Smartphone Based Plant Disease Classification using Convolutional Neural Networks

Meneses, Brymer Bernard R.

12 - STEM - Newton

Contents

Chapter 1

Introduction

1.1 Background of the Study

Agriculture plays a major role in feeding billions of people all across the globe. It has also provided a means for people to grow revenue and sustain their lives. In fact according to the World Bank Organization, "Agricultural development is one of the most powerful tools to end extreme poverty, boost shared prosperity and feed a projected 9.7 billion people by 2050. Agriculture plays an integral part of the Philippine Economy, in fact according to the Philippine Statistics Authority, the Gross Value Added (GVA) in agriculture of the year 2019 is valued at 1.78 Trillion Pesos. It also contributes to the diet of Filipinos producing various ingredients of that are necessary for the Filipino Cuisine such as rice being the staple ingredient and other leafy vegetables such as cabbage, saluyot and ampalaya.

Each year, farmers spend a lot of money on disease management, they often do so without proper technical support leading poor disease control, pollution and harmful result. Plant diseases cause substantial loss to farmers resulting to large economic loss. That's why accurate identification and diagnosis of plant diseases is essential; especially now in the era of globalization and climate change.

During the past few years, there has been a rapid advancement to the era of computer vision. There are now self-driving cars that can fully control a vehicle with minimal efforts required by the driver. Engineers and researchers have achieved this through the use of Convolutional Neural Networks (CNN). And several researchers have also applied it to the field of biology as well, (Hadush et al.) has used to detect CNNs to detect breast cancer from mammogram(MG) images achieving a detection accuracy 91.86%.

This study aims to develop a mobile application that uses a CNN Model to detect and classify potential diseases in plants. Due to the current pandemic, the researcher decided to use publicly available image datasets — which will be used to train several CNN Models that would be embedded in the application — instead of making one. Because the model inference will happen on the device itself without relying on a cloud-based server, farmers and other people would be able to input an image to this application and use this app without having to rely on an existing data or internet connection.

1.2   Conceptual Framework

1.3   Theoretical Framework

1.4   Statement of the Problem

1. How effective is the model in predicting the type of disease present in the given image? In terms of:

   (a) F1 Score

   (b) Speed of Classification

2. How comparable is the model made by the researcher to other methods of detecting plant disease?

1.5 Hypotheses

1. The F1 Score and the Speed of Classification implies that the model is not effective in predicting the tyoe of disease present in the given image.

2. The model is not comparable to other methods of detecting plant disease.

1.6 Importance of the Study

The significance of the study lies on the fact that plant disease diagnosis plays a huge role in minimizing the monetary and material loss in agriculture caused by various plant diseases. That's why proper plant disease diagnosis is a detrimental process in securing the value of crops grown. The study proposes an application that can be used to scan images of plant leaves which would then be used to detect various plant diseases which may be present in an image, this application is designed to run the disease diagnosis without the need of internet or data connection. This study would certainly benefit farmers who have no access to internet connection to diagnose their plants and prevent further crop losses.

1.7 Definition of Terms

- Model

  - A machine learning model is a file that has been trained to recognize certain types of patterns. You train a model over a set of data, providing it an algorithm that it can use to reason over and learn from those data.

- Convolutional Neural Networks

  - A Convolutional neural network (CNN) is a neural network that has one or more convolutional layers and are used mainly for image processing, classification, segmentation and also for other auto correlated data.

- Gross Value Added

  - Gross value added is the measure of the value of goods and services produced

in an area, industry or sector of an economy.

- F1 Score

  - The F-score, also called the F1-score, is a measure of a model's accuracy on a dataset. A good F1 score means that you have low false positives and low false negatives, so you're correctly identifying real threats and you are not disturbed by false alarms

- Machine Learning

  - Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves.

- Classification

  - classification refers to a predictive modeling problem where a class label is predicted for a given example of input data.

Chapter 2

Methodology

The researcher followed different methods and steps to conduct this study. The data gathering procedures, experimental design, methods and steps and the statistical treatment of data are also included in this chapter.