

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's Thesis



整合影像辨識與 SLAM 技術建立 UAV 橋梁劣化即時
檢測系統

Real-time Bridge Deterioration Detection System Using
UAVs with Integrated Image Recognition and SLAM

陳羿清

Yi-Ching Chen

指導教授：曾惠斌 博士

Advisor: Hui-Ping Tserng, Ph.D.

中華民國 114 年 7 月

July, 2025





謝辭

完成這篇論文的過程，是一段充滿挑戰與學習的旅程。在這段日子裡，有許多人給予我鼓勵、支持與啟發。在此，我想誠摯表達我的感謝。

首先我要感謝正在閱讀這篇論文的您，因為有您的關注，讓我有動力在學術的道路上持續前行。感謝曾在我迷惘之時，引導我找到方向的人，讓我能猶豫之中做出選擇，朝著理想穩步前進。感謝那些始終在背後支持我、以日常的耐心與關懷陪伴我成長的人，是這些無聲的付出與接納，讓我能安心面對挑戰。感謝在我難過時願意陪我沉默、在我開心時一起分享的人，是這些陪伴讓日常變得柔軟，也讓壓力變得可以承受。感謝每一次願意聽我說話、和我討論甚至辯論的時刻，那些交流讓我學會反思與調整，也讓我持續進步。感謝在低潮時那句簡單的鼓勵，曾經讓我重新拾起信心；也感謝那些不順利的經驗，讓我學會接受不完美，並從中找到自己的步調。最後，謝謝曾經咬牙撐過去的自己。希望未來的我，能記得這一路走來的努力與體會，繼續踏實前行。

每一段經歷都是此刻的養分，願這份感謝能傳遞給那些不曾被我說出口的人們。謝謝你們，讓我能完成這段旅程。

謹誌

陳羿清

中華民國一〇四年七月



摘要

橋梁作為交通運輸系統中關鍵的基礎建設，其結構安全攸關公眾生命財產之保障與基礎設施之永續運作。隨著橋梁使用年限逐漸提高，結構劣化問題日益嚴重，傳統仰賴人力進行的橋梁檢測方法，存在檢測效率低落、人力成本高昂、作業環境危險及檢測結果主觀性高等諸多限制，難以因應大量橋梁設施的日常維護需求。面對此一挑戰，如何發展具即時性、自動化、低成本與高準確率之橋梁檢測技術，已成為當前土木工程領域的重要研究議題。

為解決上述問題，本研究提出一套結合深度學習影像辨識與同步定位與地圖建構（Simultaneous Localization and Mapping, SLAM）之無人機（Unmanned Aerial Vehicle, UAV）智慧檢測系統。系統架構包含前端影像蒐集、中段即時推論與定位、以及後端三維建模與資料整合等模組。前端部分以無人機搭載光學相機進行橋梁影像拍攝，並以 NVIDIA Jetson Orin Nano 作為邊緣運算平台，即時執行影像辨識及同步定位與地圖建構。

影像辨識模組採用 YOLOv9 進行多類別劣化偵測，針對常見缺陷類型如裂縫、剝落與白華進行訓練與評估，訓練資料來自公開資料集，測試資料則為實地拍攝之影像，並透過灰階轉換與超參數調整優化模型性能。同步定位與地圖建構部分，採用 ORB-SLAM3 作為核心演算法，於飛行過程中估算無人機姿態並同步建立稀疏地圖。後續則結合其輸出的關鍵幀與相機外參資訊，匯入 COLMAP 進

行稠密點雲重建。為實現缺陷資訊於三維空間中的準確標註與視覺化展示，本研究進一步整合建築資訊模型（Building Information Modeling, BIM），並透過點雲配準技術進行空間對齊，最終輸出包含缺陷位置、構件對應、TWD97 坐標與影像標註之結構化檢測報告。

實驗結果顯示，最終 YOLOv9 模型於本研究自建測試集上達到 $mAP@0.5$ 為 0.657，於道南橋進行之現地飛行測試中，可穩定完成即時辨識與定位作業，成功標註 25 處劣化並建立完整三維點雲模型，Micro F1-score 達 0.656，驗證系統整合效能與即時性優勢。整體而言，本研究所提出之系統具備高度模組化、即時運作、低成本與高擴充性等特性，能有效協助橋梁管理單位進行結構檢測與缺陷紀錄，提升傳統檢測流程之智慧化與作業效率，並提供具實務應用潛力之技術解決方案。

關鍵字：無人機、橋梁檢測、影像辨識、同步定位與地圖建構、邊緣運算、建築資訊模型



Abstract

Bridges are critical components of transportation infrastructure, and their structural integrity is essential for public safety and long-term serviceability. As aging structures face increasing deterioration, traditional manual inspections have proven inefficient, costly, hazardous, and prone to subjective errors—making them unsuitable for large-scale routine monitoring. Addressing these challenges requires an automated, real-time, and accurate inspection solution.

This study proposes a UAV-based intelligent inspection system integrating deep learning and Simultaneous Localization and Mapping (SLAM). The system consists of three modules: image acquisition via UAVs, real-time inference and localization using the NVIDIA Jetson Orin Nano, and 3D reconstruction and defect annotation. YOLOv9 is employed to detect common defects such as cracks, spalling, and efflorescence. The model is trained on public datasets and tested with field-collected images, with performance enhanced through grayscale preprocessing and hyperparameter tuning.

For localization, ORB-SLAM3 estimates the UAV's pose and generates sparse maps in real time. Keyframes and camera poses are then used in COLMAP to reconstruct dense point clouds. These are aligned with Building Information Modeling (BIM) via point cloud registration to produce accurate 3D annotations and structured inspection reports containing defect locations, component IDs, and TWD97 coordinates.

Experimental results show that the optimized YOLOv9 model achieved an mAP@0.5 of 0.657. Field tests at Daonan Bridge demonstrated real-time detection and mapping of 25 defects with a micro F1-score of 0.656. The proposed system is modular, low-cost, and scalable, offering an effective solution for automating bridge inspection and enhancing maintenance workflows.

Keywords: Unmanned Aerial Vehicle (UAV), Bridge Inspection, Image Recognition, Simultaneous Localization and Mapping (SLAM), Edge Computing, Building Information Modeling (BIM)





目次

	Page
謝辭	i
摘要	ii
Abstract	iv
目次	vii
圖次	xiii
表次	xv
符號列表	xvii
第一章 緒論	1
1.1 研究背景與動機	1
1.2 研究目的	3
1.3 論文架構與研究流程	5
第二章 文獻回顧	7
2.1 橋梁檢測定義與國內規範	7
2.1.1 檢測制度與評估指標	7
2.1.2 橋梁檢測類型分類	8
2.1.3 RC 橋梁表面劣化項目與辨識標準	8
2.2 劣化辨識技術的發展	9



2.2.1	離線影像辨識與多目標劣化檢測	9
2.2.2	即時物件檢測與多重缺陷辨識	10
2.2.3	性能比較與挑戰	11
2.3	無人機影像辨識應用於土木工程	12
2.3.1	橋梁缺陷檢測與裂縫辨識	13
2.3.2	三維測繪與缺陷可視化	13
2.3.3	災後即時應變與邊緣運算技術	14
2.3.4	技術整合與實務應用推動	15
2.3.5	系統比較與應用挑戰	15
2.4	視覺 SLAM 技術之發展	16
2.4.1	視覺 SLAM 的核心概念與主要演算法	16
2.4.2	動態環境與複雜場景下的 SLAM 應用	17
2.5	SLAM 在土木工程等領域中的應用	18
2.5.1	橋梁與隧道巡檢	18
2.5.2	室內建築與施工場域	18
2.5.3	災後快速應變	19
2.6	小結	19
第三章 系統建構方法		21
3.1	硬體與系統架構	22
3.1.1	硬體設備組成	22
3.1.2	系統流程概述	24
3.2	橋梁劣化資料集	27
3.2.1	劣化型式說明	27

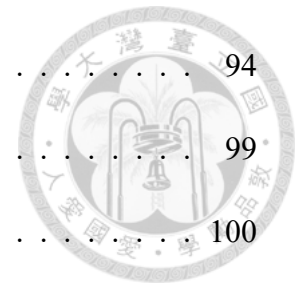
3.2.2	資料蒐集與分割	29
3.3	影像擷取與前處理	30
3.3.1	影像來源與 ROS 輸入	30
3.3.2	影像增強與轉換	31
3.4	劣化偵測：YOLOv9	33
3.4.1	模型結構概述	33
3.4.2	YOLOv9 之 ROS 模組	35
3.4.3	訓練方法與流程	37
3.4.4	評估指標	38
3.5	同步定位與地圖構建：ORB-SLAM3	39
3.5.1	系統架構概述	40
3.5.2	相機內參標定	42
3.5.3	ORB-SLAM3 之 ROS 模組	42
3.6	相機外參與時間戳對應	45
3.7	點雲獲取	46
3.7.1	稠密點雲重建 (COLMAP)	46
3.7.2	BIM 模型轉點雲	48
3.8	點雲配準與相機位姿變換	49
3.8.1	點雲配準目的與流程	49
3.8.2	相機位姿變換與更新	51
3.9	劣化資訊展示與互動應用	52
3.9.1	三維視覺化展示	53
3.9.2	三維量測與位置查驗	53



3.9.3	劣化標註表與報告書	54
3.10	系統應用情境分析	57
3.11	小結	59
第四章	系統驗證與實測	61
4.1	劣化辨識模型	61
4.1.1	實驗環境與系統設定	61
4.1.2	模型選擇與實驗設計	62
4.1.3	基礎訓練參數設定	62
4.1.4	測試集設計與評估原則	63
4.1.5	模型比較與超參數調整	64
4.1.6	模型訓練結果彙整	68
4.1.7	橋梁劣化分類性能與案例	72
4.1.8	灰階與原彩影像之辨識效能探討	75
4.1.9	結果分析與小結	77
4.2	前導實驗：系統可行性測試	79
4.2.1	驗證目的與測試條件	80
4.2.2	現地實驗紀錄	81
4.2.3	實驗觀察與問題分析	82
4.3	正式實驗：系統應用與資料取得	85
4.3.1	實驗設計與作業流程	86
4.3.2	系統運行結果	88
4.4	實驗資料後處理與結果展示：道南橋	93
4.4.1	稠密點雲重建成果	93



4.4.2	空間配準與劣化定位成果	94
4.4.3	橋梁劣化檢測報告書範例	99
4.4.4	三維劣化展示與互動	100
4.5	實驗結果與討論	101
4.5.1	實驗結果	101
4.5.2	問題與討論	102
第五章	結論與建議	104
5.1	結論	104
5.2	研究限制	105
5.3	未來研究建議	106
	參考文獻	109







圖次

1.1	研究流程圖	6
3.1	系統流程圖	24
3.2	現地施測所需硬體示意圖	25
3.3	YOLOv9 劣化偵測成果 (左) 與 ORB-SLAM3 運行畫面 (右) . . .	26
3.4	橋梁常見劣化類型範例	28
3.5	影像前處理流程圖	31
3.6	原始影像、YOLO 使用影像、SLAM 使用影像的對比示例	32
3.7	YOLOv9 之可編程梯度資訊 (左) 與廣義高效層聚合網路 (右) . . .	34
3.8	劣化辨識模組運作流程圖	35
3.9	辨識時為灰階影像 (左) 與帶有標註框的原始影像 (右)	36
3.10	YOLOv9 訓練流程圖	38
3.11	無 IMU 的 ORB-SLAM3 運作流程圖	40
3.12	SLAM 模組運作流程圖	43
3.13	ORB-SLAM3 飛行軌跡與即時劣化標註結果	44
3.14	時間戳對應方法示意	45
3.15	COLMAP 稠密點雲重建流程圖	48
3.16	BIM 模型轉點雲流程圖	49
3.17	點雲配準與轉換流程圖 (point-pair 與 ICP)	50
3.18	劣化資訊展示與對應流程圖	52
3.19	使用 PyVista 將劣化標註資訊展示於不同來源點雲	53
3.20	於 CloudCompare 中進行劣化點雲之三維互動檢查	53
3.21	點雲射線碰撞結果視覺化	54
3.22	劣化檢測報表範例	56

4.1	模型訓練過程中 Precision、Recall 及 mAP 等指標隨訓練步驟之變化	69
4.2	驗證集混淆矩陣	70
4.3	驗證集各類別 Precision-Recall 曲線，框內數值為各類別 mAP@0.5	71
4.4	不同置信度閾值下 Precision、Recall 與 mAP@0.5 之變化趨勢	72
4.5	測試集混淆矩陣	73
4.6	本研究最佳模型於現地影像多類別劣化辨識成果實例。	74
4.7	灰階與原彩影像之劣化辨識結果對照	76
4.8	寶橋位置圖，方框內為本次實驗飛行範圍	79
4.9	系統驗證現場操作實況 - 寶橋	81
4.10	ROS 節點圖 (rqt_graph)	82
4.11	寶橋實驗，系統即時運行畫面	83
4.12	寶橋空拍畫面，橋下區域存在懸空電纜、鋼索等多項障礙物	84
4.13	影像串流於飛行過程中短暫中斷，導致 SLAM 模組追蹤失敗	84
4.14	ORB-SLAM3 MapViewer，因多次 tracking loss 導致地圖不完整	85
4.15	道南橋	85
4.16	道南橋位置圖，方框內為本次實驗飛行範圍	86
4.17	系統驗證現場操作實況 - 道南橋	88
4.18	道南橋實驗，系統即時運行畫面	89
4.19	道南橋完整地圖建構成果	89
4.20	道南橋劣化辨識成果示意	90
4.21	保留與去重影像示例。左為保留之代表影像，右為被剔除之對應影像	90
4.22	道南橋實驗中即時劣化辨識之混淆矩陣	92
4.23	道南橋 COLMAP 建立稀疏模型成果	93
4.24	道南橋 COLMAP 重建之稠密點雲成果	94
4.25	道南橋之 BIM 幾何模型點雲，作為絕對座標基準	95
4.26	CloudCompare 初步點對配準畫面	95
4.27	Open3D 執行 ICP 配準成果	95
4.28	橋梁劣化檢測報告書範例	99
4.29	使用 PyVista 平台之三維劣化位置展示	100
4.30	利用 CloudCompare 工具進行點雲量測與多重校驗	100



表次

2.1	劣化辨識技術性能比較	12
2.2	UAV 應用於橋梁檢測技術比較	16
3.1	主要硬體設備之規格	23
3.2	資料集分割概況	29
3.3	劣化影像之時間戳與辨識資訊範例	37
3.4	Zenmuse X5S 相機內參 (經標定取得)	42
3.5	影格之相機姿態資訊示意	43
3.6	相機位姿與缺陷類型對應結果示意	45
3.7	劣化位置彙整表範例 (單位：公尺)	55
3.8	相機位姿資料範例 (單位：公尺)	55
3.9	不同應用場景之特性比較	58
3.10	系統流程於不同情境之應用與處理時間	58
4.1	實驗硬體與系統環境	61
4.2	YOLOv9 基礎訓練參數設定	63
4.3	不同預訓練權重對辨識效能之影響	65
4.4	不同 Batch size 設定對辨識效能之影響 (yolov9-c)	65
4.5	不同 Optimizer 對辨識效能之影響 (yolov9-c, BS = 16)	65
4.6	不同 Learning rate (lr0) 設定對辨識效能之影響 (yolov9-c, BS = 16)	66
4.7	不同 Weight Decay 設定對辨識效能之影響 (yolov9-c, BS = 16)	66
4.8	不同 Momentum 設定對辨識效能之影響 (yolov9-c, BS = 16, WD = 0.001)	66
4.9	不同平移增強 (Translation) 對辨識效能之影響	67

4.10 不同色相增強 (Hue) 對辨識效能之影響	67
4.11 不同飽和度增強 (Saturation) 對辨識效能之影響	67
4.12 不同亮度增強 (Brightness / Value) 對辨識效能之影響 (Saturation = 0.9)	68
4.13 最終 YOLOv9 訓練參數設定 (更新處以粗體標示)	68
4.14 不同置信度閾值設定下最終模型於本研究測試集之辨識效能	71
4.15 最佳模型在本研究測試集 (置信度閾值 = 0.3) 之分類效能	73
4.16 灰階影像與原彩影像輸入下各類別 mAP@0.5 指標比較	75
4.17 寶橋現地實驗記錄	81
4.18 道南橋現地實驗記錄	88
4.19 道南橋實驗即時劣化辨識之各類別 Precision、Recall 與 F1-Score . . .	92
4.20 本次實驗之劣化相機定位成果彙整表 (節錄)	96
4.21 劣化標註與 BIM 構件名稱對應結果 (節錄)	97
4.22 各構件劣化分布彙整表	98



符號列表

P	精確率 (Precision)：模型正確預測的比例
R_c	召回率 (Recall)：模型成功找出的正確目標比例
F_1	F1 分數：精確率與召回率的調和平均
mAP	平均精確率均值 (mean Average Precision)：用於整體檢測性能評估
IoU	交併比 (Intersection over Union)：預測框與真實框的重疊比例
TP	真陽性 (True Positives)：正確偵測的目標數
FP	偽陽性 (False Positives)：誤判為正樣本的數量
FN	偽陰性 (False Negatives)：漏檢的正樣本數
R	旋轉矩陣 (Rotation Matrix)：表示相機或物體姿態之方向
t	平移向量 (Translation Vector)：描述相機或物體的位置變化
T	剛體變換矩陣 (Transformation Matrix)：綜合旋轉與平移資訊



E	外參矩陣 (Extrinsic Camera Matrix)：相機從世界座標到像素座標的轉換
C	相機光學中心 (Camera Center)：相機在空間中的參考點
q_w, q_x, q_y, q_z	四元數參數 (Quaternion Components)：姿態的旋轉表示法
K	相機內參矩陣 (Intrinsic Matrix)：包含焦距與主點位置等光學參數
d	深度 (Depth)：像素點對應的距離值，常見於視差圖或點雲



第一章 緒論

1.1 研究背景與動機

臺灣地勢特殊，山區地勢陡峭，平原地狹人稠，造就了高度集中的都市發展與完善的交通建設。根據劉翠溶的研究 [1]，臺灣都市化程度自 20 世紀中期以來顯著提升，平地地區的城市開發使得人口聚集現象越發明顯。為因應密集的人口需求，部分都會區逐步發展出了密集的公路網絡與交通系統。

此外，臺灣山區開發程度高，為了連結山區與平原的交通，興建了大量的橋梁與隧道作為基礎設施。根據交通部運輸研究所的統計，截至 2025 年 6 月 6 日，臺灣現有橋梁共計 26,338 座，其中車行橋梁 22,996 座、鐵道橋梁 1,490 座及人行天橋 1,852 座，構成了全國交通網絡的重要支柱 [2]。這些基礎建設不僅滿足了交通便利性的需求，也因應了特殊地理環境的挑戰。

然而，橋梁結構長期受到自然環境如地震、颱風與水流沖刷的影響，容易出現劣化或損壞。臺灣歷年來曾發生多起因橋梁劣化導致的事故，例如 2019 年南方澳大橋坍塌事件，因鋼索鏽蝕斷裂導致橋梁坍塌，造成重大傷亡，凸顯了傳統橋梁檢測方法的不足 [3]。因此，定期檢測橋梁結構的安全性，對於確保交通暢通與公共安全至關重要。當今實務上常用的橋梁檢測方法包括目視檢測、無損檢測技術（如超音波檢測、雷達檢測）以及結構健康監測系統（Structural Health

Monitoring, SHM) [4-6]。

目視檢測雖然成本低且操作簡單，但高度依賴檢測人員的經驗，容易導致主觀判斷錯誤 [7]。無損檢測技術能夠提供較為準確的結果，但需要昂貴的專業設備，並且部分方法僅能檢測局部區域，無法全面掌握橋梁的整體健康狀態 [5]。此外，結構健康監測系統能夠提供即時數據，但其安裝成本高昂，且長期維護需求高，難以普遍應用於橋梁結構 [6]。

隨著電腦視覺與深度學習技術的快速發展，目視檢測逐漸能藉由自動化的影像辨識系統來替代傳統人工檢測，顯著提高檢測效率與準確性 [8]。深度學習技術，例如卷積神經網絡 (Convolutional Neural Networks, CNNs)，在結構裂縫檢測、損傷分類與橋梁劣化評估中表現出色，能有效處理大量影像數據並提供精確的檢測結果 [7, 8]。

同時，邊緣運算硬體的進步 (如 NVIDIA Jetson 等嵌入式系統)，使得即時運算能力得以實現在前端裝置上，大幅降低了數據傳輸延遲與計算資源需求，為橋梁檢測的自動化與高效化提供了技術基礎 [9, 10]。

在突發災害情境下，例如地震、洪水或橋梁損壞事件，如何即時掌握橋梁結構是否安全，對緊急應變極為關鍵。Mandirola 等人於災後橋梁檢測研究中指出，無人機可於事故發生後快速部署，第一時間取得橋面高解析度影像，並快速辨識潛在劣化區域，減少人力進入危險區域的風險 [11]。該研究強調了「現場即時判斷能力」在橋梁檢測任務中的重要性，特別是在災後初期階段，若能迅速提供決策依據，將有助於有效規劃搶修與交通管制。

美國國家公路與運輸官員協會 (AASHTO) 於 Asphalt Magazine 報告中探討了無人機 (Unmanned Aerial Vehicle, UAV) 在橋梁檢測中的應用，指出無人機能夠提高檢測效率、降低成本並提升安全性。例如，密西根州交通部的研究顯示，

使用無人機進行橋面檢測的成本約為傳統方法百分之 5，且所需時間大幅減少。此外，無人機還能減少交通中斷，降低使用者延誤成本 [12]。



此外，UAV 技術的普及進一步促進了自動化橋梁檢測的應用。無人機具有靈活的操作性能，能夠抵達人工難以接近的區域，並以低成本完成橋梁的全方位影像採集。在土木工程領域，無人機已被廣泛應用於結構檢測、災害評估與建築監測等場景 [13, 14]。例如，日本東日本高速道路株式會社（NEXCO 東日本）已將無人機應用於橋梁定期檢測，以減少對交通的影響並提高檢測品質 [15]。

這些技術的發展與應用背景為本研究提供了堅實的技術基礎，亦突顯即時性與現場反應能力在橋梁劣化檢測系統中的必要性。本研究因此結合無人機、電腦視覺與邊緣運算，提出能夠應對突發事件的高效檢測架構，期望能提升橋梁安全維護作業的即時性與智慧化程度。

1.2 研究目的

傳統以人力方式進行的橋梁檢測方法存在許多限制，包括高度依賴人工判斷、檢測效率低下，以及耗費大量人力與資源。這類方法往往需要檢測人員攀爬至橋梁結構的高危險區域，存在安全隱患，同時因受檢測人員經驗與視覺條件的影響，可能導致檢測結果的不一致與主觀偏差 [16]。

隨著無人機（UAV）技術的成熟，結合電腦視覺與深度學習技術的橋梁檢測方法逐漸受到關注。無人機具有靈活性與低成本的特點，可高效完成橋梁全方位影像資料的收集，而深度學習技術則能提供準確的劣化自動辨識能力，顯著提升橋梁檢測的效率與準確性 [17]。為了解決傳統方法的不足，並滿足現代橋梁檢測的需求，本研究提出結合無人機與電腦視覺技術的新檢測方法。

本研究的主要目標如下：



1. **多類別劣化辨識**：目前多數研究僅聚焦於單一橋梁劣化類型（如裂縫、混凝土剝落），對於多類別劣化的同步辨識尚屬不足 [18]。本研究將建立能夠即時識別多類別劣化的影像辨識模型，以提升橋梁檢測的適用性與精準性。

研究問題：應採用何種深度學習架構與訓練流程，以提升多類別劣化辨識之即時性與精確性？不同劣化類型之資料蒐集與模型訓練有何挑戰？

2. **三維劣化定位與可視化**：傳統方法在記錄劣化位置時，多依賴人工目視觀測或簡單的影像套疊比對技術，耗費人力且難以準確定位劣化位置。為提升定位精度，本研究將引入同步定位與地圖建構（Simultaneous Localization and Mapping, SLAM）技術，透過無人機拍攝影像建立三維點雲模型，提供可視化的檢測結果 [19]。並結合 BIM 模型中的空間資訊，令橋檢人員知悉劣化確切位置，以利後續維修與決策作業。

研究問題：如何結合 SLAM 技術與 BIM 模型，將多類劣化資訊正確投影於三維空間中，並實現自動化座標轉換與可視化呈現？在複雜橋梁環境下，三維定位與標註流程如何設計，以兼顧精度與實用性？

3. **高機動性的邊緣運算橋梁檢測系統**：傳統的結構健康監測（SHM）系統雖具備即時監測能力，但其硬體需在橋梁建設或施工後事先佈設，部署困難且機動性差，且系統建置與維護成本高昂。本研究將透過 **NVIDIA Jetson Orin Nano** 進行邊緣運算，提升即時處理能力，結合無人機巡檢以即時進行缺陷標註，並快速提供檢測結果 [10]。

研究問題：如何於嵌入式運算平台（如 NVIDIA Jetson Orin Nano）上實現即時影像處理、劣化辨識與標註？在運算資源有限下，如何平衡模型精度、推論速度與系統穩定性？

綜上所述，本研究致力於開發一套高效率、高準確率且全面的橋梁劣化檢測系統，融合多類別劣化辨識、三維劣化位置標註技術以及邊緣運算技術，以提升檢測即時性並降低運算成本，使橋梁劣化檢測作業更加智慧化與自動化。



1.3 論文架構與研究流程

本研究共分為五個章節，各章節內容安排如下：

第一章為緒論，說明橋梁檢測的重要性及傳統方法的限制，進一步闡述本研究的動機與目標。

第二章為文獻回顧，總結橋梁檢測相關領域的研究進展，包括現行檢測規範、無人機應用、深度學習模型設計、SLAM 定位方法等關鍵技術，並整理其優勢與限制作為後續系統規劃基礎。

第三章為系統建構方法，說明本研究建立之橋梁劣化即時檢測系統的整體架構，包括硬體組成、影像辨識模型訓練流程、ROS 模組整合與後處理機制設計等內容。

第四章為系統驗證與實測，分別說明模型訓練成果、前導實驗與正式實驗之流程與目的，並呈現系統於實地飛行中所獲得之辨識結果與三維標註效果，進行後處理與結果展示。

第五章為結論與建議，總結本研究成果與技術貢獻，並針對系統之潛在侷限提出改善方向與後續應用建議。

本研究之整體流程可區分為四個階段，如圖 1.1 所示。首先進行需求分析，並針對現行規範與法規及現有技術進行全面盤點與探討，以奠定後續系統設計的

基礎。再根據前述分析，依序完成資料集建立、硬體串接、ROS 模組整合與後處理流程設計。此外，模型訓練與測試亦同步進行，以形成完整之系統架構。接著透過前導實驗與正式實驗，驗證系統於實地環境下之可行性與穩定性。最終，進行後處理與成果展示，並歸納本研究的主要貢獻與未來展望。

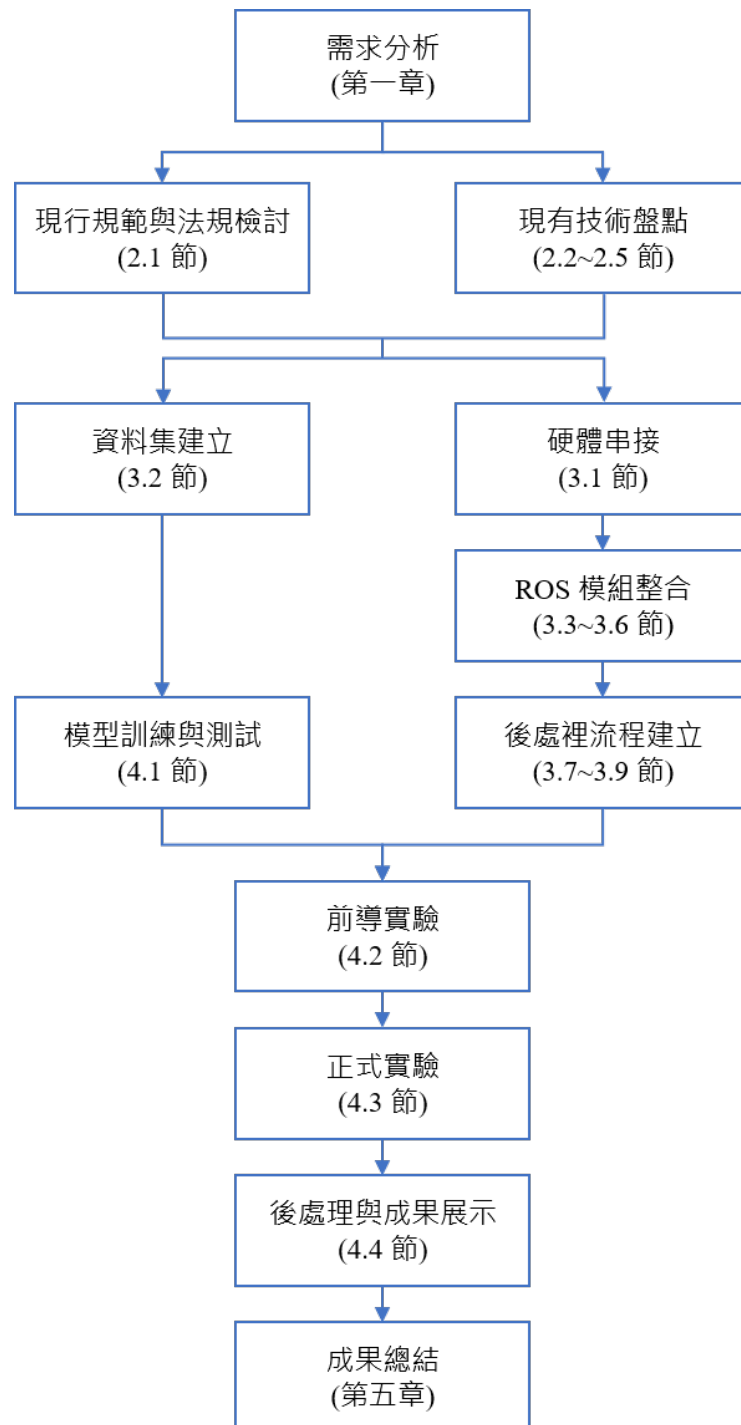


圖 1.1: 研究流程圖



第二章 文獻回顧

2.1 橋梁檢測定義與國內規範

橋梁檢測屬於專業的工程作業，為確保各單位作業一致，建立明確的檢測規範與評估標準極為重要。透過制度化的規範，不僅能指引檢測作業流程，更可作為後續缺陷辨識模型訓練與評估之基礎依據。以下將說明中華民國現行橋梁檢測制度，並針對鋼筋混凝土橋梁常見劣化項目與辨識準則進行整理。

2.1.1 檢測制度與評估指標

我國現行的橋梁檢測作業有明確的分類與定義，並訂有正式規範作為依據。例如交通部頒布之《公路橋梁檢測及補強規範》規定，橋梁檢測以目視檢查為主、儀器檢測為輔，需於定期週期以及特定時機（如重大天然災害後）對橋梁各構件進行全面檢查 [20]。定期檢測旨在及早發現橋梁構件的問題，掌握損傷劣化情況，並採取必要對策以維持橋梁結構之良好狀態。檢測結果通常以圖繪或照片詳細記錄各處劣化或損傷部位，並透過量化指標評估構件劣化程度，建立橋梁現況的基本評估資料。許多國內單位會對每一橋梁構件劣化情形給予 D 值、R 值、U 值等指標（分別代表損壞程度、修復緩急以及使用功能受影響程度），再依各構件權重計算全橋的綜合評估分數與狀況指標，作為後續維修補強優先順序的依據。



2.1.2 橋梁檢測類型分類

橋梁檢測類型依據國內規範主要可分為下列幾種 [20]：

- **日常巡查：**由養護單位人員對橋梁進行日常的巡迴檢查，以早期發現異常或損傷徵兆。巡查頻率視橋齡及橋梁等級而定，例如臺北市對橋齡 30 年以上的橋梁每天巡查一次、15~30 年每 3 天一次、15 年以下每 7 天一次 [21]。
- **定期檢測：**定期對所有橋梁構件進行例行的全面檢測評估，以確保橋梁結構安全。例如臺北市政府規定一般市區橋梁原則上每 2 年須進行 1 次結構安全檢測，重要橋梁則提高至每年 1 次 [21]。
- **特別檢測：**當橋梁遭遇特殊事件（如強烈地震、重大颱風、水患沖刷）或發現重大劣化疑慮時所進行的額外檢測，針對可能受損的構件進行深入調查與評估。

2.1.3 RC 橋梁表面劣化項目與辨識標準

在鋼筋混凝土橋梁的目視檢查中，國內規範明確定義了各種常見表面劣化類型及其評估基準 [20]。例如：

- **混凝土裂縫：**細微髮線裂縫且無滲水、無鋼筋鏽蝕痕跡者屬輕微劣化；若伴隨滲水或鏽蝕，則視為中重度劣化。
- **混凝土剝落與鋼筋外露：**如保護層脫落並露出鏽蝕鋼筋，屬嚴重劣化。
- **白華與滲水：**如混凝土表面出現白色結晶沉積伴隨水跡，則視為滲水性劣化指標。

- **構件傾斜與下陷**：若影響結構穩定與行車安全，則屬最高等級劣化。

上述國內評估準則為本研究定義**橋梁劣化辨識標準**之重要依據；各類劣化類型的判定門檻，皆參照規範所載描述與等級標準設定。此舉確保影像辨識結果與國家標準一致，具備工程實用性與政策接軌性。

2.2 劣化辨識技術的發展

橋梁結構的劣化檢測是確保基礎設施安全的重要環節。隨著人工智慧與深度學習技術的快速發展，近年來研究者紛紛嘗試利用影像辨識方法來提升檢測效率與準確度。本節將依據「離線影像辨識」與「即時物件檢測」兩大方向，回顧相關的研究成果。

2.2.1 離線影像辨識與多目標劣化檢測

早期橋梁劣化檢測主要依賴目視檢查，高度仰賴檢測人員經驗，易產生主觀誤差且效率低下。為克服這些限制，研究者開發了基於手工特徵提取的技術，如邊緣檢測與紋理分析，用於識別裂縫與剝落等缺陷 [22]。然而，這些方法在複雜環境中表現欠佳，且難以應對多樣化的劣化類型。隨著深度學習技術的進展，特別是卷積神經網絡（Convolutional Neural Networks, CNNs）的應用，劣化檢測的準確性與效率顯著提升 [8]。

Cha 等人於 2017 年提出一種基於 CNN 的橋梁裂縫檢測方法，不需手動特徵提取即可自動辨識影像中的裂縫 [8]。而 Yeum 與 Dyke 於 2015 年利用影像處理技術發展出自動偵測橋梁裂縫的方法，顯著提升了檢測效率與準確度 [7]。Zou 等人於 2019 年提出 DeepCrack，該模型結合多尺度與層級特徵來強化細小與不連續裂

縫的辨識能力。其在多個實驗資料集上皆展現出優異的像素級分割效果，為橋梁結構劣化監測提供了重要依據 [23]。另一方面，Spencer Jr. 等人於 2019 年統整了電腦視覺技術在橋梁檢測中的最新進展，強調人工智慧與機器學習在自動化監測領域未來將發揮更大潛力 [24]。

近年來，基於 Transformer 的模型在影像辨識領域表現突出。Wang 與 Su 於 2022 年開發了一種基於 Transformer 的裂縫分割模型，該模型透過多層自注意力機制捕捉影像中的長距離相依性，提升了對裂縫形態變化的適應性 [25]。相較於傳統 CNN，Transformer 在處理不同材質與光線條件下的裂縫檢測時表現更優，尤其在裂縫形態複雜或背景雜亂的情況下，能更精確地分割裂縫區域。

多目標劣化檢測方面，Hüthwohl 等人於 2019 年提出一個三階段多分類器系統：首先利用 CNN 進行初步分類，識別潛在缺陷區域；接著透過特徵提取網路細化劣化區域；最後以決策樹進行最終分類，精確辨識裂縫、剝落與銹蝕等多類混凝土缺陷，有效降低人工檢查的主觀影響 [26]。Mundt 等人於 2019 年發布了 CODEBRIM 資料集，涵蓋五種常見缺陷，並考量缺陷間的重疊與多樣性，為多目標分類提供更具挑戰性的訓練數據。該研究同時運用元學習技術最佳化 CNN 架構，使模型快速適應新型橋梁結構與損害模式 [27]。此外，CrackEmbed 方法將裂縫視為異常特徵嵌入點雲數據中，結合深度學習進行分割與檢測，顯著提升災後結構裂縫識別的精度 [28]。

2.2.2 即時物件檢測與多重缺陷辨識

即時物件檢測技術的發展為橋梁劣化檢測帶來新機遇。Redmon 等人於 2016 年提出的 YOLO (You Only Look Once) 技術，採用單一神經網路同時預測物件邊界框與類別機率，顯著提升即時檢測的效率與準確度 [29]。後續版本如



YOLOv4 引入 CSPNet 架構與馬賽克資料增強技術，在不降低速度的情況下，顯著提升了準確率 [30]。隨著技術演進，Wang 等人於 2024 年在先前版本的基礎上提出了 YOLOv9 架構，結合可編程梯度資訊 (PGI) 機制與廣義高效層聚合網路 (GELAN) 架構，使其在低延遲應用 (如無人機橋梁巡檢) 中表現更為優異 [31]。Ren 等人於 2015 年所提出的 Faster R-CNN，則代表兩階段偵測技術的主流架構，其透過 RPN (Region Proposal Network) 生成候選區域並進行分類，在準確率上表現出色，雖然推論速度略慢，卻適合高精度需求場景 [32]。

針對多重損害特徵的即時辨識，Ali 等人在 2021 年的研究證實，在 GPS 訊號受限環境下，結合無人機 (UAV) 與深度學習的 Faster R-CNN 模型，可有效進行多目標缺陷的探測與定位。此方法在無人機自主巡檢情境下能取得穩定成果，展現其在惡劣環境中的適應性與泛化能力 [33]。此外，Zhang 等人於 2020 年提出基於 YOLOv3 的單階段檢測器，並結合遷移學習與焦點損失技術，能有效偵測包括裂縫、剝落及鋼筋外露等多種類型的橋面損害，在提升檢測準確度的同時，也兼具良好的即時性與計算效率 [34]。

即時檢測技術在無人機巡檢中實現邊緣運算與即時反饋，但計算資源限制與模型運算需求將是本研究會面對的挑戰。

2.2.3 性能比較與挑戰

表 2.1 總結了劣化辨識技術特點與指標，並分析其挑戰。

雖然 Transformer、DeepCrack 等方法常應用於離線處理流程，具備優異的空間辨識能力，適合複雜裂縫特徵分析；而 YOLO 系列、Faster R-CNN 等方法則以速度見長，在多數研究中展現出良好的即時處理潛力。然而，是否能達成即時運作，仍須視實際的系統設計與部署環境而定，並非單由模型架構即可決定。

表 2.1: 劣化辨識技術性能比較

參考文獻	技術	優勢	挑戰
Wang 與 Su [25]	Transformer 分割模型	長距離特徵學習能力強	訓練成本高、需大量數據
Zou 等人 [23]	DeepCrack (多尺度 CNN)	裂縫分割精度高，適應非連續裂縫	易受背景雜訊干擾
Hüthwohl 等人 [26]	三階段分類系統	多類缺陷細分類	系統複雜、需大量計算
Mundt 等人 [27]	CODEBRIM + 元學習 CNN	多目標泛化能力佳	標註困難、資料建構成本高
Cha 等人 [8]	CNN 影像分類	自動特徵提取，結構簡單	光照敏感、依賴資料品質
Redmon 等人 [29]	YOLOv1	單階段偵測速度快	難以處理小尺寸目標
Bochkovskiy 等人 [30]	YOLOv4 + CSPNet	精度與速度平衡佳	網路架構複雜
Wang 等人 [31]	YOLOv9 + PGI/GELAN	適用低延遲場景、邊緣部署	需進一步模型精簡化
Ren 等人 [32]	Faster R-CNN + RPN	準確率高，擅長多類別辨識	推理速度較慢，不利即時應用
Ali 等人 [33]	UAV + Faster R-CNN	GPS 不足下仍可穩定辨識	對硬體與能耗要求高
Oscó 等人 [35]	深度學習於 UAV 應用總覽	多應用場景整合，強調實務潛力	模型泛化性與資料標註仍為挑戰

綜觀這些方法，儘管在準確度與多目標偵測能力上已持續進步，仍面臨數項挑戰，包括：資料集標註成本高昂、模型泛化能力不足，以及在資源受限設備上兼顧速度與精度的困難，尚未形成可即時部署的完整檢測流程。

2.3 無人機影像辨識應用於土木工程

Oscó 等人於 2021 年回顧了無人機 (UAV) 在農業、建築、災害監測等領域的深度學習應用，指出 UAV 結合 AI 具有高度自動化潛力 [35]。由於無人機具備高機動性與低成本優勢，能夠輕鬆抵達傳統人工作業難以到達的橋梁或建築結構區域，並透過深度學習技術進行劣化辨識、維修評估與監測資料的收集。相較於傳統的地面檢測與高空作業平台，無人機能顯著降低作業風險與時間成本。



2.3.1 橋梁缺陷檢測與裂縫辨識

無人機結合影像辨識技術已廣泛應用於橋梁裂縫等缺陷的檢測。王姿樺等人於 2022 年利用 Faster R-CNN 模型搭配 ResNet50 骨幹網路，對無人機拍攝的橋梁影像進行裂縫辨識。該方法在不同材質（如混凝土與鋼結構）的橋梁中均展現穩定檢測效果，平均精度（mAP）達 80.7%，其結果顯示該方法在不同材質的橋梁構造中均能保持穩定的檢測效果 [36]。張永辰等人於 2023 年採用 YOLOv4 模型進行裂縫辨識，並開發裂縫寬度計算方法，通過影像像素校準技術實現毫米級精度的測量，裂縫量測精度優於 0.22 mm，為結構安全評估提供精確數據 [37]。蔡宜真於 2024 年透過機器人作業系統（ROS）建立即時自動化檢測系統，整合無人機影像與深度學習模型，提升橋梁裂縫檢測效率與準確度 [38]。此外，饒見有等人於 2024 年採用 DeepLabV3++ 進行影像語意分割，能同時辨識剝落、鋼筋銹蝕、白華等多種劣化現象，有效減少人工介入需求 [17]。

Kao 等人於 2022 年進一步提出一套搭載雷射測距模組的無人機橋梁檢測系統，能自動辨識並量測裂縫長度與深度，並整合圖像處理與位移資料進行智慧化損傷判斷，其平均誤差在 1% 以內 [39]。該研究證實整合感測器與 AI 模型的 UAV 平台具備高精度與高效率的橋梁裂縫檢測能力，並具現地部署潛力。

2.3.2 三維測繪與缺陷可視化

無人機在三維測繪與結構建模領域亦展現出高度應用潛力。藉由搭載高解析度相機與各類感測器，無人機能快速獲取大範圍影像資料，進而生成精確的三維模型，顯著降低傳統測量所需的人力與時間成本。Ellenberg 等人於 2016 年採用無人機拍攝的熱影像來偵測橋面剝離現象，不需直接接觸結構表面即可發現內部潛在劣化，展示出紅外線影像在橋梁健康監測中的可行性與高效性 [40]。

此外，Nex 與 Remondino 於 2014 年回顧了無人機在 3D 測繪領域的應用，並強調飛行高度、相機參數及影像重疊度對三維建模精度的影響。他們指出，影像重疊度達百分之 80，並適當控制飛行高度，可有效降低模型幾何誤差；且多旋翼無人機適合局部精細掃描，而固定翼無人機則在廣域測繪中更具效率 [41]。

隨著技術進步，LiDAR 感測器亦逐步整合至無人機系統中，進一步提升三維測繪的精度與應用範疇。Bolourian 與 Hammad 於 2020 年提出一種基於 LiDAR 的路徑規劃方法，通過分析三維點雲數據來預測缺陷可能出現的位置，並結合基因演算法與 A* 演算法對自主飛行路徑進行優化。此方法不僅提升了巡檢效率與覆蓋範圍，還能針對高風險區域進行精細掃描，為結構健康監測提供更全面且可靠的數據支援 [42]。

2.3.3 災後即時應變與邊緣運算技術

針對災後應變需求，Mandirola 等人在 2022 年研究中指出，無人機能在橋梁坍塌或災後第一時間迅速部署進行損壞調查，獲取高解析度影像，減少人員進入危險區域的風險，並即時提供關鍵資訊以支援決策 [11]。Tilon 等人在 2022 年於專案中提出結合 UAV 與邊緣運算的道路與橋梁即時監測系統，整合影像擷取、現地處理與即時回傳，其回應時間低至數秒，顯著提升緊急情境下的資訊可得性與應變效率 [43]。

此外，Lin 等人於 2022 年針對橋梁結構健康監測系統提出具邊緣運算能力的感測節點，可於現場即時分析預力梁的應變數據，判別剛度變化與潛在劣化，減少資料傳輸負擔與延遲 [44]。這些研究強調「即時性」在災後橋梁檢測中的關鍵角色，並證明 UAV 與邊緣計算可作為即時應變系統的關鍵節點。

災後巡查需求不僅存在於橋梁結構，在建築物等其他基礎設施中亦同樣重

要。Jozi 等人於 2024 年針對住宅建築提出一種基於無人機的快速災後損傷評估方法，結合紋理與邊緣特徵指標進行自動分類，無需多次航拍即可辨識受損建物，準確率高達 91.3% [45]。該研究突顯 UAV 技術在災後環境中提供單次快速航拍、即時分析、全區域判斷的潛力，對於橋梁結構巡檢亦具高度借鑑價值，特別是在災後初期無法逐點人力檢查的情境下，提供快速決策支援。

2.3.4 技術整合與實務應用推動

為滿足現場即時檢測與邊緣運算需求，Hakani 與 Rawat 於 2024 年利用 YOLOv9 演算法結合 NVIDIA Jetson Nano 邊緣運算裝置，實現無人機即時偵測，展示了在邊緣平台上進行即時資料處理的能力 [46]。Chou 與 Liu 在 2024 年亦提出一個基於無人機輔助的橋梁劣化檢測邊緣運算平台，整合 AIoT 與深度學習，運用 Pilgrimage Walk Optimization (PWO)-Lite 演算法優化 YOLOv7-tiny 模型超參數，並結合 DeepSORT 演算法進行即時追蹤，為橋梁劣化檢測提供更迅速的資訊 [47]。此外，根據交通部發布的《無人機搭配 AI 影像辨識應用於橋梁檢測之研究》，此技術可顯著縮短人工檢測時間，並提升橋梁維護與管理效能 [17, 18]。

2.3.5 系統比較與應用挑戰

為比較不同 UAV 檢測系統與深度學習整合方案之特點與限制，表 2.2 彙整相關技術於橋梁檢測應用上的性能與挑戰。

整體而言，UAV 搭配不同深度學習模型與感測器技術已展現高度應用潛力。實務應用中仍需考量模型部署、電源續航、資料標註與系統穩定性等問題，未來整合低功耗 AI 模型與自動導航策略將成為關鍵趨勢。

表 2.2: UAV 應用於橋梁檢測技術比較

參考文獻	技術組合	優勢	挑戰
王姿樺等人 [36]	UAV + Faster R-CNN	穩定辨識不同材質橋梁裂縫	訓練資料需涵蓋多場景
張永辰等人 [37]	UAV + YOLOv4 + 寬度計算	即時偵測並量化裂縫尺寸	計算模型部署與精度平衡
蔡宜真 [38]	UAV + ROS + 深度學習	實現即時自動巡檢系統	系統整合度與電力續航
Kao 等人 [39]	UAV + 雷射測距 + 影像辨識	裂縫長度與深度量測誤差低於 1%	感測器整合與成本控制
饒見有等人 [17]	UAV + DeepLabV3++	多類劣化語意辨識	語意標註成本高
Ellenberg 等人 [40]	UAV + 紅外線影像	非接觸式檢測潛在剝離	對熱成像設備依賴高
Bolourian 與 Hammad [42]	UAV + LiDAR + A* 演算法	精準路徑與缺陷預測	設備成本與重量限制
Hakani 與 Rawat [46]	YOLOv9 + Jetson Nano	邊緣即時檢測能力強	資源限制下的效能優化
Chou 與 Liu [47]	YOLOv7-tiny + PWO-Lite + DeepSORT	即時追蹤與輕量模型	超參數調整與追蹤穩定性

2.4 視覺 SLAM 技術之發展

隨著無人機應用於橋梁巡檢與自主導航的需求日益增加，視覺 SLAM 技術逐漸成為構築智慧化巡檢系統的關鍵核心。本節將回顧視覺 SLAM 的基本概念與主流演算法架構，並探討其於動態環境與複雜場景中的應用挑戰與發展趨勢。

2.4.1 視覺 SLAM 的核心概念與主要演算法

視覺同步定位與地圖構建 (Visual Simultaneous Localization and Mapping, Visual SLAM)，允許機器人或無人機依賴相機視覺數據在未知環境中同時定位自身並建構地圖。在 GPS 訊號弱的環境下，如橋梁底部或隧道內，視覺 SLAM 成為無人機自主導航的核心工具。

Cadena 等人於 2016 年全面回顧 SLAM 發展，指出從傳統幾何方法如 EKF-

SLAM、Graph-SLAM，到單目與雙目視覺 SLAM 的應用演進，強調其在無 GPS 環境中的重要性，並提出未來朝「語意感知」、「多感測器融合」與「動態環境穩健性」發展的方向，為後續 SLAM 研究提供了重要的方向指引 [48]。Barros 等人於 2022 年全面回顧了三大主要視覺 SLAM 方法的代表性演算法，並從演算法類型、地圖密度、全局優化、閉環檢測、嵌入式應用支援度及實際應用案例等六項軟硬體面向，系統性整理並比較當前主流視覺 SLAM 架構，如 ORB-SLAM、DSO、LSD-SLAM、RTAB-Map 等，為後續應用挑選合適系統提供實務參考 [49]。

Rublee 等人於 2011 年所提出的 ORB 特徵點演算（Oriented FAST and Rotated BRIEF），結合了 FAST 的高效特徵點檢測和 BRIEF 的快速描述符，並針對 BRIEF 進行了旋轉不變性的改進，計算速度快且旋轉不變性強，成為即時 SLAM 的基礎 [50]。在此基礎上，Mur-Artal 等人於 2015 年提出的 ORB-SLAM 系統，採用關鍵幀（Keyframes）策略，通過特徵點追蹤、地圖建構和閉環檢測（Loop Closure）實現高精度定位，適用於單目和 RGB-D 攝影機 [51]。Campos 等人在 2021 年進一步推出 ORB-SLAM3，支援視覺-慣性 SLAM 及多地圖 SLAM，並導入了 Atlas 架構提升地圖管理機制，避免追蹤中斷或場景切換時出現大幅誤差。此外，該研究提出了一種新的地點識別方法，提高閉環檢測的效率和回溯能力，減少延遲並增強地圖合併的準確性，進一步強化在多樣化環境中的應用性。[52]。

2.4.2 動態環境與複雜場景下的 SLAM 應用

為因應動態環境與複雜場景下的實務需求，Zhang 等人於 2018 年提出了 Semantic SLAM，在 Octomap 技術中加入語意標記並降低雜訊干擾，並採用關鍵幀追蹤策略，減少動態物體對定位準確性的影響，顯著提升地圖可用性與 SLAM 整體可靠度，適用於大型室內或戶外空間 [53]。Zhong 等人在 2018 年則提出 Detect-SLAM 架構，透過深度神經網路進行物件偵測並過濾動態特徵點，有

效減少移動物體帶來的干擾，大幅提升動態場景下的定位與建圖精度；同時，Detect-SLAM 也能即時產生語意層級的環境地圖，並在特殊視角、低光照及動態模糊等極端條件下提供更穩定的目標辨識能力 [54]。



2.5 SLAM 在土木工程等領域中的應用

隨著視覺 SLAM 技術持續成熟，其應用已從機器人自主導航拓展至各類工程與環境監測之中。本節將依序介紹 SLAM 在橋梁與隧道巡檢、室內建築與施工場域、災後快速應變，以及其他工程與環境場域的應用實例，並說明各種場景下技術的優勢與限制。

2.5.1 橋梁與隧道巡檢

在橋梁與隧道等巡檢中，傳統方法往往受限於空間狹窄與 GPS 訊號衰減。Feng 等人於 2023 年提出將視覺-慣性 SLAM 與影像超解析度技術整合於無人機平台，用於橋梁裂縫的即時偵測與三維建模。實驗結果顯示，該系統在現場可快速生成帶紋理的點雲模型，將重建時間縮減約 94%，並使裂縫寬度量測精度提升 16% [55]。在隧道巡檢方面，Xu 等人在 2025 年利用 LiDAR 相機運行 SLAM 產生的關鍵幀點雲展開為平面圖，再透過深度學習模型進行裂縫分割，最終將辨識結果投影回三維空間。該方法在實驗中半徑誤差低於 3 公分，完成重建僅需 27 秒，證明了 SLAM 結合深度學習模型在封閉空間中進行高精度定位的可行性 [56]。

2.5.2 室內建築與施工場域

室內與施工現場常因空間受限與動態作業而難以部署 GPS 定位系統，SLAM 提供了一種既靈活又經濟的替代方案。Chen 等人於 2025 年開發的四足機器人平



台，結合視覺與 2D LiDAR SLAM，即時重建室內點雲並自動對齊至 BIM 模型，驗證了低成本感測器在無 GPS 室內環境中的應用可行性 [57]。另一方面，Lu 等人於 2021 年提出將主成分分析方法用於校正 SLAM 軌跡的幾何變形，並以僅需兩個對應點的最簡配準算法，將重建路徑自動對齊至 BIM，將人工干預與對位時間約減少 50% [58]。這些成果展示了 SLAM 在施工現場資訊化管理與進度監測上的實用價值，但同時也凸顯出完全自動化配準仍有待攻克的技術難題。

2.5.3 災後快速應變

在災後救援與結構損傷評估中，現場環境往往複雜且 GPS 多失效，小型無人機搭載 SLAM 技術可迅速進場並提供關鍵三維資訊。Pan 等人於 2025 年設計了一款微型 UAV 平台，結合輕量 LiDAR 與熱像儀，用於室內災區中的生還者定位與環境點雲建立，實驗顯示該系統在無 GPS 下仍能穩定重建環境並偵測熱源 [59]。

以上案例顯示，儘管視覺 SLAM 在土木及其他工程場域已取得豐碩成果，但在全自動化配準、動態干擾強健性、端到端即時性整合，以及多模態深度學習融合等方面，仍存在諸多待解決的技術難題。

2.6 小結

綜合上述文獻與我國橋梁檢測制度，可歸納出以下幾項研究趨勢與挑戰：

- **橋梁檢測規範與制度需求：**臺灣目前橋梁檢測制度以定期檢測、經常巡查與特別檢測為主軸，並輔以劣化等級評估，已建立完善的管理架構。然實務上仍高度仰賴人工操作，資料紀錄方式亦以紙本或照片為主，亟需導入自動化影像辨識與數位化管控，以提升作業一致性與效率。

- 
- **深度學習技術發展：**從傳統 CNN、Faster R-CNN 到 YOLO 系列與 Transformer 等架構，以及結合元學習與語意分割的演算法，已廣泛應用於裂縫、剝落、銹蝕等劣化型態之辨識，劣化辨識精度已顯著提升。
 - **無人機影像應用拓展：**UAV 輔以影像、三維建模（如 COLMAP）、即時邊緣運算平台（Jetson Nano 等）與即時追蹤技術，能有效突破人工作業侷限，提升檢測範圍與效率，並具備災後應變與現場決策輔助潛力。
 - **視覺 SLAM 系統演進：**從 Detect-SLAM、Semantic SLAM 到 ORB-SLAM 系列，持續提升動態場景下的定位穩定性、地圖建構品質與多源感測器融合能力，特別適用於無 GPS 環境下的結構巡檢與三維定位。
 - **即時性整合瓶頸：**先前關於物體偵測與定位相關的研究大多結合了 LiDAR、IMU 或熱像儀等感測器實現偵測與定位的同步作業，或透過後處理方式進行。未能在邊緣端以純視覺方式同步完成即時辨識、定位與建圖，難以滿足現場快速回饋需求。
 - **實務部署挑戰：**即使技術已趨成熟，實際應用仍面臨資料標註成本高、模型泛化能力不足、惡劣環境穩定性與設備資源受限等挑戰。未來須強化模型輕量化、穩健性、跨場景適應與系統整合能力，才能真正落地部署。

綜上所述，橋梁檢測需兼顧平時定期維護與突發災後快速應變兩大面向。在既有制度中，特別檢測對於突發事件後的快速掌握與判斷尤為重要，惟其現場風險高、時效性要求強，使**即時影像辨識與自動化巡檢技術**成為關鍵輔助手段。劣化辨識、無人機影像處理與視覺 SLAM 已逐步融合為橋梁檢測系統的核心架構。透過整合高精度辨識模型、UAV 與多感測資料、視覺 SLAM 三維定位能力，本研究將致力於開發兼具效率、準確與實用性的自動化橋梁劣化檢測平台，為未來基礎設施維護決策提供更全面且可靠的技術支援。



第三章 系統建構方法

前兩章已闡述了臺灣橋梁結構檢測的重要性以及現存傳統方法之侷限，並從文獻回顧中了解到無人機（UAV）結合電腦視覺與深度學習技術，在自動化橋梁檢測領域具有相當潛力。特別是在應對多種類型的劣化現象（如裂縫、剝落、白華等）時，深度學習可提供即時且準確的缺陷辨識；而視覺 SLAM 技術則能為無人機在 GPS 訊號有限或複雜的環境中，提供自動化定位與三維地圖資訊。綜合上述研究可知，若能在嵌入式平台上整合這些模組，並進一步在三維模型中標示缺陷位置，即能顯著提升橋梁檢測的效率與品質。

基於此背景，本研究著手建構一套整合式檢測系統，以 DJI Inspire2 搭載 Zenmuse X5S 雲台單眼相機拍攝影像，並以 NVIDIA Jetson Orin Nano 作為邊緣運算平台。系統於該平台上部署 Robot Operating System (ROS) 環境，以協調各模組間的資料流通，實現即時的 YOLOv9 物件偵測與 ORB-SLAM3 同步定位與地圖構建。在後處理階段，使用桌上型電腦透過 COLMAP 運行三角測量與多視角立體重建 (Multi-View Stereo, MVS) 產生稠密點雲，並透過點對標註與 ICP (Iterative Closest Point) 演算法將該點雲與帶有真實坐標資訊的 BIM 點雲對齊。最終透過 PyVista 套件將劣化偵測結果於三維空間中可視化顯示，並輸出橋梁檢測報告書。

本章將詳細介紹系統建構方法，包括硬體配置、影像與資料流的前處理方式、劣化偵測與視覺 SLAM 之核心技術，以及後處理與獲取空間資訊的步驟。



3.1 硬體與系統架構

3.1.1 硬體設備組成

為實現即時橋梁劣化檢測，本研究整合多項硬體設備，分述如下：

1. **DJI Inspire2 空拍機**：搭載 Zenmuse X5S 單眼鏡頭，深入橋檢人員難以到達的區域並蒐集影像以進行後續作業。支援 2.4/5.8 GHz 無線影像傳輸。
2. **DJI 遙控器**：接收空拍機傳回之即時影像訊號，並透過 HDMI 2.0 介面輸出影像。
3. **影像擷取盒 (AVerMedia)**：透過 HDMI 2.0 介面輸入 720p 影像至嵌入式平台，以 USB 3 連線以滿足頻寬需求。
4. **NVIDIA Jetson Orin Nano**：攜帶式邊緣運算平台，安裝 Ubuntu 20.04 + ROS Noetic，接收空拍影像並執行 YOLOv9 與 ORB-SLAM3。
5. **行動電源**：為開發板提供穩定電力 (15V 輸出)，利於在戶外環境作業。
6. **桌上型電腦**：後處理所使用之設備，搭載 Ubuntu 20.04 作業系統，提供充足的效能以運行 Python 腳本、COLMAP 與 PyVista 套件，以及 CloudCompare 等應用程式。

表 3.1 列出本研究所使用的空拍機、雲台相機及嵌入式開發板之關鍵規格，說明其在影像蒐集與運算方面的效能。可以看見 DJI Inspire2 具備相當的續航能力、穩定性與無線傳輸距離，足以因應大面積橋梁巡檢需求；搭載的 Zenmuse X5S 單眼雲台相機具備高解析度與靈活度，能夠清晰的獲取各個角落的影像；

NVIDIA Jetson Orin Nano 則具備強大的 GPU 資源，在邊緣端即可實現 YOLOv9 與 ORB-SLAM3 的即時運算，而無需仰賴雲端服務。



表 3.1: 主要硬體設備之規格

設備	關鍵規格
DJI Inspire2	• 構造形式：多旋翼無人機 • 飛行時間：最長約 27 分鐘 • 重量：3440 g • 最大起飛重量：4000 g • 無線傳輸距離：最遠可達 7 公里 • 最大可承受風速：10 m/s
Zenmuse X5S	• 重量：461 g • 傳感器：4/3" CMOS • 有效像素 2080 萬 • 解析度 5280 × 3956 (4:3) • 可控轉動範圍：俯仰 +30° 至 -90°
NVIDIA Jetson Orin Nano	• 作業系統：Ubuntu 20.04 + ROS Noetic • GPU 架構：NVIDIA Ampere (1024 CUDA 核心、32 Tensor 核心) • CPU：6 組 Arm Cortex-A78AE v8.2 64-bit 核心 (1.5MB L2 + 4MB L3) • 記憶體：8 GB LPDDR5 • 功耗：7W/15W 模式



3.1.2 系統流程概述

圖 3.1 展示本研究系統之整體資料流與功能模組。系統可分為「影像資料取得」、「即時偵測與定位」、「空間資訊獲取與缺陷可視化」三大階段，各階段在執行時彼此串聯，各自負責不同任務，細節將於後續各節分別說明。

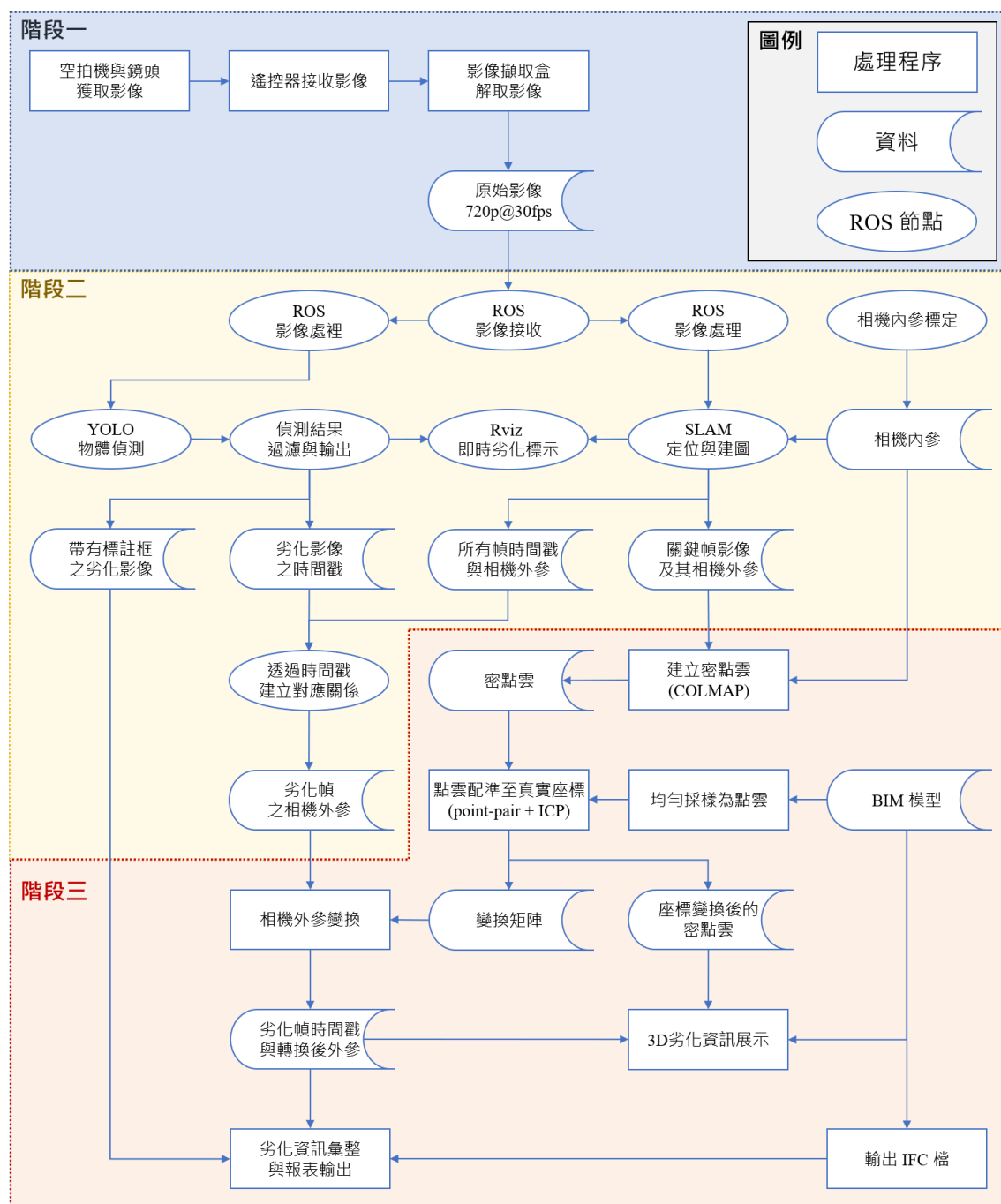
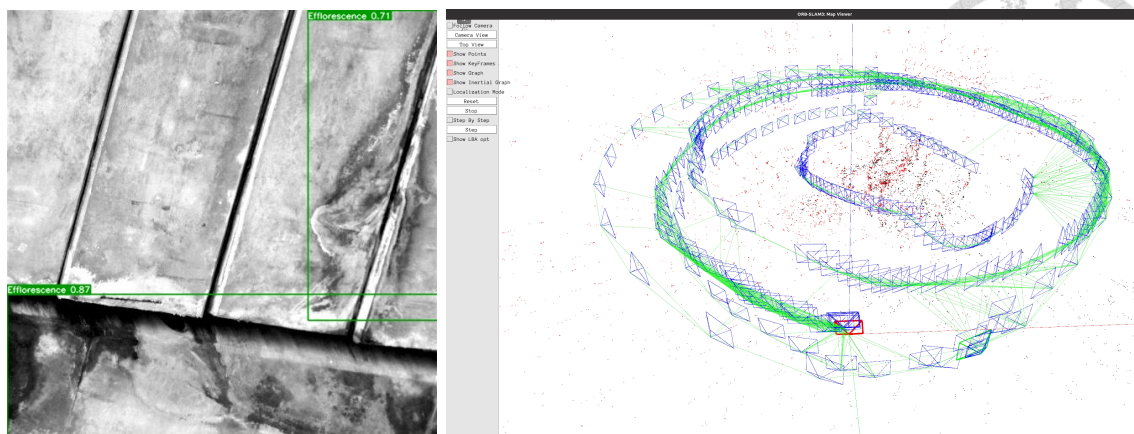


圖 3.1: 系統流程圖



圖 3.2: 現地施測所需硬體示意圖

1. 階段一：影像資料取得：如圖 3.2 所示，本系統透過空拍機（DJI Inspire2 + Zenmuse X5S）進行橋梁區域的空拍影像蒐集，影像經由遙控器及擷取盒輸入至 Jetson Orin Nano。
2. 階段二：即時偵測與定位：Jetson Orin Nano 收到影像後，將影像導入 ROS，在 ROS 環境下由 OpenCV 進行影像前處理，包括裁切、灰階轉換或對比增強，並以 ROS 話題的形式發佈。在完成影像處理後，系統同步執行兩大核心套件：
 - **YOLOv9 缺陷偵測**：偵測橋梁常見的劣化類別（裂縫、剝落、白華），輸出包含邊界框及置信度的推論結果，如圖 3.3a，並在過濾雜訊後，輸出帶有標註框的影像，並將辨識成果以文字檔的方式儲存以利後續資料處理。
 - **ORB-SLAM3 同步定位與地圖構建**：以單眼相機之影像為基礎，估計相機姿態與場景結構，並動態生成地圖，如圖 3.3b。系統會將每張影像對應的相機外參 (extrinsic) 及時間戳 (Timestamp) 作為之後三維重建與標註的基礎。



(a) YOLOv9 檢測結果示意圖

(b) ORB-SLAM3 地圖構建示意圖

圖 3.3: YOLOv9 劣化偵測成果（左）與 ORB-SLAM3 運行畫面（右）

由於此步驟直接在 Jetson Orin Nano 上即時執行，可大幅降低資料傳輸量與雲端依賴，同時保留即時性。詳細演算法流程與參數設置，請見第 3.4 與第 3.5 節。

3. 階段三：空間資訊獲取與缺陷可視化：單眼相機之影像無法透過 SLAM 即時產生稠密點雲，為了呈現缺陷在三維空間中的樣貌與位置，在邊緣運算裝置獲取初步成果後，再將資料移至桌上型電腦作進一步運算。本研究以 COLMAP 將 ORB-SLAM3 之關鍵幀（Keyframes）與對應的外參進行 MVS，建立稠密點雲；再透過 CloudCompare 與 ICP 演算法將點雲對齊至帶有真實尺度與絕對坐標資訊的 BIM 模型，取得變換矩陣並同步更新各劣化幀姿態。最後以 PyVista 在三維點雲與上疊加缺陷標註，視覺化呈現其絕對位置與結構構件對應，並輸出相關數據資訊。詳見第 3.6 節至 3.9 節。

透過以上階段的串聯，本研究系統可同時兼顧即時偵測與定位，以及高精度三維建模與缺陷標註。後續各節將針對每一部分之技術選用、實作流程與參數設定等細節進行說明。



3.2 橋梁劣化資料集

3.2.1 劣化型式說明

為因應橋梁檢測與維護之實務需求，本研究聚焦於三種類型的橋梁劣化，分別為「裂縫 (Crack)」、「剝落 (Spalling)」以及「白華 (Efflorescence)」。

根據 Hühwohl 等人提出的多階段分類架構 [26]，結合交通部於 2023 年頒布之公路橋梁檢測及補強規範 [60] 之實務指引，若能正確辨識上述三種常見缺陷，即可滿足多數橋梁檢測場景的基本需求。本研究對其定義與可能影響範圍敘述如下：

1. **裂縫**：混凝土或橋面鋪面產生的縱向、橫向、斜向或龜裂等破損狀態，包含單條裂縫與多分支裂縫。由於裂縫常是構件內部受力失衡或環境侵蝕的初期徵兆，若未及時修補，可能導致滲水、鋼筋鏽蝕與結構強度下降 [8, 26]。
2. **剝落**：混凝土表層因脫落或破損而出現的局部剝落現象，常伴隨鋼筋外露。成因包含載重疲勞、化學侵蝕或施工缺陷等。由於鋼筋外露後更易受侵蝕，故若不及時處理，剝落範圍將持續擴大，影響橋梁耐久度 [26]。
3. **白華**：混凝土基材中的可溶性鹽類隨水分遷移至表面並結晶，形成白色粉狀或結晶狀沉積物 [17]。雖不一定立即危及結構安全，但顯示結構可能存在滲漏或濕度問題，長期積累會加速其他類型劣化的產生。

圖 3.4 展示了橋梁檢測中常見的三種類型劣化：(a) 與 (b) 顯示「裂縫」，其中 (a) 為單一裂縫，(b) 則呈現較為分散且交錯的網狀裂縫；(c) 與 (d) 則顯示「剝落」，(c) 為普通剝落，而 (d) 可見鋼筋外露的情況；(e) 與 (f) 則展示「白華」，表現為混凝土表面出現鹽類結晶。



(a) 單一裂縫 (攝於福和橋)



(b) 網狀裂縫 (攝於一壽橋)



(c) 剝落 (攝於永福橋)



(d) 剝落伴隨鋼筋外露 (攝於一壽橋)



(e) 白華 (攝於永福橋)



(f) 白華 (攝於福和橋)

圖 3.4: 橋梁常見劣化類型範例

綜觀上述文獻 [26, 60]，相較於其他類型之損傷，裂縫、剝落與白華的出現頻率及可偵測性均較高，也較具即時檢測意義。後續章節將著重於此三類缺陷之自動化偵測與三維標註流程。



3.2.2 資料蒐集與分割

本研究所使用之資料集總計 30248 張影像。考量模型訓練需求以及避免過度擬合 (Overfitting)，本研究按常見做法將資料分為 訓練集、驗證集與 測試集。其中，訓練集與驗證集皆取自 Roboflow 平台上的兩組開放影像資料集 [61, 62]；測試集則分為兩組，一組取自與前述資料集相同之來源 (測試集 A)，用於初步評估模型性能，另一組則由本研究使用空拍機與手機於 2024 年 11 月至 2025 年 1 月間於雙北地區拍攝之混凝土橋梁影像 (測試集 B，當中包含萬壽橋、恆光橋、一壽橋、寶橋、福和橋、永福橋、舊中正橋、華中橋之影像)，並於 Roboflow 提供的網頁介面進行標註 [63]，以評估模型的泛用性。表 3.2 展示了各子集之影像張數、標籤數統計及其佔比。

表 3.2: 資料集分割概況

資料集	影像張數	標籤總數	白華標籤		裂縫標籤		剝落標籤	
			數量	佔比	數量	佔比	數量	佔比
訓練集	26,090	56,308	14,896	26.5%	25,546	45.4%	15,866	28.2%
驗證集	2,514	5,368	1,432	26.7%	2,465	45.9%	1,471	27.4%
測試集 A	1,447	3,318	945	28.5%	1,433	43.2%	940	28.3%
測試集 B	197	353	78	23.2%	103	29.2%	172	48.7%
總計	30,248	65,347	17,351	26.6%	29,547	45.2%	18,449	28.2%

由表 3.2 可知，本研究整合來自 Roboflow 的開放資料集及本研究拍攝影像，以涵蓋橋梁劣化檢測中最常見的三類缺陷 (白華、裂縫、剝落)。

考量標籤標註的一致性，本研究在資料合併時進行標籤歸類，原始資料集內的多種類裂縫（如縱向裂縫、龜裂等）皆統一歸類為「Crack」；剝落與鋼筋外露通常伴隨發生，因此統一歸類為「Spalling」；白華 (Efflorescence) 作為獨立類別。此舉可簡化標籤結構，亦更符合本研究的橋梁劣化檢測目標（參見第 3.2.1 節）。

本研究的測試集除了使用 Roboflow 資料，也特別加入實拍影像，涵蓋不同環境條件（光照變化、角度、影像解析度等），以確保模型可以在真實場景中適用。

值得注意的是，相較於 Roboflow 資料集，本研究拍攝的測試集中「剝落 (Spalling)」標籤比例最高（48.7%），反映真實橋梁巡檢場景中，剝落相較於其他兩者更為常見。此外，測試集中的「白華 (Efflorescence)」標籤比例最低（23.2%），顯示其在實際場景中的可見性較低，亦會受到光照條件與拍攝角度的影響。

3.3 影像擷取與前處理

3.3.1 影像來源與 ROS 輸入

本研究所使用之影像來自 DJI Inspire2 + Zenmuse X5S 空拍機。影像經由無線電頻段 (2.4/5.8 GHz) 傳送至遙控器，並透過 HDMI 2.0 介面輸出至影像擷取盒，最終透過 USB 3 介面傳輸至 Jetson Orin Nano 進行處理。其內部執行 ROS 環境，使用 ROS 的好處在於，每個功能皆能作為獨立模組，如影像擷取、轉換、發送等，彼此間以話題 (Topic) 形式進行資料傳遞，方便擴充、除錯，並讓系統在後續整合 YOLO 和 SLAM 模組時更具彈性。

影像輸入階段由 camera_publisher 節點將擷取之影像發布為 ROS Topic，格式為 sensor_msgs/Image，解析度為 720p @ 15 fps。該影像數據可由其他 ROS 節點訂閱，以進行進一步的處理與分析。



3.3.2 影像增強與轉換

影像處理在物件偵測與 SLAM 系統中扮演關鍵角色。Redmon 等人 (2016) 指出，適當的影像標準化與灰階轉換，可提升 YOLO 的辨識精度與模型泛化性 [29]；Krizhevsky 等人 (2012) 亦證實影像尺寸統一可減少計算資源消耗 [64]。另一方面，Mur-Artal & Tardós (2017) 在 ORB-SLAM2 的研究中證明，對比度增強及亮度調整可顯著提升特徵點匹配成功率 [65]，而 LSD-SLAM 研究指出直方圖均衡化與 Gamma 校正等技術同樣有利於大範圍地圖建構 [66]。

為提升物件偵測與 SLAM 系統的運行效果，camera_publisher 所發布的影像分別經由兩個 ROS 節點訂閱並進行「YOLO 影像預處理」與「SLAM 影像增強」分別進行前處理，以提升 YOLOv9 偵測與 ORB-SLAM3 的效能。整體影像處理流程如圖 3.5 所示。

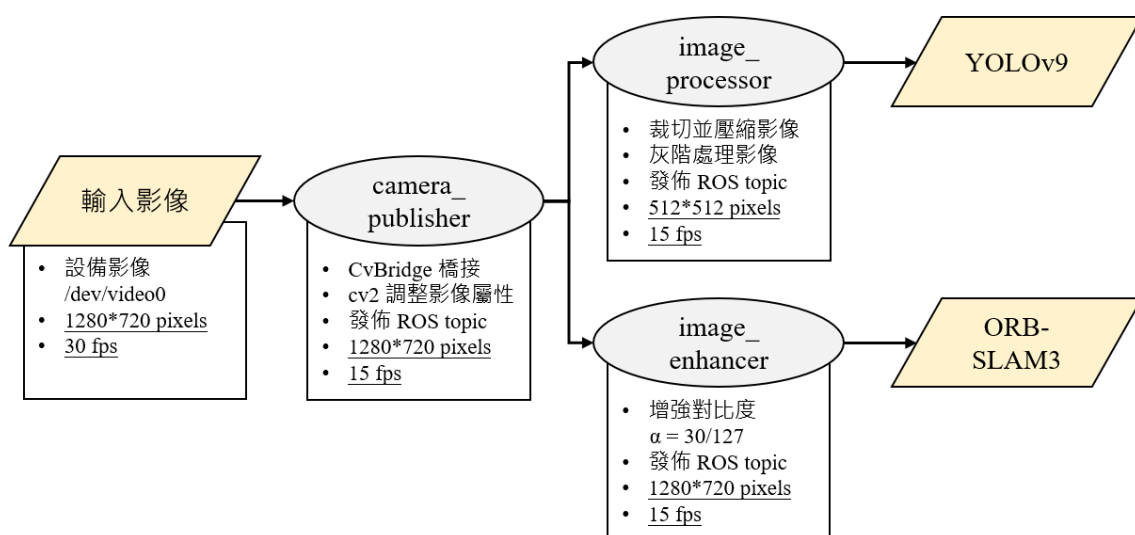


圖 3.5: 影像前處理流程圖

圖 3.6 則示範了相同橋梁影像在不同處理策略下的最終輸出：左上為原始影像，右上為 YOLO 使用的影像（灰階化與裁切），右下則是 SLAM 使用的影像（對比度增強）。

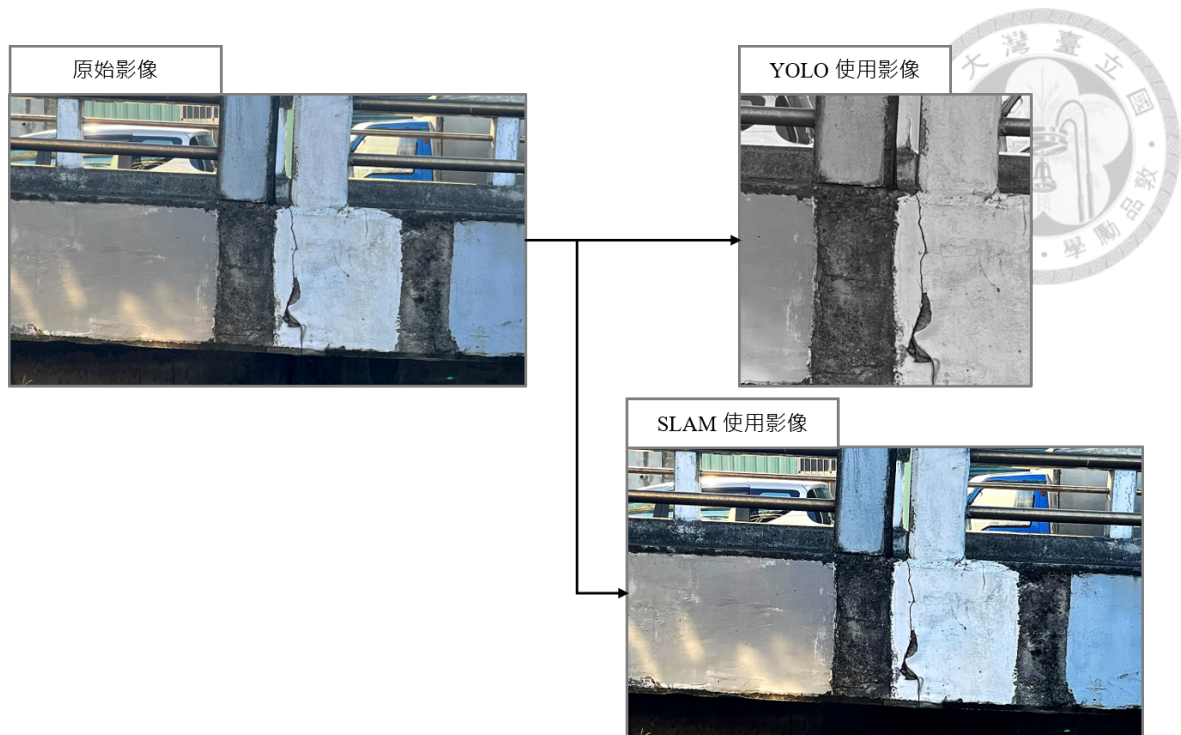


圖 3.6: 原始影像、YOLO 使用影像、SLAM 使用影像的對比示例

- **YOLO 影像預處理：**

輸入來源 由 image_processor 節點訂閱 camera_publisher 的影像，並執行裁切、灰階轉換與尺寸標準化。

灰階轉換 混凝土橋梁表面受到油漆、光線等影響，會有表面顏色不一的狀況，故將影像轉換為灰階格式，從而提升推論精度。

尺寸標準化 影像尺寸統一為 512×512 ，與訓練集一致，同時降低推論計算量 [29, 67]。

- **SLAM 影像增強：**

輸入來源 由 image_enhancer 節點訂閱 camera_publisher 的影像，並執行亮度/對比度調整、Gamma 校正及直方圖均衡化。

影像增強 能凸顯影像細節，讓影像特徵點更加清楚，使 ORB-SLAM3 的特徵點偵測更穩定，提升定位與地圖構建的精度 [65, 66]。



3.4 劣化偵測：YOLOv9

YOLOv9 (You Only Look Once, version 9) 為一套開源且高效能的單階段物件偵測演算法，由王建堯等人於 2024 年提出 [31]。其架構延續 YOLO 系列強調即時性與準確性的設計理念，並融合兩項核心創新：可編程梯度資訊 (Programmable Gradient Information, PGI) 與廣義高效層聚合網路 (Generalized Efficient Layer Aggregation Network, GELAN)，顯著提升模型於資源受限場域中的適應性與對微小目標之辨識能力。本研究將 YOLOv9 應用於橋梁劣化檢測任務中，結合 ROS 環境實現即時偵測，並於自建資料集上進行訓練與評估。本節將說明 YOLOv9 模型架構、推論流程、訓練方法與評估指標。

3.4.1 模型結構概述

YOLOv9 採用模組化設計，整體結構可劃分為 Backbone、Neck、Head 與訓練期間專用之 Auxiliary 分支。為因應現代深度學習模型在即時性與準確性間的權衡需求，其主幹架構融合可編程梯度資訊 (PGI) 與廣義高效層聚合網路 (GELAN) 兩項關鍵技術，有效強化多尺度特徵提取能力並提升模型於邊緣裝置之推論效率。以下將分別說明各模組之功能與設計細節。

圖 3.7 展示了 PGI 與 GELAN 交互強化：PGI 模組 (左) 可於資源受限場景彈性開關動態單元；GELAN 主幹 (右) 則以多層下／上採樣路徑與可逆分支交匯，提升微小目標 (如橋梁裂縫) 的辨識效能。

- **可編程梯度資訊 (PGI)**：在每個 Stage 以可學習 Gate 動態開關殘差支路，訓練時增加額外梯度路徑以減少資訊瓶頸，推論時自動裁剪冗餘計算，尤其利於邊緣裝置。

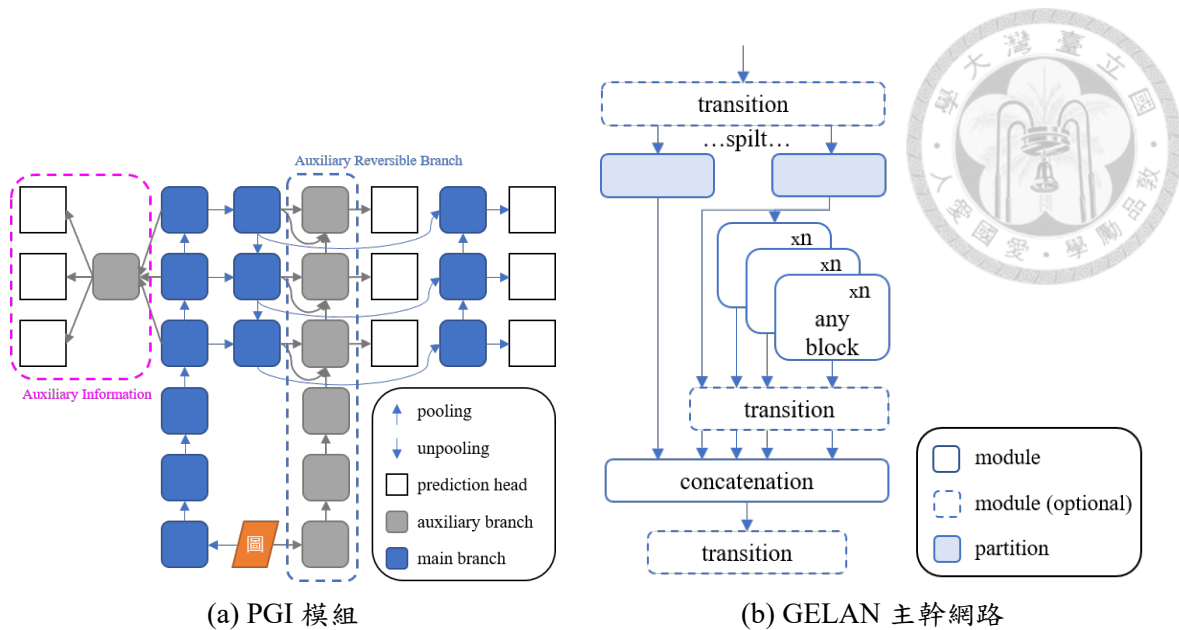


圖 3.7: YOLOv9 之可編程梯度資訊 (左) 與廣義高效層聚合網路 (右)。

- **廣義高效層聚合網路 (GELAN)**：以多分支-再聚合拓撲堆疊 *RepNCSPPELAN* 與 *ADown*，在維持輕量計算的同時提升多尺度特徵利用率；分支結構經重參數化後可於推論期折疊為單卷積核，延遲更低。

以下介紹 YOLOv9 整體架構的四大部分：

1. **Backbone**：採用 GELAN 結構，作為特徵提取的核心模組。Backbone 包含多個 *RepNCSPPELAN4* 模組，利用深度可分離卷積與殘差連接（Residual Connection）提升特徵表達能力與計算效率。此外，加入 *ADown* 模組進行空間降維與多尺度特徵提取，使模型能兼顧細節與語意層級的資訊，強化對不同大小物件的感知能力。
2. **Neck**：以 PANet（Path Aggregation Network）為基礎設計，結合自訂的 *SPPELAN* 模組，負責多層級特徵的融合與增強。Neck 利用上採樣（Upsample）與特徵串接（Concat）操作，有效整合高解析度與低解析度的特徵圖，強化物件定位的精度與分類的辨識能力。
3. **Head**：負責物件的邊界框回歸與類別分類。Head 使用多分支結構，分別預

測邊界框位置、大小及物件類別機率，並結合非極大值抑制（NMS）等後處理技術，輸出最終檢測結果。

4. **Auxiliary**：僅於訓練階段啟用，提供額外的梯度監督訊號，協助模型更快收斂並避免陷入局部最優。Auxiliary 模組通常包含輔助分類器或特徵監督層，透過多任務學習的方式提升主模型的泛化能力與穩定性。

3.4.2 YOLOv9 之 ROS 模組

本研究使用 YOLOv9 於 ROS 環境中部署即時劣化辨識模組，此模組基於 yolov9_ros 開源套件進行開發 [68]。模組運作流程如圖 3.8 所示。

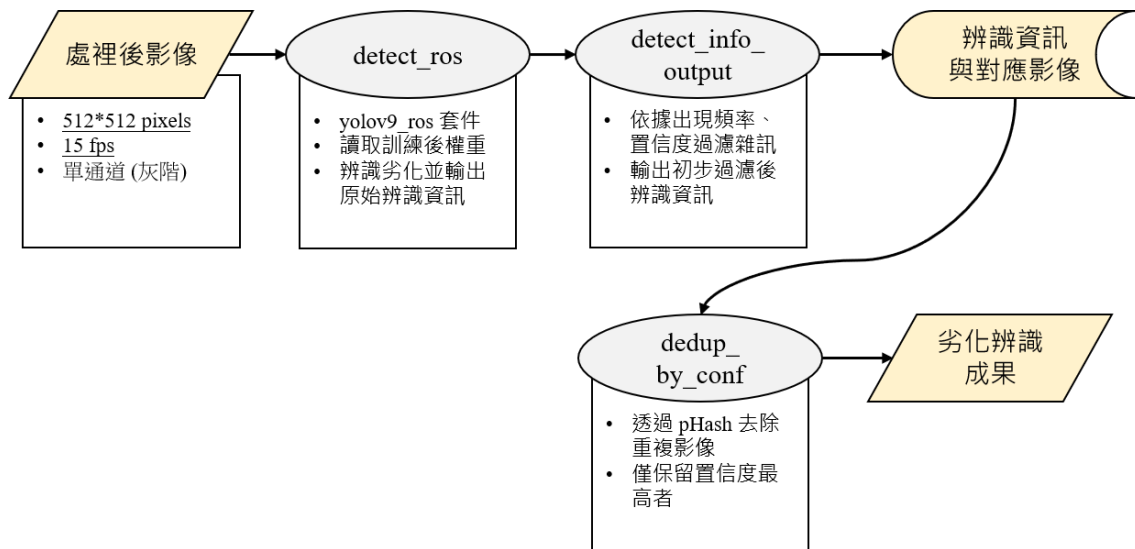


圖 3.8: 劣化辨識模組運作流程圖

本模組首先由 detect_ros 節點透過訂閱處理後之影像，進行物件偵測並即時辨識劣化情形，並發布即時辨識的初步結果。

於初步推論階段，考量到即時影像串流之影格數量龐大，且辨識模型可能產生誤檢 (false positive)，因此於 detect_info 訂閱節點中，設置了雙重過濾機制：一為置信度閾值 (confidence threshold)，二為出現頻率閾值 (frequency threshold)。僅當特定類別的偵測結果在短時間內（2 秒內）累積達到指定頻率（5 次以上），

且每次置信度皆高於設定門檻 (0.4)，才認定該次偵測成立，進而啟動影像儲存與標註程序。此外，為避免短時間內重複儲存同一劣化類型之影格，本系統設置冷卻機制，於每次成功儲存後，對該類別啟動固定冷卻期 (2 秒)，在冷卻期間再次偵測到同類型劣化不會重複儲存。

為保留劣化現象之原始彩色樣貌，系統在偵測過程中同步暫存原始影像。當劣化事件確認成立後，依據辨識結果中的標註框 (bounding box) 坐標資訊，將標註框重新繪製於對應的原始影像上，並以劣化類型與置信度作為檔名進行儲存，如圖 3.9 所示。同時，每筆劣化偵測資訊 (包含時間戳、類型、置信度、坐標點及影像檔案名稱) 皆會被記錄，供後續分析使用。



(a) 原始辨識結果

(b) 透過標註框坐標將標註框繪製於原始影像

圖 3.9: 辨識時為灰階影像 (左) 與帶有標註框的原始影像 (右)

完成初步儲存後，由於連續影格間高度重疊，易造成同一劣化特徵被重複偵測。因此系統自動呼叫去重節點 `dedup_by_conf`，進行影像去除冗餘作業。該節點基於感知哈希 (perceptual hash, pHash) 演算法進行影像相似度過濾。pHash 會將影像轉為灰階，將圖片尺寸壓縮，執行離散餘弦轉換 (DCT) 後提取低頻特徵，並生成 64 bit 長度的哈希指紋。再透過計算 Hamming distance 判斷影像間的視覺相似度，有效移除冗餘影格，僅保留置信度最高之代表性影像 [69]。

此節點藉由 pHash 方法計算每張影像的低頻視覺特徵哈希值，並將哈希距離

在閾值範圍內（本研究以 10 作為預設值）之影像歸為同一群組。於每一群相似影像中，僅保留置信度最高者作為代表，刪除其他重複影格，以減少資料冗餘並提升後續處理效率。去重後之影像將被儲存，並同步產生去重後的紀錄表，如表 3.3 所示。

表 3.3: 劣化影像之時間戳與辨識資訊範例

Timestamp	Type	Confidence	Coordinates (x1,y1,x2,y2)
1741745114.514	Efflorescence	0.712	(19, 140, 457, 275)
1741745135.269	Spalling	0.652	(89, 112, 124, 301)
1741745181.353	Spalling	0.717	(60, 12, 160, 263)
1741745203.674	Crack	0.504	(4, 141, 58, 495)
...	...	...	...

3.4.3 訓練方法與流程

本研究採用 YOLOv9 原作者於 GitHub 開源之訓練腳本 [70]，並以 yolov9 系列模型 (yolov9-m, yolov9-c) 為基礎進行訓練。訓練流程中，使用 yolov9 作者提供之預訓練權重，搭配官方提供之訓練腳本與對應的配置檔。資料集劃分為訓練集與驗證集，透過驗證集監控模型於不同 epoch 的收斂情況，並防止過度擬合。

損失函數方面，邊界框回歸採用 Complete IoU Loss (CIOU-Loss)，能同時考慮目標與預測框之中心距離、尺寸比例與重疊面積，提高定位準確性；分類則使用 Binary Cross Entropy Loss (BCE Loss) [71]，適應多標籤物件偵測需求。學習率調整方面，訓練起始階段採用 Warm-up 策略，並於訓練初期設定較高學習率以快速收斂，隨後以 Cosine Annealing 策略平滑下降，避免收斂後期因學習率震盪而影響性能穩定性。

本研究依循原著之訓練方法，整體訓練流程如圖 3.10 所示，後續驗證階段針



對部分超參數進行微調之內容詳述於 4.1 節。

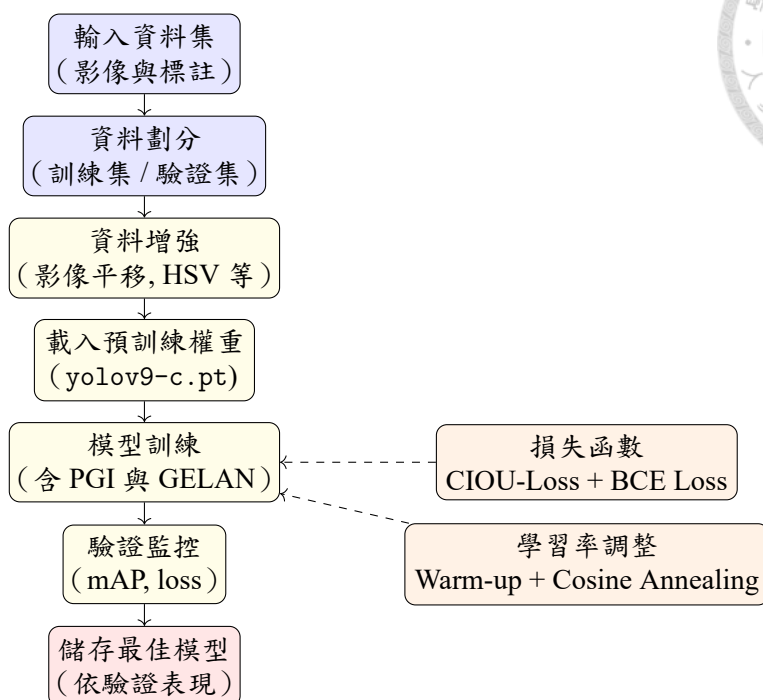


圖 3.10: YOLOv9 訓練流程圖

3.4.4 評估指標

為評估 YOLOv9 於橋梁缺陷偵測任務的表現，本研究採用以下評估指標：

- 精確率 (Precision)：

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.1)$$

其中， TP (True Positives) 為正確偵測的目標數， FP (False Positives) 為錯誤偵測的目標數。Precision 代表模型在預測為缺陷時，實際為缺陷的比例，可衡量模型的誤檢率 [72]。

- 召回率 (Recall)：

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.2)$$

其中， FN (False Negatives) 為未能偵測出的缺陷數。Recall 代表實際存在

的缺陷中，有多少成功被偵測出來，適用於衡量漏檢情況 [72]。

- F1-score：

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.3)$$

F1-score 為 Precision 與 Recall 之調和平均數，綜合考量 Precision 與 Recall，以衡量模型的整體效能。

- mAP (Mean Average Precision)：

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3.4)$$

其中， N 為所有類別數（本研究為 3）， AP_i 為第 i 類別的 Average Precision (AP)。mAP 衡量模型在不同 IOU 閾值下的精度，為物件偵測任務的重要評估標準 [72]。

- mAP@50：

$$\text{mAP@50} = \frac{1}{N} \sum_{i=1}^N AP_i^{\text{IoU}=0.5} \quad (3.5)$$

其中， $AP_i^{\text{IoU}=0.5}$ 代表在 IoU 閾值設為 0.5 的條件下計算出的 AP。本研究將 mAP@50 作為主要評估指標，因其能有效衡量模型對於橋梁缺陷的準確率，並較 mAP@0.5:0.95 更適合實務應用 [72]。

3.5 同步定位與地圖構建：ORB-SLAM3

ORB-SLAM3 是一套開源的即時同步定位與建圖 (SLAM) 系統，具備高度模組化與多執行緒設計。其系統架構包含追蹤 (Tracking)、局部建圖 (Local Mapping) 與閉環檢測 (Loop Closing) 三大模組，能有效擷取場景中具區辨性的 ORB 特徵點進行位置估計與環境重建。針對動態場景與長時間運行環境，



ORB-SLAM3 並引入多地圖管理與重定位機制，提升了系統的穩健性 [52]。本模組使用 ORB-SLAM3 搭配單眼鏡頭，以純視覺進行 SLAM 運算。本節將詳細介紹模組的架構與運作流程。

3.5.1 系統架構概述

ORB-SLAM3 採用特徵式視覺定位與建圖方法，整體流程如圖 3.11 所示。系統以 ORB 特徵點為基礎，依據輸入影像提取場景資訊，並透過前端追蹤與後端最佳化以確保定位與地圖的穩定性與精度。核心流程包含以下四大模組：

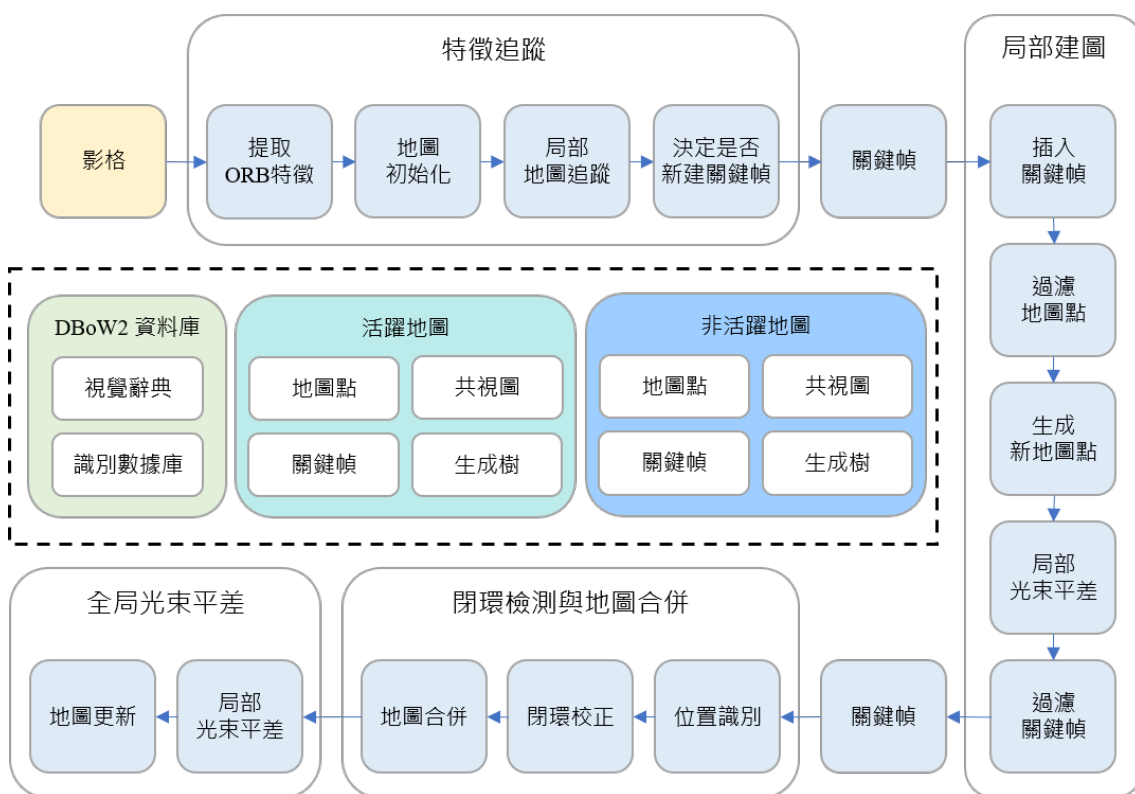


圖 3.11: 無 IMU 的 ORB-SLAM3 運作流程圖

1. **Tracking**：此階段負責從輸入影像中提取 ORB 特徵點，並進行地圖初始化。

系統會持續追蹤特徵點以進行局部地圖追蹤，並根據特徵匹配的穩定度與視角變化自動決定是否新增 Keyframe，以確保地圖的密度與精度。

2. **Local Mapping**：新建的 Keyframe 會被插入至本地地圖中，此階段會過濾

掉不穩定的地圖點，確保地圖品質。接著執行局部光束平差（Local Bundle Adjustment，Local BA）來最佳化相機姿態與地圖點的位置，以減少追蹤誤差。



3. **Loop Closing & Map Merging**：為了修正長期累積誤差，系統會進行閉環偵測（Loop Closure），當發現已經拜訪過的場景時，會進行閉環校正來調整整個地圖的拓撲結構。此外，在多段軌跡的情境下，系統還會進行地圖合併（Map Merging），以構建一致且完整的環境模型。

4. **Full Bundle Adjustment (Full BA)**：在完成閉環偵測後，系統會進一步執行全域光束平差（Full BA），此步驟會同時最佳化所有關鍵影格（Keyframes）與三維地圖點，進一步提升整體地圖的精度與一致性，尤其適用於大範圍場景的精細建模。

此外，ORB-SLAM3 亦具備完善的資料管理機制，用以協助地圖構建與歷史資訊的高效存取。其可劃分為以下三大模組：

- **DBoW2 資料庫**：用於影像回溯與重定位的識別機制，包含視覺辭典（Visual Vocabulary）與識別資料庫（Recognition Database），協助系統在動態場景中快速比對歷史影像。
- **活躍地圖（Active Map）**：儲存目前追蹤與最佳化所使用的資料，包含地圖點、共視圖、關鍵幀與生成樹等，為 SLAM 系統即時運作的核心資料結構。
- **非活躍地圖（Inactive Map）**：保存未使用或暫時脫離主要軌跡的資料，供未來重定位或地圖合併時使用，提升系統於長時間運行下的穩定性。



3.5.2 相機內參標定

為確保 ORB-SLAM3 系統能準確執行特徵點投影與三維重建，本研究首先針對所使用之 Zenmuse X5S 相機進行內參標定。標定程序於 ROS 環境下執行，透過內建工具 cameracalibrator.py 搭配棋盤格標定板與標定影像序列，計算相機之投影矩陣與畸變參數，供 ORB-SLAM3 模組初始化使用。

Zenmuse X5S 相機之標定結果如表 3.4 所示，包含影像解析度、相機矩陣（camera matrix）及畸變係數（distortion coefficients）等資訊。

表 3.4: Zenmuse X5S 相機內參（經標定取得）

參數	數值
image_width	1280
image_height	720
camera_matrix	$\begin{pmatrix} 1212.10609 & 0 & 672.41489 \\ 0 & 1209.0734 & 333.27649 \\ 0 & 0 & 1 \end{pmatrix}$
distortion_model	plumb_bob
distortion_coeffs	[0.057274, -0.063448, -0.007219, 0.010737, 0]

3.5.3 ORB-SLAM3 之 ROS 模組

本研究在 ROS 環境中部署同步定位與地圖構建模組，基於 orb_slam3_ros 開源套件進行開發 [73]。該模組整合 ORB-SLAM3 演算法，實現即時的單目視覺 SLAM 功能。為實現與 YOLOv9 模組之間的時間同步與空間對應，本研究進一步修改該套件，串接劣化資訊與相機姿態，實現劣化位置的即時三維標註功能。整體模組運作流程如圖 3.12 所示。

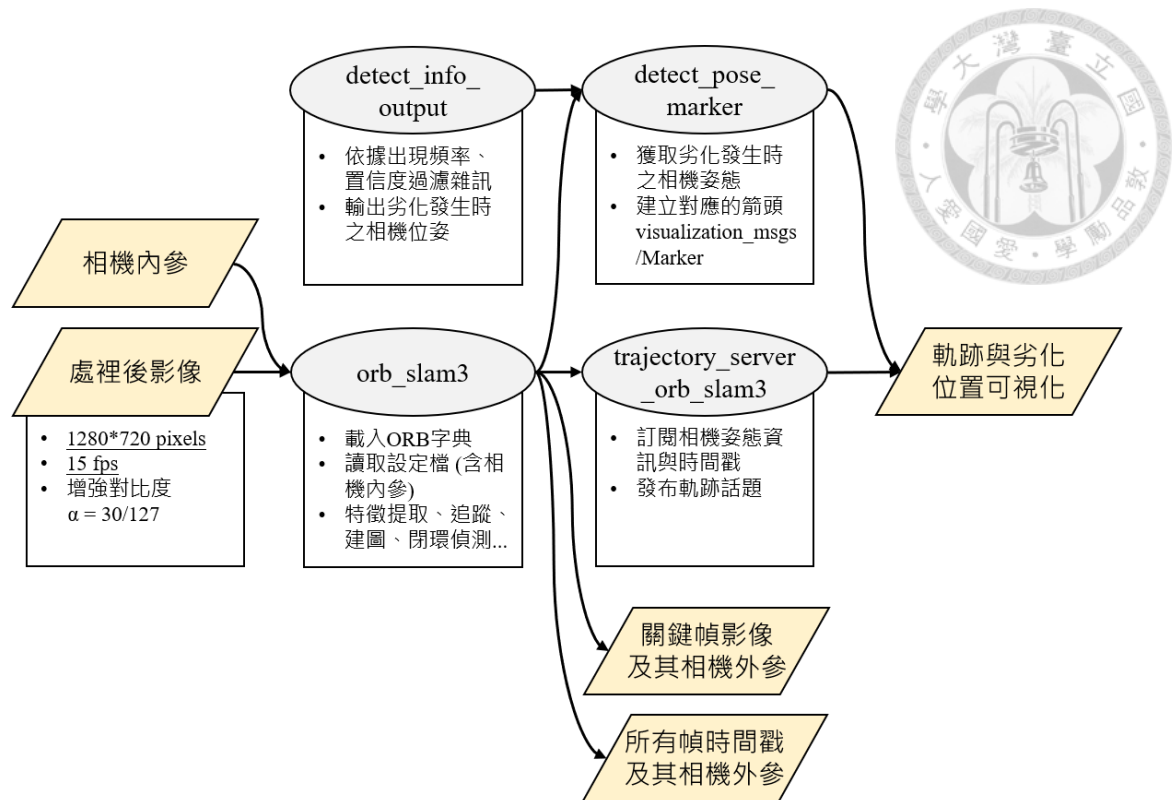


圖 3.12: SLAM 模組運作流程圖

本模組主要包含以下 ROS 節點：

- /orb_slam3：在系統初始化階段，orb_slam3 節點載入相機內參與 ORB 字典（ORBvoc.txt），並根據設定檔（.yaml）配置前述之相機內參等相關參數。運行過程中，該節點訂閱 image_enhancer 處理後之影像，執行 ORB 特徵提取、追蹤、地圖建構與閉環偵測，輸出 /orb_slam3/camera_pose、/orb_slam3/kf_markers 等資訊。在節點結束運行時，orb_slam3 將分別輸出所有影格與關鍵幀之相機姿態資訊，包含時間戳（**Timestamp**）、位置座標（**px, py, pz**）與對應的旋轉四元數（**qx, qy, qz, qw**），如表 3.5 所示。

表 3.5: 影格之相機姿態資訊示意

Timestamp (nsec)	px (m)	py (m)	pz (m)	qx	qy	qz	qw
1741748630808501058	0.956904	0.068050	0.728101	0.007236	-0.434805	-0.004299	0.900485
1741748630900607717	0.955512	0.067988	0.726118	0.007308	-0.436861	-0.004261	0.899489
1741748630012521704	0.955184	0.067916	0.725492	0.007401	-0.436859	-0.004377	0.899489
1741748630120001118	0.954531	0.067634	0.725391	0.007355	-0.437208	-0.004410	0.899319
...	...	...	...	...	...	...	...

- /trajectory_server_orb_slam3：訂閱相機姿態資訊與時間戳，發布完整軌跡於 /orb_slam3_ros/trajectory 話題，供 rviz 與後續模組使用。

以下節點被用於整合 YOLOv9 劣化資訊與 ORB-SLAM3 相機姿態，實現即時劣化定位：

- /detect_info_output：整合 detect_info 中 YOLOv9 劣化資訊與 ORB-SLAM3 相機姿態資訊，輸出至 /deterioration_camera_pose。
- /defect_pose_marker：根據上述輸出，繪製三維坐標系中的劣化位置。如圖 3.13 所示，ORB-SLAM3 成功建構出完整無人機飛行軌跡（綠線），其中每個綠點為關鍵幀拍攝位置，並以箭頭標示對應劣化位置與方向。這些箭頭為 /defect_pose_marker 所輸出的即時劣化定位結果，有助於理解劣化相對於橋梁結構之空間關係。

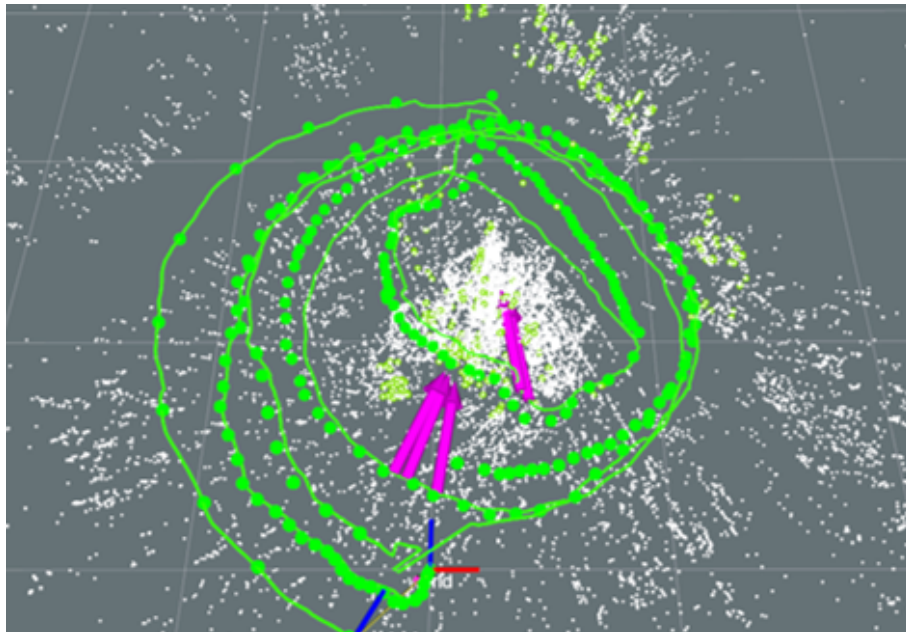


圖 3.13: ORB-SLAM3 飛行軌跡與即時劣化標註結果



3.6 相機外參與時間戳對應

為實現缺陷位置的三維標註，本研究將 YOLOv9 輸出之劣化辨識資訊與 ORB-SLAM3 輸出之相機姿態進行整合。YOLOv9 最終輸出的偵測結果中，均附帶時間戳 (Timestamp) 如表 3.3 所示，用於與 ORB-SLAM3 輸出之相機姿態 (表 3.5) 當中的資訊進行對齊。雖然劣化偵測與 SLAM 兩模組在開發版上同步運行，但因為資料處理的時間差，所以紀錄的時間戳會有毫秒級的微小差異，足以被忽略 (± 1 幀 ≈ 66 ms@15 fps)。因此可透過時間戳的對應，從 SLAM 輸出的外參資料中找出時間差最小者，作為劣化影格之相機姿態。圖 3.14 展示了時間戳對應的方法，表 3.6 展示了劣化類型 (Type) 與對應的相機姿態。

YOLO 輸出			SLAM 輸出	
時間戳	劣化類型	置信度, 角點	時間戳	外參 (x, y, z, qx, qy, qz, qw)
1741731989.604	白華	0.65, 25.3 ...	1741731989.598	
1741732095.791	裂縫	0.75, 64.0 ...	1741731989.608	0.78, 0.35, 0.86, 0.14,
1741732114.514	剝落	0.81, 68.3 ...	1741731989.619	
.....			...	...
			1741732095.780	
			1741732095.790	0.22, 0.65, 0.07, 0.98,
			1741732095.801	
			1741732095.811	

圖 3.14: 時間戳對應方法示意

表 3.6: 相機位姿與缺陷類型對應結果示意

px	py	pz	qx	qy	qz	qw	Type
0.7712	0.1178	1.2943	0.2060	-0.8152	0.2160	0.4964	Efflorescence
0.6320	0.1015	1.2354	0.2102	-0.8109	0.2232	0.4984	Spalling
0.7243	0.1503	1.5626	0.1802	-0.8549	0.2174	0.4352	Crack
0.6573	0.1320	1.6206	0.1789	-0.8546	0.2134	0.4383	Spalling
...	...	...	...	...	...	...	...



3.7 點雲獲取

本節說明本研究中所使用之點雲資料來源、生成流程與應用情境。本研究使用之點雲分為兩類：其一是由 UAV 空拍影像與 SLAM 輸出之相機外參，透過 COLMAP 重建而成之稠密點雲 (dense point cloud)，帶有相對位置資訊；其二是由 BIM 模型轉換而得之幾何點雲，帶有絕對位置的坐標資訊。

由於稠密點雲重建過程涉及大量圖像處理與三維重構計算，為確保運算效率與成果品質，後續流程將透過高效能桌上型電腦執行，而非於開發板上進行。以下將依點雲來源依序說明其建構方式與應用場景。

3.7.1 稠密點雲重建 (COLMAP)

單眼相機於 SLAM 任務中雖可實現同步定位與建圖，但所構建的地圖僅為稀疏特徵點雲 (sparse point cloud)，難以提供完整的場景幾何資訊。為補足視覺化需求並提供三維標註與比對所需之幾何基礎，本研究透過 COLMAP 進行稠密點雲重建。

COLMAP [74] 為一套功能完整的三維重建系統，整合了運動重建結構 (Structure-from-Motion, SfM) [75] 與多視角立體重建 (Multi-View Stereo, MVS) [76] 技術，能從多張影像及其相對位姿中自動重建場景的三維點雲，並提供點雲、相機參數與模型輸出等結果以供後續分析使用。

本研究採用已知之相機姿態進行稠密點雲重建，其流程如圖 3.15 所示，主要包含以下步驟：

1. **SLAM 運行與資料擷取**：透過 ORB-SLAM3 執行同步定位與建圖 (SLAM)，

擷取關鍵幀影像 (Keyframes) 及其相機外參 (旋轉四元數 (q_x, q_y, q_z, q_w) 與相機中心 C)。



2. **相機內參標定**：使用 ROS 工具完成相機內參 (intrinsics) 標定，包括相機矩陣與畸變係數，詳見 3.5.2。

3. **位移向量計算**：將 SLAM 提供之旋轉四元數轉換為旋轉矩陣 $\mathbf{R}$ ，並利用相機中心 $\mathbf{C}$ 計算 COLMAP 所需之平移向量 $\mathbf{t}$ 。其轉換關係式如下：

$$\mathbf{t} = -\mathbf{RC} \quad (3.6)$$

此關係可確保世界坐標系與相機坐標系的一致性。

4. **建立稀疏模型骨架**：將上述計算得到之 $(\mathbf{R}, \mathbf{t})$ 與相機內參匯入 COLMAP，建立與 SLAM 坐標系對齊的 sparse model。

5. **特徵擷取與匹配**：使用 COLMAP 的 feature_extractor 與 exhaustive_matcher 模組，從對齊後的影像中萃取特徵描述子並匹配，將結果寫入資料庫。

6. **三角化重建稀疏點雲**：透過 point_triangulator 模組，結合 sparse model 與特徵匹配結果，三角化生成稀疏點雲。

7. **多視角立體重建 (MVS)**：在執行 image_undistorter 影像去扭曲預處理後，使用 patch_match_stereo 模組估算稠密視差圖，再透過 stereo_fusion 模組融合多視角點雲，濾除不一致匹配，最終將稠密點雲輸出。此結果提供完整場景幾何資訊，後續可用於缺陷分析與 BIM 點雲配準。

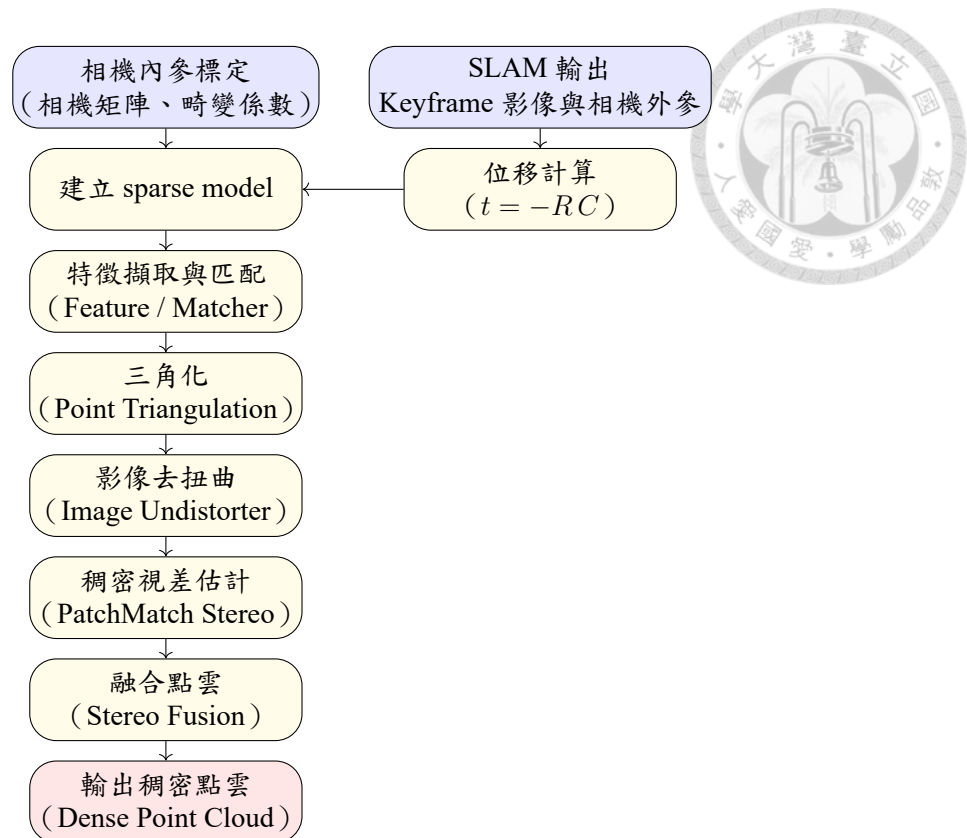


圖 3.15: COLMAP 稠密點雲重建流程圖

3.7.2 BIM 模型轉點雲

前述 COLMAP 方法以純視覺方式重建密點雲，僅能獲取三維空間中的相對位置資訊。倘若導入帶有真實世界坐標資訊與構件資訊之 BIM 模型，方能進一步提升系統的實用性。為方便進行後續點雲配準作業，會將 BIM 當中的實體採樣為點雲形式，以利後續運算。本研究之 BIM 點雲轉換流程如圖 3.16 所示，主要包含以下步驟：

1. **模型匯出**：將 BIM 軟體中的模型以中性格式（如 IFC 或 OBJ）匯出，使得模型包含每個構件的幾何資訊與世界坐標。
2. **轉換為網格模型**：將 IFC 或 OBJ 檔案匯入 Blender，轉換為三角網格模型（.ply 檔）。
3. **網格模型轉點雲**：使用 Python 與 Open3D 函式庫將 Blender 匯出的三角網格

模型 (.ply 檔) 轉換為幾何點雲。本研究採用均勻採樣 (uniform sampling) 方式，於網格表面隨機生成點，產生密度適中的點雲資料。

4. 資料儲存與格式轉換：將轉換後之點雲以 .ply 格式儲存，以利後續與稠密點雲進行配準 (registration)。

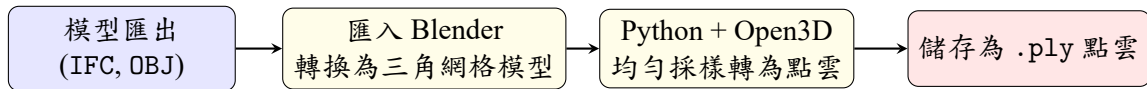


圖 3.16: BIM 模型轉點雲流程圖

透過上述轉換流程，BIM 模型得以點雲表示形式，保留了原始模型的空間佈局與尺度資訊，後續亦能與實測取得之稠密點雲進行配準。

3.8 點雲配準與相機位姿變換

本節說明如何將由 SLAM 與 COLMAP 重建之稠密點雲，與 BIM 模型所轉換之幾何點雲進行空間配準 (point cloud registration)，並同步計算相機位姿於真實坐標系中的轉換關係。透過配準與姿態變換，可將劣化偵測結果映射至 BIM 模型空間，作為後續缺陷比對與定位之基礎。

3.8.1 點雲配準目的與流程

由於 SLAM 與 COLMAP 所重建之點雲為相對坐標系，與 BIM 模型所使用之絕對坐標系不一致，需透過配準將兩者轉換至統一空間。其目的包括：

- 建立稠密點雲與 BIM 點雲之間的空間對應關係。
- 求解旋轉矩陣 $\mathbf{R}$ 、平移向量 $\mathbf{t}$ 與縮放係數 s ，實現相機位姿與缺陷點坐標的轉換。

- 將所有劣化偵測結果映射至絕對坐標系，供後續巡檢使用。

本研究所採用之配準流程如圖 3.17 所示，包含以下步驟：

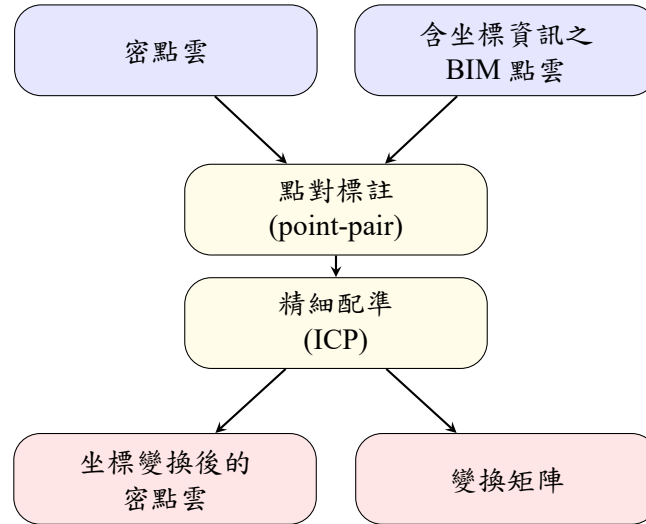


圖 3.17: 點雲配準與轉換流程圖 (point-pair 與 ICP)

1. 輸入點雲準備：一組為 COLMAP 重建之稠密點雲，另一組為由 BIM 模型三角網格均勻採樣轉換而得之點雲（含真實坐標資訊）。
2. 初步配準：於 CloudCompare 應用程式 [77] 中手動標註對應點對 (point-pair registration)，取得粗略剛體變換初始值，記為矩陣 $\mathbf{T}_{\text{initial}}$ 。
3. 精細配準 (ICP)：透過 Open3D 套件 [78] 執行 ICP (Iterative Closest Point) 演算法 [79]，採用 point-to-point 模式，以最小化對應點雲距離來求解最佳剛體變換矩陣 $\mathbf{T}_{\text{ICP}}$ ：

$$\mathbf{T}_{\text{ICP}} = \begin{bmatrix} \mathbf{R}_{\text{ICP}} & \mathbf{t}_{\text{ICP}} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3.7)$$

4. 最終整合變換矩陣計算：將初始配準矩陣 $\mathbf{T}_{\text{initial}}$ 與 ICP 精細配準矩陣 $\mathbf{T}_{\text{ICP}}$ 透過矩陣相乘方式整合成最終變換矩陣 $\mathbf{T}_{\text{total}}$ ：

$$\mathbf{T}_{\text{total}} = \mathbf{T}_{\text{ICP}} \cdot \mathbf{T}_{\text{initial}} = \begin{bmatrix} \mathbf{R}_{\text{total}} & \mathbf{t}_{\text{total}} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3.8)$$



3.8.2 相機位姿變換與更新

完成點雲配準後，可將所得剛體變換矩陣 $\mathbf{T}$ 套用至表 3.6 當中的相機外參，將相機位置與方向轉換至 BIM 模型所使用之世界坐標系。本研究使用四元數 (quaternion) 與旋轉矩陣相互轉換方式進行坐標更新，其步驟如下：

1. 原始外參重建：由每筆資料中之四元數 (q_x, q_y, q_z, q_w) 計算旋轉矩陣 $\mathbf{R}$ ：

$$\mathbf{R} = \begin{bmatrix} 1 - 2(q_y^2 + q_z^2) & 2(q_x q_y - q_z q_w) & 2(q_x q_z + q_y q_w) \\ 2(q_x q_y + q_z q_w) & 1 - 2(q_x^2 + q_z^2) & 2(q_y q_z - q_x q_w) \\ 2(q_x q_z - q_y q_w) & 2(q_y q_z + q_x q_w) & 1 - 2(q_x^2 + q_y^2) \end{bmatrix} \quad (3.9)$$

2. 齊次外參矩陣重建：旋轉矩陣 $\mathbf{R}$ 與平移向量 $\mathbf{t}$ 組成 4×4 齊次外參矩陣：

$$\mathbf{E} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3.10)$$

3. 應用剛體變換：將配準所得變換矩陣 $\mathbf{T}$ 左乘於每筆外參 $\mathbf{E}$ ，得到更新後之姿態：

$$\mathbf{E}' = \mathbf{T} \cdot \mathbf{E} \quad (3.11)$$

4. 旋轉矩陣轉回四元數：由 $\mathbf{E}'$ 擷取旋轉部分 $\mathbf{R}'$ ，透過以下公式轉回四元數：

$$s = \frac{1}{2\sqrt{\text{tr}(\mathbf{R}') + 1}} \quad (3.12)$$



$$\begin{cases} q'_w = \frac{1}{4s} \\ q'_x = (\mathbf{R}'_{21} - \mathbf{R}'_{12})s \\ q'_y = (\mathbf{R}'_{02} - \mathbf{R}'_{20})s \\ q'_z = (\mathbf{R}'_{10} - \mathbf{R}'_{01})s \end{cases} \quad (3.13)$$

5. **輸出格式整理**：最終將每筆更新後相機姿態 $(x', y', z', q'_x, q'_y, q'_z, q'_w)$ 以及對應缺陷類型整合輸出成 CSV 檔。

此轉換流程已整合至 Python 程式中實作，確保每一筆劣化標註點皆可精確對應至真實坐標，後續可作為三維可視化、比對分析與巡檢引導之依據。

3.9 劣化資訊展示與互動應用

在完成點雲對齊與相機位姿轉換後，本研究進一步整合 BIM、相機外參與影像資料，進行視覺化、構件對應與互動查驗，產出三種輸出成果進行整合與展示，協助使用者掌握劣化分布、對應影像與所屬構件。整體流程如圖 3.18 所示。

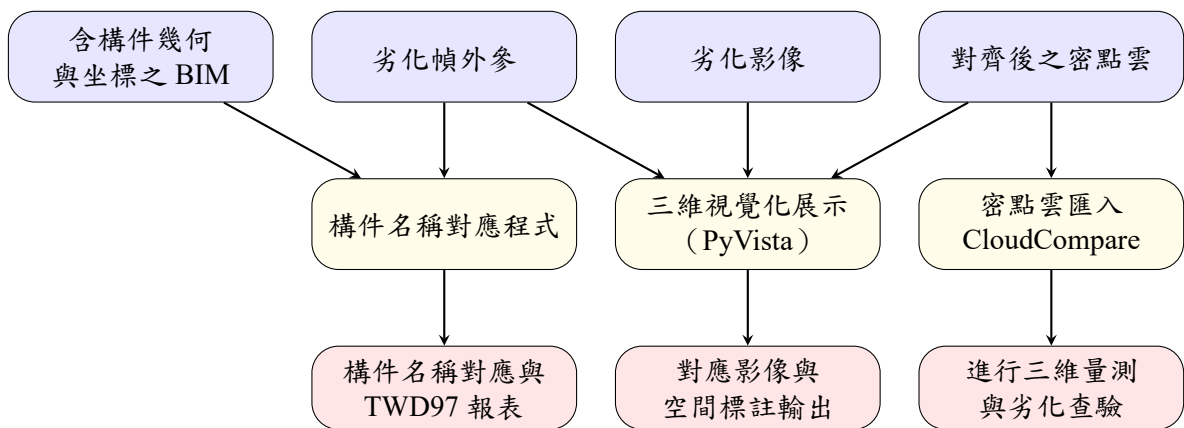


圖 3.18: 劣化資訊展示與對應流程圖

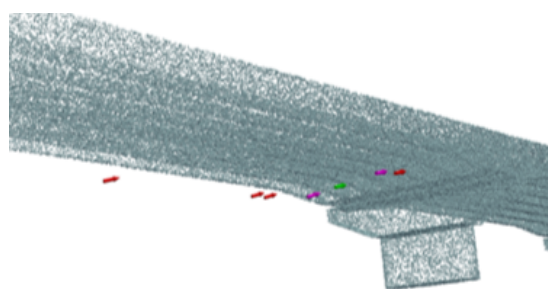


3.9.1 三維視覺化展示

於 Python 環境中，本研究採用 PyVista 套件 [80]，將轉換後的稠密點雲、BIM 點雲與劣化標註資訊整合顯示。輸入資料包括稠密點雲與 BIM 幾何點雲、每筆劣化的相機外參與時間戳等，最終可於同一視窗中以箭頭呈現劣化位置，並搭配文字與影像資訊進行空間查驗。圖 3.19 顯示了展示結果於稠密點雲與幾何點雲上的應用，劣化點以箭頭與標註框呈現於對應位置。



(a) COLMAP 建立之稠密點雲及劣化標註

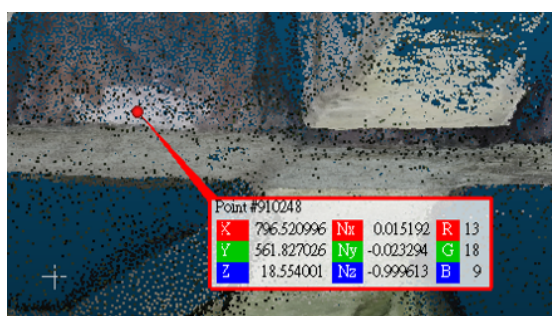


(b) BIM 模型轉換之點雲及劣化標註

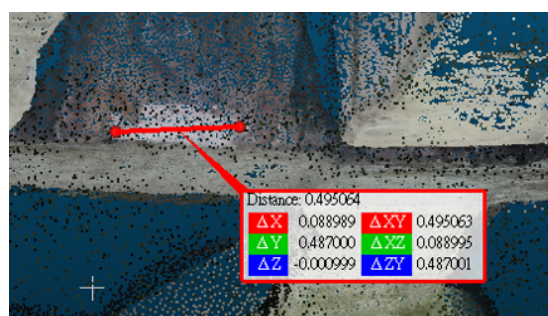
圖 3.19: 使用 PyVista 將劣化標註資訊展示於不同來源點雲

3.9.2 三維量測與位置查驗

針對細部位置與距離分析，本研究亦支援透過 CloudCompare 進行互動式量測與定位，如圖 3.20 所示。使用者可點選任一點雲點查看坐標、法向量與 RGB，亦可量測兩點間之空間距離及軸向分量，協助缺陷範圍評估與修補計畫制定。



(a) 查看劣化坐標資訊



(b) 量測劣化尺度

圖 3.20: 於 CloudCompare 中進行劣化點雲之三維互動檢查



3.9.3 劣化標註表與報告書

系統針對每一筆劣化標註結果，整合其三維空間位置、影像檔名與所屬構件資訊，並輸出結構化表格。首先，系統載入 IFC 模型，透過 ifcopenshell 將模型中 IfcProduct 的幾何頂點轉換至世界座標系，並建立 KDTree 索引結構，以利於快速進行劣化點與構件之最近鄰查詢 [81]。

為進一步驗證劣化點與構件的空間對應關係，系統也實作射線碰撞偵測 (ray tracing)。此方法從每筆相機外參提供的位置與方向產生射線，並於 KDTree 建立的點雲資料中逐步檢查距離，若某採樣點與最近的點雲頂點距離小於設定閾值，即判定為命中。系統隨後記錄命中位置，並將結果進行視覺化，圖 3.21 展示相機外參產生之射線（顏色依劣化類型區分）及其對應編號。

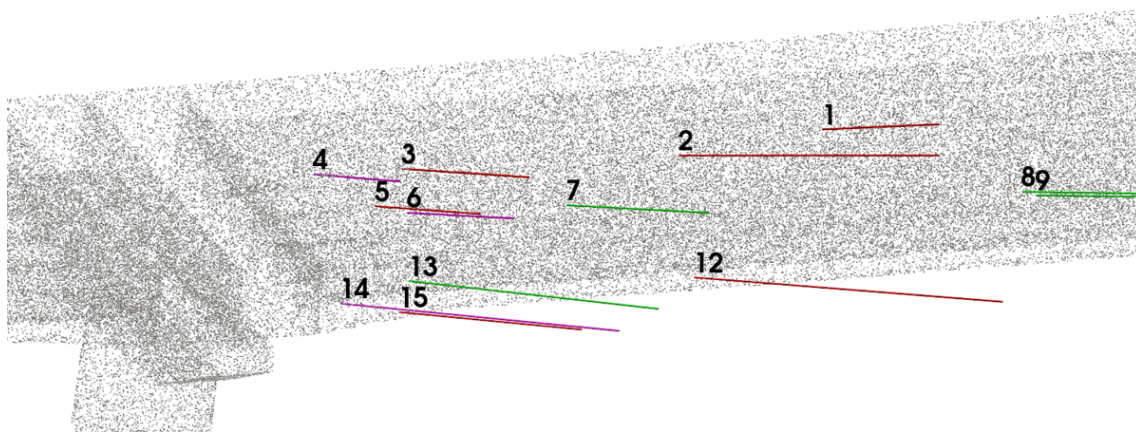


圖 3.21: 點雲射線碰撞結果視覺化

完成構件比對與射線碰撞分析後，系統將命中點的空間位置轉換為臺灣常用的 TWD97 平面坐標系，並整合構件名稱、劣化類型等資訊輸出完整報表（如表 3.7）。報表記錄每筆劣化點的流水編號、所屬構件、劣化類型，以及其在 TWD97 坐標系下的東坐標 $E(X)$ 、北坐標 $Y(N)$ 與高程 (Elevation)，以提供後續工程分析及定位使用。

表 3.7: 劣化位置彙整表範例 (單位：公尺)

Defect ID	Element	Type	E(X)	N(Y)	Elevation
1	Span2_MainGirder2	Efflorescence	307780.83	2764564.17	16.33
2	Span2_DeckSlab	Spalling	307795.74	2764568.31	16.51
3	Span2_MainGirder4	Efflorescence	307762.48	2764550.96	15.29
4	Pier2_CapBeam	Crack	307792.60	2764560.37	15.44
5	Span1_DeckSlab	Spalling	307794.13	2764558.76	15.88

另會透過 3.6 節得到的相機位姿資訊，經運算後輸出相機位姿（表 3.8），記錄相機於拍攝時的精確位置與朝向，包括縱坐標、橫坐標、高程，以及方向角（Azimuth）與俯仰角（Pitch）。以利維修人員根據相機拍攝位置於現場尋找劣化。

表 3.8: 相機位姿資料範例 (單位：公尺)

Defect ID	Type	E(X)	Y(N)	Elevation	Azimuth	Pinch
1	Efflorescence	307780.08	2764563.98	16.52	91.96°	+14.78°
2	Spalling	307780.50	2764563.14	16.06	86.79°	+14.53°
3	Efflorescence	307791.35	2764567.01	16.74	87.39°	+17.50°
4	Crack	307794.97	2764568.11	17.13	87.93°	+18.70°
5	Spalling	307794.08	2764565.59	16.91	88.86°	+17.81°

最終產出之劣化標註報表如圖 3.22 所示，以一筆劣化為單位生成頁面，內容綜合前述空間坐標與相機位姿資訊，並搭配對應的劣化影像及劣化位置平面圖。此報表結構化呈現劣化點基本屬性、空間定位及拍攝姿態，方便使用者進行點對點之現場查驗與修繕規劃。

Defect ID: [ID]

Type: [Defect Type]
 Element: [Element Name]

Defect Coordinates (TWD97)	Camera Pose
E(X): -- Y(N): -- Elevation: --	E(X): -- Y(N): -- Elevation: -- Azimuth: -- Pinch: --

劣化影像

劣化位置平面圖

圖 3.22: 劣化檢測報表範例

此種報表結合影像與空間資訊，有助於提升劣化檢測結果的實用性與現場應用價值，並支持後續自動化分析與決策流程。



3.10 系統應用情境分析

本章節自前端至後端展示了完整的研究方法，當中包含可即時取得的資訊以及需透過後處理獲得的資訊。這些內容建立在理想的情境基礎上，但在實務應用中往往會面臨一些限制，例如老舊橋梁通常缺乏 BIM 模型（本研究所使用之道南橋 BIM 模型亦是為了研究需求而根據圖資所建立）；此外，在災害發生時往往時間緊迫，使用 MVS 技術生成稠密點雲所需之處理時間可能造成救災上的延誤。

為了應對上述的實務挑戰，本系統在設計時具備一定的彈性，能依據現場的實際條件調整運作模式，特別是在後處理階段。以下將針對四種典型情境的系統運作流程進行詳細說明。

1. **具備 BIM 模型且時間寬裕：**此為最理想之狀態，即現場已有可用的 BIM 模型，且有充裕時間進行 MVS 重建。本系統會先透過即時辨識獲得初步損傷資訊，並即時透過 SLAM 生成疏點雲與 BIM 點雲進行初步疊合。之後再透過 MVS 技術重建稠密點雲，提升後處理分析的精確性。
2. **缺乏 BIM 模型但時間寬裕：**在老舊橋梁缺乏 BIM 模型的情況下，系統將會依賴 SLAM 生成之疏點雲以及後續 MVS 技術重建稠密點雲。儘管缺乏 BIM 參考，本方法仍能達到三維精確辨識與展示橋梁劣化的目標。
3. **具備 BIM 模型但時間緊迫：**此情境常見於災害現場，時間緊迫無法等待 MVS 重建程序。本系統可即時運用 SLAM 生成之疏點雲與現有 BIM 點雲進行快速疊合，即時提供現場工程師足夠準確的初步劣化分析。
4. **缺乏 BIM 模型且時間緊迫：**此為最嚴苛的情境，系統僅能即時辨識，無法進行後處理。此時即時生成的 SLAM 疏點雲與 YOLOv9 辨識結果將直接作為現場初步評估依據，提供即時損傷資訊，協助緊急應變決策。

表 3.9 整理了上述四種典型情境的特性，包含急迫性、是否具備 BIM 模型，以及常見發生情境等。



表 3.9: 不同應用場景之特性比較

場景編號	1	2	3	4
急迫性	低	低	高	高
BIM 模型	有	無	有	無
常見情境	新建橋梁 定期檢測	老舊橋梁 定期檢測	新建橋梁 緊急檢測	老舊橋梁 緊急檢測

表 3.10 整合不同情境下，各項子模組於系統流程中的適用性與處理時間，提供工程師依據需求與現地條件進行應變調整。

表 3.10: 系統流程於不同情境之應用與處理時間

步驟	目的	應用場景				處理所需時間	建議設備
		1	2	3	4		
現地系統部署	架設空拍機與邊緣運算設備以進行巡檢	●	●	●	●	設備運送至定點後，預計數分鐘內可完成	邊緣平台
空拍機巡檢作業	透過空拍機與邊緣運算設備即時獲取劣化偵測成果	●	●	●	●	視無人機飛行時間而定，疏點雲與相機相對姿態會即時被建立	邊緣平台
點雲配準	獲取劣化之絕對位置及所在構件資訊	●		●		僅需數分鐘即可完成手動配準，其餘自動化流程僅需數秒	邊緣平台或高效運算設備
三維劣化展示：BIM 點雲	令使用者知悉劣化位置	●		●		渲染流程僅需數秒	邊緣平台或高效運算設備
橋檢報告輸出	作為即時檢測的階段成果，令使用者快速掌握橋梁狀況	●		●		自動化流程僅需數秒即可完成	邊緣平台或高效運算設備
稠密點雲重建	作為後續劣化展示與測量的關鍵工具	●	●			視飛行時長、點雲品質、硬體設備而定，通常需數小時	高效運算設備
三維劣化展示：稠密點雲	令使用者知悉劣化位置與劣化情況	●	●			渲染流程僅需數秒	高效運算設備
劣化尺度量測	令使用者了解劣化尺度資訊以利後續檢修作業	●				渲染流程僅需數秒	高效運算設備



3.11 小結

本章節說明了本研究建立之橋梁劣化檢測工作流程，綜合各節之內容，可歸納以下五項關鍵要點與方法論特色：

1. **端到端整合式架構：**系統自「影像取得」起，於邊緣裝置即時完成劣化偵測（YOLOv9）與相機定位（ORB-SLAM3），透過 ROS 串連各模組，確保任務流程即時且可擴充。後端則執行稠密點雲重建（COLMAP）、點雲配準（point-pair+ICP）、三維標註（PyVista）與報告書輸出，建立完整橋檢工作流程。
2. **異質硬體協同運作：**
 - 空拍側：DJI Inspire2 + Zenmuse X5S 提供清晰之橋梁影像。
 - 邊緣側：Jetson Orin Nano 搭配輕量化 YOLOv9 模型，在 15W 功耗下實現即時推論與同步定位。
 - 後端側：高效能桌機處理稠密點雲重建與可視化運算，強化整體流程效能並避免單點瓶頸。
3. **劣化資料完整且評估機制嚴謹：**建構涵蓋三類劣化（白華、裂縫、剝落）共計 30,248 張影像之多來源資料集，並區分訓練、驗證與測試集；模型效能則以 Precision、Recall、F1 與 mAP@50 評估，確保其實用性與泛化能力。
4. **空間對位與自動標註機制強化：**系統透過點雲配準求得剛體變換矩陣 T ，統一 SLAM、BIM 與稠密點雲坐標系統，進而將劣化結果（類型與外參）精準對位至結構構件。本研究進一步導入射線投射與 IFC 模型解析，自動標註每筆劣化所屬構件資訊，並轉換為 TWD97 坐標格式，利於工程後續定位與應

用整合。

5. **互動式視覺化與報告產出**：整合 PyVista 與 CloudCompare 提供即時 3D 劣化展示、距離量測與坐標查詢功能，並支援結果匯出與自動化報告書產出。另結合 pHash 感知哈希進行影像去重與劣化篩選，有效提升資料品質與後處理效率。

6. **多場景應用之彈性設計**：參考橋梁年齡、災後急迫性與 BIM 資訊完整度，設計四種典型應用情境。本系統可依據場景選擇對應流程組合，如災後情境可強調即時辨識與簡報輸出，而 BIM 現有者可進一步執行構件對位與 IFC 標註，展現良好之模組彈性與適應能力。

綜上所述，本研究方法兼顧即時性、精確性與模組化特性，自空拍影像擷取、邊緣辨識到三維定位與視覺化均建立具體實作流程，證實商用 UAV 搭配低功耗裝置即可達成「多類劣化即時辨識＋精準空間定位」目標，對橋梁巡檢數位轉型具實質助益。



第四章 系統驗證與實測

本章節將介紹本研究所採用的模型訓練、測試集驗證、以及實地測試結果，並詳細說明點雲配準、BIM 資訊整合、與劣化資訊展示的應用。

4.1 劣化辨識模型

本研究採用 YOLOv9 作為劣化辨識任務的主體架構，進行混凝土結構物中三種常見劣化型態：白華、裂縫與剝落之自動化偵測。考量橋梁檢測任務對即時性與準確性的需求，YOLOv9 能在高速推論的同時兼顧準確率，為本研究之首選。

4.1.1 實驗環境與系統設定

本研究之模型訓練與測試皆於單一工作站上進行，系統硬體與軟體環境如表 4.1 所示。所有訓練均於單張 GPU 上進行，模型訓練過程包含資料擴增、前處理、監督學習訓練與推論，於同一環境內完成所有測試以確保一致性。

表 4.1: 實驗硬體與系統環境

項目	規格與版本
中央處理器 (CPU)	AMD Ryzen 5 7600
圖形處理器 (GPU)	NVIDIA GeForce RTX 4080 (16GB VRAM)
記憶體 (DRAM)	32GB DDR5
作業系統	Ubuntu 20.04 LTS
深度學習框架	PyTorch 2.4.1 (CUDA 12.1)



4.1.2 模型選擇與實驗設計

本研究初步選定中階版權重 yolov9-m 與輕量化模型 yolov9-c 比較效能差異，以確認模型容量對劣化辨識任務之影響。

在超參數調整方面，本研究採用「逐步優化調參策略」，每輪實驗僅針對一組超參數進行變化，其餘設定保持不變，並保留該輪表現最佳之結果作為後續實驗的基礎。此策略類似單一變數調整法（One-at-a-Time, OAT），但進一步納入「逐步沿用最優組合」的概念，使調參過程更貼近實務中常見的參數疊加優化方式，亦有助於控制實驗成本與提升整體效能。

本研究依此策略，依序調整以下三大類參數：

1. **模型與訓練參數**：預訓練權重、batch size、優化器、學習率（learning rate）、動量（momentum）、權重衰減（weight decay）；
2. **資料增強參數**：影像平移（translation）、色相（hue）、飽和度（saturation）、明度（brightness）；
3. **其他設定**：其餘參數以 YOLOv9 官方建議參數作為初始設定，測試時固定初期置信度閾值為 0.3。

4.1.3 基礎訓練參數設定

為確保各項超參數調整實驗具可比性，本研究以表 4.2 所展示之 YOLOv9 官方建議參數作為初始設定（baseline），所有變項調整皆基於此組參數進行。

本研究所有超參數調整實驗皆以此組基礎參數為控制設定，僅針對個別變數進行調整，並於後續各表中清楚標示變動項目及其對應辨識效能。除此之外，本

表 4.2: YOLOv9 基礎訓練參數設定

參數	值	參數	值
初始學習率 (lr0)	0.01	最終學習率比例 (lrf)	0.01
動量 (momentum)	0.937	權重衰減 (weight_decay)	0.0005
Warmup 次數 (warmup_epochs)	3.0	Warmup 動量 (warmup_momentum)	0.8
初始 bias lr (warmup_bias_lr)	0.1	Box loss 權重 (box)	7.5
類別 loss 權重 (cls)	0.5	類別正樣本加權 (cls_pw)	1.0
物件正樣本加權 (obj_pw)	1.0	DFL loss 權重 (df)	1.5
IoU 訓練門檻 (iou_t)	0.20	Anchor 門檻 (anchor_t)	5.0
Focal loss gamma (fl_gamma)	0.0	HSV 色相增強 (hsv_h)	0.015
HSV 飽和度增強 (hsv_s)	0.7	HSV 明度增強 (hsv_v)	0.4
平移增強 (translate)	0.1	縮放增強 (scale)	0.9
左右翻轉 (fliplr)	0.5	混合增強 (mixup)	0.15
Copy-Paste 增強 (copy_paste)	0.3	上下翻轉 (flipud)	0.0
旋轉 (degrees)	0.0	透視 (perspective)	0.0
錐形變形 (shear)	0.0	馬賽克增強 (mosaic)	0.0

研究所有模型訓練均採固定訓練輪數 (epoch) 為 **100**，以確保各組超參數調整實驗在相同的訓練條件下進行，提升比較之公平性與一致性。於訓練過程中，對每個 epoch 進行驗證，並保留驗證 mAP@0.5 表現最佳之模型作為最終輸出。

4.1.4 測試集設計與評估原則

為全面評估模型於不同場景之辨識能力，本研究設計兩組獨立測試資料集，分述如下：

- **測試集 (Roboflow)**：與訓練資料來源相同，主要用於觀察模型在標準資料下的學習效果；
- **測試集 (本研究拍攝)**：由本研究實地以空拍機與手機拍攝雙北地區混凝土橋梁，並人工標註所得，用以模擬實務應用場景，檢驗模型泛用能力與實務部署效能。

為確保訓練、測試與實際應用流程一致，兩組測試集之所有影像皆於測試前統一 resize 為 512×512 並轉為灰階 (grayscale)，與模型訓練及部署階段的前處

理完全相同。此設計有助於排除影像尺寸與色彩資訊帶來的變異，專注評估模型對結構劣化特徵的辨識能力。



影像轉為灰階之細節與合理性已於 3.3.2 節詳述，主要考量於橋梁現場檢測時，主要目標為混凝土材質，灰階化能進一步提升模型對紋理與形狀特徵的敏感度。後續小節亦將比較「有無轉為灰階」對辨識效能之影響，以驗證該預處理策略之適用性。

雖然兩組測試集均進行完整測試，但本研究以本研究測試集（測試集 B）之 $mAP@0.5$ 表現最佳者作為最終模型選擇依據，以確保模型適用於真實場域。Roboflow 測試集（測試集 A）則作為輔助觀察模型學習穩定性與訓練成效之參考依據。

此外，為合理比較不同超參數組合的影響，所有超參數調整階段測試均統一採用置信度閾值 0.3。此數值考量了橋梁檢測實務需求，為確保潛在劣化不被漏檢，故選用較低的置信度作為超參數調整階段的參數。

在選出最終採用模型權重後，本研究進一步對該模型進行置信度閾值調整實驗，以觀察不同偵測敏感度對 precision、recall 與 mAP 的影響，提供實務應用上最適合的門檻建議。

4.1.5 模型比較與超參數調整

本節依「逐步優化調參策略」之流程，分別展示各超參數調整實驗的成果。為利比較，所有試驗皆固定置信度閾值為 0.3，並以 $mAP@0.5$ 、Precision、Recall 為主要評估指標。表格依參數類別排列，先列模型／訓練參數，次列資料增強參數，最後獨立展示置信度閾值微調結果。

表 4.3: 不同預訓練權重對辨識效能之影響

預訓練權重	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
yolov9-m	0.788	0.355	0.576	0.581	0.588	0.606
yolov9-c	0.791	0.391	0.596	0.669	0.539	0.628

最佳值：yolov9-c

表 4.4: 不同 Batch size 設定對辨識效能之影響 (yolov9-c)

Batch size	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
4	0.824	0.335	0.577	0.665	0.449	0.566
8	0.805	0.374	0.590	0.607	0.544	0.591
12	0.806	0.381	0.600	0.648	0.530	0.611
16	0.791	0.391	0.596	0.669	0.539	0.628

最佳值：Batch size (BS) = 16

表 4.5: 不同 Optimizer 對辨識效能之影響 (yolov9-c, BS = 16)

Optimizer	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
SGD	0.791	0.391	0.596	0.669	0.539	0.628
Adam	0.923	0.221	0.572	0.889	0.318	0.607
AdamW	0.768	0.390	0.585	0.585	0.528	0.584

最佳值：SGD，與預設值相同

表 4.6: 不同 Learning rate (lr0) 設定對辨識效能之影響 (yolov9-c, BS = 16)

Learning rate	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
0.005	0.759	0.411	0.590	0.628	0.573	0.615
0.010	0.791	0.391	0.596	0.669	0.539	0.628
0.020	0.820	0.311	0.569	0.661	0.443	0.580
0.040	0.853	0.243	0.549	0.818	0.360	0.600

最佳值：Learning rate = 0.01，與預設值相同

表 4.7: 不同 Weight Decay 設定對辨識效能之影響 (yolov9-c, BS = 16)

Weight Decay	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
0.0001	0.726	0.418	0.581	0.624	0.542	0.583
0.0005	0.791	0.391	0.596	0.669	0.539	0.628
0.0010	0.852	0.319	0.587	0.741	0.513	0.637
0.0050	0.954	0.166	0.561	0.927	0.201	0.568

最佳值：Weight Decay (WD) = 0.001，優於預設值 0.0005

表 4.8: 不同 Momentum 設定對辨識效能之影響 (yolov9-c, BS = 16, WD = 0.001)

Momentum	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
0.937	0.852	0.319	0.587	0.741	0.513	0.637
0.907	0.822	0.345	0.587	0.696	0.460	0.591
0.867	0.765	0.398	0.587	0.522	0.595	0.574
0.817	0.745	0.425	0.590	0.529	0.551	0.553

最佳值：Momentum = 0.937，與預設值相同

根據模型與訓練參數調整後的測試結果，yolov9-c 搭配 Batch size = 16 與 Weight Decay = 0.001 為目前整體表現最佳之組合，後續所有超參數調整實驗皆以此為基礎設定。

表 4.9: 不同平移增強 (Translation) 對辨識效能之影響

Translation	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
0.0	0.807	0.339	0.576	0.616	0.516	0.578
0.1	0.852	0.319	0.587	0.741	0.513	0.637
0.2	0.873	0.301	0.589	0.710	0.402	0.573
0.3	0.875	0.295	0.587	0.766	0.468	0.625

最佳值：Translation = 0.1，與預設值相同

表 4.10: 不同色相增強 (Hue) 對辨識效能之影響

Hue	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
0.0075	0.848	0.325	0.587	0.685	0.489	0.601
0.0150	0.824	0.341	0.585	0.741	0.513	0.637
0.0225	0.839	0.333	0.589	0.661	0.506	0.604
0.0300	0.831	0.328	0.582	0.700	0.460	0.581

最佳值：Hue = 0.015，與預設值相同

表 4.11: 不同飽和度增強 (Saturation) 對辨識效能之影響

Saturation	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
0.3	0.867	0.309	0.588	0.719	0.459	0.595
0.5	0.833	0.332	0.586	0.722	0.526	0.627
0.7	0.852	0.319	0.587	0.741	0.513	0.637
0.9	0.825	0.337	0.584	0.700	0.599	0.657

最佳值：Saturation = 0.9，優於預設值 0.7

表 4.12: 不同亮度增強 (Brightness / Value) 對辨識效能之影響 (Saturation = 0.9)

Brightness	Roboflow 測試集			本研究測試集		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
0.0	0.832	0.331	0.584	0.677	0.529	0.615
0.2	0.824	0.341	0.585	0.710	0.572	0.656
0.4	0.825	0.337	0.584	0.700	0.599	0.657
0.6	0.865	0.304	0.588	0.783	0.452	0.630

最佳值：Brightness = 0.2，優於預設值 0.4

4.1.6 模型訓練結果彙整

根據前述模型與訓練參數調整後的測試結果，yolov9-c 搭配 Batch size = 16、Weight Decay = 0.001 與 Momentum = 0.937 為整體表現最佳之組合。而資料增強相關參數調整結果顯示，Saturation = 0.9 能在不過度犧牲 Precision 的前提下大幅提升 Recall 與整體 mAP 表現，因此在進行置信度閾值調整前，將以此組資料增強設定為訓練基礎。表 4.13 整理目前實驗中所採用之 YOLOv9 訓練超參數組合，表中已將經過調整與更新的項目以粗體標示：

表 4.13: 最終 YOLOv9 訓練參數設定 (更新處以粗體標示)

參數	值	參數	值
初始學習率 (lr0)	0.01	最終學習率比例 (lrf)	0.01
動量 (momentum)	0.937	權重衰減 (weight_decay)	0.001
Warmup 次數 (warmup_epochs)	3.0	Warmup 動量 (warmup_momentum)	0.8
初始 bias lr (warmup_bias_lr)	0.1	Box loss 權重 (box)	7.5
類別 loss 權重 (cls)	0.5	類別正樣本加權 (cls_pw)	1.0
物件正樣本加權 (obj_pw)	1.0	DFL loss 權重 (df)	1.5
IoU 訓練門檻 (iou_t)	0.20	Anchor 門檻 (anchor_t)	5.0
Focal loss gamma (fl_gamma)	0.0	HSV 色相增強 (hsv_h)	0.015
HSV 飽和度增強 (hsv_s)	0.9	HSV 明度增強 (hsv_v)	0.4
平移增強 (translate)	0.1	縮放增強 (scale)	0.9
左右翻轉 (fliplr)	0.5	混合增強 (mixup)	0.15
Copy-Paste 增強 (copy_paste)	0.3	上下翻轉 (flipud)	0.0
旋轉 (degrees)	0.0	透視 (perspective)	0.0
錐形變形 (shear)	0.0	馬賽克增強 (mosaic)	0.0

圖 4.1 展示本研究最佳模型於訓練過程中的指標變化，包括 Precision、Recall、mAP@0.5、mAP@0.5:0.95、學習率 (lr) 與 loss。由圖可見，各項指標自訓練初期迅速提升，於 60 至 80 步左右趨於穩定，表示模型已有效收斂。最終 Precision、Recall、mAP@0.5 均趨於穩定，且損失函數持續下降，顯示過擬合現象不明顯，訓練過程穩健。

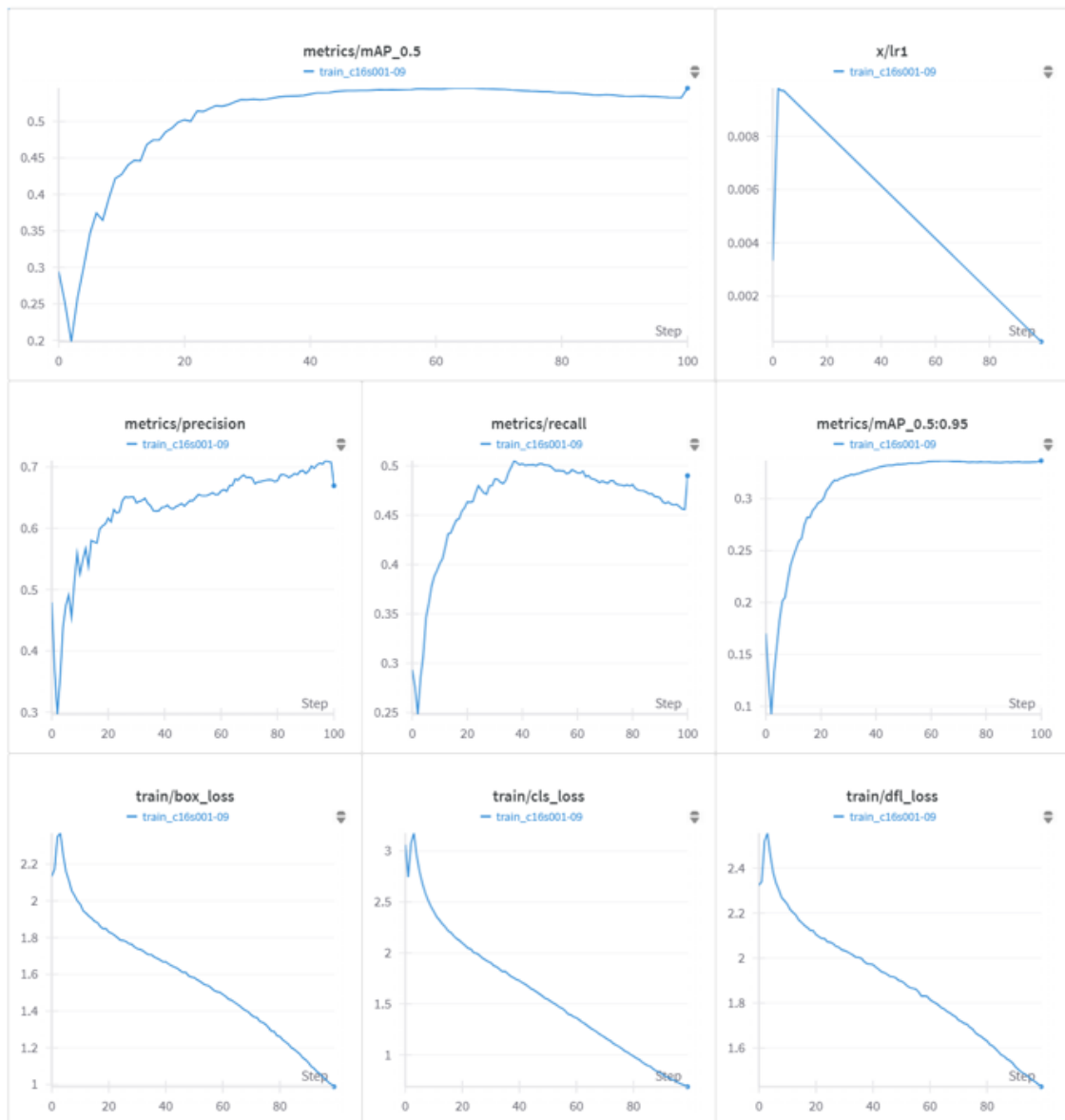


圖 4.1: 模型訓練過程中 Precision、Recall 及 mAP 等指標隨訓練步驟之變化

圖 4.2 為驗證集的混淆矩陣，橫軸為真實標籤，縱軸為預測標籤。從矩陣中可觀察到，模型對裂縫 (Crack) 和剝落 (Spalling) 類別的判斷較為準確，分別有

0.58 與 0.56 的正確預測比率；白華（Efflorescence）則有較多漏檢的情況（0.60），說明其外觀特徵與背景相近，是日後改善重點。

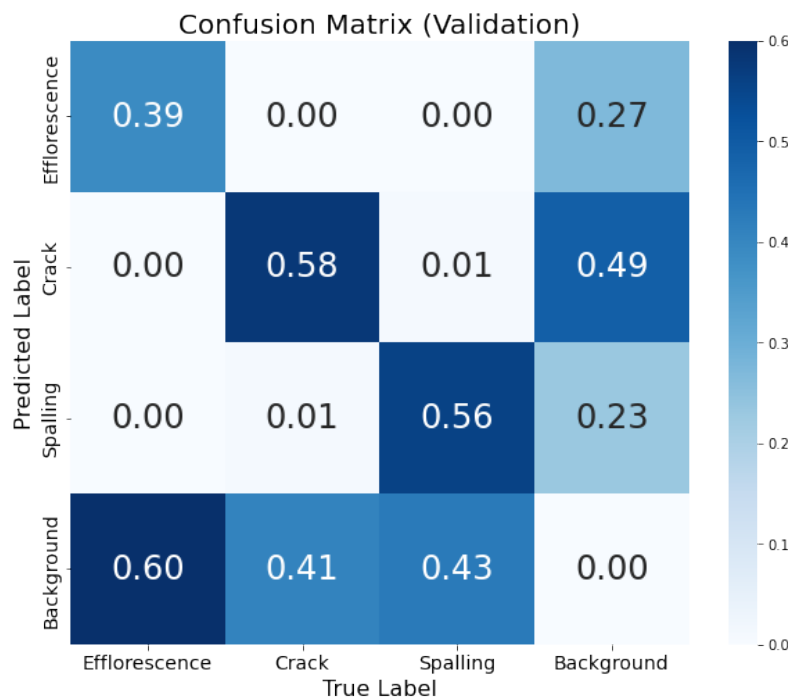


圖 4.2: 驗證集混淆矩陣

圖 4.3 呈現驗證集三類別及總體的 Precision-Recall 曲線。曲線下方面積（AUC）越大代表辨識能力越好。剝落（Spalling）與裂縫（Crack）類在大部分召回率區間內 Precision 維持較高，分別有 $mAP@0.5$ 為 0.628 與 0.614，驗證模型針對這兩類劣化的判斷較為精確。白華（Efflorescence） $mAP@0.5$ 只有 0.395，說明在多數閾值下精確率較難提升，符合前述混淆矩陣的結果，總體 $mAP@0.5$ 則達到 0.546。

綜合上述結果，模型於訓練及驗證過程均展現良好的學習趨勢與泛化能力，針對常見劣化類型（裂縫、剝落）具備高辨識率。未來可針對白華等難以區分的類別，增強資料多樣性及提升特徵敏感度，以提升實務應用成效。

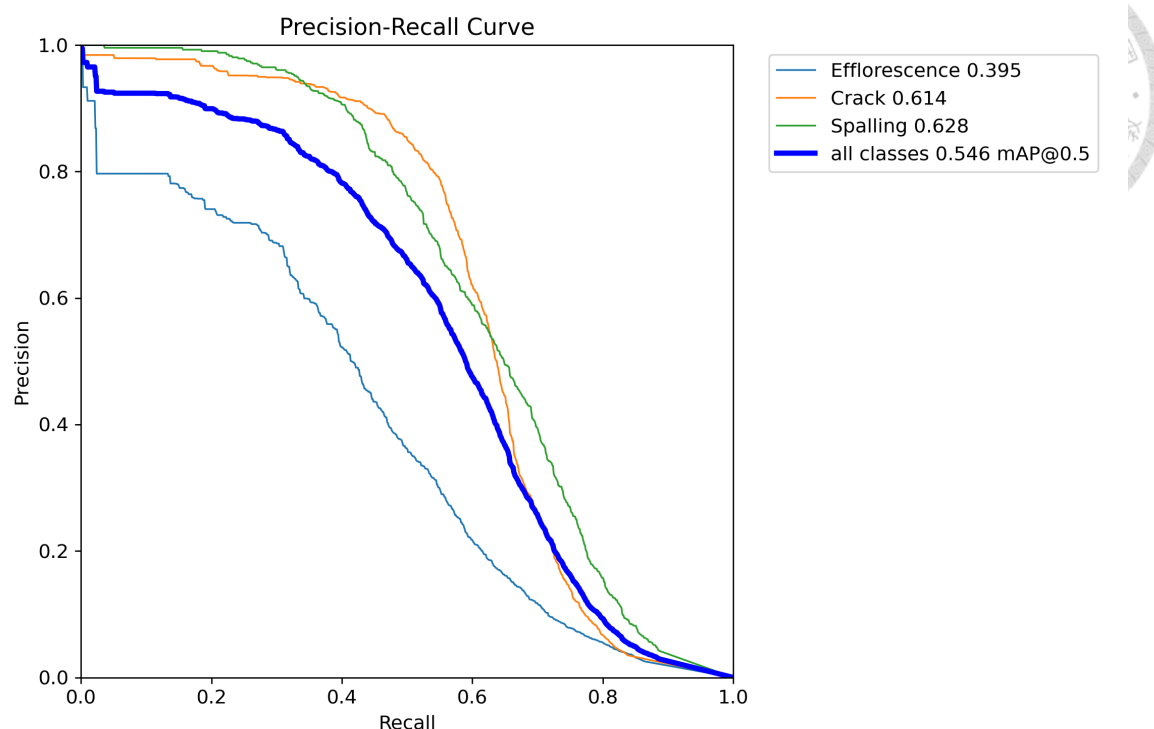


圖 4.3: 驗證集各類別 Precision-Recall 曲線，框內數值為各類別 mAP@0.5

由表 4.14 及圖 4.4 可觀察到，隨著置信度閾值提高，Precision 呈現明顯上升、Recall 則逐漸下降，顯示模型預測更為嚴格，但同時可能漏檢部分目標。mAP@0.5 則在 0.3 附近達到最高值（0.657），為 Precision 與 Recall 之間取得最佳平衡的區間。

表 4.14: 不同置信度閾值設定下最終模型於本研究測試集之辨識效能

置信度閾值	Precision	Recall	mAP@0.5
0.15	0.690	0.603	0.650
0.20	0.690	0.603	0.656
0.25	0.690	0.603	0.653
0.30	0.700	0.599	0.657
0.35	0.732	0.531	0.639
0.40	0.759	0.478	0.627
0.45	0.810	0.430	0.622
0.50	0.853	0.354	0.604

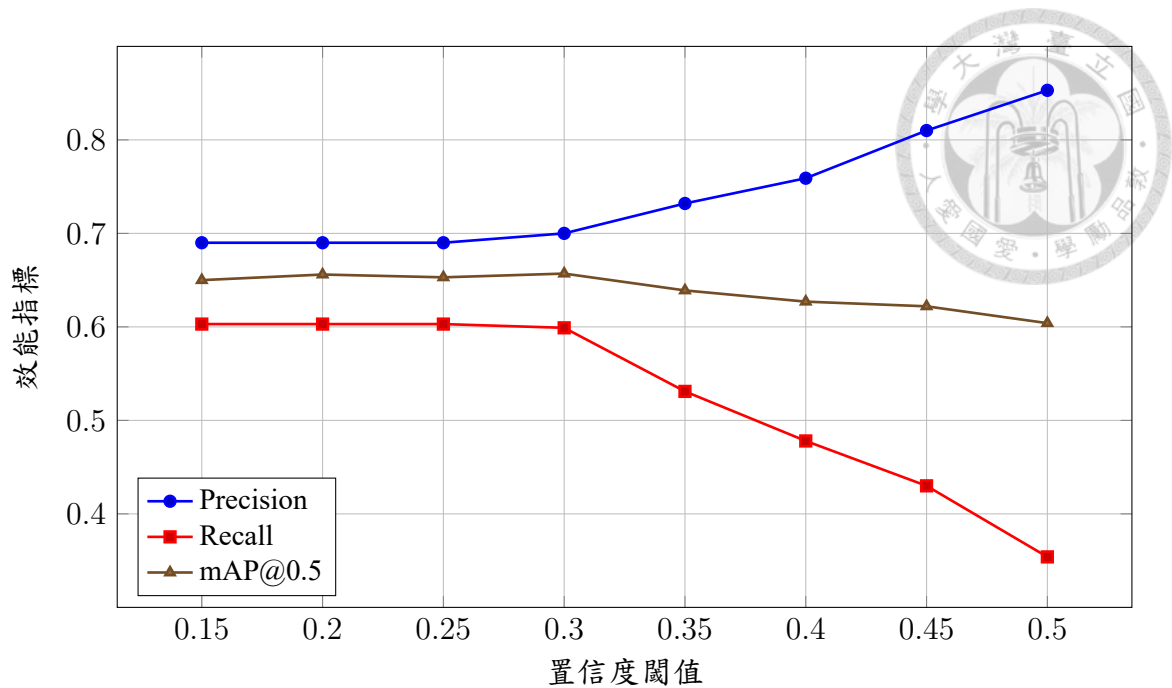


圖 4.4: 不同置信度閾值下 Precision、Recall 與 mAP@0.5 之變化趨勢

綜合三項指標，當置信度閾值設定為 0.3 時，模型同時保有較高的準確率與召回率，並使整體辨識效能（mAP@0.5）達到最佳。故本研究後續劣化辨識結果與分析皆採用置信度閾值 0.3 作為最終判定基準。

4.1.7 橋梁劣化分類性能與案例

如圖 4.5 與表 4.15 所示，模型在本研究測試集展現出不錯的分類能力：

- 白華（Efflorescence）類別有 78 筆標註，模型的 Precision 為 0.613，Recall 為 0.590，mAP@0.5 為 0.582。混淆矩陣中可見部分白華樣本被誤分類為背景或其他類別，反映出白華特徵辨識的挑戰。
- 裂縫（Crack）類別有 103 筆標註，Precision 高達 0.764，Recall 亦達 0.660，mAP@0.5 為 0.714，表現最佳。混淆矩陣顯示裂縫誤判情況較少，模型對裂縫辨識能力較為穩健。
- 剝落（Spalling）類別有 172 筆標註，Precision 0.723，Recall 0.547，mAP@0.5

為 0.676。混淆矩陣指出剝落與背景或白華間存在部分誤判，未來可透過增強資料與模型改進降低錯誤率。



- 整體 (all) 而言，模型 Precision 與 Recall 分別為 0.700 與 0.599，mAP@0.5 為 0.657，說明模型對多類別劣化目標具優秀識別能力。

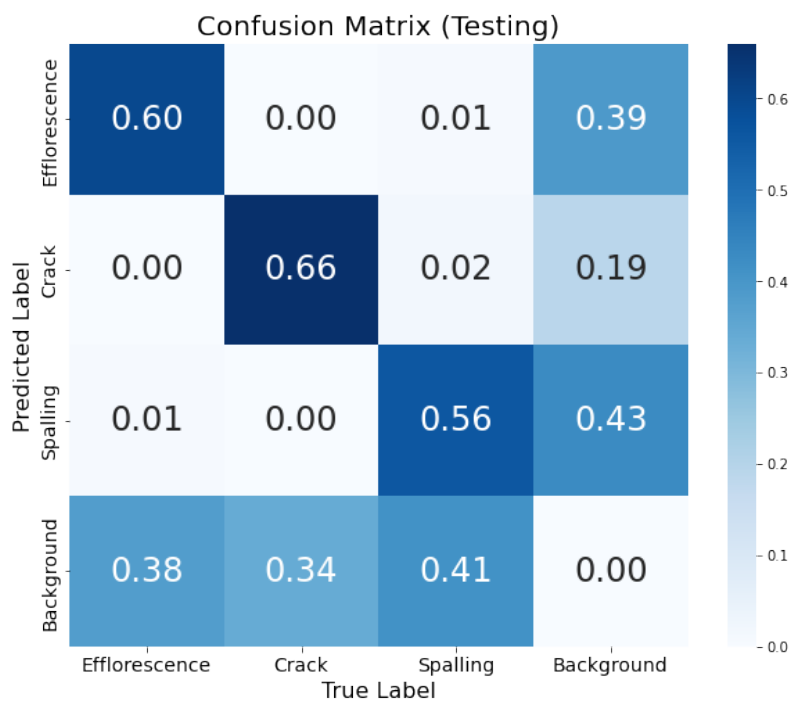
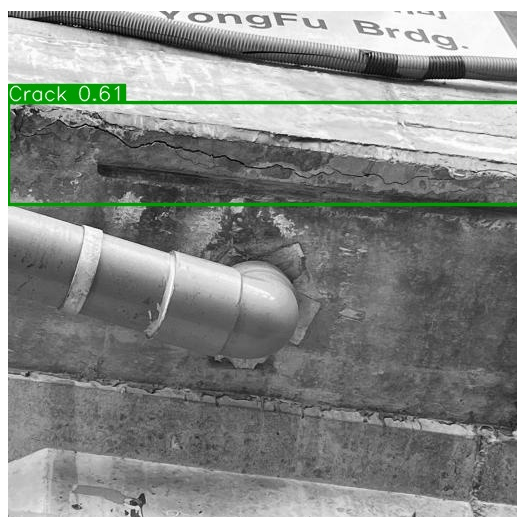


圖 4.5: 測試集混淆矩陣

表 4.15: 最佳模型在本研究測試集（置信度閾值 = 0.3）之分類效能

類別	影像數	標註數	Precision	Recall	mAP@0.5
全部 (all)	197	353	0.700	0.599	0.657
白華 (Efflorescence)	197	78	0.613	0.590	0.582
裂縫 (Crack)	197	103	0.764	0.660	0.714
剝落 (Spalling)	197	172	0.723	0.547	0.676

圖 4.6 展示六組代表性測試案例，呈現模型於各類橋梁劣化的辨識效果。



(a) 裂縫正確偵測 (攝於永福橋)



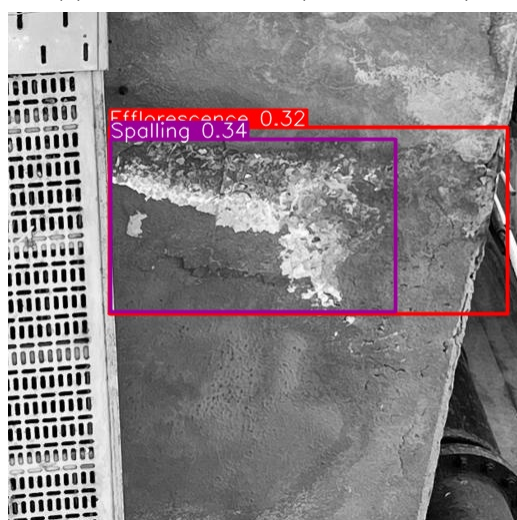
(b) 白華漏檢 (攝於永福橋)



(c) 白華正確偵測 (攝於永福橋)



(d) 三類型劣化同時出現，部分漏檢 (攝於永福橋)



(e) 白華被誤檢為剝落 (攝於福和橋)



(f) 剝落多處皆正確辨識 (攝於永福橋)

圖 4.6: 本研究最佳模型於現地影像多類別劣化辨識成果實例。

上圖中，(a) 為單一明顯裂縫，模型能準確框選並給予高置信度；(b) 為一組白華特徵，由於紋理與背景差異較小，且受橋下護網影響，出現漏檢現象，顯示模型對低對比目標仍有提升空間；(c) 為白華區域正確被辨識之情形，證明經灰階化預處理後模型對細微紋理具敏感性；(d) 為白華、剝落與裂縫共存場景，其中剝落與裂縫兩類劣化均正確辨識，惟白華出現漏檢情形；(e) 為白華被誤判為剝落，但白華仍被正確辨識；(f) 為剝落分布於多處且皆被正確偵測，顯示本研究模型對剝落類具有較高準確率與召回率。綜上，模型在主要劣化型態辨識展現良好性能，針對複雜與低對比劣化仍有進一步改善空間。

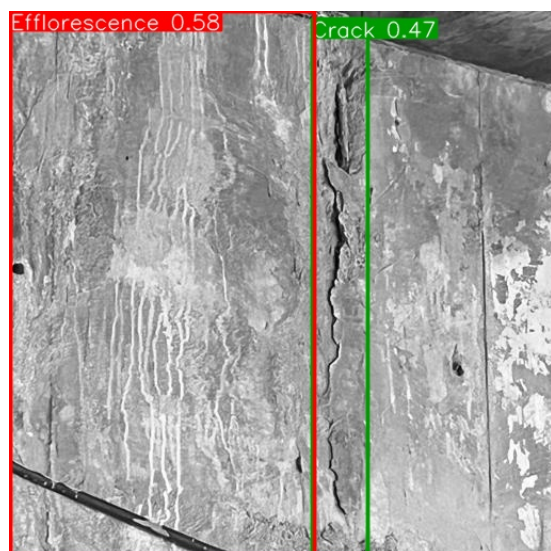
4.1.8 灰階與原彩影像之辨識效能探討

為驗證影像灰階化對模型效能之實質影響，本研究特別針對本研究測試集，分別以灰階影像（預設）與原彩影像作為輸入，並於其餘設定完全一致（包含影像尺寸 512×512 、預處理流程等）下，評估兩者之辨識效果。其結果彙整如表 4.16 所示。

表 4.16: 灰階影像與原彩影像輸入下各類別 mAP@0.5 指標比較

類別	灰階 mAP@0.5	原彩 mAP@0.5	提升值
全部 (all)	0.657	0.521	+0.136
白華 (Efflorescence)	0.582	0.475	+0.107
裂縫 (Crack)	0.714	0.679	+0.035
剝落 (Spalling)	0.676	0.409	+0.267

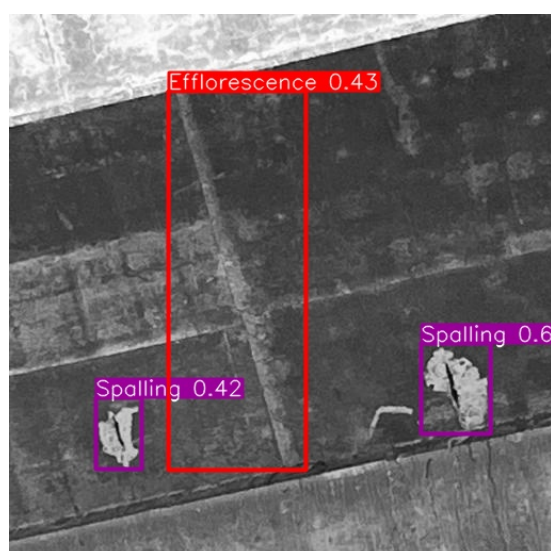
上表可見灰階化處理能顯著提升模型於剝落與白華類別的辨識能力，其中剝落類 mAP@0.5 提升超過 0.26，白華類亦有 0.13 以上的提升。裂縫類雖提升幅度較小，但仍有助於整體辨識表現。



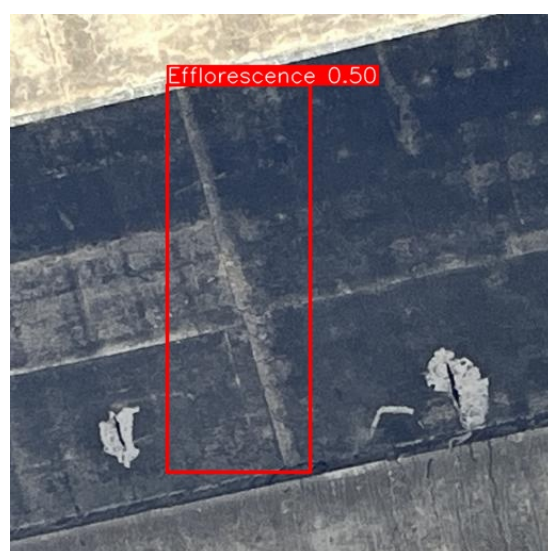
(a) 灰階影像辨識結果（永福橋）



(b) 原彩影像辨識結果（永福橋）



(c) 灰階影像辨識結果（福和橋）



(d) 原彩影像辨識結果（福和橋）

圖 4.7: 灰階與原彩影像之劣化辨識結果對照

如圖 4.7 所示，灰階影像（左欄）明顯較原彩影像（右欄）能更準確標註出白華（Efflorescence）與剝落（Spalling）區域，尤其是剝落分布面積或形狀較不明顯時，灰階化能使模型聚焦於紋理與結構變異。原彩影像雖可辨識大部分裂縫區域，但對於低對比的白華與剝落則常有誤判或漏檢。

這驗證本研究將影像預處理統一為灰階的設計是有效且合理的。推測原因在於剝落及白華劣化特徵色彩與背景相近，彩色影像易受干擾，灰階處理反而能突

顯結構紋理，提升偵測準確性。未來若遇到色彩對比明顯之場域或新型態劣化，亦可再評估保留色彩資訊之必要性。



綜合上述，本研究模型於測試集展現良好多類別辨識能力。對於明顯的裂縫和剝落具高度偵測精度與召回率，白華類則仍易與背景混淆。整體而言，模型已能應對大多數實務場景，未來可針對混合型與低對比度劣化目標進行進一步改進，強化模型的實地適應力與全面性。

4.1.9 結果分析與小結

本研究針對 YOLOv9 劣化辨識模型之設計、訓練與驗證，進行一系列系統性實驗與調參，重點分析結果如下：

- **模型架構選擇：**實驗顯示，輕量化 yolov9-c 權重在本研究測試集之 mAP@0.5 與 precision 表現優於 yolov9-m，且推論速度快、資源需求低，更適合橋梁現場部署。說明本研究資料集規模及目標特徵複雜度適合中小型模型，能有效避免過擬合並提升泛化能力。
- **訓練參數影響：**經系統性調整，Batch size 以 16 為佳，兼顧模型穩定訓練及效能提升。優化器方面，SGD 展現最穩定的精度與召回率；Adam/AdamW 雖偶有較高 precision，但在 recall 及 mAP 上波動大，不利於工程部署下的可靠性。學習率 (lr) 維持 YOLOv9 官方預設 0.01，最能確保收斂速度與最終效能；Weight Decay 微調至 0.001，較預設值能提升泛化性，抑制過擬合。
- **資料增強與預處理：**飽和度 (Saturation) 調高至 0.9，可顯著提升 mAP 及 recall，特別對於剝落與白華類劣化，證明高飽和度資料擴增有助於模型學習低對比目標特徵。此外，統一將所有影像轉換為灰階，能顯著提升對劣化

的敏感度，尤其剝落與白華類別 mAP 分別提升 0.267 與 0.136，顯著優於原彩輸入。



- **測試集設計與評估：**兩組測試集皆經標準化前處理（resize、灰階化），排除因資料型態差異產生的評估偏誤。最終以「本研究測試集」之 mAP@0.5 作為模型選擇依據，強調模型對現地場域的適應力。
- **置信度閾值調整：**系列測試顯示，將閾值設為 0.3 能兼顧 precision 與 recall，mAP@0.5 達最高（0.657），推薦做為劣化自動辨識之預設門檻。現場應用上可依實務風險容忍度彈性調整。

綜合分析，本研究最終模型組合如下：

yolov9-c, SGD, Batch size = 16, Weight Decay = 0.001, Saturation = 0.9

（其餘參數採官方預設，影像統一灰階化，置信度閾值 0.3）

該組合於自製測試集多類別指標（mAP@0.5 = 0.657, Precision = 0.700, Recall = 0.599）均達到最優，顯示模型具備應用潛力，特別在裂縫與剝落辨識上有穩定且優異表現。白華雖辨識仍具挑戰，但灰階化及資料增強策略已帶來明顯進步。而儘管本研究模型在裂縫與剝落辨識上表現優異，仍有數項待改善之處，歸納如下：

- **白華易與背景混淆：**白華類 mAP@0.5 僅 0.582，常被誤判為背景。建議擴充不同環境下的白華樣本，並導入注意力模組 [82] 以提升辨識能力。
- **Recall 偏低，潛在漏檢：**總體 recall 為 0.599，部分劣化在複雜背景或遮蔽情境下未被偵測。可透過 hard negative mining [83] 等策略提升召回率。
- **固定輸入尺寸限制細部辨識：**單一尺度輸入對小面積劣化表現有限，建議未來加入多尺度訓練或 super-resolution [84] 等技術增強影像細節。

綜上，本研究模型已具實用性，但針對白華、遮蔽、細節與多任務整合等問題，仍有進一步優化空間。未來朝向資料多樣化與架構強化發展，可提升模型在實地橋梁檢測中的穩定性與全面性。



4.2 前導實驗：系統可行性測試

本次初步驗證實驗選擇於臺北市與新北市交界之「寶橋」進行，目的為測試本研究所建構之無人機檢測系統在實際橋梁環境下的整合運行可行性。寶橋橫跨景美溪之都市型道路橋梁，於民國 67 年（1978 年）8 月竣工，為四跨三墩之梁式橋，全長 103.24 公尺，橋寬 19.5 公尺，連接臺北市文山區木新路與新北市新店區寶橋路（如圖 4.8 所示），為兩地之重要交通樞紐，長年承受高交通流量與各類載重，具備劣化監測需求，故本研究以其作為初期系統驗證場地。



圖 4.8: 寶橋位置圖，方框內為本次實驗飛行範圍



4.2.1 驗證目的與測試條件

本階段實驗以寶橋作為場域，進行本研究所建構之橋梁劣化檢測系統的首次實地驗證。主要目的為確認無人機搭載 YOLOv9 劣化辨識模組與 ORB-SLAM3 定位與建圖模組，在實際橋梁環境中的整合運作可行性，並蒐集實測資料供後續分析與系統修正。具體實驗目的包括：

- 確認無人機飛行過程中，影像能穩定傳入 YOLO 模組進行即時劣化偵測。
- 確認 ORB-SLAM3 模組能同步進行影像追蹤與初步定位建圖，並與影像辨識流程順利整合。
- 確認整體系統能完成 ROS 架構下之任務串接，包含影像擷取、推論發布與相機位姿記錄等。
- 初步觀察實地操作所需的飛行策略、視角設置與資料擷取條件，以作為後續實驗之參考。
- 錄製 rosbag 檔案，以利後續於離線環境中進行系統調整與模組改進。

本次驗證使用設備為 DJI Inspire 2 無人機搭載 Zenmuse X5S 單眼雲台相機，搭配 NVIDIA Jetson Orin Nano 作為邊緣運算平台，基於 ROS，整合 YOLOv9 與 ORB-SLAM3 模組，同步記錄影像與姿態資訊。詳細軟硬體資訊同表 3.1。

測試作業選於天氣良好、無強風之日間時段進行，系統架設地點位於寶橋橋下自行車道，施作前進行基本場勘並設置飛行邊界。操作過程由一名飛手與一名安全人員共同執行，確保飛行過程之設備與人員安全。所有影像與姿態資料將用於後續稽核與系統調校。



4.2.2 現地實驗紀錄

本次寶橋現地實驗記錄摘要如表 4.17 所示，統整拍攝時間、氣候條件、操作人員與影格資料量等基本資訊，並整理劣化偵測結果數量。

表 4.17: 寶橋現地實驗記錄

日期時間	2024/11/6 下午 3 時
天氣	多雲
實驗地點	寶橋（臺北市文山區與新北市新店區交界）
無人機飛手	陳羿清
參與作業人數	2 人（飛手、安全人員）
紀錄時長	14 分 52 秒（892 秒）
影格資料數	13380 筆（892 × 15fps）
SLAM 關鍵幀數	429 幀（有追蹤失敗情況）
劣化幀數（經自動過濾）	63 幀

圖 4.9 展示操作人員進行飛行控制與影像串流作業之現場照片，可看出本場域具備典型的都市河岸橋梁特徵，並存在部分飛行障礙（如濃密植栽），對系統整合測試具代表性。



圖 4.9: 系統驗證現場操作實況 - 寶橋

此外，圖 4.10 為本次實驗期間以 rqt_graph 工具擷取之 ROS 節點架構圖，呈現各模組在 ROS 架構下之實際串接關係與資料流向。圖中可見，相機影像經由影像處理節點後，同步輸入 YOLO 劣化辨識與 ORB-SLAM3 模組，再由後端節點整合辨識結果與相機位姿，並於 RViz 中以箭頭與文字標註方式即時呈現，形成完整的 ROS 工作流程。

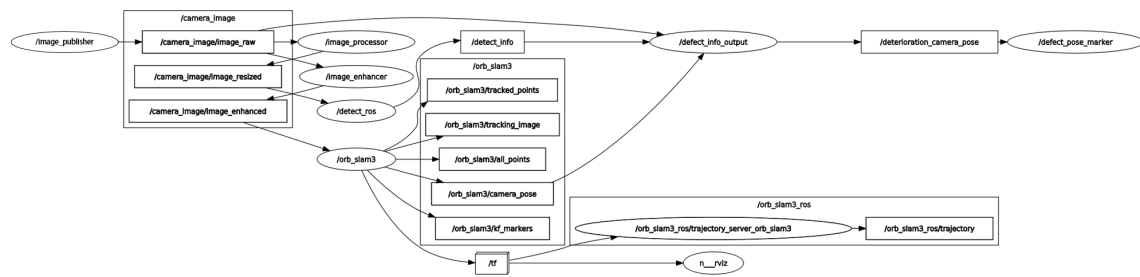


圖 4.10: ROS 節點圖 (rqt_graph)

4.2.3 實驗觀察與問題分析

本次寶橋實驗聚焦於整合系統於實地場域中的運作可行性，包含影像串流、劣化推論、SLAM 定位與資料同步等模組是否能穩定執行。劣化辨識成果將於後續正式實驗中進一步探討與分析，本節重點將聚焦於系統運行層面之實際觀察與問題歸納。

本次寶橋實驗雖順利完成系統整合測試與資料收集，但亦觀察到數項影響系統效能與操作安全的潛在問題，說明如下：

1. 橋下光照變化劇烈，影響影像穩定性：當日拍攝時部分區域因陽光穿透橋體結構產生強烈陰影，造成影像明暗差異過大，易使 YOLO 偵測結果出現誤判，亦會影響 SLAM 模組的特徵匹配準確度。系統實際運行畫面如圖 4.11 所示，包含即時 YOLO 劣化辨識（左下）、ORB-SLAM3 追蹤與 RViz 中劣化箭頭展示（中）、以及地圖建構狀態（右上）。

2. 橋下環境障礙物密集，影響飛行安全與通訊穩定性：圖 4.12 為寶橋空拍影像，圖中可見寶橋橋下區域存在懸空電纜、鋼索等多項障礙物，對無人機飛行操作造成實質干擾與潛在風險。另於繞行至橋墩背面或橫向飛行時，機體可能暫時超出操作人員視距，導致影像鏈路中斷（如圖 4.13），進而影響即時影像辨識與追蹤穩定性。

3. 純視覺 SLAM 容易發生追蹤失敗 (Tracking Loss)：ORB-SLAM3 在無 GNSS 或 IMU 輔助的情況下，僅依賴單眼相機進行影像特徵追蹤與位姿估計，對於特徵不足或畫面劇烈變動的情境極為敏感。在本次實驗中，當飛行軌跡出現快速轉向、視角變化過大，或影像中出現大面積低紋理區域（如天空、水面）時，系統即容易失去追蹤，導致 SLAM 建圖中斷（如圖 4.14 所示）。

整體而言，寶橋實驗成功驗證了系統運作可行性，但亦暴露出在複雜橋下場域中，系統仍存在穩定性與操作風險。後續規劃中，須考量障礙物密度、視距限制、相機角度控制與通訊品質等因素，以提升 SLAM 穩定性與影像資料品質。

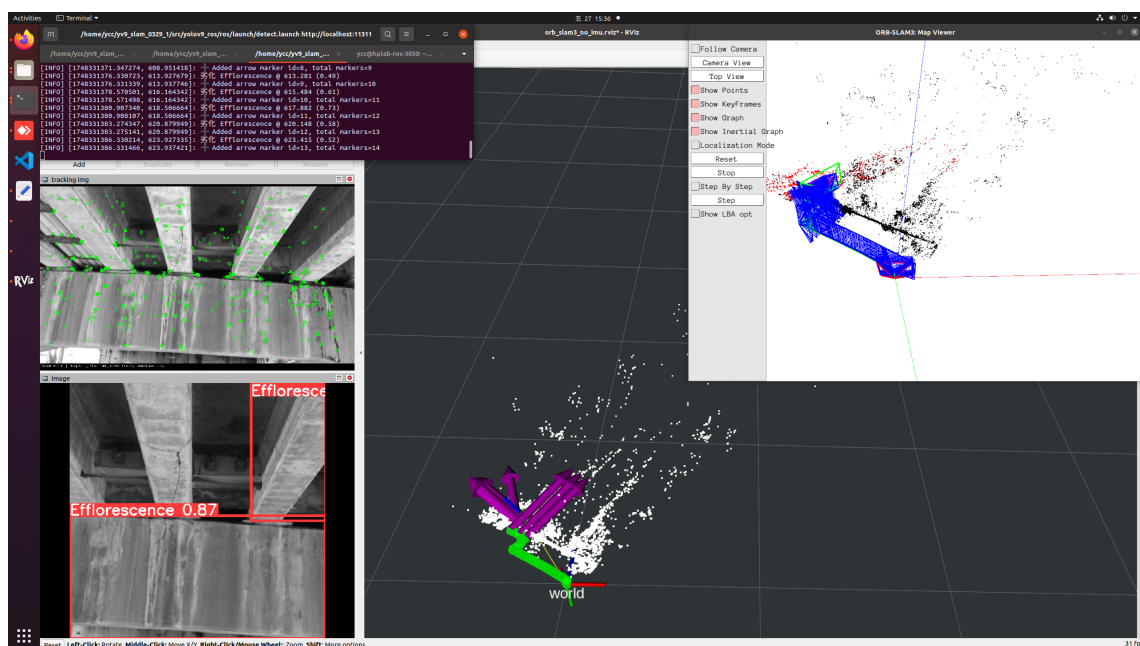


圖 4.11: 寶橋實驗，系統即時運行畫面



圖 4.12: 寶橋空拍畫面，橋下區域存在懸空電纜、鋼索等多項障礙物

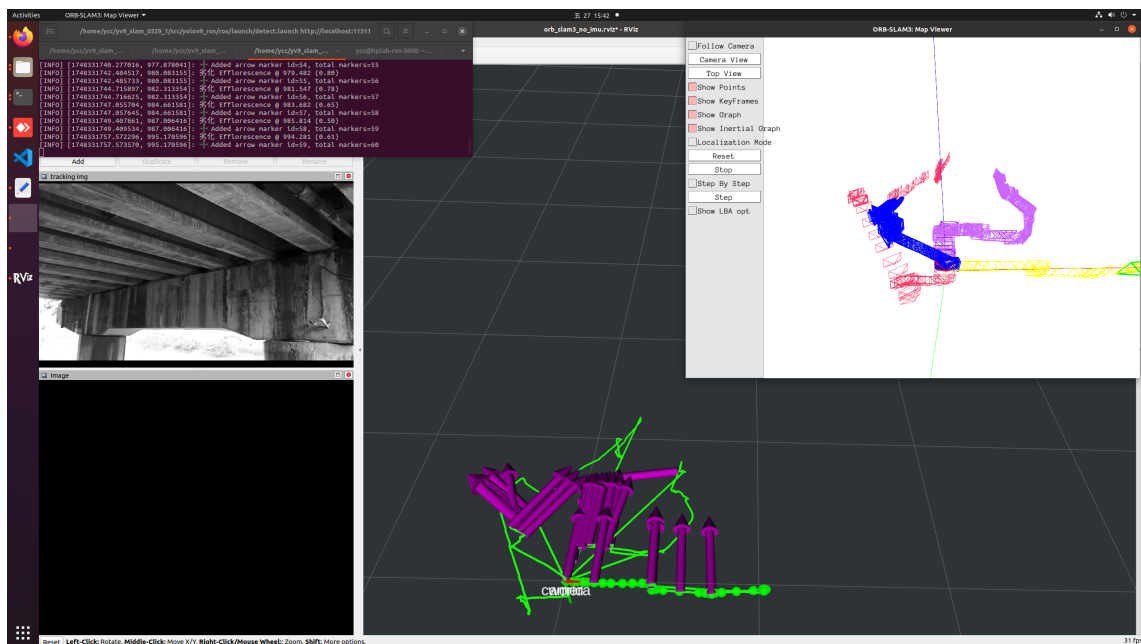


圖 4.13: 影像串流於飛行過程中短暫中斷，導致 SLAM 模組追蹤失敗

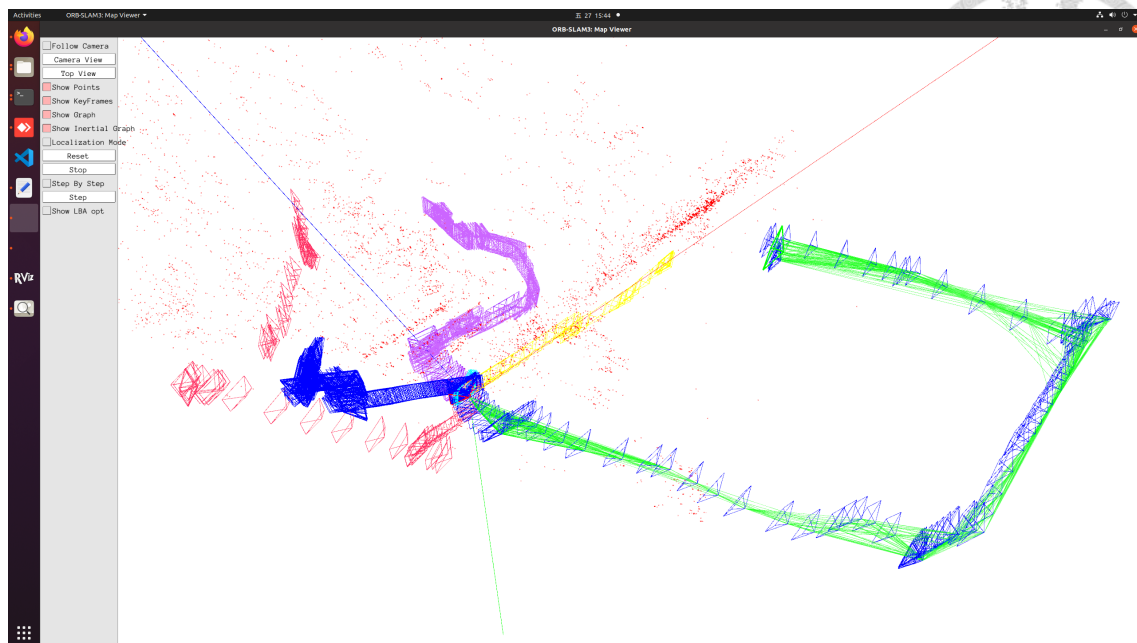


圖 4.14: ORB-SLAM3 MapViewer，因多次 tracking loss 導致地圖不完整

4.3 正式實驗：系統應用與資料取得

繼寶橋之初步測試後，本研究進一步選擇同為臺北市文山區的「道南橋」（圖 4.15）作為正式實驗場地，針對整體系統進行更完整與深入的實地應用與資料採集。道南橋為橫跨景美溪之都市型道路橋梁，於民國 73 年（1984 年）8 月竣工，位處車輛通行密集區域，橋體全長 120 公尺，為三跨兩墩之預鑄 PC 梁式橋。橋寬 11.15 公尺，銜接臺北市文山區的指南路一段與指南路二段（如圖 4.16 所示），具備典型城市橋梁構造特徵，且橋下空間相對開闊，便於飛行作業與影像觀測。



圖 4.15: 道南橋



圖 4.16: 道南橋位置圖，方框內為本次實驗飛行範圍

4.3.1 實驗設計與作業流程

道南橋實驗為本研究系統進行整合後之正式應用場景，目的在於實地驗證整體系統於橋梁結構下之劣化偵測、視覺定位與資料記錄功能是否具備實用性與穩定性。本次實驗全面執行包含無人機作業規劃、飛行策略調整、資料採集流程、影像串流與位姿同步記錄、以及後處理等完整流程，並將前次實橋實驗中觀察到的系統問題加以修正與強化。

本次驗證使用設備與前次相同，可參考表 3.1。檢測流程包含：

- **任務規劃與場勘：**於施作前完成飛行區域界定與障礙物盤點，確認適合飛行路徑並依據結構類型規劃拍攝策略。
- **飛行作業與資料同步記錄：**操作人員於橋側進行飛行操控，並透過無線連線將相機畫面同步傳入 Jetson Orin Nano，進行 YOLO 劣化辨識與

ORB-SLAM3 定位作業，並同步記錄各筆影像與相機姿態資訊。

- **後處理資料重建與整合：**實驗結束後，彙整各關鍵幀影像與相機位姿，導入 COLMAP 執行稠密點雲重建，並結合 BIM 模型進行點雲對齊與缺陷空間定位。詳細內容將在 4.4 節中詳細說明。

針對前次寶橋測試所揭露之操作困難與技術瓶頸，本次道南橋實驗特別針對下列面向進行修正與強化：

- **人員編組調整：**前次實驗僅配置飛手與安全人員，本次道南橋實驗則新增「系統監控人員」，協助即時檢查 YOLO 與 SLAM 模組運行狀態，並確認訊號串流是否穩定，確保整體作業流暢度與系統穩定性。
- **橋下障礙物密度與飛行安全性提升：**相較於寶橋，道南橋橋下空間開闊、無懸吊電纜與結構突出物，可有效降低飛行風險，提升任務安全性。
- **穩定視角與路徑設計：**根據前次 tracking loss 問題，調整無人機飛行軌跡與相機拍攝角度，避免快速視角切換，提升 SLAM 模組追蹤穩定性。
- **拍攝時段選擇：**拍攝時段選擇於接近正午時進行（橋下皆為陰影），以減少光線頻繁變化對 YOLO 與 ORB-SLAM3 的影響。

圖 4.17 為道南橋現地實驗之執行畫面，可見現場配置三名人員分工合作以確保飛行任務順利執行。圖中左側著反光背心者為無人機飛手，負責操控飛行路徑與相機視角；右側為安全人員，負責觀測周遭環境與提醒障礙物風險；中央人員則負責系統監控及器材移動。透過明確分工與協作，確保整體作業流程穩定、資訊同步完整。



圖 4.17: 系統驗證現場操作實況 - 道南橋

4.3.2 系統運行結果

為佐證本系統於複雜橋梁場域中之應用效能，道南橋實驗完整呈現了從影像擷取、即時辨識、同步定位至地圖建構的全流程整合成果。實驗記錄摘要如表 4.18 所示，統整拍攝時間、氣候條件、操作人員與影格資料量等基本資訊，並整理劣化偵測結果數量。

表 4.18: 道南橋現地實驗記錄

日期時間	2025/3/12 上午 11 時
天氣	多雲
實驗地點	道南橋（臺北市文山區）
無人機飛手	陳羿清
參與作業人數	3 人（飛手、安全人員、系統監控）
紀錄時長	11 分 03 秒（663 秒）
影格資料數	9945 筆（663 × 15fps）
SLAM 關鍵幀數	746 幀
原始劣化幀數	52 幀

圖 4.18 為實驗期間系統即時運行畫面，可見 RViz 中已同步顯示 SLAM 所生

成之空間點雲與辨識模組所偵測之劣化類型與信心分數；下方影像則顯示 SLAM 特徵點追蹤與 YOLOv9 推論結果。

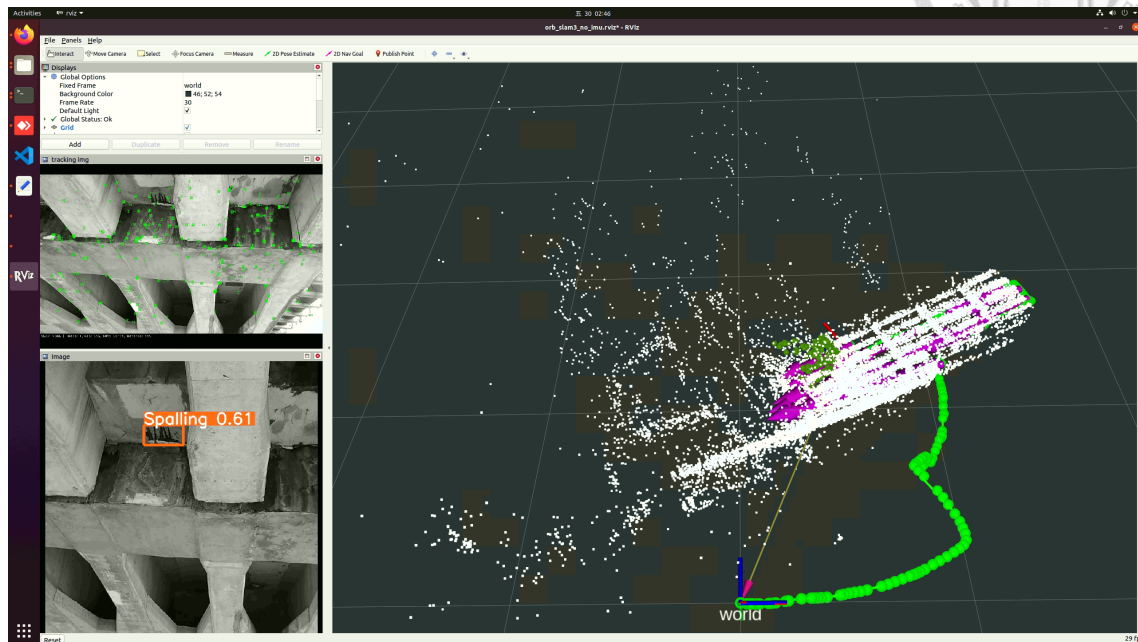


圖 4.18: 道南橋實驗，系統即時運行畫面

SLAM 運行過程中，所有影像與軌跡資料成功建構為無追蹤中斷之地圖，整體軌跡連續、點雲覆蓋範圍完整，顯示系統穩定性顯著優於前次寶橋測試，詳如圖 4.19 所示。RViz 中顯示追蹤軌跡、SLAM 建圖與劣化箭頭疊加成果。

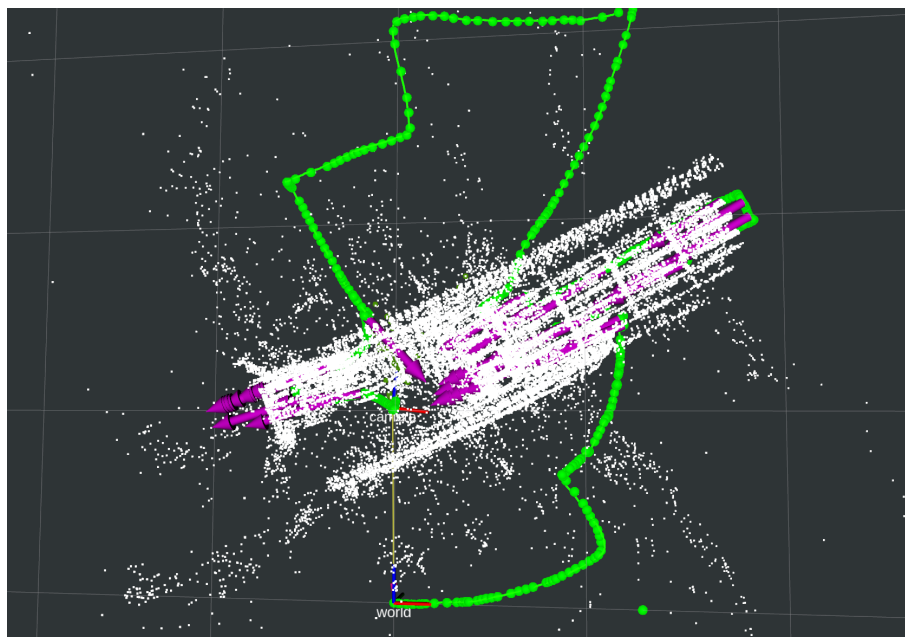


圖 4.19: 道南橋完整地圖建構成果

此外，本系統亦同步記錄實際辨識畫面，圖 4.20 展示了部分現場辨識成果：圖中紅色框為偵測之白華區域，紫色框則標註剝落位置，均為系統於橋下拍攝過程中即時辨識與標註所得。這些畫面可協助工程人員即時了解缺陷位置與類型，並輔助後續影像篩選與檢修規劃。



圖 4.20: 道南橋劣化辨識成果示意

圖 4.21 展示了去重程序的運作成果，左側為最終保留的劣化影像，右側則為經去重程序剔除的影像。上列為系統使用感知哈希（pHash）自動辨識為重複的劣化，兩者外觀幾乎一致，右圖因置信度低而被自動排除；下列則為經人工複查後判定為重複的案例，兩張影像有顯著光影差異，屬於同一劣化但未被自動剔除。

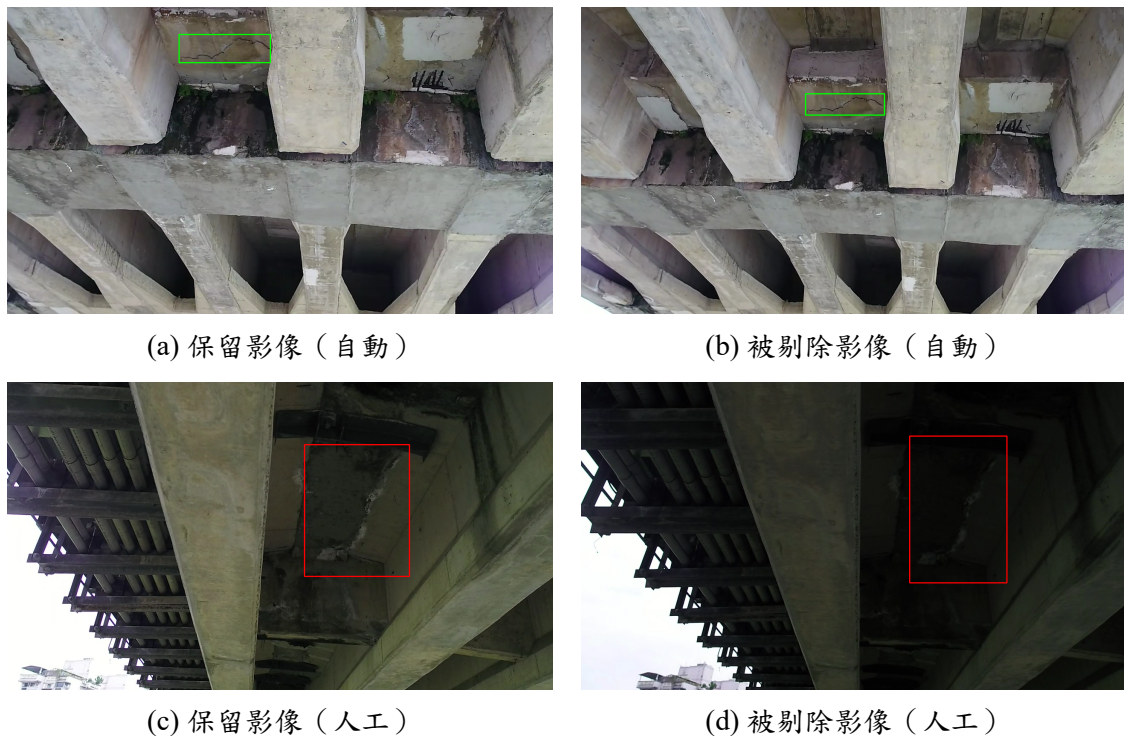


圖 4.21: 保留與去重影像示例。左為保留之代表影像，右為被剔除之對應影像

本次實驗即時劣化偵測模組初步偵測 52 幀劣化影像，經 pHash 技術自動去重後減至 33 幀，爾後透過人工複查再剔除 8 幀重複影像，最終共保留 25 幀具代表性的劣化影像，並據此統計實驗中總計偵測到 25 處劣化位置。



實驗顯示部分重複的影像並未被 pHash 自動去重，主因可能包括以下幾點：

- **拍攝角度或構圖略有差異：**雖為同一處劣化，但因 UAV 或攝影裝置角度略微偏移，使得畫面中劣化位置位移或比例不同，導致 pHash 特徵值差異偏大，無法達到自動判定閾值。
- **光照變化或遮蔽影響：**天候變化或結構陰影造成畫面亮度、對比改變，使得感知哈希產生的二進位特徵有顯著差異，誤判為非重複影像。
- **去重閾值設置較保守：**為避免誤刪真正不同的劣化影像，本系統將 pHash 相似度門檻設定為較保守值（如 Hamming 距離 ≤ 5 ），在提高精準度的同時，可能會漏判部分實際重複的影像。

因此，人工後處理仍為提升資料品質之必要手段。尤其在實地應用中，考量各種外在條件干擾，自動去重雖可大幅減少冗餘資料，但尚難完全取代人工辨識的彈性與準確性。

經過自動與人工去重後，劣化個數如圖 4.22 所示，圖中行表示模型預測類別，列表示真實類別，數字為對應樣本數。可見：

- 白華（Efflorescence）共 14 筆樣本，其中 11 筆正確預測，3 筆誤判為背景。
- 裂縫（Crack）共 4 筆樣本，均正確辨識。
- 剝落（Spalling）共 7 筆樣本，5 筆正確辨識，1 筆誤判為背景，1 筆誤判為白華。

- 背景類別中有 15 筆樣本被誤判為三種劣化之一。

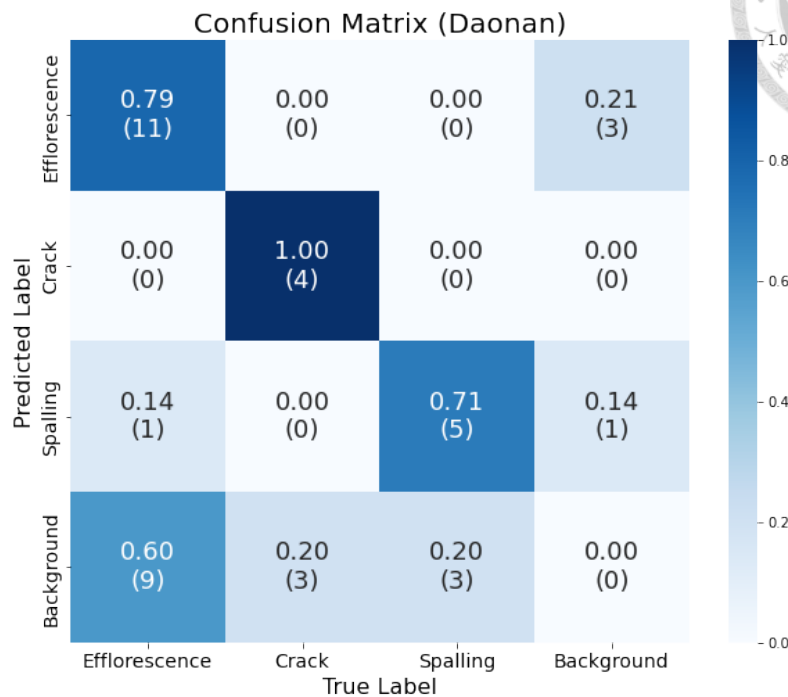


圖 4.22: 道南橋實驗中即時劣化辨識之混淆矩陣

表 4.19 彙整了道南橋實驗中三大劣化類別的辨識績效指標，包含 Precision（精確率）、Recall（召回率）與 F1-Score。整體而言，道南橋現地實驗展現了較模型測試階段（如表 4.15）更佳的 Precision（0.700 提升至 0.800），主因在於系統採用 pHash 等自動去重機制，有效過濾重複與雜訊影像，降低誤報率，提升精準度。然而，由於部分樣本被去重剔除，Recall 相對略有下降（0.599 下降至 0.556），但仍保持合理平衡，展現即時劣化辨識於實地複雜環境下的應用潛力。

表 4.19: 道南橋實驗即時劣化辨識之各類別 Precision、Recall 與 F1-Score

類別	Precision	Recall	F1-Score
白華 (Efflorescence)	0.786	0.550	0.644
裂縫 (Crack)	1.000	0.571	0.727
剝落 (Spalling)	0.714	0.625	0.667
Micro 平均	0.800	0.556	0.656



4.4 實驗資料後處理與結果展示：道南橋

本節展示本研究於道南橋實驗中所蒐集之資料，經由後處理流程所產出之多項成果，包含稠密點雲建構與配準、缺陷定位與三維資訊可視化。各步驟不僅驗證本系統從資料擷取到空間標註的完整性，亦展現其在實務應用中的可行性。

4.4.1 稠密點雲重建成果

本研究利用 ORB-SLAM3 所記錄之相機關鍵幀與姿態資料，依據 3.7.1 小節所述之流程，導入 COLMAP 工具進行影像比對與稠密點雲重建。其第一步為建立稀疏模型骨架 (sparse model)，以驗證相機軌跡與對應影像是否正確匯入，並檢查是否與 SLAM 系統所生成之姿態一致。圖 4.23 顯示建立完成之稀疏模型，可見 746 筆影像皆成功註冊，共重建出 202217 個三維點，平均每張影像有超過 1758 個對應觀測點，重投影誤差控制於 0.71 px，整體成果良好，顯示資料品質可支撐後續 MVS 重建。

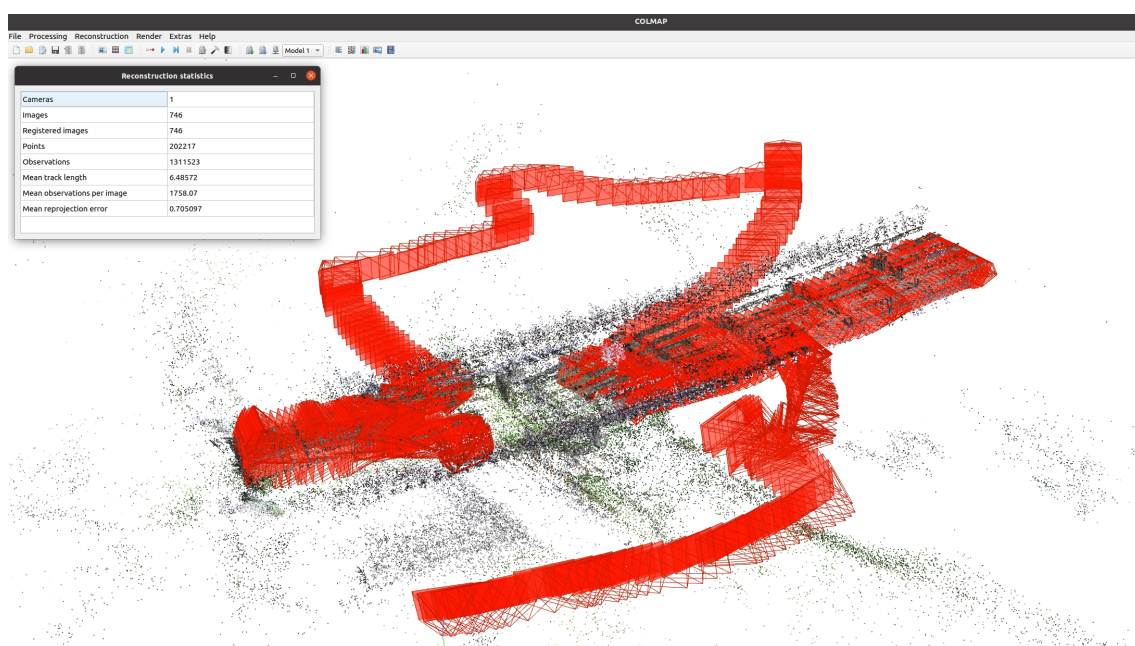


圖 4.23: 道南橋 COLMAP 建立稀疏模型成果

在完成 stereo_fusion 模組的深度融合後，COLMAP 成功輸出完整稠密點雲，圖 4.24 為本研究於道南橋所建立之稠密點雲成果，可清楚辨識橋梁結構形貌與周邊環境，包括橋台、橋墩、護欄與兩側堤岸等。此成果為後續與 BIM 模型對齊、缺陷定位與視覺展示之基礎資料。



圖 4.24: 道南橋 COLMAP 重建之稠密點雲成果

4.4.2 空間配準與劣化定位成果

為將稠密點雲對齊至 BIM 模型之絕對座標系，本研究使用 CloudCompare 執行初始配準，並搭配 Open3D 套件進行精細 ICP 演算法處理。初始點對由使用者選定數對對應點，建立剛體初始轉換，接續以 point-to-point ICP 進行誤差最小化，最終取得稠密點雲與幾何模型間的剛體轉換矩陣與尺度係數，詳細流程已於 3.8 節中說明。

圖 4.25 為由 BIM 模型轉換而得之幾何點雲，可視為對齊目標之絕對座標基準。圖 4.26 為使用 CloudCompare 工具進行初步點對配準之畫面，使用者手動對應稠密點雲與幾何模型上的對應結構節點（如橋墩轉角、翼牆邊緣等），計算剛體初始轉換與尺度。最後，圖 4.27 為經過 ICP 精細配準後之視覺化結果，點雲成功對齊至 BIM 幾何模型，顯示整體空間變換準確，後續劣化標註位置亦可正確映射於模型中。

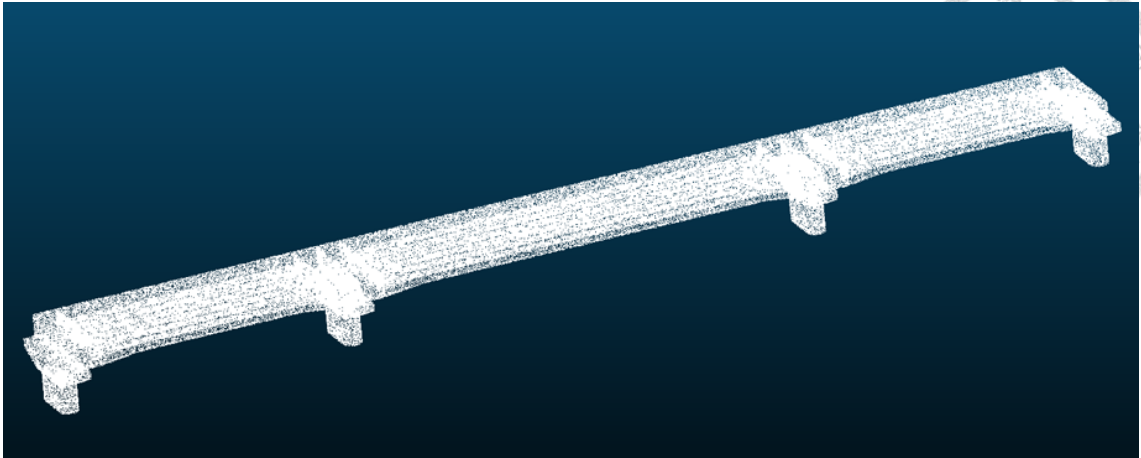


圖 4.25: 道南橋之 BIM 幾何模型點雲，作為絕對座標基準

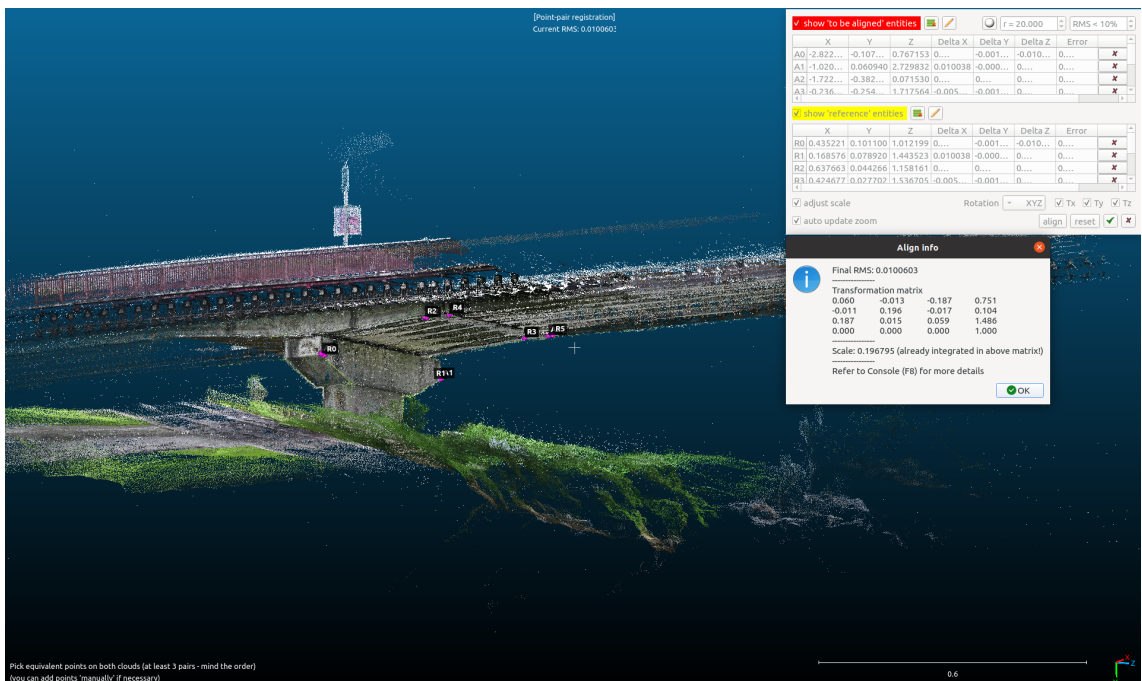


圖 4.26: CloudCompare 初步點對配準畫面

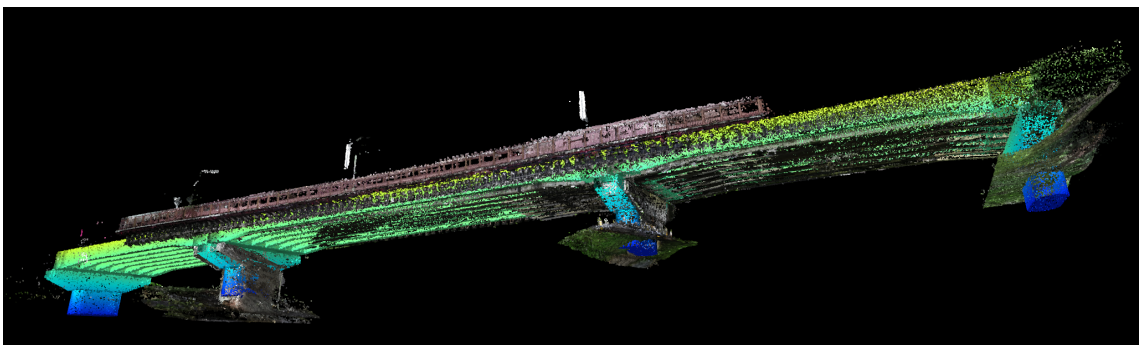


圖 4.27: Open3D 執行 ICP 配準成果

經歷初步配準與 ICP 精細配準後，依據前述方法分別取得兩組 4×4 變換矩陣，分別對應 CloudCompare 初始配準與 Open3D ICP 配準之剛體轉換。最終，透過矩陣相乘方式合併兩階段的轉換結果，產出一組完整對應於稠密點雲與 BIM 絕對座標系之總變換矩陣，該矩陣將作為後續相機外參轉換之依據。

每筆劣化標註均對應一組相機外參與時間戳，本研究將其姿態透過前述剛體變換矩陣轉換至真實世界座標，完成劣化標註於結構模型中的三維定位。轉換過程可參考 3.8.2 小節。

本階段輸出的劣化相機定位部分結果彙整於表 4.20，包含劣化標註於 BIM 空間中的三維位置 (p_x, p_y, p_z) 、旋轉四元數 (q_x, q_y, q_z, q_w) 以及劣化類型。該資料經由先前取得的剛體變換矩陣轉換後產出，作為後續缺陷位置視覺化的依據。

表 4.20: 本次實驗之劣化相機定位成果彙整表（節錄）

ID	Type	E(X)	Y(N)	Elevation	Azimuth	Pinch
1	Efflorescence	307780.08	2764563.98	16.52	91.96	14.78
2	Efflorescence	307780.5	2764563.14	16.06	86.79	14.53
3	Efflorescence	307791.35	2764567.01	16.74	87.39	17.50
4	Spalling	307794.97	2764568.11	17.13	87.93	18.70
5	Efflorescence	307794.08	2764565.59	16.91	88.86	17.81
6	Spalling	307793.15	2764565.22	16.67	89.18	17.73
7	Crack	307787.83	2764563.11	16.29	86.43	16.68
8	Crack	307770.74	2764556.34	15.00	85.52	14.48
9	Crack	307770.46	2764555.75	15.08	84.91	14.03
10	Efflorescence	307761.09	2764550.82	14.44	82.01	13.51

另外，系統亦會依據構件資訊進行比對，將劣化標註與 BIM 模型中的構件名稱對應，並彙整成結構化報表，如表 4.21 所示，提供構件層級的缺陷位置與類型資訊，以輔助工程師後續檢修與分析。具體方法在 3.9.3 小節中已詳細說明。

表 4.21: 劣化標註與 BIM 構件名稱對應結果 (節錄)

ID	Element	Type	E(X)	Y(N)	Elevation
1	Span2_DeckSlab	Efflorescence	307783.68	2764563.85	18.48
2	Span2_DeckSlab	Efflorescence	307788.52	2764563.58	19.15
3	Span2_MainGirder6	Efflorescence	307795.71	2764567.18	19.15
4	Span2_MainGirder6	Spalling	307798.03	2764568.26	19.15
5	Span2_MainGirder4	Efflorescence	307797.90	2764565.70	19.15
6	Span2_MainGirder4	Spalling	307797.00	2764565.29	18.86
7	Span2_DeckSlab	Crack	307792.57	2764563.41	18.70
8	Span2_DeckSlab	Crack	307780.39	2764557.09	18.50
9	Span2_DeckSlab	Crack	307780.13	2764556.61	18.51
10	Span2_MainGirder5	Efflorescence	307770.72	2764552.17	17.78

最終，系統進一步將各構件是否有缺陷以及其類型數量進行彙整，產出如表 4.22 所示之劣化分布表。該表提供橋檢人員快速檢索，亦作為報表生成與決策建議的依據。



表 4.22: 各構件劣化分布彙整表

Element	HasDefect	Efflorescence	Crack	Spalling
Abutment1_CapBeam	no	0	0	0
Abutment1_CrossBeam	yes	1	0	0
Abutment1_Wall	no	0	0	0
Abutment2_CapBeam	no	0	0	0
Abutment2_CrossBeam	no	0	0	0
Abutment2_Wall	no	0	0	0
Pier1_CapBeam	no	0	0	0
Pier1_Column	no	0	0	0
Pier1_CrossBeam1	no	0	0	0
Pier1_CrossBeam2	no	0	0	0
Pier2_CapBeam	no	0	0	0
Pier2_Column	no	0	0	0
Pier2_CrossBeam1	no	0	0	0
Pier2_CrossBeam2	no	0	0	0
Span1_DeckSlab	yes	1	0	2
Span1_MainGirder1	no	0	0	0
Span1_MainGirder2	yes	2	0	1
Span1_MainGirder3	no	0	0	0
Span1_MainGirder4	yes	1	0	0
Span1_MainGirder5	no	0	0	0
Span1_MainGirder6	no	0	0	0
Span1_MainGirder7	yes	0	0	1
Span2_DeckSlab	yes	4	4	1
Span2_MainGirder1	no	0	0	0
Span2_MainGirder2	yes	1	0	0
Span2_MainGirder3	no	0	0	0
Span2_MainGirder4	yes	1	0	1
Span2_MainGirder5	yes	1	0	0
Span2_MainGirder6	yes	1	0	1
Span2_MainGirder7	no	0	0	0
Span3_DeckSlab	no	0	0	0
Span3_MainGirder1	no	0	0	0
Span3_MainGirder2	no	0	0	0
Span3_MainGirder3	no	0	0	0
Span3_MainGirder4	no	0	0	0
Span3_MainGirder5	no	0	0	0
Span3_MainGirder6	no	0	0	0
Span3_MainGirder7	no	0	0	0



4.4.3 橋梁劣化檢測報告書範例

本節展示本系統產出的橋梁劣化檢測報告書範例，每頁記載一筆劣化資訊。

圖 4.28 展示四筆劣化，每筆記錄包含缺陷類型、構件名稱、空間座標、相機位姿、辨識影像以及平面圖中的缺陷位置標示，協助工程人員進行後續檢修作業。

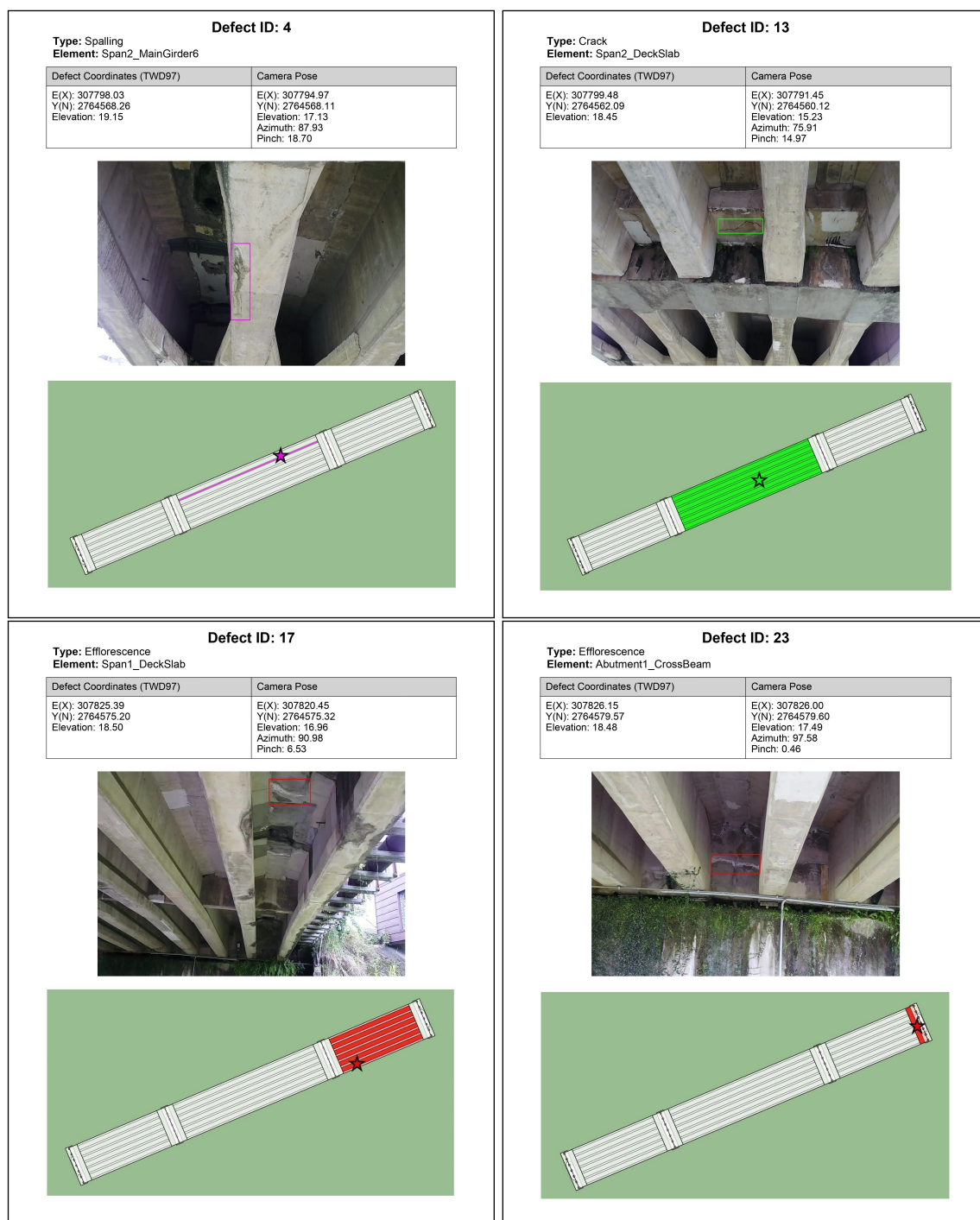


圖 4.28: 橋梁劣化檢測報告書範例



4.4.4 三維劣化展示與互動

整合上述空間配準與定位成果，本研究利用 PyVista 平台進行三維劣化資訊之視覺化展示。圖 4.29 所示即為劣化點於三維空間中的位置與姿態，可即時檢視各類缺陷之分佈，輔助工程人員更直觀地掌握劣化狀況。

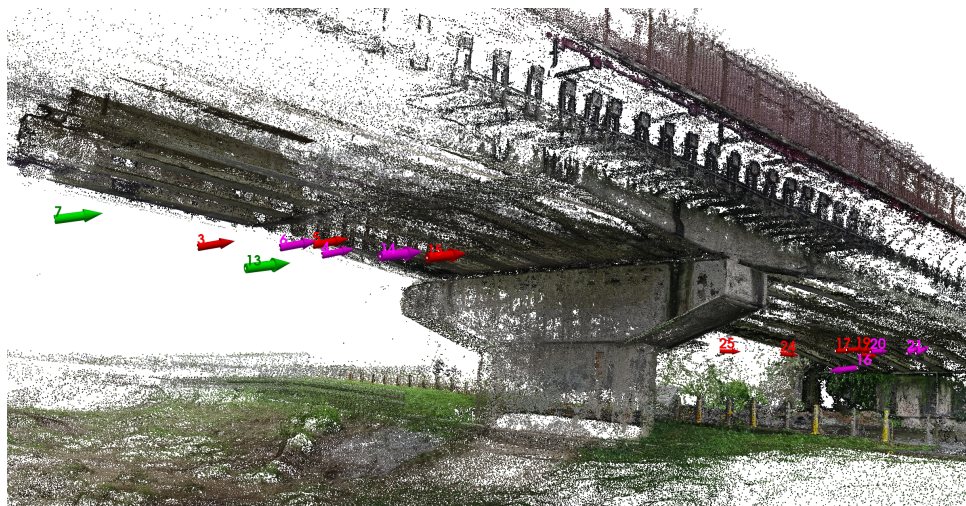


圖 4.29: 使用 PyVista 平台之三維劣化位置展示

此外，稠密點雲經轉換後帶有真實尺度與座標資訊，使用者可透過 CloudCompare 工具，進行點雲瀏覽、量測等動作，提升資料應用的靈活性與準確性。如圖 4.30 所示，人員透過事前的劣化量測，可在前往現地作業前掌握劣化情況，從而擬定修繕計畫。

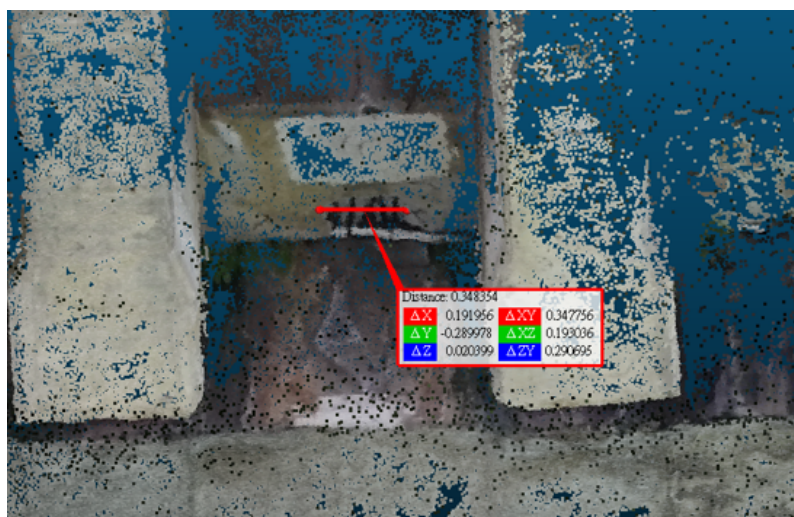


圖 4.30: 利用 CloudCompare 工具進行點雲量測與多重校驗



4.5 實驗結果與討論

本節將彙整橋梁劣化辨識模型訓練及測試、系統實地驗證（寶橋與道南橋）及三維重建的實驗成果。

4.5.1 實驗結果

劣化辨識模型（節 4.1）經測試確立以 yolov9-c 為核心。關鍵成果包括：灰階預處理顯著提升剝落（mAP@0.5 提升 0.267）與白華（mAP@0.5 提升 0.136）辨識力；最終模型於「本研究測試集」達 mAP@0.5 0.657、Precision 0.700、Recall 0.599，對裂縫與剝落辨識穩健。

在不同測試資料集上的表現存在差異。Roboflow 測試集 recall 值普遍偏低，顯示資料多樣性不足可能導致模型傾向過擬合至訓練樣本分佈。反之，自製測試集包含現地拍攝資料，雖拍攝條件多樣，但 mAP 整體表現更佳，且更符合我國的實務應用場景。

進一步分析三類劣化之分類表現，裂縫與剝落類別精度與召回率較高，因其特徵明顯，如裂縫具線性斷裂、剝落有邊界及陰影變化；白華則因與背景灰度接近、特徵邊緣模糊，辨識仍具挑戰。

初步系統可行性測試（4.2節）於寶橋驗證 YOLOv9 與 ORB-SLAM3 於無人機 ROS 平台整合的可行性，辨識出光照、障礙物及純視覺 SLAM 穩定性等作業挑戰。

正式實驗（4.3節）於道南橋進行，克服飛行與通訊障礙，實現 746 關鍵幀的連續 SLAM 建圖。即時劣化辨識共記錄 25 處劣化，現地辨識效能 Precision

0.800、Recall 0.556、F1-Score 0.656。高 precision 得益於感知哈希與人工複查流程，降低誤判與重複資料。

實驗資料後處理（4.4節）利用道南橋 SLAM 數據，以 COLMAP 重建出高清晰度（20.2 萬點，重投影誤差 0.71 px）稠密點雲模型，並透過 ICP 對齊至 BIM 模型，完成缺陷定位與報告輸出流程。

4.5.2 問題與討論

本研究所提出之橋梁劣化檢測系統整合 YOLOv9、ORB-SLAM3 與 ROS 架構，並於寶橋與道南橋實地驗證整體流程之可行性。然而，從模型效能與實地作業層面觀察，系統仍面臨數項挑戰與改進空間，說明如下：

- **白華辨識效能仍有限：**雖透過灰階處理與高飽和度資料增強，白華類別 mAP 有顯著提升，但與裂縫、剝落相比仍偏低。其原因在於白華邊界模糊、灰階分布接近背景。未來可考慮導入注意力模組，或增加高變異性場景資料以提升模型對細微變異的辨識能力。
- **SLAM 容易失去追蹤 (Tracking Loss)：**寶橋實驗中觀察到，單眼 ORB-SLAM3 在低紋理、快速視角轉換或光照劇烈變化下容易失去追蹤。後續可考慮整合 IMU 輔助 (Visual-Inertial SLAM)，或採用更具穩健性的 SLAM 架構以提升穩定性。
- **現場飛行風險與操作限制：**寶橋橋下電纜密集、通訊遮蔽、視距限制等問題，均對作業安全與資料傳輸穩定性造成影響。道南橋雖已改善場域條件與人員分工，但未來建議針對不同橋梁場域制定標準化飛行規範與安全評估機制。

- 
- **自動去重機制具漏判風險：**感知哈希 (pHash) 可初步過濾明顯重複影像，但在構圖、角度或光照差異下可能漏判。為確保資料品質，人工複查仍不可或缺。未來可考慮結合 BIM 資訊或深度特徵匹配提升自動去重準確性。
 - **三維定位精度受限於初始配準準確性：**稠密點雲對齊至 BIM 模型需仰賴手動點對配準與 ICP 精細化，若初始配對誤差大，恐影響後續劣化定位準確度。未來可探討導入自動配準演算法或輔助定位標記，以提高整體空間標註精度與效率。

綜合而言，本研究系統已展現實用性與可行性，惟於低對比特徵辨識、SLAM 穩健性、資料去重與三維標註流程等面向仍有精進空間。後續研究可朝多感測器融合、模型多任務學習與自動化後處理流程發展，以提升系統於實務場域中的穩定性與應用價值。



第五章 結論與建議

本研究成功整合影像辨識與同步定位與地圖建構 (SLAM) 技術，並應用於無人機 (UAV) 平台，建立了一套橋梁劣化即時檢測系統。透過本系統，期望能提升橋梁檢測的效率、安全性與準確性，並為橋梁的維護管理提供更智慧化的解決方案。本章將針對研究成果進行總結，探討研究過程中遭遇的限制，並對未來的研究方向提出建議。

5.1 結論

本研究的主要貢獻與結論如下：

- **成功開發並驗證 UAV 橋梁劣化即時檢測系統：**本研究整合了 YOLOv9 物件偵測演算法與 ORB-SLAM3 技術，成功在 UAV 平台上實現了對橋梁常見劣化（如裂縫、剝落、白華等）的即時辨識與三維定位。實驗結果顯示，系統能夠在實際橋梁環境中運行，並初步達到預期的檢測效果。
- **實現多類別劣化辨識與初步定位：**透過訓練 YOLOv9 模型，系統能夠辨識多種橋梁劣化類型。結合 ORB-SLAM3 技術，能夠在 UAV 飛行過程中，初步記錄並視覺化劣化發生時的相機姿態，為後續的精確三維定位與點雲重建提供基礎數據。

- 
- **建立劣化資訊後處理流程：**本研究提出了一套後處理流程，能自動輸出包含劣化座標、平面位置圖、對應影像及所屬構件的檢測報告書。並利用 COLMAP 進行稠密點雲重建，讓使用者透過 CloudCompare 與 PyVista 等工具，將辨識出的劣化資訊與 BIM 模型或重建的三維點雲進行對位與可視化展示，有助於橋檢人員直觀理解劣化位置與情況。
 - **邊緣運算與系統整合：**本系統以 NVIDIA Jetson Orin Nano 作為邊緣運算平台，在 ROS 環境下整合各項模組，驗證了在資源受限的嵌入式設備上執行複雜 AI 模型的潛力，並提升了系統的即時反應能力。

5.2 研究限制

儘管本研究達成預期目標，但在研究過程中仍遭遇部分限制，條列如下：

- **無人機電量限制：**無人機的電池續航力有限，可能影響單次檢測的範圍與時長，特別是對於大型橋梁或需長時間作業的場景。
- **飛航法規與區域限制：**無人機的飛行需遵守相關民航法規，部分區域（如高速公路、鐵路周邊）需持有相關證照並額外申請許可方能進行檢測。
- **影像辨識模型的泛化能力：**雖然本研究已使用多來源資料集進行影像辨識模型訓練，涵蓋多種劣化情況，但現階段主要仍以日常巡檢中常見的劣化型態為主，如混凝土裂縫、剝落與白華等，對於地震、水災等災害情境下可能出現的特殊破壞類型（如構件斷裂、橋面隆起、墊層滑移等）尚未納入訓練資料，限制了模型在災後緊急應變時的應用效能。
- **SLAM 定位精度與穩定性：**純視覺 SLAM 在環境特徵較少的區域（如大面積單調的橋面或水面反射區域），定位精度和穩定性可能會下降，甚至出現

追蹤失敗的情況。寶橋實驗中就觀察到多次 tracking loss 導致地圖斷裂。

- **環境因素的影響：**無人機的飛行與影像擷取易受風速、光照、橋下障礙物（如寶橋實驗中的懸空電纜）等環境因素影響，可能導致影像品質不佳或飛行安全問題，進而影響檢測結果的穩定性與完整性。
- **點雲重建與對位的複雜性：**雖然本研究建立了點雲重建與 BIM 對位的流程，但單眼相機進行的稠密點雲重建在精度和完整性上仍有限制。且點雲與 BIM 模型的自動化精準對位仍具挑戰，目前流程中仍需部分人工介入（如 CloudCompare 中的手動點對配準）。
- **邊緣運算平台的效能限制：**雖然 Jetson Orin Nano 展現了不錯的邊緣運算能力，但同時執行高複雜度的影像辨識與 SLAM 運算，現階段仍可能面臨運算資源的極限，影響系統的即時性與可擴展性。

5.3 未來研究建議

基於本研究的成果與限制，未來可從以下幾個方向進行深入研究：

- **提升影像辨識模型泛化能力：**擴充涵蓋多橋型、材質、光照與災後情境（如地震斷裂、洪災沖刷等）的劣化影像資料，建構具代表性的訓練集。亦可結合歷年巡檢紀錄與航拍影像進行標註，搭配少量標註學習、遷移學習與自監督方法，提升模型於災後情境的適應力。另可探索資料增強與多模態學習（如融合 LiDAR 資訊）技術，以強化模型整體的泛化與實用性。
- **增強 SLAM 系統定位精度與穩定性：**應用視覺慣性 SLAM（Visual-Inertial SLAM, VI-SLAM）技術，整合 IMU 數據以提升在動態環境或特徵稀疏區域的定位穩健性。

- 
- **UAV 自主飛行與路徑規劃：**結合 BIM 模型或初步重建的稀疏地圖，進行 UAV 自主飛行路徑規劃，以確保檢測的覆蓋率與影像品質；並研究基於強化學習的 UAV 自主避障與視角調整策略，以適應複雜的橋下環境。
 - **應用數位孿生技術：**深化 BIM 模型與實測點雲的整合應用，建立更精確的橋梁數位孿生（Digital Twin），實現劣化演變追蹤與預測性維護。
 - **雲端協同運算：**探索雲端協同運算架構，將部分非即時性的高複雜度運算（如大規模點雲處理）移至雲端，實現雲端與邊緣裝置協同作業。

透過上述方向的持續研究與改進，期望能進一步完善 UAV 橋梁劣化即時檢測系統，使其在橋梁檢測工作中發揮更大的應用價值。





參考文獻

- [1] 劉翠溶. 八十年來臺灣的都市發展. 經濟論文, 20(2):1–35, 1991.
- [2] Ministry of Transportation and Communications (MOTC). 全國橋梁統計資訊網.
<https://bss.iot.gov.tw/bss/public/public/login/>, 2025. [資料取得日期：2025 年 6 月 6 日].
- [3] 中央通訊社. 南方澳大橋斷裂事故調查報告出爐監造單位疏失釀禍, 11 2020.
最後查閱日期：2025-02-14.
- [4] P. C. Chang, A. Flatau, and S. C. Liu. Review paper: Health monitoring of civil infrastructure. Structural Health Monitoring, 2(3):257–267, 2003.
- [5] C. R. Farrar and K. Worden. An introduction to structural health monitoring. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1851):303–315, 2007.
- [6] J. M. W. Brownjohn. Structural health monitoring of civil infrastructure. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1851):589–622, 2007.
- [7] C. M. Yeum and S. J. Dyke. Vision-based automated crack detection for bridge

inspection. Computer-Aided Civil and Infrastructure Engineering, 30(10):759–770, 2015.



- [8] Y. J. Cha, W. Choi, and O. Büyüköztürk. Deep learning-based crack damage detection using convolutional neural networks. Computer-Aided Civil and Infrastructure Engineering, 32(5):361–378, 2017.
- [9] Mahta Zakaria, Enes Karaaslan, and F. Necati Catbas. Advanced bridge visual inspection using real-time machine learning in edge devices. Advances in Bridge Engineering, 3(1):27, 2022.
- [10] NVIDIA Corporation. Edge computing with nvidia jetson: Enabling real-time ai at the edge, 2024.
- [11] M. Mandirola, C. Casarotti, S. Peloso, I. Lanese, E. Brunesi, and I. Senaldi. Use of uas for damage inspection and assessment of bridge infrastructures. International Journal of Disaster Risk Reduction, 72:102824, 2022.
- [12] American Association of State Highway and Transportation Officials. Drones: Improving safety and cutting costs for bridge inspections. Technical report, Asphalt Magazine, 2016.
- [13] Alon Oring. Fast and robust structural damage analysis of civil infrastructure using uav imagery. arXiv preprint arXiv:2110.04806, 2021.
- [14] Sina Tavasoli, Xiao Pan, T. Y. Yang, Saudah Gazi, and Mohsen Azimi. Autonomous damage assessment of structural columns using low-cost micro aerial vehicles and multi-view computer vision. arXiv preprint arXiv:2308.16278, 2023.

[15] NEXCO 東日本. 介紹使用無人機 (uav) 拍攝的影片進行鋼橋檢查, 2024.
Accessed: 2025-02-14.

[16] 交通部運輸研究所. 橋梁檢測工具效能提升計畫, 2024.

[17] 交通部. 無人機搭配 ai 影像辨識應用於橋梁檢測之研究 (2/2). Technical report, 交通部, 2024.

[18] 交通部. 無人機搭配 ai 影像辨識應用於橋梁檢測之研究 (1/2). Technical report, 交通部, 2022.

[19] Emesent. Routine bridge inspection using slam, 2024.

[20] 交通部. 公路橋梁檢測及補強規範, 2018.

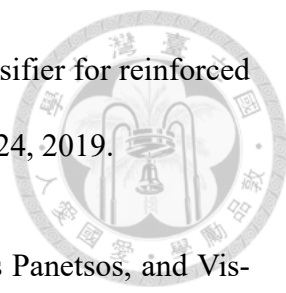
[21] 臺北市政府工務局新建工程處. 橋梁檢測成果, 2024.

[22] Ikhlas Abdel-Qader, Osama Abudayyeh, and Michael E. Kelly. Analysis of edge-detection techniques for crack identification in bridges. Journal of Computing in Civil Engineering, 17(4):255–263, 2003.

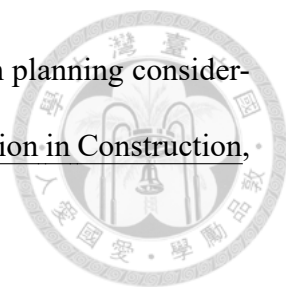
[23] Qin Zou, Zheng Zhang, Qingquan Li, Xianbiao Qi, Qian Wang, and Song Wang. Deepcrack: Learning hierarchical convolutional features for crack detection. IEEE Transactions on Image Processing, 28(3):1498–1512, 2019.

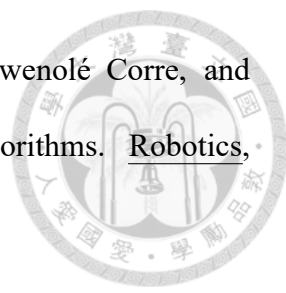
[24] Billie F. Spencer, Vedhus Hoskere, and Yasutaka Narazaki. Advances in computer vision-based civil infrastructure inspection and monitoring. Engineering, 5(2):199–222, 2019.

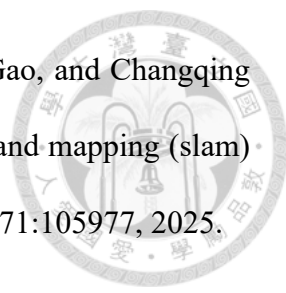
[25] Wenjun Wang and Chao Su. Automatic concrete crack segmentation model based on transformer. Automation in Construction, 139:104275, 2022.

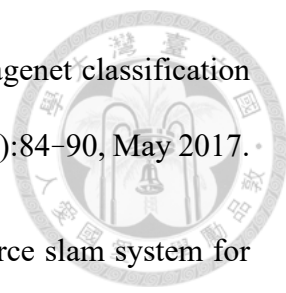
- 
- [26] Philipp Hühnwahl, Ruodan Lu, and Ioannis Brilakis. Multi-classifier for reinforced concrete bridge defects. Automation in Construction, 105:102824, 2019.
- [27] Martin Mundt, Sagnik Majumder, Sreenivas Murali, Panagiotis Panetsos, and Visvanathan Ramesh. Meta-learning convolutional neural architectures for multi-target concrete defect classification with the concrete defect bridge image dataset. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 11188–11197, 2019.
- [28] Jingdao Chen and Yong Kwon Cho. Crackembed: Point feature embedding for crack segmentation from disaster site point clouds with anomaly detection. Advanced Engineering Informatics, 52:101550, 2022.
- [29] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2016.
- [30] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection, 2020.
- [31] Chien-Yao Wang, I-Hau Yeh, and Hong-Yuan Mark Liao. Yolov9: Learning what you want to learn using programmable gradient information, 2024.
- [32] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks, 2016.
- [33] Rahmat Ali, Dongho Kang, Gahyun Suh, and Young-Jin Cha. Real-time multiple damage mapping using autonomous uav and deep faster region-based neural networks for gps-denied structures. Automation in Construction, 130:103831, 2021.

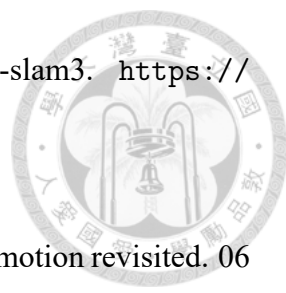
- 
- [34] Chaobo Zhang, Chih-chen Chang, and Maziar Jamshidi. Concrete bridge surface damage detection using a single-stage detector. Computer-Aided Civil and Infrastructure Engineering, 35(4):389–409, 2020.
- [35] Lucas Prado Osco, José Marcato Junior, Ana Paula Marques Ramos, Lúcio André de Castro Jorge, Sarah Narges Fatholahi, Jonathan de Andrade Silva, Edson Takashi Matsubara, Hemerson Pistori, Wesley Nunes Gonçalves, and Jonathan Li. A review on deep learning in uav remote sensing. International Journal of Applied Earth Observation and Geoinformation, 102:102456, 2021.
- [36] 王姿樺 (Zi-Hua Wang), 高書屏 (Szu-Pyng Kao), and 林志憲 (Jhih-Sian Lin). 應用深度學習技術輔助橋梁裂縫辨識. 航測及遙測學刊, 27(4):247–257, Dec 2022.
- [37] 張永辰 (Yung-Chen Chang), 高書屏 (Szu-Pyng Kao), 王豐良 (Feng-Liang Wang), and 林志憲 (Jhih-Sian Lin). 應用 uav 影像及深度學習技術輔助橋梁裂縫量化分析. volume 11, pages 15–33. 中華民國地籍測量學會, Jan 2023.
- [38] 蔡宜真. 以無人機建立即時自動化橋梁裂縫影像辨識系統. 碩士論文, 國立臺灣大學, 2024.
- [39] Szu-Pyng Kao, Feng-Liang Wang, Jhih-Sian Lin, Jichiang Tsai, Yi-De Chu, and Pen-Shan Hung. Bridge crack inspection efficiency of an unmanned aerial vehicle system with a laser ranging module. Sensors, 22(12), 2022.
- [40] A. Ellenberg, A. Kontsos, F. Moon, and I. Bartoli. Bridge deck delamination identification from unmanned aerial vehicle infrared imagery. Automation in Construction, 72:155–165, 2016.
- [41] Francesco Nex and Fabio Remondino. Uav for 3d mapping applications: a review. Applied Geomatics, 6(1):1–15, 03 2014.

- 
- [42] Neshat Bolourian and Amin Hammad. Lidar-equipped uav path planning considering potential locations of defects for bridge inspection. Automation in Construction, 117:103250, 2020.
- [43] Sofia Tilon, Francesco Nex, George Vosselman, Irene Llave, and Norman Kerle. Towards improved unmanned aerial vehicle edge intelligence: A road infrastructure monitoring case study. Remote Sensing, 14:4008, 08 2022.
- [44] Yiching Lin, Chin-Yu Hsiao, Jian-Hua Tong, Chih-Pin Liao, Shin-Tai Song, Hsin-Chu Tsai, and Jui-Lin Wang. Application of edge computing in structural health monitoring of simply supported pci girder bridges. Sensors, 22:8711, 11 2022.
- [45] Daniel Jozi, Nima Shirzad-Ghaheroudkhani, Garvit Luhadia, Shaghayegh Abtahi, and Mustafa Gül. Rapid post-disaster assessment of residential buildings using unmanned aerial vehicles. International Journal of Disaster Risk Reduction, 111:104707, 2024.
- [46] Raj Hakani and Abhishek Rawat. Edge computing-driven real-time drone detection using yolov9 and nvidia jetson nano. Drones, 8(11), 2024.
- [47] Jui-Sheng Chou and Chi-Yun Liu. Optimized lightweight edge computing platform for uav-assisted detection of concrete deterioration beneath bridge decks. Journal of Computing in Civil Engineering, 39(1):04024045, 2025.
- [48] Cesar Cadena, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, José Neira, Ian Reid, and John J. Leonard. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Transactions on Robotics, 32(6):1309–1332, 2016.

- 
- [49] Andréa Macario Barros, Maugan Michel, Yoann Moline, Gwenolé Corre, and Frédérick Carrel. A comprehensive survey of visual slam algorithms. Robotics, 11(1), 2022.
- [50] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In 2011 International Conference on Computer Vision, pages 2564–2571, 2011.
- [51] Raúl Mur-Artal, J. M. M. Montiel, and Juan D. Tardós. Orb-slam: A versatile and accurate monocular slam system. IEEE Transactions on Robotics, 31(5):1147–1163, 2015.
- [52] Carlos Campos, Richard Elvira, Juan J. Gómez Rodríguez, José M. M. Montiel, and Juan D. Tardós. Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam. IEEE Transactions on Robotics, 37(6):1874–1890, 2021.
- [53] Liang Zhang, Leqi Wei, Peiyi Shen, Wei Wei, Guangming Zhu, and Juan Song. Semantic slam based on object detection and improved octomap. IEEE Access, 6:75545–75559, 2018.
- [54] Fangwei Zhong, Sheng Wang, Ziqi Zhang, China Chen, and Yizhou Wang. Detect-slam: Making object detection and slam mutually beneficial. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pages 1001–1010, 2018.
- [55] Chu-Qiao Feng, Bao-Luo Li, Yu Fei Liu, Zhizhong Fu, Yonghong Yan, and Junqi Fan. Crack assessment using multi-sensor fusion simultaneous localization and mapping (slam) and image super-resolution for bridge inspection. Automation in Construction, 155:105047, nov 2023.

- 
- [56] Huitong Xu, Meng Wang, Cheng Liu, Yongchao Guo, Zihan Gao, and Changqing Xie. Tunnel crack assessment using simultaneous localization and mapping (slam) and deep learning segmentation. Automation in Construction, 171:105977, 2025.
- [57] Zhengyi Chen, Changhao Song, Boyu Wang, Xingyu Tao, Xiao Zhang, Fangzhou Lin, and Jack C.P. Cheng. Automated reality capture for indoor inspection using bim and a multi-sensor quadruped robot. Automation in Construction, 170:105930, 2025.
- [58] Tao Lu, Sonja Tervola, Xiaoshu Lü, Charles Kibert, Qunli Zhang, Tong Li, and Zhi-tong Yao. A novel methodology for the path alignment of visual slam in indoor construction inspection. Automation in Construction, 127:103723, 07 2021.
- [59] Xiao Pan, Sina Tavasoli, T. Y. Yang, and Sina Poorghasem. Post-disaster building indoor damage and survivor detection using autonomous path planning and deep learning with unmanned aerial vehicles, 2025.
- [60] 交通部. 公路橋梁檢測及補強規範. https://www.motc.gov.tw/ch/app/divpubreg_list/view?module=divpubreg&id=740&serno=425, 2023.
- [61] Roboflow. Defects-6bfhc dataset [data set]. <https://universe.roboflow.com/x-i0m9x/defects-6bfhc>, 2024.
- [62] Roboflow. Bridge-ssz3a dataset [data set]. <https://universe.roboflow.com/peccunb/bridge-ssz3a>, 2024.
- [63] B. Dwyer, J. Nelson, and T. Hansen. Roboflow (version 1.0). <https://roboflow.com>, 2024.

- 
- [64] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. Commun. ACM, 60(6):84–90, May 2017.
- [65] Raul Mur-Artal and Juan D. Tardos. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. IEEE Transactions on Robotics, 33(5):1255–1262, October 2017.
- [66] Jakob Engel, Thomas Schöps, and Daniel Cremers. Lsd-slam: Large-scale direct monocular slam. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, Computer Vision – ECCV 2014, pages 834–849, Cham, 2014. Springer International Publishing.
- [67] Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, 2022.
- [68] Sanghyun Park. Yolov9 ros. https://github.com/SanghyunPark01/yolov9_ros, 2024.
- [69] Christoph Zauner. Implementation and benchmarking of perceptual image hash functions. 2010.
- [70] WongKinYiu. Yolov9: Learning what you want to learn using programmable gradient information. <https://github.com/WongKinYiu/yolov9>, 2024.
- [71] Zhaohui Zheng, Ping Wang, Wei Liu, Jinze Li, Rongguang Ye, and Dongwei Ren. Distance-iou loss: Faster and better learning for bounding box regression, 2019.
- [72] Mark Everingham, S. M. Ali Eslami, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. The pascal visual object classes challenge: A retrospective. International Journal of Computer Vision, 111(1):98–136, 2015.

- 
- [73] Thien Nguyen. orb_slam3_ros: A ros implementation of orb-slam3. https://github.com/thien94/orb_slam3_ros, 2021.
- [74] Johannes Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. 06 2016.
- [75] S. Ullman and Sydney Brenner. The interpretation of structure from motion. Proceedings of the Royal Society of London. Series B. Biological Sciences, 203(1153):405–426, 1979.
- [76] Michael Goesele, Noah Snavely, Brian Curless, Hugues Hoppe, and Steven Seitz. Multi-view stereo for community photo collections. pages 1–8, 11 2007.
- [77] CloudCompare Development Team. CloudCompare [GPL software]. <http://www.cloudcompare.org/>, 2025.
- [78] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Open3d: A modern library for 3d data processing, 2018.
- [79] P.J. Besl and Neil D. McKay. A method for registration of 3-d shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2):239–256, 1992.
- [80] C. Bane Sullivan and Alexander A. Kaszynski. Pyvista: 3d plotting and mesh analysis through a streamlined interface for the visualization toolkit (vtk). Journal of Open Source Software, 4(37):1450, May 2019.
- [81] Jon Louis Bentley. Multidimensional binary search trees used for associative searching. Commun. ACM, 18(9):509–517, September 1975.
- [82] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones,

Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2023.

[83] Abhinav Shrivastava, Abhinav Gupta, and Ross Girshick. Training region-based object detectors with online hard example mining, 2016.

[84] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks, 2015.