Capstone Proposal

Allyson Julian

December 30, 2019

1 Introduction

Advances in agricultural technology over the past 30 years have made it easier for farmers to manage their farms, particularly when the farms are comprised of multiple fields greater than 1000 acres in size. The use of GPS on aerial imagery, drones, etc. have been instrumental in precision agriculture[1].

But one of the challenges of modern farming is the detection of plant diseases. For large farms in particular, it is time-consuming for a farmer to manually check each growing plant for disease. It can potentially be more cost-effective to diagnose plant diseases with automated tools [2].

2 Problem Statement

Farmers need a way to detect plant disease early on in the planting season without investing too much time manually checking everything themselves.

They need software and hardware tools to do plant disease detection at scale to be more cost-effective in the planting process. In other words, there must be a classifier built that can process an image of a plant and detect whether a plant is diseased or not and what disease it has.

Access to reliable cellular and/or wifi data is often limited in rural areas, so any kind of tool for this problem area must also be able to run offline on cheap embedded systems built specifically for machine learning, e.g. NVIDIA's Jetson Nano:

`https://developer.nvidia.com/embedded/jetson-nano-developer-kit`

3 Dataset and Input

The main dataset used for training and validation will be the PlantVillage Dataset obtained from GitHub:

`https://github.com/spMohanty/PlantVillage-Dataset`

The dataset consists of 20,638 images each of which have the dimensions 256 by 256 pixels and are all colored, non-greyscale JPEGs. The images are split up into several folders, labelled according to the plant species (e.g. Pepper Bell) and whether they are healthy or not (e.g. Pepper bell healthy, Pepper bell Bacterial spot).

4 Solution Statement

To address the need for a plant disease detecting tool, I will be building a Python library with an API that will have two major functions: image preprocessing necessary to prepare the data for classification, and the classifier itself. This solution is based on the work done in some previous studies [3][4].

The classifier will be built with using a Convolutional Neural Network (CNN) with TensorFlow. This model will be based on the CNNs described in previous studies [5][6].

The library will be used in a web app that will act as a UI for the trained classifier using the Flask framework. It will allow the user to upload photos to be preprocessed in the backend and then sent to the endpoint for classification.

A functioning prototype of the plant disease detector will be built using a Jetson Nano robot equipped with a camera. This will act as a proof-of-concept as to what can be achieved using open source machine learning and cheap embedded systems.

5 Benchmark Model

The benchmark for the classifier will be the results obtained by previous studies by Toda et al, Fuentes et al, on plant detection which all utilize a CNN as a classifier[6] [5].

6 Evaluation Metrics

The metrics used to evaluate the performance of the model will be the accuracy of the model when it comes to the predicted classes vs the actual ones, confusion matrix, and F-score.

7 Project Design

7.1 Preprocessing Data

The images will be preprocessed according to methodologies similar to those described in previous studies[7][6][5]: (1) color transformation of the RGB values in the image, (2) masking of the non-diseased parts of the leaf.

7.2 Training Model

The plant disease classifier will be built using TensorFlow, with the model being a Convolutional Neural Network (CNN) trained on the image dataset of 256x256 images in the PlantVillage dataset. This type of model was chosen because CNNs have shown promising results in disease classification in previous studies on this topic[4][6].

7.3 Deployment

The trained model will be imported into AWS SageMaker so that it can be deployed as an endpoint. The Flask web app will act as a front-end for that endpoint. It will have an upload screen to allow the user to submit a photo of a plant leaf and determine what disease it has, if any.

References

[1] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: A Review," *Sensors*, vol. 18, no. 8, p. 2674, Aug. 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/8/2674

[2] E. Fujita, Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi, "Basic Investigation on a Robust and Practical Plant Diagnostic System," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*. Anaheim, CA, USA: IEEE, Dec. 2016, pp. 989–992. [Online]. Available: http://ieeexplore.ieee.org/document/7838282/

[3] M. K. Tripathi and D. D. Maktedar, "Recent machine learning based approaches for disease detection and classification of agricultural products," in *2016 International Conference on Computing Communication Control and automation (ICCUBEA)*. Pune, India: IEEE, Aug. 2016, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/7860043/

[4] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," 2016. [Online]. Available: https://www.hindawi.com/journals/cin/2016/3289801/

[5] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition," *Sensors*, vol. 17, no. 9, p. 2022, Sep. 2017. [Online]. Available: https://www.mdpi.com/1424-8220/17/9/2022

[6] Y. Toda and F. Okura, "How Convolutional Neural Networks Diagnose Plant Disease," 2019. [Online]. Available: https://spj.sciencemag.org/plantphenomics/2019/9237136/

[7] H. Al Hiary, S. Bani Ahmad, M. Reyalat, M. Braik, and Z. ALRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases," *International Journal of Computer Applications*, vol. 17, no. 1, pp. 31–38, Mar. 2011. [Online]. Available: http://www.ijcaonline.org/volume17/number1/pxc3872754.pdf