

二、構想說明

(一)作品宗旨及目的

人工智慧下的深度學習(deep learning)，在 2012 年發表的 AlexNet 深度網路結構後掀起了全世界對於深度學習的重視，2018 年 Google DeepMind 以透過標註上萬份視網膜掃描影像中的病變區，訓練出能正確辨識青光眼的 AI 模組，讓影像的智慧化提升到新的應用層次。

放眼目前影像的 AI 部分，從協助人員決策支援到車輛自動控制的部分都看到 AI 涉入的身影，回首會發現 AI 能提供人力與效能提升上具有強大的功效，不但節省了大量的人力成本，並且確保了產出品質，對於目前農產品的應用上 AI 實際落地狀況有限，畢竟農田裡環境、生物、與人為操作等的變因太多了，就研究發展進度而言，目前仍停留在資料累積的階段。

YOLO 是目前 AI 方法論中被廣泛使用的物件偵測模型，其主要在於輸入整張圖片，直接預測物體類別、信心度與座標位置，目前已發展至 YOLOv4。透過此技術，可用來預測本題圖片中的產品以及產品中是否存在瑕疵，根據本題的需求，準確地框出產品與瑕疵範圍是首要目標。另一個主要目標是計算面積大小，由於要計算面積大小，因此得知框列產品的長度變成必須達成的目標。在現有的攝影技術中，能夠估計攝影圖像中的技術包括：

1. 視覺慣性測距(Visual Inertial Odometry, VIO)透過記錄行動裝置中的陀螺儀資訊與影像的變化，可估計出目標物的長度，目前相關的技術的應用框架為 ARKit。
2. 部分行動裝置的攝像頭具備 ToF 模組，用於測量物體的距離，達到快速對焦與提升強化拍攝出的影像。
3. 具備多鏡頭裝置可拍攝出深度的影像。
4. 以上技術可取得影像中的距離資訊，對於本系統預估面積來說，有直接的幫助。
5. 本題的構想主要基於 YOLO 模型，除了預測物體類別與座標位置外，加入物品的實際長度，如此便可順利進行面積預測。配合許多有偵測距離技術的行動裝置，讓訓練好的模型搭配行動裝置拍攝時所產生的距離資料，進行面積的預測。

(二)技術架構及進行步驟

1. 採用方法與原因:

I. AI 模型技術(用於預測與物體與物體面積):

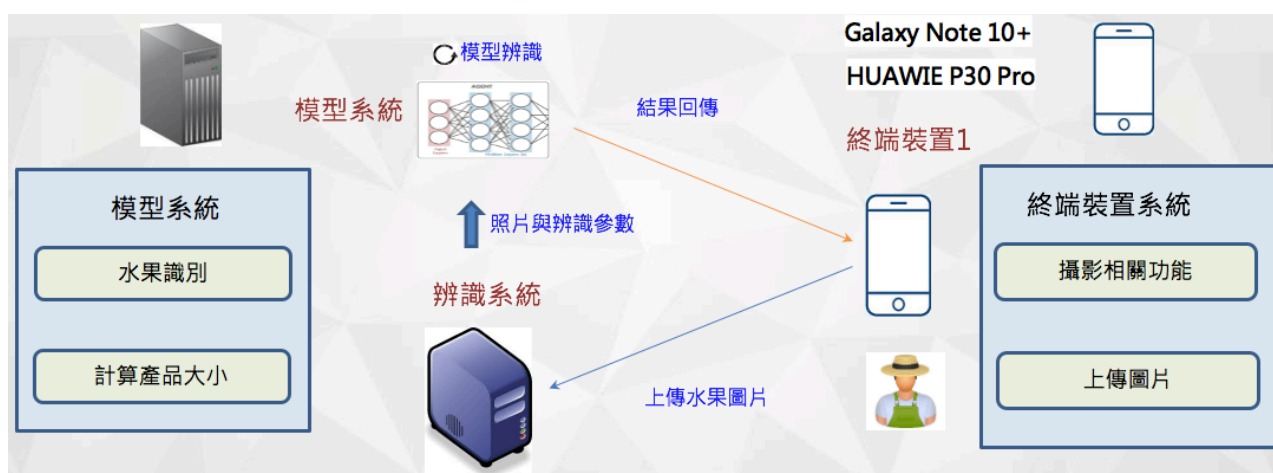
- 利用 YOLO 物體識別技術，識別出影像水果區域。
- 改良 YOLO 物體識別模型的輸出，加入水果的長度與面積的資訊訓練 AI 模型。

II. 相機與裝置相關技術(用於增加預測面積的準確率):

- 使用視覺慣性測距(Visual Inertial Odometry, VIO)技術預估距離，透過手機上的慣性測量單元(Inertial measurement unit, IMU)資訊與影像預測物體距離。
- 使用飛時測距技術(Time of Flight, ToF)技術預估拍攝物體距離。
- 多鏡頭影像深度解析技術預估影像深度。

本作品主要採用物體識別技術(YOLOv4)改良模型，加入距離面積以及上述相機技術的資訊進行模型訓練，預測結果為識別水果的區域面積，輔以裝置內本身所具備的相機技術校正預測結果，增加預測準確性。

2.系統環境架構說明:

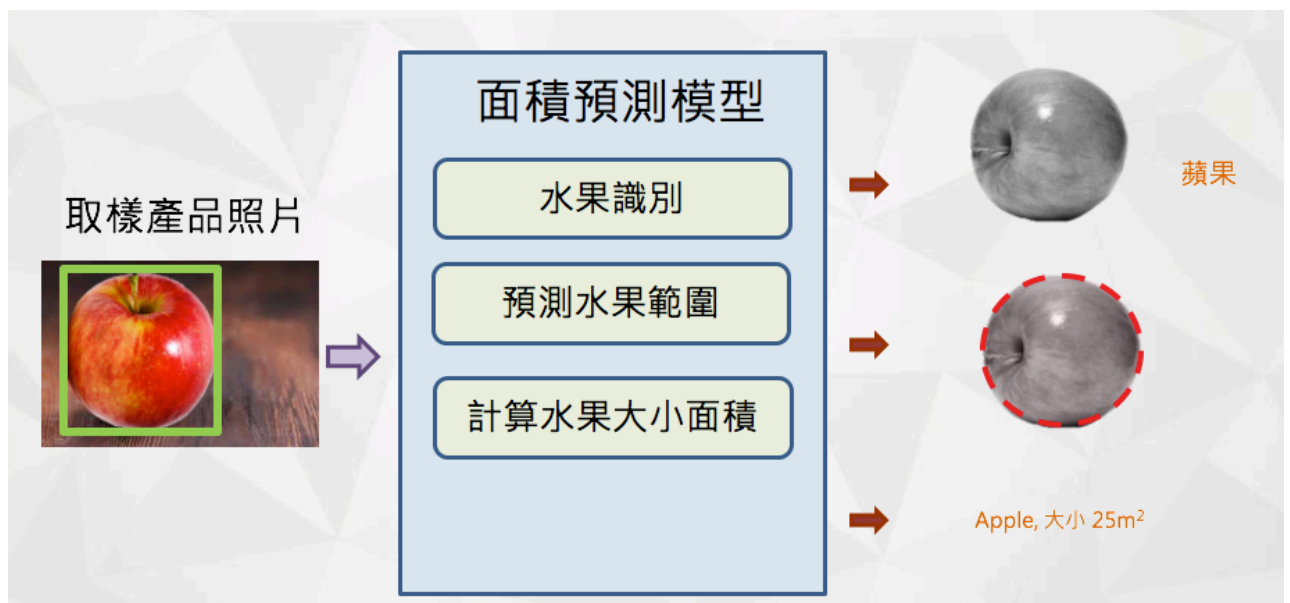


本作品分為主要部分為雲端模型系與終端裝置系統。透過手持的終端裝置進行水果的拍照，上傳圖片與該圖片相關攝影圖樣資訊(例如拍攝圖片當下的IMU、ToF相關資訊等)至雲端資料系統，資料系統根據各裝置建立相關資訊，並將訓練資料(包括圖片、可能實際大小與圖像資訊等)透過雲端模型系統進行水果的辨識，大小預測與計算，系統也包含讓使用者上傳測量後的實際面積，讓資料集的內容不斷增強與優化。

3.執行方法:

- 透過終端裝置的照片上傳雲端。
- 利用 YOLOV4 進行水果識別。
- 透過訓練好的模型以及終端裝置取得的數據，預測水果大小與範圍。
- 計算出水果大小面積。

4.系統流程:



我們的 AI 模型架構，主要先從圖片預測出水果的種類，並匡列出產品在圖片中的範圍，每個被匡選的範圍中顯示水果的種類，以及該種類的預測的面積大小。在我們的系統中，一張圖片水果被識別出並匡選，並回傳水果的種類與預測的面積大小。