

# Automatic Assessment of Public Open Spaces Using Street View Imagery

Shuting Chen<sup>a</sup>, Filip Biljecki<sup>a,b,\*</sup>

<sup>a</sup>*Department of Architecture, National University of Singapore, Singapore*

<sup>b</sup>*Department of Real Estate, National University of Singapore, Singapore*

---

## Abstract

Public Open Space (POS) is essential to urban areas. Assessing them usually requires tedious approaches such as fieldwork and manual processes. Street View Imagery (SVI) and Computer Vision (CV) have been adopted in some urban environment research, bringing fine granularity and human perspective. However, limited aspects have been subject in these studies, and SVI and CV have not been used for holistic POS assessment. This research introduces a novel approach of employing them in conjunction with traditionally used geospatial and remote sensing data for automating POS assessment and doing so extensively. Indicators from both subjective and objective perspectives are developed, and CV algorithms are adopted for retrieving visual features. In a case study spanning 800 POS in Hong Kong and Singapore, a method is designed to predict both subjective and objective scores. The results demonstrate the perceptual models achieved acceptable to high accuracy scores, and suggest that SVI reflects different aspects of POS compared to previous approaches. The paper concludes that SVI can be adopted in POS assessment as a new instrument, extending their research scope to rarely considered off-road areas, and contributing with a new approach for the design and allocation of POS in urban planning.

**Keywords:** Urban perception, Deep learning, Urban Analytics, GeoAI, Google Street View, Participatory planning

---

\*Corresponding author

*Email addresses:* [chenshuting@u.nus.edu](mailto:chenshuting@u.nus.edu) (Shuting Chen), [filip@nus.edu.sg](mailto:filip@nus.edu.sg) (Filip Biljecki)

*Preprint submitted to Cities*

*April 23, 2023*

## 1. Introduction

Public Open Spaces (POS), such as parks, sports fields, playgrounds (Figure 1), are a crucial component of the urban built environment that provides spaces for recreational and social activities (Koohsari et al., 2015; Volenec et al., 2021; Giles-Corti et al., 2005; Lamb et al., 2019; Hoffmann et al., 2018; McCormack et al., 2010; Kaźmierczak, 2013). It is beneficial to human health in both mental and physical aspects (Taylor et al., 2011; Koohsari et al., 2015; Francis et al., 2012; Lamb et al., 2019; McCormack et al., 2010) and provides ecological benefits (Maruani and Amit-Cohen, 2007; Tang, 2017). Evidence suggested that the quality of POS is more crucial than the quantity in providing positive impacts on the human psychology (Francis et al., 2012). Well planned and designed open spaces will potentially contribute to human well-being and provide socio-ecological benefits to the city (Davern et al., 2016; Wang and Foley, 2021; Villanueva et al., 2015; Giles-Corti et al., 2005). Investigating the rationale behind the design, planning and maintenance of POS is of great importance for contemporary urban development (Davern et al., 2016).

Existing tools for auditing POS usually employ onsite fieldwork, census data analysis, survey, focus group and document analysis, and so on, which are time-consuming and labour-intensive (Zhu et al., 2021; Hidayat and Ridwan, 2018; Wang and Foley, 2021; Campos-Sánchez et al., 2019; Edwards et al., 2013).

Increasing the efficiency of the assessment, some studies adopted geospatial tools with 2D digital maps, aerial and remote sensing imagery, automating the process of data collection (Villanueva et al., 2015; Mavoia et al., 2015). Nevertheless, this approach also suffers from some disadvantages. For example, open remote sensing imagery has limited resolution and only provides a bird's eye view, in contrast with the ground-level perspective of POS users.

In recent years, Street View Imagery (SVI) and Computer Vision (CV) have been gaining a foothold in research related to urban built environment (Biljecki and Ito, 2021), potentially providing new opportunities of investigating the POS in fine granularity, and from the perspective of humans.

However, existing urban built environment studies that employed SVI have been almost entirely confined to the streetscape, as widely available imagery from commonly used large-scale services such as Google Street View and Baidu Maps is predominantly collected on driveable roads. Therefore, SVI has not gained popularity in assessing POS, which are often not plied by cars. In general, the use of SVI for applications that focus on areas farther from roads has been sparse. There have been rare and limited instances of POS assessment research involving SVI, such as the tool developed by Edwards et al. (2013). However, as their approach requires manual observations by raters, the auditing process is still yet to be auto-



Figure 1: This illustration indicates different types of POS and highlights the infrastructures that are related closely to human activities and play a major role in the quality of POS. These features may be well captured by street-level imagery, while usually not observable from GIS and remote sensing data. And this study aims to take this advantage to assess the quality of POS.

mated (Taylor et al., 2011; Edwards et al., 2013). That is, no POS assessment has adopted CV techniques, and the usability of using SVI and CV in automating the POS auditing process has yet to be explored.

In addition, SVI provide human eye-level observation and are often used in modelling the subjective perception of urban space. It presents an opportunity for exploring the POS assessment from subjective perspective, bringing in new insight to the POS assessment.

For the first time, this study aims to bridge the research gaps by answering can Street View Imagery (SVI) and Computer Vision (CV) be used in constructing an indicator system and automating the assessment of Public Open Space (POS)? If yes, how is the performance that SVI are used in subjective indicator modelling in POS assessment? And what is the relationship between the scores of objective and subjective perspectives of POS assessment?

This topic is timely, as we have noticed the recent increase of ‘off-road’ imagery in services such as Google Street View (GSV), which are now including

spaces beyond traditional venues, e.g. hiking trails (Middel et al., 2019) and urban waterbodies (Luo et al., 2022), giving way for introducing novel applications of this increasingly relevant urban data source. Further, the topic is spurred by the growth of crowdsourced SVI platforms such as Mapillary and KartaView, which allow anyone to collect photos on any platform, resulting in a diversity of alternative venues such as cycleways, parks, and pedestrian zones (Yap et al., 2022a; Hou and Biljecki, 2022).

This study contributes to the body of knowledge in four main aspects: first, it investigates the use SVI and CV for automating the assessment of POS, as an extension of the traditional time- and resource-intensive approach of manual data collection and visual interpretation. Second, on a broader scope, previous urban built environment research that adopted SVI was mainly focused on roads, and this research extends the scope of SVI applications to ‘off-road’ areas, a rarity in urban studies. Third, past efforts usually adopt single-source data in auditing POS, but multi-sourced data utilised in this research provides multiple perspectives, reducing the potential bias of a single data source. Fourth, this SVI-driven study complements previous POS auditing tools that only assess POS from objective aspects. As this research proposes subjective indicators to obtain human perception in POS, also, a comparison between the objective and subjective indicators is conducted to understand the relationship between both perspectives.

## **2. Background and Related work**

### *2.1. Definition and categories of public open space*

The definition of POS typically describes open spaces that can be accessed by the public (Paul et al., 2020; UN-Habitat, 2020). The specific categories of POS have been classified in various ways in previous literature (Davern et al., 2016). The categories of POS proposed by Campos-Sánchez et al. (2019) mainly divide POS into two categories, one for optional activities (e.g. recreation and sport), such as parks, gardens, plazas, playgrounds and sports fields; and the other refers to streets and pedestrian, for essential functions (e.g. transport and mobility).

Davern et al. (2016) suggested that the POS and green space are two overlapping but different constructs. POS includes vegetated and non-vegetated landscapes; green space consists of public-accessible and private spaces. The categories of vegetated POS are overlapping with public-accessible green space, these usually include parks, gardens, natural reserves, green walls and community gardens, and so on (Davern et al., 2016). Besides, POS also include non-vegetated spaces such as playgrounds, civic squares, plaza (Davern et al., 2016).

This study mainly focuses on the categories of POS in urban areas, and those used for optional activities (i.e. urban streets are not in the scope of this study): parks, gardens, sporting fields, plazas, playgrounds and waterfront promenades.

## *2.2. Studies on auditing POS*

Some auditing tools and indicators have been developed to assess the quality of POS. Building on the experience of 70 cities, UN-Habitat has developed 20 indicators in five dimensions to assess the quality of POS in the Global Public Space Programme (UN-Habitat, 2020)<sup>1</sup>. The five dimensions are accessibility, green environment, comfort and safety, amenities, use and user.

Public Open Space Tool (POST)<sup>2</sup> developed in 1996, has been a keystone for many subsequent studies related to POS assessment. POST focuses on the attributes that may affect human physical activity. This tool evaluates the POS in four components: activity, environmental quality, amenities, and safety (Broomhall et al., 2004). Many studies have adopted POST to audit the quality of POS. Zhu et al. (2021) employed the POST to assess 160 POS and examine its association with residents' mental health. However, POST requires manual work such as fieldwork and survey, which is time- and resource-intensive. To overcome this shortcoming, Taylor et al. (2011) extended the scope of POST by employing data such as Google Maps and GSV to provide the possibility of remote auditing and automating the process of data collection. Based on this work, Edwards et al. (2013) further developed this tool into a desktop application, Public Open Space Desktop Auditing Tool (POSDAT). Mygind et al. (2016) re-evaluated the usefulness of POSDAT in the assessment of 171 POS in Australia, the result achieved high intrarater reliability of 87% agreement. Szczepańska and Pietrzyk (2020) also adopted the remote sensing and street view data from Google Earth to conduct remote POS evaluation. These research demonstrated the opportunity that SVI, GIS and remote sensing data did reflect the quality of POS and can be adopted in POS auditing studies. Nevertheless, despite that such data can be an alternative for fieldwork, the process of data analyses and interpretation is still entirely manual (Taylor et al., 2011; Edwards et al., 2013; Mygind et al., 2016; Szczepańska and Pietrzyk, 2020).

## *2.3. Studies on application of SVI in urban built environment research*

The advent of applications of SVI in conjunction with CV in urban environment research is potential to overcome the shortcoming of traditional POS auditing tools.

---

<sup>1</sup><https://unhabitat.org/programme/global-public-space-programme>

<sup>2</sup><https://www.science.uwa.edu.au/centres/cbeh/projects/post>

SVI has several advantages in terms of data availability and data characteristics. Firstly, compared with on-site visits which require substantial human and financial input, SVI can be obtained remotely and at a low-cost from various platforms (Biljecki and Ito, 2021; Kang et al., 2020; Anguelov et al., 2010). Secondly, street-level imagery reflects human perspective with fine-grained observations in the individual level, which provides details of ground-level information that GIS and remote sensing data lack (Helbich et al., 2019; Ibrahim et al., 2020).

The applications of SVI and CV in urban space studies are ample, and studies demonstrated that numerous visual features of built environments can be extracted from SVI using CV techniques such as semantic segmentation, object detection and scene classification algorithms. Several metrics have been developed to characterise the urban space for a variety of domains. For example, to capture the ground-level sense of greenery, Green View Index (GVI) is used, which is calculated by measuring the proportion of the identified greenery in SVI (Li et al., 2015). Some researchers have developed methods to quantify and classify the morphology of the street canyon (Hu et al., 2020; Gong et al., 2019; Li et al., 2018). A well-used metric is the Sky View Factor (SVF), which describes the openness of street canyons (Li et al., 2018). Such metrics are used widely in a variety of studies from health to transportation (Dong et al., 2018; Lu, 2019; Ki and Lee, 2021; Basu and Sevtsuk, 2022; Wang et al., 2022a; Zhang et al., 2022b).

In addition to such objective metrics extracted from SVI, there are also many studies focused on modelling human perception using diverse information derived from SVI (Cheng et al., 2017; Dubey et al., 2016; Kruse et al., 2021; Moreno-Vera et al., 2021; Salesses et al., 2013; Tang and Long, 2019; Verma et al., 2020; Yao et al., 2021; Zhang et al., 2018; Ito and Biljecki, 2021). For example, previous studies have employed CV algorithms to obtain perceptions of urban environments from a human perspective based on six perceptual indicators, namely, safe, lively, beautiful, wealthy, depressing, and boring (Dubey et al., 2016; Zhang et al., 2018; Wang et al., 2019). Surveys are usually conducted via online platforms to obtain the scoring of human perception for SVI, combined with features extract from SVI, deep learning models are constructed to predict the human perception of the urban built environment.

Various topics related to urban social and economic issues have been investigated using SVI in conjunction with CV and perception methods mentioned above. These include social inequalities (Lin et al., 2021), the relationship between the built environment and crime rates (Zhou et al., 2021; Hipp et al., 2021; He et al., 2017), street vitality (Li et al., 2022b,a; Wang and Vermeulen, 2021; Liu et al., 2021; Jiang et al., 2022), the association between the quality of street greenery and economic status (Li et al., 2015). Among these, urban green space relevant issues are frequently explored using SVI (Li et al., 2015; Helbich et al., 2019; Lu, 2019;

Xia et al., 2021; Dong et al., 2018; Larkin and Hystad, 2019). Growing evidence suggests that visual features from SVI reflect certain relationships between greenery and human psychological and physical health (Helbich et al., 2019; Larkin and Hystad, 2019; Larkin et al., 2021; Wang et al., 2019). Green space may facilitate physical activities (Lu, 2019; Ki and Lee, 2021; Koohsari et al., 2015).

Many studies have developed composite quantitative indicators using SVI for holistic urban built environment auditing (Vanwolleghem et al., 2014; He et al., 2017; Kelly et al., 2013; Griew et al., 2013; Rundle et al., 2011; Ito and Biljecki, 2021; Cleland et al., 2021), covering topics such as liveability (Cleland et al., 2021), bikeability (Ito and Biljecki, 2021; Vanwolleghem et al., 2014), neighbourhood environment (Rundle et al., 2011), built form (Pang and Biljecki, 2022; Yan and Huang, 2022), street environment (Wang et al., 2022b; Griew et al., 2013; He et al., 2017), as well as overall built environment (Kelly et al., 2013). It has been demonstrated that the tools using SVI to audit the urban environment were able to achieve moderate to high inter-raters agreement (Rundle et al., 2011; Kelly et al., 2013; Griew et al., 2013), providing a reliable alternative to on-site audits. A variety of indicators such as sidewalk, bike lane, cleanliness, aesthetics were selected in these studies, to assess the quality of urban environment in diverse aspects, which affirm that SVI is a trustworthy reflection of urban space and that a wide range of indicators can be quantified using it. Although these tools have considerably reduced the time and cost required for fieldwork, in most of these research, raters still need to directly observe and record all the data from SVI manually. Some adopting CV to automate the data processing have demonstrated the usability of SVI in conjunction with CV to conduct comprehensive assessments (Ito and Biljecki, 2021; Wang et al., 2022b). Despite intense research activities, urban studies that adopted SVI to evaluate urban space are largely limited to urban street areas. Even though POS-related features, e.g. sidewalk and greenery, have been widely investigated using SVI and CV, they are confined to roads and their vicinity, and to the best of our knowledge, there has been no comprehensive POS assessment that adopted such method, a gap we seek to bridge with this paper. Therefore, this study aims at expanding the scope of previous research in developing audit tools to assess the quality of POS using SVI in conjunction with CV, automating the comprehensive POS audit process and further enhancing auditing efficiency. On a broader scope, it brings attention of SVI taken in alternative venues, and it contributes to the saturated SVI landscape of research by demonstrating a use case of SVI not taken in the traditional venue of urban streets. This is important for a variety of reasons, e.g. insights obtained only from streets will not give a full picture of a city, especially when analysing other distinct and dissimilar areas such as public open spaces.

### 3. Study area and data description

#### 3.1. Study area

The planning of POS originated in the 18th and 19th centuries in Western countries with the aim of improving people's crowded living environment (Maruani and Amit-Cohen, 2007; Giles-Corti et al., 2005). Evidence suggested that POS plays an important role in perceptions of the quality of high-density urban spaces (Mitrany, 2005). The density of Asian cities is generally higher than that of Western cities (Motomura et al., 2022; Schneider et al., 2015), and as two such instances, both Singapore and Hong Kong are facing the issue of incorporating user-friendly and high-quality POS in high-density urban fabric during urbanisation (Xue et al., 2017; Tang, 2017; Tang and Wong, 2008). As the population grows and urban sprawls, POS is facing the threat of diminution (Tang and Wong, 2008; Hee and Ooi, 2003). Growing demand and limited land supply have led planners to consider the efficiency of land use. This also calls for developing audit tools of POS quality to assist in decision-making.

Despite the similarities that Singapore and Hong Kong are facing, there are significant differences in the planning and challenges of urban POS in these two cities (Xue et al., 2017), ensuring diversity to validate the research. In Hong Kong, with its hilly topography, the urban area takes up less than 30 percent of the land area, thus apart from large areas of country parks, urban POS are small and compactly integrated into the urban texture. Whereas Singapore is known for its *City in a Garden* policy (Ng, 2019), which incorporates a much larger range of green and POS into the urban built environment (Xue et al., 2017).

#### 3.2. Data collection

The locations of POS for both cities have been a key dataset in this study. The POS layer for Singapore is retrieved from the land use map of Master plan 2019<sup>3</sup> published by Urban Redevelopment Authority (URA, Singapore), where the extracted categories include: Open Space, Parks, Sports and Recreation. The layer of Hong Kong was retrieved from the Land Utilisation Map<sup>4</sup> published by the Planning Department of Hong Kong, and the extracted category was Institutional or Open Space. The POS layers are instrumental to download SVI from Google Street View (GSV) — locations are sampled within the spatial extent of each POS, for which GSV was queried. At the locations where SVI are available, SVI is downloaded in four heading directions (0°, 90°, 180°, 270°).

---

<sup>3</sup><https://www.ura.gov.sg/maps/>

<sup>4</sup>[https://www.pland.gov.hk/pland\\_en/info\\_serv/open\\_data/landu/index.html](https://www.pland.gov.hk/pland_en/info_serv/open_data/landu/index.html)

In total, images of 9694 points in Singapore and 5870 points in Hong Kong are obtained. The four direction images of every sampling point are stitched into a panorama. Some noise is removed from the downloaded SVI dataset, for example, images taken indoors, images taken at angles that did not reflect their surroundings, and overexposed or too dark images. Most importantly, to filter images taken from road areas, the road layer in the land use plan is used to extract SVI that fall within the road network polygons. The cleaned SVI dataset includes 7708 panoramas covering 275 POS in Singapore and 3368 panoramas covering 525 POS in Hong Kong.

GIS data for indicator extraction are mainly downloaded from governmental open data websites of Singapore<sup>5</sup> and Hong Kong<sup>6</sup>. Data not available on the government websites are downloaded from OpenStreetMap, which are of high quality in the study areas (Biljecki and Chow, 2022). The remote sensing data used in this study is Landsat 8 imagery downloaded from the USGS EarthExplorer<sup>7</sup>, the spatial resolution is 30 m. The remote sensing data is adopted in computing the green coverage.

#### **4. Methodology**

This study adopts SVI in conjunction with GIS and remote sensing datasets, computes two categories of indicators, namely subjective and objective ones. Perception modelling using SVI is introduced to measure subjective indicators. CV techniques are employed to extract the visual features from SVI, a survey is conducted to obtain the perception scores from SVI, and supervised machine learning (Random Forest) is adopted to train the predictive models for subjective indicators. GIS data (e.g. POI, land use, location of facilities) and remote sensing data (i.e. Landsat satellite imagery for computing vegetation metric) in conjunction with SVI visual features are used to quantify the objective indicators.

Figure 2 illustrates the analytical framework. The research opts for Hong Kong and Singapore as study areas, a selection elaborated in Section 3.

Given the comprehensive nature of this research, some methodological aspects and findings are included in the appendix for further reading.

##### *4.1. Indicator selection*

Traditional POS assessments focused only on objective indicators. In this study, subjective indicators are proposed together with objective indicators to form

---

<sup>5</sup><https://data.gov.sg/>

<sup>6</sup><https://data.gov.hk/en/>

<sup>7</sup><https://earthexplorer.usgs.gov/>

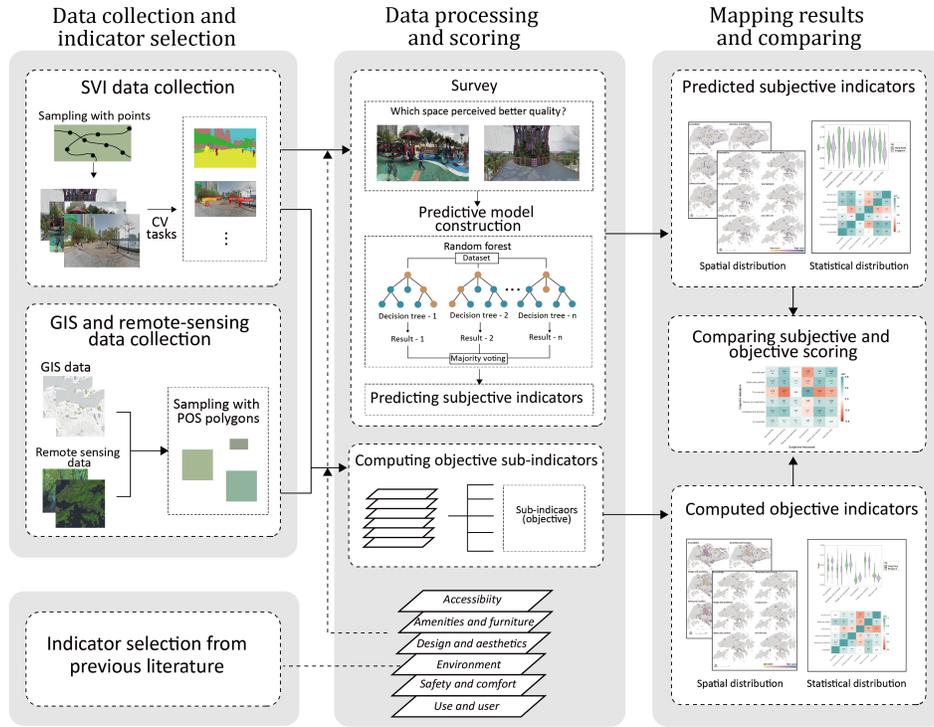


Figure 2: Analytical framework of this study.

a indicator system for POS assessment.

Taking the indicators selected by UN-Habitat (2020) as a reference, combining the auditing tools that have been frequently used in past literature (Wang and Foley, 2021; Broomhall et al., 2004; Mei and Qi, 2020; Edwards et al., 2013), six-dimension indicators are selected in this study, namely *Accessibility*, *Amenities and furniture*, *Design and aesthetics*, *Environment*, *Safety and comfort*, *Usage and user*.

The completed table of indicator filtering is showed in Table A.1 in Appendix A. After screening the indicators, some are excluded by evaluating the data availability from SVI, GIS and remote sensing data sources. For example, water purification and water regulation in the evaluation research conducted by Wang and Foley (2021) are difficult to be assessed using SVI and GIS tools, thus these indicators are not included here. Table 1 lists the indicators selected after screening. In total, 24 indicators are selected and categorised into the six dimensions. It is notable that there are duplicate items in different dimensions, as the score of each dimension will be evaluated separately. It is desirable that each dimension is ex-

Table 1: Composition of indicators in this study.

Dimension	Objective Indicators	Subjective Indicators
<i>Accessibility</i>	Parking area Bike lanes Public transport (bus stops) Sidewalk Fence	Accessibility
<i>Amenities and furniture</i>	Dust bins Signage and emergency items Seating Public access toilets Lighting	Amenities and furniture
<i>Design and aesthetics</i>	Aesthetics features Diversity of landscape elements Variety in colour	Design and aesthetics
<i>Environment</i>	Life and animals Water body Green coverage	Environment
<i>Safety and comfort</i>	Lighting Surrounding building Surrounding road Sidewalk Vandalism	Comfort and safety
<i>Use and user</i>	Number of users Type of activities Surrounding economic activities Land use diversity	Use and user

amined as comprehensively as possible.

To investigate the relationship between subjective and objective indicators, this study proposes to introduce subjective indicators corresponding to the objective in each dimension, allowing *apples to apples* comparison. The scores of subjective indicators are derived from human perception predicted by machine learning models, hence there are no sub-categories (each category is characterised by one subjective indicator).

#### 4.2. Extraction of visual features from SVI

CV techniques are used to extract visual features from SVI. These tasks are adopted to extract high-level features (HLF) and low-level features (LLF). LLF refers to the basic elements in images, e.g. edges, corners, and colour. While HLF is the semantical information in images including scenes, behaviour, etc.

High-level feature extraction tasks employed in this study are Semantic Segmentation (SS), Object Detection (OD) and Scene Classification (SC). Semantic segmentation is used to extract the ratio of features (e.g. sky, grass, tree, ground),

the pre-train model adopted DeepLabv3 (Chen et al., 2017) and ResNet-269 (Zhang et al., 2022a) network based on ADE20K dataset (Zhou et al., 2017b) is selected. The task is implemented by the GluonCV toolkit (Guo et al., 2020). Object detection is conducted to calculate the presence of certain facilities (e.g. benches, dust bins) and the number of users. SparseR-CNN model (Sun et al., 2020) trained on MS COCO dataset (Lin et al., 2014) is adopted. The detection is conducted with the MMDetection toolbox developed by Chen et al. (2019). Classification task is implemented by adopting pre-trained ResNet-50 (He et al., 2016) models on Places365-Standard dataset (Zhou et al., 2017a). Figure 3 shows the example of the high-level feature extraction result.

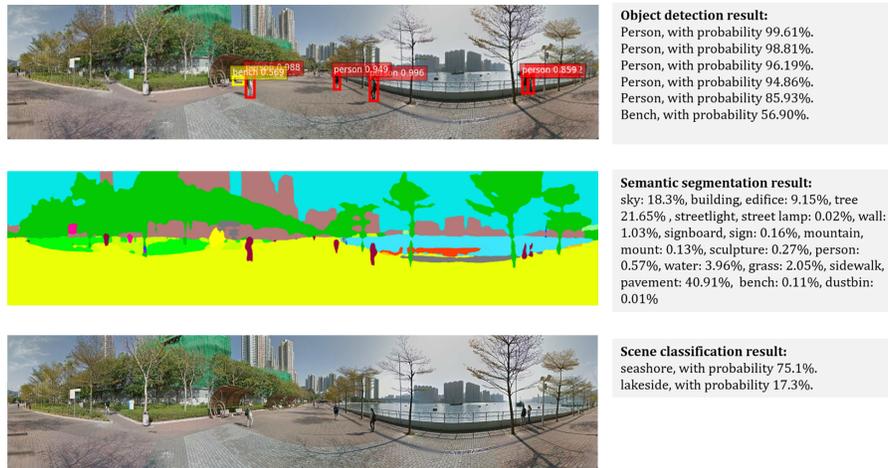


Figure 3: Example of SVI high-level feature extraction. Source: Google Street View.

Low-level feature extraction tasks employed in this study include edge detection, blob detection, HSB (hue, saturation, brightness) statistics, and colourfulness statistics. Evidence from previous research shows that low-level features are correlated to the human perception of place (Rossetti et al., 2019). The low-level feature extraction tasks are implemented by the Python package OpenCV<sup>8</sup>.

In addition to the high-level features and low-level features, there are additional three metrics derived from landscape measures, namely Shannon’s diversity index (SHDI), Shannon’s evenness index (SHEI) (Shannon, 1948), Simpson’s diversity index (SIDI) (Simpson, 1949) and richness (Spellerberg and Fedor, 2003) are calculated using the segmentation patches, the details related to these metrics can be found in Appendix B.

<sup>8</sup><https://opencv.org/>

Table C.1 in Appendix C summarises the visual features extracted from SVI using CV techniques and their definitions.

#### *4.3. Subjective indicators: modelling human perception*

In this sub-section, for the first time, this study proposes a set of analytical workflow to bring the subjective indicators into POS assessment. It quantifies the quality of POS from human perception.

To obtain the perception evaluation scores of POS from SVI, a survey is conducted, combined with SVI features, the machine learning algorithm is employed to construct models to predict the perception scores of spaces. The Institutional Review Board of the National University of Singapore has reviewed and approved the ethical aspects of this research.

##### *4.3.1. Survey on human perception*

A survey investigating human perceptions of the POS environment is designed and distributed on Amazon Mechanical Turk.

There are mainly two types of survey formats were designed in such kind of perception modelling in previous research, one is for participants to rate the space quality of a single street view Ito and Biljecki (2021), and the other is to compare different images and have participants select the one with better space quality (Salesses et al., 2013; Zhang et al., 2018). This study uses the latter, considering the perception of high or low scores may vary greatly from person to person, while the relative perception may be more stable.

The survey presents two random street view panoramas from the image library and asks participants to choose the space that they felt is of better quality across the six dimensions.

There are 400 images in the image library, where 230 images are from POS in Singapore and 170 from Hong Kong. According to previous research, each image requires 22–32 comparison to obtain a robust scoring (Salesses et al., 2013), hence the survey is designed to have every participant rating for 10 pairs of images, totally 600 respondents contribute to 6000 comparisons for each dimension.

To maximise the quality of the survey, a timer provided in the survey platform is used to filter out participants who took too little time to complete the survey. After filtering those, there are 4590 comparisons, 9180 votes for 400 images in total.

##### *4.3.2. Perception score calculation*

The perception scores are calculated following the formula employed in the research conducted by Salesses et al. (2013) as below:

$$W_{i,d} = \frac{w_{i,d}}{w_{i,d} + l_{i,d} + e_{i,d}}, L_{i,d} = \frac{l_{i,d}}{w_{i,d} + l_{i,d} + e_{i,d}} \quad (1)$$

Where

- $W_{i,d}$  is the win rate of image  $i$  in  $d^{th}$  dimension.
- $L_{i,d}$  is the loss rate of image  $i$  in  $d^{th}$  dimension.
- $w_{i,d}$ ,  $l_{i,d}$  and  $e_{i,d}$  represent image  $i$  has been selected as win, loss or equal in  $d^{th}$  dimension respectively.

$$S_{i,d} = \frac{10}{3} \left( W_{i,d} + \frac{1}{n_i^w} \sum_{j_1=1}^{n_i^w} W_{j_1,d} - \frac{1}{n_i^l} \sum_{j_2=1}^{n_i^l} L_{j_2,d} + 1 \right) \quad (2)$$

where

- $j_1$  and  $j_2$  are the images that image  $i$  has won over and lost to in  $d^{th}$  dimension respectively.
- $n_i^w$  and  $n_i^l$  are the total number of images that image  $i$  has won over and lost to in  $d^{th}$  dimension respectively.
- $W_{j_1,d}$  is the win rate of image  $j_1$  in  $d^{th}$  dimension.
- $L_{j_2,d}$  is the lose rate of image  $j_2$  in  $d^{th}$  dimension.

Equation 1 Calculates the win and loss rates for each image in the survey. Equation 2 corrects the win rate of image  $i$  with the average win rate of images that lost the comparison and the average loss rate of images that won the comparison (Salesses et al., 2013).  $\frac{10}{3}$  is the factor that adjusts the range of the score  $S$  to 0 to 10 (Salesses et al., 2013).

#### 4.3.3. Testing robustness of perception scores

The unstable of human perception has been a challenge for data cleaning of perception modelling. However, the test of robustness of collected perceptual data and computed scores has often been overlooked in past studies.

This study attempts to test the robustness of perception scores by measuring the consistency of disjoint subsets of votes. An inter-rater reliability test is conducted by testing the Cohen's kappa coefficient (Cohen, 1960),

Considering the high uncertainty of human perception, the consistency is tested in binary values rather than continues values. Zhang et al. (2018) suggested that

the scores in the middle range are more unstable, this study employed Equation 3 to filter out different range of middle values as noise, then the Cohen's kappa coefficients are computed for different  $\delta$  values to see the best performance.

$$L_{i,d} = \begin{cases} 1, S_{i,d} < \mu_d + \delta\sigma_d \\ -1, S_{i,d} > \mu_d - \delta\sigma_d \end{cases} \quad (3)$$

where

- $\mu$  and  $\sigma_d$  is the mean value and standard deviation of S score in  $d^{th}$  dimension, respectively;
- $\delta$  is the factor that determine the threshold to classify S as positive or negative.

In the robustness test, the SVI that has been compared less than 10 times in either subset are filtered as noise. After filtering, the number of SVI in the tested dataset is 200.

#### 4.3.4. Predictive model construction and subjective scores prediction

Before constructing the model, to better understand the relationship between the visual feature and perception scores and conduct the feature selection process, Backward Elimination method (Derksen and Keselman, 1992) base on regression models is adopted to select a subset of variables for a better performance. Multiple linear regression models are constructed over a range of  $\delta$  value from 0.8 to 1.2 using Equation 3. Here the  $\delta$  donates different thresholds to filter the unstable perception scores in the middle range. For each  $\delta$  value, a subset of variables is selected as predictors.

Then, the Random Forest Classifier algorithm is employed to construct predictive models with visual feature predictors. Based on the robustness test, the input scores are classified as positive and negative values by Equation 3, the Random Forest Classifiers are also constructed over a range of  $\delta$  values from 0.5 to 1.5 to understand the performance of models with different  $\delta$  values. Since human perception is highly uncertain, this step is important for identifying the best parameter for filtering the noise in the dataset, and achieving better performance.

The Random Forest models are constructed using the 5-fold cross-validation method by the Python package scikit-learn (Pedregosa et al., 2011). The number of estimators is set as 1000, and the max depth of trees is set ranging from 6 to 8, the adjustments are subject to each model to achieve better performance.

The results of predictive models simply classify the SVI into positive or negative values, indicating the good or bad quality of space, respectively. Zhang et al. (2018) demonstrated the feasibility of using the probability of being classified as

positive as the score for each image to form a continuous scoring scale rather than simply scoring the image as positive or negative. This study follows this approach

#### *4.4. Objective indicators calculation*

Objective indicators are calculated and evaluated through SVI features in conjunction with GIS and remote sensing data and methods. All sub-indicators under the six dimensions are normalised by the min-max normalisation method and finally synthesised into six scores for each POS with equal weights.

The objective indicators are calculated at the POS level rather than at the sampling points level, which is different from the subjective indicators.

Since the sampling points are at locations where SVI are available, which is biased towards areas with roads and paths, the GIS and remote sensing data are not suitable to sample with these points. Therefore, For GIS and remote sensing data covering the entire POS, all data are sampled directly with POS polygons, and indicators are calculated for each POS as a whole. For SVI data that cannot cover the entire POS, to keep consistent with GIS and remote sensing data, the sampling points should be aggregated into each POS. The basic concept is to take the average value of all sampling points, adjustments are subject to different indicators, the detailed calculation and aggregation methods are presented in this sub-section.

Table 2 summarises the description, data type, data source and scale of indicators that have been investigated in this study.

##### *4.4.1. GIS and remote sensing indicators calculation*

The GIS-based indicators are mainly calculated by computing the presence or number of facilities within or around POS. For the features that are difficult to be counted as numbers (e.g. parking spaces, bike lanes), a dummy value of 1 or 0 is recorded to indicate the presence or absence of the feature in the POS.

Regarding the countable features, the facilities are counted as numbers. In addition, since the scale of the POS may affect the number of features, for example, the larger POS generally has more facilities compared with smaller ones. To eliminate the factor of the scale of POS, the numbers are normalised by the area of POS.

When identifying whether a feature is to be counted into the facility of a certain POS, it is noted that some facilities serve the POS even though outside the boundary of POS. for instance, the car parks are not necessarily inside the POS, and bus stops are surrounding the POS. For these features, a radius of 400 m is selected to aggregate the feature, this radius is determined by considering the general 5-minute walking distance based on previous research (UN-Habitat, 2020). For the road network and bike lanes, the aggregate radius is set as 50 m, simply to identify the linear features surrounding POS. Except for the two cases mentioned above,

other features are only counted if inside the POS, for instance, toilets and activity facilities will only be taken into consideration when they are within the boundaries of POS.

In addition, there are two GIS-based indicators are calculated using specific criteria or formulas. The scores of indicator *Surrounding road* in *Safety and comfort* dimension are measured by identifying the categories of the surrounding road network, the higher the category of the road network, the higher the score, this is to take into consideration that crimes are more likely to happen on minor roads.

The land use diversity index is calculated using the formula derived from SHDI (see Equation B.1 in Appendix B). Previous research have adopted the SHEI for measuring the evenness of land use (Ito and Biljecki, 2021; Frank et al., 2005), and SHDI for measuring the diversity of land use (Hao et al., 2012). In this study, the diversity of land use distribution is considered better aligned with the objective of quantifying the diversity of visitors to POS. In this context, the  $p_i$  represents the proportion of  $i_{th}$  land use category against the total land area, and  $n$  is the number of the land use type. The radius for calculating the land use diversity index is also set as 400 m.

#### 4.4.2. SVI indicators calculation

The SVI indicators are metrics calculated from visual features extracted from SVI using CV. Most of them are calculated from the result of semantic segmentation task. These metrics are calculated based on two main principles. For visual features where the coverage may affect the spatial quality (e.g. green coverage and water body), this study refers to the GVI developed by Li et al. (2015), using the pixel ratio of visual features out of the pixel of the whole image for measurement, and the score on the POS scale is determined by the average pixel ratio of visual features among all SVI captured at this POS.

Another category of visual features is those not suitable to be calculated as pixel ratio (e.g. dust bins, streetlights and seatings). These features are simply recorded as dummy variables, where 1 represents presence and 0 represents absence on each SVI. The score on the POS scale are calculated as the proportion of SVI where the measured visual feature is present in each POS, it can be interpreted as the probability of a visual feature present when randomly selecting an SVI from all SVI captured in each POS.

In addition, the indicator *diversity of landscape elements* and *variety in pattern* under the *design and aesthetics* dimension are simply calculated as the average value of SHDI and colourfulness metric among all the SVI captured in each POS.

Table 2: Summary of objective indicators (SS: Semantic segmentation, OD: Object detection, LLF: Low-level features extraction).

Dimensions	Indicators	Description	Data	Data Source		Scale
				Singapore	Hong Kong	
<i>Accessibility</i>	Parking area	Presence of parking space around POS (400 m buffer)	GIS	Nparks, OSM	OSM	1 if present, 0 if absent
	Bike lane	Presence of bike lane around POS(50 m buffer)	GIS	Nparks, LTA	Data.gov.hk	1 if present, 0 if absent
	Public transport(bus station)	No. of public transport bus station per sqm around POS (400 m buffer)	GIS	LTA	Data.gov.hk	0-1 (min-max normalisation)
	Sidewalk	Average pixel ratio of sidewalk/pavement in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Fence	Average pixel ratio of fence in the POS	SVI (SS)	GSV	GSV	0-1 (reverse scale min-max normalisation)
<i>Amenities and furniture</i>	Dust bins	Proportion of SVI where dustbins was present in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Signage and emergency items	Proportion of SVI where signboard was present in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Seating	Proportion of SVI where seating was present in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Public access toilets	No. of public access toilets per sqm within POS	GIS	OSM, Data.gov.sg	OSM	0-1 (min-max normalisation)
	Lighting	Proportion of SVI where streetlight was present in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
<i>Design and aesthetics</i>	Public space identity	Proportion of SVI where aesthetics features (sculpture, fountain, waterfall, flower pot, vase) was present in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Diversity of landscape elements	Mean value of Simpson diversity index of SVI in POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)

*Continued on next page*

Continued from previous page

Dimensions	Indicators	Description	Data	Data Source		Scale
				Singapore	Hong Kong	
	fence	Average pixel ratio of fence in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Variety in colour (colourfulness)	Mean value of colourfulness value of SVI in POS	SVI (LLF)	GSV	GSV	0-1 (min-max normalisation)
<i>Environment</i>	Life and animals	Proportion of SVI where life/ animal was present in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Water body	Average pixel ratio of water body of all SVI in the POS	SVI (SS)	GSV	GSV	0-2 (min-max normalisation)
	Green coverage	Average pixel ratio of green coverage of all SVI in the POS	SVI (SS)	GSV	GSV	0-1 (mean value of SVI-green coverage and normalised NDVI)
		NDVI (normalised to 0-1)	Remote sensing	Landsat 8	Landsat 8	
<i>Safety and comfort</i>	Lighting	Proportion of SVI where streetlight was present in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Surrounding building	Average pixel ratio of built features (building, fence, wall) in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Surrounding road	Categories of surrounding roads (50 m buffer)	GIS	OSM	OSM	1 if primary road, 0.66 if secondary road, 0.33 if tertiary road
	Sidewalk	Average pixel ratio of sidewalk/pavement in the POS	SVI (SS)	GSV	GSV	0-1 (min-max normalisation)
	Vandalism	Average pixel ratio of dirt track of all SVI in the POS	SVI (SS)	GSV	GSV	0-1 (reverse scale min-max normalisation)
<i>Use and user</i>	Number of users	Average number of person counted in all SVI in the POS	SVI (OD)	GSV	GSV	0-1 (min-max normalisation)
	Type of activities	No. of BBQ facility/ basketball court/ badminton court/ tennis court per sqm within POS	GIS	OSM, Data.gov.sg	OSM, Data.gov.hk	0-1 (min-max normalisation)

Continued from previous page

Dimensions	Indicators	Description	Data	Data Source		Scale
				Singapore	Hong Kong	
	Surrounding economic activities	No. of POI around POS (400 m buffer)	GIS	OSM	OSM	0-1 (min-max normalisation)
	Land use diversity	Land use diversity index	GIS	URA	Planning Department of Hong Kong	0-1 (min-max normalisation)

## 5. Results

### 5.1. Perception modelling

#### 5.1.1. Survey results and robustness of perception scores

Table 3: Cohen kappa coefficient of perception scores calculated from disjoint subsets of raters.

Indicator Categories	Cohen Kappa Coefficient						Mean
	$\delta=0.5$	$\delta=0.6$	$\delta=0.7$	$\delta=0.8$	$\delta=0.9$	$\delta=1$	
<i>Accessibility</i>	0.42	0.47	0.48	0.66	0.67	0.60	0.55
<i>Amenities and furniture</i>	0.17	0.13	0.24	0.52	0.44	0.35	0.31
<i>Design and aesthetics</i>	0.15	0.10	0.16	0.24	0.22	0.47	0.22
<i>Environment</i>	0.25	0.25	0.33	0.24	0.37	0.35	0.30
<i>Safety and comfort</i>	0.20	0.32	0.37	0.46	0.61	0.60	0.43
<i>Use and user</i>	0.29	0.28	0.32	0.33	0.33	0.06	0.27

Figure 4 shows the SVI that obtained high and low perception scores from the survey. The standard deviations are calculated from different subset of the survey results.

The result of the Cohen’s kappa coefficient test in Table 3 shows moderate consistency between scores of disjoint subsets across the six dimensions. Among them, the *Accessibility* shows the highest mean value of the kappa coefficient, and the *Design and aesthetics* obtains the lowest consistency across raters. In addition, a trend that can be observed from the table 3 is that the kappa coefficient increases with the growth of  $\delta$ . Almost all of the six dimensions achieve moderate to high consistency between raters when the  $\delta$  is greater than 0.8, this may indicate that the scores around the median of the scale can be relatively unsound and inaccurate.

It is worth mentioning that this study have also tested the consistency of perception scores with continues values using R-squared of of the Pearson correlation between disjoint subsets follow the study by Salesses et al. (2013). Even though

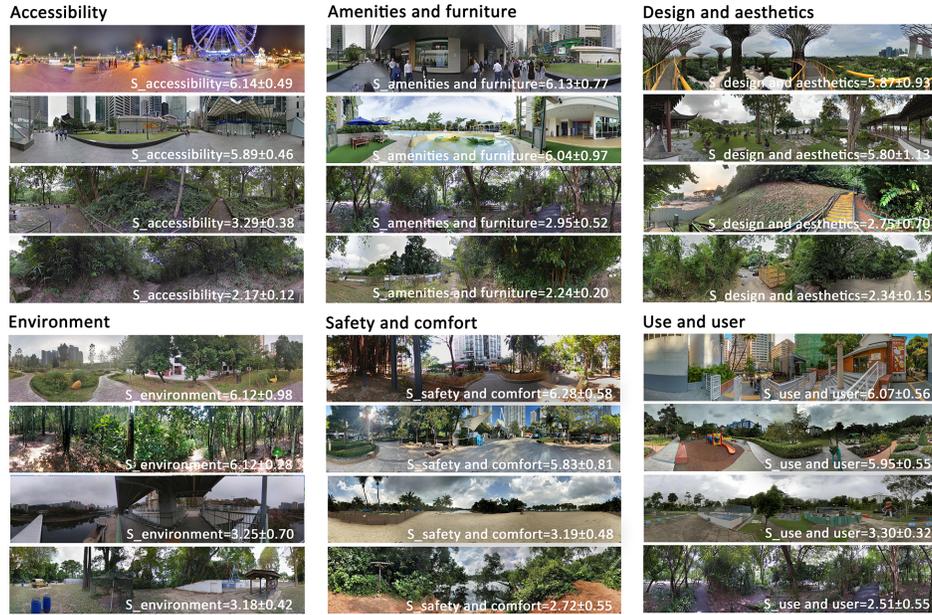


Figure 4: Examples of SVI with perception scores of six dimensions. The examples are at the minimum and maximum boundaries to indicate a wide range situations included in this research. Source: Google Street View.

previous research achieve high R-squared values range from 56.0% to 87.76% (Sallees et al., 2013), this study did not achieve such high consistency. Nevertheless, considering the highly uncertainty of human perception, the moderate to high consistency tested with binary values using Cohen’s kappa coefficient is acceptable. In addition, this result has also revealed that filtering the scores in middle range may effectively increase the quality of perception dataset for predictive model construction, which is a contribution.

### 5.1.2. Identifying input predictors

Table 4 summarises the adjusted R-squared under different values of factor  $\delta$  using Equation 3 to filter the noise. It can be observed from the table that when  $\delta$  is smaller than 1, the R-squared and adjusted R-squared across six dimensions show upward trends with the increase of  $\delta$  value. However, downward trends can be observed when  $\delta$  further increases to 1.2.

Table D.1 shows the result of Back Elimination method base on regression models when  $\delta$  is set as 1.0. The table lists the identified sensitive variables that significantly correlated to the perception scores. These variables are input as independent variables in the subsequent predictive model construction. Overall, the

adjusted R-squared are ranging from 0.187 to 0.508, the visual features explain 50.8% of variance of *Amenities and furniture* perception. Further descriptive analysis and tables are summarised in Appendix D.

In addition, it's notable that the result also highlights that the dimension *Design and Aesthetics* and *Use and user* suffer from the highest probability of unstable scoring in perception data. Also, these two dimensions obtain lowest adjusted R-squared values 0.216 and 0.187, respectively. This indicates that the stability of perception scores is important for the performance of models, which reaffirms the necessity of cleaning the unstable perception score dataset.

Table 4: R-squared of regression models under different  $\delta$  value.

Indicators	$\delta=0$			$\delta=0.8$		
	Observations	R <sup>2</sup>	Adj. R <sup>2</sup>	Observations	R <sup>2</sup>	Adj. R <sup>2</sup>
<i>Accessibility</i>	400	0.177	0.162	166	0.327	0.302
<i>Amenities and furniture</i>	400	0.257	0.242	170	0.408	0.382
<i>Design and aesthetics</i>	400	0.144	0.131	181	0.225	0.198
<i>Environment</i>	400	0.180	0.173	166	0.375	0.356
<i>Safety and comfort</i>	400	0.163	0.148	161	0.226	0.206
<i>Use and user</i>	400	0.149	0.132	165	0.184	0.169

Indicators	$\delta=1.0$			$\delta=1.2$		
	Observations	R <sup>2</sup>	Adj. R <sup>2</sup>	Observations	R <sup>2</sup>	Adj. R <sup>2</sup>
<i>Accessibility</i>	117	0.356	0.321	82	0.351	0.309
<i>Amenities and furniture</i>	120	0.537	0.508	90	0.514	0.479
<i>Design and aesthetics</i>	124	0.254	0.216	87	0.251	0.214
<i>Environment</i>	129	0.408	0.384	94	0.448	0.423
<i>Safety and comfort</i>	125	0.295	0.272	89	0.275	0.250
<i>Use and user</i>	120	0.208	0.187	90	0.203	0.184

### 5.1.3. Constructed predictive model for perception scores

Feature importance scores in Figure 5 shows the selected input predictors and the importance in the models in six dimensions. Interesting patterns can be observed is that the scene classification results occupied the first place of feature importance ranking in all Random Forest models, except for the *Safety and comfort* dimension, and the *Design and aesthetics* was the most closely related to the low-level features among all six dimensions.

Figure 6a demonstrates the average accuracy scores calculated using 5-fold

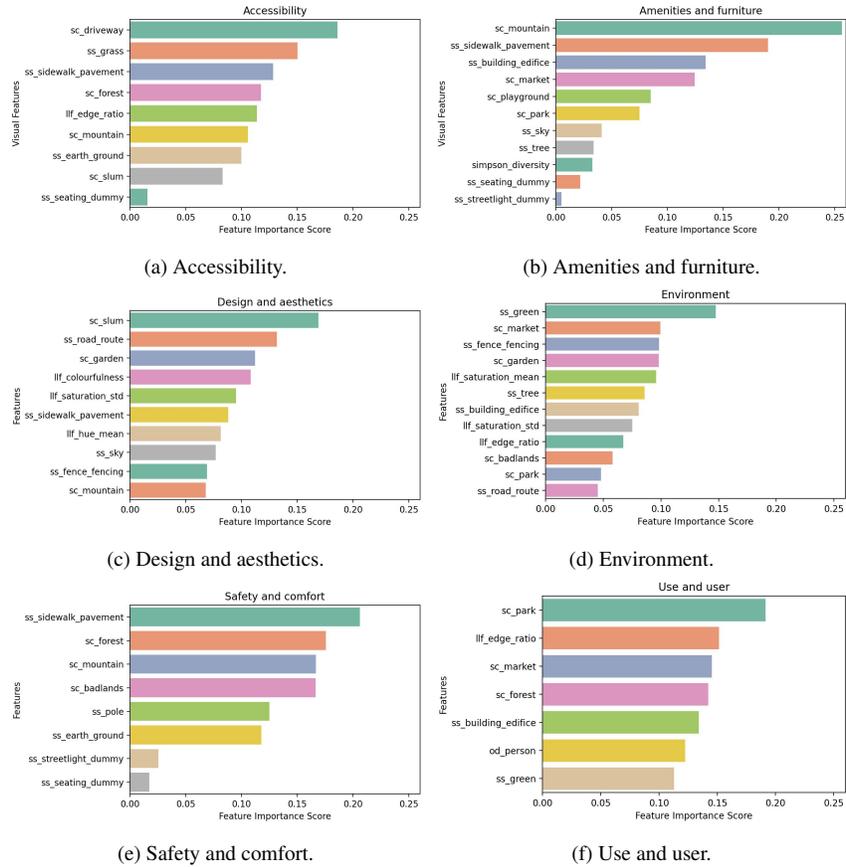


Figure 5: Feature importance score in the Random Forest models across six dimensions.

cross-validation. It is clear that models achieve best performance when the value of  $\delta$  is around 1.0, this is consistent with the kappa test and the regression results in Sections 5.1.1 and Section 5.1.2.

However, previous research noted that the accuracy scores of classifier models keep increasing when the  $\delta$  value grew up to 1.8 (Zhang et al., 2018). The reason for this gap may be due to the limitation of the sample size in this study. Based on the experimental results, the sample size for each dimension is approximately 115-125 images (i.e. more than half of the SVI have been filtered as noise) when the  $\delta$  value is equal to 1.0 in this study. Whereas the previous study has a much larger sample size exceeding 2000 images (Zhang et al., 2018).

Therefore, it can be concluded that the performance of training models starts to degrade when the sample size was less than 115-125 in this study. Hence the  $\delta$

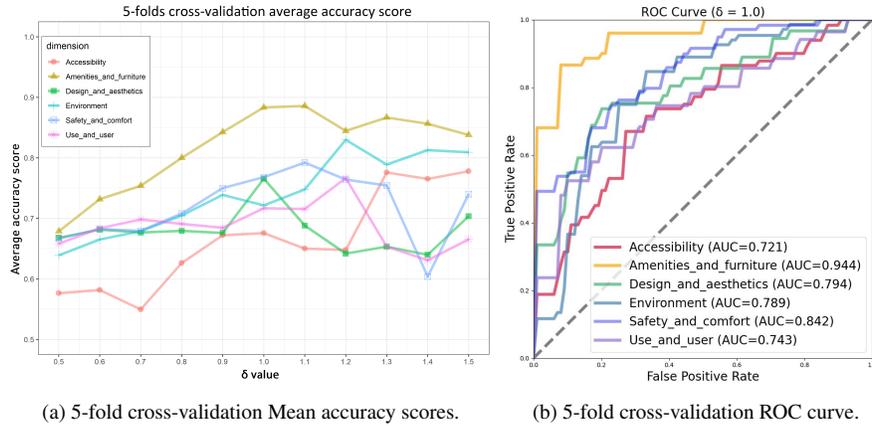


Figure 6: Evaluation of Random Forest models.

value 1.0 was selected to determine the threshold to filter the noise of perception dataset and construct the predictive models.

Figure 6b shows the Receiver Operating Characteristic (ROC) curve of models under the 5-fold cross-validation across six dimensions under the  $\delta$  value 1.0. Of the six indicators, *Amenities and furniture* achieved the highest Area under the Curve (AUC) at 0.944, followed by *Safety and comfort* at 0.842. The AUC for all models are greater than 0.71, indicating models across six dimensions achieve acceptable to outstanding performance. This result demonstrates the usability of the SVI in predicting perceptual scores on POS.

## 5.2. Mapping and comparing the subjective and objective scores

### 5.2.1. subjective scores from perception modelling results

Figure 7 and Figure 8 illustrate the visualisation of perceptual scores across six dimensions using hexagonal grid map. The map provides insights into the spatial distribution of scores in different dimensions. For example, it can be observed that more environmentally attractive locations are usually relatively less accessible, have poorer amenities, are more likely to be unsafe, and attract less diversity of users. This trend is more pronounced in Singapore compared with Hong Kong. This pattern can be further evidenced by Figure 9, where the *Environment* shows negative correlation with *Amenities and furniture*, *Safety and comfort* and *Use and user*.

In addition, specific hot spots and cold spots for the scoring can be identified from maps, for example, the high accessibility scores in the downtown area and east coast, and good environmental quality in the central catchment area in Singapore.

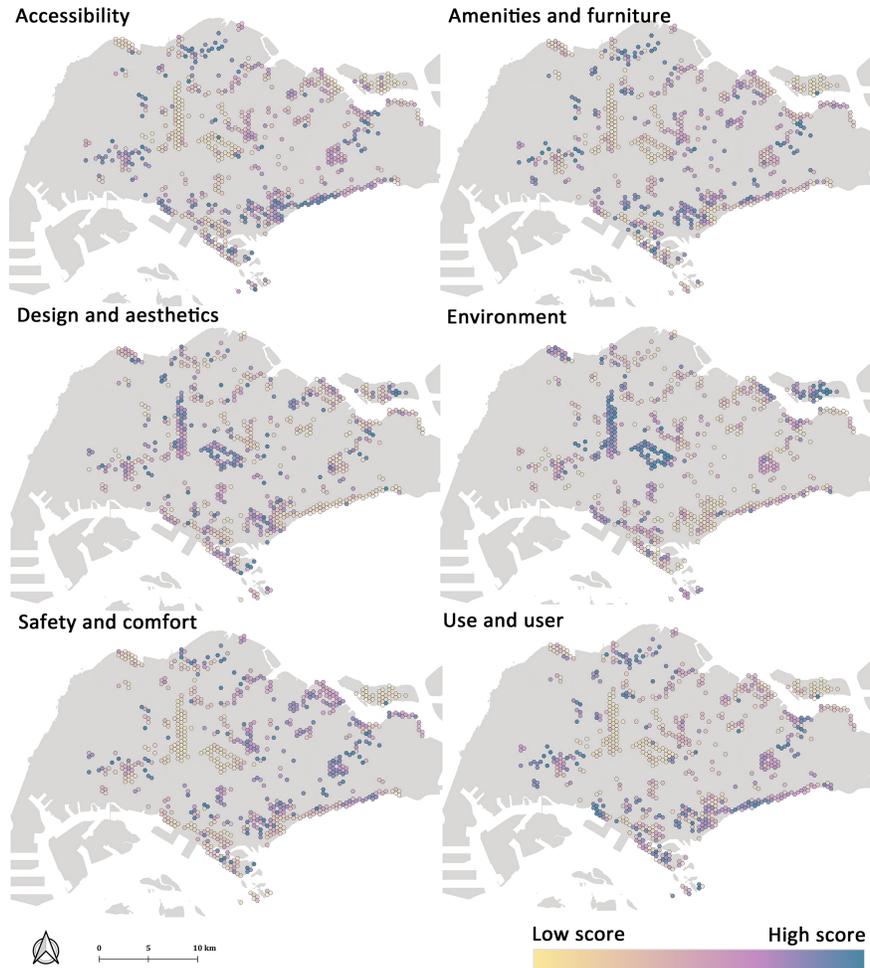


Figure 7: Maps of predicted perceptual scores in public open spaces in Singapore.

Figure 10 shows the statistical distribution of predicted perceptual scores across the six dimensions. We can observe that Singapore has a better environment, which aligns with the *Cities in a Garden* policy. While Hong Kong are evaluated safer, has better facilities and more diversity of users.

### 5.2.2. Objective scores

Figure 11 and Figure 12 presents the spatial distribution of objective indicator scores. Overall, the objective scores see similar trends with subjective indicators, with some specific differences. For example, the environmental quality of POS

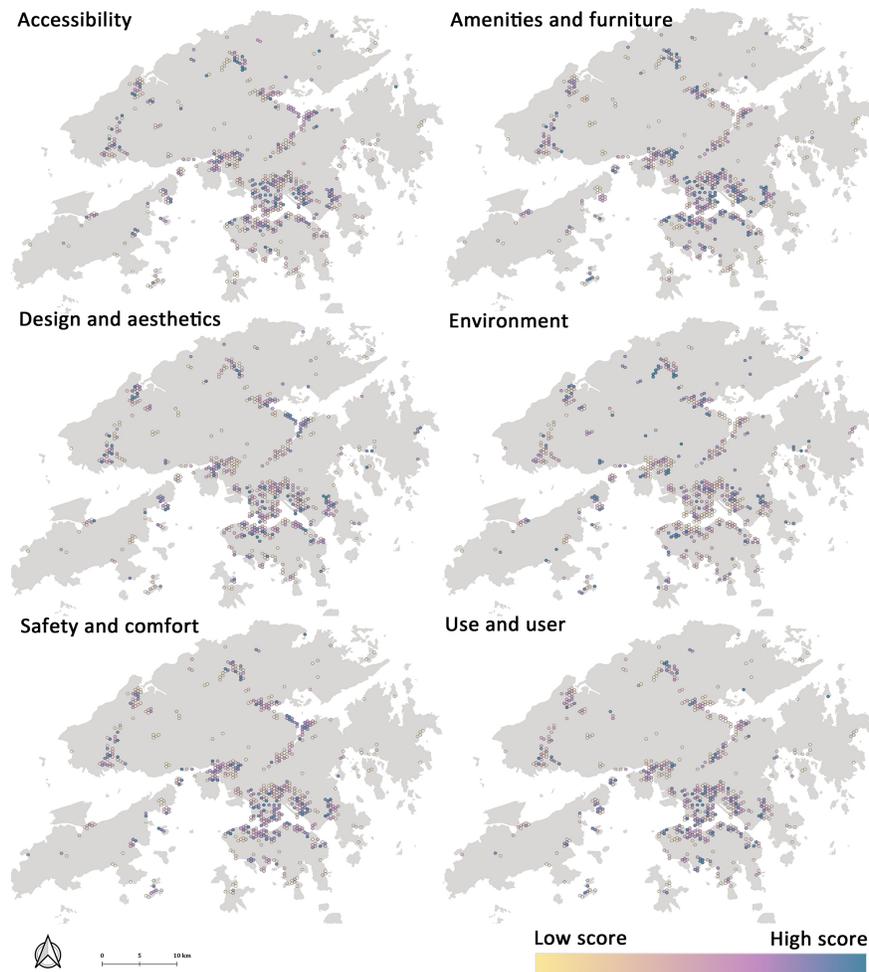


Figure 8: Maps of predicted perceptual scores in public open spaces in Hong Kong.

shows opposite pattern with aesthetic dimension in terms of spatial distribution, which is not observed in perceptual results. And this trend is also further proven in the correlation matrix (Figure 13), where the *Design and aesthetics* shows negative correlation with *Environment*.

For the statistical distribution of objective scores (Figure 14), corresponding to subjective scores in the previous section, Singapore shows a better quality in the environment, while Hong Kong has a higher score in use and user. While in the other four dimensions, the two cities show comparable levels of quality, this is different from subjective scores.

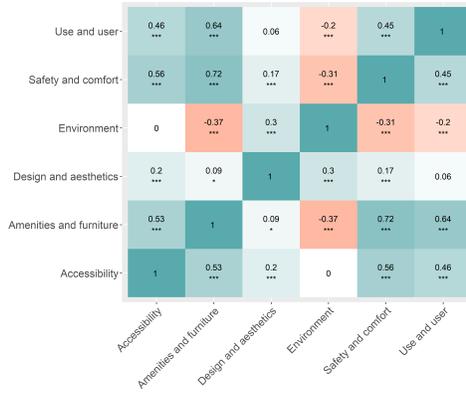


Figure 9: Correlation matrix of subjective indicators (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001).

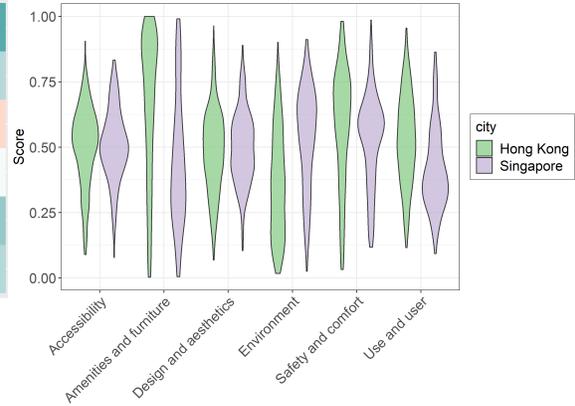


Figure 10: Distributions of subjective indicators.

Under the objective categories, sub-indicators can be calculated to further identify the specific aspects that contribute to the high or low score in each dimension of objective indicator (Appendix E).

### 5.2.3. Overall distribution and comparison of subjective and objective scores

To conclude the results of subjective and objective indicators, the visualisation using hexagonal grid maps provides insight into the quality of POS across six dimensions in spatial scale. From observation, the subjective and objective scores are mostly show similar trends. Generally speaking, less urbanised areas witness higher scores in environmental aspects, and and vice versa for accessibility, amenities, safety and user aspects. Whereas aesthetics shows different spatial pattern in subjective and objective aspects.

For the overall statistical distribution of scores, the scores across the six dimensions in Hong Kong are flatter, and those of Singapore are more peaked, this may indicate that the quality of POS in Hong Kong is more unevenly distributed. In comparison, the quality of POS in Singapore is more clustered in the middle range.

An interesting finding is that based on both the subjective and objective indicators, POS in Singapore showed an overall higher quality than Hong Kong in terms of *Environment*, while Hong Kong scores higher in the *Use and user* aspect. This finding provides further evidence to previous studies on the characteristics of POS in the two cities — as a garden city, Singapore has a more extensive green landscape coverage, while Hong Kong’s POS is more compactly integrated into the urban fabric and therefore gains more user diversity (Xue et al., 2017). The result may also suggest a direction for the future POS planning to improve shortcomings

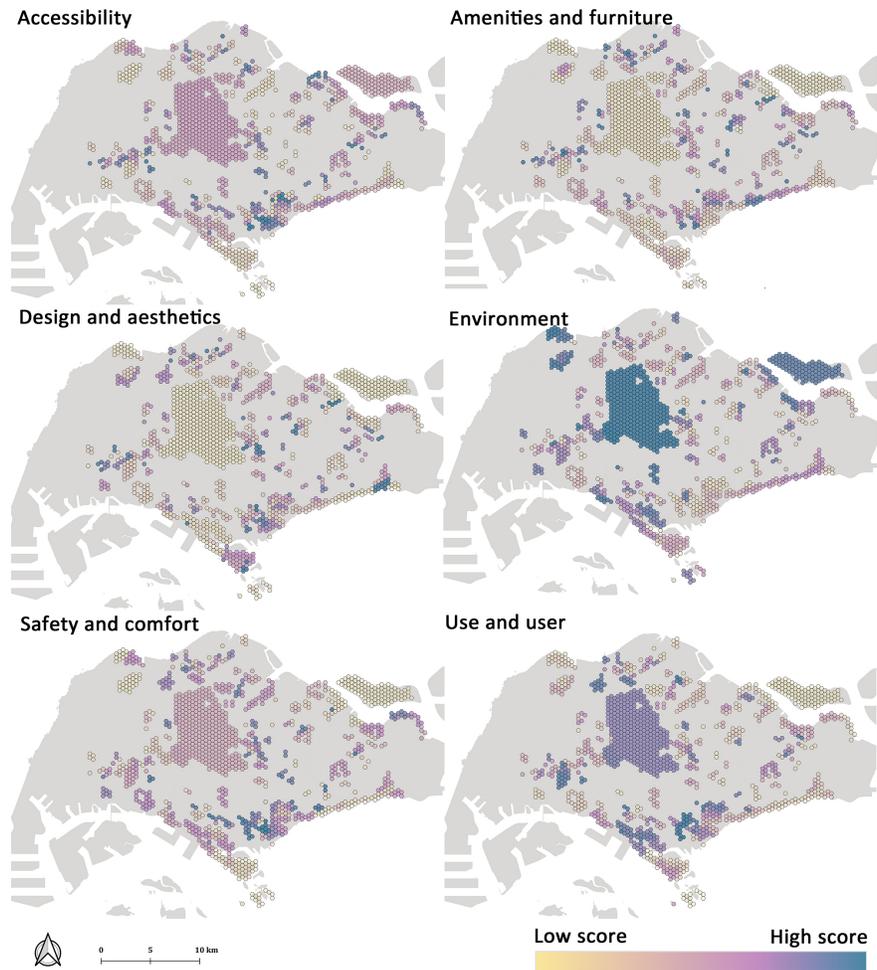


Figure 11: Maps of objective scores in Singapore.

for both cities.

Overall, subjective indicators show positive correlations with objective scores across all six dimensions, all correlation coefficients are significant. The subjective scores of *Environment* shows strongest correlation with objective scores ( $r = 0.68$ ,  $p < 0.001$ ).

The *Design and Aesthetics* shows a weakest but significant correlation ( $r = 0.17$ ,  $p < 0.001$ ). The aesthetic quality of space is intrinsically subjective and is relatively difficult to be described in terms of objective facts (Dhar et al., 2011; Zhang et al., 2018).

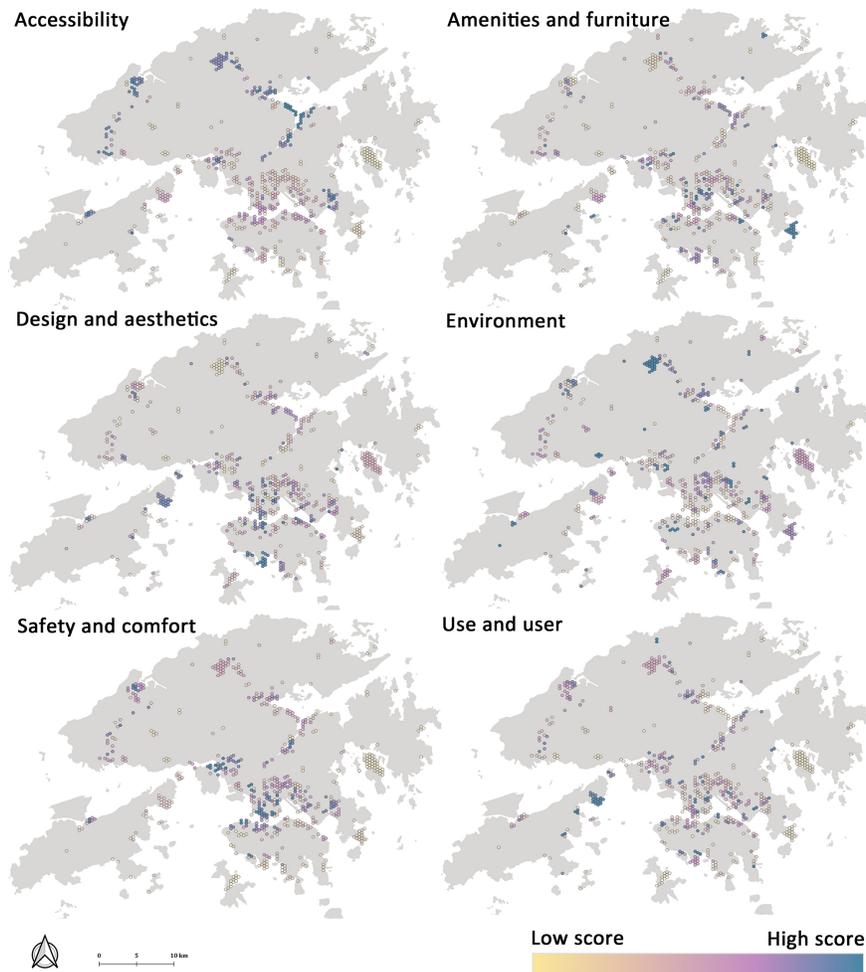


Figure 12: Maps of objective scores in Hong Kong.

The subjective aspect of *Accessibility* also witnesses a weak correlation with the objective indicators ( $r = 0.25$ ,  $p < 0.001$ ). This may be because the objective scores for measuring accessibility are predominately derived from GIS data, which is from the bird's eye perspective and notably different from the human perspective using SVI. From a bird's eye view, GIS tools capture not only the facilities within the POS, but also those surrounding the POS, including car parks, public transport stops, etc. However, SVI captures accessibility in individual points, such as the presence or quality of pavements, bike lanes, etc. That is to say, these two types of metrics emphasise different aspects of accessibility, perceptual metrics are more

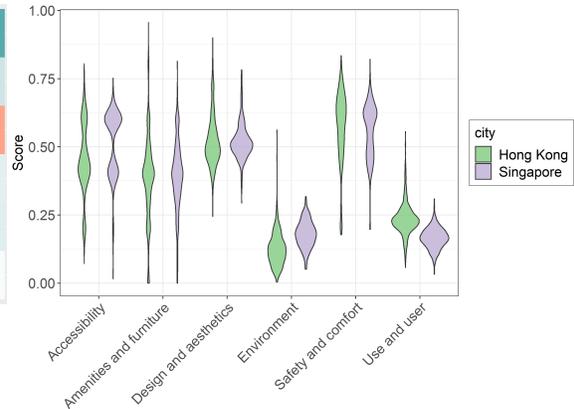
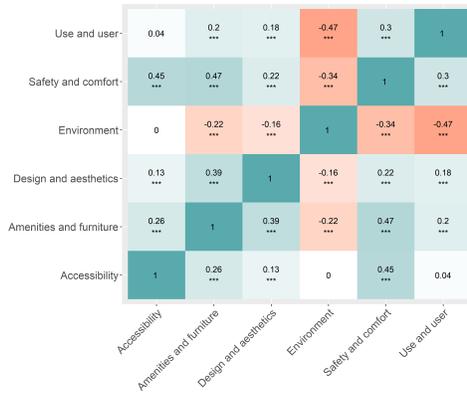


Figure 13: Correlation matrix of objective indicators (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001).

Figure 14: Distributions of objective indicators.

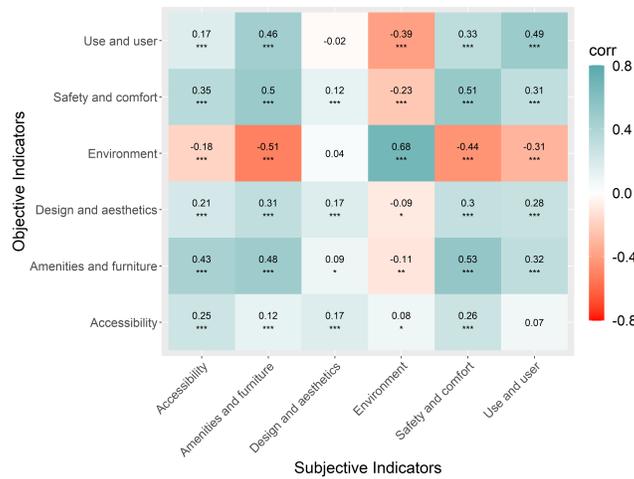


Figure 15: Correlation matrix between subjective and objective indicators (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001)

closely related to human perception, while objective metrics consider across the board.

## 6. Discussion

The experimental results of perception modelling demonstrated the Random Forest Classifier for subjective indicator modelling in POS assessment achieved acceptable to outstanding performance of models, suggesting the feasibility of us-

ing SVI to predict the quality of POS from the human perspective. This finding answers the first sub-question of this study, extending the scope of previous research aimed at modelling the perception of urban built environment using SVI to the off-road area (Zhang et al., 2018; Verma et al., 2020; Yao et al., 2021; Rossetti et al., 2019; Kruse et al., 2021; Salesses et al., 2013; Cheng et al., 2017), bringing in the subjective perspective measurement to the POS assessment, a rarity. The Cohen Kappa coefficient result showed acceptable to moderate consistency between disjoint subsets for perception scores. Even though the test did not achieve high consistency as Salesses et al. (2013) reported, the result revealed that filtering the unstable scores in the middle range can effectively improve the perception data quality and achieve better model performance.

This study also compared the objective and subjective indicators in POS assessment, demonstrating that objective and subjective are correlated with each other but still reflect different aspects. The *Design and Aesthetics* and *Accessibility* showed the weakest correlations between objective and subjective indicators, revealing that the SVI data and GIS reflect the different perspectives of urban space, and the human perception may differ from the objectively assessed scores, especially in the aesthetics aspect which may be highly subjective in nature as previous studies noted (Dhar et al., 2011; Zhang et al., 2018).

In summary, this study demonstrated that Street View Imagery (SVI) and Computer Vision (CV) can be used in constructing an indicator system for POS assessment. In terms of automating the POS assessments, the method proposed by this research adopted SVI, GIS and remote sensing data, the data collection and processing can be automated by software. The method greatly improves the efficiency of traditional fieldwork and manual assessment in previous research (Edwards et al., 2013; Broomhall et al., 2004).

It should be noted that this study still suffered from some limitations and should be address in the future research. Firstly, a great challenge is to secure data quality. Survey data inevitably contain noise, human perception are unstable. Taking the approach of filtering noise in this study as a start, future research should investigate more on improving the quality of survey results and filtering noise. In addition, SVI data can be inconsistent due to the varying time and locations it was taken — for example, some photos were taken during the day and others at night — and the sheer volume makes the issue of cleaning the data a particular challenge. Besides, the quality of data sourced from OpenStreetMap also needs to be further explored as it may be heterogeneous.

Secondly, bias of data is a potential factor leading to inaccurate results. Survey respondents are not live in the cities that have been assessed in this study, which may cause the bias of perception, this bias was also reported by Ito and Biljecki (2021). Furthermore, SVI are biased to street (Middel et al., 2019), the uneven

distribution resulted in the sampling data reflecting only part of the POS, which is a particular issue that should be considered in this study due to the relatively low coverage of the SVI in the POS area. Nevertheless, with the development of SVI platforms, the coverage is improving (Middel et al., 2019). Meanwhile, future research may introduce crowdsourced street view data to overcome this bias. In addition, it is worth mentioning that the objective indicator values calculated by SVI can only represent the characteristics of a specific point rather than the whole space. Errors in the location and angle of observation points, etc., may cause significant fluctuations in SVI feature data values (Wang et al., 2021), and such bias is often overlooked and deserves further exploration in future studies.

Thirdly, current computer vision models for high-level feature extraction are not perfectly adequate for POS auditing, many facilities in POS (e.g. the children's play equipment and aesthetics features) are difficult to be detected by existing models. Therefore, a future direction is to customise state-of-the-art deep learning models for higher accuracy recognition. Given the rapid deployment of deep learning in urban planning (Yap et al., 2022b), we expect an increased performance and usability in the future.

## **7. Conclusion**

This study revisited the topic of POS assessment with a contemporary set of technologies to automate and improve this perennial challenge in urban planning. The findings provided empirical evidence for the feasibility of automating the POS assessment using a novel method we introduced. Street-level imagery, in conjunction with GIS and remote sensing data, were used to establish a more time- and labour-efficient framework that could potentially be scaled up and replicated in other cities.

For the first time, perception modelling was introduced to the POS assessment, to understand the quality of POS closer to the human perspective.

In addition, from a standpoint of data application, using SVI as the primary data, this study explored the urban built environment with SVI of off-road areas, a rarity in urban studies.

The takeaways of this study are:

- SVI used in this research added unparalleled insights to traditionally used GIS or remote sensing data in comprehensive POS assessment, because it reflects more detailed information related to the quality of POS from human scale.
- Visual features derived from SVI are associated with the human perception of the quality of POS. It is possible to model the quality of POS from human

perspective using computer vision techniques and machine learning algorithms.

- The objective indicators are correlated but reflected different aspects from perception indicators. The different perspectives of SVI (human perspective) and GIS or remote data (bird's eye perspective) to some extent has led to a discrepancy between subjective and objective scores.
- The quality of the survey results is an important factor affecting the accuracy of the perception models. In this study, it was evidenced that the calculated scores were more unstable in the middle range. Filtering the data in the middle band appropriately can significantly improve the accuracy of the model training.

Assessing the quality of POS and targeting policy-making with recommendations for improvement is timely and of great importance in urban studies. Quantifying the quality of POS using SVI in conjunction with GIS and remote sensing data may be developed into a universal tool in future exploration, and this study would serve as a first step in exploring such a tool. At the same time, future research should also aim to improve the accuracy and reliability of the assessment using SVI. Possible directions include, firstly, exploring approaches to improve the reliability of survey results. Secondly, adopting more diverse data sources of SVI to mitigate the bias of spatial coverage. Thirdly, experimenting with various machine learning models to improve the accuracy of predictive models as much as possible. Meanwhile, as the accuracy improves, future research should also aim at developing a standalone method using only SVI.

### **Acknowledgements**

We are grateful to the study participants. We appreciate the comments by the editor and reviewers that have helped improve the quality of the manuscript. We thank the members of the NUS Urban Analytics Lab for the discussions, and April Zhu for the design of the illustration. The Institutional Review Board of the National University of Singapore has reviewed and approved the ethical aspects of this research (reference code NUS-IRB-2021-932). This research is part of the project Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start Up Grant R-295-000-171-133.

### **Appendix A. Summary of indicators from previous research**

Table A.1: Summary of indicators from previous research.

Dimension	UN-Habitat (2020)	Broomhall et al. (2004)	Wang and Foley (2021)
Categories of POS	Streets, sidewalks and cycling lanes, squares, waterfront areas, gardens and parks	All	Parks
<i>Accessibility</i>	Parking area. Bike lanes. Pedestrian crossing. Ramps for wheelchairs. Public transports.	Walking paths or cycleways. Placement of paths. Access to public transport.	External accessibility. Internal connections and way-finding. Entrances.
<i>Amenities and furniture</i>	Lighting. Amenities for recreational . Seating. Waste bins. Bike racks. Signage and emergency items. Water and toilets facilities.	Children’s play equipment. Items of play equipment. Playground surface. Picnic tables. Parking facilities. Public access toilets. Kiosk or café. Seating. Club rooms or meeting rooms. Dog litter bags. Taps or other water sources. Drinking fountains.	
<i>Design and aesthetics</i>	Public space identity.	Aesthetic features.	Diversity of landscape elements. Variety in pattern, colour, style and textures. Variety of topography. Coherence and continuity of the built environment.
<i>Environment</i>	Biodiversity (ratio of green coverage). Environmental and community resilience. Energy efficient elements.	Beach / river foreshore. Water features. Trees. Grass. Animals.	Water regulation. Water purification. Climate regulation. Carbon sequestration and storage. Adaptation of extreme events. Fauna diversity. Flora diversity. Habitats diversity. Coverage of permeable surface. Functional connection.

*Continued on next page*

*Continued from previous page*

Dimension	UN-Habitat (2020)	Broomhall et al. (2004)	Wang and Foley (2021)
<i>Safety and comfort</i>	Perception of safety. & level of security. Quality of censorial experience. Overall comfort (waste management system,vandalism).	Shade along paths. Litter. Lighting. Visibility of surrounding roads and buildings. Number of houses overlook the POS. Zebra crossing. Pedestrian crossing. Roads surrounding the POS are minor roads or cul-de-sacs.	Internal inter-visibility. External views and scenery.
<i>Use and user</i>	Number and variety of users. Number and variety of activities. Amount of mixed use in frontage building. Presence of formal and informal economic activities. Restriction rules.	Type of usage / type of activities.	Tourism.

## Appendix B. Description and formulas of SHDI, SHEI and SIDI

SHDI and SIDI describe the diversity of visual features in SVI, taking into account both the homogeneity and richness of features (Nagendra, 2002). SHEI is normalised by the richness from the SHDI, it measures the equality of the distribution of different visual features (Nagendra, 2002).

Equation B.1, B.2, B.3 are the definitions of the indices, and the richness is calculated as the number of visual feature types in an SVI.

$$SHDI = - \sum_i^n P_i * \ln P_i \quad (B.1)$$

$$SHEI = - \frac{\sum_i^n P_i * \ln P_i}{\ln(n)} \quad (B.2)$$

$$SIDI = \sum_i^n P_i * P_i \quad (B.3)$$

where

- n is number of visual feature types computed in semantic segmentation task.
- $P_i$  is the proportion of  $i^{th}$  visual features against the total pixels.

## Appendix C. Visual features extracted from SVI

Table C.1: Visual features extracted from SVI.

Visual features	Descriptions
ss_waterbody	Ratio of pixels classified as water / waterfall, falls / lake / sea / river in the semantic segmentation task.
ss_tree	Ratio of pixels classified as tree in the semantic segmentation task.
ss_streetlight_dummy	Presence of pixels classified as streetlight in the semantic segmentation task.
ss_sky	Ratio of pixels classified as sky in the semantic segmentation task.
ss_signboard_dummy	Presence of pixels classified as signboard in the semantic segmentation task.
ss_sidewalk_pavement	Ratio of pixels classified as sidewalk in the semantic segmentation task.
ss_seating_dummy	Ratio of pixels classified as bench/seat in the semantic segmentation task.
ss_sculpture	Ratio of pixels classified as sculpture in the semantic segmentation task.
ss_road_route	Ratio of pixels classified as road/route in the semantic segmentation task.
ss_pole	Ratio of pixels classified as pole in the semantic segmentation task.
ss_green	Ratio of pixels classified as tree/grass/plant, flora, plant life in the semantic segmentation task.
ss_grass	Ratio of pixels classified as grass in the semantic segmentation task.
ss_fence_fencing	Ratio of pixels classified as fence/fencing in the semantic segmentation task.
ss_earth_ground	Ratio of pixels classified as earth/ground in the semantic segmentation task.
ss_dustbin_dummy	Presence of pixels classified as dustbin in the semantic segmentation task.
ss_building_edifice	Ratio of pixels classified as building/edifice in the semantic segmentation task.
simpson_diversity	The Simpson's diversity index calculated from the semantic segmentation patches.
shannon_diversity	The Shannon's diversity index (Shannon, 1948) calculated from the semantic segmentation patches.
shannon_evenness	The Shannon's evenness index (Shannon, 1948) calculated from the semantic segmentation patches.
richness	The richness index (Spellerberg and Fedor, 2003) calculated from semantic segmentation patches.
sc_slum	Probability of an SVI being classified as slum in the scene classification task.
sc_playground	Probability of an SVI being classified as playground in the scene classification task.
sc_park	Probability of an SVI being classified as park/amusement park in the scene classification task.
sc_mountain	Probability of an SVI being classified as mountain / mountain / river / valley / mountain snowy / creek / field, wild / marsh / lake / natural / canyon path in the scene classification task.

*Continued on next page*

*Continued from previous page*

Visual features	Descriptions
sc_market	Probability of an SVI being classified as drugstore / department store / shopping mall, indoor / supermarket / market outdoor / pet shop in the scene classification task.
sc_garden	Probability of an SVI being classified as roof garden/topiary garden/botanical garden in the scene classification task.
sc_forest	Probability of an SVI being classified as rainforest/forest, broadleaf in the scene classification task.
sc_driveway	Probability of an SVI being classified as driveway, broadleaf in the scene classification task.
sc_badlands	Probability of an SVI being classified as badlands, broadleaf in the scene classification task.
od_truck	Number of truck being detected in the objective detection task.
od_person	Number of person being detected in the objective detection task.
od_car	Number of car being detected in the objective detection task.
l1f_saturation_std	Standard deviation of saturation values for all pixels.
l1f_saturation_mean	Mean of saturation values for all pixels.
l1f_hue_std	Standard deviation of hue values for all pixels.
l1f_hue_mean	Mean of hue values for all pixels.
l1f_edge_ratio	Number of pixels classified as edge.
l1f_edge_no_pixel	Ratio of pixels classified as edge.
l1f_colourfulness	Colourfulness metric (Hasler and Suesstrunk, 2003) calculated from SVI.
l1f_brightness_std	Standard deviation of hue values for all pixels.
l1f_brightness_mean	Mean of hue values for all pixels.
l1f_blob_no	Number of blobs detected in SVI.

#### **Appendix D. Relationship between visual features and human perceptual scores**

Overall, visual feature *ss\_sidewalk\_pavement* is significant and positively correlated with all indicator dimensions except *Environment*, indicating that the pedestrian facilities contribute to the quality of POS except for the environmental aspect. In comparison, visual variable *sc\_mountain* shows a significant and negative correlation with all indicators except *Environment*, indicating that POS classified with higher probability as mountain are generally of worse quality except *Environment* dimension. The visual feature *ss\_streetlight\_dummy*, *ss\_seating\_dummy* and *ss\_building\_edifice* share the same trend, i.e. a significant positive correlation with the four dimensions, namely *Accessibility*, *Amenities and furniture*, *Safety and comfort*, *Use and users*, this is reasonable because these features are all relevant to the physical facilities of POS.

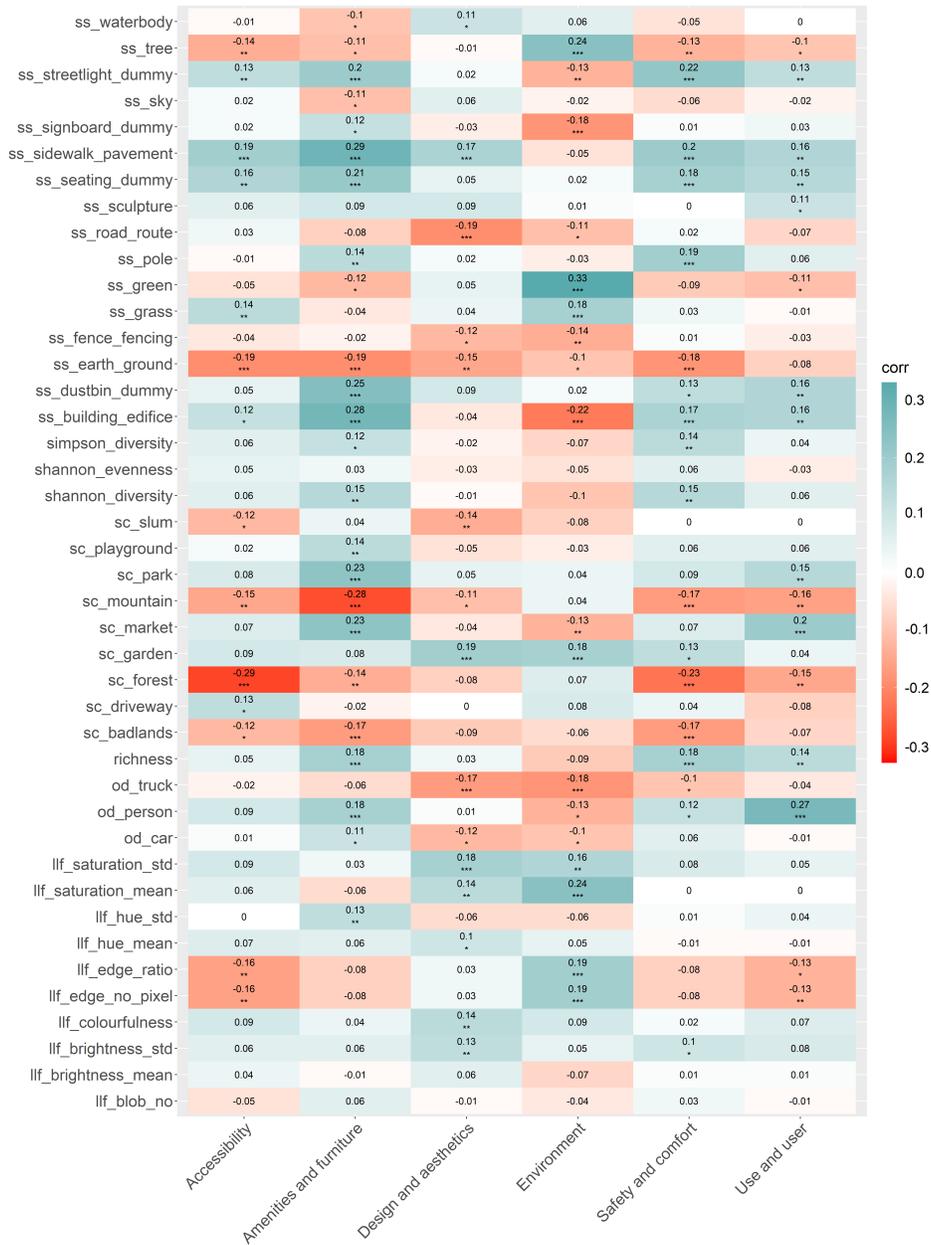


Figure D.1: Pearson correlation coefficients matrix between the survey scores and visual feature statistics (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001).

Table D.1: Regression models under the  $\delta$  value of 1.0.

	Accessibility	Amenities	Aesthetics	Environment	Safety/ Comfort	Use/user
Intercept	4.593*** (0.178)	4.351*** (0.104)	2.791*** (0.483)	3.982*** (0.114)	4.327*** (0.079)	4.308*** (0.084)
llf_edge_ratio	-2.289* (1.050)					
llf_hue_mean			0.012* (0.005)			
llf_saturation_std			0.015** (0.006)			
od_person						0.035*** (0.010)
od_truck			-0.220* (0.098)	-0.207* (0.082)		
sc_driveway	3.154* (1.255)					
sc_forest	-5.420** (1.771)				-3.903** (1.481)	
sc_garden			1.120* (0.483)	1.843*** (0.410)		
sc_mountain	-1.095* (0.539)	-1.461*** (0.411)			-1.515* (0.635)	-1.050* (0.482)
sc_park		1.093* (0.467)				
ss_building_edifice		1.454** (0.447)				
ss_dustbin_dummy						0.301* (0.135)
ss_earth_ground		-1.276** (0.479)				
ss_fence_fencing				-3.455* (1.704)		
ss_grass	1.596** (0.483)			1.279** (0.387)		
ss_pole					74.365*** (21.167)	
ss_seating_dummy		0.258* (0.123)				
ss_sidewalk_pavement	1.478** (0.504)	1.471** (0.545)	0.992* (0.480)		1.518** (0.465)	
ss_tree				1.797*** (0.348)		
ss_waterbody		-2.114* (0.935)	2.518* (1.137)			
Observations	117	120	124	129	125	120
$R^2$	0.356	0.537	0.254	0.408	0.295	0.208
Adjusted $R^2$	0.321	0.508	0.216	0.384	0.272	0.187
Residual Std. Error	0.626(df = 110)	0.558(df = 112)	0.656(df = 117)	0.567(df = 123)	0.590(df = 120)	0.635(df = 116)
F Statistic	10.140*** (df = 6.0; 110.0)	18.521*** (df = 7.0; 112.0)	6.637*** (df = 6.0; 117.0)	16.952*** (df = 5.0; 123.0)	12.559*** (df = 4.0; 120.0)	10.125*** (df = 3.0; 116.0)

Note:

\*p<0.05; \*\*p<0.01; \*\*\*p<0.001  
(The values in brackets show the variance of variables)

Some visual features contribute in particular to a specific aspect of POS quality. For example, there is a moderate and positive correlation between visual feature *ss\_green* and indicator *Environment* ( $r = 0.33$ ,  $p < 0.001$ ), which is in line with intuition. *sc\_forest* is negatively correlated to *Accessibility* ( $r = 0.29$ ,  $p < 0.0001$ ), indicating those with higher probability as forest are perceived have bad accessibility. The *od\_person* shows positive correlation with dimension *Use and user* ( $r = 0.27$ ,  $p < 0.001$ ).

Of the six indicator dimensions, *Design and aesthetics* is less correlated with visual characteristics, however, it is also the indicator that is associated closest with lower-level characteristics. It was significantly and positively correlated with the mean and standard deviation of saturation, the mean of hue, the colourfulness and the standard deviation of brightness.

#### **Appendix E. Analysis of sub-categories of objective indicators**

Table E.1 summarises the mean value and standard deviation of the sub-indicators. From the table, we can identify the specific aspect that contribute to the high or low score in each dimension of objective indicator. For example, it can be observed that Singapore has more green coverage and higher NDVI value, contributing to a higher average score of *Environment* dimension shown in Figure 14. For the sub-indicators under *Use and user* dimension, Hong Kong got a higher mean value in all categories, hence an overall higher score of the *Use and user* dimension can be observed from Figure 14 compared with Singapore.

Table E.1: Mean value and standard deviation of sub-indicators of objective indicators.

Sub-indicators	Hong Kong		Singapore	
	Mean	Std. Dev.	Mean	Std. Dev.
Parking area	0.89	0.313	0.936	0.246
Bike lanes	0.262	0.44	0.62	0.486
Public transport (bus stops)	0.055	0.094	0.012	0.012
Sidewalk	0.179	0.185	0.076	0.088
Fence	0.886	0.128	0.937	0.089
Dust bins	0.224	0.302	0.283	0.295
Signage and emergency items	0.674	0.351	0.679	0.29
Seating	0.267	0.325	0.196	0.238
Public access toilet	0.03	0.078	0.004	0.015
Lighting	0.902	0.297	0.976	0.155
Diversity of landscape elements	0.763	0.114	0.763	0.113
Aesthetic features	0.203	0.3	0.131	0.207
Variety in colour (colourfulness)	0.233	0.121	0.256	0.084
Green coverage (SVI)	0.317	0.177	0.495	0.171
Green coverage (NDVI)	0.367	0.204	0.548	0.165
Life and animals	0.002	0.037	0.002	0.017
Water body	0.03	0.079	0.018	0.055
Surrounding building	0.295	0.176	0.176	0.129
Surrounding road_cate	0.502	0.382	0.574	0.437
Vandalism	0.999	0.023	0.994	0.057
Number of users	0.122	0.154	0.022	0.05
Type of activities	0.031	0.091	0.007	0.027
Surrounding economic activities	0.023	0.064	0.004	0.01
Land use diversity	0.766	0.126	0.627	0.142

## References

- Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., Weaver, J., 2010. Google Street View: Capturing the World at Street Level. *Computer* 43, 32 – 38. doi:10.1109/mc.2010.170.
- Basu, R., Sevtsuk, A., 2022. How do street attributes affect willingness-to-walk? City-wide pedestrian route choice analysis using big data from Boston and San

Francisco. *Transportation Research Part A: Policy and Practice* 163, 1–19. doi:10.1016/j.tra.2022.06.007.

Biljecki, F., Chow, Y.S., 2022. Global Building Morphology Indicators. *Computers, Environment and Urban Systems* 95, 101809.

Biljecki, F., Ito, K., 2021. Street view imagery in urban analytics and gis: A review. *Landscape and Urban Planning* 215, 104217. doi:10.1016/j.landurbplan.2021.104217.

Broomhall, M., Giles-Corti, B., Lange, A., 2004. Quality of public open space tool (post). Perth, Western Australia: School of Population Health, The University of Western Australia .

Campos-Sánchez, F.S., Abarca-Álvarez, F.J., Reinoso-Bellido, R., 2019. Assessment of open spaces in inland medium-sized cities of eastern andalusia (spain) through complementary approaches: Spatial-configurational analysis and decision support. *European Planning Studies* 27, 1270–1290. doi:10.1080/09654313.2019.1579302.

Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., Zhang, Z., Cheng, D., Zhu, C., Cheng, T., Zhao, Q., Li, B., Lu, X., Zhu, R., Wu, Y., Dai, J., Wang, J., Shi, J., Ouyang, W., Loy, C.C., Lin, D., 2019. MMDetection: Open mmlab detection toolbox and benchmark. *arXiv preprint arXiv:1906.07155* .

Chen, L.C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* .

Cheng, L., Chu, S., Zong, W., Li, S., Wu, J., Li, M., 2017. Use of tencent street view imagery for visual perception of streets. *ISPRS International Journal of Geo-Information* 6, 265. doi:10.3390/ijgi6090265.

Cleland, C.L., Ferguson, S., Kee, F., Kelly, P., Williams, A.J., Nightingale, G., Cope, A., Foster, C., Milton, K., Kelly, M.P., et al., 2021. Adaptation and testing of a microscale audit tool to assess liveability using google street view: Maps-liveability. *Journal of Transport & Health* 22, 101226.

Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20, 37–46.

Davern, M., Farrar, A., Kendal, D., Giles-Corti, B., 2016. Quality green space supporting health, wellbeing and biodiversity: A literature review. Report prepared for the National Heart Foundation. University of Melbourne: Victoria .

- Derksen, S., Keselman, H.J., 1992. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology* 45, 265–282.
- Dhar, S., Ordonez, V., Berg, T.L., 2011. High level describable attributes for predicting aesthetics and interestingness, in: *CVPR 2011, IEEE*, pp. 1657–1664.
- Dong, R., Zhang, Y., Zhao, J., 2018. How green are the streets within the sixth ring road of beijing? an analysis based on tencent street view pictures and the green view index. *International Journal of Environmental Research and Public Health* 15, 1367. doi:10.3390/ijerph15071367.
- Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A., 2016. Deep learning the city: Quantifying urban perception at a global scale, in: Leibe, B., Matas, J., Sebe, N., Welling, M. (Eds.), *Computer Vision – ECCV 2016*. Springer International Publishing, Cham. volume 9905, pp. 196–212. doi:10.1007/978-3-319-46448-0\_12.
- Edwards, N., Hooper, P., Trapp, G.S.A., Bull, F., Boruff, B., Giles-Corti, B., 2013. Development of a public open space desktop auditing tool (posdat): A remote sensing approach. *Applied Geography* 38, 22–30. doi:10.1016/j.apgeog.2012.11.010.
- Francis, J., Wood, L.J., Knuiman, M., Giles-Corti, B., 2012. Quality or quantity? exploring the relationship between public open space attributes and mental health in perth, western australia. *Social science & medicine* 74, 1570–1577.
- Frank, L.D., Schmid, T.L., Sallis, J.F., Chapman, J., Saelens, B.E., 2005. Linking objectively measured physical activity with objectively measured urban form: findings from smartraq. *American journal of preventive medicine* 28, 117–125.
- Giles-Corti, B., Broomhall, M.H., Knuiman, M., Collins, C., Douglas, K., Ng, K., Lange, A., Donovan, R.J., 2005. Increasing walking: how important is distance to, attractiveness, and size of public open space? *American journal of preventive medicine* 28, 169–176.
- Gong, F.Y., Zeng, Z.C., Ng, E., Norford, L.K., 2019. Spatiotemporal patterns of street-level solar radiation estimated using google street view in a high-density urban environment. *Building and Environment* 148, 547–566.
- Griew, P., Hillsdon, M., Foster, C., Coombes, E., Jones, A., Wilkinson, P., 2013. Developing and testing a street audit tool using google street view to measure environmental supportiveness for physical activity. *International Journal of Behavioral Nutrition and Physical Activity* 10, 1–7.

- Guo, J., He, H., He, T., Lausen, L., Li, M., Lin, H., Shi, X., Wang, C., Xie, J., Zha, S., Zhang, A., Zhang, H., Zhang, Z., Zhang, Z., Zheng, S., Zhu, Y., 2020. Gluoncv and gluonnlp: Deep learning in computer vision and natural language processing. *Journal of Machine Learning Research* 21, 1–7. URL: <http://jmlr.org/papers/v21/19-429.html>.
- Hao, P., Geertman, S., Hooimeijer, P., Sliuzas, R., 2012. The land-use diversity in urban villages in shenzhen. *Environment and Planning A: Economy and Space* 44, 2742–2764. URL: <https://doi.org/10.1068/a44696>, doi:10.1068/a44696. eprint: <https://doi.org/10.1068/a44696>.
- Hasler, D., Suesstrunk, S.E., 2003. Measuring colorfulness in natural images, in: *Human vision and electronic imaging VIII*, International Society for Optics and Photonics. pp. 87–95.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.
- He, L., Páez, A., Liu, D., 2017. Built environment and violent crime: An environmental audit approach using google street view. *Computers, Environment and Urban Systems* 66, 83–95. doi:10.1016/j.compenvurbsys.2017.08.001.
- Hee, L., Ooi, G.L., 2003. The politics of public space planning in singapore. *Planning Perspectives* 18, 79–103.
- Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., Wang, R., 2019. Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in beijing, china. *Environment International* 126, 107–117. doi:10.1016/j.envint.2019.02.013.
- Hidayat, J.T., Ridwan, M., 2018. Assessment of the quality of public green open space (gos) in the urban fringes in response to urban sprawl phenomenon (case study district of tanah sereal, bogor city). *IOP Conference Series: Earth and Environmental Science* 179, 012027. doi:10.1088/1755-1315/179/1/012027.
- Hipp, J.R., Lee, S., Ki, D., Kim, J.H., 2021. Measuring the built environment with google street view and machine learning: Consequences for crime on street segments. *Journal of Quantitative Criminology* doi:10.1007/s10940-021-09506-9.
- Hoffmann, E., Campelo, D., Hooper, P., Barros, H., Ribeiro, A.I., 2018. Development of a smartphone app to evaluate the quality of public open space for

physical activity. an instrument for health researchers and urban planners. *Landscape and Urban Planning* 177, 191–195.

Hou, Y., Biljecki, F., 2022. A comprehensive framework for evaluating the quality of street view imagery. *International Journal of Applied Earth Observation and Geoinformation* 115, 103094. doi:10.1016/j.jag.2022.103094.

Hu, C.B., Zhang, F., Gong, F.Y., Ratti, C., Li, X., 2020. Classification and mapping of urban canyon geometry using google street view images and deep multitask learning. *Building and Environment* 167, 106424.

Ibrahim, M.R., Haworth, J., Cheng, T., 2020. Understanding cities with machine eyes: A review of deep computer vision in urban analytics. *Cities* 96, 102481. doi:10.1016/j.cities.2019.102481.

Ito, K., Biljecki, F., 2021. Assessing bikeability with street view imagery and computer vision. *Transportation Research Part C: Emerging Technologies* 132, 103371. doi:10.1016/j.trc.2021.103371.

Jiang, Y., Han, Y., Liu, M., Ye, Y., 2022. Street vitality and built environment features: A data-informed approach from fourteen chinese cities. *Sustainable Cities and Society* 79, 103724. doi:10.1016/j.scs.2022.103724.

Kang, Y., Zhang, F., Gao, S., Lin, H., Liu, Y., 2020. A review of urban physical environment sensing using street view imagery in public health studies. *Annals of GIS* 26, 261–275. doi:10.1080/19475683.2020.1791954.

Kaźmierczak, A., 2013. The contribution of local parks to neighbourhood social ties. *Landscape and urban planning* 109, 31–44.

Kelly, C.M., Wilson, J.S., Baker, E.A., Miller, D.K., Schootman, M., 2013. Using google street view to audit the built environment: inter-rater reliability results. *Annals of Behavioral Medicine* 45, S108–S112.

Ki, D., Lee, S., 2021. Analyzing the effects of green view index of neighborhood streets on walking time using google street view and deep learning. *Landscape and Urban Planning* 205, 103920. doi:10.1016/j.landurbplan.2020.103920.

Koohsari, M.J., Mavoa, S., Villanueva, K., Sugiyama, T., Badland, H., Kaczynski, A.T., Owen, N., Giles-Corti, B., 2015. Public open space, physical activity, urban design and public health: Concepts, methods and research agenda. *Health & place* 33, 75–82.

- Kruse, J., Kang, Y., Liu, Y.N., Zhang, F., Gao, S., 2021. Places for play: Understanding human perception of playability in cities using street view images and deep learning. *Computers, Environment and Urban Systems* 90, 101693. doi:10.1016/j.compenvurbsys.2021.101693.
- Lamb, K.E., Mavoa, S., Coffee, N.T., Parker, K., Richardson, E.A., Thornton, L.E., 2019. Public open space exposure measures in Australian health research: A critical review of the literature. *Geographical Research* 57, 67–83.
- Larkin, A., Gu, X., Chen, L., Hystad, P., 2021. Predicting perceptions of the built environment using GIS, satellite and street view image approaches. *Landscape and Urban Planning* 216, 104257. doi:10.1016/j.landurbplan.2021.104257.
- Larkin, A., Hystad, P., 2019. Evaluating street view exposure measures of visible green space for health research. *Journal of Exposure Science & Environmental Epidemiology* 29, 447–456. doi:10.1038/s41370-018-0017-1.
- Li, X., Li, Y., Jia, T., Zhou, L., Hijazi, I.H., 2022a. The six dimensions of built environment on urban vitality: Fusion evidence from multi-source data. *Cities* 121, 103482. doi:10.1016/j.cities.2021.103482.
- Li, X., Ratti, C., Seiferling, I., 2018. Quantifying the shade provision of street trees in urban landscape: A case study in Boston, USA, using Google Street View. *Landscape and Urban Planning* 169, 81–91. doi:10.1016/j.landurbplan.2017.08.011.
- Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening* 14, 675–685. doi:10.1016/j.ufug.2015.06.006.
- Li, Y., Yabuki, N., Fukuda, T., 2022b. Exploring the association between street built environment and street vitality using deep learning methods. *Sustainable Cities and Society* 79, 103656. doi:10.1016/j.scs.2021.103656.
- Lin, J., Wang, Q., Li, X., 2021. Socioeconomic and spatial inequalities of street tree abundance, species diversity, and size structure in New York City. *Landscape and Urban Planning* 206, 103992. doi:10.1016/j.landurbplan.2020.103992.
- Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft COCO: Common objects in context, in: *European conference on computer vision*, Springer. pp. 740–755.

- Liu, M., Jiang, Y., He, J., 2021. Quantitative evaluation on street vitality: A case study of zhoujiadu community in shanghai. *Sustainability* 13, 3027. doi:10.3390/su13063027.
- Lu, Y., 2019. Using google street view to investigate the association between street greenery and physical activity. *Landscape and Urban Planning* 191, 103435. doi:10.1016/j.landurbplan.2018.08.029.
- Luo, J., Zhao, T., Cao, L., Biljecki, F., 2022. Water view imagery: Perception and evaluation of urban waterscapes worldwide. *Ecological Indicators* 145, 109615. doi:10.1016/j.ecolind.2022.109615.
- Maruani, T., Amit-Cohen, I., 2007. Open space planning models: A review of approaches and methods. *Landscape and urban planning* 81, 1–13.
- Mavoa, S., Koohsari, M.J., Badland, H.M., Davern, M., Feng, X., Astell-Burt, T., Giles-Corti, B., 2015. Area-level disparities of public open space: A geographic information systems analysis in metropolitan melbourne. *Urban Policy and Research* 33, 306–323.
- McCormack, G.R., Rock, M., Toohey, A.M., Hignell, D., 2010. Characteristics of urban parks associated with park use and physical activity: A review of qualitative research. *Health & place* 16, 712–726.
- Mei, L., Qi, L., 2020. Central urban open space system and green economy planning based on spatial clustering algorithms and ahp model. *Journal of Intelligent & Fuzzy Systems* 39, 5785–5795. doi:doi.org/10.3233/JIFS-219218.
- Middel, A., Lukasczyk, J., Zakrzewski, S., Arnold, M., Maciejewski, R., 2019. Urban form and composition of street canyons: A human-centric big data and deep learning approach. *Landscape and Urban Planning* 183, 122–132.
- Mitrany, M., 2005. High density neighborhoods: Who enjoys them? *GeoJournal* 64, 131–140.
- Moreno-Vera, F., Lavi, B., Poco, J., 2021. Quantifying urban safety perception on street view images. *perception* 15, 38.
- Motomura, M., Koohsari, M.J., Lin, C.Y., Ishii, K., Shibata, A., Nakaya, T., Kaczynski, A.T., Veitch, J., Oka, K., 2022. Associations of public open space attributes with active and sedentary behaviors in dense urban areas: A systematic review of observational studies. *Health & Place* 75, 102816.

- Mygind, L., Bentsen, P., Badland, H., Edwards, N., Hooper, P., Villanueva, K., 2016. Public open space desktop auditing tool—establishing appropriateness for use in Australian regional and urban settings. *Urban Forestry & Urban Greening* 20, 65–70. doi:10.1016/j.ufug.2016.08.001.
- Nagendra, H., 2002. Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Applied Geography* 22, 175–186.
- Ng, L., 2019. A city in a garden, in: *Dense and Green Building Typologies*. Springer, pp. 5–6.
- Pang, H.E., Biljecki, F., 2022. 3D building reconstruction from single street view images using deep learning. *International Journal of Applied Earth Observation and Geoinformation* 112, 102859. doi:10.1016/j.jag.2022.102859.
- Paul, A., Nath, T.K., Noon, S.J., Islam, M.M., Lechner, A.M., 2020. Public open space, green exercise and well-being in Chittagong, Bangladesh. *Urban Forestry & Urban Greening* 55, 126825.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.
- Rossetti, T., Lobel, H., Rocco, V., Hurtubia, R., 2019. Explaining subjective perceptions of public spaces as a function of the built environment: A massive data approach. *Landscape and Urban Planning* 181, 169–178.
- Rundle, A.G., Bader, M.D., Richards, C.A., Neckerman, K.M., Teitler, J.O., 2011. Using Google Street View to audit neighborhood environments. *American Journal of Preventive Medicine* 40, 94–100.
- Salesses, P., Schechtner, K., Hidalgo, C.A., 2013. The collaborative image of the city: Mapping the inequality of urban perception. *PLOS ONE* 8, e68400. doi:10.1371/journal.pone.0068400.
- Schneider, A., Mertes, C.M., Tatem, A., Tan, B., Sulla-Menashe, D., Graves, S., Patel, N., Horton, J.A., Gaughan, A., Rollo, J.T., et al., 2015. A new urban landscape in East–Southeast Asia, 2000–2010. *Environmental Research Letters* 10, 034002.
- Shannon, C.E., 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 379–423.

- Simpson, E.H., 1949. Measurement of diversity. *nature* 163, 688–688.
- Spellerberg, I.F., Fedor, P.J., 2003. A tribute to claude shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘shannon–wiener’ index. *Global ecology and biogeography* 12, 177–179.
- Sun, P., Zhang, R., Jiang, Y., Kong, T., Xu, C., Zhan, W., Tomizuka, M., Li, L., Yuan, Z., Wang, C., Luo, P., 2020. SparseR-CNN: End-to-end object detection with learnable proposals. *arXiv preprint arXiv:2011.12450* .
- Szczepańska, A., Pietrzyk, K., 2020. An evaluation of public spaces with the use of direct and remote methods. *Land* 9, 419. doi:10.3390/land9110419.
- Tang, B.s., 2017. Is the distribution of public open space in hong kong equitable, why not? *Landscape and Urban Planning* 161, 80–89.
- Tang, B.s., Wong, S.w., 2008. A longitudinal study of open space zoning and development in hong kong. *Landscape and Urban Planning* 87, 258–268.
- Tang, J., Long, Y., 2019. Measuring visual quality of street space and its temporal variation: Methodology and its application in the hutong area in beijing. *Landscape and Urban Planning* 191, 103436. doi:10.1016/j.landurbplan.2018.09.015.
- Taylor, B.T., Fernando, P., Bauman, A.E., Williamson, A., Craig, J.C., Redman, S., 2011. Measuring the quality of public open space using google earth. *American journal of preventive medicine* 40, 105–112.
- UN-Habitat, 2020. Public space site-specific assessment: Guidelines to achieve quality public spaces at neighbourhood level. UN-Habitat.
- Vanwollegem, G., Van Dyck, D., Ducheyne, F., De Bourdeaudhuij, I., Cardon, G., 2014. Assessing the environmental characteristics of cycling routes to school: a study on the reliability and validity of a google street view-based audit. *International Journal of Health Geographics* 13, 1–9.
- Verma, D., Jana, A., Ramamritham, K., 2020. Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. *Building and Environment* 186, 107340. doi:10.1016/j.buildenv.2020.107340.
- Villanueva, K., Badland, H., Hooper, P., Koohsari, M.J., Mavoa, S., Davern, M., Roberts, R., Goldfeld, S., Giles-Corti, B., 2015. Developing indicators of public open space to promote health and wellbeing in communities. *Applied geography* 57, 112–119.

- Volenc, Z.M., Abraham, J.O., Becker, A.D., Dobson, A.P., 2021. Public parks and the pandemic: How park usage has been affected by covid-19 policies. *PLoS one* 16, e0251799.
- Wang, J., Foley, K., 2021. Assessing the performance of urban open space for achieving sustainable and resilient cities: A pilot study of two urban parks in dublin, ireland. *Urban Forestry & Urban Greening* 62, 127180. doi:10.1016/j.ufug.2021.127180.
- Wang, M., Chen, Z., Rong, H.H., Mu, L., Zhu, P., Shi, Z., 2022a. Ridesharing accessibility from the human eye: Spatial modeling of built environment with street-level images. *Computers, Environment and Urban Systems* 97, 101858. doi:10.1016/j.compenvurbsys.2022.101858.
- Wang, M., He, Y., Meng, H., Zhang, Y., Zhu, B., Mango, J., Li, X., 2022b. Assessing street space quality using street view imagery and function-driven method: The case of xiamen, china. *ISPRS International Journal of Geo-Information* 11, 282.
- Wang, M., Vermeulen, F., 2021. Life between buildings from a street view image: What do big data analytics reveal about neighbourhood organisational vitality? *Urban Studies* 58, 3118–3139. doi:10.1177/0042098020957198.
- Wang, R., Liu, Y., Lu, Y., Zhang, J., Liu, P., Yao, Y., Grekousis, G., 2019. Perceptions of built environment and health outcomes for older chinese in beijing: A big data approach with street view images and deep learning technique. *Computers, Environment and Urban Systems* 78, 101386. doi:10.1016/j.compenvurbsys.2019.101386.
- Wang, Z., Tang, G., Lü, G., Ye, C., Zhou, F., Zhong, T., 2021. Positional error modeling of sky-view factor measurements within urban street canyons. *Transactions in GIS* 25, 1970–1990.
- Xia, Y., Yabuki, N., Fukuda, T., 2021. Development of a system for assessing the quality of urban street-level greenery using street view images and deep learning. *Urban Forestry & Urban Greening* 59, 126995. doi:10.1016/j.ufug.2021.126995.
- Xue, F., Gou, Z., Lau, S., 2017. The green open space development model and associated use behaviors in dense urban settings: Lessons from hong kong and singapore. *Urban Design International* 22, 287–302.

- Yan, Y., Huang, B., 2022. Estimation of building height using a single street view image via deep neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing* 192, 83–98. doi:10.1016/j.isprsjprs.2022.08.006.
- Yao, Y., Wang, J., Hong, Y., Qian, C., Guan, Q., Liang, X., Dai, L., Zhang, J., 2021. Discovering the homogeneous geographic domain of human perceptions from street view images. *Landscape and Urban Planning* 212, 104125. doi:10.1016/j.landurbplan.2021.104125.
- Yap, W., Chang, J.H., Biljecki, F., 2022a. Incorporating networks in semantic understanding of streetscapes: Contextualising active mobility decisions. *Environment and Planning B: Urban Analytics and City Science* .
- Yap, W., Janssen, P., Biljecki, F., 2022b. Free and open source urbanism: Software for urban planning practice. *Computers, Environment and Urban Systems* 96, 101825. doi:10.1016/j.compenvurbsys.2022.101825.
- Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., Ratti, C., 2018. Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning* 180, 148–160. doi:10.1016/j.landurbplan.2018.08.020.
- Zhang, H., Wu, C., Zhang, Z., Zhu, Y., Lin, H., Zhang, Z., Sun, Y., He, T., Mueller, J., Manmatha, R., et al., 2022a. Resnest: Split-attention networks, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2736–2746.
- Zhang, Y., Zhang, F., Chen, N., 2022b. Migratable urban street scene sensing method based on vision language pre-trained model. *International Journal of Applied Earth Observation and Geoinformation* 113, 102989. doi:10.1016/j.jag.2022.102989.
- Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., Torralba, A., 2017a. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* .
- Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., Torralba, A., 2017b. Scene parsing through ade20k dataset, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 633–641.
- Zhou, H., Liu, L., Lan, M., Zhu, W., Song, G., Jing, F., Zhong, Y., Su, Z., Gu, X., 2021. Using google street view imagery to capture micro built environment

characteristics in drug places, compared with street robbery. *Computers, Environment and Urban Systems* 88, 101631. doi:10.1016/j.compenvurbsys.2021.101631.

Zhu, W., Wang, J., Qin, B., 2021. Quantity or quality? exploring the association between public open space and mental health in urban china. *Landscape and Urban Planning* 213, 104128. doi:10.1016/j.landurbplan.2021.104128.