

逢甲大學學生報告 ePaper

食品瑕疵檢測

Food Defect Detection

作者：吳佖諒

系級：電子四甲

學號：D0846959

開課老師：林峰正,蔡明翰,梁詩婷,王通溫

課程名稱：智慧辨識檢測與應用

開課系所：資電學院綜合班

開課學年：111 學年度 第 1 學期



## 一、中文摘要

精緻農產品及智慧化生產為台灣農業及工業重點發展對象，透過自動化流水線，結合深度學習及計算機視覺。本專題開發出一套適用於工廠內的自動化胡蘿蔔辨識系統，具體來說我們模擬出一個食品業實際情況，也就是隨機放置及不同光線下的狀態。在這種極具挑戰的前提下，我們完成了精準辨識出胡蘿蔔位置及範圍，透過語義分割及物件偵測。這項工作中考慮的農產品是胡蘿蔔，胡蘿蔔為許多國家大量生產的農產品之一。透過徹底的實驗及分析，我們將對 20 條的胡蘿蔔進行檢測評估，總共蒐集到 3120 張照片，其語意分割結果平均 wIoU 為 0.9899，物件偵測結果平均 F1-score 為 0.9549。以其穩定及精確度達到足以應付流水線上自動化生產的標準。預計這項研究可以統合到自動化流水線分級分類系統中，可以協助未來自動化農業的發展。

**關鍵字：**胡蘿蔔、檢測、工業 4.0、深度學習



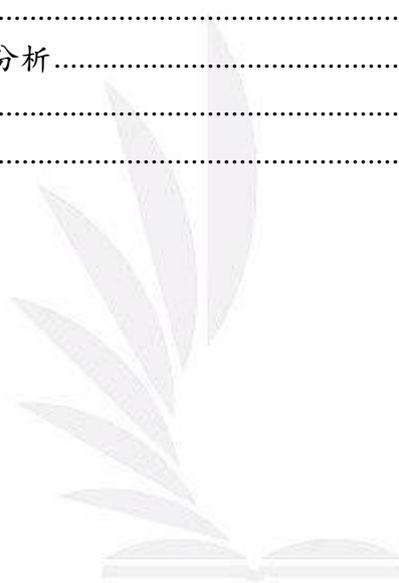
## 二、 Abstract

Delicate agricultural products and intelligent production are the key development targets of Taiwan's agriculture and industry. Through automated assembly lines, combined with deep learning and computer vision. This topic develops a set of automatic carrot identification system suitable for factories. Specifically, we simulate an actual situation in the food industry, that is, the state of random placement and different light conditions. Under this extremely challenging premise, we have completed the precise identification of the location and range of carrots, through semantic segmentation and object detection. The agricultural product considered in this work is carrot, which is one of the agricultural products produced in large quantities in many countries. Through thorough experiments and analysis, we will conduct detection and evaluation on 20 carrots. A total of 3120 photos have been collected. The average wIoU of semantic segmentation results is 0.9899, and the average F1-score of object detection results is 0.9549. With its stability and accuracy, it can meet the standard of automatic production on the assembly line. It is expected that this research can be integrated into the automated assembly line classification system, which can assist the development of automated agriculture in the future.

Keyword : carrot, Industry4.0, inspection, deep learning

## 目 次

一、	中文摘要.....	1
二、	Abstract.....	2
三、	研究動機.....	4
四、	文獻回顧及探討.....	4
五、	研究方法及步驟.....	5
	機台架設及控制.....	5
	相機擷取影像.....	6
	影像標記.....	7
	模型訓練.....	8
	語義分割.....	9
	物件偵測.....	9
六、	實驗設置.....	12
七、	實驗結果與分析.....	12
八、	結論.....	14
九、	參考文獻.....	15



### 三、研究動機

由於台灣正在發展精緻農產品的階段，其中三大主軸之一便是卓越農業(陳祈睿, 2011)，其中智慧財產加值運用提到「強化農業科技研發成果管理制度，建構產學研合作環境及配套措施，進行農業科技研發成果盤點分析與專利佈局」。在 agricultural marketing resource center(AgMRC) (Carrots, 2021)中有研究調查新鮮的胡蘿蔔價格高於加工過後的胡蘿蔔，所以胡蘿蔔是一種需時間賽跑的蔬果，且在 Food and Agriculture Organization of the United Nation(FAO) (Nations, 2022)調查指出，胡蘿蔔總產量從 2000 年的 2200 萬噸逐步提升到 2021 年的 4100 萬噸，如圖 1。到目前傳統蔬果分級主要還是依靠人力，於是本專題便想利用影像辨識系統來辨識出農產品，並結合滾輪機械來構成自動化流水線系統，在減少人力需求的同時並且加速蔬果從田地到消費者手中的時間，保持胡蘿蔔的新鮮程度。

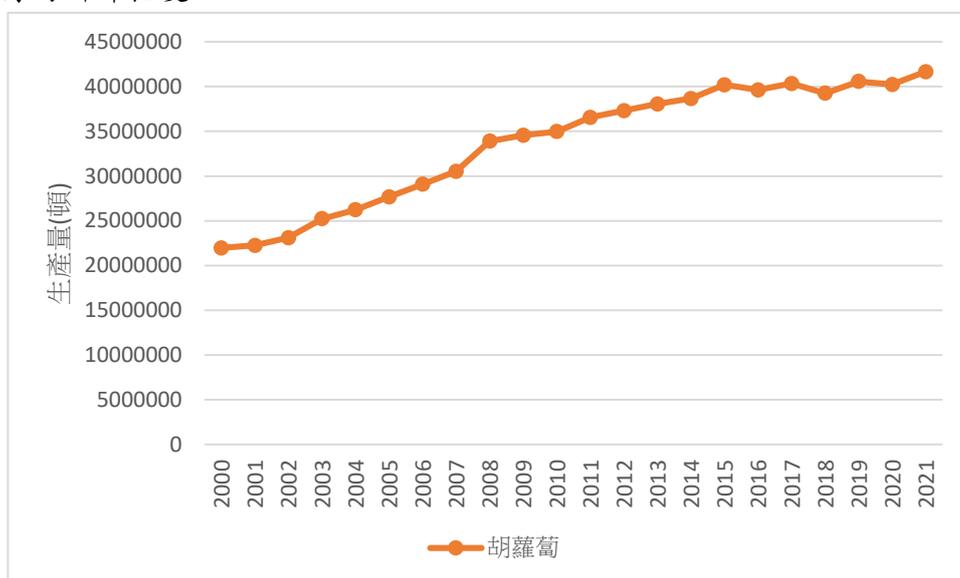


圖 1 FAO 統計胡蘿蔔年產量折線圖

### 四、文獻回顧及探討

至今為止，深度學習及計算機視覺已經普遍被利用於農產品預測及檢測系統中。其中包括品級分類，體積預測等。

最新型的研究 (Örnek, 2021)收集了 500 美金的胡蘿蔔，並以 7:2:1 分成訓練、驗證、測試三個部分，在研究中利用七層的 Multilayer Perceptron (MLP) 網路及四層的 Long Short-Term Memory (LSTM) 網路，其結果用三種指標來衡量，即均方根誤差(RMSE)、平均絕對誤差(MAE)、和判定係數( $R^2$ )，總結其結果，這兩個網路是有機會被採用的，但此方法沒有利用計算機視覺，皆透過人工量測胡蘿蔔特定位置的直徑，還需人力介入，無法達到流水線自動化生產。

## 五、研究方法及步驟

本專題將透過計算機視覺系統來辨識出胡蘿蔔位置及圖像訊息，用以蒐集胡蘿蔔形狀資訊以利後續研究做觀察。我們將會使用的 MATLAB 軟體以及 arduino，撰寫一套基於模擬工廠流水線實際環境下的計算機視覺辨識系統，如圖 2。



圖 2 實驗所需軟硬體設施

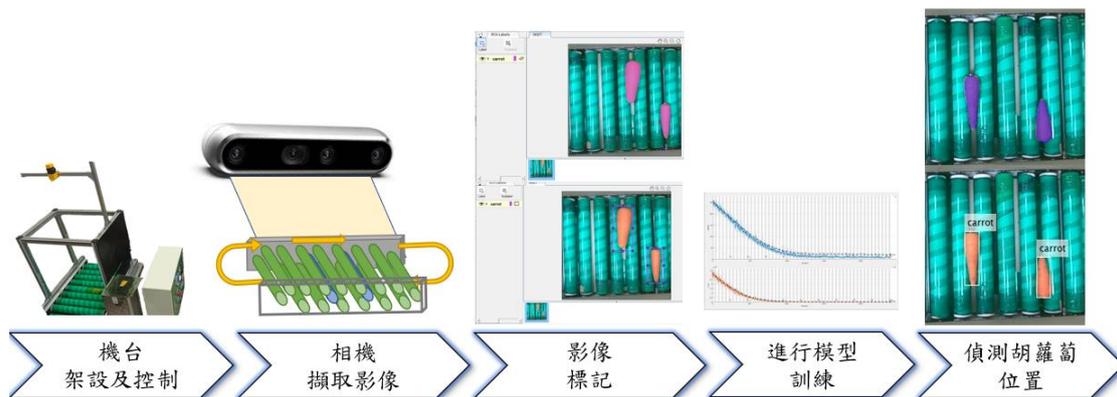


圖 3 流程圖

本專題整套流程如圖 3 中所顯示，總共分為五個步驟，分別為機台架設及控制、相機擷取影像、影像標記、模型訓練、偵測胡蘿蔔位置。

以下將介紹各步驟詳細內容。

### 機台架設及控制

本專題我們透過滾輪來模擬工廠中實際的生產線，讓目標物體沿著滾動，如圖 4。

簡單來說，滾輪輸送帶是由主控台進行操作，可以進行運轉、停止、轉速等設定。輸送帶是由 18 根等間距的滾輪所組成，分為兩個種模式，前進、定點旋轉，如圖 5。具體來說，前進模式，就是滾輪透過平行逆時針轉動，將物體向前輸送。定點旋轉模式，透過抬升下方兩個滾輪將其接觸到的四根滾輪進行向外旋轉，至此就可以達到定點旋轉。

我們透過安裝相機在滾輪上方，從上至下的角度進行拍攝，安裝位置距離滾輪上方 80cm 確保鏡頭能完整拍攝到整個滾輪範圍。

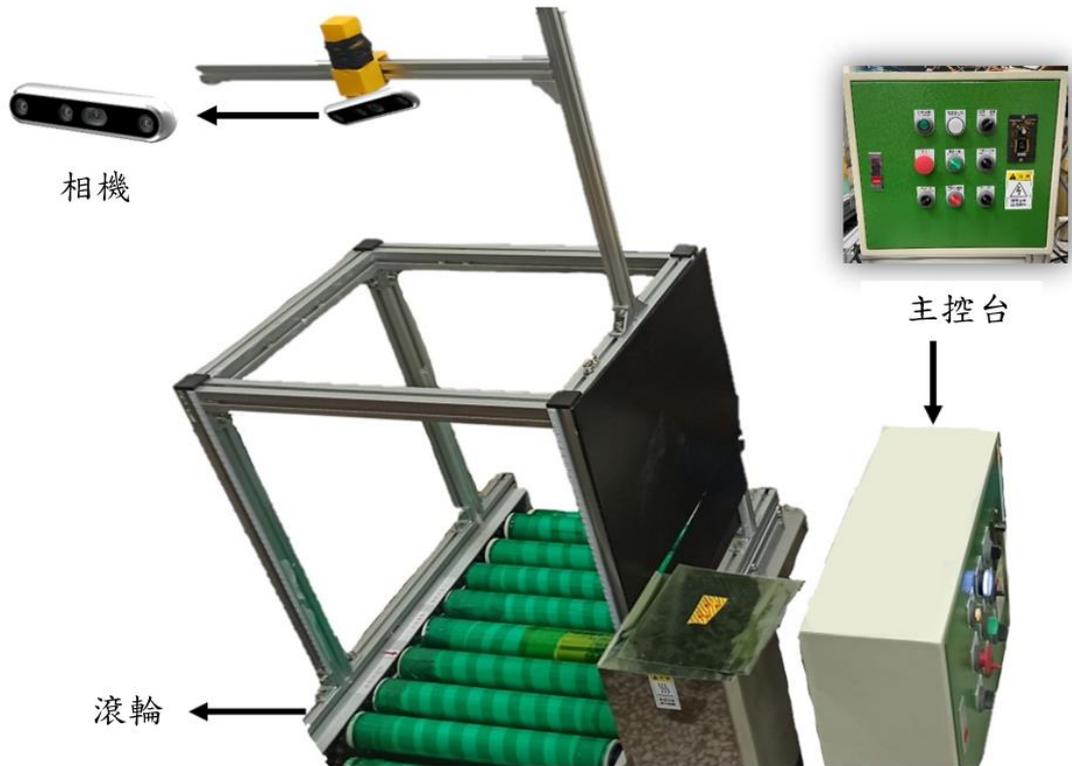


圖 4 由滾筒輸送帶、主控台、深度攝像頭組成的硬件設備設置

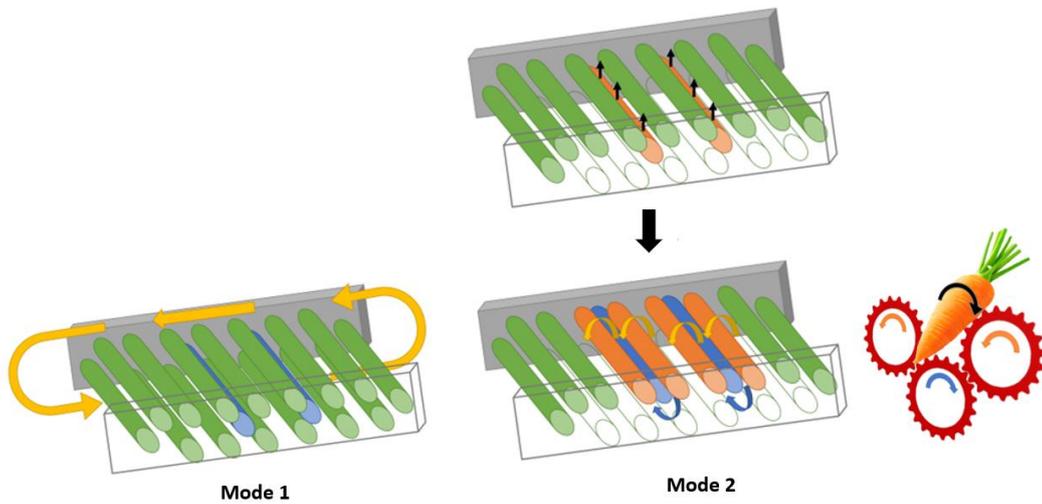


圖 5 滾輪兩種操作模式，前進、定點旋轉

## 相機擷取影像

為了模擬工廠內真實情況，胡蘿蔔將被隨機擺放在滾輪上，以模擬真實工廠內胡蘿蔔被放置情況，這邊我們蒐集了不同光線下的資料，分別是開燈與關燈，具體來說，本專題事先準備了 20 條胡蘿蔔做為檢測對象，胡蘿蔔將會由右至左經過鏡頭視角，由於胡蘿蔔具有不規則形狀，有部分胡蘿蔔在轉動時會滾

動橫置滾輪上方，並非原本垂直於滾輪凹槽中。最終我們蒐集了 10 隻影片總共 3120 張，範例如圖 6。

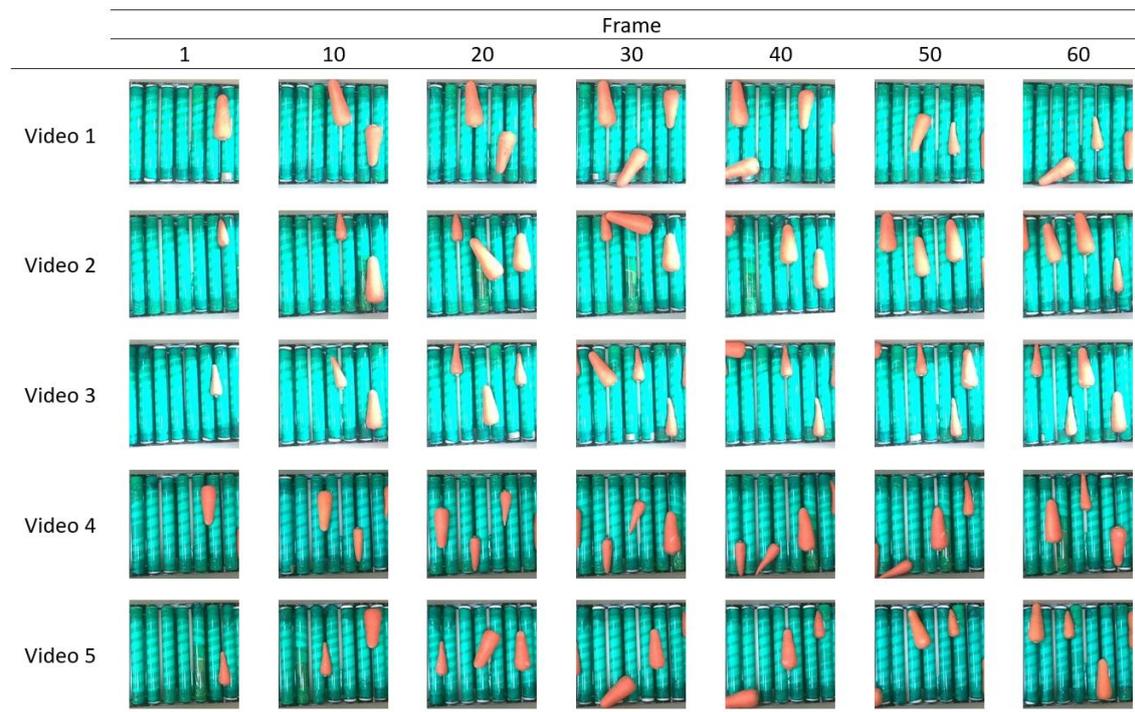


圖 6 在不同光單條件下蒐集的胡蘿蔔範例

## 影像標記

我們透過相機擷取影像中，蒐集到的 3120 張影像透過 MATLAB 中的 Image Labeler 進行影像標記，由於我們需要進行語義分割及物件偵測，所以我們需要標記兩種資料檔，分別為方框標記，如圖 7、及像素標記，如圖 8。

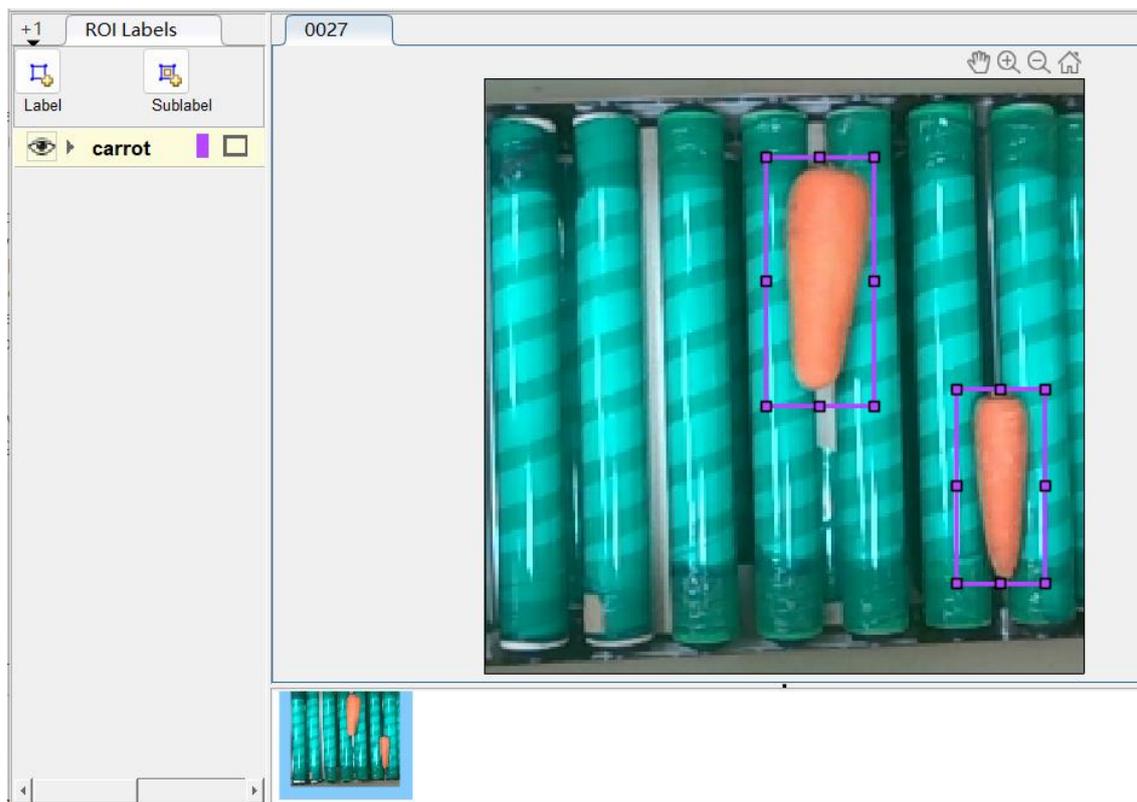


圖 7 圖中為 Image Labeler 進行方框標記

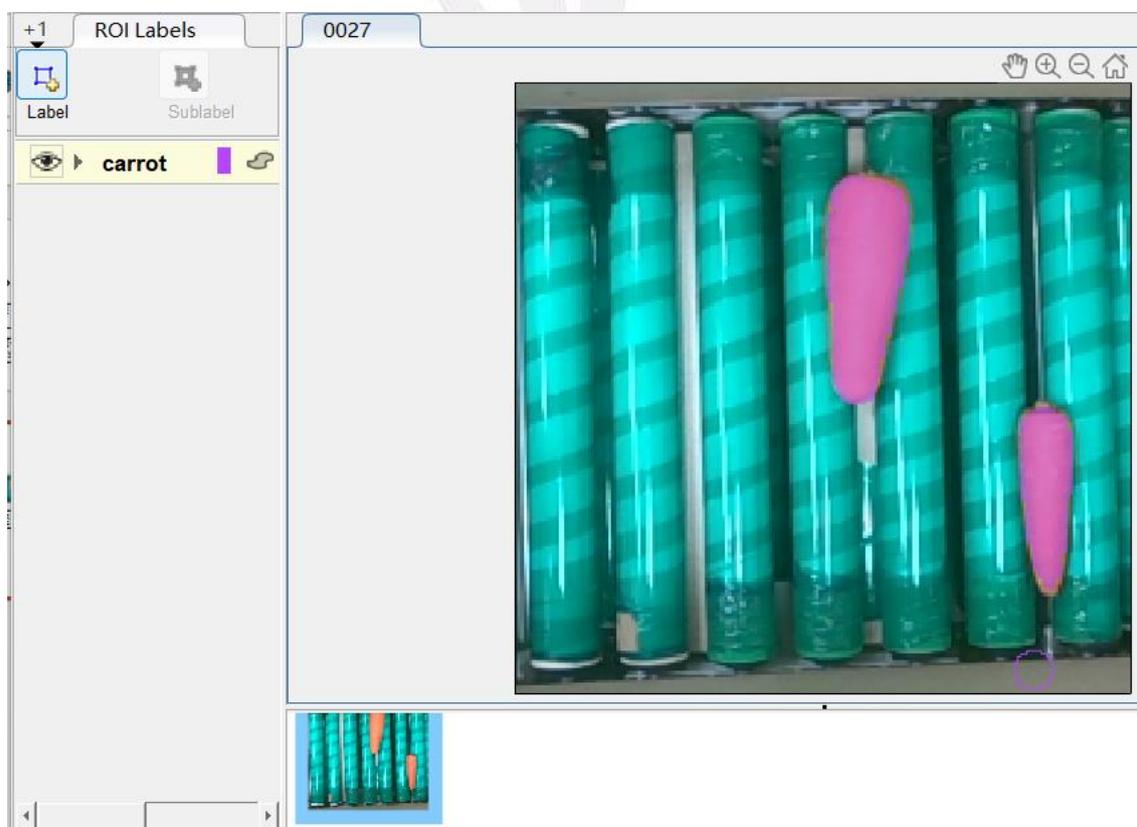


圖 8 圖中為 Image Labeler 進行像素標記

## 模型訓練

本章節將介紹模型訓練，本階段將影像標記中標記好的影像進行模型訓練，也就是語義分割及物件偵測，及其使用模型之介紹。

## 語義分割

這個過程是將影像中的每一點像素都歸類進相應的類別，這邊有兩個類別：背景及胡蘿蔔。

根據 (Gan, 2022)中所採用的算法是 DeepLabv3+ (Chen L.-C. a., 2018) ，如圖 9。簡而言之，該架構採用 Atrous Spatial Pyramid Pooling (ASPP)來編譯多尺度的信息。此外，這個架構網路中還採用了一個 encoder-decoder，用以恢復空間信息，目的是為了更精確的捕捉到物體邊界。

編碼器的結構參考了 DeepLabv3 (Chen L.-C. a., 2017)中的結構。

這邊我們所採用的主幹結構是 ResNet-18 (He, 2016)，因此所有圖像被先調整為  $211 \times 211 \times 3$ 。

## 物件偵測

儘管語義分割能區分像素位置，但很多情況是胡蘿蔔脫離了滾輪凹槽，碰巧與另一個胡蘿蔔相撞。

這就導致了無法區分兩個胡蘿蔔的問題，因為相同種類的像素是相連的。因此我們利用物件偵測來克服這個問題。如圖 10 中描繪了像素劃分區域(紫色)，這邊利用物件偵測來區分個別胡蘿蔔。這個過程的目標是為了找到胡蘿蔔的空間位置，並使用邊界框來進行分類。我們採用 YOLOv4-tiny (Jiang, 2020)結構，他是 YOLOv4 (Bochkovskiy, 2020)的簡化版本，如圖 11 所示。

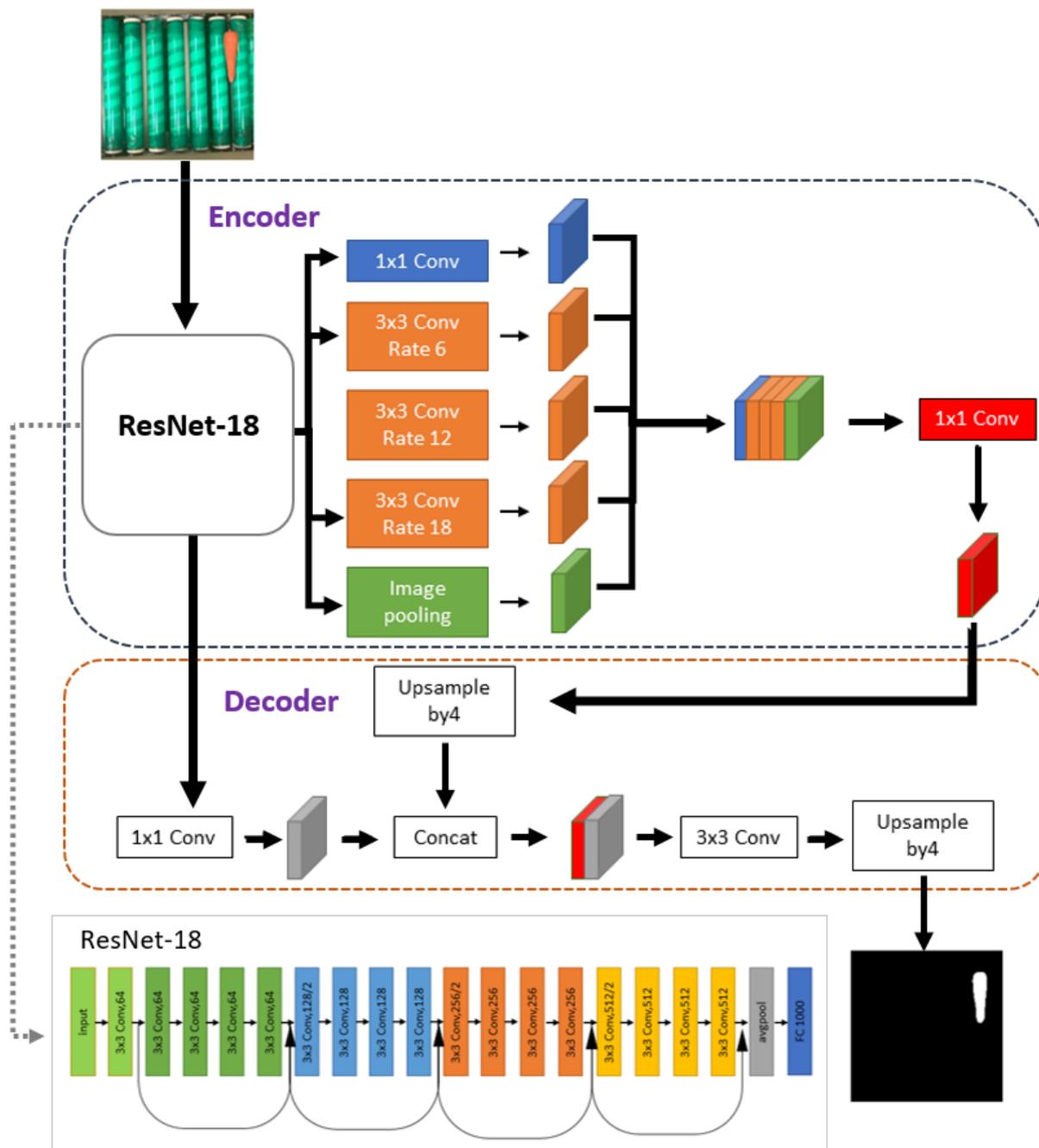


圖 9 語義分割中 DeepLabv3+架構

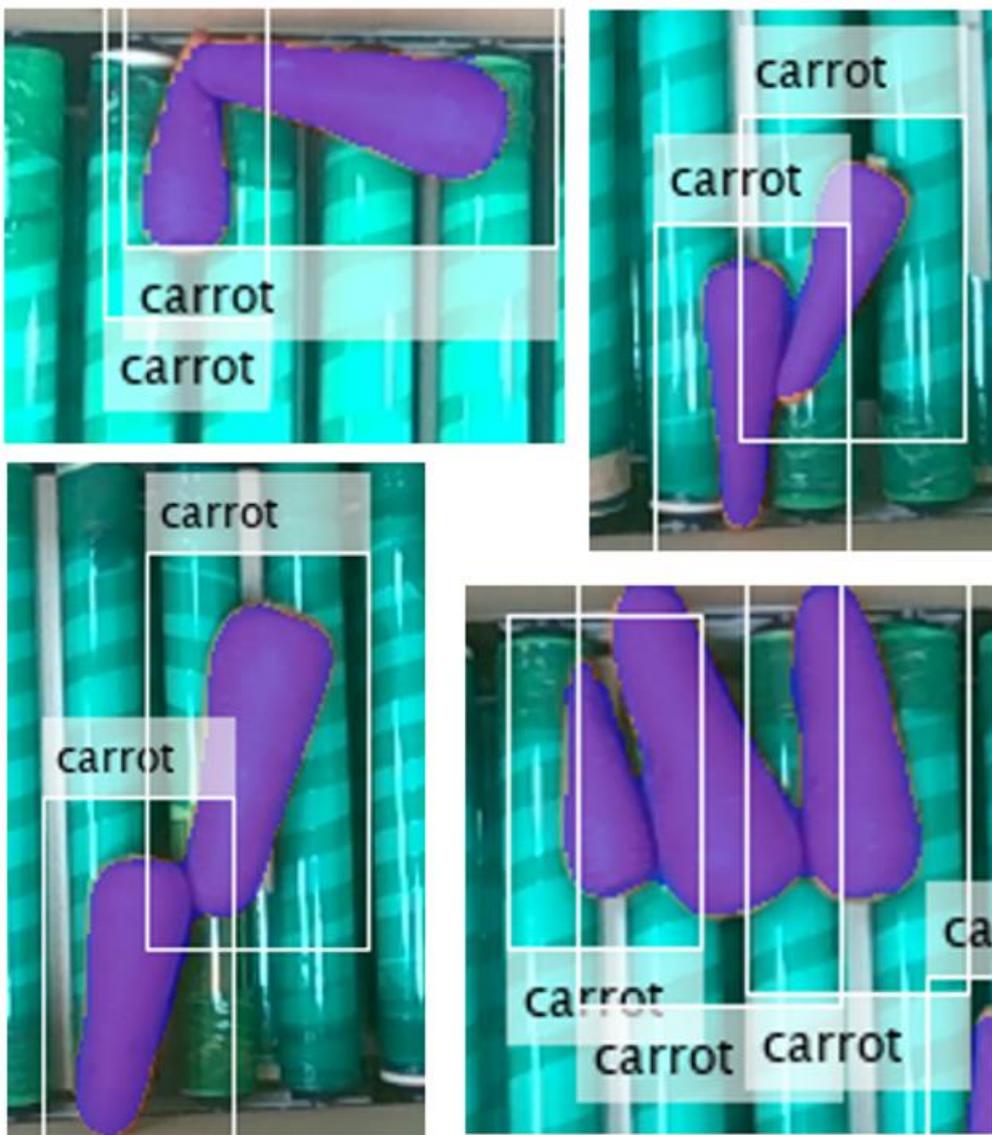


圖 10 由於使用語義分割時無法正確的識別某些胡蘿蔔，引此採用物件偵測的方法。

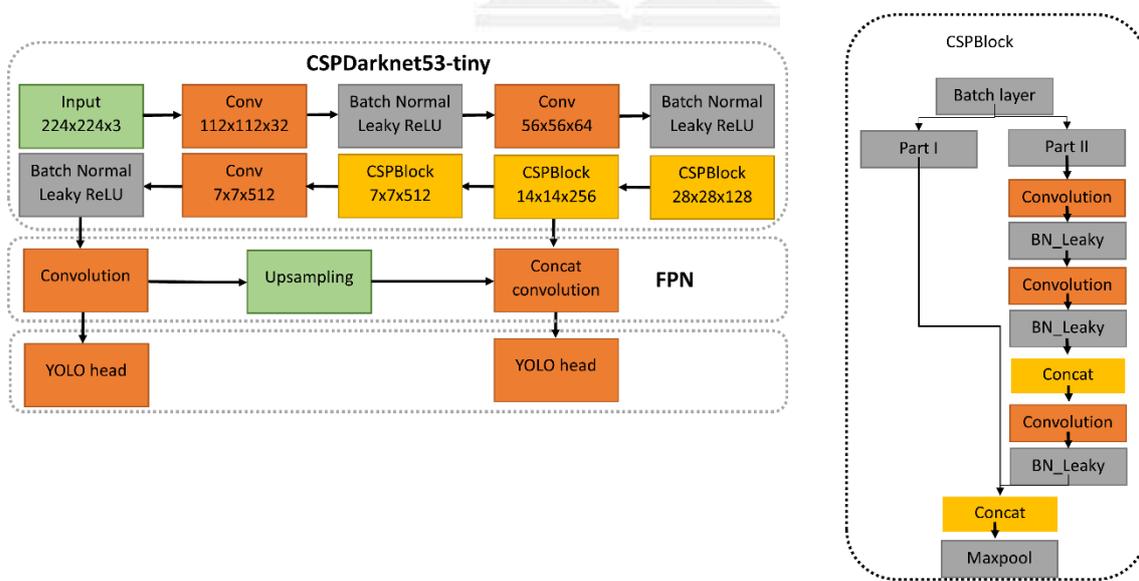


圖 11 物件偵測中 YOLOv4-tiny 架構

## 六、實驗設置

這邊所有模型訓練皆在 MATLAB 2022a、Intel Core i5-11400、32GB 內存和 NVIDIA GeForce RTX 3060、Window10 系統下進行。

訓練設置

Parameter	Semantic Segmentation	Object Detection
	DeepLabv3+(ResNet18)	YOLOv4-tiny(CSPDarknet53)
Input size	211×211×3	224×224×3
Optimizer	Adam	Adam
Epoch	30	100
Learning rate	0.00001	0.00001
Mini batch size	50	50
Anchor size	-	[51,78;82,48;86,35] [80,29;36,45;65,22]

表 1 訓練配置

所有模型都是使用預訓練模型，如 ResNet-18 和 CSPDarknet53 都是先從 ImageNet 數據集中學習特徵，所以每種訓練模式 epoch 都比較低(<100)，對於優化器及學習率分別固定為 Adam 和 0.00001。這邊為了確保公平及穩定性，採用交叉驗證，以 90% 作為訓練集 10% 作為測試集，重複 10 次，所有資料都會被測試過。

## 七、實驗結果與分析

透過訓練好的模型進行預測，如圖 12。其語義分割結果，如表 2。物件偵測結果，如表 3。可以看到語義分割結果中 IoU 平均為 0.9335，全體高於 0.9 表示模型預測已經相當準確。另一方面，物件偵測中 F1-score 平均為 0.9549，全體高於 0.9，表示各種情況下都能準確偵測胡蘿蔔位置。

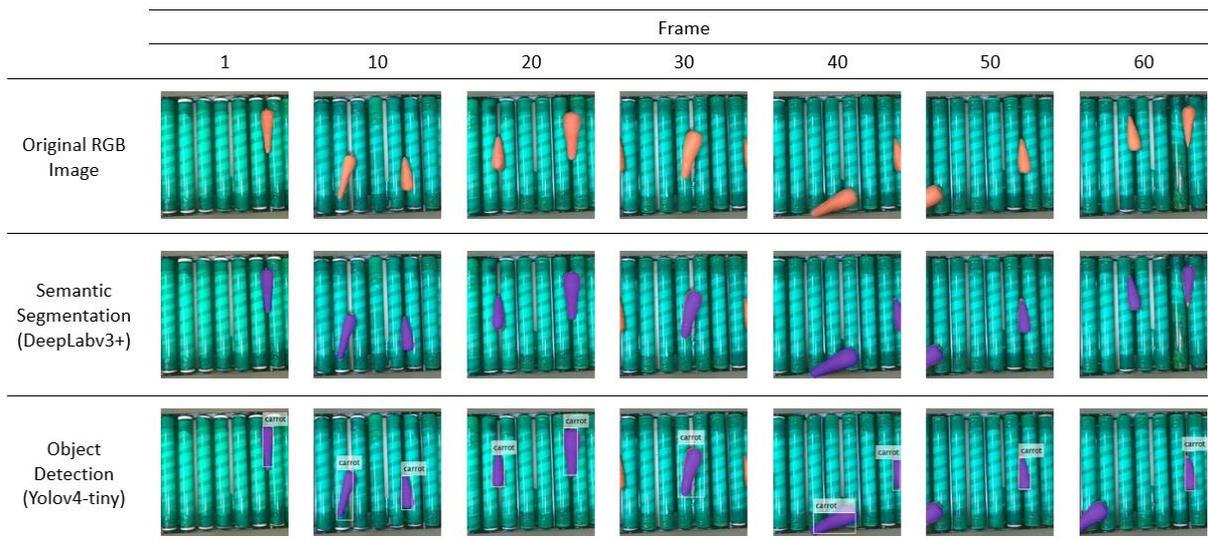


圖 12 模型預測結果範例

Video	$ACC_{carr}$	$IoU_{carr}$	$DSC_{carr}$	$gAcc$	$mAcc$	$mIoU$	$wIoU$
1	0.9373	0.9102	0.9530	0.9921	0.9673	0.9508	0.9844
2	0.9315	0.9178	0.9503	0.9946	0.9652	0.9560	0.9893
3	0.9551	0.9178	0.9475	0.9938	0.9760	0.9556	0.9878
4	0.9711	0.9403	0.9701	0.9950	0.9841	0.9674	0.9901
5	0.9673	0.9422	0.9673	0.9956	0.9826	0.9687	0.9913
6	0.9791	0.9454	0.9526	0.9952	0.9879	0.9701	0.9907
7	0.9668	0.9406	0.9656	0.9939	0.9818	0.9669	0.9880
8	0.9529	0.9325	0.9558	0.9947	0.9755	0.9634	0.9895
9	0.9682	0.9478	0.9728	0.9970	0.9835	0.9723	0.9941
10	0.9681	0.9402	0.9671	0.9967	0.9832	0.9684	0.9935
Average	0.9597	0.9335	0.9602	0.9949	0.9787	0.9640	0.9899
Minimum	0.9315	0.9102	0.9475	0.9921	0.9652	0.9508	0.9844
Maximum	0.9791	0.9478	0.9728	0.9970	0.9879	0.9723	0.9941

表 2 語義分割交叉驗證評估指標

Video	Precision	Recall	F1-score
1	0.9880	0.8853	0.9338
2	0.9923	0.9489	0.9701
3	0.9710	0.9445	0.9576
4	0.9698	0.9406	0.9550
5	0.9826	0.9601	0.9712
6	0.9777	0.9286	0.9525
7	0.9545	0.8825	0.9171
8	0.9539	0.9200	0.9366
9	0.9826	0.9579	0.9701
10	0.9899	0.9799	0.9849
Average	0.9762	0.9348	0.9549
Minimum	0.9539	0.8825	0.9171
Maximum	0.9923	0.9799	0.9849

表 3 物件偵測交叉驗證評估指標

## 八、結論

透過我們所蒐集到的 10 隻影片，我們對其分別進行了交叉驗證，在語義分割中我們得到了平均 0.9335 的 IoU，物件偵測部分也得到了 0.9549 的 F1-score。證明了透過 DeepLabv3+ 及 YOLOv4-tiny 是可以準確辨識出胡蘿蔔位置及點位。

期盼未來可以透過影像辨識蒐集到的資訊做更深度的研究。

## 九、參考文獻

- Bochkovskiy and Wang, Chien-Yao and Liao, Hong-Yuan Mark Alexey. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.
- Carrots. (2021 年 8 月). 擷取自 AgMRC:  
<https://www.agmrc.org/commodities-products/vegetables/carrots>
- Chen and Papandreou, George and Schroff, Florian and Adam, Hartwig Liang-Chieh. (2017). Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.
- Chen and Zhu, Yukun and Papandreou, George and Schroff, Florian and Adam, Hartwig Liang-Chieh. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the European conference on computer vision (ECCV), 801--818.
- Gan and Luo, Shi-Hao and Li, Chih-Hsueh and Chung, Shih-Wei and Liang, Sze-Teng and Tan, Lit-Ken Yee-Siang. (2022). An automated cucumber inspection system based on neural network. Journal of Food Process Engineering, e14069.
- He and Zhang, Xiangyu and Ren, Shaoqing and Sun, Jian Kaiming. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 770--778.
- Jiang and Zhao, Liquan and Li, Shuaiyang and Jia, Yanfei Zicong. (2020). Real-time object detection method based on improved YOLOv4-tiny. arXiv preprint arXiv:2011.04244.
- Nations and Agriculture Organization of the United Food. (2022). Statistics.  
<http://www.fao.org/faostat>.
- Örnek N., & Kahramanlı Örnek, H.M. (2021). Developing a deep neural network model for predicting carrots volume. Journal of Food Measurement and Characterization, 3471--3479.
- 陳祈睿. (2011 年 3 月). 精緻農業健康卓越方案推動成果. 擷取自 行政院農業委員會: <https://www.coa.gov.tw/ws.php?id=23055>