land

MDPI

Article

Measuring Perceived Walkability at the City Scale Using Open Data

Yang Ye [1,2,*], Chaozhi Jia [1,2] and Stephan Winter [3]

1 School of Architecture and Design, Harbin Institute of Technology, Harbin 150001, China
2 Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology, Harbin 150001, China
3 Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
* Correspondence: yeyang@hit.edu.cn

Abstract: The walkability of the built environment has been shown to be critical to the health of residents, and open data have been widely used to assess walkability. However, previous research has focused on the relationship between the built environment and walking behavior rather than perceived walkability, and there is a lack of systematic research on walkability at the urban scale using open data. This paper presents a methodological framework for systematically measuring and assessing perceived walkability at the urban scale, considering general and specific features. The walkability indices are obtained using variables from open data or calculated automatically through machine learning and algorithms to ensure they can be evaluated at a larger urban scale. The proposed method is applied to Harbin, China, to assess the perceived walkability of streets using hundreds of thousands of street view images and points of interest obtained from open data. The results are compared with a subjective evaluation of walkability to validate the proposed method. The results demonstrate that measures of the urban built environment can describe perceived walkability. Thus, the proposed framework shows promise for assessing the walkability of urban spaces, supporting policy proposals, and establishing design guidelines for optimising urban spaces.

Keywords: perceived walkability; perception; built environment; open data

check for updates

Citation: Ye, Y.; Jia, C.; Winter, S. Measuring Perceived Walkability at the City Scale Using Open Data. *Land* **2024**, *13*, 261. https://doi.org/10.3390/land13020261

Academic Editors: Seung-Nam Kim, Gi-Hyoug Cho, Hee-Chul Kim and Maria Rosa Trovato

Received: 1 January 2024
Revised: 16 February 2024
Accepted: 17 February 2024
Published: 19 February 2024

1. Introduction

Walkability is an important characteristic of the built environment since it is a quantitative and qualitative measurement of how inviting or un-inviting an area is to pedestrians [1]. Traditionally, research frameworks of walkability were based on studies of the relationship between the built environment and walking behavior [2,3]. Studies from a public health perspective have also investigated the relationship between the built environment and the frequency and type of physical activities. In addition, objective quantitative assessment methods and models have been used to evaluate factors influencing walking activities [4]. Although the results of objective quantitative assessments are relatively accurate and reliable, the indices may not match residents' perceptions, and it is difficult to reflect the street level walkability from the perspective of the pedestrians' perception. In other words, the subjective dimension or perceived walkability has often been overlooked, until recently.

Perceived walkability, i.e., how walk-friendly people experience a certain urban area or street space [5], has received limited attention. Only recent studies have shown that environmental elements affect people's willingness to walk, and objective and perceived walkability are highly correlated but not always in the same direction [6,7]. In addition, Jensen et al. [8] found that walkability measurements typically did not include micro-scale streetscape features associated with perceived walkability [9,10]. Therefore, a framework for assessing perceived walkability based on existing frameworks is required to conduct an objective assessment of perceived walkability.

Another shortcoming of current perceived walkability research is a lack of methods applicable to the urban scale and with a balance between universality and local identity. Previous studies of human scale walkability have used field surveys and questionnaires, but these methods are time consuming and cannot be carried out on a large scale. In addition, although many studies have been carried out, it is difficult to compare the results in terms of the impact of the built environment on perceived walkability because they have been conducted in different times and spaces using different audit tools. Open data, as an emerging data source, has a wide variety of data, many of which are closely linked to the built environment at the micro-scale, such as POI data or road network data, and street view image data can well simulate the built environment from the human perspective. In addition, its data coverage is wide, dense, and frequently updated, and can be rapidly collected on a large scale. Therefore, the shortcomings of current perceived walkability research can be addressed by combining built environment elements information obtained from open data with the weights of each element in terms of perceived walkability in general or in a particular dimension obtained through Analytic Hierarchy Process (AHP) or Principal Component Analysis (PCA), etc. Thus, it can become a potential tool for large-scale acquisition of urban built environment features for perceived walkability research.

Based on the above analysis, the following three research objectives are proposed:

1. Develop a framework for assessing perceived walkability. Specifically, identify the elements of the built environment that are relevant to perceived walkability and the extent to which they influence perceived walkability.
2. Establish a methodology for using open data to assess perceived walkability in a built environment within the above framework. Specifically, the relationship between perceived walkability and elements of the built environment will be expressed through mathematical formulae and algorithms based on open data, forming a perceived walkability audit scale.
3. Validate the perceived walkability audit scale. Specifically, the evaluation results of this scale through open data can reflect the pedestrian's walking experience.

2. Literature Review

This section provides the background on walkability research with a special focus on the relationship between users' perception of the built environment and data-based measurements of walkability in the urban built environment.

2.1. The Relationship between Perceived Walkability and the Built Environment

Recent research on perceived walkability has shown that the built environment, especially elements related to the human scale, significantly influences pedestrians' walking experiences. The results of these studies demonstrate both the generality and specificity of urban environmental elements in relation to the walking experience. The generality aspect is reflected in the fact that factors of the built environment that influence perceived walkability are similar across studies, e.g., the number and density of sidewalks and intersections influence a pedestrian's perception of a destination's accessibility [11,12], the number of everyday amenities within walking distance is correlated with a measure of convenience [13,14], the aesthetic elements increase the pedestrian's comfort [15–17], and the traffic volume, sidewalk width, and the level of street lighting [18–20] can affect the sense of safety. This suggesting that a generalised research framework can be established to assess perceived walkability. On the other hand, the specificity aspect is reflected in the degree of influence of the built environment's elements on perceived walkability and the correlation between the elements and perceived walkability differ for different climates, geographic areas, demographics, cultural backgrounds, urban designs, and walkability policies [21–23]. Overall, the built environment has a significant impact on perceived walkability. Although the elements of the built environment that affect perceived walkability are generally the same, the degree to which the elements of the built environment affect perceived walkability may vary in different research contexts. Therefore, research on perceived walkability requires an assessment methodology with a general framework that reflects the relationship between perceived walkability and the built environment, and with elements and their weights that reflect the characteristics of different research contexts.

2.2. Utility of Open Data for Walkability and Built Environment Research

Open data, including street view imagery (SVI), OpenStreetMap (OSM) topographic data, and point of interest (POI) data, have been widely used to evaluate the relationship between walkability and the built environment. The data are freely available in large quantities from several data providers, such as Google, Tencent, and Baidu. Open data have been used in two main research areas of walkability and the built environment. The first area has focused on identifying attributes of the built environment and their distribution, and their relationship with objective walkability at large spatial and temporal scales, i.e., a district, city, or a comparison of multiple cities. These include extracting a specific element from open data and reflecting the pedestrian friendliness of the urban environment through its distribution pattern [24], or the walking-related characteristics of the urban road network [25] at the city scale. Alternatively, the identification results of several environmental elements can be combined and calculated to describe [26] or predict the distribution pattern of walking-related features in a given urban form pattern [27]. These studies confirmed that identifying and extracting built environment attributes from SVI through semantic segmentation reflected their real distribution at a large scale and described correlations between the built environment and walking trip characteristics. However, the walking preference from the subjective perspective of pedestrians was ignored.

The second type of research has focused on establishing relationships between the built environment and human perception. These studies analysed the relationship between human perception and characteristics of the built environment. These studies characterised people's preferences for particular urban environments through differences in shortest paths and actual behaviour between two locations [28], or measured the relationship between people's perceptions and urban environments through mass online questionnaires [29]. Data collection and comparative analyses have been carried out in a number of cities [30,31], and the results provide insights into the relationship between the built environment and subjective human perceptions such as boring, lively, and safe. The methodology and results of these studies provide guidance for investigating the relationship between perceived walkability and the urban environment, although the environmental elements associated with perceived walkability and the similarities and differences in their influence on other perceptions need to be further explored.

In conclusion, current research has demonstrated a strong link between the human-scale built environment and pedestrians' walking experience, and the potential of open data and algorithms in assessing perceived walkability. However, there are still several research gaps in perceived walkability assessment research, including the establishment of a research framework that is both general and adapted to the specificities of different research contexts, the elements of the built environment that are relevant to perceived walkability, and the methodology for collecting and calculating these elements through these open data. Filling these gaps is a contribution of this study.

3. Development and Formulation of a Walkability Scale

3.1. A General Perceived Walkability Framework

This paper proposes a perceived walkability measurement framework based on the research described in the literature review and specifically on the framework Jacobs used to describe the effect of the built street environment on people's perceptions of a liveable city [32]. The proposed framework considers five pedestrian experiences: accessibility, convenience, visual comfort, safety, and climate adaptation. The focus is on built environment features at the microscale. The built environment elements for assessing perceived walkability are derived from audit tools used in former studies, including NEWS [33], MAPS [34], EAST-HK [35], FASTVIEW [36], PEAT [37], and REAT [38]. The following factors were evaluated:

1. Accessibility. A well-connected road network encourages walking due to shorter distances and less time for typical daily trips while meeting the needs of a diverse range of travellers and providing various convenient road options.
2. Convenience. This refers to the ease of reaching destinations (or places of daily amenities) within walking distance.
3. Climate adaptation. This refers to the existence of facilities and their spatial design to reduce pedestrian vulnerability to the effects of extreme weather conditions. Many cities suffer from long, cold winters or hot summers, and urban spatial designers should consider the walking experience and reduce the negative impact of extreme weather conditions on pedestrians.
4. Visual comfort. This refers to the degree to which the built environment provides a visually pleasant experience for pedestrians. A pleasant visual environment provides visual appeal on walking journeys and keeps people interested in walking.
5. Safety. This factor includes social and traffic safety.

3.2. Development of a Perceived Walkability Scale for Harbin City

A two-round expert survey with questionnaires was conducted to develop the perceived walkability scale. The questionnaires included a scale of 5 factors and 20 indices, which were selected based on the audit tools mentioned in Section 3.1. The built environment factors and indices related to walking behaviour or perception included in each audit tool were first listed, from which the factor categories used in most of the scales were selected, as well as the built environment indices more commonly associated with medium-to-high density urban environments. The factors and indices included in the scale also took into account the specificities of the city of Harbin. Harbin is China's northernmost provincial capital city, with a population of 4.6 million and five hundred square kilometres of a built-up urban area. Since Harbin is located inland and at a high latitude, the area has a very cold climate with mild summers and long, cold winters. Central heating is required for six months of the year. The average temperature is −19 °C in January and 23 °C in July.

The questionnaire was administered to 20 experts in the fields of urban design, architecture, and urban management, all of whom have been working in their respective fields for more than 10 years, 15 of whom work and live in Harbin for more than 10 months a year, and another 5 of whom work in Beijing, Shanghai, and Tianjin but have lived in Harbin for more than 5 years. All experts reported being "very familiar" or "familiar" with walkability studies and the built environment of Harbin. Of the 20 experts, 8 were female and 12 were male. The average age was 42.8 years. Each expert participated in both rounds of the survey.

The purpose of the first round of questionnaires was to screen for factors and indices that should be included in the perceived walkability scale. We sent the questionnaire to each expert and asked them to select whether these factors and indices were appropriate for assessing perceived walkability and whether any factors and indices should be added. As a result, all 19 experts agreed on the five general context factors, three of the 20 indices were deleted because they were considered to be weakly related to perceived walkability, and four were considered to be similar to the other indices and were therefore combined into one. This resulted in a perceived walkability scale with 5 factors and 12 indices.

The purpose of the second round of the questionnaire was to determine the weight of each factor and index. Here we used the Analytical Hierarchy Process (AHP) [39], which is a common analytical method used in sociological research for complex multi-objective decision problems. The AHP method usually treats the objective problem as a system and calculates the hierarchical ranking by building a hierarchical model, constructing a judgement (pairwise comparisons) matrix and fuzzy quantitative methods for qualitative indicators. This in turn generates multi-level, multi-indicator decision making and evaluation. We asked each expert to compare the importance of each factor and index in pairs, using a five-point scale ranging from 1—very unimportant to 5—very important.

3.3. Formulation of the Indices Based on Open Data

In this section, we define the indices relevant to perceived walkability in the results of the expert questionnaire and develop equations for them. It is assumed that the variables can be obtained from open data sources or calculated automatically using machine learning and algorithms. Subsequently, the indices were evaluated at the larger urban scale.

3.3.1. Accessibility

Two indices were used for the analysis of road networks (Table 1). Spatial design network analysis (sDNA) is a convenient method for calculating these indices [40]. As the method in sDNA, we describe roads as a network of nodes and edges and use intersections as endpoints. The betweenness of a node is defined as the degree to which this node is selected, and the closeness is defined as the ease of access from one point to the rest of the network within a walking radius.

Table 1. Equation and computation for indices of accessibility factor.

Index	Equation	Computation Process
Betweenness	$\mathrm{Bte}(x) = \sum_{s \neq t \neq v} \frac{\partial st(x)}{\partial st}$ $\mathrm{Bte}(x)$ denotes the value of betweenness index, $\partial st(x)$ denotes the number of times the shortest path between point s and t passes through point x within the radius R, and ∂st denotes the number of shortest paths between point s and t.	Betweenness index and closeness index will be calculated using the sDNA plug-in in GIS and the R-value represents the interrelationship of paths within the corresponding distance of activity for a particular path, in this case 400 m, which is the empirical distance considered acceptable for most pedestrians in walkability studies.
Closeness	$\mathrm{Clo}(x) = \sum_{y \in R_x} \frac{P(y)}{d_\theta(x,y)}$ $\mathrm{Clo}(x)$ denotes the value of Closeness index, p (y) is the weight of node y within the radius R of the computational range, $d\theta(x,y)$ denotes the shortest path distance from point x to point y	

3.3.2. Convenience

The convenience factor is calculated by using POI and street network data (Table 2). Access to public transit is represented by a connection. The function diversity index is a measure of the richness of the variety of travel destinations, and the function-density represents the density of basic services on both sides of the street, i.e., the ratio of the number of different types of residential service POIs in the street section to the length of the walking space.

Table 2. Equation and computation for indices of convenience factor.

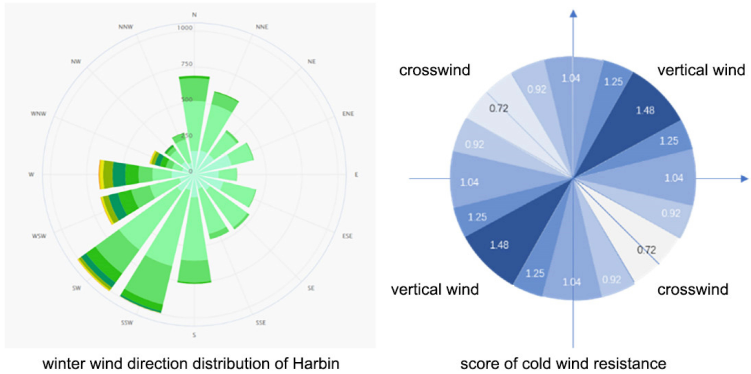
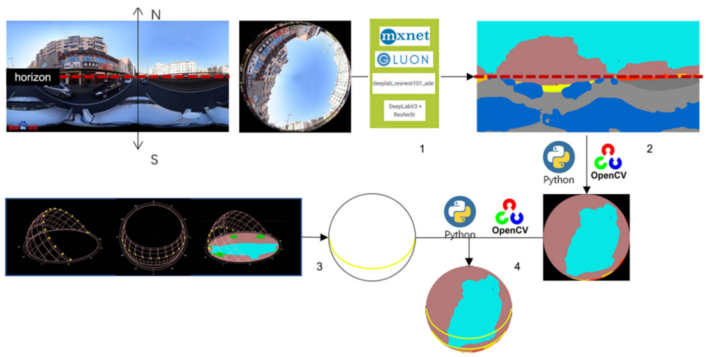
Index	Equation	Computation Process
Connection	$Trans = \sum Ni \times wi$ Ni denotes the number of type i public transport facilities within the buffer zone of the road segment, wi denotes the weight of type i public transport facilities	1. Select POI of all types of traffic facilities within the buffer zone of the street segment. (The buffer zone is the area within a radius of 400 metres * from the centre of the street segment) 2. Calculate the product of the number of POIs and weights for each type of traffic facility and add them together. (Transport facility weighting: bus stop = 0.1, metro stop = 0.7 BRT stop = 0.2)
Function diversity	$\mathrm{Ser} = -\sum_{i=1}^{n} P_i \ln P_i$ n denotes the total number of POI types, P_i denotes the share of the i-th POI type in the total. If there is only one POI type, the Shannon index is the minimum value of 0.	1. Collect POI within main city area of Harbin City, and divide POI into 11 categories: eating and drinking, shopping, health care, amenities, public space, sports and leisure, education and culture, public transport, government agencies, corporations, public facilities. 2. Select POI within the buffer zone of the street segment. (The buffer zone is the area within a radius of 400 metres * from the centre of the street segment) 3. Calculate function diversity index and function density index as the equation.
Function density	$\mathrm{Qua} = \frac{N_{poi}}{L}$ N_{poi} denotes the total number of facility POIs within the buffer of the street segment, L denotes the length of the street segment.	

* A 400 m buffer zone is the empirical value considered acceptable for most pedestrians in walkability studies.

3.3.3. Climate Adaptation

Three indices are used to assess the adaptation to climate: snow load, direct solar radiation, and winter wind environment (Table 3). The snow load index is the capacity of the sidewalk buffer space to accommodate snow generated by snow shovelling. A study of the subjective perceptions of the thermal environment quality in winter found no significant correlation between residents' perceptions of the temperature when walking outdoors and the air temperature in the street. However, residents were more sensitive to sunlight conditions and wind speed [41]. Therefore, the winter wind environment index and the direct solar radiation index are included.

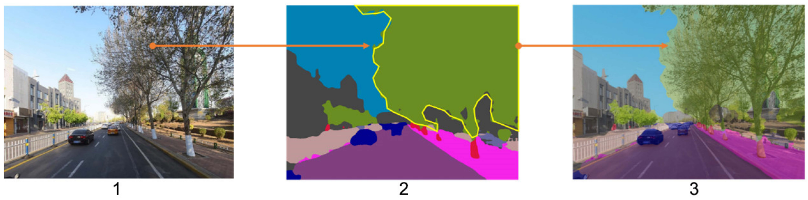
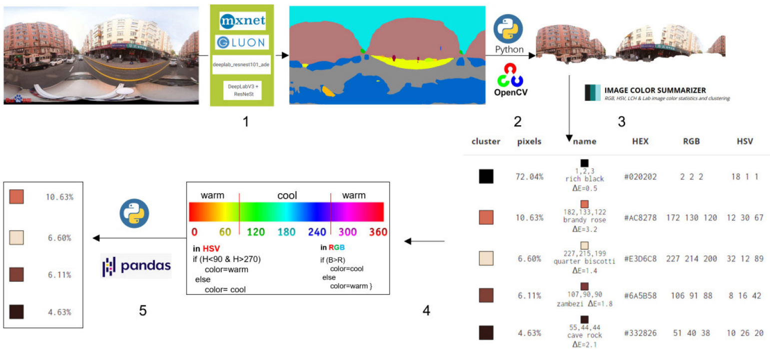
Table 3. Equation and computation for indices of climate adaptation factor.

Index	Equation	Computation Process
Winter wind environment	$\mathrm{Tow} = S_{towards} + S_{width}$ $S_{towards}$ denotes the street orientation evaluation score, S_{width} denotes the street width evaluation score.	 1. The street orientation is scored according to the angle of the street facing the prevailing winter wind direction in Harbin, reflecting the ability of the street to withstand cold winter wind crossing. 2. Street width is scored according to the street classification, reflecting the ability of the street to reduce the wind speed of cold winter winds.
Direct solar radiation	$\mathrm{Sol} = \frac{1}{n}\sum_{i}^{n} \frac{\mathrm{pdirect}}{p} h_{sunshine}$ pdirect denotes the number of pixels of the sun track in the image within the sky area, p is the total number of pixels of the sun track in the image, and h_{sunshine} denotes the number of hours of sunlight in selected day (such as winter solstice day), which can be obtained by looking up a table of hours of sunlight	 1. Semantic segmentation of the image. 2. Crop the image and convert it to a fisheye image. 3. Create a map of the sun's trajectory for selected date in winter and project it horizontally. 4. Superimpose the trajectory projection onto the fisheye image and calculate the percentage of sky area within the trajectory.

3.3.4. Visual Comfort

We chose two indices describing visual comfort (Table 4). The level of greenery is a critical element in assessing urban environments and walkability. In this study, the vegetation view index is the ratio of the area occupied by different vegetation types in the SVI, as identified by semantic segmentation, to the total area of the street interface. The facade colour index is a measure of warm colours on the façade because they provide psychological comfort to pedestrians in a cold urban area [42].

Table 4. Equation and computation for indices of visual comfort factor.

Index	Equation	Computation Process
Vegetation view	$Veg = \frac{S_{veg}}{S_{full}}$ S_{veg} denotes the vegetation area identified by semantic segmentation in the full street view image. S_{full} denotes the area of the full image.	 1. Image semantic segmentation to identify various types of vegetation. 2. The area of each type of vegetation is summed to obtain the total vegetation area. 3. Calculate the proportion of the vegetation area in the image.
Façade color	$\text{War} = \frac{S_{war}}{S_{full}}$ S_{war} denotes the warm pixels identified in the full street view image. S_{full} denotes the area of the full image.	 1. Conduct semantic segmentation of street view images; 2. Extracting building façade areas using the OpenCV library in the Python program, with the semantic segmentation result image as mask. 3. Analyse the top five main colour categories of the building facade and output RGB and HSV values; 4. Filtering according to the RGB and HSV values of the clusters and retaining the warm colour cluster categories. 5. Derive statistics on the proportion of warm colours in the panoramic image, to derive the street building warm colour index.

3.3.5. Safety

Two indices are selected to describe pedestrian safety (Table 5). The vehicle index quantifies the areas of motorised vehicles in the SVI. The sidewalk index is the proportion of the sidewalk area in the SVI.

Table 5. Equation and computation for indices of safety factor.

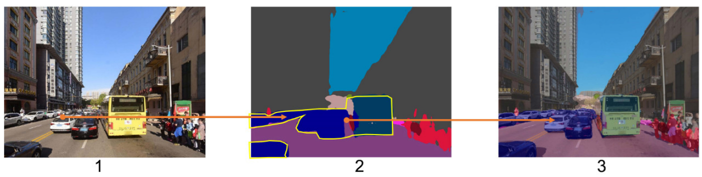
Index	Equation	Computation Process
Vehicle	$Non = \frac{s_{full}}{s_{veh}}$ s_{veh} denotes the total motor vehicle area identified by semantic segmentation, including cars, buses, trucks, and motorcycle. S_{full} denotes the area of the full image. A higher score presents that walking behaviour in the street segment is less disturbed by motor vehicles	 1. Image semantic segmentation to identify various types of motor vehicles. 2. The area of each type of motor vehicle is summed to obtain the total area occupied by motor vehicles. 3. Calculate the proportion of the area of motor vehicles in the image.

Table 5. *Cont.*

Index	Equation	Computation Process
Sidewalk	$\text{Sid} = \frac{S_{sid}}{S_{full}}$ s_{sid} denotes the total sidewalk area identified by semantic segmentation. S_{full} denotes the area of the full image. A higher score presents larger walkway area	 1. Image semantic segmentation to identify sidewalks. 2. The areas of sidewalks are summed to obtain the total sidewalk area. 3. Calculate the proportion of the sidewalk area in the image.

4. Application and Validation of the Walkability Scale

We applied the walkability scale to Harbin, China, and used open data to evaluate the walkability of the streets. The results of this evaluation were compared with the results of a manual evaluation to validate the walkability scale.

4.1. Measuring and Visualising Walkability in Harbin

4.1.1. Study Area

This study uses Harbin, a cold-region city in China, as a case study and develops a walkability measurement tool for urban built-up areas with similar urban scales and extreme climates. The geographic scope is the urban area within the third ring road of Harbin, which includes the main built-up areas of the city. The outer suburbs and villages are not included.

4.1.2. Data Collection and Processing

POI data: We used the free API interface of the AutoNavi Online Map to obtain basic data. AutoNavi Map provides 20 categories of POIs. We used Python to collect web page data within Harbin's urban administrative divisions and added it to the map. The data collection period is March 2022. The final POI data consist of 86,710 daily service facilities in eleven categories (Table 6). The POI data distribution is shown in Figure 1.

Table 6. Classification and number of POIs in the study area.

Classification of POI	Content	Quantity	Proportion
Eating and Drinking	Chinese restaurants, international restaurants, fast food restaurants, casual dining venues, cafes, tea houses, beverage shop, pastry, bakery shops, etc.	17,100	29%
Shopping	shopping malls, convenience stores, home appliance stores, supermarkets, flower, bird and fish markets, furniture markets, general markets, etc.	1331	2%
Health Care	general hospitals, specialist hospitals, clinics, emergency centres, pharmacies, etc.	5935	10%
Amenities	beauty salons, hairdressers, repair stations, laundries, post offices, logistics and courier services, telecommunication offices, etc.	7668	13%
Public Space	parks, squares	70	0.10%
Sports and Leisure	sports venues, entertainment venues, leisure venues, cinemas	2789	5%
Education and Culture	schools, museums, exhibition halls, convention centres, art galleries, libraries, science and technology centres, planetariums, etc.	5827	10%
Public Transport	bus stations, subway stations	4099	7%
Government Agencies	government agencies, social organizations office space	6281	11%
Corporations	companies, enterprises	7262	12%
Public Facilities	public toilets, accessible facilities, emergency shelters	501	1%
Total	-	58,863	100%

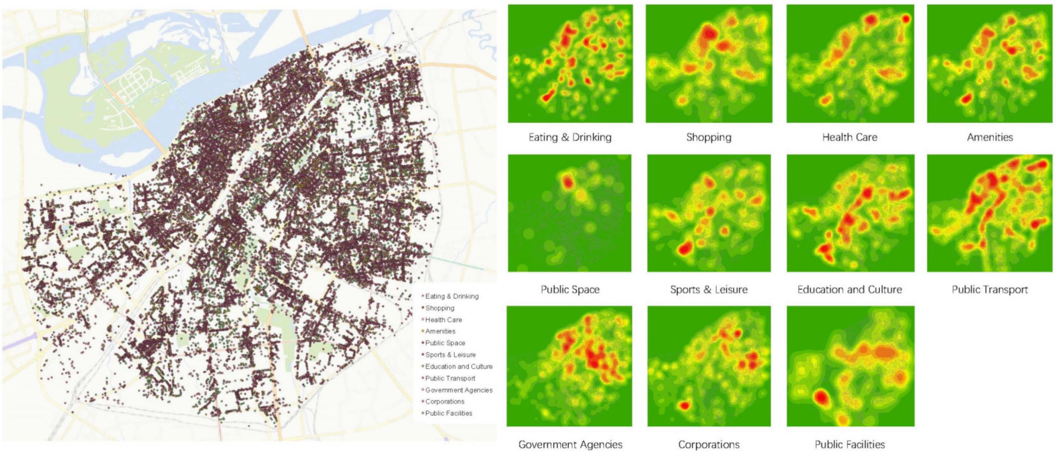


Figure 1. Distribution of POIs within the study area.

Road network: A basic road network dataset in the main urban area of Harbin was established using OSM as the data source and layers from the Harbin urban vector road network in the National Geographic Information public service platform Sky Map.

Street View Imagery: The street view images used in this study were sourced from Baidu Maps and were collected between August and October 2017. According to the Harbin City Statistical Yearbook, from 2017 to 2022, the population, the road and street area, the landscape and greenspace area, and the new built area within the study area only increased by 0.7%, 0.08%, 0.6%, and 0.2%, so although the street view imagery data were collected six years ago, they could still reflect the present street environment as there were few new developments and renovations in the study area.

The road network is divided by breakpoints every 50 metres. The breakpoints were converted to coordinates in the WGS84 coordinate reference system, and four images were acquired at each coordinate (at 0°, 90°, 180°, and 270°, respectively), thus constituting a panoramic view of the streetscape at a single breakpoint. A total of 131,044 panoramic images from 32,761 break points were obtained in the study area in Harbin (Figure 2).

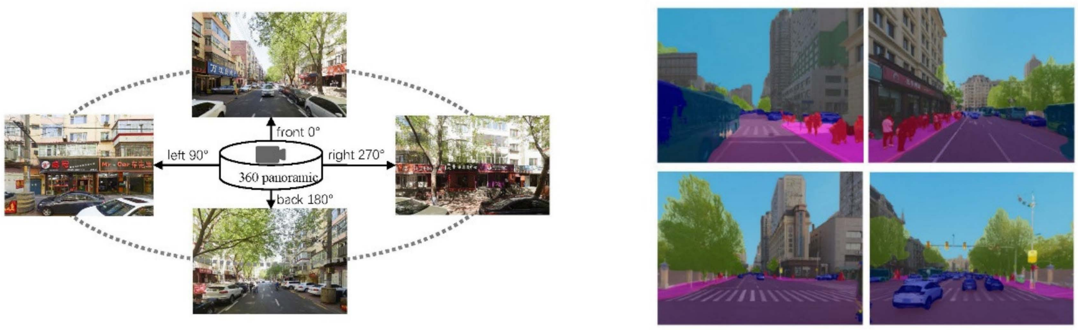


Figure 2. Example of the street view image capture method and semantic segmentation results.

Following this method, we established a breakpoint every 50 m on the road network in order to calculate the indices. The average of the scores for each breakpoint was the score for that street segment. Six of the twelve indices were analysed using semantic segmentation of the street view images. We used the Cityscapes dataset [43] and the Deeplabv3+ algorithm model trained by the MXNET deep learning framework [44] for the semantic segmentation of the images. The other six indices were calculated using the road network and POI data. The results of the indices and the product of the scores and weights for the indices were added and represented the walkability index of the street.

4.1.3. Results

As shown in Table 7, the results of the expert survey showed that the weights of the five factors ranged from 0.1 to 0.32, with Climate adaptation and Convenience having high impacts of 0.32 and 0.27, respectively, and Safety having a low impact of 0.1. The weights of the 12 indices ranged from 0.04 to 0.13, with Betweenness and Street direction having high impacts, both above 0.1, and Vegetation view, Vehicle level, and Sidewalk all having values of 0.05 or less. The rating results were tested for consistency, and the consistency of the comparison results for all factors and indices ranged from 0.65 to 0.88.

The measurement results of walkability index for Harbin City are shown in Figure 3. The results for the factors and indices are shown in Figure 4, and a typical street space with high, medium, and low scores of the walkability index and the factors are shown in Figure 5.

Table 7. Perceived walkability Scale.

Factors	Weight	Indices	Weight	Description
Accessibility	0.18	Betweenness	0.11	Route numbers in a certain radius (variety of route choice)
		Closeness	0.07	The difficulty, on average, of navigating to all possible destinations in a radius from each link
Convenience	0.27	Access to public transportation	0.09	Ease of access to public transportation
		Function diversity	0.09	Variety of infrastructures and amenities
		Function density	0.09	Number of infrastructures and amenities
Climate adaption	0.32	Snow load ability	0.10	Side space of sidewalk (usually vegetation which could pile snow and prevent melt snow hazard to sidewalk)
		Street direction	0.13	Influence of micro-climate, avoiding direct west and northern wind in winter
		Direct solar radiation	0.08	Influence of micro-climate; direct sunshine is the main source of warmth for pedestrian in winter
Visual comfort	0.13	Vegetation view	0.04	Presence of green plants
		Facade color	0.09	Warm or cool tones
Safety	0.10	Vehicle level	0.05	Motor vehicle presence and volume
		Sidewalk index	0.05	Presence and width of sidewalk

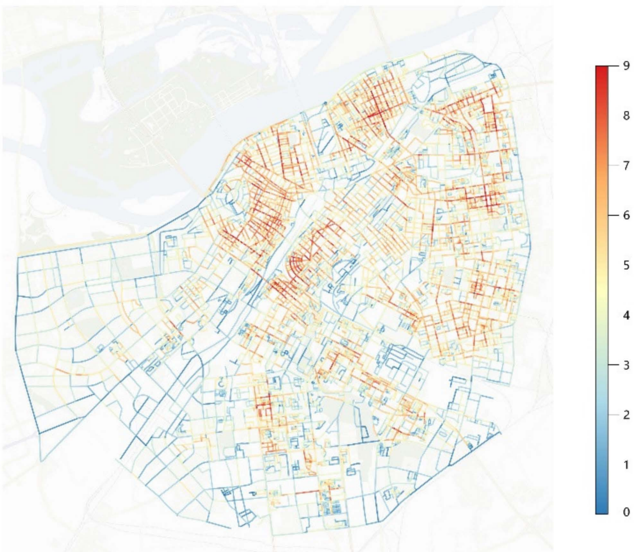


Figure 3. Visualisation of the results of total walkability of Harbin City.

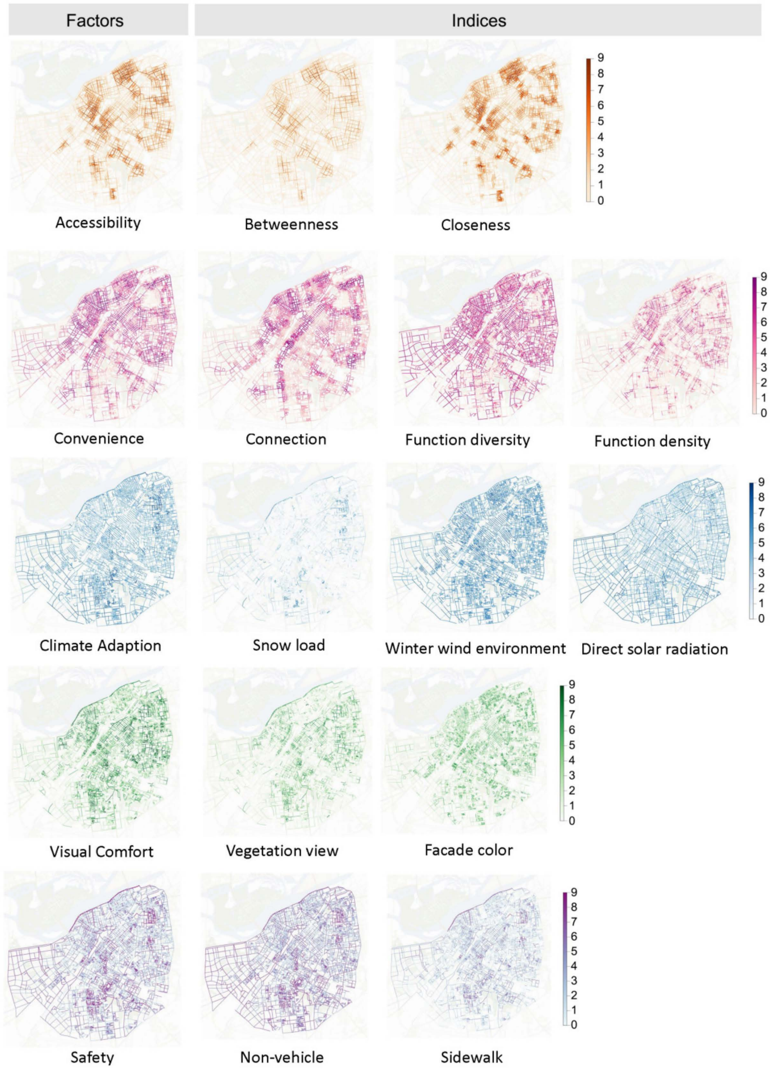


Figure 4. Visualisation of the results of factors and indices of Harbin City.

Figure 5. SVI of typical street space with high, medium, and low scores of perceived walkability and each factor.

The perceived walkability is high in the central city and low in the surrounding urban areas, and a clustered distribution is observed in the central city. Streets with high scores are mainly located in the centre of the three old urban areas and on primary and secondary streets in mixed residential and office areas. A concentration of high-score streets occurs in the southwest of the city, where several universities are located. The eastern side of the city has a more even distribution of perceived walkability, with most of the area scoring medium to high. This area has been being developed as a mixed residential and commercial/office urban area for 20 years. Areas with low perceived walkability scores are concentrated at the city's edge, along the railway line through the city centre, and on the west side of the city in newer, predominantly residential urban areas that have been developed since 2010.

Among the distributions of the scores for five factors, Accessibility, Convenience, Visual comfort, and Safety are similar to those of perceived walkability. In contrast, high scores for the Climate adaptation are observed in the southwest of the city and in the new development district to the east. The distributions of the scores for the five indices (Closeness, Function diversity, Wind environment, Facade colour, and vehicle level) are similar to those of perceived walkability. The distributions of betweenness, function density, and solar radiation are relatively high, whereas the distributions of Snow load and Sidewalk are relatively low. High scores of Connection and Vegetation view are concentrated around major arterial roads and scenic spots.

4.2. Validation of the Walkability Scale: User Feedback

To validate whether the proposed walkability assessment method could reflect human perceptions of walkability, a comparison experiment was designed between the results of open data measurements and of a personal questionnaire.

4.2.1. Validation Process

The walkability results derived from the open data were categorised into five levels from 1 to 5 (lowest to highest). An equal proportion of streets in each level was selected, with three points on each street where street view images were taken. A total of 158 streets, 416 points, and 1664 images (four images for each point) were used for the validation (Table 8). Images of the location on each sampled street and a map with the address and orientation of the street were placed on a single slide (Figure 6). Each slide was scored five times during the experiment, and the average of each score was used as the walkability of the point and the score for the factors. A total of 21 participants took part in the experiment. Each participant scored 40 breakpoints on one slide.

Table 8. Street selection.

Open Data Measurement Score Level	Total Number of Streets	Number of Streets Selected
1	1302	15
2	2872	44
3	3267	35
4	2699	39
5	1167	25
total	11,307	158

Figure 6. Example of a slide shown to participants.

4.2.2. Validation Result

The results were tallied, and 768 valid scored slides were collected, followed by an additional round of experiments to ensure that each slide had five valid scores. After averaging the scores for each point in the street on the slide, Pearson correlation analysis was performed with the results of the open data measurements. The results are listed in Table 9.

Table 9. Correlation between open data measurement and questionnaire.

	Accessibility	Convenience	Climate Adaption	Visual Comfort	Safety	Perceived Walkability
Pearson Correlation	0.393 **	0.518 **	0.026	0.509 **	0.462 **	0.459 **
Sig (2-tailed)	0.000	0.000	0.748	0.000	0.000	0.000

** $p < 0.01$.

The experimental results showed significant correlations in walkability and the factors, except for Climate adaptation, between the open data measurement results and the questionnaire results. Overall walkability and five factors (Accessibility, Convenience, Visual comfort, and Safety) had correlation coefficients from 0.393 to 0.518. The results suggest that the proposed method using open data is suitable to assess perceived walkability and the factors, with the exception of Climate adaptation. We suspected that the reason for the irrelevance of the climate adaptation factor might be related to the method used in the validation experiment, where participants were asked to rate slides rather than onsite. We then conducted a supplementary experiment in January 2024, in which 47 sample points on 15 streets were surveyed in situ for the Climate adaptation factor, and each point was rated by five investigators; the results of the ratings were significantly correlated with the results of the open data, with a correlation of 0.612.

5. Discussion

People's perceptions of walkability are highly subjective. Traditionally, these perceptions were measured using questionnaires. This paper proposed a new approach to evaluate perceived walkability. A scale for measuring perceived walkability using open data was developed. It differs from traditional methods in two ways. First, we used indices related to the three-dimensional built environment in addition to indices typically used in walkability studies utilising two-dimensional map data. We incorporated a factor describing climate adaptation, which is critical for measuring perceived walkability in cities with extreme climates, such as Harbin, where the case study was conducted. The inclusion of additional indices brought the assessment closer to the real walking experience of pedestrians. Second, all indices were measured using open data. Thus, they could be calculated on a large scale, saving labour and time compared to traditional interviews and questionnaires. We applied this method to the main urban areas of Harbin City and obtained the index, factor, and perceived walkability scores for each street.

A questionnaire was used to compare the perceived walkability results of the open data calculations with a subjective assessment of walkability to validate the proposed method. The results showed moderate and significant correlations between the two approaches for overall walkability and all five factors, and suggested that the proposed method using open data is suitable to assess perceived walkability. The moderate correlations could be attributed to the following factors: firstly, this may be related to the virtual audits methods which the validation experiment used, as although the participants of the experiment claimed that they were familiar with the street points shown in the slides, photos and brief descriptions of the street could not fully replace the onsite built street environment. Secondly, the moderate correlation may reflect differences in the perceived walkability of the build environment for different age groups. The weights of the indices used for the open data rating were based on a survey of experts with an average age of 42.8 years, whereas the average age of the volunteers who participated in the validation experiment was 24.9 years. Age might explain the differences in perceived walkability between these two groups. Thirdly, street function was not assessed in this study, although streets with different functions have different environmental characteristics. People have different perceptions of walkability for streets with different functions.

There are possible differences in perceptions of walkable environments because we did not consider people walking in different environments. The street function was not assessed in this study, although streets with different functions have different environmental characteristics. People have different perceptions of walkability for streets with different functions. In addition, the indices and weights were derived from a questionnaire administered by an expert panel and validated by another group of participants representing the general public. The low-to-moderate correlations may be due to differences in perceived walkability between the experts and the people in the validation group.

The open data measurement and questionnaire results were not correlated for the adaptation to weather conditions. The reason may be the method used in the validation experiment, because the participants were asked to look at street view images to rate walkability. Although previous studies have confirmed that virtual assessments are highly reliable and valid for evaluating the built environment characteristics [45,46], further discussion is required to determine the ability of virtual assessments to replace on-site assessments in high-density urban environments and in extreme climate urban environments.

6. Conclusions

Open data are being generated at an unprecedented rate and are increasingly used in studies of the urban built environment. We established a multidimensional perceived walkability audit scale that uses open data. The results and validation demonstrated that open data can be used to characterise perceived walkability on a city scale. The three objectives proposed in this paper have been achieved as follows:

Firstly, we have developed a scale of perceived walkability with 5 factors and 12 indices, which has been weighted and tested for consistency by an expert survey and can be used to describe perceived walkability. At the same time, the scale takes into account the generality and specificity of the built environment, so that it can be used to conduct comparative studies of perceived walkability in multiple cities to determine similarities and differences in the relationship between the built environment and perceived walkability.

Secondly, we have developed a method that combines open data from multiple data sources to simulate and represent the built environment, conducted a simulation analysis of perceived walkability, and established formulas for calculating each index in the scale using open data. This provides a new tool for investigating and understanding human behaviour and perception in urban spaces. In addition, this study demonstrates that street view images are highly suitable for fine-scale built environment observations and social sensing and can model the micro-scale built environment and describe the relationship between the environment and human behaviour, which is a promising direction for future research.

Third, we validated the perceived walkability scale, confirming that open data can be used to assess and predict perceived walkability. By comparing the results of open data and manual scoring of each index in the perceived walkability scale, it was confirmed that measures of the urban built environment can describe perceived walkability. This finding provides evidence for understanding the interaction between the urban built environment and human behaviour.

The limitations to this study are twofold. The first is that the Street View Images used in this paper are not comprehensive in terms of seasonal elements of the built environment, e.g., snow or trees with fallen leaves are not covered. Although this does not affect the calculation of the indices during the semantic segmentation, the lack of winter scenes may affect the participants' feedback for the validation experiment. Secondly, the validation experiment was conducted under the influence of the epidemic, and most of the participants were teachers and students from the same university. With a limited number of samples and a limited diversity of subjects, their feedback on the images may differ from the feedback from a sample survey conducted across the city.

The results of this study provide several opportunities for future research. First, this study has established a method for collecting and calculating elements of the built environment based on open data, which can be applied to other studies of the relationship between the built environment and human behaviour and perception. Second, this study has developed an open data-based method for assessing perceived walkability, which makes it possible to obtain the distribution of perceived walkability at a large scale, and thus to study the spatial and temporal heterogeneity of perceived walkability. Third, based on the audit scale used in this study, population-specific scales could be developed to investigate population differences in the effects of the built environment on perceived walkability.

Author Contributions: Y.Y.: conceptualization, methodology, resources, formal analysis, funding acquisition, writing—original draft, writing—review and editing. C.J.: data curation, validation, investigation, visualization, writing—original draft. S.W.: supervision, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been supported by National Natural Science Foundation of China (No.52008128) for the first author.

Data Availability Statement: The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: The authors value the contributions of all researchers, staff, and volunteers that participated in the delivery of the initiative.

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

References

1. Forsyth, A. What is a walkable place? the walkability debate in urban design. *Urban Des. Int.* **2015**, *20*, 274–292. [CrossRef]
2. Calthorpe, A. *Transit-Oriented Development Design Guidelines (Resolution No. r-280480)*; Planning Department: San Diego, CA, USA, 1992.
3. Ewing, R.; Cervero, R. Travel and the built environment: A synthesis. *Transp. Res. Rec.* **2001**, *1780*, 87–114. [CrossRef]
4. Carr, L.J.; Dunsiger, S.I.; Marcus, B.H. Walk score™ as a global estimate of neighborhood walkability. *Am. J. Prev. Med.* **2010**, *39*, 460–463. [CrossRef] [PubMed]
5. De Vos, J.; Lättman, K.; Van der Vlugt, A.L.; Welsch, J.; Otsuka, N. Determinants and effects of perceived walkability: A literature review, conceptual model and research agenda. *Transp. Rev.* **2023**, *43*, 303–324. [CrossRef]
6. Fonseca, F.; Papageorgiou, G.; Tondelli, S.; Ribeiro, P.; Conticelli, E.; Jabbari, M.; Ramos, R. Perceived walkability and respective urban determinants: Insights from Bologna and Porto. *Sustainability* **2022**, *14*, 9089. [CrossRef]
7. Koohsari, M.J.; McCormack, G.R.; Shibata, A.; Ishii, K.; Yasunaga, A.; Nakaya, T.; Oka, K. The relationship between walk score®and perceived walkability in ultrahigh density areas. *Prev. Med. Rep.* **2021**, *23*, 101393. [CrossRef] [PubMed]

8. Jensen, W.A.; Stump, T.K.; Brown, B.B.; Werner, C.M.; Smith, K.R. Walkability, complete streets, and gender: Who benefits most? *Health Place* **2017**, *48*, 80–89. [CrossRef] [PubMed]
9. Bornioli, A.; Parkhurst, G.; Morgan, P.L. Affective experiences of built environments and the promotion of urban walking. *Transp. Res. Part A Policy Pract.* **2019**, *123*, 200–215. [CrossRef]
10. Wallmann, B.; Bucksch, J.; Froboese, I. The association between physical activity and perceived environment in German adults. *Eur. J. Public Health* **2012**, *22*, 502–508. [CrossRef]
11. Blecic, I.; Canu, D.; Cecchini, A.; Congiu, T.; Fancello, G. Walkability and street intersections in rural-urban fringes: A decision aiding evaluation procedure. *Sustainability* **2017**, *9*, 883. [CrossRef]
12. Schlossberg, M.; Johnson-Shelton, D.; Evers, C.; Moreno-Black, G. Refining the grain: Using resident-based walkability audits to better understand walkable urban form. *J. Urban. Int. Res. Placemaking Urban Sustain.* **2015**, *8*, 260–278. [CrossRef] [PubMed]
13. Ewing, R.; Hajrasouliha, A.; Neckerman, K.M.; Purciel-Hill, M.; Greene, W. Streetscape features related to pedestrian activity. *J. Plan. Educ. Res.* **2016**, *36*, 5–15. [CrossRef]
14. Rosenberg, D.; Ding, D.; Sallis, J.F.; Kerr, J.; Norman, G.J.; Durant, N.; Harris, S.K.; Saelens, B.E. Neighborhood Environment Walkability Scale for Youth (NEWS-Y): Reliability and relationship with physical activity. *Prev. Med.* **2009**, *49*, 213–218. [CrossRef] [PubMed]
15. Ensari, E.; Akbay, S. Walkability and colour experience: Facade colours and pedestrian walking preferences on urban streets. In Proceedings of the AIC Interim Meeting, Lisbon, Portugal, 25–29 September 2018.
16. Li, X.; Ratti, C. Mapping the spatial distribution of shade provision of street trees in Boston using Google Street View panoramas. *Urban For. Urban Green.* **2018**, *31*, 109–119. [CrossRef]
17. Wang, R.; Lu, Y.; Zhang, J.; Liu, P.; Yao, Y.; Liu, Y. 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]
18. Abley, S.; Wade-Brown, C.; Thomas, L.; Linton, L.; Shuttleworth, K. *Guide to Undertaking Community Street Reviews*; NZ Transport Agency: Auckland, New Zealand, 2010.
19. Frank, L.D.; Schmid, T.L.; Sallis, J.F.; Chapman, J.; Saelens, B.E. Linking objectively measured physical activity with objectively measured urban form: Findings from, S.M.A.R.T.R.A.Q. *Am. J. Prev. Med.* **2005**, *28*, 117–125. [CrossRef] [PubMed]
20. Van Cauwenberg, J.; Van Holle, V.; Simons, D.; Deridder, R.; Clarys, P.; Goubert, L.; Nasar, J.; Salmon, J.; De Bourdeaudhuij, I.; Deforche, B. Environmental factors influencing older adults' walking for transportation: A study using walk-along interviews. *Int. J. Behav. Nutr. Phys. Act.* **2012**, *9*, 85. [CrossRef] [PubMed]
21. Azegami, Y.; Imanishi, M.; Fujiwara, K.; Kusaka, H. Effects of solar radiation in the streets on pedestrian route choice in a city during the summer season. *Build. Environ.* **2023**, *235*, 110250. [CrossRef]
22. Cerin, E.; Conway, T.L.; Cain, K.L.; Kerr, J.; De Bourdeaudhuij, I.; Owen, N.; Reis, R.S.; Sarmiento, O.L.; Hinckson, E.A.; Salvo, D.; et al. Sharing good NEWS across the world: Developing comparable scores across 12 countries for the Neighborhood Environment Walkability Scale (NEWS). *BMC Public Health* **2013**, *13*, 309. [CrossRef]
23. Lee, J.; Park, S. Current design guidelines' streetscape improvement for visual perception and walkability: A case study of Sejong city, Republic of Korea. *Front. Archit. Res.* **2023**, *12*, 423–443. [CrossRef]
24. Liu, D.; Jiang, Y.; Wang, R.; Lu, Y. Establishing a citywide street tree inventory with street view images and computer vision techniques. *Comput. Environ. Urban Syst.* **2023**, *100*, 101924. [CrossRef]
25. Ning, H.; Li, Z.; Wang, C.; Hodgson, M.E.; Huang, X.; Li, X. Converting street view images to land cover maps for metric mapping: A case study on sidewalk network extraction for the wheelchair users. *Comput. Environ. Urban Syst.* **2022**, *95*, 101808. [CrossRef]
26. Li, X.; Ratti, C. Mapping the spatio-temporal distribution of solar radiation within street canyons of boston using Google Street View panoramas and building height model. *Landsc. Urban Plan.* **2019**, *191*, 103387. [CrossRef]
27. Gong, F.Y.; Zeng, Z.C.; Ng, E.; Norford, L.K. Spatiotemporal patterns of street-level solar radiation estimated using Google Street View in a high-density urban environment. *Build. Environ.* **2019**, *148*, 547–566. [CrossRef]
28. Miranda, A.S.; Fan, Z.; Duarte, F.; Ratti, C. Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment. *Comput. Environ. Urban Syst.* **2021**, *86*, 101563. [CrossRef]
29. 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]
30. Dubey, A.; Naik, N.; Parikh, D.; Raskar, R.; Hidalgo, C.A. Deep learning the city: Quantifying urban perception at a global scale. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Part I 14. Springer: Berlin/Heidelberg, Germany, 2016; pp. 196–212.
31. Zhang, F.; Zhou, B.; Liu, L.; Liu, Y.; Fung, H.H.; Lin, H.; Ratti, C. Measuring human perceptions of a large-scale urban region using machine learning. *Landsc. Urban Plan.* **2018**, *180*, 148–160. [CrossRef]
32. Jacobs, J. *The Death and Life of Great American Cities*; Vintage Books: New York, NY, USA, 2012.
33. Cerin, E.; Conway, T.L.; Saelens, B.E.; Frank, L.D.; Sallis, J.F. Cross-validation of the factorial structure of the Neighborhood Environment Walkability Scale (NEWS) and its abbreviated form (NEWS-A). *Int. J. Behav. Nutr. Phys. Act.* **2009**, *6*, 1–10. [CrossRef] [PubMed]
34. Millstein, R.A.; Cain, K.L.; Sallis, J.F.; Conway, T.L.; Geremia, C.; Frank, L.D.; Chapman, J.; Van Dyck, D.; Dipzinski, L.R.; Kerr, J.; et al. Development, scoring, and reliability of the Microscale Audit of Pedestrian Streetscapes (MAPS). *BMC Public Health* **2013**, *13*, 403. [CrossRef] [PubMed]

35. Cerin, E.; Chan, K.W.; Macfarlane, D.J.; Lee, K.Y.; Lai, P.C. Objective assessment of walking environments in ultra-dense cities: Development and reliability of the environment in asia scan tool—Hong Kong version (EAST-HK). *Health Place* **2011**, *17*, 937–945. [CrossRef]
36. Griew, P.; Hillsdon, M.; Foster, C.; Coombes, E.; Jones, A.; Wilkinson, P. Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. *Int. J. Behav. Nutr. Phys. Act.* **2013**, *10*, 1–7. [CrossRef] [PubMed]
37. Troped, P.J.; Cromley, E.K.; Fragala, M.S.; Melly, S.J.; Hasbrouck, H.H.; Gortmaker, S.L.; Brownson, R.C. Development and reliability and validity testing of an audit tool for trail/path characteristics: The Path Environment Audit tool (PEAT). *J. Phys. Act. Health* **2006**, *3*, S158–S175. [CrossRef] [PubMed]
38. Dunstan, F.; Weaver, N.; Araya, R.; Bell, T.; Lannon, S.; Lewis, G.; Patterson, J.; Thomas, H.; Jones, P.; Palmer, S. An observation tool to assist with the assessment of urban residential environments. *J. Environ. Psychol.* **2005**, *25*, 293–305. [CrossRef]
39. Saaty, T.L. *What Is the Analytic Hierarchy Process?* Springer: Berlin/Heidelberg, Germany, 1988; pp. 109–121.
40. Cooper, C.H.; Chiaradia, A.J. Sdna: 3-d spatial network analysis for gis, cad, command line & python. *SoftwareX* **2020**, *12*, 100525.
41. Jin Yumeng, K.J.; Hong, J. Thermal environment of traditional residential streets in Harbin in winter. *Build. Sci.* **2016**, *32*, 34–38.
42. Xu, L.; Han, Q.; Zhibin, W.; Si, L. Urban color characteristics of cold city Harbin. In Proceedings of the 2016 International Conference on Architectural Engineering and Civil Engineering, Shanghai, China, 9–11 December 2016; Atlantis Press: Paris, France, 2016; pp. 61–64.
43. 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 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 3213–3223.
44. Mellon, J. Mxnet: A Growing Deep Learning Framework. Carnegie Mellon University, Software Engineering Institute's Insights (Blog). 2022. Available online: http://insights.sei.cmu.edu/blog/mxnet-a-growing-deep-learning-framework/ (accessed on 15 March 2023).
45. Bethlehem, J.R.; Mackenbach, J.D.; Ben-Rebah, M.; Compernolle, S.; Glonti, K.; B'ardos, H.; Rutter, H.R.; Charreire, H.; Oppert, J.M.; Brug, J.; et al. The spotlight virtual audit tool: A valid and reliable tool to assess obesogenic characteristics of the built environment. *Int. J. Health Geogr.* **2014**, *13*, 1–8. [CrossRef] [PubMed]
46. 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]