國立勤益科技大學工業工程與管理系

專　題　製　作

影像辨識-利用 YOLOv8 辨別道路危機

組員：

四工三甲　3B115006　林琮祐

四工三甲　3B115011　蔡暘鏞

四工三甲　3B115013　范梓葳

四工三甲　3B115035　盧奕安

四工三甲　3B015041　江浩維

指導教授：洪永祥 教授

中　華　民　國　一　一　四　年　五　月　一　日

摘要

隨著城市化進程加速，道路交通事故和基礎設施問題日益嚴重，即時偵測和預警道路危機變得至關重要。在本研究中，我們採用了先進的物件偵測模型YOLOv8，針對道路障礙物及設計不良（如坑洞、裂縫和水溝蓋）進行自動化檢測是否有無障礙物(如:坑洞、水溝蓋、裂縫)。

在本研究中，我們收集了大量的道路障礙物影像數據集並一一標註關鍵特徵，使用 YOLOv8 進行模型訓練和優化(超參數調整)。開發過程中，我們利用 Anaconda 進行機器學習，並選擇 Spyder 作為開發環境進行影像辨識。透過數據增強技術，如圖像翻轉、旋轉和亮度調整，使本研究提高了模型的泛化能力。

同時，使用預訓練模型進行微調，不僅縮短了訓練時間，還提升了辨識精度。然而，在光線不足或背景複雜的環境中，模型表現有所下降。未來研究可著重於提升模型在各種環境下的穩定性，擴充數據集以增強模型的泛化能力，並探索將模型應用於移動設備或無人機，以實現即時巡檢，進一步提高道路維護效率。

這項研究展示了 YOLOv8 在道路危機辨識中的巨大潛力。透過技術改進和數據優化，我們相信未來能使模型更高效，降低道路危機對交通安全的影響，並促進基礎設施管理與維護的便利性。

關鍵詞: YOLOv8、Spyder、道路障礙物、影像辨識

致謝

首先，我們要由衷地感謝洪永祥教授。洪教授在專題初期即為我們提供了明確且具前瞻性的研究方向，讓我們能夠在一開始就建立起清晰的目標與實作方向。在整個專題進行的每個階段，教授都不辭辛勞地參與我們的討論會，對我們提出的想法與執行方式給予具體而深入的建議，讓我們能夠少走許多彎路，避免重複不必要的錯誤。教授嚴謹而溫和的教學態度，讓我們在面對問題與挑戰時學會用更全面的角度去思考，培養了解決問題的能力，也提升了我們對研究的熱情。正因為有洪教授在學術上與精神層面上的全力支持與鼓勵才能夠達到目前的深度與品質。

此外，我們也誠摯感謝范淳凱學長在整個專題進行期間的協助與指導。學長以自身的實作經驗與專業知識，在我們遭遇技術瓶頸與困惑時，總是能夠即時提供實用的解決方案，幫助我們釐清許多程式設計與模型訓練上的關鍵問題。特別是在資料前處理、模型參數調整以及實驗設計等方面，范學長提出了許多務實可行的建議，使我們能夠更有效率地完成實驗並取得可靠的成果。學長不僅技術能力優異，也總是耐心地傾聽我們在研究過程中的困難與想法，並給予我們莫大的鼓勵與支持，這對我而言是非常寶貴的經驗與助力。

目錄

摘要.....I

致謝.....II

目錄.....III

表目錄.....V

圖目錄.....VI

第一章 緒論.....1

1.1 研究背景.....1

1.2 研究動機.....2

1.3 研究目的.....2

1.4 研究架構與流程.....3

1.5 研究限制.....4

第二章 文獻探討.....5

2.1 不同版本的 YOLO(YOU ONLY LOOK ONCE)模型介紹.....5

2.2 卷積神經網路(CONVOLUTIONAL NEURAL NETWORK,CNN)介紹.....7

2.3 卷積神經網路結構.....10

2.4 卷積神經網絡(CNN)的演進與架構比較.....11

2.5 YOLOV8 模型中的激活函數分析.....12

2.6 背景及重要性.....15

2.7 相關研究比較.....18

第三章 研究方法與步驟.....20

3.1 實驗方法.....20

3.2 數據蒐集（DATA GATHERING）.....21

3.3 數據預處理(DATA PREPROCESSING).....21

3.4 選擇模型(MODEL SELECTION).....22

3.5 模型訓練(MODEL TRAINING).....24

3.6 測試模型(MODEL TESTING).....24
3.7 數據增強、超參數調整(DATA AUGMENTATION & HYPERPARAMETER TUNING).....25
第四章 實驗結果與分析.....27
4.1 資料及介紹.....27
4.2 環境設置.....28
4.3 評估指標.....29
4.4 超參數設定.....31
4.5 實驗結果.....32
第五章 結論與未來研究方向建議.....44
5.1 結論.....44
5.2 未來展望.....45
參考文獻.....46

表目錄

表 2-1 YOLO 各版本特點整理 6
表 2-2 CNN 介紹 8
表 2-3 CNN 架構演進 11
表 2-4 SiLU（Swish）激活函數特性表 13
表 2-5 YOLOv8 系列激活函數演進 13
表 2-6 實際性能比較 15
表 3-1 模型比較表 22
表 4-1 數據蒐集來源比較表 29
表 4-2 混淆矩陣 29
表 4-3 各類別表現比較表 33
表 4-4 各類別的召回率變化趨勢表 35
表 4-5 分類模型或偵測模型的性能分析指標表 36
表 4-6 各類別的 AP 值表 38
表 4-7 類別曲線解析表 40
表 4-8 損失函數觀察（左六張圖）表 42
表 4-9 評估指標變化（右四張圖）表 43

圖目錄

圖 1-1 研究架構圖 3
圖 2-1 物件偵測示意圖 .....錯誤! 尚未定義書籤。
圖 2-2 LeNet 卷積神經網路架構.....8
圖 2-3 卷積層示意圖 9
圖 2-4 池化層示意圖 9
圖 2-5 全連結層示意圖 10
圖 2-6 事故傷害原因死亡人數占比 17
圖 3-1 實驗流程架構圖 20
圖 3-2 預測標註圖 25
圖 4-1 訓練集資料夾 28
圖 4-2 測試集資料夾 28
圖 4-3 模型訓練 PYTHON 程式碼 31
圖 4-4 模型測試 PYTHON 程式碼.....31
圖 4-5 精確率與置信度曲線圖 32
圖 4-6 召回率與置信度曲線圖 34
圖 4-7 精確率與召回率曲線圖 37
圖 4-8 F1 分數與置信度曲線圖 39
圖 4-9 訓練與驗證集指標視化圖 41

第一章 緒論

1.1 研究背景

隨著全球經濟和人口的快速增長，大家對交通運輸的需求也越來越大。與此同時，道路安全的問題也變得更加嚴重了，對個人生命和整個社會經濟都帶來了很大的影響。根據世界衛生組織（World Health Organization,WHO）的報告，每年有超過 119 萬人因交通事故身亡，也有數百萬人因此受傷甚至殘疾。這些事故不僅僅是生命的損失，還造成了巨大的經濟壓力，全球範圍內因事故造成的損失大約佔 GDP 的 3%。[1]

面對這些挑戰，科技創新成為解決問題的重要工具，尤其是人工智慧（Artificial Intelligence,AI）[2]的應用，為了改善道路安全帶來了新的希望。運用 AI 不僅能快速處理大量數據，還能通過深度學習[3]從而辨識實現高精度的預測和即時反應。在 AI 技術中，它可以從監控影像、行車記錄器等數據來源中擷取有用資料，用於危險辨識，為交通管理者和駕駛者提供即時有效的支援。

影像辨識技術[4]之所以能發展得這麼快，全是因為深度學習和電腦視覺技術的進步快速。透過構建卷積神經網絡（Convolutional Neural Network,CNN）[5][35]這樣的深度學習模型，影像辨識在目標檢測、物體追蹤和模式分類等方面取得了很好的成果。目前，這些技術已進一步應用於交通系統中，主要集中於兩個方向：一、偵測道路違規行為。二、辨識道路現況（如路面損壞），這些狀況可能導致交通事故或擁堵。透過影像辨識，系統能迅速發現問題，並協助相關部門即時採取行動，例如清除障礙物或修復路面，從而降低事故發生的風險。

雖然影像辨識技術的潛力很大，但在實際應用中也有不少挑戰需要克服：如數據多樣性不足，如果影像數據的質量和多樣性不夠高，像是光線不足、天氣影響或者攝影機角度不妥，都可能導致模型準確性下降。所以我們需要收集更多不同情境的影像數據，並用數據增強技術來模擬真實場景，提升模型的穩定性。

總之，影像辨識技術在改善道路安全和交通管理方面的應用前景進一步廣

泛運用。未來，這項技術還可以和物聯網、自動駕駛、智慧信號燈等智慧城市技術結合起來，打造更高效、安全的交通系統。它還可以為政策制定提供支持，比如根據事故數據優化基礎設施，或者進行安全教育。與保險業務結合也很有潛力，可以根據駕駛行為為用戶設計個性化的保險方案，促進更安全的駕駛行為。雖然前路仍有挑戰，但隨著技術不斷進步，影像辨識必然會成為提升道路安全的重要支柱，讓全球交通系統變得更安全、更高效。

1.2 研究動機

在現代社會中，交通安全日益重要，隨著車輛數量的增加，造成交通事故頻率不斷上升，也會因為天雨路滑視線不佳或是道路上的設計不良(水溝蓋)導致騎士自摔。影像辨識技術作為人工智慧的一項重要應用，能有效識別道路危機，提升道路安全並減少事故損失。我們對影像辨識的興趣源於相關課程的學習，期望在專題中深入探討其在道路危機管理的應用，分析不同情境下的影像資料，以有效檢測和預警潛在危機。

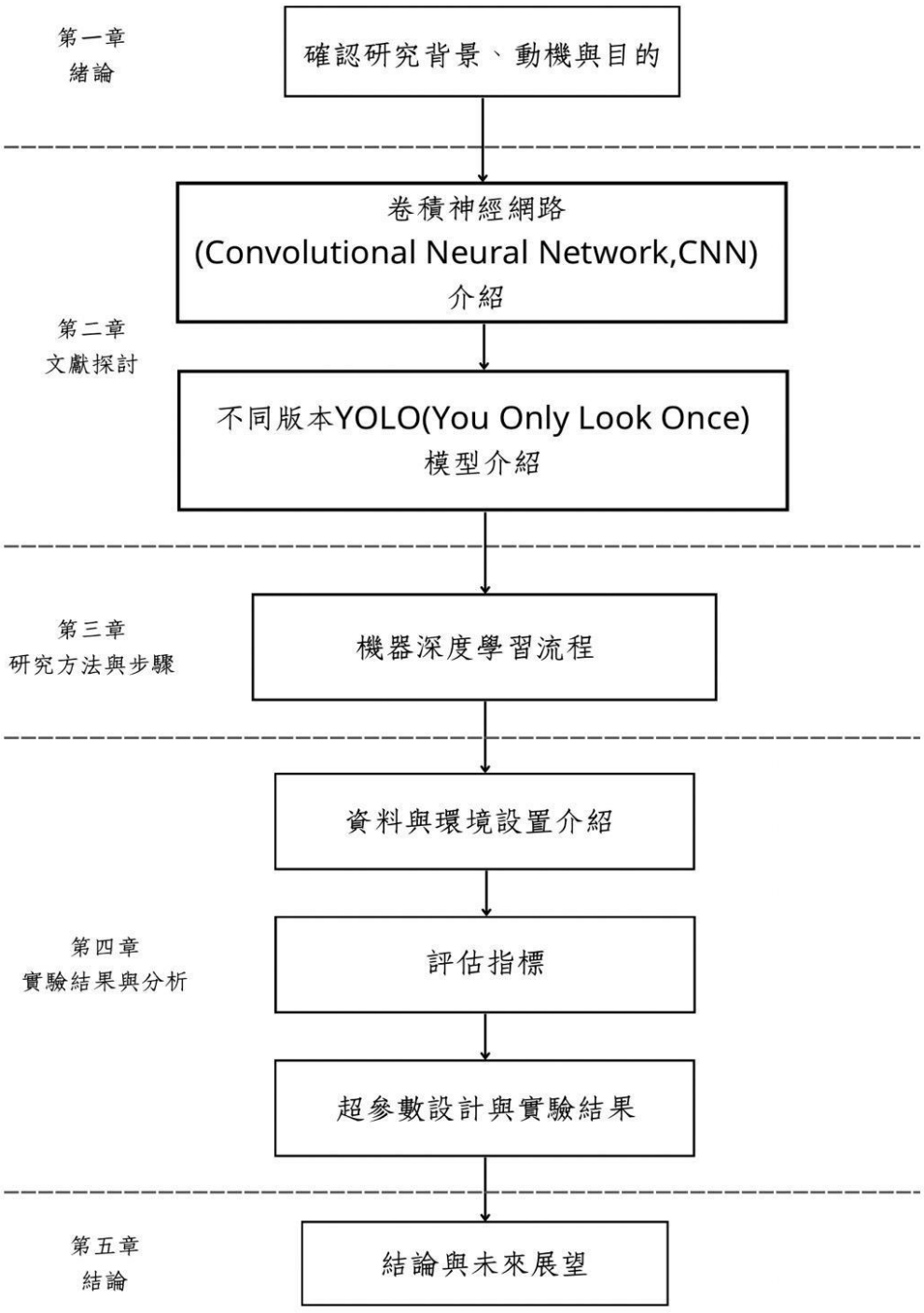
台灣的交通事故率長期居高不下，每年因道路事故導致數千人傷亡，對家庭和社會造成巨大衝擊。現有的道路設施與管理制度不周，容易造成道路危機。作為經常騎行機車的民眾，經常親眼目睹許多危險的交通情況，深感交通規範和基礎設施的不足。因此，希望透過專題研究，深入探討台灣道路危機的成因，並提出系統性改進建議且配合道路危機辨識系統，希望成為政府交通政策的參考，改善台灣道路安全環境。[6][7][8]

1.3 研究目的

本研究旨在探索利用人工智能影像辨識技術識別道路危機，以提高道路安全性和交通管理效率。通過分析道路影像，AI 系統能夠即時識別潛在的道路障礙危機因素，並為相關管理部門提供及時預警。研究目標主要為開發高效的影像處理模型，能夠準確識別各種道路危險。最終，本研究旨在為智能交通系統的發展提供理論支持和實踐指導，推動道路安全和交通流暢性提升。

1.4 研究架構與流程

圖 1-1 研究架構圖



資料來源:本專題綜合整理

1.5 研究限制

本研究可能面臨以下限制:

技術限制:設備版本老舊，舊設備在處理速度、資料儲存容量及感測精度方面可能影響資料的準確性與完整性。

環境限制:台灣道路地形挑戰，山區道路蜿蜒，增加了資料收集的難度。

第二章 文獻探討

本章節旨在探討卷積神經網路（Convolutional Neural Network,CNN）與 YOLO 物件偵測技術之原理與發展，並分析其於影像辨識領域中的應用價值。首先介紹 CNN 的結構與特性，說明其在圖像處理上的優勢；接著詳述 YOLO 系列模型的演進過程，從 YOLOv1 至 YOLOv8，闡述其即時性與準確性如何促進智慧監控與自動檢測技術的應用。最後，回顧台灣傳統道路檢測方式的限制，並透過日本交通安全制度與 AI 應用案例，作為本研究發展方向的重要參考依據。

2.1 不同版本的 YOLO(You Only Look Once)模型介紹

YOLO 是一種革命性的即時物體檢測系統，其名稱「You Only Look Once」體現了它只需對圖像「看一眼」就能進行檢測的特性。[9][36]

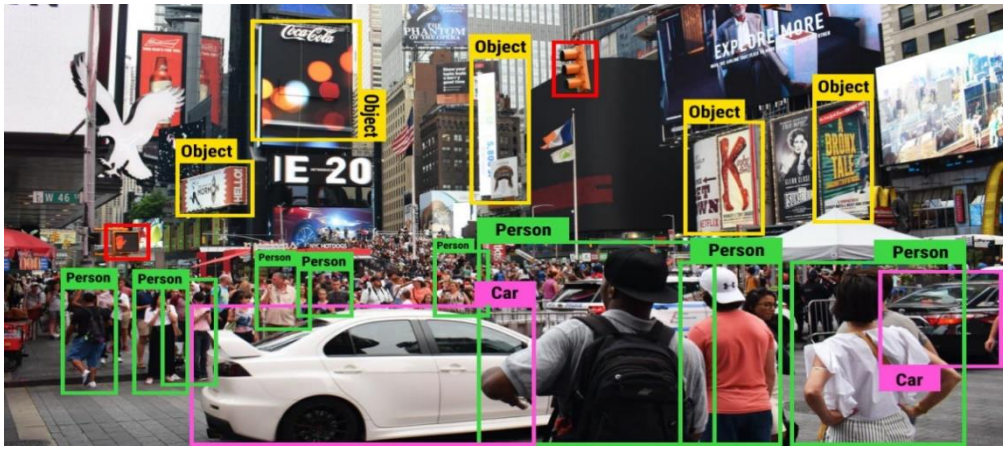


圖 2-1 物件偵測示意圖 **[30]**

它將物體檢測視為單一回歸問題，使用統一的神經網絡從完整圖像直接預測邊界框和類別概率，與傳統兩階段檢測方法形成鮮明對比。YOLO 的工作原理是將輸入圖像分割成網格，每個網格單元負責預測包含物體中心的邊界框及其類別，之後通過非極大值抑制移除重複檢測。雖然現代 YOLO 架構更為複雜，但我們可以類比早期的卷積神經網絡如 LeNet 來理解其基本結構：如同 LeNet 使用卷積層提取特徵、池化層降維、最後通過全連接層進行分類，YOLO 同樣採用卷積層提取視覺特徵，但擴展為更深更寬的網絡，並將最終輸出轉換為網格

化的空間預測。自 2015 年問世以來，YOLO 經歷了多次迭代，從 YOLOv1 到 YOLOv8，每個版本都引入了重要改進，如更深的卷積架構、錨框機制、多尺度特徵融合、改進的損失函數和創新的訓練策略等。YOLO 的主要優勢在於極快的處理速度、較低的背景誤檢率和良好的泛化能力，但在檢測密集小物體和不尋常比例物體方面仍有局限。由於其高效性和實用性，YOLO 被廣泛應用於自動駕駛、監控系統、機器人視覺、醫學影像分析等眾多領域，成為物體檢測領域的一個重要里程碑，重新定義了速度與準確度的權衡關係，為實時視覺應用開闢了新的可能性。[9][29]

表 2-1 YOLO 各版本特點整理

版本	發表年份	主要特點
YOLOv1	2015	1.首個 YOLO 版本，由 Joseph Redmon 發出 2.單個卷積網路同時進行分類與定位 3.對重疊目標的處理較弱 4.快速但對小物體偵測較差
YOLOv2/YOLO9000	2016	1. 引入批次標準化（Batch Normalization）2. 改進錨框（anchor boxes）3.使用 Darknet-19 作為骨幹網路 4.支援 9000 類別的物件偵測 5.加入多尺度訓練（multi-scale training）
YOLOv3	2018	1.使用更新的 Darknet-53 作為骨幹網路 2.結合三層不同尺度的特徵圖進行預測 3.最後層不再使用 softmax4.精度高於前幾代，但速度略低
YOLOv4	2020	1.由 Alexey Bochkovskiy 開發 2.引入新骨幹網：CSPDarknet53，結合 PANet、Mish 激活函數 3.使用數據增強技術：Mosaic、CutMix4. 引入自由錨框機制（Anchor-Free）

YOLOv5	2020	1.由 Ultralytics 公司開發，於 PyTorch 實現 2.提供多種不同尺寸的模型（ nano,small,medium,large,xlarge ） 3.模型部署方便 4.加入自動學習錨框功能
YOLOv6	2022	1.由美團發表2.強化分類性能3.引入更高效的骨幹網與輕量化設計 4.對小型目標偵測優化
YOLOv7	2022	1.進一步改進模型架構 2.引入可擴展的高效設計 3. 強化訓練方法 4.在處理小物體方面有顯著改進
YOLOv8	2023	1.由 Ultralytics 推出 2.全新架構與自由錨點（anchor-free）的檢測頭3.支持多種任務：偵測、分類、分割、姿態估計 4.整合追蹤與部署功能

資料來源：[14][36]

2.2 卷積神經網路(Convolutional Neural Network,CNN)介紹

卷積神經網路（Convolutional Neural Network, CNN）是一種深度學習架構，專門設計用來處理具有類似網格結構的數據，例如影像。其核心概念在於利用卷積層自動擷取影像中的特徵，再透過池化層與全連接層進行分類或預測。CNN 在圖像分類、目標檢測、語音辨識等多個領域取得重大突破，其中經典架構如 LeNet、AlexNet、VGGNet、ResNet 等，皆為後續模型發展奠定了基礎。

在影像辨識任務中，CN 能有效提取局部空間特徵，進而理解整體影像語義，這使其成為自動駕駛、醫療影像分析、工業瑕疵檢測等應用的關鍵技術。

表 2-2 CNN 介紹

階段	圖片區塊說明	功能
輸入層(Input)	原始圖片	資料輸入
卷積層 1（Convolution Layer 1）	多個卷積後的特徵圖（feature maps）	特徵提取 + 非線性
池化層 1（Pooling 1）	pooled feature maps	特徵壓縮
卷積層 2（Convolution Layer 2）	第二層更抽象的特徵圖	深層特徵提取
池化層 2（Pooling 2）	更小尺寸的 pooled features	特徵壓縮
全連接層（Fully-connected 1）	一長串神經元	綜合特徵
輸出層（Output）	softmax 機率分佈	分類結果

資料來源：[10] [25][26]

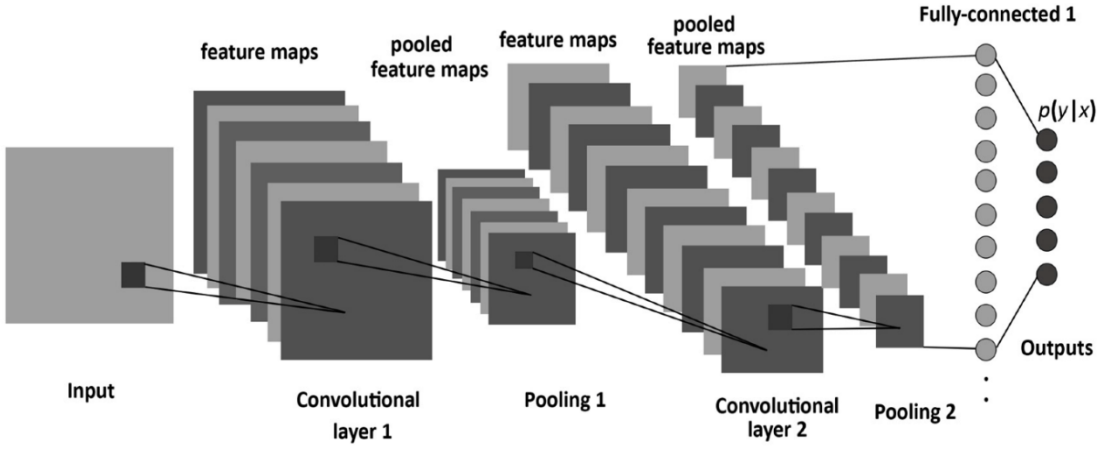


圖 2-2 LeNet 卷積神經網路架構**[10]**

卷積層（Convolutional Layer）

卷積層是 CNN 的核心組件，用於從輸入資料中自動提取局部特徵。它透過卷積核（也稱濾波器）在輸入上滑動，計算內積，從而產生特徵圖（Feature Map）。每個卷積核可以學習圖像中的特定結構，如邊緣、角落或紋理，進而

逐層建構出更抽象的特徵表示。卷積操作保留了圖像的空間關係，使模型能夠有效辨識不同區域的圖形模式。

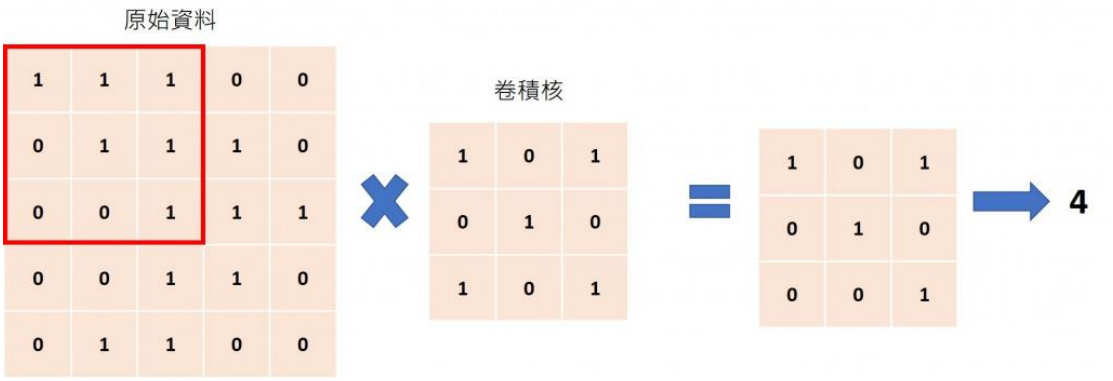


圖 2-3 卷積層示意圖**[11]**

池化層（Pooling Layer）

池化層的主要功能是降低特徵圖的空間尺寸，從而減少參數數量與計算量，同時提高模型對小幅度平移、旋轉等變化的容忍度。最常見的是最大池化（Max Pooling），其會從一個區塊中選出最大值，保留最重要的特徵訊號。透過池化，模型可更專注於重要特徵，忽略不必要的細節與雜訊。

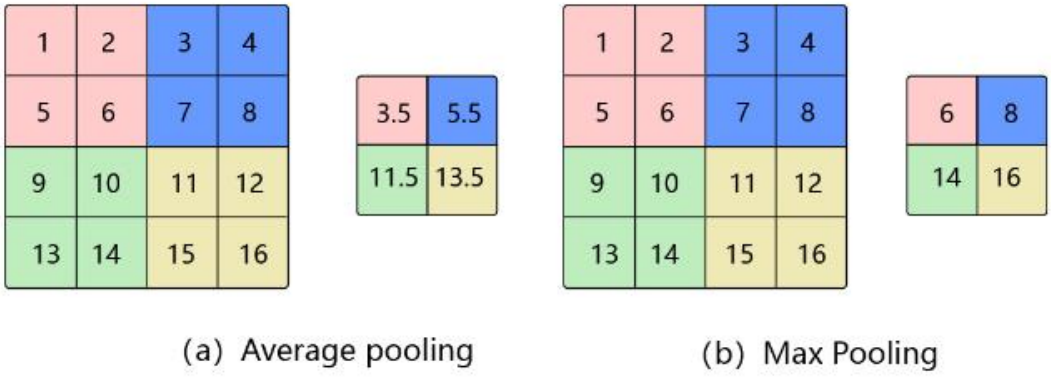


圖 2-4 池化層示意圖**[12]**

全連結層（Fully Connected Layer）

全連結層是 CNN 的最後幾層，負責將前面提取的高階特徵整合起來並進行分類或預測。在這一層中，每個神經元將與前一層的所有輸出單元相連。輸入通常會先被展平成一維向量，再傳入一或多個全連結層，最後透過像是 Softmax 的激活函數輸出各類別的機率分佈。

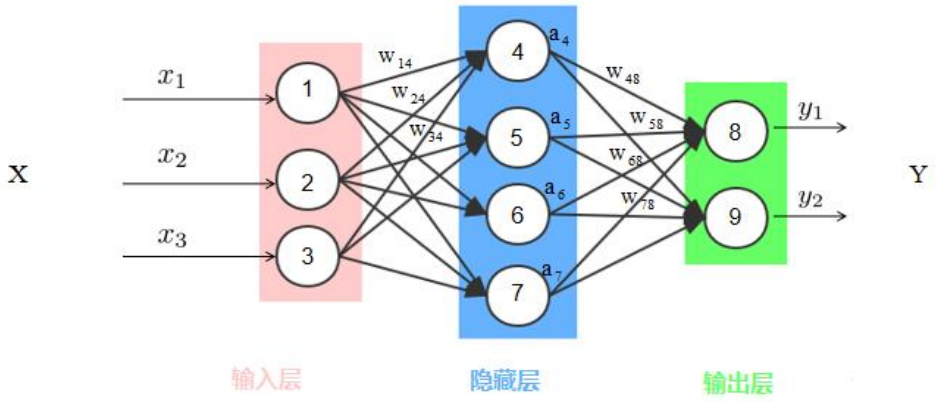


圖 2-5 全連結層示意圖**[13]**

2.3 卷積神經網路結構

卷積神經網路（CNN）應用於即時物件偵測時，通常可分為三大結構：骨幹網路（Backbone）、頸部網路（Neck）與頭部網路（Head），並結合輔助模塊以強化模型效能。

骨幹網路（Backbone）

功能：特徵提取

骨幹網路是整個模型的起點，主要目的是從原始輸入圖片中提取出豐富的多層次特徵。這些特徵包含了圖像的邊緣、紋理、形狀等資訊，為後續目標辨識打下基礎。

頸部網路（Neck）

功能：特徵融合與強化

從 Backbone 傳來的特徵圖會進入 Neck 網路，這個階段的目的是將不同解析度與語意層級的特徵進行融合。

頭部網路（Head）

功能：輸出預測結果

Head 是模型的最後一段，負責將前面處理過的特徵圖轉換為具體的輸出結果，包括每個物體的邊界框座標（bounding box）、分類（class label）以及置信度分數（confidence score）。

2.4 卷積神經網絡(CNN)的演進與架構比較

卷積神經網絡（CNN）自從 1998 年 LeNet-5 的提出以來，經歷了多次架構的演進，並在圖像分類、目標檢測等領域取得了驚人的成果。不同的 CNN 架構在特定的任務中展現出了各自的優勢，以下是幾種重要的卷積神經網絡架構的介紹及比較。

表 2-3 CNN 架構演進

架構	發表年份	特點與技術	備註
LeNet-5	1998	最早的 CNN 架構之一 應用於手寫數字識別（MNIST）層數較少	Yann Le Cun 提出
AlexNet	2012	使用 ReLU、Dropout 使用 GPU 訓練 ImageNet 冠軍	深度學習熱潮起點
VGGNet	2014	小卷積核（3x3）堆疊結構簡單但參數多常見 VGG16/VGG19	適合影像分類任務
ResNet	2015	殘差學習（Residual Learning）跳接連接（Skip Connections）可訓練 100+層深度網路	ResNet-18/34/50/101/152
Darknet	YOLOv2/v3 使用	為 YOLO 設計的骨幹網路使用 3x3 卷積核+x1 支援殘差結構	YOLOv2 用 Darknet-19YOLOv3 用 Darknet-53
CSPDarknet	2019 至今	引入 Cross Stage Partial（CSP）結構減少重複計算與參數量訓練穩定	用於 YOLOv4/YOLOv5/ YOLOv8 等

資料來源：[15][16][17][24]

架構比較與發展趨勢

這些 CNN 架構代表了深度學習在計算機視覺領域的關鍵進展：

1.LeNet-5：奠定了 CNN 的基本結構。

2.AlexNet：證明了深度 CNN 的實用性和強大能力。

3.VGGNet：標準化了網絡設計，強調簡潔統一的架構。

4.ResNet：解決深層網絡訓練問題，開創超深網絡時代。

5.Darknet 和 CSPDarknet：將 CNN 優化應用於實時目標檢測，平衡速度與精度。

從這些發展可以看出，CNN 架構的演進趨勢包括：增加網絡深度、改進連接方式、優化計算效率、以及針對特定任務的專門化設計。每一代架構都在前人基礎上解決關鍵問題，並為特定應用場景帶來重要突破。

2.5 YOLOv8 模型中的激活函數分析

激活函數在目標檢測中的作用與選擇

在目標檢測模型中，激活函數對於特徵提取和非線性映射至關重要。YOLOv8 作為目標檢測領域的最新進展，其激活函數的選擇直接影響著模型的性能表現。[22]

YOLOv8 採用的激活函數：SiLU / Swish

YOLOv8 採用了 SiLU（Sigmoid Linear Unit）作為其主要激活函數，也被稱為 Swish 激活函數。這是對之前 YOLO 版本中使用的 Leaky ReLU 的一次重要升級。[23]

表 2-4 SiLU（Swish）激活函數特性表

特性	描述
數學表達式	$f(x) = x \cdot \sigma(x)$ (1) 其中$\sigma(x) = 1/1 + e^{-x}$為 sigmoid 函數。
導數形式	$f'(x) = \sigma(x) + x \cdot \sigma(x) \cdot \big(1 - \sigma(x)\big)$
值域	$(-\infty, \infty)$
曲線特性	近似線性但帶有平滑的彎曲，允許負值小幅度通過。
在 YOLOv8 中的應用位置	骨幹網絡（CSPDarknet）、頸部網絡（PANet）和檢測頭中的所有卷積層之後。

資料來源：[15]

表 2-5 YOLOv8 系列激活函數演進

YOLO 版本	主要激活函數	數學形式	特點	貢獻
YOLOv1–v3	Leaky ReLU	$f(x) = max\,(0.1x, x)$ (2)	解決了 ReLU 的神經元死亡問題	穩定性與簡單性兼具
YOLOv4	Mish	$f(x) = x \cdot tanh\big(ln(1 + e^{x})\big)$ (3)	平滑、無界、自正則化	提升特徵表達能力
YOLOv5	SiLU/Swish	$f(x) = x \cdot \sigma(x)$ (4)	計算效率高、性能佳	達到性能與速度平衡
YOLOv8	SiLU/Swish	$f(x) = x \cdot \sigma(x)$ (5)	繼承 YOLOv5 設計並優化	持續提升精度與效率

資料來源: [15][17]

表 2-6 實際性能比較

激活函數	mAP（COCO 資料集）	推理速度	訓練穩定性
ReLU	基準線	最快（+5%）	中等
Leaky ReLU	+0.5%	相當	良好
ELU	+0.7%	較慢（-3%）	良好
Mish	+1.5%	較慢（-7%）	優異
SiLU/Swish	+1.4%	中等（-2%）	優異

資料來源:[18]

SiLU/Swish 在 YOLOv8 中的優勢

非單調性：不同於 ReLU 系列的單調增加，SiLU 在負值區域有輕微下降，允許更豐富的特徵表達。

平滑激活曲線：相比 ReLU 的銳角轉折，SiLU 沒有不連續點，有利於梯度傳播。

邊界效應：

當 x 遠大於 0 時，$f(x)\approx x$

當 x 遠小於 0 時，$f(x)\approx 0$

這種平滑過渡有利於模型訓練。

優化收斂：幫助模型更穩定地收斂至更佳的局部最優解。

自注意力特性：因為 sigmoid 可解釋為「權重」，所以 $x\sigma(x)$類似於為輸入賦 予重要性。

表 2-6 實際性能比較

激活函數	mAP（COCO 資料集）	推理速度	訓練穩定性
ReLU	基準線	最快（+5%）	中等
Leaky ReLU	+0.5%	相當	良好
ELU	+0.7%	較慢（-3%）	良好
Mish	+1.5%	較慢（-7%）	優異
SiLU/Swish	+1.4%	中等（-2%）	優異

資料來源: [19]

YOLOv8 採用 SiLU/Swish 激活函數是一項合理且高效的選擇。它在訓練穩定性、特徵學習能力及運算效率間取得了良好平衡。此選擇延續了 YOLO 系列在追求實時應用中的一貫設計理念，使得 YOLOv8 成為目前最具競爭力的即時目標檢測模型之一。

2.6 背景及重要性

傳統的檢測方法

傳統的道路路況觀察方法主要依賴人力和傳統設備，缺乏現代技術的輔助。通過定期人工巡檢，養護人員按路線徒步或駕車沿線路巡查，肉眼識別坑洞、裂縫、水溝堵塞等問題，並使用書面表格或 3C 設備記錄問題位置與嚴重程度。這些人員通常需依靠個人經驗判斷問題的嚴重性和修復優先順序。在記錄方面，工作人員會使用捲尺、水平儀等簡單測量工具來量化問題大小，用相機拍攝照片作為證據。傳統系統還高度依賴市民報告，通過電話熱線、信件或當面報告向當局反映道路問題。資料管理則主要採用紙質檔案系統，將巡檢表格、照片和維修記錄存檔，缺乏有效的數字化管理。維修決策往往基於經驗判斷和有限預算，而非數據分析，導致資源分配可能不夠合理。傳統方法的缺點明顯，包括效率低下、人力資源需求大、觀察不連續、數據不精確、難以建立長期趨勢

分析，且受主觀因素影響較大。

台灣近年事故

近年來，因道路破損（如裂縫、坑洞）與設計不良的水溝蓋所導致的交通事故逐漸受到重視。根據交通部公路局統計，台灣每年約有超過 3,000 件機車自摔事故與路面不平或坑洞有關，佔整體單一車輛事故的約 5%。其中雨季期間，因積水掩蓋坑洞、裂縫等危險，事故數量更有上升趨勢。另外，監察院曾指出，部分地區的水溝蓋設計與施工不良，包括高低落差過大、縫隙過寬、表面材質過滑等，也造成輪胎打滑或卡住，導致騎士摔倒或車輛損壞。在 2022 年台北市的資料中，就有超過 400 件機車事故與「路面障礙物」或「設施缺陷」有直接關聯。這些問題反映出道路基礎設施管理的漏洞，需透過高效率巡檢與科技輔助（如影像辨識、AI 道路檢測）來即時掌握異常，降低交通事故風險。

台灣 2022 年年事故分析

根據交通部與警政署資料，2022 年台灣共發生超過 30 萬件交通事故，造成約 2,560 人死亡、超過 40 萬人受傷，顯示整體死傷情形較前一年上升。其中，機車事故占比高達 92%，仍為主要肇事與死傷來源。高齡者（65 歲以上）在事故中死亡人數達 910 人，約占總死亡人數 35%，是高風險族群之一；行人死亡人數則為 394 人，平均每日近 47 名行人在道路上發生死傷事故，凸顯行人安全不足。

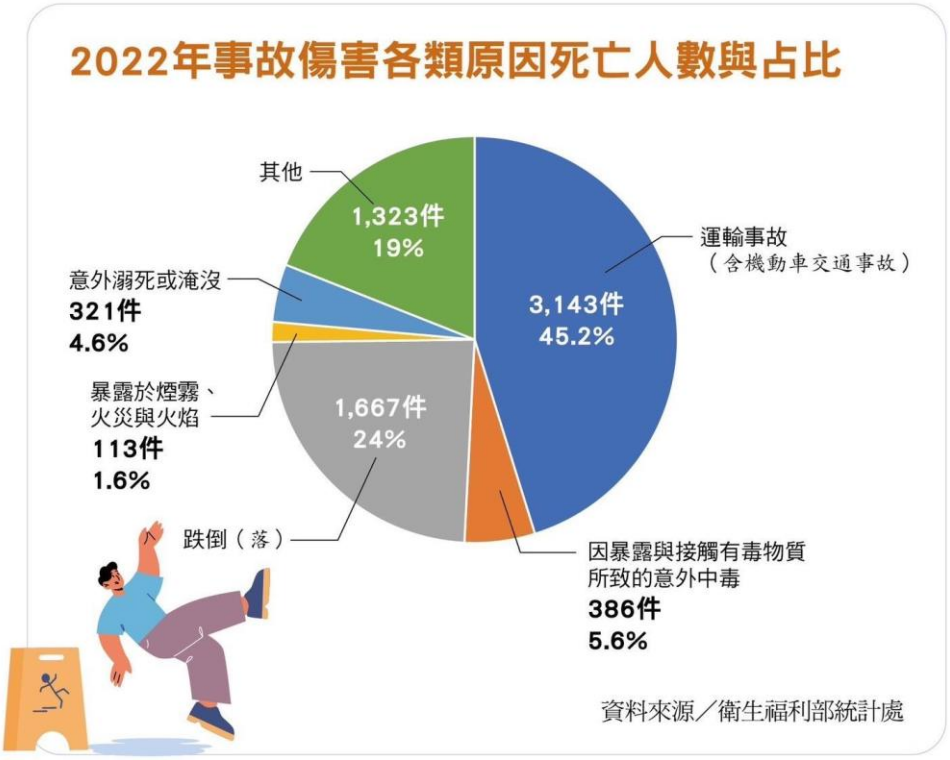


圖 2-6 事故傷害原因死亡人數占比**[20]**

亞洲地區-日本交通事故

1.駕駛人守法意識強

日本人普遍很遵守交通規則，不論是開車、騎車還是走路，紅燈不闖、酒駕極少見。

2.嚴格的駕照制度

日本考駕照非常難，不只要上駕訓班，還有筆試、實際路考等，費用高、過程長，能拿到駕照的人通常駕駛技術和安全意識都不錯。

3.罰則重、執法嚴

違規會被高額罰款，還會扣分甚至吊銷駕照。警察取締很認真，酒駕、超速、用手機開車這些行為都會被嚴厲處理。

4.交通設施完善

像是人行道、交通號誌、斑馬線、監視器設計得都很人性化，也有很多先

進的交通安全科技，保護行人與駕駛。

影像辨識在日本交通上的應用

1.智慧型監視器與紅綠燈

很多十字路口安裝了具備影像辨識功能的監視器，可以即時偵測車流、人流，甚至違規行為（像闖紅燈、違規轉彎）自動取證。

2.自駕與先進駕駛輔助系統（ADAS）

像 Toyota、Honda 等車廠，早就把影像辨識應用在車輛的安全系統上，例如車道偏移警示、自動剎車、行人辨識、交通號誌識別等。

3.智慧停車系統

透過影像辨識來確認車輛是否停在格內、有沒有違規停車，有的城市甚至可以用手機 App 查即時空位。

4.鐵路與平交道監控

日本也會在鐵路與平交道裝設影像系統來偵測異常情況，比如有人跌倒或車卡住時，能立即通報。

5.行人與高齡者安全輔助

有些地方會裝設影像系統辨識行人（尤其是老人或小孩）靠近馬路時，自動延長紅綠燈時間或發出語音警示。

2.7 相關研究比較

目前已有許多研究將深度學習技術應用於道路損壞檢測或交通物件辨識領域。例如：

Zhang et al.（2020）使用 Faster R-CNN 模型辨識高速公路坑洞，雖具有高準確率，但計算量大、處理速度較慢。

Li et al.（2021）將 YOLOv5 應用於城市道路裂縫檢測，證實 YOLO 系列在準確度與效率間取得良好平衡。

Nguyen et al.（2023）引入 YOLOv7 對多類型道路異常進行即時辨識，顯示其在邊緣裝置部署方面具備潛力。

然而，大多數研究仍面臨以下挑戰：模型對不同光照與天氣條件的適應性不足、資料標註成本高、邊緣部署限制等。因此，本研究選擇最新 YOLOv8 模型，嘗試在上述限制中尋求突破。

第三章 研究方法與步驟

3.1 實驗方法

本章節將介紹本研究所採用的機器學習方法，並詳細描述實驗的設計和實施步驟。我們將探討有關於機器學習的應用-影像辨識，並介紹深度學習和電腦視覺技術在影像辨識中的應用。本研究運用 YOLOv8 模型進行自動化檢測，並透過輕量化設計以提升模型的效率與準確率。在實驗設計方面，將說明資料集的建構方式、訓練環境的設定，以及模型效能的評估流程，目的是確保即使在資源有限的硬體平台上也能穩定運作。

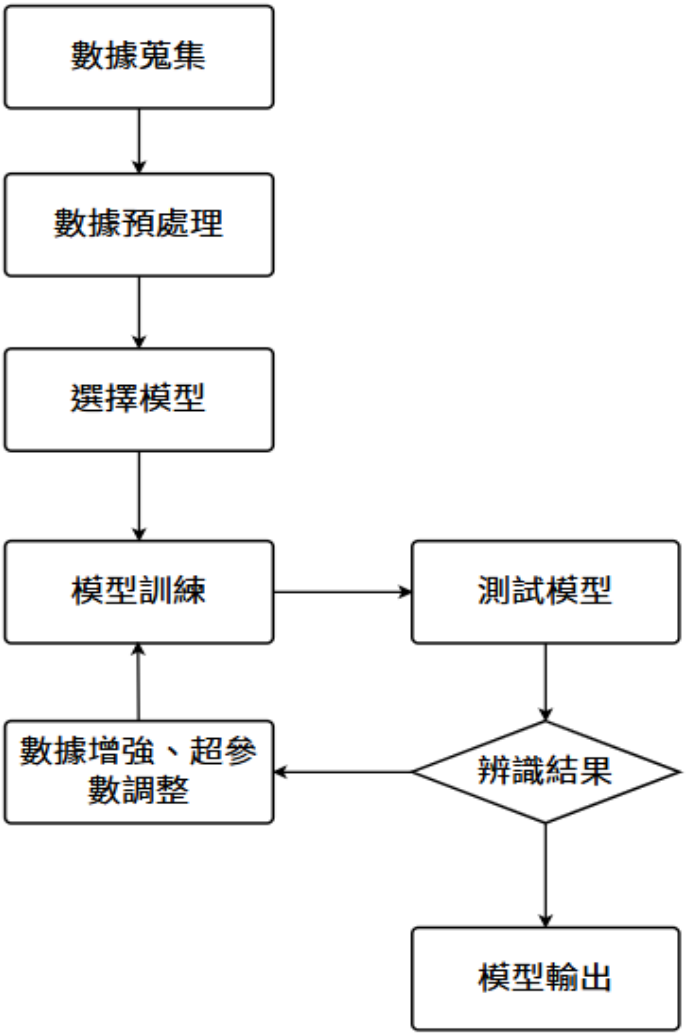


圖 3-1 實驗流程架構圖

資料來源:本專題綜合整理

3.2 數據蒐集（Data Gathering）

數據搜集的定義是透過各種手段和工具，有系統性的搜集、研究或應用相關目標的數據，方便後續預處理、訓練跟測試，也是人工智慧、數據分析和機器學習中重要的步驟。數據搜集的核心為:

1.多樣性

數據搜集可以來自多重來源，例如：Google Maps、社交媒體、開放數據平台、自行拍攝等，這些不同來源的數據有助於提供更全面的資訊，並增加結果的準確性。

2.過程系統化

數據搜集需要明確的規劃，包括選擇合適的工具與方法或是確保數據的質量跟一致性，也需要遵守倫理範圍，尊重著作權。

3.數據質量要求

數據搜集需要足夠全面、準確、具相關性，及有足夠的數量以保證後續分析的精確性、有效性和可靠性。

3.3 數據預處理(Data Preprocessing)

數據預處理的定義是在進行影像數據分析或模型訓練之前，把影像數據清理與整理的過程，目的是為了提高數據的品質，以保證分析結果與模型的準確性和穩定性。在影像數據預處理中，比較常見的操作大致有以下幾個步驟：

1.去除模糊影像

拍照時，有可能因為拍攝時不停的晃動，導致鏡頭對焦不準等原因，導致拍攝出來的影象模糊不清楚，有些細節就沒有辦法展現在影像上，這時候就要整理這些沒有用的影像。

2.剔除異常影像

造成影像問題有許多原因，像是檔案毀損或是沒有我們需要的內容物，為了避免讓這些異常影像影響到數據分析的準確性，所以需要人工來檢查、整理

後刪除。

3.移除重複影像

過多重複的影像數據會導致訓練數據失去平衡而且還會增加不需要的計算負擔。我們使用了 ImageFinder 這個程式來檢測並且刪除與數據重複的影像。

4.影像擴充

網路上所能搜集到的資料有限，為了能夠增加我們影像辨識的準確率，我們利用各種方法來擴充我們的資料集，像是旋轉、翻轉、裁剪、調整顏色或者改變影像尺寸等方式，來增加我們資料集的多樣性，以提高模型的穩定性。

3.4 選擇模型(Model Selection)

相較於其他物件偵測模型（如 SSD、Faster R-CNN），YOLOv8 在速度與準確率間取得最佳平衡，尤其適合需即時反應之道路影像辨識任務。以下為模型比較摘要：

表 3-1 模型比較表

模型	偵測速度（FPS）	精確度（mAP）	複雜度	是否適合即時應用
Faster R-CNN	7–10	高	高	否
SSD	20–30	中	中	是
YOLOv5	40–60	高	中	是
YOLOv8	60–80	極高	低	是

資料來源:[21]

YOLOv8 模型特點

YOLOv8 是由 Ultralytics 於 2023 年推出的最新版本，是 YOLO 系列至今最成熟且多功能的一代。YOLOv8 不僅支援即時目標檢測任務，還整合了影像分割（Segmentation）、姿態估測（Pose Estimation）等功能，具備高度彈性與擴展性。與前幾代相比，YOLOv8 在準確度與速度上取得更佳平衡，主要創新點包括以下幾項：

1.Anchor-Free 檢測頭（Decoupled Head）

YOLOv8 拋棄傳統 anchor-based 設計，採用 anchor-free 模式，大幅簡化了訓練流程並提升小物體辨識能力。此外，輸出 head 採用 decoupled 設計，分別處理分類與邊界框回歸，使特徵學習更加獨立與精準。

2.改進的骨幹網路（Backbone）

YOLOv8 採用輕量化與高效能兼具的骨幹網路結構，其核心模組為 C2f（Concatenate-2-Fusion），相較於以往的 C3 模組可節省更多參數與運算資源，同時保有豐富特徵表達力。該骨幹網路具備優良的提取能力，適用於高解析輸入。

3.功能模組與特徵融合（Neck: FPN + PAN）

YOLOv8 延續 FPN（Feature Pyramid Network）與 PAN（Path Aggregation Network）的多層特徵融合策略，使模型能同時兼顧大、中、小物體的偵測，提升在密集場景中的辨識效果。

4.改進的損失函數

YOLOv8 採用改良版的 Distribution Focal Loss（DFL）來精確回歸邊界框，並結合 CIoU（Complete IoU）損失進一步提升定位準確度，對於物體位置的預測更加穩健。

5.自動圖像增強與後處理技術

YOLOv8 支援 Mosaic、MixUp 等先進的資料增強策略，並內建 NMS（Non-Max Suppression）、Class-Aware NMS 等後處理技巧，有效濾除重複檢測並提升推論品質。

6.多任務支援與模組化訓練

YOLOv8 將多種任務（Detection,Segmentation,Pose）整合進統一架構，使模型可根據應用需求彈性調整。使用者可透過 Ultralytics 的 API 進行高效部署、微調與模型轉換，降低研究與開發門檻。

綜上所述，YOLOv8 在本研究中為最合適選擇，能有效處理多變的道路影像資料，並實現即時異常偵測的目標。

3.5 模型訓練(Model Training)

當我們進行模型建立時，首先需對原始資料集進行切分，依照常見比例（如 8:2 或 7:3）區分為訓練集與驗證集，以確保模型在未見資料上具備良好的泛化能力。數據切分可採隨機方式，必要時亦可依資料特性採分層抽樣（stratified sampling）。接著進行超參數設計，針對所選模型設定適當的超參數（hyperparameters），如學習率（learning rate）、批次大小（batch size）、訓練輪數（epochs）、激活函數（activation function）與損失函數（loss function）等，以提升訓練效率並降低過擬合風險[28]。模型訓練階段透過誤差反向傳播（backpropagation）與優化器（optimizer）不斷調整權重，同時可搭配正則化（regularization）與早停機制（early stopping）以強化模型穩定性。每輪訓練後，利用驗證集監控模型表現，如準確率（accuracy）與損失值（loss）變化，以判斷模型是否收斂。最終將表現最佳的模型儲存（model saving），以利後續測試與應用部署。

3.6 測試模型(Model Testing)

在此階段採用於驗證集上表現最佳的模型版本，透過儲存於訓練過程中的最佳權重（best weights），以確保後續評估能反映模型的最強辨識能力與穩定性。

測試重點包含分類準確率（accuracy）、精確率（precision）、召回率（recall）、F1 分數（F1-score）等評估指標，並輔以混淆矩陣（confusion matrix）觀察各類別間的預測表現。此外，也檢視模型在不同資料類型下的反應，以模擬實際應用場景。[31]

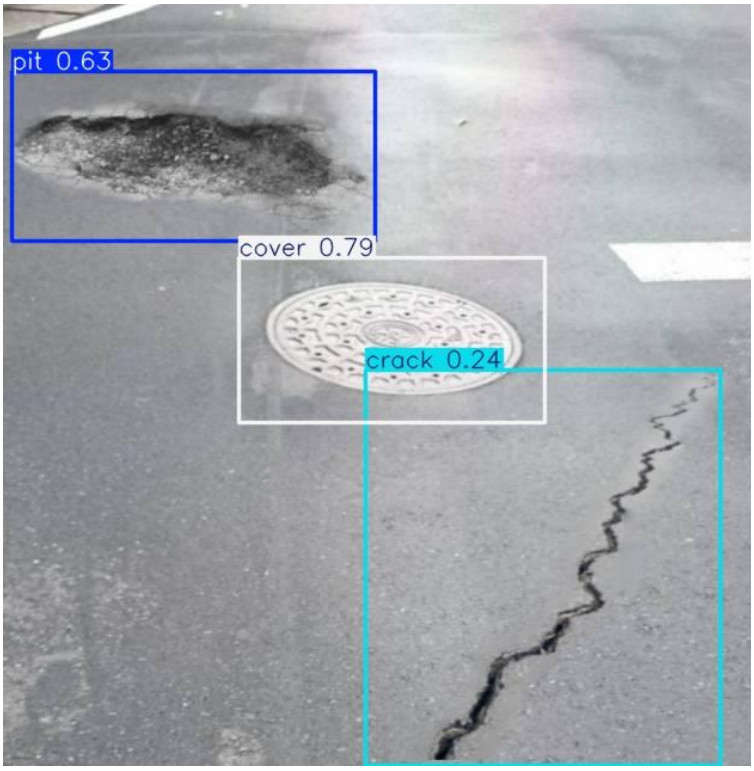


圖 3-2 預測標註圖

資料來源:本專題綜合整理

圖中有三個被辨識出來的區域，每個區域都有對應的標籤與置信度（confidence score）：

1.pit 0.63（藍框）：代表模型偵測到一個坑洞（pothole），置信度為 0.63，表示模型有 63%的把握這個區域是坑洞。

2.cover 0.79（白框）：代表一個人孔蓋（manhole cover），置信度是 0.79。

3.crack 0.24（青藍框）：代表一條裂縫（crack），置信度為 0.24，表示模型對這個偵測比較不確定。

3.7 數據增強、超參數調整(Data Augmentation & Hyperparameter Tuning)

機器學習中，數據增強的技術被廣泛運用來提升模型的泛化能力，尤其是在面對數據有限的時候，研究者通常會圖像翻轉、提高照片亮度、鏡像增強等

調整，來增加數據的多樣性。

圖像翻轉和鏡像增強是最常見的方法，這些技術能模擬不同角度或方向，讓模型學會在各種視角能識別目標，能有效擴充訓練集的多樣性，進而提高模型的穩定性和準確度。

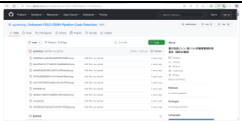
但過度增強圖像可能會帶來誤判問題，特別是對於細節較為敏感的物體檢測中，過度旋轉或過度改變亮度會使得圖像中關鍵特徵的關聯性被模糊，從而影響預測影像準確性。

第四章 實驗結果與分析

4.1 資料及介紹

本研究自 google 街景截圖、google 圖片、開放行數據庫與自行拍攝蒐集總共約五千張裂縫、坑洞與水溝蓋照片作為資料集。如表 4-1 所示，以數據增強將資料及擴充到三萬張。

表 4-1 數據蒐集來源比較表

數據蒐集來源	Google 圖片	Google 街景截圖	自行拍攝	開放型數據庫
圖片				
優點	較容易搜尋得到相關圖片	可以簡單截取到不易到達目的地的照片	可自行選擇拍攝地點、角度、時間.....等	會有人工篩選後的相關照片、數量較多
缺點	容易重複、較會出現台灣以外拍攝的圖片	圖片較模糊	較耗時間	個別數據庫裡的圖片較為單一且重復，若進行訓練前不進行挑選，AI 容易過度擬合

資料來源：本研究綜合整理

隨後，將蒐集完的照片使用 Labelimg 進行人工標註，得到其每張照片物件的位置與類別，並將資料劃分成 9:1，作為訓練、驗證與測試集。

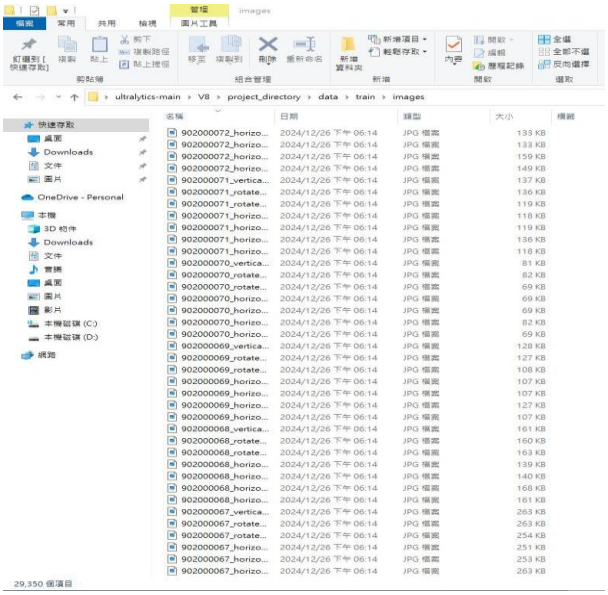


圖 4-1 訓練集資料夾

資料來源:本專題綜合整理

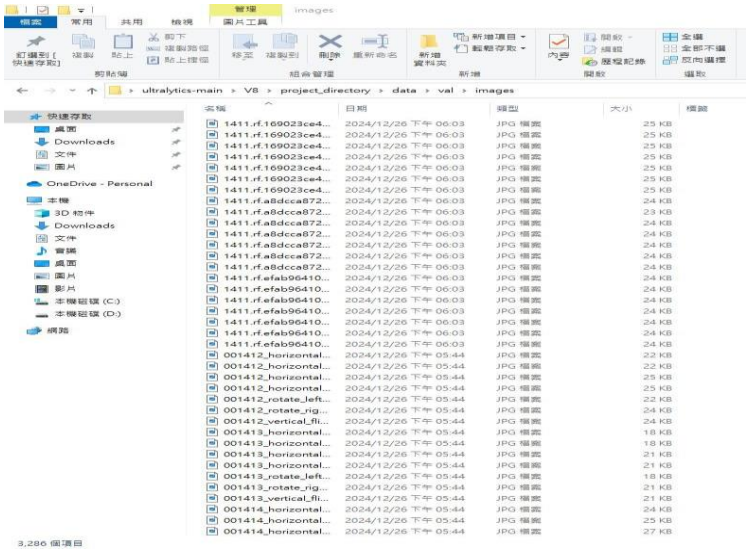


圖 4-2 測試集資料夾

資料來源:本專題綜合整理

4.2 環境設置

GPU:NVIDIA GeForce RTX2080

CPU:Intel64 Family6 Model158

Python:3.10.16

Pytorch:2.7.0

4.3 評估指標

本專題為了全面的評估 YOLOv8 模型在道路破損檢測任務的效能，我們採用多向且常見並且具代表性的評估指標，包括準確率(Accuracy)、精確率(Precision)、召回率(Recall)、F1 分數(F1 score)、平均準確率 AP(Average Precisoin)[37]等。

mAP 為評估物件偵測相關方法是否優秀的常用指標，由於均值精度同時考量精度與召回率，所以此指標可以同時呈現是否有好的偵測精度與是否容易漏檢物體。以下是相關詞的解釋與公式[31][32][34]:

表 4-2 混淆矩陣

實際預測	positive	Negative
positive	真陽性(True Positive,TP)	偽陽性(False Positive,FP)
Negative	真陰性(True Negative,TN)	偽陰性(False Negative,FN)

資料來源： [27]

精確率(Precision):反正模型預測為正類中實際為正類的比例，較能衡量誤報的狀。相關公式如下:

$$\text{Precision}=\frac{TP}{(TP+FP)} \tag{6}$$

召回率(Recall):表示實際正類中被正確預測出來得比例，較能衡量誤報的狀況。相關公式如下:

$$\text{Recall}=\frac{TP}{(TP+FN)} \tag{7}$$

平均準確率(Mean Average Precision,mAP):在物件偵測任務中使用最廣泛的指標之一，綜合考量了模型在所有類別和不同 IoU 閾值下的精確與召回表現，能有效衡量偵測框的準確性。相關公式如下:

$$\frac{1}{N}\sum_{q=0}^{N} AP(q) \tag{8}$$

準確率(Accuracy):衡量模型預測正確的整體比例，但在類別不平衡的情況下，容易產生誤導。相關公式如下:

$$\frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{9}$$

均值精度(Average Precision,AP)為精度，主要依賴於 P-R 曲線的面積總和，代表模型在不同召回率的精度表現。相關公式如下:

$$\text{AP}=\int_0^1 Precision(Recall)d(Recall) \tag{10}$$

置信度(Confidence):指模型對某一預測結果的信心程度，通常以概率形式表達，範圍介於 0 到 1 之間。置信度越高，代表模型越確信該預測是正確的，常用於篩選預測結果以減少誤判。

F1 值(F1 Score):是精確率與召回率的調和平均數，用來綜合評估模型在精確率和召回率間的平衡。其公式為：

$$\frac{2\times Precision\times Recall}{Precision+Recall} \tag{11}$$

4.4 超參數設定



圖 4-3 模型訓練 python 程式碼

資料來源:本專題綜合整理



圖 4-4 模型測試 python 程式碼

資料來源:本專題綜合整理

4.5 實驗結果

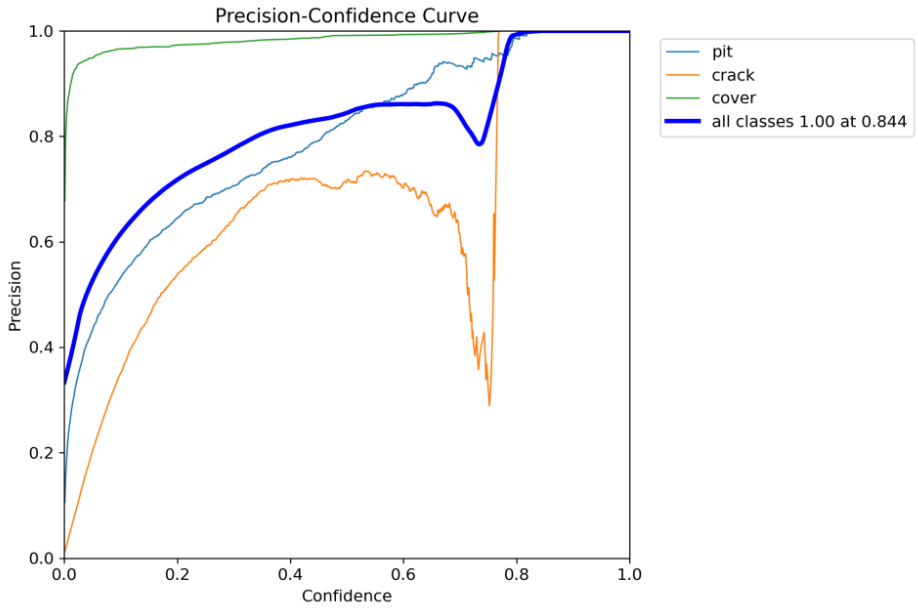


圖 4-5 精確率與置信度曲線圖

資料來源:本專題綜合整理

這張圖展示的是不同類別（pit、crack、cover）在不同置信度閾值下的精確率變化，用來協助設定一個最佳的置信度閾值，以達到高精度預測。

圖中藍色粗線為所有類別整體平均的精確率曲線，最佳置信度設定為 0.844，在此點所有類別平均精確率為 1.00（即預測結果完全正確）。

意義：將模型的預測結果只保留置信度大於 0.844 的部分，能有效排除錯誤預測，提高預測品質，尤其適用於對誤判容忍度低的情況。

表 4-3 各類別表現比較表

類別	顏色	精確率走勢分析
pit（坑洞）	淺藍色	整體表現良好，精確率穩定上升，0.6 之後逐步接近 1.0。適合用於實際應用。
crack（裂縫）	橘色	表現波動大，在 0.6 置信度後大幅下降。**模型對 crack 的判斷不穩定或數據不足**，可能需要更多標註或資料增強。
cover（水溝蓋)	綠色	表現優秀，從低置信度就維持接近 100%精確率。**模型對此類別的辨識最穩定且可靠。**

資料來源：本研究綜合整理

高置信度（如 0.844）：高精度，低召回率（少預測但幾乎不錯）

低置信度（如 0.4）：高召回率，低精度（預測更多但誤判增加）

因此若應用場景對誤判容忍低（如安全監測），建議使用此圖建議的 0.844 作為閾值。後續方向:

1. 針對 crack 類別進行優化

增加 crack 類別的訓練樣本

檢查 crack 標註是否一致與準確

應用資料增強技術如翻轉、對比度變化等。

2. 考慮多閾值策略

為每個類別設定不同的最佳置信度閾值，而不是用單一值 0.844，例如：cover 可使用更低置信度（如 0.5）仍有高精度；crack 可提高閾值避免誤判。

3. 補充其他指標

召回率：衡量找出所有正確目標的比例。

F1 值：精確率與召回率的調和平均，更全面反映模型性能。

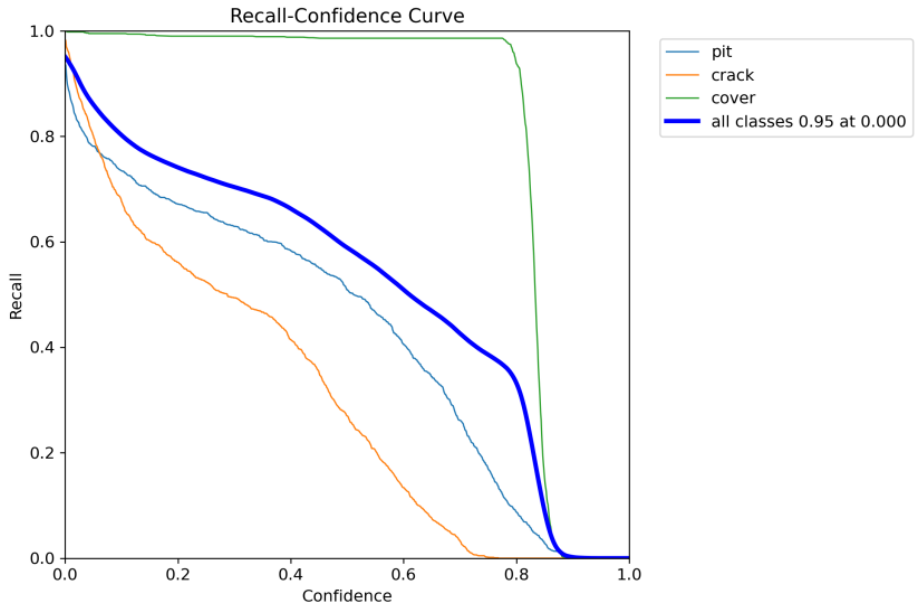


圖 4-6 召回率與置信度曲線圖

資料來源:[33]

這張圖呈現的是在不同的置信度設定下，各類別的召回率表現情況，幫助我們了解模型在不同信心水準下，能夠正確捕捉到多少實際存在的目標（TP/(TP+FN)）。

圖中藍色粗線代表所有類別的總體平均召回率，在置信度為 0.000 時達到最高召回率 0.95。

意義：若完全不設置信度閾值（即讓所有預測都被接受），模型能抓到最多的真實目標，但此時預測的準確性（精確率）會大幅下降，增加誤報。

表 4-4 各類別的召回率變化趨勢表

類別	顏色	召回率分析
pit（坑洞）	淺藍色	隨著置信度提升，召回率穩定下降。表示 pit 預測在各階段都有良好分佈，但高置信度下會漏掉一些實例。
crack（裂縫）	橘色	起始召回率偏低、下降速度快，代表模型對 crack 敏感度低，容易漏檢，這與 Precision 曲線中的不穩定表現相呼應。
cover（水溝蓋）	綠色	幾乎維持 1.0，直到高達 0.9 的置信度才急遽下降。代表 cover 是模型中辨識最穩定、可靠的類別。

資料來源：[31]

與精確率的對比分析:

高置信度時精確率上升，召回率下降：模型更嚴格，只保留最有信心的預測，低置信度時召回上升，精確率下降：模型預測更多，但也包含許多錯誤。

因此若結果希望錯誤越少則要選擇置信度 0.844（Precision 最高），遺漏檢測越少則保留低置信度預測（Recall 最高）。後續方向:

1.找出最佳平衡點（F1-score）

可視覺化 F1-score 與 Confidence 曲線，選擇 F1 分數最高的點作為最終閾值。

2.根據情況需求調整策略

安全導向場景（如交通、工地）：錯過一次代價大須提高召回率（降低信心門檻）。

品質導向場景（如產品檢驗）：誤報代價高須提高精確率（提升門檻）

通用場景：根據 F1-score 或 mAP 分析平衡點。

3. 強化 crack 類別表現

從兩張圖都能看到 crack 的辨識表現明顯落後。

增加 crack 類資料量。

強化 crack 的特徵學習（例如使用特徵融合、注意力機制）。

運用資料增強方法（如模擬裂縫、對比增強）。

表 4-5 分類模型或偵測模型的性能分析指標表

指標	最高值	對應置信度	註解
精確率	1.00	0.844	非常高，但召回率會下降
召回率	0.95	0.000	捕捉率高，但精度低

資料來源：[31][32][33]

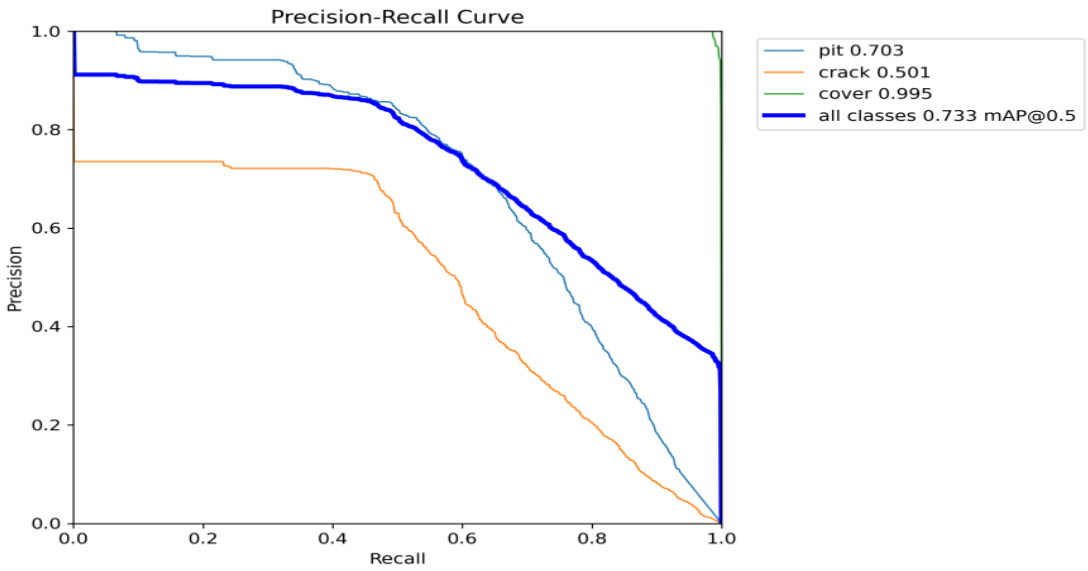


圖 4-7 精確率與召回率曲線圖

資料來源:[31][32][33]

此圖展示的是模型在不同判斷閾值下的精確率與召回率變化，用以評估模型在各類別下的偵測平衡表現，圖中的每條線代表一個類別的表現，藍色粗線代表所有類別的加權平均。

模型整體平均表現：mAP@0.5=0.733

mAP（mean Average Precisionqu,平均平均精度）是常用來評估物件偵測模型的指標。

本圖為 IoU≥0.5 門檻下的平均精度，0.733 表示模型整體表現良好，大多數預測都能兼顧精確與完整。

表 4-6 各類別的 AP 值表

類別	精確率(AP@0.5)	解釋說明	表現
pit（坑洞）	0.703	模型對坑洞的偵測能力佳，召回與精度有合理平衡。	穩定中等表現
crack（裂縫）	0.501	模型對裂縫的判別力較差，常出現誤判或漏判。	須優化
cover(水溝蓋)	0.995	幾乎完美的辨識能力，是模型表現最佳的類別。	幾近完美

資料來源：本研究綜合整理

曲線解讀分析:

cover（綠線）幾乎維持在高精確率區域，即使召回率上升也幾乎不失誤，代表模型對 cover 類極有信心且極少誤判。

crack（橘線）呈現一條下滑曲線且精確率偏低，說明模型在辨識 crack 上信心不足，預測結果雜訊高。

pit（淺藍線）處於中間水準，隨著召回率上升，精確率穩定下降，為典型預期型態。後續方向:

1.提升 crack 類別的表現

增加 crack 樣本數量（提升樣本多樣性）。

加強 crack 的資料增強（模擬真實裂縫條件）。

使用注意力機制或更深層模型來提升辨識力。

2.根據場景需求設定召回/精度平衡點

若錯判成本高（如誤認障礙物）則提高 precision（犧牲部分 recall）。

若漏檢成本高（如漏檢裂縫）則提高 recall（容忍部分誤判）。

可視情況需求設定判斷門檻。

3.搭配 Confidence 曲線做綜合決策

前兩張圖為「PrecisionvsConfidence」、「RecallvsConfidence」，此圖為「PrecisionvsRecall」。

三圖合併觀察可找出最佳的 threshold 與 trade-off 點，設定在 F1-score 最大處較佳。

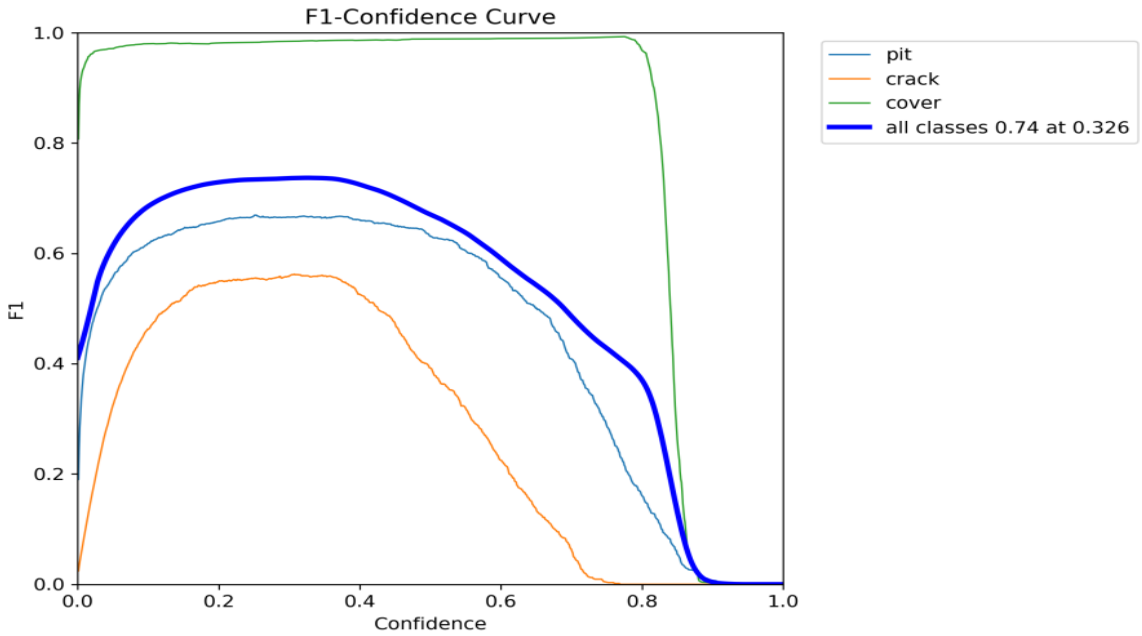


圖 4-8 F1 分數與置信度曲線圖

資料來源:本專題綜合整理

本圖展示模型在不同置信度（confidence threshold）下的 F1 分數變化，用來評估模型對各類別的偵測能力是否達到理想的精確率與召回率平衡。

模型最佳判斷門檻

模型的整體 F1 分數在置信度為 0.326 時達到最大值，F1=0.74，代表在此設定下，精度與召回率達到最佳平衡點，適合作為部署模型時的預設 threshold。

表 4-7 類別曲線解析表

類別	曲線表現	解讀說明
cover（水溝蓋）	曲線極高且平穩（綠色），F1 分數接近 1。	模型對此類極有信心，辨識能力近乎完美，誤判率極低。
pit（坑洞）	中段穩定（淺藍色），高 F1 分數但略低於 cover。	屬於穩定表現，精確與召回維持平衡，模型對 pit 有基本理解。
crack（裂縫）	曲線低且快速下降（橘色）。	模型對 crack 辨識信心不足，預測波動大、常出現誤判與漏判。

資料來源：本研究綜合整理

後續方向:

1 .crack 類須優化

表現明顯較弱透過以下方式改善

增加 crack 樣本數與類型，模擬不同環境與材質裂縫。

使用資料增強技術（模糊、對比、旋轉等）提升泛化能力。

導入注意力機制或更新骨幹網路(backbone)以強化特徵萃取。

2.根據場景調整 threshold 策略

如果漏檢風險高（如漏判裂縫影響安全），可選較低 confidence 閾值。

如果誤判代價高（如誤認障礙物），則選擇較高閾值保精確率。

搭配其他圖進行多面評估。

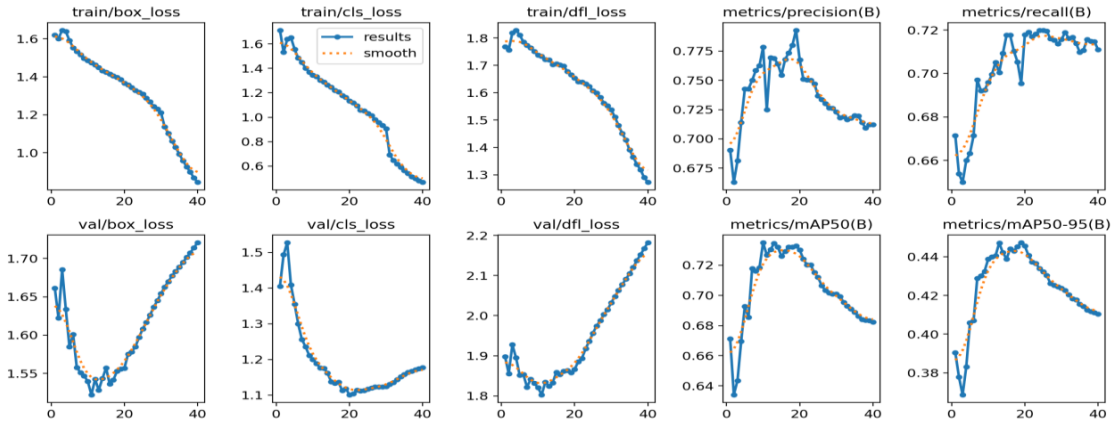


圖 4-9 訓練與驗證集指標視化圖

資料來源:本專題綜合整理

為確保模型學習過程中不僅在訓練集上表現良好，也能有效泛化到未知資料，我們透過下列圖表（共 10 個子圖）監控模型表現，分別顯示:損失函數下降情形（訓練與驗證）、關鍵評估指標變化（Precision、Recall、mAP）。

表 4-8 損失函數觀察（左六張圖）表

圖表名稱	說明
train/box loss	顯示預測邊界框位置的誤差，持續下降，代表模型定位能力提升。
train/cls loss	分類損失穩定下降，說明模型在學習不同類別判斷上表現良好。
train/dfl loss	分布式回歸損失（YOLOv8 特有），也穩定下降，表示模型空間預測能力逐步增強。
val/box loss	驗證集邊界損失略有波動，於約第 20epoch 後趨於平穩，顯示模型泛化尚可。
val/cls loss	初期下降快速，但在約第 20~30epoch 間達到最小後反彈，可能出現過擬合跡象。
val/dfl loss	出現類似情形，下降後略有上升，亦表示模型可能開始過度擬合訓練資料。

資料來源：本研究綜合整理

表 4-9 評估指標變化（右四張圖）表

圖表名稱	解釋
metrics/precision(B)	約第 20epoch 後趨勢反轉，有下降趨勢，表示誤判率可能增加。
metrics/recall(B)	整體穩定略升，代表模型仍有偵測能力，但可能增加召回時造成精度下降。
metrics/mAP50(B)	IoU≥0.5 的 mAP，在第 20~25epoch 附近達高點後開始下滑。
metrics/mAP50-95(B)	mAP@0.5~0.95 全範圍，最高點出現在約 20epoch 後開始下降，泛化能力可能降低。

資料來源：本研究綜合整理

多項評估指標均在第 20epoch 附近達到最佳值，之後開始下降或波動，顯示模型訓練效果進入瓶頸甚至過擬合。

第五章 結論與未來研究方向建議

5.1 結論

本研究針對 YOLOv8 模型在道路影像分析上的應用進行探討，主要針對道路上的破損及道路設計不良----坑洞、水溝蓋與裂縫進行分類與偵測。本研究透過數據收集、圖像標註、數據增強與模型訓練等步驟。本研究成功並建立一套完整能夠準確辨識道路問題的深部學習模型。在研究過程中，模型於測試集上展現出不錯的精確率(Precision)與召回率(recall)，mAP 值也達到了實用水準，顯示 YOLOv8 在此應用中具備高度的可行性。

此外，本研究結果顯示，若能進一步擴充數據集來源並調整優化模型參數，有機會實際運用在城市道路之中，可即時監測與維修系統中，提供道路安全與維護效率。未來可以朝著即時辨識、移動設備佈署與多類別道路異常識別等方向進行發展。

1.效能評估與比較

本研究透過實驗驗證 YOLOv8 模型在不同類型道路缺陷（特別是細小裂縫與遮蔽水溝蓋）上的辨識能力。模型展現出良好的準確率與 mAP 值，與傳統 YOLO 系列進行比較，顯示 YOLOv8 在速度與精度間達到平衡。

2.技術改進與優化

在數據預處理階段將圖像旋轉、亮度調整與鏡像等數據增強技術，提升模型在多樣化影像情境下的泛化能力，並減少過度增強所導致的誤判率。此外，合理調整超參數也有助於優化訓練效率與穩定性。

3.實驗結果與分析

實驗顯示此模型能有效辨識不同類型的道路缺陷，特別是在複雜背景與不規則形狀下仍具辨識能力。在分析中發現，裂縫類因為邊界模糊及長條型態(相似如牆上裂痕)，偵測相對困難；而坑洞因深度問題在陰暗處或夜間難以辨識。

4.應用範圍與局限性

本研究成果可應用於智慧交通、道路巡檢與基礎設施維護系統。但實際部署中，仍然需要考量影像品質不一、車速影響與環境變因（如光線與遮蔽物）所顯示的偵測準確度下降。未來可結合即時辨識提升實用性。

5.2 未來展望

未來的研究可以在以下方面進一步深入探討和改進:

1.精度提升與模型優化

未來可針對目標物體在不同背景與尺寸變化下的表現進行調整，透過引入更細緻的特徵提取策略與改進的錨點生成方式，提升模型對微小或部分遮蔽目標的辨識能力，進一步強化 YOLOv8 在複雜場景下的應用效能。

2.增強數據與增強泛化能力

未來可強化資料多樣性，特別是在不同光照、天氣及路面狀況下的影像樣本，以模擬實際應用情境，提升模型在多變環境中的穩定性與適應力，以強化模型在真實場景中的辨識效能與實用性。

3.及時影片與監控系統整合

未來可將 YOLOv8 模型整合至即時影片串流系統中，透過攝影機或行車記錄器實現連續畫面的目標偵測，達到即時監控的應用需求。結合即時推論技術與警示機制，能即時回報異常路面情況，提升道路維護的效率與反應速度。

參考文獻

[1] 世界衛生組織:道路交通傷害
https://www.who.int/zh/news-room/fact-sheets/detail/road-traffic-injuries

[2] 蘇豐文，111 年 5 月，人工智慧-基本觀念與導論 ，高立圖書有限公司，p1-14

[3] 陳兵旗，2022 年 3 月，機器視覺技術，崧燁文化事業有限公司，p210~231

[4] 張宗彥，2023 年 05 月，一本精通-OpenCV 與 AI 影像辨識，深智數位股份有限公司，p1-1~1-4

[5] 張志勇，廖文華，石貴平，王勝石，游國忠，2021 年 1 月，AI 人工智慧 Artificial Intelligence ，全華圖書股份有限公司，p5-10~5-16

[6] 行人死傷與高齡者統計-公視新聞網（2023）
https://news.pts.org.tw/article/629801

[7] 營業用車、大客車事故與地區差異分析 READr 專題報導
https://www.readr.tw/post/2931

[8] 交通事故總數與死傷統計-公視新聞網（2023）
https://news.pts.org.tw/article/616135

[9] You Only Look Once: Unified,Real-Time Object Detection
https://arxiv.org/pdf/1506.02640

[10] CNN 的結構,包括卷積層,集合層和完全連接的層
https://www.researchgate.net/figure/Structure-of-a-CNN-consisting-of-convolutional-pooling-and-fully-connected-layers-70_fig2_343715313

[11] 卷積神經網路(Convolutional neural network)
https://ithelp.ithome.com.tw/m/articles/10304942

[12] 池化層（Pooling Layer）
https://blog.csdn.net/m0_63007797/article/details/128729049

[13] 全連結層（Fully Connected Layer）
https://blog.csdn.net/qq_52445967/article/details/126364115

[14]YOLO-V1 TO YOLO-V8, THE RISE OF YOLO AND ITS COMPLEMENTARY NATURE TOWARD DIGITAL MANUFACTURING AND INDUSTRIAL DEFECT DETECTION
https://www.mdpi.com/2075-1702/11/7/677
[15] ActNAS : Generating Efficient YOLO Models using Activation NAS
https://arxiv.org/abs/2410.10887
[16] YOLOv4: Optimal Speed and Accuracy of Object Detection
https://arxiv.org/abs/2004.10934
[17] Jon Krohn、Grant Beyleveld、Aglaé Bassens，2022 年 1 月，深度學習的 16 堂課 CNN・RNN・GAN・DQN・DRL 看得懂、學得會、做得出！，黃駿、哈雷，旗標科技股份有限公司，p1-8~1-17
[18] 活化函數改良的 YOLOv8 模型對血球和尿液顆粒檢測的效果
https://www.authorea.com/doi/full/10.22541/au.173173879.92891429/v1
[19] Searching for Activation Functions
https://arxiv.org/abs/1710.05941
[20] 事故傷害原因死亡人數占比
https://www.advisers.com.tw/?p=14924
[21] Faster R-CNN vs YOLO vs SSD — Object Detection Algorithms
https://medium.com/ibm-data-ai/faster-r-cnn-vs-yolo-vs-ssd-object-detection-algorithms-18badb0e02dc
[22] 目標檢測算法-YOLOv8 解析（附論文和源碼）
https://blog.csdn.net/sxlsxl119/article/details/131411336
[23] Ultralytics Docs p14
https://docs.ultralytics.com/models/yolov8/
[24] CSPNet: A New Backbone that can Enhance Learning Capability of CNN
https://arxiv.org/pdf/1911.11929
[25] 卷積神經網路(Convolutional Neural,CNN)
https://hackmd.io/@allen108108/rkn-oVGA4

[26] Gradient-Based Learning Applied to Document Recognition
http://www.dengfanxin.cn/wp-content/uploads/2016/03/1998Lecun.pdf
[27] 機器學習筆記常見的二元分類評估指標-----混淆矩陣、ROC 曲線
https://medium.com/@SCU.Datascientist/python 學習筆記-常見的二元分類評估指標-混淆矩陣-roc-曲線-f214ecd84dab
[28] 湯曉鷗，陳玉琨，2019，人工智慧基礎，五南圖書出版公司，第 68 頁
[29] 鄧文淵，2020 年 9 月，Python 機器學習超進化：AI 影像辨識跨界應用實戰，文淵工作室，p5-2 -5-4
[30] TensorFlow 中圖像識別功能
https://pctechmag.com/2024/11/best-use-cases-for-tensorflow-in-actual-world-applications/
[31] 準確率、召回率、精確度相關指標
https://developers.google.com/machine-learning/crash-course/classific
[32] Precision、Recall 的定義、公式與圖解
ation/precision-and-recall?hl=zh-tw
[33] Precision 與 Recall 的取捨問題，以及 F1 分數的意義
https://towardsdatascience.com/precision-vs-recall-386cf9f89488
[34] 提供了 Precision、Recall 等指標的詳細說明與計算方式
https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall
[35] [機器學習 ML NOTE]Convolution Neural Network 卷積神經網路
https://medium.com/雞雞與兔兔的工程世界/機器學習 -ml-note-convolution-neural-network-卷積神經網路-bfa8566744e9
[36] YOLO 目標偵測詳解：演化、演算法與應用
https://encord.com/blog/yolo-object-detection-guide
[37] 深度學習系列: 什麼是 AP/mAP?
https://chih-sheng-huang821.medium.com/深度學習系列-什麼是 ap-map-aaf089920848

[38] 使用 TensorFlow 學習 Softmax 回歸 (Softmax Regression)
https://medium.com/ 手 寫 筆 記 / 使 用 -tensorflow- 學 習 -softmax- 回 歸 -softmax-regression-41a12b619f04