract these features without manual intervention. The purpose of feature extraction is to identify patterns that distinguish healthy leaves from diseased ones and to recognize specific diseases such as rust and angular leaf spot.

DATA PREPROCESSING

Data preprocessing is performed to convert raw image data into a clean and structured format suitable for deep learning models. In this step, unnecessary noise is removed, images are resized to a fixed dimension, and pixel values are normalized. Data augmentation techniques such as rotation, flipping, and zooming may also be applied to increase dataset diversity. Proper preprocessing improves model performance, reduces overfitting, and ensures consistency in input data.

DATASET SPLITTING

After preprocessing, the dataset is divided into training and testing sets. Typically, 80% of the images are used for training the model, while the remaining 20% are used for testing. The training dataset helps the model learn patterns related to different diseases, while the testing dataset is used to evaluate how well the model performs on unseen images.

FEATURE SELECTION

Feature selection in deep learning is handled automatically by the model. The CNN identifies and prioritizes the most important features from the images while ignoring irrelevant information. This reduces model complexity and improves accuracy. By focusing on meaningful features such as infected areas and lesion patterns, the system becomes more efficient in detecting diseases.

MODEL TRAINING

In this stage, the deep learning model (CNN combined with Transformer) is trained using the training dataset. The model learns to identify patterns associated with different diseases and severity levels. Optimization algorithms such as Adam optimizer are used to minimize loss and improve accuracy. The training process involves multiple epochs, where the model continuously updates its weights to achieve better performance.

DISEASE CLASSIFICATION

Disease classification is the final function of the system. When a user uploads a bean leaf image, the system processes the image and extracts features. These features are then passed to the trained model, which predicts the disease type (such as rust or angular leaf spot) and its severity level (mild, moderate, or severe). The result is displayed to the user, helping in early detection and proper disease management.

4.2 METHOD OF IMPLEMENTATION

1. Technology Selection:

The system is developed using Python as the primary programming language. FastAPI is used to build the backend API, while the frontend can be developed using HTML, CSS, and JavaScript. Firebase or other databases may be used to store prediction results and user data.

2. Data Collection:

A dataset of bean leaf images is collected, including both healthy and diseased leaves. These images are obtained from agricultural datasets or captured manually. The dataset forms the foundation for training the deep learning model.

3. Feature Extraction:

Important visual features such as color, texture, and disease patterns are extracted automatically using CNN layers. These features help the model identify disease characteristics effectively.

4. Data Preprocessing:

The collected images are cleaned and prepared by resizing, normalization, and augmentation. This ensures that the data is consistent and suitable for training the model.

5. Dataset Splitting:

The dataset is divided into training and testing sets, usually in an 80:20 ratio. This helps in training the model and evaluating its performance.

6. Model Training:

The hybrid CNN-Transformer model is trained on the dataset. The model learns to classify diseases and predict severity levels using optimization techniques.

7. Prediction and Result Display:

When a user uploads a leaf image, the system processes it and sends it to the trained model. The model predicts the disease type and severity, and the result is displayed to the user through the interface.

4.2.1 FORMS

Forms provide an interface for users to interact with the system. These forms are used to upload images and manage system operations.

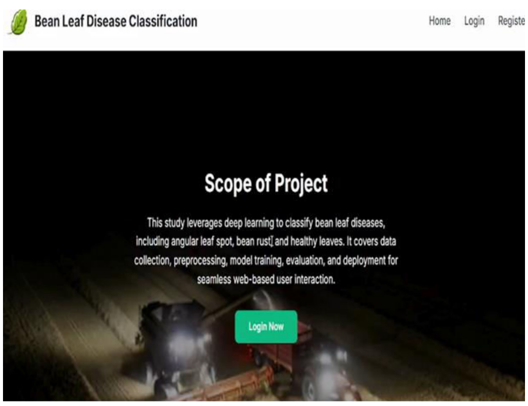
- **Image Upload Form:** This form allows users to upload bean leaf images. Once submitted, the image is sent to the backend for prediction.
- **User Login Form:** Registered users can log in using their credentials to access the system features.
- **Admin Login Form:** The administrator logs in to manage the system and monitor predictions.
- **Admin Management Form:** This form allows the admin to view and manage uploaded images and prediction results.

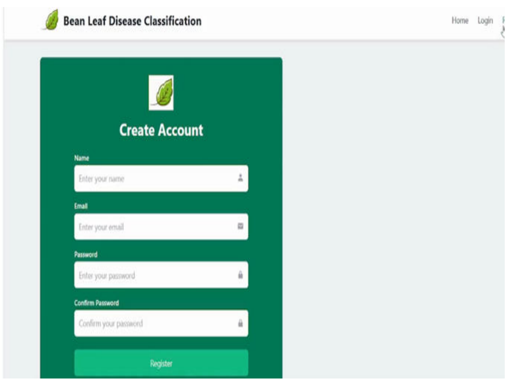
5.2.2 OUTPUT SCREENS

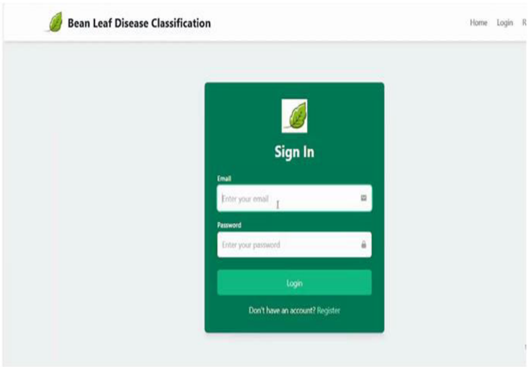
When a user uploads a bean leaf image, the output screen displays the prediction results generated by the system. The deep learning model processes the image and analyzes important features such as color, texture, and infected regions. Based on this analysis, the system predicts whether the leaf is healthy or affected by a disease.

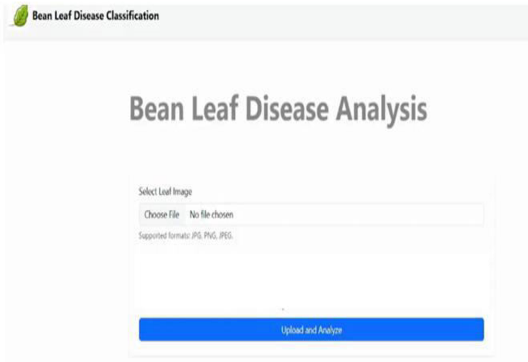
If the leaf is healthy, the system displays a message indicating that no disease is detected. If a disease is present, the system shows the disease name (such as rust or angular leaf spot) along with its severity level (mild, moderate, or severe). Additionally, visualization techniques like Grad-CAM may highlight the affected areas on the leaf image.

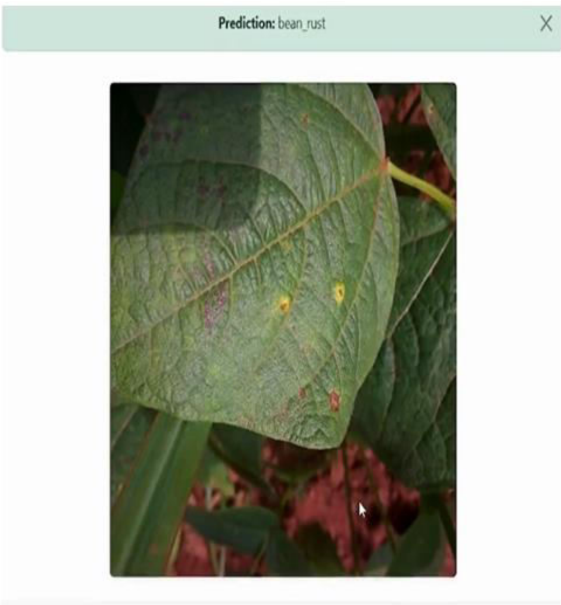
The output screen provides clear and easy-to-understand results, helping users take timely action for disease control and crop management.











5.2.3 RESULTS ANALYSIS

The result analysis is performed to evaluate the performance of the bean leaf disease detection system. The deep learning model, which combines Convolutional Neural Network (CNN) and Transformer techniques, is trained using the training dataset and tested using the testing dataset to measure its prediction accuracy. The system analyzes the features extracted from the leaf images, such as color variations, texture patterns, and infected regions, to classify the leaf as healthy or diseased.

Performance metrics such as accuracy, precision, recall, and F1-score are used to evaluate the effectiveness of the model. Accuracy measures the overall correctness of predictions, while precision indicates how many predicted disease cases are actually correct. Recall measures the model's ability to detect all diseased leaves, and the F1-score provides a balance between precision and recall. A confusion matrix is also used to visualize the classification performance.

The results show that the system can successfully detect diseases such as rust and angular leaf spot with high accuracy and reliability. Additionally, the model is able to estimate the severity level of the disease, which helps in better crop management. The inclusion of Grad-CAM visualization further improves trust by showing the affected areas on the leaf. Overall, the system proves to be effective for real-time agricultural applications and supports farmers in early disease detection and decision-making.

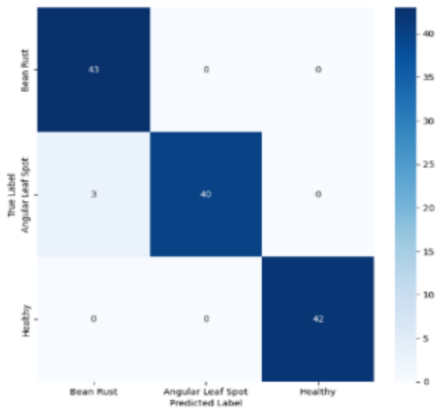
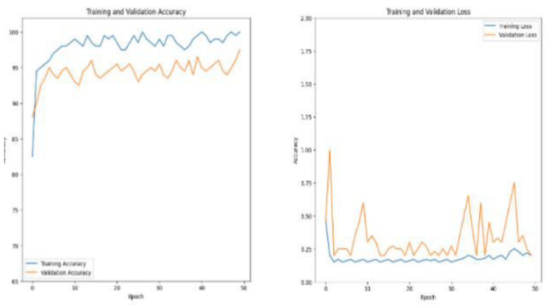


Fig 5.2.1 CONFUSION MATRIX



(a) (b)

Fig 5.2.2 a)Accuracy b) Loss

6. CONCLUSION

The bean leaf disease detection system developed in this project provides an effective and intelligent solution for identifying plant diseases using deep learning techniques. Plant diseases are a major challenge in agriculture, leading to reduced crop yield and financial losses for farmers. Early detection and proper diagnosis of diseases are essential to prevent their spread and ensure healthy crop production. This project addresses this problem by using image-based disease detection with the help of a Convolutional Neural Network (CNN) model.

In this system, a dataset of bean leaf images is collected and preprocessed to improve image quality and consistency. The dataset is then divided into training and testing sets, where the training data is used to train the CNN model, and the testing data is used to evaluate its performance. The model learns to extract important features from the images, such as color, texture, and patterns, which are useful for identifying different types of diseases. After training, the system is capable of taking a new leaf image as input and predicting the type of disease present in the leaf.

One of the important features of this project is the use of Grad-CAM (Gradient-weighted Class Activation Mapping) for visualization. This technique highlights the infected regions of the leaf, allowing users to understand how the model makes its predictions. This improves the transparency and interpretability of the system, making it more useful for practical applications.

The performance of the system is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. The results indicate that the model performs well and can accurately classify different bean leaf diseases. The system is also efficient, user-friendly, and capable of providing quick results, which makes it suitable for real-time usage.

Overall, this project demonstrates the effectiveness of deep learning in solving real-world agricultural problems. The bean leaf disease detection system can assist farmers, researchers, and agricultural experts in identifying diseases at an early stage, thereby reducing crop damage and improving productivity. In the future, the system can be enhanced by increasing the dataset size, adding more disease categories, and developing a mobile or web-based application for easier access. Integration with real-time monitoring systems and IoT devices can further improve its usability. Thus, this project provides a strong foundation for developing advanced smart agriculture solutions using artificial intelligence.

REFERENCES

[1] R. Karthik, R. Aswin, K. S. Geetha, and K. Suganthi, "An explainable deep learning network with transformer and custom CNN for bean leaf disease classification," IEEEAccess, vol. 13, pp. 38562–38573, Jan. 2025. doi: 10.1109/ACCESS.2025.3546017.

[2] U. Mwaibale, N. Mduma, H. Laizer, and B. Mgawe, "Enhancing detection of common bean diseases using Fast Gradient Sign Method– trained Vision Transformers," Frontiers in Artificial Intelligence, vol. 8, p. 1643582, Aug. 2025. doi: 10.3389/frai.2025.1643582

[3] H. Chen, H. Zhai, J. Hu, H. Chen, C. Wen, Y. Feng, K. Wang, Z. Li, and G. Wang, "YOLOv8-RBean: Runner Bean Leaf Disease Detection Model Based on YOLOv8," Agronomy, vol. 15, no. 4, p. 944, Apr. 2025. doi: 10.3390/agronomy15040944.

[4] M. S. Krishna, P. Machado, R. I. Otuka, S. W. Yahaya, F. N. dos Santos, and I. K. Ihianle, "Plant leaf disease detection using deep learning: A multi-dataset approach," J, vol. 8, no. 1, pp. 4–24, Jan. 2025. doi: 10.3390/j8010004.

[5] H. C. Laizer, N. Mduma, D. Machuve, and R. Maganga, "Common beans imagery dataset for early detection of bean rust and bean anthracnose diseases," Data in Brief, vol.54, p. 110508, May 2024. doi: 10.1016/j.dib.2024.110508.

[6] D.-C. Rodríguez-Lira, D.-M. Córdova-Esparza, J. M. Álvarez-Alvarado, J.-A. Romero-González, J. Terven, and J. Rodríguez-Reséndiz, "Comparative analysis of YOLO models for bean leaf disease detection in natural environments," AgriEngineering, vol. 6, no. 4, pp. 4585–4603, Nov. 2024. doi: 10.3390/agriengineering6040262.

[7] N. E. Ghannam, O. M. El Zein, D. R. Fathy, and H. Mancy, "Enhanced detection of bean leaf diseases using a stacked CNN ensemble with transfer learning," Int. J. Intell. Eng. Syst., vol. 16, no. 1, pp. 1–12, Jan. 2025. doi: 10.22266/ijies2025.0228.01.

[8] E. Önler, "Feature fusion based artificial neural network model for disease detection of bean leaves," Electronic Research Archive, vol. 31, no. 5, pp. 2409–2427, Feb. 2023. doi: 10.3934/era.2023122.

[9] S. Serttaş and E. Deniz, "Disease detection in bean leaves using deep learning," Communications Faculty of Sciences University of Ankara Series A2-A3: Physical Sciences and Engineering, vol. 65, no. 2, pp. 115–129, Dec. 2023. doi: 10.33769/aupse.1247233.

[10] S. Sagar, M. Javed, and D. S. Doermann, "Leaf-based plant disease detection and explainable AI," arXiv preprint arXiv:2404.16833, Dec. 2023. Available: https://arxiv.org/abs/2404.16833.

[11] H. Slimani, J. El Mhamdi, and A. Jilbab, "Artificial intelligence-based detection of fava bean rust disease in agricultural settings: An innovative approach," Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 6, pp. 119–128, Jun. 2023. doi: 10.14569/IJACSA.2023.0140614.

[12] R. I. Omara, S. M. Kamel, S. M. El-Ganainy, R. A. Arafa, Y. S. Mostafa, S. A. Alamri, S. A. Alrumman, M. Hashem, and M. M. Elsharkawy, "Host resistance to Uromyces appendiculatus in common bean genotypes," Plants, vol. 11, no. 5, p. 628, May 2022. doi: 10.3390/plants11050628.

[13] P. V. Yeswanth, M. Saisanika, and S. Deivalakshmi, "Self-governing feedback network (SGFN) based super-resolution for bean leaf disease detection," J. Adv. Appl. Sci. Res., vol. 6, no. 3, pp. 950–957, May 2024. doi: 10.46947/joaasr632024950.

[14] F. Girma, C. Fininsa, H. Terefe, and B. Amsalu, "Evaluation of common bean (Phaseolus vulgaris) genotypes for resistance to common bacterial blight and angular leaf spot diseases, and agronomic performances," Heliyon, vol. 8, no. 8, p. e10425, Aug. 2022. doi: 10.1016/j.heliyon.2022.e10425.

[15] P. Sahu, A. Chug, A. P. Singh, D. Singh, and R. P. Singh, "Deep learning models for beans crop diseases: Classification and visualization techniques," Int. J. Modern Agric., vol. 10, no. 1, pp. 796–812, 2021. Available: https://modern-journals.com/index.php/ijma/article/view/670.