AIoT Coding, Engineering and Entrepreneurial Skills Education for Gifted Students

Defect Detection in Irwin mangoes (*Mangifera indica)*

Student(s): 42 Zachary

Project ID: P30

Abstract

Farmers, wholesalers and retailers around the world often rely on human labor to verify a fruit's saleability, leading to high percentages of costly food waste, supply chain carbon emission and unnecessary packaging. This project aims to develop a defect detection model in YOLOv8 to identify predetermined defects such as discolouration and bruising, initially using Irwin mangoes (*Mangifera indica*) common in Taiwan, to help address this issue by identifying defects faster, cheaper, and more accurately.

Objective/Background/Motivation

- Nearly ½ of all fruits and vegetables produced globally are wasted every year (Feeding Hong Kong)
- Large quantities of food at the retail level are classified as 'ugly fruit': fruit that are consumable but not fit for sale
- This results in high levels of waste and profit for retailers, wholesalers and farmers
- In addition, the enormous plastic and carbon footprint associated with the packing and shipping of the produce creates a **large net negative effect on the Earth's environment**

Why mangoes?

- Largest producers of mangoes are from Asia (India produces ~19m tonnes, China ~5m tonnes and Thailand ~3m tonnes) (S&P Global)
- American wholesalers/retailers like Costco and Amazon have implemented similar solutions for other fruit but are not in the Asian market yet
- Obtainable dataset from BIIC LAB's AI CUP 2020

- Motivation: Had started my small social enterprise at a young age to manually collect and repurpose 'ugly fruits' into snacks. The CityU EE GEF program motivated me to use AI and AIoT to address the problem at its source and scale my impact.

Methodology

- The model is developed with YOLOv8, the newest iteration of the original YOLO (You Only Look Once)
- Preparation of the dataset included Python modules such as PIL, rembg, and os, including the conversion of bounding boxes from FiftyOne to YOLO format
- The model classifies the defects into 5 categories, "乳汁吸附", "機械傷害", "炭疽病", "著色不佳", and "黑斑病" then draws bounding boxes around them to display it (Fig. 1)

Results/Application

While Irwin mangoes are not a species of mango with high produce and demand, and mangoes only comprise a small portion of the global food waste problem, this project aims to serve as a proof of concept for the classification and advancement of other similar Asian fruits on the market.

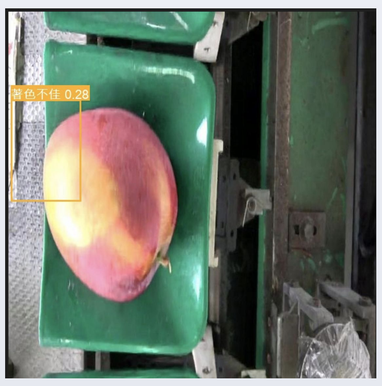

Figure 1: Example of bounding boxes on image

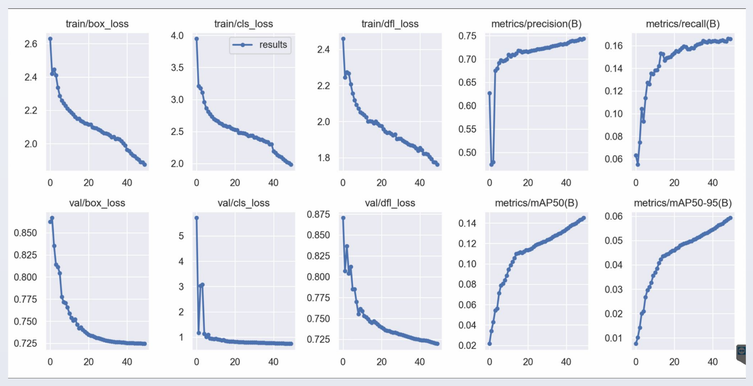

Figure 2: Results from first model

GitHub Link

QR Code for Video Demo

YouTube Link

Discussion/Conclusion

From this experience, I learned:

- the importance of preparing the right data before training
- the complexity regarding finding patterns within images with background noise
- not to underestimate the time it takes to process and train with large datasets
- how far machine learning has improve from being exclusive to experts to being access

Improvements:

- starting earlier to allow for more leeway in terms of training the model
- applying Gaussian blur to reduce amount of noise and speckles in the image
- greyscaling to see if it can help isolate certain defining aspects of the defects