Problem Statement: Accurate and efficient detection of plant diseases is crucial for effective crop management and yield optimization. Traditional methods often rely on manual inspection, which can be time-consuming and require expert knowledge. The paper aims to develop a hybrid model that leverages the strengths of both Vision Transformers (ViTs) and Convolutional Neural Networks to improve plant disease detection while reducing the number of trainable parameters.

Dataset Used: The study employed a publicly available dataset on potato plant diseases. The dataset consists of 4072 items, categorized into 3 categories based on type of potato disease. The data is divided into a training set, a validation set, and a test set, comprising 80%, 10%, and 10% of the dataset respectively. This allocation amounts to 3251 items for the training set, 416 for the validation set, and 405 for testing the model. Potato Disease Leaf Dataset(PLD)

Methodologies Used: By integrating convolutional layers with transformer architectures, the model effectively captures both detailed local patterns and broader contextual information. This design aims to enhance accuracy while maintaining a manageable number of trainable parameters, making the model more efficient.

Findings and Contributions: The hybrid model demonstrated state-of-the-art performance, achieving a training accuracy of 92.83% and a test accuracy of 88.86% on the dataset obtained with the K850 lens. These results highlight the model's effectiveness in accurately identifying plant diseases. The study's key contributions include:

- Introducing a hybrid architecture that combines ViTs and CNNs for plant disease detection.

- Utilizing multispectral imaging to capture a comprehensive range of data across different wavelengths.
- Achieving high accuracy with a reduced number of trainable parameters enhances the model's efficiency.

Relevance to our Project: For a project focused on AI-based plant disease detection, this paper offers valuable insights into model architecture and data acquisition strategies. The hybrid approach provides a balanced solution that leverages the strengths of both ViTs and CNNs, potentially leading to more accurate and efficient disease detection systems.