Automated Leaf Classification using Machine Learning and Image Processing

Md Nahin Islam
U00899990
Date: 23rd April, 2025

THE UNIVERSITY OF MEMPHIS
Herff College of Engineering

Introduction

- Importance in botany, agriculture, environmental monitoring
- Manual identification is time-consuming and requires expertise
- Aim: Use AI to automate species recognition from leaf images

Objectives

- Build a dataset of leaf species
- Apply preprocessing & feature extraction
- Train SVM, Random Forest, and CNN models
- Evaluate models using standard metrics
- Develop a real-time identification system

Dataset Overview

- UCI Leaf Dataset
- Contains images of 40 plant species
- Two types of data:
  - RGB images for CNN-based modeling
  - CSV file with 14 pre-extracted features (shape, texture, etc.) for SVM & RF

Methodology

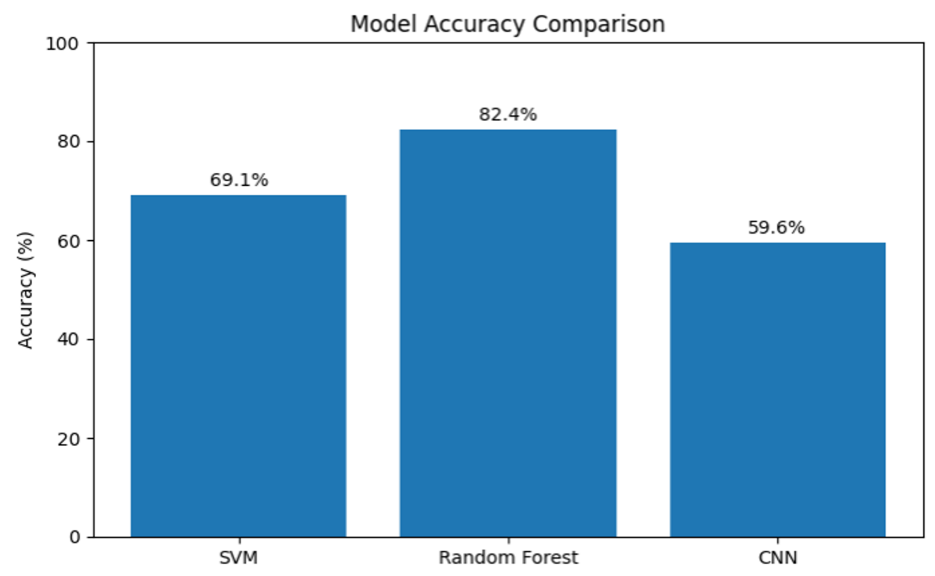
- SVM
  - Used 14 numerical features from CSV
- Random Forest
  - Ensemble of decision trees
- CNN
  - Trained directly on RGB images
  - Included data augmentation to reduce overfitting

Data Preprocessing

- For SVM & Random Forest (CSV Data)
  - Applied StandardScaler to normalize feature values
- For CNN
  - Resize all images to 128x128 pixels
  - Applied Data Augmentation to combat overfitting

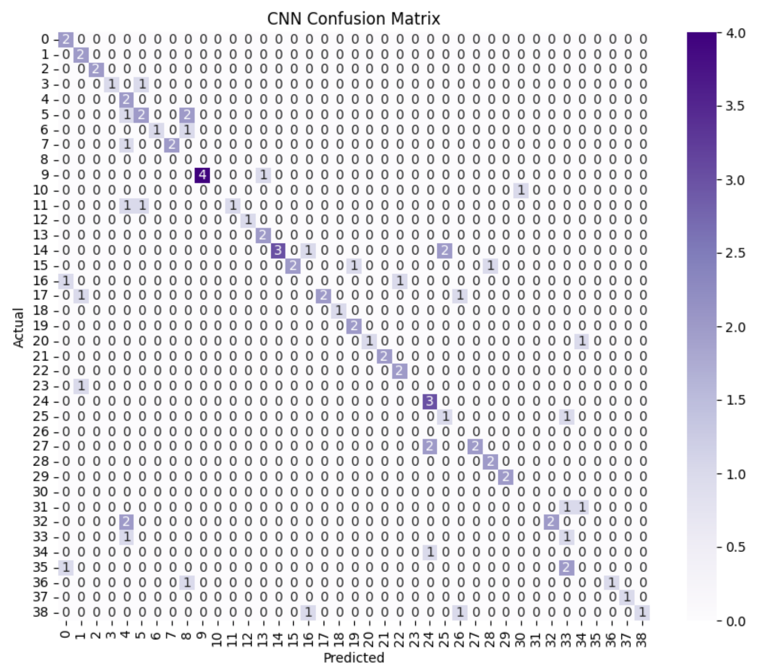
Results & Model Comparison

- Random Forest: 82.4% (Best Performer)
- SVM: 69.1%
- CNN: 71%



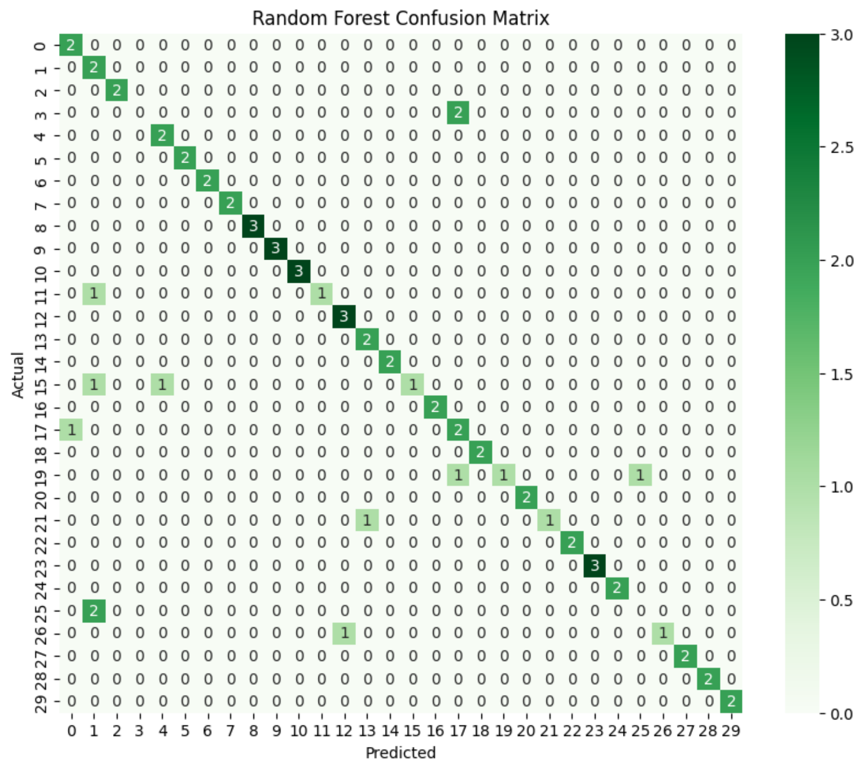
Results & Model Comparison

- CNN Confusion Matrix

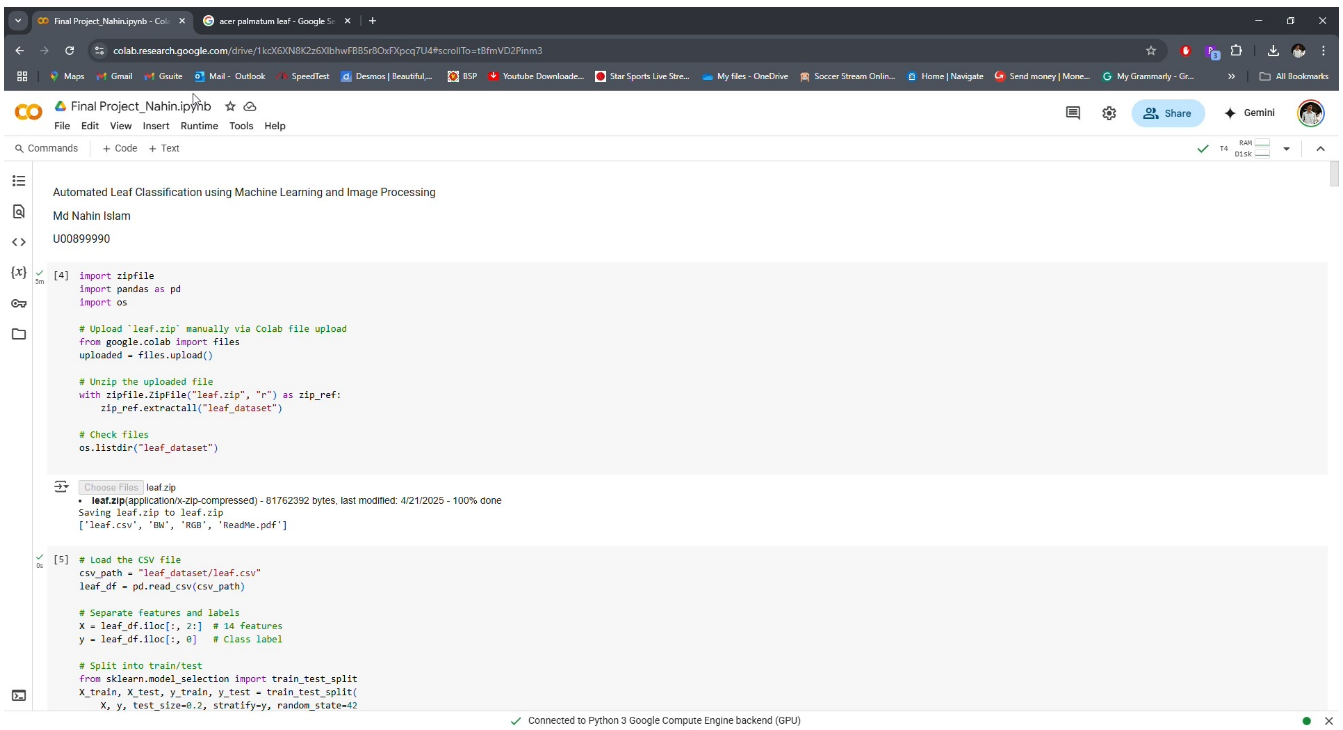


Results & Model Comparison

- Random Forest Confusion Matrix



Real-Time Leaf Identification System



Conclusion & Future Work

- Implemented SVM, Random Forest, and CNN for automated leaf classification
- Random Forest achieved highest accuracy using structured features
- CNN performance improved to 71% with data augmentation techniques
- Developed a real-time identification system
- Expand dataset with diverse, real-world leaf images

THANK YOU!