

利用深度學習模型 Mask R-CNN 辨識鋼結構橋梁 鏽蝕之研究

黎俊成^{1*} 高書屏² 王豐良³ 林志憲³

摘要

近幾年，鋼結構橋梁在臺灣逐漸盛行，但臺灣為容易發生鏽蝕的環境，因此橋梁的鏽蝕檢測必須投入大量成本進行維護。而本研究利用 UAV 結合 Mask R-CNN 進行鏽蝕辨識實驗，而結果表示該方法可以改善傳統橋梁檢測的缺點。而目前臺灣無適合本研究的鏽蝕數據集，故本研究自行建立了鏽蝕數據集；其匯入 Mask R-CNN 進行訓練，並經過本研究的實驗，揀選出最佳的超參數配置，即「優化器 SGD 搭配學習率 1×10^{-3} 」，而模型訓練的評估指標結果為：Recall 可達 97.1%、Precision 可達 90.4%、mAP 可達 91.0% 以及 mIoU 可達 89.0%；再經過辨識結果的分析，發現影像中的背景雜訊會影響鏽蝕的辨識。

關鍵詞：鋼結構橋梁、橋梁檢測、無人機、鏽蝕、Mask R-CNN

1. 前言

臺灣有著多高山及多河川的地形，而為了連結分隔兩地的交通，故需要橋梁的建設來維繫交通。橋梁建設目前在國內除了混凝土橋梁之外，還有鋼結構組立的橋梁，而鋼結構因為有著加工組建快速容易、節省砂石、自重輕盈以及材料強度較高等優點，因此常被作為都市高架橋以及跨河橋等 (李家順及張英發, 2019)，且其的數量在國內也有逐年增長的趨勢。

但由於臺灣地區在氣候以及環境污染的影響之下，鋼結構橋梁容易發生鏽蝕劣化之情形。而「鏽蝕」即為金屬材料的「腐蝕」現象，其中金屬在空氣中受到化學或電化學的作用之下，所產生的產物就稱為「鏽」，而在「鏽」長期的侵蝕下，會導致金屬材料產生退化，大幅降低疲勞強度，並產生疲勞裂縫，此現象會縮短鋼結構橋梁壽命，因此，為了避免鏽長期破壞鋼結構橋梁的結構性能，國家必須花費大量成本及時間進行檢測以及維護鋼結構橋梁 (李家順等, 2013、陳正平, 2019、Zhang *et al.*,

2021)，若維護不當則會發生公共安全之虞，例如 2019 年 10 月發生的宜蘭南方澳大橋斷裂事件，即是因為橋梁構件積水而導致鋼絞線及端錨構件發生鏽蝕斷裂，破壞了橋梁的結構性能 (國家運輸安全調查委員會, 2020)。以此事件為借鏡，在進行鋼結構橋梁檢測時，針對鏽蝕劣化的檢測是非常重要的。

而近年來，無人航空載具 (Unmanned Aerial Vehicle, UAV) 快速發展，UAV 也開始投入橋梁的檢測工作，並且在臺灣 UAV 已是符合規範的檢測工具 (中華民國交通部, 2020)，且對於橋梁檢測，有著降低人力成本、減少檢測時間、提高安全性、提高檢測範圍以及避免交通控管等優點 (Khedmatgozar Dolati *et al.*, 2021、Mandirola *et al.*, 2022)，除此之外，針對檢測人員不易檢測之結構部位，其可藉由拍攝的方式收集影像資料 (Wang *et al.*, 2020)，並利用影像處理和深度學習的方式檢測橋梁的鏽蝕劣化。

隨著電腦視覺技術快速發展，相關的結構劣化檢測開始利用自動的影像處理技術取代傳統的人

¹ 國立中興大學土木工程學系 碩士生

² 國立中興大學土木工程學系 教授

³ 國立中興大學土木工程學系 博士候選人

* 通訊作者, 電話: 04-22840437 轉 306, E-mail: b0976245966@gmail.com

收到日期: 民國 112 年 01 月 30 日

修改日期: 民國 112 年 02 月 23 日

接受日期: 民國 112 年 03 月 28 日

工目視檢測，以鏽蝕檢測為例：Shen *et al.* (2013) 使用傅立葉轉換 (Fourier transform) 技術結合影像處理，利用色彩以及紋理特徵檢測鋼結構橋梁的鏽蝕劣化；Vorobel *et al.* (2021) 利用色彩空間技術，將影像的色彩空間轉換為 HSV 色彩空間，後續再利用影像處理方法分割塗料上的鏽蝕顏色特徵，以達到鏽蝕檢測的任務。

而深度學習在近幾年亦開始投入結構損傷檢測的研究，其結合了影像處理的技術，以訓練方式建立損傷辨識的框架，可以大量且批次的模式進行辨識作業，與傳統的影像處理辨識法相比，深度學習可以降低人工作業，以及減少辨識時間，有著節省時間以及人力成本的優勢。而深度學習在鏽蝕檢測方面，已經有相關學者利用全卷積神經網路 (Fully Convolutional Networks, FCN) 以及卷積神經網路 (Convolutional Neural Networks, CNN) 進行金屬結構的鏽蝕辨識，使辨識偵測達到自動化，提升了工作的效率。其中 CNN 模型在鏽蝕辨識上，因為其受到影像雜訊的影響比較小，因此有比較良好的適用性 (Munawar *et al.*, 2022)，而 Katsamenis *et al.* (2020) 的實驗比較了 FCN、U-Net 以及 Mask R-CNN 三套模型對於鏽蝕辨識的性能，最後結果為 Mask R-CNN 精度最高。

而在 CNN 系列中，對於鏽蝕辨識相關的研究：Pirie & Moreno-García (2021) 比較三套深度學習模型 (CNN、YOLO、Mask R-CNN) 辨識水中物體鏽蝕的性能，在最後其研究結果發現 Mask R-CNN 模型的表現最佳，mAP 可以達到 77.1%，而 YOLO 的 mAP 僅達到 57%，而此文獻也提及 CNN 需過於仰賴人工干預來達成預判辨識目標位置；而 YOLO 因為鏽蝕對於其而言不為一種物體，而是物體上的特徵，並且 YOLO 亦不擅長辨識較小或過於近距離的目標，故 YOLO 在辨識鏽蝕上相較不適用。除了以上 Mask R-CNN 對於鏽蝕辨識的優點，其目前已為發展成熟的網路模型，再加上其對於結構的鏽蝕劣化有著優良的檢測以及分割效果 (Wu & Kalfarisi, 2019)，故 Mask R-CNN 在鏽蝕辨識上有一定的優勢。而其相關的文獻還包含：Fondevik *et al.* (2020)

利用 Mask R-CNN 模型訓練 UAV 收集的影像，並透過兩個階段的數據增強，避免模型訓練出現過度擬合之情況；而其最後訓練結果顯示 Mask R-CNN 模型的平均交併比 (Mean Intersection over Union, mIoU) 可達到 77.5%。Xi *et al.* (2020) 開發了一套深度的 Mask R-CNN 模型，並將其應用於齒輪點蝕辨識和量測，而該方法可有效的分割影像中齒輪點蝕，其分割影像的平均精度可達 88.2%。Ta *et al.* (2022) 利用 Mask R-CNN 模型辨識鋼結構的鏽蝕螺栓，除此之外，研究中發現影像的拍攝距離和角度皆會影響模型精度，其中距離目標 1 m 以及拍攝視角 15° 可以達到最佳的總體準確率 97.5%。

針對臺灣目前在鋼結構橋梁的鏽蝕檢測是用傳統的目視方式進行檢查，而此檢測法不僅會有人力成本過高的情形，還會有檢測範圍受限以及檢測效率不佳的問題，甚至有檢測人員安全的疑慮，儘管可以搭載橋檢車進行檢測，但檢查範圍仍是不足，例如橋梁拱圈高處或是橋底板皆為橋檢人員不容易檢測之結構部位 (圖 1)。



圖 1 松鶴橋，鋼結構橋梁示意圖

根據以上之因素，本研究將利用 UAV 結合 Mask R-CNN 模型進行實驗，一方面藉由高機動性 UAV 替代傳統的人為目視檢測法，增加檢測的範圍，另一方面 UAV 結合深度學習模型使鏽蝕辨識的流程可以自動化，大幅減少了橋梁檢測的人力成本。

2. 研究方法

2.1 研究流程

本研究詳細的研究流程可詳圖 2。使用 UAV 進行鋼結構橋梁鏽蝕的拍攝，後續再利用 Mask R-CNN 進行鏽蝕辨識，而由於國內目前尚無適合用於本研究的鏽蝕數據集，故本研究自行建立了鏽蝕數據集，並匯入 Mask R-CNN 進行訓練，再經過超參數的調整配置，以及分析影響鏽蝕辨識的因素，評估 Mask R-CNN 進行鏽蝕辨識的效果，以及評估本研究自行建立的數據集品質。

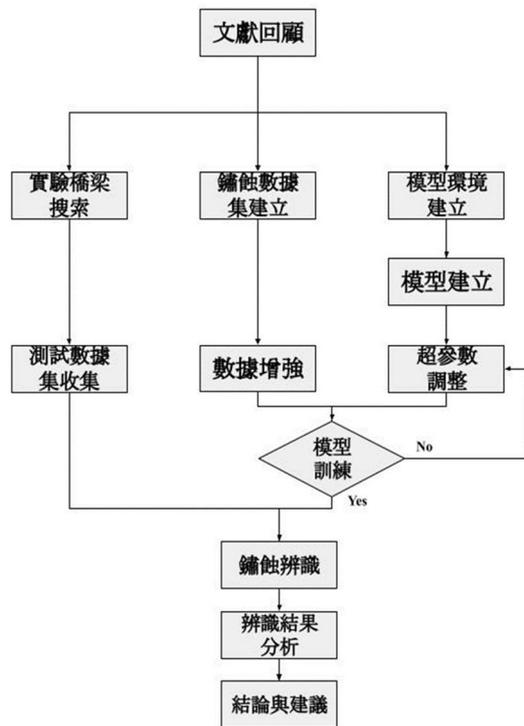


圖 2 本研究之流程圖

2.2 研究資料收集與處理

2.2.1 鏽蝕數據集建立

由於目前國內尚無適合用於本研究的鏽蝕數據集，故本研究自行以手機 ASUS Zenfone 8 進行拍攝收集，而拍攝目標皆為日常生活中的金屬鏽蝕，例如金屬欄杆、鋼柱、金屬坐椅、汽車鋼板以及鐵橋構件等，而每張影像的尺寸為 4576×3432，後續再以 512×512 的尺寸將影像進行裁切，並以人工揀

選出含有鏽蝕的影像，而本數據集的鏽蝕影像可詳圖 3，其中有包含了不同條件的資料，例如：不同顏色的鏽蝕(a)、不同塗料顏色為背景的鏽蝕影像(b)，以及富含雜訊背景的鏽蝕影像(c)。

為了增加模型訓練的數據數量，本研究藉由影像翻轉和調整影像明亮度的方式進行數據增強(Data Augmentation)，最後將模型訓練的鏽蝕數據集增強為 4,068 筆影像數據。

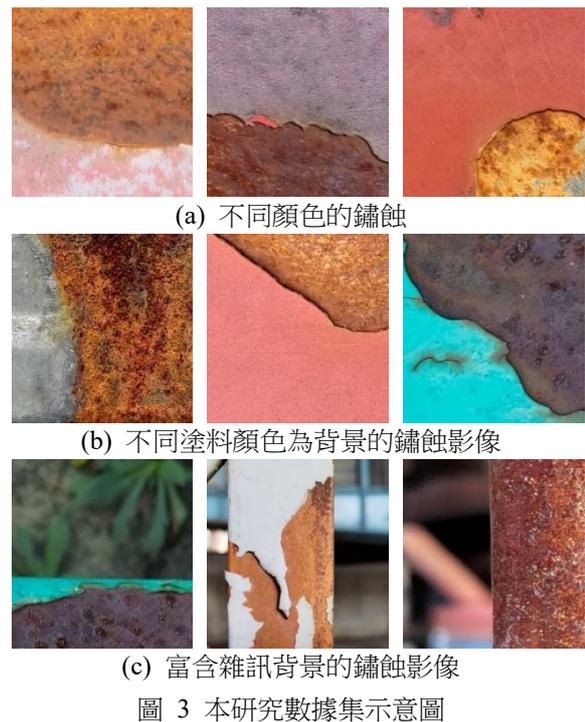


圖 3 本研究數據集示意圖

2.2.2 測試數據

本研究測試影像由無人機 DJI Mavic 2 Pro 搭載攝影機 L1D-20c 進行拍攝，而拍攝的實驗橋梁選定位於南投市的綠美橋(圖 4)，以及位於彰化縣福興鄉的福興橋(圖 5)，前者於西元 2022 年 10 月 21 日星期五大約早上 10:00 至 11:00 以 UAV 進行拍攝，後者則是於西元 2022 年 11 月 17 日星期四大約早上 10:00 至 11:00 進行拍攝，而拍攝的影像尺寸皆為 5472×3648，拍攝目標主要鎖定在橋梁拱圈高處的鏽蝕劣化；除了使用原始的影像尺寸，本研究還另外以正方形的形狀(與訓練數據相似)進行影像裁切，再進行辨識，以此比較影像裁切前後的辨識效果。



圖 4 綠美橋現地示意圖



圖 5 福興橋現地示意圖

2.2.3 影像資料標註

本研究的影像標註是利用開源的標註工具—LabelMe (Russell *et al.*, 2008)，其操作是以人工方式判釋並標註，而標註結果經過一系列的轉檔作業之後，會產生三個檔案，最後再將檔案放置於相對應的資料夾以後，即可藉由模型讀取資料夾的路徑進行訓練。而經過 LabelMe 標註之後產生的遮罩示意圖如圖 6 所示。

2.3 本研究 Mask R-CNN 架構

Mask R-CNN (Mask Region-based Convolutional Neural Network) 模型為 Faster R-CNN 模型延伸而出，由 He *et al.* (2017) 文獻中的 Mask R-CNN 架構示意圖 (圖 7) 所示，其配置了 ROIAlign，解決了 ROI Pooling 像素偏差的缺點，令所有 ROI 的尺寸相同，並保留了確切的空間位置，使辨識精確率提升；另外，模型中為每個 ROI 設置微型的全卷

積層 (FC Layer)，以像素到像素的方式進行預測，並使每個預測為獨立發生的事件，達到實例分割 (Instance Segmentation) 的效果。

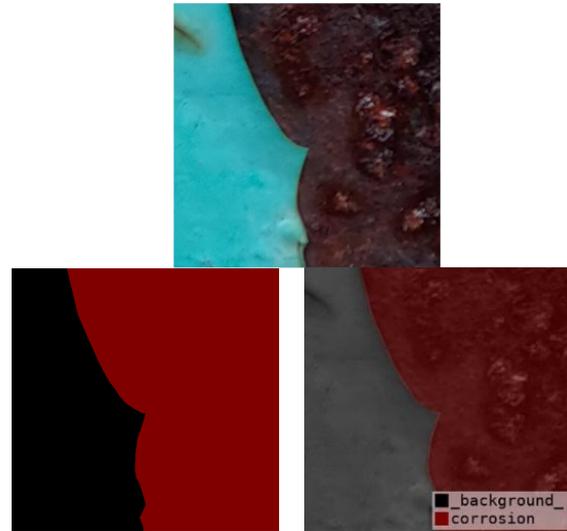
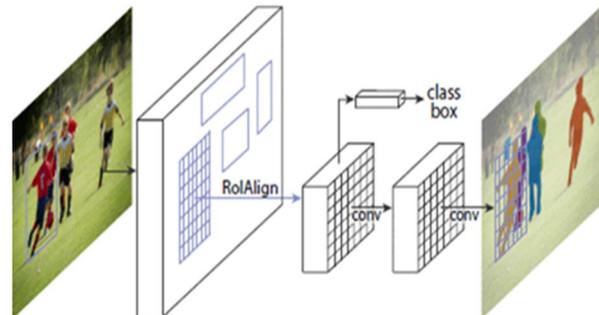


圖 6 LabelMe 標註的遮罩示意圖

圖 7 Mask R-CNN 架構示意圖 (He *et al.*, 2017)

而本研究所使用的 Mask R-CNN 的詳細流程可以詳見圖 8，其流程大致如下：

- (1) 首先，利用殘差網路 (ResNet) 模型結合特徵金字塔網路 (FPN) 萃取影像特徵，並產製出特徵圖。
- (2) 將產製出的特徵圖輸入候選區域網路 (RPN)，生成感興趣區域 (ROI)。
- (3) 利用上一階段生成的 ROI 經過 ROIAlign 處理，使每個 ROI 的尺寸相同，並將原始影像和特徵圖的像素相互對應，使預測結果更為準確。
- (4) 最後利用全卷積神經網路 (FCN) 對調整完成的 ROI 生成遮罩 (Mask)，以及利用微型的全卷積層 (FC Layer) 分別生成辨識目標的類別 (Class) 和辨識框 (Bounding Box, BBox)，以完成 Mask R-CNN 的實例分割工作。

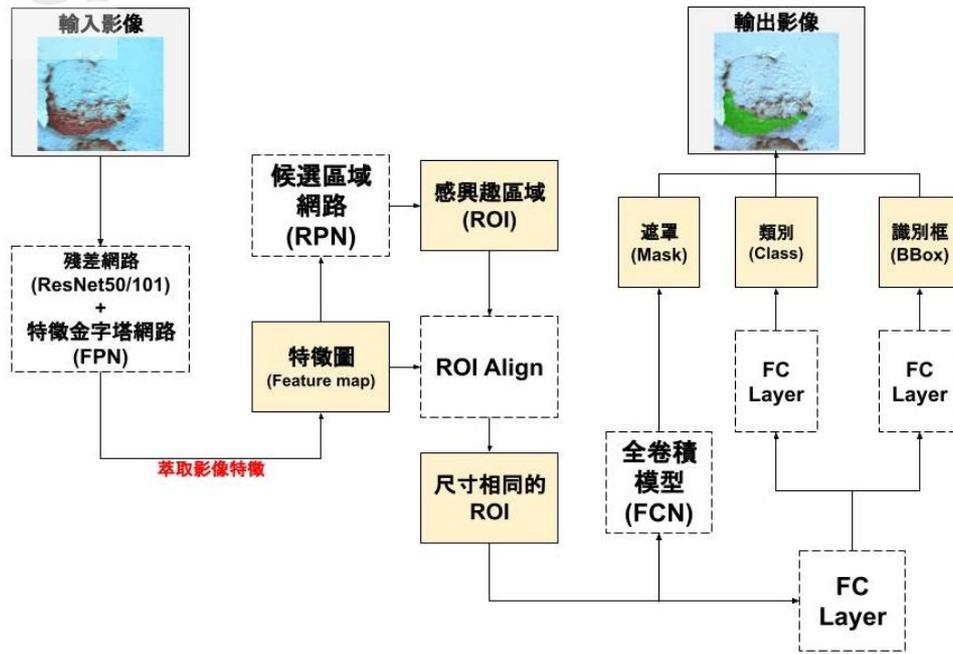


圖 8 本研究 Mask R-CNN 運作流程圖

2.4 Mask R-CNN 損失函數

Mask R-CNN 的損失函數為三種不同的損失函數的總和，其計算公式如式 (1) 所示：

$$L_{Mask\ R-CNN} = L_{cls} + L_{box} + L_{mask} \dots \dots \dots (1)$$

其中， L_{cls} 為分類上的損失分數， L_{box} 為識別框的損失函數， L_{mask} 為遮罩的損失分數。 L_{cls} 是由 RPN 中的邊界框分類層(box-classification layer,cls)計算而得，其使用了二分類的交叉熵，而詳細的計算公式如式 (2)：

$$L_{cls} = -\log [p_i p_i^* + (1 - p_i)(1 + p_i^*)] \dots \dots \dots (2)$$

其中， p_i 為錨點的預測機率， p_i^* 為錨點被分類的結果。 L_{box} 則是代表辨識框與真實情況之間的偏移，計算公式如式(3)，辨識框偏移量經過 Smooth_{L1} Function 計算後即是此損失分數：

$$L_{box} = p_i^* L_{smo}(t_i, t_i^*) \dots \dots \dots (3)$$

其中， t_i 為預測錨點的辨識框之參數化坐標， t_i^* 為真實情況的參數化坐標， L_{smo} 為 Smooth_{L1} Function，其公式如式(4)：

$$L_{smo}(t_i, t_i^*) = Smooth_{L1}(t_i - t_i^*) \dots \dots \dots (4)$$

其中式(4) 中的 Smooth_{L1}：

$$smooth_{L1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \dots \dots \dots (5)$$

此外，雖然 L_{mask} 為 FCN 計算而出，但有別於一般 FCN 常用的 Softmax 函數搭配交叉熵 (Cross Entropy) 的組合，Mask R-CNN 則是改用 Sigmoid 函數搭配平均二元交叉熵 (Average Binary Cross Entropy)，此方式可以使每個類別的損失分數獨立計算，即類別預測不互相干擾 (Ren *et al.*, 2015、He *et al.*, 2017)。

2.5 激勵函數

Mask R-CNN 的激勵函數使用了兩套不同的演算法，在分類上，使用了歸一化指數函數 (Softmax Function)，其可以將輸入值限制在 0~1 之間，而其輸出結果為模型的預測值的機率，且輸出值的總和為 1。其演算公式如式(6)：

$$y_i = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}} \dots \dots \dots (6)$$

其中， C 為類別的總數， y 為預期輸出值， z 為模型訓練的預測值， i 代表了模型訓練第 i 個輸出。而遮罩分支則是使用了 Sigmoid 函數，其公式可詳式(7)

，Sigmoid 函數會將輸出數值限制在 0~1 之間，且 0 和 1 分別代表了兩個類別，適合應用在二元分類的網路模型。

$$f(x) = \frac{1}{1+e^{-x}} \dots\dots\dots (7)$$

2.6 優化器

優化器 (Optimizer) 在深度學習中，利用調整權重的方式降低損失分數，而本研究選擇常見的兩套優化器進行實驗，其一為隨機梯度下降法 (Stochastic Gradient Descent, SGD)，以一個批次 (Batch) 為單位進行計算梯度值，並重新更新模型權重，其公式如式 (8)。

$$W = W - \eta \frac{\partial L}{\partial W} \dots\dots\dots (8)$$

其中，W 為權重，L 為損失分數，η 為學習率。

而另外一套選擇目前深度學習領域流行的 Adam 優化器，其會以自適應的方式調整學習率，並且計算效率較快以及對於記憶體容量要求較低，而其計算過程可詳式(9) 至式 (13)。

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \frac{\partial L_t}{\partial w_t} \dots\dots\dots (9)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \left(\frac{\partial L_t}{\partial w_t}\right)^2 \dots\dots\dots (10)$$

m_t 為梯度的一階動差函數， v_t 二階動差函數， β_1 為一階動差函數的指數衰減率， β_2 為二階動差函數的指數衰減率，通常 β_1 和 β_2 會分別設置為 0.9 和 0.999。

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \dots\dots\dots (11)$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \dots\dots\dots (12)$$

\hat{m}_t 為偏離校正後梯度的一階動差函數、 \hat{v}_t 為偏離校正後二階動差函數。

$$W \leftarrow W - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \dots\dots\dots (13)$$

ε 為防止分母為零的參數，通常會設置相當微小的數值。

2.7 模型性能評估指標

評估模型的指標利用了混淆矩陣 (Confusion Matrix) (圖 9)，其依據真實情形和預測結果進行分類，而矩陣內部分別為真陽性 (True Positive, TP)、偽陽性 (False Positive, FP)、真陰性 (True Negative, TN) 以及偽陰性 (False Negative, FN)；在本研究中，TP 代表「預測的結果與實際情形皆是鏽蝕」，FP 為「預測的結果為鏽蝕，但實際上卻不是鏽蝕」，TN 為「預測的結果與實際情形皆不為鏽蝕」，FN 為「預測的結果不為鏽蝕，但實際上卻是鏽蝕」。而混淆矩陣可以計算模型的評估指標：精確率 (Precision)、召回率 (Recall)。

	真實為陽性	真實為陰性
預測為陽性	TP (True Positive)	FP (False Positive)
預測為陰性	FN (False Negative)	TN (True Negative)

圖 9 混淆矩陣示意圖

精確率為「TP」與「TP+FP」之間的比率關係，即是確認「被預測的資料中，正確的資料所佔的比例」，其公式如式 (14)：

$$Precision = \frac{TP}{(TP+FP)} \dots\dots\dots (14)$$

召回率又被稱為查全率，其為「TP」與「TP+FN」之間的比率關係，即是確認「真實為正確的資料中，被預測為正確的資料所佔的比例」，其公式如式 (15)：

$$Recall = \frac{TP}{(TP+FN)} \dots\dots\dots (15)$$

交並比 (Intersection over Union, IoU) 為評估語意分割的指標，其計算了「預測結果」與「真實情況」兩集合之間的交集與聯集之間的比例關係，詳細計算概念如圖 10，而 IoU 的數值越高，代表預測的結果與真實狀況越接近。至於 mIoU 即為每一種類別的 IoU 總和的平均，其計算公式用混淆矩陣表示如式 (16)：

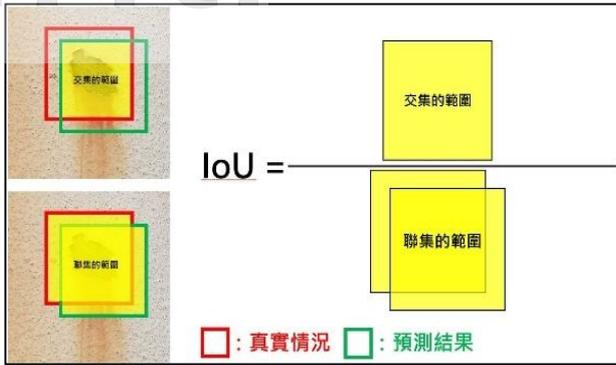


圖 10 IoU 計算示意圖

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN+FP+TP} \dots\dots\dots (16)$$

其中， k 為類別的數量。

Mean Average Precision (mAP)為 AP (Average Precision) 之平均，而 AP 為精確率與召回率所繪製的 Precision-Recall 曲線構成的面積，計算公式如式 (17)：

$$AP = \int_0^1 p(r)dr \dots\dots\dots (17)$$

其公式代表著變量 r 沿著精確率 (Precision) 曲線進行積分。

而 mAP 是將所有類別所計算的 AP 進行總和再加以平均而出，其計算公式如式 (18)：

$$mAP = \frac{1}{k} \sum_i^k AP_i \dots\dots\dots (18)$$

其中， k 為類別的數量。

2.8 模型訓練及實驗設計

2.8.1 遷移學習

本研究使用的 Mask R-CNN 由函式庫 Tensorflow (2 版) 構建而成，其中為了提升模型訓練的效率，本研究利用遷移學習 (Transfer Learning) 的方式，選擇由 Microsoft 發布的「COCO」預訓練權重 (Lin et al., 2014) 進行遷移學習，而選擇原因即是因為 COCO 數據集訓練的預訓練模型有著豐沛且詳細的標註資訊，相當適合應用於物件偵測相關的模型訓練進行遷移學習。

2.8.2 訓練數據集分割

本研究在模型訓練時，將訓練數據集通常會拆分為訓練集 (Training Set)、驗證集 (Validation Set)，而測試集 (Testing Set) 為 UAV 所拍攝的實驗橋鏽蝕影像，不納入訓練集進行訓練；而本研究是利用 K 折驗證法 (K-Fold Validation) 進行驗證，其會經過多次的驗證再判斷模型訓練的成效，以避免訓練結果因為過度仰賴特定的資料拆分驗證結果，而導致的偏差。其會依照 K 值的設置，將數據集分成 K 個部分，且每次提取一部份作為驗證集，其餘作為訓練集，待模型完成 K 次的訓練之後，會依據 K 的次數計算出平均的模型分數。K 折驗證法的示例圖可見圖 11，而示例圖以 K = 3 為例。

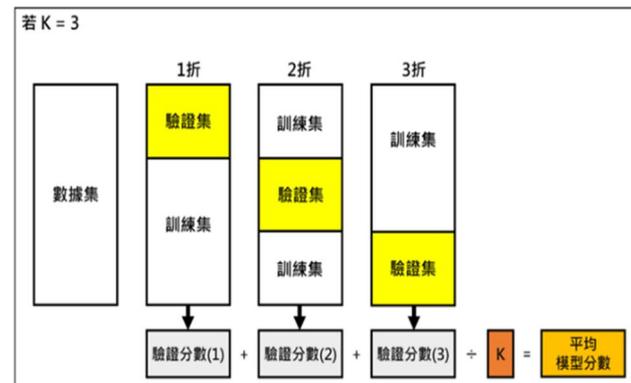


圖 11 K 折驗證法示例圖

2.8.3 模型超參數調整實驗

本研究第一個設計的實驗為「模型超參數調整實驗」，其目的是為了找到適合用於本研究建置的鏽蝕數據集訓練的 Mask R-CNN 超參數，分別為優化器、學習率以及訓練週期配置，即表 1 中紅色框選部分，而此實驗步驟如圖 12 所示，首先會分別利用兩套優化器進行訓練，並設置三種不同的學習率，而選定這三個學習率的原因，即是 1×10^{-2} 的設置在本研究先前的實驗中已出現了梯度爆炸的現象，而 1×10^{-6} 的設置的成果與 1×10^{-5} 的成果大致相同，但訓練時間相當耗時，故本研究選定學習率 1×10^{-3} 至 1×10^{-5} 的區間進行研究；後續會經過暫時設定的 150 次訓練週期，最後會產製 6 套損失分數結果，以此結果可以分析哪一套優化器配合多少學習率為最佳的組合，此外，該結果結合 K 折驗證法，可以分析出適當的訓練週期配置。

表 1 模型訓練超參數調整實驗表

超參數名稱	設置值
優化器	SGD/Adam
學習率	$1 \times 10^{-3} / 1 \times 10^{-4} / 1 \times 10^{-5}$
LEARNING MOMENTUM	0.9
WEIGHT DECAY	0.0001
訓練週期	待實驗結果
STEPS PER EPOCH	4068
VALIDATION STEPS	100

2.8.4 測試影像數據實驗

而本研究第二個設計的實驗為「測試影像數據實驗」，此部分將會進行不同條件的測試影像在鏽蝕辨識上的成果，其條件包含不同影像尺寸 (512×512、1024×1024、2048×2048)、背景雜訊的多寡 (本研究的背景雜訊，即為影像中鏽蝕以外的背景資訊) 以及影像明亮度等不同條件。此實驗會展示不同條件下的辨識結果，並分析影響鏽蝕辨識的原因。

3. 研究成果與分析

3.1 超參數調整實驗成果與分析

本章節將分析模型超參數調整實驗的結果，其結果可詳圖 13，首先觀察優化器 Adam 搭配學習率 1×10^{-3} (紅色曲線)，此僅僅經過兩次訓練週期就發生了梯度爆炸的情形，代表本模型在訓練此數據集時，對於優化器 Adam 而言，學習率 1×10^{-3} 的配置為過大的數值，此會使模型的訓練無法正常收斂。

除了上述組合有梯度爆炸的問題，其餘搭配的實驗結果皆為正常，但收斂的程度不盡相同，從圖 13 中可以發現優化器 SGD 搭配學習率 1×10^{-5} (綠色曲線) 以及優化器 SGD 搭配學習率 1×10^{-4} (藍色曲

線) 的收斂結果比其它結果略差一點，前者最佳的損失分數大約為 0.07，後者則是大約為 0.04。至於優化器 Adam 搭配學習率 1×10^{-4} (粉紅色曲線)、優化器 Adam 搭配學習率 1×10^{-5} (灰色曲線) 以及優化器 SGD 搭配學習率 1×10^{-3} (黃色曲線) 之損失分數最好都可以達到大約 0.02，且從各個訓練及驗證損失函數圖可觀察到三種組合皆達到完美擬合的狀態，而其中 SGD 與學習率 1×10^{-3} 的組合在第 93 次訓練週期達到大約 0.02 的損失分數，而另外兩種組合在第 93 次訓練週期時僅能達到大約 0.03，而根據以上分析，SGD 與學習率 1×10^{-3} 的組合的收斂速度是最快的。而根據以上結果，將 SGD 與學習率 1×10^{-3} 的組合設定訓練週期 93 次，並利用 K 折驗證法進行驗證，而其結果將 4 折的驗證損失分數取平均並繪製成曲線圖，如圖 14 所示，發現模型經過訓練之後，分數有穩定收斂，並在 93 次訓練週期內皆未發生發生分數上升的情況，代表未發生過度擬合。

3.2 測試影像數據實驗成果及分析

3.2.1 模型之評估指標成果

模型經過訓練之後，其訓練成果會以召回率 (Recall) (圖 15)、精確率 (Precision) (圖 16)、mAP (圖 17) 以及 mIoU (圖 18) 作為指標進行評估，模型訓練之後產生的權重 (weight)，會經過評比選擇最佳的訓練結果，以提供模型進行測試數據集的辨識作業，而本次模型實驗比較了 93 次的訓練結果，最後選擇了第 87 次訓練週期的訓練結果，其召回率 (Recall) 可達 97.1%、精確率 (Precision) 可達 90.4%、mAP 可達 91.0% 以及 mIoU 可達 89.0%，而其他的訓練分數較為平均，並且普遍偏高。

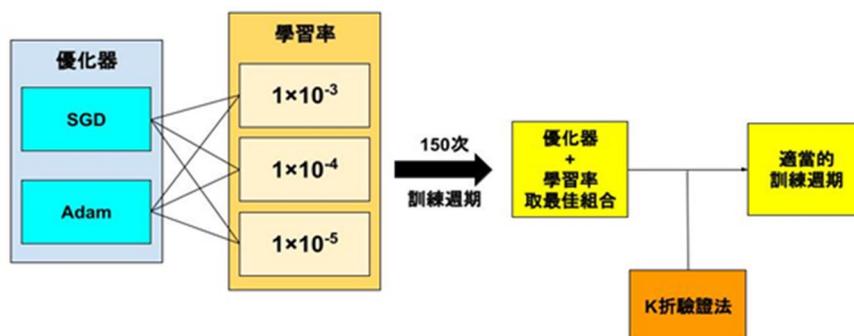


圖 12 模型超參數調整實驗步驟示意圖

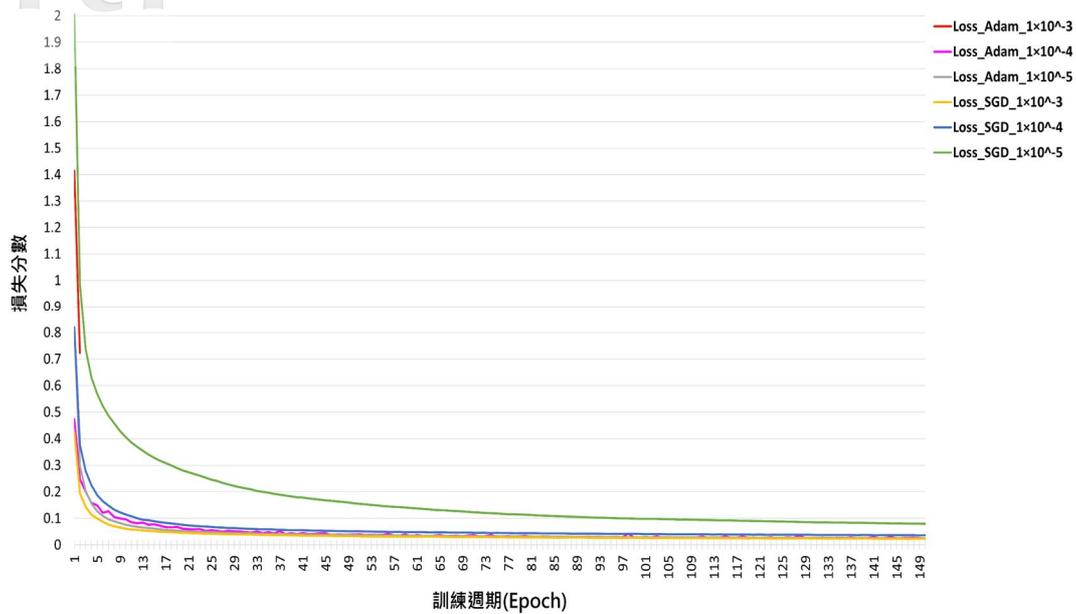


圖 13 優化器及學習率配置損失函數比較圖

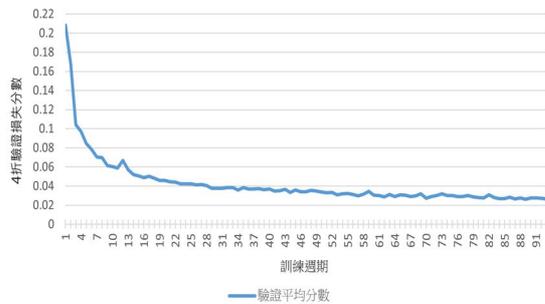


圖 14 K 折驗證平均損失分數曲線圖

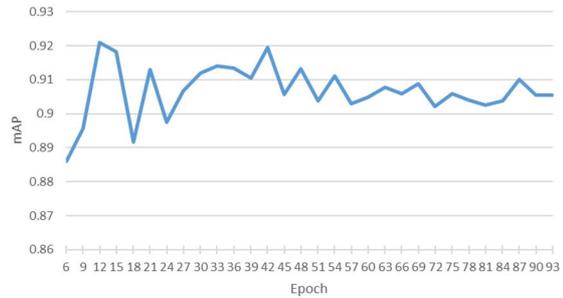


圖 17 mAP 分數曲線圖

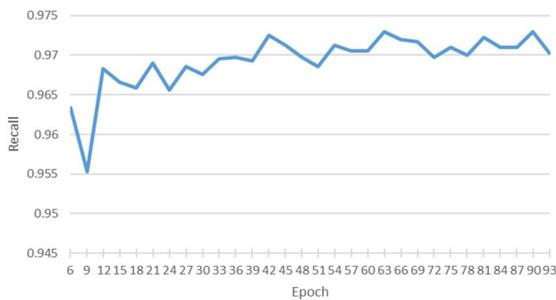


圖 15 召回率分數曲線圖

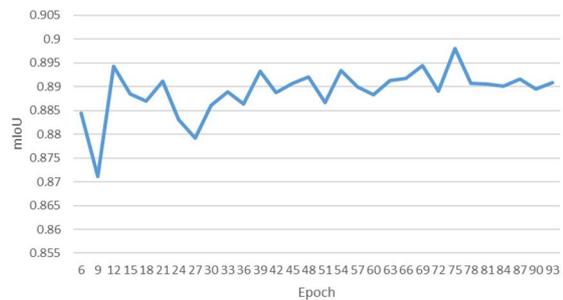


圖 18 mIoU 分數曲線圖

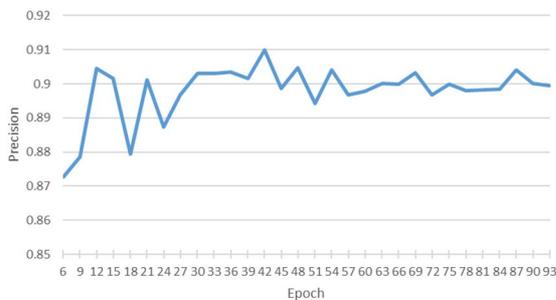


圖 16 精確率分數曲線圖

3.2.2 結果分析—不同測試影像尺寸

此實驗的結果可詳圖 19，影像尺寸包含由 UAV 拍攝的原始影像 (5472x3648) 圖 19 (a)，以及與訓練影像相同尺寸的影像 (512x512) 圖 19 (b) 和其倍數的影像尺寸，分別為 1024x1024 圖 19 (c) 的測試影像和 2048x2048 圖 19 (d) 的測試影像；從實驗的結果可以觀察到原始影像在鏽蝕辨識上並無法穩定被辨識，而經過實驗之後，發現三個種類尺寸的辨識效果皆明顯上升。但若是裁切影像中包含了

複雜的背景雜訊，或是鏽蝕範圍較小等情形，則模型就無法穩定辨識鏽蝕；反之，若在影像裁切時，特別針對鏽蝕區域進行裁切，而此結果顯示如圖 20 所示，無論影像尺寸為何，鏽蝕皆可被辨識。

根據以上實驗結果所示，「測試影像的尺寸」不為影響鏽蝕辨識的條件，反而「測試影像的背景雜訊」對於鏽蝕辨識有一定程度的影響。

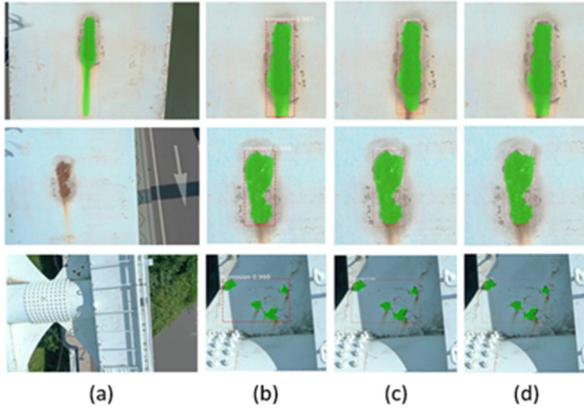


圖 19 不同測試影像尺寸之鏽蝕辨識結果圖

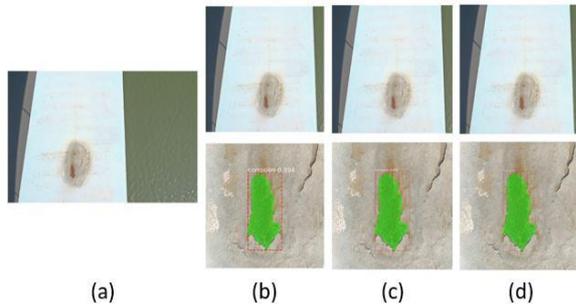


圖 20 不同裁切範圍之鏽蝕辨識結果比較圖

3.2.3 結果分析—測試影像背景雜訊

本實驗為分析「測試影像的背景雜訊」對於鏽蝕辨識的影響，其辨識結果可詳圖 21，其中綠框部分即是有成功辨識鏽蝕的成果，但由於鏽蝕具有較複雜的幾何形狀和顏色特徵，導致語意分割部分無法完美符合真實的鏽蝕情況。藍框部分則是經過本研究訓練的模型成功分辨紅色底漆和鏽蝕，並無誤判之情形。紅框部分為沒有辨識成功的測試影像，而其影像的背景雜訊與綠框和藍框的影像相比，確實更為複雜。

經過本實驗得出：測試影像會因為 UAV 在拍攝辨識目標時所造成的背景雜訊而影響鏽蝕辨識，導致誤判或無法辨識的情形，儘管為相同的結構部

位，但不同拍攝角度和距離皆會產生不同的辨識結果。

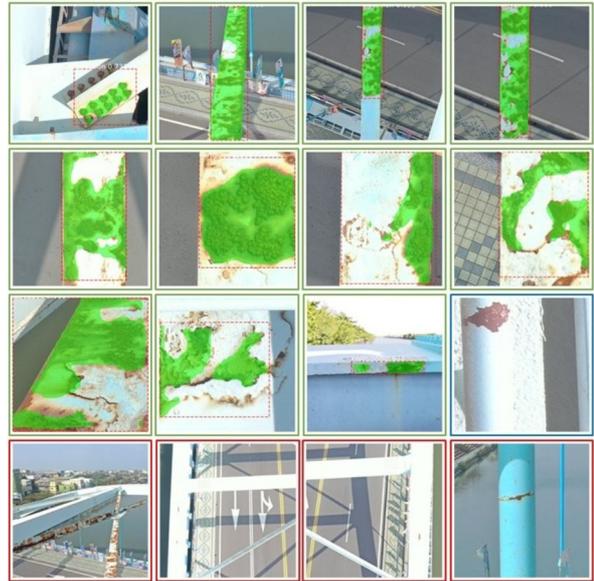


圖 21 背景雜訊對於鏽蝕辨識之實驗成果圖

3.2.4 結果分析—測試影像明亮度

本實驗為分析「不同的影像明亮度」對於鏽蝕辨識的影響，本實驗將測試影像的明亮度分別調整為「+25」、「+50」、「-25」以及「-50」四種不同的灰階值，再使用模型進行鏽蝕的辨識，而其實驗的辨識結果可詳圖 22。

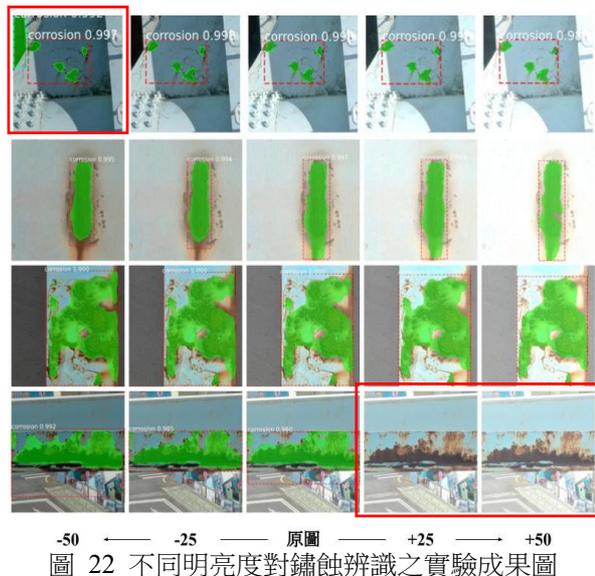


圖 22 不同明亮度對鏽蝕辨識之實驗成果圖

從鏽蝕辨識結果可以觀察到在亮度的調整之下，「背景雜訊較複雜」的影像比較有顯著的影響，在紅框標示的部分，為經過亮度調整之後產生了辨

識錯誤以及辨識遺漏的狀況；而「背景雜訊較單純」的測試影像受到明亮度調整的影響較小，大致上的鏽蝕皆可成功被辨識。

4. 結論與建議

4.1 結論

本研究利用 UAV 結合 Mask R-CNN 進行鋼結構橋梁的鏽蝕辨識，透過本次的實驗，發現 UAV 可以輕易的拍攝橋梁高處的構件，並且可以針對容易積水的結構部位進行檢測，解決直接目視檢測的限制，提高了檢測範圍，此外，本實驗在以 UAV 收集橋梁鏽蝕影像時，僅出動了兩人進行作業，故此檢測方法可以較少的人力成本即可完成檢測作業。

而 Mask R-CNN 部分，藉由本研究自行建立了鏽蝕數據集進行訓練，而數據集總共 4,068 筆鏽蝕影像數據，雖然相較於其他數據集數量較少，但經過本研究的模型超參數調整實驗，得出「優化器 SGD 搭配學習率 1×10^{-3} 」的搭配最適合，且模型訓練的評估指標的 Recall 可達 97.1%、Precision 可達 90.4%、mAP 可達 91.0% 以及 mIoU 可達 89.0%，而此結果顯示：Mask R-CNN 經過本研究建立的數據集訓練之後，已經有相當高的辨識率，並且可以有有效的辨識 UAV 影像的鏽蝕劣化；另外，本研究的訓練數據和測試數據為不同的拍攝工具，而經過一系列的實驗與觀察評估指標的分數，發現此方法有一定的可行性，對於鏽蝕辨識無特別的影響。

至於測試影像對於鏽蝕辨識的影響實驗，可以發現影像尺寸對於辨識結果無特別影響，而背景雜訊對於鏽蝕辨識具有最大的影響力，導致在針對相同的構件時，不同的拍攝距離和角度所產生不同的背景雜訊，會導致不同的辨識結果。

最後，本研究是聚焦在 Mask R-CNN 「鏽蝕辨識」的能力，而結果顯示其在鏽蝕辨識上有尚佳的應用性。

4.2 建議

(1) 本研究建立的鏽蝕數據集匯入 Mask R-CNN 訓練後，已經有相當高的辨識率，而未來的相關

研究仍可持續增加鏽蝕影像數據，持續強化模型判斷鏽蝕的能力。

- (2) 本研究發現影像的背景雜訊會影響鏽蝕辨識的結果，故以 UAV 拍攝測試影像時，要盡量減少背景雜訊的產生，謹慎控制 UAV 的拍攝距離和角度。
- (3) 本研究發現影像的背景雜訊會影響鏽蝕辨識的結果，故以 UAV 拍攝測試影像時，要盡量減少背景雜訊的產生，謹慎控制 UAV 的拍攝距離和角度。
- (4) 本研究發現影像的背景雜訊會影響鏽蝕辨識的結果，故以 UAV 拍攝測試影像時，要盡量減少背景雜訊的產生，謹慎控制 UAV 的拍攝距離和角度。

致謝

本研究感謝國科會計畫編號：NSTC111-2622-M005-001 之經費支持。另外也感謝交通部公路總局第二區養護工程處提供的協助及建議。

參考文獻

- 中華民國交通部，2020。公路橋梁檢測及補強規範，部頒規範。[Ministry of Transportation and Communications, R.O.C., 2020. Specification for inspection and reinforcement of highway bridges, ministerial norm. (in Chinese)]
- 李家順、延允中、孟伯鈞，2013。鋼橋塗裝常見缺失及後續改善對策-以 102 年考評縣府鋼橋為例，臺灣公路工程，39(12)：2-31。[Li, J.S., Yan, Y.C., and Meng, B.J., 2013. Common defects in steel bridge coating and subsequent improvement measures - Take the steel bridge of the county government evaluated in 102 as an example, Taiwan Highway Engineering, 39(12): 2-31. (in Chinese)]
- 李家順、張英發，2019。鋼橋螺栓接合設計與施工探討，臺灣公路工程，45(6)：2-42。[Li, J.S., and Zhang, Y.F., 2019. Discussion on design and

- construction of bolted joint of steel bridge, Taiwan Highway Engineering, 45(6): 2-42. (in Chinese)]
- 國家運輸安全調查委員會，2020。重大運輸事故調查報告第一冊。[Taiwan Transportation Safety Board, 2020. Major Transportation Accident Investigation Report Volume I. (in Chinese)]
- 陳正平，2019。從懸索鋼橋倒塌談鋼橋防蝕的重要，技師期刊，87：73-80。[Chen, Z.P., 2019. Discussion on the importance of corrosion prevention of steel bridges from the collapse of suspension steel bridges, Professional Engineer Journal, 87: 73-80. (in Chinese)]
- Fondevik, S.K. Stahl, A., Transeth, A.A., and Knudsen, O.O., 2020. Image segmentation of corrosion damages in industrial inspections, Proceedings of the IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), Baltimore, MD, USA, pp.787-792, DOI: 10.1109/ICTAI50040.2020.00125.
- He, K., Gkioxari, G., Dollár, P., and Girshick, R.B., 2017. Mask R-CNN, Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp.2980-2988.
- Katsamenis, I., Protopapadakis, E., Doulamis, A., Doulamis, N., and Voulodimos, A., 2020. Pixel-level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, DOI: 10.1007/978-3-030-64556-4_13.
- Khedmatgozar Dolati, S.S., Caluk, N., Mehrabi, A., and Khedmatgozar Dolati, S.S., 2021. Non-destructive testing applications for steel bridges, Applied Sciences, 11(20): 9757, DOI: 10.3390/app11209757.
- Lin, T., Maire, M., Belongie, S.J., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C.L., 2014. Microsoft coco: Common objects in context, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 8693, pp.740-755, DOI: 10.1007/978-3-319-10602-1_48.
- Mandirola, M., Casarotti, C., Peloso, S., Lanese, I., Brunesi, E., and Senaldi, I., 2022. Use of UAS for damage inspection and assessment of bridge infrastructures, International Journal of Disaster Risk Reduction, 72: 102824, DOI: 10.1016/j.ijdr.2022.102824.
- Munawar, H.S., Ullah, F., Shahzad, D., Heravi, A., Qayyum, S., and Akram, J., 2022. Civil infrastructure damage and corrosion detection: An application of machine learning, Buildings, 12(2): 156, DOI: 10.3390/buildings12020156.
- Pirie, C., and Moreno-García, C.F., 2021. Image pre-processing and segmentation for real-time subsea corrosion inspection, Proceedings of the 22nd Engineering Applications of Neural Networks Conference (EANN2021), Halkidiki, Greece, pp.220-231, DOI: 10.1007/978-3-030-80568-5_19.
- Ren, S., He, K., Girshick, R.B., and Sun, J., 2015. Faster R-CNN: Towards real-time object detection with region proposal networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39: 1137-1149, DOI: 10.1109/TPAMI.2016.2577031.
- Russell, B.C., Torralba, A., Murphy, K.P., and Freeman, W.T., 2008. LabelMe: A database and web-based tool for image annotation, International Journal of Computer Vision, 77(1-3): 157-173, DOI: 10.1007/s11263-007-0090-8.
- Shen, H.K., Chen, P.H., and Chang, L.M., 2013. Automated steel bridge coating rust defect recognition method based on color and texture feature, Automation in Construction, 31: 338-356, DOI: 10.1016/j.autcon.2012.11.003.

- Ta, Q.B., Huynh, T.C., Pham, Q.Q., and Kim, J.T., 2022. Corroded bolt identification using mask region-based deep learning trained on synthesized data, *Sensors*, 22(9): 3340, DOI: 10.3390/s22093340.
- Vorobel, R., Ivasenko, I., Berehulyak, O., and Mandzii, T., 2021. Segmentation of rust defects on painted steel surfaces by intelligent image analysis, *Automation in Construction*, 123: 103515, DOI: 10.1016/j.autcon.2020.103515.
- Wang, D., Zhang, Y., Pan, Y., Peng, B., Liu, H., and Ma, R., 2020. An automated inspection method for the steel box girder bottom of long-span bridges based on deep learning, *IEEE Access*, 8: 94010-94023, DOI: 10.1109/ACCESS.2020.2994275.
- Wu, Z.Y., and Kalfarisi, R., 2019. Deep learning-based defect detection and assessment for engineering structures, *Proceedings of the 5th International Conference on Smart Monitoring, Assessment and Rehabilitation of Civil Structures (SMAR 2019)*, Potsdam, Germany, pp.1-7.
- Xi, D., Qin, Y., and Wang, Y., 2020. Vision measurement of gear pitting under different scenes by deep mask R-CNN, *Sensors*, 20(15): 4298, DOI: 10.3390/s20154298.
- Zhang, Y., Zheng, K., Zhu, J., Lei, M., and Feng, X., 2021. Research on corrosion and fatigue performance of weathering steel and High-Performance steel for bridges, *Construction and Building Materials*, 289: 123108, DOI: 10.1016/j.conbuildmat.2021.123108.

Research on Identifying Corrosion of Steel Structure Bridges Using Deep Learning Model Mask R-CNN

Jyun-Cheng Li ^{1*} Szu-Pyng Kao ² Feng-Liang Wang ³ Jhih-Sian Lin ³

Abstract

In recent years, steel structure bridges have become increasingly popular in Taiwan, but Taiwan is prone to corrosion, so a lot of costs must be invested in the maintenance of bridge corrosion detection. In this study, UAV and Mask R-CNN are used to conduct corrosion identification experiments, and the results show that this method can improve the shortcomings of traditional bridge detection. At present, there is no corrosion dataset suitable for this study in Taiwan, so this study has established a corrosion data set by itself; It is imported into Mask R-CNN for training, and through the experiment of this study, the best super-parameter configuration is selected, that is, “optimizer SGD collocation learning rate 1×10^{-3} ”, and the evaluation index results of model training are: Recall can reach 97.1%, Precision can reach 90.4%, mAP can reach 91.0% and mIoU can reach 89.0%; After the analysis of the identification results, it is found that the background noise in the image will affect the identification of corrosion.

Keywords: Steel Structure Bridge, Bridge Detection, UAV, Corrosion, Mask R-CNN

¹ Master Student, Department of Civil Engineering, National Chung Hsing University

² Professor, Department of Civil Engineering, National Chung Hsing University

³ Ph.D. Candidate, Department of Civil Engineering, National Chung Hsing University

* Corresponding Author, Tel: 886-4-22840437 ext.306, E-mail: b0976245966@gmail.com

Received Date: Jan. 30, 2023

Revised Date: Feb. 23, 2023

Accepted Date: Mar. 28, 2023