A DEEP LEARNING APPROACH FOR PLANT LEAF DISEASE DETECTION USING VGG16

A. Geethanjali[1], C.Manikanta[2], A.Rihana[3], E.Chandu[4], E.Hanumanth Reddy[5], A.Durvasulu[6]

1.Assistant Professor, Department of Electronics and Communication Engineering, Sri Venkatesa Perumal college of Engineering and Technology, Andhra Pradesh, India

2.UG scholar,Department of Electronics and Communication Engineering, Sri Venkatesa Perumal college of Engineering and Technology, Andhra Pradesh, India

3.UG scholar,Department of Electronics and Communication Engineering, Sri Venkatesa Perumal college of Engineering and Technology, Andhra Pradesh, India

4.UG scholar,Department of Electronics and Communication Engineering, Sri Venkatesa Perumal college of Engineering and Technology, Andhra Pradesh, India

5.UG scholar,Department of Electronics and Communication Engineering, Sri Venkatesa Perumal college of Engineering and Technology, Andhra Pradesh, India

6.UG scholar,Department of Electronics and Communication Engineering, Sri Venkatesa Perumal college of Engineering and Technology, Andhra Pradesh, India

ABSTRACT

Plant leaf diseases present a serious threat to global food security, as they reduce both the quality and quantity of agricultural output. Consequently, identifying such diseases promptly and accurately is crucial to preventing crop damage and addressing the rising food demands. Traditional methods often rely on laboratory tests and expert analysis, which can be expensive and not readily available to all farmers. In contrast, modern Deep Neural Networks have shown great promise in image-based classification tasks.

This study introduces an automated system for detecting plant diseases, which includes several stages such as image pre-processing, disease classification, feature extraction, and selection. The classification process is enhanced through the application of transfer learning using pre-trained models like VGG16, which improves efficiency and reduces the need for extensive training.

Keywords: *Plant leaf diseases, Deep Neural Networks, Image Classifications ,Pre processing, Feature Selection , Transfer Learning , Precision*

1.INTRODUCTION:

Plant diseases are a major threat to agricultural productivity and food availability. If not identified and addressed promptly, these diseases can spread rapidly and devastate large farming areas. Many of these plant diseases are not commonly recognized, making early detection essential to avoid widespread crop damage. Detecting diseases manually is labor-intensive and impractical, especially on farms that span many acres. Automation in this area is increasingly necessary. Although laboratory techniques are used to diagnose plant diseases, they are often too costly and time-consuming to be viable for early detection, which involves examining a vast number of plants. Agriculture plays a critical role in India's economy, contributing significantly to its GDP and employing a large share of the population. People across the world rely on agriculture for food and livelihood. In the past, farmers depended on expert agronomists to diagnose plant diseases, a process that was slow and not always reliable. In many instances, machine learning models can deliver more accurate results than human experts. Recently, image processing has been widely applied in agriculture to improve product quality and assist farmers in protecting their crops. These technologies extract useful information from plant images to identify diseases more efficiently.Despite India's economic progress, over 90% of its population still relies directly or indirectly on agriculture. Pests and plant diseases can severely reduce crop yields and quality. Since many of these diseases are crop-specific, regional, and seasonal, they require complex and integrated management strategies. The vastness and diversity of Indian farmlands make it difficult to identify these diseases early, especially since their symptoms are often too subtle to be seen with the naked eye. Limited visibility and unpredictable patterns of disease make it hard to gather, store, and analyze information for effective prevention.

1.2 MOTIVATION

Farmers frequently turn to online resources or consult experienced gardeners to identify plant diseases. In some cases, they carry a piece of the affected plant or a photograph to nearby agricultural support centers in search of expert advice. However, diagnosing diseases based solely on visual inspection of leaves is often inefficient and requires considerable time and effort. Moreover, such methods are prone to errors.Even skilled agronomists and plant pathologists may misidentify diseases due to the subtle differences in symptoms among various infections. Since many diseases manifest with similar visual signs, it becomes challenging to pinpoint the exact issue without advanced tools. While genetic testing provides a more accurate diagnosis, it is generally expensive and time-intensive, making it impractical for everyday use by most farmers.As a result, there is a growing need for a fast, accurate, and cost-effective approach to plant disease detection. Modern technologies like machine learning and image processing offer a promising alternative by automating the recognition of disease symptoms through leaf imagery.This project is inspired by the need to support farmers with accessible and reliable tools for early disease identification. Such solutions can help prevent crop damage, improve yield, and modernize agricultural practices through the integration of AI-driven diagnostics.

1.3 DOMAIN OVERVIEW

In recent years, deep learning has revolutionized numerous technological domains. A prominent area of innovation is computer vision—the capability of machines to interpret and understand visual data such as images and videos. Technologies like autonomous vehicles, biometric systems, and facial recognition depend heavily on computer vision. At the heart of this field lies image processing, the method through which visual data is analyzed and interpreted. Images are defined by their resolution, i.e., their width and height in pixels. For instance, a 500x400 image has 200,000 pixels.

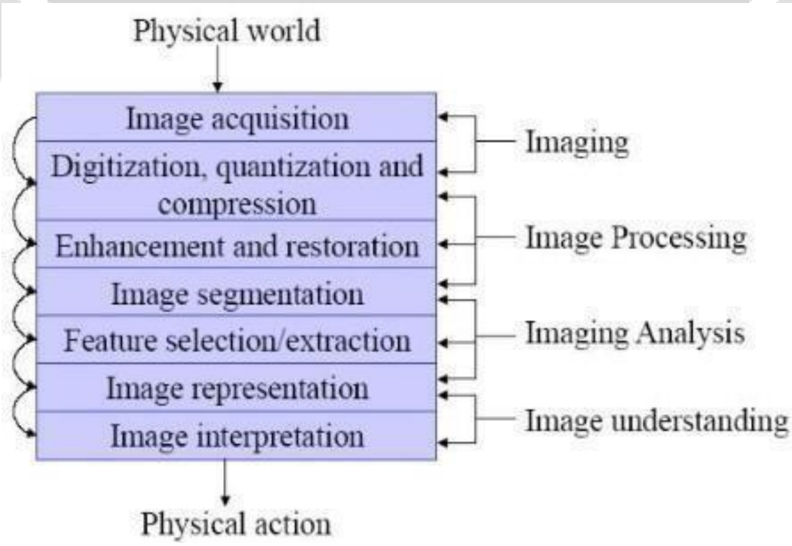

Fig 1: Fundamental steps in Image Processing

Each pixel holds a specific visual value, represented in several formats.Image processing involves a sequential set of operations executed on every pixel. These operations can be carried out iteratively to transform or enhance the image. Digital image processing generally follows three stages: pre-processing, enhancement, and information extraction. While the term "RGBA" is commonly used, it does not clarify the colour space used or whether alpha values have been pre-multiplied, which can be critical when interpreting image data formats in byte-oriented systems. This is the foundational step, where the image is captured from a hardware device. Also considered part of pre-processing, it may include operations like dilation and erosion, which use structuring elements and set theory. This technique improves the visual appeal of an image, often by adjusting contrast, brightness, or sharpness to emphasize important features. The Fourier Transform breaks down images into their frequency components, facilitating filtering, compression, and analysis in the frequency domain. The Discrete Fourier Transform (DFT) captures a sample set sufficient to reconstruct the spatial image entirely.

2: METHODOLOGY

The methodology adopted in this study is centered around the use of deep learning techniques, specifically leveraging the VGG16 convolutional neural network architecture, to detect and classify plant leaf diseases. Initially, a comprehensive dataset comprising images of both healthy and diseased plant leaves was collected. These images represented a wide range of plant species and included various categories of common leaf diseases. The dataset served as the foundational input for training and evaluating the deep learning model.Before the images could be fed into the VGG16 model, several preprocessing steps were carried out to standardize and optimize the data. All images were resized to 224x224 pixels, the input size expected by the VGG16 architecture. Normalization was applied to the pixel values to ensure consistency in data scale, typically transforming them to a range between 0 and 1. In order to improve the robustness of the model and to address the potential problem of overfitting due to a limited number of images, data augmentation techniques were also employed. These included transformations such as horizontal and vertical flipping, rotation, and slight zooming to synthetically expand the dataset and provide the model with more varied input conditions.The core of the methodology lies in the application of transfer learning through the VGG16 model. VGG16, originally trained on the ImageNet dataset, was used for its proven effectiveness in image classification tasks. The pre-trained model's convolutional layers were retained for feature extraction, while the top classification layers were replaced with custom fully connected layers tailored to the specific classes in the leaf disease dataset. This allowed the model to learn disease-specific patterns while benefiting from the generalized feature representations learned from the much larger ImageNet dataset.The training of the modified model was carried out using a supervised learning approach. The dataset was typically split into training and validation subsets, and the model was trained using a loss function such as categorical cross-entropy, which is suitable for multi-class classification problems.

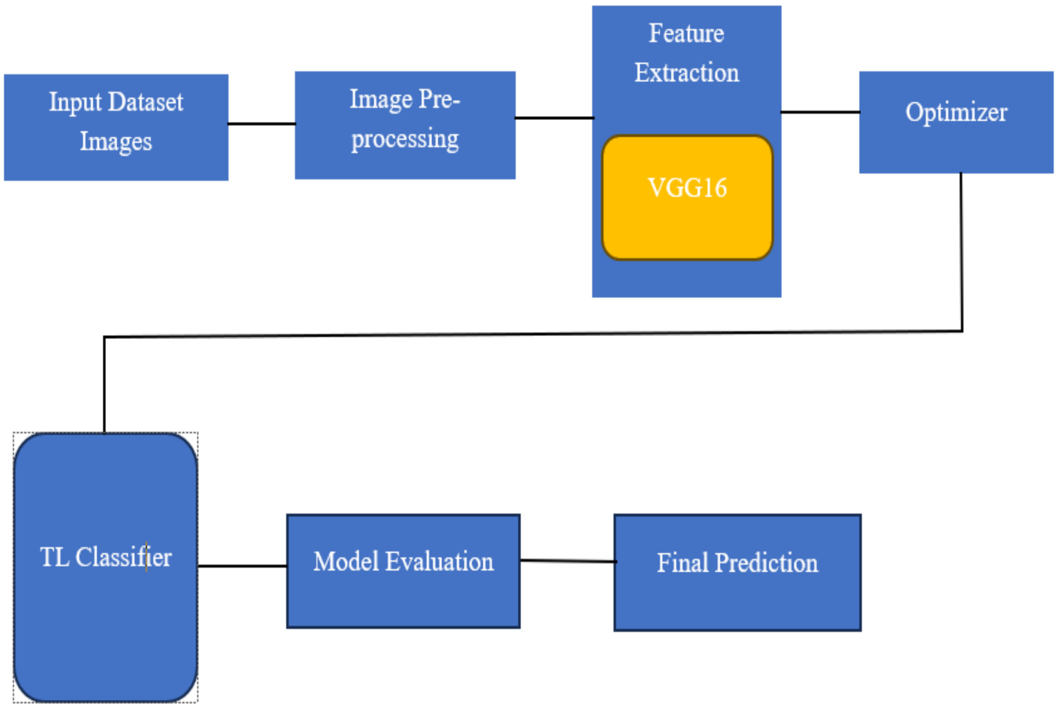

Fig 2: Block Diagram of CNN built on VGG16

An optimizer like Adam or stochastic gradient descent (SGD) was employed to update the weights of the network during training. Model performance was monitored through metrics such as accuracy, precision, recall, and F1-score, and adjustments were made through techniques like learning rate scheduling and early stopping to optimize training.Finally, the trained model was evaluated using a separate test set to measure its generalization performance. A confusion matrix was used to visualize and assess the classification outcomes for each class. In some cases, comparative analysis was performed with other deep learning models like ResNet or simple CNNs to highlight the performance advantages of the VGG16-based approach. Depending on the scope of the study, the model could also be integrated into a practical application, such as a mobile app or web platform, enabling real-time detection and diagnosis of plant leaf diseases in agricultural settings.

3.EXPERIMENTAL RESULTS

The experimental phase of this study was designed to rigorously train, validate, and evaluate a deep learning model for the purpose of identifying plant leaf diseases. The experiment began with the selection of an appropriate dataset, which is critical to the performance of any machine learning model. A widely recognized public dataset, such as the Plant Village dataset, was utilized due to its comprehensive collection of high-quality images of plant leaves categorized into multiple disease classes as well as healthy samples. This dataset includes thousands of labeled images spanning several plant species and disease types, providing a robust foundation for model development. Once the dataset was acquired, the images underwent a series of preprocessing operations aimed at ensuring compatibility with the VGG16 architecture and enhancing the learning process. All images were resized to a resolution of 224x224 pixels, which corresponds to the input dimensions required by VGG16. To improve training stability and convergence, pixel values were normalized. In addition, data augmentation techniques such as random rotations, flips, scaling, and translations were applied to artificially increase the diversity of the training data, thereby improving the model's ability to generalize to unseen data.The experiment employed transfer learning through the use of the VGG16 convolutional neural network, pre-trained on the ImageNet dataset. This approach allowed the model to benefit from the general features learned from a large and diverse dataset, while adapting the higher-level layers to the specific task of plant leaf disease classification. The top layers of the VGG16 model were removed and replaced with custom layers, typically including one or more dense (fully connected) layers followed by a softmax output layer corresponding to the number of disease categories in the dataset.Training was conducted on a workstation equipped with a suitable GPU to accelerate the computation process. The dataset was split into training, validation, and test sets—commonly in a ratio such as 70:15:15—to allow for both model optimization and unbiased evaluation. During training, categorical cross-entropy was used as the loss function, given the multi-class nature of the classification task. The Adam optimizer was selected due to its efficiency and adaptive learning capabilities. Several hyperparameters such as learning rate, batch size, and the number of epochs were tuned experimentally to achieve optimal performance.Model performance was continuously monitored on the validation set to prevent overfitting. Techniques such as early stopping and dropout were implemented to improve generalization. After training, the model was evaluated on the reserved test set. The accuracy, precision, recall, F1-score, and confusion matrix were computed to assess the classification performance across different disease categories.The experiment concluded with a comparison of the obtained results to existing models or baseline classifiers to demonstrate the effectiveness of the proposed approach. In some cases, performance visualization tools like ROC curves and training-validation loss plots were also generated to provide deeper insights into the model's behaviour during training.

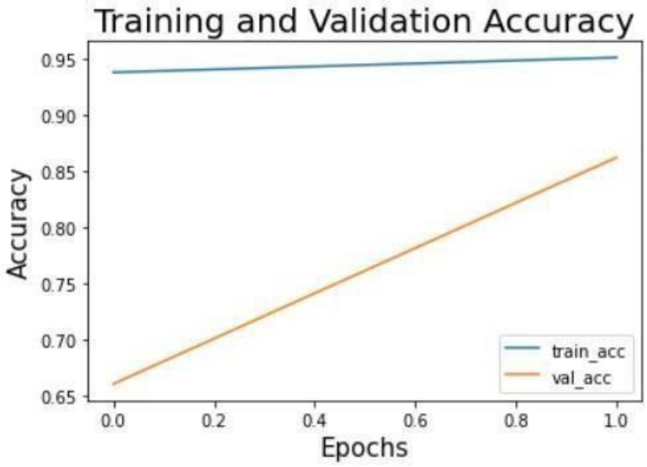

Fig 3: Accuracy of Training and Validation

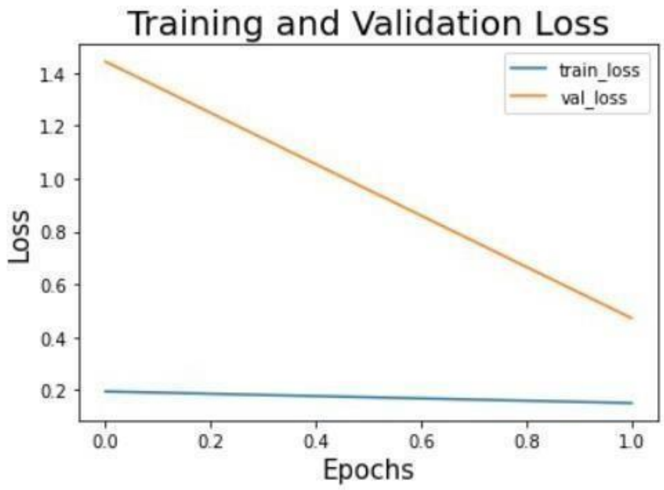

Fig 4: Loss of Training and Validation

4.CONCLUSION

Agricultural crops like fruits and vegetables are vital sources of food, oil, and fiber, and play a key role in a nation's economy. Consequently, crop losses caused by bacterial diseases can lead to substantial economic challenges for both farmers and the country. Timely identification and classification of plant leaf diseases are essential to improving crop yield and ensuring food security.This project introduces a system for detecting leaf diseases based on Convolutional Neural Networks (CNN), VGG16, and Transfer Learning. Leveraging the pre-trained VGG16 model for feature extraction minimizes the need for manual feature design and speeds up the training phase. With a high accuracy rate of 95% on test data, the system outperforms conventional methods and demonstrates strong potential for real-world agricultural applications.The system's scalability makes it versatile, enabling its use in various settings, from mobile apps for small-scale farmers to drone-based monitoring systems for large agricultural fields. Even with a limited dataset, the robustness of the model highlights the effectiveness of transfer learning in agriculture-related machine learning applications.This method not only offers a practical solution for detecting multiple plant diseases but also reduces reliance on agricultural experts for manual inspections. It equips farmers with fast and accurate information, facilitating timely interventions and minimizing the overuse of chemicals or wrong treatments.In conclusion, this research lays the groundwork for the development of AI-powered plant health monitoring systems. By integrating deep learning with agriculture, it shows how technology can enhance farming practices, improve efficiency, and contribute to global food security.

5.REFERENCES

[1] Khalid M. Hosny, Walaa M. El-Hady, Farid M. Samy, Eleni Vrochidou and George A. Papakostas Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern, 2023

[2] Wasswa Shafik, Ali Tufail, Abdallah Namoun, Liyanage Chandra tilak De Silva and Rosyzie Anna Awg Haji Mohd Apong A Systematic Literature Review on Plant Disease Detection: Motivations, Classification Techniques, Datasets, Challenges, and Future Trends ,2023

[3] K. P. Asha Rani and S. Gowrishankar Pathogen-Based Classification of Plant Diseases: A Deep Transfer Learning Approach for Intelligent Support Systems,2023

[4] Dehuan Luo , Yueju Xue , Xinru Deng , Bin Yang , Haifei Chen and Zhujiang title was Citrus Diseases and Pests Detection Model Based on Self-Attention YOLOV8,2023

[5] Emmanuel Moupojou, Appolinaire Tagne , Florent Retraint , Anicet adonkemwa, Dongmo Wilfried , Hyppolite Tapamo and Marcellin Nkenlifack FieldPlant: A Dataset of Field Plant Images for Plant Disease Detection and Classification with Deep Learning

[6] Ruchi Rani,Jayakrushna Sahoo,Sivaiah Bellamkonda,Sumit Kumar,Sanjeev Kumar Pippal Role of Artificial Intelligence in Agriculture: An Analysis and Advancements With Focus on Plant Diseases IEEE Access, 2023

[7] Vasileios Balafas , Emmanouil Karantoumanis , Malamati Louta and Nikolaos Ploskas Machine Learning and Deep Learning for Plant Disease Classification and Detection ,2023

[8] Serosh Karim Noon , Muhammad Amjad , Muhammad Ali Qureshi and Abdul Mannan Handling Severity Levels of Multiple Co-Occurring Cotton Plant Diseases Using Improved YOLOX Model,2022

[9] Sunil C. K. , Jaidhar C. D. and Nagamma Patil Cardamom Plant Disease Detection Approach Using EfficientNetV2 ,2022

[10] Lili Li , Shujuan Zhang and Bin Wang , Plant Disease Detection and Classification by Deep Learning A Review ,2021

[11] C. Sun, X. Zhou, M. Zhang and A. Qin, SE-VisionTransformer: Hybrid network for diagnosing sugarcane leaf diseases based on attention mechanism, Sensors, vol. 23, no. 20, pp. 8529, Oct. 2023.

[12] H. Tian, T. Wang, Y. Liu, X. Qiao and Y. Li, Computer vision technology in agricultural automationâ€”A review, Inf. Process. Agricult., vol. 7, no. 1, pp. 1- 19, 2020.

[13] D. P. Hughes and M. Salathe, An open access repository of images on plant health to enable the development of mobile disease diagnostics, arXiv:1511.08060, 2015.

[14] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat and N. Batra, PlantDoc: A dataset for visual plant disease detection, Proc. 7th ACM IKDD CoDS 25th COMAD, pp. 249-253, Jan. 2020.

[15] M. Adi, A. K. Singh, H. Reddy A. Y. Kumar, V. R. Challa, P. Rana and U. Mittal, An overview on plant disease detection algorithm using deep learning, Proc. 2nd Int. Conf. Intell. Eng. Manage. (ICIEM), pp. 305-309, Apr. 2021.

[16] M. Sharif, M. A. Khan, Z. Iqbal, M. F. Azam, M. I. U. Lali and M. Y. Javed, Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection, Comput. Electron. Agricult., vol. 150, pp. 220-234, Jul. 2018.

[17] Y. Mekonnen, S. Namuduri, L. Burton, A. Sarwat and S. Bhansali, Machine learning techniques in wireless sensor network based precision agriculture, J. Electrochem. Soc., vol. 167, no. 3, Jan. 2020.

[18] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris and D. Bochtis, Machine learning in agriculture: A comprehensive updated review, Sensors, vol. 21, no. 11, pp. 3758, May 2021.

[19] L. Li, S. Zhang and B. Wang, Plant disease detection and classification by deep learningâ€”A review, IEEE Access, vol. 9, pp. 56683-56698, 2021.

[20] J. Liu and X. Wang, Plant diseases and pests detection based on deep learning: A review, Plant Methods, vol. 17, no. 1, pp. 1-18, Dec. 2021.