iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: VIII Month of publication: August 2024

DOI: https://doi.org/10.22214/ijraset.2024.63940

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VIII Aug 2024- Available at www.ijraset.com

Coffee Leaf Disease Detection Using CNN

Anilkumar Shetty[1], Dr. Kusumadhara S[2], Dr. Bhagya H K[3], Dr. Savitha M[4]
[1]*Department of Digital Electronics and Communication Engineering KVG College of Engineering, Sullia*
[2]*Dr. Kusumadhara S. Professor & Head, Dept. of E&C Engineering, KVGCE, Sullia*
[3]*Dr. Bhagya H K. Professor, Dept. of E&C Engineering, KVGCE, Sullia*
[4]*Dr. Savitha M. Professor, Dept. of E&C Engineering KVGCE, Sullia*

Abstract: The project presents a comprehensive approach to detecting coffee leaf diseases through deep learning techniques, specifically utilizing Convolutional Neural Networks (CNNs). The system is built upon the VGG16 architecture, which has been fine-tuned to recognize and classify four types of coffee leaf conditions: Minor, Nodisease, Phoma, and Rust. A comprehensive preprocessing pipeline has been established to improve image quality, guaranteeing that the model is trained on well-optimized data. Data augmentation is used during training to avoid overfitting and enhance the model's ability to generalize to new data. The model is validated on a separate test set, where it achieves an accuracy of approximately 88%, indicating a high level of reliability in disease detection. Additionally, the project includes tools for evaluating the model's performance through metrics like accuracy, True Positive Rate(TPR), True Negative Rate(TNR), and False Positive Rate(FPR), providing a detailed analysis of its effectiveness. The final system is also equipped with a user-friendly interface for real-time disease detection, which can significantly benefit coffee farmers by enabling early intervention and reducing the impact of diseases on crop production. This project underscores the potential of CNNs in precision agriculture, paving the way for more intelligent and automated farming solutions.

Keywords: CNN, VGG16, TPR, TNR and FPR

I. INTRODUCTION

Coffee is one of the most favoured beverages globally and stands as a major commercial product, grown in over 50 countries around the world. Each day, an estimated 2.25 billion cups of coffee are consumed globally. Its cultivation and trade support millions of livelihoods, from farmers to distributors. Coffee's economic impact is substantial, with global markets constantly evolving to meet rising consumer demand.

In addition to its economic significance, coffee has cultural and social importance, often serving as a central element in daily routines and social interactions. India is one of the leading coffee producers in the world, recognized for its top-quality coffee grown under shade rather than in direct sunlight.

The country's unique growing conditions, including its diverse climate and soil types, contribute to the distinctive flavors and characteristics of its coffee. Additionally, India’s traditional methods of cultivation, such as intercropping with spices and other plants, enhance the sustainability and richness of its coffee production.

In the 2023-24 season, India produced approximately 350,000 metric tons of coffee, with around 80% of it being exported internationally. The revenue generated from the coffee industry was projected to reach USD 1.2 billion in 2024, with the market expected to continue growing at an annual rate of 7.4%. This growth reflects increasing global demand for high-quality coffee and India's strengthening position in the international market. Moreover, investments in sustainable practices and technology are anticipated to further enhance the industry’s profitability and environmental impact[11]. However, coffee plants are highly susceptible to pests and diseases, leading to crop losses estimated at 20-30% of the country's total production[12]. Coffee is a valuable global commodity but faces threats from leaf diseases like Phoma and Rust. Early detection and accurate diagnosis are essential to manage these diseases and prevent significant crop losses. Conventional manual inspection methods are both time-intensive and susceptible to mistakes.

In recently deep learning and computer vision offer automated solutions. This project utilizes Convolutional Neural Networks (CNNs), specifically the VGG16 architecture, to detect and classify coffee leaf diseases. By applying Transfer Learning and Data Augmentation, the model aims to enhance disease identification and support farmers in improving coffee yield and reducing economic impacts[1].

[1]https://en.wikipedia.org/wiki/Economics_of_coffee

II. LITERATURE REVIEW

Agriculture is vital to sustaining civilization, comparable to technological progress. A major challenge in agriculture is safeguarding plants from diseases caused by insects and natural enemies. Historically, this has involved manual plant inspection. However, advancements in Data Science and technology have introduced new approaches to disease detection. Convolutional Neural Networks (CNNs) offer a powerful method for diagnosing plant diseases by analyzing images of plant leaves, achieving an accuracy rate of 86% in disease detection. While alternative methods exist, CNNs enhance the diagnostic process by simplifying it and increasing efficiency [2].

The effectiveness of chemical control for a disease is influenced by the stage of disease development at which the fungicide is administered. Applying the fungicide at the right time can significantly improve its efficacy in managing the disease. Conversely, using it at an inappropriate stage may reduce its effectiveness and lead to suboptimal control of the disease [4]. Improper and negligent use of pesticides can lead to a buildup of resistance in pathogens over time, which can significantly diminish plants' ability to resist diseases. This resistance can result in increased treatment costs and reduced crop yields. Therefore, accurately and promptly identifying plant diseases is vital for effective precision agriculture. Timely detection allows for targeted interventions, minimizing the risk of resistance development and improving overall crop health [5].

Soft computing [13] is vital for the precise and prompt detection of biotic stresses affecting coffee plant leaves. Techniques like genetic algorithms [6], machine learning, and computer vision are actively researched for disease detection. Image processing [7] and expert systems [8] are also employed to enhance diagnosis accuracy. These methods collectively improving the ability to identify and manage diseases, ensuring timely intervention and better crop management. Recent advancements in these areas further support the development of effective solutions for monitoring and addressing plant health issues. These techniques have experienced limited adoption because of their constrained efficiency and success rates. According to [9], integrating image processing with artificial neural networks achieved a success rate of 90%. Although some plant diseases lack prominent visual symptoms, many common coffees leaf diseases exhibit clear and recognizable signs. Effectively managing these diseases necessitates methods that can precisely detect visible symptoms, thereby enhancing overall disease control and crop health. Advances in technology and methodology are essential for enhancing the detection and management of coffee leaf diseases[5].

[5]https://www.britannica.com/science/coffee-rust

III. METHODOLOGY

The project methodology involves the use of Convolutional Neural Networks (CNNs) for detecting and classifying coffee leaf diseases. The process begins with data collection, where images of coffee leaves with various disease conditions are gathered and organized into categories. The images undergo preprocessing, including resizing, normalization, and augmentation, to enhance the dataset and improve model performance is observed in fig1. The CNN model is built using the VGG16 architecture, a well-known deep learning model pre-trained on a large image dataset (ImageNet). Transfer learning is employed, where the VGG16 model is fine-tuned with the coffee leaf dataset, allowing the network to learn specific features related to the diseases. The model's layers are modified to fit the new classification task, and the final layers are trained to output the probability of each disease class.

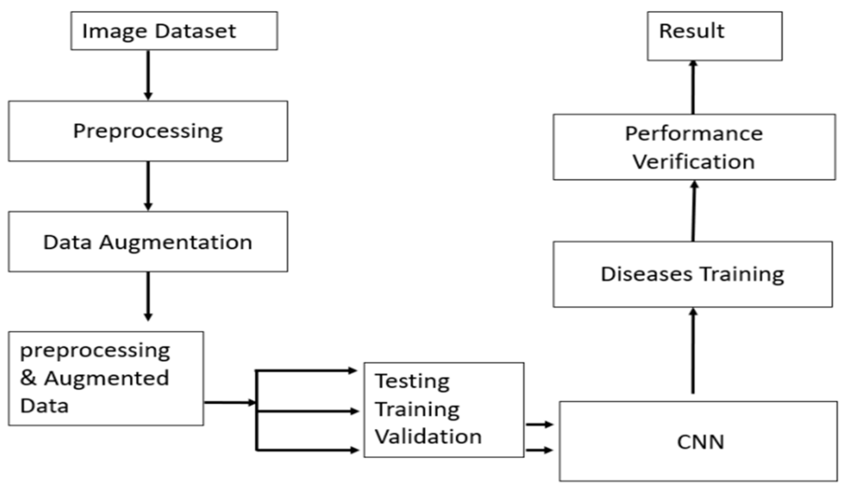


Fig 1. Methodology Flow-chart

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VIII Aug 2024- Available at www.ijraset.com

Data augmentation techniques, such as rotation, zoom, and horizontal flipping, are applied to the training set to increase the diversity of the images and prevent overfitting. The model is then trained on the augmented dataset, and its performance is evaluated using metrics like accuracy, True Positive Rate (TPR), and False Positive Rate (FPR). The final trained model is tested on a separate validation set to assess its ability to generalize to unseen data. The methodology also includes the implementation of a system that can process new images of coffee leaves and predict the disease type, providing a practical tool for farmers and agricultural stakeholders to diagnose diseases early and take necessary action to protect their crops.

IV. DATA SET

The dataset[10] for the coffee leaf disease detection project is organized into distinct folders for training, testing, and preprocessing. Each folder contains images classified into four categories: Minor, Nodisease, Phoma, and Rust. This classification is essential for training and evaluating the performance of the Convolutional Neural Network(CNN) model.

A. Training Data-set

The training dataset consists of images from four distinct categories of coffee leaf diseases: Minor with 332 images, Nodisease with 284 images, Phoma with 388 images, and Rust with 260 images, totaling 1264 images. These images are used to train the CNN model by teaching it to recognize and differentiate between the disease categories. During the training process, the model adjusts its parameters to minimize the discrepancy between its predictions and the actual labels, effectively learning the unique patterns and characteristics linked to each disease type. Some example leaf images from data set are in below fig 2.

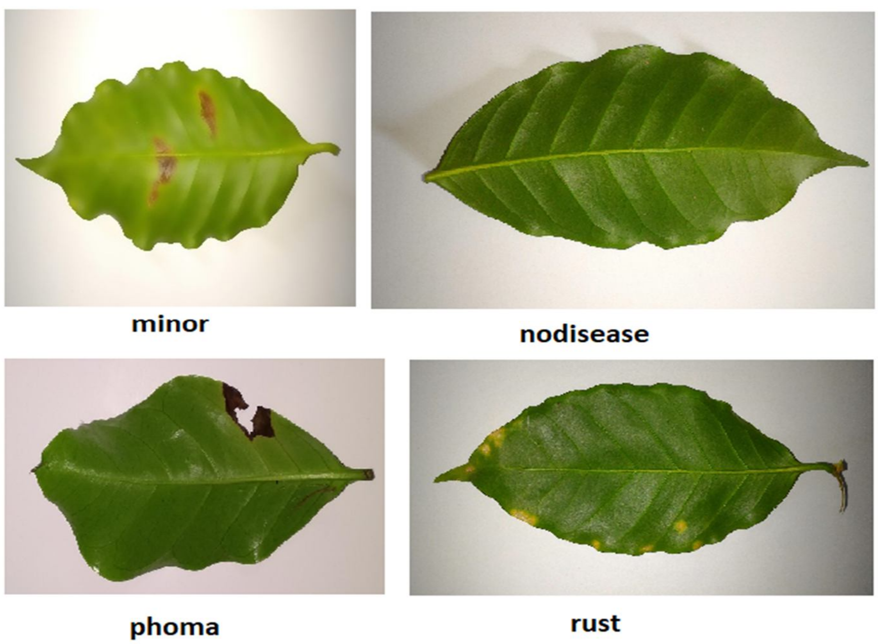


Fig2. Sample of coffee leaves in Leaf Dataset

B. Testing Dataset

The testing dataset, comprising 400 images, is used to evaluate the model's performance after training. It includes 128 images of Minor, 116 images of Nodisease, 96 images of Phoma, and 60 images of Rust. This dataset is crucial for assessing the model’s ability to generalize to new, unseen data and determines its accuracy and effectiveness in correctly classifying images that were not included in the training process.

C. Preprocessed Dataset

The preprocessed dataset consists of 1264 images, categorized as follows: 332 images of Minor, 284 images of Nodisease, 388 images of Phoma, and 260 images of Rust. These images have been resized to 256x256 pixels and filtered using a bilateral filter to reduce noise while maintaining essential features. This standardized dataset is used for both training and testing, ensuring consistency in format and quality across all images.

D. Image Processing

To prepare the images for the model, several preprocessing techniques are applied. Each image is resized to 256x256 pixels to maintain uniform input size. A bilateral filter is used to enhance image quality by smoothing out noise while preserving edges. This preprocessing is crucial for improving the model's ability to extract meaningful features from the images.

E. Data Augmentation

To enhance the model's robustness, data augmentation techniques are applied to the training images. These techniques include rotation, shifting, shearing, zooming, and horizontal flipping. Data augmentation helps the model generalize better by simulating various conditions and variations that it might encounter in real-world scenarios. This process increases the diversity of the training data and improves the model's ability to handle different types of image distortions and variations.

F. Model Training and Evaluation

The CNN model, built on the VGG16 architecture, is trained using the processed and augmented images. The VGG16 model, pre-trained on the ImageNet dataset, serves as the base for feature extraction, while additional layers are fine-tuned for the specific task of coffee leaf disease classification. The model is trained using categorical crossentropy as the loss function and the Adam optimizer.

After training, the model's performance is evaluated using metrics such as accuracy, True Positive Rate (TPR), True Negative Rate (TNR), and False Positive Rate (FPR). These metrics provide insights into the model's ability to correctly classify images and its performance across different disease categories.

V. RESULT AND DISCUSSION

The results of the coffee leaf disease detection model will be showcased on the web page, featuring detailed performance metrics for each disease category: Minor, Nodisease, Phoma, and Rust. The display will include overall classification accuracy and specific insights into how well the model identifies each condition. This information reflects the effectiveness of the model in distinguishing between various coffee leaf diseases based on the preprocessed dataset used for training and evaluation. The web page result is shown in fig 3.

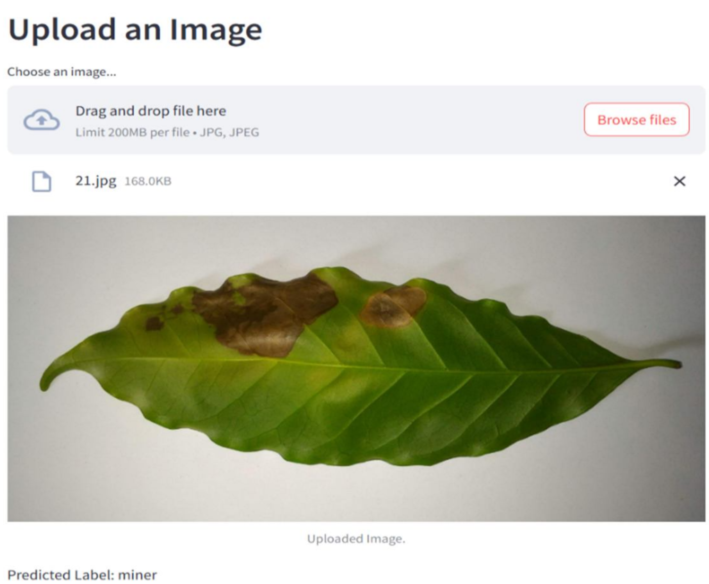


Fig3. Web page output

From fig.4 the confusion matrix for coffee leaf disease detection model evaluates the model's performance in classifying each disease category by comparing predicted versus actual labels. It reveals True Positives Rate (TPR) where diseases are correctly predicted, False Positives Rate(FPR) where diseases are incorrectly identified, and True Negatives Rate (TNR) where non-diseases are correctly recognized. This matrix is crucial for understanding the model's accuracy and effectiveness, helping to pinpoint strengths and areas needing improvement in the disease classification process.

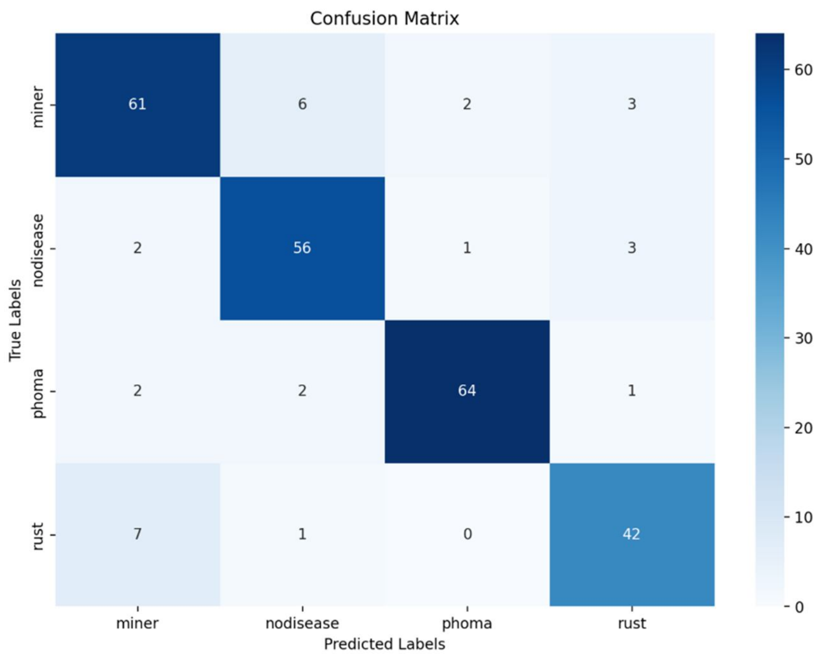


Fig4. Confusion Matrix

The True Positive Rate (TPR) measures the proportion of actual positive cases correctly identified by the model, showing how well it detects each disease. The False Positive Rate (FPR) indicates the proportion of negative cases incorrectly classified as positive, highlighting potential misclassifications. The True Negative Rate (TNR) reflects the proportion of actual negative cases accurately recognized, demonstrating the model's effectiveness in identifying healthy leaves and minimizing false alarms. The Result is shown in fig 5.

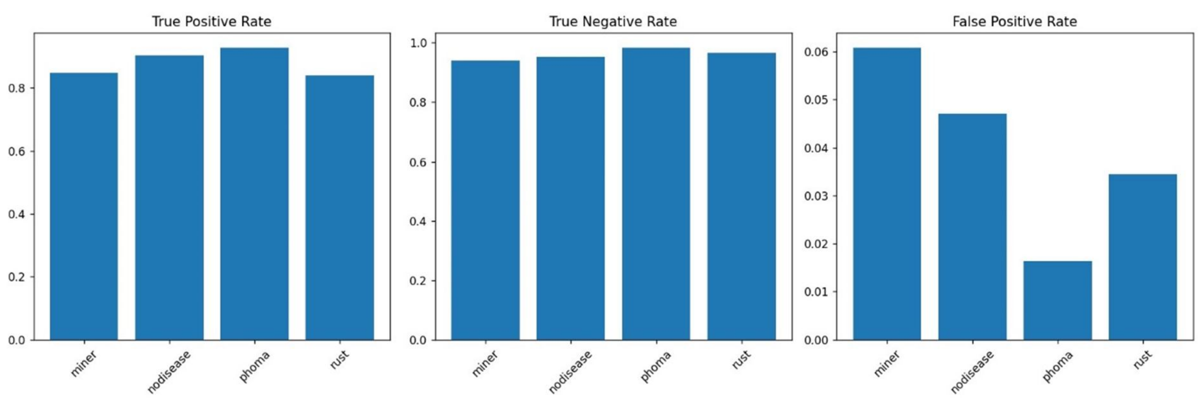


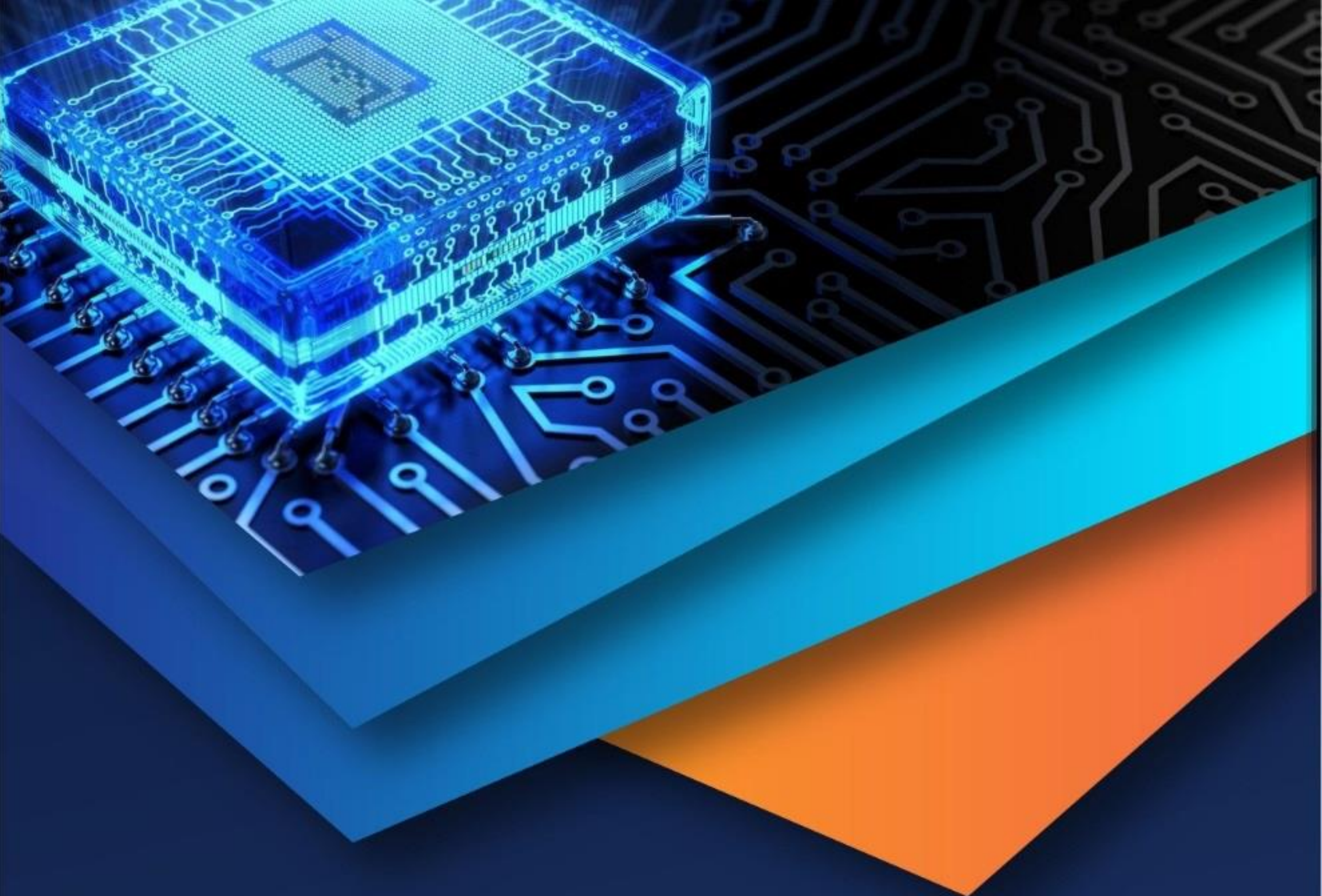
Fig5. TPR, TNR, FPR

VI. CONCLUSION

The coffee leaf disease detection model, utilizing a fine-tuned VGG16 architecture, demonstrates robust performance in classifying different coffee leaf diseases. The model achieves an accuracy of approximately 88.1%, with commendable True Positive Rates (TPR) across all categories, indicating effective detection of diseases. The True Negative Rate (TNR) is high, reflecting the model's reliability in identifying healthy leaves. Additionally, the False Positive Rate (FPR) is relatively low, signifying minimal misclassification of non-disease cases. These metrics affirm the model's effectiveness and potential for practical deployment in diagnosing coffee leaf diseases.

REFERENCES

[1] Disease Detection in Coffee Plants Using Convolutional Neural Network(CNN) - Manoj Kumar, Pranav Gupta, Puneet Madhav, Sachin Computer Engineering Dept. Delhi Technological University, New Delhi , India.
[2] Plant Disease Detection Using Convolution Neural Network(CNN) - M. Shobana; Vaishnavi S; Gokul Prasad C; Pranava Kailash SP; Madhumitha K P; Nitheesh C; Kumaresan N- 2022
[3] Early Disease Detection in Plants using CNN - Tejaswinia,*,Priyanka Rastogia , Swayam Duaa , Manikantaa , Vikas Dagara – 2023
[4] Ricardo Balardin. Factors affecting fungicide efficacy in the tropics, fungicides, odile carisse, intechopen, doi: 10.5772/13583.
[5] S. A. Miller, F. D. Beed, and C. L. Harmon, “Plant disease diagnostic capabilities and networks,” Annual Review of Phytopathology, vol. 47, pp. 15–38.
[6] P. Marcos, n. L. Silva rodovalho and a. R. Backes, "coffee leaf rust detection using genetic algorithm," 2019 xv workshop de visão computacional (wvc), são bernardo do campo, brazil, 2019, pp. 16-20.
[7] L. Han, m. S. Haleem and m. Taylor, "a novel computer visionbased approach to automatic detection and severity assessment of crop diseases," science and information conference (sai), london, pp. 638-644.
[8] Mettleq, A. S. A., & Abu-Naser, S. S. (2019). A Rule Based System for the Diagnosis of Coffee Diseases. International Journal of Academic Information Systems Research (IJAISR), 3(3), 1-8.
[9] Suhartono, Derwin & Aditya, Wahyu & Lestari, Miranty & Yasin, Muhammad. Expert System in Detecting Coffee Plant Diseases. International Journal of Electrical Energy. 156- 162. 10.12720/ijoee.1.3.156-162.
[10] https://www.kaggle.com/datasets/noamaanabdulazeem/jmuben-coffee-dataset
[11] https://www.statista.com/outlook/30010000/119/coffee/india#marketglobalrevenue.
[12] https://economictimes.indiatimes.com/news/economy/agriculture/india-toget-594349-to-control-leaf-rust-in-coffee/articleshow/2466942.cms
[13] Kumara, V., & Ganesan, E. (2024). Efficient Wastewater Treatment Optimisation with Solow-Polasky-JAYA Algorithm and Self-Organising Fuzzy Sliding Mode Control. Pro-cess Integration and Optimization for Sustainability, 1-13.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY