CHAPTER 1
INTRODUCTION

1.1. Background

Fruits are the edible fleshy parts of plants that contain seeds. The three main classifications of fruits are drupes, pomes, and berries. Drupes are fruits with a hard stony covering surrounding the seed, such as a peach. Pomes are fruits with several seeds enclosed in a fleshy envelope, such as an apple. Berries are fruits that have many seeds and a thin skin, such as a grape.

There are many different types of fruits, with different flavors, textures, and nutritional values. Some common fruits include apples, bananas, oranges, strawberries, and grapes. Fruits are an important part of a healthy diet as they are a good source of vitamins, minerals, and fiber. They can also help to lower cholesterol and blood pressure, and reduce the risk of some chronic diseases such as heart disease, stroke, and cancer.

Rotten fruits are a reality in the fruit and vegetable industry. While consumers expect to see an occasional bad apple or bruised banana, they don't expect to find a significant number of rotten fruits in their shopping bags. Unfortunately, that's often the case. In fact, a significant percentage of the fruits and vegetables that are grown each year never make it to market because they are rotten.

There are a number of reasons why fruits and vegetables rot. The most common cause is damage during harvesting or transportation. Rough handling can bruise fruits and vegetables, which makes them more susceptible to rot. In addition, fruits and vegetables that are not properly refrigerated during transport can also rot.

Once fruits and vegetables arrive at the grocery store, they must be properly stored to prevent rot. Most fruits and vegetables should be stored in a cool, dry place. However, some fruits and vegetables, such as tomatoes, need to be stored in a warm, humid environment. If fruits and vegetables are not stored properly, they will rot.

Rotten fruits can cause loss for the fruit seller, because if a fruit rot it can spread out to other fruits. To prevent and minimize the damage caused by the rotten fruit, the rotten fruit must be removed immediately. Its easy to do this if the fruits are in a small group , the problem is to find the rotten fruit in a large group of fruits.

To help solve this problem, I propose the use of machine learning in order to detect rotten fruit. The machine learning i propose is CNN because CNN is a common machine learning algorithm that can be trained to detect and recognize image. To train and test the CNN dataset consisting of images of different types of fruit in fresh and rotten states is used. The dataset will be splitted 80% for training and 20% for testing.

To make CNN detect an object in an image , the image needs to be divided into parts and then processed with cnn. The results of this is a visualization of the object that the algorithm thinks is the fresh or rotten fruits, and show the evaluation result of the algorithm using Mean Average Precision.

1.2. Problem Formulation

1. How to classify image data and train cnn using classified image data
2. How to turn cnn into object detection
3. How to visualize prediction result
4. Calculate MaP(Mean average Precision) from the object detection

1.3. Scope

In this project author uses javascript and html to create the program , Fruit fresh and rotten for classification dataset that were used for this project was taken from kaggle website totaling of 13.600 pictures of fresh and rotten fruits such as oranges, apples and bananas, the author of the dataset already split the data in two parts for training and testing . The next step is to determine how much images will be used by to train the cnn and to test the final program, after training the cnn the next step is to create image pyramid of the tested image, then create roi(region of interest) boxes then run it through every area tested image and pyramid of tested image, classify roi area, collect (coordinate, label and prediction percentage) of the classified area, run the collected data to non max suppression to eliminate overlapping coordinate area, then visualize the collected data based on non max suppression result. After the data has been visualized the next step is to calculate precision, recall, average precision and mean average precision.

1.4. Objective

The main objective of this project is to create a program to find damaged or rotten oranges that are hard to see inside a picture. The second objective of this project is to create a web based object detection program using cnn algorithm.