

計畫編號：ARDSWC-112-060

建置 UAV 巡檢影像與發展智慧型影像分析
判識水土保持設施混凝土泥表面異常之
可行性評估研究

**A study of feasibility assessment on
building UAV inspection images and
developing intelligent image analysis for
automatic recognition of surface damage
of concrete structure of water and soil
conservation facilities**
成果報告

執行單位：國立中興大學

執行期間：112 年 02 月 08 日至 112 年 12 月 31 日

計畫主持人：蔡鴻旭 教授

農業部農村發展及水土保持署 編印

中華民國 112 年 12 月

(本報告書內容及建議純屬執行單位意見，僅供本署施政參考)

建置 UAV 巡檢影像與發展智慧型影像分析判識水土保持設施混凝土表面異常之可行性評估研究

摘要

本計畫發展自適應影像語意分割技術可自動判識影像水土保持設施混凝土表面異常應用於無人機(unmanned aerial vehicle, UAV)定期巡檢，主要檢測標的以水土保持設施中護岸混凝土表面之裂隙。首先，本計畫基於 UAV 拍攝參數發展由影像語意分割裂隙來估算混凝土表面實際裂隙寬度演算法，接著，建置時間序 UAV 巡檢影像歷程資料庫及提供 UAV 定期巡檢影像拍攝參數與操作流程。然後，在 UAV 影像樣本少及不平衡資料集條件下，發展多尺度自適應影像語意分割方法可識別巡檢影像中不同解析度混凝土表面裂隙，同時也發展出抗陰影干擾判識裂隙自適應生成式影像語意分割方法，提升裂隙識別效能，最後，建置響應式互動網站，讓使用者可體驗以上述方法。

關鍵詞：水土保持設施、UAV 巡檢影像、影像分析、影像分割、深度學習、卷積神經網路、生成對抗網路、表面瑕疵、轉移學習、領域自適應、多尺度影像解析度學習、不平衡樣本、抗陰影干擾裂隙偵測

A study of developing intelligent image segmentation based on deep learning for detection on surface defects of concrete structure for water and soil conservation facilities

Abstract

The project proposes the domain adaptation based semantic image segmentations which can be used in detecting the cracks on the concrete structure of water and soil conservation facilities. These methods can be also employed in inspecting the surface damage of images which are obtained by UAVs with an automatic cruise control system. The inspection targets of water and soil conservation facilities in the project aims at the cracks in the concrete revetment.

In the first issue of the project is to develop an algorithm to estimate the width of cracks on the concrete from the images processed in the image segmentation techniques according to the relationships of photo parameters of UAVs and the distance between camera of UAVs and the crack on the concrete. Consequently, images are collecting from cameras of UAVs, which can be constructed in a portfolio database in a way of the temporal (time) sequences. Additionally, a multi-resolution learning technique based on domain adaption and deep convolutional network is developed for semantic image segmentation on the crack detection for the surface of the concrete structure of target facilities. The domain

adaption approach is utilized in the design of the model for semantic image segmentation to overcome the imbalanced pixel sets or less samples problems. Moreover, a semantic image segmentation technique based on domain adaptation and conditional generative adversarial networks is devised in the project, which can be exploited in recognize cracks to resist on the interface of shadows on the surface of the concrete structure of target facilities. Therefore, these image-segmentation models mentioned above can be applied in the detection of abnormal image contents for the portfolio database. Finally, the project builds the web site followed by Responsive Web Design (RWD) protocol, which offers the operations for detecting the abnormal UAV images by these two models proposed in the project. This realizes so-called the user-experience playing.

Keywords : Soil and Water facilities, UAV Cruise Images, Image Analysis, Image Segmentation, Semantic Image Segmentation, Deep Learning, Convolutional Neural Networks, Generative Adversarial Network, Surface Defects, Transfer Learning, Domain Adaptation, Image Resolution Learning with Multiple scales, Cracks with Shadows Inspection, Imbalance data

目次

摘要.....	ii
Abstract.....	iii
目次.....	v
圖次.....	vii
第一章 計畫內容	1-1
第一節 計畫動機	1-1
第二節：計畫目的	1-3
第三節：背景	1-4
一、UAV 攝影參數說明	1-4
二、影像分割相關技術研究.....	1-8
三、多尺度 UAV 影像分割相關研究.....	1-9
四、領域自適應相關研究.....	1-10
五、影像分割效能量化指標.....	1-11
第二章 工作執行方法與步驟	2-1
第一節 工作執行方法	2-1
一、UAV 拍攝影像參數設定及量測裂縫寬度之研究.....	2-1
二、UAV 巡檢航線參數設定與巡檢歷程影像資料庫建置之研究	2-1
三、多尺度影像分割應用於偵測 UAV 影像混凝土表面裂隙之研究	2-1
四、生成式混凝土表面影像分割應用於偵測 UAV 影像混凝土表面 裂隙之研究	2-1
五、建置響應式網頁模型識別體驗網站	2-2
第二節 工作執行步驟	2-3
一、UAV 拍攝影像參數設定及量測裂縫寬度之研究.....	2-3
二、UAV 巡檢航線參數設定與巡檢歷程影像資料庫建置之研究	2-18
三、多尺度影像分割應用於偵測 UAV 影像混凝土表面裂隙之研究	2-24
四、生成式混凝土表面影像分割應用於偵測 UAV 影像混凝土表面 裂隙之研究	2-43
五、建置響應式網頁模型識別體驗網站	2-52
第三章 執行成果及效益	3-1

第一節 執行成果	3-1
第二節 執行效益	3-1
第三節 交付項目	3-3
第四章 結論.....	4-1
參考文獻.....	參 1
附錄一、112 年水保局創新研究計畫審查意見暨回覆說明	附-1
附錄二、112 年水保局創新研究期中報告審查意見暨回覆說明	附-3
附錄三、空拍機參數說明	附-5
附錄四、無人機地面採樣距離（Ground sample distance，GSD）概念，與無人機鏡頭與混凝土表面之距離建議表	附-10
附錄五、資料擴增用於模擬現實情況之成果	附-11
附錄六、拍攝花絮	附-14
附錄七、期末審查意見暨回覆說明	附-146

圖次

圖 1-1:本計畫各主題關聯圖	1-2
圖 1-2: 實景拍攝範圍與影像對應、參數位置示意圖	1-4
圖 1-3:VFOV 於實景與影像對應之示意圖	1-4
圖 1-4: HFOV 於實景與影像對應之示意圖	1-5
圖 1-5: DFOV 於實景與影像對應之示意圖	1-5
圖 2-1: 真實裂隙寬成像所需像素數流程	2-3
圖 2-2: 對應不同寬度的裂隙手繪圖	2-6
圖 2-3: 拍攝不同 WD 下裂隙影像.....	2-6
圖 2-4:影像重新建模寬度(左圖)0.06M(60MM); (右圖)0.10M(100MM)之裂縫	2-10
圖 2-5:RESIZE 影像(I')與原影像(I)像素比例關係	2-11
圖 2-6:建模以取得裂縫現實長度	2-13
圖 2-7:人工標註裂隙影像。(紅線)60MM 寬裂隙;(綠線)100MM 寬裂縫..	2-13
圖 2-8: 影像命名規則說明.....	2-18
圖 2-9: 單次巡檢 CSV 表格.....	2-19
圖 2-10: UAV 影像儲存架構	2-19
圖 2-11:UAV 實驗影像關係概念	2-20
圖 2-12: API 資料庫概念。(紅框)同次巡檢拍攝(藍框)相近 GPS 拍攝影像.	2-20
圖 2-13: 眉溪觀音橋護岸模擬巡檢影像	2-23
圖 2-14:模擬巡檢 GPS 飛行軌跡(藍)手動拍攝(紅)第一次復現飛行軌跡(綠) 第二次復現飛行軌跡	2-23
圖 2-15 : P/O 率模型概念架構圖.....	2-24
圖 2-16 : 不同 P/O 率之影像範例	2-25
圖 2-17 : P/O 率回歸預測架構圖	2-27
圖 2-18 : P/O 率回歸與語意分割模型架構圖	2-29
圖 2-19 : C-SE 模組架構圖	2-29
圖 2-20 : 水平裂縫_2~3M 範例影像.....	2-33
圖 2-21 : 水平拉距_距起降點 7M 範例影像.....	2-33
圖 2-22 : 下游斜拍_起降點高 6M 範例影像.....	2-34
圖 2-23 : 垂直拉距_起降點高 10M 範例影像.....	2-34
圖 2-24 : 與 111 年之效能比較.....	2-36
圖 2-25 : 視覺化模擬干擾比較	2-38
圖 2-26 : 水平裂縫_2~3M 範例影像.....	2-39
圖 2-27 : 水平拉距_距起降點 7M 範例影像.....	2-39

圖 2-28：下游斜拍_起降點高 6M 範例影像.....	2-40
圖 2-29：垂直拉距_起降點高 10M 範例影像.....	2-40
圖 2-30：學習 μ 和 $\mu - 1$ 之架構	2-44
圖 2-31：源影像和 MASK 特徵對應架構	2-44
圖 2-32：(A)KAGGLE 裂縫影像，(B)路面裂縫影像.....	2-46
圖 2-33：(A)ISTD 影像，(B)陰影遮罩	2-46
圖 2-34：混凝土和路面裂縫分割結果。(A)CNN，(B)OURS (w/o SHADOW)， (C)OURS (w SHADOW)，(D)[TSAI ET AL., 2018]。	2-49
圖 2-35：空拍混凝土裂縫分割結果。(A)水平裂縫_2~3M; (B)水平拉距_起降點 7M; (C)垂直拉距_起降高 10M; (D)下游斜拍_起降高 6M。	2-50
圖 2-36：多尺度 UAV 影像分割應用於混凝土表面裂隙偵測網站架構.....	2-52
圖 2-37：多尺度 UAV 影像分割應用於混凝土表面裂隙偵測網頁效果展示。	2-53
圖 2-38：多尺度 UAV 影像分割應用於混凝土表面裂隙偵測網頁模擬手機展 示.....	2-53
圖 2-39：生成式混凝土表面影像裂縫分割網頁架構圖	2-54
圖 2-40：生成式混凝土表面影像裂縫分割網頁效果展示	2-55
圖 2-41：生成式混凝土表面影像裂縫分割網頁模擬手機展示	2-55
圖 附錄 3-1：起降點高 6M、12M 上游向下游斜拍示意圖.....	附-6
圖 附錄 3-2：起降點高 6M、12M 下游向上游斜拍示意圖	附-6
圖 附錄 3-3：起降點高 2M、10M 垂直拉距示意圖	附-7
圖 附錄 3-4：起降點水平距離 7M、12M、18M 示意圖	附-7
圖 附錄 3-5：起降點高 6M(左)、12M(右)上游向下游斜拍圖.....	附-8
圖 附錄 3-6：起降點高 6M(左)、12M(右)下游向上游斜拍圖.....	附-8
圖 附錄 3-7：起降點垂直高 2M(左)、10M(右)拍攝圖.....	附-8
圖 附錄 3-8：起降點水平距離 7M(左上)、12M(右上)、18M(下)拍攝圖	附-9
圖 附錄 6-1：實驗場域環境照.....	附-14
圖 附錄 6-2：眉溪觀音吊橋附近護岸拍攝	附-15

表次

表 2-1:手機鏡頭估測最遠工作距離	2-8
表 2-2: UAV 鏡頭估測最遠工作距離	2-8
表 2-3 : 於空拍影像測試資料集上的效能衡量	2-32
表 2-4 : 於昏暗空拍影像測試資料集上的效能衡量	2-36
表 2-5 : 於起霧空拍影像測試資料集上的效能衡量	2-36
表 2-6 : 於下雨空拍影像測試資料集上的效能衡量	2-37
表 2-7 : 其他多尺度模型之效能比較	2-41
表 2-8: 路面分割量化結果.....	2-50
表 2-9: 空拍混凝土裂縫語意分割之量化結果	2-51
表 附錄 3-1:無人機拍攝參數表	附-8
表 附錄 4-1:無人機安全感測系統	附-10
表 附錄 4-2:與表面之最短距離建議表	附-10
表 附錄 5-1: 水平裂縫_2~3M 資料擴增成果.....	附-11
表 附錄 5-2: 下游斜拍_起降點高 6M 資料擴增成果.....	附-12
表 附錄 5-3: 水平拉距_距起降點 7M 資料擴增成果.....	附-12
表 附錄 5-4: 垂直拉距_起降點高 10M 資料擴增成果.....	附-13

第一章 計畫內容

第一節 計畫動機

近期以 UAV 來實施攝影監測已逐漸運用於許多實務應用領域，利用 UAV 來巡檢水土保持設施也逐漸興起熱門的研究，主要是利用 UAV 拍攝影像(影片)後再觀察水土保持設施表面否有異常或裂隙。可是目前上述工作，仍須人工介入收集整理 UAV 影像及觀察水土保持設施表面否有裂隙。因此，本計畫欲挑戰上述問題。首先，本計畫發展一個 UAV 水土保持設施影像態樣歷程系統，利用演算法來記錄時序巡檢資料庫系統，接著，利用人工智慧的深度學習影像分析技術分析單次巡檢及連續巡檢的影像差異。但是收集水土保持設施表面是否有裂隙的影像不易，此引發僅有少量影像問題，另外，因大多數的設施都是正常只有少部分有異常，如裂縫、裂隙等，此引發巡檢影像集為類別不平衡，因此，克服上述少樣本及類別不平衡問題，本計畫已研究發展人工影像自動產生技術及領域適性學習與對比學習來解決少樣本及類別不平衡發展智慧型影像分類及分割技術應用於自動辨識混凝土護岸設施表面是否裂隙，另外因為 UAV 影像尺度與地面拍攝有所不同，本計畫已研究多尺度影像辨識。圖 1-1 說明本計畫主題關聯圖。

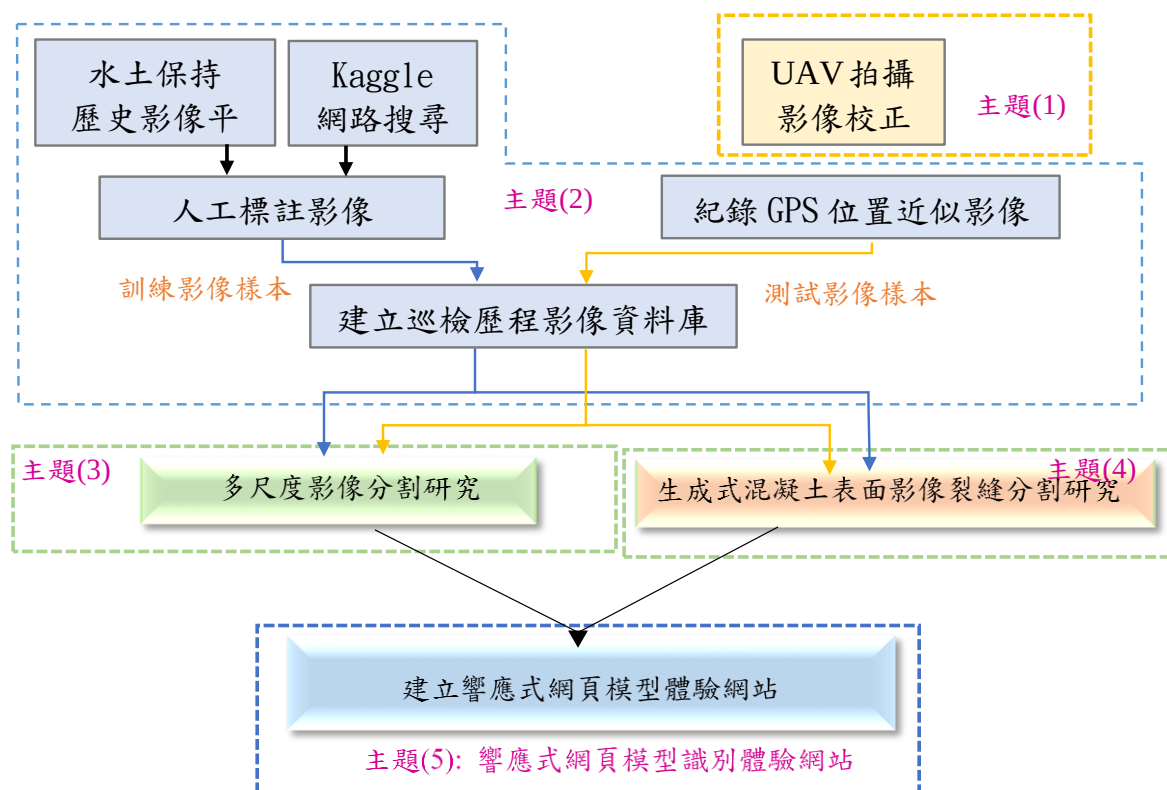


圖 1-1:本計畫各主題關聯圖

第二節：計畫目的

本計畫目標為完成下列工作項目。

- UAV 拍攝影像參數設定及量測裂縫寬度之研究
- UAV 巡檢航線參數設定與巡檢歷程影像資料庫建置之研究
- 多尺度影像分割應用於偵測 UAV 影像混凝土表面裂隙之研究
- 生成式混凝土表面影像分割應用於偵測 UAV 影像混凝土表面裂隙之研究
- 建置響應式網頁模型識別體驗網站

第三節：背景

一、UAV攝影參數說明

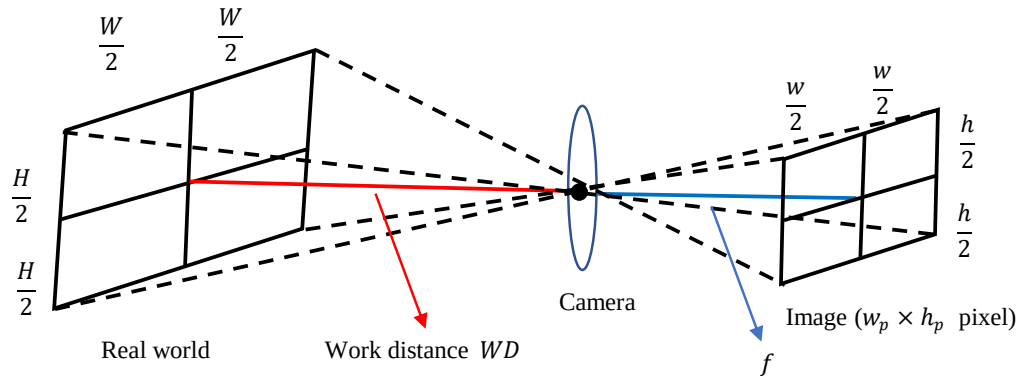


圖 1-2: 實景拍攝範圍與影像對應、參數位置示意圖

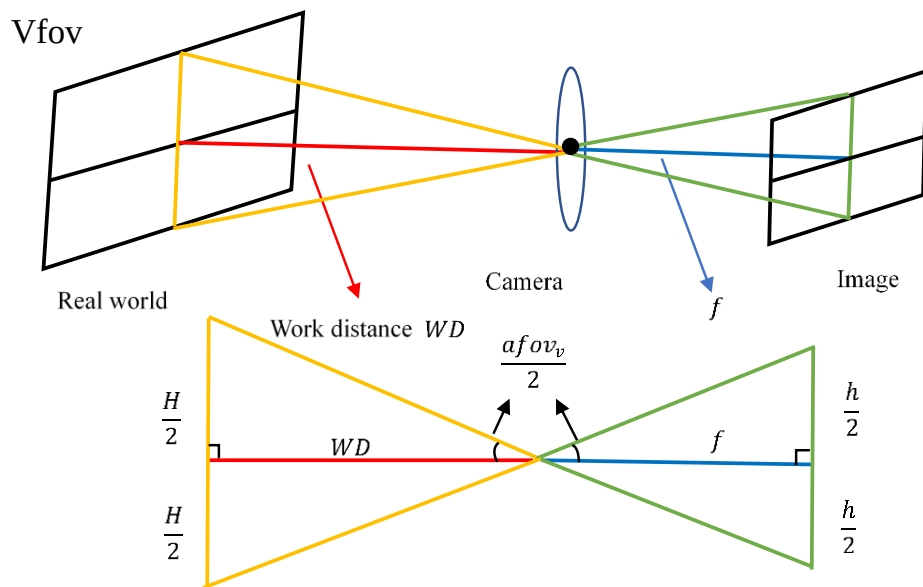


圖 1-3: Vfov 於實景與影像對應之示意圖

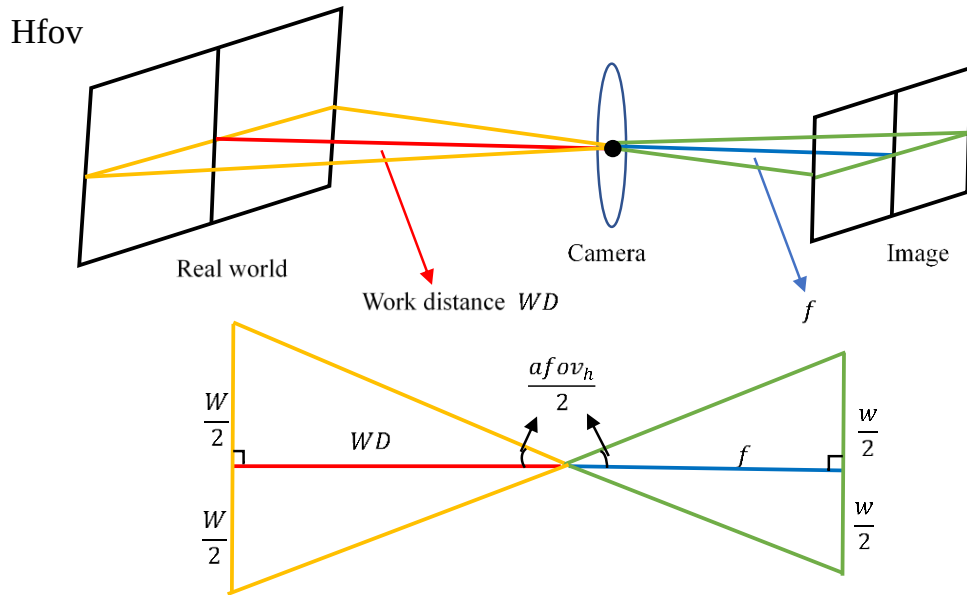


圖 1-4: Hfov 於實景與影像對應之示意圖

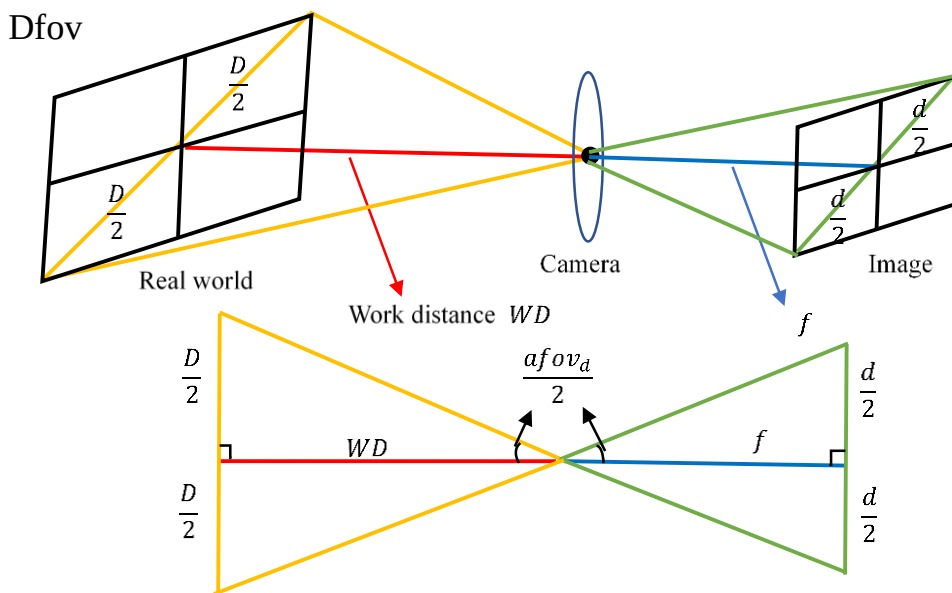


圖 1-5: Dfov 於實景與影像對應之示意圖

附錄三說明本計畫實驗所使用 UAV 及其攝影規格。

Fov(Field of View) 在一般情況下指的是視場，意思是人眼或是設備所看到的觀察區域；如果是在相機參數，則指的是視角場 AFov(Angular Field of View)，意思是以度(Degree)為單位的

全角，即設備捕捉光線之間的角度。而 Fov 在大部分的情況都是表示 DFov(Diagonal Field of View)，在相機參數中的角度也都是指 DFov 的 AFov，而圖 1-3 表示 VFoV、圖 1-4 表示 HFov、圖 1-5 表示 DFov，下方公式可以從角度方面簡單表示，實景在垂直、水平及對角經過鏡頭與影像的關聯。

$$fov_v = 2 \tan^{-1} \left(\frac{h}{2f} \right) = 2 \tan^{-1} \left(\frac{H}{2WD} \right) ;$$

$$H = 2 WD \tan \left(\frac{fov_v}{2} \right) ;$$

$$fov_h = 2 \tan^{-1} \left(\frac{w}{2f} \right) = 2 \tan^{-1} \left(\frac{W}{2WD} \right) ;$$

$$W = 2 WD \tan \left(\frac{fov_h}{2} \right) ;$$

$$fov_d = 2 \tan^{-1} \left(\frac{d}{2f} \right) = 2 \tan^{-1} \left(\frac{D}{2WD} \right)$$

下列說明公式中各符號說明。

H : Real world height. (mm)

W : Real world width. (mm)

D : Real world diagonal. (mm)

WD : Work distance, the vertical distance from the lens to the shooting surface. (mm)

γ : Real world crack width(mm)

f : Focal length. (mm)

h : Image height. (mm)

w : Image width. (mm)

d : Image diagonal. (mm)

w_p : Pixel width of the image

h_p : Pixel height of the image

fov_v : Vertical angular field of view. (degree)

fov_h : Horizontal angular field of view. (degree)

fov_d : Diagonal angular field of view. (degree)

二、影像分割相關技術研究

今年來基於深度神經網路的影像語意分割方法。如[Long, et al., 2015]中提出，可將分類用的卷積神經網路（如 AlexNet、VGG 或 ResNet）轉化為用於語意分割的全卷積網路（FCN）。為改進模型，目前已經開發多種方法，包括利用上下文資訊[Hung, et al., 2017]或擴大感知域[Yu & Koltun, 2016]。然而，訓練這些高級網路，需要收集大量的密集像素標註，以匹配深度卷積神經網路的模型容量。因此，最近提出弱監督和半監督方法[Pathak, et al., 2015]，以降低收集分割真實標註的高昂成本。然而，在大多數實際應用中，很難獲得弱標註，而且訓練模型可能無法很好地適應未見影像領域。另一種應對標註問題的方法是基於渲染構建合成數據集，例如 GTA5[Richter, et al., 2016]和 SYNTHIA[Ros, et al., 2016]。雖然這些數據集的數據收集成本較低，因為可以通過部分自動化過程進行像素級標註，但通常這些數據集與真實世界數據集一起用於聯合學習以提高性能。[Chen, et al., 2023]利用了重建遮罩實現影像分割，使得標註的遮罩能夠以 RGB 影像的形式輸入進自編碼器並和原影像的特徵對齊並由解碼器輸出重建後的 RGB 影像。

三、多尺度UAV影像分割相關研究

HED (Xie & Tu, 2015)、DeepCrack (Liu et al., 2019)將各層輸出特徵分別進行分類預測，確保各層的特徵對分類皆有貢獻，以學習多尺度特徵。RCF (Liu et al., 2017)延伸 HED，融合同一尺度下之卷積層輸出以獲得更豐富的特徵資訊。FPN (Lin et al., 2017)逐步由 high level 往 low level 進行特徵融合。FPHBN (Yang et al., 2019)、DeepCrack (Zou et al., 2018)改良 HED，結合類似 FPN 方式融合多尺度特徵，並同時各別進行語意分割。GCN(Peng et al., 2017)透過拆解卷積核增加有效感受域(effective receptive field)且同時保持低計算複雜度。DANet (Fu et al., 2019)引入雙注意力機制，使模型能夠學習到全域的上下文關聯性，增加有效感受域。DeepLabv3+(Chen et al., 2018)使用 ASPP (Atrous Spatial Pyramid Pooling)學習多尺度特徵，ASPP 透過不同膨脹率的空洞卷積(Yu & Koltun, 2015; Yu & Koltun, 2017)學習多尺度特徵。DenseASPP (Yang et al., 2018)透過稠密連接(densely connect)不同膨脹率的空洞卷積改良 ASPP。DMA-Net (Sun et al., 2022)改良 DeepLabv3+，學習 attention mask，並用於計算 low-level 以及 high-level 特徵的權重。PCANet(Sang et al., 2020)提出結合 ASPP 與 self-attention(Vaswani et al., 2017)的 Vortex Atrous Attention 模組，同時捕捉全域的上下文關聯性以及相對位置訊息。MFPA-Net (Jiang

et al., 2022)結合 DANet、FPN 與 ASPP，使其多尺度特徵擷取能力，與 UAV 影像之性質契合。訓練與測試影像的解析度不一致容易導致 vision transformer 的效能降低，而 SegFormer (Xie et al., 2021)透過移除 positional embedding 並加入 Mix-FFN 解決此問題。CrossViT (Chen et al., 2021)將影像分割成兩組：large patch 與 small patch，並利用 cross attention 使兩個組交換訊息，進而學習多尺度特徵。CrackFormer (Liu et al., 2021; Liu et al., 2023)使用 local self-attention 模組捕捉大感受域訊息、scaling-attention 模組使得裂縫邊界更清晰。然而以上方法僅透過設計良好的架構期待模型能夠處理多尺度的問題，卻缺乏對多尺度資訊的監督限制，導致模型可能無法有效利用訓練影像集固有的多尺度資訊。ResFormer(Tian et al., 2023)透過 scale consistency loss，使得同一張影像在不同解析度下能夠相同的預測結果。但是此方法僅限制不同解析度之模型預測應一致，並沒有提取同一張影像在不同解析度下所蘊含的額外資訊。

四、領域自適應相關研究

影像分類的領域適應方法已經發展以應對源領域和目標領域之間的域偏移問題。基於 CNN 分類器的性能增益。這些方法背後的主要見解是通過調整來源影像和目標影像之間的特徵分

布來解決問題。[Ganinand, et al., 2015] 提出 Domain Adversarial Neural Network(DANN)用於轉移特徵分布。之後提出多種變體，使用不同的損失函數[Long, et al., 2015]或分類器[Long, et al., 2016]。最近，PixelDA 方法[Hoffman, et al., 2016]通過將源影像轉移到目標領域，從而獲得目標影像的模擬訓練集，從而解決影像分類的域適應問題。注意到，針對像素級預測任務的域適應任務尚未被廣泛探索。Hoffman 等人[Hoffman, et al., 2016]引入在語意分割上應用對抗學習的域適應任務，採用完全卷積的方式處理特徵表示和其他類別約束，類似於受約束的 CNN。其他方法則專注於通過採用分類別對抗學習或標籤轉移，將合成到真實或跨城市影像進行適應。與 PixelDA 方法類似，一個同時進行的工作，CyCADA[Hoffman, et al., 2017]使用 CycleGAN[Zhu, et al., 2017]將源領域影像轉移到目標領域並進行像素對齊，從而生成額外的訓練數據，並結合特徵空間對抗學習。[Tsai, et al., 2018]利用識別遮罩是否正常來將遊戲交通影像的識別能力遷移至真實影像。

五、影像分割效能量化指標

令 y 與 $\hat{y}$ 分別是語意分割標註與語意分割預測，其中 y_i 表示 y 在 i 上的像素值、 i 為影像上的座標。實驗結果的衡量指標定義

如下：

- Precision

$$\text{precision}(\hat{y}, y) = \frac{\sum_i y_i \hat{y}_i}{\sum_i y_i \hat{y}_i + \sum_i (1 - y_i) \hat{y}_i}$$

- Recall

$$\text{recall}(\hat{y}, y) = \frac{\sum_i y_i \hat{y}_i}{\sum_i y_i \hat{y}_i + \sum_i y_i (1 - \hat{y}_i)}$$

- F1-score

$$\text{F1-score}(\hat{y}, y) = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

- mIoU (mean Intersection over Union)

$$\text{mIoU}(\hat{y}, y) = \frac{1}{2} \left(\frac{\sum_i y_i \hat{y}_i}{\sum_i y_i + \sum_i \hat{y}_i - \sum_i y_i \hat{y}_i} + \frac{\sum_i (1 - y_i)(1 - \hat{y}_i)}{\sum_i 1 - y_i + \sum_i 1 - \hat{y}_i - \sum_i (1 - y_i)(1 - \hat{y}_i)} \right)$$

- Accuracy:

Accuracy($\hat{y}, y$)

$$= \frac{\sum_i y_i \hat{y}_i + \sum_i (1 - y_i)(1 - \hat{y}_i)}{\sum_i y_i \hat{y}_i + \sum_i (1 - y_i)(1 - \hat{y}_i) + \sum_i y_i (1 - \hat{y}_i) + \sum_i (1 - y_i) \hat{y}_i}$$

- DICE:

$$\text{Dice}(y, \hat{y}) = \frac{2 \sum_i y_i \hat{y}_i}{\sum_i y_i + \sum_i \hat{y}_i} = \frac{2|y \cap \hat{y}|}{|y| + |\hat{y}|}$$

第二章 工作執行方法與步驟

第一節 工作執行方法

本計畫各主題的工作執行方法說明如下。

一、UAV拍攝影像參數設定及量測裂縫寬度之研究

- (一) 拍攝影像校正及距離量測實驗
- (二) 實驗模擬之研究
- (三) 量測裂縫與估算拍攝距離
- (四) 估計真實裂隙寬度測試之研究
- (五) 以模型轉換影像估計真實裂縫寬度

二、UAV巡檢航線參數設定與巡檢歷程影像資料庫建置之研究

- (一) UAV 巡檢路線歷程設計規劃
- (二) 提取 UAV 影像描述資訊
- (三) 建立水土保持護岸設施巡檢歷程資料庫
- (四) 水土保持護岸設施巡檢模擬實驗

三、多尺度影像分割應用於偵測UAV影像混凝土表面裂隙之研究

- (一) 多尺度影像分割模型設計
- (二) 實驗影像集說明
- (三) 空拍影像測試資料集實驗結果
- (四) 模型效能比較

四、生成式混凝土表面影像分割應用於偵測UAV影像混凝土表面裂隙之研究

- (一) 裂縫分割模型建模方法
- (二) 實驗影像集說明

(三) 實驗結果

五、建置響應式網頁模型識別體驗網站

(一)網頁呈現生成式混凝土表面影像裂縫分割模型

(二)網頁呈現多尺度 UAV 影像分割應用於混凝土表面裂隙偵測模型

第二節 工作執行步驟

一、UAV拍攝影像參數設定及量測裂縫寬度之研究

本主題目標估計適合 UAV 攝影鏡頭距標的物的拍攝距離 (Work distance, WD)，當 UAV 在 WD 的範圍內進行拍攝，模型可偵測拍攝的裂隙影像。接著，研究裂隙寬度與合適拍攝距離的關係。最後，研究如何透過影像中裂隙像素寬度(pixel)，估算真實裂隙寬度 (mm)。

(一) 拍攝影像校正及距離量測實驗

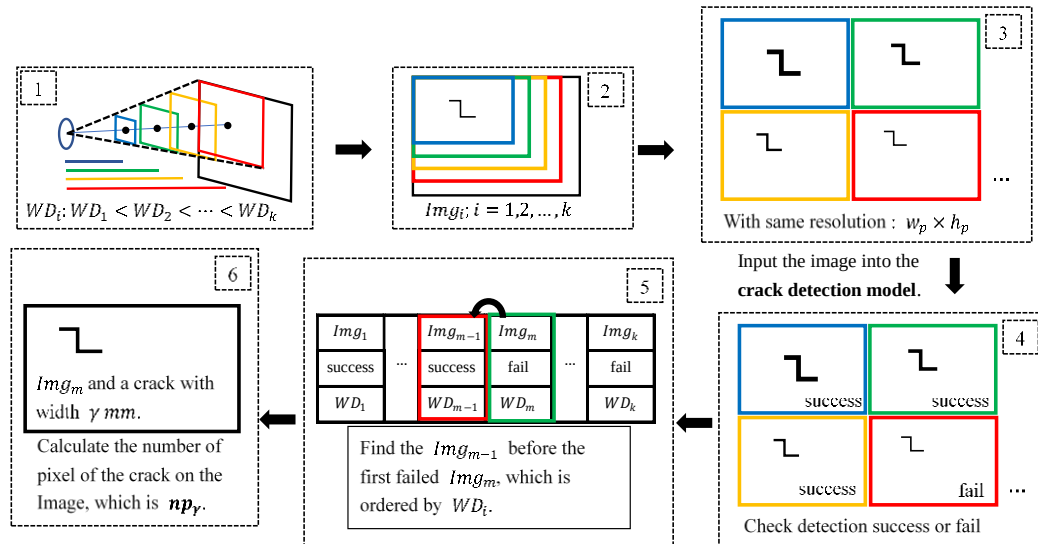


圖 2-1: 真實裂隙寬度成像所需像素數流程

為取得 WD 可以透過使用 W 、 w 或是 H 、 h 配合 f 進行比例換算，而一般情況難以直接從影像去計算對應現實的 W 跟 H ，所以這裡介紹一個參數 np_γ (Number of pixels)， np_γ 為實際寬度為 γ (毫米 mm) 的裂隙需要多少像素 (number of pixels) 才能夠被模型 (在實驗室中以人眼對影像進行識別) 偵測到， np_γ 的單位為 pixel，

如果已知像素長度(length of pixel)、 γ 、 f 以及 np_γ 就可以用比例算出 WD ，以下演算法可以估測 np_γ ，並於圖 2-1 說明計算流程。

可偵測裂隙之像素寬(np_γ)演算法

Input: Img_i 、 f 、 h 、 w 、 h_p 、 w_p 、 γ

Output: np_γ

Step 1: 在不同 WD 下拍攝實際裂隙。 $(WD_1 < WD_2 < \dots < WD_k)$ ，

如圖 2-1(1)。

Step 2: 展示不同距離裂縫影像大小，如圖 2-1(2)。

Step 3: 影像皆為相同像素，如圖 2-1(3)可知。

Step 4: 排列影像準備進行偵測，如圖 2-1(4)。

Step 5: If $np_{\gamma_m} = 0$ ，取其前一張影像 Img_{m-1} ，使 $np_\gamma \leftarrow np_{\gamma_{m-1}}$ 。

Else $m \leftarrow m + 1$ ，until $m > k$ ，如圖 2-1(5)。

Step 6: 將 WD_i 拍攝之影像 Img_i 進入模型中計算其 np_{γ_i} ； $i =$

$1, 2, \dots, k$ ，如圖 2-1(6)。

估算 np_γ 後，式(1-1)計算影像裂隙寬度 $\Delta\gamma$ (mm)，然後，利用實際裂隙寬度 γ ，計算影像與現實比例 $\frac{\gamma}{\Delta\gamma}$ ，接著，再利用鏡頭焦距 f ，計算式(1-2)，即可估算 WD 。只要估算 np_γ 即可根據不同鏡頭的焦距來取得不同鏡頭所需之工作距離

以下為相機鏡頭參數(規格):

- f : Focal length. (mm)
- h_p : Pixel height of the image
- w_p : Pixel width of the image
- h : Sensor size height (mm)
- w : Sensor size width (mm)

以下為影像中裂隙寬度:

- np_γ : Pixel of crack width on image

以下為人工模擬裂隙寬度的參數:

- γ : Crack width(mm)

以下為式(1-1)計算之參數說明:

- $\Delta\gamma$: Crack width on image. 影像中裂隙寬
- WD : Work distance, the vertical distance from the lens to the shooting surface (mm)

$$np_\gamma \times \frac{h}{h_p} = np_\gamma \times \frac{w}{w_p} = \Delta\gamma \quad (1-1)$$

$$f \times \frac{\gamma}{\Delta\gamma} = WD \quad (1-2)$$

(二) 實驗模擬之研究

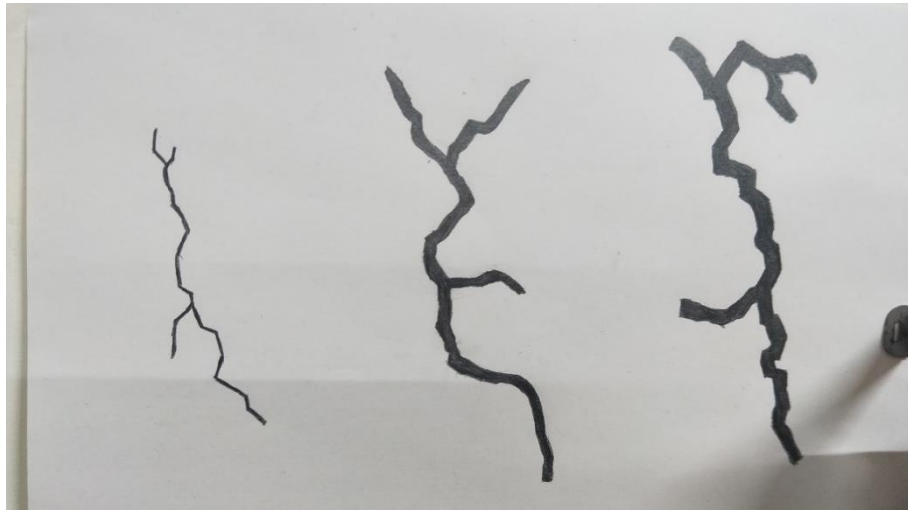


圖 2-2: 對應不同寬度的裂隙手繪圖

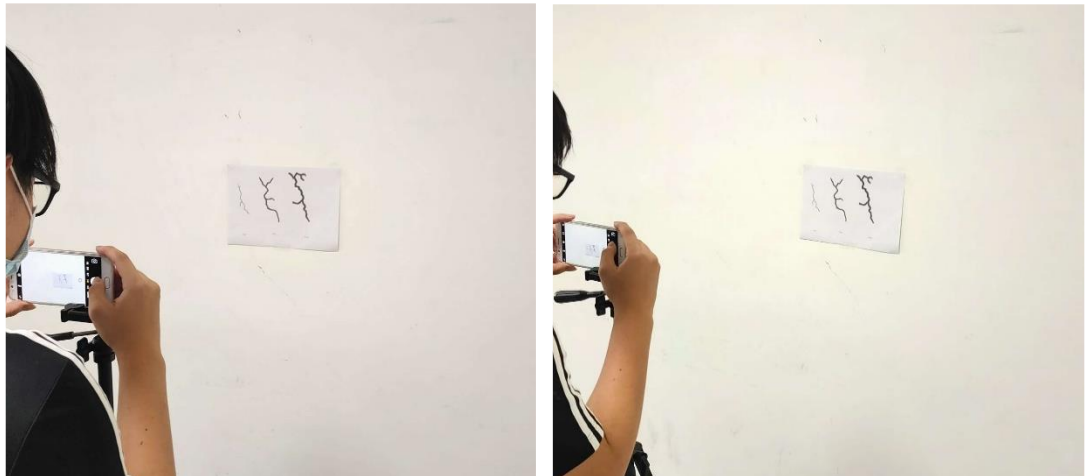


圖 2-3: 拍攝不同 WD 下裂隙影像

本議題模擬實驗，使用手機拍攝不同距離(工作距離 WD)的裂隙圖案。圖 2-2 說明測試手繪裂隙測試影像，圖 2-3 說明拍攝不同 WD 裂隙影像之實驗過程。議題(一)提到之偵測裂隙寬度模型，以手動標註代替，使用方法為使用影像標註工具 labelme 來手動標註，當不同距離的裂隙開始難以辨識時，即不再標註。使用標註點的像素座標計算標註線的像素長，依照流程圖所示找

到每一種寬度的第一次裂隙偵測失敗，取其前一項之像素長 np_γ 。

經觀察 1mm、3mm、5mm 各自可顯示的極限距離分別為 300cm、

600cm、700cm，其 np_γ 數值計算如下：

- 1mm(裂隙寬)於 300cm(WD)

$$np_1 = \frac{1.91050 + 2.90689}{2} \approx 2.408$$

- 3mm(裂隙寬)於 600cm(WD)

$$np_3 = \frac{2.50200 + 3.10161}{2} \approx 2.802$$

- 5mm(裂隙寬)於 700cm(WD)

$$np_5 = \frac{2.81603 + 2.74591}{2} \approx 2.781$$

即模型可識別裂隙寬之最小像素數量，最後本主題選取最大的 np_3 作為推估 UAV 影像拍攝所使用的 np_γ 。以實驗手機鏡頭為例，其鏡頭傳感器的大小為 $5.22 \times 3.92\text{mm}(w \times h)$ ，影像像素為 $4608 \times 3456(w_h \times h_p)$ ，手機焦距為 $4\text{mm}(f)$ ，經式(1-1), (1-2)計算可得出不同裂隙寬度所需 WD，結果如表 2-1。

$$\Delta\gamma = 2.802 \times \frac{5.22}{4608} \approx 0.003174$$

$$WD_1 = 4 \times \frac{1}{0.003174} \approx 1260.24(\text{mm})$$

$$WD_3 = 4 \times \frac{3}{0.003174} \approx 3780.72(\text{mm})$$

$$WD_5 = 4 \times \frac{5}{0.003174} = 6301.20(\text{mm})$$

表 2-1:手機鏡頭估測最遠工作距離

實際裂隙寬度	手機鏡頭拍攝最遠距離 (鏡頭焦距 4mm)
1mm	1260.24mm
3mm	3780.72mm
5mm	6301.20mm

以 UAV 為例，其鏡頭傳感器的大小為 $6.17 \times 4.55\text{mm}(w \times h)$ ，影像像素為 $4000 \times 3000(w_h \times h_p)$ ，焦距為 $4.64\text{mm}(f)$ ，經式(1-1)、(1-2)計算可得出不同裂隙寬度所需 WD 如表 2-2。

表 2-2: UAV 鏡頭估測最遠工作距離

實際裂隙寬度	UAV 鏡頭拍攝最遠距離 (鏡頭焦距 4.6m)
1mm	925.48mm
3mm	2776.44mm
5mm	4627.40mm

例 1: 若 UAV 欲拍攝實際裂隙寬度為 1mm 之影像，其拍攝距離 WD 範圍須介於[300 mm, 925 mm]，超過 925mm 將導致影像中裂隙辨識準確較低，UAV 防撞系統允許最小拍攝距離為 300mm，以避免碰撞風險。

例 2: 若 UAV 欲拍攝實際裂隙寬度為 5mm 之影像，其拍攝距離 WD 範圍須介於[300 mm, 4620mm]，超過 4620mm 將導致影像中裂隙辨識準確較低。

本議題結論：室內手機模擬有其限制，鏡頭距離標的物距離有

限，此外，室內模擬裂縫標註寬度較實際裂隙較小，僅發展測試流程，以後續議題說明實驗場域測試結果為主。

(三) 量測裂縫與估算拍攝距離

上述議題(二)中所提到之可偵測裂縫所需最小像素寬 $np_\gamma = 2.802$ ，鏡頭傳感器的大小為 $13.3 \times 8.88\text{mm}(w \times h)$ ，影像像素為 $5472 \times 3648(w_h \times h_p)$ ，焦距為 $\frac{29}{\frac{36}{13.3}} = 10.7 \text{ mm}(f)$ ，經過式(1-1)計算影像裂隙寬度 $\Delta\gamma(\text{mm})$ ，然後，利用實際裂隙寬度 γ ，計算影像與現實比例 $\frac{\gamma}{\Delta\gamma}$ ，接著，再利用鏡頭焦距 f ，計算式(1-2)，即可取得 WD。其計算公式如下，分別計算真實裂縫寬為 60mm 及 100mm 之 WD。(只要取得 np_γ 即可根據不同鏡頭的焦距來取得不同鏡頭所需之工作距離)

$$2.802 \times \frac{13.3}{5472} \approx 0.00681$$

$$WD_{60} = 10.7 \times \frac{60}{0.00681} \approx 94273.1(\text{mm})$$

$$WD_{100} = 10.7 \times \frac{100}{0.00681} \approx 157121.9(\text{mm})$$



圖 2-4:影像重新建模寬度(左圖)0.06m(60mm); (右圖)0.10m(100mm)之裂縫

圖 2-4 說明取較短之工作距離(WD_{60})，此工作距離經計算為 94273.1mm(毫米)，表示拍攝距離只需小於 94.2731m(公尺)即可拍攝到 60mm(毫米)之裂縫，此次模擬實驗無人機拍攝距離最遠為 18m(公尺)，其小於 94.27m(公尺)，確實在議題(二)的估算範圍內。

本議題結論：空拍機鏡頭正視標地物並距離 2m 至 18m 均可以清楚拍攝 60mm 裂縫的影像並量測到其裂縫寬。

(四) 估算真實裂隙寬度之研究

原始裂隙影像(I)在模型計算中會經過 resize，得到 resize 影像(I')，使經過模型辨識的裂隙影像的像素不同於原始像素，導致

無法透過比例計算實際場景的裂隙寬度，因此，需要先計算出 resize 影像之解析度 $w'_p \times h'_p$ 與原影像之解析度 $w_p \times h_p$ 之長寬換算比例，再使用式(1-3)將 $\Delta w'_p$ 、 $\Delta h'_p$ 計算為原影像裂隙像素寬的 Δw_p 、 Δh_p ($\Delta w'_p$ 、 $\Delta h'_p$ 、 Δw_p 、 Δh_p 以圖 2-5 表示)，接著，再以式(1-4)計算出 np_γ ，最後，使用式(1-1)算出 $\Delta\gamma(\text{mm})$ 再經過式(1-5)即可獲得實際裂隙寬度 $\gamma(\text{mm})$ 。模型計算裂隙影像與拍攝距離 (WD) 計算裂隙寬度演算法說明如下。

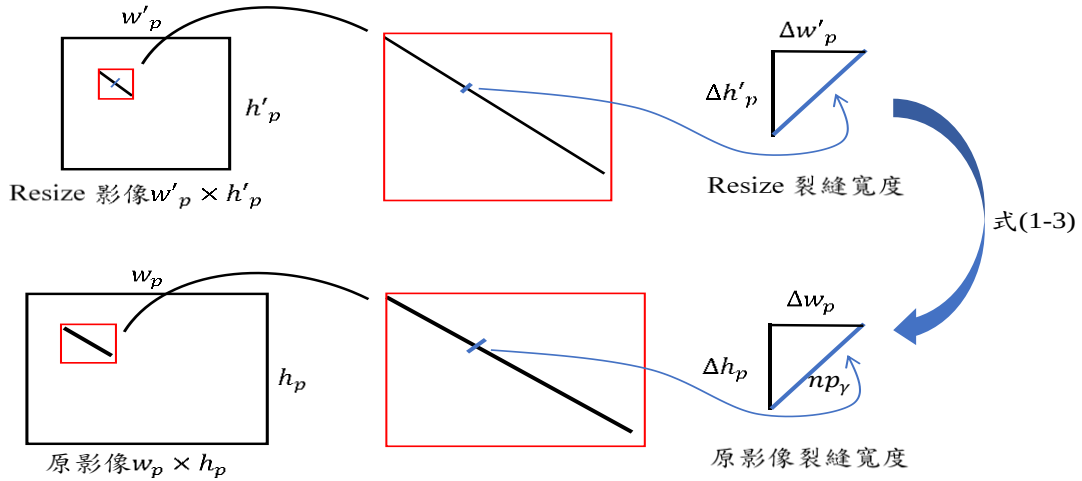


圖 2-5: Resize 影像(I')與原影像(I)像素比例關係

$$\Delta w_p = \Delta w'_p \cdot \frac{w_p}{w'_p}; \Delta h_p = \Delta h'_p \cdot \frac{h_p}{h'_p} \quad (1-3)$$

$$np_\gamma = \sqrt{\Delta w_p^2 + \Delta h_p^2} \quad (1-4)$$

$$\gamma = \frac{WD \times \Delta\gamma}{f} \quad (1-5)$$

模型計算裂隙影像與拍攝距離(WD)計算裂隙寬度演算法:

Input: WD 、 w_p 、 h_p 、 w'_p 、 h'_p 、 $\Delta w'_p$ 、 $\Delta h'_p$ 、 f

Output: 實際裂隙寬度 $\gamma(mm)$

Step 1:從模型計偵測的 I' ，估計 resize 裂隙寬在 x 軸 y 軸上的差異分別為 $\Delta w'_p$ 與 $\Delta h'_p$ 。

Step 2:利用 I 及 I' 的解析度與 resize 裂隙寬的 $\Delta w'_p$ 及 $\Delta h'_p$ 使用式(1-3)計算原影像裂隙寬在 x 軸 y 軸的大小 Δw_p 及 Δh_p 。

Step 3:使用式(1-4)可以計算出裂隙的像素寬 np_γ 。

Step 4:最後利用式(1-1)計算 $\Delta\gamma$ ，配上 WD 及 f 使用式(1-5)即可計算出實際裂隙寬度 γ 。

下面兩例來說明上述演算法的計算結果。

例 1: 若已知影像中裂隙寬為 2.4 pixels，且拍攝地點距離裂隙 3000mm，則實際裂隙寬為約 **1mm**。

例 2: 若已知影像中裂隙寬為 2.8 pixels，且拍攝地點距離裂隙 7000mm，則實際裂隙寬為約 **5mm**。

(五) 以模型轉換影像估計真實裂縫寬度

經過測試後，整理出以下注意事項。

- 巡航路線校正：需標立顯著的起飛地點，確保巡航路線一致，礙於 GPS 定位隨著地球自轉產生偏差，維持相同起飛地點有助於維持飛行軌跡重現的精度。
- 飛行注意事項：大部分 UAV 沒有側向避障功能，拍攝時應確認路線淨空。

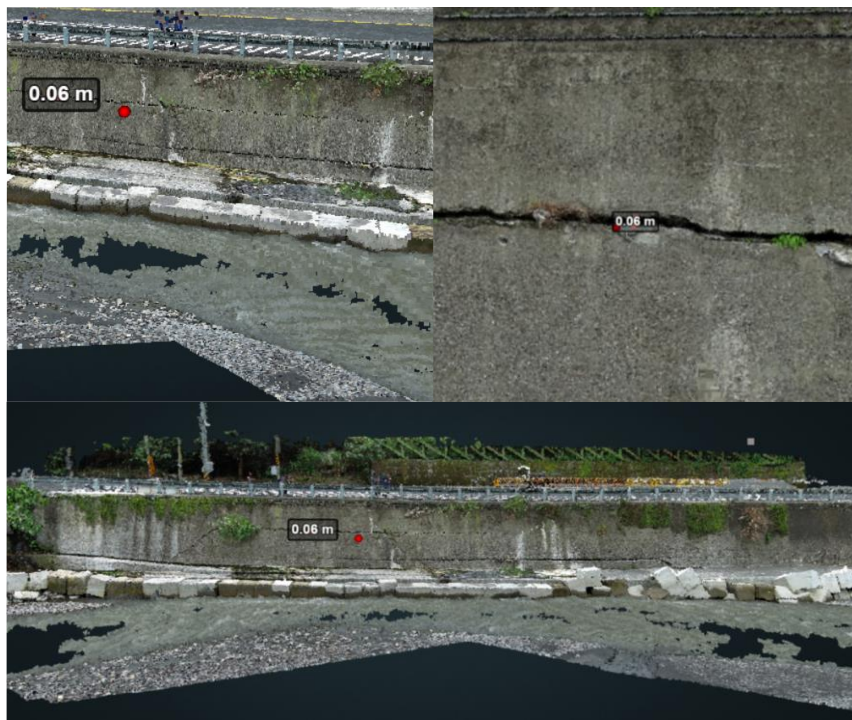


圖 2-6:建模以取得裂縫現實長度



圖 2-7:人工標註裂隙影像。(紅線)60mm 寬裂隙;(綠線)100mm 寬裂縫

以模型計算之裂隙影像與拍攝距離(WD)計算實際裂隙寬度

演算法:

Input: WD 、 x'_1 、 x'_2 、 y'_1 、 y'_2 、 w'_p 、 h'_p 、 w_p 、 h_p 、 f

Output: 實際裂隙寬度 $\gamma(mm)$

Step 1:取出經模型計算出的 resize 影像，擷取出 resize 影像中裂隙寬(線段)，其兩端點座標 (x'_1, y'_1) 及 (x'_2, y'_2) 。

Step 2:利用原影像解析度 $w_p \times h_p$ 及 resize 影像解析度 $w'_p \times h'_p$ 配合 resize 影像中裂隙寬的 $\Delta w'_p$ 及 $\Delta h'_p$ 使用式(1-6)計算對應到原影像裂隙寬兩端點座標 (x_1, y_1) 及 (x_2, y_2) 。

Step 3:使用式(1-7)可以計算出裂隙的像素寬 np_γ 。

Step 4:最後利用式(1-1)計算 $\Delta\gamma$ ，配上 WD 及 f 使用式(1-5)即可計算出實際裂隙寬度 γ 。

$$x_i = x'_i \times \frac{w_p}{w'_p}; y_i = y'_i \times \frac{w_p}{w'_p} \quad (1-6)$$

$$np_\gamma = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1-7)$$

圖 2-7 表示以解析度 5472×3648 鏡頭拍攝，水平距離 18m 的混凝土表面裂隙，經過模型轉換解析度為 912×912 的縮小影像，實際為 γ 裂縫寬在縮小影像中兩點座標為 $(x'_1, y'_1), (x'_2, y'_2)$ 。且已知原影像解析度與縮小影像解析度，經式(1-6)算出兩點座標在

原影像所對應座標 $(x_1, y_1), (x_2, y_2)$ ，再經過式(1-7)計算裂縫在原始影像上之像素寬 $np_\gamma(pixel)$ ，再與鏡頭感知器與影像像素 $(pixel)$ 進行式(1-1)比例換算，計算 $\Delta\gamma$ (裂縫在感知器上之寬度)，再用工作距離(WD)與焦距(f)再次進行比例換算，最後得出真實裂縫寬度。

例(1): 估計實際裂縫 $\gamma = 60mm$ 之裂縫，其 Resize 影像中裂縫寬座標為 $(x'_1, y'_1) = (393.2, 324.3), (x'_2, y'_2) = (393.1, 328.2)$:

$$w = 13.3mm ; w_p = 5472pixel ; h_p = 3648pixel ;$$

$$w'_p = 912pixel ; h'_p = 912pixel ;$$

$$WD = 18000mm ; f = 10.7mm$$

$$x_1 = 393.2 \times \frac{5472}{912} \approx 2359.2 ; y_1 = 324.3 \times \frac{3648}{912} \approx 1297.2$$

$$x_2 = 393.1 \times \frac{5472}{912} \approx 2358.6 ; y_2 = 328.2 \times \frac{3648}{912} \approx 1312.8$$

$$np_\gamma = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \approx 15.6$$

$$15.6 \times \frac{13.3}{5472} \approx 0.03791 (mm)$$

$$0.03791 \times \frac{18000}{10.7} \approx 63.77mm$$

縮小影像中裂縫影像座標轉換為原始影像中裂隙列座標，最後，

實際 60mm 裂縫寬經演算法估算結果為 63.77mm。

例(2):估計實際裂縫 $\gamma = 100\text{mm}$ 之裂縫，其 Resize 影像中裂縫

座標為 $(x'_1, y'_1) = (453.1, 328.9)$, $(x'_2, y'_2) = (453, 335.2)$:

$$w = 13.3\text{mm}; w_p = 5472\text{pixel}; h_p = 3648\text{pixel};$$

$$w'_p = 912\text{pixel}; h'_p = 912\text{pixel};$$

$$WD = 18000\text{mm}; f = 10.7\text{mm}$$

$$x_1 = 453.1 \times \frac{5472}{912} = 2718.6; y_1 = 328.9 \times \frac{3648}{912} = 1315.6$$

$$x_2 = 453 \times \frac{5472}{912} = 2718; y_2 = 335.2 \times \frac{3648}{912} = 1340.8$$

$$np_\gamma = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \approx 25.2$$

$$25.2 \times \frac{13.3}{5472} \approx 0.06125 \text{ (mm)}$$

$$0.06125 \times \frac{18000}{10.7} \approx 103.03\text{mm}$$

縮小影像中裂縫影像座標轉換為原始影像中裂隙列座標，最

後，實際 100mm 裂縫寬經演算法估算結果為 103.03mm。

本議題結論:經計算 60mm 裂縫和 100mm 裂縫分別估計為 63.77mm 和 103.03mm，誤差約為 4mm 其在可接受範圍內，此方法可以使用於預測裂縫寬度。

二、UAV巡檢航線參數設定與巡檢歷程影像資料庫建置之研究

(一) UAV 巡檢路線歷程設計規劃

本議題模擬勘察水土保持護岸設施巡檢實驗場域，請教專家(中興大學水土保持系教授)提供實驗場域及巡檢路線，選定於埔里眉溪觀音吊橋附近護岸，拍攝本計畫之水土保持設施裂隙識別所需 UAV 影像，模擬適當 UAV 巡檢參數 (例如：高度、標的設施水平距離、拍攝間格與拍攝角度等)。相關影像已呈現在**附錄三**。圖 2-8 說明資料庫儲存影像命名格式。

檔名: <YYYY>_<MM>_<DD>_<hr>_<min>_<sec>.png

範例

時間: 2022/11/01

位置: 24.04184, 120.69902

儲存檔名: 2022_11_01_18_06_12.png

圖 2-8: 影像命名規則說明

(二) 提取 UAV 影像描述資訊

當 UAV 影像拍攝後，經由程式提取影像的拍攝時間、拍攝高度、裝置紀錄的度分秒格式座標經緯度(DMS)與轉換後的十進位格式座標經緯度(DD)，最後匯出成單次巡檢的 CSV 表格，如圖 2-9 所示。圖 2-10 說明影像儲存的架構，第一層資料夾為拍攝地點名稱，可以人工提供，或由程式將影像彙整至紀錄中最接近的地點，第二層為拍攝日期，資料夾裡存放當日拍攝的影像。最後經由**議題(四)**的 API 程式，提供使用者查詢服務。

	A	B	C	D	E	F	G
1	ori_img_name	DateTime	lat	lng	lat_ori	lng_ori	altitude
2	IMG_3345.JPG	2023:04:03 16:17:41	24.70138611	121.3739917	(24.0, 42.0, 4.99)	(121.0, 22.0, 26.37)	1147.844
3	IMG_3346.JPG	2023:04:03 16:17:42	24.70138611	121.3739917	(24.0, 42.0, 4.99)	(121.0, 22.0, 26.37)	1147.844
4	IMG_3347.JPG	2023:04:03 16:17:43	24.70138611	121.3739917	(24.0, 42.0, 4.99)	(121.0, 22.0, 26.37)	1147.844
5	IMG_3348.JPG	2023:04:03 16:17:44	24.70138611	121.3739917	(24.0, 42.0, 4.99)	(121.0, 22.0, 26.37)	1147.844
6	IMG_3349.JPG	2023:04:03 16:17:44	24.70138611	121.3739917	(24.0, 42.0, 4.99)	(121.0, 22.0, 26.37)	1147.844

圖 2-9: 單次巡檢 CSV 表格

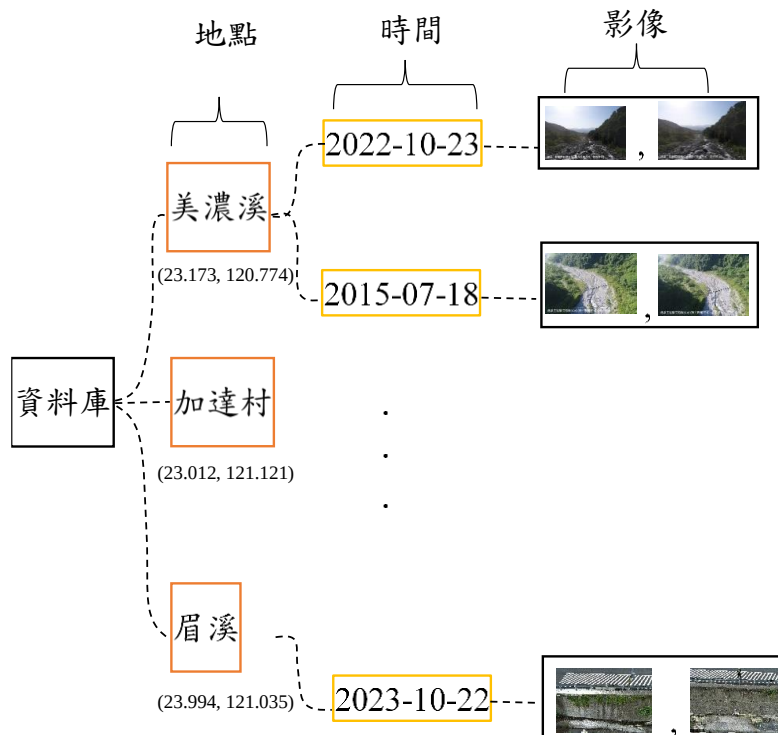


圖 2-10: UAV 影像儲存架構

(三) 建立水土保持護岸設施巡檢歷程資料庫

本議題模擬三次巡檢，已所蒐集的影像來建置水土保持設施護岸歷程資料庫，本議題也已設計 API 程式，提供介接取得資料庫的影像。圖 2-11 說明本議題影像資料庫架構，以同一巡檢路線之歷程影像，針對巡檢路線中 n 個拍攝紀錄點，持續追蹤紀錄點之設施影像內容，其中色塊代表 UAV 拍攝影像。圖 2-12 說明歷程資料庫架構，紅色虛框表示同一次巡檢拍攝 UAV 影像，

藍色虛框表示同紀錄點(近似)不同次巡檢拍攝 UAV 影像。

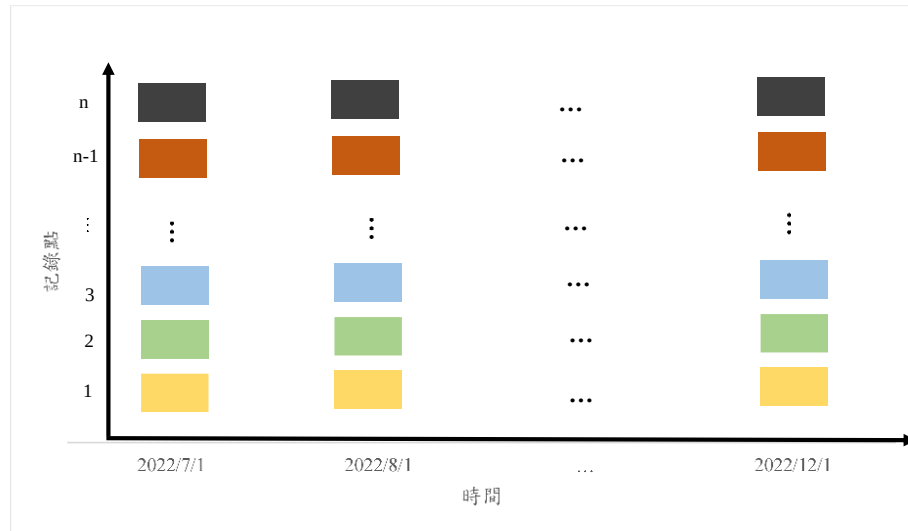


圖 2-11: UAV 實驗影像關係概念

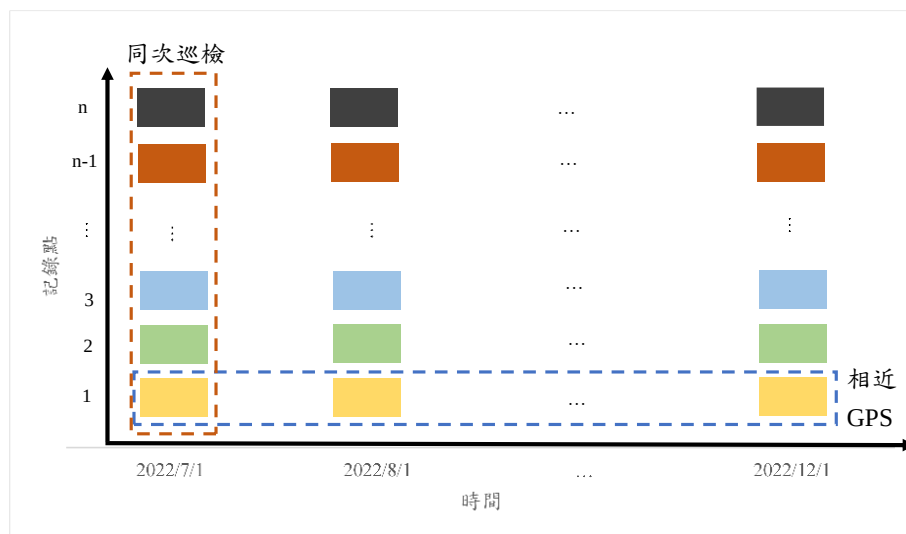


圖 2-12: API 資料庫概念。(紅框)同次巡檢拍攝(藍框)相近 GPS 拍攝影像

(四) 水土保持護岸設施巡檢模擬實驗

本議題使用附錄三說明的 Autel 無人機，實作議題(一)到議題(三)設計，模擬巡檢歷程拍攝，並以埔里觀音橋附近的護岸為

巡檢實驗場域。圖 2-13 說明模擬任務巡檢拍攝 10 個紀錄點，並選取部分影像至於表格中，其中第一次飛行使用手動拍攝影像，第二次與第三次為無人機以第一次飛行參數，例如：巡航高度、拍攝間隔、標的物距離等。

圖 2-14 說明任務錄製與 2 次復現 GPS 軌跡圖，其中橫軸為經度，縱軸為緯度。藍線為手動飛行之 GPS 軌跡，綠線與紅線為使用無人機任務錄製復現 GPS 軌跡，黑點表示校準點，懸停在定點但不拍攝影像，從橘點才開始拍攝紀錄。在實驗中，可以觀察到手動拍攝、第二次與第三次的復現軌跡相當靠近，透過計算同一記錄點的經緯度歐式距離誤差，在 $5.58 \times 10^{-6} \sim 30.1 \times 10^{-6}$ 區間，估算真實距離大約在 70~300cm。未來建立巡檢歷程資料庫，建議記錄點間隔 5m，以利系統區隔不同檢查點拍攝影像。

本議題結論：因拍攝環境影響變因，導致不同次巡檢在同一紀錄點之 GPS 定位座標取得產生誤差，但影像中內容可能大致相同。因此未來建立巡檢歷程資料庫，可以採用以下巡檢參數的來設定空拍機巡檢參數，其偵測標的設施水平距離 10m，垂直 10m，拍攝記錄點間隔 5m。

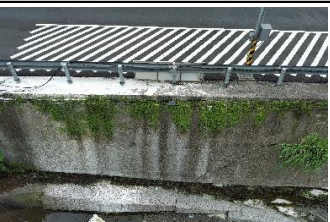
記錄點	第 1 次巡檢	第 2 次巡檢	第 3 次巡檢
1			
2			
4			
6			
8			
9			
10			

圖 2-13: 眉溪觀音橋護岸模擬巡檢影像

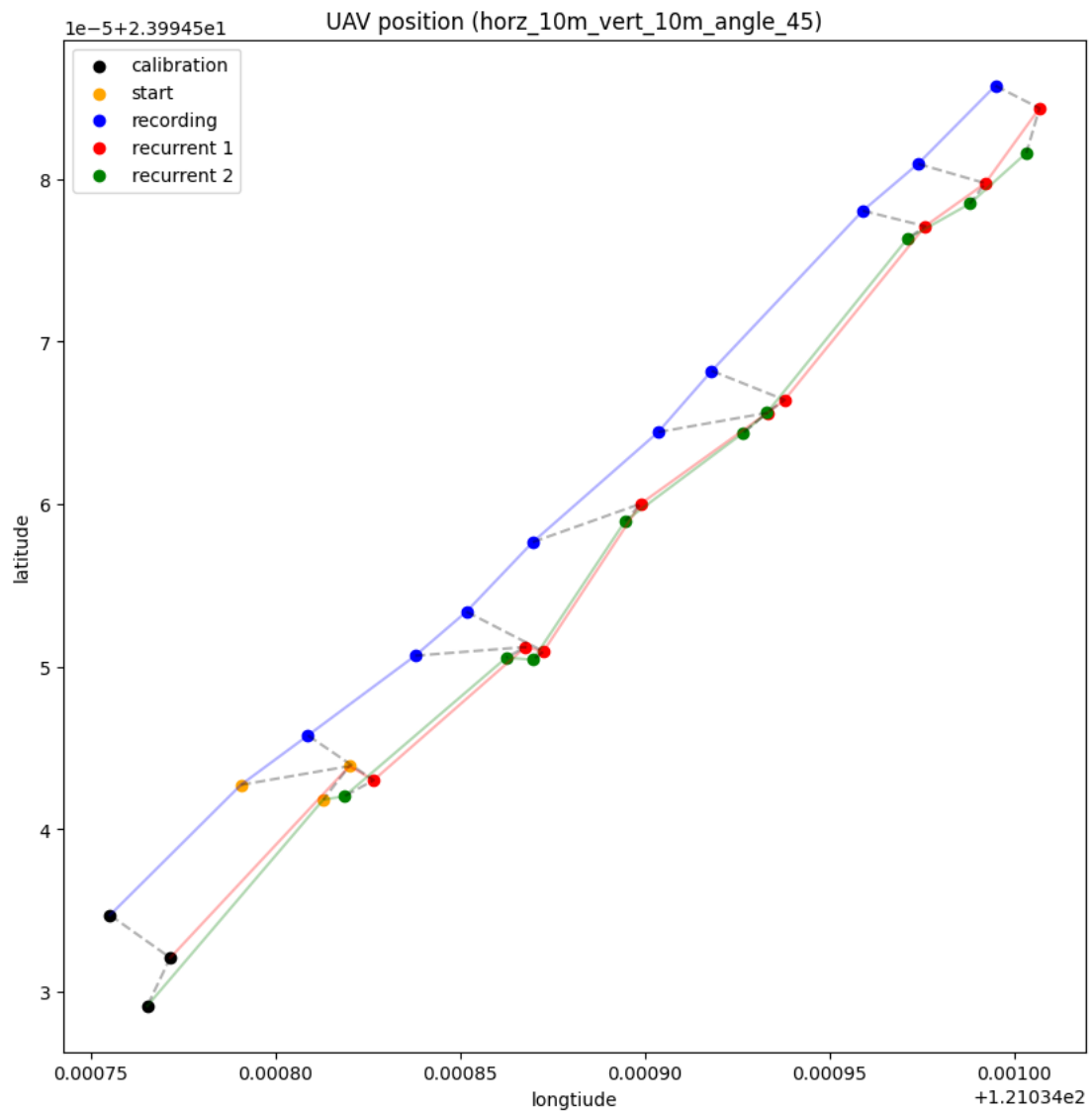


圖 2-14:模擬巡檢 GPS 飛行軌跡(藍)手動拍攝(紅)第一次復現飛行軌跡(綠)第二次復現飛行軌跡

三、多尺度影像分割應用於偵測UAV影像混凝土表面裂隙之研究

(一) 多尺度影像分割模型設計

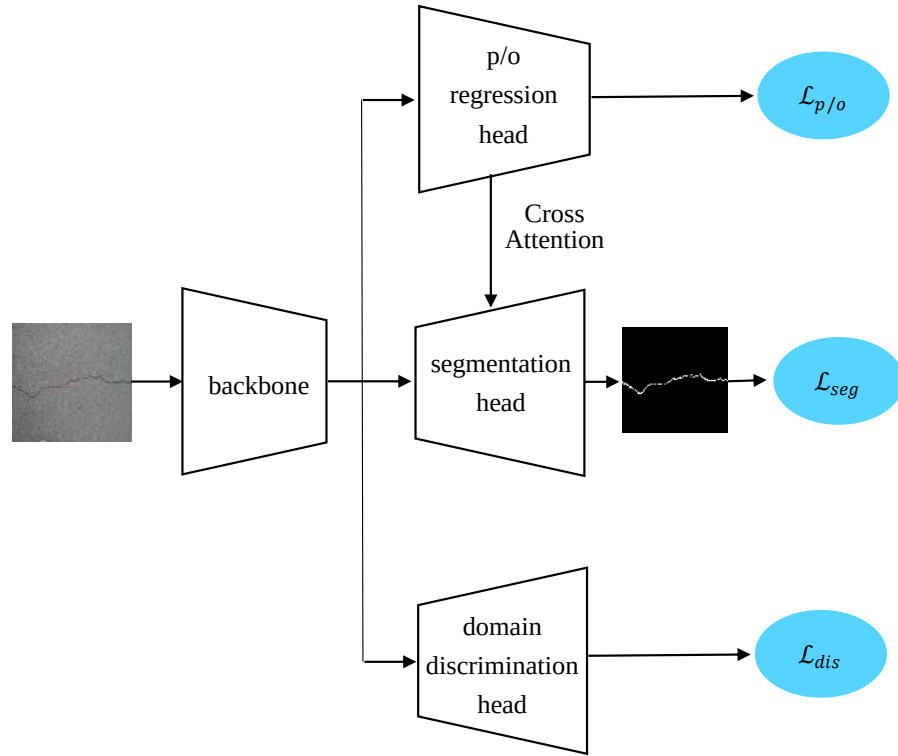


圖 2-15：p/o 率模型概念架構圖。

圖 2-15 展示本主題之模型架構，此架構藉由學習影像解析度特徵提升影像切割效能。為了方便量化解析度，在此主題中解析度資訊將以 p/o 率衡量。此架構包含三個分支：p/o regression head、segmentation head 與 domain discrimination head，影像先經由 backbone 提取特徵，再分別送入 p/o regression head、segmentation head 與 domain discrimination head 分別預測 p/o 率、影像分割預測與 domain 判別預測。其中 $\mathcal{L}_{seg}$ 為影像分割任務的

損失函數、 $\mathcal{L}_{p/o}$ 為 p/o 率 regression 損失函數、 $\mathcal{L}_{dis}$ 為領域自適應損失函數。在此定義 backbone、p/o regression head、segmentation head 與 domain discrimination head 分別為 $\mathcal{E}_{enc}$ 、 $\mathcal{D}_{p/o}$ 、 $\mathcal{D}_{seg}$ 與 $\mathcal{D}_{dis}$ 。

(1) 定義 p/o 率



(a) 原圖、p/o 率大



(b) 縮圖、p/o 率小



(c) 裁切子圖、p/o 率大

圖 2-16：不同 p/o 率之影像範例

p/o 率是 pixel per object 的縮寫，其定義是一張影像中的某物件(裂隙)由多少像素所描述。圖 2-16 展示不同 p/o 率之影像，

圖 2-16 (a)為解析度為 384×256 之影像、圖 2-16 (b)為將圖 2-16 (a)縮小至解析度 192×128 之影像、圖 2-16 (c)為將圖 2-16 (a)裁切至 384×256 解析度之影像。由此例可得知影響 p/o 率的因素並非解析度，而是物件(裂縫)在影像中的占比。故本主題提出透過學習不同 p/o 率的影像以提取多尺度特徵表示，並應用於裂隙之影像切割任務。

(2)損失函數： $\mathcal{L}_{seg}$ 、 $\mathcal{L}_{p/o}$ 與 $\mathcal{L}_{dis}$

此議題之損失函數包含兩個： $\mathcal{L}_{seg}$ 與 $\mathcal{L}_{p/o}$ 。其中 $\mathcal{L}_{seg}$ 為傳統影像切割之損失函數，在此議題中使用 focal Tversky loss，其定義如下：

$$\begin{aligned}\mathcal{L}_{seg} &= \left(\frac{TP}{TP + \alpha FP + \beta FN} \right)^\gamma \\ TP &= \sum_i y_i \hat{y}_i \\ FP &= \sum_i (1 - y_i) \hat{y}_i \\ FN &= \sum_i y_i (1 - \hat{y}_i)\end{aligned}\tag{3-1}$$

其中 y 為 ground truth segmentation label、 $\hat{y}$ 是預測 segmentation mask、 i 是座標位置、超參數設置為 $\alpha = 0.7, \beta = 0.3, \gamma = 0.75$ 。

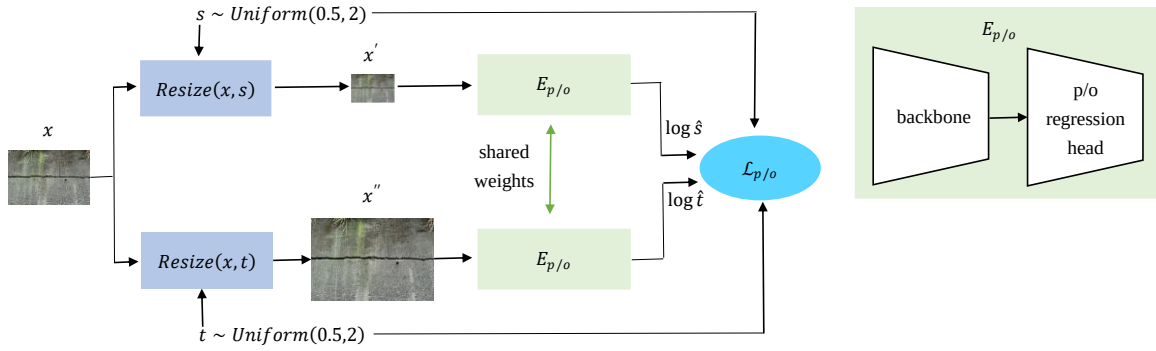


圖 2-17：p/o 率回歸預測架構圖

根據議題(1)所敘述，雖然 p/o 率能夠合適地描述物件在影像中的大小，然而相較於解析度，p/o 率之定義更為抽象且難以衡量。故本議題透過將同一張影像調整大小來衡量 p/o 率的相對關係。圖 2-17 說明其概念流程。其中 x 為訓練影像、 s, t 控制 x 調整大小後的尺寸、 x', x'' 分別為 x 調整 s 與 t 倍大小後之影像。接著 x', x'' 經過 p/o 率回歸類神經網路 $E_{p/o}$ 預測得到 $\log \hat{s}$ 與 $\log \hat{t}$ ，其中 $E_{p/o}$ 由 backbone 與 p/o regression head 組合而成。最後，利用 $s, t, \log \hat{s}$ 與 $\log \hat{t}$ 計算 p/o 率回歸損失函數 $\mathcal{L}_{p/o}$ ，其中 $\mathcal{L}_{p/o}$ 的定義如式(3-2)。

$$\begin{aligned}\mathcal{L}_{po} &= \|d - \hat{d}\|^2 \\ d &= \log s - \log t \\ \hat{d} &= \log \hat{s} - \log \hat{t}\end{aligned}\tag{3-2}$$

此損失函數透過隱含地擬和兩張影像之 p/o 率的對數差來學習 p/o 率之特徵表示，並且不需要人工標註 p/o 率。為了減緩

訓練資料與測試資料的領域差距(domain gap)，本議題使用對抗式領域自適應方法，其詳細說明參照(Ganin et al., 2015; Ganin et al.,2016)。令來源資料集與目標資料集分別為 $D_s = (X_s, Y_s)$ 與 $D_t = X_t$ ，其中 X_s 、 Y_s 與 X_t 分別為來源影像集合、來源標註集合與目標影像集合。領域自適應的損失函數定義如下：

$$\mathcal{L}_{dis}(x) = \begin{cases} -\log \hat{d} & , \text{if } d = 1 \\ -\log(1 - \hat{d}) & , \text{if } d = 0 \end{cases} \quad (3-3)$$

$$d = \begin{cases} 0, & \text{if } x \in X_s \\ 1, & \text{if } x \in X_t \end{cases}$$

$$\hat{d} = \mathcal{D}_{dis}(\mathcal{E}_{enc}(x))$$

其中 d 與 $\hat{d}$ 分別為領域標註(domain label)以及領域預測(domain prediction)。 $\hat{d}$ 透過將 backbone 提取的特徵經過 domain discrimination head 預測其所屬領域。此外，當 $d = 0$ 時表示影像 x 來自來源領域(source domain)、當 $d = 1$ 時表示影像 x 來自目標領域(target domain)。至於對抗式領域自適應的最佳化目標函數定義如下：

$$\max_{\mathcal{E}_{enc}} \min_{\mathcal{D}_{dis}} \mathcal{L}_{dis} \quad (3-4)$$

式(3-4)希望 $\mathcal{D}_{dis}$ 能夠正確判別領域標註而 $\mathcal{E}_{enc}$ 能夠學習到足以混淆 $\mathcal{D}_{dis}$ 的特徵表示。

(3)模型架構

上述所提及的 backbone 使用 SegFormer (Xie et al., 2021) 所提出的 MiT-B3，為輕量且適合用於進行影像分割之 transformer encoder。Segmentation head 以及 p/o regression head 則是根據 SegFormer 的 decoder 進行小幅改動。圖 2-18 說明 p/o 率回歸與語意分割模型的模型架構。

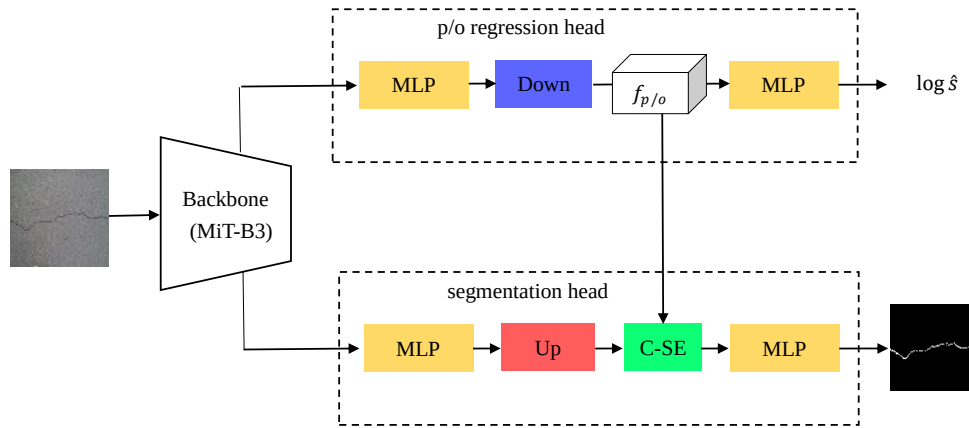


圖 2-18：p/o 率回歸與語意分割模型架構圖

在 p/o regression head 中間的特徵 $f_{p/o}$ 經過 Cross Squeeze and Excitation (C-SE) 模組融合至 segmentation head，使得語意分割預測得以利用到 p/o 率的特徵表示。

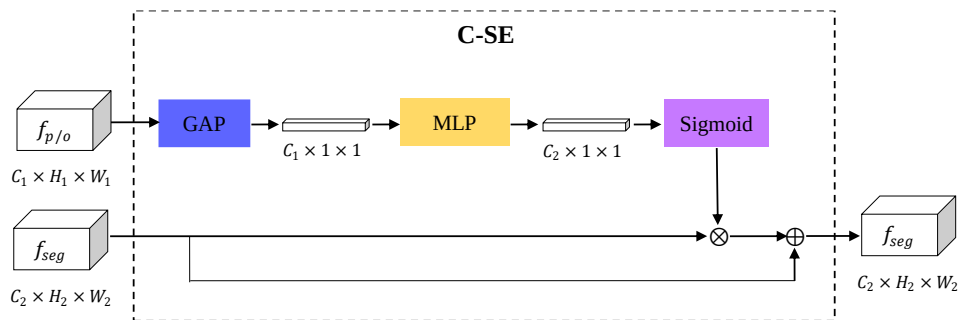


圖 2-19：C-SE 模組架構圖

圖 2-19 說明 C-SE 模組修改自 SENet (Hu et al., 2018) 的 SE 模組。此模組將 $f_{p/o}$ 與語意分割中間特徵 f_{seg} 進行融合。首先 $f_{p/o}$ 經過 global average pooling (GAP) 轉換成特徵向量。接著，經過 MLP 將通道數量轉換成與 f_{seg} 一致。最後，經過 sigmoid 對 f_{seg} 之通道實施注意力機制。透過此方法， $f_{p/o}$ 將指示 f_{seg} 輸入影像的 p/o 資訊，進而達到多尺度影像語意分割。

(4) 訓練演算法

本議題的演算法之定義如下。首先，根據式(3-2-4)，3-11 行先得到最佳的 $\mathcal{D}_{dis}$ 。接著，12-17 行計算用於更新 $\mathcal{E}_{enc}$ 的 $\mathcal{L}_{dis}$ 。再來，18-23 行計算 p/o 率的損失函數 $\mathcal{L}_{p/o}$ 、24-27 行計算語意分割的損失函數 $\mathcal{L}_{seg}$ 。最後，計算整體的損失函數並更新 $\mathcal{E}_{enc}$ 、 $\mathcal{D}_{p/o}$ 與 $\mathcal{D}_{seg}$ 的權重。Algorithm 3-1 表示訓練 $\mathcal{E}_{enc}$ 、 $\mathcal{D}_{p/o}$ 與 $\mathcal{D}_{seg}$ 的流程。

Algorithm 3-1. The training algorithm**Input:** source domain D_s , target domain D_t ;**Output:** model $\{\mathcal{E}_{enc}, \mathcal{D}_{p/o}, \mathcal{D}_{seg}, \mathcal{D}_{dis}\}$;

```

1. Initialize weights of  $\mathcal{E}_{enc}, \mathcal{D}_{p/o}, \mathcal{D}_{seg}, \mathcal{D}_{dis}$ ;
2. For  $i = 1, \dots, N_{iter}$  do //  $N_{iter}$  is the number of iterations
3.   // The domain adaptation part: train  $\mathcal{D}_{dis}$ 
4.   For  $j = 1, \dots, N_{dis}$  do //  $N_{dis}$  is the total steps for updating  $\mathcal{D}_{dis}$ 
5.      $(x_s, y_s) \sim D_s$ ;
6.      $x_t \sim D_t$ ;
7.      $\hat{d}_s \leftarrow \mathcal{D}_{dis}(\mathcal{E}_{enc}(x_s))$ ;
8.      $\hat{d}_t \leftarrow \mathcal{D}_{dis}(\mathcal{E}_{enc}(x_t))$ ;
9.     Compute  $\mathcal{L}_{dis}$  according to (3-2-3);
10.    Update weights of  $\mathcal{D}_{dis}$  by minimizing  $\mathcal{L}_{dis}$ ;
11.  end
12.  // The domain adaptation part: train  $\mathcal{E}_{enc}$ 
13.   $(x_s, y_s) \sim D_s$ ;
14.   $x_t \sim D_t$ ;
15.   $\hat{d}_s \leftarrow \mathcal{D}_{dis}(\mathcal{E}_{enc}(x_s))$ ;
16.   $\hat{d}_t \leftarrow \mathcal{D}_{dis}(\mathcal{E}_{enc}(x_t))$ ;
17.  Compute  $\mathcal{L}_{dis}$  according to (3-2-3);
18.  // The p/o ratio part
19.   $x \sim X_s \cup X_t$ ;
20.   $s, t \sim Uniform(0.5, 2)$ ;
21.   $x', x'' \leftarrow resize(x, s), resize(x, t)$ ;
22.   $\log \hat{s}, \log \hat{t} \leftarrow \mathcal{D}_{p/o}(\mathcal{E}_{enc}(x')), \mathcal{D}_{p/o}(\mathcal{E}_{enc}(x''))$ ;
23.  Compute  $\mathcal{L}_{p/o}$  according to (3-2-2);
24.  // The segmentation part
25.   $(x_s, y_s) \sim D_s$ ;
26.   $\hat{y}_s \leftarrow \mathcal{D}_{seg}(\mathcal{E}_{enc}(y_s))$ ;
27.  Compute  $\mathcal{L}_{seg}$  according to (3-2-1);
28.  // Compute total loss and update the weights
29.   $\mathcal{L} \leftarrow \mathcal{L}_{seg} + \mathcal{L}_{p/o} - \mathcal{L}_{dis}$ ;
30.  Update weights of  $\mathcal{E}_{enc}, \mathcal{D}_{p/o}, \mathcal{D}_{seg}$  by minimizing  $\mathcal{L}$ ;
31. end

```

(二) 實驗影像集說明

本議題所採用的訓練資料分成來源資料集以及目標資料集，

其中來源資料及取自(khanhha, 2020)所整理的裂隙語意分割資

料集、目標資料集取自歷史影像平台(ARDSWC, 2015)並人工

標註其裂隙語意分割標註。其中訓練來源資料集與驗證來源資

料集分別包含 7988 與 1695 筆影像-語意分割標註配對、訓練目標資料集與驗證目標資料集分別包含 609 與 152 筆影像-語意分割標註配對。測試資料集來自眉溪空拍影像，其詳細敘述請參照附錄三。

(三) 空拍影像測試資料集實驗結果

表 2-3：於空拍影像測試資料集上的效能衡量

測試資料集 \ 衡量指標	precision	recall	F1-score	mIoU	accuracy
水平裂縫_2~3m (115 張)	0.692	0.487	0.534	0.681	0.981
水平拉距_距起降點 7m (48 張)	0.714	0.339	0.455	0.665	0.989
下游斜拍_起降點高 6m (22 張)	0.636	0.425	0.504	0.645	0.994
垂直拉距_起降點高 10m (43 張)	0.559	0.229	0.313	0.591	0.995

表 2-3 顯示模型在不同角度下所拍攝之空拍影像測試資料集的效能。模型在「水平裂縫_2~3m」上的表現是最好的，這是因為在這種尺度下所拍攝的影像與訓練資料集相近。以上衡量指標中，precision 為本議題欲關注的重點。因為由於訓練資料集往往是近距離對裂隙拍攝，故不存在除了裂隙與水泥表面以外的特徵。以上問題使得模型容易將樹枝、雜草與苔癬等物件誤判成裂隙，導致 precision 表現低下。然而在本議題的實驗結果中，precision 依然保持著不錯的水準，表示模型有能力保證其偵測出的裂隙確實是裂隙，而不是其他物件。圖 2-20、圖 2-21、圖 2-22 與圖 2-23 分別展示不同角度下所拍攝之空拍範

例影像，左欄為輸入影像、中間是估計、右欄是標註。

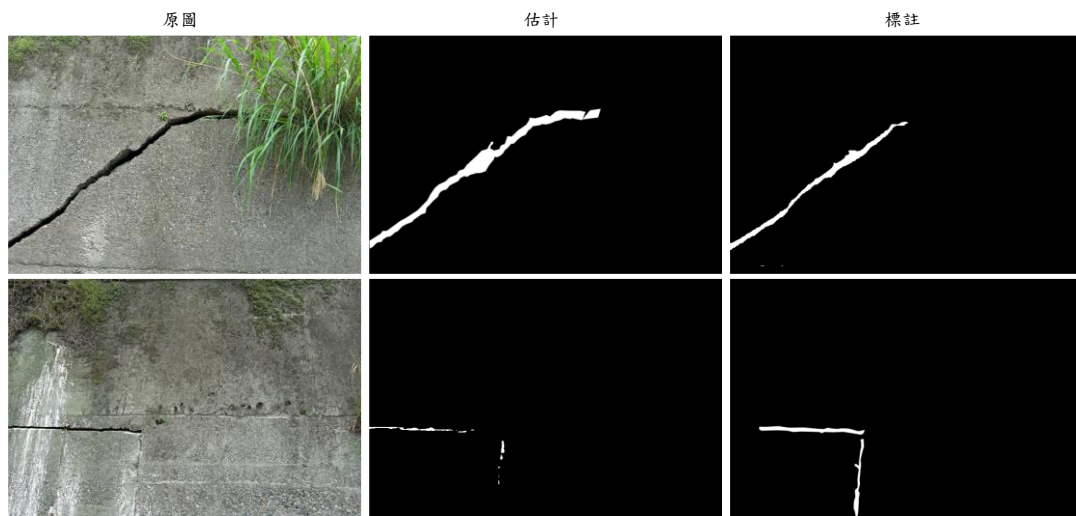


圖 2-20：水平裂縫_2~3m 範例影像

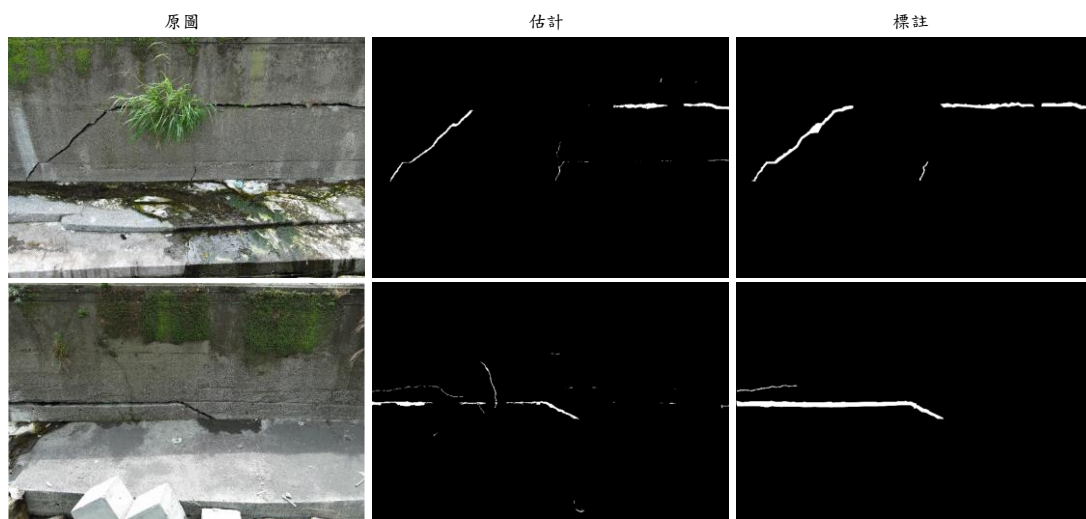


圖 2-21：水平拉距_距起降點 7m 範例影像



圖 2-22：下游斜拍_起降點高 6m 範例影像



圖 2-23：垂直拉距_起降點高 10m 範例影像

(四) 模型效能比較

(1) 與 111 年模型效能比較

圖 2-24 展示 111 年模型在本議題之測試資料集的效能比較，其中圖 2-24 (a-d)分別為 precision、recall、f1-score 與 mIoU，藍、黃、綠與紅分別為今年之多尺度領域自適應 UAV 影像語意分割

(多尺度分割)、111 年之兩階段影像裂隙分割(兩階段分割)、111 年之領域自適應裂隙影像分割(自適應分割)與 111 年之領域自適應裂隙影像偵測(自適應偵測)。大致上今年的模型效能相較於 111 年皆有顯著的提升。至於在圖 2-24 (a)、(c)與(d)中，可發現今年之多尺度分割在「水平裂縫_2~3m」上略遜於 111 年之兩階段分割。然而在其他尺度下的測試資料集依然能勝過 111 年之兩階段分割，這是因為今年之多尺度分割利用多尺度學習策略，提升了模型在多尺度影像上的泛化能力。

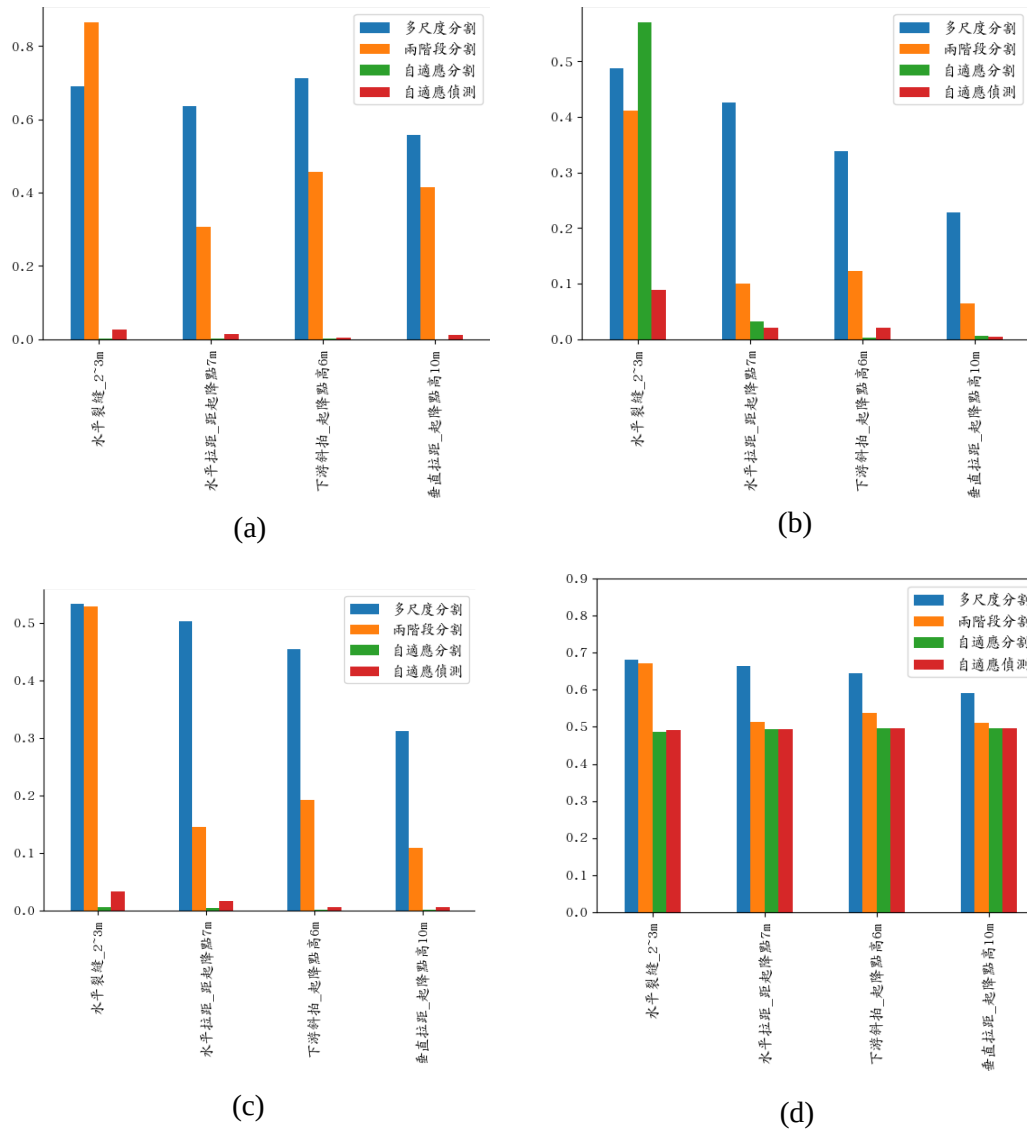


圖 2-24：與 111 年之效能比較

(2)於模擬干擾下之實驗結果

表 2-4：於昏暗空拍影像測試資料集上的效能衡量

測試資料集 \ 衡量指標	precision	recall	F1-score	mIoU	accuracy
水平裂縫_2~3m (115 張)	0.621	0.402	0.434	0.643	0.978
水平拉距_距起降點 7m (48 張)	0.753	0.288	0.413	0.627	0.994
下游斜拍_起降點高 6m (22 張)	0.631	0.364	0.457	0.645	0.989
垂直拉距_起降點高 10m (43 張)	0.553	0.215	0.298	0.587	0.995

表 2-5：於起霧空拍影像測試資料集上的效能衡量

測試資料集 \ 衡量指標	precision	recall	F1-score	mIoU	accuracy
水平裂縫_2~3m (115 張)	0.559	0.4	0.382	0.62	0.975
水平拉距_距起降點 7m (48 張)	0.671	0.298	0.409	0.627	0.994
下游斜拍_起降點高 6m (22 張)	0.638	0.361	0.453	0.644	0.988
垂直拉距_起降點高 10m (43 張)	0.497	0.203	0.271	0.579	0.994

表 2-6：於下雨空拍影像測試資料集上的效能衡量

測試資料集 \ 衡量指標	precision	recall	F1-score	mIoU	accuracy
水平裂縫_2~3m (115 張)	0.62	0.415	0.418	0.639	0.977
水平拉距_距起降點 7m (48 張)	0.734	0.265	0.384	0.617	0.994
下游斜拍_起降點高 6m (22 張)	0.635	0.303	0.389	0.62	0.988
垂直拉距_起降點高 10m (43 張)	0.556	0.217	0.3	0.588	0.995

表 2-4 表 2-5 表 2-6 分別展示將測試資料集進行三種模擬干擾後的實驗結果，其中表 2-4 為昏暗，表 2-5 為起霧表 2-6 為下雨，在結果中可以發現在起霧的干擾下，模型效能在 precision 的表現下降得最多，表示模型對抗起霧干擾的能力較為低下。圖 2-25 展示視覺化模擬干擾比較，其中圖 2-25 (a-d)分別為 precision、recall、f1-score 與 mIoU，藍、黃、綠與紅分別表示原始影像、昏暗干擾、起霧干擾與下雨干擾。由圖 2-25 中發現效能不因干擾產生顯著降低，表示模型對以上干擾皆有較強的泛化能力。

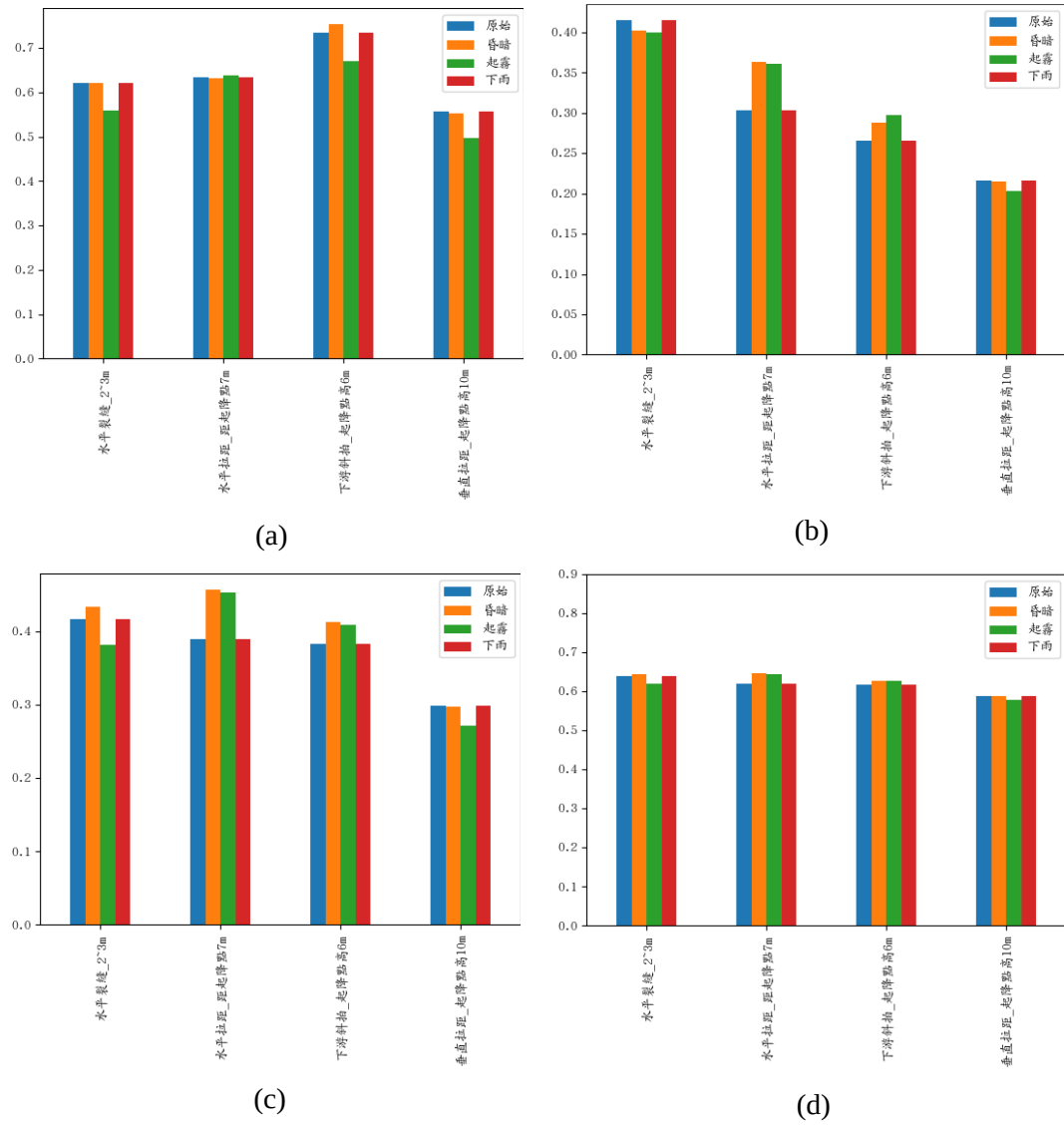


圖 2-25：視覺化模擬干擾比較

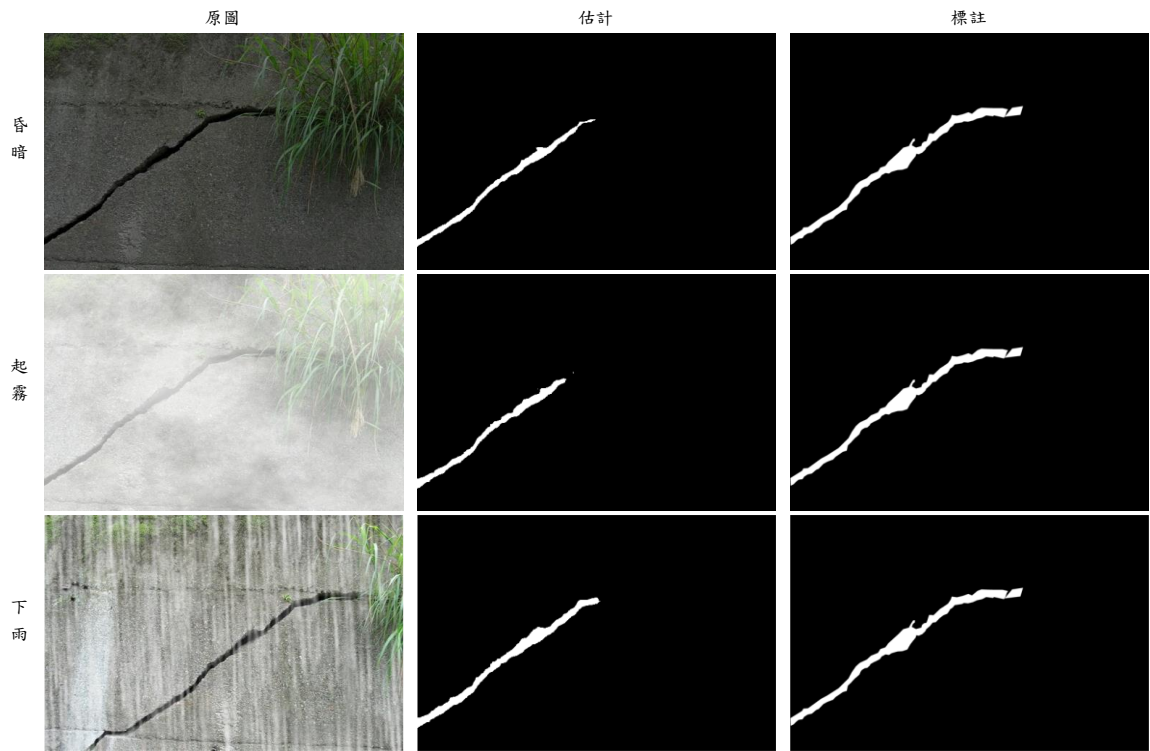


圖 2-26：水平裂縫_2~3m 範例影像

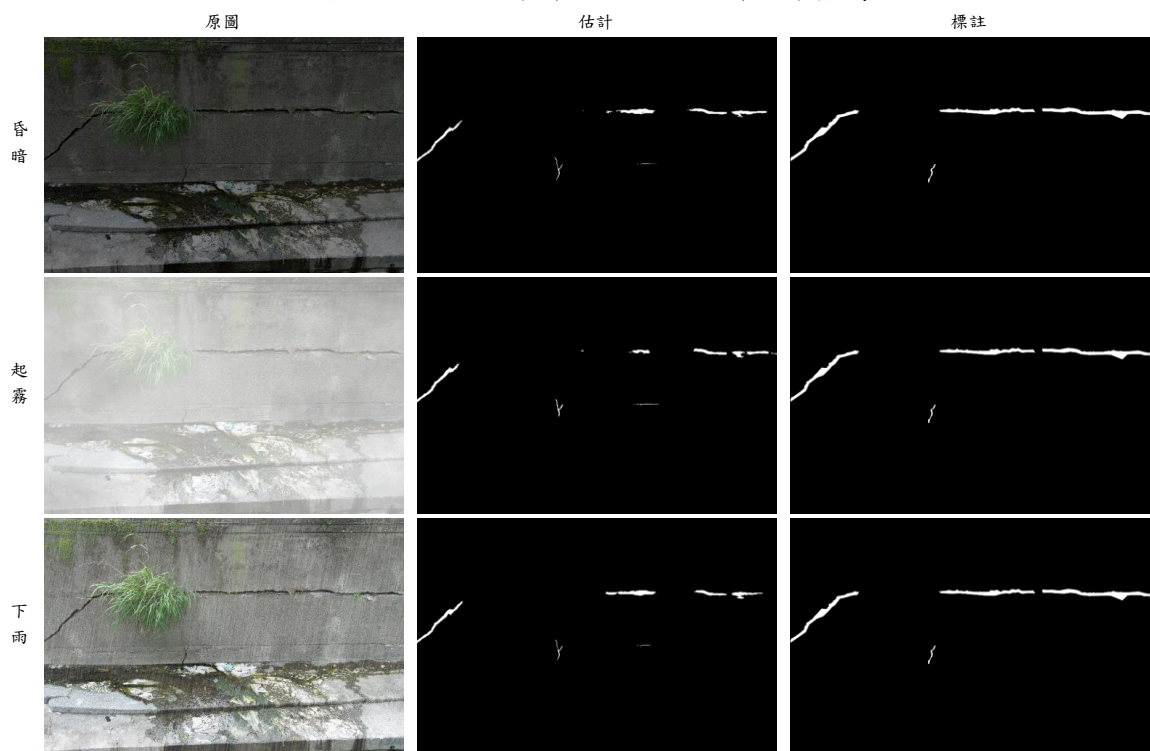


圖 2-27：水平拉距_距起降點 7m 範例影像



圖 2-28：下游斜拍_起降點高 6m 範例影像

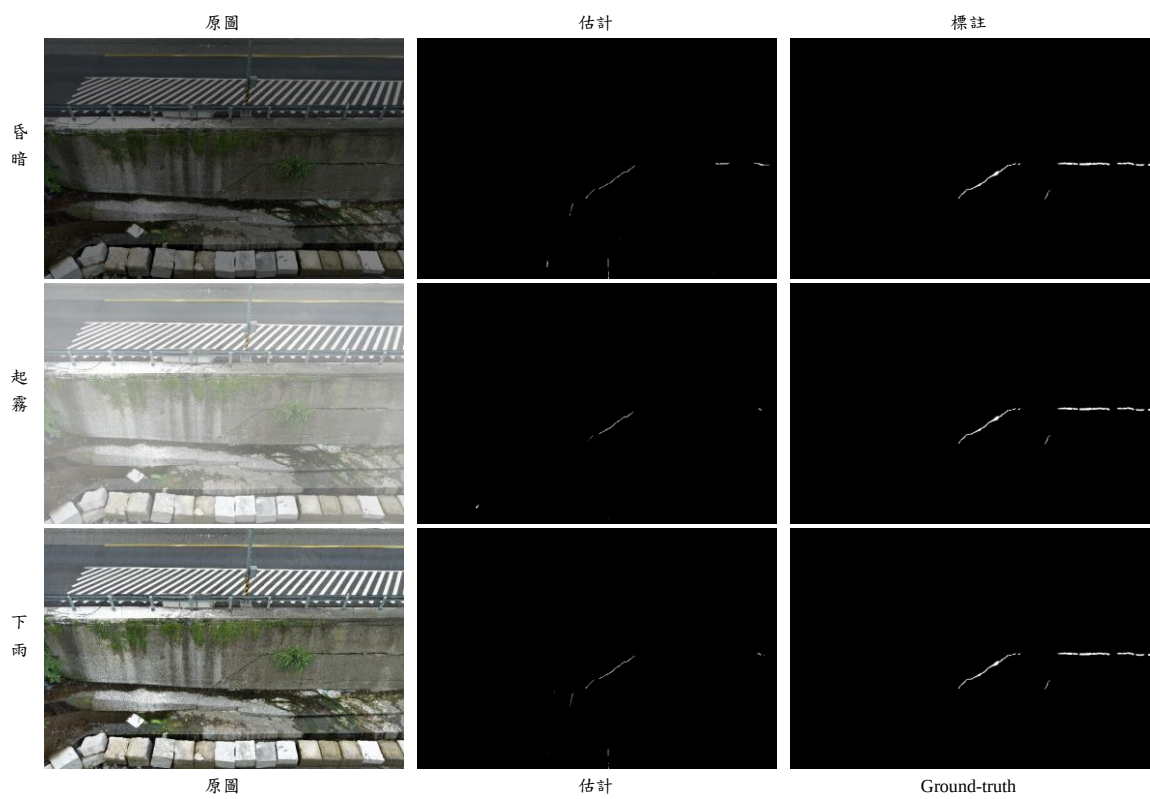


圖 2-29：垂直拉距_起降點高 10m 範例影像

圖 2-27、圖 2-28 及圖 2-29 是模擬干擾測試資料集在三種干擾下的範例影像，其中左欄是影像、中間是預測、右欄是標註。上、中、下列分別是昏暗、起霧與下雨的干擾模擬。

(3)其他多尺度模型效能比較

表 2-7 中，左邊 Models 顯示其他多尺度模型的名稱。最上方一系列的格式為 X-Y，其中 $X \in \{P, R, F\}$ 表示衡量指標、P 表示 precision、R 表示 recall、F 表示 f1-score、 $Y \in \{2, 7, 6, 10\}$ 表示測試資料集、2 表示「水平裂縫_2~3m」、7 表示「水平拉距_距起降點 7m」、6 表示「下游斜拍_起降點高 6m」、10 表示「垂直拉距_起降點高 10m」。

表 2-7：其他多尺度模型之效能比較

Models	P-2	R-2	F-2	P-7	R-7	F-7	P-6	R-6	F-6	P-10	R-10	F-10
deeplabv3	0.655	0.334	0.42	0.485	0.342	0.372	0.62	0.322	0.415	0.206	0.163	0.166
hrnet32	0.684	0.302	0.41	0.3	0.183	0.202	0.395	0.144	0.201	0.05	0.086	0.059
segformer	0.587	0.362	0.416	0.445	0.334	0.367	0.547	0.296	0.368	0.409	0.261	0.302
proposed	0.692	0.487	0.534	0.714	0.339	0.455	0.636	0.425	0.504	0.559	0.229	0.313

表 2-7 中，左邊 Models 顯示其他多尺度模型的名稱。最上方一系列的格式為 X-Y，其中 $X \in \{P, R, F\}$ 表示衡量指標、P 表示 precision、R 表示 recall、F 表示 f1-score、 $Y \in \{2, 7, 6, 10\}$ 表示測試資料集、2 表示「水平裂縫_2~3m」、7 表示「水平拉距_距起降點 7m」、6 表示「下游斜拍_起降點高 6m」、10 表示「垂直

拉距_起降點高 10m」。表 2-7 顯示，本方法在大部分指標上皆能勝過其他多尺度模型。

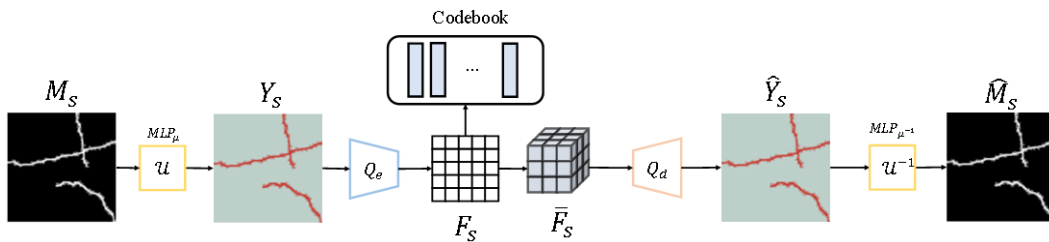
四、生成式混凝土表面影像分割應用於偵測UAV影像混凝土表面裂隙之研究

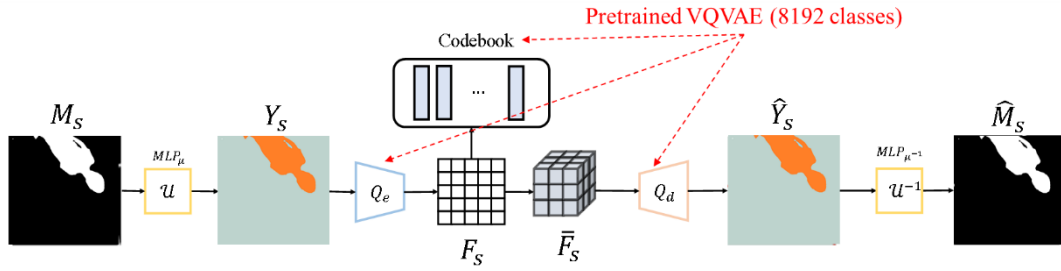
(一) 裂縫分割模型建模方法

(1) Maskiage 和重建 mask

一個 pretrained 的 VQVAE 模型包含了編碼器 Q_e 和解碼器 Q_d ，其輸入必須為一張 RGB 影像。因為 mask 不是 RGB 影像，無法作為 VQVAE 的輸入，因此本議題需要學習一個 MLP_μ 將一張有 k 個通道的 mask 轉換為一張 RGB 的 mask (maskiage) 使得 mask 可以輸入到 VQVAE，解碼出來的影像再經過另一個 $MLP_{\mu^{-1}}$ 將生成的 RGB 影像轉回為 mask。圖 2-30 為學習 MLP_μ 和 $MLP_{\mu^{-1}}$ 之架構。 M_s 為源影像(來自 kaggle)的 mask， M_s 透過 MLP_μ 轉為 maskiage(Y_s)。接著，將 Y_s 輸入到 Q_e 得到特徵 F_s ， F_s 每個位置的特徵被對應到編碼簿(codebook)裡計算出最相似的特徵，取代 F_s 以得到特徵 $\bar{F}_s$ 。最後，將 $\bar{F}_s$ 輸入到 Q_d 得到預測出來的 $\hat{Y}_s$ ，經過 $MLP_{\mu^{-1}}$ 後得到預測裂縫的 $\hat{M}_s$ 。透過損失函數公式(4-1)來最佳化訓練 MLP_μ 和 $MLP_{\mu^{-1}}$ 。

$$\min_{\mu, \mu^{-1}} \|\hat{M}_s - M_s\|_2 \quad (4-1)$$



圖 2-30: 學習 μ 和 μ^{-1} 之架構

(2)特徵對應

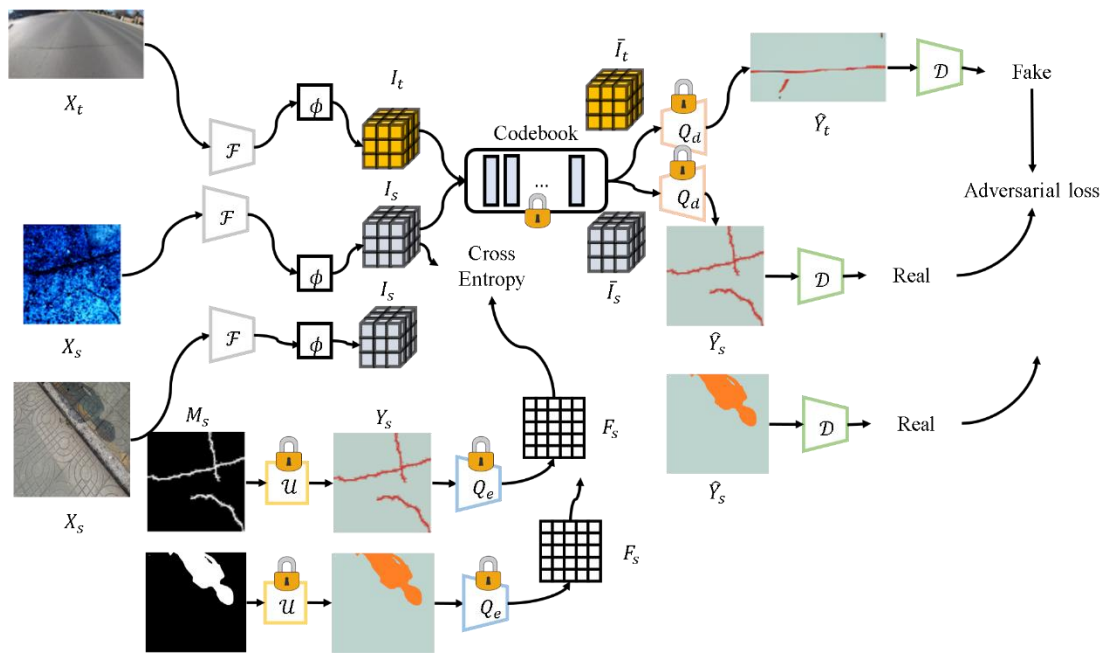


圖 2-31: 源影像和 mask 特徵對應架構

圖 2-31 為源影像和其 mask 特徵對應架構。 X_t 為路面影像， $\mathcal{F}$ 為影像的特徵提取器， ϕ 為一個 MLP 用來預測影像特徵對應編碼簿索引， $\mathcal{D}$ 為鑑別器。將 X_s 和 X_t 輸入到 $\mathcal{F}$ 和 ϕ 得到預測索引 I_t 和 I_s ， M_s 透過 μ 和 Q_e 得到索引 F_s 作為 I_s 的標準答案，因此可由(4-2)來計算 I_s 和 F_s 的 Cross entropy 使得 kaggle 影像的特徵和 mask 的特徵可以對應。透過最佳化式子(4-2)以更新 $\mathcal{F}$ 和 ϕ 。

$$\min_{\mathcal{F}, \phi} \mathbb{E}_{F_s} [-\log I_s] \quad (4-2)$$

接著， I_s 和 I_t 每位置的 patch 特徵被匹配編碼簿裡最相似的特徵，分別得到特徵 $\bar{I}_s$ 和 $\bar{I}_t$ ，將 $\bar{I}_s$ 和 $\bar{I}_t$ 輸入至 Q_d 得到預測出 maskiage(mask 的 RGB 影像，通道數同類別數量)。最後， $\hat{Y}_s$ 和 $\hat{Y}_t$ 輸入至 $\mathcal{D}$ 並分別給予其標準答案 1(real)和 0(fake)以計算式(4-3)。透過最佳化式(4-3)可以更新 $\mathcal{F}$ 、 ϕ 和 $\mathcal{D}$ 的權重。下列 Algorithm 4-1 表訓練上述三個模型($\mathcal{F}$ 、 ϕ 和 $\mathcal{D}$)的流程。

$$\min_{\mathcal{F}, \phi} \max_{\mathcal{D}} \mathbb{E}_{x_s} [\log \mathcal{D}(\hat{Y}_s)] + \mathbb{E}_{x_t} [\log (1 - \mathcal{D}(\hat{Y}_t))] \quad (4-3)$$

Algorithm 4-1. The training algorithm of the crack detection model

Input: Source image dataset D_s and target image dataset D_t ;

Output: The weights of learned networks, $\theta_{\mathcal{F}}$, θ_{ϕ} and $\theta_{\mathcal{D}}$;

1. Initialize network weights $\theta_{\mathcal{F}}$, θ_{ϕ} and $\theta_{\mathcal{D}}$;
 2. **For** iteration $\leftarrow 1$ to N_{iter} **do** // N_{iter} is maximal iteration
 3. $(x_s, M_s) \sim D_s$; $x_t \sim D_t$; // “ $\sim$ ” means sampling.
 4. $I_s \leftarrow \mathcal{F}(x_s)$; $I_t \leftarrow \mathcal{F}(x_t)$;
 5. $F_s \leftarrow Q_e(\mu(M_s))$;
 6. $\bar{I}_s \leftarrow \text{Codebook}(I_s)$; $\bar{I}_t \leftarrow \text{Codebook}(I_t)$;
 7. $\hat{Y}_s \leftarrow Q_d(\bar{I}_s)$; $\hat{Y}_t \leftarrow Q_d(\bar{I}_t)$;
 8. Compute the cross entropy loss;
 9. Compute the adversarial loss;
 10. Update $\theta_{\mathcal{F}}$, θ_{ϕ} and $\theta_{\mathcal{D}}$ via adversarial loss & cross entropy loss
 11. **end**
-

(二) 實驗影像集說明

本主題使用的資料集為 kaggle 上混凝土裂縫和 EdmCrack600(Mei et al., 2020)資料集。如圖 2-32 (a)和(b)所示，圖 2-32 (a)的裂縫影像較為單調而圖 2-32 (b)路面裂縫影像包含

如陰影、行人等額外干擾物件。混凝土裂縫集包含 8818 張裂縫影像，EdmCrack600 包含 600 張路面裂縫影像，本研究挑選其中 400 張為訓練影像其餘 200 張為測試影像。



圖 2-32: (a)kaggle 裂縫影像，(b)路面裂縫影像

此外，本主題除裂縫識別外，為處理路面陰影造成裂縫誤判問題，使用 ISTD[Wang, et al., 2018]資料集訓練模型陰影分割，圖 2-33 表示其影像和陰影遮罩。



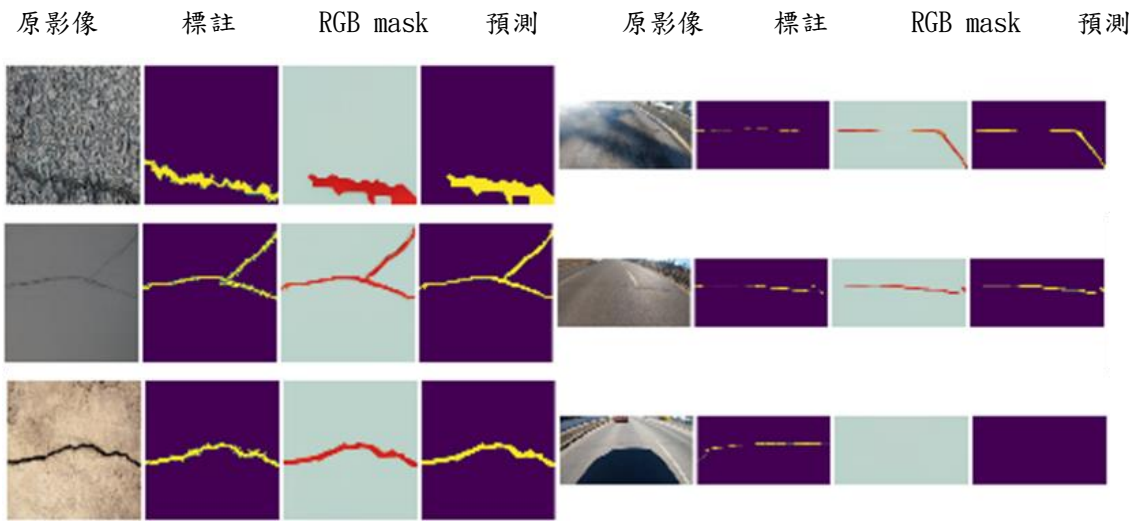
圖 2-33: (a)ISTD 影像，(b)陰影遮罩

(三) 實驗結果

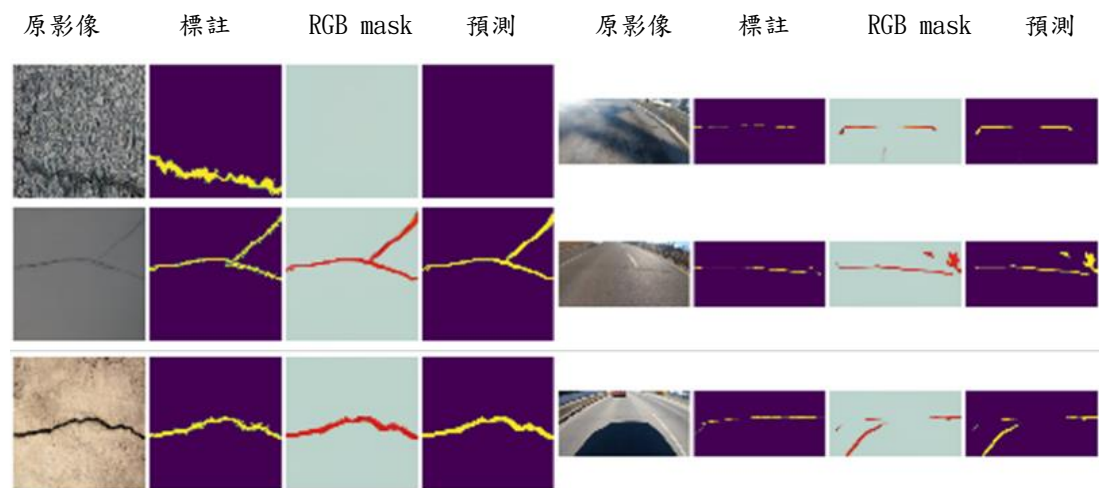
(1) 實驗結果視覺呈現

圖 2-34 為混凝土和路面裂縫分割結果。圖 2-34 (a)為傳統 CNN 實施裂縫分割結果，圖 2-34 (b)為本主題提出無陰影偵測模型(ours w/o shadow)的分割結果，圖 2-34 (c)為有陰影偵測模型(ours w shadow)的分割結果，圖 2-34 (d)為(Tsai et al.,2018)的分割結果。圖 2-34 中的混凝土裂縫皆可完整分割因為有標註的 mask。當識別的影像為路面影像時，可發現沒有做陰影偵測會將光影交叉處誤判(箭頭處)，如圖 2-34 (a)、(b)和(d)的結果都有誤判發生。相較之下，圖 2-34 (c)可識別陰影並可更有效地分割出陰影中和光影交界處的裂縫。

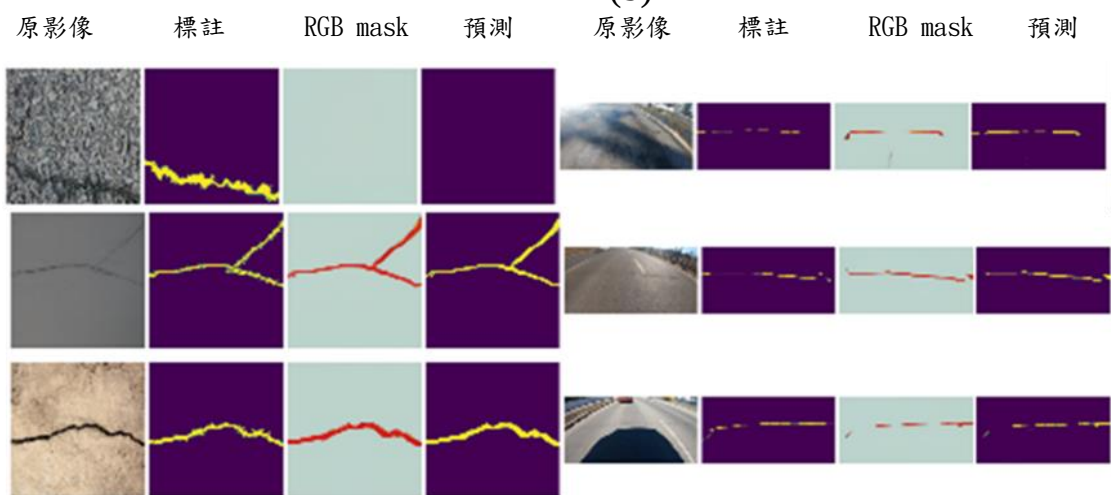
圖 2-35 為 UAV 混凝土裂縫影像分割結果。由圖 2-35 結果可觀察到本主題提出的模型在不同飛行拍攝尺度下可將目標裂縫分割。圖 2-35 (a)和(b)為近距的裂縫分割結果，結果顯示模型會將岩石交叉處誤判為裂縫。圖 2-35 (c)和(d)為較遠的裂縫影像分割結果，觀察發現模型在對付干擾及障礙物上有不錯的表現，因為較遠的影像更多的干擾及障礙物，但圖 2-35 (c)、(d)與圖 2-35 (a)、(b)的結果一樣，僅將岩石交叉處誤判。



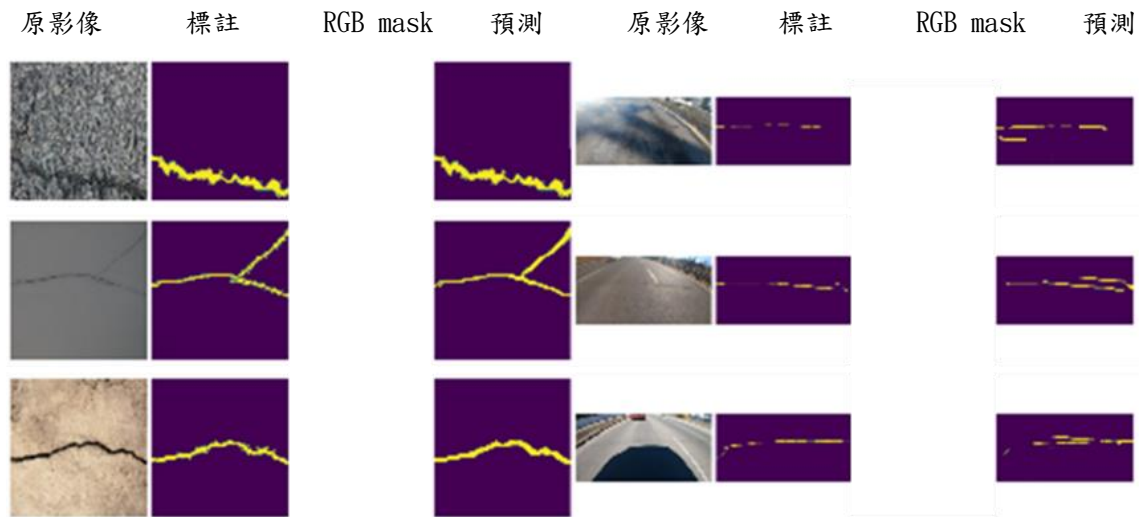
(a)



(b)



(c)

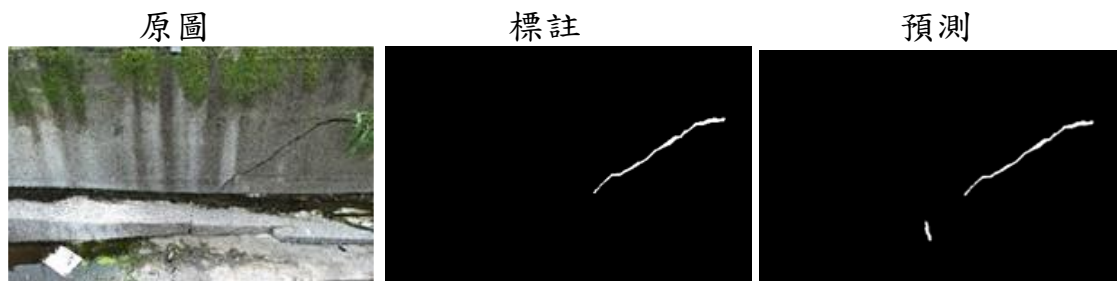


(d)

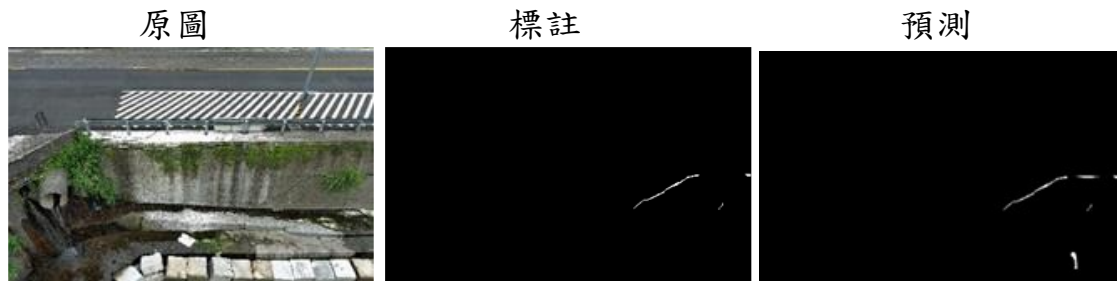
圖 2-34: 混凝土和路面裂縫分割結果。(a)CNN，(b)Ours (w/o shadow)，(c)Ours (w shadow)，(d)[Tsai et al., 2018]。



(a)



(b)



(c)

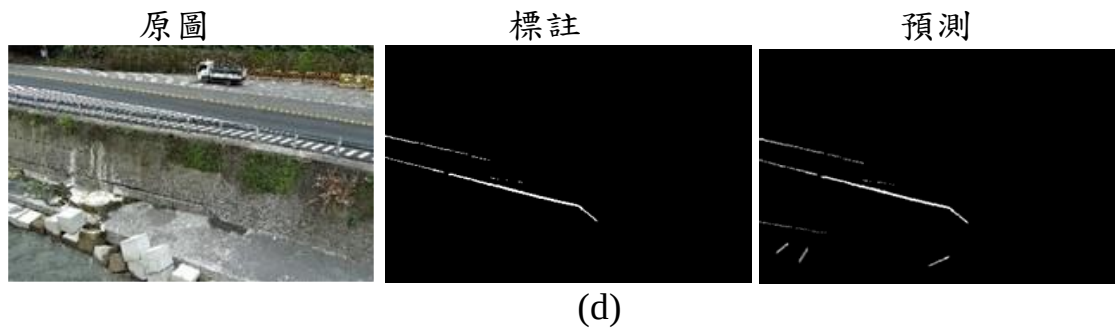


圖 2-35:空拍混凝土裂縫分割結果。(a)水平裂縫_2~3m; (b)水平拉距_起降點 7m; (c)垂直拉距_起降高 10m; (d)下游斜拍_起降高 6m。

(2)模型效能比較

本議題使用 IoU 和 Dice 為衡量模型效能的指標，分別針對下列模型做量化比較:CNN、(Tsai et al., 2018)、Ours (w/o $\mathcal{D}$)和 Ours(w $\mathcal{D}$)。表 2-8 為路面分割量化結果。觀察表 2-8 可發現，在無陰影的情況下，本主題提出的方法(ours w & w/o $\mathcal{D}$)和(Tsai et al., 2018)的量化結果差不多且比傳統的 CNN 有更好的效能。在有陰影的干擾的情況下，傳統的 CNN 的效能非常差，但本主題提出的模型加入 $\mathcal{D}$ 後的模型效能就提升許多，並超過 (Tsai et al., 2018) 和沒有加入 $\mathcal{D}$ 的效能。由此結果可得知，若影像中有複雜的地形及遮蔽物，可加入 $\mathcal{D}$ 也許有效地解決問題。

表 2-8: 路面分割量化結果

	w/o shadow		w shadow	
	IoU	Dice	IoU	Dice
CNN	0.73	0.72	0.46	0.42
[Tsai et al., 2018]	0.78	0.75	0.58	0.55
Ours (w/o $\mathcal{D}$)	0.77	0.74	0.62	0.6
Ours (w $\mathcal{D}$)	0.78	0.75	0.69	0.65

表 2-9: 空拍混凝土裂縫語意分割之量化結果

測試資料集 \ 衡量指標	mIoU	Dice
水平裂縫_2~3m	0.772	0.724
水平拉距_距起降點 7m	0.705	0.681
下游斜拍_起降點高 6m	0.637	0.620
垂直拉距_起降點高 10m	0.570	0.559

表 2-9 為空拍混凝土裂縫語意分割之量化結果。觀察表 2-9 可發現模型的效果會隨著拍攝距離而降低，這是因為距離越遠會包含到越多的干擾物。當拍攝距離為 2~3m 時，裂縫的影像跟 kaggle 混凝土影像很接近因此模型有很不錯的效果。模型最低的效果為起降高 10m，可以由圖 2-35 的結果看出影像中有很多岩石交叉處會導致誤判。

(四) 結論

本主題提出了裂縫分割模型使用自適應學習將 kaggle 的混凝土裂隙影像的識別能力遷移到路面裂隙影像，得出裂縫分割影像，再偵測 UAV 混凝土裂縫影像。本研究考慮了陰影干擾並將陰影視為除裂縫以外的第二類別。在路面裂縫影像上，實驗結果顯示本主題提出的方法可有效的處理陰影干擾問題。另外，在空拍影像分割上，實驗結果發現拍攝距離越遠模型效果會因為包含更多干擾而比較容易誤判導致模型效能較差，後續研究可針對此問題提出數種方法解決。

五、建置響應式網頁模型識別體驗網站

(一) 網頁呈現多尺度 UAV 影像分割應用於混凝土表面裂隙偵測模型

呈現多尺度 UAV 影像分割應用於混凝土表面裂隙偵測模型

操作，方式分為兩個部分：

上傳影像：上傳本地影像，即可對此影像進行分析。圖 2-36 說明多尺度 UAV 影像分割應用於混凝土表面裂隙偵測操作概念圖，提供即時分析影像，以響應式網頁呈現。

歷史資料查詢：使用者可透過網頁查詢建模實驗影像，包含混凝土表面裂隙實驗影像及其經過模型分析後之影像，圖 2-37 展示網站。

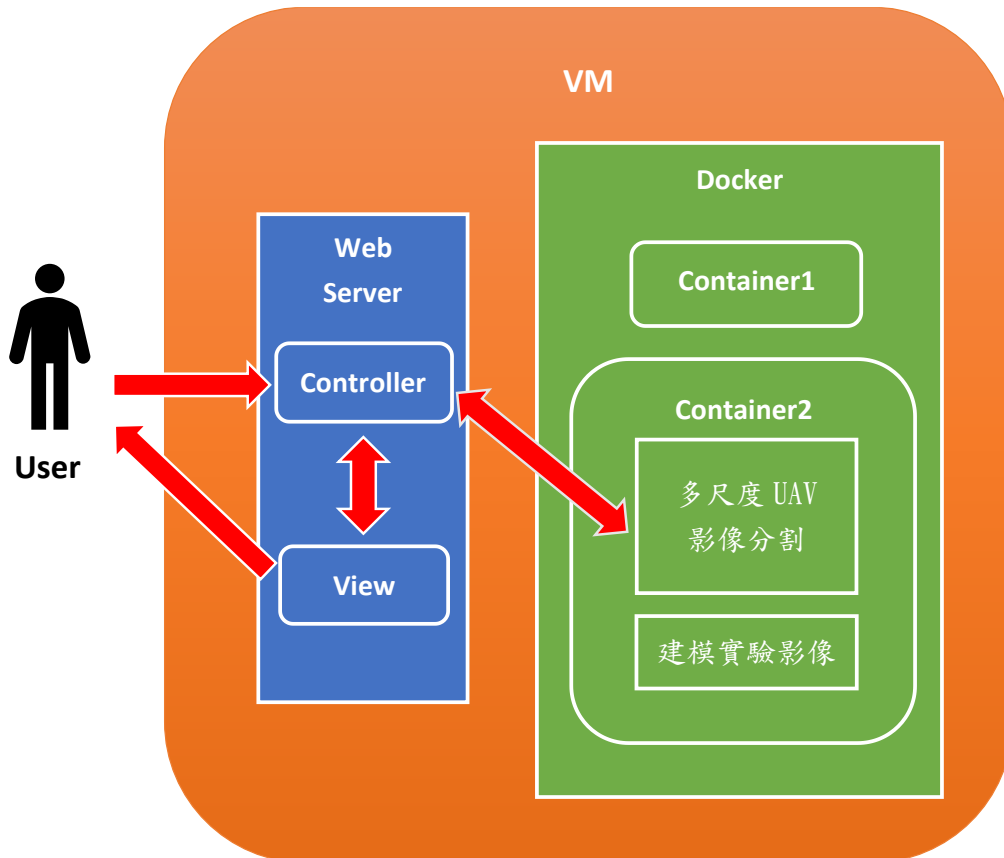


圖 2-36: 多尺度 UAV 影像分割應用於混凝土表面裂隙偵測網站架構

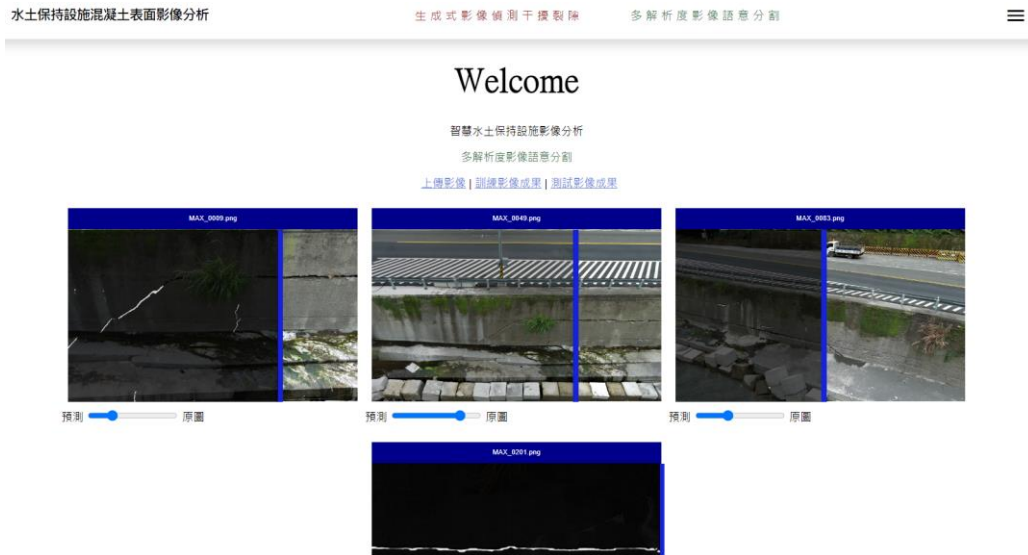


圖 2-37: 多尺度 UAV 影像分割應用於混凝土表面裂隙偵測網頁效果展示

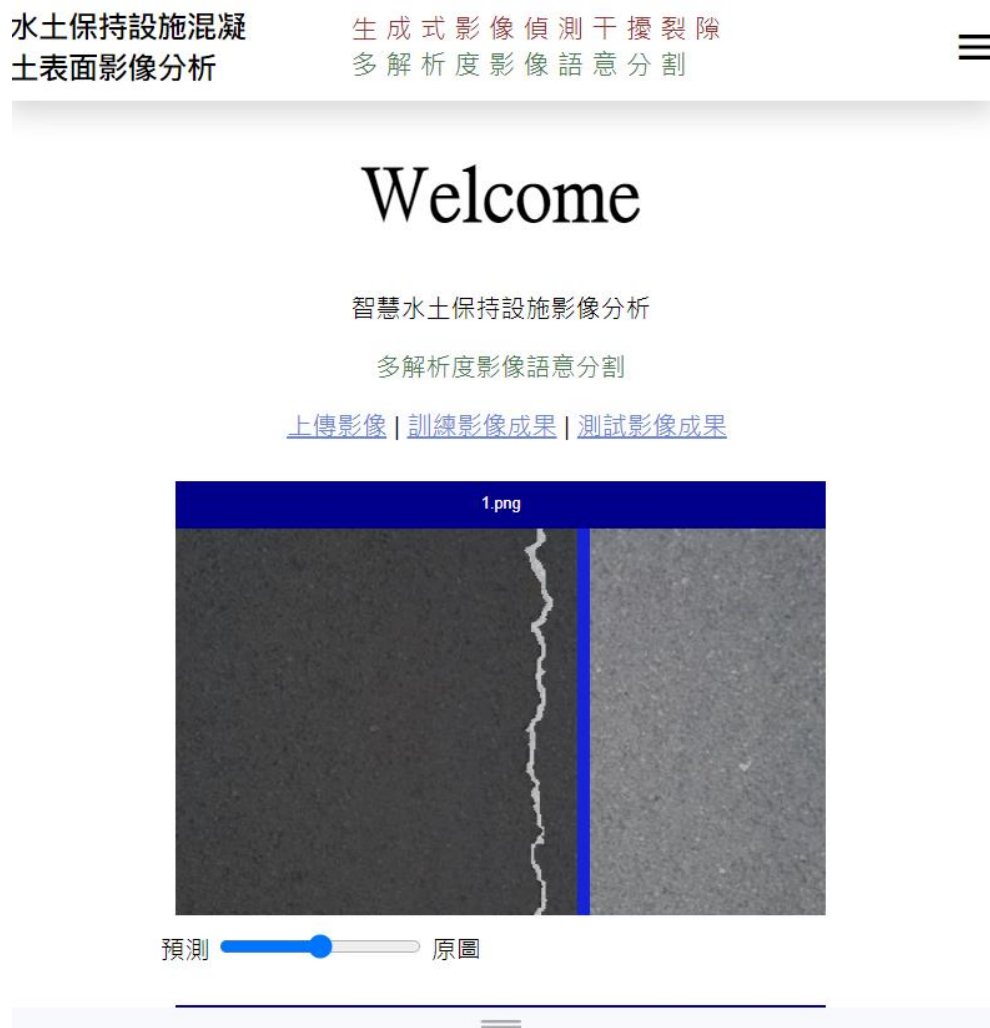


圖 2-38: 多尺度 UAV 影像分割應用於混凝土表面裂隙偵測網頁模擬手機展示

(二) 網頁呈現生成式混凝土表面影像裂縫分割

呈現生成式混凝土表面影像裂縫分割模型操作，方式分為兩個部分：

上傳影像：上傳本地影像，即可對此影像進行分析。圖 2-39 說明生成式混凝土表面影像裂縫分割操作概念圖，提供即時分析影像，以響應式網頁呈現。

歷史資料查詢：使用者可透過網頁查詢建模實驗影像，包含混凝土表面裂隙實驗影像及其經過模型分析後之影像，圖 2-40 展示網站。

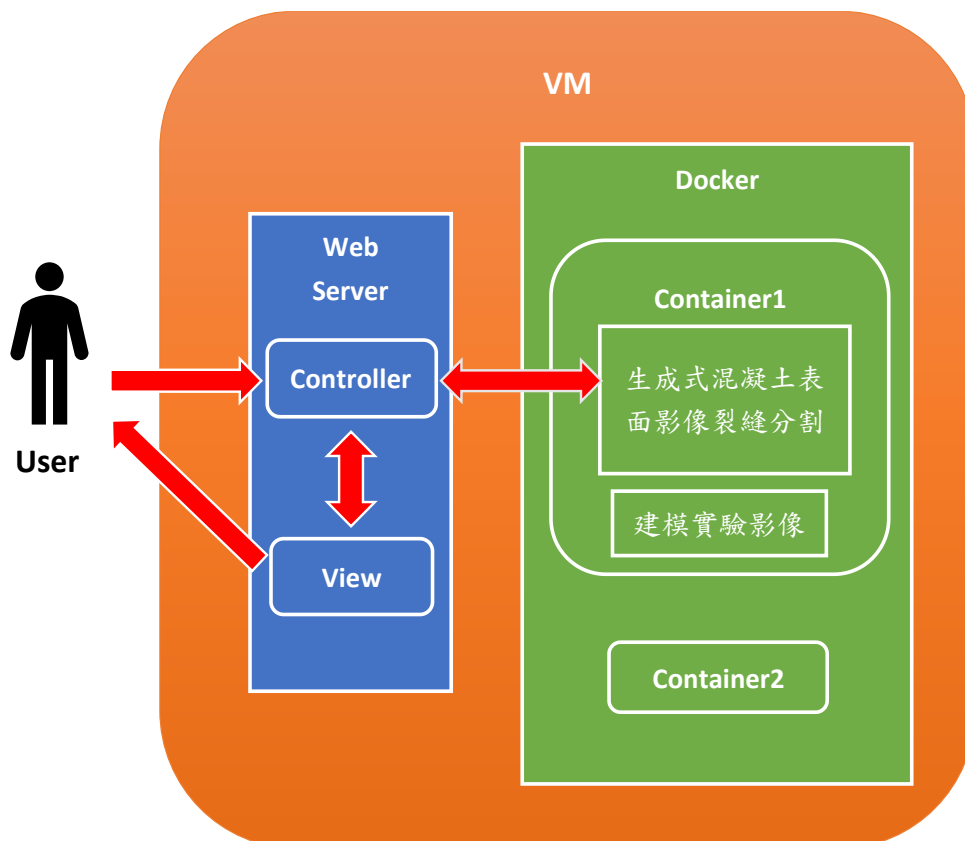


圖 2-39: 生成式混凝土表面影像裂縫分割網頁架構圖

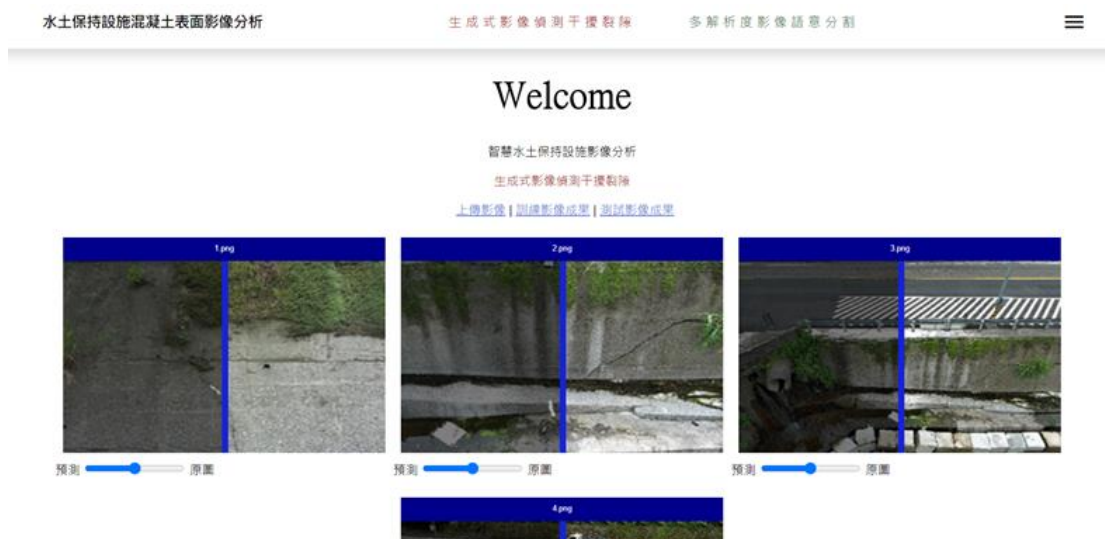


圖 2-40: 生成式混凝土表面影像裂縫分割網頁效果展示



圖 2-41: 生成式混凝土表面影像裂縫分割網頁模擬手機展示

第三章 執行成果及效益

第一節 執行成果

聚焦於使用 UAV 拍攝水土保持混凝土設施，實驗場域為埔里眉溪觀音吊橋附近拍攝，驗證本計畫執行成果，並設計影像分割模型來辨識 UAV 影像中是否有裂隙？主要分析水土保持混凝土設施表面影像是否有裂隙，本計畫已完成下列成果。

- UAV 拍攝影像參數設定及量測裂縫寬度。
- UAV 巡檢航線參數設定與巡檢歷程影像資料庫建置。
- 多尺度影像分割應用於偵測 UAV 影像混凝土表面裂隙及效能分析研究報告。
- 生成式混凝土表面影像分割應用於偵測 UAV 影像混凝土表面裂隙及效能分析研究報告。
- 建置響應式網頁模型識別體驗網站成果展示。

第二節 執行效益

一、提供標註資料庫一份之效益，提供發展相關系統的實驗影像集。

二、資料庫建置流程研究，可以運用於建置歷史巡檢影像，長期追蹤定點水土保持設施，提供發展設施異常檢測。

三、多尺度 UAV 影像分割應用於混凝土表面裂隙偵測模型，可以使用近距離拍之水土保持設施混凝土表面裂隙影像訓練

模型，並遷移至UAV拍攝遠距離之水土保持設施影像，來偵測混凝土表面裂隙，並應用於自動化設施巡檢。

四、利用生成式混凝土表面影像裂縫分割模型，可以準確判識遭光線不足之設施影像，提高巡查效率。除了上述效益之外，可以在少量影像限制之下，訓練深度學習影像偵測模型。

第三節 交付項目

本計畫可交付下列項目。

- UAV 拍攝影像參數設定及量測裂縫寬度說明一份。
- UAV 巡檢航線參數設定與巡檢歷程影像資料庫建置說明一份
- 多尺度影像分割應用於偵測 UAV 影像混凝土表面裂隙及效能分析研究報告說明一份。
- 生成式混凝土表面影像分割應用於偵測 UAV 影像混凝土表面裂隙及效能分析研究報告說明一份。
- 建置響應式網頁模型識別體驗網站成果展示說明一份。

第四章 結論

本計畫已完成項目說明如下

- UAV 拍攝影像參數設定及量測裂縫寬度之研究：針對量測裂縫與估算拍攝距離之研究，得到結論是空拍機鏡頭正視標地物並距離 2m 至 18m 均可以清楚拍攝 60mm 裂縫的影像並量測到其裂縫寬，影像分割模型預測出裂隙可估計真實裂縫寬度，本計畫發展出一個演算法可用模型預測出裂隙與拍攝距離(WD)估算實際裂隙寬度，假設已知原始影像為距離裂隙 18m 拍攝，解析度為 5472x3648 的影像，在模型中計算時的解析度縮小為 912x912 影像，經實驗顯示實際 60mm 裂縫寬，演算法估算為 63.77mm，另一例，實際 100mm 裂縫寬，演算法估算結果為 103.03mm，兩者誤差約 4mm，其在可接受範圍內，此方法可用於預測實際裂縫寬度。
- UAV 巡檢航線參數設定與巡檢歷程影像資料庫建置之研究：圖 2-14 說明任務錄製的實驗，第一次手動拍攝、第二次與第三次的復現 GPS 軌跡圖，其中橫-縱軸為經-緯度。藍線為手動飛行，綠線與紅線為使用 UAV 任務錄製復現 GPS 軌跡，黑點表示校準點，懸停在定點但不拍攝影像，

從橘點才開始拍攝紀錄。在實驗中，可以觀察到手動拍攝、第二次與第三次的復現軌跡相當靠近，透過計算同一記錄點的經緯度歐式距離誤差，在 $5.58 \times 10^{-6} \sim 30.1 \times 10^{-6}$)區間，估算真實距離大約在 70~300cm。未來建立巡檢歷程資料庫，建議記錄點間隔 5m，以利系統區隔不同檢查點拍攝影像。結論因拍攝環境影響變因，導致不同次巡檢在同一紀錄點之 GPS 定位座標取得產生誤差，但影像中內容可能大致相同。因此未來建立巡檢歷程資料庫，可以採用以下巡檢參數的來設定空拍機巡檢參數，其偵測標的設施水平距離 10m，垂直 10m，拍攝記錄點間隔 5m。

- 多尺度影像分割應用於偵測 UAV 影像混凝土表面裂隙之研究：使用 111 年蒐集的資料集作為測試及三個影像分割模型對本項模型多尺度影像分割做效能比較，並以 precision、recall、f1-score 與 mIoU 作為量化指標。本項模型效能有顯著提升勝過其他三個模型。本項模型在本計畫收集 UAV 影像集，於不同角度及不同距離裂隙下所拍攝。本項模型在「水平裂縫_2~3m」上表現最好，因為在這種尺度下所拍攝影像與訓練資料集相近。另外，本項實驗針對三種環境干擾：昏暗、起霧、下雨，來實施韌性測驗，結

果發現在起霧干擾下，模型效能在 precision 下降最多，表示模型對抗起霧干擾的能力較為弱，不過在三種環境干擾下本模型的偵測裂隙結果，還是維持不錯的效能，此外，本項模型也比較其他 3 種多尺度模型，本項模型效能都優於這 3 種模型。

- 生成式混凝土表面影像分割應用於偵測 UAV 影像混凝土表面裂隙之研究：本項模型設計除偵測裂隙外，也可同時也解決陰影造成誤判裂隙的問題，以路面陰影資料集來訓練模型後，再用來偵測有陰影干擾裂隙，可明顯地提升傳統偵測裂隙方法效能。本計畫已完成本項模型對原影像、無陰影路面裂縫與有陰影路面裂縫的效能比較，結果顯示本項模型對有無陰影干擾的裂縫偵測有不錯成果，針對本計畫收集 UAV 影像集中在「水平裂縫_2~3m」影像偵測裂隙的 mIoU 與 DICE 分別為 0.772 及 0.724，優於其他三種拍攝距離。
- 建置響應式網頁模型識別體驗網站成果展示：已完成網站架構及模型封裝流程設計，及上述兩種模型佈署，可提供線上體驗操作。

參考文獻

- [1] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
- [2] W.-C. Hung, Y.-H. Tsai, X. Shen, Z. Lin, K. Sunkavalli, X. Lu, and M.-H. Yang. Scene parsing with global context embedding. In *ICCV, 2017*.
- [3] F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. In *ICLR, 2016*.
- [4] D. Pathak, P. Krahenbuhl, and T. Darrell. Constrained convolutional neural networks for weakly supervised segmentation. In *ICCV, 2015*.
- [5] S. R. Richter, V. Vineet, S. Roth, and V. Koltun. Playing for data: Ground truth from computer games. In *ECCV, 2016*.
- [6] G. Ros, L. Sellart, J. Materzynska, D. Vazquez, and A. Lopez. The SYNTHIA Dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *CVPR, 2016*.
- [7] Chen, J., Lu, J., Zhu, X., & Zhang, L. (2023). Generative semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7111-7120).
- [8] Xie, S., & Tu, Z. (2015). Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 1395-1403).
- [9] Liu, Y., Yao, J., Lu, X., Xie, R., & Li, L. (2019). DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing*, 338, 139-153.
- [10] Liu, Y., Cheng, M. M., Hu, X., Wang, K., & Bai, X. (2017). Richer convolutional features for edge detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3000-3009).
- [11] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2117-2125).

- [12]Yang, F., Zhang, L., Yu, S., Prokhorov, D., Mei, X., & Ling, H. (2019). Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Transactions on Intelligent Transportation Systems*, 21(4), 1525-1535.
- [13]Zou, Q., Zhang, Z., Li, Q., Qi, X., Wang, Q., & Wang, S. (2018). Deepcrack: Learning hierarchical convolutional features for crack detection. *IEEE transactions on image processing*, 28(3), 1498-1512.
- [14]Peng, C., Zhang, X., Yu, G., Luo, G., & Sun, J. (2017). Large kernel matters--improve semantic segmentation by global convolutional network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4353-4361).
- [15]Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., & Lu, H. (2019). Dual attention network for scene segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 3146-3154).
- [16]Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).
- [17]Yu, F., & Koltun, V. (2015). Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*.
- [18]Yu, F., Koltun, V., & Funkhouser, T. (2017). Dilated residual networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 472-480).
- [19]Yang, M., Yu, K., Zhang, C., Li, Z., & Yang, K. (2018). Denseaspp for semantic segmentation in street scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3684-3692).
- [20]Sun, X., Xie, Y., Jiang, L., Cao, Y., & Liu, B. (2022). DMA-Net: DeepLab with multi-scale attention for pavement crack segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 23(10), 18392-18403.
- [21]Sang, H., Zhou, Q., & Zhao, Y. (2020). PCANet: Pyramid convolutional attention network for semantic segmentation. *Image and Vision Computing*, 103, 103997.
- [22]Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez,

- A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [23]Jiang, X., Mao, S., Li, M., Liu, H., Zhang, H., Fang, S., ... & Zhang, C. (2022). MFPA-Net: An efficient deep learning network for automatic ground fissures extraction in UAV images of the coal mining area. *International Journal of Applied Earth Observation and Geoinformation*, 114, 103039.
- [24]Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, 34, 12077-12090.
- [25]Chen, C. F. R., Fan, Q., & Panda, R. (2021). Crossvit: Cross-attention multi-scale vision transformer for image classification. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 357-366).
- [26]Liu, H., Miao, X., Mertz, C., Xu, C., & Kong, H. (2021). Crackformer: Transformer network for fine-grained crack detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 3783-3792).
- [27]Liu, H., Yang, J., Miao, X., Mertz, C., & Kong, H. (2023). CrackFormer Network for Pavement Crack Segmentation. *IEEE Transactions on Intelligent Transportation Systems*.
- [28]Tian, R., Wu, Z., Dai, Q., Hu, H., Qiao, Y., & Jiang, Y. G. (2023). Resformer: Scaling vits with multi-resolution training. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 22721-22731).
- [29]Y. Ganin and V. Lempitsky. Unsupervised domain adaptation by backpropagation. In *ICML*, 2015.
- [30]M.Long, Y. Cao, J. Wang, and M.Jordan. Learning transfer able features with deep adaptation networks. In *ICML*, 2015.
- [31]M. Long, H. Zhu, J. Wang, and M. I. Jordan. Unsupervised domain adaptation with residual transfer networks. In *NIPS*, 2016.
- [32]J.Hoffman, D.Wang, F.Yu, and T.Darrell. Fcns in the wild: Pixel-level adversarial and constraint-based adaptation. *CoRR*,abs/1612.02649,2016.
- [33]J.Hoffman,E.Tzeng,T.Park,J.Y.Zhu,P.Isola,K.Saenko,A.A.Efros,and

- T.Darrell.Cycada: Cycle-consistent adversarial domain adaptation. *CoRR,abs/1711.03213*,2017.
- [34]J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image to-image translation using cycle-consistent adversarial net works. *ICCV*, 2017.
- [35]Tsai, Y. H., Hung, W. C., Schalter, S., Sohn, K., Yang, M. H., & Chandraker, M. (2018). Learning to adapt structured output space for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7472-7481).
- [36]Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).
- [37]khanhha. (2020). Crack Segmentation. Retrieved from github: https://github.com/khanhha/crack_segmentation
- [38]<https://www.kaggle.com/datasets/yatata1/crack-dataset>
- [39]Q. Mei, M. Gül, and M.R. Azim, Densely connected deep neural network considering connectivity of pixels for automatic crack detection. *Automation in Construction*, 2020. 110: p. 103018.
- [40]Wang, J., Li, X., & Yang, J. (2018). Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1788-1797).

- 諮詢專家：詹勳全教授兼系主任（已口頭同意擔任本計畫諮詢專家）
國立中興大學 水土保持學系
國立中興大學 大數據中心 研究員

附錄

附錄一、112 年水保局創新研究計畫審查意見暨回覆說明

項次	審查意見	回覆說明
為配合本局業務需求，計畫內容建議調整(或納入後續執行評估參考)部分：		
一	水保局 UAV 影像成果上傳平台之 UAV 的影像均已串接到本局歷史影像平台逾 2.2 萬張，無需申請帳號即可下載。	感謝委員建議，計畫初期會以爬蟲技術介接歷史影像平台下載 UAV 影像分析是否可以用於本計畫。
二	如以 UAV 自動巡檢為情境，判釋出異常處時，宜自動記錄其位置坐標及方位，以便後續維護處理時確認其位置。	感謝委員建議，UAV 影像或視訊應有存其位置坐標及方位(GPS 資料)，判釋異常影像後可以利用程式自動讀取其位置坐標及方位，提供異常位置回報，上述說明會納入計畫執行。
三	研究以相似影像演算法，依據 GPS 飛航軌跡找尋數個相近的拍攝地點，再設計演算法從中挑選匹配的影像，但是這只有解決空間上的問題，在不同時間因為不同拍攝條件以及不同大氣狀況，所得到的影像，彼此之間要如何比較呢？建議思考類似平均亮度或輻射強度正規化的演算法。	感謝委員建議，在計畫書中「議題 1-3: 整理不同時序的 UAV 影像」由提及此問題，相鄰連續時間但在同一(近似)GPS 位置的影像如何處理，計畫執行期間會遵照委員建議似平均亮度或輻射強度正規化的演算法來研究與實驗，觀察期效果。也會注意 UAV 尋檢定位 GPS 飛航前，是

項次	審查意見	回覆說明
		否需要前置 calibration? 也是研究議題。
四	人工智慧訓練成功的關鍵，在於有高品質且龐大數量的訓練資料，本計劃考慮水土保持設施異常的案例較少，準備利用少樣本與不平衡資料及以及遷移學習的建模演算法，但是最原始的影像標註問題，要如何解決？建議尋找國際上相近的資料庫，可以節省人力與時間，所得到的成果也比較有公信力。	感謝委員建議，本計畫以三種方式來尋找國際上相近的資料庫，(1)survey papers 中有提供資料庫；(2)至 Kaggle 尋找；(3)至 Github 尋找。如有找到與本計畫問題相近的國際水土保持資料集，將研究自適應領域學習(Domain Adaption)，此為遷移學習的一種演算法，使得少量資料之訓練模型可從其他資料集訓練並應用到本計畫的影像異常偵測問題上。另一種方式可運用生成影像技術來生成相似的影像異常的影像集來訓練偵測異常影像模型。
五	本計劃所使用的 USB 影像，應該沒有太多人、車牌、門牌等敏感資訊，因此 YOLO 的使用應該再加思考，適當調整。	感謝委員建議，如影像中沒有人、車牌、門牌等敏感資訊可以不用 YOLO 來當前處理辨識敏感資訊。

附錄二、112 年水保局創新研究期中報告審查意見暨回覆說明

	審查意見	回覆說明
為配合本局業務需求，計畫內容建議調整(或納入後續執行評估參考)部分：		
一	期中報告內容多描述理論及方法，請補充實際完成之成果，包括已收集多少水保設施之 UAV 影像、影像辨識之 Recall, precision 及運算時間等。	感謝委員建議 目前共收集共 464 張水保設施之 UAV 影像，已人工標註 228 張，並分成以下四類，下游斜拍、垂直拉距、水平拉距、水平裂縫。影像辨識之 Recall、precision、dice coefficient、MIOU 與 F1-score。提供於第二節--三、四的效能比較。
二	請針對特殊狀況做測試，包括光線不足、下雨或部分遮蔽狀況之因應方式。	感謝委員建議。 本計畫於測試階段對測試資料使用 image augmentation 模擬現實各種情況，除了光線不足，下雨和部分遮蔽等常見情況。此外，因應山中常有起霧等特殊情況，因此本計畫也有模擬起霧情況。提供於第二節--三的效能比較。
三	P2-8 UAV 欲拍攝 1mm 的裂縫，拍攝距離必須小於 925mm，如此近的距離，對 UAV 的安全恐有疑慮。	感謝委員建議。 據所知部分 UAV 有防撞功能，可以避免拍攝時有安全疑慮。此外，因主要拍攝的裂縫會是 20mm 以上的裂縫，可以離 2m 至 18m，因此實際上不會以 925mm 這麼近的距離進行飛行拍攝。
四	如欲以 UAV 作為構造物巡檢的工具，應先確認巡檢的項目、重點及評估標準為何？	感謝委員建議。 本計畫研究重點為 UAV 拍攝混凝土表面影像裂縫偵測實驗與裂縫寬度量測。本計畫實驗場域埔里眉溪觀音吊橋之護岸，本計畫模擬實驗結果可提供水土

	審查意見	回覆說明
		保持專家來訂定評估標準。
五	P2-7，偵測範圍建議可納入無人機地面採樣距離（Ground sample distance, GSD）概念，提供無人機鏡頭與混凝土表面之距離建議表，並建議參考符合現行資安規定之無人機型進行表列呈現(部分代表機型即可)。	感謝委員建議。 本計畫模擬實驗中無人機空拍說明，請參閱附錄四（附-10）。實驗之無人機製作廠商(Autel)符合規定，無資安疑慮。
六	大部分驗證與訓練圖片，位於室內或環境可控制，當資料轉為戶外設施、防砂設施，或現地日照陰影等干擾時，建議可於後續現地拍照後，進行評估是否直接套用，或受其他干擾時之調整流程或方法等。	感謝委員建議。 本計畫之模型使用 domain adaption 技術以 Kaggle 資料集作為 source data，路面陰影影像資料集作為 target data，因此也可以使用於判斷混凝土表面的裂縫。另外以程式模擬場景昏暗、下雨及起霧，三種干擾。其模型效能於第二章-第二節-三-模型效能比較中詳述。
七	為因應機關組織改造，機關名稱修正為「農業部農村發展及水土保持署」，報告書封面及內文等請依本署最新範本格式撰寫，相關範本檔案請至本署電子圖書管理中心下載(https://elibrary.ardswc.gov.tw/)。	期末報告以更正，感謝委員提醒

附錄三、空拍機參數說明

本計畫使用此品牌空拍機 EVO II Pro V3 如下圖



規格如下所示：

影像感測器	1 英寸 CMOS ； 2000 萬像素
鏡頭	FOV 82° 29 mm (35 mm 格式等效) f/2.8-f/11 對焦點：0.5 m 至無窮遠；
ISO 範圍	ISO 100 ~ ISO 44000 (影片) ； ISO 100 ~ ISO 6400 (影像)
影像最大解析度	5472×3648(3:2) 5472×3076(16:9) 3840×2160(16:9)
影像拍攝模式	單張拍攝(默認) 連拍/AEB 連拍/定時拍/HDR 拍照
錄影解析度	5472x3076P30/P25/P24 3840x2160P60/P50/P48/P30/P25/P24 2720x1528P60/P50/P48/P30/P25/P24 1920x1080P60/P50/P48/P30/P25/P24
影片存儲最大碼流	120 Mbps
圖片格式	JPEG, DNG, JPG+DNG

影片格式 MP4, MOV (MPEG-4
AVC/H.264, HEVC/H.265)

工作環境溫度 5°C to 45°C

無人機飛行軌跡示意圖

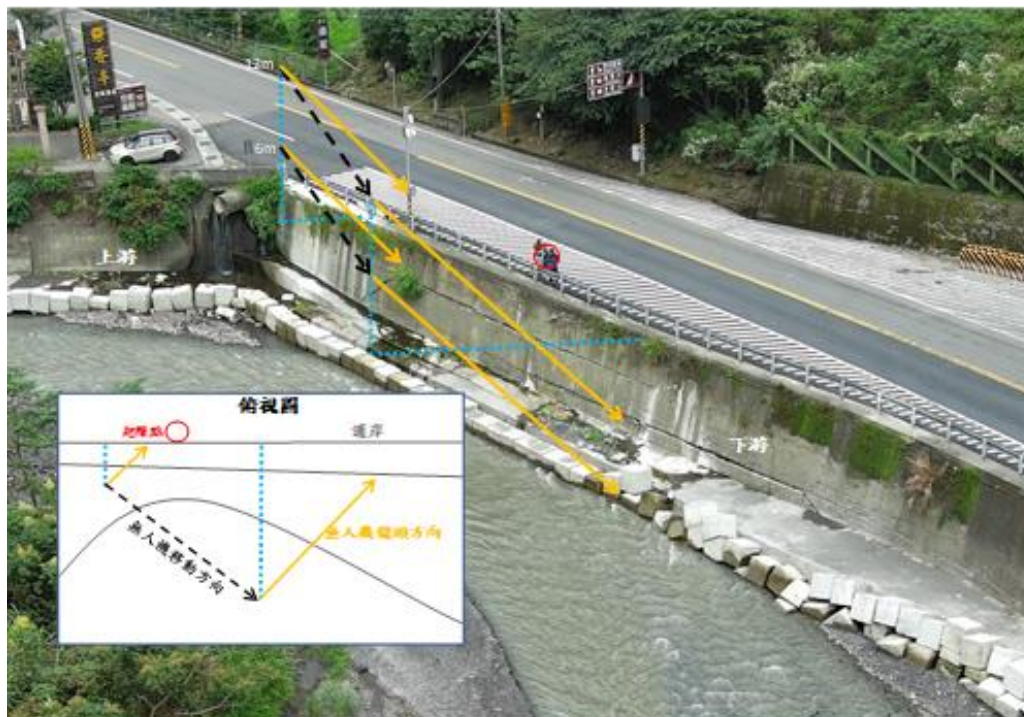


圖 附錄 3-1: 起降點高 6m、12m 上游向下游斜拍示意圖



圖 附錄 3-2: 起降點高 6m、12m 下游向上游斜拍示意圖

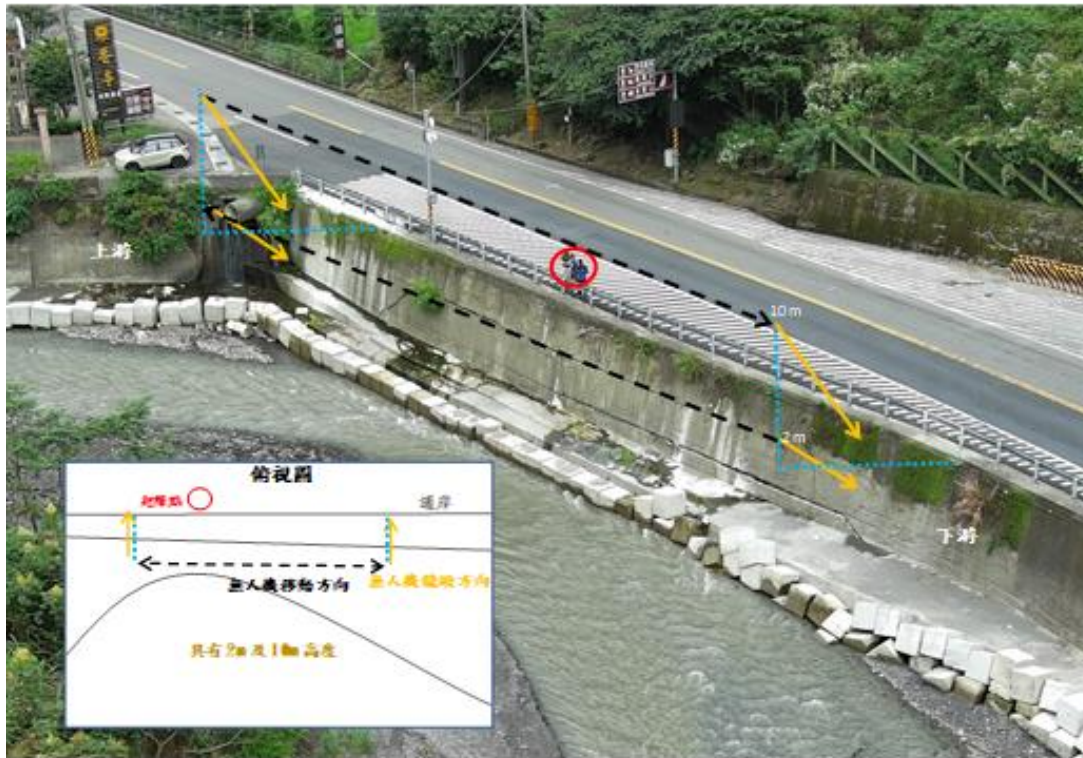


圖 附錄 3-3:起降點高 2m、10m 垂直拉距示意圖

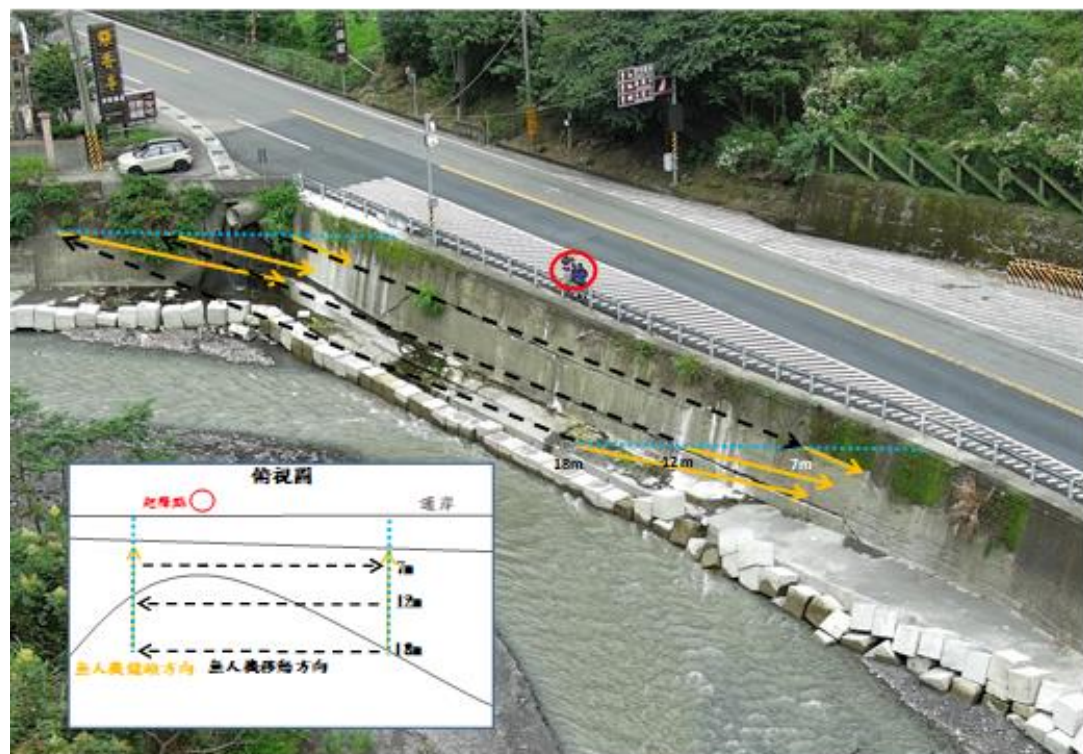


圖 附錄 3-4:起降點水平距離 7m、12m、18m 示意圖

表 附錄 3-1:無人機拍攝參數表

拍攝方式	高度	距離	角度	數量
上游斜拍	6/12 (m)	NaN	45°	43
下游斜拍	6/12 (m)	NaN	-45°	54
垂直拉距	2/10 (m)	12	0°	82
水平拉距	0 (m)	7/12/18 (m)	0°	165



圖 附錄 3-5:起降點高 6m(左)、12m(右)上游向下游斜拍圖



圖 附錄 3-6:起降點高 6m(左)、12m(右)下游向上游斜拍圖



圖 附錄 3-7:起降點垂直高 2m(左)、10m(右)拍攝圖



圖 附錄 3-8:起降點水平距離 7m(左上)、12m(右上)、18m(下)拍攝圖

附錄四、無人機地面採樣距離（Ground sample distance，GSD）概念，與無人機鏡頭與混凝土表面之距離建議表

只要障礙物在上表無人機感測範圍內，並且速度小於有效感應速度，即可感測到與障礙物距離，可以人為控制以避免碰撞，減少安全疑慮。

表 附錄 4-1:無人機安全感測系統

感測系統方向	準確感測範圍	有效感應速度
向前	0.5~18m	<12m/s
向後	0.5~16m	<12m/s
向上	0.5~10m	<5m/s
向下	0.5~10m	<5m/s
側邊	0.5~10m	<5m/s

表 附錄 4-2:與表面之最短距離建議表

感測系統方向	速度=12m/s	5m/s	懸停
向前	13m	6m	1~1.5m
向後	13m	6m	1~1.5m
向上	NaN	6m	1~1.5m
向下	NaN	6m	1~1.5m
側邊	NaN	6m	1~1.5m

因無人機之感測速度限制，以有效感應速度來提供距離，給予 1 秒時間緩速，並給予 1m 之距離緩衝，考慮到準確感測距離最近為 0.5m，懸停以 1m 為安全距離，操作時以人為時刻檢測距離避免操作至安全距離以內。

附錄五、資料擴增用於模擬現實情況之成果

為驗證模型對於現實情境是否具有強健性，本研究使用實際拍攝之水土保持設施影像為原影像集，並對所有影像依光線不足，下雨及起霧等情境做資料擴增(data augmentation)。

表 附錄 5-1: 水平裂縫_2~3m 資料擴增成果

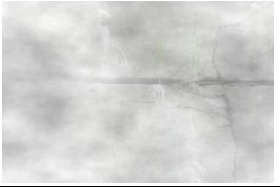
2m	弱	中	強
原影像			
昏暗			
下雨			
起霧			

表 附錄 5-2: 下游斜拍_起降點高 6m 資料擴增成果

6m	弱	中	強
原影像			
昏暗			
下雨			
起霧			

表 附錄 5-3: 水平拉距_距起降點 7m 資料擴增成果

7m	弱	中	強
原影像			
昏暗			
下雨			
起霧			

表 附錄 5-4: 垂直拉距_起降點高 10m 資料擴增成果

10m	弱	中	強
原影像			
昏暗			
下雨			
起霧			

表 附錄 5-1~表 附錄 5-4，四個表格分別表示不同飛行距離的測試集套用不同情境和強度的示意圖。

附錄六、拍攝花絮



圖 附錄 6-1: 實驗場域環境照



圖 附錄 6-2: 埔里眉溪觀音吊橋附近護岸拍攝

附錄七、期末報告審查意見暨回覆說明

	審查意見	回覆說明
計畫內容建議調整(或納入後續執行評估參考)部分：		
一	如遇植被茂盛之構造物是否會影響判釋？有無解決方法？	構造物表面植被但裂隙沒有遭到遮蔽不影響判釋。但若是裂隙大部分區域遭到植被遮擋，則無法判釋。
二	距離太遠無法判釋，最佳距離 2 至 3m，UAV 拍攝似有一定難度，有無改善方法？	依據期末報告中表 2-3 及表 2-9 說明，模型效能在離構造物水平距離 5~7M，雖有不錯的效能。但是會拍攝到更多不相關物件內容，造成模型誤判。。
三	照片之色彩、明暗、對比等是否會影響判釋結果？	在期末報告中中表 2-4，輕微的色彩、明暗、對比變化，不會影響判釋結果。至於更劇烈的變化，造成裂隙與背景嚴重混淆，則模型偵測效能會明顯下降。