

# 基於多源城市地理大數據的都市區騎樓地圖建置方法及應用研究

## Research on the Construction Method and Application of Arcade Map in Metropolitan Area Based on Multi-source Urban Geography Big Data

葉宜儒\* 賴進貴† 郭巧玲‡

### 1. 研究背景及意義

#### 1.1 研究背景

騎樓建築，源於因應亞熱帶氣候烈日、驟雨之特性，而逐漸發展形成的特殊建築構造形式 [19]，廣泛分佈於中國華南（圖 1-a）、台灣（圖 1-b），新加坡（圖 1-c）和馬來西亞等地，甚至在歐洲、日本等國亦有類似的建築形式出現。在台灣，騎樓建築最早出現於 1886 年於台北城興建的“亭仔腳”建築樣式 [22]，歷經兩百多年的發展，已成為了市鎮型聚落的典型建築文化景觀 [15]。就功能而言，騎樓作為一種半戶外的空間，除避雨外，夏季可以有效阻擋太陽輻射熱，為行人在提供舒適度較佳的活動空間；冬季則可以減少冷空氣的進入，起到保溫的作用 [21]，可謂行人在城市中最佳的避風港 [18]。再者所謂步行者友善的可步行城市（Walkable City）的規劃思潮正成為主流，許多國家亦開始步入老齡社會，步行者對安全步行空間的需求正逐漸浮現，騎樓空間則是步行空間系統的重要組成元素之一 [20]。



(a) 光緒年間位於廈門市的騎樓



(b) 目前台灣的常見騎樓



(c) 位於新加坡的騎樓

圖 1 騎樓建築形式<sup>1</sup>

近年來，隨著資訊技術的蓬勃發展，為地理學研究帶來了新的機遇。Google 街景圖、興趣點（Point of Interest, POI），OSM（OpenStreetMap）等數據源的出現，大幅降

\*國立台灣大學地理環境資源學系碩士生

†國立台灣大學地理環境資源學系教授

‡中央研究院人文社會科學研究中心助研究員

<sup>1</sup>資料來源 (a)<https://kknews.cc/travel/p4233op.html> (b)<https://udn.com/news/story/7323/4532666> (c)<https://cn.nytimes.com/real-estate/20160327/t27iht-resingapore/zh-hant/>

低了資訊收集的成本 [17]，但也造成了資料冗餘的問題，如何對龐大數據進行處理與分析並加以應用亦成爲了新的課題。以機器學習 (Machine Learning) 爲代表的人工智慧 (Artificial Intelligence) 技術與高性能計算 (High Performance Computing) 的日趨成熟，使得提取隱藏在大數據背後的資訊和知識成爲可能 [8]，例如在同樣依賴大數據分析的醫療和氣象領域中：NVIDIA 實驗室 Myronenko[11] 提出了一個根據大量病例數據訓練的新型胶质瘤分割網路，可以輔助醫師進行神經膠質瘤的診斷；Li-Chiu Chang 等人 [4]，使用歷年颱風氣象觀測資料訓練神經網路，實現了根據颱風路徑對可能產生的風暴增水進行預測……

而關於都市規劃和都市景觀的相關研究，劍橋大學的研究團隊在 2015 年首次提出的 SegNet 模型，可以高效地識別出人眼視角圖像數據中的天空、人行道、車道、建築、綠化等共計 12 種要素 [2]。受到其啓發，後續還有 DeepLab[5]，OCRNet[13]，等一系列模型相繼問世 [7][5]，基於這類深度學習模型，可以高效而準確地從街景圖像中提取需要的元素，以滿足各項城市研究的需要。

Gaurav Mittal 等 [10] 和魏書法，程章林 [23] 先後通過卷積神經網路 (Convolutional Neural Network, CNN) 和改進 Faster R-CNN 的模型，實現了從城市場景中自動識別垃圾，可以爲市容管理降低人力成本；張澤軒等 [16] 則使用改進後的“YOLO”模型，實現對城市道路中車輛與行人的檢測，爲分析交通安全和隱患提供了信息……由此可見，以機器學習爲代表的人工智慧技術已經被廣泛的應用到醫療、氣象、都市規劃等多個行業中，並已取得了良好的成效，其在地理大數據分析上的潛力值得期待。

## 1.2 問題的提出

騎樓空間在包括台灣在內的許多東南亞地區是重要的文化符號，同時也是都市人行環境的重要組成部分，如果管理良好，可以爲行人提供遠優於露天人行道的行走體驗，尤其在酷暑或暴雨等天候下。然而在包括 Google 地圖在內的常見電子地圖中，均缺乏對騎樓位置的完整標識。在 OSM 數據中，雖然可以使用如：

building = roof<sup>1</sup>

的句法 (key=value)，通過標籤搜索出近似於騎樓的遮簷建物。但據此參數在全台範圍內也只檢索出不足一千筆<sup>2</sup> 資料，顯然存在大量遺漏，並且其中大多數係離散的面狀資料。另筆者結合 Google 地圖和其街景圖對檢索結果進行複查發現，其中大多數並不屬於騎樓，而是加油站、涼亭等半開放式建築 (圖2)，可以說幾乎沒有騎樓被正確的標註。這也就造成了目前針對徒步的路徑規劃，只考慮了鋪裝路面的可達性，並未針對天候、季節和個人喜好等因素做特別的優化。

<sup>1</sup>OpenStreetMap Wiki. (2019, February 23). Zh-hant:Tag:building=roof - OpenStreetMap Wiki. Zh-hant:Tag:building=roof. Retrieved November 5, 2020, from <https://wiki.openstreetmap.org/wiki/Zh-hant:Tag:building%3Droof>

<sup>2</sup>筆者使用 <https://overpass-turbo.eu/> 網頁進行檢索，檢索日期：2020.11.05



圖 2 景隆街 36 巷內一處<sup>3</sup>

而針對台灣地區的現行法規而言，不論是中央法規還是各縣市政府根據自身情況在都市計畫中作出的規定，大都屬“騎樓或無遮簷人行道”這樣的選擇性陳述，因此單純的根據都市計畫和法規等進行推算騎樓的分佈存在盲點。並且在台灣，騎樓通行空間被挪作他用（停放汽車、機車，擺放貨物，甚至安裝卷閘門）的情況極為普遍，嚴重影響行人的使用體驗。

綜上所述，騎樓地圖的存在可以成為步行路線導航優化的重要參考因素之一。從宏觀尺度上看，也可以成為都市計畫的參考。然而因為騎樓空間的特殊性，難以從航照或者衛星影像中直接判釋，而傳統依託人力的調查，在龐大的資料量體面前也顯得力不從心，目前尚缺乏一種可以高效調查騎樓分佈及使用情況的有效方法，本研究將試圖填補這個空缺。

### 1.3 研究意義

本課題擬利用實驗區內街景圖像結合其他地理大數據（如 OSM，社群網路評價等），通過機器學習等城市大數據主流計算技術，探索從街景圖像中辨識出騎樓並評估騎樓空間有無被占為他用的方法。最後期待能使用此方法，為實驗區建置一套都市騎樓地圖，並標記出騎樓空間的不當使用現狀。在動線設計，城市管理及城市規劃等領域，將發揮重要的作用。具體而言，本研究的意義可分為以下兩方面：

#### (1) 學術價值

本研究涉及地理學，計算機科學，建築學和都市計畫學等多學科的融合。研究側重於開發從街景圖像為代表的城市地理大數據中獲取都市中騎樓位置及其使用現狀的方法，提出相應的數據集和模型，使之更適合用於騎樓空間的辨識，擴展地理資訊系統在城市管理中的應用。

同時本研究也在試圖探討地理資訊萃取（Geographic Information Extraction）的議題，即從海量以街景圖像為代表的地理大數據中，萃取騎樓資訊並加以利用，以求能進一步擴展地理信息的價值。

#### (2) 現實意義

從宏觀視角看來，本研究成果將可以輔助政府的施政決策，作為規劃“可步行城市”的參考。也將使之成為評估騎樓不當使用的高效工具，還騎樓空間於行人，一定程

<sup>3</sup>筆者製作

度上也可以為身障人士創造更為友善的通行環境。就個人而言，可以優化當前步行導航的算法，為行人在驟雨或酷暑的天氣下，規劃更舒適更安全的徒步動線，增進人們選擇步行作為出行方式的意願，間接促進永續發展。

## 2. 研究流程

總體上本研究將分為三階段，預計耗時 18~20 個月，具體流程如下：

### 2.1 前期研究階段（~2021.03）

本階段將側重於對研究動機、目的等進行梳理，了解相關領域的研究進展，整理文獻。因考慮到本計畫是針對技術方法進行開發和驗證的研究，故實驗區域範圍需要有一定的規模，因此我們選擇了城市環境多樣的台北市作為本次研究的區域。同時因為桃園市于 2020 年 8 月已經完成了全市騎樓建築的人工踏勘作業，我們也會將桃園市納入實驗區域，用以進一步驗證本研究的可信度。

### 2.2 試驗設計及執行（2021.04~2021.11）

本階段首先將使用 Google 官方提供之 API 獲取研究區域內的街景圖像數據，之後進行本研究最為核心的街景圖像分析階段。預計將從街景圖像抽取一定比例的圖片，通過算法標記和人工標記相結合的方法構建數據集，用以訓練物件辨識或語意分割模型，比較二者的各項性能參數以選擇合適的技術路線，再反覆檢核模型的各項參數以求達到最優的識別準確率。最後將疊加 OSM、基本地形圖和政府測繪圖資等資訊，通過 GIS 技術於地圖上加以標記，形成台北、桃園二市的騎樓分佈地圖，如果時間和能力條件允許，將在此基礎上進一步探討對騎樓空間使用情況的進行評價的可能性，構建使用狀況地圖。

### 2.3 成果呈現與歸納（2021.12~）

最後一階段試圖與研究意義進行呼應，嘗試討論本研究成果的應用價值，並結合進入 AR/VR 等空間資訊視覺化技術，對成果加以呈現。

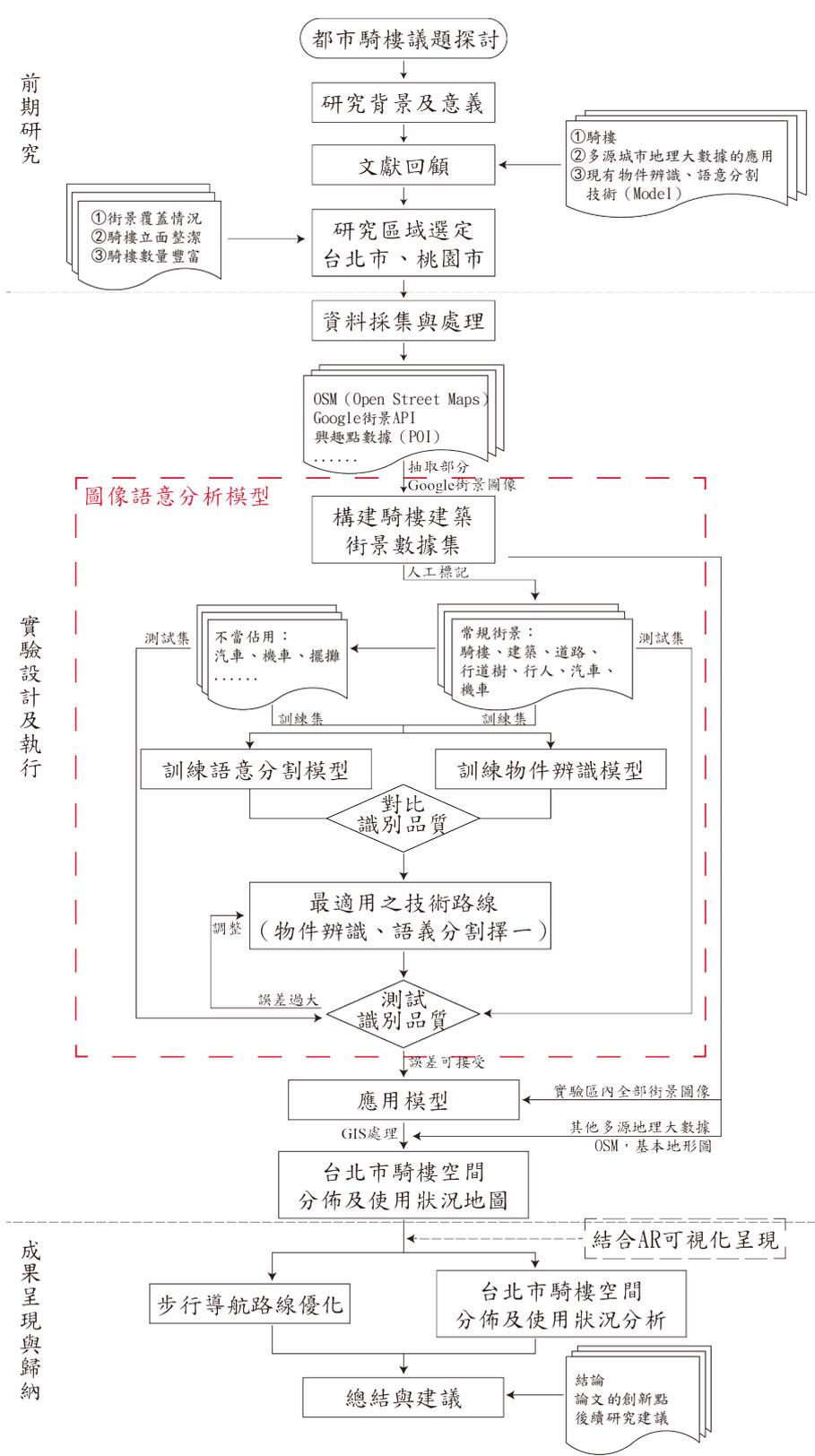


圖 3 研究流程示意圖<sup>1</sup>

<sup>1</sup>筆者自繪

### 3. 擬採用方法

通過前期的文獻梳理和研究，綜合考量算法性能及開源性等因素，我們選取 Yolo V5 (Pytorch 環境)，OCRNet (Pytorch 環境，PaddlePaddle 環境) 二者進行進一步的研究。

#### 3.1 Yolo 模型目標檢測

You Only Look Once(Yolo) 模型目標檢測的基本原理可概括為是將輸入圖像劃分成一定大小的格網，然後檢測格網中是否包含有目標的中心點，據此預測目標邊界框的位置、大小、定位置信度以及分類的概率向量；之後對每個邊界框進行類別預測，計算每個預測邊界模型與其真實值的交並比 (Intersection over union, IoU)，通過非極大值抑制 (Non-Maximum Suppression, NMS)，篩選出置信度最高的候選框，達到輸出目標邊界的目的 (如圖 4) [12]。

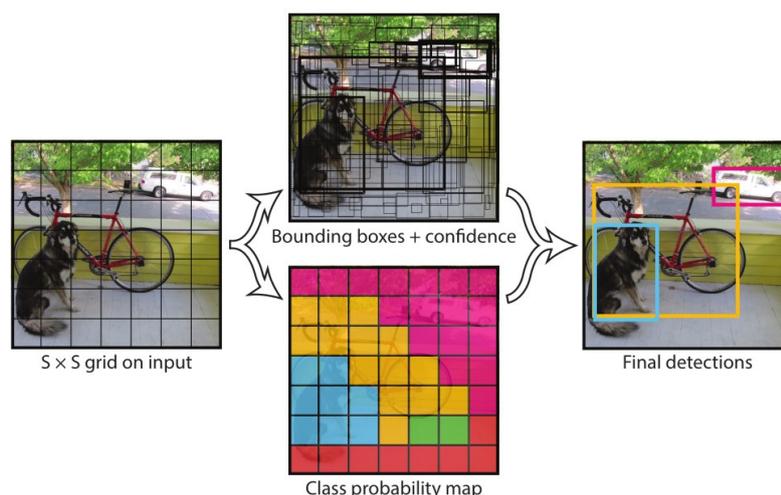


圖 4 Yolo 預測示意圖 [12]

自 2016 年 Yolo 算法首次被提出以來，因為其在運算效能和準確度方面取得了良好平衡，一直備受關注，原作者 Joseph Redmon 先後更新了 V1~V3 三個版本，而 Yolo V5 算法是 Ultralytics 團隊對 Yolo 系列算法改進的產品，於 2020 年 5 月首次開源於 GitHub 平台<sup>1</sup>。根據目前公開的測試數據，Yolo V5 算法在 COCO 數據集上的性能遠優於 Yolo V3 和 EfficientDet (D0~D4)<sup>2</sup>，與 Yolo V4 準確度接近，但速度優於 Yolo V4。同時 Ultralytics 研究團隊依然在對算法進行不斷的更新和完善，目前已更新至 4.0 版本，未來的潛力值得期待。

<sup>1</sup>截止本文撰寫時，yolov5 團隊尚未發表正式論文，只開源了相關算法，URL：<https://github.com/ultralytics/yolov5>

<sup>2</sup>截止本文撰寫時，該算法官方未開源，故不列入考慮範圍

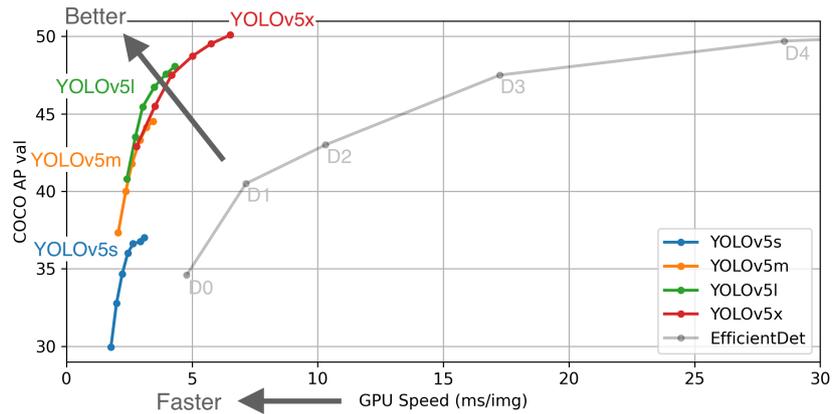


圖 5 Yolo V5 性能測試<sup>1</sup>

Yolo V5 根據網絡深度的不同存在有 s,m,l,x 四個版本，從圖 5 可以看出，YoloV5x 的推理準確度最高，亦最為符合本研究的目標，故我們將選取 YoloV5x 執行後續的研究。

然而更深的網絡結構在提升預測精度的同時，也對計算效能提出了更高的要求，我們使用相同的數據集和參數在 Tesla V100 和 GTX 1080Ti 上分別對 YoloV5x 進行訓練，相比後者，Tesla V100 可以縮短一半以上的訓練時間，同時 Tesla V100 更大的 GPU 記憶體容量也可以調試更為激進的訓練參數提供可能，在最新版本的 Yolo V5 算法也完善了多 GPU 的支援，因此，我們希望能申請高效能計算服務為研究助力。

### 3.2 OCRNet 模型語義分割

語義分割任務即輸入一張圖像，經過模型的運算輸出圖像中每個像素的語義類別。下圖展示了選自 Cityscapes[9], LIP[6], ADE20K[14] 和 COCO-Stuff[3] 的示例，第一行是原始的輸入圖像，第二行是對應於這些圖像的 Ground-Truth 的語義分割結果，其中不同的顏色代表了不同的語義類別。



圖 6 語義分割任務示例<sup>2</sup>

<sup>1</sup>資料來源：<https://github.com/ultralytics/yoloV5>

<sup>2</sup>資料來源：[https://zhuanlan.zhihu.com/p/196685234?utm\\_source=wechat\\_session](https://zhuanlan.zhihu.com/p/196685234?utm_source=wechat_session)

微軟亞洲研究院提出的 OCR 方法分為四個實現階段 [13]：

- (1) 軟物體區域 (Soft Object Regions)：根據網絡中間層的特征表示估測一個粗略的語義分割結果作為 Ocr 方法的輸入；
- (2) 物件區域表示 (Object Region Representations)：根據粗略的語義分割結果和網絡最深層的特征表示計算出 K 組向量，其中每一個向量對應一個語義類別的特征表示；
- (3) 物件上下文特征表示 (Object Contextual Representation, OCR)：計算網絡最深層輸出的像素特征表示 (Pixel Representations)，與計算得到的物件區域特征表示 (Object Region Representation) 之間的關係矩陣，然後根據每個像素和物件區域特征表示在關係矩陣中的數值把物體區域特征加權求和，最後得到 OCR；
- (4) 當把 OCR 與網絡最深層輸入的特征表示拼接之後作為上下文資訊增強的特征表示 (Augmented Representation)，即可基於增強後的特征表示預測每個像素的語義類別。

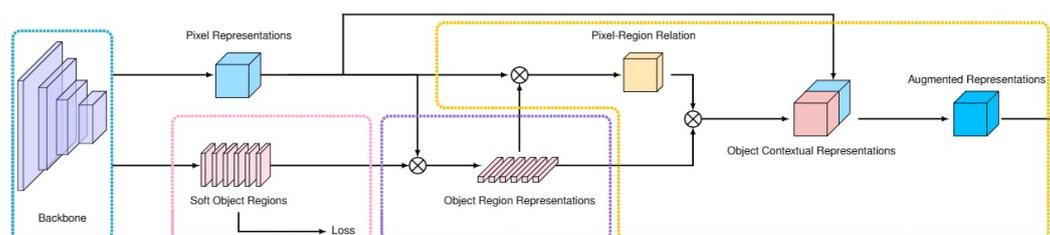


圖 7 OCRNet 框架示意圖 [13]

根據論文中所提出的測試結果，OCRNet 方法在 5 個主流的語義分割數據集上都取得了不錯的效果。截止本文撰寫時，NVIDIA 團隊基於 HRNet+OCR 主幹網絡結構設計的高效多尺度融合方法 [1]，在針對街景的 Cityscapes 數據集中取得的預測精度是目前開源算法中最高 (85.4%)<sup>1</sup>。

## 4. 小結

目前相關領域的研究中，目標檢測和語義分割兩條技術路線都有得到一定程度的應用，也都存在有不同程度的優勢和局限，因此本研究將分別測試、對比兩條技術路線，選擇其中較為合適者進行進一步的發展，亦或是探討將其加以融合的可能性。

眾所周知，深度學習不論是在訓練還是在預測、推理階段均對 GPU 的效能提出了很高的要求，而本研究亦屬大數據性質，僅台北市一市就存在有約 400~600 萬張街景照片需要分析，為求能如期完成研究目標，希望中心能不吝提供高效能運算資源的協助。

<sup>1</sup>資料來源：<https://www.cityscapes-dataset.com/benchmarks/> 檢索日期：2021.02.04

## 参考文献

- [1] Andrew Tao, Karan Sapra, and Bryan Catanzaro, “Hierarchical Multi-Scale Attention for Semantic Segmentation,” *arXiv:2005.10821 [cs]*, May 2020, arXiv: 2005.10821. [Online]. Available: <http://arxiv.org/abs/2005.10821>
- [2] V. Badrinarayanan, A. Handa, and R. Cipolla, “Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling,” *arXiv:1505.07293 [cs]*, 2015.
- [3] H. Caesar, J. Uijlings, and V. Ferrari, “COCO-Stuff: Thing and Stuff Classes in Context,” *arXiv:1612.03716 [cs]*, Mar. 2018, arXiv: 1612.03716. [Online]. Available: <http://arxiv.org/abs/1612.03716>
- [4] L.-C. Chang, F.-J. Chang, S.-N. Yang, F.-H. Tsai, T.-H. Chang, and E. E. Herricks, “Self-organizing maps of typhoon tracks allow for flood forecasts up to two days in advance,” *Nature Communications*, vol. 11, no. 1, p. 1983, 2020.
- [5] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs,” *arXiv:1606.00915 [cs]*, 2017.
- [6] K. Gong, X. Liang, D. Zhang, X. Shen, and L. Lin, “Look into Person: Self-supervised Structure-sensitive Learning and A New Benchmark for Human Parsing,” *arXiv:1703.05446 [cs]*, Jul. 2017, arXiv: 1703.05446. [Online]. Available: <http://arxiv.org/abs/1703.05446>
- [7] A. Kundu, V. Vineet, and V. Koltun, “Feature space optimization for semantic video segmentation,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Conference Proceedings, pp. 3168–3175.
- [8] Y. Li and W. Zhenxin, “Measuring visual enclosure for street walkability: Using machine learning algorithms and google street view imagery,” *Applied Geography*, vol. 76, pp. 147–153, 2016.
- [9] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, “The Cityscapes Dataset for Semantic Urban Scene Understanding,” *arXiv:1604.01685 [cs]*, Apr. 2016, arXiv: 1604.01685. [Online]. Available: <http://arxiv.org/abs/1604.01685>
- [10] G. Mittal, K. B. Yagnik, M. Garg, and N. C. Krishnan, “Spotgarbage: smartphone app to detect garbage using deep learning,” ser. UbiComp ’16. New York, NY, USA: Association for Computing Machinery, Conference Proceedings, p. 940–945.

- [11] A. Myronenko, “3d mri brain tumor segmentation using autoencoder regularization,” *arXiv:1810.11654 [cs, q-bio]*, 2018.
- [12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” *arXiv:1506.02640 [cs]*, May 2016.
- [13] Y. Yuan, X. Chen, and J. Wang, “Object-Contextual Representations for Semantic Segmentation,” *arXiv:1909.11065 [cs]*, Jul. 2020, arXiv: 1909.11065. [Online]. Available: <http://arxiv.org/abs/1909.11065>
- [14] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba, “Scene Parsing through ADE20K Dataset,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 5122–5130, iSSN: 1063-6919.
- [15] 吳欽憲, “騎樓的景觀與社會空間識覺——以臺北市為例,” 碩士學位論文, 2009.
- [16] 張澤軒, 陳虎, 吳志紅, and 王敏, “基於深度學習的道路實景行人車輛檢測,” *現代電腦 (專業版)*, no. 10, pp. 32–36, 2018.
- [17] 張麗英, 裴韜, 陳宜金, 宋辭, and 劉小茜, “基於街景圖像的城市環境評價研究綜述,” *地球資訊科學學報*, vol. 21, no. 01, pp. 46–58, 2019.
- [18] 施宥毓, “騎樓法制之研究,” 碩士學位論文, 2015.
- [19] 王怡青, “從步行者觀點探討使用者對「臺南市騎樓暢通計畫」政策認知及滿意度-以北門路為例,” 碩士學位論文, 2017.
- [20] 蔡宜姘, “騎樓空間使用分析-以臺南市東區騎樓暢通計畫示範區為例,” 碩士學位論文, 2020.
- [21] 蘇瑛敏 and 張惠婷, “亞熱帶騎樓形式對於戶外行人舒適度影響之研究,” *建築學報*, no. 101, pp. 39–58, 2017.
- [22] 賴裕鵬 and 聶志高, “台灣街屋與中國廣東騎樓之比較研究：以建築法規對傳統街屋騎樓影響為例,” *都市與計畫*, vol. 38, no. 1, pp. 73–98, 2011.
- [23] 魏書法 and 程章林, “基於圖像的城市場景垃圾自動檢測,” *集成技術*, vol. 6, no. 01, pp. 39–52, 2017.