

國立臺灣大學工學院土木工程學系

碩士論文

Department of Civil Engineering

College of Engineering

National Taiwan University

Master Thesis



應用電腦視覺技術於袋裝水泥倉儲管理系統自動化

Automated vision-based warehouse management

system for cement bag

曹裕

Yu Tsao

指導教授：林之謙 博士

Advisor: Jacob J. Lin, PhD.

中華民國 112 年 7 月

July, 2023

# 國立臺灣大學碩士學位論文

## 口試委員會審定書

MASTER'S THESIS ACCEPTANCE CERTIFICATE  
NATIONAL TAIWAN UNIVERSITY

應用電腦視覺技術於袋裝水泥倉儲管理系統自動化

Automated vision-based warehouse management system for cement bag

本論文係曹裕(姓名)P09521709(學號)在國立臺灣大學土木工程學系(系/所/學位學程)完成之碩士學位論文，於民國 112 年 07 月 03 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department / Institute of Department of Civil Engineering, National Taiwan University on 03 (date) 07 (month) 2023 (year) have examined a Master's thesis entitled above presented by Tsao Yu (name) P09521709 (student ID) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

林之謙

(指導教授 Advisor)

林之謙

曾惠斌

曾惠斌

陳柏翰

陳柏翰

紀乃文

紀乃文

系主任/所長 Director:

葛宇甯



## 誌謝

謹以此文獻研究之成果，我要向許多人致謝，表達我深深的感激與敬意。

首先，我要感謝我的指導老師 林之謙教授，您的專業知識、耐心指導、以及對研究的熱情，深深影響了我的觀念，對於學習的態度更加嚴謹。同樣地，我要衷心感謝 曾惠斌教授，您的悉心指導、寶貴建議以及深入的學術見解，使我受益匪淺，對我在研究領域起到了重要的推動作用。

接著，我要感謝研究助理銘順，研究室的政宏、承耘、懷恩及梓航等學長學弟，你們在我研究過程中的無私奉獻與幫忙令人欽佩。我也要感謝共同奮鬥的研究所同學，感謝你們平時的合作、討論和共享。在我們互相學習中，你們的熱情、批評精神和深入的洞察力，讓我知道人外有人，天外有天。有幸能與你們當同學，這段求學期間的收穫，對我未來的人生有著極大的意義。

此外，我要感謝我的母親、老婆、小孩，你們的無私、支持和信任是我前進的力量源泉。感謝您們對我的付出和無盡的鼓勵，讓我有勇氣面對各種挑戰。

最後，我要感謝自己。感謝自己的毅力、耐心和奮鬥精神，讓我能夠克服挑戰、跨越障礙，完成這段研究旅程。這個經歷將成為我人生中寶貴的一部分，成為我將來的成長和發展奠定堅實的基礎。未來，我將以更高的標準對待每一件事，為社會做出更多的貢獻。

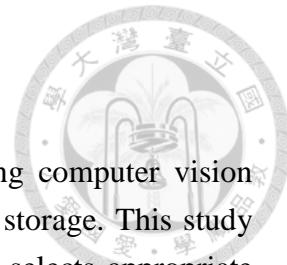
中華民國 112 年 7 月



## 摘要

本論文旨在探討應用電腦視覺技術於袋裝水泥倉儲管理系統自動化的可行性。本研究主要採用點雲(Point Cloud)技術建立 3D 模型，並選擇適當的監視器架設位置，藉此監控並識別水泥品名，進行數量計算與進出貨追蹤，並將攝影機訊息自動傳送到資料庫中，以實現自動化管理。此外，本研究採用物體偵測(Object detection)技術、OCR 光學字元辨識 (Optical Character Recognition)技術、人工智能(Artificial intelligence)等，透過標註現有的資料來進行訓練和學習，然後建立深度學習模型來進行預測，並由機器學習辨識監視器中水泥資訊，包含：進出貨、品牌、數量等資料來進行貨物管理，最後結合庫存管理系統，將貨物資訊自動寫入資料庫中，提供銷貨管理系統使用。經由實驗結果顯示，本研究所提出的方法能夠有效地辨識各種廠牌的水泥之高準確率，及高度誤差，並且能夠在不同的營造業和水泥經銷商等領域廣泛應用。本研究所提出的自動化管理方法，將為營建倉儲管理系統的現代化升級提供有力的支持。

**關鍵字：**點雲、電腦視覺、深度學習、袋裝水泥、倉儲管理



## Abstract

The purpose of this paper is to explore the feasibility of applying computer vision technology to automate the management system of bagged cement storage. This study primarily utilizes point cloud technology to build a 3D model and selects appropriate positions for installing monitors to monitor and identify cement types, perform quantity calculations, and track incoming and outgoing shipments. The camera information is automatically transmitted to a database to achieve automated management. In addition, this research employs object detection technology, Optical Character Recognition (OCR), artificial intelligence, and other techniques. By annotating existing data, training and learning are conducted to establish deep learning models for prediction. Machine learning is used to recognize cement information captured by the monitors, including shipment records, brands, quantities, etc., for effective inventory management. Finally, by integrating with an inventory management system, the cement information is automatically written into the database, providing support for sales management systems. Experimental results demonstrate that the proposed method can accurately identify various cement brands with high precision and minimal error. It can be widely applied in different construction industries and cement distributors. The automation management approach proposed in this study provides strong support for the modernization of construction storage management systems.

**Keywords:** point cloud, computer vision, deep learning, bagged cement, storage management



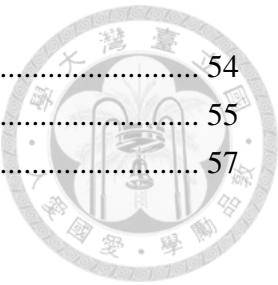
## 目錄

|                    |     |
|--------------------|-----|
| 誌謝 .....           | i   |
| 摘要 .....           | ii  |
| Abstract .....     | iii |
| 目錄 .....           | iv  |
| 圖目錄 .....          | v   |
| 表目錄 .....          | vii |
| 第一章 緒論 .....       | 1   |
| 1-1 研究背景與動機 .....  | 1   |
| 1-2 研究目的與重要性 ..... | 2   |
| 1-3 研究方法與流程 .....  | 4   |
| 1-4 論文架構 .....     | 5   |
| 第二章 文獻探討 .....     | 7   |
| 2-1 監控工地進度 .....   | 7   |
| 2-2 保障工人安全 .....   | 8   |
| 2-3 監視材料的使用 .....  | 9   |
| 第三章 系統架構與設計 .....  | 12  |
| 3-1 資料收集 .....     | 13  |
| 3-2 資料前處理 .....    | 15  |
| 3-3 影像辨識模型設計 ..... | 16  |
| 3-4 袋裝水泥辨識 .....   | 18  |
| 第四章 系統實作與評估 .....  | 26  |
| 4-1 倉庫點雲建立 .....   | 26  |
| 4-2 建立影像辨識模型 ..... | 36  |
| 4-3 進貨出貨判斷 .....   | 41  |
| 4-4 水泥高度計算 .....   | 44  |
| 4-5 訓練模型數據 .....   | 48  |
| 4-6 系統優點與限制 .....  | 56  |
| 第五章 結論與建議 .....    | 58  |
| 5-1 研究結果總結 .....   | 58  |
| 5-2 未來研究方向建議 ..... | 59  |
| 第六章 參考文獻 .....     | 61  |

## 圖目錄



|  |    |
|--|----|
| 圖 1-1 水泥倉庫現場 .....                     | 2  |
| 圖 1-2 水泥與棧板 .....                      | 3  |
| 圖 1-3 倉庫俯視圖 .....                      | 4  |
| 圖 2-1 方法概述 .....                       | 8  |
| 圖 2-2 基於視覺的進展監測的宏觀概念化 .....            | 10 |
| 圖 3-1 系統架構圖 .....                      | 13 |
| 圖 3-2 倉庫現況 .....                       | 14 |
| 圖 3-3 倉庫監視器配置圖 .....                   | 15 |
| 圖 3-4 單視立體高度測量透視圖 .....                | 19 |
| 圖 3-5 交叉比(ross ratio)示意圖 .....         | 19 |
| 圖 3-6 針孔相機模型圖 .....                    | 22 |
| 圖 3-7 將相片座標轉換成世界空間座標 .....             | 22 |
| 圖 3-8 照片平面的 x 軸 .....                  | 23 |
| 圖 3-9 相機空間座標轉換為世界空間座標 .....            | 24 |
| 圖 4-1 倉庫實體照片 1 .....                   | 28 |
| 圖 4-2 倉庫實體照片 2 .....                   | 28 |
| 圖 4-3 倉庫點雲圖 .....                      | 29 |
| 圖 4-4 倉庫拍照路線圖 1 .....                  | 30 |
| 圖 4-5 倉庫內部點雲 1 .....                   | 30 |
| 圖 4-6 倉庫內部點雲 2 .....                   | 31 |
| 圖 4-7 倉庫拍照路線圖 2 .....                  | 32 |
| 圖 4-8 倉庫內部點雲 3 .....                   | 32 |
| 圖 4-9 倉庫內部點雲 4 .....                   | 33 |
| 圖 4-10 以環繞方式拍攝水泥樣本 .....               | 34 |
| 圖 4-11 水泥點雲圖 .....                     | 34 |
| 圖 4-12 YOLOv5 水泥與堆高機訓練影像收集 .....       | 37 |
| 圖 4-13 YOLOv5 水泥與堆高機測試影像收集 .....       | 38 |
| 圖 4-14 偵測到的水泥、堆高機、輪胎等物件 .....          | 40 |
| 圖 4-15 偵測到出貨 .....                     | 41 |
| 圖 4-16 偵測到進貨 .....                     | 42 |
| 圖 4-17 偵測到水泥包裝上的文字 .....               | 43 |
| 圖 4-18 加入 CCTV 點雲圖 .....               | 46 |
| 圖 4-19 相片座標圖 .....                     | 47 |
| 圖 4-20 混淆矩陣 (Confusion Matrix) 圖 ..... | 50 |
| 圖 4-21 F1 分數 (F1 score) 圖 .....        | 52 |
| 圖 4-22 P 曲線 (P-curve) 圖 .....          | 53 |



|                               |    |
|-------------------------------|----|
| 圖 4-23 R 曲線 (R-curve) 圖 ..... | 54 |
| 圖 4-24 PR 曲線 (PR-curve) ..... | 55 |
| 圖 4-25 YOLOv5 檢測到的邊界框圖 .....  | 57 |

## 表目錄

|                         |    |
|-------------------------|----|
| 表 4-1 YOLOv5 訓練結果 ..... | 49 |
| 表 4-2 模型所犯的錯誤類型 .....   | 51 |



# 第一章 緒論



在現代倉儲管理中，有效的物料進出控制和貨物管理是營建業運營成敗的關鍵要素，傳統的人工物料盤點和管理方式存在著效率低下，有著容易出錯和耗時費力的問題。然而，隨著科技的不斷進步，倉儲管理系統自動化成為提升效率和準確性的重要手段。本研究旨在探索應用 3D 模型和電腦視覺技術於倉儲管理系統的自動化，以改善傳統倉儲管理方式的不足之處，通過利用先進的電腦視覺技術，本研究在探討是否可以實現對貨物的自動識別、數量統計和位置追蹤，以提高倉儲管理的效率和準確性。

## 1-1 研究背景與動機

在營造業中，倉儲管理是一個重要的環節。由於營造業的物料種類眾多、數量龐大，而且常常涉及到多個工地、供應商以及不同的項目，因此要做好倉儲管理，需要具備高度的準確性、即時性和有效性。一個好的倉儲管理系統可以大大提高物料的運輸效率，減少物料損耗和人力成本，保證工程進度和品質的穩定。

當今如何在倉庫中使用電腦視覺技術的一些示例：亞馬遜<sup>[30]</sup>正在使用電腦視覺技術來跟蹤其倉庫中貨物的移動。該數據用於提高庫存水平的準確性，識別庫存不足的區域，並防止缺貨。沃爾瑪正在使用電腦視覺技術來自動化訂單揀選過程。這導致訂單揀選的準確性和效率顯著提高。UPS 正在使用電腦視覺技術來檢查貨物是否存在缺陷，這有助於提高產品質量並減少退貨次數。隨著技術的不斷發展，可以期待看到更多關於電腦視覺技術如何在倉庫中使用的例子。這項技術有可能顯著提高倉庫運營的效率和準確性，並導致成本降低、客戶滿意度提高和品牌聲譽提高。

然而，目前大多數營造企業的倉儲管理仍然是手動作業和人工監控，存在著一些問題，如容易出現人為失誤、監控範圍受限、資訊收集不全面等。為了解決這些問題，本



研究希望利用現代科技，將電腦視覺技術和物聯網技術應用到營造業的倉儲管理中，實現自動化、智能化、資訊化的倉儲管理。

因此，本論文旨在研究如何利用相機和監視器等電腦視覺技術，結合網際網路技術，實現營造業倉儲管理的自動化，從而提高物料管理的準確性和效率。具體來說，本研究希望探討如何利用影像辨識來對於水泥倉庫自動盤點，將庫存可以即時傳回系統統的資料庫中，減少人為盤點錯誤，進而研究電腦視覺技術應用在倉儲系統中的機會與效率。圖 1-1 為一般水泥倉庫的樣貌。



圖 1-1 水泥倉庫現場

## 1-2 研究目的與重要性

本研究方向是探索如何利用電腦視覺技術，來準確測量追蹤水泥的進出貨。實驗將使用偵測到的水泥、推高機、輪胎等物件作為參考，並融合物體檢測和高度測量技術，以提供更準確、可靠的水泥高度資訊。透過本研究，本文將探索這些方法在倉庫管理系統中的應用。

### 1. 研究目的：

基於實務狀況與創新技術發展目標，本研究旨在開發一個自動化袋裝水泥盤點系統，利用影像辨識與電腦視覺三維重建技術，以解決人工耗力費時的問題。該系統將能



能夠自動識別並計數袋裝水泥，並實現倉庫盤點的自動化，從而提高盤點的效率和準確性。在實際上目前水泥在搬運及入庫皆是如圖 1-2 使用棧板上放置水泥，由推高機移動放置於卡車上或倉庫內。



圖 1-2 水泥與棧板

使用自動視覺化的倉庫管理系統來管理水泥袋有幾個好處。首先，這些系統可以幫助提高庫存記錄的準確性，通過實時追蹤水泥袋的位置和數量，這些系統可以幫助防止導致庫存不足或存貨過多的錯誤。其次，自動視覺化的倉庫管理系統可以提高倉庫運營的效率。通過自動化計算和定位水泥的任務，這些系統可以讓倉庫工人專注於其他任務，如裝卸卡車。第三，自動視覺化的倉庫管理系統可以幫助降低勞動成本。通過自動化計算和定位水泥袋的任務，這些系統可以幫助減少對人工勞動的需求。

## 2. 研究重要性：

在現今高度自動化、智能化的社會中，倉儲管理是企業不可或缺的一環，影響到整個生產鏈上下游供應鏈的運作效率，而自動化的盤點技術對於倉儲管理而言更是必不可少的。然而，目前常見的盤點方式多仰賴人工，不僅成本高昂且容易出現錯誤，因此本研究的解決方案具有重要意義，可以提高倉儲管理效率、降低成本及減少錯誤發生的風



險，同時也可提升營運流程的可靠性與穩定性，有助於企業增強競爭力。此外，本研究所探討的技術也可應用在其他產業領域，具有相當的普遍性及實用性，因此本研究亦對相關產業的發展與進步具有一定的貢獻。圖 1-3 為倉庫的俯視圖，倉庫前方皆會有一個廣場供卡車或聯結車進出的空間，倉庫內為放置貨物的空間，如何架設監視器，讓整個倉庫的貨物都可被監視到，也是一個重要課題。



圖 1-3 倉庫俯視圖

### 1-3 研究方法與流程

本研究將採取以下方法與流程，建構人工智慧倉儲管理系統，讓倉庫水泥的進出貨，得以透過數據化呈現，藉由監視器的錄影。簡述內容如下：

1. 倉庫管理中實現電腦視覺技術的研究方法如下：

- 照片識別：照片識別是一種可用於識別照片中對象的技術。這可用於識別貨架上的物品、跟蹤進出倉庫的貨物移動以及檢查貨物是否有缺陷。
- 對象跟蹤：對象跟蹤是一種可用於跟蹤對象隨時間移動的技術。這可用於跟蹤倉庫中貨物的移動、識別瓶頸並優化倉庫佈局。
- 機器學習：機器學習是一種人工智慧，可用於訓練電腦從數據中學習。這可用於提高照片識別和對象跟蹤算法的準確性。

這些只是可用於在倉庫管理中實施電腦視覺技術的幾種方法。使用的具體方法將取決於倉庫的具體條件及需要。



## 2. 本研究將採取以下流程簡述內容如下：

- ① 資料收集：收集相關文獻及研究報告，並參考相關業界專家的意見，以確立研究方向和研究問題。
- ② 需求分析：針對研究問題進行需求分析，確定系統功能和性能要求。
- ③ 技術選擇：選擇合適的點雲建模技術、電腦視覺技術及相關軟體工具，以實現系統的動化功能。
- ④ 系統設計：設計基於 3D 建模技術和電腦視覺技術的自動化倉儲管理系統，包括系統架構、功能模塊、軟體介面等。
- ⑤ 系統實現：實現系統設計，開發系統模塊，並進行系統整合和測試。
- ⑥ 系統評估：對系統進行評估，測試系統的功能和性能是否達到預期效果。
- ⑦ 結果分析：分析實驗結果，評估系統的效果和性能，提出改進方案。

## 1-4 論文架構

首先在第一章，本研究將介紹本論文的研究背景、研究目的和重要性。文中將討論研究方法和研究流程，以便讀者能夠理解論文的整體架構和進程。在第二章中，筆者將回顧與本論文研究相關的文獻。文中將探討監視器的使用、電腦視覺技術以及倉儲管理在營造業中的應用。這些文獻將為本次的研究提供理論基礎和相關背景知識。第三章將介紹本論文所使用的系統架構和設計。文中將討論相機和監視器的選擇和架設位置，並探討電腦視覺技術的整合以及倉儲管理系統的使用。這將為筆者的研究提供實際操作和實現的基礎。在第四章中，文中將介紹本論文所設計的系統的實現和評估。本研究將討論系統的硬體實現和軟體實現，並進行系統的測試和驗證。這將幫助本研究評估系統的性能和效能。在最後一章，本研究將總結本論文的研究成果和結論。文中將討論系統的優點

和限制，並提出未來研究的方向和建議。這將有助於指導進一步的研究和發展。透過以上六個章節的內容，本論文將全面探討與研究目標相關的背景知識、技術應用、系統設計和實現等各個方面，最終提供全面且有價值的研究結果和建議。



## 第二章 文獻探討



在建設專案中，有效運用先進的技術和創新的解決方案對於提高工地管理效率、確保工人安全以及優化資源利用至關重要。監視器在營造業中的使用可以提高工地的安全性、管理效率和生產力。它提供了實時的視覺監控，幫助管理人員追蹤和管理各方面的工作，從而確保項目的順利進行。然而，基於人的手動監視器需要額外的時間和成本來處理大量影片、照片等數據。為了克服這些限制，必須引入基於視覺的自動化監控系統和遠端監控技術，以提高生產力和安全性分析。許多研究已經提出了基於視覺的設備活動識別方法，用於設備操作和生產率分析。

近年來，電腦視覺技術在營建業中的應用逐漸受到關注，因其能夠從影片或照片數據中自動識別和分析物體，提供寶貴的資訊和洞察力。本研究將討論這些技術在營建業不同的應用案例，包括工地安全監控、材料管理、工作進度追蹤和資源分配等。目前研究側重於面向半自動化或操作級視覺監控自動化的視覺化資料分析。基於上述的許多用途，監視器的使用對視覺系統的性能有著巨大的影響。

### 2-1 監控工地進度

Golparvar-Fard et al. (2013)，從現場行動裝置的影片流，提出了一個識別單個車輛行為的框架。Rezazadeh Azar et al. (2013)<sup>[3]</sup> 根據確定的活動，挖土機的泥土裝載循環時間，對一輛沙石卡車進行了分析。. Bügler et al. (2017)<sup>[4]</sup>還提出了一種通過攝影測量和影片分析自動測量土方作業生產率的方法。Kim et al. (2018)<sup>[2]</sup>提出的方法確定了考慮挖土機和砂石卡車之間相互作用的活動類型。此研究跟本研究此實驗過程像似包含；資料收集、資前處理、影像辨識模型建立、最後提出系統報告與分析。本章特別將它流程圖呈現做為參考如圖 2-1 方法概述，這與本次研究方法有類似之處，都是用 CCTV 閉路電視將影像紀錄後，以電腦視覺技術加以分析，然後再將其模型建立。

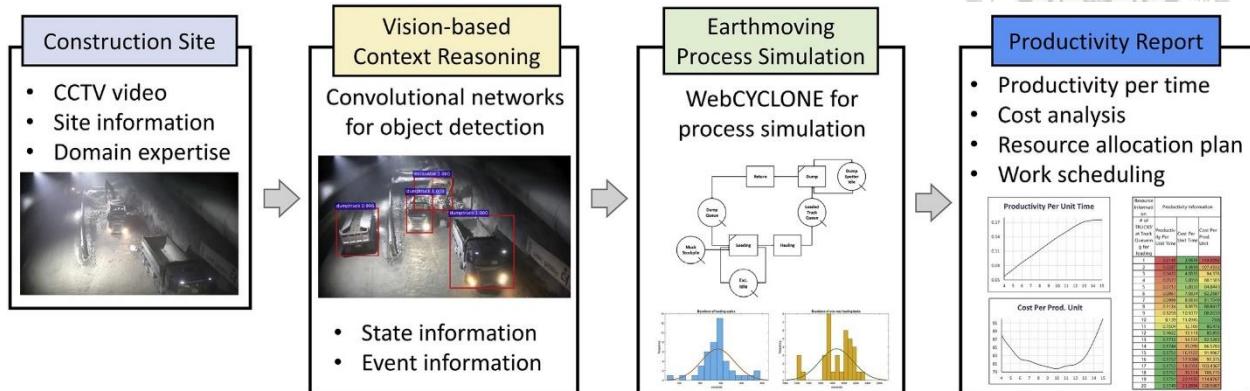


圖 2-1 方法概述

Lin (2021) 從檢討土木工程進度監測現況，介紹電腦視覺技術之機會，討論視覺技術應用影像於工程進度控管，包含工項及施作兩方面的監控。實境建模將現實世界中的物體，例如建築物、工地、機具等，透過影像或掃描方式以三維點雲或網格（mesh）模型數位化呈現，工程師可以取代以往傳統測量方式進行快速及準確的量測及紀錄。

## 2-2 保障工人安全

Han 和 Lee (2013)<sup>[5]</sup>從事故統計數據和導致人體工程學傷害的清單中確定了關鍵的身體姿勢和動作清單。他們開發了一種方法，可以從立體照片中識別工人不安全的姿勢和動作。Ding et al. (2017)<sup>[6]</sup>的方法還通過使用卷積神經網路和長短期記憶來檢測不安全的工人行為。由於佩戴個人防護設備（PPE）是瞭解工人不安全行為的主要指標，Park et al. (2015)<sup>[7]</sup>和 Fang 等人 (2018)<sup>[8]</sup>提出了一種自動監控方法，可以從監控攝影機檢測未戴安全帽的工人。

Yusun et al. (2022)<sup>[17]</sup>使用室內閉路電視（CCTV）監控開發了一種基於電腦視覺技術的早期火災探測模型（EFDM）。所提出的 EFDM 通過實際火災測試得出火災探測時間。



基於視覺的技術的影響最初是有限的，因為數據採集設備的複雜性和處理所需的計算能力。

## 2-3 監視材料的使用

Golparvar-Fard et al. (2010/2015)<sup>[12][13]</sup> 將影像式點雲模型分割為三維空間單位體素（Voxel）與 BIM 模型比較，發展出以支援向量機（Support Vector Machine, SVM）之監督式學習（Supervised learning）模型判斷物件的存在與否。

Bosché et al. (2015)<sup>[14]</sup> 則就機電模型與雷射掃描點雲模型進行比較，討論較為精細的物件進度監測方法。Turkan et al. (2012)<sup>[15]</sup> 進一步針對施工施作步驟細項進行空間幾何分析，並加入實獲值（Earned Value）以成本分析進度。

Han and Golparvar-Fard (2015)<sup>[16]</sup> 提出應用電腦視覺方式反投影 BIM 模型中的元件至相機影像分割出施工物件，透過這些分割影像進行材質分析達到施作細節步驟判斷，如木作材質為模板裝設中、水泥材質為完成灌漿。

然而，隨著設備和高性能計算的日益普及，在自動化施工流程的應用中實施基於視覺的技術已經變得可行。Varun et al. (2022)<sup>[18]</sup> 研究其在建築自動化進度監控中的適用性的，從數位照片、影片、深度照片和 3D 點雲中獲取數位資訊，處理資訊並採取行動。如圖 2-2 顯示了構成基於視覺的進度監控的六個宏觀過程。這些過程分為三個階段：①使用 3D 重建技術採集竣工數據並將其轉換為點雲。②從點雲生成竣工模型（參考計劃模型，如果可用）③比較竣工和按計劃模型，通過視覺化和/或量化已完成的工作來評估進度。

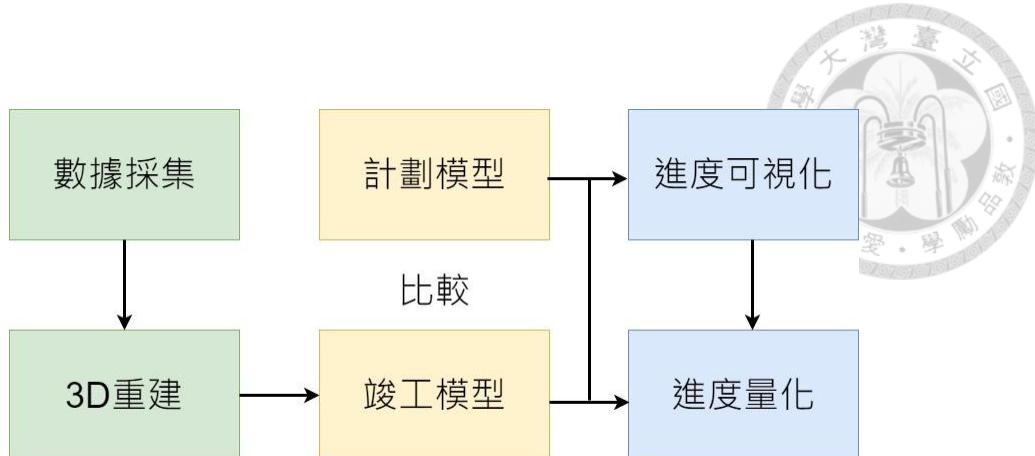


圖 2-2 基於視覺的進展監測的宏觀概念化

雖然企業資源規劃 ERP 起源於製造業，基於經營建築企業的需求，ERP 顯示了其在建築行業的潛力。但是，該建築行業的獨特性質阻止了現有 ERP 系統的直接實施。

Jonathan et al. (2003)<sup>[19]</sup> 強調了建立基本理論對於開發企業資源規劃系統 CERP 的重要性，CERP 必須解決一般此行業慣例的性質。由於建築項目本質上是短暫的，跨越不同的時間段，需要不同的資源，建築企業通常很難在生產能力和手頭的實際工作量之間取得平衡。為方便業界不同使用者接受及使用系統，CERP 系統應為：以項目為導向、集成、並行和分散式、開放和可擴展來評估預期收益。

Mirosław et al. (2009)<sup>[20]</sup> 開發一種統一的方法來定義關鍵績效指標 KPI 的資訊特異性，解決了與企業資源規劃 ERP 系統在工程建設行業中的應用有關的關鍵問題：工程建設行業中哪些業務流程領域不能使用 ERP 來收集與業務流程相關的 KPI。

材料管理是一個集成的過程，由人員、組織、技術和程式組成，用於在整個資本專案的生命週期中有效地識別、量化、獲取、加速、檢查、運輸、接收、存儲和保存材料、設備和關聯資訊。目標是確保以有效的方式採購材料和設備的正確質量和數量，以合理的成本獲得，並在需要時可用。這使得物料管理在資本專案中的早期和全面參與變得非常重要。Carlos et al. (2014)<sup>[21]</sup> 該研究清楚地表明，材料管理的成熟度、形式性和系統性方法顯著提高。



建設專案由構成項目供應鏈的大量承包商、分包商和供應商組成。建築供應鏈管理作為一個研究領域，關注的是改善這些公司的協調，以提高項目績效，包括速度、成本、可靠性、品質等各種指標。Yu-Cheng et al. (2003)<sup>[22]</sup> 介紹了 e-Hub 建築供應鏈管理 (eHub-ConSCM) 系統的原型和系統，用於執行建設專案的資訊共用服務。說明 eHubConSCM 系統的資訊需求，描述了一般的系統架構，並介紹了 eHubConSCM 系統的實現。

監視器在工地、倉庫以及各種公共場合，已非單純儲存影像或照片的功能，實際上監視器已經變成人工智慧的眼睛，透過這眼睛能分析各種資料，訓練人工智慧達到人們想要的功能。在電腦視覺技術的發展為營建業帶來了許多新的應用和解決方案。但是，目前還未看到運用在水泥倉庫貨物管理中的自動化。我想利用電腦視覺技術，應該能夠開發出能夠自動識別、計算和追蹤水泥倉庫中的貨物的系統。這項技術若能開發成功，在倉儲管理的應用能夠大幅提高貨物管理的效率和準確性，同時減少人力成本和錯誤。透過自動化的人工智慧計算，電腦視覺技術成為水泥倉庫管理的重要工具，為業界帶來更高的效益和競爭力。

### 第三章 系統架構與設計



本章節闡述了此研究整合了許多方法來架構與設計此系統，其中包含了四個部分：1.資料收集 2.資料前處理 3.影像辨識模型 4.袋裝水泥辨識結果，流程如圖 3-1 說明整個系統研究的過程，在以下說明。透過「1.資料收集」的過程，本實驗可以利用各個不同角度的相機或監視器，來統計收集本研究想要偵測的水泥廠牌圖片、堆高機與輪胎的圖片、倉庫空間的大小與水泥擺放的位置。這些收集的資料可以做後續研究的基礎。透過對資料的研究來讓業者知道需要放多少支監視器、架設的位置、含蓋的空間以及需要多少經費。為了後續的目標偵測，必須事先處理一些已經收集的資料作「2.資料前處理」。這部分包含收集倉庫影像、收集 CCTV 影像、最後收集倉庫點雲資料，讓本研究可以從這些倉庫點雲資料找到所有監視器的參數，以利後續與偵測到物件的高度做測量計算。

傳統的目標檢測方法使用基於滑動視窗的區域選擇策略<sup>[23]</sup>，它不是有針對性的，並且具有很高的複雜性。他手工設計的特徵提取器對目標的多樣性不是很完整。由於電腦性能的提高，特別是 GPU，深度學習方法的提高<sup>[24]</sup>這需要很多計算，再次吸引了研究人員的青睞。特別是卷積神經網路在電腦視覺技術領域的成功應用。這些方法的優點是不需要手動標特徵，可以自動偵測提出照片特徵，近年來，許多研究人員提出了一系列基於深度學習的目標檢測演算法。本研究將它放在「3.影像辨識模型」。

最後，本實驗將資料前處理中得到監視器在倉庫空間的參數，結合影像辨識模型中偵測到的物件，使用單視立體高度計算與多視立體高度計算兩種方法，將水泥的高度計算出來，然後轉換成數量，並且偵測進貨與出貨以及偵測出水泥的廠牌，這部份我放在「4.袋裝水泥辨識結果」說明。

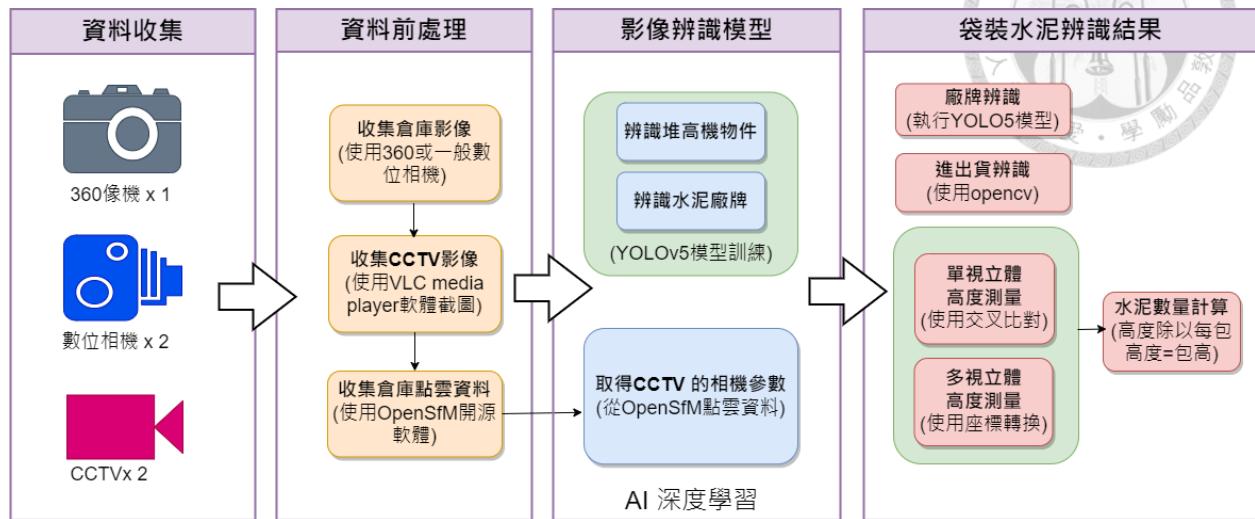


圖 3-1 系統架構圖

### 3-1 資料收集

監視器架設：監視器架設是為了實現袋裝水泥倉儲管理系統的自動化。透過適當的監視器位置，可以捕捉到水泥的進出貨活動、水泥品名、數量等資訊，並將這些資訊用於庫存管理和貨物追蹤。

- 倉庫尺寸和布局：

倉庫的大小和形狀會影響監視器的數量和位置。通常需要監控整個倉庫的入口和出口，以及存放區域，因此需要根據倉庫的尺寸和布局來決定監視器的分布。

- 存放物品的種類和價值：

倉庫存放的物品種類和價值也是決定監視器數量和位置的考量因素。存放價值高的物品，需要加強監控，防止盜竊等意外發生。

- 監控需求和目的：

監控的目的不同，需要設置的監視器數量和位置也不同。如果是監控出入口，需要考慮人車通行的情況，如果是監控存放區域，需要考慮物品的堆疊情況。

- 照明和視野：

監視器的照明和視野也是決定位置的重要因素。監視器的視野需要考慮遮擋物的情況，以及光線的影響。



- 監視器的解析度和覆蓋範圍：

監視器的解析度和覆蓋範圍也需要考慮。解析度高的監視器可以提供更清晰的影像，覆蓋範圍大的監視器可以監控更多的區域，每個倉庫的形狀大小空間都不一樣，必須依照各例去安排監視器，圖 3-2 為一般常看到的倉庫模式。



圖 3-2 倉庫現況

基於現場的插座設備以及倉庫的條件考量，本研究說服業者在水泥儲存區加設日光燈管兩座以加強燈光照明，有利於資料收集與日後影像監視器的明亮。本實驗將監視器架設如圖 3-3 所示：1、2 架設在鐵門兩側，用以探測水泥物件的進入門口或者出去門口，用以初

步偵測進貨及出貨方向。3、4 號攝影機為原來業者已架設在倉庫中堅的兩側天花板上，原來是用來監視貨物是否遭竊與平時保安公司監看為用，此為舊有之設備。本實驗也將之運用在電腦影像技術開發前端收集資料以節省成本。這四支監視器可以同步進行計算，用以補足各監視器有死角，不能全角度見到的問題。



圖 3-3 倉庫監視器配置圖

## 3-2 資料前處理

### 1. 收集 360 或一般數位相機的影像:

先用 360 或一般相機建立倉庫的點雲。選擇適合的相機：市場上有各種 360 度或一般數位相機可供選擇，因此選擇一款符合需求的數位相機很重要。一些要考慮的因素包括照片的解析度、視野範圍和價格。設置相機：選擇好相機後，需要在倉庫中設置相機。確保相機穩定放置並且能夠清晰地觀察整個區域。拍攝照片：相機設置好後，可以開始拍攝照片。務必從倉庫的各個角度拍攝照片。處理照片：拍攝完照片後，需要對照片進行處理以建立點雲，有各種軟體可供使用。檢查點雲：點雲建立完成後，需要仔細檢查確保其準確性。可以進入倉庫內並將點雲與實際環境進行比對。

### 2. 收集閉路電視(CCTV)影像:

再建立點雲的過程，加入兩台 CCTV 的影像，以求得 CCTV 位於倉庫點雲的相對位置。辨識倉庫中的 CCTV 攝影機：穿越倉庫，辨識所有的 CCTV 攝影機。記下每個攝影機的位置和它所朝向的方向。拍攝 CCTV 攝影機的照片：使用相機拍攝每個 CCTV 攝影機的照片。確保照片清晰且焦點正確。處理照片：使用軟體程式處理照片，並建立 CCTV 攝影機的點雲。



### 3. 將 1 與 2 步驟一起建立倉庫點雲。

結合數據。下一步是將步驟 1 和步驟 2 的數據結合起來。這可以使用各種軟體工具完成。一旦數據被結合，可以使用各種軟體工具將其可視化。這可以幫助本研究識別倉庫中感興趣的區域。由這以上 3 步就可知所有影像的相對位置與 CCTV 的相對位置，及獲得整個倉庫的點雲。

## 3-3 影像辨識模型設計

### 1. 收集足夠數量的照片：包含所有的水泥與推高機的照片。

為了建立一個強大而準確的模型，收集足夠數量的照片非常重要，這是因為照片數量越多，模型所能學習的數據就越多，這有助於模型更好地理解所訓練的物體的不同特徵。在訓練識別水泥和推高機的模型時，收集這兩種物體的照片是很重要的，並且要從不同的角度和光線條件下進行拍攝。這將幫助模型學習在不同呈現方式下識別這些物體。一旦收集到足夠數量的照片，就需要對它們進行標註，本研究需要為每張照片標註出其中包含的物體，這將幫助模型學習照片和標籤之間的關係。完成照片標註後，可以對模型進行驗證，這意味著需要在模型之前從未見過的一組照片上進行測試，這將有助於評估模型的準確性。

如果模型的準確性不夠高，可能需要收集更多的照片或更仔細地進行標註。如果模型的準確性足夠高，就可以使用它來在新的照片中識別水泥和推高機。以下是收集足夠數量照片的一些原則：

使用多種來源來收集照片，本研究拍攝自己的照片，來確保照片的高品質，照片應清晰且聚焦，仔細標註每一張照片，標籤應準確且一致，定期對模型進行驗證，這將有助於確保模型的準確性。

### 2. YOLOv5 模型：

近年來，許多研究人員提出了一系列基於深度學習的目標檢測演算法。2015 年 Redmon et al. 提出了 YOLO 演算法<sup>[25]</sup>，它比其他演算法快得多。Joseph et al.<sup>[26]</sup> 在 2018 年提出了 YOLOv3 目標檢測演算法，進一步提高了檢測速度和準確性。到 2020 年，YOLO 系列演算法<sup>[27]</sup> 已經發展到 YOLOv5。

YOLOv5 模型基於回歸的目標檢測演算法。將照片或影片輸入深度網路後，YOLO 根據損失函數的計算完成對目標分類和位置資訊的預測，使目標檢測問題轉化為回歸問題解<sup>[28]</sup>。物體檢測涉及通過邊界框為特定物體設置區域，例如人、汽車、建築物或布料。這種電腦視覺技術執行自動識別。YOLO 是一個應用深度學習的物件檢測模型，在此將照片劃分為一定的區域，並將權重分配給「物體的概率」以出現在每個區域中。此外，它具有即時影片記錄的優勢，因為它是一種統一的檢測方法，可以找到物體的位置並對其進行分類，這一優勢可以提高物體識別的速度。

在這項研究中，本實驗使用 YOLO 模型使用自定義各個廠牌水泥照片數據集進行品 牌及物件檢測訓練。之所以選擇具有出色即時性能的 YOLO，是因為本研究的目標是快速準確地檢測倉庫中的袋裝水泥。因此，YOLOv5s<sup>[29]</sup> 被選為開始訓練的預訓練模型。要訓練模型之前，首先本實驗要有一些照片，本實驗在資料前期處理階段已經準備了不少的圖片，這些圖片在 YOLOv5 裡面需要 2 個資料夾，包含訓練與測試。本實驗必須要將這些圖片中想要被偵測出的照片給它們貼上標籤、註釋。然後放到兩個檔案夾 images 與 labels，這兩個檔案夾分別是訓練(train) & 驗證(val)。本次研究我們拍了各大品牌的袋裝水泥，包含太空包。還拍攝了堆高機，目的是要用已知的堆高機高度或者是輪胎，來推測判斷棧板上的水泥高度。標註的工具可使用 labelme 或者 cvat。當本實驗快速從監視器影像中真偵測出水泥的邊界框(bounding box)，本研究可以將它與倉庫點雲做計算，算初其邊界框的長度。加上 YOLOv5 模型的深度訓練，讓 AI 可以自動偵測水泥的廠牌。

### 3. 取得 CCTV 的相機參數：



使用點雲技術可以將 CCTV 所抓取的水泥照片與現有的倉庫點雲結合，這個新的點雲將使本實驗能夠計算出所有水泥物件的座標，進而確定水泥堆疊的高度。因此，可以計算出水泥的數量。我們建立點雲主要的目的就是要得到所有監視器的參數，以便未來計算監視器所拍出來的影片，進而分析影片中的物體的座標，而且只需要建立一次即可。

### 3-4 袋裝水泥辨識

#### 1. 單視立體高度測量

單眼立體高度測量是一種從單張照片中估計物體高度的技術。該技術通過使用攝影機與物體之間的已知距離來計算物體相對於攝影機的角度。然後可以使用三角函數來計算物體的高度。

① 要使用單眼立體高度測量，本研究需要以下兩個資訊：

物體的單張照片與攝影機與物體之間的已知距離，一旦獲得這兩個資訊，可以按照以下步驟估計物體的高度：

1. 在照片中識別物體。
2. 測量攝影機與物體之間的距離。
3. 計算物體相對於攝影機的角度。
4. 使用三角函數計算物體的高度。

② 以倉庫進貨圖 3-4 當範例來說明，要算出水泥的高度可以用以下方法：

1. 從平行線中找到消失點。透視圖中的平行線都匯聚在一個稱為消失點的點上。要找到消失點，請通過任意兩條平行線畫一條線。線與地平線相交的點是消失點。
2. 畫出消失線。找到消失點後，通過它畫一條平行於地平線的線。這是消失線。



### 3. 使用交叉比來計算水泥的高度。



圖 3-4 單視立體高度測量透視圖

### 4. 交叉比(cross ratio)

交叉比是一個數學概念，可用於比較射影空間中的四個點。定義為四點坐標之差乘積之比。交叉比是一個射影不變量，這意味著如果四個點進行射影變換，它不會改變。如圖 3-5 中表示其  $A, B, C, D$  四個點用來定義它們實體的距離，而  $a, b, c, d$  四個點用來定義它們照片的距離。

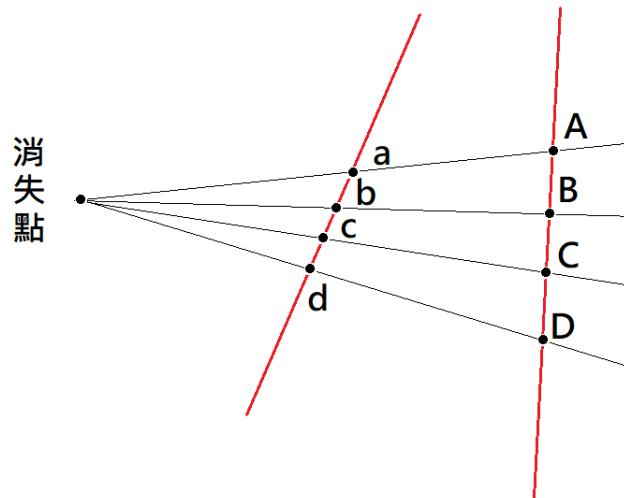


圖 3-5 交叉比(cross ratio)示意圖

cross ratio 如下：

$$CR(A, B, C, D) = \frac{AC}{AD} : \frac{BC}{BD} = \frac{AC * BD}{BC * AD} \quad (3.1)$$



$CR(A,B,C,D)$ 在投影變換下保持不變，得到以下的等式

$$\frac{AC * BD}{BC * AD} = \frac{ac * bd}{bc * ad} \quad (3.2)$$

在這個實驗中  $AD$  為堆高機高度， $CD$  為輪胎高度， $BD$  為水泥標高高度。

所以計算可得：

$$BD = \frac{\frac{(ac * bd)}{(bc * ad)} * CD * AD}{\frac{(ac * bd)}{(bc * ad)} * AD - 1} = \frac{\frac{CR(a,b,c,d)}{AD-CD} * CD * AD}{\frac{CR(a,b,c,d)}{AD-CD} * AD - 1} \quad (3.3)$$

③ 以下是使用單視立體高度測量的一些優點：

- 使用簡單：單視立體高度測量是一種相對簡單的技術，任何具備基本三角學知識的人都可以使用。
- 成本較低：單眼立體高度測量不需要任何特殊設備，對於企業和個人來說是一個價格合理的選擇。
- 高精度：單眼立體高度測量可以高精度地估計物體的高度。

如果正在尋找一種簡單、經濟實惠且準確的估計物體高度的方法，那麼單視立體高度測量是適合的解決方案。

④ 以下是使用單視立體高度測量時的一些限制：

- 距離影響測量精度：攝影機與物體的距離越近，測量結果越精確。
- 角度影響測量精度：當視角過小或物體在視野中過於接近或遠離相機時，測量的精度將受到影響。

單視體高度測量是一種相對簡單的技術，可以用於估計各種環境中物體的高度，該技術特別適用於無法使用多個照片或立體攝影機的情況。

## 2. 多視立體高度測量



多視立體高度測量是一種通過從不同視點拍攝多個立體照片來測量場景中物體高度的技術，該技術通過比較不同照片中物體的相對位置來計算其深度。

要使用多視立體高度測量，系統需要一個立體攝像系統，其中包含兩個或更多個攝影機，這些攝影機彼此之間的距離不同。還需要一個可以用於處理立體照片並計算場景中物體深度的軟體應用程式。

① 一旦具備了必要的設備，就本研究可以按照以下步驟來測量場景中物體的高度：

- A. 從不同視點拍攝多個立體照片。
- B. 將立體照片導入軟體應用程式。
- C. 選擇要測量高度的物體。
- D. 軟體應用程式將自動計算物體的深度。
- E. 物體的高度可以通過將深度乘以兩個攝影機之間的距離來計算。

多視立體高度測量是一種強大的技術，可用於測量各種場景中物體的高度。該技術相對易於使用，並且可以與各種立體攝像系統一起使用。

② 設計的方法如下：

本研究用拍攝的相片建庫的點雲後重新加入 CCTV 所拍攝的相片一起加入 openSfM 建立點雲後，會得到 CCTV 的參數及座標，用這個來計算後面的影像所在的座標，這要用到針孔相機模型來解釋及推導，如圖 3-6 是個針孔相機模型，我們可以利用此模型圖來推導相片的座標，相機的物理座標。相機空間座標以及世界空間座標的關係。然後將它簡化成圖 3-7 的三步驟，如何從將相片座標轉換成世界空間座標的方法及過程，然後將它公式寫出來，最後再將它用電腦語言寫出程式。

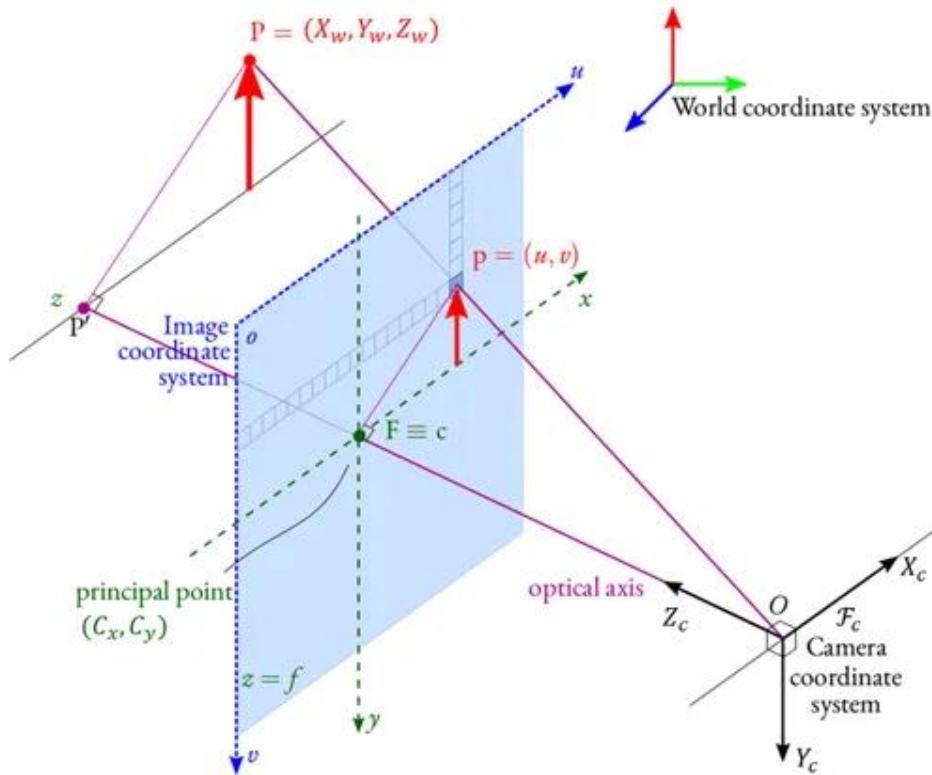


圖 3-6 針孔相機模型圖

(from: [https://docs.opencv.org/3.4.15/d9/d0c/group\\_\\_calib3d.html](https://docs.opencv.org/3.4.15/d9/d0c/group__calib3d.html))

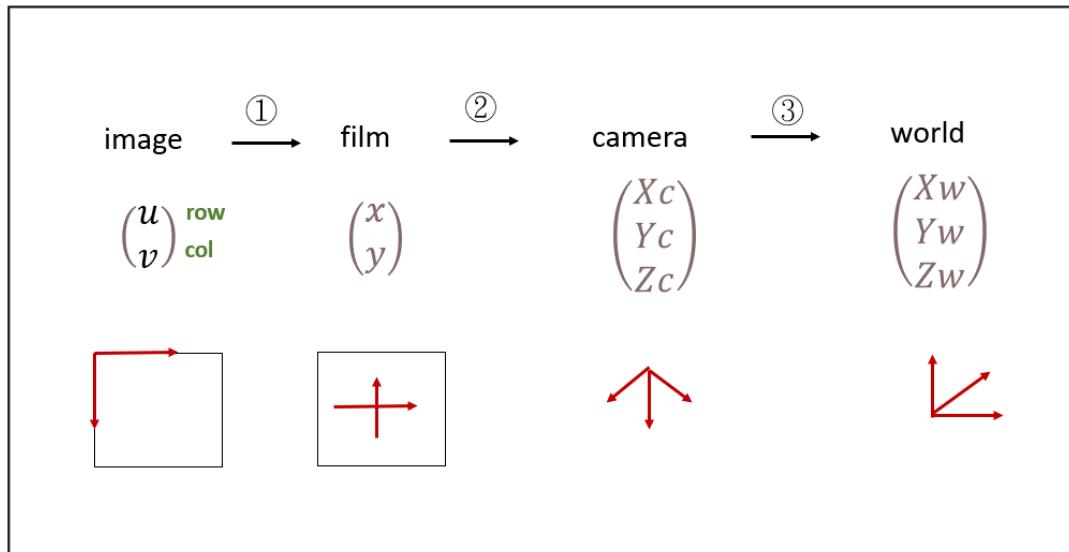


圖 3-7 將相片座標轉換成世界空間座標

本研究要把相片平面座標轉換成世界空間的座標，要以下 3 步驟；



### A. 相片像素座標(u, v)轉換為相機物理座標中 P 點的座標(uc, vc, f)

像素座標系的原點是在左上角，而相機的物理座標系的中心點是在光學中心，是在相機的正中間，由此正中心來抓取感光元件(CCD)進來的光，透過光柵來存取獲取光線。所以，只需要把起始點從相片的左上角移動到相片的正中心即可取得相機物理座標中 P 點的座標。

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} - \begin{bmatrix} u_0 \\ -v_0 \end{bmatrix} \quad (3.4)$$

### B. 相片座標轉換為相機空間座標

以下簡要說明參數所代表的意思，如圖 3-8 所示，O 代表相機中心，F 代表相片中心，F 與 O 的距離稱為焦距(Focal Length) f。把物點 P(Xw, Yw, Zw) 投影至透過相似三角形計算的轉換投影至相片上 P(x,y)點，本研究可以透過三角函數來計算出(Xc, Yc, Zc)。

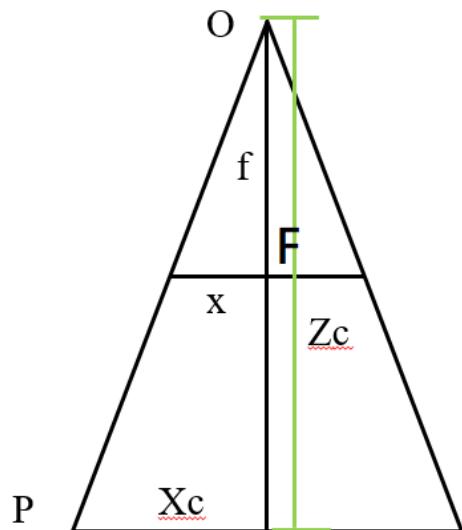


圖 3-8 照片平面的 x 軸

以三角函數計算：

$$\frac{x}{X_c} = \frac{f}{Z_c} \rightarrow X_c = \frac{Z_c}{f} x \quad (3.4)$$



$$\frac{y}{Y_c} = \frac{f}{Z_c} \rightarrow Y_c = \frac{Z_c}{f} y \quad (3.5)$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} \frac{Z_c}{f} x \\ \frac{Z_c}{f} y \\ Z_c \end{bmatrix} = Z_c \begin{bmatrix} \frac{1}{f} & 0 & 0 \\ 0 & \frac{1}{f} & 0 \\ 0 & 0 & \frac{1}{f} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.6)$$

### C. 把相機空間座標轉換為世界空間座標

如圖 3-9 所示將相機空間座標(Xc, Yc, Zc)轉換成世界空間座標(Xw, Yw, Zw)，分為旋轉角的轉換(R)與位移(T)，R 是一個 3x3 的矩陣，實際意義則記錄了 XYZ 軸旋轉的角度，T 的部分則是在相機空見座標系統中，相機空間座標中心與世界空間座標中心的

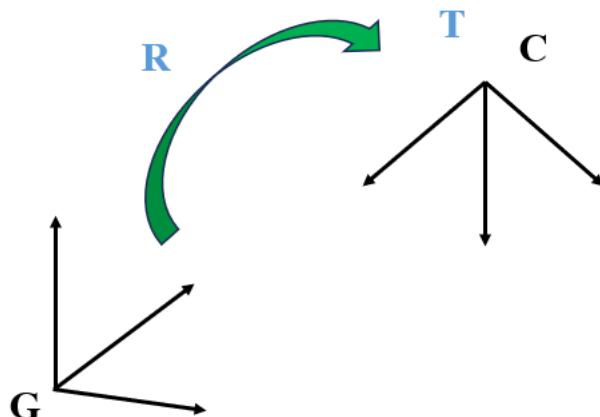


圖 3-9 相機空間座標轉換為世界空間座標

位移量，

以下為其數學式

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = R^T \begin{bmatrix} X_c - T_X \\ Y_c - T_Y \\ Z_c - T_Z \end{bmatrix} \quad (3.7)$$



將此數學式用使用 Python 程式寫出，再將點雲圖中所的的相機參數帶入，算出世界空間座標( $X_w, Y_w, Z_w$ )。

以下是使用多視立體高度測量的一些原則：

1. 使用高品質的立體攝像系統以獲得最佳效果。
2. 從不同的視點拍攝立體照片以提高測量精確度。
3. 使用專為多視立體高度測量設計的軟體應用程式。
4. 校準立體攝像系統，以確保照片對齊正確。

遵循這些，本研究可以從立體照片中獲得準確的高度測量結果。

這裡本實驗有特別觀察，在所有廠牌的水泥，其包裝、重量都是一樣大小的，所以不會有不同廠牌高度不一樣的問題，在長期重量的壓力下幾乎每包水泥的高度都一致，所以本實驗可以使用上述兩種方法測量高度，然後去計算包數。

## 第四章 系統實作與評估



### 4-1 倉庫點雲建立

#### 1. 使用 OpenSfM 開源軟體套件

在此實驗中，使用了 652 張照片並運用 OpenSfM 進行處理，以生成一個 3D 模型。

OpenSfM 是一個功能強大的開源軟體套件，能夠從這些 2D 照片中重建出精確而逼真的 3D 模型。這些照片涵蓋了對象的多個角度和視角，我遵循了一些最佳實踐，如使用高分辨率照片和確保相機穩定性，以獲得最佳的結果。

以下是對此軟體的簡介：

OpenSfM 是一個用於運動結構（Structure from Motion，簡稱 SfM）的開源軟體套件。

SfM 是一種電腦視覺技術技術，可以從一組 2D 照片中重建 3D 模型。OpenSfM 可以用於重建各種對象的 3D 模型，包括建築物、景觀，甚至人物。

要使用 OpenSfM 從 2D 照片中重建 3D 模型，需要：

- ① 收集一組要重建的對象的 2D 照片。
- ② 將照片轉換為 OpenSfM 可讀取的格式。
- ③ 運行 OpenSfM 處理照片。
- ④ OpenSfM 將生成對象的 3D 模型。

以下步驟提供了如何使用 OpenSfM 從 2D 照片中重建 3D 模型的詳細概述：

收集一組要重建的對象的 2D 照片。照片數量越多，3D 模型的品質越好。

將照片轉換為 OpenSfM 可讀取的格式。OpenSfM 可以讀取各種照片格式，包括 JPEG、PNG 和 TIFF 等。總之，OpenSfM 是一個強大的工具，可以從一組 2D 照片中重建 3D 模型。通過一點練習，本研究可以使用 OpenSfM 來創建各種對象的 3D 模型。



## 2. 使用 CloudCompare 免費開源軟體

CloudCompare 是一款用於 3D 點雲處理、配準和分析的免費開源軟件。它適用於 Windows、macOS 和 Linux。CloudCompare 最初由 Telecom ParisTech 和 EDF R&D 的 Daniel Girardeau-Montaut 於 2003 年開發。CloudCompare 是適用於各種應用程式的強大工具，包括：

3D 測量、3D 建模、3D 打印、3D 考古學、3D 地質、3D 地震學、3D 醫學成像，

CloudCompare 是一款免費的開源軟體，這意味著它可供所有人使用。

以下是 CloudCompare 的一些主要功能：

- 點雲處理：可用於以多種方式處理點雲，包括配準、對齊、過濾和分割。
- 點雲可視化：提供了多種可視化點雲的工具，包括 3D 渲染、顏色映射和點雲註釋。
- 3D 網格生成：可用於從點雲生成 3D 網格。
- 3D 表面重建：可用於從點雲重建 3D 表面。
- 3D 數據導出：可用於將 3D 數據導出為多種格式，包括 OBJ、PLY 和 STL。

總之，CloudCompare 是一款功能強大且用途廣泛的 3D 點雲處理工具。它是免費和開源的，這使它成為任何使用 3D 數據的人的寶貴工具。

## 3. 倉庫 3D 建模結果

### ① 手機拍照試建 3D 全景

為了初步建立倉庫的點雲圖，我決定使用手機拍攝照片並嘗試運行 Opensfm 程式。這個程式是一個基於結構化從影像重建三維場景的工具，它可以根據照片中的特徵點和相機的位姿估計，生成場景的點雲模型。

手機是人們日常生活中常用的工具，具有便攜性和易於操作的特點。使用手機拍照建立點雲可以隨時隨地進行，無需額外的設備或複雜的設置。手機相機的技術不斷進



步，像素和照片品質不斷提高，這些高品質的照片已經可以提供更準確、更真實的點雲數據。相比專業的三D照相機，手機拍照建立點雲的成本相對較低。一開始拍攝各個角落和視角的照片。為確保拍攝倉庫的不同部分，包括貨架、貨物、牆壁和天花板等。同時，也需要拍攝了倉庫外部的場地，以獲取更多的參考點。

拍攝完畢後，將這些照片複製到電腦中，並下載了 Opensfm 程式。照片例如圖 4-1 和圖 4-2 都是用手機拍的照片的特徵點和相機位置，一張是由裡面向外面拍，另一張是由外面向裡面拍，最後產的的結果，如圖 4-3 的點雲圖，這顯示將不同角度的相片放在一起計算，可以得出相片裡物體的相對位置，並且可以在點雲中呈現精確的位置。



圖 4-1 倉庫實體照片 1



圖 4-2 倉庫實體照片 2

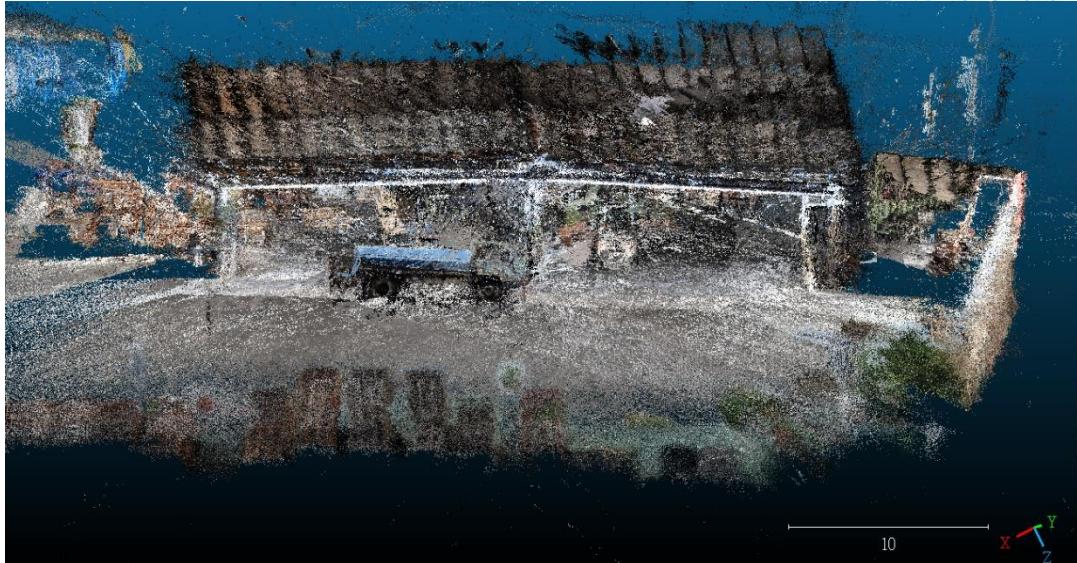


圖 4-3 倉庫點雲圖

## ② 拍攝倉庫內部影像

為了利用 3D 點雲來呈現倉庫內部的大小和佈局，本研究決定採三種方式去拍攝，包含：環繞倉庫內部拍攝方式、用蛇行的拍照方式以及環繞水泥棧板與推高機拍攝方式。我們用這三種拍攝方式來產生照片，拍攝完成後，我們將把照片傳輸到電腦中，使用 OpenSfM 進行處理，它會分析照片中的特徵點和相機位置，生成倉庫環境的點雲模型。

### • 環繞倉庫內部拍攝方式

首先，本研究先以直線進行環繞倉庫內部的方式去拍照片，拍攝路徑如圖 4-4，這是最方便也最容易拍攝照片的方式，順利建出點雲如下；可以很清楚清楚的看到倉庫內水泥放置的位置，與倉庫的空間利用，包含推高機可以行走的路線（圖 4-5、圖 4-6），是效率非常高的蒐集點雲資料方式。

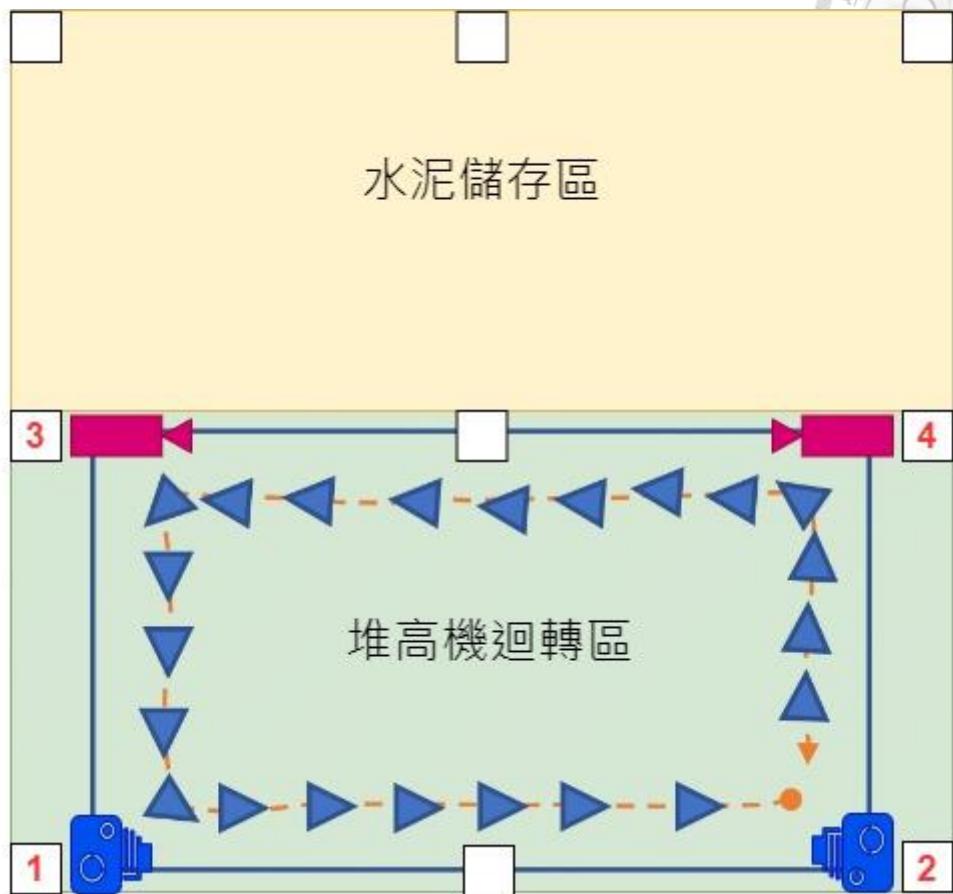


圖 4-4 倉庫拍照路線圖 1

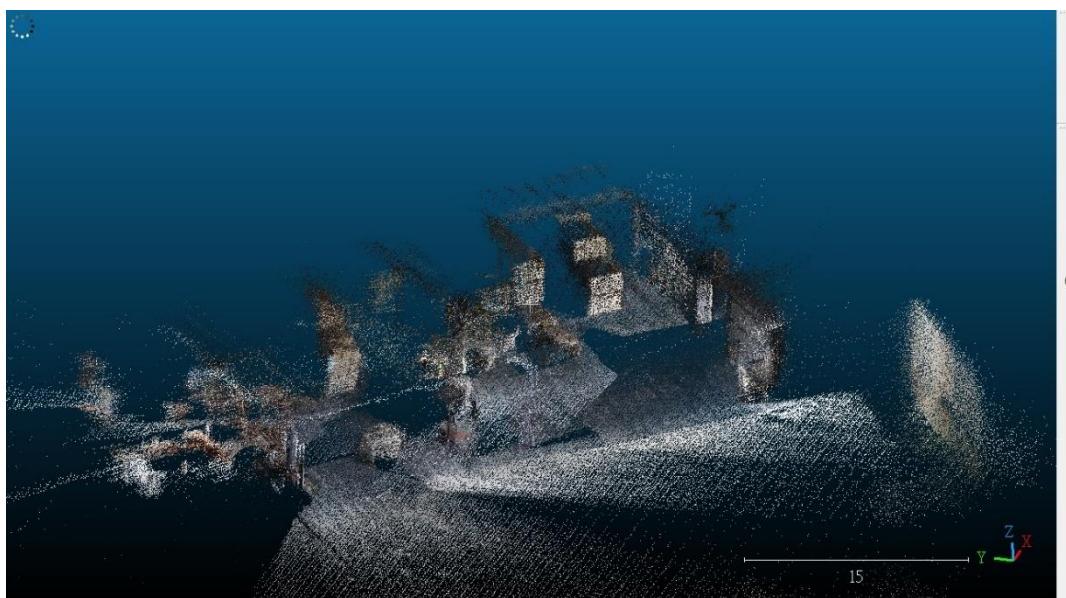


圖 4-5 倉庫內部點雲 1

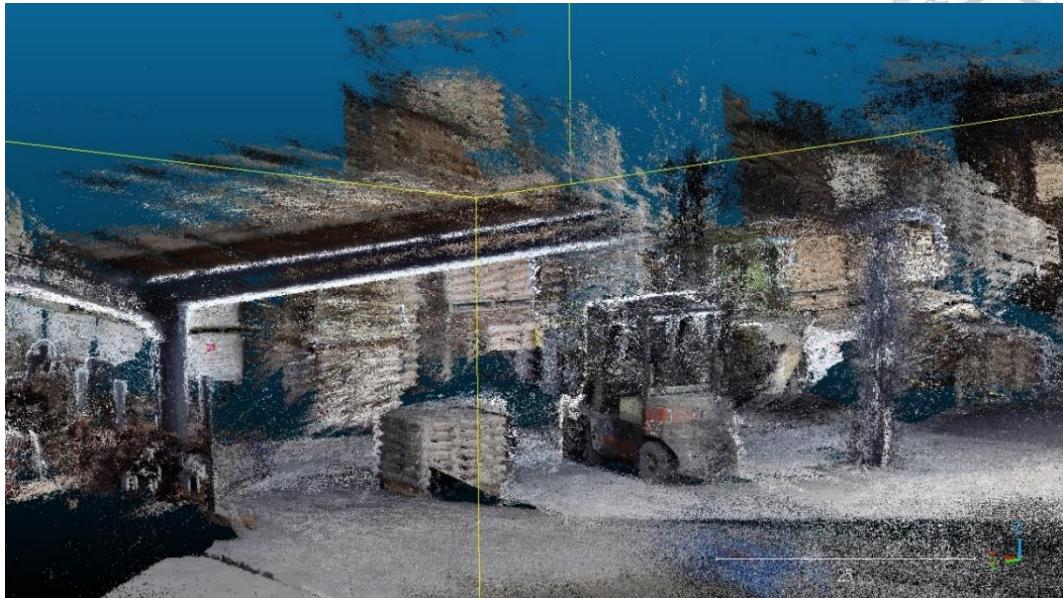
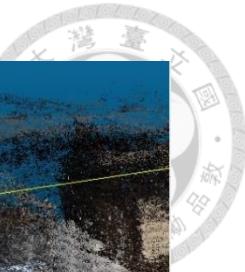


圖 4-6 倉庫內部點雲圖 2

- 用蛇行的拍照方式

本研究測試用不同的拍攝行徑路線去跑點雲，試圖找到拍攝照片的最佳方式，能讓計算出來的點雲更為清楚，這裡提出蛇行的拍照方式如圖 4-7，用此方式希望能將倉庫內的各種角度都能拍到，讓建出來的點雲更為清楚。實驗所得到新的點雲圖如下圖 4-8、圖 4-9果然比原本直線進行拍攝的方式得到倉庫的點雲圖更細緻。放大之後可以看到更清楚的相對空間位置。

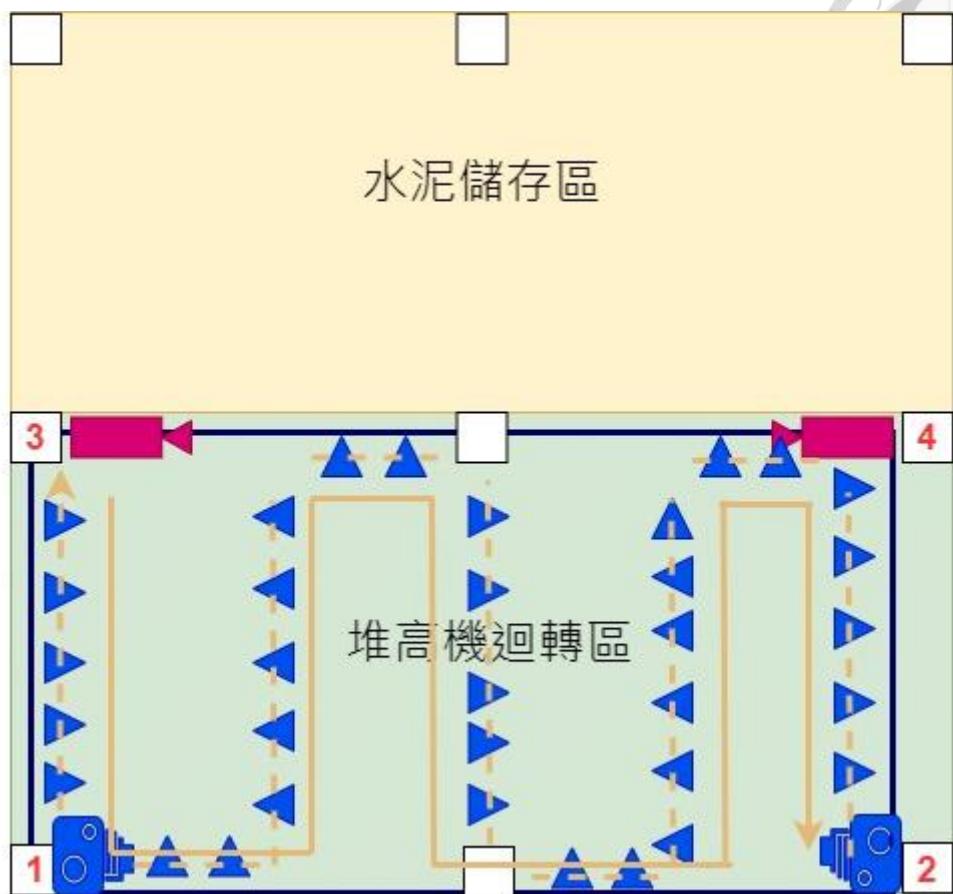


圖 4-7 倉庫拍照路線圖 2

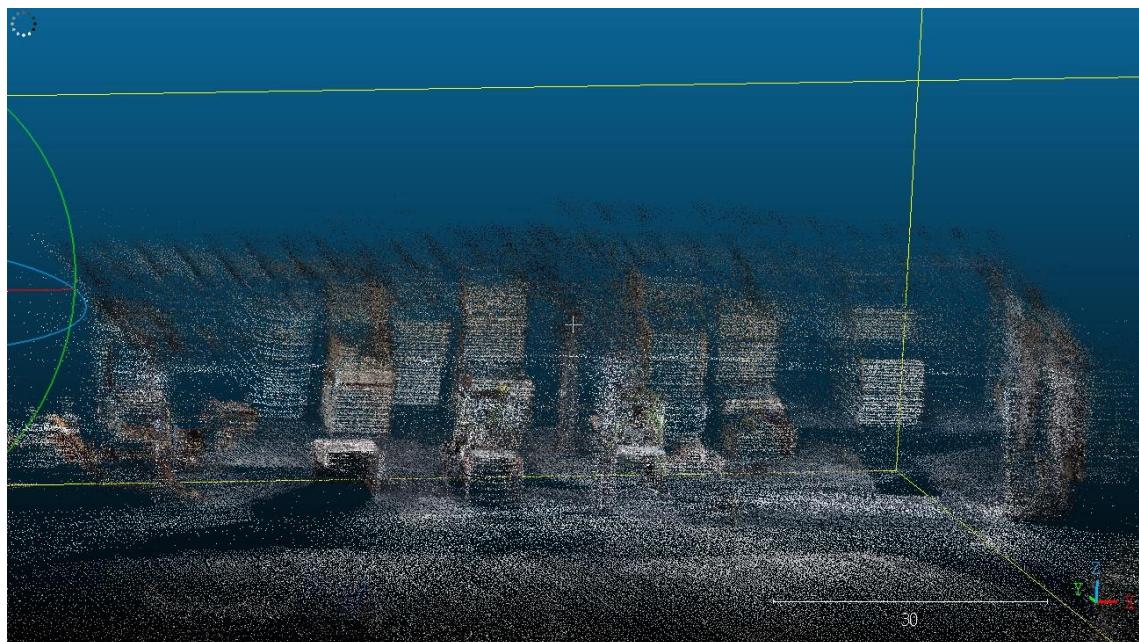


圖 4-8 倉庫內部點雲圖 3



圖 4-9 倉庫內部點雲圖 4

- 環繞水泥棧板與堆高機拍攝方式

由於本研究目的是要偵測水能的高度，本研究在實作過程中也對一個單位的棧板水泥做了點雲的建立，意即相機環繞著物件一圈一圈地拍攝，如圖 4-10。目的在了解水泥的高度與堆高機之間的高度比對。得到水泥與堆高機的 3D 點雲圖，如圖 4-11，利用此點雲圖得到堆高機高度、輪胎的高度以及水泥高度的比例，對上實踐測量高度的比是否一樣。未來在測量水泥高度可以用此數據及座標來做參考。這個點雲圖上的水泥物件的直線長度，各個頂點的座標，都是我們研究的方向。

。

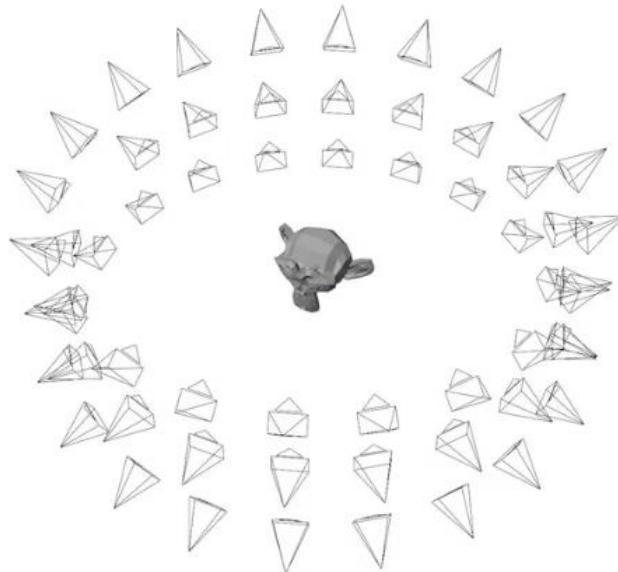


圖 4-10 以環繞方式拍攝水泥樣本

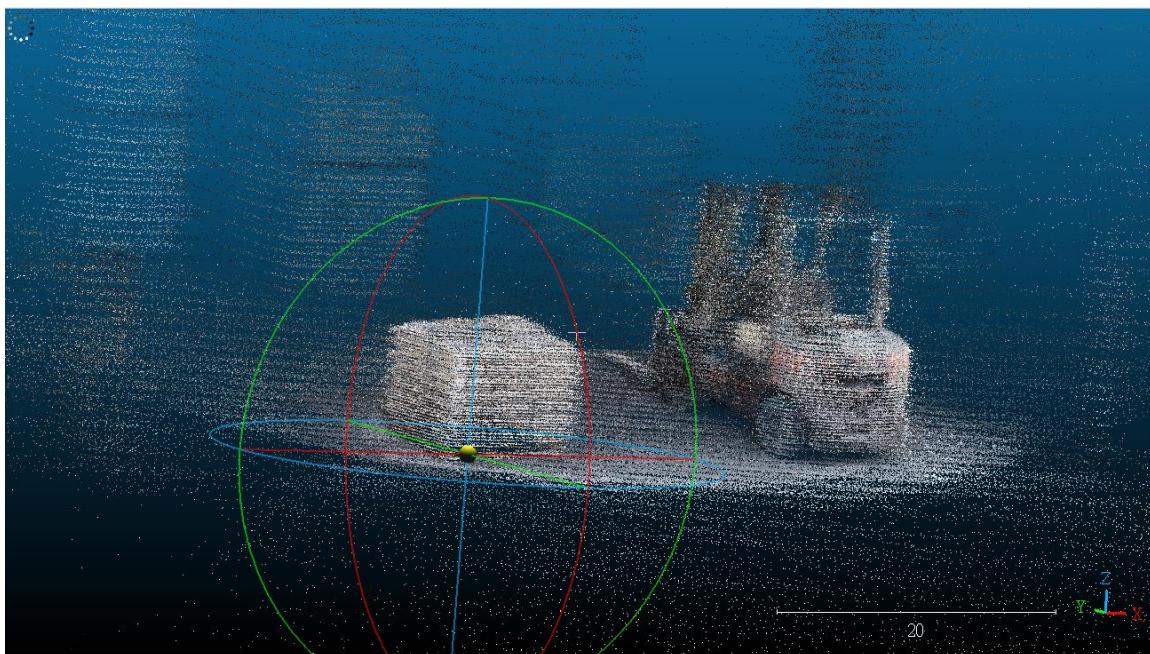


圖 4-11 水泥點雲圖

### ③ 加入 CCTV 圖片去建立點雲圖

通過將 CCTV 拍攝的照片添加到點雲地圖中，您可以更準確地呈現倉庫空間。相機參數將幫助您對照片進行對齊，並創建一個更連貫的點雲地圖。這對於安全、庫存管理

和品質控制等各種目的都非常有用。本研究用以上各種方式的點雲圖，用來分析得到本研究想知道的數據。相機標定是確定相機內部參數（如焦距、主點位置）和外部參數（如相機的位置和姿態）的過程。



OpenSfM 執行結果的重構文件是一個 JSON 文件，這是本研究所要的重要數據，這些參數只要跑一次點雲就可得到，後續可一直使用，不必再跑點雲圖了。包含以下數據：

- 相機：此對象包含有關重建中使用的相機的資訊。每個相機都有以下屬性：
  - projection\_type：相機使用的投影類型。可以是“透視”、“棕色”、“魚眼”或“等距柱狀”。
  - width：攝影機拍攝的照片的寬度。
  - height：攝影機拍攝照片的高度。
  - focal：相機的焦距。
  - k1：相機的第一畸變係數。
  - k2：相機的二次畸變係數。
- 鏡頭：此對象包含有關相機拍攝的鏡頭的資訊。每個鏡頭具有以下屬性：
  - image\_name：用於創建鏡頭的照片文件的名稱。
  - 旋轉矩陣：拍攝照片的相機的旋轉，這表示為角度軸向量。
  - 位移向量：拍攝鏡頭的相機的位移，這表示為向量。
  - camera：拍攝照片的相機 ID。
  - orientation：照片的 EXIF 方向標記，這可以是 1、3、6 或 8。
  - capture\_time：捕獲照片的時間，這表示為 UNIX 時間格式。
  - 頂點：從鏡頭中重建的頂點列表。
  - faces：從鏡頭中重建的面孔列表。

- scale：應用於重建頂點的比例因子。
- 共變異數：重建頂點的共變異數矩陣。
- merge\_cc：合併以創建重建頂點的連接組件的數量。



重建文件可用於可視化重建場景、創建場景的 3D 模型或執行需要訪問重建數據的其他任務。

## 4-2 建立影像辨識模型

### 1. 建立袋裝水泥資料集

水泥與堆高機訓練影像收集：

本研究需要用於驗證的照片數量取決於幾個因素，包括照片的品質、驗證任務的複雜性以及所需的準確性。一般而言，訓練所需的照片數量通常比測試所需的照片數量要多，一個常用的指標是將訓練數據的至少 10% 用於測試。因此，如果有 400 張訓練照片，則至少需要 40 張照片進行測試。如果希望達到更高的準確性，可能也需要更多的照片。所以。確定測試所需的照片數量的最佳方法是進行實驗。使用高品質的照片，照片的品質越好，測試的準確性就越高。以下是收集訓練和測試照片的一些注意事項：

- 確保照片的品質高。
- 照片應該光線充足且清晰。
- 照片應標記物體類別。

收集完照片後，本研究使用 YOLOv5 訓練工具來訓練模型，最終將生成能夠測試照片中物體的模型。本研究我總共拍了 581 張圖片，如圖 4-12 所示水泥與堆高機訓練影像收集，包含各種廠牌、各角度包裝、各個車載狀況來與預先標示其特徵，作為訓練模型的基礎資料。而水泥與堆高機測試影像收集共有 65 張照片，用來測試模型之用，如圖 4-132 所示，標誌的越清楚，未來辨識影像物件的檢測精度越高。



圖 4-12 YOLOv5 水泥與堆高機訓練影像收集

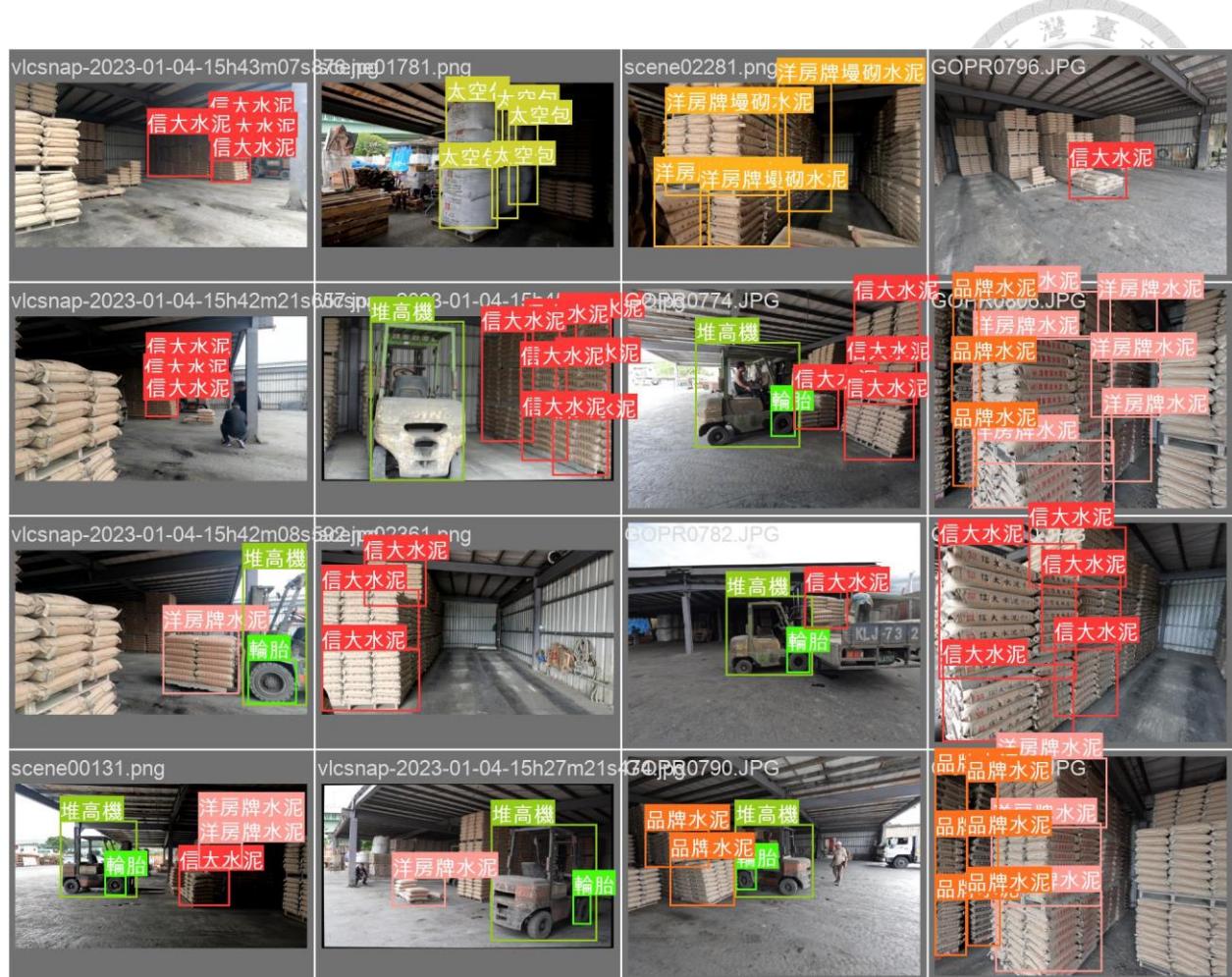


圖 4-13 YOLOv5 水泥與堆高機測試影像收集

## 2. 進行 YOLOv5 訓練

本實驗下了以下參數於 YOLOv5 模型訓練程式；

```
!python train.py --img 640 --batch 16 --epochs 120 --data custom_data.yaml --weights  
yolov5s.pt --cache
```

其定義如下：

- img: 照片尺寸，預設值為 640。
- batch: 批次大小，預設值為 8。
- epochs: 訓練時期的次數，預設值為 300。
- data: 數據文件的路徑，預設值為 data/coco128.yaml。
- weights: 權重文件的路徑，預設值為 weights/yolov5s.pt。

- cache: 是否緩存數據，預設值為 True。



該模型具有 214 個層、7038508 個參數、7038508 個梯度和 16.0 GFLOPs。

輸出表格中的第一列顯示層的編號。第二列顯示該層中的參數數量。第三列顯示實現該層的模組的名稱。第四列顯示傳遞給模組的引數。

模型以一系列卷積層開始，從輸入照片中提取特徵。這些層之後是一系列殘差塊，有助於提高模型的準確性。模型之後有一系列上採樣層，增加特徵圖的大小。這些層之後是一個最終的卷積層，用於預測照片中物體的邊界框和類別概率。

模型摘要顯示，該模型相對較大且複雜。然而，模型的大小和複雜性是實現其高準確度所必需的。總的來說，本研究提供的輸出顯示了 YOLOv5 模型是一個強大的物體檢測模型，該模型能夠從照片數據集中學習檢測物體，並能夠達到高準確度。

輸出結果展示了在包含 211 個物體的 50 張照片數據集上訓練 YOLOv5 模型的結果。該模型進行了 119 個訓練時期，並且輸出結果按照每個時期進行了顯示。輸出表格中的列代表以下內容：

- Epoch (時期)：時期的編號。
- GPU\_mem (GPU 記憶體)：模型使用的 GPU 記憶體量。
- box\_loss (框盒損失)：用於邊界框回歸的損失。
- obj\_loss (物體性質分類損失)：用於物體性質分類的損失。
- cls\_loss (類別分類損失)：用於類別分類的損失。
- Instances (實例)：照片中檢測到的物體數量。
- Size (大小)：照片的大小。



### 3. CCTV 影片運行此模型

本研究根據已經訓練了一個可以偵測倉庫水泥的 AI 模型。現在，我使用 CCTV 錄製的影片來執行偵測模型。該模型可以清楚地偵測出水泥廠牌、堆高機和輪胎的邊界框。首先，我將載入已經訓練好的模型並準備 CCTV 錄製的影片作為輸入。

在程式碼中，我載入了訓練好的模型，並通過 cv2.VideoCapture 從 CCTV 錄製的影片中讀取每一幀。然後，本研究對每一幀進行預處理，將其轉換為模型所需的格式。接下來，本研究使用載入的模型對輸入照片進行預測，並解析輸出結果。最後，我們將繪製邊界框的程式碼添加到相應的位置，在影片幀上繪製出檢測到的水泥廠牌、堆高機和輪胎的邊界框，如圖 4-14 所顯示。



圖 4-14 偵測到的水泥、推高機、輪胎等物件

本研究可以利用偵測到的水泥、推高機、輪胎等物件，透過單視立體高度測量或多視角立體高度測量來準確測量水泥的高度。



在單視立體高度測量中，我們使用單一攝影機，並利用場景中物體的已知尺寸。透過分析物體的透視和相對大小，本研究可以估算出水泥的高度。這種方法需要準確的物體偵測和校準良好的攝影機。

另一方面，多視角立體高度測量利用多個攝影機或視角從不同角度捕捉場景。透過在不同視角中三角測量偵測到的物體位置，本研究可以準確計算其三維位置。憑藉這些資訊，本研究可以精確確定水泥的高度。

這兩種方法各有優勢和考量。單視立體高度測量計算量較小，所需攝影機較少，更具成本效益。然而，它可能對光照條件和遮擋較敏感。多視角立體高度測量雖然較為複雜，但通過利用多個視角，提供更高的準確性和抗干擾的穩定性。

### 4-3 進貨出貨判斷

#### 1. 進出貨的判斷

本研究的系統在偵測到水泥後進行了一個判斷過程，並使用了 x 軸座標來辨別出貨或進貨。這裡有一個簡單的邏輯可以描述你們的流程：

偵測到水泥物件並取得其 x 軸座標。將水泥座標與參考線進行比較。如果水泥座標由小於參考線變成大於參考線，則判定為出貨，如圖 4-15 所示。如果水泥座標由大於參考線變成小於參考線，則判定為進貨，如圖 4-16 所示。這是一個簡單的邏輯，你可以根據你



圖 4-15 偵測到出貨



們的需求和系統細節進一步擴充和調整。請確保在實作過程中考慮到各種可能的情況和邊界條件，以確保準確判斷出貨或進貨的結果。



圖 4-16 偵測到進貨

這個過程相對簡單，但可以根據特定系統的需求進一步擴充和調整。例如，系統可能希望追蹤堆高機進入或離開特定區域的次數。在這種情況下，系統需要記錄堆高機每次進出區域時的 X 軸座標。系統可以利用這些資訊來確定已發生的出貨或進貨次數。

使用此過程時應考慮到一些研究限制。首先，此過程假設每次堆高機進出特定區域時，都是在出貨或進貨水泥。這個假設可能不適用於所有情況。例如，堆高機可能進出區域以卸貨或裝貨其他材料，而不僅僅是水泥。在這些情況下，系統需要能夠區分水泥和其他材料。另一個研究限制是此過程未考慮同一區域中可能有多台堆高機操作的情況。在這種情況下，系統需要能夠追蹤多台堆高機的 X 軸座標，以確定哪台堆高機正在出貨或進貨水泥。

儘管存在這些限制，使用 X 軸座標來辨別水泥出貨或進貨的過程是一個有價值的工具，可提高水泥管理系統的效率和準確性。通過追蹤水泥的移動，系統可以更好地了解水泥的使用情況，並對未來的進貨做出明智的決策。



## 2. OCR 光學字元識別 OCR 光學字元識別(Optical Character Recognition , OCR )

使用 OCR 技術來檢測水泥包裝上的字體，以區分不同的品牌。由於圖片可能不一定  
是正面拍攝的，這對中文字的檢測帶來了一些困難。因此，如果無法檢測到字體，需要  
將圖片進行 90 度的旋轉，甚至可能需要旋轉多次，以幫助檢測。

使用 OCR 技術的目的是為了輔助 YOLO 模型在檢測物體廠牌時的工作。通過檢測水泥包  
裝上的字體，可以識別不同品牌的產品。這樣的過程是為了增強 YOLO 模型在廠牌檢測  
方面的準確性。通過使用 OCR 技術來輔助或檢測字體，如圖 4-17 所示，可以提供更多關  
於物體廠牌的資訊，從而改善 YOLO 模型的識別能力。

本研究在訓練模型的初期，偵測廠牌的準確度還不夠，我們在推高機上架設一台 360 相  
機，使用 OCR 來輔佐水泥廠牌偵測的不足。不過，隨著模型訓練的準確率越來越高，後  
續已經不需要用到 OCR 技術了。



圖 4-17 偵測到水泥包裝上的文字



## 4-4 水泥高度計算

### 1. 單視立體高度測量

本研究根據由第三章系統架構與設計之 3-3 影像辨識模型設計中，參考圖 3-5 交叉比(ross ratio)示意圖，在本研究這個實驗中 AD 為堆高機高度，CD 為輪胎高度，BD 為水泥標高高度，參考公式(3.3)。

- ① 在這個實測中本研究已知；AD 堆高機高度為 210 公分，CD 輪胎高度為 65 公分。
- ② 我用尺測量相片的像素；ac 為 93 像素、bd 為 68 像素、bc 為 26 像素、ad 為 124 像素。

$$CR(a,b,c,d) = (ac * bd) / (bc * ad) = 93 * 68 / (26 * 124) = 1.962$$

所以可推算水泥高度為；

$$\begin{aligned} BD &= \frac{\frac{(ac * bd)}{bc * ad} * CD * AD}{\frac{(ac * bd)}{AD - CD}} = \frac{CR(a,b,c,d) * CD * AD}{\frac{CR(a,b,c,d)}{AD - CD} * AD - 1} \\ BD &= \frac{\frac{1.962}{145} * 65 * 210}{\frac{1.962}{145} * 210 - 1} = 100.297 \end{aligned}$$

本研究用此公式去驗證，得出圖中水泥高度為 100.297 公分，實際測量現場的高度是 100 公分，準確率非常高。本實驗經過實地測量得到每包水泥高度是 12.5 公分，所以本實驗可得到  $(100.297 / 12.5 = 8.02376 \approx 8)$  高度是 8 包高，我們已知棧板上一層是 6 包，所以一個棧板計算出是 48 包。由此驗證出本實驗可以用此方法計算出水泥包數，未來我們也可以用此方法撰寫 python 程式自動計算。

由於本實驗的照片中的消失點與消失線都是使用手工去畫，但由此計算準確度已經很高了。水泥現場的量測也會因為堆疊與擠壓高度有些微差距，儘管如此，計算還是很



準確的。目前，本研究是在照片中手動繪製消失點和消失線。將來，希望使用影像辨識的方式，用一些工具來自動化此執行過程。

## 2. 多立體視角高度測量

我對相機用於精確 3D 測量的令人難以置信的準確性和建模能力感到驚訝，特別是因為我剛剛開始探索這個領域。雖然網上有大量理論文章，但只有少數實用文章展示了這些理論在實際問題中的應用。攝影測量領域是理論與實踐之間的鮮明對比。這主要是因為攝影測量中的相機可以達到極高的精度，能夠測量低至每個像素微米級的細節。為了彌合理論與實踐之間的差距，需要復雜的校準算法來補償這些差異。實際實施還涉及對硬件規格、參數調整、掌握校準技術以捕獲準確的照片。

本研究用拍攝的相片建庫的點雲後重新加入 CCTV 所拍攝的相片一起加入，使用 openSfM 建立新的點雲圖，如圖 4-18 所示。會得到 CCTV 的參數及座標，用這個來計算後面的影像所在的座標。

我們根據由第三章系統架構與設計之 3-3 影像辨識模型設計中的 5.多立體視角高度測量中所述，本研究可以將相片的座標利用建立點雲取得 CCTV 的相機參數，然後依照 3 步驟轉換座標；

① 相片座標( $u, v$ )轉換為相機空間座標中 P 點的座標( $u_c, v_c, f$ )

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} - \begin{bmatrix} u_0 \\ -v_0 \end{bmatrix}$$

② 相片座標轉換為相機空間座標

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} \frac{Z_c}{f} x \\ \frac{Z_c}{f} y \\ Z_c \end{bmatrix} = Z_c \begin{bmatrix} \frac{1}{f} & 0 & 0 \\ 0 & \frac{1}{f} & 0 \\ 0 & 0 & \frac{1}{f} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$



③ 把相機空間座標轉換為世界空間座標

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = R^T \begin{bmatrix} X_c - T_X \\ Y_c - T_Y \\ Z_c - T_Z \end{bmatrix}$$

1. 本研究從 OpenSfM 點雲資料中找到 reconstruction 檔案到 CCTV 相機參數如下；

其代碼是描述相機的 JSON。具有以下字段：

- projection\_type (相機使用的投影類型)：。在本例中，投影類型為 perspective。
- width (照片的寬度)： 3840 (以像素為單位)。
- height (照片的高度)： 2160 (以像素為單位)。
- focal (相機的焦距)：: 0.8280991993834689 (以像素為單位)。
- k1 (第一徑向畸變係數)： -0.00782496600079536。
- k2 (第二徑向畸變係數)： -0.0006648348429029573。
- 由 GLscene02251.png 資訊得知；

rotation: [1.506410335130914, -0.48622784176665107, 0.11291310323342098],

translation: [-22.45745824096986, 55.08099607695289, 426.2839224491021],

rotation 和字段 translation 一起代表相機在世界中的姿態。

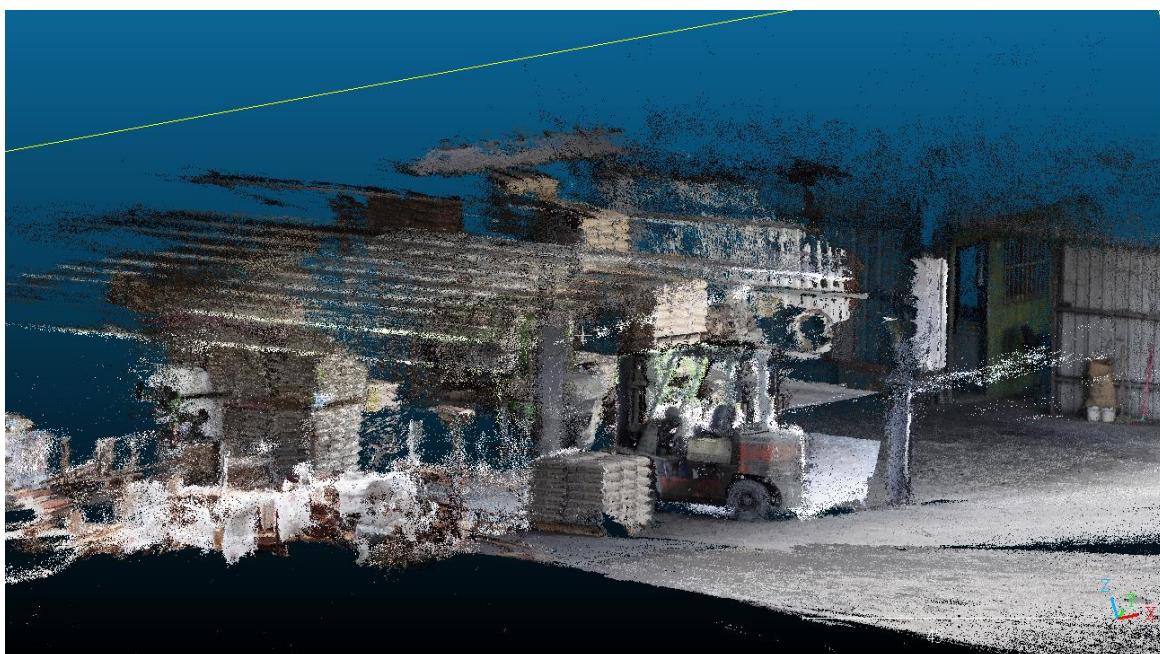


圖 4-18 加入 CCTV 的點雲圖



2. 由影像中找出水泥邊界框的座標  $\min(u_1, v_1)$  與  $\max(u_2, v_2)$ ，然將它轉換世界空間的座標  $(x_1, y_1, z_1)$  與  $(x_2, y_2, z_2)$ ，如圖 4-19 所示，水泥的座標分別為邊界框的左上角為  $P1(x_1, y_1, z_1)$ ，右下角為  $P2(x_2, y_2, z_2)$ ，推高機的輪胎座標分別為左上角  $P3(x_3, y_3, z_3)$ ，右下角為  $P4(x_4, y_4, z_4)$ 。



圖 4-19 相片座標圖

影像座標轉換成世界空間座標，計算結果如下：

$$P1(384,495) \rightarrow (x_1, y_1, z_1) (0.332933, 0.273682, 426.283922)$$

$$P2(495,584) \rightarrow (x_2, y_2, z_2) (0.600806, 0.41347, 426.283922)$$

$$P3(609,675) \rightarrow (x_3, y_3, z_3) (1.001112, 0.584599, 426.283922)$$

$$P4(675,745) \rightarrow (x_4, y_4, z_4) (1.11868, 0.67404, 426.283922)$$

在這個實測中我們已知輪胎高度為 65 公分，所以：

$$y_2 - y_1 = 0.41347 - 0.273682 = 0.139788$$

$$y_4 - y_3 = 0.67404 - 0.584599 = 0.089441$$

$$\text{水泥高度} = (y_2 - y_1) / (y_4 - y_3) * 65 \text{ 公分} = 101.589 \text{ 公分}$$

其結果  $(101.589 / 12.5 = 8.12712 \approx 8)$  與單視立體高度計算一樣是 8 包高。



## 4-5 訓練模型數據

輸出的前幾行顯示了訓練開始前的結果。模型尚未能夠檢測到任何物體，且各損失非常高。隨著模型的訓練，損失開始下降，檢測到的物體數量開始增加。在訓練結束時，模型能夠檢測到照片中的大多數物體，而且損失相對較低。

代表以下內容：

- Class (類別)：物體類別的名稱。
- Images (照片)：數據集中包含指定類別物體的照片數量。
- Instances (實例)：數據集中指定類別物體的數量。
- P (精確率)：指定類別的模型對於該類別的精確預測比例。精確率是指預測物體中實際存在於照片中的比例。
- R (召回率)：指定類別的模型對於該類別的召回比例。召回率是指模型預測出的實際物體數量與數據集中該類別物體數量之比。
- mAP50 (平均準確率)：指定類別的模型在 IoU 臨界值為 50% 時的平均準確率 (mAP)。

mAP50 指標是衡量模型在檢測物體方面的準確度的一個指標。它通過平均 IoU 臨界值為 50% 時的精確率和召回率來計算。mAP50 分數越高，模型的性能越好。在這個案例中，模型實現了 0.691 的 mAP50 分數，這是一個不錯的結果。

表 4-1 YOLOv5 訓練結果顯示了在包含 277 個物體的 65 張照片數據集上訓練 YOLOv5 模型的結果。各列模型的整體結果顯示在“all”行中。該模型的精確率為 0.868，召回率為 0.705，mAP 為 0.754。這意味著模型在檢測到的物體中有 86.8% 的正確預測，它能檢測到照片中 70.5% 的物體，並且在 IoU 臨界值為 50% 時的平均準確度為 75.4%。



每個單獨類別的結果顯示模型對於不同類別的表現不同。例如，對於 A 廠牌水泥，模型的精確率為 0.937，召回率為 0.57，但對於 D 廠牌水泥，僅有 0.87 的精確率和 0.60 的召回率。這意味著模型在檢測 A 廠牌水泥方面表現優於 D 廠牌水泥。

表 4-1 YOLOv5 訓練結果

| Class | Images | Instances | P     | R     | mAP50 |
|-------|--------|-----------|-------|-------|-------|
| all   | 65     | 277       | 0.868 | 0.705 | 0.754 |
| A廠牌水泥 | 65     | 52        | 0.937 | 0.57  | 0.626 |
| B廠牌水泥 | 65     | 39        | 0.866 | 0.665 | 0.785 |
| C廠牌水泥 | 65     | 44        | 0.778 | 0.659 | 0.696 |
| D廠牌水泥 | 65     | 5         | 0.87  | 0.6   | 0.73  |
| E廠牌水泥 | 65     | 49        | 0.718 | 0.633 | 0.579 |
| 輪胎    | 65     | 38        | 0.93  | 0.895 | 0.926 |
| 堆高機   | 65     | 50        | 0.979 | 0.912 | 0.934 |

由平均準確率來看，E 廠牌水泥的辨識率比較低，本研究歸納了原因。E 廠牌是太空包，其型態是不一樣的，他不像 A~D 廠牌水泥都是固定大小紙袋裝的水泥，他沒有其他可辨識的文字與包裝 logo，太空包是一噸的散裝水泥，每天出貨不多，只有個位數字，所以辨識率最低是能理解的，準確度只有 0.579。除此之外，其他廠牌都有不錯的準確率。總體而言，結果表明 YOLOv5 模型對於袋裝水泥檢測是一個不錯的選擇。

#### 4. 訓練模型精確度

在 YOLOv5 訓練中，混淆矩陣 (Confusion Matrix)、F1 分數 (F1 score)、P 曲線 (P-curve)、PR 曲線 (PR-curve) 和 R 曲線 (R-curve) 被用於評估模型的性能，(圖 4-20、圖 4-21、圖 4-22)所示，。

混淆矩陣是用於評估分類模型性能的表格，顯示正確分類的實例數量以及被錯誤分類的實例數量。混淆矩陣可以幫助識別模型所犯的錯誤類型。

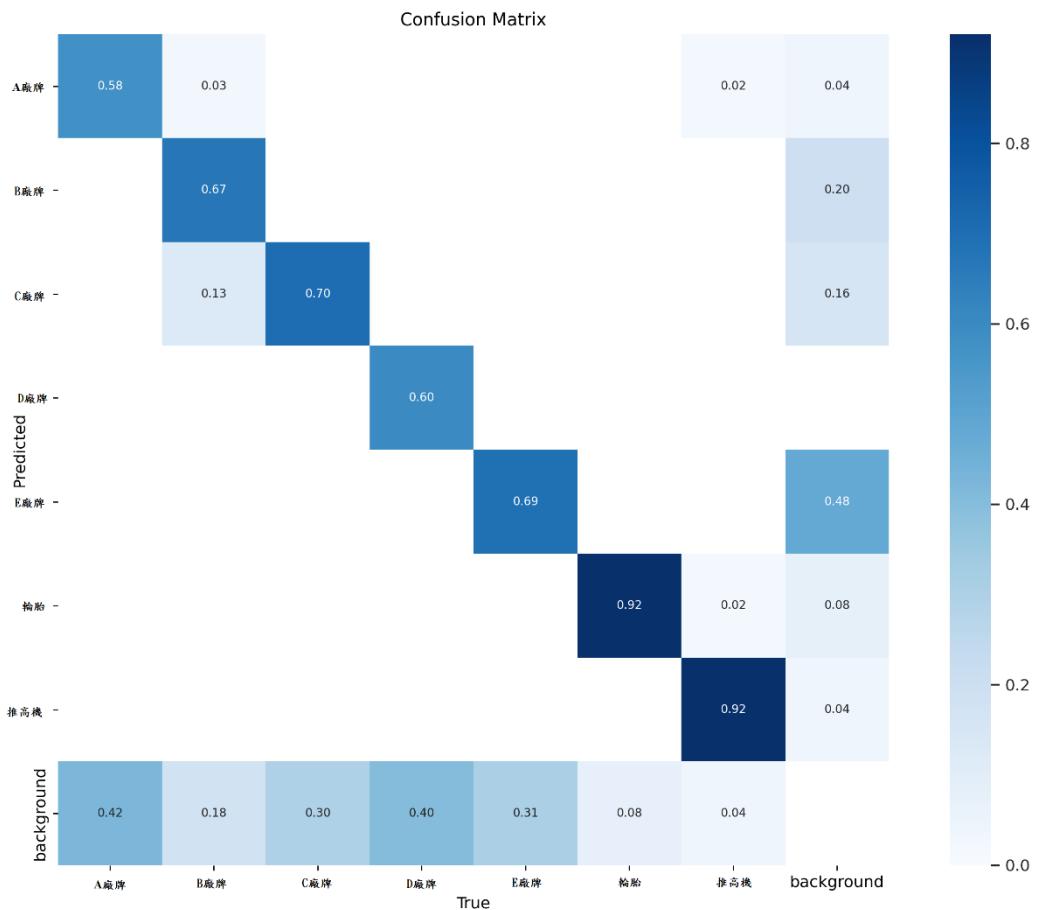


圖 4-20 混淆矩陣 (Confusion Matrix) 圖



表 4-2 模型所犯的錯誤類型

| 預測   | 實際 |
|--|----|
| 正例   | TP |
| 負例   | FP |
| TP 代表真正例 (True Positive)，表示模型正確地將實例分類為正例。  |    |
| FP 代表假正例 (False Positive)，表示模型將實例錯誤地分類為正例。 |    |
| TN 代表真負例 (True Negative)，表示模型正確地將實例分類為負例。  |    |
| FN 代表假負例 (False Negative)，表示模型錯誤地將實例分類為負例。 |    |

① F1 分數是評估分類模型準確性的一個指標。

- A. 它通過計算精確度和召回率的調和平均值來得到。相比僅使用準確度，F1 分數更全面地考慮了精確度和召回率。
- B.  $F1\text{ 分數} = 2 * (\text{精確度} * \text{召回率}) / (\text{精確度} + \text{召回率})$

② 精確度是預測為正例的實例中實際為正例的比例。

$$\text{精確度} = \text{TP} / (\text{TP} + \text{FP})$$

③ 召回率是實際為正例的實例中被預測為正例的比例。

$$\text{召回率} = \text{TP} / (\text{TP} + \text{FN})$$

④ P 曲線是精確度和召回率之間的曲線。P 曲線可用於識別分類的最佳臨界點值。最佳臨界點值是最大化 F1 分數的臨界值。

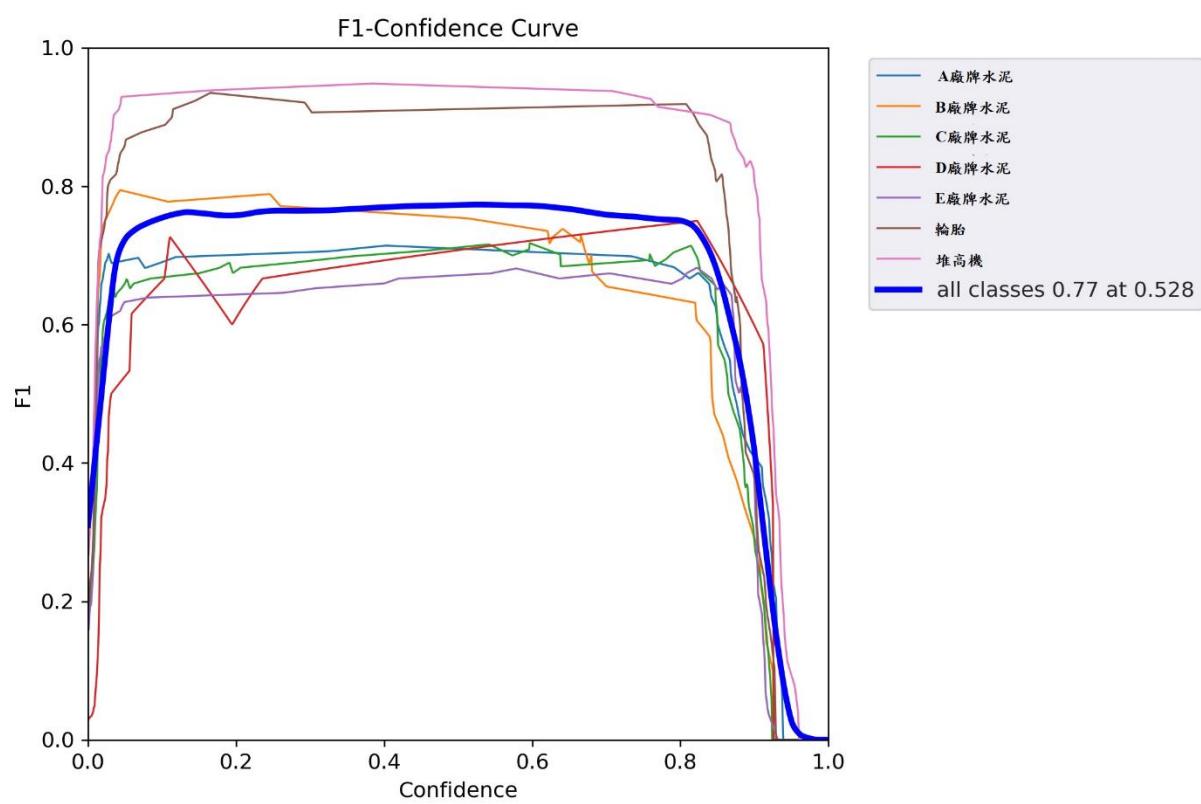


圖 4-21 F1 分數 (F1 score) 圖

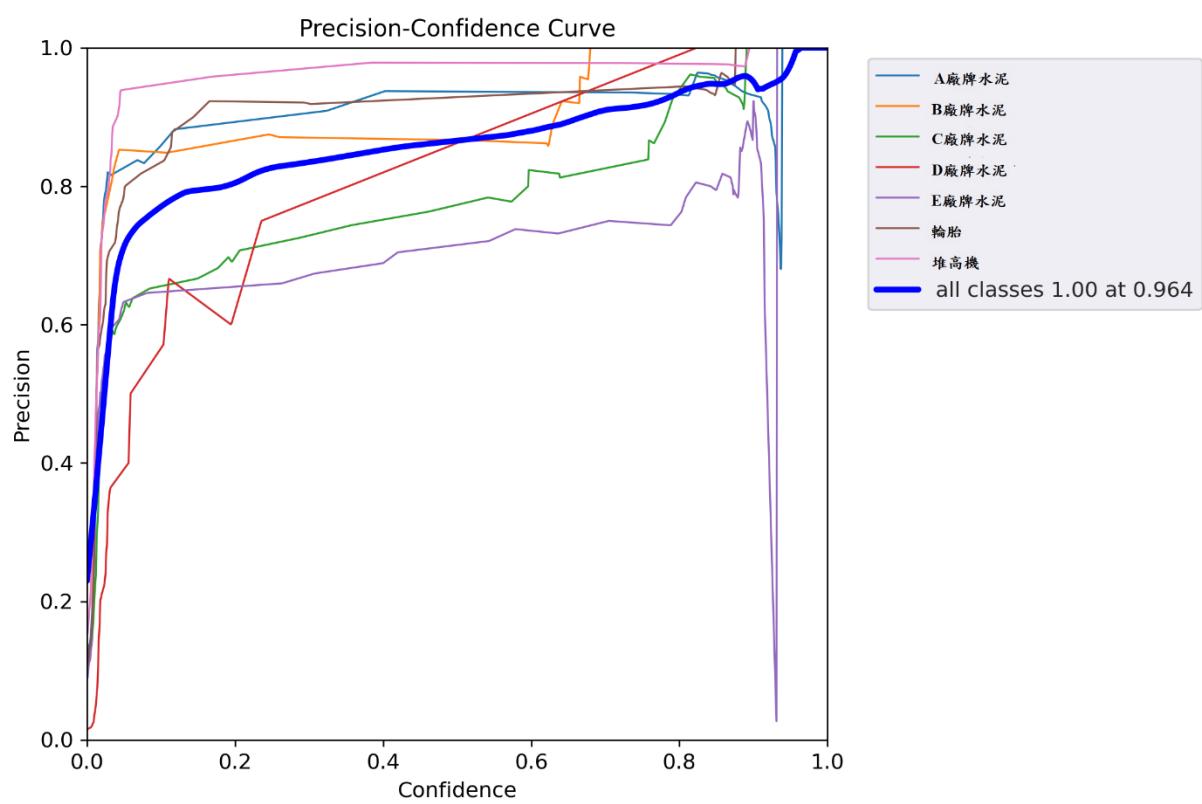


圖 4-22 P 曲線 (P-curve) 圖

- ⑤ 如圖 4-23 所示，R 曲線是召回率和 F1 分數之間的曲線。R 曲線也可用於識別分類的最佳臨界值。最佳臨界點是最大化 F1 分數的臨界值。

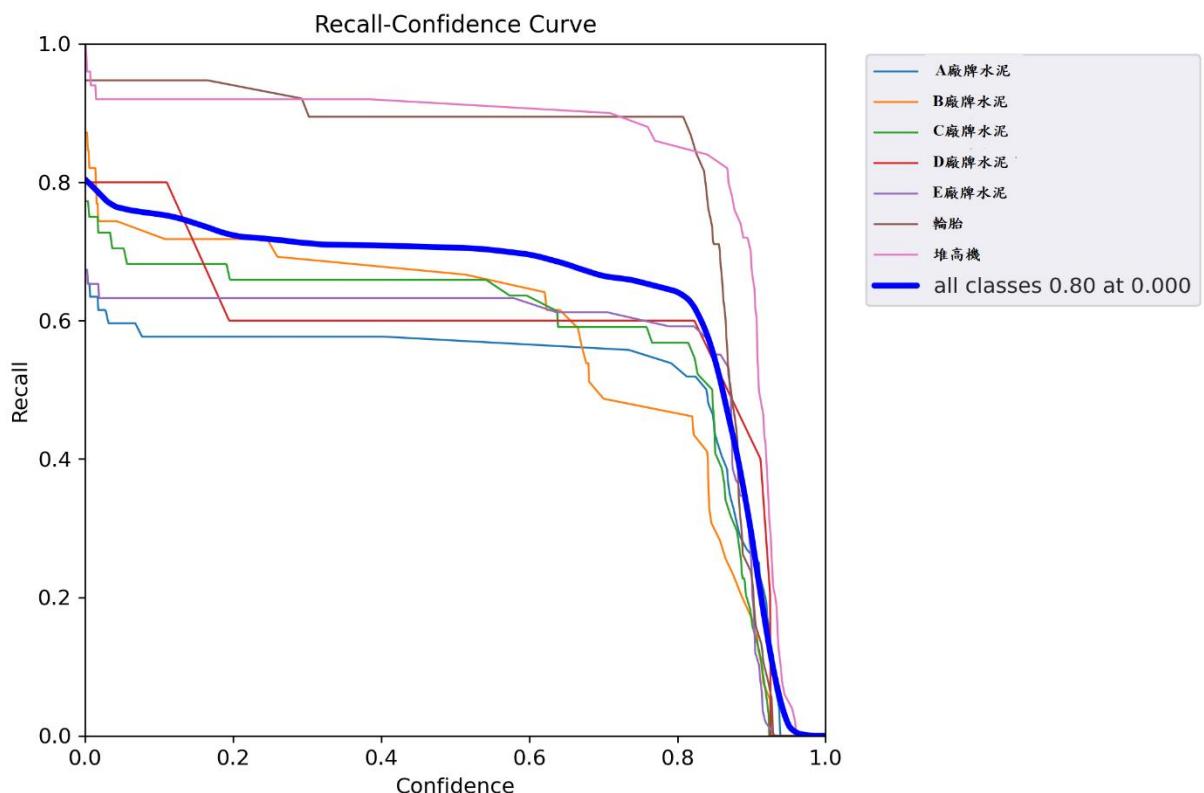


圖 4-23 R 曲線 (R-curve) 圖

- ⑥ 如圖 4-24 所示，PR 曲線是精確度與召回率之間的曲線。PR 曲線可用於識別分類的最佳臨界值。最佳臨界值是最大化

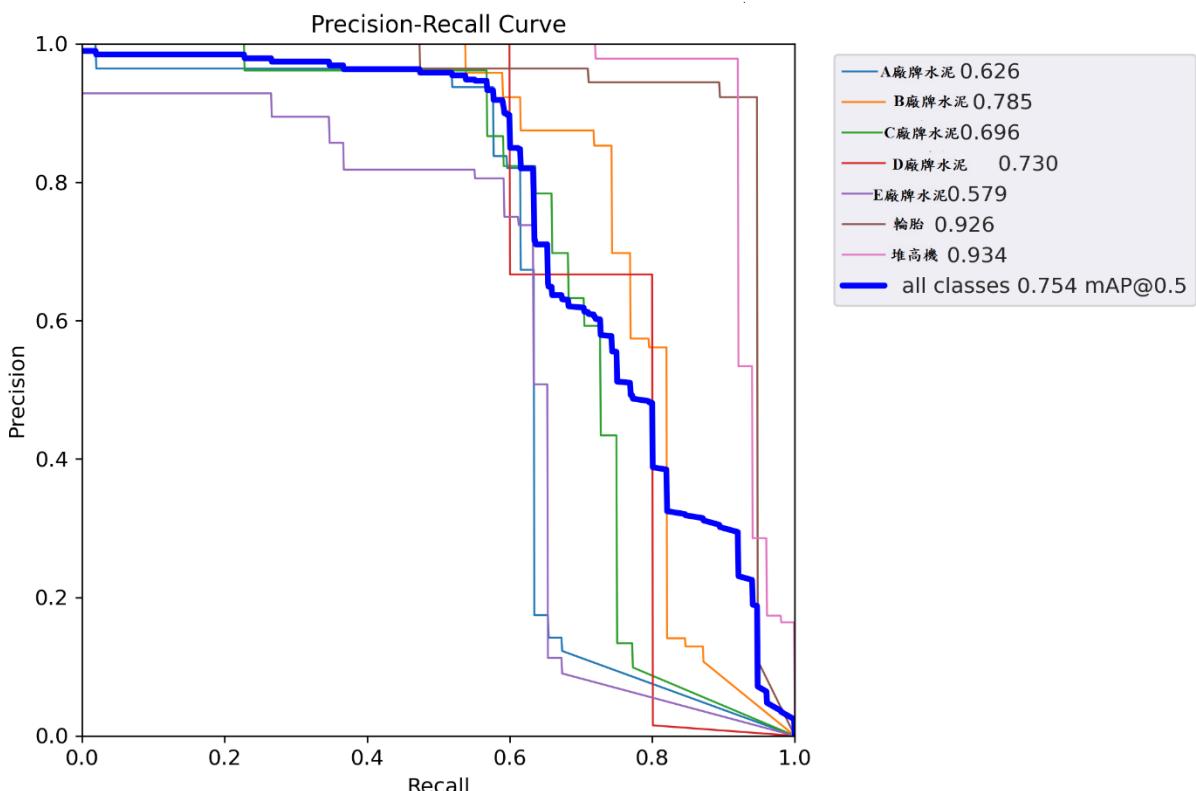


圖 4-24 PR 曲線 (PR-curve)

YOLOv5 訓練的結果通常報告 F1 分數、P 曲線、PR 曲線和 R 曲線。F1 分數是評估 YOLOv5 模型性能最常用的指標。P 曲線、PR 曲線和 R 曲線較少使用，但可以幫助識別分類的最佳臨界值。

透過這些指標來改進 YOLOv5 模型性能的一些示例：

1. 如果模型產生了大量的假陽性，可以增加臨界值以減少假陽性的數量。
2. 如果模型產生了大量的假陰性，可以降低臨界值以減少假陰性的數量。
3. 如果模型在某個特定類別上的表現不好，可以增加該類別的權重。
4. 通過使用這些指標，我們可以改進 YOLOv5 模型得到更好的結果。



## 4-6 系統優點與限制

本研究之系統具有多個優點。首先，它是自動化的，這意味著它可以在無需人工干預的情況下辨識出貨和進貨。這節省了時間和人力成本，同時提高了效率。其次，該系統快速且準確。它可以快速而準確地判斷水泥是否正在出貨或進貨，從而減少了人為判斷的主觀性和誤差。第三，該系統具有即時性。它可以實時監控水泥的運輸狀態，並提供即時的出貨或進貨資訊。最後，該系統提高了效率。它幫助管理人員追蹤和管理水泥的運輸，使整個水泥管理流程更加高效，有助於減少庫存並避免供應鏈中斷的風險。

然而，該系統也存在一些限制。首先，它需要一條參考線。該參考線用於比較水泥的 x 軸座標，如果參考線的位置不準確或頻繁變化，可能會影響系統的正確性。其次，該系統可能無法區分同一區域中的多個物件。如果同一區域中有多台堆高機操作，或者同一區域中有多个水泥物件，這可能會成為問題。第三，該系統需要合適的硬體設備和運算資源。另外一個限制是本研究目前適用高度去計算每次進出貨的包數，實際上每次進出貨並非都是完整的一個棧板的數量，也並非所有的一堆水泥都一樣高，本研究並沒有討論欠包的情形。以上都是我們未來要加克服的問題，以及未來必須再發展的部分。

再者，雖然本研究可以通過 1. 單目立體測高和 2. 多視角測高來計算水泥的高度，但是多立體視角測高需要考慮角度的問題。這是因為 YOLOv5 檢測到的邊界框是平面的，如圖 4-25 所示，它們的高度往往會大於實際的 3D 高度。這種局限性來自於算法本身，需要在未來解決，以便準確計算。



圖 4-25 YOLOv5 檢測到的邊界框圖



## 第五章 結論與建議

### 5-1 研究結果總結

在本研究中，我們成功地採用影像分析技術來計算水泥倉庫的進貨和出貨動態。透過物件偵測技術，本研究能夠即時辨識推高機上的水泥品牌，並有效地進行庫存查詢。此外，即使只有單支監視器，本研究也發展了一種交叉比對的方式來計算水泥的高度。而在多支監視器同步的情況下，本研究能夠將拍攝的影像迅速轉換成實際空間的座標，進一步精確地計算水泥的高度。

這些成果確認了本研究所提出的系統開發的可行性。透過影像分析技術，本研究能夠快速且準確地獲取水泥倉庫的進出貨資訊，使得倉儲管理更加高效和自動化。同時，物件偵測技術的應用使得本研究能夠辨識不同廠牌的水泥，有效地儲存和追蹤庫存，提供準確的數據供日後查詢使用。

在單支監視器情況下，本研究提出了交叉比對的方式來計算水泥的高度。這種方法充分利用影像之間的對應關係，提高了高度計算的精度。而在多支監視器同步的情況下，本研究成功地將拍攝的影像轉換成實際空間的座標，這一方法不僅計算迅速，而且精確度也得到有效提升。

然而，本研究也意識到本研究還存在一些限制。例如，在多支監視器同步的情況下，尚需更進一步解決不同監視器間的影像校準問題，以確保計算出的水泥高度的一致性。另外，在使用物件偵測技術時，部分特殊情況下可能會有誤判的情況出現，這需要進一步改進和優化算法。

未來，我們將持續改進本研究的系統，解決存在的限制和問題。進一步的研究方向



包括提高監視器間的影像校準準確性，加強物件偵測算法的穩定性，並進行更多實地測試，以驗證系統在實際應用中的效能。

總結而言，本研究的研究成果證明了利用影像分析技術和物件偵測技術進行水泥倉儲管理的自動化的可行性。這項研究為營建倉儲管理系統的現代化升級提供了有力的支持，並為未來相關領域的研究和應用提供了有價值的參考。

## 5-2 未來研究方向建議

未來的研究方向可以著重於以下四個主題：將由本實驗的素材庫影片轉成線上串流 LIVE 監控器影片、點雲計算時間的減少與即時、多輛堆高機同時運行的情形、以及倉庫其他商品的測量。

首先，將由本實驗的資料庫影視轉成線上串流 LIVE 監控器視覺是一個很有潛力的研究方向。這樣的系統可以讓使用者透過網絡即時監控水泥輸送的情況，而不僅僅是觀看已發生的影片。透過將資料庫影片轉成線上串流，使用者可以即時了解水泥的輸送情況，一步提高管理的效率和及時性。

第二個研究方向是減少一點雲計算的時間並實時性。在目標之前的系統中，點雲計算可能需要相當長的時間，這可能限制了系統的實時性。未來的研究可以探索優化點雲計算方法，減少計算時間，以實現更高的即時性。這將使系統能夠更快地處理大量的資料點，提出高知識別和判別的效率。

第三個研究方向是處理多堆高機同時運行的情況。目前的系統可能無法正確區分多個堆高機同時運行的情況。這可能會導致混淆和錯誤的識別結果。未來的研究可以著眼於開發更先進的演算方法和技術，以分區和追蹤多個堆高機的運動軌跡和操作行為，從



而實際發現對多個堆高機的同步監控控制和管理。

最後，研究人員可以擴展系統的功能，實際對倉庫其他商品的舉例。除了水泥之外，倉庫可能還存在放著其他種類的商品。未來的研究可以將系統擴展到其他商品的偵測和監控，以實現更全面的庫房管理。

並結合相關的物理知識和算法，以確保庫內不同商品的類型和位置。這將有助於庫房管理人員更好地追蹤跟蹤和管理倉庫庫存，提供高庫存的準確性和效率。

除了上面提到的四個研究方向，還有其他潛在的擴展方向值得探索。響應速度，以響應不同環境條件下的挑戰。同時，也可以研究如何結合機器學習和深度學習技術，以提高系統的自主化和智能化水平穩。此外，進一步研究系統的可擴展性和可靠性，應對大型模型庫環境和長期運營的需求。

總結而言，未來的研究方向包括將資料庫影視轉成線上串流 LIVE 監控器影片、減少點雲計算時間與即時、多堆高機同時運行的情況、以及倉庫其他商品的測算。這些研究方法將逐步提出提升水泥管理系統的效率和準確性，並為倉庫管理提供更全面和智慧的解決方案。展望未來，若能將此系統開發出來，我將會驗證此系統真正在業界的使用情況，希望這些研究的結果能為倉庫管理帶來新的突破和進步。

## 第六章 參考文獻



1. Golparvar-Fard, M., Heydarian, A., & Niebles, J. C. (2013). Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. *Advanced Engineering Informatics*, 27(4), 652-663.
2. Kim, H., Bang, S., Jeong, H., Ham, Y., & Kim, H. (2018). Analyzing context and productivity of tunnel earthmoving processes using imaging and simulation. *Automation in Construction*, 92, 188-198.
3. Rezazadeh Azar, E., Dickinson, S., & McCabe, B. (2013). Server-customer interaction tracker: computer vision-based system to estimate dirt-loading cycles. *Journal of Construction Engineering and Management*, 139(7), 785-794.
4. Bügler, M., Borrmann, A., Ogunmakin, G., Vela, P. A., & Teizer, J. (2017). Fusion of photogrammetry and video analysis for productivity assessment of earthwork processes. *Computer-Aided Civil and Infrastructure Engineering*, 32(2), 107-123.
5. Han, S., & Lee, S. (2013). A vision-based motion capture and recognition framework for behavior-based safety management. *Automation in Construction*, 35, 131-141.
6. Ding, L., Fang, W., Luo, H., Love, P. E., Zhong, B., & Ouyang, X. (2018). A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Automation in construction*, 86, 118-124.
7. Park, M. W., Elsafty, N., & Zhu, Z. (2015). Hardhat-wearing detection for enhancing on-site



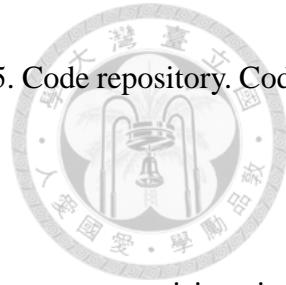
- safety of construction workers. *Journal of Construction Engineering and Management*, 141(9), 04015024.
8. Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., & An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in construction*, 85, 1-9.
9. Seo, J., Han, S., Lee, S., & Kim, H. (2015). Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*, 29(2), 239-251.
10. Yang, J., Park, M. W., Vela, P. A., & Golparvar-Fard, M. (2015). Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future. *Advanced Engineering Informatics*, 29(2), 211-224..
11. 林之謙. (2021). 電腦視覺技術於自動化工程進度管控. *土木水利*, 48(2), 22-31.
12. Golparvar-Fard, M., Peña-Mora, F., & Savarese, S. (2015). Automated progress monitoring using unordered daily construction photographs and IFC-based building information models. *Journal of Computing in Civil Engineering*, 29(1), 04014025.
13. Golparvar-Fard, M., Peña-Mora, F., & Savarese, S. (2010, June). D4AR—4 Dimensional augmented reality-tools for automated remote progress tracking and support of decision-enabling tasks in the AEC/FM industry. In Proc., The 6th Int. Conf. on Innovations in AEC Special Session-Transformative machine vision for AEC.
14. Bosché, F., Ahmed, M., Turkan, Y., Haas, C. T., & Haas, R. (2015). The value of integrating



- Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction*, 49, 201-213.
15. Turkan, Y., Bosché, F., Haas, C. T., & Haas, R. (2013). Toward automated earned value tracking using 3D imaging tools. *Journal of construction engineering and management*, 139(4), 423-433.
16. Han, K. K., & Golparvar-Fard, M. (2015). Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs. *Automation in construction*, 53, 44-57.
17. Ahn, Y., Choi, H., & Kim, B. S. (2023). Development of early fire detection model for buildings using computer vision-based CCTV. *Journal of Building Engineering*, 65, 105647.
18. Reja, V. K., Varghese, K., & Ha, Q. P. (2022). Computer vision-based construction progress monitoring. *Automation in Construction*, 138, 104245.
19. Shi, J. J., & Halpin, D. W. (2003). Enterprise resource planning for construction business management. *Journal of construction engineering and management*, 129(2), 214-221.
20. Skibniewski, M. J., & Ghosh, S. (2009). Determination of key performance indicators with enterprise resource planning systems in engineering construction firms. *Journal of construction engineering and management*, 135(10), 965-978.
21. Caldas, C. H., Menches, C. L., Reyes, P. M., Navarro, L., & Vargas, D. M. (2015). Materials



- management practices in the construction industry. Practice Periodical on Structural Design and Construction, 20(3), 04014039.
22. LINa, Y. C., & TSERNG, H. P. The Development of e-Hub Supply Chain Management Portal System for Construction Projects. ISARC2003 The Future Site, 531.
23. Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 886-893). Ieee.
24. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.
25. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
26. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.
27. Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).
28. Liu, X. P., Li, G., Liu, L., & Wang, Z. (2019). Improved YOLOV3 target recognition algorithm based on adaptive eged optimization. Microelectron. Comput, 36, 59-64.



29. Jocher, G., Nishimura, K., Mineeva, T., & Vilariño, R. (2020). yolov5. Code repository. Code repository.
30. Shannon Bond (2019). in San Francisco,"Amazon introduces computer vision into warehouses" FINANCIAL TIMES, JULY 2 2019. <https://www.ft.com/content/ce0a7828-97bd-11e9-8cfb-30c211dcd229>