Review

Street View Imagery (SVI) in the Built Environment: A Theoretical and Systematic Review

Yongchang Li [1], **Li Peng** [1], **Chengwei Wu** [1] and **Jiazhen Zhang** [2,*]

1   College of Art and Design, Nanjing Forestry University, No.159, Longpan Road, Nanjing 210037, China
2   Faculty of Architecture and Urban Planning, University of Mons, Rue d' Havre, 88, 7000 Mons, Belgium
*   Correspondence: jiazhen.zhang@umons.ac.be

Abstract: Street view imagery (SVI) provides efficient access to data that can be used to research spatial quality at the human scale. The previous reviews have mainly focused on specific health findings and neighbourhood environments. There has not been a comprehensive review of this topic. In this paper, we systematically review the literature on the application of SVI in the built environment, following a formal innovation–decision framework. The main findings are as follows: (I) SVI remains an effective tool for automated research assessments. This offers a new research avenue to expand the built environment-measurement methods to include perceptions in addition to physical features. (II) Currently, SVI is functional and valuable for quantifying the built environment, spatial sentiment perception, and spatial semantic speculation. (III) The significant dilemmas concerning the adoption of this technology are related to image acquisition, the image quality, spatial and temporal distribution, and accuracy. (IV) This research provides a rapid assessment and provides researchers with guidance for the adoption and implementation of SVI. Data integration and management, proper image service provider selection, and spatial metrics measurements are the critical success factors. A notable trend is the application of SVI towards a focus on the perceptions of the built environment, which provides a more refined and effective way to depict urban forms in terms of physical and social spaces.

Keywords: adoption strengths; artificial intelligence; building environment; critical success factors; computer vision; deep learning; street view image; visual analytics

1. Introduction

SVI is an innovative type of geographic data used for sensing the physical environment of cities [1]. SVI enables users to remotely explore realistic streetscapes by providing 360° panoramic spatial information and real-time observations of the real world from the perspective of pedestrians, encompassing natural settings and artificial landscapes [2]. Furthermore, the rapid development of deep learning and image analysis technologies has facilitated the processing of fine-scale streetscape data. The emergence of such vast data sources has provided an unprecedented opportunity for digitisation, enabling researchers to conduct large-scale studies on the urban environment and human activities.

SVI provides an emerging source of data for the research on the urban built environment, allowing for a more accurate and comprehensive audit by sensing the elements and scenes captured in SVI. The existing research in this area has primarily focused on the architectural characteristics and health implications [3]. The evaluation of built-environment exposures is a well-established field of health research that may be applied to mental and physical health outcomes [3]. In addition, the domains defined by order, e.g., building tops and façade elements [4,5], are well-established properties of the built environment. They can be used to evaluate critical architectural attributes, including a building's type, condition, and function [6–9]. Conversely, disorderly neighbourhoods, e.g., broken windows and graffiti, might imply poor socioeconomic conditions, such as high crime rates [10]. SVI provides an opportunity to virtually audit the study area and evaluate the built environment in numerous locations with little effort or financial cost [11]. Meanwhile, since SVI

systems are now available in most countries and regions around the world, including many areas without existing footprints or 3D building data, AI-based algorithms can quickly and cost-effectively be used to obtain the 3D urban morphology from SVI data, without any existing building information [12]. In addition, SVI can be used to measure urban canyon impact mechanisms such as the radiation temperature [13], buoyancy effects [14], and short-wave irradiance [15] at large scales. Metric community characteristics such as safety [16], housing prices [17], and demographic statistics [18] are derived from the mesoscale. At the microscopic scale, the habitats, resident health [19,20], and greening ornamentation in buildings have been researched [21]. The findings of these studies have suggested that the use of SVI and artificial intelligence technology to investigate the quantification and image expression of built environment factors can help to excavate additional geospatial information from the city, as well as provide more complex or specific indicators and enable large-scale and quantitative urban built environment evaluations. This approach positively improves urban resilience towards low-carbon cities [22,23] and contributes to life cycle assessments for buildings and building refurbishment [24].

Since the early days of services providing large-scale SVI, researchers have recognised that this approach is highly suitable for evaluating the characteristics of the built environment [25]. However, the few attempts to review the scope of this research area have focused on GSV only and on narrowly defined specific public health areas and micro-neighbourhood environmental aspects, or the studies have not been systematically reviewed. The previous literature has suggested a strong link between the physical urban environment and various health behaviours of citizens. Researchers in epidemiology, psychology, and geography have increasingly examined the effects of the built environment on various health outcomes. Still, few studies have examined the perceptions of the building environment at the geographic scales required for population-based studies [26–29]. The previous studies have examined the subjective perceptions of the urban environment and the role of sensations in mental health. Some studies have examined the composition of perception-related images in favour of safer, greener, or more beautiful environments [30], and such studies, while contributing to the study of SVI in health, only focus on one of the many applications in built environment research. Some rapidly emerging articles have reviewed the use of SVI to quantify the features of building environments [31,32] and to explore their feasibility of use [33]. Still, these approaches have focused on physical features, such as trees and sidewalks [26], or specific environmental exposures, such as air pollution [34]. Such research is primarily application-oriented and lacks a systematic formal framework. Additionally, the research on SVI adoption in the built environment is far behind its actual development status. A systematic review and assessment of the existing building applications of SVI has not been conducted yet. This review will fill this gap and indicate areas for future research to capitalise on this new and expanding big data source.

Given the current pace of implementation, this paper systematically and comprehensively reviews the application of SVI in the built environment. This research follows the innovation–decision progress framework [35]. The following overall research question guides this research: How should SVI be adopted and implemented in the built environment? To answer this question, this review aims to identify and summarise the relevant image platforms, data extraction and analysis methods, research applications, advantages, and limitations. The key findings are summarised, highlighting the potential value of SVI for a wide range of urban built environment research applications. This review not only supplements the deficiencies of the latest assessments of SVI in the built environment, but also provides essential guidance for using SVI technology to improve the built environment. This research is discussed as follows:

- Section 2 describes the main research methods. The adoption and implementation of SVI in the built environment are explored using the systematic review method and the innovation–decision progress mechanism.

- Section 3 explores the general characteristics of SVI and the needs and main application areas of SVI in the knowledge phase, based on the innovation–decision process.
- Section 4 analyses the potential benefits of SVI as a new data source and identifies the dilemma of its adoption in the built environment during the persuasion phase. Critical success factors (CSFs) are proposed for SVI implementation based on the reviewed publications and guidance for building environment practitioners.
- Sections 5 and 6 summarise the current trends and discuss the focus of future research on SVI-based urban environmental assessments.

2. Data Sources and Research Methods

2.1. Data Sources

In this study, the first step was to determine the selection of reference journal articles from the WoS and Scopus databases to create a unified analysis database. WoS and Scopus are still the primary sources of citation data that have authority and representativeness [36]. In addition, Scopus has a broader range of journals and WoS can enable a more comprehensive citation analysis [37], and they complement each other in this process to obtain a comprehensive view of the current state of international research and the research frontiers. The search was restricted by creating search strings to make this study more scientific [38].

The second step was to retrieve articles from the databases. Relevant topic papers were selected from journals using search terms in two academic databases (Figure 1). The first screening phase was performed by searching for titles, abstracts, authors, and keywords. Then, we excluded articles that were not available in full-text form as a second check. Multiple keywords were used to conduct a third eligibility check to capture the different trending items in building environment research, such as house prices, historical sites, and neighbourhood activities. Finally, the acquired literature was analysed based on the content, methodology, and year, among other factors (Figure 1). Since Google Maps first launched SVI in 2007, this paper captured the academic literature on the use of panoramic street images for urban built environment research published from 1 January 2007 to 7 April 2022, a span of fifteen years. Figure 1 shows the whole process of review in detail, and 263 related pieces of literature (n = 263) were ultimately obtained.

2.2. Research Methods

Firstly, a systematic review approach was identified as the primary evaluation framework for the systematic review, which consists of three main phases: literature collection, identification, and analysis. This method is flexible and accurate [39]. The data sources section completes the literature collection phase of the systematic review method.

The articles' contents are identified and categorised in the second stage. The Rogers innovation–decision process [35] was used as the formal conceptual framework to categorise the articles (Figure 2), which helps to systematically grasp the application of SVI in the built environment, including in relation to feasibility and research fields. The innovation–decision process includes the time frame from the first awareness of the innovation to its adoption or rejection by the potential adopters. This model is based on many empirical studies and contains a set of research methods, data collection approaches, and analytical models; it can be applied to studies on the diffusion of innovation and it is predictable [40]. The innovation–decision process facilitates a systematic grasp of the application of SVI in the built environment, integrates the strengths and hindrances of researchers in the specific use case, and evaluates the use of SVI in the built environment. The knowledge stage includes streetscape technology sources, requirements, and applications. At the persuasion stage, the benefits and dilemmas of panoramic images are analysed to obtain a better understanding and perspective. The decision stage occurs mainly to adopt or refuse the innovative actions caused by SVI. At the implementation stage, SVI is deployed and used in the built environment.
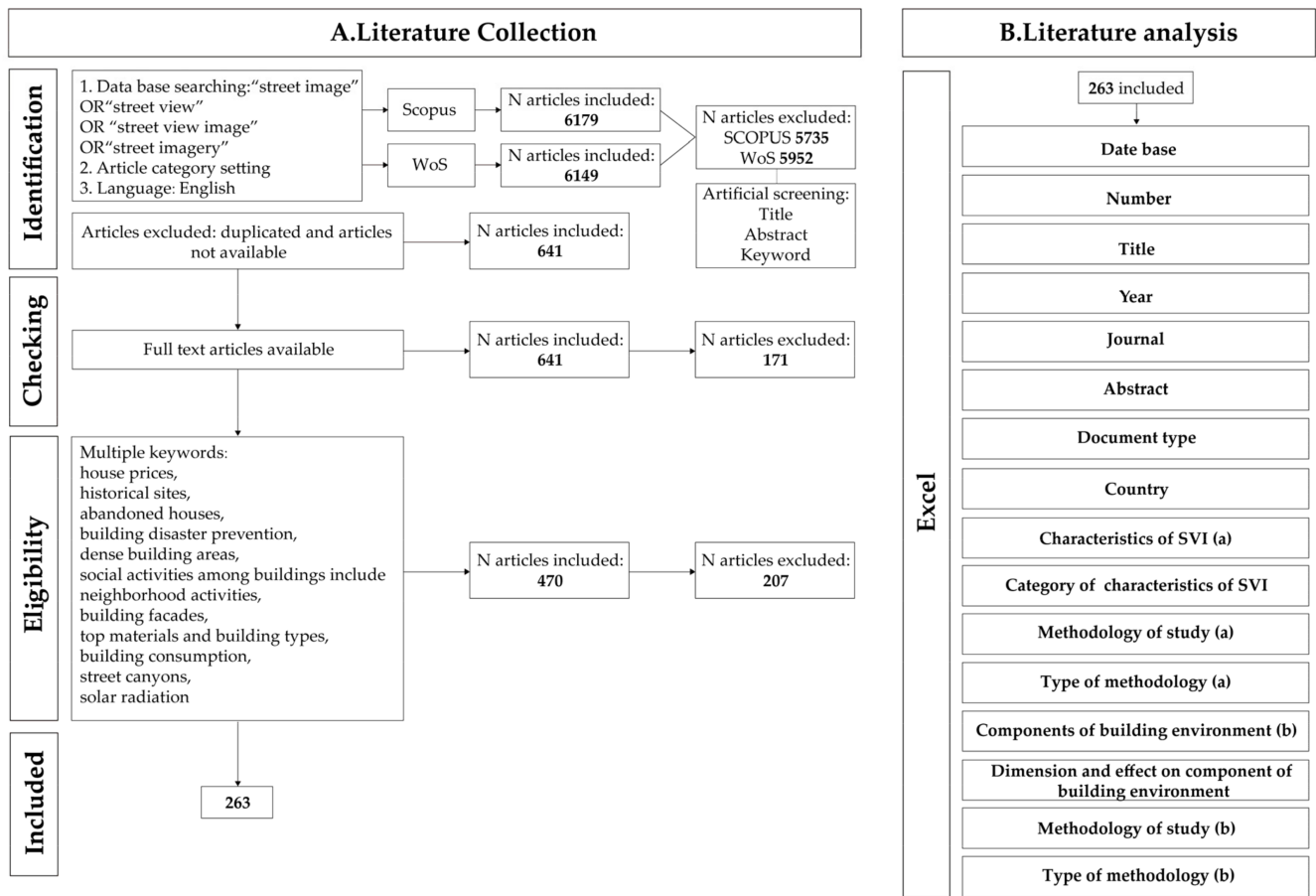
A. Literature Collection

B. Literature analysis

Figure 1. Literature collection process and analysis.
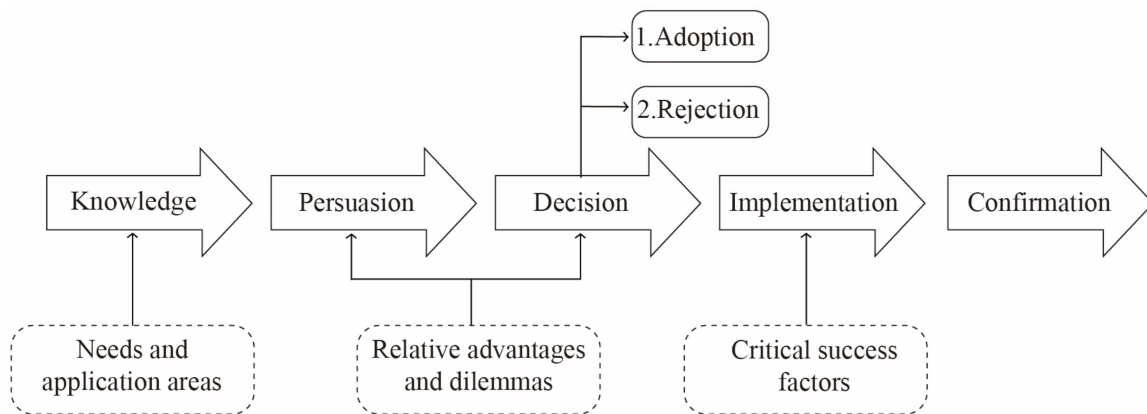
Figure 2. SVI adoption and implementation process in the built environment.

3. Application Status Analysis of SVI

3.1. Source of SVI

Currently, dozens of street view services serve as sources of SVI data, most of which are regional, covering one or a few countries. Google Street View (GSV) is a street view service provider from Google, an American company that services the entire world. However, some countries have their own local SVI services. For instance, GSV is not available in Morocco, but the local service Carte.ma Street View covers about ten major cities in Morocco. Additionally, Google services have been banned in some countries, such as China. This country has two local data sources: Baidu Street View (BSV) and Tencent Street View (TSV). In recent years, Apple integrated the "Look Around" feature into the Apple Maps app on

iOS-enabled devices. To some extent, the new feature is fairly similar to the long-standing Street View feature in Google Maps: it enables users to zoom in on a particular area. Apple Maps has enabled the new "Look Around" feature in a growing number of cities. Similar to GSV, it offers a method for interacting with maps, enabling cities to be rendered in 3D [41]. Therefore, this section includes Apple Maps within the scope of the key service providers (Table 1).

Table 1. Overview of the major service providers of SVI.

Service Provider	Territory Covered	Acquisition Method	Maximum Resolution (Width × Height)
GSV	1. Covers more than 90 countries and extends to indoor spaces (2017); 2. Not available in some countries, such as China.	1. APIs can be accessed through the web interface integrated within Google Maps, smartphone applications, and other APIs; 2. Historical images are not allowed.	2048 × 2048
BSV	296 cities in China.	The same as GSV (cancelled service in 2011)	1024 × 512
TSV	1. 652 city streets in China, covering 2,295,000 km; 2. The internal panoramic data for buildings and scenic spots are collected but only cover the categories of scenic spots, hotels, shopping malls, food, etc.	The same as GSV	960 × 960
Apple map	Covers 13 countries and 47 regions (2022)	API	-
Mapillary	Has over 1.5 billion images and over 10 million kilometres.	API	It depends on the uploader
KartaView	Over 7.6 million kilometres have been recorded by contributors all around the world.	API	It depends on the uploader

All three types of SVI data are saved as panoramas, preserving the 360° panoramic visual information of the shooting location. In practical acquisitions and applications, the visual environment of each location can be described via multiple SVIs facing distinct natural view angles. Compared with BSV and TSV, GSV is superior in its coverage and resolution. Secondly, CS can provide images from sidewalks, bike lanes, and walkways at the micro-scale compared to GSV. At large scales, CS possesses a broader coverage and temporal resolution at locations not accessible by GSV [42]. The temporal resolution of CS images will be finer in some locations than those of GSV, for which the images are typically acquired every few years and there is limited access to older images. However, CS image resources come from user uploads; thus, the image quality and field of view are often limiting factors, and the positioning accuracy of the CS images is also a cause for concern when compared to GSV. As a new supplier of street view data, Apple Maps' "Look Around" feature is more vibrant and fluid, and the photographs are of high quality. The data acquired by Apple and their high-resolution 3D photos have enabled users to obtain more accurate overall information and expansive views of highways, buildings, parks, airports, shopping malls, and other public locations [41]. In terms of privacy protection, Apple Maps has an edge over Google Maps. However, as a new supplier of street view imagery, whether Apple's new "Look Around" program is as precise and accurate as GSV is yet to be proven; therefore, more testing is required. It is also noteworthy that Google Maps can be accessed on almost any device or computer, while Apple Maps is limited to just Apple's own devices. Compared to Apple Maps, GSV is more stable, more dependable, and has greater coverage.

Overall, the spatial coverage rates of these current user-contributed services are far less comprehensive than those of GSV, which tends to have complete coverage of cities and relatively uniform sampling [43]. GSV is the most famous and extensive service to provide SVI worldwide. Table 2 compares the three types of services.

Table 2. The main comparison of the three types of services.

Similarities	Differences
1. Stored in the form of a panorama. 2. Contains 360° panoramic visual information of the shooting location. 3. Location information collected. 4. The parameters can be adjusted, such as the position, panorama, size, features, heading, spacing, radius, and light source.	Comparing GSV with BSV and TSV. Coverage: GSV > BSV and TSV; Resolution: GSV > BSV and TSV Comparing GSV with Apple Maps: Coverage: GSV > Apple Maps; Resolution: uncertain Comparing GSV with CS. (1) At the micro-scale, CS can include images from sidewalks, bike lanes, and walkways, while at large scales, CS has a broader coverage and temporal resolution at locations that are not reachable by GSV; (2) At some locations, the temporal resolution of the images will be finer than with GSV, for which images are typically acquired every few years and there is limited access to older images; (3) CS usually provides a narrower field of view than GSV images, and the extracted elements are limited; (4) There are biases in the locational accuracy of CS, which may lead to problems in map applications; (5) The spatial coverage rates of these user-contributed services are much less comprehensive than for GSV.

In addition, social media photos are crowdsourced photos shared by users on social media platforms that capture urban indoor and outdoor landscapes. Unlike SVI, streetscape photos are disseminated precisely according to the road network, while social media photos are dispersed in the city's primary locations for employment, recreation, and tourism. The former reflects the objective urban street landscape, while the latter, to some extent, express the specific groups' subjective experiences of the city. As a complement to the collection of streetscape photographs, social media photos may be used as a source. Due to the particularity of the images, social media photos have certain advantages in urban image perception [44,45].

3.2. SVI Analysis Methods

3.2.1. Computer Vision

Computer vision aims to replace human eyes with imaging equipment, to recognise and measure objects, and to extract information from pictures or high-dimensional data [46,47]. Traditional computer vision methods mostly use shallow, medium-level, and manually designed features to express images, such as colour spectrum, texture, shape, scale-invariant feature transform (SIFT) [48], the histogram of oriented gradient (HOG) [49], and generalized search trees (GIST) [50] data. These features require a substantial amount of specialist knowledge for feature engineering, have limited picture representation efficiency, and do not apply to various tasks. The introduction of AlexNet in 2012 addressed the difficulty of using feature representation in deep learning when processing high-dimensional data such as images, enabling the application of deep learning methods to image interpretation tasks, through which AlexNet can learn task-related visual characteristics autonomously.

3.2.2. Deep Learning

According to the different task types and model principles, deep learning can be divided into automatic encoder (auto-encoder), generative adversarial neural network (GAN), recursive neural network (RNN), and deep convolution neural network (DCNN) approaches. Among them, deep convolution neural networks are mainly used for image data analysis, and the landmark model AlexNet is the most extensively used classic DCNN.

For the deep learning model of computer vision tasks, the representative structures used for image object classification are AlexNet, VGG, GoogLeNet, ResNet, DenseNet, and others. Before the practical tasks are performed, this structure is often utilised to extract picture features, and then the particular network structure related to the tasks is used for the analysis. On the other hand, the training set heavily influences the model's capacity to generalise and the number of categories detected. The present deep learning model can be trained end-to-end with a high accuracy using open-source datasets (Table 3). The trained model can be directly applied to various street scenes, social media photos, and other data, providing a research basis for the image-based quantitative analysis of the urban built environment.

Table 3. Major open-source training sets.

Type	Dataset	Number of Labels
Object-oriented detection and recognition [51,52]	VOC 2012	Objects: 27,000 objects; Categories: 20.
	Microsoft COCO	Objects: 200,000 images; Categories: 80 types of objects
Object-oriented Semantic segmentation [53,54]	ADE20K	Objects: Pixel-level annotation of 435,000 objects; Categories: 150.
	Cityscapes	Objects: Pixel-level annotation of 65,000 objects; Categories: 30.
Scene-oriented type and Attribute classification [55–58]	SUN	Objects: 14,000 scene images; Categories: 102 scene attributes; Measure: closed/open, indoor/outdoor, natural/man-made, etc.
	ImageNet	Objects: Over 14 million images; Categories: 1000 categories.
	Places	Objects: 10 million nature images; Categories: Hundreds of categories of scenes.

To date, deep learning has been the primary method for analysing streetscapes. The artificial-intelligence-based SVI analysis methods apply deep learning, computer vision, and other cutting-edge artificial intelligence fields to the processing and analysis of SVI and to city-focused application practices. Compared with traditional methods, most of the methods based on digital image processing and traditional computer vision use shallow- and medium-level visual features and manually defined features, with which it is difficult to express deep semantic information in picture scenes completely and efficiently, limiting the large-scale use of SVI in urban research fields. The current computer vision technology supported by deep learning can identify semantic objects and scene contents in pictures more efficiently, providing powerful tools for extracting semantic information from street scenes alongside tools for understanding and quantitatively expressing the contents of the built environment.

3.3. Needs of SVI

The examination of building environments is not an emerging or unfamiliar field, as there are many previous studies documenting the characteristics of building environments, such as accessibility, physical barriers, accessibility to public transportation and recreational spaces, and greenery [59,60]. By detecting and understanding the elements and scenarios of the built environment, researchers can quantitatively study the urban built environment. However, in this field, most studies rely on in-person assessments and field surveys to collect data on the relevant characteristics of building environments [27,61]. Traditional urban spatial studies usually utilise self-reports, questionnaires, and field surveys. Questionnaires and self-reports are the most prevalent data sources for evaluating various neighbourhood characteristics [53]. Previous data collection methods have faced deficiencies in terms of their high labour intensity, lengthy update cycles, and geographic restrictions. With the acceleration of urbanisation, it is difficult for traditional theoretical methods to cope with rapid urban development and to describe the dynamic evolution of the urban built environment as a complex system with accurate quantitative data.

Currently, most data used for field observations are collected by walking or driving about the study region using predetermined questionnaires to record and characterise the surroundings [27], which is time-consuming and impractical for large-scale applications. The traditional approach uses a field research-based methodology, which makes it difficult to evaluate on a large, fine-grained scale [62]. Although remotely sensed images can provide a bird's eye view of cities from macro- and high altitudes, they are expensive, and their high resolution is susceptible to atmospheric influence, environmental interference, sensor jitter and other factors, making the acquired data uncertain. The secondary sources include those based on the spatial analysis and modelling of predefined environmental measures, such as spatial accessibility measures [63,64]. However, this approach fails to characterise the built environment in detail and may be limited to particular environmental factors [64].

Without adequate support from appropriate technologies, these challenges are generally inevitable. SVI data provide a visual record of a building's environmental features and can support more effective and scalable alternatives to site-based approaches. SVI systems can collect images in multiple directions to create panoramic views, and image users can observe the features contained in the built environment using audit instruments through a virtual "driving" community. With the use of SVI systems to provide object visibility and broad access to data, researchers can improve their workflow efficiency, review multiple cities simultaneously, and obtain micro-scale streetscape elements more effectively. Accuracy and coherence have been shown between observational field audits using SVI and image-based interpretations [11]. Therefore, this new paradigm should be proposed to solve these problems and guide studies on sensing the urban built environment.

3.4. Main Application Areas

SVI has been extensively used in various environmental perception practices to allow for the quantitative representation and analysis of physical spaces and to extrapolate the semantic information related to socioeconomic and human activities embedded beyond the physical space. SVI is widely adopted in building environment quantification, spatial emotion perception, and spatial semantic speculation.

3.4.1. Quantification of the Built Environment

- Element identification

Visual object recognition, scene type, and attribute classification applications are the most prevalent applications when measuring a building environment. In terms of visual object recognition, vegetation is the most sophisticated area type. SVI can capture the vegetation in the street at different height levels with an extremely high resolution. Furthermore, SVI also provides high-resolution and multi-layered information about trees, shrubs, lawns, and other forms of vegetation in the street and allows for vegetation assessments [65]. The green view index (GVI), the sky view factor (SVF), the tree view factor (TVF), and street-tree visual audit methods are often used to quantify urban greening and analyse the visibility of urban forests. In addition to cross-sectional comparative analyses, SVI offers the possibility of longitudinal studies, facilitating the analysis of the temporal changes in GVI in cities [66].

On the other hand, SVI provides multidimensional information about the form, colour, material, and other aspects of a building, which can be extracted to present the building type [67], the building's condition [68], and the building's age [69], as well as the height and number of floors of the building [70]. Other studies focused on extracting building features have involved detecting building façade features (including the building façade's colour and other features), graffiti artwork [56], and window-to-wall ratios [71]. This research field also seems to be currently focused to a large extent on more minor urban features and street facilities, or those that are often overlooked in spatial datasets, such as traffic signs and traffic signals [72,73], utility poles [74], and access holes [75].

- Physical environment assessment

SVI is applied for thermal environment simulations, the detection of sound and light environment, and air quality evaluations.

Firstly, SVI is mainly utilised for radiation and temperature simulations by combining meteorological data computations and numerical modelling calculations with information about the SVI system's shooting position and geometric characteristics [15]. With deep learning techniques, the SVI system can extrapolate the SVF in the environment to evaluate the urban heat island effect and thermal comfort [76]. In addition, the influence of vegetation on the thermal environment has become a research hotspot, and the utilisation of SVI to extract different types of vegetation on the roadside can help to analyse the spatial relationship between the layout and thermal environment. Regarding the radiance, by projecting the solar trajectory onto a fisheye image of the streetscape, SVI can be used to calculate the solar duration [77] and to quantify the total street-level shortwave irradiance [15]. In contrast to the expensive and limited use of 3D building models to calculate solar radiation, SVI fisheye images are a highly desired supplemental data source for simulating solar radiation within street canyons. However, the existing SVI-based radiation estimation models require a combination of dynamic weather conditions in practical cases and in the analysis of separated radiation direction maps.

Secondly, SVI systems are equipped with ground-based photographic equipment to capture the physical urban environment in a three-dimensional profile view, conveying more detailed visual content and calculating the various effects of elemental indicators on people's behaviour and perceptions. Therefore, SVI systems can further estimate the PV potential in densely populated metropolitan regions, areas where vehicle traffic may cause solar glare, and for human perceptions of noise. For instance, SVI systems can quantify the impacts of building façades, courtyards, and streetscapes on noise annoyance and stress levels [78], and SVI can also be used to detect traffic noise in urban environments.

Finally, deep learning methods can be used to analyse the features extracted from SVI data, and SVI can be applied to assess air quality. Mobile monitoring (either bicycle-based or GSV-based) has been frequently used to gather real-time air quality measurements to evaluate local air quality and air pollutant exposures [79–81], including black carbon [82] and particle count concentrations [83]. Meanwhile, architectural elements such as greenery and buildings in the built environment are gradually becoming crucial points in air quality research [84].

3.4.2. Emotional Perception

Individuals develop unique sensations of place based on their unique visual surroundings, experiences, and resident activities in the environment. Deep learning models trained with datasets can simulate individuals' emotions about scenes in the built environment to further evaluate the built environment with respect to three main areas: a sense of security, health, and the quality of life.

- Community safety

A sense of security is a high-level attribute of people's perceptions of urban scenes. By revealing the environmental factors associated with crash data, including road conditions [85] and road characteristics [86], the analysis of the SVI can provide valuable information for pedestrian and driver safety. In addition, the neighbourhood environmental disorder level has been considered a strong predictor of neighbourhood crime rates and residents' fear of crime. This involves physical features related to the spatial layout of buildings, street design, and the diversity of land use. Therefore, the use of SVI enables research into the relationship between crime and the physical characteristics of the built environment [87].

- Public health

SVI data represent a significant, publicly available data source that can be utilised to create metrics for the characterisation of the physical environment through machine learning techniques [88]. The current research has suggested that built environments'

characteristics are correlated with mental health and chronic disease. Further research includes concerns regarding well-being [89] and obesity [90]. In addition, the architectural characteristics may have an indirect impact on the psychological health of the occupants through factors such as the walkability [91], greenery [92], and public open spaces [93]. Stress and mental health are the primary research focal points. Infectious illness research is also vital to health and well-being since disease transmission is directly connected to environmental variables. SVI provides an excellent opportunity to examine the environments in which infectious agents breed, with current studies covering potential dengue breeding environments [94], areas of high risk for COVID-19 [95], and pathogenic environmental factors and their transmission pathways [19].

- Environmental behaviour

The building form and function and human-scale features in the built environment are the main factors influencing the vitality of a street. SVI is similar to the human perspective. Hence, it has been utilised in a wide range of urban perception studies, with the main extracted features including sidewalk quality [96], recreational facilities [97], and street interface fencing [98] features. Using SVI to analyse the quality of life in the built environment also includes identifying potential urban congestion points [99], understanding measures to mitigate near-road pollution [100], predicting the difficulty of driving a car [101], and identifying garbage dumps [102]. Moreover, the built environment can affect the behaviour of people who engage in physical activity [103]. Utilising SVI enables the measurement of residential environments related to walking infrastructure and traffic safety, such as the effect of greenery on walking behaviour [104] and walking infrastructure [105]. Cycling is another type of physical activity that has health and environmental benefits [106]. The images captured via SVI can be used to evaluate the environmental factors influencing cycling behaviour and to determine road recyclability.

3.4.3. Spatial Semantic Speculation

The urban scenes depicted in the streetscape not only convey the visual information in the scene but also implicitly express the information about the city's function, history, culture, and the socioeconomic and human activities behind the visual scenes. SVI records the city's physical environment, and the characteristics of the physical environment can predict the non-visual aspects of the city. This information can be combined with spatial material attribute data, such as household income and house price data, to check the prediction and evaluate the economic environment. The income level, education level, and even the political orientation of a neighbourhood can be inferred by identifying parked cars [107], neighbourhood store signs [108], and even vegetation [109]. The relationship between changes in urban physical space and socioeconomic levels can be studied by quantifying how places in neighbourhoods change [110]. Based on the broken window theory of the built environment, house photos, and the condition of a house's surroundings, streetscape pictures can predict crime in the neighbourhood to some extent [111]. Streetscape pictures can be used to predict house price information and perform electricity consumption assessments [112].

4. Development Outlook of SVI

4.1. Perceived Benefits

As the element that links the street to the city as a whole, the quality of the built environment is essential to the urban environment. SVI is an excellent way to observe the built environment and to examine the relationships between the built environment and its parts. SVI has numerous perceived benefits in the built environment (Table 4).

Table 4. Summary of the perceived benefits of SVI in the built environment.

Benefits	Findings (Empirical Research, Opinion-Based)
Wide coverage	360° panoramic views [113]; Views of the entire city at street level [11].
High coverage density	SVI was sampled for the road network [114]; The images of the sampling points contain street scenes of buildings, people, vehicles, trees, roads, billboards, telegraph poles, etc. [115].
Detailed content	Parameters can be adjusted [116]; Measuring multiple variables in the micro-build environment [10].
Highly efficient acquisition	Only 7.3 s to rate each item with 360° GSV scenes [117]; Automatically extracts information for a more consistent, objective, and large-scale collection [118].
Anthropomorphic perspective	Images are captured by a camera mounted on a car, bike, or backpack [118]; A rich sense of reality and strong messages [119].
Others	Remote access to location capability at a low cost [11]; NZ $0.70 per km for a field researcher and NZ $0.02 per MB for a virtual audit [11]; Comparative data at the international level [11]; Security considerations [10]; Virtual surveys can be conducted year-round, regardless of the season or weather conditions [120].

The significant benefits of SVI include its comprehensive coverage, high coverage density, complex expression level, acquisition efficiency, and anthropomorphic perspective.

Firstly, SVI systems already cover most cities within the coverage area and can be viewed in 360° [121], allowing researchers to analyse the data from a worldwide perspective.

Secondly, in terms of the coverage density, SVI provides high-density coverage of all levels of the road network in the built environment. The visual images between sampling points can be seamlessly combined, giving a complete picture of the physical spaces of urban streets.

Thirdly, regarding the expression content, the SVI provides an exhaustive and detailed representation of the actual state of the urban built environment from a human perspective. The continuous availability of high-definition images ensures the fineness of the SVI representations of the physical space in the urban built environment. With the further support of relevant artificial intelligence technologies, the precise extraction of semantic targets and the efficient understanding of a scene's content can be achieved.

Fourthly, regarding the data acquisition efficiency, Google, Baidu, and other map service providers provide commercial and free street view data under certain conditions, which can be accessed and downloaded through applicable APIs, thereby simplifying the procedure and encouraging the use of automated techniques. In addition, the use of artificial intelligence technology dramatically increases the overall audit speed [117], enabling quick and efficient evaluations of large amounts of image data, and allowing researchers to audit more streets in almost half the number of days using SVI [66].

Fifthly, SVI can capture objective cityscapes from a human perspective. The information contained in SVI data can be used to explore the intangible aspects of urban life and people's perceptions of the environment [122]. SVI captures a three-dimensional profile view of the urban streetscape and can record the views or perceived scenes from the ground.

Lastly, this review revealed that SVI is consistently secure, protecting researchers [10] and enabling them to conduct research at a low cost [120], in addition to enabling worldwide data comparisons [11].

4.2. Dilemmas of SVI Adoption

The adoption and application of SVI in the built environment presents several challenges related to image acquisition, image quality, data spatial distribution, data timing, and analysis methods (Table 5).

Table 5. Summary of dilemmas when using SVI in the built environment.

Dilemmas	Findings (Empirical Research; Opinion-Based)
Image acquisition	The images of adjacent acquisition points have similarities [118]; User-uploaded images are not available in GSV [123]; User permission restrictions [117].
Image quality	Blurred and inadequate 2D pixels result in low reliability and detection rates [113]; Instability and potential bias in extracting relevant streetscape variables [10]; The vector data registration process is complex [118].
Spatial distribution of data	Uneven distribution of countries and urban areas [124]; Open spaces and backyards are not covered [97]; Vehicles collecting GSV data may not reach every location [114].
Data time	Uncertain image update time: 2–4 years [125]; The information extracted from the images is difficult to match with weather or time data [118].
Analysis method	The depth of the neighbourhood features extracted by the computer vision model may be limited [95]; Small containers are difficult to identify, such as jars and bottles [95,97].
Cost Limitations	It takes time and is unsuitable as an online method [126];
	Computational cost and processing speed become issues [115].
	GSV images cost a maximum of USD 7 per 1000 panoramic images [97]
Others	Privacy concerns [62,97]. Technical restrictions [31].

4.2.1. Image Acquisition Challenges

The obstacles to the acquisition of SVI sources can hinder the method's application.

First, regarding the restricted access, most SVIs currently originate from map service providers such as Google and Baidu. The accessibility of street view data relies heavily on such companies' business development directions and data provision policies. Still, some service providers require that users pay to use their service, thereby increasing the acquisition cost [10]. In addition, the Google website only contains information about the devices used to capture images, the areas currently covered, and the areas they are currently imaging. Although the recent inclusion of user-uploaded data (including images) in Google Maps may increase the variability in the image quality and authenticity, user-uploaded data are included in a separate unlinked "photo sphere" that is subject to acceptance criteria [123].

Second, user permission restrictions require that users obtain prior written authorisation to publish any content provided on the map and that they do not advertise or provide instructional information about illegal activities. These regulations may severely limit the opportunities to implement SVI, such as in the field of criminology or for historical update studies of the built environment.

Third, the availability of GSV images varies worldwide because of different political, economic, legal, and technical factors. For instance, no GSV service is available in most parts of Africa, South America, the Middle East, India, China, Southeast Asia, and Russia. GSV mostly comprises sporadic, unofficial coverage in several nations. This is noteworthy, since it resembles certain other crowdsourced datasets (e.g., Mapillary).

Finally, more minor characteristics, such as door numbers, are occasionally lost or erroneous owing to the "noise" in the acquired images, rendering certain elements unsuitable for recognition using GSV images.

4.2.2. Image Quality Issues

Mapping services have established assumed quality assurance mechanisms. There will be inevitable deficiencies in the quality of images caused by factors such as lighting issues and weather, given the number of images, environmental conditions, and geographic coverage [127,128]. In addition, objects in the user's focus are often obstructed in the

images, such as passing cars and people [129]. Vegetation seems to be the main obstacle, often obscuring buildings and other objects. While this, on the one hand, facilitates the research of greenery in the built environment, it hinders the range of the images. For example, large objects such as buildings tend to be completely obscured.

4.2.3. Uneven Spatial Distribution

SVI services tend to have geographically dense coverage, but the coverage is unevenly distributed. The spatial distribution of SVI is hotspot-shaped, occurring heavily and frequently in some localised areas or cities, and remaining unavailable in about half of the countries worldwide. In contrast, smaller towns and rural areas may not always be included in areas where such features are available. This means that the research is tailored to the urban built environment [130]. In addition, the image availability and capture frequency rates vary across cities, with more affluent communities having higher image availability rates and better capture times.

4.2.4. Temporal Instability of Data

In addition to limitations in geographic coverage, the temporal instability of SVI has been criticised as a weakness in its systematic use for observing the built environment. First, the frequency of updates, which seems to be a common problem [125], may be higher because some elemental situations and features of the streetscape environment may change over time, while exhibiting random variations and regular day-, season-, or weather-related fluctuations in measurement errors. These can include the number, features, and activities of pedestrians; parked or moving vehicles; and many physically disordered markers such as litter. Thus, in some areas, images are collected infrequently (often out of date) and insufficiently to research the current conditions and perform an updated analysis or longitudinal temporal analysis (e.g., change detection). Second, the image acquisition time is frequently highlighted as a concern. The image capture rate may not match the desired research period, and inconsistencies in the time of day, season, and weather are present in field observations. SVI may also lead to bias or may not match the periods of other datasets used in the research [116]. For example, collecting streetscape videos early in the morning may result in lower levels of observed social and pedestrian activity, which may (depending on the timing of trash removal) affect the degree of physical disorder in the measured streetscape. In addition, it is noteworthy that different city sections can be captured over different periods. Finally, there is often a dispersion in the timing of the image collection, where some images are taken in winter and others are taken in summer, spring, or fall. Differences in the temporal distribution can easily lead to bias [131]. For example, in studies evaluating green spaces in the built environment, which require images taken during the same period, it is necessary to examine images and exclude data that are not from the same period to maintain the temporal consistency.

4.2.5. Analysis Method Deficiencies

There are two significant trends in the current use of street view images.

The first trend is to directly use pre-trained models based on deep learning to classify or regress street scenes. Such methods can predict explicit semantic information in street scenes, such as identifying objects and scene types. However, the pre-trained model's training set is not similar to the application set's distribution, and the "domain adaptation" diminishes the model's accuracy and affects the statistical analysis. Most research ignores rigorous statistical analyses and causal inferences, such as correlations between SVI visual items, spatial dependencies, and SVI visual objects.

The second trend is to employ deep learning models to extract generic scene features, such as the 512 features retrieved by the ResNet model based on the Places dataset [58], to represent the scene's visual similarity and specificity to other scenes in different places and regions. Since the high-dimensional features used in such methods are extracted using deep learning "black box" models, it is still difficult to interpret the semantic content

expressed by the features, which can lead to a lack of interpretability of the conclusions during research.

4.2.6. Cost Limitations

Currently, some service providers require that users pay to use their services, making the acquisition cost higher [10]. As for GSV, the Mapbox API is free of charge if the number of dynamic maps the JavaScript API calls is less than 50,000 per month [132]. To date, the cost of GSV starts at USD 0.007 each (USD 7.00 per 1000), with a usage limit of 30,000 maximum queries per minute [133]. The BSV system based on SDK (including Location SDK, Map SDK, Navigation SDK, Eagle Eye SDK, etc.) and the JavaScript API are free. The WebAPI services exceed the free quota and need to be purchased for an additional fee. Although BSV offers a variety of service purchase options, 60,000 RMB per month for multiple service options is not cheap [134].

In addition, computer vision models using supervised learning methods typically require large training datasets consisting of tens of thousands of manually labelled images to train the models adequately. Thus, the research teams must have enough time and resources to create these large training datasets. For example, the architecture for Faster R-CNN and ResNet-101, which has a near-maximum accuracy on the Microsoft COCO object detection dataset, still requires excellent runtime performance [52]. On a PC with a 3.6 GHz i7-7700 processor, 32 GB RAM, and 1080 Ti graphics, it took 95 h of processing time to perform object detection on 1 million images [97]. Therefore, the dataset processing time and cost are affected by various aspects, including not only the number of datasets but also the technical facilities required for researchers, and there is a lack of research on the comparative cost of each dataset. Although the related literature explores the differences in categories between different datasets [52], it mainly focuses on the characterisation of datasets and concentrates on the computer domain. It is noteworthy that some work has been carried out to predict the required execution times for a wide range of the most frequently used components of neural networks [135,136]. Although these approaches cannot be used to compare diverse data sets, they can be used not only to infer the execution time for a batch or entire epoch, but they can also support making a well-informed choice for the appropriate hardware and model [137]. This will contribute to future AI-based SVI research.

4.2.7. Other Dilemmas

Various other dilemmas are associated with the current use of SVI to analyse the built environment, such as privacy issues and technical costs [115,126]. Images featuring human features must be erased or hidden to protect people's privacy. This may lead to underestimating the neighbourhood environment and issues in urban safety research, which may bias study conclusions. In addition, upon review, it was found that there is a partial lack of data sharing (e.g., code and trained models), which can also lead to unavailability in some of the corresponding studies (e.g., replication or duplication).

4.3. Critical Success Factors

Even though this review shows that SVI can be used in a wide range of ways to evaluate the built environment and provide useful urban information that was unknown beforehand, there are still some challenges to be addressed in the current application of SVI. To weaken the barriers to the implementation of SVI, the CSFs developed based on a literature review of case studies deserve attention.

- Selecting an accurate image service provider

Currently, the use of SVI is in the early adoption stage. Although there are several picture resource providers (as mentioned in Section 3.1), their correctness in practical applications must be considered. OpenStreetMap (OSM) can have open data problems such as an insufficient coverage or irregular alignments, which must be handled via validation masks. These masks filter the samples used to train the model in a proper

way [18]. Therefore, selecting a reliable and image-rich provider of SVI resources is crucial for the research on the whole urban built environment.

- Appropriate spatial metrics

The quantitative measures of the built environment are mainly used for the components of the street space, including the street pavement, interfaces, and the enclosed sky view and streetscape [138], although SVI provides a vast array of features and scenarios from which to pick, which implies that unless the urban building space is evaluated quantitatively using a specified measure, the image recognition will result in a mismatch of the feature points falling on these elements. The use of appropriate geographic areas for estimating environmental exposures is essential to studying the determinants of the built environment; an uncertain geography will highlight the spatial extent to which individuals experience their environment, as well as the temporal uncertainty in the timing and duration of these experiences [139]. Consequently, it is difficult to assess whether the elements identified through SVI as spatial metrics are a true reflection of the environments exposed in everyday life. Furthermore, it is equally vital to identify appropriate metrics for activity spaces. Using various methods to define neighbourhoods and activity spaces can cause different results, such as measuring the GVI of an area through vegetation and street characteristics, which can ultimately affect the quality of the spatial perceptions within the area.

- Data integration and management

The SVI data are large in volume and dynamic, and the identification of SVI elements often requires the evaluation of a massive number of street images. Even though deep learning improves the detection efficiency compared to machine learning, it still requires the use of the researcher's equipment. Additionally, the investigation team must have sufficient time and resources to process the dataset. Therefore, the researchers need to identify targets based on the data they expect to generate. Meanwhile, there is a need to provide an effective mechanism to convert the segmented semantic SVI data into accurate and meaningful information, and more importantly to apply the acquired dataset in quantitative evaluation studies of the urban built environment. The statistical analysis mechanisms based on data and space fulfil this need. The statistical analysis techniques can be used to examine causal linkages between evaluation themes and SVI components; the spatial analysis methods can be used to visualise the spatial distribution patterns of the urban built environment's elements and to display the relationships between the urban built environment and these elements (Table 6).

Table 6. Major elemental data analysis methods.

Type	Methods	Sub-Methods
Statistical analysis	Correlation analysis method [33,140]	Spearman's coefficient; Analysis of Variance; Intraclass correlation coefficient (ICC); Pearson correlation coefficient; Kendall correlation coefficient.
	Analytic hierarchy process [141]	None
	Regression analysis [104,142]	Stepwise regression; Ridge regression; Lasso regression; Confusion Matrix.
Spatial analysis	Graph-based spatial analysis [62]	Overlay analysis; Buffer analysis; Network analysis.
	Data-based spatial analysis [143,144]	Getis'G; Moran's I; Poisson regression.

5. Discussion

In this research, we analysed the present applications of SVI in several respects. Firstly, the systematic description of the numerous applications highlights the adaptability and variety of SVI systems. Secondly, the innovative decision process framework can help in systematically reviewing the research challenges in the built environment, which is why

SVI is required as a novel method to acquire data, offering a comprehensive clarification of how SVI is adopted and implemented. This provides valuable guidance for understanding the adoption of and the decision-making process for SVI in the built environment, which is relatively uncommon in other studies, as most are application oriented. Thirdly, this paper summarises the present benefits and challenges of SVI, allowing researchers to make quick judgments. These advantages and limitations are equally instructive in studies of SVI promoting healthy cities, walkability, urban planning, and other related issues.

In this review, we found that experiments and simulations are the primary tools for evaluating the urban built environment. Deep learning is the standard and most sophisticated approach to image processing. Deep learning is commonly utilised in research, and it accelerates the extraction of features and the segmentation of pictures, which is crucial for much of the research discussed in Section 3. In addition, due to the advantageous nature of deep learning in terms of semantic speculation, the research on urban environment perception has received increasing attention. The application of SVI analysis methods in urban research has advanced beyond scene categorisation, object backdrop distinction, and position detection to the physical environments of streets and to spatial perception.

One of the most prevalent current uses of SVI involves plants and greenery. The following factors were considered herein: (1) Street trees, shrubs, lawns, and other forms of greenery have long been known to be essential elements of urban landscape design. SVI provides multiple benefits in urban environments, meeting diverse and overlapping goals. (2) Street greenery significantly contributes to the beauty and walkability of residential streets. The presence of plants often enhances the aesthetic assessment of urban settings. (3) Remote sensing imagery has been used to calculate green space percentages, green space/building area ratios, green space densities, and other factors so as to analyse, evaluate, and visualise urban greenery [145]. At the same time, SVI provides an entirely new perspective when assessing the profile view of street greenery. The integration of both data types bridges the gap between the previous studies and provides new research perspectives, with additional research opportunities for urban greenery studies.

In addition, the detection of temporal variation in SVI is becoming increasingly attractive [146], and temporal variation has been intensely discussed in relation to the recent data infrastructure for urban architecture. However, most street view services (including GSV as the most popular service) do not allow the retrieval of historical images through APIs. The only time-series studies have also collected data from GSV's web interface (including historical images) or other means, rather than through APIs [147,148]. The current SVI providers gradually continue to restrict access, which may favour the development of crowdsourcing services. This may alleviate these problems but could substantially limit the current research in the field.

Lastly, the real value of SVI can only be seen when combined with existing semantic segmentation techniques. It is difficult to segregate SVI applications and papers into meaningful groups, since some cover more than one domain, but this shows that the research topic is multidisciplinary.

5.1. Future Studies

The future studies on the evaluation of the built environment based on SVI should also focus on the following areas:

- The integration of various data sources, such as remote sensing images, geotagged social media data, cell phone signalling, and bus cards. Attention should also be paid to the use of new methods, including deep learning and big data analysis, to conduct multiangle and multilevel research within a fine-scale perspective on the urban environment and to improve the reliability and accuracy of the evaluation based on SVI in the built environment.
- With the advent of the 5G era, the real-time uploading of SVI data recorded via webcasts, geotagged social media, and traffic loggers will be faster and easier. Street view data stored in the cloud will be more diverse and available in real-time. Com-

puter vision technologies and web-based developments also offer the possibility for interactive platforms that enable street-view uploading and analysis. The growing CS platforms, developments in autonomous driving, and urban infrastructure are expected to address the current spatial coverage and temporal sampling frequency issues of street view data with crowdsourced street view sharing platforms such as OpenStreetCam and Mappilary.

- With the recent progress that has been made, the deep learning technology applied in urban research appears to be more accurate. The trend of semantic segmentation to achieve more rapid and higher-resolution images was identified by combing the previous related research. The current semantic segmentation approach mainly involves low-resolution representation or recovering high-resolution representation learning. With the progress and development of deep learning technology, high- and low-resolution parallel learning is gaining more attention. In addition, the latest semantic segmentation model can provide more possibilities for street view images in urban research.

- As the dense coverage of indoor data at the microscale becomes more available (e.g., the extension of voluntary SVI), we predict that this might bring about enhancements and novelties for applications such as change detection for indoor data.

- In addition to using street maps for elemental measurements [8], the current research advances show the feasibility of generating 3D models from street maps [149–151], which can be combined with a model database to quickly generate virtual cities with certain style requirements and high accuracy.

5.2. Restrictions

The limitations of this research are worth discussing and focusing on in future research work:

- Advancements in computer vision and processing capacity are crucial for the future development of SVI. However, the chosen publications do not address the technical research elements, and the investigation of the semantic segmentation approaches for SVI is beyond the scope of this review.

- The concept of the "built environment" in this review may still limit the applications of SVI. For example, this paper does not consider studies of urban parks, trails, and urban agricultural areas. This review also deliberately excludes direct traffic observations through SVI, which may affect the assessment of the perceived quality of life in the built environment.

- The growing interest in SVI research and the corresponding increase in the number of publications has created a need for other researchers to follow this area and contribute to the knowledge base.

In conclusion, SVI research must be given continual and dynamic attention. Next-generation information technologies such as big data, artificial intelligence (AI), cloud computing (5G), and Internet of Things (IoT) technologies all have dynamic development rules that need to be taken into consideration. SVI research in the built environment should be closely tracked across disciplines to ensure that the literature is complete and to better clarify the development of SVI research in general.

6. Conclusions

This paper provides a comprehensive answer to adopting and implementing SVI in the built environment by summarising and analysing the papers contained in the Scopus and the WoS databases.

Primarily, SVI can capture elements of the built environment at the line-of-sight level for assessment at a lower cost. With the considerable development of this urban data source and the establishment of supporting infrastructure (e.g., services), the use of SVI for urban analysis has become a trend that will continue to grow for the foreseeable future, as seen in the number of SVI-related studies and applications.

Secondly, with the support of artificial intelligence technology, using the street-level landscape and three-dimensional profile information provided by SVI, representative evaluation elements, such as roads, pedestrians, trees, and buildings, can be selected to analyse and evaluate a specific environmental element or comprehensive environmental elements within the spatial scope of streets, communities, and cities. This enables the quantification of the urban environment, environmental perception detection, and semantic speculation.

Thirdly, SVI adoption is not an easy task, presenting obstacles from both the imagery and technology facets. The data integration and management, the selection of appropriate imagery service providers, and spatial metrics are the critical success factors that can reduce these barriers.

Finally, the supporting infrastructure (e.g., services, volume, coverage of data, and computer vision technology) needs to be further developed and enhanced. The future application trends for SVI are mainly focused on the perception of the urban environment. The research shows that the emergence of SVI provides a practical aid for analysing urban environmental perceptions in terms of the spatiotemporal coverage and granularity, offering the possibility for more refined, efficient, and large-scale depictions of urban forms from physical and social spaces. At the same time, this study provides planning and design policy advice and assistance for sustainable smart city development and the management of residents' health regarding urban design teaching, research, and practice.

Author Contributions: L.P. developed the research topic. Y.L. wrote the original draft. C.W. and J.Z. were responsible for the review and editing. The project administration was undertaken by Y.L. All the authors contributed to writing the paper. All the authors have read and agreed to the published version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tang, Y.; Zhang, J.; Liu, R.; Li, Y. Exploring the Impact of Built Environment Attributes on Social Followings Using Social Media Data and Deep Learning. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 325. [CrossRef]
2. Ye, Y.; Richards, D.; Lu, Y.; Song, X.; Zhuang, Y.; Zeng, W.; Zhong, T. Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landsc. Urban Plan.* **2019**, *191*, 103434. [CrossRef]
3. Li, X.; Zhang, C.; Li, W.; Ricard, R.; Meng, Q.; Zhang, W. Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban For. Urban Green.* **2015**, *14*, 675–685. [CrossRef]
4. Zhang, G.; Pan, Y.; Zhang, L. Deep learning for detecting building façade elements from images considering prior knowledge. *Automat. Constr.* **2022**, *133*, 104016. [CrossRef]
5. Zhong, T.; Ye, C.; Wang, Z.; Tang, G.; Zhang, W.; Ye, Y. City-scale mapping of urban façade color using street-view imagery. *Remote Sens.* **2021**, *13*, 1591. [CrossRef]
6. Yu, Q.; Wang, C.; McKenna, F.; Yu, S.X.; Taciroglu, E.; Cetiner, B.; Law, K.H.J.E.E.; Vibration, E. Rapid visual screening of soft-story buildings from street view images using deep learning classification. *Earthq. Eng. Eng. Vib.* **2020**, *19*, 827–838. [CrossRef]
7. Laupheimer, D.; Tutzauer, P.; Haala, N.; Spicker, M. Neural networks for the classification of building use from street-view imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *4*, 177–184. [CrossRef]
8. Kang, J.; Körner, M.; Wang, Y.; Taubenböck, H.; Zhu, X.X. Building instance classification using street view images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 44–59. [CrossRef]
9. Gonzalez, D.; Rueda-Plata, D.; Acevedo, A.B.; Duque, J.C.; Ramos-Pollan, R.; Betancourt, A.; Garcia, S. Automatic detection of building typology using deep learning methods on street level images. *Build. Environ. Behav.* **2020**, *177*, 106805. [CrossRef]
10. Zhou, H.; Liu, L.; Lan, M.; Zhu, W.; Song, G.; Jing, F.; Zhong, Y.; Su, Z.; Gu, X. Using Google Street View imagery to capture micro built environment characteristics in drug places, compared with street robbery. *Comput. Environ. Urban Syst.* **2021**, *88*, 101631. [CrossRef]

11. Badland, H.M.; Opit, S.; Witten, K.; Kearns, R.A.; Mavoa, S. Can virtual streetscape audits reliably replace physical streetscape audits? *J. Hered.* **2010**, *87*, 1007–1016. [CrossRef] [PubMed]

12. Pang, H.E.; Biljecki, F. Geoinformation. 3D building reconstruction from single street view images using deep learning. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102859.

13. Kim, E.S.; Yun, S.H.; Park, C.Y.; Heo, H.K.; Lee, D.K. Estimation of Mean Radiant Temperature in Urban Canyons Using Google Street View: A Case Study on Seoul. *Remote Sens.* **2022**, *14*, 260. [CrossRef]

14. Zhao, Y.; Li, H.; Kubilay, A.; Carmeliet, J. Buoyancy effects on the flows around flat and steep street canyons in simplified urban settings subject to a neutral approaching boundary layer: Wind tunnel PIV measurements. *Sci. Total Environ.* **2021**, *797*, 149067. [CrossRef]

15. Carrasco-Hernandez, R.; Smedley, A.R.D.; Webb, A.R. Using urban canyon geometries obtained from Google Street View for atmospheric studies: Potential applications in the calculation of street level total shortwave irradiances. *Energ. Build.* **2015**, *86*, 340–348. [CrossRef]

16. Porzi, L.; Rota Bulò, S.; Lepri, B.; Ricci, E. Predicting and understanding urban perception with convolutional neural networks. In Proceedings of the 23rd ACM International Conference on Multimedia, Brisbane, Australia, 26–30 October 2015; pp. 139–148.

17. Law, S.; Paige, B.; Russell, C. Take a look around: Using street view and satellite images to estimate house prices. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19. [CrossRef]

18. Ayala Lauroba, C.; Sesma Redín, R.; Aranda, C.; Galar, M. A deep learning approach to an enhanced building footprint and road detection in high-resolution satellite imagery. *Remote Sens.* **2021**, *13*, 3135. [CrossRef]

19. Zhang, Y.; Chen, N.; Du, W.; Li, Y.; Zheng, X. Multi-source sensor based urban habitat and resident health sensing: A case study of Wuhan, China. *Build. Environ.* **2021**, *198*, 107883. [CrossRef]

20. Qiu, L.; Zhu, X. Housing and community environments vs. Independent mobility: Roles in promoting children's independent travel and unsupervised outdoor play. *Int. J. Environ. Res. Public Health* **2021**, *18*, 2132. [CrossRef]

21. Ringland, J.; Bohm, M.; Baek, S.R.; Eichhorn, M. Automated survey of selected common plant species in Thai homegardens using Google Street View imagery and a deep neural network. *Earth Sci. Inform.* **2021**, *14*, 179–191. [CrossRef]

22. Wu, C.; Cenci, J.; Wang, W.; Zhang, J. Resilient City: Characterization, Challenges and Outlooks. *Buildings* **2022**, *12*, 516. [CrossRef]

23. Zhu, Y.; Koutra, S.; Zhang, J. Zero-Carbon Communities: Research Hotspots, Evolution, and Prospects. *Buildings* **2022**, *12*, 674. [CrossRef]

24. Mayer, M.; Bechthold, M. Data granularity for life cycle modelling at an urban scale. *Arch. Sci. Rev.* **2020**, *63*, 351–360. [CrossRef]

25. Kelly, C.M.; Wilson, J.S.; Baker, E.A.; Miller, D.K.; Schootman, M. Using Google Street View to audit the built environment: Inter-rater reliability results. *Ann. Behav. Med.* **2013**, *45*, S108–S112. [CrossRef] [PubMed]

26. Kang, Y.; Zhang, F.; Gao, S.; Lin, H.; Liu, Y. A review of urban physical environment sensing using street view imagery in public health studies. *Ann. GIS* **2020**, *26*, 261–275. [CrossRef]

27. Rzotkiewicz, A.; Pearson, A.L.; Dougherty, B.V.; Shortridge, A.; Wilson, N. Systematic review of the use of Google Street View in health research: Major themes, strengths, weaknesses and possibilities for future research. *Health Place* **2018**, *52*, 240–246. [CrossRef]

28. Charreire, H.; Casey, R.; Salze, P.; Simon, C.; Chaix, B.; Banos, A.; Badariotti, D.; Weber, C.; Oppert, J.-M. Measuring the food environment using geographical information systems: A methodological review. *Public Health Nutr.* **2010**, *13*, 1773–1785. [CrossRef]

29. Charreire, H.; Mackenbach, J.D.; Ouasti, M.; Lakerveld, J.; Compernolle, S.; Ben-Rebah, M.; McKee, M.; Brug, J.; Rutter, H.; Oppert, J.-M. Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: A systematic review (the SPOTLIGHT project). *Health Place* **2014**, *25*, 1–9. [CrossRef]

30. Larkin, A.; Gu, X.; Chen, L.; Hystad, P.J.L.; Planning, U. Predicting perceptions of the built environment using GIS, satellite and street view image approaches. *Landsc. Urban Plan.* **2021**, *216*, 104257. [CrossRef]

31. Biljecki, F.; Ito, K.J.L.; Planning, U. Street view imagery in urban analytics and GIS: A review. *Landsc. Urban Plan.* **2021**, *215*, 104217. [CrossRef]

32. He, N.; Li, G. Urban neighbourhood environment assessment based on street view image processing: A review of research trends. *Environ. Chall.* **2021**, *4*, 100090. [CrossRef]

33. Rundle, A.G.; Bader, M.D.; Richards, C.A.; Neckerman, K.M.; Teitler, J.O. Using Google Street View to audit neighborhood environments. *Am. J. Prev. Med.* **2011**, *40*, 94–100. [CrossRef] [PubMed]

34. Weichenthal, S.; Hatzopoulou, M.; Brauer, M. A picture tells a thousand . . . exposures: Opportunities and challenges of deep learning image analyses in exposure science and environmental epidemiology. *Environ. Int.* **2019**, *122*, 3–10. [CrossRef] [PubMed]

35. Rogers, E.M.; Singhal, A.; Quinlan, M.M. Diffusion of innovations. In *An Integrated Approach to Communication Theory and Research*; Routledge: London, UK, 2014; pp. 432–448.

36. Mongeon, P.; Paul-Hus, A. The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics* **2016**, *106*, 213–228. [CrossRef]

37. Falagas, M.E.; Pitsouni, E.I.; Malietzis, G.A.; Pappas, G. Comparison of PubMed, Scopus, web of science, and Google scholar: Strengths and weaknesses. *FASEB J.* **2008**, *22*, 338–342. [CrossRef]

38. Birkle, C.; Pendlebury, D.A.; Schnell, J.; Adams, J.J. Web of Science as a data source for research on scientific and scholarly activity. *Quant. Sci. Stud.* **2020**, *1*, 363–376. [CrossRef]

39. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
40. Seligman, L.J. Sensemaking throughout adoption and the innovation—Decision process. *Eur. J. Innov. Manag.* **2006**. [CrossRef]
41. Apple Map Usage. Available online: https://www.apple.com/maps/ (accessed on 1 June 2022).
42. Mahabir, R.; Schuchard, R.; Crooks, A.; Croitoru, A.; Stefanidis, A.J. Crowdsourcing street view imagery: A comparison of mapillary and OpenStreetCam. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 341. [CrossRef]
43. Alvarez Leon, L.F.; Quinn, S.J.G. The value of crowdsourced street-level imagery: Examining the shifting property regimes of OpenStreetCam and Mapillary. *GeoJournal* **2019**, *84*, 395–414. [CrossRef]
44. Callau, A.À.; Albert, M.Y.P.; Rota, J.J.; Giné, D.S. Landscape characterization using photographs from crowdsourced platforms: Content analysis of social media photographs. *Open Geosci.* **2019**, *11*, 558–571. [CrossRef]
45. Oteros-Rozas, E.; Martín-López, B.; Fagerholm, N.; Bieling, C.; Plieninger, T. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecol. Indic.* **2018**, *94*, 74–86. [CrossRef]
46. Lürig, M.D.; Donoughe, S.; Svensson, E.I.; Porto, A.; Tsuboi, M. Computer vision, machine learning, and the promise of phenomics in ecology and evolutionary biology. *Front. Ecol. Evol.* **2021**, *9*, 148. [CrossRef]
47. Ren, Z.; Fang, F.; Yan, N.; Wu, Y. State of the art in defect detection based on machine vision. *Int. J. Precis. Eng. Manuf. Technol.* **2021**, *9*, 661–691. [CrossRef]
48. Lowe, D.G. Object recognition from local scale-invariant features. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; pp. 1150–1157.
49. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; pp. 886–893.
50. Oliva, A.; Torralba, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vis.* **2001**, *42*, 145–175. [CrossRef]
51. Everingham, M.; van Gool, L.; Williams, C.K.; Winn, J.; Zisserman, A.J.I.j.o.c.v. The pascal visual object classes (voc) challenge. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [CrossRef]
52. Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In Proceedings of the 2014 European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 740–755.
53. Wang, R.; Yuan, Y.; Liu, Y.; Zhang, J.; Liu, P.; Lu, Y.; Yao, Y. Using street view data and machine learning to assess how perception of neighborhood safety influences urban residents' mental health. *Health Place* **2019**, *59*, 102186. [CrossRef]
54. Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The cityscapes dataset for semantic urban scene understanding. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 3213–3223.
55. Patterson, G.; Hays, J. Sun attribute database: Discovering, annotating, and recognizing scene attributes. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2751–2758.
56. Novack, T.; Vorbeck, L.; Lorei, H.; Zipf, A. Towards detecting building facades with graffiti artwork based on street view images. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 98. [CrossRef]
57. Zhang, F.; Zhou, B.; Liu, L.; Liu, Y.; Fung, H.H.; Lin, H.; Ratti, C.J.L.; Planning, U. Measuring human perceptions of a large-scale urban region using machine learning. *Landsc. Urban Plan.* **2018**, *180*, 148–160. [CrossRef]
58. Zhou, B.; Lapedriza, A.; Khosla, A.; Oliva, A.; Torralba, A. Places: A 10 million image database for scene recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 1452–1464. [CrossRef]
59. Macintyre, S.; Ellaway, A. Ecological approaches: Rediscovering the role of the physical and social environment. *Soc. Epidemiol.* **2000**, *9*, 332–348.
60. Masoumi, H.E. Associations of built environment and children's physical activity: A narrative review. *Rev. Environ. Health* **2017**, *32*, 315–331. [CrossRef] [PubMed]
61. McCurley, J.L.; Gutierrez, A.P.; Gallo, L.C. Diabetes prevention in US Hispanic adults: A systematic review of culturally tailored interventions. *Am. J. Prev. Med.* **2017**, *52*, 519–529. [CrossRef] [PubMed]
62. Hao, X.; Long, Y. Street greenery: A new indicator for evaluating walkability. *Shanghai Urban Plan. Rev.* **2017**, *1*, 32–36.
63. Leslie, E.; Cerin, E. Are perceptions of the local environment related to neighbourhood satisfaction and mental health in adults? *Prev. Med.* **2008**, *47*, 273–278. [CrossRef]
64. Pliakas, T.; Hawkesworth, S.; Silverwood, R.J.; Nanchahal, K.; Grundy, C.; Armstrong, B.; Casas, J.P.; Morris, R.W.; Wilkinson, P.; Lock, K. Optimising measurement of health-related characteristics of the built environment: Comparing data collected by foot-based street audits, virtual street audits and routine secondary data sources. *Health Place* **2017**, *43*, 75–84. [CrossRef]
65. Xia, Y.; Yabuki, N.; Fukuda, T.J.U.F.; Greening, U. Development of a system for assessing the quality of urban street-level greenery using street view images and deep learning. *Urban For. Urban Green.* **2021**, *59*, 126995. [CrossRef]
66. Li, X.J.E.; Analytics, P.B.U.; Science, C. Examining the spatial distribution and temporal change of the green view index in New York City using Google Street View images and deep learning. *Environ. Plan. B Urban Anal. City Sci.* **2021**, *48*, 2039–2054. [CrossRef]
67. Doersch, C.; Singh, S.; Gupta, A.; Sivic, J.; Efros, A. What makes Paris look like Paris? *Commun. ACM* **2012**, *31*, 103–110. [CrossRef]

68. Koch, D.; Despotovic, M.; Sakeena, M.; Döller, M.; Zeppelzauer, M. Visual estimation of building condition with patch-level ConvNets. In Proceedings of the 2018 ACM Workshop on Multimedia for Real Estate Tech, Yokohama, Japan, 11 June 2018; pp. 12–17.

69. Zeppelzauer, M.; Despotovic, M.; Sakeena, M.; Koch, D.; Döller, M. Automatic prediction of building age from photographs. In Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval, Yokohama, Japan, 11–14 June 2018; pp. 126–134.

70. Kim, H.; Han, S. Interactive 3D building modeling method using panoramic image sequences and digital map. *Multimed. Tools Appl.* **2018**, *77*, 27387–27404. [CrossRef]

71. Szcześniak, J.T.; Ang, Y.Q.; Letellier-Duchesne, S.; Reinhart, C.F. A method for using street view imagery to auto-extract window-to-wall ratios and its relevance for urban-level daylighting and energy simulations. *Build. Environ.* **2022**, *207*, 108108. [CrossRef]

72. Nassar, A.S.; Lefevre, S. Automated Mapping of Accessibility Signs with Deep Learning from Ground-level Imagery and Open Data. In Proceedings of the 2019 Joint Urban Remote Sensing Event, Vannes, France, 22–24 May 2019; pp. 1–4.

73. Lu, Y.; Lu, J.; Zhang, S.; Hall, P. Traffic signal detection and classification in street views using an attention model. *Comput. Vis. Media* **2018**, *4*, 253–266. [CrossRef]

74. Barranco-Gutiérrez, A.I.; Martínez-Díaz, S.; Gómez-Torres, J.L. An approach for utility pole recognition in real conditions. In Proceedings of the Pacific-Rim Symposium on Image and Video Technology, Guanajuato, Mexico, 28 October–1 November 2013; pp. 113–121.

75. Vishnani, V.; Adhya, A.; Bajpai, C.; Chimurkar, P.; Khandagle, K. Manhole detection using image processing on google street view imagery. In Proceedings of the 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 20–22 August 2020; pp. 684–688.

76. Johansson, E. Influence of urban geometry on outdoor thermal comfort in a hot dry climate: A study in Fez, Morocco. *Build. Environ.* **2006**, *41*, 1326–1338. [CrossRef]

77. Middel, A.; Lukasczyk, J.; Maciejewski, R. Sky view factors from synthetic fisheye photos for thermal comfort routing—A case study in Phoenix, Arizona. *Urban Plan.* **2017**, *2*, 19–30. [CrossRef]

78. Zhang, K.; Qian, Z.; Yang, Y.; Chen, M.; Zhong, T.; Zhu, R.; Lv, G.; Yan, J. Using street view images to identify road noise barriers with ensemble classification model and geospatial analysis. *Sustain. Cities Soc.* **2022**, *78*, 103598. [CrossRef]

79. Idso, C.D.; Idso, S.B.; Balling, R.C., Jr. The urban CO_2 dome of Phoenix, Arizona. *Phys. Geogr.* **1998**, *19*, 95–108. [CrossRef]

80. Brantley, H.; Hagler, G.; Kimbrough, E.; Williams, R.; Mukerjee, S.; Neas, L. Mobile air monitoring data-processing strategies and effects on spatial air pollution trends. *Atmos. Meas. Tech.* **2014**, *7*, 2169–2183. [CrossRef]

81. Hankey, S.; Marshall, J.D. On-bicycle exposure to particulate air pollution: Particle number, black carbon, PM2. 5, and particle size. *Atmos Environ.* **2015**, *122*, 65–73. [CrossRef]

82. Alexeeff, S.E.; Roy, A.; Shan, J.; Liu, X.; Messier, K.; Apte, J.S.; Portier, C.; Sidney, S.; van den Eeden, S.K. High-resolution mapping of traffic related air pollution with Google street view cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. *Environ. Health* **2018**, *17*, 1–13. [CrossRef]

83. Alizadeh Kharazi, B.; Behzadan, A.H. Flood depth mapping in street photos with image processing and deep neural networks. *Comput. Environ. Urban Syst.* **2021**, *88*, 101628. [CrossRef]

84. Wu, D.; Gong, J.; Liang, J.; Sun, J.; Zhang, G. Analyzing the Influence of Urban Street Greening and Street Buildings on Summertime Air Pollution Based on Street View Image Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 500. [CrossRef]

85. Hanson, C.S.; Noland, R.B.; Brown, C. The severity of pedestrian crashes: An analysis using Google Street View imagery. *J. Transp. Geogr.* **2013**, *33*, 42–53. [CrossRef]

86. Mooney, S.J.; DiMaggio, C.J.; Lovasi, G.S.; Neckerman, K.M.; Bader, M.D.; Teitler, J.O.; Sheehan, D.M.; Jack, D.W.; Rundle, A.G. Use of Google Street View to assess environmental contributions to pedestrian injury. *Am. J. Public Health* **2016**, *106*, 462–469. [CrossRef] [PubMed]

87. Amiruzzaman, M.; Curtis, A.; Zhao, Y.; Jamonnak, S.; Ye, X. Classifying crime places by neighborhood visual appearance and police geonarratives: A machine learning approach. *J. Comput. Soc. Sci.* **2021**, *4*, 813–837. [CrossRef]

88. Keralis, J.M.; Javanmardi, M.; Khanna, S.; Dwivedi, P.; Huang, D.; Tasdizen, T.; Nguyen, Q.C. Health and the built environment in United States cities: Measuring associations using Google Street View-derived indicators of the built environment. *BMC Public Health* **2020**, *20*, 1–10. [CrossRef]

89. Hart, E.A.C.; Lakerveld, J.; McKee, M.; Oppert, J.-M.; Rutter, H.; Charreire, H.; Veenhoven, R.; Bárdos, H.; Compernolle, S.; De Bourdeaudhuij, I. Contextual correlates of happiness in European adults. *PLoS ONE* **2018**, *13*, e0190387. [CrossRef]

90. Yang, Y.; Lu, Y.; Yang, L.; Gou, Z.; Zhang, X. Urban greenery, active school transport, and body weight among Hong Kong children. *Travel Behav. Soc.* **2020**, *20*, 104–113. [CrossRef]

91. Wang, R.; Lu, Y.; Zhang, J.; Liu, P.; Yao, Y.; Liu, Y.J. The relationship between visual enclosure for neighbourhood street walkability and elders' mental health in China: Using street view images. *J. Transp. Health* **2019**, *13*, 90–102. [CrossRef]

92. Helbich, M.; Yao, Y.; Liu, Y.; Zhang, J.; Liu, P.; Wang, R. Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China. *Environ. Int.* **2019**, *126*, 107–117. [CrossRef]

93. Taylor, B.T.; Fernando, P.; Bauman, A.E.; Williamson, A.; Craig, J.C.; Redman, S. Measuring the quality of public open space using Google Earth. *Am. J. Prev. Med.* **2011**, *40*, 105–112. [CrossRef]

94. Haddawy, P.; Wettayakorn, P.; Nonthaleerak, B.; Su Yin, M.; Wiratsudakul, A.; Schöning, J.; Laosiritaworn, Y.; Balla, K.; Euaungkanakul, S.; Quengdaeng, P. Large scale detailed mapping of dengue vector breeding sites using street view images. *PLoS Neglect. Trop. D* **2019**, *13*, e0007555. [CrossRef] [PubMed]

95. Nguyen, Q.C.; Huang, Y.; Kumar, A.; Duan, H.; Keralis, J.M.; Dwivedi, P.; Meng, H.-W.; Brunisholz, K.D.; Jay, J.; Javanmardi, M. Using 164 million google street view images to derive built environment predictors of COVID-19 cases. *Int. J. Environ. Res. Public Health* **2020**, *17*, 6359. [CrossRef] [PubMed]

96. Gustat, J.; Anderson, C.E.; Chukwurah, Q.C.; Wallace, M.E.; Broyles, S.T.; Bazzano, L.A. Cross-sectional associations between the neighborhood built environment and physical activity in a rural setting: The Bogalusa Heart Study. *BMC Public Health* **2020**, *20*, 1–10. [CrossRef] [PubMed]

97. Kruse, J.; Kang, Y.; Liu, Y.-N.; Zhang, F.; Gao, S. Places for play: Understanding human perception of playability in cities using street view images and deep learning. *Comput. Environ. Urban Syst.* **2021**, *90*, 101693. [CrossRef]

98. Meng, L.; Wen, K.-H.; Zeng, Z.; Brewin, R.; Fan, X.; Wu, Q.J.S. The impact of street space perception factors on elderly health in high-density cities in Macau—analysis based on street view images and deep learning technology. *Sustainability* **2020**, *12*, 1799. [CrossRef]

99. Qin, K.; Xu, Y.; Kang, C.; Kwan, M.P. A graph convolutional network model for evaluating potential congestion spots based on local urban built environments. *Trans. GIS* **2020**, *24*, 1382–1401. [CrossRef]

100. Gabbe, C.; Oxlaj, E.; Wang, J. Residential development and near-roadway air pollution: Assessing risk and mitigation in San Jose, California. *J. Transp. Health* **2019**, *13*, 78–89. [CrossRef]

101. Skurowski, P.; Paszkuta, M. Saliency map based analysis for prediction of car driving difficulty in Google street view scenes. In *AIP Conference Proceedings*; AIP Publishing LLC: Melville, NY, USA, 2018; p. 110003.

102. Conley, G.; Zinn, S.C.; Hanson, T.; McDonald, K.; Beck, N.; Wen, H. Using a deep learning model to quantify trash accumulation for cleaner urban stormwater. *Comput. Environ. Urban Syst.* **2022**, *93*, 101752. [CrossRef]

103. Ewing, R.; Handy, S.J. Measuring the unmeasurable: Urban design qualities related to walkability. *J. Urban Des.* **2009**, *14*, 65–84. [CrossRef]

104. Li, X.; Zhang, C.; Li, W.; Kuzovkina, Y.A.; Weiner, D.J.U.F.; Greening, U. Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' socioeconomic conditions in Hartford, Connecticut, USA. *Urban For. Urban Green.* **2015**, *14*, 751–759. [CrossRef]

105. Yin, L.; Wang, Z. Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Appl. Geogr.* **2016**, *76*, 147–153. [CrossRef]

106. Lu, Y.; Yang, Y.; Sun, G.; Gou, Z. Associations between overhead-view and eye-level urban greenness and cycling behaviors. *Cities* **2019**, *88*, 10–18. [CrossRef]

107. Gebru, T.; Krause, J.; Wang, Y.; Chen, D.; Deng, J.; Aiden, E.L.; Fei-Fei, L. Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13108–13113. [CrossRef] [PubMed]

108. Ma, R.; Wang, W.; Zhang, F.; Shim, K.; Ratti, C. Typeface reveals spatial economical patterns. *Sci. Rep.* **2019**, *9*, 1–9. [CrossRef] [PubMed]

109. Wang, R.; Feng, Z.; Pearce, J.; Yao, Y.; Li, X.; Liu, Y. The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images. *Sustain. Cities Soc.* **2021**, *66*, 102664. [CrossRef]

110. Kreft, H.; Jetz, W. Global patterns and determinants of vascular plant diversity. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 5925–5930. [CrossRef] [PubMed]

111. Arietta, S.M.; Efros, A.A.; Ramamoorthi, R.; Agrawala, M. City forensics: Using visual elements to predict non-visual city attributes. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 2624–2633. [CrossRef]

112. Rosenfelder, M.; Wussow, M.; Gust, G.; Cremades, R.; Neumann, D. Predicting residential electricity consumption using aerial and street view images. *Appl. Energy* **2021**, *301*, 117407. [CrossRef]

113. Xue, F.; Li, X.; Lu, W.; Webster, C.J.; Chen, Z.; Lin, L. Big Data-Driven Pedestrian Analytics: Unsupervised Clustering and Relational Query Based on Tencent Street View Photographs. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 561. [CrossRef]

114. Zhai, W.; Peng, Z.-R. Damage assessment using Google street view: Evidence from hurricane Michael in Mexico beach, Florida. *Appl. Geogr.* **2020**, *123*, 102252. [CrossRef]

115. Zhi, L.; Xiao, Z.; Qiang, Y.; Qian, L. Street-Level Image Localization Based on Building-Aware Features via Patch-Region Retrieval under Metropolitan-Scale. *Remote Sens.* **2021**, *13*, 4876. [CrossRef]

116. Byun, G.; Kim, Y. A street-view-based method to detect urban growth and decline: A case study of Midtown in Detroit, Michigan, USA. *PLoS ONE* **2022**, *17*, e0263775. [CrossRef] [PubMed]

117. Plascak, J.J.; Rundle, A.G.; Babel, R.A.; Llanos, A.A.; LaBelle, C.M.; Stroup, A.M.; Mooney, S.J. Drop-and-spin virtual neighborhood auditing: Assessing built environment for linkage to health studies. *Am. J. Prev. Med.* **2020**, *58*, 152–160. [CrossRef] [PubMed]

118. Yin, L.; Cheng, Q.; Wang, Z.; Shao, Z. 'Big data'for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts. *Appl. Geogr.* **2015**, *63*, 337–345. [CrossRef]

119. Liu, M.; Han, L.; Xiong, S.; Qing, L.; Ji, H.; Peng, Y. Large-scale street space quality evaluation based on deep learning over street view image. In Proceedings of the 2019 International Conference on Image and Graphics, Beijing, China, 23–25 August 2019; pp. 690–701.

120. Berland, A.; Lange, D.A. Google Street View shows promise for virtual street tree surveys. *Urban For. Urban Green.* **2017**, *21*, 11–15. [CrossRef]

121. Goel, R.; Garcia, L.M.; Goodman, A.; Johnson, R.; Aldred, R.; Murugesan, M.; Brage, S.; Bhalla, K.; Woodcock, J. Estimating city-level travel patterns using street imagery: A case study of using Google Street View in Britain. *PLoS ONE* **2018**, *13*, e0196521. [CrossRef]

122. Zhang, Q.-S.; Zhu, S.-C. Visual interpretability for deep learning: A survey. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 27–39. [CrossRef]

123. Whitehead, J.; Smith, M.; Anderson, Y.; Zhang, Y.; Wu, S.; Maharaj, S.; Donnellan, N.J. Improving spatial data in health geographics: A practical approach for testing data to measure children's physical activity and food environments using Google Street View. *Int. J. Health Geogr.* **2021**, *20*, 1–15. [CrossRef]

124. Nguyen, T.T.; Nguyen, Q.C.; Rubinsky, A.D.; Tasdizen, T.; Deligani, A.H.N.; Dwivedi, P.; Whitaker, R.; Fields, J.D.; DeRouen, M.C.; Mane, H.; et al. Google Street View-Derived Neighborhood Characteristics in California Associated with Coronary Heart Disease, Hypertension, Diabetes. *Int. J. Environ. Res. Public Health* **2021**, *18*, 10428. [CrossRef]

125. Suel, E.; Bhatt, S.; Brauer, M.; Flaxman, S.; Ezzati, M. Multimodal deep learning from satellite and street-level imagery for measuring income, overcrowding, and environmental deprivation in urban areas. *Remote Sens. Environ.* **2021**, *257*, 112339. [CrossRef]

126. Zhang, Y.; Siriaraya, P.; Kawai, Y.; Jatowt, A. Automatic latent street type discovery from web open data. *Inf. Syst.* **2020**, *92*, 101536. [CrossRef]

127. Li, X.; Ratti, C.; Seiferling, I. Quantifying the shade provision of street trees in urban landscape: A case study in Boston, USA, using Google Street View. *Landsc. Urban Plan.* **2018**, *169*, 81–91. [CrossRef]

128. Lauko, I.G.; Honts, A.; Beihoff, J.; Rupprecht, S. Local color and morphological image feature based vegetation identification and its application to human environment street view vegetation mapping, or how green is our county? *Geo-Spatial Inf. Sci.* **2020**, *23*, 222–236. [CrossRef]

129. Bin, J.; Gardiner, B.; Li, E.; Liu, Z. Multi-source urban data fusion for property value assessment: A case study in Philadelphia. *Neurocomputing* **2020**, *404*, 70–83. [CrossRef]

130. Szczepańska, A.; Pietrzyk, K. An evaluation of public spaces with the use of direct and remote methods. *Land* **2020**, *9*, 419. [CrossRef]

131. Larkin, A.; Hystad, P. Evaluating street view exposure measures of visible green space for health research. *J. Expo. Sci. Environ. Epidemiol.* **2019**, *29*, 447–456. [CrossRef]

132. Mapbox Pricing. In: Mapbox [Internet]. Available online: https://www.mapbox.com/pricing/ (accessed on 2 June 2022).

133. Street View Static API Usage and Billing | Street View Static API | Google Developers. In: Google Developers [Internet]. Available online: https://developers.google.com/maps/documentation/streetview/usage-and-billing (accessed on 20 May 2022).

134. Street View Static API Usage and Billing | Street View Static API | Baidu Map. Available online: https://lbsyun.baidu.com/products/panoramic (accessed on 13 June 2022).

135. Qi, H.; Sparks, E.R.; Talwalkar, A. Paleo: A Performance Model for Deep Neural Networks. In Proceedings of the International Conference on Learning Representations (ICLR), San Juan, Puerto Rico, 2–4 May 2016. Available online: https://openreview.net/forum?id=SyVVJ85lg (accessed on 8 June 2022).

136. Coleman, C.; Narayanan, D.; Kang, D.; Zhao, T.; Zhang, J.; Nardi, L.; Bailis, P.; Olukotun, K.; Ré, C.; Zaharia, M. Dawnbench: An end-to-end deep learning benchmark and competition. *Training* **2017**, *100*, 102.

137. Justus, D.; Brennan, J.; Bonner, S.; McGough, A.S. Predicting the computational cost of deep learning models. In Proceedings of the 2018 IEEE international conference on big data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 3873–3882.

138. Jianguo, W. *Urban Design*; Southeast University Press: Nanjing, China, 2011.

139. Kwan, M.-P. The uncertain geographic context problem. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 958–968. [CrossRef]

140. Duoqian, M.; Qinghua, Z.; Yuhua, Q.; Ji-Ye, L.; Guo-Yin, W.; Wei-Zhi, W.; Yang, G.; Lin, S.; Shenming, G.; Hongyun, Z. From human intelligence to machine implementation model: Theories and applications based on granular computing. *Caai Trans. Intell. Syst.* **2016**, *6*, 743–757.

141. Xie, J.P. *Green Design Evaluation and Optimization*; China University of Geosciences Press: Beijing, China, 2004.

142. Li, X.; Zhang, C.; Li, W. Does the visibility of greenery increase perceived safety in urban areas? Evidence from the place pulse 1.0 dataset. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 1166–1183. [CrossRef]

143. Salesses, P.; Schechtner, K.; Hidalgo, C.A. The collaborative image of the city: Mapping the inequality of urban perception. *PLoS ONE* **2013**, *8*, e68400. [CrossRef] [PubMed]

144. He, L.; Páez, A.; Liu, D. Built environment and violent crime: An environmental audit approach using Google Street View. *Comput. Environ. Urban Syst.* **2017**, *66*, 83–95. [CrossRef]

145. Faryadi, S.; Taheri, S. Interconnections of urban green spaces and environmental quality of Tehran. *Int. J. Environ. Res.* **2009**, *3*, 199–208.

146. Revaud, J.; Heo, M.; Rezende, R.S.; You, C.; Jeong, S.-G. Did it change? Learning to Detect Point-of-Interest Changes for Proactive Map Updates. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4086–4095.

147. Najafizadeh, L.; Froehlich, J.E. A Feasibility Study of Using Google Street View and Computer Vision to Track the Evolution of Urban Accessibility. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; pp. 340–342.

148. Cândido, R.L.; Steinmetz-Wood, M.; Morency, P.; Kestens, Y. Reassessing urban health interventions: Back to the future with Google street View time machine. *Am. J. Prev. Med.* **2018**, *55*, 662–669. [CrossRef]

149. Kim, S.; Kim, D.; Choi, S. CityCraft: 3D virtual city creation from a single image. *Vis. Comput.* **2020**, *36*, 911–924. [CrossRef]

150. Wang, X.; Tang, P.; Shi, X. Analysis and Conservation Methods of Traditional Architecture and Settlement Based on Knowledge Discovery and Digital Generation—A Case Study of Gunanjie Street in China. In Proceedings of the 24th CAADRIA Conference, Wellington, New Zealand, 15–18 April 2019. [CrossRef]

151. Toker, A.; Zhou, Q.; Maximov, M.; Leal-Taixé, L. Coming down to earth: Satellite-to-street view synthesis for geo-localization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 6488–6497.