A Project Report on

**Cross-Modal Therapy Companion**

**Submitted by:**

**Sidra Aiman - 100522733042**

**Einstein Ellandala - 100522733089**

**Rohith Kodam - 100522733093**

**Under the supervision of**

**PROF. P. V. SUDHA**

**Assistant Professor, Department of CSE,**

**UCE(A), OU, Hyderabad-500007**

**A blue and white logo

AI-generated content may be incorrect.**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,**

**UNIVERSITY COLLEGE OF ENGINEERING(A),**

**OSMANIA UNIVERSITY, HYDERABAD-500007**

**JUNE-2025**

2022 – 2026

UNIVERSITY COLLEGE OF ENGINEERING (AUTONOMOUS)

OSMANIA UNIVERSITY, HYDERABAD – 500007

**A blue and white logo

AI-generated content may be incorrect.**

**CERTIFICATE**

This is to certify that the Mini Project work titled "**Cross-Modal Therapy Companion**" submitted by **Sidra Aiman, Einstein Ellandala, Rohith Kodam** students of the Department of Computer Science and Engineering, University College of Engineering, Osmania University, is a record of the bonafide work carried out by them during the academic year 2025-26. The project is submitted in partial fulfilment of the requirements for the award of the degree of Bachelor of Engineering in Computer Science and Engineering. The work has been carried out under the supervision of **PROF. P. V. SUDHA**, and it is to the best of my knowledge that the work is original and reflects the sincere efforts of the students.

Signature of the Supervision Signature of the Head of the Dept.

**PROF. P. V. SUDHA PROF. P. V. SUDHA**

Department of CSE, Department of CSE,

University College of Engineering, University College of Engineering,

Osmania University Osmania University

**DECLARATION**

We **Sidra Aiman, Einstein Ellandala, Rohith Kodam** students of the Department of Computer Science and Engineering, University College of Engineering, Osmania University, hereby declare that the work presented in this Mini Project titled "**Cross-Modal Therapy Companion**" is an original contribution carried out by us during the academic year 2025-26. This project report is submitted in partial fulfilment of the requirements for the degree of Bachelor of Engineering in Computer Science and Engineering. The project work has not been submitted elsewhere for the award of any degree or diploma.

We affirm that no part of this report is plagiarized, and wherever references have been made, they have been appropriately cited. The findings and analysis presented in the report are based on our genuine and authentic work under the guidance of **PROF. P. V. SUDHA**.

We further declare that we have adhered to ethical practices throughout the research and project development process, maintaining academic integrity at every stage. The data, analysis, and outcomes of this report are factual to the best of our knowledge, and we take full responsibility for the contents of this submission.

**Sidra Aiman - 100522733042**

**Einstein Ellandala - 100522733089**

**Rohith Kodam - 100522733093**

**ACKNOWLEDGEMENT**

We would like to express our deep sense of gratitude and whole-hearted thanks our project guide **PROF. P. V. SUDHA**, Professor at Department of Computer Science and Engineering, University College of Engineering Osmania University, for giving us the privilege of working under his guidance, with tremendous support and cogent discussion, constructive criticism, and encouragement throughout this dissertation work carrying out the project work. We also thank him for his support from the department and for making all the resources available to us students. We also extend our thanks to the entire faculty of the Department of Computer Science and Engineering, University College of Engineering, Osmania University who encouraged us throughout the course of our Bachelor's degree and allowed us to use the many resources present in the department. Our sincere thanks to our parents and peers for their valuable suggestions, morals, strength, and support for the completion of our project.

**ABSTRACT**

The increasing prevalence of mental health awareness has catalyzed the development of digital tools aimed at promoting well-being. However, a significant gap exists in the market, with most applications relying on subjective, single-modality inputs like manual mood logging. These methods often lack the objectivity and holistic perspective necessary for profound self-reflection. This project, titled "Mindful Companion," addresses this gap by presenting a novel, cross-modal web application designed as a private, data-driven dashboard for personal emotional analysis. The system provides users with a comprehensive and insightful tool to track, visualize, and understand the intricate connections between their emotional state, biometric expressions, and lifestyle habits.

The core innovation of the "Mindful Companion" lies in its multi-modal data integration framework. The application captures data from three distinct sources to build a rich, contextualized user profile. Firstly, it employs a **Computer Vision** module for real-time **Facial Emotion Recognition**, utilizing the robust DeepFace library to analyze images captured from a user's webcam. This provides an objective assessment of expressed emotion. Secondly, an **Audio Processing** module analyzes the user's vocal tone through a state-of-the-art **Wav2Vec2 Transformer model**, a deep learning architecture that can discern emotional cues from the prosodic features of speech. Thirdly, the system incorporates **self-reported data**, allowing users to log crucial lifestyle factors such as hours of sleep and daily physical activity level.

The system's backend is developed using **Python** with the **Flask** micro-framework, creating a RESTful API to handle all data processing and analysis tasks. The frontend is a responsive, single-page application built with standard **HTML, CSS, and JavaScript**, ensuring broad accessibility across all modern devices and browsers. The communication between client and server is handled asynchronously via the fetch API, with data exchanged in JSON format. For data persistence in this prototype, the system utilizes the Pandas library to manage a local CSV file, which serves as a secure, on-device logbook.

A key deliverable of this project is a custom **heuristic assessment algorithm** that synthesizes the multi-modal inputs into a single, easy-to-understand "Potential Stress Score." Based on this score, the application generates a detailed feedback report and provides personalized, actionable recommendations to the user, encouraging mindful adjustments to their daily routines. The user's historical data is rendered on an interactive dashboard powered by **Chart.js**, featuring dynamic line charts to track stress and sleep trends over time, and doughnut charts to visualize the spectrum of detected facial and vocal emotions.

Testing was conducted at multiple levels, including unit testing for the heuristic algorithm, functional testing to validate user workflows, and integration testing to ensure seamless communication between the frontend and backend modules. The final result is a stable, fully functional prototype that successfully demonstrates the feasibility and utility of applying a cross-modal AI approach to personal well-being.

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S. No** | **Section Title** | **Page No.** |
| 1 | Introduction | 2 |
| 2 | Literature Survey | 2 |
| 2.1 | VGG19-Based Paddy Leaf Detection | 2 |
| 2.2 | CNN + Random Forest Hybrid Approach | 2 |
| 2.3 | Comparison with Existing Models | 3 |
| 3 | System Design | 4 |
| 3.1 | System Architecture | 4 |
| 3.2 | System Modules | 5 |
| 4 | Implementation | 5 |
| 4.1 | Tools and Technologies | 5 |
| 4.2 | Dataset | 5 |
| 4.3 | Model Training | 6 |
| 4.4 | Leaf Validation | 7 |
| 4.5 | Disease Area Detection | 7 |
| 4.6 | Severity Estimation22 | 7 |
| 4.7 | Web Application | 7 |
| 4.8 | Sample Workflow | 7 |
| 4.9 | Sample Outputs | 8 |
| 5 | Results and Discussions | 8 |
| 5.1 | Model Evaluation Metrics | 8 |
| 5.2 | Accepted Paddy Leaf Image (Input) | 10 |
| 5.3 | Disease Prediction Output | 10 |
| 5.4 | Summary Table | 11 |
| 6 | Conclusion | 11 |

# 1.Introduction

In the 21st century, the discourse surrounding mental health has shifted from a reactive to a proactive stance. There is a growing global recognition that maintaining mental and emotional well-being is as crucial as physical health. This paradigm shift has been accompanied by a surge in digital tools designed to support this journey. However, while technology has offered new avenues for connection and information, it has also been identified as a significant source of stress and digital fatigue. The challenge, therefore, lies in leveraging technology not as a distraction, but as a focused, beneficial tool for introspection and self-awareness. This project, the "Mindful Companion," is born from this challenge, aiming to create a deeply personal and insightful digital space for users to reflect on their emotional state.

The "Mindful Companion" is a sophisticated, web-based application that serves as a private well-being dashboard. It moves beyond the limitations of traditional mood trackers by employing a **cross-modal analysis** framework. This approach is founded on the understanding that human emotion is a complex tapestry woven from multiple threads of expression. A single data point, such as a manually logged mood, provides an incomplete and often biased picture. Our system, by contrast, integrates three distinct data modalities to create a richer, more objective snapshot of a user's well-being. It analyzes visual cues through **facial emotion recognition**, auditory signals through **vocal tone analysis**, and contextualizes this biometric data with **user-reported lifestyle factors**, namely hours of sleep and daily activity level.

The primary goal of this project is to empower individuals with data-driven insights into their own lives. By consistently capturing and visualizing this multi-faceted data, the application helps users identify subtle yet powerful correlations: "Does a lack of sleep consistently lead to a more negative vocal tone?" or "How does a high-activity day impact my facial expression?" The answers to such questions can foster a profound sense of self-awareness and encourage positive behavioral changes. The system culminates its analysis in a "Potential Stress Score," a heuristic metric designed to provide a simple, at-a-glance summary of a user's current state, supplemented with personalized feedback and actionable recommendations.

It is imperative to state that the "Mindful Companion" is designed and intended as a **tool for self-reflection and informational purposes only**. It is not a medical device, nor is it a substitute for professional psychological advice, diagnosis, or treatment. The algorithms and AI models, while powerful, provide estimations and should be used to encourage curiosity and mindfulness about one's well-being, not to draw definitive clinical conclusions. The application's core philosophy is to provide a safe, private space for users to engage with their emotional data, fostering a proactive and informed approach to personal mental health management.

This report will detail the journey of bringing the "Mindful Companion" from concept to a fully functional prototype. We will explore the existing landscape of well-being applications, outline the system's technical architecture, and delve into the specific algorithms and machine learning models that power its analytical capabilities. Furthermore, we will discuss the design principles, implementation details, testing methodologies, and the ethical considerations inherent in developing such a personal tool.

#### 1.1 PROBLEM STATEMENT

Post-Traumatic Stress Disorder (PTSD) and chronic anxiety affect millions worldwide, characterized by debilitating episodes or attacks that can feel unpredictable and overwhelming. A significant challenge in managing these conditions is the difficulty in recognizing the subtle physiological and emotional shifts that precede a crisis. This "pre-episode" state often includes changes in vocal tone, facial affect, and sleep patterns, but these signs can be missed by the individual experiencing them.

Existing digital mental health tools often lack the capability for passive, objective monitoring. They rely on manual self-reporting, which is unreliable during periods of high distress, or focus on a single data modality, which provides an incomplete picture. This creates a critical gap: there is no integrated system that can continuously and non-invasively monitor for a confluence of early warning signs and provide immediate, actionable support.

This project aims to address this problem by developing a **Cross-Modal Therapy Companion**, a system designed to detect the early signs of PTSD episodes or anxiety attacks by analyzing vocal diaries, facial expressions, and sensor-derived lifestyle data, and to suggest timely, calming interventions.

#### 1.2 SCOPE OF THE PROJECT

The scope of this project is to build a functional prototype that demonstrates the feasibility of the proposed system. This includes:

* Developing a web interface for capturing a user's voice diary (audio), facial emotion (image), and manual logs for sleep/activity.
* Integrating pre-trained AI models to analyze vocal and facial data for emotional and stress-related biomarkers.
* Designing a heuristic algorithm to synthesize the data from all modalities and calculate a "distress score."
* Implementing a mechanism to suggest calming interventions when the distress score exceeds a set threshold.
* Creating a dashboard to visualize trends in the collected data over time.

**Out of Scope:** This project is a proof-of-concept and does **not** include:

* Clinical trials or validation with real patients.
* Real-time processing from wearable sensor data (this is simulated via manual logs).
* A production-ready, HIPAA-compliant secure architecture.

#### 1.3 OBJECTIVES

1. To design and implement a system capable of capturing and analyzing multi-modal data (voice, face, lifestyle).
2. To utilize machine learning models to extract emotional and stress-related features from audio and image inputs.
3. To develop an algorithm that identifies potential pre-episode states by detecting anomalous patterns across the data streams.
4. To create a user interface that provides both a long-term dashboard and immediate, context-aware intervention suggestions.
5. To demonstrate the potential of cross-modal AI as a supportive tool in managing PTSD and anxiety.

### 2. SYSTEM REQUIREMENTS AND ANALYSIS

#### 2.1 PROPOSED MODULES

1. **Data Capture Module:** A user-facing interface for recording audio diaries, capturing facial snapshots, and logging sleep/activity data.
2. **Vocal Biomarker Analysis Module:** A backend service that processes audio to detect stress indicators like vocal jitter, shimmer, pitch variability, and emotional tone using a Wav2Vec2 model.
3. **Facial Affect Analysis Module:** A service utilizing DeepFace to classify the user's dominant facial emotion (e.g., fear, anger, sad, neutral).
4. Behavioral Pattern Module: Analyzes logged sleep and activity data to detect significant deviations from the user's baseline (e.g., sudden insomnia, sharp drop in activity).
5. **Risk Assessment & Intervention Module:** The core logic engine. It uses a heuristic algorithm to fuse the outputs from all other modules into a single risk score. If the score surpasses a threshold, it triggers the intervention suggestion component.
6. **Dashboard & Visualization Module:** A secure interface (conceptually for both user and therapist) to view historical data, track trends, and identify potential trigger patterns.

#### 2.2 SOFTWARE & HARDWARE REQUIREMENTS

* **Hardware:** Standard computer with Webcam/Microphone, 8GB+ RAM recommended.
* **Software:** Python 3.9+, Flask, Pandas, PyTorch, Transformers, Librosa, DeepFace, OpenCV, FFmpeg, modern web browser.

#### 2.3 FEASIBILITY ANALYSIS

* Technical Feasibility: High. The core technologies (AI models for voice/face analysis) are mature and accessible. The primary challenge is the intelligent fusion of the data, which is addressed by the heuristic algorithm in this prototype.
* **Operational Feasibility:** High. The system is designed for ease of use, requiring the user to perform a simple daily check-in. The dashboard and interventions are straightforward.
* **Economic Feasibility:** High. The prototype relies exclusively on open-source software and pre-trained models, making development costs minimal.

**2. Literature Survey**

Several researchers have proposed methods for paddy leaf disease detection using traditional Convolutional Neural Networks (CNNs) and hybrid machine learning techniques.

**2.1 VGG19-Based Paddy Leaf Detection**

* **Source:** VGG19 Enhanced Convolutional Neural Network for Paddy Leaf Disease Detection
* **Summary:** Utilized VGG19 for paddy leaf disease classification with three disease classes and one healthy class. Achieved an accuracy of 91.19%.
* **Limitation:** High computational cost, no severity estimation, no real-time validation for input images.

**2.2 CNN + Random Forest Hybrid Approach**

* **Source:** A Comprehensive Study on Paddy Leaf Disease Detection using CNN and Random Forest
* **Summary:** Combined CNN feature extraction with Random Forest classification. Achieved 92.19% accuracy across five paddy diseases.
* **Limitation:** No severity estimation, higher computational complexity due to hybrid model, no leaf validation module.

**2.3 Comparison with Existing Models**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **VGG19 Paper** | **CNN + Random Forest Paper** | **Proposed System** |
| Model | VGG19 | CNN + Random Forest | EfficientNetB0 |
| Accuracy | 91.19% | 92.19% | 95% |
| Number of Classes | 4 | 5 | 10 |
| Severity Estimation | No | No | Yes |
| Leaf Validation | No | No | Yes |
| Real-Time Capability | No | - | Yes |
| Lightweight Model | No | - | Yes |

The proposed system outperforms existing works by covering more diseases, providing severity estimation, real-time web interface, and leaf validation.

**3. System Design**

**3.1 System Architecture**

User Uploads image

Not a paddy leaf

Paddy Leaf Validation

No

Yes

Disease Classification

Healthy Leaf

Yes

No

Disease Area Detection

Severity Estimation

Result Displayed

**3.2 System Modules**

* **Image Upload Module:** Validates and stores uploaded images.
* **Paddy Leaf Validation Module:** Uses green pixel detection, contour analysis, and aspect ratio checks to verify whether the uploaded image contains a paddy leaf.
* **Disease Classification Module:** Uses EfficientNetB0 to predict one of ten disease classes.
* **Disease Area Detection Module:** Uses image masking to isolate diseased regions and removes irrelevant backgrounds (sky, dark spots).
* **Severity Estimation Module:** Calculates the disease-affected area and classifies severity as Low, Medium, or High.
* **Results Display Module:** Displays the disease name, severity, percentage affected, original image, and disease-highlighted image.

**4. Implementation**

**4.1 Tools and Technologies**

| **Tool** | **Purpose** |
| --- | --- |
| Python | Core programming |
| Flask | Web framework |
| TensorFlow | Model training and prediction |
| OpenCV | Image processing and masking |
| NumPy | Numerical operations |
| HTML/CSS | Web interface |

**4.2 Dataset**

* The dataset consists of **paddy leaf images classified into 10 categories**: Bacterial Leaf Blight, Bacterial Leaf Streak, Bacterial Panicle Blight, Blast, Brown Spot, Dead Heart, Downy Mildew, Hispa, Normal, Tungro.
* <https://www.kaggle.com/datasets/dasa7753912/new-paddy-doctor-paddy-disease-classification?select=paddy-disease-classification>

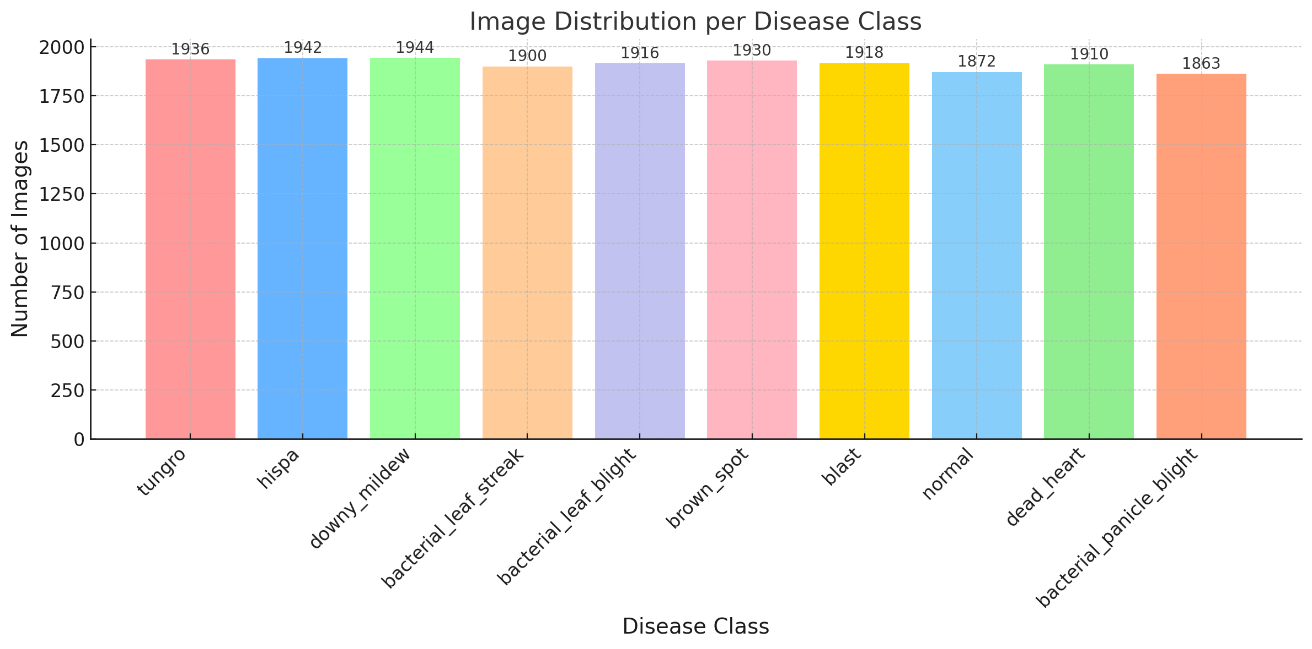


Figure 1: Training dataset image distribution

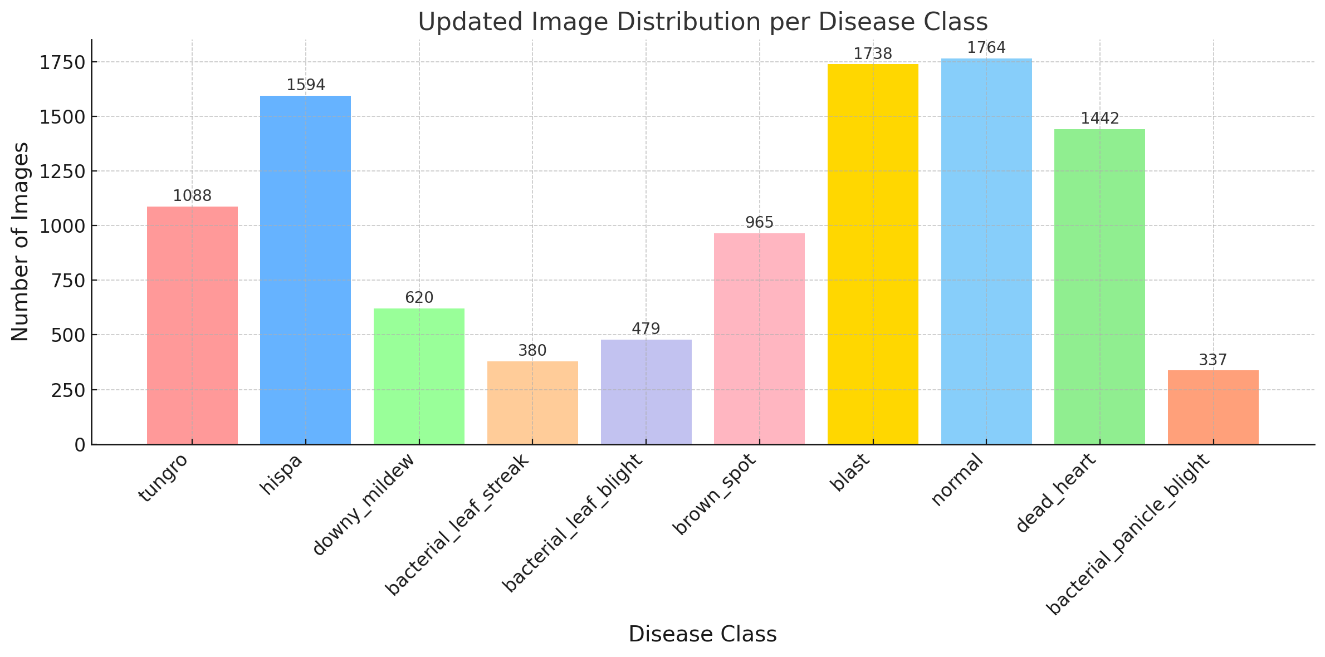
****

Figure 2: Testing dataset image distribution

**4.3 Model Training**

* Model: **EfficientNetB0**
* Dataset Preprocessing: Image resizing to 224x224, data augmentation, class balancing using class weights.
* Initial Training: Top layers trained with frozen base model.
* Fine-tuning: Entire model unfrozen and trained with a lower learning rate.
* Performance Evaluated with Accuracy, Confusion matrix and Classification Report.

**4.4 Leaf Validation**

* **Green Pixel Check:** Ensures minimum 5% green area.
* **Contour Detection:** Confirms presence of leaf-like structures.
* **Aspect Ratio Check:** Ensures paddy leaf’s typical elongated shape.

**4.5 Disease Area Detection**

* Green leaf area is isolated using HSV colour masking.
* Sky blue and dark regions are excluded.
* Diseased areas are detected based on colour ranges commonly associated with disease symptoms.

**4.6 Severity Estimation**

* Disease area vs. total leaf area is calculated pixel-wise.
* Severity is classified as:
  + **Low:** <30% affected
  + **Medium:** 30%-60% affected
  + **High:** >60% affected

**4.7 Web Application**

* Built with Flask.
* Supports real-time image upload and prediction.
* Provides instant visual feedback.
* Cleans up old uploaded files automatically to manage storage.

**4.8 Sample Workflow**

1. User uploads an image of a paddy leaf.
2. System validates whether it is a paddy leaf.
3. Model predicts the disease.
4. If diseased, severity is estimated and affected areas are visualized.
5. Results are displayed on the web page.

**4.9 Sample Outputs**

* Disease Name: Example: "Tungro"
* Severity: High
* Visual Outputs:
  + Original Image
  + Disease Affected Region

**5. Results and Discussions**

**5.1 Model Evaluation Metrics**

**5.1.1 Confusion Matrix**

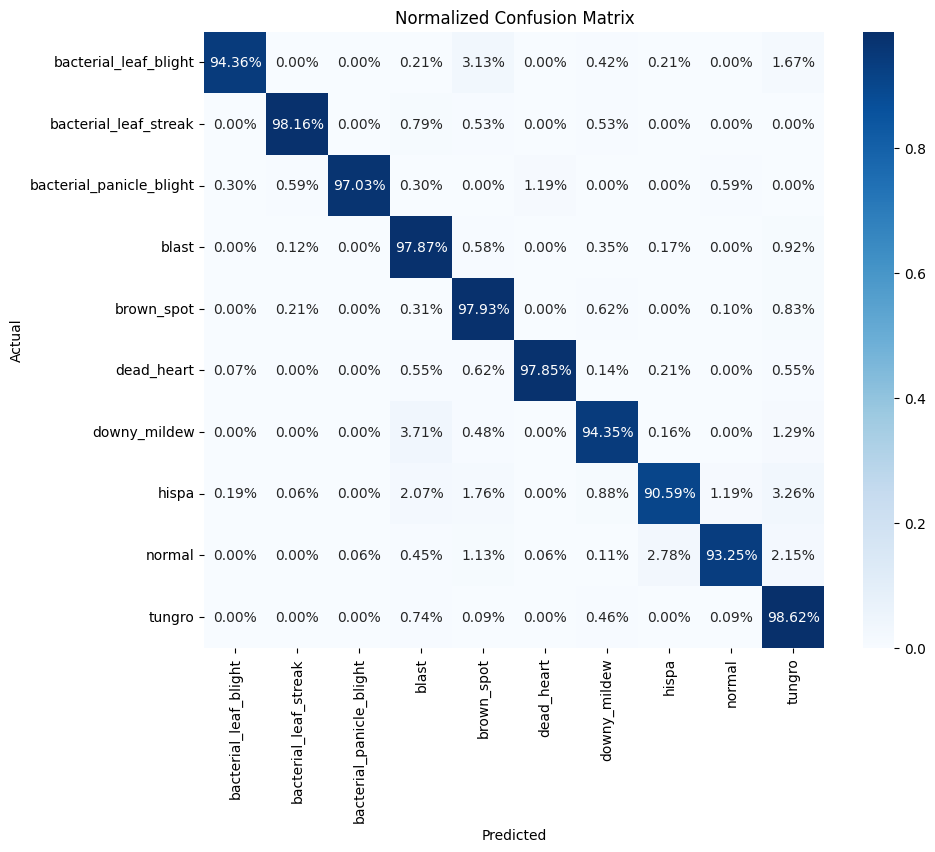
****

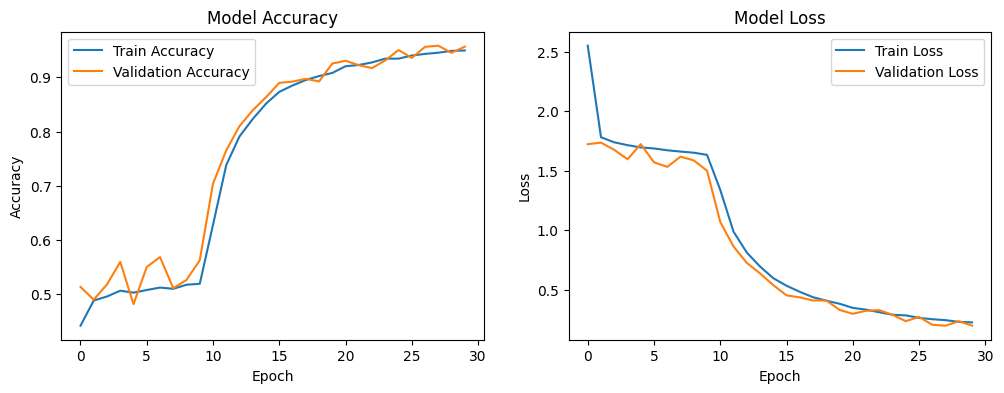
Figure 3: Confusion Matrix of the Paddy Leaf Disease Classification Model.

**5.1.2 Classification Report**

| **Class Name** | **Precision** | **Recall** | **F1-Score** | **Support** |
| --- | --- | --- | --- | --- |
| Bacterial Leaf Blight | 0.99 | 0.94 | 0.97 | 479 |
| Bacterial Leaf Streak | 0.98 | 0.98 | 0.98 | 380 |
| Bacterial Panicle Blight | 1.00 | 0.97 | 0.98 | 337 |
| Blast | 0.95 | 0.98 | 0.96 | 1738 |
| Brown Spot | 0.91 | 0.98 | 0.95 | 965 |
| Dead Heart | 1.00 | 0.98 | 0.99 | 1442 |
| Downy Mildew | 0.94 | 0.94 | 0.94 | 620 |
| Hispa | 0.96 | 0.91 | 0.93 | 1594 |
| Normal | 0.99 | 0.93 | 0.96 | 1764 |
| Tungro | 0.89 | 0.99 | 0.93 | 1088 |
|  |  |  |  |  |
| **Metric** | **Precision** | **Recall** | **F1-Score** | **Support** |
| Accuracy |  |  | 0.96 | 10407 |
| Macro Avg | 0.96 | 0.96 | 0.96 | 10407 |
| Weighted Avg | 0.96 | 0.96 | 0.96 | 10407 |

Table 5.1: Precision, Recall, and F1-Score for each Paddy Leaf Disease Class.

**5.1.3 Accuracy and Loss Curves**

Figure 4: Model Accuracy Curve and Model Loss Curve.

**5.2 Accepted Paddy Leaf Image (Input)**

* Image uploaded by the user.

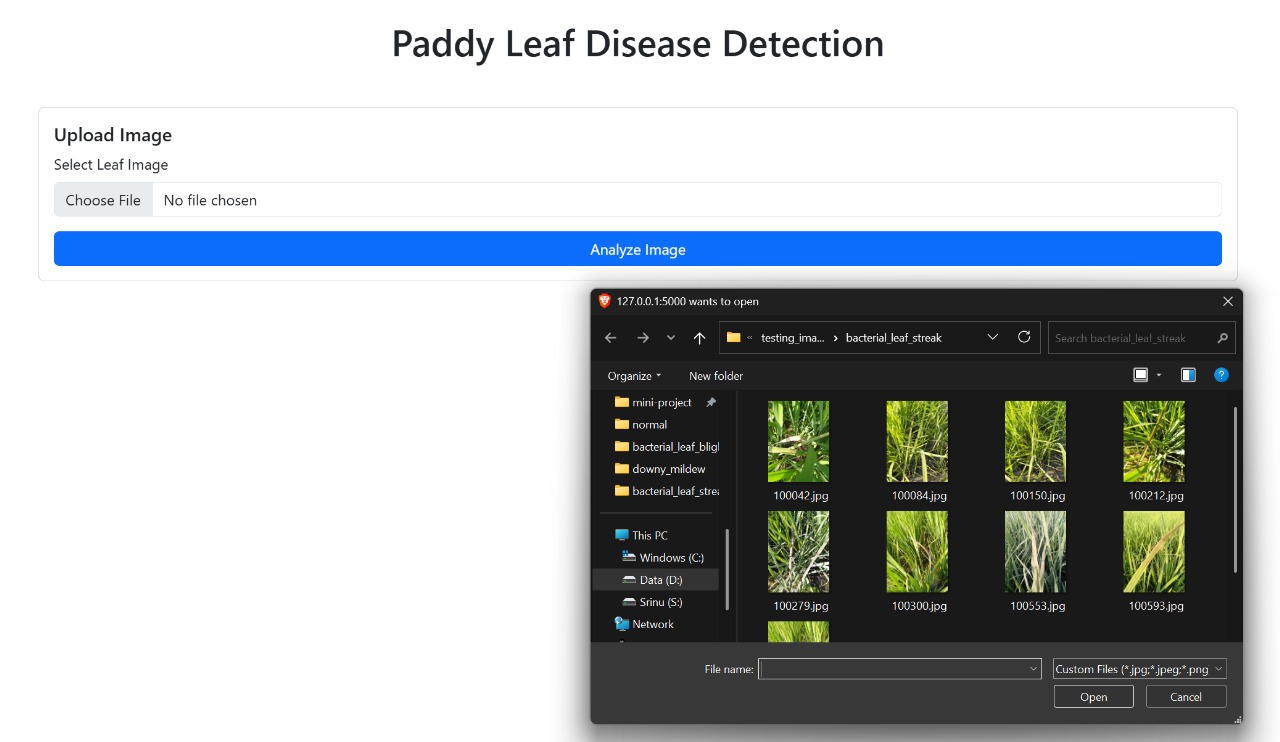
Example:  


Figure 5: Uploaded Paddy Leaf Image

**5.3 Disease Prediction Output**

* Display the predicted disease result from the system.

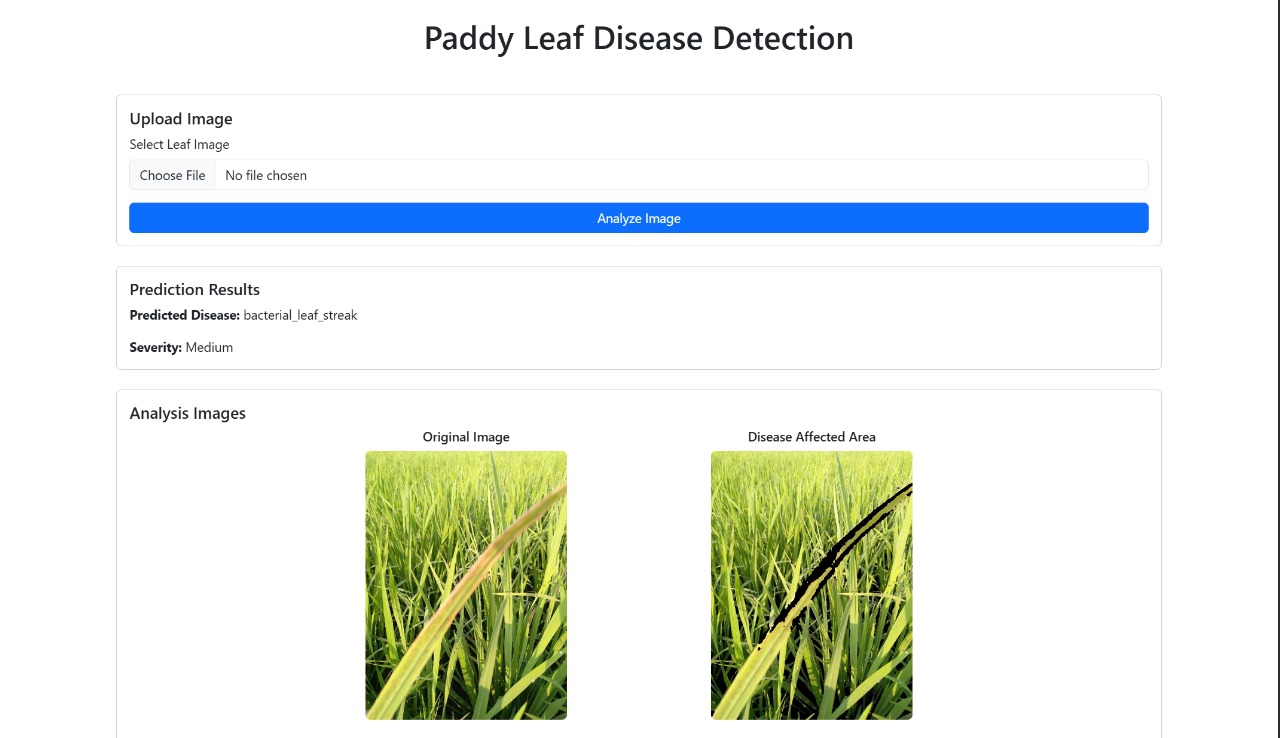
Example:  


Figure 6: Predicted Disease and Severity on Web Interface

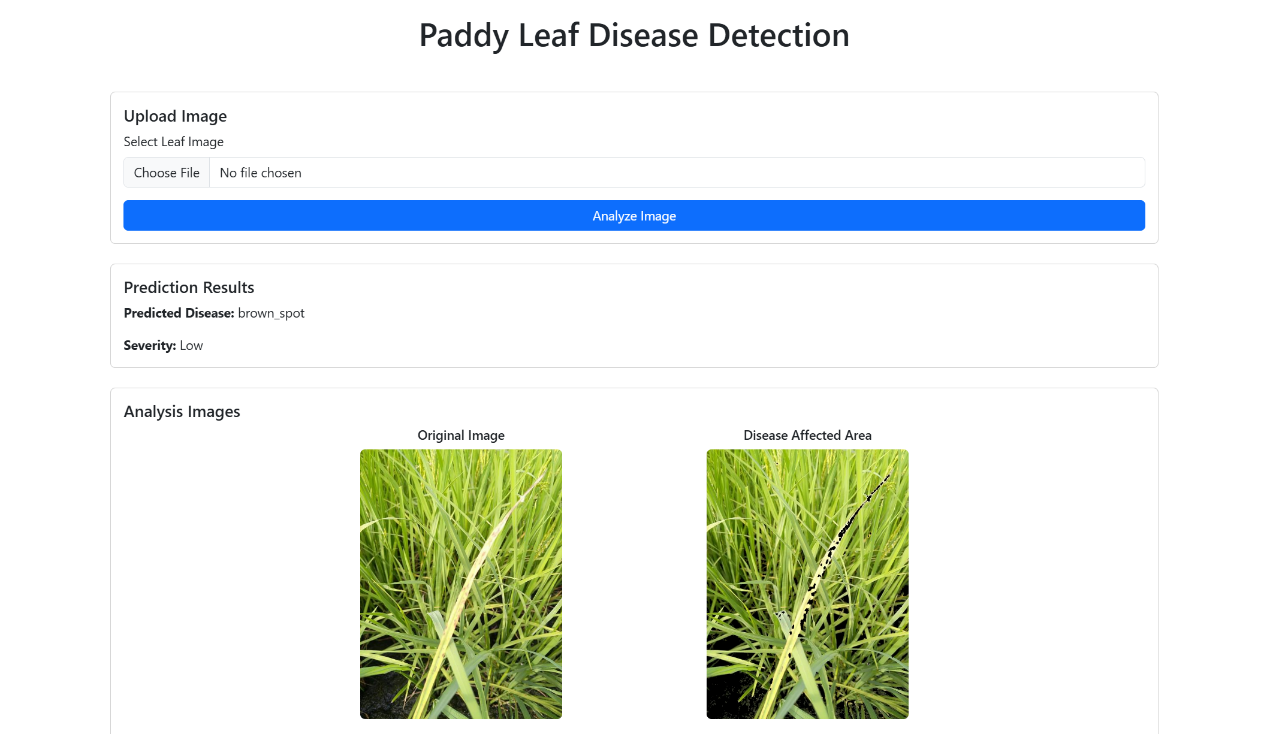


Figure 7: Predicted Disease and Severity on Web Interface

**5.4 Summary Table**

| **Image Name** | **Predicted Disease** | **Severity** |  |
| --- | --- | --- | --- |
| Fig. 6 | Bacterial leaf streak | Medium |  |
| Fig. 7 | Brown spot | Low |  |

**6. Conclusion**

The proposed system efficiently detects paddy leaf diseases and estimates severity using a lightweight deep learning model combined with image processing techniques. It offers:

* Real-time disease prediction.
* Visual feedback on diseased areas.
* Practical deployment potential for mobile and web platforms.

The system outperforms existing models in terms of disease coverage, speed, accuracy, and usability