

深度學習之影像分析甜椒

指導老師：潘國興老師

組員：機械三丁 U0911315 謝立峻、機械三丁 U0911317 廖育宸

機械三丁 U0911312 蕭榮彬、機械三丁 U0911333 廖彥儒

摘要

因為現代化的農業逐漸成為主流，而搭載人工智慧的自動採收機也逐漸開始嶄露頭角，因此我們嘗試以YOLOv5來訓練人工智慧的判斷農作物影像的能力，並希望能夠讓人工智慧擁有自主判斷甜椒的位置、成熟度、生長狀況等數據的能力，並在未來能夠利用此技術，提高甜椒或是其他作物的品質與產量。

關鍵詞:YOLOv5、深度學習、甜椒辨識

1. 簡介

在現今的科技產品中，使用人工智慧已經不是非常稀有的事情，大多數的新型產品中都已經有人工智慧在進行輔助工作與運算，像是特斯拉的自動駕駛、GOOGLE的語音助理、掃地機器人的路線規劃等等，都是有運用到人工智慧的項目。因此利用人工智慧所檢測農作物，並藉由搭載人工智慧的機械手臂進行採收就可以完成自動化農業的產線，並且可以顯著的降低農場栽培作物所需花費的人力與時間，並提升農作物的品質與利潤。

而對於甜椒此種常見的經濟作物來說，確保作物的品質與產量是一件非常重要的事情，在消費者的挑選下，只有長相與味道良好的甜椒才會是有價值的商品，所以藉由訓練好的人工智慧有規則的施肥、噴藥與灑水下，就能夠產出品質與產量兼具的良好作物。

但現今主要的問題則是出在人工智慧的訓練上，藉由機械學習法所花費的時間與物力實在還是太多，且長期訓練的人工智慧也不能運用在其他的物體上，導致大部分的農場還是使用大量人力的傳統採收法。

對此我們希望能夠以深度學習的方法，藉由讓人智慧自動學習的方法來提升訓練的速度，並降低訓練人工智慧成本，藉此使現代化農業的普及率能得到提升。

2. 研究方法

2.1 資料收集

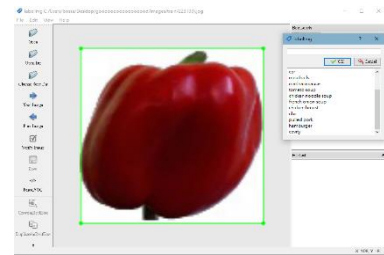
本次專題中，我們打算使用YOLOv5來進行深度學習，讓電腦能夠分辨甜椒的成熟程度，此研究的主要重點為數據集的製作，為了讓電腦能夠正確的分辨甜椒，而為此我們就則需要先創建一個以甜椒為主的資料庫。

首先我們需要蒐集甜椒的照片，於是我們先去了新竹關西的金勇DIY農場進行拍攝，以手機拍攝甜椒的照片，並且將照片裁切成不同大小來做為訓練集的素材，且為了提升模型的辨識能力，我們將照片裁剪完後如圖一，將照片分成了正常、反轉與夜晚濾鏡等，以此增加訓練的樣本數，也讓程式能夠進行夜晚照片的訓練。



圖一 不同狀態的甜椒

全部調整完後就如圖二，使用labelImg將照片進行標記，並分類成train、test、val檔，總和約為4.836張的照片作為訓練的資料，然後使用YOLOv5獲得訓練的結果後再以數據評估訓練模型是否符合預期成果，而後以該結果再設定重新調整的訓練模型，以此獲得提升模型的精準度。



圖二 labelImg甜椒的方法

2.2 YOLOv5的架構

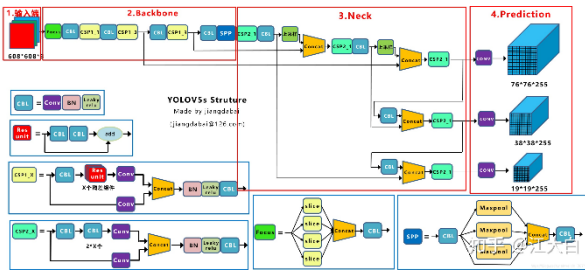
YOLOv5s與YOLOv5m的架構如圖三、圖四，總共分為四個部分:1.輸入端：Mosaic資料增強、自適應錨框計算。2.Backbone：Focus結構，CSP結構。3.Neck：FPN+PAN結構。4.Prediction：GIOU_Loss。裡面的基本結構。

YOLOv5的主幹網絡內容基本一樣，只在深度和寬度上不同，如圖三左下角:

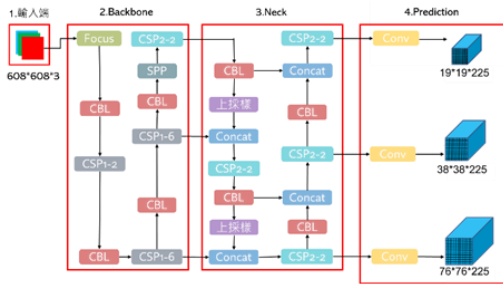
1.Focus結構:切片操作，將608×608×3的三通圖像經過切片處理，變成304×304×12的特徵圖，再經過使用32個捲積核的捲積操作，最終變成304×304×32的特徵圖。

2.Res unit：參考Resnet網路中的殘差結構，讓網路可以建構的更深。

3.CSP結構:CSP1_X結構在Backbone主幹網路中，另一種CSP2_X結構在Neck中CSP模塊中的捲積核大小都是3×3，步進值為2，假設輸入圖像尺寸是608×608，那麼特徵圖變化規律是：608×608→304×304→152×152→76×76→38×38→19×19，最終得到了19×19大小的特徵圖。



圖三 YOLOv5s的網路架構[1]



圖四 YOLOv5m的網路架構[2]

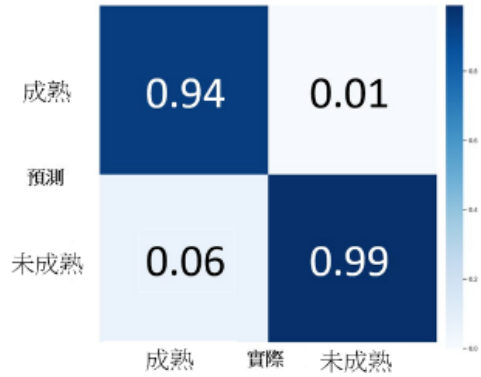
3. 結果與討論

將數據庫以Yolov5訓練後獲得的數據如圖五

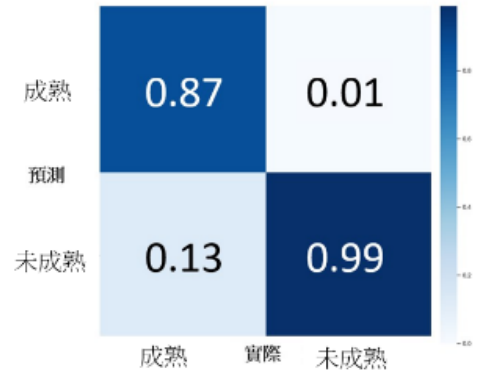
模型	圖片尺寸	Batch	層數	mAP-50	mAP 0.5:0.95	Time(h:m)
YOLOv5 m	270×270	16	200	0.983	0.909	01:35
YOLOv5 s	270×270	16	200	0.982	0.901	01:30

圖五 YOLOv5的訓練數據

圖六、圖七為YOLOv5訓練出來的混淆矩陣，橫向為實際值，縱向為預測值，若預測與實際相似度越高顏色越深，代表準確率越高，而根據圖形可分成四類，並藉由四格的配比讓我們得知電腦分類圖片的情況，就圖六而言，判斷成功的成熟有94%而判斷失敗的成熟則有6%，以此我們就能夠了解模型能否正常的判斷未成熟與成熟的甜椒，而圖六在成熟甜椒的精準度判斷上比圖七還要精準，這是因為圖六使用的模型相比於與圖七的模型、雖然使用的資源較多、但因為模型較大的關係所以，所以精準度也會比模型較小得圖七來的高。



圖六 YOLOv5m 混淆矩陣



圖七 YOLOv5s 混淆矩陣

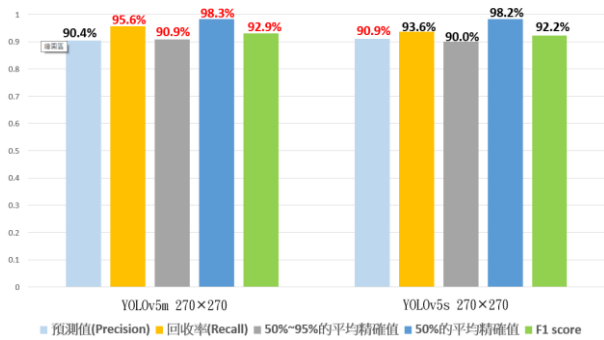
而藉由這些數據，我們可以分別計算出圖八的Precision精確率(1)、Recall召回率(2)、F1-score(3)，然後藉由這些數據來判斷這次模型的好壞。而後由再藉由混淆矩陣，我們即可得出，真陽性(TP)、偽陽性(FP)、偽陰性(FN)、真陰性(TN)這四類數據，然後再藉由這些數據就可以計算我們的Precision預測值(1)、Recall回收率(2)、然後再利用預測值與回收率來計算F1-Score(3)，得到圖八之後並以此來評斷這次模型的訓練成果，我們可以藉由預測值與回收率來觀察圖片的回收狀況，而平均精確值則能夠讓我們了解模型判斷的成功率與精確值是多少，而F1-score則是能夠同時比較預測值與回收率，使我們能夠知道模型的綜合評價，而由圖八則可以得知，YOLOv5m的的回收率相比於YOLOv5s來的更為優秀，這表示在輸出的圖片中YOLOv5m回收的圖片比YOLOv5s來的多，且相對較高的精確值與F1-score也代表YOLOv5m能比YOLOv5s判斷更多的未知的圖片，因此可以得知，YOLOv5m的訓練結果比YOLOv5s的訓練結果來的優秀。

	實際	肯定	否定
預測			
肯定	TP	FP	
否定	FN	TN	

$$precision = \frac{TP}{TP + FP} \tag{1}$$

$$recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1-score = 2 \times \frac{precision \times recall}{precision + recall} \quad (3)$$



圖八 綜合數據圖

由圖九中可以看出，我們的模型已經具有基礎的辨識能力，但對於有裂痕或是在葉子中的未成熟甜椒來說還是有些費力，如果重新調整訓練的圖片或是增加訓練圖片的總量，或許能夠讓準確率獲得進一步的提升。從圖九中可以看出我們的模型已經能夠辨識出甜椒的成熟狀態了，但還是能夠看出一部分紫色甜椒的精準度還是偏低，而主要的原因則是有夜晚濾鏡的紫色甜椒會比較容易導致誤判。



圖九 YOLOv5 的訓練結果

4. 結論

這次我們使用YOLOv5的深度學習模型分類甜椒的成熟度，並加入了反轉或是夜晚濾鏡的圖片來增

加訓練難度與樣本多樣性，而從我們訓練出來的模型來看，精確率都能夠達到90%以上，而在實際的偵測中也能夠達到94%，由此可看出我們所訓練出來的模型已有了能夠判斷甜椒成熟度的能力，但對於紫色甜椒的辨識還是有一些問題，因為我們添加了夜晚的濾鏡，因此顏色較暗的紫色甜椒會比其他甜椒來得不明顯，而導致電腦難以判斷是甜椒還是背景，而要解決這項問題，就需要有更多相同的樣本，或是實際在晚上拍攝照片才可以改善。

而以目前研究結果而言，我們的模型已經能夠輔助農民分辨甜椒，但希望未來能夠將再將影像辨識的技術套用到機械手臂或是無人機上，藉由機械來輔助農民管理農地，以此讓現代化農業能夠得到普及，就能夠有效降低農場所需的勞動人口，並更妥善的利用所有的土地。

5. 誌謝

在此感謝潘國興老師以及學長指導，藉由老師初步的帶領，我們才能漸漸的了解深度學習的模型與種類，也感謝學長能對我們所選用的模型提出建議，使我們能夠更快的做出我們想要成果，這裡也要感謝金勇農場與老師能夠提供我們照片，使我們能夠減少蒐集甜椒圖片的難度，最後感謝這一年老師的耐心指導我們才能完成這次的專題。

6. 參考文獻

1. 黃寶興，張劍平，基於YOLOv5演算法進行衛星影像目標檢測之研究，學術論文，資訊工程系資訊科技應用碩士在職專班，110
2. Y. Luo, Y. Zhang, X. Sun, H. Dai, X. Chen, Intelligent Solutions in Chest Abnormality Detection Based on YOLOv5 and ResNet50, Journal of Healthcare Engineering 2021, 2267635 (2021)