

基於深度學習影像辨識與機械手臂實現智能物件取放任務

Object Pick-and-Place Tasks Using Robotic Arm and Deep Learning

大四 谷苡瑄 郭昕叡 國立中興大學電機工程學系



摘要

本專題旨在透過機械手臂結合深度學習影像辨識技術，實現夾取指定顏色的方塊並將其移動至指定目標位置上方進行放置。首先，我們利用固定在架子上的RAPOO C200攝影機拍攝影像，並將其回傳至電腦進行深度學習運算與影像辨識處理。接著，根據辨識結果，ESP32控制器將指揮機械手臂進行夾取操作，夾取的物體為三種顏色的方塊（紅色、綠色、藍色）。機械手臂將準確地將方塊移動至目標板上方，該目標板具有三種形狀（圓形、三角形、五邊形），並搭配編號1至5，共計15種組合的圖形紙板。

研究動機與目的

隨著工業4.0的推進，工業自動化和智能製造成為現代工廠的重要發展方向。六軸機械手臂作為工業自動化中的核心組成部分，憑藉其靈活性和高效性，已被廣泛應用於生產線上的各種操作。其中，智能物件取放成為本專題研究的重點。

本次專題研究目的在實現以下創新與突破：

1. 技術創新：

首次使用模組化機械手臂進行二次開發，在Windows系統下使用micropython編程開發測試。

2. 提高分揀效率及準確性：

影像端透過YOLOv8影像辨識技術的應用，實現快速準確的物體辨識。手臂端利用上位機調整手臂姿態，使其能運作順暢。

3. 系統整合：

透過通訊架構起影像識別和機械手臂控制之間的溝通橋樑，使系統能高效精確的實現夾取的動作。

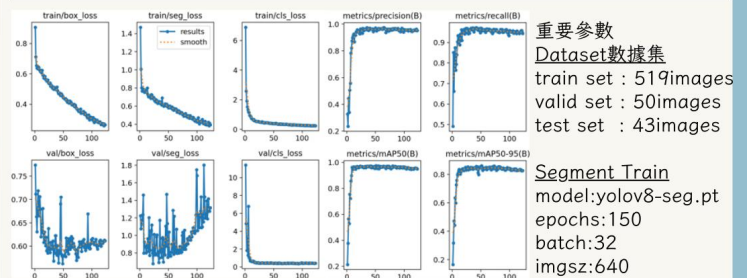
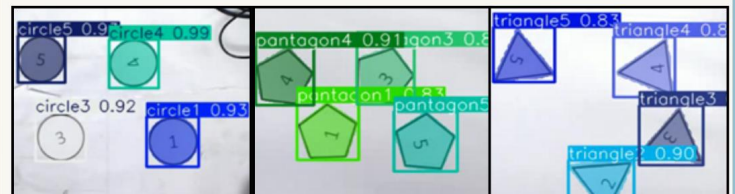
4. 優化程式碼：

模組化、精簡化程式碼，往後可應用於其他專案。

研究方法

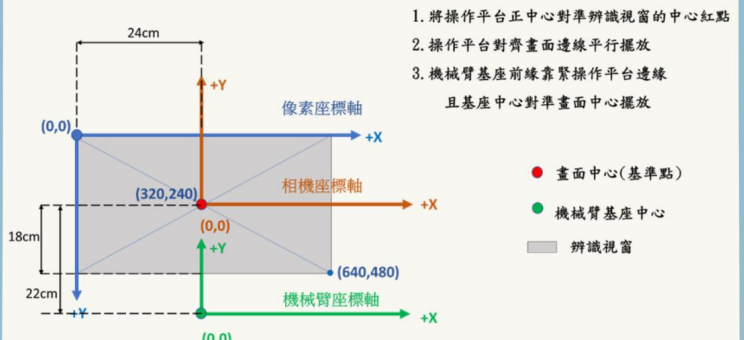
• 影像辨識

首先使用Labelme手動分割標記圖片，接著使用Roboflow進行圖片預處理和轉檔，最後使用YOLOv8訓練數據集得到最佳模型。



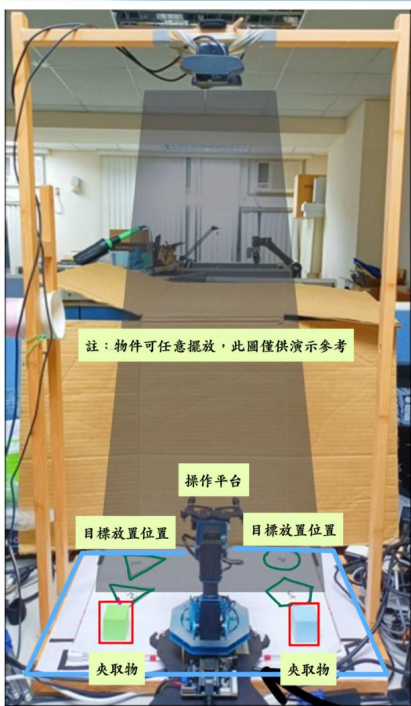
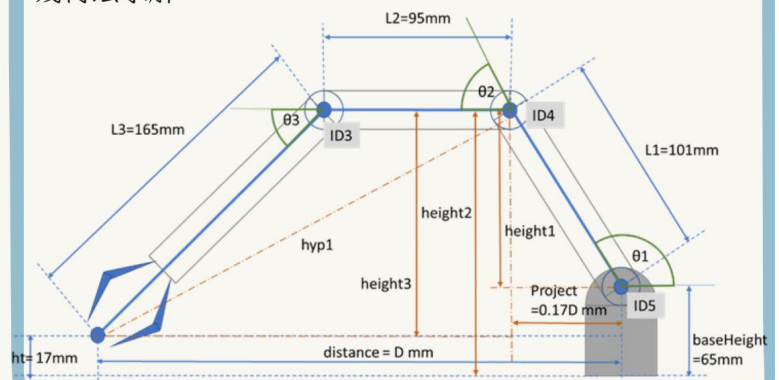
• 座標轉換(三維重建)

從二維座標系轉成三維座標系過程中的轉換順序:像素座標系→圖像座標系→相機座標系→機械手臂座標系。

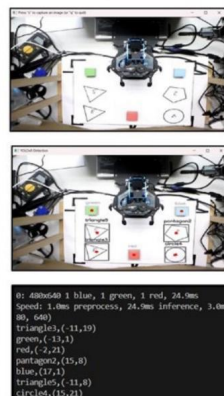


• 逆向運動學

給定機械手臂末端執行器（夾具）所需到達的空間位置，根據此座標推算各關節軸所需轉動的角度。此專題中採用幾何法求解。



拍攝工作區域的影像，並將影像傳回至電腦進行深度學習運算與顏色、圖形、數字辨識。



註：物件可任意擺放，此圖僅供演示參考

操作平台

目標放置位置

目標放置位置

夾取物

夾取物

0: Assistant 1 blue, 1 green, 1 red, 24.9ms
Speed: 1.0ms preprocess, 24.9ms inference, 3.0ms
ms, 640
triangle, (-11,19)
green, (-11,1)
red, (-2,21)
pentagon, (15,8)
blue, (17,1)
triangle, (-11,8)
circle, (15,21)

基於深度學習影像辨識與機械手臂實現智能物件取放任務

Object Pick-and-Place Tasks Using Robotic Arm and Deep Learning

大四 谷苡瑄 郭昕叡 國立中興大學電機工程學系



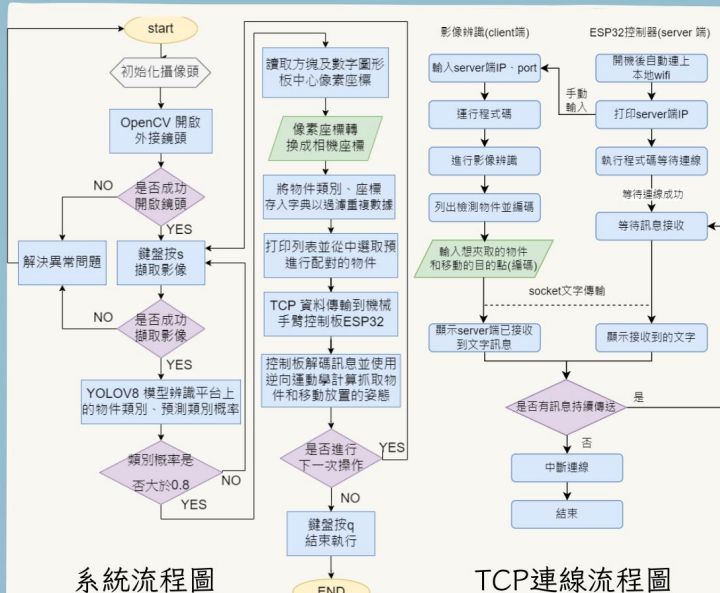
系統整合

• 影像辨識(Client端)運作流程

1. 檢測到不同類別的物件所包含的資訊(類別名稱&中心相機座標)進行編碼，並以列表形式打印出來。
2. 使用者輸入夾取物件的編碼以及目標位置的編碼，被選擇的資訊將會合併成一個字串。
3. 把字串轉換成手臂端可接收的格式，轉換後將再次將訊息合併成新字串進行資料傳輸。
4. 創建一個TCP client，設定server端的IP地址和端口，嘗試與server端連線，若成功連線，server端回傳成功連線訊息並可進行資料傳輸。
5. 接收並解碼server端成功接收訊息的回應(the server has received your msg{次數})。

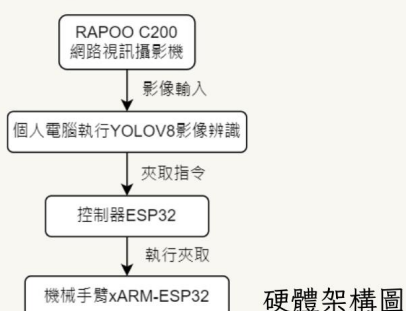
• Server端程式碼(ESP32開機後自動運行)

1. 連上無線網路
2. 創建TCP server，綁定IP位址和端口後開始監聽
3. Server端回應表示同意建立連線並確認收到client端請求
4. 接收client端的目標物件資訊(包含物件類別以及中心點相機座標)並進行解碼，若訊息傳送結束則斷開連線。
5. 處理接收到的訊息，將有用訊息進行分割提取後輸入手臂控制函式。



系統流程圖

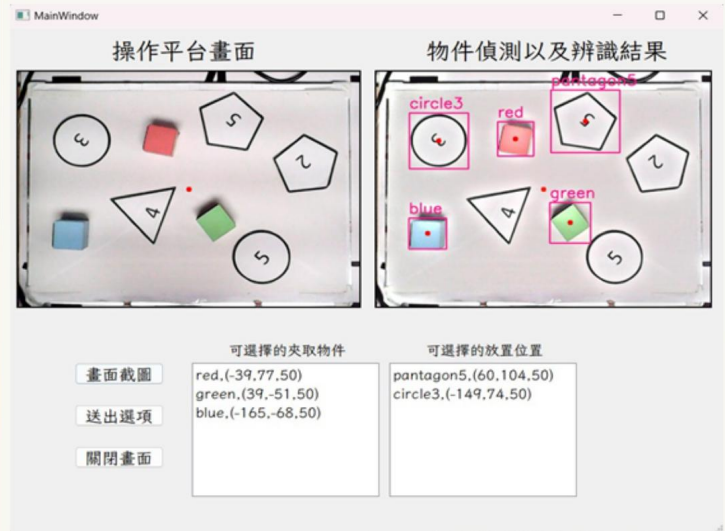
TCP連線流程圖



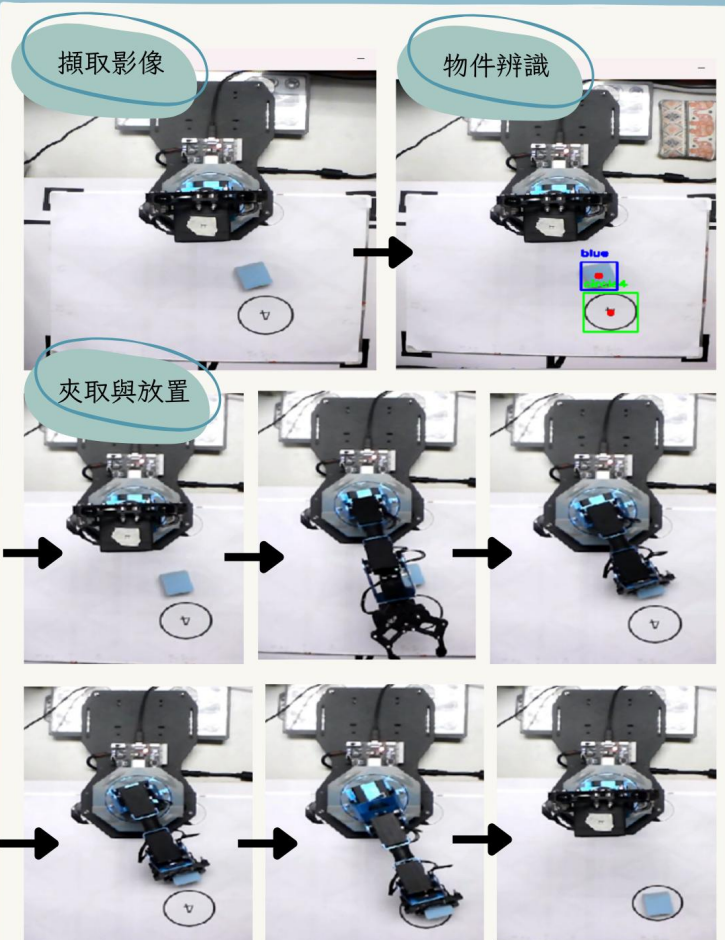
硬體架構圖

成果展示

-人機介面設計示意圖-



1. 按下"畫面擷取", 在左側及下方欄顯示辨識結果
2. 點選夾取物件、放置位置, 並按下"送出選項"
3. 按下"關閉畫面"結束執行



版本一: 運行動作組
優點-精準、快速
缺點-缺乏靈活性

版本二: 逆向運動學
優點-靈活性高
缺點-存在奇異點