**Project Report**

**Plant Disease Classification using CNN and Vision Transformer**

**1. Objective**

The objective of this project is to develop an automated system for classifying plant diseases from leaf images using deep learning techniques. The system aims to assist farmers and agricultural experts in early and accurate disease detection, thereby improving crop yield and reducing losses.

**2. Problem Statement**

Plant diseases are a major threat to food security and agricultural productivity worldwide. Manual identification of plant diseases is time-consuming, requires expert knowledge, and is prone to human error. There is a need for an automated, accurate, and scalable solution that can classify plant diseases from images, enabling timely intervention and treatment.

**3. Methodology**

**a. Dataset**

- Source: PlantVillage dataset from Kaggle (https://www.kaggle.com/datasets/emmarex/plantdisease)

- Size: >1GB, 15 classes (for the color subset)

- Structure: Images are organized into class folders and split into train, validation, and test sets.

**b. Data Preprocessing**

- Images are resized to 224x224 pixels.

- Data augmentation (random horizontal flip, rotation, color jitter) is applied to the training set.

- Images are normalized using ImageNet mean and standard deviation.

**c. Model Architectures**

**1. Convolutional Neural Network (CNN):**

- 4 convolutional layers with batch normalization and max pooling.

- Dropout for regularization.

- Two fully connected layers for classification.

**2. Vision Transformer (ViT):**

- Pretrained ViT (vit\_base\_patch16\_224) from the timm library.

- Fine-tuned on the PlantVillage dataset.

**d. Training**

- Both models are trained using the AdamW optimizer and cross-entropy loss.

- Learning rate scheduling with warmup is used.

- The best model (based on validation accuracy) is saved for each architecture.

**e. Evaluation**

- Models are evaluated on a held-out test set.

- Metrics: Accuracy and loss on train, validation, and test sets.

- Random test images are used for demonstration of prediction.

**4. Results**

**CNN Model:**

- Achieved high training and validation accuracy after several epochs.

- Test accuracy:

**Vision Transformer (ViT):**

- Achieved competitive or superior accuracy compared to CNN.

- Test accuracy:

**Sample Prediction Output:**

- The model correctly predicts the disease class for random test images with high confidence.

**Training Curves:**

- Loss and accuracy curves show effective learning and generalization.

**5. References**

1. PlantVillage Dataset: https://www.kaggle.com/datasets/emmarex/plantdisease

2. Vision Transformer (ViT): Dosovitskiy, A., et al. (2020). "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." https://arxiv.org/abs/2010.11929

3. PyTorch Documentation: https://pytorch.org/docs/stable/index.html

4. timm Library: https://github.com/huggingface/pytorch-image-models

5. Plant Disease Detection using Deep Learning: Mohanty, S.P., Hughes, D.P., Salathé, M. (2016). "Using Deep Learning for Image-Based Plant Disease Detection." https://arxiv.org/abs/1604.03169

**6. Conclusion**

This project demonstrates the effectiveness of deep learning, specifically CNNs and Vision Transformers, in classifying plant diseases from images. The system can be further improved by using larger datasets, more advanced models, or deploying as a mobile/web application for real-world use.