



# 無人機測繪技術：智慧人孔偵測之精度評估

張智安<sup>1,\*</sup> 招博允<sup>2</sup> 謝文傑<sup>2</sup>

<sup>1</sup> 國立陽明交通大學土木工程學系教授

<sup>2</sup> 國立陽明交通大學土木工程學系碩士

## 摘要

人孔是管線資料庫中的必要圖資，本研究之目的是為了提高人孔調查的效率，提出採用無人飛行載具（Unmanned Aerial Vehicle, UAV）攝影測量方式進行人孔調查，並使用深度學習技術來自動偵測和分類人孔；同時也對無人機測繪技術和智慧製圖技術在人孔調查的精度上進行評估。研究中採用 Mask Region-Based CNN（Mask-RCNN）深度學習技術對無人機正射圖像進行人孔智慧識別，以獲取圖像中的人孔區域，並從正射圖像的地理坐標獲得人孔所在的地理坐標，以達成人孔測繪的目的。實驗成果顯示，採用 2 cm 空間解析度的無人機多光譜正射影像，利用影像分割深度學習進行人孔智慧識別，獨立驗證區的整體 F1-Score 可達 92%、召回率為 90%、準確率為 94%，偵測成果之誤授及漏授均小於 10%，而自動化萃取及人工量測成果之坐標均方根誤差小於 10 cm。因此，結合無人機攝影測量和人工智慧技術，具有達成人孔智慧判識應用的良好潛力。

**關鍵詞：**人孔偵測、深度學習、無人機正射影像、攝影測量、製圖

---

\* 通訊作者：張智安

電子郵件：tateo@nycu.edu.tw

（收件日期：2023 年 2 月 9 日；修正日期：2023 年 4 月 1 日；接受日期：2023 年 4 月 7 日）



# UAV Photogrammetry Technology: Accuracy Assessment of Intelligent Manhole Detection

Tee-Ann Teo<sup>1,\*</sup>, Po-Yun Chao<sup>2</sup>, Wen-Kiat Chia<sup>2</sup>

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

<sup>2</sup>Master, Department of Civil Engineering, National Yang Ming Chiao Tung University

## Abstract

This study aims to improve the efficiency of manhole extraction by proposing the use of Unmanned Aerial Vehicle (UAV) photogrammetry and deep learning techniques to detect and classify manholes automatically. We also evaluate the accuracy of intelligent manhole extraction using UAV photogrammetry in an independent check area. The study uses Mask Region-Based CNN (Mask-RCNN) deep learning technology to identify manholes in UAV orthoimages and to obtain the geographic coordinates of the manholes. The experiment used UAV multispectral orthoimages with a 2 cm spatial resolution for manhole extraction. The experimental results achieved an overall F1-Score of 92%, a recall rate of 90%, and a precision of 94%, with a detection error of less than 10%. The root-mean-square-error of the automated extraction and manual measurement results are less than 10 cm. Therefore, the combination of UAV photogrammetry and artificial intelligence technology has great potential for achieving the goal of automatic manhole extraction.

**Keywords:** manhole extraction, deep learning, UAV orthoimage, photogrammetry, mapping

---

\* Corresponding Author: Tee-Ann Teo  
E-mail: tateo@nycu.edu.tw

## 壹、緒論

隨著數位科技的發展，測繪業務的作業方法和程序也隨之改變。提高圖資品質、增加更新頻率、減少人力需求是測量技術發展的重要方向之一。無人飛行載具系統（Unmanned Aircraft Systems, UAS）是新興的量測資料收集技術，無人飛行載具（Unmanned Aerial Vehicle, UAV）的機動性強，已廣泛應用於測量等領域。利用 UAV 搭配攝影測量技術可以快速獲得地面正射拼接影像，結合無人機影像和智慧判識技術，可以快速提供測量製圖所需的資訊。

「三維管線」是國土資訊系統重要的發展工作之一，而人孔（Manhole）是三維管線資料庫中必要的圖資，在內政部國土測繪中心（2022）《一千分之一數值航測地形圖測製作業規定》、內政部營建署（2020）《公共設施管線資料標準》，以及臺灣資通產業標準協會（2020）《高精地圖圖資內容及格式標準》中，人孔是必要的圖徵，由於路面人孔數量龐大，需花費許多人力進行數化，因此有必要發展自動化人孔偵測程序。

本研究整理三項人孔調查測量作業相關之規範，分析人孔圖徵類別及其量測精度規範，以探討使用無人機攝影測量進行人孔智慧判識的可行性。

一、依內政部國土測繪中心（2022）《一千分之一數值航測地形圖測製作業規定》，道路上（含人行道）之電力人孔、電信人孔、自來水人孔、雨污水下水道人孔、電線桿（電力、電信）、路燈等均需測繪，手孔及管線則免予施測。依內政部規定，探討人孔測量所需的量測精度，地形地物之幾何精度檢查中，抽查點位重複量測之平面位置與原平面位置之均方根誤差不大於 25 cm，或地物點間之相對距離與原距離之均方根誤差不大於 40 cm。

二、內政部營建署（2020）《公共設施管線資料標準》第二版之適用範疇為國內具有二維或三維數值坐標描述之設施與管線資料，包括歸屬於公共管線之電信管線、電力管線、自來水管線、下水道、瓦斯管線、水利管線、輸油管線、綜合管線等八大類管線，以及非屬公共設施管線之工業用管線。孔蓋中心坐標施測成果之坐標誤差應小於  $\pm 20$  cm；孔蓋蓋頂高程測量，其高程為間接高程，施測成果之誤差應小於  $\pm 20$  cm。

三、在《高精地圖圖資內容及格式標準》（臺灣資通產業標準協會，2020）中，人孔的類別包含孔蓋（holeCover）及水溝蓋（drainCover）兩大類，其精度要求為孔蓋中心平面坐標誤差應小於  $\pm 20$  cm，孔蓋中心三維坐標誤差應小於  $\pm 30$  cm。

綜合以上三項人孔測量相關製圖的規範，現行無人機攝影測量製圖作業可獲取空間解析度小於 5 cm 的影像，且空中三角測量作業可達到 10 cm 以內的均方根誤差（內政部國土測繪中心，2020），因此有機會結合無人機攝影測量與人工智慧技術，納入人孔測量實施工作之智慧整合應用。

自動化人孔偵測可分為傳統機器學習及深度學習技術，Alshaiba, Núñez-Andrés, and Lantada（2020）利用車載地面光達反射特性及機器學習分類人孔，經由設定反射值門檻選取人孔蓋候選點，再把候選點合併為人孔區域，在良好的光達反射值條件下，可以達到 90% 以上的精度。Yu, Guan, and Ji（2015）亦利用車載地面光達反射特性，但使用深度學習技術萃取高階的特徵，並利用隨機森林分類器分類人孔，其人孔分類精度亦可達到 90% 以上。除了車載地面光達資料，航照影像也可以應用於人孔偵測，Commandre, En-Nejjary,

Pibre, Chaumont, Delenne, and Chahinian (2017) 將航照影像截切為許多  $40 \times 40$  pixels 的區塊，以 Convolutional Neural Network (CNN) 分類方法對大量的區塊進行分類，以區別人孔及非人孔區塊，因受限於航照影像空間解析度，其分類精度約為 60%。由於航照影像空間解析度會影響人孔偵測成果，Liu et al. (2019) 發展了 Multi-level Convolution Matching Network (M-CMN)，以多尺度影像偵測航照影像中人孔所在之矩形框位置，其人孔分類精度可達到 90% 以上，由於是採用矩形框偵測人孔，仍缺乏人孔邊界資訊。

由於空載影像空間解析度是影響人孔偵測成果的關鍵因子，本研究採用無人機高空間解析度影像，並以深度學習影像分割演算法自動化萃取人孔邊界。本研究之目的為無人機測繪技術及智慧製圖技術於人孔偵測之精度評估，研究中應用深度學習模型進行智慧人孔偵測，以獨立檢核區評估無人機正射影像之人孔偵測率及人孔中心點定位精度。

## 貳、研究資料

本研究實驗範圍位於臺中水湳經貿園區，實驗區域總面積約 84 公頃。無人機正射影像空間解析度為 2 cm，空中三角測量定位精度優於 10 cm，定位精度上滿足人孔調查測量作業之規範要求。實驗區共分割為 36 個圖幅（如圖 1），每個圖幅為  $155 \text{ m} \times 155 \text{ m}$ ，其中 35 幅作為深度學習模型之訓練及驗證資料，將訓練過程中，訓練與驗證影像數量比例為 8:2；另 1 幅則用於獨立驗證分析。本研究採用 1 張未參與任何訓練參數計算的影像作為獨立檢測區域 (Independent Check Area)，以驗證訓練模型的精度。



圖 1 研究範圍：36 幅無人機正射影像

註：紅框為獨立驗證區。

資料來源：本研究自行產製。

## 參、研究方法

本研究採用 Mask Region-Based CNN (Mask-RCNN) 深度學習技術進行人孔智慧判識，此實體分割 (Instance Segmentation) 技術應用於無人機正射影像，以獲取影像中人孔區域，並從正射影像的地理坐標獲得人孔所在的地理坐標，以達成人孔測繪之目的。

### 一、建立人孔偵測訓練資料集

本研究深度學習所需的人孔標籤經由人工數化取得，因受限於無人機影像的空間解析度，難以在影像上辨別不同孔蓋之紋理，故本研究僅以形狀區分孔蓋與水溝蓋。孔蓋部分細分為道路區域之孔蓋及非道路區域之孔蓋 (holeCover2)，因此人孔共有三個類別 (如表 1)。透過人工數化無人機正射影像之坐標位置，獲得原始人孔資料共 4,539 筆。為增加訓練資料，在訓練深度學習模型前將每幅影像切割成  $512 \times 512$  pixels 大小的區塊 (Patch) 影像，並使區塊影像之間重疊 50%，以 50% 重疊率進行資料增益 (Data Augmentation) 後，完成人孔訓練資料建置。「區塊影像之間重疊 50%」是為了避免 1 個人孔被分割在兩個不同訓練影像中，在進行訓練時，會對輸入的訓練影像集進行資料增益，資料增益參數包含平移、旋轉、尺度等參數。人孔資料增益成果如圖 2 所示，36 幅影像共產生 12,888 張區塊影像，其中 36 幅影像之 10,973 張區塊影像作為訓練及檢驗影像，另 1 幅影像之 1,915 張區塊影像作為獨立驗證之精度評估影像。

### 二、人孔偵測

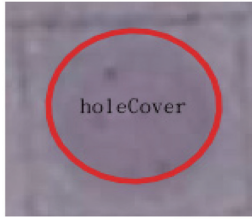
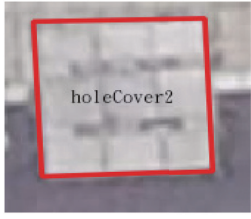
Mask-RCNN (He, Gkioxari, Dollár, and Girshick, 2017) 是基於 RCNN (Girshick, Donahue, Darrell, and Malik, 2014) 及 Faster-RCNN (Girshick, 2015) 演進而來，通過在卷積神經網路中預測目標邊界框和分割遮罩來提高目標檢測精度。RCNN 使用 Selective Search 演算法，從輸入影像中隨機萃取大量的候選區塊 (Region Proposal) (如：2,000 個)，再將萃取的候選區塊重新取樣為長寬比一致的矩形影像 (如： $224 \times 224$  pixels)。接著，利用 CNN 深度學習模型，對重新取樣後的候選區塊進行特徵 (Feature) 萃取，進一步利用支持向量機 (Support Vector Machine, SVM) 分類器對特徵進行分類，並對分類為目標物的外框進行線性回歸 (Bounding-Box Regression)，最終獲得目標物的外框 (Boundary Box)。

由於 RCNN 的計算費時，Fast-RCNN 是為了提升 RCNN 運算效能而發展的演算法，RCNN 的大量重疊候選區塊在 CNN 特徵萃取過程中，需要個別運算許多重複的區域。而 Fast-RCNN 只進行一次整張影像的 CNN 特徵萃取，再將 CNN 萃取到的影像特徵與 2,000 個候選區塊整合，利用 Region of Interest Pooling (RoI Pooling) 將取出的候選區塊對應到 CNN 的 Feature Map，輸出後各自接上全連接層 (Fully Connected Layer) 進行 Softmax 分類及線性回歸，最終獲得目標物的外框。

Fast-RCNN 已提升 RCNN 的 CNN 特徵萃取之運算效能，而 Faster-RCNN 則是進一步提升 RCNN 的 Selective Search 演算法萃取 Region Proposal 區塊之運算效能。相較於 RCNN 的 Selective Search 演算法，Faster-RCNN 使用 Region Proposals Network (RPN) 搭配 Anchor box



表 1 人孔類別

| 項目          | 類別 (classes)                                                                      |                                                                                    |                                                                                     |
|-------------|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
|             | 水溝蓋                                                                               | 孔蓋                                                                                 | 非道路區域之孔蓋                                                                            |
| 特徵樣本        |  |  |  |
| 原始數量 (筆)    | 706                                                                               | 1,569                                                                              | 2,264                                                                               |
| 增益後人孔資料 (張) | 1,953                                                                             | 4,064                                                                              | 6,871                                                                               |

資料來源：本研究自行產製。

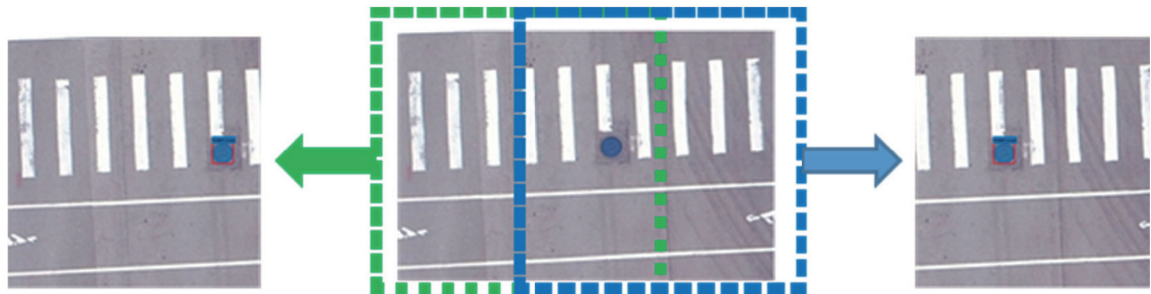


圖 2 影像切割後人孔資料將獲得增益

資料來源：本研究自行產製。

提升 RCNN 的效能。Mask-RCNN 利用 RoI Align 技術對邊界框內的資訊重新取樣，RoI Align 則是從 RoI Pooling 演進而來。RoI Pooling 在計算外框時會直接取整數位置，造成外框位置的錯位；RoI Align 則不直接取整數位置，而是採用雙線性內插的方式來獲取位置，故有較佳的效果，並使用多重損失函數 (Loss Function) (包括 Lcls 分類損失、Lloc 回歸損失及 Lmask 遮罩損失) 訓練網路。<sup>1</sup> 此圖是使用訓練資料在不同疊代次數的 Total Loss (如圖 3)，展示經過 5,000 個 epoch 的訓練參數後趨於收斂，訓練及驗證整體損失函數均下降且趨於平緩，最終的 Total Loss 低於 0.5。圖 4 為本研究 Mask-RCNN 訓練模型預測人孔偵測之成果範例。

### 三、人孔精度評估

人孔偵測精度評估以人工數化的人孔為基準，以不同量化精度指標評估偵測成果，並計算準確率 (Precision)、召回率 (Recall)、漏授率 (Omission)、誤授率 (Commission) 及 F1-Score (如公式 1~5)。

<sup>1</sup> 更多關於 Mask-RCNN 的細節，請參考 He et al. (2017)。

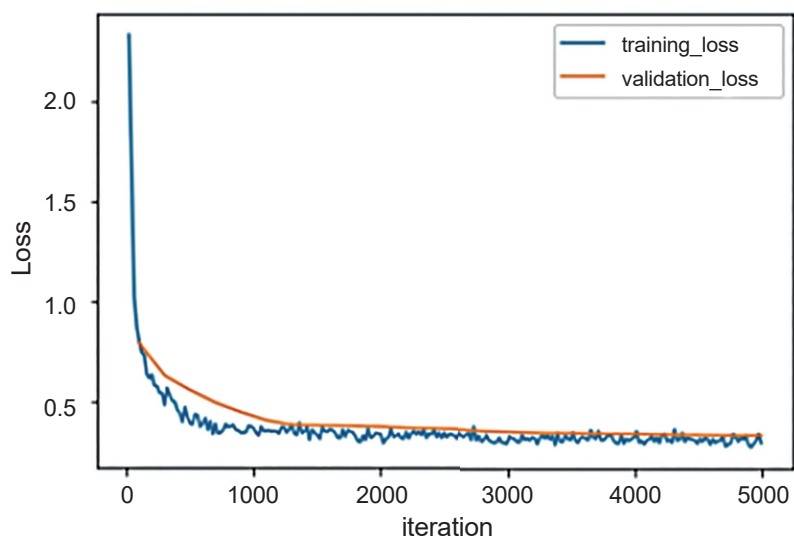


圖 3 深度學習訓練成果

資料來源：本研究自行產製。

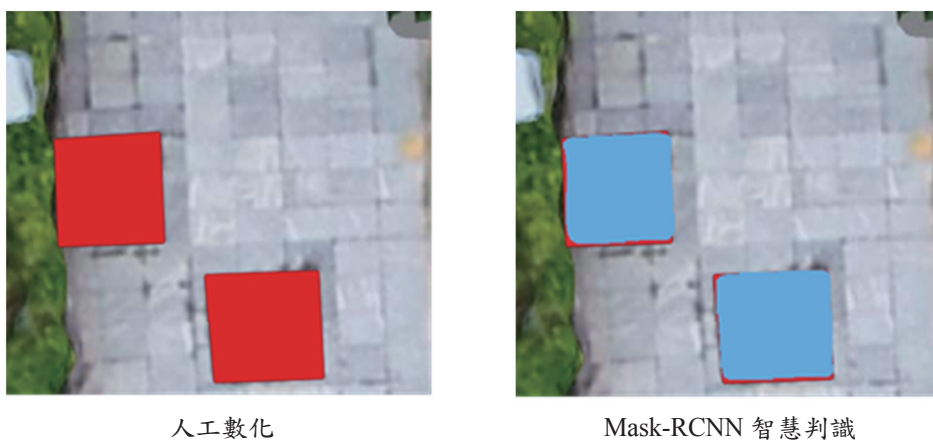


圖 4 人孔智慧判識成果示意圖

資料來源：本研究自行產製。

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (1)$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (2)$$

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

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

$$\text{F1 - Score} = \frac{2}{\left(\frac{1}{\text{Precision}}\right) + \left(\frac{1}{\text{Recall}}\right)} \quad (5)$$

其中，*TP* (True Positive) 代表人孔被正確偵測為人孔；*FP* (False Positive) 代表非人孔被錯誤偵測為人孔；*FN* (False Negative) 則代表人孔被錯誤偵測為非人孔。

## 肆、實驗成果

本研究使用 Mask-RCNN 深度學習模型對 35 張圖片進行模型訓練，通過資料擴充技術增加訓練資料，並利用遷移學習訓練深度學習模型，再以 1 個獨立檢測圖幅進行預測及精度分析。表 2 為深度學習模型的預測結果，在獨立檢測圖幅預測中，所有人孔類別的漏授和誤授率均小於 10%；在人孔類別中，水溝蓋的識別結果最佳；而道路孔蓋的結果較差。分析人孔漏授的情況，多數發生在人孔表現有被其他物體遮蔽的情況（例如：沙塵），而誤授則是連續人孔分割不完全所造成。

表 2 人孔偵測之成果（獨立驗證）

| 類別       | 人孔是否被成功偵測     |                 |                 | 自動萃取人孔是否正確       |                  |                  | F1-Score |
|----------|---------------|-----------------|-----------------|------------------|------------------|------------------|----------|
|          | 召回率<br>Recall | 漏授率<br>Omission | 漏授率<br>Omission | 準確率<br>Precision | 誤授率<br>Commision | 誤授率<br>Commision |          |
| 所有類別     | 196/217       | 21/217          | 9.68%           | 196/209          | 13/209           | 6.22%            | 92.02%   |
| 水溝蓋      | 67/69         | 2/69            | 2.90%           | 67/69            | 2/69             | 2.90%            | 97.10%   |
| 孔蓋       | 9/14          | 5/14            | 35.71%          | 9/16             | 7/16             | 43.75%           | 60.00%   |
| 非道路區域之孔蓋 | 118/134       | 16/134          | 11.94%          | 118/124          | 6/124            | 4.84%            | 91.47%   |

資料來源：本研究自行產製。

無人機正射影像使用 TWD97 橫麥卡托二度分帶 (TM2) 投影系統，通過智慧判識可成功偵測人孔區域，並可計算人孔在物空間的坐標位置。通過比較人工數化和智慧判識人孔之間的中心點坐標，計算中心點坐標之間的差值以分析識別精度。以人工數化的人孔形心作為參考，統計利用深度學習模型自動偵測和人工數化的人孔形心之間的距離差異量（如表 3），從統計結果得知，自動化探測與人工數化結果的均方根誤差小於 5 cm，此誤差可以滿足人孔圖徵之精度要求。



表 3 智慧判識與人工數化資料之人孔形心的距離差異量

| 類別       | Mean    |         | RMSE    |         |
|----------|---------|---------|---------|---------|
|          | dE (cm) | dN (cm) | dE (cm) | dN (cm) |
| 所有類別之人孔  | 1.00    | -1.22   | 3.57    | 3.62    |
| 水溝蓋      | -0.46   | -1.35   | 3.30    | 3.64    |
| 孔蓋       | 1.26    | -1.15   | 2.94    | 2.14    |
| 非道路區域之孔蓋 | 1.65    | -1.05   | 4.45    | 3.57    |

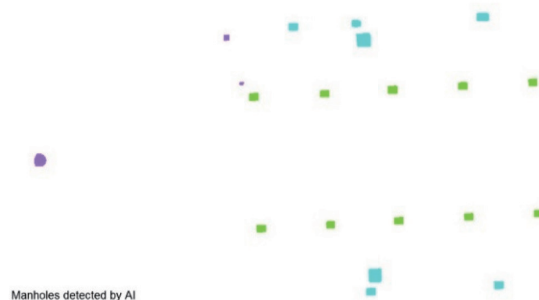
註：RMSE：均方根誤差（Root Mean Squared Error）。

資料來源：本研究自行產製。

總體實驗結果與分析，無人機影像與人工智慧技術的智慧整合可輔助「人孔調查」測量施測作業，經由人工智慧識別人孔區域及類別，再進一步以邊緣線正規化技術獲取矩形及圖形的人孔向量資料，以建立人孔圖徵的幾何及屬性資料（如圖 5）。兩者具有很高的一致性，因此智慧判識人孔之成果可供後續測量製圖運用。



(a) 無人機正射影像



(b) 人工智慧預測之網格式人孔



(c) 人工智慧預測之人孔多邊形



(d) 人工數化之人孔多邊形

圖 5 人孔智慧判識成果

資料來源：本研究自行產製。

## 伍、結論

本研究利用無人機正射影像與人工數化人孔資料建立人孔 AI 訓練資料集，以 Mask-RCNN 影像分割深度學習技術進行人孔智慧判識應用，並利用獨立檢核區量化深度學習模型偵測之精度。就實驗成果可知：採用 2 cm 空間解析度的無人機多光譜正射影像，以影像分割深度學習進行人孔偵測成果之整體 F1-Score 可達 92%，召回率達到 90%，準確率達到 94%。實驗成果顯示獨立驗證區的偵測成果之誤授及漏授均小於 10%，自動化萃取及人工量測成果之坐標均方根誤差小於 10 cm，結合無人機攝影測量及人工智慧技術，有很好的潛力達成人孔智慧判識之應用目的。

## 致謝

本研究承蒙內政部計畫支持（112PL024A）及臺灣世曦公司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 年 12 月 20 日，《「109 年度多旋翼機航拍與影像處理及設備維護作業採購案」工作總報告》。
- 內政部國土測繪中心，2022 年 12 月 14 日，《一千分之一數值航測地形圖測製作業規定》。
- 內政部營建署，2020 年 1 月 13 日，《公共設施管線資料標準（文件編號：NGISTD-ANC-028-2020.01.13）》。
- 臺灣資通產業標準協會，2020 年 3 月 16 日，《高精地圖圖資內容及格式標準（文件編號：TAICS TS-0024 v1.0）》。
- Alshaiba O., Núñez-Andrés M. A., and Lantada N., 2020, “Automatic Manhole Extraction from MMS Data to Update Basemaps,” *Automation in Construction*, 113, 103110. doi:10.1016/j.autcon.2020.103110
- Commandre B., En-Nejjary D., Pibre L., Chaumont M., Delenne C., and Chahinian N., 2017, “Manhole Cover Localization in Aerial Images with a Deep Learning Approach,” paper presented at *ISPRS Hannover Workshop: HRIGI 17–CMRT 17–ISA 17–EuroCOW 17*, Hannover, Germany. doi:10.5194/isprs-archives-XLII-1-W1-333-2017
- Girshick R., 2015, “Fast R-CNN,” in *2015 IEEE International Conference on Computer Vision*, Santiago, Chile: IEEE, 1440-1448. doi:10.1109/ICCV.2015.169
- Girshick R., Donahue J., Darrell T., and Malik J., 2014, “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation,” paper presented at *Proceedings of the 2014 IEEE conference on computer vision and pattern recognition*, Columbus, OH.

- He K., Gkioxari G., Dollár P., and Girshick R., 2017, “Mask R-CNN,” paper presented at *Proceedings of the 2014 IEEE international conference on computer vision*, Venice, Italy.
- Liu W., Cheng D., Yin P., Yang M., Li E., Xie M., and Zhang L., 2019, “Small Manhole Cover Detection in Remote Sensing Imagery with Deep Convolutional Neural Networks,” *ISPRS International Journal of Geo-Information*, 8(1), 49. doi:10.3390/ijgi8010049
- Yu Y., Guan H., and Ji Z., 2015, “Automated Detection of Urban Road Manhole Covers Using Mobile Laser Scanning Data,” *IEEE Transactions on Intelligent Transportation Systems*, 16(6), 3258-3269. doi: 10.1109/TITS.2015.2413812