Indo/US Collaborative Research Grants

National Science Foundation of US and Technology Innovation Hubs of India

सत्यमेव जयते

Title: Machine Learning-based Crop Diseases Detection on the Edge
Indian PI: Prof. Maryam Shojaei Baghini, Prof. Rajbabu Velmurugan, IIT Bombay and TIH-IoT, India
Indian Collaborator: Dr. R. Ramesh, ICAR-CCARI, Goa, india
US PIs: Prof. Jingtong Hu, University of Pittsburgh, Pittsburgh, PA, US.

This collaborative project aims at experimental steps towards on-the-edge AI-based diagnosis of diseases common among plants such as tomato and eggplant without sending data to the cloud. The main motivation is to reduce the cost on reliable internet access, cloud services and remote server pertaining to the budget-constrained farming. Moreover, due to the shortage of manpower and facilities for generating data for the early-stage diagnosis, controlled data generation with optimal processing and labeling are necessary.

There are several soil and ambient weather-related parameters that impact the spread of plant diseases. Hence, it is crucial to measure these parameters and analyze their temporal variation in affordable ways. Another required modality is cost-effective visible light and infrared imaging supported by established image analysis algorithms and AI/ML-based modeling and inference. Hence, we have brought multi-modal edge computing as the ultimate goal based on which the this project was defined. **The core target novelty lies in the combination of controlled plant cultivation, and controlled plant infection at the small scale, efficient data labeling, developing appropriate algorithms for quality and selective processing of the low/medium-size data and guiding the direction of optimal model generation with explainable results (TIH IIT Bombay, India). The evaluation of such algorithms in a systematic manner, model discovery and optimization, transfer learning, success/failure analysis and efficient multi-modality neural networks implemented on the edge devices (University of Pittsburgh) completes the project aim.**

The AI/ML and edge computing for the plant disease detection requires locally produced datasets from plant cultivation under controlled conditions, followed by the data processing, analysis, multi-modal modeling and implementation for the inference on the edge devices. This is a well-defined extension of both the current NSF project focusing on the multi-modal and edge computing hardware and the TIH IIT Bombay funded project on the AI/ML based effective and efficient agriculture.

We investigate and analyze various models to choose a backbone model achieving promising accuracy as per the literature and later optimize it preserving the accuracy. Since the data for plants are very limited for both public datasets and our produced datasets, dataset processing, segmentation and optimization are necessary. We transfer the model pre-trained on the publicly available dataset to our target dataset. The overall binary classification accuracy achieved on the public PlantVillage dataset of 99.92%-100%. By applying the transfer learning using the custom collected dataset, the preliminary selected backbone model exhibits inference accuracy of around 84% with multiple leaves and 89% with single leaf. We then implement the model on the edge devices.

We have implemented the backbone model on the embedded GPU Jetson TX2. There is trade-off between latency and power consumption for different batch sizes. We will explore applying pruning and quantization to further optimize the implementation efficiency.