

Project Proposal

Plant Disease Detection Using Leaf Images

Problem Statement:

Plant diseases cause significant losses in agriculture worldwide. Early and accurate detection of plant diseases can help farmers take timely action to protect crops and improve yield. This project aims to develop a machine learning-based image classifier that can identify common diseases in plants from leaf images, enabling easy and fast diagnosis.

Objectives:

- Train an image classification model to detect various plant diseases from leaf images.
- Develop a user-friendly web application allowing users to upload leaf images and receive disease diagnosis instantly.
- Achieve a classification accuracy greater than 85% on the test dataset.

Proposed Methodology:

Dataset Selection:

Use the PlantVillage dataset, a large, publicly available dataset with over 50,000 labeled images of healthy and diseased leaves across multiple crops and diseases.

Dataset link: [PlantVillage on Kaggle](#)

Data Preprocessing:

- Resize images to a uniform size (e.g., 224x224 pixels) suitable for CNN input.
- Normalize pixel values.
- Split dataset into training, validation, and test sets (e.g., 70%-15%-15%).
- Apply data augmentation (rotations, flips) to improve model robustness.

Model Building:

- Start with a Convolutional Neural Network (CNN) architecture.
- Use transfer learning with pretrained models like MobileNetV2 or ResNet50.
- Add a softmax layer for multi-class classification.

Project Proposal

Model Training and Evaluation:

- Train the model using Adam optimizer and categorical cross-entropy loss.
- Use early stopping based on validation accuracy and loss.
- Evaluate model performance using accuracy and confusion matrix.

Deployment:

- Build a Streamlit web app for image upload and disease prediction.
- Display predicted disease name and confidence score.
- Optionally show disease details and recommended actions.

Dataset Details:

- Source: PlantVillage Dataset (Kaggle)
- Size: ~54,000 images
- Classes: Over 38 including healthy and diseased leaves of multiple crops
- Image type: Color images of leaves, various disease conditions visible

Expected Outcomes:

- An image classification model that accurately identifies plant diseases with >85% accuracy.
- A Streamlit app enabling prediction of plant disease from leaf images.
- Documentation explaining the model and user instructions for the app.

Timeline (3-4 Days):

- Day 1-2: Download dataset, preprocess images, perform data augmentation, setup environment.
- Day 3-4: Train the CNN model, validate, tune hyperparameters, evaluate results.
- Day 5-7: Build Streamlit frontend, integrate model, enable prediction.
- Day 7-10: Final testing, debugging, documentation, and packaging.