

# 應用深度學習於航照正射影像之房屋偵測

張智安<sup>1\*</sup> 傅于洳<sup>2</sup>

## 摘要

為提升辨識房屋偵測效率，本研究以深度學習技術建立智慧辨識方法，使用臺灣通用電子地圖搭配航照正射影像建立訓練資料集，萃取影像中房屋區域並偵測前後期房屋變遷區域。研究策略先偵測前後期房屋區域，再利用前後期房屋區域進行變遷分析。分析房屋區域偵測成果，影像中房屋高差移位會增加誤授的比例，智慧辨識的房屋面積大於臺灣通用電子地圖的房屋範圍，房屋偵測成果之準確率約 74%，召回率達 90%。比較使用不同年度或範圍之訓練與預測資料後，發現利用前期圖資作為深度學習模型的訓練資料，預測相同範圍之後期房屋區域，有較佳的偵測精度。

**關鍵詞：**房屋偵測、變遷分析、深度學習、語意分割

## 1. 前言

### 1.1 研究背景與動機

隨著人工智慧的發展，深度學習 (Deep Learning, DL) 技術廣泛應用在各個不同領域，以提升資料識別或預測之能力。航遙測領域也有許多深度學習技術的應用發展，例如提升觀測資料品質、影像識別能力等。從圖資更新 (Map Updating) 的角度出發，若能在修測前獲得變遷的範圍或變遷比例，將有助於判斷圖資更新的急迫性或優先順序，進而提升更新的效益。為能提升測量製圖的效率，發展應用人工智慧技術的智慧製圖是一項重要的議題。

### 1.2 相關研究

建築物是電子地圖中重要的圖層之一，對於災害管理、人口估算、城市規劃等許多地理空間相關應用至關重要。房屋圖層的產製主要為使用高解析衛星立體對影像、航照立體對影像、空載光達等資料。由於航照影像有較佳的空間解析度及影像重疊率，常使用立體測量產製精確的房屋圖層。在人工智慧產製房屋區域的研究發展中，為能直接獲取房

屋區域進行製圖，大多研究使用正射影像配合深度學習直接萃取房屋區域，例如Yuan (2017) 提出學習逐像元分類的卷積類神經網路，從空間解析度0.3 m 的航空影像中自動化萃取房屋區域與輪廓；Wu *et al.* (2018) 使用空間解析度0.075 m 的航空影像建立房屋訓練及檢核資料集，以評估不同萃取房屋輪廓 (Building Footprint) 深度學習模型的效能；Yang *et al.* (2018) 使用1 m 航空影像大範圍偵測美國的房屋，應用於都市變遷 (Urban Dynamic) 研究；SpaceNet (Van Etten *et al.*, 2018) 使用高解析衛星幾何改正影像建立全球多個城市的房屋偵測資料集，應用於建立房屋偵測演算法。雖然正射影像中的房屋輪廓線精度有限，針對不同的應用目的有不同程度的房屋偵測精度需求，正射影像中的房屋輪廓線仍可以滿足不同的應用需求。例如Google Open Building (2021) 計畫使用高解析衛星幾何改正影像萃取非洲的房屋輪廓，應用於非洲地區社會公益 (Social Good) 相關計畫；在開放街圖的緊急繪圖 (Crisis Mapping) 計畫，數化房屋輪廓亦是使用高解析衛星幾何改正影像；Ghaffarian *et al.* (2019) 利用深度學習模型訓練災前房屋輪廓資料及高解析衛星幾何改正影像以預測災後建物狀況，應用於災後

<sup>1</sup> 國立陽明交通大學土木工程學系 教授

<sup>2</sup> 國立陽明交通大學防災與水環境研究中心 副工程師

\* 通訊作者, E-mail: tateo@nycu.edu.tw

收到日期：民國 110 年 11 月 08 日

修改日期：民國 110 年 11 月 29 日

接受日期：民國 110 年 12 月 02 日

快速更新建物地圖，可協助判斷救援行動及重建決策。綜合以上國際發展，使用正射影像直接萃取房屋輪廓的研究有其應用價值。

房屋變遷偵測鎖定有變遷的區域，再經由人工介入進行修測，故使用正射影像萃取房屋區域，有機會滿足變遷分析之應用需求。以深度學習進行房屋變遷偵測主要有兩種策略，第一種策略為建立房屋偵測深度學習模型，分別從前後期正射影像偵測房屋區域，再利用前後期房屋區域偵測變遷 (Xu *et al.*, 2019)；第二種策略為整合前後期正射影像建立變遷偵測模型，直接預測變遷區域 (Chen & Shi, 2020、Diakogiannis *et al.*, 2021)。雖然第二種策略具有end-to-end直接預測變遷區域的優勢，但取得大量具有房屋變遷的訓練資料集困難度較高，而房屋偵測的訓練資料集較普及。本研究欲探討的議題是整合前期電子地圖及後期正射影像偵測之房屋區域，利用這兩組資料進行變遷偵測，以探究在都會區偵測變遷的效益。

### 1.3 研究目的

本研究之目的為利用臺灣通用電子地圖與對應航照正射影像建立房屋訓練資料集，並應用深度學習技術偵測房屋區域及房屋變遷。探討與分析使用通用電子地圖建立深度學習訓練資料集於都會區偵測房屋區域及房屋變遷之成效。

## 2. 研究資料

研究實驗範圍位於新竹縣市區域，包含有明顯房屋變遷之竹北市，如新竹高鐵站周圍的建設發展等，研究資料為國土測繪中心提供之 106 年及 108 年臺灣通用電子地圖，地圖之圖幅編號及圖號沿用五千分之一基本地形圖之分幅及圖號編碼 (內政部國土測繪中心，2020)。臺灣通用電子地圖包含 10 大類圖資，如道路、鐵路及捷運、水系、行政界、建物等，本研究僅選用建物類別中的建物圖層，及正射影像類別中的彩色正射影像。研究區域共包含 23 幅五千分之一圖幅 (圖 1)，總面積約為 161.477 km<sup>2</sup>。選定 106 年 22 幅臺灣通用電子地圖及正射影像作為訓練資料集，預測資料集則包括 1 幅「跨圖

幅預測」之 106 年涵蓋新竹高鐵站之正射影像 (圖幅號 96224072)，及 23 幅「跨年度預測」之 108 年正射影像。

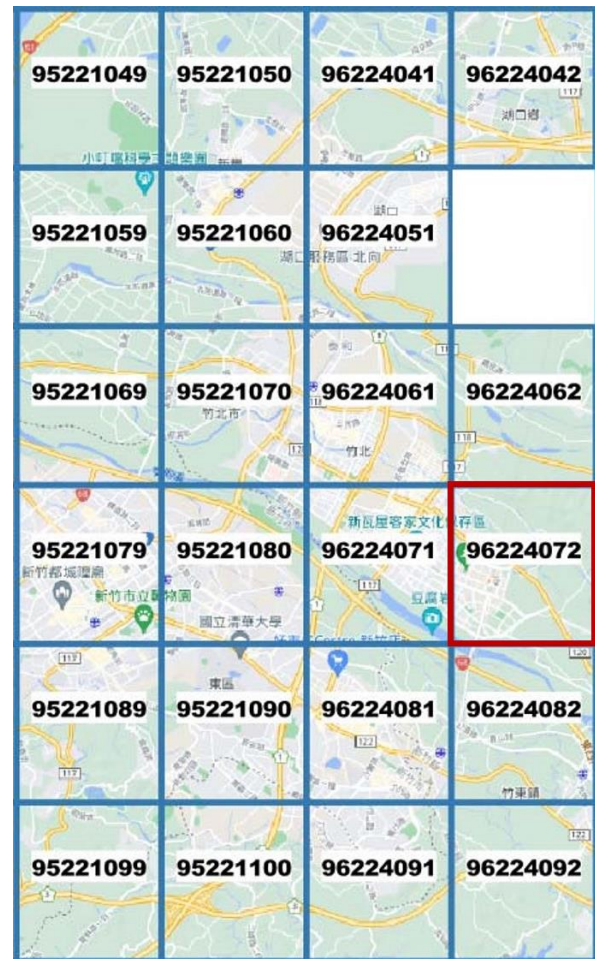

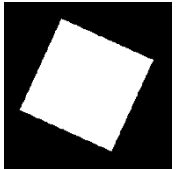


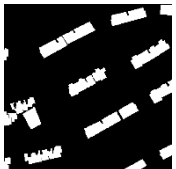

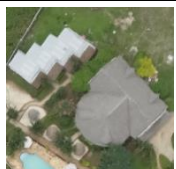
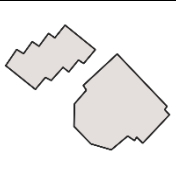



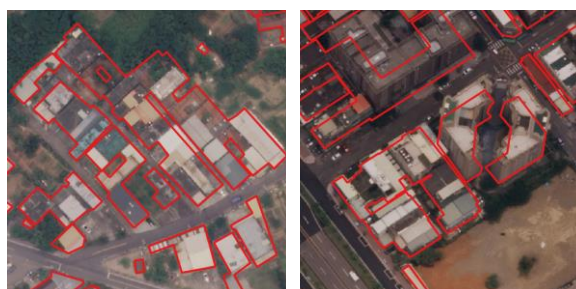
圖 1 研究範圍：23 幅五千分之一圖幅

建立深度學習所需的房屋標籤時，依框選建物範圍及位置不同，主要分為三種 (表 1)，第一種為框選建物地籍範圍，即建築物基地 (Footprint) (Maggiore *et al.*, 2017)；第二種為框選訓練影像中建物屋頂 (Ji *et al.*, 2019)；第三種為框選訓練影像中建物範圍 (Open AI Tanzania, 2018)。本文主要研究議題為探討是否可直接使用臺灣通用電子地圖產生房屋偵測訓練資料集，及分析此方法可能會遇到的問題，故採用之房屋標籤及影像皆來自臺灣通用電子地圖之建物圖層及對應航照正射影像。電子地圖的房屋範圍與第一種房屋標籤相似，以立體測圖方式繪製多邊形向量，圖徵範圍以整體的外框圈選建物區表示房屋區域，連棟屋未分戶 (圖 2)。彩色

航照正射影像的空間解析度為 25 cm，由於航照正射影像非真實正射改正影像 (True Orthophoto) (Chen *et al.*, 2007)，故正射影像的高樓會有嚴重的高差移位現象 (圖 2(b))，影像中房屋區域面積比實際房屋多邊形向量大。

表 1 房屋標籤種類

	正射影像	房屋標籤	正射影像套疊房屋標籤
第一種			
第二種			
第三種			



(a) 連棟屋 (b) 高差移位  
圖 2 臺灣通用電子地圖建物圖層套疊正射影像

### 3. 研究方法

本研究房屋辨識與分類方法採用 DeepLabV3+ (Chen *et al.*, 2018) 深度卷積類神經網路 (Deep Convolutional Neural Network, DCNN)，此影像語意分割 (Semantic Segmentation) 應用於偵測航照正射影像之房屋區域，變遷分析則比對前期臺灣通用電子地圖房屋區域及後期深度學習偵測房屋區域，以智慧辨識房屋變遷區域，整體流程如圖 3 所示。

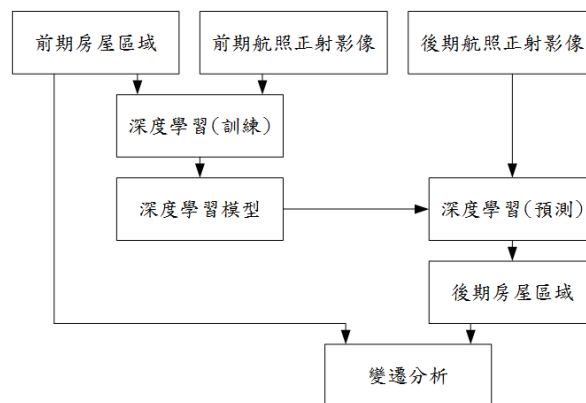


圖 3 智慧辨識房屋變遷區域流程圖

### 3.1 建立房屋偵測訓練資料集

本研究採用 DeepLabV3+模型以語意分割深度學習技術偵測房屋區域，語意分割屬於像元式分類，針對每個像元進行分類，因此訓練資料集須為像元等級 (pixel-level) 的標籤。為建立偵測房屋區域所需之訓練資料集，本研究僅選用臺灣通用電子地圖中的建物圖層於建立訓練資料集。資料集建立方法主要分為兩大部分，首先，把建物圖層由向量式 (Vector) 圖資轉換成空間解析度 25 cm 之網格式 (Raster) 圖資以對應航照正射影像，並使用數字儲存資料屬性，將網格式資料分為 0 及 1 兩類，0 代表非房屋 (non-building)；1 代表房屋 (building region)。接著，將每張航照正射影像等間隔裁切成許多 512 × 512 像元大小的區塊 (patch)，每個區塊之間上下左右皆重疊 50%，以避免房屋的邊緣被切除。此外，為剔除區塊範圍剛好落在河道或水庫等無房屋的區域，每一區塊須至少有 10% 的面積屬於房屋 (building region)，若區塊內房屋面積比例少於 10% 則不納入訓練。

經由前述處理，106 年 22 幅臺灣通用電子地圖及航照正射影像經過裁切，共產生 27,927 個區塊，其中 80% 作為訓練資料 (Training)，10% 作為驗證資料 (Validation)，其餘 10% 則作為獨立檢核資料 (Test)。1 幅 (圖幅號 96224072) 106 年及 23 幅 108 年度資料為獨立檢核資料，故無需進行裁切。

### 3.2 房屋區域偵測

房屋區域偵測適合以語意分割 (Semantic

Segmentation) 進行像元 (Pixel) 偵測, 本研究使用 DeepLab v3+ (Chen *et al.*, 2018) 偵測房屋區域。DeepLab v3+ 使用 FCN (Fully Convolutional Networks) 的概念, 並利用 Encoder-Decoder 架構萃取輸入資料的特徵, 以空洞卷積 (Atrous Convolutional) 得到多尺度特徵, 再融合多尺度資訊, 以提升輸出成果邊界之品質。圖 4 是使用訓練資料在不同疊代次數的精度 (Accuracy) 及損失函數 (Loss Function), 此圖展示經過 100 個時期 (epoch) 的訓練參數趨於收斂, 訓練、驗證及檢核的損失函數均下降且趨於平緩, 最終的檢核資料集 (Test) 的精度達 82%。

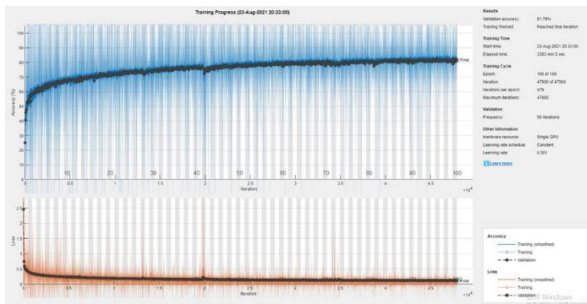


圖 4 房屋偵測深度學習訓練成果

### 3.3 房屋變遷偵測

房屋區域偵測以語意分割進行像元偵測, 偵測成果最小單元為像元, 然而考慮到現實生活中房屋變遷通常以物件 (Object) 為單位, 故需將屬於房屋區域的像元經過區域成長 (Region Growing) 後, 產生以物件為單元的房屋區域。

房屋變遷偵測為比較前期臺灣通用電子地圖的房屋物件及深度學習模型偵測後期的房屋物件。房屋變遷的型態可分為三類 (表 2), 若通用電子地圖的房屋物件與深度學習偵測到的房屋物件重疊, 即代表房屋沒有變遷, 反之, 兩者沒有重疊, 則代表房屋有變遷, 而變遷可能是拆除或新增房屋。由於通用電子地圖的房屋物件與深度學習偵測到的房屋物件面積不同, 不會完全一致重疊, 因此兩者重疊比例需達到 50% 以上, 方能判斷為無變遷。換言之, 本研究無法判識增建的區域, 主要偵測目標為獨立區域之拆除或新增。

表 2 房屋變遷型態

臺灣通用電子地圖 房屋物件	深度學習偵測 房屋物件	判斷
有	有	無變遷
有	無	變遷：拆除
無	有	變遷：新增

### 3.4 精度評估指標

房屋偵測精度評估以人工數化的臺灣通用電子地圖建物圖層為基準, 比較自動偵測成果; 房屋變遷偵測則是以臺灣通用電子地圖前後期變遷為基準, 比較自動變遷判識成果。

為量化深度學習模型偵測房屋區域與變遷之精度, 針對人工數化與深度學習模型偵測成果建立混淆矩陣 (Confusion Matrix) (圖 5), 並計算準確率 (Precision)、召回率 (Recall)、漏授率 (Omission)、誤授率 (Commission)、F1-Score 及 IoU (Intersection-over-Union) (式 1~6)。

用於房屋偵測精度分析的混淆矩陣中, TP (True Positive) 代表房屋被正確偵測為房屋; TN (True Negative) 代表非房屋被正確偵測為非房屋; FN (False Negative) 代表房屋被錯誤偵測為非房屋; FP (False Positive) 代表非房屋被錯誤偵測為房屋。

用於變遷偵測精度分析的混淆矩陣中, TP 代表人工數化變遷物件, 深度學習模型有偵測到改變; TN 代表人工數化無變遷之物件, 深度學習模型無偵測到改變, 此類物件不屬於變遷, 故不納入後續精度分析; FN 代表人工數化變遷物件, 深度學習模型無偵測到改變; FP 代表人工數化無變遷物件, 深度學習模型有偵測到改變。

$$\text{Precision} = \text{TP}/(\text{TP}+\text{FP}) \dots\dots\dots(1)$$

$$\text{Recall} = \text{TP}/(\text{TP}+\text{FN}) \dots\dots\dots(2)$$

$$\text{Commission} = 1-\text{Precision} \dots\dots\dots(3)$$

$$\text{Omission} = 1-\text{Recall} \dots\dots\dots(4)$$

$$\text{F1-Score} = 2/(1/\text{Precision} + 1/\text{Recall}) \dots\dots\dots(5)$$

$$IoU = TP/(TP+FP+FN) \dots\dots\dots (6)$$

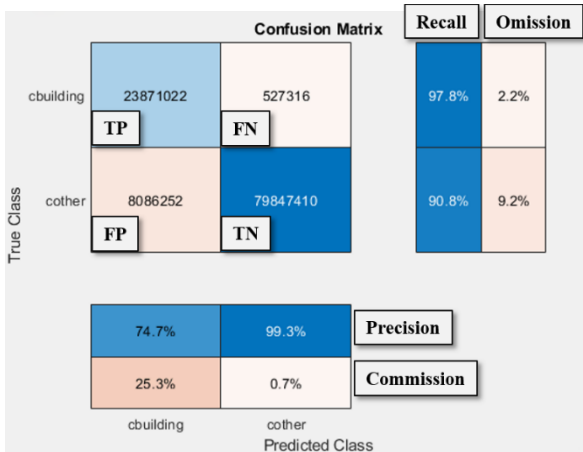


圖 5 混淆矩陣

與臺灣通用電子地圖 (含圖幅號 95221070, 不含 96224072) 進行訓練, 以 108 年圖幅號 95221070 正射影像預測 108 年圖幅號 95221070 範圍內之房屋區域。此種預測方式使用不同年度但範圍相同的訓練及預測資料, 目的為分析跨年度預測房屋的成效。方式(2) 使用 106 年 22 幅航照正射影像與臺灣通用電子地圖 (含圖幅號 95221070, 不含 96224072) 進行訓練, 以 108 年圖幅號 96224072 正射影像進行獨立檢驗, 預測 108 年圖幅號 96224072 範圍內之房屋區域。此預測方式之目的為分析獨立驗證區跨年度預測房屋的成效。

表 3 房屋區域偵測方式

		訓練資料	預測資料	預測成果
單圖幅	跨圖幅預測	106 年 22 幅臺灣通用電子地圖與航照正射影像	106 年圖幅號 95221070 正射影像	DL 預測 106 年圖幅號 95221070 房屋區域
			106 年圖幅號 96224072 正射影像	DL 預測 106 年圖幅號 96224072 房屋區域
	跨年度預測		108 年圖幅號 95221070 正射影像	DL 預測 108 年圖幅號 95221070 房屋區域
			108 年圖幅號 96224072 正射影像	DL 預測 108 年圖幅號 96224072 房屋區域
22 幅圖幅	跨圖幅預測	106 年 22 幅正射影像	DL 預測 106 年 22 幅房屋區域	
	跨年度預測	108 年 22 幅正射影像	DL 預測 108 年 22 幅房屋區域	

## 4. 實驗成果

### 4.1 房屋區域偵測分析

房屋區域偵測精度分析以臺灣通用電子地圖建物圖層為基準, 比較深度學習預測單圖幅及 22 幅圖幅的房屋區域, 以評估深度學習模型應用於偵測房屋之成效。

經由使用不同年度或圖幅範圍之訓練與預測資料, 可將房屋區域偵測方式分為「跨圖幅預測」及「跨年度預測」(表 3)。偵測單圖幅房屋依輸入的預測資料不同, 將「跨圖幅預測」再分為兩種預測方式, 方式(1) 使用 106 年 22 幅航照正射影像與臺灣通用電子地圖 (含圖幅號 95221070, 不含圖幅號 96224072) 進行訓練, 並以 106 年圖幅號 95221070 正射影像預測同年度圖幅號 95221070 範圍內之房屋區域。使用相同的圖資作為訓練及預測資料, 目的為探討正射影像中房屋高差移位現象對偵測精度的影響程度。方式(2) 使用 106 年 22 幅航照正射影像與臺灣通用電子地圖 (含圖幅號 95221070, 不含圖幅號 96224072) 進行訓練, 以 106 年圖幅號 96224072 正射影像進行獨立檢驗, 預測同年度圖幅號 96224072 範圍內之房屋區域。

「跨年度預測」同樣依預測資料不同分為兩種預測方式, 方式(1) 使用 106 年 22 幅航照正射影像

單一圖幅偵測成果中 (表 4、圖 6), 使用「跨圖幅預測」方式(1) 的漏授率為 2.16%, 誤授率為 25.30%。漏授率低的原因是訓練與預測資料重疊, 而誤授率高的原因是影像中房屋高差移位造成深度學習模型預測的房屋區域面積大於人工數化。由此可知, 房屋高差移位會造成約 25%的誤授。「跨圖幅預測」方式(2) 的漏授率提高至 10.48%, 誤授率

達 34.86%。深度學習模型找到的房屋區域比實際多，大部分誤授的範圍同樣來自影像中房屋高差移位，故可知獨立驗證區（指未參與任何訓練）的漏授會達到 10% 左右。

運用「跨年度預測」方式(1) 的誤授率及漏授率分別為 34.05% 及 7.74%，跨年度預測的漏授率低於 10%。比對「跨圖幅預測」方式(2) 的成果後可發現，同一區域在訓練模型時使用前期（106 年）資料，對預測該區域後期（108 年）房屋有正面效益。「跨年度預測」方式(2) 的誤授率高達 47.34%，漏授率提高至 12.98%。從精度分析成果可知，使用 22 幅 106 年臺灣通用電子地圖與航照正射影像訓練之深度學習模型的泛用性仍不足，故造成誤授的原因除了高差移位現象外，房屋訓練資料樣態不足也是原因之一。

表 4 單圖幅房屋偵測精度分析

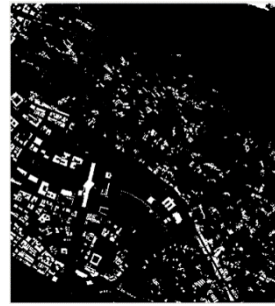
	跨圖幅預測		跨年度預測	
	(1)	(2)	(1)	(2)
TP/ FP/ FN/ TN (pixel)	23871022/ 8086252/ 527316/ 79847410	7890317/ 4222863/ 923696/ 99295124	23985096/ 12382406/ 2012607/ 77077171	8408594/ 7560130/ 1253857/ 97499951
準確率	74.70%	65.14%	65.95%	52.66%
誤授率	25.30%	34.86%	34.05%	47.34%
召回率	97.84%	89.52%	92.26%	87.02%
漏授率	2.16%	10.48%	7.74%	12.98%
F1-Score	84.72%	75.41%	76.92%	65.61%
IoU	73.48%	60.52%	62.49%	48.82%



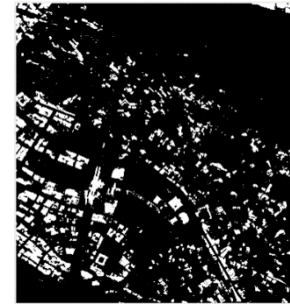
(a) 106 年臺灣通用電子地圖房屋區域 (圖幅號 95221070)



(b) 深度學習模型偵測 106 年房屋區域 (圖幅號 95221070)



(c) 106 年臺灣通用電子地圖房屋區域 (圖幅號 96224072)



(d) 深度學習模型偵測 106 年房屋區域 (圖幅號 96224072)



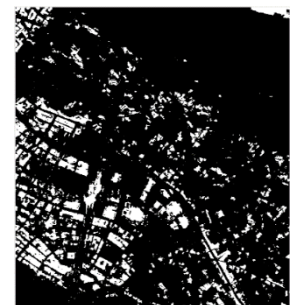
(e) 108 年臺灣通用電子地圖房屋區域 (圖幅號 95221070)



(f) 深度學習模型偵測 108 年房屋區域 (圖幅號 95221070)



(g) 108 年臺灣通用電子地圖房屋區域 (圖幅號 96224072)



(h) 深度學習模型偵測 108 年房屋區域 (圖幅號 96224072)

圖 6 臺灣通用電子地圖與深度學習模型偵測房屋區域 (黑色：非房屋；白色：房屋區域)

綜上所述，觀察單圖幅「跨圖幅預測」及「跨年度預測」之成果可知，使用 106 年圖資進行訓練，跨年預測 108 年房屋區域，這樣跨年度預測的成果所計算的誤授率高於漏授率。另外，比較四種偵測方式之成果，「跨圖幅預測」(1) 及「跨年度預測」(1) 這兩種方式預測成果的準確率與召回率等皆優於其他兩種方式。因此，建議房屋偵測之實務作業應選用與欲預測區域相同範圍且跨年度之訓練資料。

22 幅圖幅的房屋偵測成果中，使用 106 年 22

幅臺灣通用電子地圖與正射影像（含圖幅號 95221070，不含 96224072）進行訓練，以 108 年 22 幅正射影像預測 108 年房屋區域後（表 3），逐像元評估 22 幅偵測成果，精確率達 61.39%，召回率達 94.36%，漏授率為 5.64%，誤授率為 38.61%（表 5），與單圖幅偵測中「跨年度預測」方式(2) 相比，前者預測精度較佳。雖然兩者使用同樣的訓練資料，但後者所預測的圖幅範圍與訓練資料的圖幅範圍無重疊。這也表示本研究訓練之深度學習模型的泛用性仍有改善空間，建議於模型訓練階段增加房屋訓練資料的樣態。另外，為降低預測精度受模型推演能力的影響，若欲以後期（108 年）正射影像偵測房屋區域時，應使用相同範圍之前期（106 年）圖資作為深度學習模型的訓練資料。

表 5 22 幅圖幅房屋偵測精度分析

	跨圖幅預測	跨年度預測
TP/ FP/ FN/ TN (pixel)	379561047/ 126713288/ 7572620/ 1958446164	382916072/ 240818181/ 22896197/ 1879856138
準確率	74.97%	61.39%
誤授率	25.03%	38.61%
召回率	<b>98.04%</b>	<b>94.36%</b>
漏授率	1.96%	5.64%
F1- Score	84.97%	74.39%
IoU	73.87%	59.22%

## 4.2 房屋變遷偵測分析

房屋變遷偵測精度分析以臺灣通用電子地圖前後期變遷為基準，比較 106 年臺灣通用電子地圖與深度學習模型偵測 108 年房屋區域間之變遷區域（圖 7）。

單圖幅房屋變遷偵測分析是使用 106 年 22 幅航照正射影像與臺灣通用電子地圖（含圖幅號 95221070，不含 96224072）作為訓練資料，分別預測圖幅號 95221070 及 96224072 兩個不同圖幅範圍內 108 年房屋區域，再比較 106 年電子地圖及深度學習模型偵測 108 年房屋區域間之改變。22 幅圖幅房屋變遷偵測分析則以 106 年 22 幅航照正射影像與臺灣通用電子地圖（含圖幅號 95221070，不含

96224072）作為訓練資料，預測 22 幅圖幅範圍內 108 年房屋區域，再比較 106 年電子地圖及深度學習模型偵測 108 年房屋區域間之變化，以評估 22 幅圖幅房屋前後期變遷偵測之精度（表 6）。

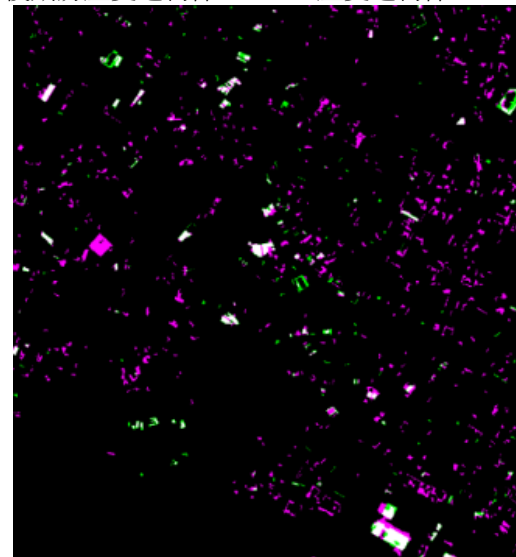
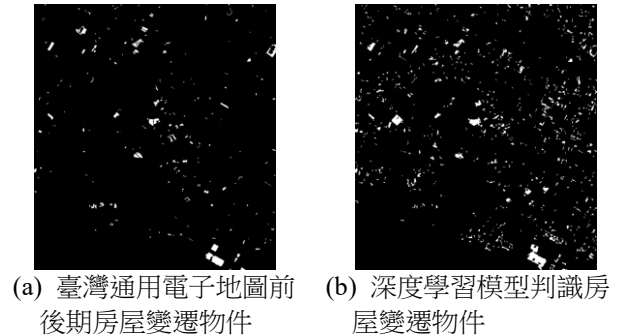


圖 7 人工數化及深度學習模型判識房屋變遷物件

本研究的精度分析僅針對變遷區域，並未納入無變遷區域，以避免高估精度。分析單圖幅或 22 幅圖幅房屋變遷偵測精度，以深度學習模型偵測圖幅號 95221070 範圍內房屋變遷的漏授率為 48.89%，誤授率為 83.38%；偵測圖幅號 96224072 範圍內房屋變遷的漏授率為 45.19%，誤授率為 86.42%（表 7、圖 8），可見不論訓練資料及預測資料的圖幅範圍是否重疊，前後期房屋變遷偵測的誤授率皆高於 80%。深度學習模型偵測 22 幅圖幅房屋變遷的誤授率及漏授率分別為 88.07% 及 47.10%，整體而言，誤授的比例高於漏授。研究成果顯示，雖然房屋偵測的召回率 (Recall) 達 90%，但變遷分析的召回率僅有 50% 左右。

表 6 房屋區域偵測方式

	訓練資料	預測資料	預測成果 (106 年臺灣通用電子地圖/ 108 年深度學習模型預測)
單圖幅	106 年 22 幅 臺灣通用 電子地圖 與航照正 射影像	108 年圖幅號 95221070 正射影像	圖幅號 95221070 房屋前後期變遷物件
		108 年圖幅號 96224072 正射影像	圖幅號 96224072 房屋前後期變遷物件
22 幅圖 幅		108 年 22 幅 正射影像	22 幅圖幅 房屋前後期變遷物件

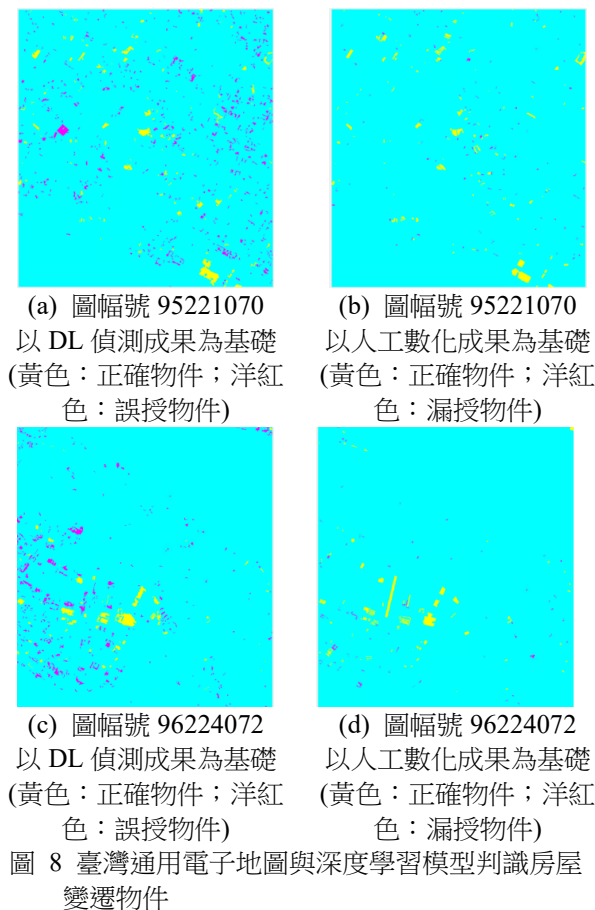
表 7 房屋變遷偵測精度分析

	單圖幅		22 幅圖幅
	95221070	96224072	
TP/ FP/FN (object)	115/ 577/ 110	74/ 471/ 61	1763/ 13013/ 1570
準確率	16.63%	13.58%	11.93%
誤授率	83.38%	86.42%	88.07%
召回率	51.11%	54.18%	52.90%
漏授率	48.89%	45.19%	47.10%
F1- Score	25.08%	21.76%	19.47%
IoU	14.34%	12.21%	10.79%

近一步探討房屋變遷偵測成果精度不佳的原因，主要可歸為四類 (表 8)。第一類為後期 (108 年) 正射影像中混凝土路面造成深度學習模型偵測錯誤，誤將混凝土路面當作房屋；第二類為後期正射影像因拍攝角度與光線等產生陰影，使深度學習模型將陰影認為是房屋變遷區域；第三類為後期影像中建物因高差移位效應，房屋範圍大於通用電子地圖的房屋範圍，導致深度學習模型將高差位移量視為變遷；第四類為深度學習模型有成功偵測到房屋變遷，但在精化變遷偵測成果過程中，面積小 (<16 m<sup>2</sup>) 的變遷區域會被過濾刪除，造成變遷偵測失敗。

表 8 房屋變遷偵測錯誤之類型

	106 年 正射影像	108 年 正射影像	變遷偵測成果 (以 DL 偵測成 果為基準) (黃色：正確； 洋紅色：誤授)
第一類			
第二類			
第三類			
第四類			



此外，觀察前後期房屋變遷偵測精度，及物件數量與面積統計成果 (表 7、表 9)，可發現變遷判識精確率與物件面積相關，實驗成果趨勢與 EuroSDR 房屋變遷偵測成果有一致的行為 (Champion *et al.*, 2009)。統計全區前後期房屋變遷偵測之物件數量與面積，多數被正確偵測的物件面



積較大，平均面積約 795.026 m<sup>2</sup>；面積中值約 312.937 m<sup>2</sup>，而多數被遺漏的房屋物件面積較小，平均面積約 147.719 m<sup>2</sup>，面積中值約 102.125 m<sup>2</sup>。由此可知，大面積物件被正確判識的比例較高，面積較小的區域容易被遺漏。

表 9 房屋變遷偵測成果之面積統計

	變遷偵測：正確		變遷偵測：遺漏	
	單圖幅	全區	單圖幅	全區
物件數量	115	1837 (1763+74)	110	1631 (1570+61)
物件平均面積 (m <sup>2</sup> )	592.635	795.026	123.858	147.719
物件面積中值 (m <sup>2</sup> )	255.625	312.937	102.750	102.125

推測原因為正射影像中建物受高差移位影響，使影像中房屋範圍大於以立體測繪方式產生之通用電子地圖房屋範圍，故利用深度學習模型偵測 108 年房屋區域之成果中，預測的房屋物件面積會大於真實房屋面積。考量到通用電子地圖的房屋物件與深度學習偵測到的房屋物件範圍不會完全重疊，再加上為減少將建物高差移位量誤判為變遷之機率，在判斷房屋有無變遷時，若深度學習偵測到的房屋物件與通用電子地圖的房屋物件面積重疊 50% 以上，該物件被判斷為無變遷。然而若新增或拆除之房屋物件面積小，容易與影像中受高差移位效應影響之區域重疊大於 50%，或是落在影響範圍內，如此將無法被正確判斷為變遷物件。

建議在深度學習模型訓練階段加入使用正射影像匹配產生之數值地表模型 (Digital Surface Model, DSM)，以減少因高差移位效應產生的問題。除此之外，倘若臺灣通用電子地圖的彩色正射影像改為以影像密匹配產生之近真實正射影像，消除地形與地物所產生之高差移位，應可大幅改善此問題並提升變遷偵測精度 (Ham *et al.*, 2021)。

## 5. 結論與建議

本研究利用航照正射影像與臺灣通用電子地圖建物圖層建立訓練資料集，以語意分割深度學習技術應用於房屋區域及變遷偵測，並建立混淆矩陣，量化深度學習模型偵測之精度。

就實驗成果可知採用臺灣通用電子地圖與航照正射影像，以語意分割深度學習進行房屋偵測成果之準確率 (F1-Score) 約 74%，召回率 (Recall) 達到 90%。房屋變遷偵測成果與房屋面積及高差移位相關，呈現使用非真實正射改正影像進行變遷分析有其極限。變遷偵測準確率不高的主要因為正射影像中房屋高差移位，使深度學習模型在訓練時因正射影像中房屋範圍與臺灣通用電子地圖中建物輪廓有落差 (圖 9(c)、(d))，降低深度學習模型偵測房屋成果之精度，進而影響前後期房屋變遷偵測成果。此外，不同年度之正射影像因高差移位量不同，前後期房屋範圍不一致 (圖 9(a)、(b))，導致深度學習模型在偵測變遷時誤將高差移位量的變化視為物件改變，也會降低變遷偵測成果之精度。使用單一影像資料進行偵測有其精度的限制，從資訊融合的角度出發，未來可整合數值地表模型，以降低高差移位的影響 (Teo, 2019)，以提升偵測的精度。



圖 9 房屋高差移位示意圖

## 致謝

本研究承蒙內政部計畫 (110CCL031C) 支持及內政部國土測繪中心提供測試資料得以順利完成，謹此致謝。

## 參考文獻

- 內政部國土測繪中心，2020。臺灣通用電子地圖測製更新作業說明(109 年度版)。[National Land Surveying and Mapping Center, 2020. Instructions for Taiwan electronic map updating (2020 Edition). (in Chinese) ]
- Champion, N., Rottensteiner, F., Matikainen, L., Liang, X., Hyypä, J., and Olsen, B.P., 2009. A test of automatic building change detection approaches, Proceedings of CMRT09, International Archives of Photogrammetry, Remote Sensing and Spatial Information Systems, Paris, France, 38 (Part 3-W4), pp. 145-150.
- Chen, L.C., Teo, T.A., Wen, R.Y., and Rau, J.Y., 2007. Occlusion-compensated true orthorectification for high resolution satellite images, Photogrammetric Record, 22(117): 39-52.
- Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11211, pp. 833-851.
- Chen, H., and Shi, Z., 2020. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection, Remote Sensing, 12(10): 1662.
- Diakogiannis, F.I., Waldner, F., and Caccetta, P., 2021. Looking for change? Roll the Dice and demand Attention, Remote Sensing, 13(18): 3707.
- Ghaffarian, S., Kerle, N., Pasolli, E., and Jokar Arsanjani, J., 2019. Post-disaster building database updating using automated deep learning: An integration of pre-disaster OpenStreetMap and multi-temporal satellite data, Remote sensing, 11(20): 2427.
- Google Open Building, 2021. A dataset of building footprints to support social good applications, Available at: <https://sites.research.google/open-buildings/>, Accessed November 4, 2021.
- Ham, S., Kim, H., Park, K., Han, S., Chung, D., and Lee, I., 2021. Micro-level building change detection from true orthoimages using spatio-temporal attention mechanism, Proceedings of International Symposium on Remote Sensing (ISRS 2021), pp. 222-225.
- Ji, S., Wei, S., and Lu, M., 2019. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set, IEEE Transactions on Geoscience and Remote Sensing, 57(1): 574-586.
- Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P., 2017. Can semantic labeling methods generalize to any city? The inria aerial image labeling benchmark, Proceedings of 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, pp. 3226-3229.
- Open AI Tanzania, 2018. Open AI Tanzania Building Footprint Segmentation Challenge, Available at: <https://competitions.codalab.org/competitions/20100>, Accessed June 24, 2021.
- Teo, T.A., 2019. Deep-Learning for LOD1 building reconstruction from airborne Lidar data, IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019), Yokohama, Japan, pp.86-89.

- Van Etten, A., Lindenbaum, D., and Bacastow, T.M., 2018. Spacenet: A remote sensing dataset and challenge series, arXiv preprint arXiv:1807.01232.
- Wu, G., Guo, Z., Shi, X., Chen, Q., Xu, Y., Shibasaki, R., and Shao, X., 2018. A boundary regulated network for accurate roof segmentation and outline extraction, *Remote Sensing*, 10(8): 1195.
- Xu, J.Z., Lu, W., Li, Z., Khaitan, P., and Zaytseva, V., 2019. Building damage detection in satellite imagery using convolutional neural networks, arXiv preprint arXiv:1910.06444.
- Yang, H.L., Yuan, J., Lunga, D., Laverdiere, M., Rose, A., and Bhaduri, B., 2018. Building extraction at scale using convolutional neural network: Mapping of the united states, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8): 2600-2614.
- Yuan, J., 2017. Learning building extraction in aerial scenes with convolutional networks, *IEEE transactions on pattern analysis and machine intelligence*, 40(11): 2793-2798.

# Building Detection from Aerial Orthoimage Using Deep Learning Technology

Tee-Ann Teo <sup>1\*</sup> Yu-Ju Fu <sup>2</sup>

## Abstract

Building model is an essential element in a topographic map. In order to improve the automation of building extraction, this research uses deep learning technology to identify building regions and changed areas from multi-temporal aerial orthoimage. The building polygons from Taiwan e-Map and corresponding aerial orthoimage are combined to generate the training dataset for deep learning. The building regions are automatically predicted by multispectral orthoimage using convolutional neural network. Then, the change detection compares bi-temporal building regions from deep learning in two seasons. The experimental results indicated that the F1-score and recall in building detection were 74% and 90%, respectively. The error is mainly caused by the relief displacement of building. Moreover, the accuracy of change detection is mainly related to the size of building area.

**Keywords:** Building Detection, Change Detection, Deep Learning, Semantic Segmentation

---

<sup>1</sup> Professor, Department of Civil Engineering, National Yang Ming Chiao Tung University

<sup>2</sup> Associate Engineer, Disaster Prevention and Water Environment Research Center,  
National Yang Ming Chiao Tung University

\* Corresponding Author, E-mail: tateo@nycu.edu.tw

Received Date: Nov. 08, 2021

Revised Date: Nov. 29, 2021

Accepted Date: Dec. 02, 2021