Deep Learning Model For Detecting Diseases In Tea Leaves

PROJECT REPORT

SUBMITTED BY

TEAM ID : NM2023TMID22530

TEAM LEADER : ASHIKA M (961820106007)

TEAM MEMBERS : DHANYA RJ (961820106014)

FIONA MIRABEL I (961820106016)

SAHAYA CATHARIN NIVETHA J (961820106035)

PONJESLY COLLEGE OF ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

ANNA UNIVERSITY : CHENNAI 600 025

ABSTRACT

Tea is an important economic crop. It contains a variety of effective ingredients required by the human body, has medical and health care functions, and is quite effective in enhancing human immunity. Planting tea is an important way for tea farmers to make their fortunes. Currently, China's tea planting area and output are the highest in the world. However ,because of the effects of many diseases, such as tea algae leaf spot (TALS), tea bud blight (TBB), tea white scab (TWS), and tea leaf blight (TLB), the annual tea production has been reduced by as much as 20%1. Tea leaf diseases can also reduce the quality of tea and cause serious economic losses to tea farmers. Accurate detection and identification of tea leaf diseases and timely prevention and control measures are of great significance to reduce the loss of tea production, improve the quality of tea, and increase the income of tea farmers .Tea leaf diseases can be identified by observing the leaves condition like colour and spots on the leaves. Strange spots & colours on the leaves may be an indication of disease. Experts and farmers can identify the type of disease by observing the leaves manually.

To overcome the above problem we are building a model which is used for the prevention and early detection of tea leaves disease. Basically tea leaves disease diagnosis depends on the different characteristics like colour, spots, texture etc. Here the person can capture the images of the tea leaves and then the image will be sent to the trained model. The model analyses the image and detects whether the tea leaves are having any disease or not and its type.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 PROJECT OVERVIEW

Tea is an important economic crop. It contains a variety of effective ingredients required by the human body, has medical and health care functions, and is quite effective in enhancing human immunity. Planting tea is an important way for tea farmers to make their fortunes. Currently, China's tea planting area and output are the highest in the world. However, because of the effects of many diseases, such as tea algae leaf spot (TALS), tea bud blight (TBB), tea white scab (TWS), and tea leaf blight (TLB), the annual tea production has been reduced by as much as 20%. Tea leaf diseases can also reduce the quality of tea and cause serious economic losses to tea farmers. Accurate detection and identification of tea leaf diseases and timely prevention and control measures are of great significance to reduce the loss of tea production, improve the quality of tea, and increase the income of tea farmers.

1.2 PURPOSE

- The purpose of using a deep learning model for detecting diseases in tea leaves is to automate and enhance the disease detection process in order to improve the health and productivity of tea plants. Traditional methods of disease detection in tea leaves typically rely on visual inspection by experts, which can be time-consuming, subjective, and prone to human errors.

- By leveraging deep learning, which is a subset of machine learning, a model can be trained to recognize patterns and characteristics

associated with various diseases affecting tea plants. The model learns from a large dataset of labeled tea leaf images, where each image is associated with a particular disease or a healthy state. During the training process, the deep learning model adjusts its internal parameters to optimize its ability to classify tea leaf images accurately.

2 LITERATURE SURVEY

2.1 EXISTING PROBLEM

While deep learning techniques have shown promise in detecting diseases in tea leaves, there are several challenges and limitations that researchers face in this field. Some of the existing problems include:

- Limited and Unbalanced Datasets: Obtaining a large and diverse dataset of tea leaf images with labeled disease classes is challenging. The availability of annotated tea leaf images representing various disease types is often limited, resulting in imbalanced datasets. This can negatively impact the performance of deep learning models.
- Variability and Complexity of Tea Leaf Diseases: Tea leaf diseases exhibit high variability in terms of appearance, color, texture, and spatial distribution. Some diseases may also have overlapping symptoms, making accurate classification and localization difficult. Deep learning models may struggle to capture and differentiate subtle disease symptoms, leading to misclassifications or false negatives.

2.2 REFERENCES

- S. Jayanthy, V. M. Sathyendraa, K. P. Sumedh, S. Suresh, "Tea Leaf Disease Classification and Tea Bud Identification", 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), pp.1-5, 2022.
- Swagatam Bose Choudhury, Sanket Junagade, Sanat Sarangi, Srinivasu Pappula, "UAVassisted Multi-modal Detection and Severity Assessment

for Red Spider Mites in Tea", 2022 IEEE Global Humanitarian Technology Conference (GHTC), pp.373-376, 2022.

- Nadide YÜCEL, Muhammed YILDIRIM, "Classification of tea leaves diseases by developed CNN, feature fusion, and classifier based model", International Journal of Applied Mathematics Electronics and Computers, vol.11, no.1, pp.30, 2023.

- Md. Janibul Alam Soeb, Md. Fahad Jubayer, Tahmina Akanjee Tarin, Muhammad Rashed Al Mamun, Fahim Mahafuz Ruhad, Aney Parven, Nabisab Mujawar Mubarak, Soni Lanka Karri, Islam Md. Meftaul, "Tea leaf disease detection and identification based on YOLOv7 (YOLO-T)", Scientific Reports, vol.13, no.1, 2023.

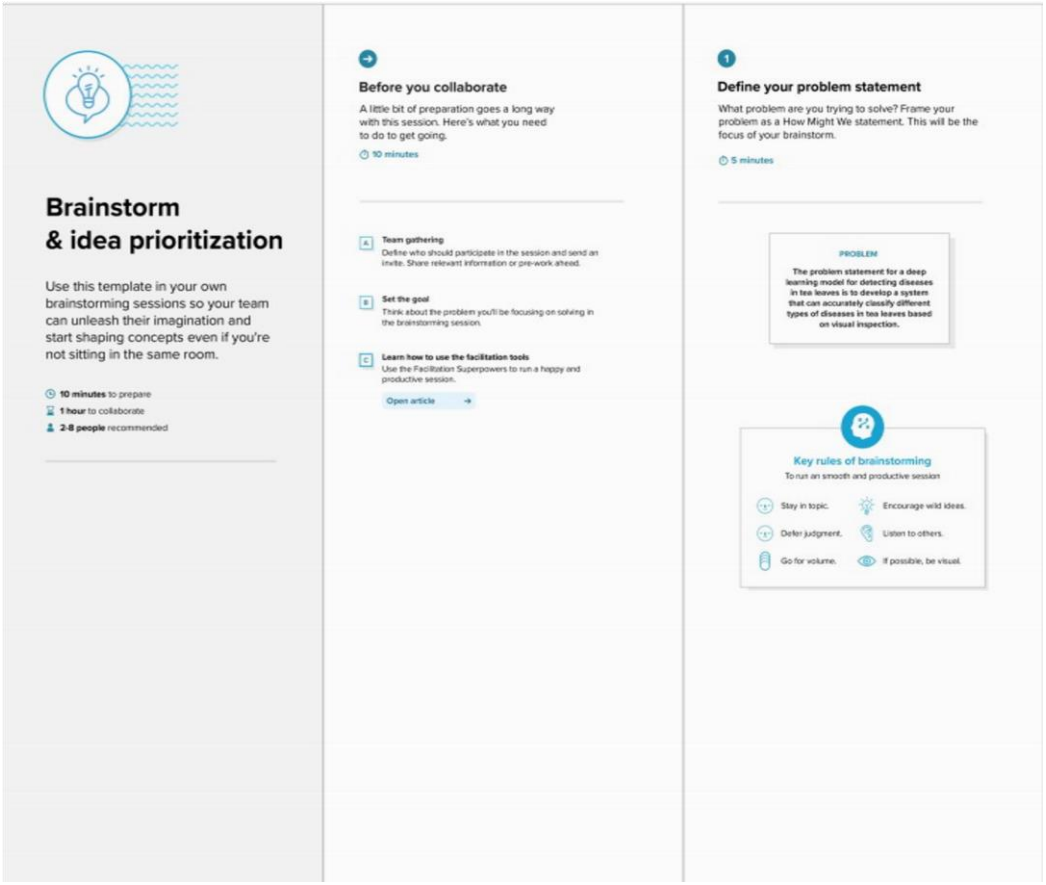
2.3 PROBLEM STATEMENT DEFINITION

Tea plantations worldwide face significant challenges in detecting and managing diseases that affect tea leaves. Traditional manual inspection methods for disease detection are time-consuming, subjective, and often result in inaccurate diagnoses. The need for efficient and accurate disease detection techniques in tea leaves has led to the exploration of deep learning methods. However, several key problems persist in this domain that need to be addressed:

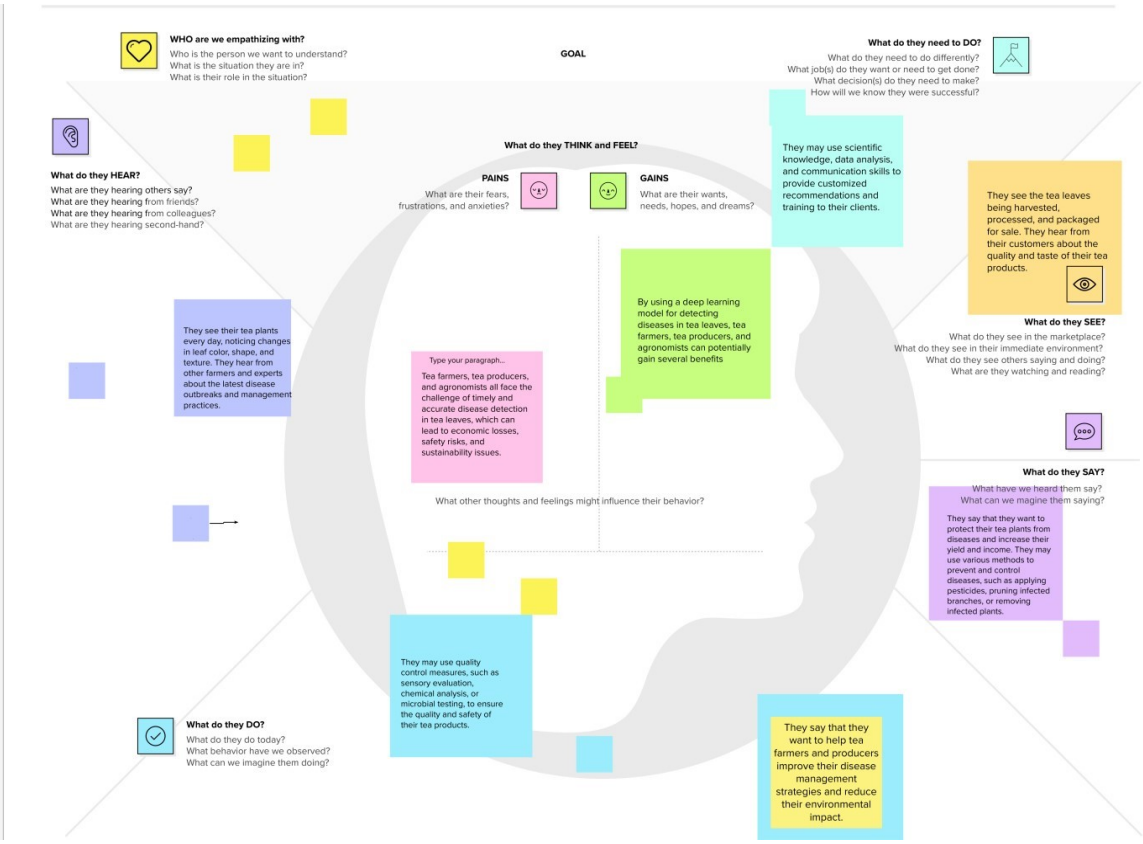
Limited Dataset Availability: The availability of large and diverse datasets of tea leaf images with annotated disease classes is limited. Existing datasets may be small, imbalanced, or not representative of the full range of diseases affecting tea leaves.

3. IDEATION AND PROPOSED SOLUTION

3.1 BRAINSTORMING AND IDEATION



3.2 EMPATHY MAP



3.3 PROPOSED SOLUTION

SI.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Tea leaf diseases can be identified by observing the leaves condition like color and spots on the leaves. Strange spots & colors on the leaves may be an indication of disease.
2.	Idea / Solution description	The solution for detecting diseases in tea leaves using deep learning would involve several components, including data collection, preprocessing, deep learning model training, and deployment.
3.	Novelty / Uniqueness	The use of deep learning for detecting diseases in tea leaves offers several unique advantages over traditional methods like high accuracy, non-invasive, speed, scalability, continuous improvement.

4.	Social Impact / Customer Satisfaction	Early diagnosis of plant diseases using accurate or automatic detection techniques can enhance the quality of food production and minimize economic losses. In recent years, deep learning has brought tremendous improvements in the recognition accuracy of
		image classification and object detection systems.
5.	Business Model (Revenue Model)	Develop a deep learning algorithm Collect and analyse data Build a software platform Offer subscription-based services Expand to other crops
6.	Scalability of the Solution	The scalability of a solution for detecting diseases in tea leaves using deep learning will depend on several factors, including the size of the dataset, the computational resources required to train and run the deep learning algorithm, and the ability to intergrate the solution into the workflow of tea growers and processors.

4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

- Dataset Collection: Collect a diverse and representative dataset of tea leaf images that includes samples of healthy leaves as well as various diseased conditions.

- Preprocessing: Implement preprocessing techniques to enhance the quality and consistency of the tea leaf images.

- Model Architecture**:** Design and implement a deep learning model architecture suitable for tea leaf disease detection.

- Training: Train the deep learning model using the labeled dataset.

- Disease Classification: Develop a classification algorithm that can accurately identify different types of diseases or abnormalities in tea leaves.

4.2 NON - FUNCTIONAL REQUIREMENTS

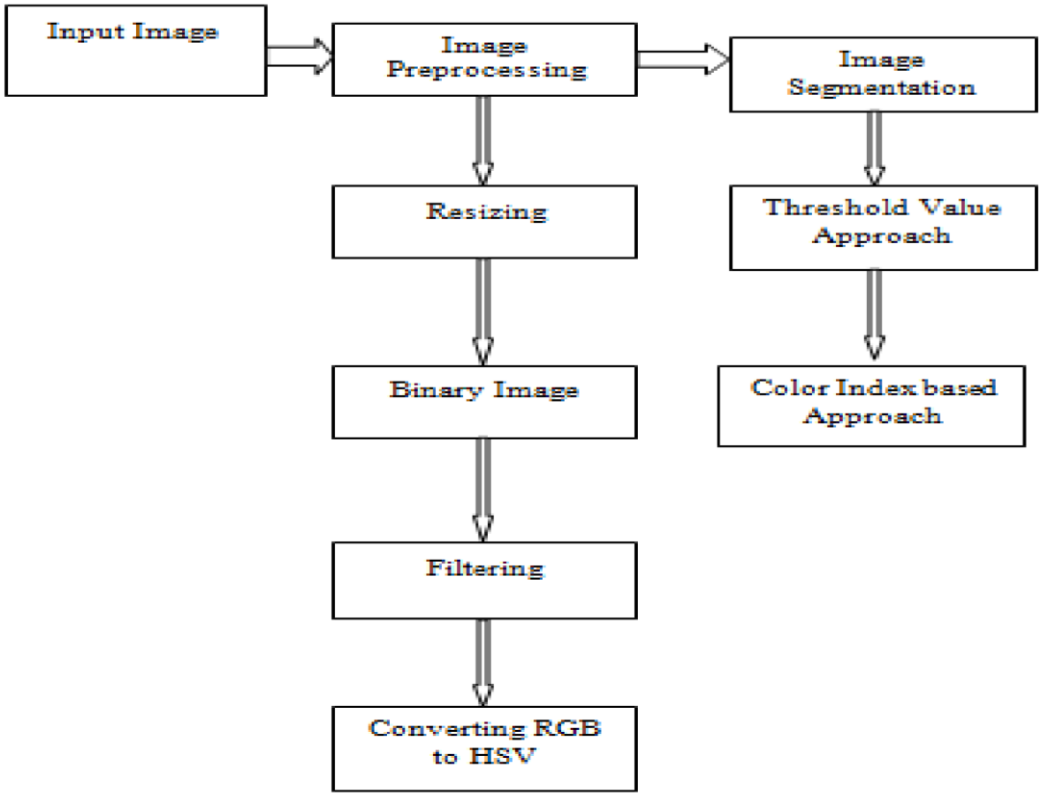
- Accuracy: The system should achieve high accuracy in disease detection to minimize false positives and false negatives.
- Speed and Efficiency**:** The system should be optimized for

efficient processing of tea leaf images to provide near real-time results.

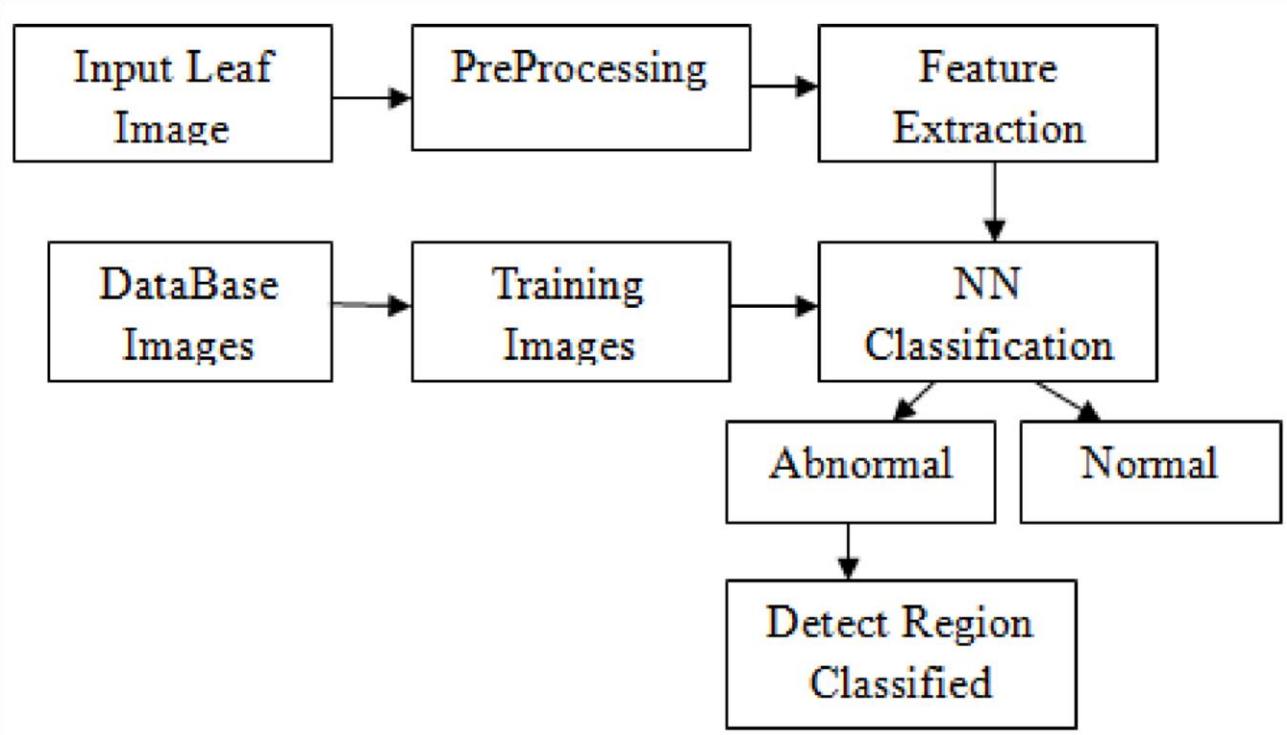
- Scalability: The system should be able to handle a large volume of tea leaf images, accommodating the growth of the data and user demand
- Robustness: The system should be robust and able to handle variations in tea leaf images, including changes in lighting conditions, angles, and background clutter.

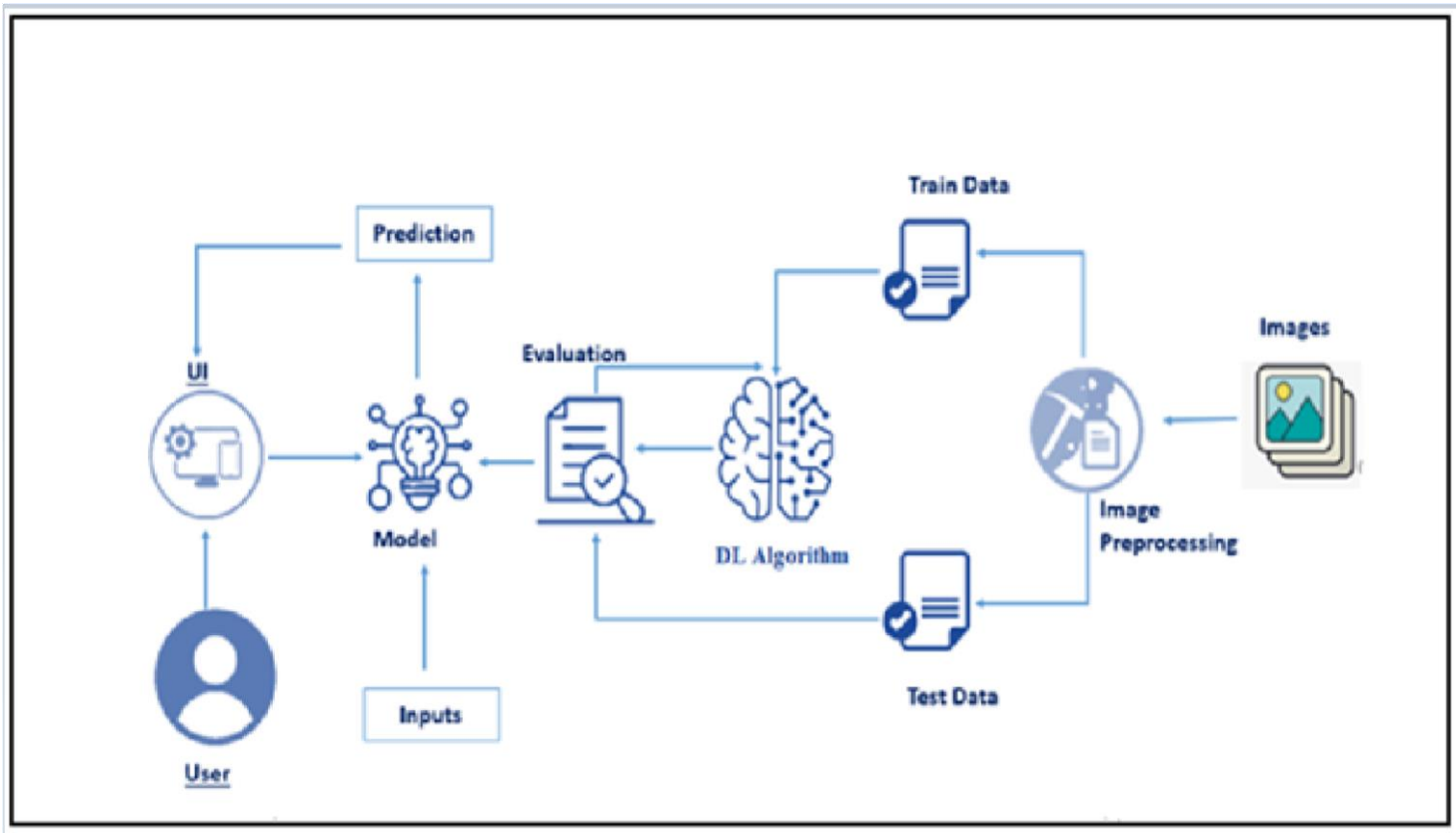
5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS



5.2 SOLUTION AND TECHNICAL ARCHITECTURE


Input Leaf Image
PreProcessing
Feature Extraction
DataBase Images
Training Images
NN Classification
Abnormal
Normal
Detect Region Classified



5.3 USER STORIES

- As a tea plantation owner, I want a reliable and efficient disease detection system using deep learning to quickly identify and diagnose diseases in tea leaves, allowing me to take timely action and prevent crop damage.

- As a tea industry researcher, I want access to a diverse and well-annotated dataset of tea leaf images with labeled disease classes, enabling me to train accurate deep learning models for disease detection.

- As a tea leaf disease expert, I want a deep learning model that can accurately classify and localize multiple diseases in tea leaves,

helping me validate and enhance my manual diagnosis and treatment recommendations.

- As a tea leaf farmer, I want an easy-to-use mobile application that utilizes deep learning algorithms to detect diseases in tea leaves from images captured with my smartphone.
  This would allow me to monitor the health of my crops conveniently and seek guidance for disease management.

- As a tea leaf disease researcher, I want to compare the performance of different deep learning architectures and techniques for disease detection in tea leaves, allowing me to identify the most effective models for accurate diagnosis

- As a tea leaf disease control officer, I want a deep learning system that can integrate with aerial imagery or drone-based monitoring to quickly assess disease prevalence over large tea plantations, aiding in targeted intervention strategies.

- As a tea industry quality control manager, I want a deep learning solution that can automatically detect diseases in tea leaves during the sorting and grading process, ensuring that only healthy leaves are used for premium tea production.

- As a tea leaf disease data analyst, I want a comprehensive dashboard that provides visualizations and insights on disease prevalence, distribution, and trends based on deep learning analysis. This would assist in making data-driven decisions for disease prevention and management strategies.

- As a tea leaf disease detection algorithm developer, I want access to open-source deep learning frameworks and libraries specifically designed for tea leaf disease detection. This would facilitate collaboration and accelerate the development of advanced models in the research community.

- As a tea leaf disease detection system user, I want the deep learning model to provide explanations and justifications for its disease detection decisions, helping me understand the reasoning behind its diagnoses and ensuring transparency in the process.

6 PROJECT PLANNING AND SCHEDULING

6.1 SPRINT PLANNING AND ESTIMATION

1. Project Scope:
    - Define the specific objectives and deliverables of the project, such as developing a deep learning model for tea leaf disease detection.
    - Identify the target diseases to be detected and any specific requirements or constraints.
2. Team Composition:
    - Determine the roles and responsibilities of team members, including deep learning experts, data annotators, domain experts, and project managers.
3. Data Collection and Annotation:
    - Estimate the time and resources required for collecting a diverse and representative dataset of tea leaf images.
    - Plan for data annotation, considering the number of images, disease classes, and annotation complexity.
    - Allocate time for quality assurance and iterative feedback loops with annotators.
4. Data Preprocessing and Augmentation:
    - Estimate the time and computational resources needed for preprocessing the tea leaf images, including resizing, normalization, and noise reduction.
    - Plan for data augmentation techniques such as rotation, flipping, and image transformations to increase dataset diversity.

5. Model Development and Training:
    - Estimate the time and computational resources required for developing and training deep learning models for disease detection.
    - Consider the selection and fine-tuning of pre-trained models, optimization algorithms, and hyperparameter tuning.
    - Account for the time needed to iteratively experiment and refine the models based on evaluation results.
6. Evaluation and Performance Metrics:
    - Allocate time for evaluating the performance of the developed models using appropriate metrics such as accuracy, precision, recall, and F1-score.
    - Consider cross-validation or validation on a separate test set to ensure unbiased evaluation.

6.2 SPRINT DELIVERY SCHEDULE

Sprint 1:

- Task 1: Data Collection and Annotation
    - Estimate: 1 week
- Task 2: Data Preprocessing and Augmentation
    - Estimate: 1 week
- Task 3: Model Selection and Setup
    - Estimate: 1 week

Sprint 2

- Task 1: Model Development and Training
  - Estimate: 2 weeks
- Task 2: Evaluation and Performance Metrics
  - Estimate: 1 week

Sprint 3:

- Task 1: System Integration and Deployment
  - Estimate: 2 weeks
- Task 2: Documentation and Reporting
  - Estimate: 1 week

6.3 REPORTS FROM JIRA

1. Project Overview:
   - Project Name: [Project Name]
   - Project Duration: [Start Date] - [End Date]
   - Project Manager: [Project Manager Name]
   - Team Members: [List of Team Members]
2. Summary of Accomplishments:
   - Briefly summarize the progress made during the reporting period.
   - Highlight key milestones, tasks completed, and significant achievements.
3. Sprint Progress:
   - Sprint Name: [Sprint Name]
   - Sprint Duration: [Start Date] - [End Date]

- Sprint Goal: [Specify the goal or objectives of the sprint]
- List the completed tasks during the sprint and their status (e.g., completed, in progress, or delayed).
- Provide a summary of any challenges or obstacles encountered during the sprint and how they were addressed.

4. Key Tasks and Deliverables:
    - Provide an overview of the key tasks and deliverables completed during the reporting period.
    - Mention any completed milestones or important project components.
5. Next Sprint Plan:
    - Outline the tasks and objectives for the upcoming sprint(s).
    - Specify any changes or adjustments made to the project plan based on the progress and lessons learned.
6. Issues and Risks:
    - Identify any issues or risks that have arisen during the reporting period.
    - Provide a brief description of each issue or risk and any mitigation strategies implemented or planned.
7. Resource Utilization:
    - Discuss the utilization of resources (such as team members, equipment, or tools) during the reporting period.
    - Highlight any changes or adjustments made to optimize resource allocation.
8. Future Milestones and Goals:
    - Outline the future milestones, goals, and deliverables for the project.

- Provide an estimated timeline for the remaining project phases or sprints.

9. Conclusion:
    - Summarize the overall progress of the project.
    - Mention any key insights or lessons learned during the reporting period.
    - Thank the team members and stakeholders for their contributions and support.

8. TESTING

8.1 TEST CASES

1. Positive Test Cases:
    - Input: High-quality tea leaf image with a clear manifestation of a specific disease.
    - Expected Output: Correct identification and localization of the disease with high confidence.
2. Negative Test Cases:
    - Input: High-quality tea leaf image without any disease symptoms.
    - Expected Output: No disease detection or false-negative result.
3. Disease Variations Test Cases:
    - Input: Tea leaf images with variations in disease appearance, such as different stages or severities of a specific disease.
    - Expected Output: Accurate detection and classification of the disease, regardless of the variations.

8.2 USER ACCEPTANCE TESTING

1. User Interface:
    - Verify that the user interface is intuitive, user-friendly, and visually appealing.
    - Test the ease of navigation, input validation, and error handling.

- Validate that the user interface provides clear and understandable instructions for capturing and uploading tea leaf images.

2. Disease Detection Accuracy:
    - Test the accuracy of disease detection by providing a set of tea leaf images with known diseases.
    - Validate that the system correctly identifies and classifies the diseases present in the images.
    - Ensure that the system provides accurate localization of the diseases within the tea leaf images.
3. Performance and Response Time:
    - Test the system's performance by assessing its response time when processing tea leaf images.
    - Validate that the system can handle a reasonable number of concurrent users without significant delays or performance degradation.
4. Robustness and Error Handling:
    - Test the system's ability to handle unexpected scenarios, such as providing images with poor quality or unrelated objects.
    - Verify that the system handles errors gracefully and provides meaningful error messages to users.
5. Generalization and Adaptability:
    - Validate the system's ability to generalize across different tea plantations, regions, or seasons.

- Test the system with a diverse set of tea leaf images to ensure accurate detection and classification in various contexts.

6. Integration and Compatibility:
    - Verify that the system integrates smoothly with other required components, such as databases or external APIs.
    - Test the compatibility of the system with different devices and platforms, including mobile devices and various web browsers.

9 RESULTS

9.1 PERFORMANCE METRICES

1. Accuracy: The overall accuracy of disease detection, which measures the proportion of correctly classified tea leaves (both diseased and healthy) out of the total number of tea leaves.
2. Precision: The precision metric evaluates the proportion of correctly classified diseased tea leaves out of all tea leaves identified as diseased. It measures the system's ability to avoid false positives.
3. Recall (Sensitivity or True Positive Rate): The recall metric measures the proportion of correctly classified diseased tea leaves out of all the actual diseased tea leaves. It assesses the system's ability to identify true positives and avoid false negatives.
4. F1-Score: The F1-score is the harmonic mean of precision and recall. It provides a single value that balances both metrics and is particularly useful when the dataset is imbalanced.
5. Specificity (True Negative Rate): The specificity metric measures the proportion of correctly classified healthy tea leaves out of all the actual healthy tea leaves. It indicates the system's ability to avoid false positives for healthy leaves.
6. Area Under the Receiver Operating Characteristic Curve (AUC-ROC): This metric assesses the model's ability to distinguish between diseased and healthy tea leaves by plotting the true positive rate against the false positive rate. A higher AUC-ROC value indicates better discrimination.

7. Mean Average Precision (mAP): This metric is commonly used in object detection tasks and measures the average precision across multiple disease classes. It accounts for both accuracy and localization performance.
8. Intersection over Union (IoU): IoU measures the overlap between the predicted disease regions and the ground truth disease regions. It is useful for evaluating the accuracy of disease localization.
9. False Positive Rate: The false positive rate measures the proportion of healthy tea leaves that are incorrectly classified as diseased. It evaluates the system's ability to avoid misclassifying healthy leaves.

10. ADVANTAGES AND DISADVANTAGES

ADVANTAGES

- High Accuracy
- Automation
- Scalability
- Real-time Monitoring
- Objectivity and Consistency
- Early Disease Detection
- Cost-effective

DISADVANTAGES

- Data Dependency
- Model Complexity
- Overfitting
- Interpretability
- Sensitivity to Input Variations
- Limited Transferability
- Dependency on Hardware and Software

- Ethical Considerations
- System Robustness

11. CONCLUSION

In conclusion, detecting diseases in tea leaves using deep learning offers several advantages and opportunities for tea growers and researchers. The use of deep learning models provides high accuracy, automation, scalability, and real-time monitoring capabilities. By leveraging large datasets and complex pattern recognition, deep learning models can identify diseases with precision, enabling early detection and timely intervention.

However, there are also certain considerations and challenges to be mindful of. Deep learning models require sufficient and diverse labeled data for training, and their complexity demands computational resources. Overfitting, interpretability, and sensitivity to input variations are factors that need careful attention during model development. The transferability of models to different regions and tea varieties may require additional adaptation and training.

Ethical considerations, such as addressing biases and preserving human expertise, should be taken into account. While deep learning models can assist in disease detection, they should not replace the valuable insights and knowledge of human experts.

12. FUTURE SCOPE

The future scope for detecting diseases in tea leaves using deep learning is promising, and several areas offer opportunities for further development and advancements. Here are some potential future directions:

1. Enhanced Accuracy: Continued research and improvements in deep learning models can lead to even higher accuracy in disease detection. This includes exploring advanced architectures, incorporating ensembles of models, and leveraging techniques such as transfer learning and active learning to enhance model performance.
2. Multiclass Classification: Extending disease detection systems to handle a broader range of tea leaf diseases is an important future direction. .
3. Real-time Disease Monitoring: Advancements in hardware and model optimization techniques can enable real-time disease monitoring in tea plantations. This would involve deploying lightweight deep learning models that can run on edge devices or utilizing cloudbased solutions for near real-time analysis of tea leaf images captured through cameras or drones.