

# 應用 YOLOv8 與 NCC 協助鹽倉影像三維建模之 控制點坐標量測

陳律志<sup>1\*</sup> 饒見有<sup>2</sup> 洪慶忠<sup>3</sup>

## 摘要

本研究針對大型圓柱形鹽倉，提出一套整合攝影測量與深度學習之自動化三維建模流程，以提升惡劣環境下之建模效率與量測精度。於鹽倉中央懸臂吊車架設雙工業相機，取得高重疊影像，並以 CLAHE 進行影像增強。利用 YOLOv8 偵測地面控制點標靶位置，結合 Normalized Cross-Correlation 進行次像素精度定位。為提升空中三角測量品質，採用 LightGlue 進行特徵匹配與影像篩選，並透過 Metashape Python 腳本自動完成空三解算與稠密點雲建構。實驗以 401 筆 GCP 為基礎，結果顯示定位誤差約為 2%，證明本方法具良好穩定性與精度，可有效降低人工量測負擔並提升鹽堆體積估算效率。

**關鍵詞：**攝影測量、正規化互相關、深度學習、三維建模

## 1. 前言

隨著工業製程與儲運需求的提升，工業鹽在化工、能源與製造業中扮演關鍵角色，大型鹽倉成為其主要儲存與流通設施。傳統對鹽堆體積與庫存量的測量方式多仰賴目視估算或人工標記，面對昏暗光線、高粉塵、堆疊不規則等複雜環境條件，常造成測量結果不穩定，甚至伴隨安全風險。雖然地面雷射掃描(Terrestrial Laser Scanning, TLS)能提供高精度量測結果，但操作需動員大型設備與人力，且無法涵蓋鹽堆背面陰影區域，導致資料缺損或建模不完全。

### 1.1 研究背景

本研究旨在解決傳統鹽倉量測方式存在的人工干預高、效率低與資料不完整等問題，提出一套融合深度學習與攝影測量技術的全自動化三維建模流程，以提升建模品質並降低現場作業風險。

本研究的具體目的如下：(1)發展一套自動化 GCPs (Ground Control Points)偵測與定位流程：透過

CLAHE (Contrast Limited Adaptive Histogram Equalization)提升影像品質，使用 YOLOv8 偵測倉壁標籤，再以 NCC 執行次像元精度比對，取得高準確度的 GCP 影像坐標。(2)應用深度特徵匹配以優化影像選取策略：利用 LightGlue 自動分析影像間重疊程度，剔除過度重疊影像以穩定 SfM 結果。(3)建構可穩定應用於鹽倉內部之自動建模流程：整合上述流程於 Metashape 腳本中，完成影像導入、相機定位、誤差剔除、模型生成與體積推估。(4)驗證該流程於惡劣環境條件下的可行性與精度表現：針對實際拍攝的鹽倉影像，評估 YOLOv8、NCC 的準確率與整體建模成果，提供作業自動化之量化依據。最終目標是建立一套可應用於封閉或危險工業空間(如鹽倉、煤堆、料場)的低成本、高效率自動量測系統，促進三維建模在實務工程與盤點管理中的推廣應用。

### 1.2 研究動機與目的

在大宗物料的儲存與管理作業中，三維建模技術已成為重要的數位工具，不僅能用於建立完整的

<sup>1</sup> 國立成功大學測量及空間資訊學系 碩士

<sup>2</sup> 國立成功大學測量及空間資訊學系 教授

<sup>3</sup> 乾坤測繪科技有限公司 總經理

\* 通訊作者, E-mail: p66131082@gs.ncku.edu.tw

收到日期：民國 114 年 07 月 15 日

修改日期：民國 115 年 01 月 21 日

接受日期：民國 115 年 03 月 23 日

場域模型，亦可進行如體積計算、變形分析、空間規劃與結構監測等延伸應用 (Kraus, 2007)。尤其在工業用鹽的儲存場域，例如鹽倉、鹽堆或臨海堆場，其物理形狀受限於堆積方式與場地幾何，且可能隨著時間產生不規則變化，因此倚賴準確的三維模型來執行庫存盤點與物流預測，是當前產業管理之重要課題。然而傳統的鹽倉測量方式，例如使用全站儀手動標記 GCP、運用 GPS 量測表面特徵點，或直接目視估算堆高，均面臨諸多限制與挑戰。在實務操作中，鹽倉內部常具有如下特性：

- (1) 光線微弱或分佈不均：光線自屋頂通氣孔或側牆，照度低且無法補光。
- (2) 粉塵濃度高：鹽堆裝卸與移動過程產生大量粒子，影響視線與感測。
- (3) 作業環境封閉、潮濕且具腐蝕性：影響儀器壽命與操作者安全。

上述特性不僅導致傳統測量需大量人力進場，還可能因視距受阻或測量盲區而造成資料缺漏或重建失真。部分單位雖已導入地面雷射掃描 (Terrestrial Laser Scanning, TLS)，但 TLS 雖可提供高精度資料，仍需耗費大量時間進行站點佈設，且在鹽堆背側、吊車遮蔽區或臨邊地形等位置，容易形成資料遺失 (Westoby *et al.*, 2012)，使得模型完整性與安全性受限。

在近年來發展中，以攝影測量為基礎的 SfM (Structure from Motion) 技術因成本低、靈活性高，逐漸成為室內或受限場域建模的替代方案。然而，SfM 技術要能產出具備工程級精度的三維模型，其關鍵在於必須引入高精確度的地面控制點 (Ground Control Points, GCPs) 進行絕對方位 (Absolute Orientation) 與尺度定義。若控制點在影像中的像點坐標 (Image Coordinates) 量測產生誤差，將導致後續空三平差 (Aerial Triangulation) 的模型產生整體扭曲或比例尺錯誤，進而嚴重影響最終的體積計算結果。因此，如何在鹽倉高粉塵、低照度且標籤易受遮蔽的惡劣環境下，自動且精確地提取控制點坐標，成為本研究亟欲解決的核心問題。

針對此一需求，本研究採用 YOLOv8 (You Only Look Once version 8) 系列模型作為核心偵測

工具。自 (Redmon *et al.*, 2016) 提出 YOLO 架構後，憑藉其端對端的一階物件偵測能力，YOLOv8 在速度與準確度之間取得優異平衡，已廣泛應用於工業場域之標籤辨識 (Jocher *et al.*, 2023)。透過深度學習模型，系統能快速在複雜背景中定位 GCP 編號，大幅改善傳統依賴人工標記的低效率流程。近年亦有研究將 YOLO 架構應用於 UAV 測量影像中的 GCP 自動偵測，顯示深度學習在控制點辨識上的可行性與潛力 (Naufal *et al.*, 2024)。

為進一步達成高精度定位需求，確保控制點量測精度達到次像素 (Sub-pixel) 等級，本研究在 YOLOv8 初步偵測後，輔以 NCC (Normalized Cross-Correlation) 方法進行模板匹配。模板匹配技術在機器視覺中常被用於邊緣與形狀之精細辨識，在光照變化或部分遮蔽情況下仍具有良好穩定性 (Heo *et al.*, 2010、Zhang *et al.*, 2017)。近年研究亦指出，模板匹配方法可有效應用於影像邊緣與形狀辨識任務 (Zhao *et al.*, 2024)，因此適合用於對 GCP 進行圖樣辨識與精細定位。

此外，為解決影像重疊過高導致相機姿態解算不穩的問題，本研究引入 LightGlue 特徵點匹配演算法。LightGlue 透過 Transformer 架構進行特徵匹配，能有效提升局部特徵點匹配效率與可靠性，並改善傳統特徵匹配方法在複雜場景中的匹配品質 (Lindenberger *et al.*, 2023)。傳統影像特徵點方法如 SIFT (Lowe, 2004) 雖具尺度與旋轉不變性，但在高粉塵與低對比環境中仍可能出現匹配失敗的情形。近年亦有研究利用深度學習方法學習影像特徵點，例如 SuperPoint (DeTone *et al.*, 2018)，顯示深度特徵在影像匹配任務中的潛在優勢。

綜合而言，本研究旨在建立一套自動化建模系統，透過解決惡劣環境下控制點坐標精確量測之瓶頸，提供一種低人力需求、高一致性且高精度的替代建模策略，進一步推動深度學習與攝影測量於工業空間中的融合應用，並為相關領域之三維建模流程自動化提供具體可行的參考框架。

## 2. 研究方法

## 2.1 研究流程

本研究旨在建立適用於鹽倉環境之自動化三維建模系統，整體處理流程如圖 1 所示。各步驟與本章節之技術架構具有對應關係，茲分述如下：

- (1) 資料獲取(對應第 2.3 節)：流程始於實驗設備及資料收集 (2.3 節)。此階段包含相機規格與佈設方式 (2.3.1 節) 的選定，以及現場控制點設置與量測 (2.3.2 節)，以確保後續建模具備真實坐標系基準。
- (2) 影像特徵篩選(對應第 2.4 節)：取得原始影像序列後，系統執行影像過濾策略 (2.4 節)。本研究採用 SuperPoint (2.4.1) 進行特徵點提取，並透過 LightGlue (2.4.2 節) 進行特徵匹配。此處設有一判斷機制(圖 1 菱形框)：
  - A.判定為 True (保留)：若相鄰影像間的特徵匹配數量充足且幾何分布良好，則保留該影像進入下一階段。
  - B.判定為 False (剔除)：若影像重疊率過高導致資訊冗餘，或特徵點過少導致幾何解算不穩定，系統將自動剔除該影像，不納入後續運算，以提升整體建模效率。
- (3) 控制點自動化識別(對應第 2.5 節)：經篩選之有效影像將進入控制點偵測 (2.5 節) 模組。首先進行 影像增顯 (2.5.1 節) 以克服現場低照度問題；接著利用 YOLOv8 模型 (2.5.2 節) 進行控制點標籤區域偵測；最後，針對偵測到的標籤範圍，執行 控制點精細匹配搜尋 (2.5.3 節)，利用 NCC 演算法取得次像素等級的精確中心坐標。
- (4) 三維重建與優化(對應第 2.6 節)：系統將二維影像特徵與三維控制點資訊結合，執行光束法平差。為確保模型精度，此過程包含自動化的剔除差錯機制 (2.6.1 節)，迭代移除高誤差點位，最終完成模型建立 (2.6.2 節)，產製出具備真實尺度之三維點雲與數值模型。

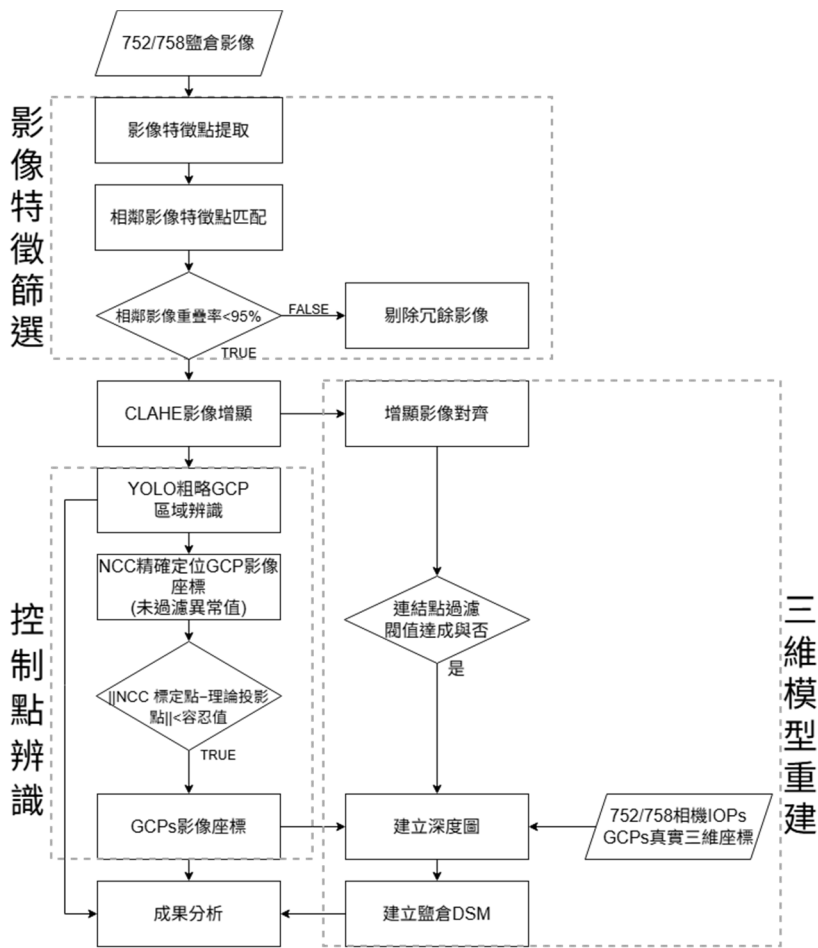


圖 1 研究流程圖

## 2.2 實驗場域

本研究的實驗場域為一座位於臺灣南部的圓柱形鹽倉，其內徑約 120 m，屋頂呈半球形拱頂結構，中央設有一座可旋轉之懸臂吊車，用以搬運工業用鹽堆(如圖 2 所示)。倉內採封閉式空間設計，具備高粉塵、低光源與高濕度等作業限制，對傳統測量技術(如全站儀、TLS、GPS)形成顯著挑戰，尤其在鹽堆背面或倉壁高處常因視距受阻導致量測不完全，進而影響模型精度與後續體積推估。

為突破上述限制本研究於鹽倉懸臂吊車底部安裝兩台工業級數位相機(命名為影像編號 752 與 758)，配置固定拍攝間隔與同步快門，透過環形軌道環繞拍攝倉內空間。每次旋轉作業可獲得涵蓋鹽

堆全區之多角度高重疊影像。後續經建模與正射處理後，產出完整之鹽倉空間展開圖(如圖 3 所示)，可觀察各控制點標籤於倉壁上的實際分布情形。

在資料蒐集中，倉壁等距設置 18 組人工標記的控制點標籤(例如 0N、20N、40N 至 340N)，提供後續 YOLOv8 偵測與 NCC 精準定位之依據。本研究共獲得 330 張影像，並進行 CLAHE 增強以強化對比，再透過 YOLOv8 模型自動辨識標籤位置。針對偵測出的子影像區塊進行 NCC 比對，取得次像元等級之 GCP 影像坐標。除控制點標記外，本研究亦以 LightGlue 對整體影像進行特徵匹配，篩除重複資訊過多之冗餘影像對，提升相機定位與建模穩定性。

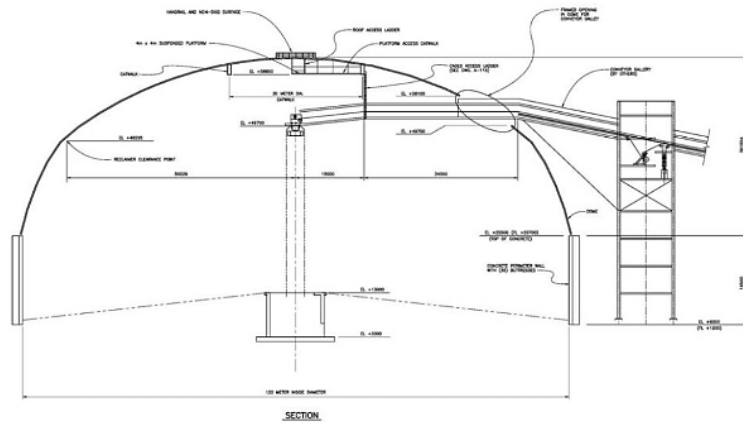


圖 2 鹽倉建築設計剖面圖

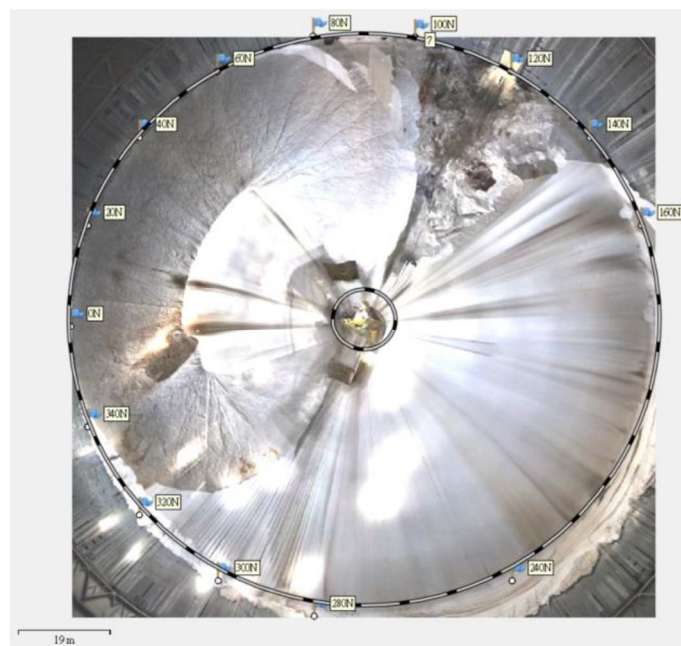


圖 3 鹽倉正射影圖

## 2.3 實驗設備及資料收集

### 2.3.1 相機規格與佈設方式

本研究採用兩台 FLIR 工業級數位相機，安裝於鹽倉中央懸臂吊車下方，垂直向下固定拍攝，隨吊臂旋轉環繞倉內取得多角度影像。相機搭載 6mm 焦距鏡頭，視野角(Field of View)為  $99^\circ \times 81^\circ$ ，影像解析度為  $4096 \times 3000$  像素(約 1200 萬畫素)，依鹽堆高度估算空間解析度約 1~2 cm。兩台相機分別命名為 752 與 758，拍攝間隔固定並具備同步快門功能，資料傳輸經網路傳至中央處理主機。為確保影像品質與後續處理穩定性，所有影像經 CLAHE 增強後才進行特徵分析與控制點偵測。

### 2.3.2 控制點設置與量測

為建立建模空間與實際場域間之尺度關係，本研究於鹽倉內壁依據方位角每隔 20 度設置一組人工控制點。標籤編號由 0N、20N 遞增至 340N，全場域共計佈設 18 組實體控制點。每組標記包含數字標籤與垂直黑條，其右上角對應位置作為 GCP 點位。

控制點三維坐標由乾坤測繪科技以全站儀量測，原始資料採 TWD97 平面坐標系與正高。為簡化處理與比對，本研究將坐標轉換為以鹽倉中心為原點的局部坐標系統(Local-E, Local-N, Local-H)，其中高程參考值以設計圖所示倉底 6m 為基準進行修正。

此轉換後之坐標除提升與模型的對齊準確性外，亦作為 YOLOv8 與 NCC 自動標記成果之精度驗證依據。需說明的是，雖現場與模型訓練定義之標籤類別共 18 類，然受限於測試資料集之隨機抽樣與部分視角遮蔽影響，後續章節(如表 3)之精確度驗證將針對其中有效觀測之 15 類樣本進行統計分析。

## 2.4 影像過濾策略

由於本研究採用固定路徑自動拍攝，所獲得影像具有極高重疊率，若不進行過濾，將導致 SfM 解算過程中出現基線過短、視差不足等問題，進而影響相機姿態與建模穩定性。

為解決此問題，本研究設計重疊率跳片策略，以每張影像為基準，透過深度學習模型 SuperPoint 擷取特徵點，再由 LightGlue 進行配對與移動量分析，依據平均特徵位移與匹配點分佈，自動判斷是否與前一張主影像過於重複。若重疊率過高則將該影像跳過，反之則視為新主影像納入建模流程。此策略可有效篩除冗餘資料，並保留空間分布良好的影像組合，提升建模效率與 SfM 解算穩定性。

### 2.4.1 SuperPoint

SuperPoint 為一種端對端的深度學習特徵點提取網路，能同時產生影像中的角點(keypoints)與其描述子(descriptors)，具備對光照變化與低品質影像的高度魯棒性。相較於傳統方法(如 SIFT)Scale-Invariant Feature Transform)、ORB(Oriented FAST and Rotated BRIEF))，SuperPoint 更適用於本研究中光線昏暗且粉塵較重的場域。每張輸入影像經適當縮放後，透過 SuperPoint 模型輸出稠密且可學習的影像特徵，作為後續匹配的基礎資料。

### 2.4.2 SuperPoint

在完成影像特徵擷取後，本研究使用 LightGlue 進行特徵點匹配與影像過濾。LightGlue 採用 Transformer 架構，透過自注意力與交叉注意力機制實現動態配對，可有效排除低信度匹配點，並搭配 RANSAC 進行幾何驗證以濾除離群值，確保每對影像間的匹配結果具有空間一致性。

為控制參與建模的影像品質與數量，本研究設計「動態主影像選擇機制」，每次以目前主影像為基準，與下一張候選影像進行特徵點比對，依平均位移判斷是否保留或跳過。若候選影像與主影像之特徵點平均位移過小，表示過度重疊，則略過該影像；否則將其視為新主影像。判斷條件如下：

$$\bar{d} = \frac{1}{N} \sum_{i=1}^N \|x_i - x'_i\| \dots\dots\dots (1)$$

其中  $\bar{d}$  為候選影像與主影像之特徵點平均位移(以像素為單位)； $N$  為有效匹配點數； $x_i$ 、 $x'_i$  分別為第  $i$  個匹配點在主影像與候選影像上的坐標位置。若  $\bar{d} < T$ ，其中  $T$  為預設閾值(如影像寬度的 5%)，則該影

像視為過度重疊，將被跳過；反之則保留，並更新為新的主影像。

圖 4 顯示使用 LightGlue 進行配對的特徵點結果。綠點為偵測特徵，綠線為有效匹配對，分布於鹽堆與倉壁區域。此結果顯示 LightGlue 可在高重疊影像間建立穩定的幾何對應。

## 2.5 控制點偵測

### 2.5.1 影像增顯

儘管主流攝影測量軟體 Agisoft Metashape 已內建自動標誌點(Marker)偵測功能，其原理主要採用基於幾何特徵匹配(Geometric Feature Matching)與灰度梯度閾值(Grayscale Gradient Thresholding)的傳統電腦視覺演算法，但在鹽倉高粉塵與光照不均的環境下，標籤常因髒汙或陰影遮蔽導致邊緣模糊，造成傳統演算法識別失敗或誤判。為此，本研究改採 YOLOv8 深度學習模型，不同於依賴硬性幾何規則的傳統方法，YOLOv8 透過卷積神經網路提取深層語義特徵，對於物件尺度變化、局部遮蔽與低信噪比影像具備極高強健性，能有效解決惡劣環境下的控制點漏判問題。

由於鹽倉內部光線微弱且分布不均，部分控制點標籤常處於陰影或低對比區域，導致後續辨識困難。為此，本研究在 YOLOv8 與 NCC 偵測前，先對原始影像套用 CLAHE (Contrast Limited Adaptive Histogram Equalization) 進行對比度增強處理 (Zuiderveld, 1994)。

實作上，本研究首先將原始影像由 RGB 色彩空間轉換至 CIE  $L^*a^*b^*$ (CIELAB)色彩空間。在此色彩模型中， $L^*$  表示亮度分量，而  $a^*$  與  $b^*$  分別代表綠-紅與藍-黃之色度分量。為避免 RGB 色彩空間中的藍色通道(Blue channel, B)與 CIELAB 色彩空間中的色度分量( $b^*$ )在符號上產生混淆，本文統一以  $L^*$ 、 $a^*$ 、 $b^*$  表示 CIELAB 各分量。

在影像增強過程中，本研究僅針對亮度分量  $L^*$  應用 CLAHE 演算法，以提升影像之局部亮度對比，同時保留原始影像之色度資訊( $a^*$  與  $b^*$  通道數值維持不變)。CLAHE 透過將影像劃分為多個局部區塊並分別進行直方圖均衡化處理，以強化低對比區域之亮度細節；同時藉由設定 clip limit 限制直方圖高度，避免對比過度增強而導致影像雜訊放大，處理完成後將強化後之  $L^*$  通道與原始  $a^*$ 、 $b^*$  通道重新合併，並將影像轉換回 BGR(Blue-Green-Red)色彩空間輸出，以符合 OpenCV 影像處理函式庫之預設格式，供後續 YOLOv8 物件偵測與 NCC 模板匹配程序使用。

透過上述影像增強流程，可顯著提升黑白控制點標籤邊緣之局部對比度，使標籤之幾何邊界更加清晰，進而提高 YOLOv8 偵測與 NCC 次像元定位之成功率。

圖 5 為 CLAHE 處理前(左)與處理後(右)之影像對照。由圖可見，處理後影像之整體亮度與局部細節皆明顯提升，特別是在控制點標籤區域(如倉壁刻度)之黑白對比邊緣更加清晰，利於後續標籤辨識與次像元定位。

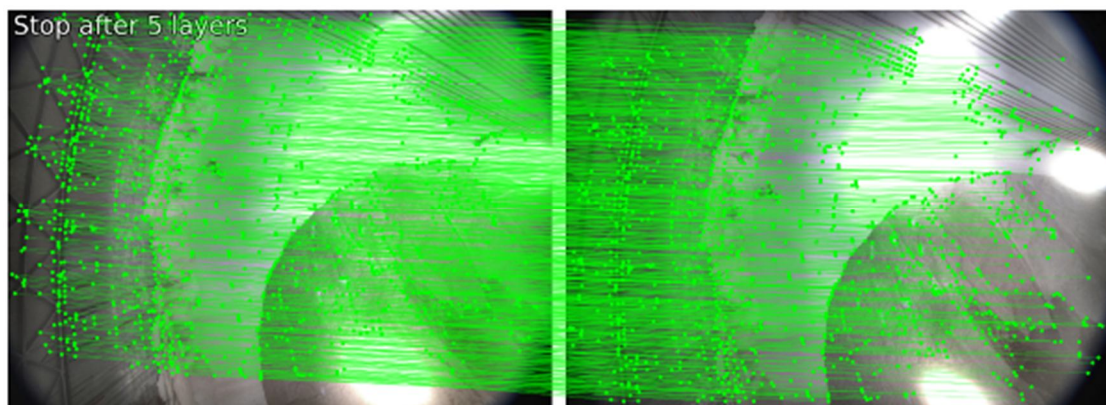


圖 4 LightGlue 影像匹配示意圖。綠色點為特徵點，線段為配對結果，顯示影像間具有連續且穩定的匹配關係

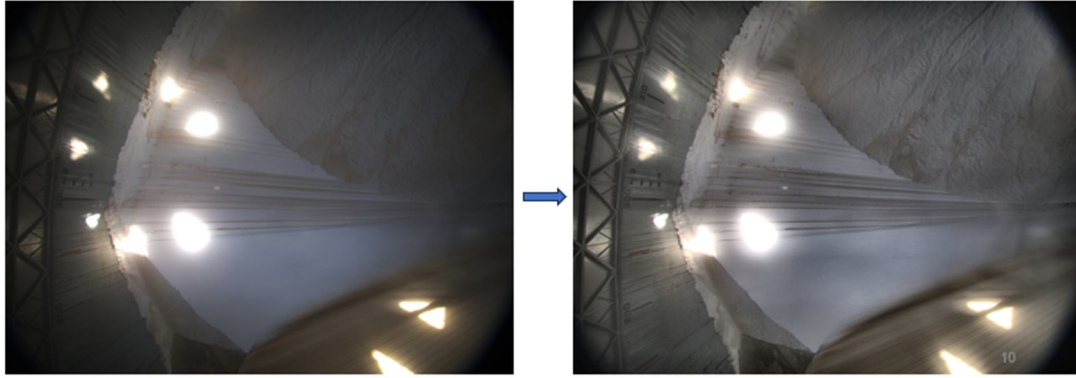


圖 5 增顯影像對比，(左)CLAHE 處理；(右)CLAHE 處理後

## 2.5.2 控制點標籤區域偵測

### 2.5.2.1 YOLOv8 模型

YOLOv8(You Only Look Once)為一種即時目標偵測架構，能於單一網路中同時預測物件的邊界框與類別，相較於傳統兩階段式方法(如 R-CNN)，YOLOv8 具備推論速度快、結構簡潔的特性，特別適用於需處理大量影像的自動化場域。

本研究採用 Ultralytics 發表的 YOLOv8 架構，屬於 YOLOv8 系列最新版本，其主要特色為 anchor-free 設計、輕量化骨幹結構，以及支援多任務(偵測、分類與分割)訓練。YOLOv8 整體由四大模組構成，分別說明如下：

#### (1) Backbone(特徵提取網路)

YOLOv8 採用改良後之 CSPDarknet 結構，透過 C2f 模組(cross-stage partial fusion)取代傳統 ResBlock，有效降低參數數量並保留語義資訊。整體網路自影像輸入後依序通過多層卷積與降採樣模組，提取由淺至深的多層特徵圖，作為後續辨識依據。最末端並接上一層 SPPF 模組，用以增強多尺度上下文特徵。

#### (2) Neck(多尺度特徵融合)

Neck 結構由 FPN(Feature Pyramid Network)與 PAN(Path Aggregation Network)組成。FPN 自深層特徵圖反向傳遞語義資訊至淺層，而 PAN 則由底部補回高解析空間資訊，使融合後的特徵兼具精細定位與語意辨識能力。此結構透過上採樣、下採樣與多層 C2f 強化模組完成特徵整合。

#### (3) Head(解耦式預測模組)

YOLOv8 採用解耦式 head 結構，於三個不同解析度的特徵圖上分別執行分類與邊界框預測，涵蓋大、中、小等不同尺寸的目標偵測任務。每一輸出層預測包含位置  $[x,y,w,h]$ 、物件置信度(objectness)與各類別機率。

#### (4) 損失函數(Loss Function)

YOLOv8 所用的損失函數結合三個分項：框選誤差(CIOU + DFL 分布式回歸)、目標置信度誤差(Binary Cross Entropy)與分類誤差(BCE 或 focal loss)。最終損失為三者加權總和，並透過動態 loss 平衡自動調整三者權重，以提升收斂穩定性與偵測準確率。

YOLOv8 整體流程如圖 6 所示：輸入影像經由 backbone 模組提取深淺特徵後，透過 neck 架構進行語義與空間資訊的多尺度融合，最終由 head 模組在三層特徵圖上進行獨立預測。此設計能同時兼顧不同尺寸與特徵層級之物體，為本研究於鹽倉環境中進行控制點標籤偵測提供高效率且穩定的解決方案。

### 2.5.2.2 鹽倉資料集

為提升 YOLOv8 模型於鹽倉場域中控制點標籤之偵測準確度，本研究自建一套專屬資料集，並結合多種資料增強方法進行監督式訓練。原始影像取自懸臂吊車繞行鹽倉過程中同步拍攝所得，經 CLAHE 處理後再進行標註。最終標註資料集包含 9716 張訓練影像、2513 張驗證影像與 178 張測試影像，涵蓋 18 類標籤(每 20 為一類，數值從 0 至 340，加上字母 "N")，均依 YOLOv8 格式標示類別與相對坐標。

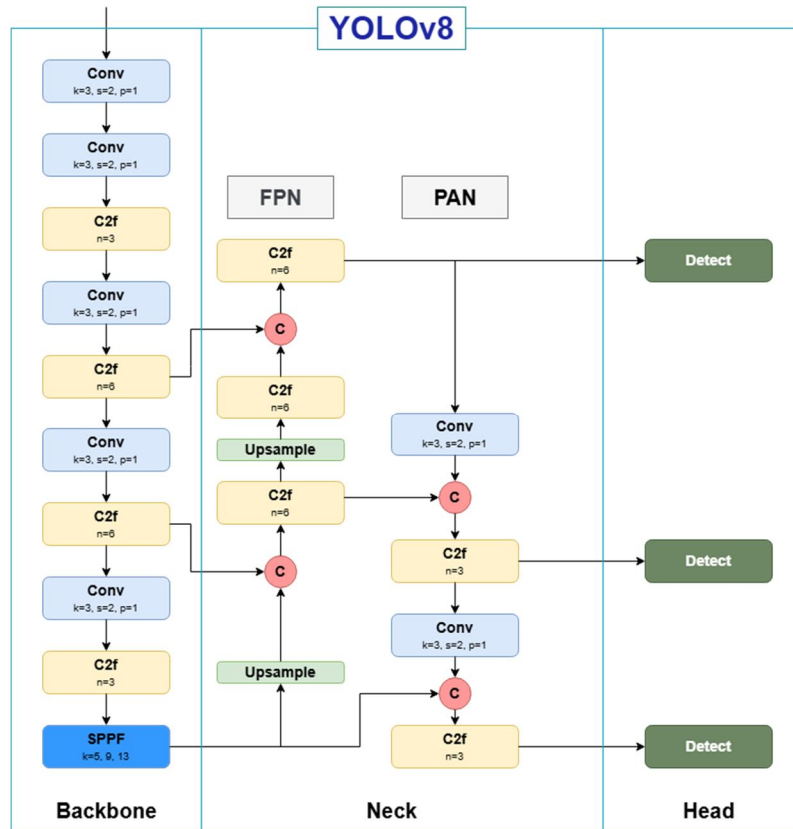


圖 6 YOLOv8 網路架構圖

為提升模型對不同光照與拍攝角度的適應性，本研究設計數種增強策略，包含：(1)影像旋轉 180°：模擬吊臂翻轉情境；(2)亮度與對比調整：強化陰暗區標籤可見性；(3)局部區域亮度校正：將候選區域亮度與模板進行匹配，調整增益與偏移值，以降低誤判；(4)同步標註修正：確保增強後標註與影像對齊。訓練過程中以 validation loss 與 mAP 為主要監控指標，並自動儲存最佳模型結果。最終模型用於鹽倉影像中標籤編號之初步定位與子影像裁切，供後續 NCC 進行亞像素級匹配。

### 2.5.2.3 子影像裁切

YOLOv8 偵測後系統會讀取每張影像的標籤預測結果，排除不合法類別(如 999)，並根據邊界框資訊裁切對應子影像作為 NCC 輸入。裁切邏輯依據框中心與大小計算左上與右下角坐標，若裁切區域超出影像邊界則自動略過。

裁切後的子影像會經灰階轉換與亮度標準化處理，調整平均值至 128、標準差至 32，以提升 NCC 對比穩定性。所有子影像依格式 {類別}\_{影像編

號}.jpg 命名，並儲存於指定資料夾供後續批次處理。

### 2.5.2.4 模型評估指標

為量化 YOLOv8 模型在控制點標籤偵測上的效能，本研究採用物件偵測領域通用之評估指標，包含精確率(Precision)、召回率(Recall)以及平均精度均值(mean Average Precision, mAP)。各指標定義如下：

- (1) 精確率 (Precision, P)：代表模型預測為正樣本的結果中，實際為正樣本的比例。計算公式如式 (2)：

$$P = \frac{TP}{TP+FP} \dots\dots\dots(2)$$

- (2) 召回率 (Recall, R)：代表在所有實際正樣本中，被模型正確預測出的比例。計算公式如式 (3)：

$$R = \frac{TP}{TP+FN} \dots\dots\dots(3)$$

其中，TP (True Positive) 為正確偵測之標籤數，FP (False Positive) 為誤判之標籤數，FN (False

Negative) 為漏判之標籤數。

(3) 平均精度均值 (mAP)：mAP 為衡量模型整體性能之綜合指標。首先計算單一類別的平均精度 (Average Precision, AP)，即 P-R 曲線 (Precision-Recall Curve) 下之面積；mAP 則為所有類別 AP 值之平均。本研究採用 mAP 0.5 (IoU 閾值設為 0.5 時的 mAP) 作為主要評估依據。

### 2.5.3 控制點精細匹配搜尋

為達成高精度控制點定位，本研究於 YOLOv8 粗略定位後導入 NCC (Normalized Cross-Correlation) 進行模板匹配，從子影像中尋找與既定標籤模板最相似的區域。NCC 為一種常用於影像對位的匹配方法，具有遮罩不敏感、尺度不變性強等特性，特別適合應用於固定格式之標籤辨識，匹配過程如下：對於子影像 I 與模板 T，於滑動視窗中計算下列 NCC 分數：

$$NCC(u, v) = \frac{\sum_{x,y} [I(x+u, y+v) - \bar{I}_{uv}] \cdot [T(x, y) - \bar{T}]}{\sqrt{\sum_{x,y} [I(x+u, y+v) - \bar{I}_{uv}]^2 \cdot \sum_{x,y} [T(x, y) - \bar{T}]^2}} \dots\dots\dots (4)$$

其中 u、v 為模板在子影像上滑動的偏移量； $\bar{I}_{uv}$  為子影像中與模板對應區域的平均值； $\bar{T}$  為模板平均值；分子為對應像素的乘積總和 (衡量相似性)；分母為各自標準差乘積 (做正規化)。

為求定位精確，NCC 目標模板設定為 51x51 像素 (奇數以定義幾何中心)，搜尋視窗則採用 YOLOv8 偵測之候選範圍。針對鹽倉內角度傾斜與照明不均等變異，本研究建構包含 76 幅不同旋轉與明暗狀態之「動態模板庫」。透過「多模板匹配 (Multi-template Matching)」策略，計算候選影像與全數模板之相似度並取最佳值，藉此克服非正射視角與局部遮蔽干擾，顯著提升坐標量測之強健性與成功率。

## 2.6 建模流程整合

為將前述自動標記成果應用於三維模型生成，本研究整合 Metashape Python API 建立一套全自動的建模流程。該流程自動導入控制點、進行相機定位與誤差篩選，最終完成模型生成與成果輸出。

### 2.6.1 剔除差錯機制

控制點標記完成後，系統將 NCC 所得之影像點坐標與實際 GCP 三維坐標進行關聯，並執行光束法平差 (Bundle Adjustment)。此程序之目的不僅在於求解相機之外方位參數 (Exterior Orientation)，亦同時進行待測物三維坐標的聯合最佳化求解。然而，受現場光影變化與局部遮蔽影響，部分 NCC 匹配結果可能偏離標籤真實中心，形成幾何上的「粗差 (Outliers)」。為避免這些錯誤觀測值汙染模型精度，本研究設計了一套基於重投影誤差 (Reprojection Error) 的自動化迭代剔除機制。

具體作法為計算每一觀測點在當前相機參數下，其「NCC 偵測位置」與「三維點反投影位置」之歐氏距離，定義為重投影誤差  $e_i$ ：

$$e_i = \|p_i^{NCC} - p_i^{proj}\| \dots\dots\dots (5)$$

其中  $e_i$  為第 i 張影像上的誤差； $p_i^{NCC}$  為第 i 張的 NCC 標定點； $p_i^{proj}$  為該控制點依目前相機模型反投影之理論位置，單位為像素。若該值大於設定容忍值，則認定該控制點於該影像上標記失敗，將其自動移除。為提升穩定性，此步驟會進行最多五輪迭代，此限制係基於實驗經驗設定：一般而言，粗差點在前三輪迭代即可有效收斂並剔除，設定五輪上限可防止系統因追求極小誤差而過度刪除有效控制點，導致幾何結構解算發散或過度擬合，從而在計算效率與模型強健性之間取得最佳平衡。

### 2.6.2 模型建立

完成 GCP 自動篩選後，系統自動觸發建模流程。操作包含：(1) 建立稀疏點雲與相機內外方位參數；(2) 建立影像深度圖；(3) 重建表面模型。建模過程中，Metashape 將自動依據輸入控制點進行空間對位與誤差最小化調整，並在地理坐標系統中輸出具備真實尺寸與方向之模型。整體建模流程可於全自動條件下完成，省去人工標記與調整流程，同時具備 GCP 精度管控與幾何穩定性，為鹽倉內部環境提供一套高效率、可重複的數位建模方案。

### 3. 實驗成果

#### 3.1 影像過濾成效分析

受限於現場硬體設備之觸發機制，相機拍攝作業僅能採用「固定時間頻率」連續擷取，無法配合吊車移動速度進行等距觸發。此限制導致原始影像序列在相鄰畫格間存在極高之重疊率，若全數投入運算將造成空三平差資源浪費與計算效率低落。因此，本研究提出一套基於特徵匹配之關鍵畫格 (Keyframe) 篩選機制。

本機制引入 LightGlue 深度學習特徵匹配技術，透過計算連續兩張影像間之特徵點平均像素位移量，作為判斷視角變化之依據。篩選邏輯設定為：當影像間之特徵位移過小 (即換算重疊率高於 95% 之過濾門檻) 時，系統視其為冗餘資料予以剔除；僅當位移量超過門檻，顯示視角已有顯著改變，方保留該影像作為新的基準畫格。

同時，為驗證深度學習模型在鹽倉惡劣環境下之優勢，本節將以傳統 SURF (Speeded Up Robust Features) 演算法作為對照分析之基準 (Baseline)。SURF 具備尺度與旋轉不變性，為廣泛應用之特徵偵測器。本研究將統計並比較 LightGlue 與 SURF 在不同重疊率條件下之匹配點對數量與篩選效能，詳細數據分析如表 1、表 2 所示。表中之『平均相對旋轉角』係指保留之相鄰影像間的拍攝視角變化量。其計算方式為：利用特徵匹配所得之連結點 (Tie Points)，解算相鄰兩張影像間的旋轉量與平移量，進而推算出相機姿態的旋轉角度變化。

LightGlue 保留影像張數顯著高於 SURF，同一

組影像中其通過率大多落在 90%~99%，顯示出匹配範圍與容錯性明顯較高。752 相機在 2023/10/12 的資料中，原始 1356 張影像中有 1340 張保留，平均方位角僅 0.18 度一張，顯示大多數連續影像皆通過 LightGlue 的匹配條件。相較之下，SURF 僅保留 105 張，平均間隔為 3.43 度，說明其匹配較為保守。

758 相機在多組資料中保留張數略低，推測與拍攝視角偏離倉儲中心、部分區域特徵不足有關，例如 2023/10/26 的影像中僅保留 1327 張 (占原始 3520 張的 37.7%)，但其平均方位角為 2.55 度，仍遠低於 SURF 的 6.92 度，匹配結果較為連續穩定。

LightGlue 在 2023/12/20 與 2024/01/10 的拍攝資料中也顯示出穩定性，在 758 相機中平均每張角度落在 1.83 至 1.84 度之間，未見明顯跳動，752 相機則保持在 0.6 至 1.88 度。整體而言 LightGlue 可維持高密度的影像匹配，有助於三維重建時獲得更完整且連續的模型輸入。

在 2023/10/26 的資料中影像保留率相對偏低，經檢視該批影像後發現，部分影像之拍攝視角偏離倉儲中心，且部分區域受鹽塵與光照不均，造成可辨識之穩定特徵點數量不足，LightGlue 在進行特徵匹配時會依據匹配品質與幾何一致性進行篩選，因此當影像之特徵匹配品質較低或視角差異過大時，系統會自動將該類影像排除。儘管該資料組之通過率較低但保留下來的影像仍維持良好的連續性，其平均相對旋轉角為 2.55°，顯示經篩選後的影像仍可提供穩定的幾何關係，對後續三維重建流程並未造成明顯影響。

表 1 SURF 過濾影像成效統計

日期	相機代號	原始影像拍攝張數	過濾後影像張數	平均相對旋轉角 (°)
2023/10/12	752	1356	105	3.43
	758	1361	50	7.20
2023/10/26	752	3521	132	2.73
	758	3520	52	6.92
2023/12/20	752	589	146	2.47
	758	1811	56	6.43
2024/01/10	752	472	107	3.36
	758	1470	66	5.45

表 2 LightGlue 過濾影像成效統計

日期	相機代號	原始影像拍攝張數	過濾後影像張數	平均相對旋轉角 (°)
2023/10/12	752	1356	1340	0.18
	758	1361	1218	2.23
2023/10/26	752	3521	1332	0.42
	758	3520	1327	2.55
2023/12/20	752	589	513	0.14
	758	1811	1518	2.12
2024/01/10	752	472	454	1.52
	758	1470	1397	1.84

### 3.2 控制點標記準確度分析

#### 3.2.1 YOLOv8 偵測成果

##### 3.2.1.1 模型精度評估

本研究首先評估 YOLOv8 模型在測試資料集上的基礎指標表現，透過 F1-Confidence 曲線、Precision-Recall 曲線及正規化混淆矩陣進行綜合分析。

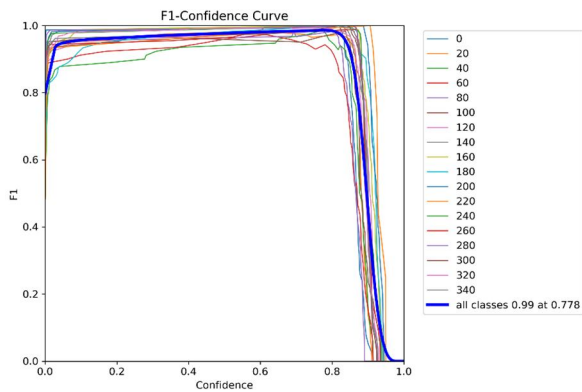


圖 7 YOLOv8 模型各類別控制點之 F1 分數與信心閾值(Confidence Threshold)關係圖，藍色曲線為全類別平均結果

F1-Confidence Curve 呈現 YOLOv8 模型於不同信心閾值下之分類效能變化，橫軸為信心閾值，縱軸為對應之 F1 分數。可觀察到各類控制點之 F1 分數大多於 0.75 至 0.80 閾值區間達最高，代表模型在分類任務中具有良好之精準率與召回率平衡。整體而言，於信心閾值設為 0.778 時可達平均 F1 分數 0.99，提供實務部署時之參考門檻依據。

Precision-Recall Curve 比較各類控制點在不同分類閾值下之精確率(Precision)與召回率(Recall)表現。本研究之模型於所有控制點標籤的分類結果，

其中對角線方格為正確分類率，非對角線區塊則代表誤分類比例。整體分布顯示模型分類準確性高，且誤判大多發生於相鄰編號之間，如標籤 20 與 40、320 與 340。此結果顯示模型在實際應用中能精準辨識標籤類別，且具良好容錯能力。

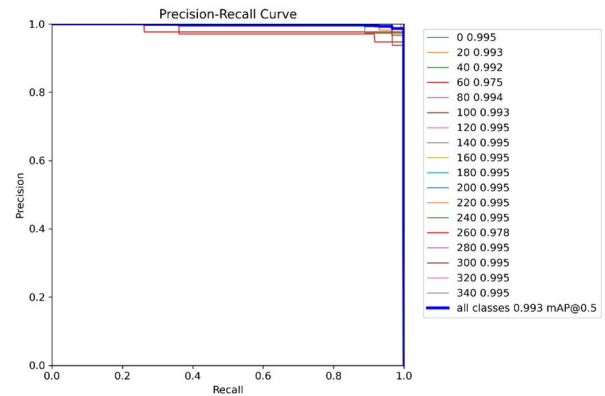


圖 8 YOLOv8 模型各類控制點之 Precision-Recall 曲線，藍色粗線為 macro average，對應 mAP 0.5 值為 0.993

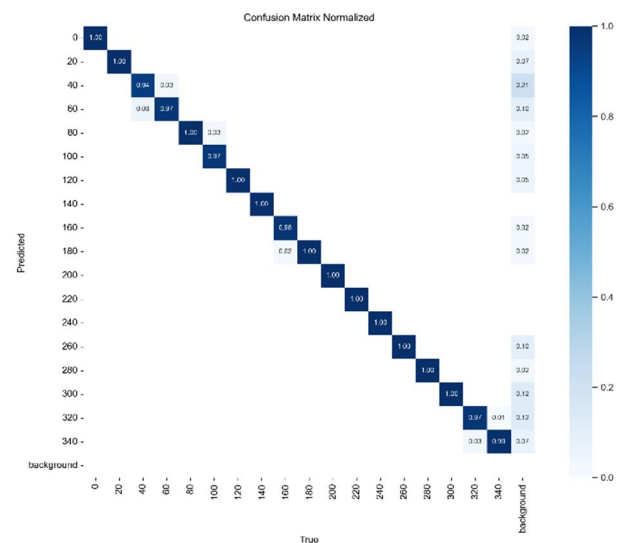


圖 9 YOLOv8 模型於測試資料集之分類結果所生成之正規化混淆矩陣(Normalized Confusion Matrix)

Normalized Confusion Matrix 顯示 YOLOv8 模型於所有控制點標籤的分類結果，其中對角線方格為正確分類率，非對角線區塊則代表誤分類比例。整體分布顯示模型分類準確性高，且誤判大多發生於相鄰編號之間，如標籤 20 與 40、320 與 340。此結果顯示模型在實際應用中能精準辨識標籤類別，且具良好容錯能力。

綜合上述分析，本研究所訓練之 YOLOv8 模型於建立之客製化 GCP 標籤資料集中展現出極高的分類準確性與穩定性。其於多數類別皆可達成 F1 分數 0.99、精確率與召回率皆接近 1.0，且整體 mAP0.5 為 0.993，顯示該模型具備高度實用性。即便在倉內光照不均與影像重疊度高等惡劣條件下，模型仍能有效分辨各類控制點標籤，並提供可靠之粗略定位資訊，作為後續 NCC 精配處理之依據。此結果驗證本研究所訓練之 YOLOv8 模型適用於鹽倉特殊環境中之自動化標記任務。

### 3.2.1.2 偵測成果展示

為提升影像中控制點標籤的自動辨識效率與

準確性，本研究採用 YOLOv8 作為目標偵測架構，負責從 CLAHE 增強後的影像中標註 GCP 位置與類別。YOLOv8 模型輸出之每筆預測包含標籤編號、邊界框坐標與信心水準(Confidence score)，其中數值代表模型對該偵測結果之預測信心。例如「120 0.9」即表示 YOLOv8 判定該區域為標籤編號 120 的機率為 90%。透過模型訓練與最佳化，本研究期望能實現高準確率、低誤判率的標籤初步自動化偵測。圖 10 與圖 11 分別對應於標籤代號為 120 與 100，以及 80 與 60 的影像片段，呈現模型在不同倉壁位置與光照條件下的辨識穩定性與準確性。

綜合圖 10 與圖 11 所示結果可得，YOLOv8 模型於本研究所建構之鹽倉場域資料集中展現出良好之偵測穩定性與準確性。無論控制點標籤位於高亮、背光、或鹽堆遮蔽區域，模型皆能準確辨識，且於多標籤同框情境下亦未出現明顯誤判。此結果反映本研究在影像前處理、資料增廣與模型訓練策略之有效性，亦驗證 YOLOv8 架構適用於大型室內場域控制點偵測任務之可行性與實用價值。



圖 10 YOLOv8 偵測結果展示(一)：標記 120 與 100 的偵測情形



圖 11 YOLOv8 偵測結果展示(二)：標記 80 與 60 的偵測情形

### 3.2.2 NCC 偵測成果

NCC(Normalized Cross-Correlation)為本研究用於標籤精定位的演算法，負責從 YOLOv8 粗略預測的區域中，以模板匹配方式找出次像元的精準對應點。圖 12 顯示部分 NCC 偵測成果，其中上半為配對準確之範例，下半為配對錯誤之案例。圖中紅點代表 NCC 偵測到的最佳對應位置。

上半部為配對正確範例，紅點位置與標籤中心對齊；下半部為配對錯誤範例，紅點明顯偏離標籤。由圖 12 可見在控制點標籤對比明顯、背景紋理不干擾的情況下，NCC 可穩定輸出接近標籤中心的結果，顯示其對於局部區域紋理變化具有良好適應性。然而，當影像遭受嚴重遮蔽、模糊或標籤油漆脫落時，NCC 可能錯判高紋理區為最佳位置，導致誤差。觀察錯誤案例可發現，其中數張存在光斑、標籤傾斜或數字變形現象，為造成配對偏移的主因。因此雖 NCC 在大多數情況下可精準定位，但仍須透過後續判別條件(如匹配信度、位置穩定性)加以篩選以確保成果品質。

### 3.2.3 YOLOv8 辨識與 NCC 匹配精確度分析

為評估 YOLOv8 與 NCC 於控制點標記任務中的實際辨識表現，本研究統計各標籤類別在測試資料集中之總數量、YOLOv8 預測結果與 NCC 精配成果，並計算各自之錯誤率(誤標或錯配比例)。YOLOv8 的任務為從原始全景影像中快速偵測出候選標籤區域，NCC 則進一步於此區域內尋找次像元的最適對應位置。

由表 3 中可見，在 401 筆控制點樣本中，YOLOv8 共成功預測 393 筆，僅有 8 筆錯漏，整體辨識錯誤率為 2%，展現出良好的初始定位能力。進一步由 NCC 進行精配處理後，總誤配件數為 9 筆，整體錯誤率亦為 2%。觀察錯配樣本可發現，錯誤多集中於標籤代碼為 180、220 與 300 等少數類別，推測其因影像模糊、遮蔽、或標籤書寫模糊所致。

在多數標籤類別中，NCC 成功率近乎 100%，顯示其對於標籤樣式穩定的類別具有高度穩定性。

反觀錯配類別，多為數量較少、或影像中位處邊緣之樣本，易受光影干擾影響匹配品質。整體而言，本研究設計之 YOLOv8 + NCC 二階段辨識架構，在維持自動化速度優勢的同時，也能有效兼顧標定精度，顯示其於鹽倉場域控制點定位應用中具有良好潛力。

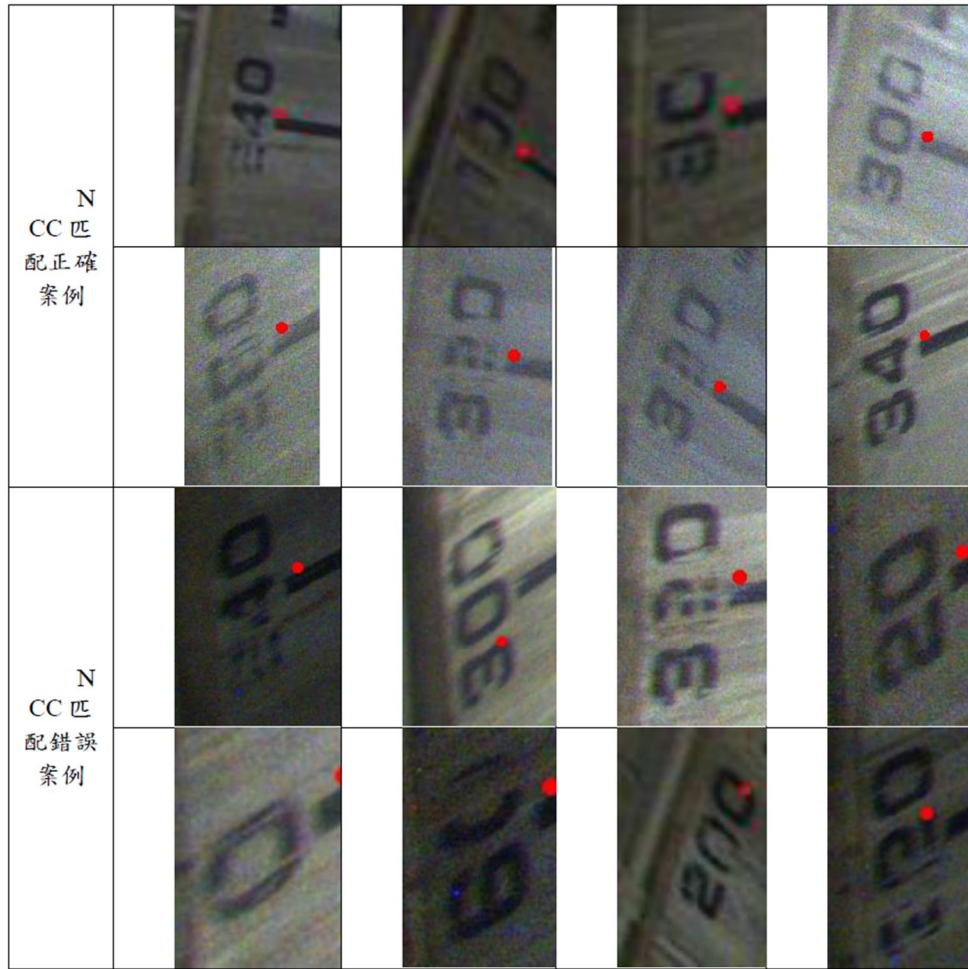


圖 12 NCC 正確與錯誤匹配案例

表 3 YOLOv8 與 NCC 於測試集上的辨識準確度統計表

標籤代號	總標籤數 (張)	YOLOv8 偵測失敗數 (張)	誤差 (%)	NCC 匹配失敗數 (張)	誤差 (%)
0	45	0	0%	1	2%
20	36	0	0%	2	6%
40	23	0	0%	0	0%
60	21	0	0%	0	0%
80	20	0	0%	0	0%
100	16	0	0%	0	0%
120	20	2	10%	0	0%
140	19	2	10%	0	0%
160	28	0	0%	0	0%
180	20	6	30%	0	0%
220	20	0	0%	3	15%
240	22	0	0%	0	0%
260	27	0	0%	0	0%
300	41	0	0%	3	7%
340	43	0	0%	0	0%
合計	401	8	2%	9	2%

### 3.3 三維建模成果與品質分析

#### 3.3.1 稀疏點雲篩選與優化策略

本研究將經 LightGlue 篩選之影像導入空三平差運算後，針對稀疏點雲(Sparse Cloud)執行了嚴格的品質檢核。在幾何結構方面，系統僅保留由大交會角構成且具備「多視角冗餘 (Multi-view Redundancy)」之特徵點，確保每一連結點至少獲得 2 張以上影像的幾何約束，有效排除單一視角誤判風險。在定位精度方面，經迭代優化後，最終留存之連結點其重投影誤差(Reprojection Error)均收斂至 0.5 像素以內。此結果證實，絕大多數三維點位經反投影回二維影像後，與原始觀測點的偏差皆小於半個像素，顯示本研究流程在惡劣環境下仍能維持次像素等級(Sub-pixel level)的高精確度匹配品質。

#### 3.3.2 建模成果

本研究透過自動控制點標記流程與高重疊率影像篩選策略，成功建立鹽倉內部完整的三維模型。圖 13 顯示模型重建後的頂視視角，可觀察倉內鹽堆形貌、堆積區及倉壁細節皆有良好重現，整體而言本研究建構之三維模型不僅具備高度完整性與細節保留，亦成功反映鹽倉內部複雜之地形變化與結構特徵，為後續進行體積估算與時間序列變化分析提供穩定基礎。

然而圖 13 中左下方區域(約 160N 至 280N 之間)出現部分控制點缺失之情形，經檢視原始影像與現場狀況，主因為該區域之牆面標籤受鹽份腐蝕嚴重，導致漆面剝落與特徵損毀 YOLOv8 無法正確框選完整標籤，且 NCC 模板匹配亦因相關係數低於閾值而判定失敗，為避免將錯誤定位的控制點強制納入計算而拉扯模型幾何，自動剔除此區域之低信賴度控制點。儘管該區缺乏控制點約束，但受惠於周圍高重疊率影像之連結點牽制，整體模型之相對精度仍維持在容許範圍內。

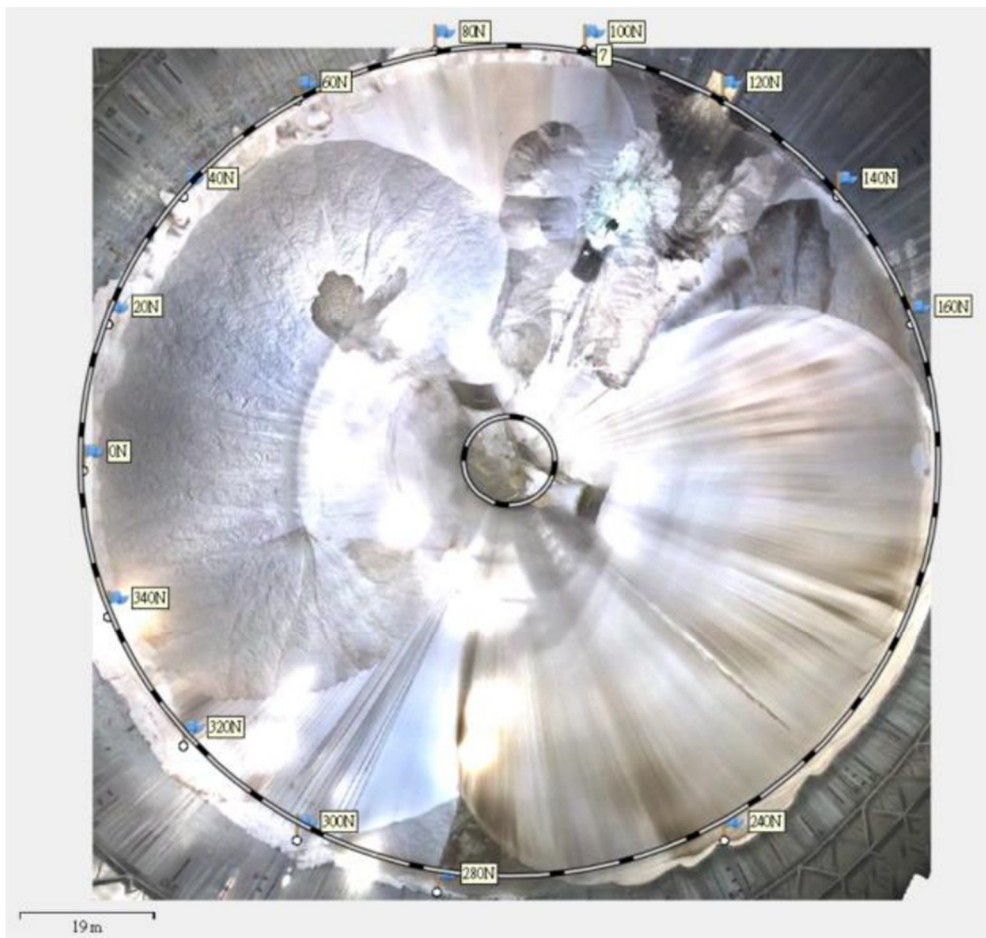


圖 13 2024/01/10 三維模型成果(頂視)

### 3.3.3 體積估算精度驗證

為驗證建模精度，本研究以採集之地面光達 (TLS) 點雲數據為基準真值。透過將光達產製之數值地表模型 (DSM) 與本研究攝影測量 DSM 進行差分運算，計算兩者在相同圓形場域內的體積差異，其高程差異分佈情形如圖 14 所示。由圖中可見，兩者模型在絕大部分區域之高程差極小，顯示空間分佈具備高度一致性。進一步計算兩者在相同圓形測試範圍內之體積差異，結果顯示光達基準體積為  $58,056.9 \text{ m}^3$ ，本研究計算之體積為  $58,960.3 \text{ m}^3$ ，兩者差異量為  $903.4 \text{ m}^3$ ，換算相對誤差僅 1.56%，證實本研究提出之自動化建模流程已具備實務應用所需之高準確度。

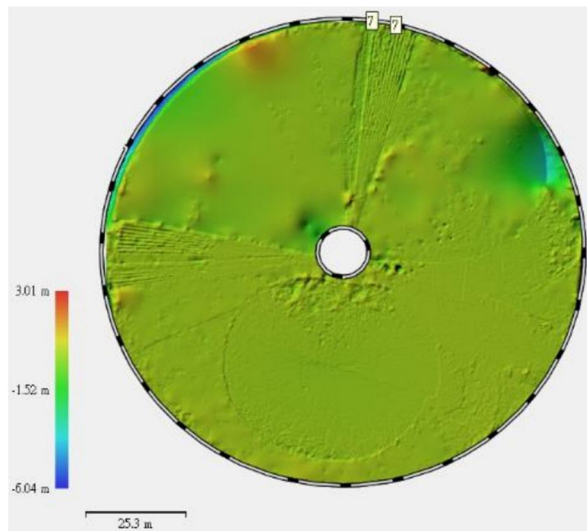


圖 14 2024/01/10 攝影測量與地面光達 DSM 相減之假色圖

## 4. 結論

本研究針對高粉塵、低光源與封閉空間等惡劣環境條件下之大型鹽倉場域，建立了一套結合深度學習與攝影測量的自動化三維建模流程。整體流程涵蓋影像前處理、控制點標記、影像過濾與模型生成四大模組。在影像篩選方面，實測顯示引入 LightGlue 機制後，可保留 90% 以上之有效影像，且相鄰影像之平均方位角間隔介於  $0.14^\circ$  至  $2.6^\circ$  之間，顯著優於傳統 SURF 方法，有效提升了 SfM

解算之幾何穩定性。在控制點識別方面，透過整合 CLAHE 影像增顯、YOLOv8 標籤偵測與 NCC 次元匹配技術，系統在測試資料中展現了 98% 的高匹配準確率，成功克服了現場標籤汙損與光照不足的挑戰。

在三維重建成果方面，本研究成功重現了鹽倉內部之倉壁標籤分佈、鹽堆堆積形狀與輸送帶結構。針對部分區域出現之點雲浮空或重影 (Ghosting) 現象，主要歸因於吊車視角單一與現場照度不足所造成之特徵匹配模糊。為進一步提升模型完整性，建議未來在硬體設備上可進行以下改進：(1) 改良取像載台：增設多角度鏡頭模組或使用可旋轉雲台 (Gimbal)，以獲取更完整的側向紋理，減少視線死角；(2) 優化照明系統：加裝與快門同步之高流明主動式補光燈 (Strobe Light)，以降低陰影干擾並提升特徵點清晰度。

在體積估算精度驗證上，以地面光達 (TLS) 數據為基準真值進行比對，本研究方法之體積計算誤差僅為 1.56% (約  $903.4 \text{ m}^3$ )。相較於傳統作業模式——即依賴全站儀進行稀疏佈點測量，或僅憑經驗進行目視估算，傳統方法常因取樣點不足或人為判斷導致體積估算誤差波動極大 (通常大於 5% 甚至更高)。本研究提出之自動化流程不僅將誤差控制在極低範圍內，且具備全場域覆蓋之特性，證實其足以取代高風險的人工測量作業，滿足工業倉儲盤點對高精度與高效率之實務需求。

綜上所述，本研究建構之自動化建模流程已證實可穩定應用於高遮蔽、高粉塵與照度不均之大型鹽倉。未來可進一步整合點雲語意分割與變異檢測技術，推進模型應用於構件辨識、結構監測等多元場景，最大化三維空間資訊於實務操作中之價值。

## 參考文獻

- DeTone, D., Malisiewicz, T., and Rabinovich, A., 2018. SuperPoint: Self-supervised interest point detection and description, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt

- Lake City, USA, pp. 224–236, DOI: 10.1109/CVPRW.2018.00060.
- Heo, Y.S., Lee, K.M., and Lee, S.U., 2010. Robust stereo matching using adaptive normalized cross-correlation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4), pp. 807–822, DOI: 10.1109/TPAMI.2010.136.
- Jocher, G., Chaurasia, A., and Qiu, J., 2023. Ultralytics YOLOv8, Available at: <https://github.com/ultralytics/ultralytics>, Accessed 20 May 2024.
- Kraus, K., 2007. *Photogrammetry- Geometry from Images and Laser Scans*, 2nd edition, Walter de Gruyter: Berlin, Germany, pp. 78–120.
- Lindenberg, P., Sarlin, P.-E., and Pollefeys, M., 2023. LightGlue: Local Feature Matching at Light Speed, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France, pp. 17627–17638, DOI: 10.1109/ICCV51070.2023.01616.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60(2): 91–110, DOI: 10.1023/B:VISI.0000029664.99615.94.
- Naufal, C., Solano-Correa, Y.T., and Marrugo, A.G., 2024. YOLO-based multi-scale ground control point detection in UAV surveying, in *Proceedings of the IEEE Colombian Caribbean Conference (C3)*, Barranquilla, Colombia, DOI: 10.1109/C358072.2023.10436290.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., 2016. You only look once: Unified, real-time object detection, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, pp. 779–788, DOI: 10.1109/CVPR.2016.91.
- Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., and Reynolds, J.M., 2012. Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications, *Geomorphology*, 179: 300–314, DOI: 10.1016/j.geomorph.2012.08.021.
- Zhang, H., Wang, X., and Li, Y., 2017. A map-based normalized cross correlation algorithm using dynamic template for vision-guided telerobot, *International Journal of Distributed Sensor Networks*, 13(9): 1687814017728839, DOI: 10.1177/1687814017728839.
- Zhao, Q., Yan, R., and Xia, L., 2024. Research on edge shape template matching algorithm based on machine vision, in *Proceedings of the 17th International Conference on Digital Image Processing (ICDIP 2025)*, Haikou, China, SPIE, Vol. 13709, DOI: 10.1117/12.3073033.
- Zuiderveld, K., 1994. Contrast limited adaptive histogram equalization. In: Heckbert, P.S. (Ed.), *Graphics gems IV*, Academic Press Professional, Inc., San Diego, CA, USA, pp. 474–485.

## Applying YOLOv8 and NCC to Assist GCP Measurement for 3D Modeling of Salt Warehouses

Lu-Chih Chen <sup>1\*</sup>      Jiann-Yeou Rau <sup>2</sup>      Ching-Jung Hung <sup>3</sup>

### Abstract

This study proposes an automated 3D modeling workflow for large cylindrical salt warehouses by integrating photogrammetry and deep learning techniques, aiming to improve modeling efficiency and measurement accuracy under harsh environmental conditions. Two industrial cameras are mounted on a central overhead crane to capture high-overlap images, which are enhanced using Contrast Limited Adaptive Histogram Equalization. Ground control point targets are detected using YOLOv8, followed by subpixel localization through Normalized Cross-Correlation. To improve the quality of aerial triangulation, LightGlue is employed for feature matching and redundant image filtering. A Metashape Python script is then used to automate bundle adjustment, error evaluation, and dense point cloud generation. Experiments conducted with 401 GCP samples show that the positioning error is approximately 2%, demonstrating the stability and accuracy of the proposed approach. The results indicate that the method effectively reduces the need for manual field measurements and improves the efficiency of salt pile volume estimation.

**Keywords:** Photogrammetry, Normalized Cross-Correlation, Deep Learning, 3D Modeling

---

<sup>1</sup> Master, Department of Geomatics, National Cheng Kung University

<sup>2</sup> Professor, Department of Geomatics, National Cheng Kung University

<sup>3</sup> General Manager, Chien Kun Surveying & Mapping Technology Co., Ltd.

\* Corresponding Author, E-mail: p66131082@gs.ncku.edu.tw

Received Date: Jul. 15, 2025

Revised Date: Jan. 21, 2026

Accepted Date: Mar. 23, 2026