Brunel University London
College of Engineering Design and Physical Sciences
Dept. of Electronic and Electrical Engineering

Department of Electronic and Electrical Engineering

Plant Disease Detection using Machine Learning

Aim

This project aims to create a **Convolution Neural Network** to detect and classify plant diseases. With artificial intelligence being integrated into many established sectors, the current artificial intelligence research techniques are outdated and resource intensive. The ultimate goal is to **establish a benchmark for achieving high accuracy with minimal computational resources**, paving the way for sustainable machine learning practices in plant disease detection and classification tasks.

Method

This project presents a **novel Convolution Neural Network (CNN) architecture** prioritising model simplicity and complying to the principles of 'green AI' [1]. In contrast to current methods, the architecture presented has a 1:1 ratio of convolution to max pooling layers, this drastically decreases the number of learnable parameters, **boasting only 184,737 learnable parameters** compared to the millions presented by current state of the art models.

Dataset

The neural network is trained using the popular open source PlantVillage dataset [2]. It consists of 15 different classes with a total of 20,639 labelled images. All images are resized to 256 x 256 and through dataset reduction to minimise bias to the larger classes, the model was trained, validated, and tested using 9,819 of the images.

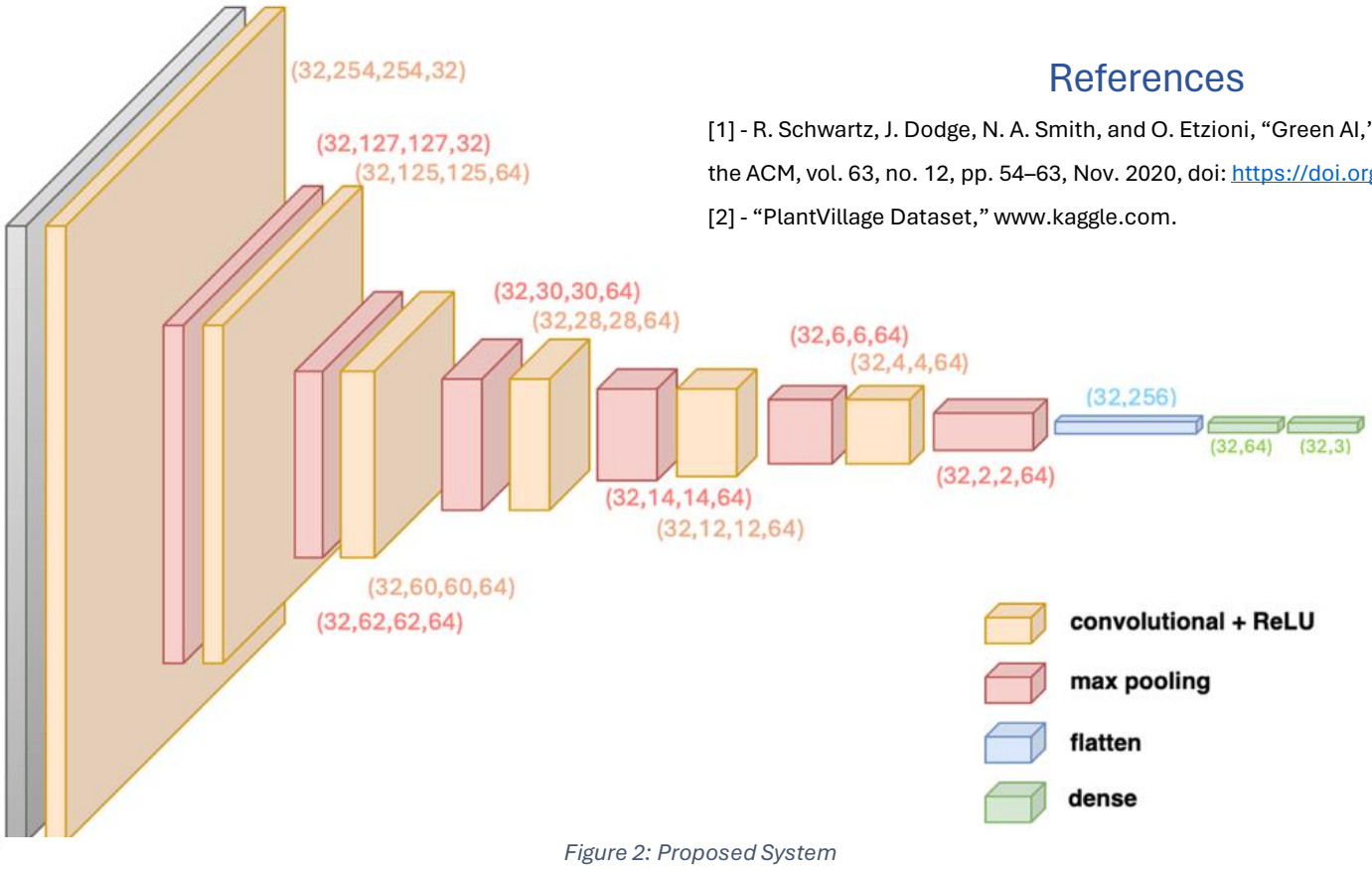

Figure 2: Proposed System

Figure 1: Example of images from the PlantVillage dataset

Results

Experiment	# images	Accuracy	Carbon emissions
D	20,639	98.03%	0.53 kg eq. CO2
E	16,880	97.64%	0.41 kg eq. CO2
G	9,819	98.61%	0.26 kg eq. CO2

By reducing the dataset to minimise the class imbalances, the training time and emission decreased, and the accuracy increased. These results show that more emphasis should be put on quality of data rather than quantity, and complexity of the system.

Future Work

The reduction of the dataset was done completely randomly, in future a more measured approach such as a PCA could be done to maximise the number of components in the minimum number of images further improving the quality of the data.

References

[1] - R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, "Green AI," Communications of the ACM, vol. 63, no. 12, pp. 54–63, Nov. 2020, doi: https://doi.org/10.1145/3381831.

[2] - "PlantVillage Dataset," www.kaggle.com.

Milo Adamski

Supervisor – Dr. Tatiana Kalganova