**** **Annexure-E**

**GANDHI INSTITUTE OF ENGINEERING AND TECHNOLOGY UNIVERSITY, ODISHA, GUNUPUR**

**Department of Computer Science & Engineering**

**Statement of Purpose – Major Project-I**

## 1. Title of the Project:

Edge-Aided Plant Species Identification Using Leaf Image Analysis and Hybrid Classification Techniques

**2. Abstract:**

## Figuring out the plant species from a simple leaf image sounds easy, but it's actually quite tricky- especially when images are taken in the real world. There might be shadows, blur, messy backgrounds, or just plain noise. In this project, we’re planning to build a plant recognition system that still works well, even when the image quality isn't great. We’ll be testing four classic edge detection methods-Sobel, Prewitt, Laplacian of Gaussian, and Canny-to see which one best captures the shape and vein structure of leaves. Before that, we’ll process the images by making them grayscale, resizing them, and adding noise at three levels: low, medium, and high, to simulate real conditions. After applying the methods, we’ll check how each one performs based on things like image clarity (PSNR), similarity to the original (SSIM), processing time, and how effectively it captures the edges. We’re mainly trying to see which method still gives decent results when the image isn’t perfect. Hopefully, what we learn here will make it easier to build basic tools that can help with plant.

## 3. Objectives:

* To find out if the uploaded image really belongs to a leaf or not.
* To build a new kind of algorithm that focuses on edge detection and leaf patterns like veins and shapes.
* To use methods like Sobel, Canny etc. to capture margins and important curves in the leaf.
* To use those extracted features to tell which plant species the leaf belongs to.
* To combine both traditional and machine learning methods for better result and less error.
* To finally give the plant name by analyzing just one single photo of its leaf.

## 4. Introduction / Problem Statement:

In agriculture and botany fields, identifying plants based on their leaf is important but it is often difficult manually. Many farmers and students face this challenge due to extensive botanical knowledge and lack of proper tools. Though some methods exist already, they are not accurate enough and don't work well in real-world noisy image conditions. Also, a significant and common problem found in most of today's machine learning-based plant detection schemes is that they fail to consistently distinguish real plant leaves from non-leaf green materials. This results in repeated misidentifications, with things like green paper, painted walls, or artificial leaves being mistakenly detected as leaves. These false positives discredit and render the system useless for real-world use. So, we thought to build a better system using edge-detection + hybrid classifiers. Our focus is on improving detection even when the image quality is low or leaves are twisted or damaged

## 5. Expected Input & Expected Output:

* Input:
* The system will take stored images of plant leaves in JPEG or PNG format, uploaded through a simple website interface. These images might be noisy, blurry, or have natural backgrounds - no need for freshly clicked or clean images.
* Output:
* Step 1: The system checks if the uploaded image contains a valid leaf or not by analyzing shape, edge, and texture patterns.
* Step 2: If confirmed, it will identify the plant species using edge-based and shape-pattern algorithms (like margins, veins, outlines)
* It will also display a confidence score or accuracy level of the result.

## 6. Proposed Methodology / Techniques to be Used:

* **Image Preprocessing:** First, we’ll clean up the leaf images by removing noise using filters like Gaussian blur or median smoothing. This helps edges and patterns to show more clearly.
* **Edge Detection:** To catch the shape and structure, we will test edge detectors like Sobel, Canny, Prewitt, and Laplacian of Gaussian (LoG). These will help to highlight the borders and inner veins of the leaf.
* **Feature Extraction:** After edges are clear, we extract features like:
* **Geometric Descriptors:** We check things like leaf roundness, length-to-width ratio, and aspect ratio to get shape clues.
* **Vein Pattern Analysis:** Using morphological thinning and something called PNR (Peak-to-Noise Ratio) to get inner vein structures.
* **Margin Contour:** We try techniques like chain code to capture how smooth or jagged the outer edge is.
* **Classification Algorithm:** For identifying which plant it is, we start simple with KNN and SVM models.

If the dataset is big enough, we’ll try using deep learning like CNNs and MobileNet.

We also plan to mix rules (like fixed edge ratios) with ML predictions - a kind of hybrid logic.

* **Validation:** Using accuracy, precision, recall and confusion matrix.

## 7. Novelty / Contribution:

* Most plant recognition models only use CNN blindly without edge features. But we are trying to combine edge detection + geometric + texture together, which improves recognition even in bad light or twisted leaves. Our new contribution is testing a hybrid algorithm which will try to do both smart filtering + machine learning based prediction.
* Also, existing works don’t first check if image is leaf or not - we are adding this step to reduce wrong inputs.

## 8. Dataset Description (if applicable):

* This collection combines over 54,000 crop leaf images taken in controlled, lab-like environments with about 18,500 images of wild plant leaves captured in real-world outdoor settings.
* The crop images are grouped into 38 categories.
* The wild plant sorted into folders by class, offering a rich and varied dataset for developing and testing plant species detection models across different conditions.
* Images include different shapes, sizes, background, rotation, and light condition.
* we are going to do the data preprocessing of the data set we have collected according to the requirement of our project and research requirement as we move along the development of project

## 9. Expected Outcomes / Deliverables:

* A working model that can take leaf image and tell the plant species.
* Code notebook or GUI interface (Tkinter or Web App with Flask).
* Accuracy report and result screenshots included for evaluation and reference.

## 10. References / Tools & Technologies:

* Python, OpenCV, Scikit-Learn
* Jupyter Notebook or VSCode IDE and Google Colab for training
* Edge Detection Algorithms: Sobel, Canny, Prewitt
* Feature extraction: Hu Moments, Contour detection, Chain Codes
* ML Models: SVM, KNN, CNN, MobileNet
* Paper: “Leaf Recognition Based on Shape Features” – IEEE, 2018
* Datasets: <https://tqwei05.github.io/PlantWild/?utm_source>

<https://www.kaggle.com/datasets/mohitsingh1804/plantvillage>

**Signature of the student Signature of the supervisor**