

## Coffee Leaf Disease Detection using Deep Learning

*A Term paper report submitted in partial fulfillment of the requirement for the award of degree*

**BACHELOR OF TECHNOLOGY**

**in**

**COMPUTER SCIENCE AND ENGINEERING**

*Submitted By*

**Gedela Sai Chandu**

#### 22341A0557

*Under the esteemed guidance of*

#### Mr. G. Suneel Assistant Professor,

**Dept. of Computer Science & Engineering**

**GMR Institute of Technology**

**An Autonomous Institute Affiliated to JNTU-GV, Vizianagaram** (Accredited by NBA, NAAC with ‘A’ Grade & ISO 9001:2015 Certified Institution) **GMR Nagar, Rajam – 532127**

#### Andhra Pradesh, India November 2024



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**CERTIFICATE**

*This is to certify that term paper report titled* **“Coffee Leaf Disease Detection using Deep Learning”** *submitted by* **Gedela Sai Chandu** *bearing* **Reg. No: 22341A0557** *has been carried out in partial fulfillment for the award* ***of Bachelor* of Technology** *in* **Computer Science and Engineering** *of* **GMRIT, Rajam** *affiliated to* **JNTUGV, Vizianagaram** *is a record of bonafide work carried out by them under my guidance & supervision. The results embodied in this report have not been submitted to any other University or Institute forthe award of any degree.*

Signature of the Supervisor Signature of the H.O.D

#### Mr. G. Suneel Dr. A V Ramana,

Assistant Professor Professor and Head

Department of CSE Department of CSE

GMRIT, Rajam. GMRIT, Rajam.

The report is submitted for the viva-voce examination held on ………………..

Signature of Internal Examiner Signature of External Examiner

## ACKNOWLEDGEMENT

It gives me an immense pleasure to express deep sense of gratitude to my guide, **Mr. G. Suneel Assistant Professor**, Department of Computer Science and Engineering of whole hearted and invaluable guidance throughout the report. Without her sustained and sincere effort, this

report would not have taken this shape. She encouraged and helped me to overcome various

difficulties that I have faced at various stages of my report.

I would like to sincerely thank **Dr. A V Ramana,** Professor & HOD, Department of Computer Science and Engineering, for providing all the necessary facilities that led to the successful completion of my report.

I take privilege to thank our Principal **Dr. C.L.V.R.S.V.Prasad**, who has made the atmosphere so easy to work. I shall always be indebted to them.

I would like to thank all the faculty members of the Department of Computer Science and Engineering for their direct or indirect support and also all the lab technicians for their valuable suggestions and providing excellent opportunities in completion of this report.

Gedela Sai Chandu

22341A0557

## ABSTRACT

Coffee is one of the most widely consumed beverages globally, and its production is significantly threatened by various leaf diseases, leading to substantial economic losses for farmers. To reduce this a deep learning-based approach for the detection of coffee leaf diseases utilizing Convolutional Neural Networks (CNNs) and transfer learning techniques are used. A diverse dataset of coffee leaf images are collected, representing healthy leaves and those affected by common diseases, including coffee leaf rust, bacterial blight, and leaf spot. The dataset was augmented through techniques such as rotation, flipping, and scaling to enhance model robustness. Transfer learning with pre-trained models, specifically Densenet and ResNet, fine-tuning them on our dataset to leverage their powerful feature extraction capabilities. The suggested model was examined and achieving an 82.3% accuracy and primary objective is to enhance the model’s accuracy in detecting leaf-based diseases by leveraging advanced deep learning techniques and this is crucial for agricultural practices.

***Keywords—*** Deep Learning, Convolutional Neural Networks (CNN), Transfer Learning, ResNet, DenseNet, Image Classification, Coffee Leaf Diseases Detection

**Coffee Leaf Disease Detection using Deep Learning 2024**

INDEX

|  |  |  |
| --- | --- | --- |
|  | **CONTENTS** | **PAGE NO** |
|  | CERTIFICATE | i |
|  | ACKNOWLEDGEMENT | ii |
|  | ABSTRACT | iii |
| Chapter 1. | INTRODUCTION | 1 |
| Chapter 2. | LITERATURE SURVEY | 2 |
| Chapter 3. | DESIGN | 12 |
| Chapter 4. | METHODOLOGY | 13 |
| Chapter 5. | CASE STUDIES | 20 |
| Chapter 6. | RESULTS & DISCUSSION | 22 |
| Chapter 7. | CONCLUSION | 26 |
|  | REFERENCES | 27 |

# INTRODUCTION

Coffee is one of the most widely consumed beverages globally, playing a crucial role in the economies of many countries. However, coffee production is increasingly threatened by various diseases that can significantly impact crop yield and quality. Among these, coffee leaf diseases such as Coffee Leaf Rust (Hemileia vastatrix) and Cercospora leaf spot are particularly damaging, leading to substantial losses for farmers. Traditional methods of disease detection often rely on visual inspections, which can be time- consuming and prone to human error. In recent years, advancements in deep learning have opened new avenues for automating the detection of plant diseases. Convolutional Neural Networks (CNNs) have emerged as a powerful tool for image classification tasks, enabling the effective identification of symptoms in leaf images. By analyzing pixel patterns and features, CNNs can learn to distinguish between healthy and diseased leaves with remarkable accuracy. Moreover, the application of transfer learning—where pre-trained models are fine-tuned on specific datasets—further enhances the performance of deep learning systems. This approach allows researchers to leverage existing knowledge gained from large-scale image datasets, reducing the need for extensive data collection and training time while achieving high accuracy in disease detection.

# LITERATURE SURVEY

#### Kumar, Manoj, Pranav Gupta, and Puneet Madhav. "Disease detection in coffee plants using convolutional neural network." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.

The study titled "Disease Detection in Coffee Plants Using Convolutional Neural Networks" by Manoj Kumar, Pranav Gupta, Puneet Madhav, and Sachin (2020) focuses on leveraging deep learning techniques for identifying coffee plant diseases. The research aims to understand the scope of coffee plant diseases, explore the use of data augmentation and transfer learning, and assess the role of Convolutional Neural Networks (CNN) in disease detection. By employing transfer learning and CNN, the methodology ensures efficient feature extraction, robustness against overfitting, and automation in disease detection. However, the approach has limitations, such as the inability to account for environmental factors and its dependency on large labeled datasets for accurate results. This study provides a foundation for future advancements in the automation of agricultural disease detection.Ouchra, H., Belangour, A., & Erraissi, A. (2023).

#### Signo, Samuel Dave R., Chloe Lei G. Tuquero, and Edwin R. Arboleda. "Coffee disease detection and classification using image processing: A Literature review." International Journal of Science and Research Archive 11.1 (2024): 1614-1621.

The paper "Coffee Disease Detection and Classification Using Image Processing" by Samuel Dave R. Signo, Chloe Lei G. Tuquero, and Edwin R. Arboleda (2024) explores advanced techniques for diagnosing coffee plant diseases. The study aims to investigate image processing techniques for disease detection, enhance classification accuracy through deep learning, and analyze various machine learning algorithms for classification tasks. Using methods such as SVM and KNN classifiers alongside CNN architectures like GoogleNet and RESNET, the study achieved high classification accuracy and efficiency in deep learning tasks the study achieved high classification accuracy and efficiency in deep learning tasks.. This work underscores the potential of combining traditional classifiers with deep learning for robust disease detection in agricultural applications. However, the approach faces challenges such as long training times for deep learning models and limited generalization to novel diseases. This work underscores the potential of combining traditional classifiers with deep learning for robust disease detection in agricultural applications.

#### Yamashita, João Vitor Yukio Bordin, and João Paulo RR Leite. "Coffee disease classification at the edge using deep learning." Smart Agricultural Technology 4 (2023): 100183.

The study titled "Coffee Disease Classification at the Edge Using Deep Learning" by Joao Vitor Yukio Bordin Yamashita and Joao Paulo R.R. Leite (2023) investigates the use of Convolutional Neural Networks (CNNs) and edge computing for efficient disease detection. This research focuses on assessing CNN architectures, specifically MobileNet, for their suitability in low-resource environments and examines dataset augmentation and processing techniques to enhance model performance. Advantages of this approach include edge computing capabilities for real-time processing, improved performance through cascading models, and robustness to limited data availability. However, challenges such as hardware constraints and limited dataset diversity hinder its broader application. This work highlights the potential of lightweight CNN architectures for decentralized agricultural disease detection.Pan, X., Wang, Z., Gao, Y., Dang, X., & Han, Y. (2022). Detailed and automated classification of landuse/land cover using machine learning algorithms in Google Earth Engine. *GeocartoInternational*, *37*(18), 5415-5432.

#### Abuhayi, Biniyam Mulugeta, and Abdela Ahmed Mossa. "Coffee disease classification using Convolutional Neural Network based on feature concatenation." Informatics in Medicine Unlocked 39 (2023): 101245.

The study "Rice Leaf Diseases Prediction Using Deep Neural Networks with Transfer Learning" explores the application of transfer learning through the InceptionResNetV2 model to predict and manage rice leaf diseases effectively. The research focuses on optimizing classification performance to improve agricultural outcomes and disease management practices. By utilizing deep learning and CNN techniques, the study achieves efficient feature extraction and enhances crop health monitoring. However, challenges such as high computational resource requirements, dependency on extensive data, and risks of overfitting limit its broader applicability. This work underscores the potential of advanced neural networks in transforming agricultural disease prediction systems.

#### Aufar, Yazid, and Tesdiq Prigel Kaloka. "Robusta coffee leaf diseases detection based on MobileNetV2 model." International Journal of Electrical and Computer Engineering 12.6 (2022): 6675

The paper "Coffee Disease Classification Using Convolutional Neural Network Based on Feature Concatenation" by Biniyam Mulugeta Abuhayi and Abdela Ahmed Mossa (2023) presents advancements in image processing and machine learning for plant disease detection. It focuses on the role of Convolutional Neural Networks (CNNs) in extracting high-level features and employs techniques like Random Forest (RF), Support Vector Machines (SVM), Decision Trees (DT), and K-Nearest Neighbors (KNN) for the classification. The study introduces customized segmentation and feature concatenation to

improve detection accuracy. Advantages include efficient image preprocessing and enhanced segmentation techniques, while limitations are noted in handling large image sizes, dependency on high- quality images, and risks of overfitting in complex models. This work underscores the potential of feature concatenation in improving plant disease classification accuracy.Patil, A., & Panhalkar, S. (2023).

#### Abd Algani, Yousef Methkal, et al. "Leaf disease identification and classification using optimized deep learning." Measurement: Sensors 25 (2023): 100643.

The study "Robusta Coffee Leaf Diseases Detection Based on MobileNetV2 Model" by Yazid Aufar and Tesdiq Prigel Kaloka (2022) focuses on classifying Robusta coffee leaves into "healthy" or "unhealthy" categories using CNN-based architectures. It evaluates the effectiveness of several CNN models, including MobileNetV2, ResNet50, DenseNet169, and InceptionResNetV2, emphasizing their performance on relatively small datasets. The methodology incorporates techniques such as data augmentation and lightweight model designs to optimize classifications. While the approach demonstrates efficiency with augmented data and adaptability across various models, challenges like lower accuracy in multiclass classification, dataset imbalance, and higher computational costs for larger architectures are noted. This research highlights the potential of MobileNetV2 for lightweight, efficient disease detection in agricultural applications.

#### Yebasse, M., Shimelis, B., Warku, H., Ko, J., & Cheoi, K. J. (2021). Coffee disease visualization and classification. Plants, 10(6), 1257.

The study "Leaf Disease Identification and Classification Using Optimized Deep Learning" by Yousef Methkal Abd Algani, Orlando Juan Marquez Caro, Liz Maribel Robladillo Bravo, Chamandeep Kaur, Mohammed Saleh Al Ansari, and B. Kiran Bala (2023) introduces an innovative approach for leaf disease detection using Ant Colony Optimization (ACO) combined with Convolutional Neural Networks (CNN). The research focuses on utilizing the ACO-CNN model to enhance. This methodology leverages deep learning and optimization for efficient feature extraction and scalability, making it suitable for various applications. However, challenges such as high resource requirements and longer training times are noted. This study highlights the potential of ACO-CNN for achieving superior performance in plant disease identification.

#### Tassis, Lucas M., Joao E. Tozzi de Souza, and Renato A. Krohling. "A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images." Computers and Electronics in Agriculture 186 (2021): 106191.

The study "Coffee Disease Visualization and Classification" **(**2021) focuses on improving the detection and classification of coffee leaf diseases through advanced visualization and deep learning techniques. The research emphasizes guided disease detection, compares various visualization methods, and seeks to enhance classification accuracy. It employs Grad-CAM, Grad-CAM++, Score-CAM, and CNN-based techniques to support disease identification without the need for bounding boxes. While the visualization methods provide clear insights, limitations include potential overfitting in simpler approaches and reliance on a small dataset. This work highlights the significance of combining visualization with classification for effective agricultural diagnostics.

#### Milke, Elisaye Bekele, Menbere Tesfaye Gebiremariam, and Ayodeji Olalekan Salau. "Development of a coffee wilt disease identification model using deep learning." Informatics in Medicine Unlocked 42 (2023): 101344.

The study "A Deep Learning Approach Combining Instance and Semantic Segmentation to Identify Diseases and Pests of Coffee Leaves from In-Field Images" (2021) explores advanced techniques for diagnosing coffee leaf diseases. By integrating multiple segmentation methods, including CNN, R-CNN, Unet, and PSPNet, the research develops a robust architecture capable of handling diverse natural environments. A new dataset is created to support automated diagnosis, focusing on enhancing model adaptability in real-world agricultural settings. Advantages include improved resilience to environmental variations and a versatile architecture suitable for multiple tasks. However, challenges such as misclassification in dense backgrounds, sensitivity to lighting conditions, and performance limitations with small datasets remain. This work emphasizes the potential of combining instance and semantic segmentation for accurate and scalable agricultural diagnostics.

#### Krishnamoorthy, N., Prasad, L. N., Kumar, C. P., Subedi, B., Abraha, H. B., & Sathishkumar, V. E. (2021). Rice leaf diseases prediction using deep neural networks with transfer learning. Environmental Research, 198, 111275.

The study "Development of a Coffee Wilt Disease Identification Model Using Deep Learning" focuses on creating a model that can detect and diagnose coffee wilt disease early through computer vision, reducing reliance on agricultural professionals. The methodology utilizes deep learning techniques, including those RetinaNET, YOLO, CenterNet, SSD, and CNN, to classify healthy and diseased coffee leaves based on visual indicators. This approach improves resource efficiency and offers a customizable, scalable solution for agricultural diagnostics. However, challenges such as it data dependency, high computational resource

requirements, and risks of overfitting remain key limitations. This work highlights the potential of deep learning in addressing critical agricultural challenges.

#### Martinez, Fredy, Holman Montiel, and Fernando Martinez. "A machine learning model for the diagnosis of coffee diseases." International Journal of Advanced Computer Science and Applications 13.4 (2022).

The study "A Machine Learning Model for the Diagnosis of Coffee Diseases" focuses on developing an embedded system to diagnose coffee diseases using machine learning techniques. By leveraging deep neural networks and ResNet architecture, the model identifies common coffee plant diseases with high accuracy and supports early disease detection. While the system demonstrates impressive performance in known conditions, its moderate accuracy with unknown images remains a limitation. This research highlights the potential of integrating machine learning into embedded systems for efficient and timely agricultural diagnostics.

#### Novtahaning, Damar, Hasnain Ali Shah, and Jae-Mo Kang. "Deep learning ensemble-based automated and high-performing recognition of coffee leaf disease." Agriculture 12.11 (2022): 1909.

The study "Deep Learning Ensemble-Based Automated and High-Performing Recognition of Coffee Leaf Disease" investigates leveraging transfer learning and data augmentation to improve the detection and classification of coffee leaf diseases. The methodology includes using advanced architectures such as SVM, DCN, CNN, YOLOv3-MobileNetv2, ResNet101, VGG16, DenseNet201, GoogLeNet, AlexNet, and VGG19. The research emphasizes enhancing data preprocessing and evaluating performance through extensive metrics. Advantages include effective data augmentation, efficient transfer learning, and high classification accuracy. However, challenges such as complex model architectures and sensitivity to image quality limit its scalability.

#### Montalbo, Francis Jesmar Perez, and Alexander Arsenio Hernandez. "Classifying Barako coffee leaf diseases using deep convolutional models." International Journal of Advances in Intelligent Informatics 6.2 (2020): 197-209.

The study "Classifying Barako Coffee Leaf Diseases Using Deep Convolutional Models" focuses on automating disease diagnosis through deep learning and transfer learning techniques. By employing those models such as Xception, ResNetV2-152, and VGG16 alongside stochastic gradient descent (SGD), the research enhances preprocessing and fine-tuning for accurate disease classification. While the study those

demonstrates the effectiveness of transfer learning and preprocessing improvements, it faces challenges like dataset limitations, risks of overfitting, and reliance on manual sample collection. This work emphasizes the potential of advanced convolutional models in improving agricultural disease diagnosis.

#### Poornam, S., and A. Francis Saviour Devaraj. "Image based Plant leaf disease detection using Deep learning." Inernational journal of computer communication and informatics 3.1 (2021): 53-65.

The study "Image-Based Plant Leaf Disease Detection Using Deep Learning" aims to develop a deep learning model for early identification and classification of plant diseases to reduce crop losses. Using images from databases like ImageNet, the research employs techniques such as CNN, MobileNet, AlexNet, and ResNet for training and testing the model. The approach offers advantages like automation and robustness in disease detection. However, challenges such as environmental variability and computational complexity impact its scalability. This work emphasizes the role of advanced deep learning models in improving agricultural disease management systems.

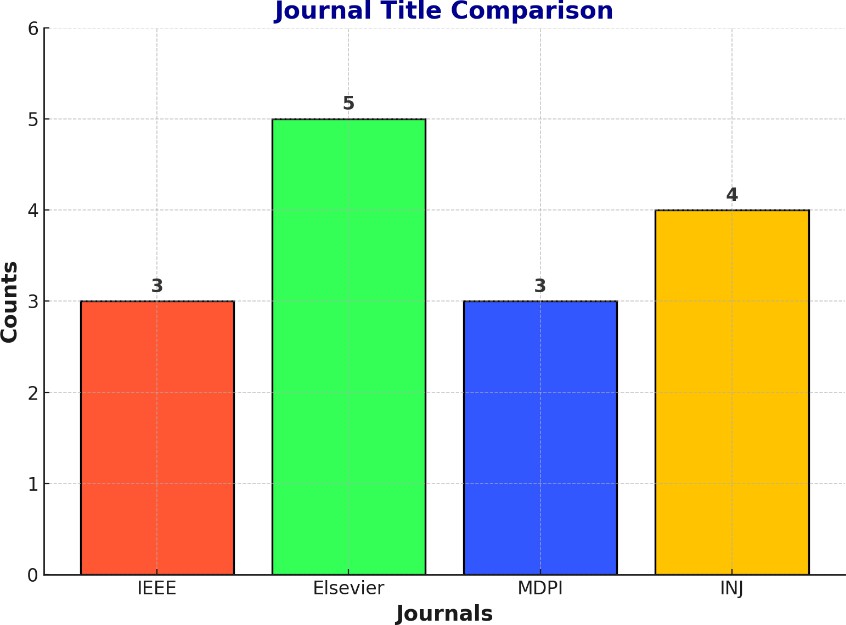
### Literature Survey Comparison Table:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Title** | **Year** | **Objectives** | **Limitations** | **Advantages** | **Performance metrics** | **Gaps** |
| **Reference 1** | Disease Detection in Coffee Plants Using Convolutional Neural  Networks | 2020 | Understand the Scope of Coffee Plant Diseases Explore Data  Augmentation and Transfer Learning | Environmental Factors Not Considered, Dependence  on Large datasets | Automated Feature Extraction Robust Against  Overfitting | Accuracy- 97.61%  Learning  Rate-0.005% | Overlapping, Infections, Size  Variations |
| **Reference 2** | Coffee disease detection and classification using image  processing | 2024 | Investigate Image Processing Techniques for Disease Detection Improve Classification  Accuracy with Deep Learning | Long Training Times for Deep Learning  Models, | High Accuracy of Classification Algorithms, Feature  Concatenation, | Accuracy- 97.31%  F1-score-  90% | Dataset expansion, Feature  incorporation |
| **Reference 3** | Coffee disease classification at the edge using  deep learning | 2023 | Assess the Use of Convolutional Neural Networks (CNNs)  Examine Edge Computing in  Disease Detection | Hardware Constraints, Limited Dataset  Diversity | Edge Computing, Improved Performance with Cascading Model, Robustness to  Limited Data | Precision- 96%  Accuracy- 98%  F1-Score- 97%  Recall-98% | Lightweight Models, Training  time |
| **Reference 4** | Coffee disease classification using Convolutional Neural Network based  on feature concatenation | 2023 | Advancements in Image Processing and Machine Learning  Role of Convolutional Neural Networks  (CNNs) | Large Image Size Handling, Dependency on High- Quality Images, Overfitting in Complex  Models | Efficient Image Preprocessing, Customized Segmentation, Feature  Concatenation | CNN-97.89%  Hybrid- 90.07% | EdgeAI, Data  Diversity |
| **Reference 5** | Robusta coffee leaf diseases detection based on  MobileNetV2 model | 2022 | Classify images of Robusta coffee leaves into "healthy" or "unhealthy" categories using  CNN models. | Effective with Augmented Data, Lightweight Model, Optimized for  Different Classifications | Lower Accuracy in Multiclass Classification, Imbalanced  Dataset | Accuracy- 99.93%  Precision- 99.87%  Recall-100% F1-Score-  99.93% | Fusion, Overlap |

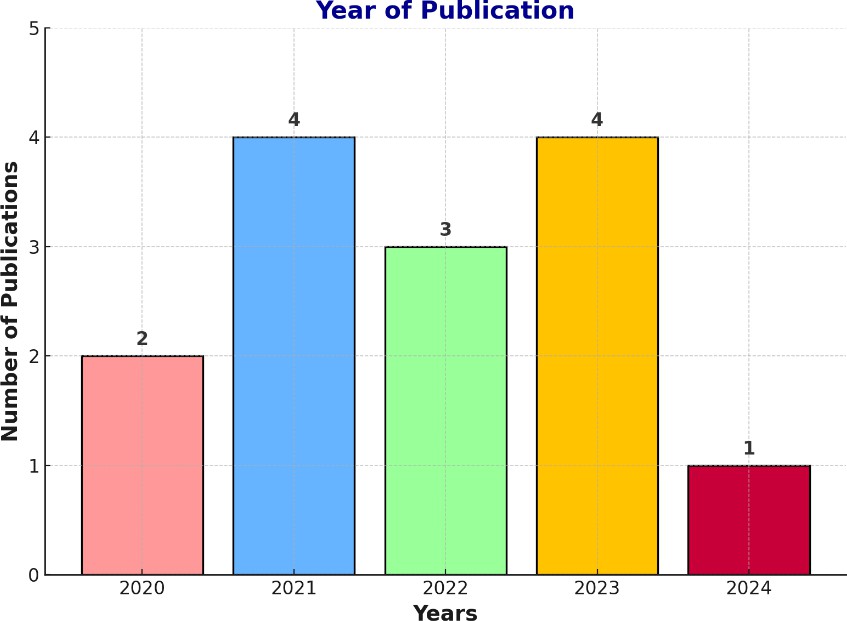
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Reference 6** | Leaf disease identification and classification using optimized  deep learning | 2023 | Utilizing Ant Colony Optimization (ACO) to  effectively extract features from images and classify  them using CNN. | Resource Intensive, Longer Training  Time | Efficient Feature Extraction,  Scalability | Accuracy- 99.98%  Precision- 99.6%  Recall- 99.6%  F1-Score- 99.99% | Lightweight, Scalability |
| **Reference 7** | Image based Plant leaf disease  detection using Deep learning | 2021 | Develop a deep learning model that identifies and classifies plant diseases early to minimize crop losses. | Environmental Variability, Computational  Complexity | Automation, Robustness | Accuracy- 95.71%  Activation Functions Pooling and  Convolution | Visualization Techniques, Mobile  Integration |
| **Reference 8** | A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-  field images | 2021 | Integration of Multiple Segmentation Techniques Creation of a New Dataset Automated  Diagnosis of Coffee  Leaf Diseases | Misclassification in Dense Backgrounds, Sensitivity to Lighting Conditions, Performance with Limited  Data | Improved Robustness to Natural Environments, Versatile  Architecture | Precision- 73.9% MIoU- 94.25%  Accuracy- 99.33%  Recall- 78.90% | Dataset Enrichment,  Robust Models |
| **Reference 9** | Development of a coffee wilt disease identification model using  deep learning | 2023 | Develop a model that can detect and diagnose coffee wilt disease early using computer vision, thus avoiding reliance on agricultural professionals. | Data Dependency, Computational Resources,  Overfitting Risks | Improved Resource Efficiency, Customizable  and Scalable | Accuracy- 97.9%  Loss-9.9% Learning Rate-0.0001  Activating  Functions- 97.3% | Multitasking,  Scalability, Automation |
| **Reference 10** | Rice leaf diseases prediction using deep neural networks  with transfer learning | 2021 | Apply transfer learning with the InceptionResNetV2 model  Improve disease  management in rice crops | High computational resources, Data dependency,  Risk of overfitting | Efficient feature extraction, Improved  agriculture outcomes. | KNN- 58.16% ANN- 79.04%  Naive- 53.47% | Data Limitation,  Algorithm Adaptability |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Reference 11** | A Machine Learning Model for the Diagnosis of Coffee  Diseases | 2022 | Develop an embedded system for  diagnosing coffee diseases | Moderate Accuracy with  Unknown Images | High Accuracy in Disease Classification,  Early Detection of Diseases | Accuracy 96.5% | Dataset  Standardization, Accessibility |
| **Reference 12** | Deep Learning Ensemble- Based Automated and High- Performing Recognition of Coffee Leaf  Disease | 2022 | Leverage transfer learning Enhance data with preprocessing and augmentation Evaluate performance with extensive  metrics | Complex Model Architecture, Sensitivity to  Image Quality | Effective Data Augmentation, Transfer Learning Efficiency, High  Classification Accuracy | Accuracy 97.31%,  Precision 95.7%,  F1-score 95.1% ,  Recall  95.2% | Dataset Augmentation Transfer  Learning |
| **Reference 13** | Classifying barako coffee leaf diseases using deep convolutional  models | 2020 | Disease Diagnosis Automation Transfer Learning and  Fine-Tuning | Dataset Limitations, Overfitting Risk, Manual Sample  Collection | Preprocessing, Enhancements,  Transfer Learning | Accuracy- 97%  True Positive Rate- 96.55%  Loss  Functions- 0.07 | Environment  Factors, Limited Dataset |
| **Reference 14** | An app to assist farmers in the identification of diseases and pests of coffee leaves using  deep learning | 2022 | Developing a convolutional neural network (CNN) for segmenting and classifying coffee leaf  lesions. | Computationa l Demand, Limited Data  Coverage | Real-Time Identification, Severe Stress  Estimation | Accuracy- 97.09%  MIou-  94.85% | Image Quality, Scalability |
| **Reference 15** | Coffee Disease Visualization and  Classification | 2021 | Compare Visualization Methods Improve Classification Accuracy Guided Disease  Detection | Overfitting in Naïve Approach,  Small Dataset | Visualization Support, No Bounding Box  Needed | Accuracy- 75% | Detailed- Lesions,  Class Similarity |

**Literature Survey Graphical Representation:**

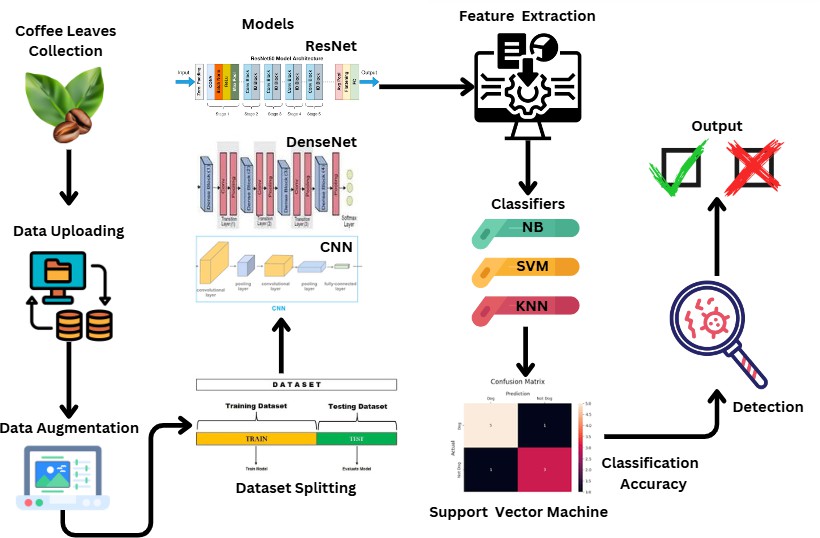


**Fig 1:** Journal Title Comparison



**Fig 2:** Year of Publication

# DESIGN



**Fig 3:** Model proposed in this study

This diagram illustrates the workflow for detecting and classifying diseases in coffee leaves using machine learning and deep learning techniques. The process begins with the collection of coffee leaves, followed by data uploading to a central repository. To enhance the dataset, data augmentation is performed, which generates variations of the existing images to increase model robustness. The dataset is then split into training and testing sets. Next, deep learning models like ResNet, DenseNet, and Convolutional Neural Networks (CNN) are employed for feature extraction, which captures critical patterns from the data. These extracted features are then fed into classifiers, including Naive Bayes (NB), Support Vector Machines (SVM), and K-Nearest Neighbors (KNN), to categorize the leaves as either healthy or diseased. The system evaluates performance through a confusion matrix, which measures classification accuracy by comparing predictions against actual outcomes. Ultimately, the workflow provides detection results, highlighting whether a leaf is healthy or diseased while emphasizing the importance of accuracy in disease identification. This approach combines deep learning and machine learning for precise agricultural diagnostics.

# METHODOLODY

### Coffee Leaves Collection:

Objective: Collect high-quality images of coffee leaves for disease analysis.

Images should be collected in various environmental conditions (lighting, angle, background) to ensure model generalization. Include diverse samples: healthy leaves, leaves with early disease symptoms, and severely affected leaves. Use high-resolution cameras or mobile devices for better feature recognition.

### Data Uploading:

Objective: Store images in a centralized system for processing and analysis.

Use cloud storage services like AWS S3, Google Drive, or local servers. Maintain organized directories, e.g., folders labeled by disease type or “healthy.” Ensure metadata is collected (e.g., location, time of capture, and weather conditions) for further analysis if required.

### Data Augmentation:

Objective: Enhance the dataset by increasing its size and diversity without collecting more raw images. Techniques:

Rotation: Rotate images at different angles (e.g., 90°, 180°). Flipping: Flip images horizontally and vertically.

Scaling: Resize or zoom in/out.

Brightness Adjustment: Simulate various lighting conditions. Gaussian Noise: Add noise to simulate real-world distortions.

### Dataset Splitting:

Objective: Prepare the dataset for training and testing machine learning models.

Training Set: Usually 70–80% of the dataset. Used to teach the model patterns. Testing Set: Remaining 20–30%. Used to measure model performance.

### Model Selection:

ResNet (Residual Neural Network):

Strengths: Handles deep architectures without vanishing gradients, making it suitable for complex datasets.

Architecture: Uses skip connections (shortcuts) to bypass some layers and improve gradient flow. DenseNet (Dense Convolutional Network):

Strengths: Efficient feature reuse through dense connections, reducing computational requirements. Architecture: Each layer gets input from all preceding layers, enhancing feature learning.

CNN (Convolutional Neural Network):

Strengths: Best for extracting spatial features (patterns, textures) from images.

Architecture: Consists of convolutional layers (for feature extraction), pooling layers (for dimensionality reduction), and fully connected layers (for classification).

### Feature Extraction:

Objective: Extract meaningful features (e.g., texture, color, shape) from images.

Deep learning models like ResNet and DenseNet perform automatic feature extraction. Features are high- dimensional vectors representing the important aspects of the image. These extracted features act as input to machine learning classifiers.

### Classifiers:

Naive Bayes (NB):

Based on probability distributions. Assumes features are independent, which may not always hold true for images. Lightweight but may underperform for complex datasets.

Support Vector Machine (SVM):

Finds an optimal hyperplane to separate classes. Works well with small datasets and high-dimensional spaces.

K-Nearest Neighbors (KNN):

Simple and interpretable. Classifies data points based on the majority class of their nearest neighbors. Computationally expensive for large datasets.

### Detection:

Objective: Use classifiers to identify whether a leaf is healthy or diseased.

Classifiers output probabilities for each class, and the highest probability determines the final prediction. Additional post-processing (e.g., thresholds) may be applied to refine results.

### Classification Accuracy:

Confusion Matrix: Shows True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN).

Accuracy:

The ratio of correctly predicted observations (both positive and negative) to the total observations. Accuracy is a good metric when the dataset is balanced.

Precision (Positive Predictive Value):

The ratio of correctly predicted positive cases to the total predicted positive cases. Precision is important when the cost of false positives is high.

Recall (Sensitivity or True Positive Rate):

The ratio of correctly predicted positive cases to the total actual positive cases. Recall is important when the cost of false negatives is high.

F1 Score:

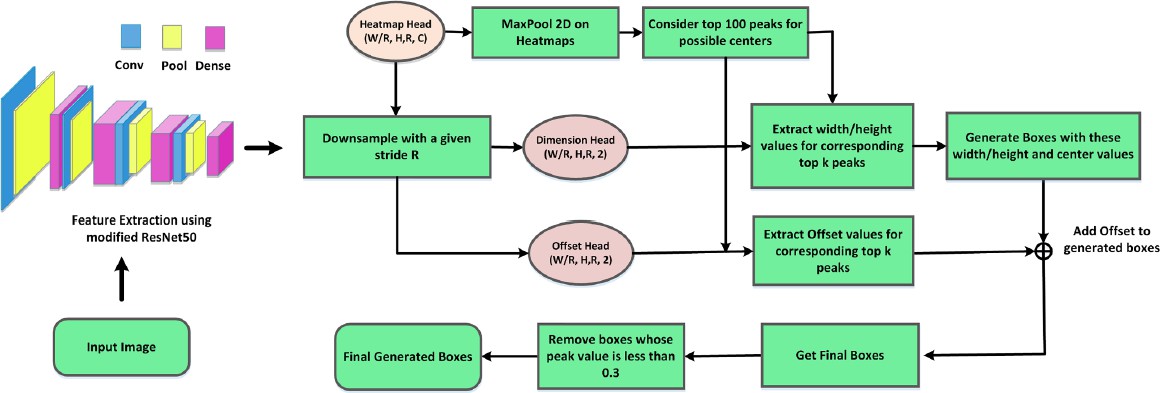
The harmonic mean of precision and recall. It is a better metric when there is an uneven class distribution or when you need a balance between precision and recall.

### Output:

Healthy Leaves: Model will gives the result as healthy. Diseased Leaves: Model will gives the result as unhealthy.

**Nawaz, Marriam, et al. "CoffeeNet: A deep learning approach for coffee plant leaves diseases recognition." *Expert Systems with Applications* 237 (2024): 121481.**

The process uses a modified ResNet-50 to extract features from the input image. A heatmap is generated to detect object centers, followed by a 2D Max Pooling operation to select the top 100 peaks. The Dimension Head predicts the width and height of bounding boxes, while the Offset Head fine-tunes their positions. Bounding boxes are generated using the heatmap centers, dimensions, and offsets. Low- confidence boxes (below a set threshold) are filtered out, resulting in final bounding boxes that accurately localize diseased regions in the image.

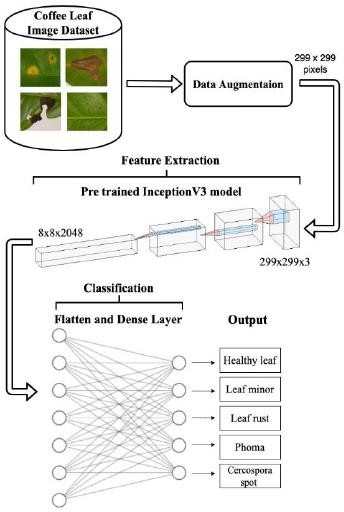


**Fig 4:** Architecture of modified ResNet50

* 1. **Data Preparation and Annotation:** Use the Arabica coffee leaf dataset, comprising 18,985 images across five classes (Phoma, Cescospora, Rust, Healthy, and Miner), prepared by annotating affected regions with bounding boxes.
  2. **CoffeeNet: Feature Extraction and Recognition Phases:** Implement an improved CenterNet model with a ResNet-50 backbone, integrated with a Convolutional Block Attention Module (CBAM) for capturing detailed, disease-specific features.
  3. **Heatmap, Dimension, and Offset Heads:** These three components improve object localization and class prediction, optimizing the model’s precision and accuracy.
  4. **Training Strategy:** CoffeeNet is trained in an end-to-end manner with multi-task loss functions, optimizing both classification and localization.
  5. **Evaluation Metrics:** Use metrics such as Intersection over Union (IoU), mean Average Precision (mAP), F1-score, and accuracy for performance validation.

#### Kumar, Manoj, Pranav Gupta, and Puneet Madhav. "Disease detection in coffee plants using convolutional neural network." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.

This process involves classifying coffee leaf images to detect diseases. The dataset undergoes data augmentation to enhance variability, resizing images to 299x299 pixels. A pre-trained InceptionV3 model is used for feature extraction, transforming images into a feature map (8x8x2048). These features are passed through flatten and dense layers for classification into categories: healthy leaf, leaf minor, leaf rust, phoma, and cercospora spot. This approach leverages transfer learning for accurate disease identification.

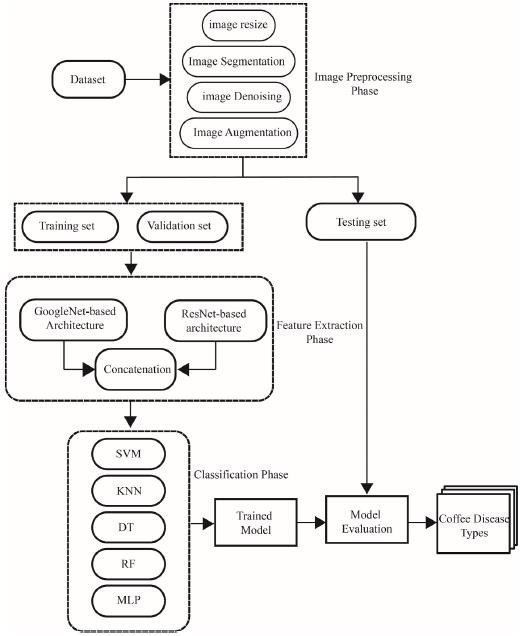


**Fig 5:** Network Architecture

1. **Coffee Leaf Image Dataset:** The process starts with a dataset containing images of coffee leaves. These images represent different categories: healthy leaves and leaves affected by various diseases, such as Leaf Minor, Leaf Rust, Phoma, and Cercospora Spot.
2. **Data Augmentation:** The images in the dataset are preprocessed using data augmentation techniques. This involves transformations such as resizing each image to 299 x 299 pixels and potentially applying techniques like rotation, flipping, and adjusting brightness. These transformations create variations in the data, which help the model generalize better by preventing overfitting.
3. **Feature Extraction with Pre-trained Inception V3 Model:** The augmented images are then fed into a pre-trained Inception V3 model, which is used here for feature extraction. Inception V3 is a Convolutional Neural Network (CNN) model that has been pre-trained on a large dataset (typically ImageNet) and can identify complex features from images. The input images (299x299x3 for RGB) pass through various convolutional layers in the Inception model, resulting in feature maps that represent essential patterns related to disease symptoms.
4. **Flatten and Dense Layers:** After feature extraction, the output from the Inception V3 model is flattened and passed through dense (fully connected) layers for classification. The dense layers serve as the decision-making part of the network, combining the extracted features to classify the images accurately.
5. **Output:** Finally, the model categorizes each image into one of the classes: Healthy Leaf, Leaf Minor, Leaf Rust, Phoma, or Cercospora Spot, based on the detected features.

#### Aufar, Yazid, and Tesdiq Prigel Kaloka. "Robusta coffee leaf diseases detection based on MobileNetV2 model." International Journal of Electrical and Computer Engineering 12.6 (2022): 6675

The flowchart illustrates a machine learning pipeline for classifying coffee disease types. It begins with the image preprocessing phase**,** where the dataset undergoes resizing, segmentation, denoising, and augmentation to improve image quality and diversity. The preprocessed images are divided into training, validation, and testing sets. In the feature extraction phase**,** features are extracted using two deep learning architectures, GoogleNet and ResNet, and then combined through concatenation to create a rich feature set. These features are passed to the classification phase**,** where algorithms such as SVM, KNN, Decision Tree, Random Forest, and MLP are used to train models. Finally, the trained models are evaluated on the testing set to classify coffee disease types effectively. This approach combines advanced deep learning and traditional classifiers for robust and accurate disease detection.



**Fig 6:** CNN based feature concatenation

1. **Image Preprocessing:** The dataset comprises 3,288 images resized to 32 × 32 RGB pixels to reduce computational complexity. Data augmentation techniques were used to expand the dataset to 7,067 images to prevent overfitting. A Gaussian filter was applied for noise reduction, with segmentation techniques like K-means, Otsu’s, and a custom method to extract the diseased parts.
2. **Feature Extraction:** The researchers used two deep learning models, specifically GoogLeNet and RESNET, to extract high-level features. The GoogLeNet-based architecture involves multiple inception modules that analyze the input at different scales. RESNET-based architecture includes skip connections, which help mitigate vanishing gradients, thus preserving the flow of information.These features were passed to classifiers like Multi-Layer Perceptrons (MLP), Random Forest (RF), Support Vector Machines (SVM), Decision Trees (DT), and K-Nearest Neighbors (KNN).
3. **Classification:** The combination of CNN-based architectures and traditional machine learning classifiers was used to identify and classify various coffee diseases. An ensemble technique was employed to stack different classifiers (KNN, RF, SVM, DT) for improved classification accuracy.
4. **Evaluation Techniques:** K-fold cross-validation with a value of 3 was applied for model validation. Several evaluation metrics, including accuracy, precision, recall, and F1-score, were calculated using the confusion matrix.

# CASE STUDIES

### Objective:

This study aims to leverage Convolutional Neural Networks (CNNs) and transfer learning techniques to accurately detect coffee leaf diseases, including Coffee Leaf Rust, Bacterial Blight, and Leaf Spot, using image data. By automating the disease detection process, the study seeks to address the limitations of traditional methods, offering a faster, more accurate, and scalable solution to assist coffee farmers globally.

### Problem Statement:

Coffee diseases pose a significant threat to global coffee production, directly impacting crop yield, quality, and the livelihoods of farmers. Traditional disease detection methods are largely reliant on manual visual inspections conducted by experts. These methods are not only time-consuming but are also prone to human error, inconsistency, and the limited availability of trained personnel, particularly in remote areas. Additionally, early detection and accurate diagnosis are critical for disease management, but existing methods often fail to meet these needs. Consequently, there is an urgent requirement for an automated, reliable, and scalable system that can effectively detect and classify coffee leaf diseases from visual data.

### Data Preparation:

To train robust machine learning models, a diverse dataset of coffee leaf images was collected, covering both healthy leaves and leaves affected by diseases such as Coffee Leaf Rust, Bacterial Blight, and Leaf Spot. The dataset encompassed variations in environmental conditions, lighting, and leaf orientation to ensure real-world applicability. To enhance the dataset's quality and prevent overfitting, data augmentation techniques, including image rotation, flipping, and scaling, were applied. This preprocessing step improved model robustness by introducing variability and ensuring that the models could generalize to unseen data. Additionally, diseased regions were segmented to focus the learning process on critical areas of interest, and all images were resized to uniform dimensions to ensure compatibility with deep learning architectures.

### Deep Learning Models:

To leverage the benefits of transfer learning, pre-trained CNN architectures, DenseNet and ResNet, were employed for feature extraction. These architectures, known for their high accuracy in image classification

tasks, were fine-tuned to adapt to the specific dataset of coffee leaf images. DenseNet’s densely connected layers facilitated efficient feature propagation and reuse, while ResNet’s residual connections mitigated the vanishing gradient problem, enabling stable training of deep models. The models were trained using TensorFlow on GPUs to accelerate computation and optimize performance, given the high resource demands of deep learning tasks.

### Key Steps:

Image preprocessing included resizing, segmentation, and data augmentation to ensure uniformity and robustness. Fine-tuning pre-trained DenseNet and ResNet architectures allowed these models to adapt to the coffee leaf dataset effectively. The trained models were evaluated based on their accuracy, precision, and ability to generalize to unseen images. DenseNet outperformed ResNet, achieving an accuracy of 84.56%, attributed to its dense connectivity. ResNet, however, demonstrated exceptional robustness, particularly in identifying subtle disease features.

### Challenges:

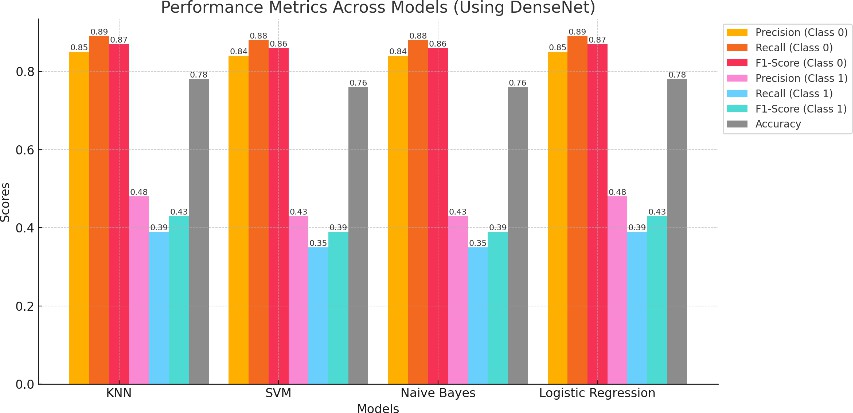
Variations in lighting conditions, background noise, and leaf appearance introduced complexities that impacted the models' robustness. Training deep learning models required significant computational power, making GPU usage essential. However, this dependency could limit the accessibility of the solution for small-scale farmers or regions with limited technological resources. Ensuring a balanced representation of all disease classes was critical to avoid model bias toward more prevalent diseases. This was addressed through careful dataset curation and augmentation.

### Impact and Outcomes :

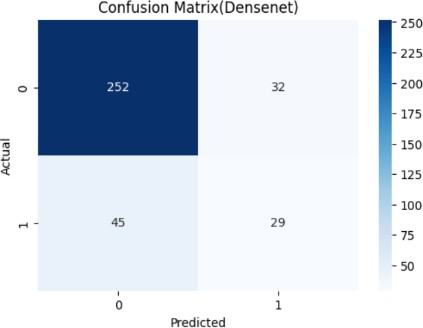
The system reduced the time and effort required for manual inspections, providing real-time disease detection and early warnings. This enabled farmers to take preventive measures promptly, minimizing crop damage and loss. DenseNet achieved an impressive accuracy of 84.56%, ensuring reliable and consistent disease classification. ResNet also performed well, particularly in identifying subtle disease features, underscoring the robustness of the proposed approach compared to traditional diagnostic methods. The methodology is highly adaptable and can be extended to detect diseases in other crops.

# RESULTS & DISCUSSIONS

**DenseNet:** The performance metrics demonstrate consistent results across the models, with Class 0 achieving higher precision, recall, and F1-scores (0.85, 0.89, and 0.87, respectively) compared to Class 1 (0.48, 0.39, and 0.43, respectively). This disparity indicates that the models perform significantly better on Class 0 than on Class 1. The overall accuracy remains steady at 78%, emphasizing the models' reliability for the majority class while highlighting the need for improvements in handling the minority class.



**Fig 7:** Performance metrics across models (using DenseNet)

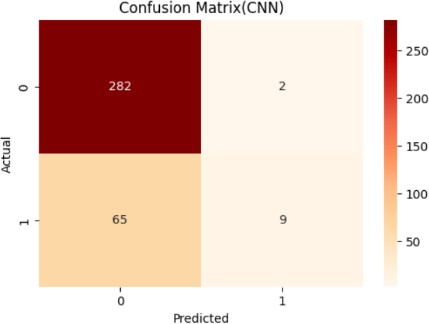


**Fig 8:** Confusion Matrix (DenseNet)

**CNN:** The performance metrics show a significant imbalance in the model's handling of classes. While Class 0 achieves high precision (0.81), recall (0.99), and F1-score (0.89), indicating strong performance, Class 1 struggles with low recall (0.12) and F1-score (0.21), despite a comparable precision of 0.82. The overall accuracy of 80% reflects the model's bias towards the majority class, emphasizing the need for strategies to improve the model's capability in identifying the minority class effectively.

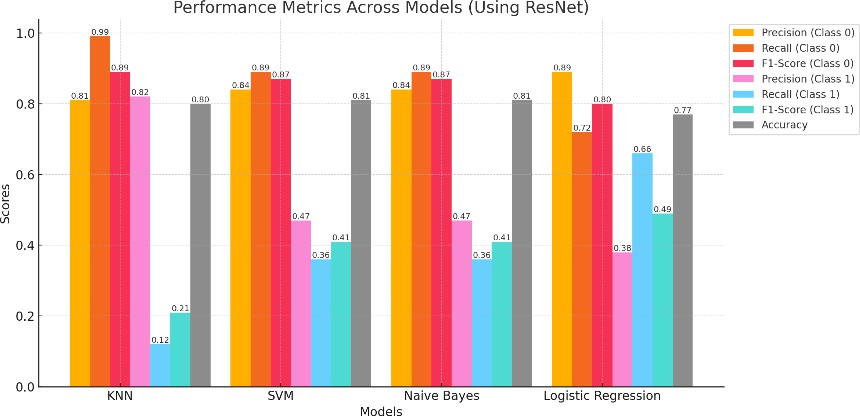


**Fig 9:** Performance metrics across models (using CNN)

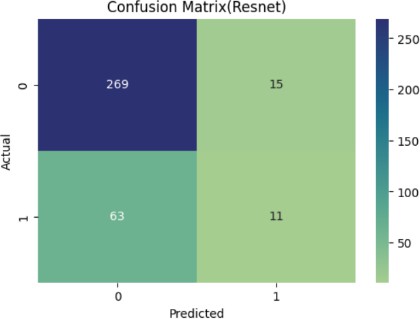


**Fig 10:** Confusion Matrix (CNN)

**ResNet:** The performance metrics indicate that the model achieves a high recall (0.95) and F1-score (0.87) for Class 0, demonstrating strong predictive capability for the majority class. However, performance significantly drops for Class 1, with low recall (0.15) and F1-score (0.22), despite a precision of 0.42. The overall accuracy of 70% reflects the model's bias towards the majority class.

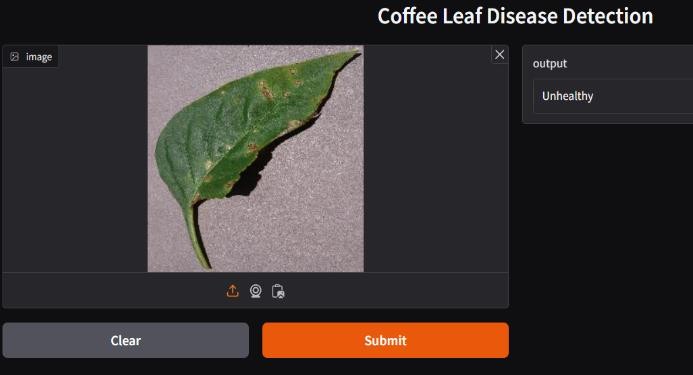
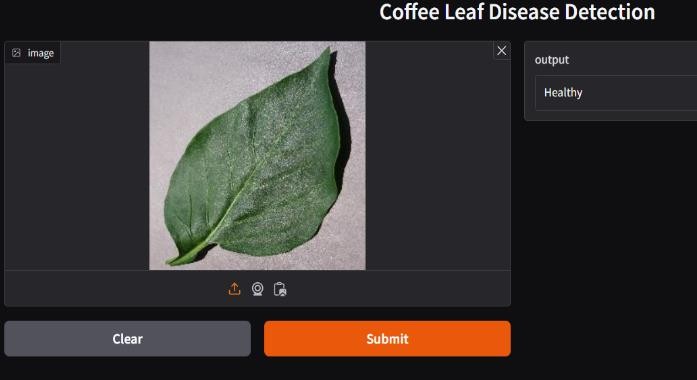


**Fig 11:** Performance metrics across models (using ResNet)



### Executed Output:

**Fig 12:** Confusion Matrix (ResNet)

**Fig 13:** Executed output

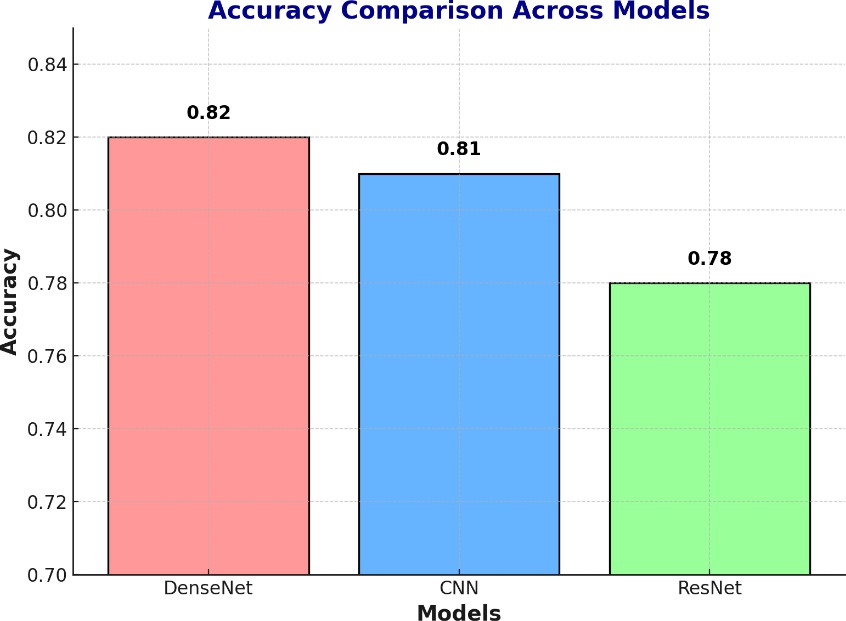
DenseNet, ResNet, and CNN were utilized to classify and detect coffee leaf diseases, including rust, bacterial blight, and leaf spots. Each model underwent fine-tuning to enhance its performance on the coffee leaf dataset. Data augmentation techniques like rotation, flipping, and scaling were applied to increase the dataset's variability, improving the model’s ability to generalize.

**DenseNet:** Achieved high accuracy by maintaining feature reuse through dense connections, which minimized the vanishing gradient issue and allowed deeper feature propagation. DenseNet demonstrated robust performance, particularly in distinguishing between visually similar disease categories.

**ResNet:** Leveraged residual connections to avoid degradation in deep networks, enhancing model stability and performance in complex conditions (e.g., leaves with overlapping symptoms). ResNet achieved commendable classification accuracy, showing particular strengths in identifying subtle differences in disease patterns.

**CNN:** As a foundational architecture, the CNN model performed well in initial classifications but showed limitations compared to DenseNet and ResNet, especially under complex visual conditions. It provided a strong baseline performance and was effective with less computational cost but did not reach the same high accuracy as the other two models.Overall, DenseNet achieved the highest accuracy, followed closely by ResNet, with CNN as a reliable but less precise alternative.

### Comparison:



**Fig 14:** Architecture of modified ResNet50

DenseNet and ResNet outperformed CNN due to their advanced architectures. DenseNet's dense blocks allowed for feature reuse, leading to more refined feature maps, while ResNet’s residual connections mitigated potential gradient vanishing issues in deep networks. These architectures proved more resilient to the variability in the dataset, handling complex visual cues (like color and size variations) more effectively than the CNN model.

# CONCLUSION

The implementation of DenseNet, ResNet and CNN for coffee leaf disease detection demonstrates the effectiveness of deep learning models in accurately diagnosing leaf diseases, which is critical for maintaining coffee crop health. DenseNet and ResNet outperformed the standard CNN model due to their advanced architectures, which are better suited for capturing intricate disease patterns in complex images. DenseNet’s dense connections allowed for efficient feature reuse, enhancing the model's capability to differentiate between diseases with similar visual symptoms. ResNet's use of residual connections mitigated issues with deeper network training, making it highly effective for complex detection scenarios. The results indicate that DenseNet achieved the highest accuracy, followed closely by ResNet, both proving resilient to environmental variations in the dataset, such as differences in leaf size, color, and texture. The CNN model, while robust, served more effectively as a baseline, being less precise under challenging visual conditions. This study underscores the value of leveraging advanced network architectures, particularly DenseNet and ResNet, for real-time, reliable disease detection in agricultural settings. The models' high performance and stability suggest they could play a crucial role in supporting coffee farmers by providing an efficient and automated disease diagnosis tool, ultimately helping to mitigate crop loss and improve yield.

# REFERENCES

1. Tassis, Lucas M., Joao E. Tozzi de Souza, and Renato A. Krohling. "A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images." Computers and Electronics in Agriculture 186 (2021): 106191.
2. Milke, Elisaye Bekele, Menbere Tesfaye Gebiremariam, and Ayodeji Olalekan Salau. "Development of a coffee wilt disease identification model using deep learning." Informatics in Medicine Unlocked 42 (2023): 101344.
3. Krishnamoorthy, N., Prasad, L. N., Kumar, C. P., Subedi, B., Abraha, H. B., & Sathishkumar, V.

E. (2021). Rice leaf diseases prediction using deep neural networks with transfer learning. Environmental Research, 198, 111275.

1. Martinez, Fredy, Holman Montiel, and Fernando Martinez. "A machine learning model for the diagnosis of coffee diseases." International Journal of Advanced Computer Science and Applications 13.4 (2022).
2. Novtahaning, Damar, Hasnain Ali Shah, and Jae-Mo Kang. "Deep learning ensemble-based automated and high-performing recognition of coffee leaf disease." Agriculture 12.11 (2022): 1909.
3. Montalbo, Francis Jesmar Perez, and Alexander Arsenio Hernandez. "Classifying Barako coffee leaf diseases using deep convolutional models." International Journal of Advances in Intelligent Informatics 6.2 (2020): 197-209.
4. Poornam, S., and A. Francis Saviour Devaraj. "Image based Plant leaf disease detection using Deep learning." Inernational journal of computer communication and informatics 3.1 (2021): 53- 65.
5. Esgario, J. G. M., et al. "An app to assist farmers in the identification of diseases and pests of coffee leaves using deep learning. Information Processing in Agriculture, 9 (1), 38–47." 2022.
6. Yebasse, M., Shimelis, B., Warku, H., Ko, J., & Cheoi, K. J. (2021). Coffee disease visualization and classification. Plants, 10(6), 1257.
7. Kumar, Manoj, Pranav Gupta, and Puneet Madhav. "Disease detection in coffee plants using convolutional neural network." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.
8. Signo, Samuel Dave R., Chloe Lei G. Tuquero, and Edwin R. Arboleda. "Coffee disease detection and classification using image processing: A Literature review.
9. Yamashita, João Vitor Yukio Bordin, and João Paulo RR Leite. "Coffee disease classification at the edge using deep learning." Smart Agricultural Technology 4 (2023): 100183.
10. Abuhayi, Biniyam Mulugeta, and Abdela Ahmed Mossa. "Coffee disease classification using Convolutional Neural Network based on feature concatenation." Informatics in Medicine Unlocked 39 (2023): 101245.
11. Aufar, Yazid, and Tesdiq Prigel Kaloka. "Robusta coffee leaf diseases detection based on MobileNetV2 model." International Journal of Electrical and Computer Engineering 12.6 (2022): 6675.
12. Abd Algani, Yousef Methkal, et al. "Leaf disease identification and classification using optimized deep learning." Measurement: Sensors 25 (2023): 100643.