Sustainable Cities and Society 93 (2023) 104480

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

Towards Human-centric Digital Twins: Leveraging Computer Vision and Graph Models to Predict Outdoor Comfort

Pengyuan Liu [a], Tianhong Zhao [a,b], Junjie Luo [a,c], Binyu Lei [a], Mario Frei [a], Clayton Miller [d], Filip Biljecki [a,e,*]

[a] Department of Architecture, National University of Singapore, Singapore
[b] School of Architecture and Urban Planning, Shenzhen University, Shenzhen, China
[c] Department of Landscape Architecture, Tianjin University, Tianjin, China
[d] Department of The Built Environment, National University of Singapore, Singapore
[e] Department of Real Estate, National University of Singapore, Singapore

ARTICLE INFO

Keywords:
Spatial analysis
Walkability
Built environment
Graph neural network
Urban study

ABSTRACT

Conventional sidewalk studies focused on quantitative analysis of sidewalk walkability at a large scale which cannot capture the dynamic interactions between the environment and individual factors. Embracing the idea of *Tech for Social Good*, Urban Digital Twins seek AI-empowered approaches to bridge humans with digitally-mediated technologies to enhance their prediction ability. We employ GraphSAGE-LSTM, a geo-spatial artificial intelligence (GeoAI) framework on crowdsourced data and computer vision to predict human comfort on the sidewalks. Conceptualising the pedestrians and their interactions with surrounding built and unbuilt environments as human-centric dynamic graphs, our model captures such spatio-temporal variations given by the sequential movements of human walking, enabling the GraphSAGE-LSTM to be spatio-temporal-explicit. Our experiments suggest that the proposed model provides higher accuracy by more than 20% than a traditional machine learning model and two state-of-art deep learning frameworks, thus, enhancing the prediction power of Urban Digital Twin. The source code for the model is shared openly on GitHub.

1. Introduction

Cities are the systems of networks and flows (Batty, 2013) of which urban sidewalks are a crucial component (Hosseini, Miranda, Lin, & Silva, 2022; Ning, Ye, Chen, Liu, & Cao, 2022). Sidewalks function not only for transportation and everyday commute but also as a carrier for social interactions and recreational physical activities (Liu, Zhang, Jin, & Liu, 2020), i.e., walking, that promote active lifestyles (Kelly, Schoot-man, Baker, Barnidge, & Lemes, 2007). In the urban environment, increased walking activities benefit the city from various perspectives, from air pollution reduction to urban spaces safety maintenance; it is also an essential measure of the life quality of a community (Ataman & Tuncer, 2022; Bicycle, 2008; Blacklock, Rhodes, & Brown, 2007; Cot-trill, Gaglione, Gargiulo, & Zucaro, 2020; Gozalo, Morillas, González, & Moraga, 2018; Patterson & Chapman, 2004). Therefore, designing and maintaining sidewalks for pedestrians is one of the key focuses of urban planners to develop a healthier and happier city.

In recent years, human perceptions have become a useful measurement to assess urban outdoor environment (Abdollahzadeh & Biloria, 2021; Bivina & Parida, 2022; Deng et al., 2021; Florio et al., 2021; Luo, Liu, & Cao, 2022; Nazarian & Lee, 2021; Zhang et al., 2018). During the walking activity on the sidewalks, people perceive multi-sensory experiences (e.g., thermal experience, surrounding traffics) interacting with a series of urban spatial objects (e.g., buildings, trees, road conditions) that impact their state of comfort when navigating and path-finding in the urban realm (Gao et al., 2022). Previous research primarily focused on the thermal experiences (i.e., thermal comfort) of the pedestrians, which are essential to understanding the relationship between urban micro-climate and spatial urban morphology (Nice et al., 2022; Vasilikou & Nikolopoulou, 2020) and also as an indicator of the sidewalk quality (Abdollahzadeh & Biloria, 2021). However, thermal comfort measures do not capture a complete comprehension of the walking experience in the environment, particularly in the outdoor settings where the inter-play between pedestrians and spatial objects along the walking is constantly changing (Bivina & Parida, 2022). The sense of the crowdedness of the road, safety, fear or willingness to walk in the sun, the slope condition of the roads, and other facets can heavily affect human comfort when walking on the routes (Guan et al., 2022; Meng & Kang, 2016; Miranda, Fan, Duarte, & Ratti, 2021; Natapov

* Corresponding author at: Department of Architecture, National University of Singapore, Singapore.
E-mail addresses: pyliu93@nus.edu.sg (P. Liu), filip@nus.edu.sg (F. Biljecki).

https://doi.org/10.1016/j.scs.2023.104480
Received 30 December 2022; Accepted 8 February 2023
Available online 8 March 2023

& Fisher-Gewirtzman, 2016). In this paper, we expound on personal comfort beyond the single-dimensional consideration of thermal experiences but as a consequence of complex interactions between people and urban spatial entities (i.e., built and unbuilt environment) along the sidewalks, comprising the visual semantics output by computer vision-based technologies, together with physiological senses of the urban spaces. Therefore, the human comfort defined in this study is a consequence driven by a set of complex factors, including the thermal environment (e.g., solar intensity of the day), the visual ambience of the changing environment, acoustics perception, and the impact of other people's activities.

For conventional (thermal) comfort studies in the outdoor environment, data used for analysing spatial objects' impact are often collected using various professional, if not, complex, equipment and sensors, such as weather stations (Vasilikou & Nikolopoulou, 2020), infrared lasers (Yoon, Kim, & Jeong, 2022), and heart stress trackers (Peng, Bardhan, Ellard, & Steemers, 2022). As accurate as the data collected, the equipment and sensors used for those studies require domain-specific experts to set up the experiments and collect data. Such technologies and devices can be a barrier for non-expert people who want to contribute their data for mass evaluation of the sidewalk or any built environment of where they are and how they perceive. This study aims to bridge such a gap. As we are shifting into an increasingly digitally-meditated world (Ash, Kitchin, & Leszczynski, 2018) in which crowdsourced data play a vital role, data used in this study (see, Section 4) can all be collected through smart and portable devices from users, thus, setting the path for using crowdsourced data in studying outdoor human comfort and sensing the built environment.

With the rapid development of digital twins (DT), a concept that refers to the digital representation of a physical entity, person, place, system, or device, enabling (near) real-time data exchange and simulations of real-world features and processes (Charitonidou, 2022; Lei, Janssen, Stoter, & Biljecki, 2023), its city-scale applications (Urban Digital Twin, UDT) are proliferating and have been increasingly adopted into various urban-related projects, including sidewalks planning (Ahn, Ham, Kim, & Kim, 2020; Zhao et al., 2022) and outdoor comfort studies (Onan Demirel, Irshad, Ahmed, & Tumer, 2021; Zaballos, Briones, Massa, Centelles, & Caballero, 2020). Artificial Intelligence (AI)-empowered models is a key in the UDT to provide simulation and prediction, thus, aiding the decision-making process for both urban planners and residents (Charitonidou, 2022; Lei et al., 2023; Li, Yu, & Shao, 2021). Inspired by the recent advance of spatial-explicit geospatial artificial intelligence (GeoAI) in urban studies (Liu & Biljecki, 2022) which incorporates spatial locations and dependencies into the AI computation process, this paper aims to develop a spatio-temporal-explicit GeoAI model through a graph-based LSTM neural network (see Section 3) that can be integrated into a UDT. It can capture the dynamic nature of the interactions between urban spatial objects and pedestrians based on a human-centric graph conceptualisation (see Section 3.2) for predicting personal comfort. It is also important to note that as a methodology-driven study rather than behavioural research, we are not seeking to explore why the pedestrian will perceive the walking experience with certain comfort feedback nor to calibrate and validate how the feedback is produced. In this paper, by leveraging the crowdsourced data to train a GeoAI model that can enhance the prediction ability of UDT, a two-way interaction between the digitally-mediated *City Brain* (Feng, Liu, & Shi, 2018) and people can be established, thus, working towards building a human-centric digital twin Batty (2018), Charitonidou (2022), Li, Yu et al. (2021), Ye et al. (2022).

Therefore, in this paper, we:

- introduce a human-centric dynamic graph construction method that brings pedestrians as the central components in the quantitative computational framework;
- advance the studies of using user-contributed (crowdsourced) data to analyse individual human comfort;
- propose a spatio-temporal-explicit GeoAI GraphSAGE-LSTM which captures the dynamic nature of pedestrian walking and supports (near) real-time personal comfort prediction, which can be integrated into UDT. Such methodological advancement would support citizen-centric applications. The source code developed in this research is shared openly on GitHub.[1]

2. Background

2.1. Perception, comfort and crowdsourced data

Human perception and comfort have long been an important measure for studying urban spaces (Zhang et al., 2018), which not only scientifically help the researchers understand urban spaces and places (Liu & De Sabbata, 2021; Zhou et al., 2021; Zhu et al., 2020) but also practically provide guidance for urban planning (Jo & Jeon, 2020; Luo et al., 2022).

In the studies of urban sidewalks, previous research shows that personal experience of the urban continuum is perceived in a dynamic nature (Potvin, 2000). The sequential manner of pedestrians experiencing the interconnected urban spaces reinforces such a dynamic perceptual mode through the actual movement between spaces (Vasilikou & Nikolopoulou, 2020), as shown in Fig. 1. Most outdoor comfort and perception studies seek to establish the correlations between environmental conditions (e.g., micro-climate) and human comfort (mostly thermal comfort) (Abdollahzadeh & Biloria, 2021; Guan et al., 2022; Peng et al., 2022; Vasilikou & Nikolopoulou, 2020), aiming to explain and validate why and how human perceive the walking experiences with various sensory data (e.g., heart rates, near-body temperature) collected. Among the data collected, visual perceptions are crucial because what people see links directly with their sense of the environment (Abu-Ghazzeh, 1999; Chen et al., 2022; Cureau, Pigliautile, Kousis, & Pisello, 2022; Lam et al., 2020; Lee, Kim, & Park, 2022; Li, Yabuki, & Fukuda, 2022; Schroeder & Anderson, 1984), and they can capture the dynamic variation of the interactions with the constantly changing surrounding spatial objects along human walking. In our study, visual sensing is also an important component that will be incorporated in the GeoAI model introduced; detailed data collection steps and description will be introduced in Sections 3 and 4.

Recent years have witnessed a growing interest in using crowdsourced data to study urban spaces (Chen, Arribas-Bel, & Singleton, 2019; deSouza et al., 2020; Kim, Ahn, & Nam, 2019; Li, Liu, Zhang, Xue, & Li, 2021; Mak & Lam, 2021; Song & Sun, 2010). Crowdsourced data, in the context of geo-spatial studies, often refers to geo-referenced content produced by non-professionals using self-location services provided by crowdsouring platforms or mobile applications. Although research in human comfort studies largely relies on the data collected through professional equipment, the use of mobile devices and applications (e.g., Fitbit, Apple Watch) is also gaining its prevalence (Abdelrahman, Chong, & Miller, 2022; Peng et al., 2022). Our study aims to drive such study practices forward with all data collected through mobile applications or cheap and widespread available devices, thus, encouraging the general public to contribute their personal comfort to the outdoor urban environment, which can serve as suggestions for designing better sidewalk services. It is worth noting that although in our study, which will be introduced in the rest of this paper, data collected through the designed experiment were not necessarily 'crowdsourcingly' contributed by the participants, we still choose to articulate the data collected as crowdsourced data. This is because one of the primary contributions of the model developed is its potential of incorporating user-generated data from readily available devices.

[1] https://github.com/PengyuanLiu1993/GSL-sidewalk-comfort

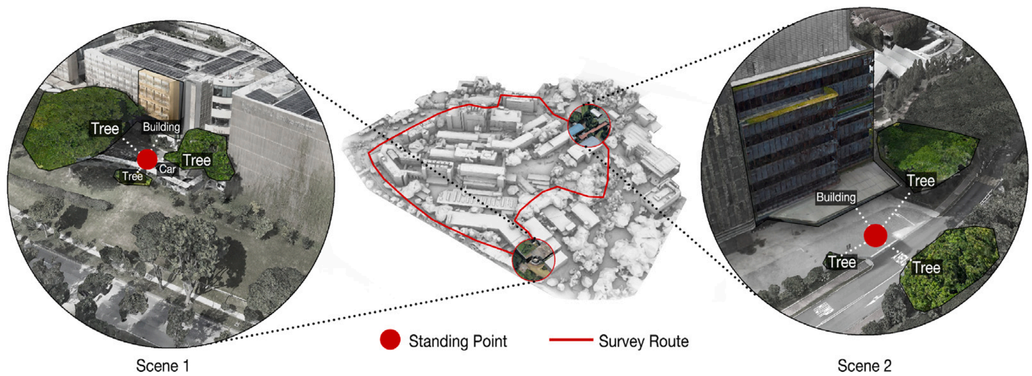

Fig. 1. Example of the dynamic surrounding environment along a sidewalk. Standing point refers to each location of a pedestrian walking. Survey route is an example of sidewalk chosen in this paper, more details will be provided in Section 4. Scene 1 and 2 showcase what are the surrounding environment when a pedestrian at the two location points.

2.2. Predicting human comfort

Although, in this paper, our research objective is to predict human comfort when walking on outdoor sidewalks, the initial idea was inspired by an indoor thermal comfort study. Abdelrahman, Chong, and Miller (2020) proposed a Build2Vec model to convert labelled property graphs transferred from a Building Information Model (BIM) into vector representations, and it captures different building objects' semantic relations and similarities. Build2Vec consists of two important components: *spatial graph constructions* and *Node2Vec* (Grover & Leskovec, 2016). The spatial graph construction step creates graphs at each timestamp of human movement in the indoor environment. Each space from a building is divided into cells of nodes using spatial discretisation. Each cell is connected to the building object's nodes (e.g., walls, windows, doors) in the graph structure when the cell is within the building objects' area of influence (AoI, e.g., fan radius). For a person in the building, the location of the cells also represents where the person is. Hence, such a conceptualisation envisions the graph structure at each timestamp where the person is human-centric. The graph varies at each timestamp when a human moves in a space transition. The original building space cells-oriented graph is converted to a human-oriented graph. Node2Vec extracts semantic similarities between different building objects, represents them in multi-dimensional vector representations (i.e., embedding) and outputs embeddings of the human movements in the space cells. Later in their study, they combined such embeddings with physiological data collected (e.g., heart rates, near-body temperatures) and fed the data into Random Forest for the thermal comfort prediction in the indoor environment (Abdelrahman et al., 2022). Their study provides a useful tool for spatial recommendations in the context of human-centric building DT applications (Kim, Schiavon, & Brager, 2018).

As elaborated in the previous section, outdoor walking faces the same challenges of dynamic variations of the surrounding environments. Therefore, we find the concept of the Build2Vec model is suitable for developing a quantitative framework for predicting human comfort in outdoor sidewalk context. However, three primary deficiencies of the model inhibit its direct employment in our study. First, the outdoor surroundings are much more complicated than the indoor environment. The indoor objects can relatively easily be extracted from the BIM models, and most objects (e.g., walls, doors, windows) are fixed in their spatial locations. In contrast, for the outdoor sidewalks, the environmental factors (e.g., rain, solar intensity), and road objects (e.g., buildings, road conditions, crowdedness of the sidewalks) are constantly changing during people's walking. Within the discipline of DT, although City Information Model (CIM) and 3D city models (Biljecki, Stoter, Ledoux, Zlatanova, & Çöltekin, 2015) are a trending topic that attracts numerous efforts to develop a model with rich semantic information about urban entities (Lei, Stouffs, & Biljecki, 2022), we can hardly find a CIM has the same level of meta information as a BIM used in Abdelrahman et al. (2022). Meanwhile, we argue that a static CIM do not fully capture the dynamics (e.g., the volumes of the traffic on the road, sound levels, weather changes) of the changing urban environment in solving real-world problems, particularly when humans as the central component to be studied. Second, Build2Vec was designed in a stack of methods instead of an end-to-end manner. That is, the model is separated into two consecutive steps that Node2Vec's output is not specifically tailored to the thermal comfort prediction task; only the step of Random Forest is a learning task that serves for the prediction. We argue that such a model may require high-quality data collected through sensors and lack the ability to take crowdsourced data as input, whose data quality is comparably lower. Thirdly, Node2Vec can only take static graph as input. Therefore, despite Abdelrahman et al. (2022) formalised the thermal comfort prediction as a time series problem, both the Node2Vec model and the Random Forest lack the ability to handle time-series data; hence, their introduced framework is *a-temporal*. Inspired by their efforts, we introduced an end-to-end deep learning framework that can capture the dynamics of the interactions between urban spatial objects and pedestrians for predicting outdoor comforts, see Section 3.

2.3. Urban digital twin and GeoAI

As a trending topic in recent urban-related studies, UDT offers urban planning with models and platforms for sustainable development that effectively combines digital technological innovations with urban operational mechanisms, and points out a viable path for urban upgrading (Lei et al., 2023, 2022). Prediction and simulation and are two of the most important characteristics of the UDT in which AI plays a vital role (Li, Yu et al., 2021). Recent advances of the AI development in urban geography stress the importance of incorporating location and spatial dependency and heterogeneity into the AI computation process, thus, formalising the initiatives of building spatial-explicitly GeoAI (Janowicz, Gao, McKenzie, Hu, & Bhaduri, 2020; Li, 2020; Liu & Biljecki, 2022). Graph representation of spatial locations and phenomenons and graph neural networks (GNN) has been recognised as a core to develop better AI frameworks that offer high-quality predictions in urban studies (Liu & Biljecki, 2022). However, most existing studies using such methods based on fixed graphs (nodes and edges do not change over time, e.g., street networks) (Zhao, Huang, Tu, He, Cao, Cao, & Li, 2022; Zhu, Liu, Yao, & Fischer, 2021), and the dynamic graphs (nodes and edges vary over time) which are suitable to capture individual level of interaction with the spatial objects are underexplored. Our study will address such a limitation by developing a GNN-based framework.

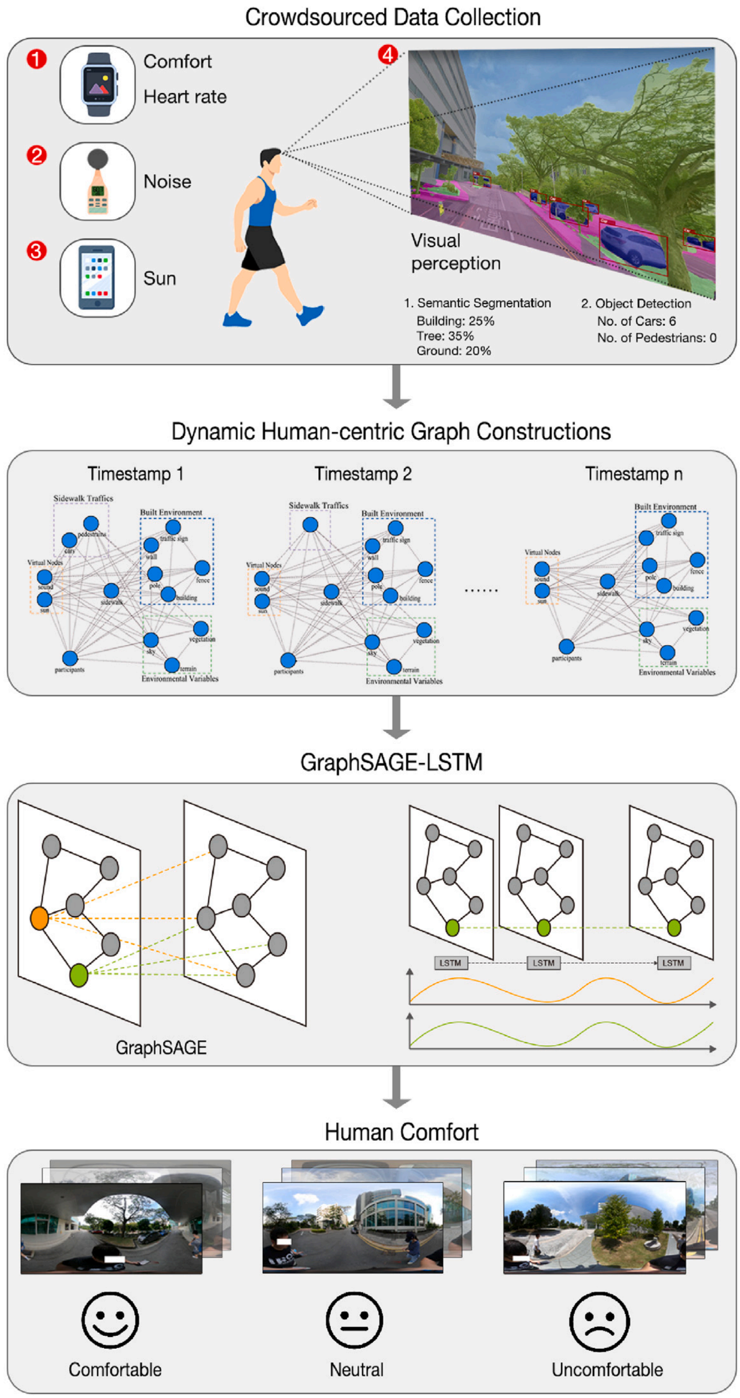

Fig. 2. The overall framework of the proposed method. The photos of the surrounding environment will be conceptualised into dynamic graphs which will be used by the GSL model for the comfort prediction.

3. Methodology

Fig. 2 illustrates our proposed approach to predict human comfort, and it consists of three components: *Crowdsourced Data Collection*, *Dynamic Spatial Graph Constructions* and *GraphSAGE-LSTM* (GSL). This section will introduce the three components in detail in their corresponding sections.

3.1. Experiment design and crowdsourced data collection

This work sets up an experiment with 15 participants (NUS students in their 20s with normal visual and walking abilities, regardless of their gender, race and cultural backgrounds) that collected dynamic status from participants about their comfort scores (at scale 1–10, 1 denotes the least comfort, and 10 is the highest comfort) walking on the selected path (see, Section 4). The 10-level comfort scale design (Syed Ahmad et al., 2022) aimed to capture in-depth correlation between human feedback and their surrounding environment, especially for people who may be sensitive to the changes along the walking. Furthermore, such a design and its correlation analysis will increase the interpretability of our proposed method (see, Sections 5.1 and 5.3). It is also important to emphasise that the human comfort defined in this paper is a consequence of complex interactions affected by the surrounding environment of pedestrians. For example, a pedestrian may prefer (i.e., give higher scores) a path that has shade but is crowded than another path in open and empty spaces, and vice versa. The comfort feedback was collected every 5 to 10 s.

As shown in Fig. 2, we used four different devices for the crowdsourced data collection. The Cozie application[2] (an open-source application for subjective feedback built on the Apple Watch platforms) was deployed to collect participants' heart rates and comfort feedback along the walking. A GoPro action camera was used to collect the panoramic photos at the locations of where the participants gave their comfort feedback. The sound metre (UT353/UT353BT mini digital sound level meters[3]) was adopted to record the level of sound surrounding the participants. A smartphone was used to collect solar intensity (through Photometer application[4]) and capture the changes of altitudes of participants' walking (using sports tracking application, Foooooot[5]). Each participant was accompanied with a research team member while walking on the sidewalk around the path. The GoPro was carried by the researcher, and all other devices were with the participant. The researcher was responsible for taking panoramic photos when the participant gave comfort feedback on Cozie, and they did not interfere or talk with the participant unless certain instructions needed (e.g., directing the path). The experiments were conducted three times a day with one participant per walk (started at 10 am, 2 pm and 5 pm) to cover as many outdoor conditions as possible, such as changes of solar intensity, crowdedness of the sidewalks because of rush hours, etc. Therefore, the experiments took place for five consecutive working days around the second half of July 2022.

For panoramic photos collected using the GoPro, as shown in Fig. 2, we conducted object detection (a computer vision technique for locating instances of objects in images) using the pre-trained fifth version of YoLo[6] (Redmon, Divvala, Girshick, & Farhadi, 2016) to count cars and pedestrians in the photo. Then, we segmented each image using DeepLabv3 (Chen, Papandreou, Kokkinos, Murphy, & Yuille, 2017) to obtain the percentages of each spatial object that occurred in each photo. The DeepLabv3 was pre-trained on the CityScape dataset (Cordts et al., 2016) using PaddleSeg (Liu et al., 2021; PaddlePaddle Authors, 2019). CityScape is a dataset designed to have a semantic understanding of urban street scenes, and it helps to classify 30 categories of urban objects in street view images (see, Cordts et al., 2016). In this paper, we only used following classes: *building*, *wall*, *fence*, *pole* (vertically oriented poles, e.g., sign pole, traffic light poles), *traffic sign*, *vegetation* (vertical vegetation, e.g., trees, hedge), *terrain* (horizontal vegetation, e.g., grass field), and *sky*.

As motivated in Sections 1 and 2, our study is intended to set the path for promoting crowdsourced data in outdoor comfort studies, thus, in designing the experiment we made sure that all devices and applications required should be readily available or cheap to purchase (as for the portable sound metre). However, compared to conventional comfort studies using professional and complicated devices (Peng et al., 2022), the quality of collected crowdsourced data is hard to measure, which brings uncertainty into the model development as well as in the output and analysis (more discussion will be provided in Sections 5 and 7). Another missing point of the data collection is micro-climate variables (e.g., wind, temperature, rain and humidity). Those micro-climate variables commonly require professional-grade sensors (e.g., weather stations) for the collection. At the time of writing this article, we have not identified any cheap and portable devices that can be easily used to collect those data; therefore, we consider such data collection beyond the scope of promoting crowdsourced data and accessible research. Further discussion will be provided in Sections 6 and 8. As a designing choice and a possible substitute, the use of solar intensity in this paper was adopted as a variable ensuring the model is robust to the changing weather (i.e., high intensity in the open spaces indicates sunny weather and vice versa).

The Institutional Review Board of our university has reviewed and approved the ethical aspects of this experiment.

3.2. Dynamic human-centric graph constructions

As mentioned in Section 2.3, graph representations and GNNs have played a vital role in developing spatial-explicit GeoAI models in urban-related studies (Liu & Biljecki, 2022). Many existing studies used fixed graphs (where nodes and edges are unchanged) to conceptualise spatial and spatio-temporal components in the models (Liu & De Sabbata, 2021; Zhao, Huang, Tu, He, Cao, Cao et al., 2022; Zhao et al., 2019; Zhu et al., 2020). However, although their nodes' values may vary subjective to the temporal differences, those fixed graphs are less intuitive in scenarios where spatial objects and environment are consistently changing (e.g., outdoor walking). In this paper, we propose a human-centred graph construction method using dynamic graphs that can better capture the dynamic nature of people's movements.

As illustrated in Fig. 2, the graph at each timestamp was constructed based on what the participant saw on the sidewalk. Fig. 3 provides a more intuitive understanding of the graph construction step. Data collected in Section 3.1 were classified into categories conceptualised as spatial objects around the participants: *Sidewalk Traffics* (number of cars and pedestrians), *Built Environment* (building, wall, fence, pole, and traffic sign), *Environmental Variables* (sky, terrain and vegetation). Those data were encoded as nodes in the graph. In addition, we created two artificial nodes (defined as *Virtual Nodes*) to take solar intensity and sound level in the graphs. Nodes' values were attributed with the data collected at each location point (see Section 4.1) when the participant walked.

Fig. 3 showcases that when a participant on the sidewalk, the traffic condition (cars and other pedestrians) are different at each timestamp. Thus, graphs vary accordingly, as well as their corresponding adjacency matrix (1 or 0 according to whether nodes are connected or not). In our study, graphs are different based on the appearance and omit of spatial objects, not only traffic conditions but also other variables, along the walking. However, it is worth to mention that although the graphs were dynamically changing, we kept the shape of the adjacency matrix same.

3.3. GraphSAGE-LSTM

The primary GeoAI framework is designed as a stack of GraphSAGE (Hamilton, Ying, & Leskovec, 2017) and Long short-term memory network (LSTM) (Hochreiter & Schmidhuber, 1997), as shown in Fig. 2. As a type of graph convolutional network, GraphSAGE is an iterative algorithm that learns embeddings (dense numerical representations of the targeted variables) for every node in a graph. Each node in the graph is represented by the aggregation of its neighbourhoods. In other words, GraphSAGE learns nodes' embeddings using the graph structures (dynamic spatio-temporal graphs constructed in Section 3.2). Meanwhile, GraphSAGE offers few options for the nodes' values aggregation (i.e., mean aggregator, pooling aggregators, and attention aggregator). In this paper, we chose the (max) pooling aggregator method because Hamilton et al. (2017) proved the pooling aggregator could achieve desired performance with reasonable time costs, which was reckoned as one of the most promising aggregation approaches in their paper. An intuitive understanding of the max pooling aggregator is that each neighbour's embedding is passed through a non-linear layer (a neural network), and an element-wise max operation is applied to their outcomes. At each location point of the corresponding timestamp, GraphSAGE produces node embeddings that are based on the graph structure and its node values, and the output embeddings will be fed

[2] https://cozie-apple.app/
[3] https://meters.uni-trend.com/product/ut353-ut353bt/
[4] https://photometer.pro/
[5] http://www.foooooot.com/
[6] https://docs.ultralytics.com/

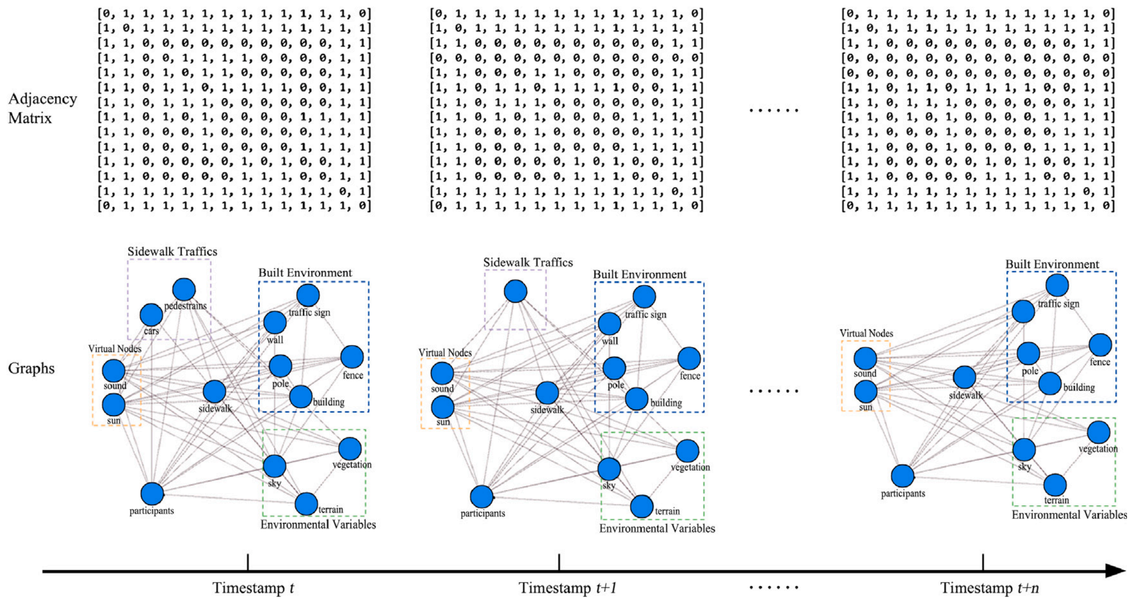

Fig. 3. Graphs vary at different location points subjective to the change of timestamps.

as input to the LSTM, formalised as a sequence classification task on human comfort.

LSTM is a type of recurrent neural network that has successfully handled sequential or time-series data in various disciplines, such as language modelling (Liu et al., 2022), traffic forecasting (Zhao et al., 2019). LSTM is used to capture the temporal dependency of personal comfort changes and the variations of the surrounding environment along the walking in this study. Following Eq. (1), each sequence of participants' walking, together with the built environment and environmental features (in the form of nodes' embeddings output by the GraphSAGE), are the input into the LSTM cell. Each cell has three designed gates (the input gate i_t, the forget gate f_t, and the output gate o_t) to regulate the information to obtain the current state of cell c_t and output hidden state h_t. σ is the sigmoid activation function, $\otimes$ is the dot product operation, and W and U are the weights of input and recurrent connections, b is the bias.

$$
\begin{aligned}
i_t &= \sigma(W_{xi}x_t + U_{hi}h_{t-1} + b_i)\\
f_t &= \sigma(W_{xf}x_t + U_{hf}h_{t-1} + b_f)\\
o_t &= \sigma(W_{xo}x_t + U_{ho}h_{t-1} + b_o)\\
c_t &= f_t \otimes c_{t-1} + i_t \otimes tanh(W_{xc}x_t + U_{hc}h_{t-1} + b_c)\\
h_t &= o_t \otimes tanh(c_t)
\end{aligned}
\tag{1}
$$

Through such a combination, GraphSAGE captures the spatial structures of each participant and the surrounding environment at each location point; LSTM addresses the dynamic nature (temporal variations) of people walking and predicts human comfort along the walking. Thus, our proposed GeoAI framework is spatio-temporal-explicit.

3.4. Data labelling and augmentation

Following Abdelrahman et al. (2022), we formalised our comfort prediction task as a ternary classification problem. That is, our proposed GeoAI model will predict human outdoor comfort on the sidewalk into three classes: *Comfortable*, *Uncomfortable* and *Neutral*. As mentioned in Section 3.1, the participants were asked to give their comfort scores at scale of 1 to 10. Thus, a data labelling step is necessary to transform numeric measures into discrete classes. However, it is worth noting that a participant's comfort feedback was a subjective feeling that differs individually. Some people are generous with the scores they give in a situation, but others may be more preservative. For example, one may score *Comfortable* as between 8 to 10 but another might only give score between 6 to 8. Therefore, it is unfair to have a set of universal categories (based on naive categorisation methods, e.g., equal intervals) but not to differentiate individual experiences. To address such an issue, following common practices in urban studies (Wei et al., 2022; Xia et al., 2022) where machine learning and deep learning involved, including those urban space perception (Ji et al., 2021; Verma, Jana, & Ramamritham, 2020), walkability (Motieyan, Kaviari, & Mostofi, 2022), and comfort-related research (Gao et al., 2022), we classified each participant's comfort using Jenks Natural Breaks method (Jenks, 1967). Jenks Natural Breaks is a data clustering method designed to optimise the arrangement of continuous values into different classes. Such a method was applied to each individual's comfort scores. As such, although the class labels of outdoor comfort are the same, those classes are tailored to each individual's subjective scoring measures. Fig. 4 shows the overall distribution of the three categories, and it demonstrates that *Neutral* is the dominating class and *Comfortable* and *Uncomfortable* share a similar amount of proportion in the labels.

In our model, intuitively, each node would only take one value (single dimension) in every graph at each location point. However, such an implementation neglects the temporal dependency of the outdoor environment during the walking activities, particularly in the scenario where participants were asked to give their comfort feedback in short temporal intervals (5 to 10 s). The impact of the surrounding environment at one location will likely be prolonged (although it may be weaker) to the next in a sequential movement of each person. To fully incorporate such a temporal dependency, we included a data enrichment step. At the model training step, each node in a graph at every location point takes values from its previous five consecutive location points, see Fig. 5. For the first five location points in the selected path, we set each node in each graph to always use the first five values. Such a step also benefits the model training by increasing the dimension of the data and their non-linearity. Note that such a design is considered as one of the most important hyperparameters in the model which required further testing in Section 5.4.

Data augmentation is another designing choice in our proposed method. It is a set of techniques to increase the diversity and the volume of the machine/deep learning training set by applying random (but realistic) transformations (Rebuffi et al., 2021; Van Dyk & Meng, 2001), and is commonly seen in various domains, such as computer vision (Shorten & Khoshgoftaar, 2019), natural language

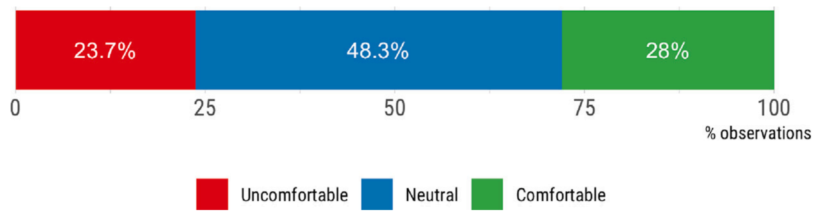

Fig. 4. The distribution of labels.

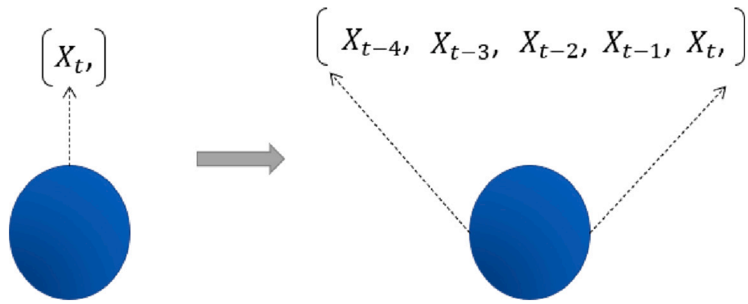

Fig. 5. Showcase of nodes' values processing.

processing (Shorten, Khoshgoftaar, & Furht, 2021), and comfort studies (Du, Sun, Zheng, Feng, & Li, 2021; Li, Dai, Cao, Liu, & Lu, 2022). As mentioned in Section 3.1, as a methodology-driven research, we did not include a large number of participants involved. To avoid the potential issue raised by insufficient training data, we adopted data augmentation strategies by using Bootstrap procedure (Efron, 1994) on the training data (70% of the whole data) to create re-sampled values and merged into the training data to increase its size, thus, improving the model performances.

3.5. Model implementations

We implemented our GSL model in Python using Deep Graph Library (Wang et al., 2019) and Pytorch (Paszke et al., 2019). We trained the model in a Google Colab[7] environment for 100 epochs (training iterations) using Adam (Kingma & Ba, 2014) (learning rate at 0.0001) as the optimisation algorithm. Meanwhile, the model adopted the loss function of cross-entropy loss. All the network parameters were randomly initialised, and the batch size was set to 128. Following Abdelrahman et al. (2020, 2022), the evaluation metrics used in the experiments is accuracy. Additionally, because the Jenks Natural Breaks created personalised labels based on the comfort scores distribution, the labels created are imbalanced as shown in Fig. 4. As such, we included F-score to be another evaluation metrics. F-score is considered a harmonic mean of the precision and recall, which is widely adopted in machine/deep learning models to evaluate model performances on imbalanced datasets (Liu & De Sabbata, 2021; Liu et al., 2022; Maxwell et al., 2017; Wang, Hu, & Joseph, 2020).

4. Study area

4.1. Selected sidewalk

As a city-state with tropical climate near the equator, Singapore has consistent temperature and humidity throughout the year. Our pilot case study and experiments were conducted around the sidewalk of the College of Design and Engineering (CDE) of the National University of Singapore (NUS). As illustrated in Fig. 6, the selected path around CDE has around 1.7 kilometres walking distance with a mixture of sidewalk conditions: *open space* (outdoor sidewalks), *semi-open space* (the sidewalk through buildings but connected to the open space, e.g., ground floor areas), *down-hill*, *up-hill*, and *flat road*. Those changing conditions of the sidewalk were recorded as the variations of solar intensity (e.g., walk from open space to the semi-open space) and altitudes. The visual sense of the surrounding objects were taken by using panoramic photos as mentioned in Section 3.1.

4.2. Location points processing

We collected 1,356 comfort feedbacks from 15 participants in the experiments. Subjective mainly to the individual's walking pace, 85 to 123 feedbacks were given by each participant. To ease the GeoAI model development (input features require the same dimension and shape, see, Section 3), we selected 85 key spatial points by applying KMeans clustering on the location points and chose clusters' centroids, as shown in Fig. 7. For participants who gave more than 85 feedback, we cut down the number of data points to match and map to those selected 85 location points. As a result, a final set of 1275 (15 participants × 85 location points) comfort feedback and corresponding data collected on the sidewalk were used in this study.

5. Results

5.1. Preliminary analysis

Fig. 8 demonstrates the data collected in the experiment described in Section 3.1, which indicates participants' comforts are influenced by the surrounding environment along the path and heart rates conditions. Fig. 9 shows a preliminary investigation of the data collected. Human comfort when walking outdoors tend to be significantly positively correlated to buildings (0.21) but negatively correlated to solar intensity (−0.41), vegetation (−0.17), terrain (−0.11), and sky (−0.15). Because of Singapore's tropical weather, with temperatures above 30 degrees most days, participants preferred shaded areas (particularly provided by the buildings) to outdoor open spaces when they walked, even if the trees can provide some shade, affirming existing findings in the literature (Jacobs et al., 2019; Siqi, Yuhong, Hien, Wu, & Xiaoli, 2022). As in our study, the temperature every day in our experiment is assumed invariant because of the tropical weather and considered the same for every participant during their walking activities. As such, the thermal

[7] https://colab.research.google.com/

Fig. 6. Study area: College of Design and Engineering (CDE) on the campus of the National University of Singapore (NUS). Map sources: Esri, HERE, Garmin, INCREMENT P, © OpenStreetMap contributors, and the GIS user community.

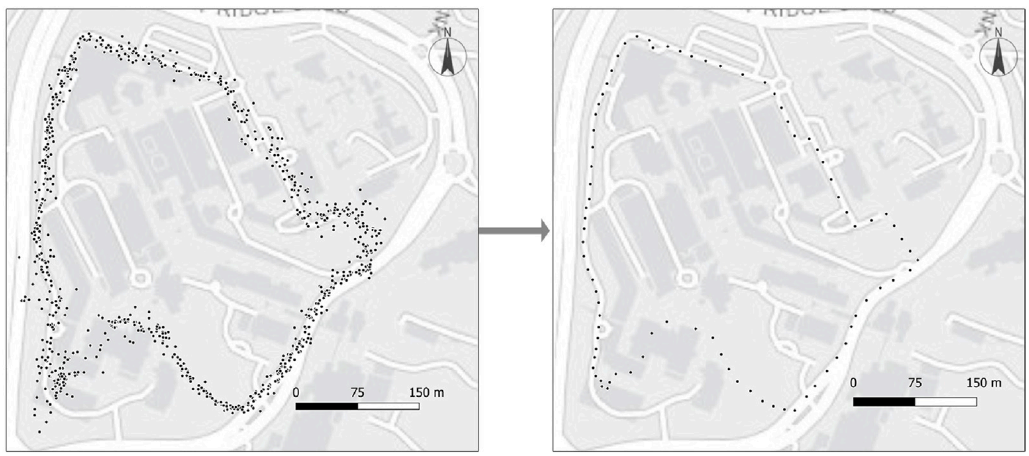

Fig. 7. Data points processing. Map sources: Esri, HERE, Garmin, INCREMENT P, © OpenStreetMap contributors, and the GIS user community.

experiences of the participants were largely decided by the solar intensity they received (i.e., exposure to the sunlight) during the walks. Therefore, Fig. 9 indicates that although, as mentioned in Section 1, we define *comfort* as a complicated consequence of inter-playing with a pedestrian's surrounding environment, thermal preferences(i.e., the unwillingness to be exposed to the sunlight) still contribute most to the human feelings (Vasilikou & Nikolopoulou, 2020). Meanwhile, Fig. 9 also shows that comfort negatively correlates with heart rate, meaning the higher the heart rate, the less comfortable the participants. Such a result is consistent with many existing comfort studies (Abdelrahman et al., 2022; Peng et al., 2022; Vasilikou & Nikolopoulou, 2020). The correlation analysis presented in this section formalises the fundamental design strategy for the ablation studies in Section 5.3, which seeks to increase the interpretability of the introduced spatio-temporal-explicit GeoAI framework.

5.2. Baseline comparisons

We set up baselines which include a traditional machine learning algorithm *Random Forest* and two neural network-based on deep learning approaches (*LSTM* and *Build2Vec*) to compare their performance with our proposed spatio-temporal-explicit GeoAI classification framework:

- Random Forest: As one of the most popular machine learning approaches, Random Forest has been intensively used in many geographic studies (including research on sidewalks) (Lee et al., 2022; Li, Kun, Shimokawa, Oyama, & Kitamura, 2013; Tribby, Miller, Brown, Werner, & Smith, 2017), and it is also the classifier used by Abdelrahman et al. (2022) to output the classification based on the embeddings produced by Build2Vec. We followed the same hyperparameters as in Abdelrahman et al. (2022) (with number of estimators as 200, max depth as 220, and max features as default) to set up the model. The Random Forest took all data collected in Section 4.1 as input and output the classification based on the labels assigned in Section 3.4. Note that such a model was naively performed on data features but not take into account the spatio-temporal information of the walking activity; thus, the random forest model is *a-spatial* and *a-temporal*.
- LSTM: the setup of this baseline method was a removal of GraphSAGE part in the model. That is, the model neglected the spatial interactions between participants and the surrounding environment but focused on the temporal dependencies of the data collected and individual's comfort. Therefore, the LSTM model used here is *a-spatial* but *temporal-explicit*.
- Node2Vec+Random Forest: as discussed in Section 2.2, Build2Vec is developed based on Node2Vec. Because in our study, we do not have a CIM that is as detailed as the BIM used in Abdelrahman et al. (2022), we cannot set up Build2Vec as a direct comparison. However, the idea of a combined method using Node2Vec and Random Forest is still a useful comparison to our proposed GSL. We employed Node2Vec at each location point to capture the spatial interactions between each pedestrian and the surrounding environment, and we extracted the embedding for the node representing the experiment's participant. Then, the output embedding was combined with data described in Section 3.1 and fed into the Random Forest, following the similar pipeline as described in Abdelrahman et al. (2022). The Node2Vec was implemented by using StellarGraph (CSIRO's Data61, 2018). As mentioned in Section 2.2, Node2Vec can only handle static graphs, such a baseline is *spatial-explicit* but *a-temporal*.

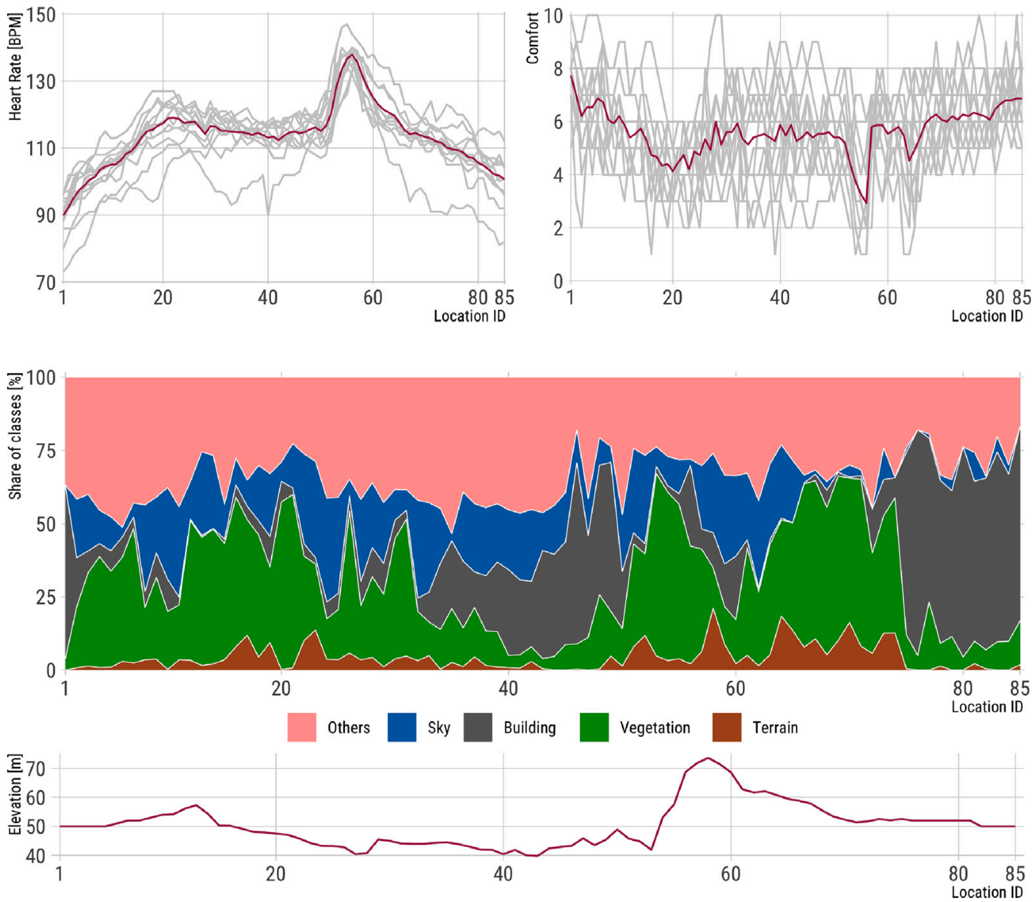

Fig. 8. Plots of data collected in this study. Top left is the plot of all participants' hearts rates along the route with the highlighted mean; top right is the plot of the collected comfort feedback with the highlighted mean; the image in the middle is the change of semantic classes of the surrounding spatial objects along the sidewalk, and the bottom image is the variations of the elevations.

As summarised in Table 1, the results reveal that GSL approach outperformed all three baselines introduced above. Our model significantly outperformed the *a-spatio-temporal* non-neural network model Random Forest by over 20% in accuracy. Also, the F-score achieved by Random Forest is significantly lower than accuracy, suggesting such a machine learning model lack the ability to process the imbalanced labels that created by the Jenks Natural Breaks (as mentioned in Section 3.4). Meanwhile, the *temporal-explicit* model LSTM achieved higher performance than Random Forest, indicating the essence of incorporating temporal dependencies of walking activities on the sidewalk into the quantitative model development; the LSTM proved to be less accurate than our model. The *spatial-explicit* Node2Vec+Random Forest approach performed much lower than our proposed GSL. As Node2Vec+Random Forest is a stacked pipeline, the parameters of the machine learning models in each part must be optimised individually. Such a result shows that an end-to-end learning framework (such as GSL) is more reasonable for outdoor comfort studies where the collected data quality is lower compared to the indoor setting (as in Abdelrahman et al. (2022)) because all the network parameters can be learnt and optimised simultaneously (LeCun, Bengio, & Hinton, 2015).

Meanwhile, the Node2Vec+Random model's performance is lower than the LSTM model but better than the Random Forest baseline. Such a comparison suggests that, although capturing the spatial interactions is important, in a sequential movement of human walking, the temporal dependencies of the walking activities play a more prominent role; thus, indicating that one's previous experiences largely influence human comfort at one location (more discussions in Section 5.4). The results demonstrate the importance of the spatio-temporal component in outdoor walking activities studies, thus, addressing the *speciality of spatial* (and temporal) within the disciplines of Spatial Science and GeoAI development (Janowicz et al., 2020; Li, 2020; Liu & Biljecki, 2022; Mai et al., 2022; Wu, Stouffs, & Biljecki, 2022).

5.3. Ablation studies

Previous experiments showed the robust capabilities of our introduced spatio-temporal-explicit GSL in the task of outdoor comfort prediction on the sidewalk. To increase the model interpretability and determine which sets of variables contribute the most to the prediction output, we designed a set of ablation studies based on the correlation analysis introduced in Section 4.2:

- GSL-visual-only: we omit data collected from smart devices and the sound metre (heart rates, solar intensity, sound and altitudes) and corresponding nodes in the dynamic graphs to examine to what extent visual perceptions contribute to the human comforts prediction. Such a test was inspired by Lee et al. (2022), which examined the relationship between visual features and sidewalk satisfaction using street view images.
- GSL-no-open-space: Section 4.2 shows how comforts are negatively correlated to variables on solar intensity (−0.41), sky (−0.15), terrain (−0.11) and vegetation (−0.17), indicating participants preferred not to walk in the open spaces where there was direct sunlight. Therefore, we design such a test excluding those data and nodes in the graphs to examine how those variables impact the model performance.

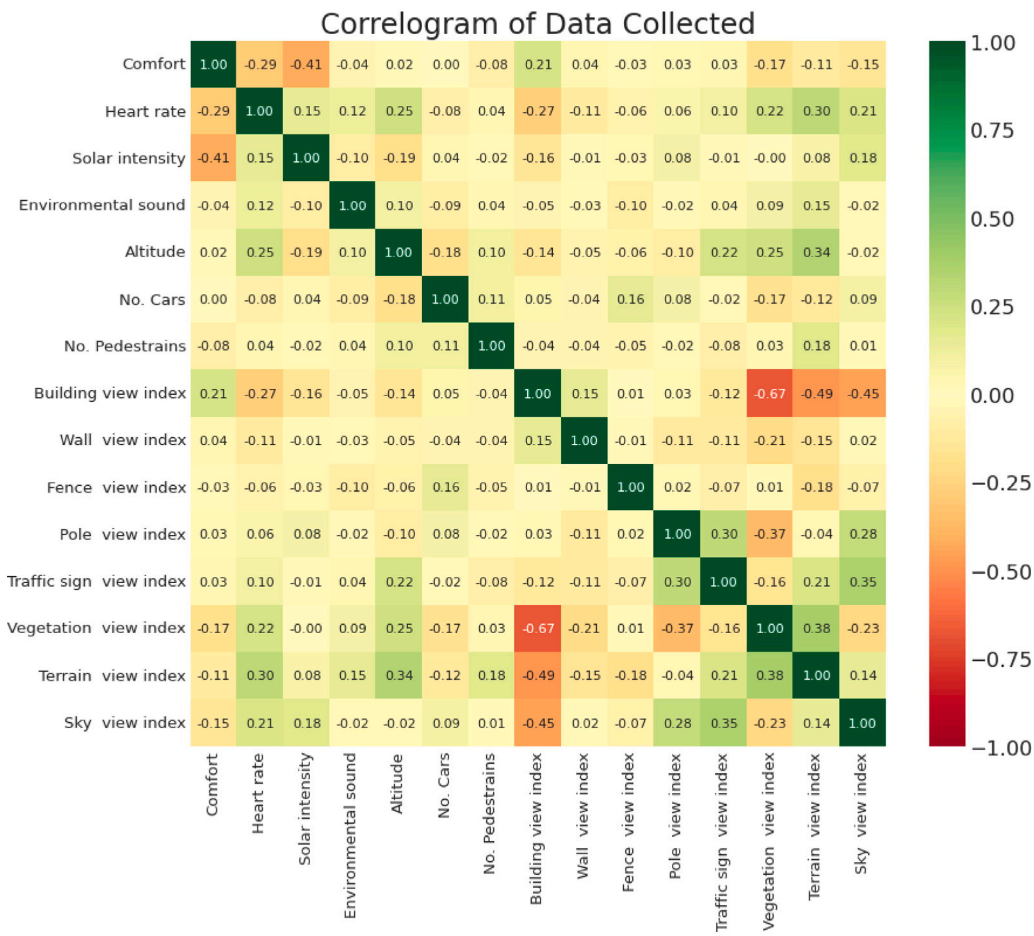

Fig. 9. Correlations of data collected. The data used in the correlation analysis were normalised between 0 and 1.

Table 1
Performances of GSL, baseline models and ablation studies. '✓' represents features that the model incorporates. Results reported as the average values by running each model 10 times.

Features	RF	LSTM	Node2Vec+RF	GSL	GSL-visual-only	GSL-no-open-space	GSL-no-build-env	GSL-no-crowd
Spatial Dependencies			✓	✓	✓	✓	✓	✓
Time Series		✓		✓	✓	✓	✓	✓
Heart Rate	✓	✓	✓	✓		✓	✓	✓
Solar Intensity	✓	✓	✓	✓			✓	✓
Sound	✓	✓	✓	✓		✓	✓	✓
Altitudes	✓	✓	✓	✓		✓	✓	✓
Cars	✓	✓	✓	✓	✓	✓	✓	
Pedestrians	✓	✓	✓	✓	✓	✓	✓	
Building	✓	✓	✓	✓	✓	✓		✓
Wall	✓	✓	✓	✓	✓	✓		✓
Fence	✓	✓	✓	✓	✓	✓		✓
Pole	✓	✓	✓	✓	✓	✓		✓
Traffic Sign	✓	✓	✓	✓	✓	✓		✓
Vegetation	✓	✓	✓	✓	✓		✓	✓
Terrain	✓	✓	✓	✓	✓		✓	✓
Sky	✓	✓	✓	✓	✓		✓	✓
Accuracy	53.52%	62.09%	58.82%	**74.53%**	64.78%	60.39%	62.73%	69.18%
F-Score	27.67%	63.21%	43.22%	**73.87%**	62.56%	58.29%	61.52%	67.58%

- GSL-no-build-env: Fig. 9 demonstrates that comforts are positively correlated with building (0.21), indicating participants were much preferred walking with buildings in sights where shades can be found. In this ablation test, we excluded all urban built environment variables (wall, fence, pole and traffic sign) in the model to examine to what extent those built environments may deteriorate the model performance. This test can be seen as a direct comparison with *GSL-no-open-space* where the latter focused more on the unbuilt environment impacts.
- GSL-no-crowd: in Section 1, one hypothesis we mentioned was the crowdedness of the sidewalks (due to other pedestrians and traffic along the road). Fig. 9 does not show a clear correlation existed in our experiments. However, we chose to include this test which excludes variables of cars and pedestrians, to investigate if they

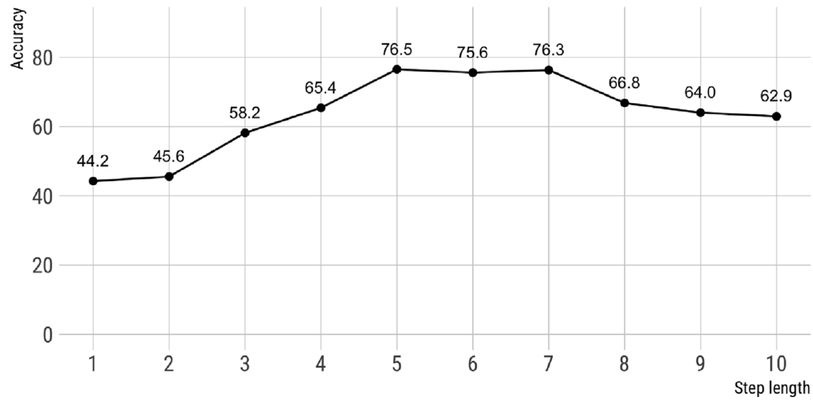

Fig. 10. Accuracy changes according to the choices of step length (best results reported).

would still be a driven factor that impacts model predictions on human comforts.

The ablation study's results are also presented in Table 1. *GSL-visual-only*, *GSL-no-open-space*, and *GSL-no-build-env* demonstrate the GSL on these three settings had similar performances among each other. Meanwhile, those results are proved to be much worse than the original model setting (see the column for GSL). Thus, it suggests all those variables are essential for the model prediction; hence, human comfort is a consequence of complex interplay with the surrounding environment. Concluded from these three tests, the variables of solar intensity, sky, terrain, and vegetation have the most impacts that worsen the model's prediction quality. Such a finding seems to echo the fact that thermal experiences contribute most to human outdoor comfort, as illustrated in Section 4.2 and existing studies (Vasilikou & Nikolopoulou, 2020). Meanwhile, cars and pedestrians have comparably less influence on the model's performance. However, as suggested by the results, the model's performance would still be more robust by including these two variables.

5.4. Hyperparameters testing

In Section 3.4, we introduced that for each location point and its corresponding graph, one node in the graph took values from its previous five consecutive location points to model the temporal impacts of the surrounding environment through a sequential movement. However, the five consecutive location points as the step length is a designing choice that, as mentioned, is a hyperparameter in the model that requires further testing. This section tests the model's robustness against the changes in step lengths from 1 to 10, and the results are shown in Fig. 10.

As suggested by Fig. 10, the settings of such a step length choice impacted the model accuracy. When the step length is one, referring to each node taking one value at the corresponding location, the model performed significantly worse than (32%) our default setting (step length is 5). The model performance increased prominently when the step length was set to 3, indicating the fact that adding temporal impact from the surrounding environment into the model is useful to improve the model's accuracy. Such a finding implies that human comfort can be impacted not only by the context in which one is standing but also is influenced by one's previous experiences. The model performance dropped after a step length set larger than 7, suggesting that longer step length introduced noises in the model, which may deteriorate the model performance. Therefore, such testing, from the model's perspective, suggests the importance of fine-tuning hyperparameters in the deep learning models (Yang & Shami, 2020); from the outdoor comfort studies perspective, implying human comfort can be more strongly impacted by short-term dependencies of the transitions in the urban continuum (Vasilikou & Nikolopoulou, 2020).

6. Showcase study

Fig. 11 shows one additional case study in which we included another participant who walked on the same sidewalk path with the same experimental settings as we described in Section 4. Having such a new experiment allows us to validate the model performance; meanwhile, it enables us to investigate the result qualitatively, particularly on those incorrectly predicted labels.

The GSL model achieved 72.94% accuracy (62 correctly predicted comforts at their corresponding location points and 23 incorrect labels) in this showcase study, which is consistent with the experiments' output presented in Section 5. As shown in Fig. 11, among the output of the incorrectly predicted results, one area where errors clustered (red circle with red marked location points) is particularly interesting. These location points and the area is around E7, a prominent building in the campus. The selected sidewalk path went through the stairs on the side of the building (semi-open space defined in Section 4.1, also shown on the top right in Fig. 11) and followed by an open space (bottom right in the figure). The errors clustered in this area can be explained through a fact of one general missing point in this article and the experiment: wind. E7 is a recently constructed building (completed in 2021) with sustainable design philosophy in mind and is a part of the university's Campus Sustainability Roadmap 2030 (National University of Singapore, 2022). The facility is designed to the highest green building standards (Architects61, 2021), which includes vertical sun-shading fins around the building facade and allows natural ventilation to flow through the building. The design of the building undoubtedly impacted human comforts along the walking; however, the model cannot capture such an influence as the wind was not incorporated in the model (discussed in Section 7). Nevertheless, the results still demonstrate the robustness and accuracy of our introduced GSL model, which illustrates its potential in sidewalk-related studies and predicts human outdoor comfort.

7. Discussion

The primary contribution of this paper is GraphSAGE-LSTM (GSL), a novel end-to-end GeoAI approach to predicting outdoor comfort on sidewalks. The GSL model takes users-contributed crowdsourced data and dynamic human-centric graphs as input, capturing the interactive nature of humans and surrounding built and unbuilt environments to predict human comforts through sequential movements. Thus, the introduced GeoAI model is spatio-temporal-explicit.

The conceptual formalisation of the human movements along the sidewalk and their interactions with urban spatial objects as dynamic graphs was established as a critical aspect in abridging the concept of human-centric AI (Lepri, Oliver, & Pentland, 2021; Nahavandi, 2019; Ruffolo, 2022), and taking into account spatio-temporal characteristics of human walking led to better predictions. The results also promote

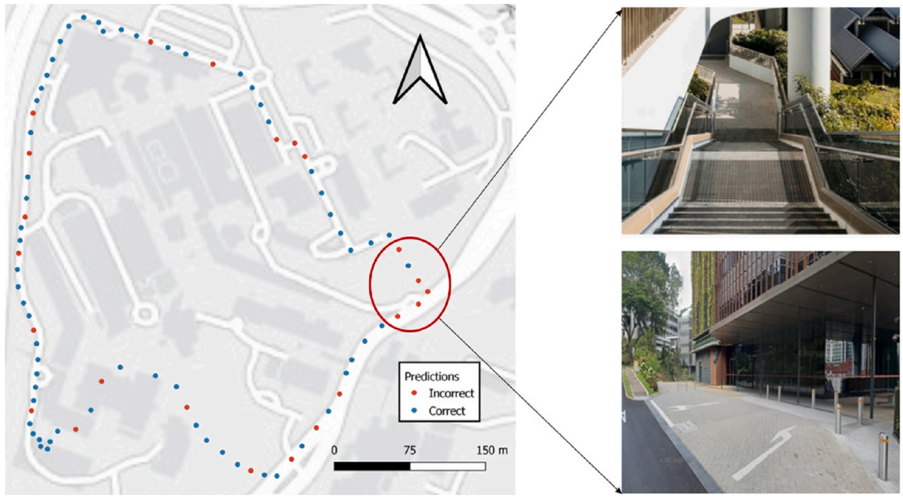

Fig. 11. Qualitative investigation of the additional experiment. Map sources: Esri, HERE, Garmin, INCREMENT P, © OpenStreetMap contributors, and the GIS user community.

the idea of human-as-sensors (Goodchild, 2007) in quantitative GeoAI-empowered urban studies (Liu & Biljecki, 2022) by encouraging the use of crowdsourced data from the general public. GSL can provide reasonable predictions even though crowdsourced data often suffer from the defect of data quality (in our study, all data were collected through non-expert mobile applications or portable and easily available devices) (Grira, Bédard, & Roche, 2010). Thus, our introduced model ease the barrier in data collection and bridge human with the digitally-mediated city brain Feng et al. (2018), particular in the context of Urban Digital Twins (UDT) (Charitonidou, 2022).

As mentioned in Section 2.3, prediction and simulation are two of the most important characteristics of the UDT in which AI plays a vital role (Li, Yu et al., 2021). Through a series of baseline comparisons, our model provides better predictions compared to conventional *a-spatio-temporal* machine learning method (Random Forest) and *a-spatial* but *temporal-explicit* deep learning-based approach (LSTM). Therefore, it has the potential to be integrated into UDT and provide better predictions of human outdoor comforts in real-life conditions.

Apart from technological innovations, our study shed light on walking activity studies. In this paper, we defined human comfort beyond a single dimension of conventional visual perception (Lee et al., 2022) or thermal comfort (Abdelrahman & Miller, 2022) but as a consequence of complex interactions with the environment that a human is in. Through the correlation analysis (Section 4.2) and the ablation studies (Section 5.3), the results indicated the fact that in Singapore, where temperatures are often above 30 degrees, people more likely to prefer not to walk in the open space where no shading areas can be found, which suggests that thermal experiences are one of the most critical factors driven to comfort experiences along the walking. Such a study suggests more covered corridors and intentional shadings from buildings' facades benefit human comforts and, therefore, might lead to increased walking activities.

8. Conclusion and outlook

This paper introduced a novel approach to predicting human outdoor comforts on the sidewalks, leveraging the use of crowdsourced data and computer vision, together with the spatio-temporal dynamics captured by the GeoAI framework, thus, breaking new ground in developing human-centric GeoAI in Urban Digital Twin. Furthermore, our experiments show that our framework can also benefit research in mobility, for example, walkability (Lo, 2009; Yap, Chang, & Biljecki, 2023) and bikeability (Ito & Biljecki, 2021), in the analysis where big data is involved, and a combination of quantitative and qualitative studies might be necessary.

We hope to continue this research in several directions in our future work. First, as indicated in Section 4.1, the experiments were conducted in the daytime. However, Singapore and many other countries have an active nightlife, which features nighttime walking, an interesting research topic. Thanks to the recent advance in low-light and nighttime image processing (Al-Ameen, 2019; Fu et al., 2022; Gu, Chen, & Zhang, 2018), a similar workflow used to extract urban spatial objects through GoPro has become feasible. Meanwhile, sample-based video analysis (Chen, Ding, & Zhang, 2018; Chen, Wu, Zhou, & Zhang, 2019; Chen, Zhou, Liu, & Zhang, 2018) can be another option to achieve real-time comfort predictions, thus, improve the framework's efficiency. Second, in our current experiment, we consider each participant as *average* person, and we did not further differentiate their other physiological factors that will enrich the data on each participant and possibly influence the model performance, such as gender, weight, and socio-cultural background. In our future work, we will incorporate those related variables to develop a more individual-tailored human-centric GeoAI model. Third, we will further integrate the model into the UDT with other datasets. Although we discussed the fact that our study aims to promote the use of crowdsourced data and deliberately not include some other environmental variables (e.g., wind) because many of those data cannot be user-contributed, the showcase study introduced in Section 6 indicates some variables might drive the model to be more accurate. Because UDT is often a well-established model with a wide range of data (e.g., wind and some other micro-climate variables) integrated and processed simultaneously, we hope to seek a practical way of integrating our model with other datasets and enable the framework to be more robust. Fourth, driven by the goal of *AI for social good* (Tomašev et al., 2020), we will use this model to study sidewalk walkabilities at an urban scale with more participants and scenarios. We hope to integrate such a model in the DT platform to achieve a near real-time comfort prediction to support downstream applications, such as human-oriented comfort-driven navigation. We also aim to extend the research objective of this research by including vulnerable groups (e.g., mobility-impaired communities) not only to include more human participants but also to increase human complexity in the model. Thus, by leveraging the prediction ability of our proposed GSL and the enhanced prediction power of UDT, we expect to build a more inclusive society to achieve the goal of *One City for All* (Beall, 1997).

In summary, this paper introduced a spatio-temporal-explicit GeoAI model to predict outdoor human comforts using crowdsourced data and dynamic graphs. We consider our GSL an valuable addition to the UDT, which, from an academic perspective, to advance research in completing UDT by incorporating humans into the model, thus evolve UDT into a human-centric digitally-mediated environment; from a technological perspective, to integrate GeoAI-empowered models in the UDT to achieve location-attentive predictions; from societal-impact perspective, to provide cutting-edge technologies and research to envision city services more tailored to its residents' needs.

CRediT authorship contribution statement

Pengyuan Liu: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation Data Curation, Writing – original draft, Writing – review & editing. **Tianhong Zhao:** Conceptualization, Methodology, Software, Writing – review & editing. **Junjie Luo:** Methodology, Investigation, Visualization, Writing – review & editing. **Binyu Lei:** Project administration, Writing – original draft, Writing – review & editing. **Mario Frei:** Project administration, Writing – review & editing. **Clayton Miller:** Resources, Writing – review & editing. **Filip Biljecki:** Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The Institutional Review Board of the National University of Singapore has reviewed and approved the ethical aspects of this research (Reference Code: NUS-IRB-2022-293). We thank the members of the NUS Urban Analytics Lab, Building and Urban Data Science Lab, and Integrated Data, Energy Analysis + Simulation (IDEAS) Lab for the discussions. We also thank the participants in the experiments. We appreciate the editor and two anonymous reviewers for the insightful comments and helpful suggestions. This research is part of the project Multi-scale Digital Twins for the Urban Environment: From Heartbeats to Cities, which is supported by the Singapore Ministry of Education Academic Research Fund Tier 1 (A-8000139-01-00).

References

Abdelrahman, M., Chong, A., & Miller, C. (2020). Build2Vec: Building representation in vector space. In *Symposium on simulation for architecture and urban design (SimAUD)*.
Abdelrahman, M. M., Chong, A., & Miller, C. (2022). Personal thermal comfort models using digital twins: Preference prediction with BIM-extracted spatial–temporal proximity data from Build2Vec. *Building and Environment*, *207*, Article 108532.
Abdelrahman, M. M., & Miller, C. (2022). Targeting occupant feedback using digital twins: Adaptive spatial–temporal thermal preference sampling to optimize personal comfort models. *Building and Environment*, *218*, Article 109090.
Abdollahzadeh, N., & Biloria, N. (2021). Outdoor thermal comfort: Analyzing the impact of urban configurations on the thermal performance of street canyons in the humid subtropical climate of sydney. *Frontiers of Architectural Research*, *10*(2), 394–409.
Abu-Ghazzeh, T. M. (1999). Communicating behavioral research to campus design: Factors affecting the perception and use of outdoor spaces at the university of Jordan. *Environment and Behavior*, *31*(6), 764–804.
Ahn, C., Ham, Y., Kim, J., & Kim, J. (2020). A digital twin city model for age-friendly communities: Capturing environmental distress from multimodal sensory data. In *Proceedings of the 53rd hawaii international conference on system sciences*.
Al-Ameen, Z. (2019). Nighttime image enhancement using a new illumination boost algorithm. *IET Image Processing*, *13*(8), 1314–1320.
Architects61 (2021). NUS E7. URL https://a61.com.sg/projects/nus-e7. (Accessed at 05.09.2022).
Ash, J., Kitchin, R., & Leszczynski, A. (2018). *Digital geographies*. Sage.
Ataman, C., & Tuncer, B. (2022). Urban interventions and participation tools in urban design processes: A systematic review and thematic analysis (1995 – 2021). *Sustainable Cities and Society*, *76*, Article 103462. http://dx.doi.org/10.1016/j.scs.2021.103462.
Batty, M. (2013). *The new science of cities*. MIT Press.
Batty, M. (2018). Digital twins. *Environment and Planning B: Urban Analytics and City Science*, *45*(5), 817–820.
Beall, J. (1997). *A city for all: Valuing difference and working with diversity*. Zed Books.
Bicycle, O. (2008). Pedestrian program. *Oregon Bicycle and Pedestrian Design Guide*.
Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., & Çöltekin, A. (2015). Applications of 3D city models: State of the art review. *ISPRS International Journal of Geo-Information*, *4*(4), 2842–2889.
Bivina, G., & Parida, M. (2022). Objective and subjective approaches towards analysis of level of service of sidewalks. In *Proceedings of the fifth international conference of transportation research group of India* (pp. 343–367). Springer.
Blacklock, R. E., Rhodes, R. E., & Brown, S. G. (2007). Relationship between regular walking, physical activity, and health-related quality of life. *Journal of Physical Activity and Health*, *4*(2), 138–152.
Charitonidou, M. (2022). Urban scale digital twins in data-driven society: Challenging digital universalism in urban planning decision-making. *International Journal of Architectural Computing*, Article 14780771211070005.
Chen, M., Arribas-Bel, D., & Singleton, A. (2019). Understanding the dynamics of urban areas of interest through volunteered geographic information. *Journal of Geographical Systems*, *21*(1), 89–109.
Chen, C., Ding, F., & Zhang, D. (2018). Perceptual hash algorithm-based adaptive GOP selection algorithm for distributed compressive video sensing. *IET Image Processing*, *12*(2), 210–217.
Chen, C., Li, H., Luo, W., Xie, J., Yao, J., Wu, L., et al. (2022). Predicting the effect of street environment on residents' mood states in large urban areas using machine learning and street view images. *Science of the Total Environment*, *816*, Article 151605.
Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 834–848.
Chen, C., Wu, Y., Zhou, C., & Zhang, D. (2019). Jsrnet: A joint sampling–reconstruction framework for distributed compressive video sensing. *Sensors*, *20*(1), 206.
Chen, C., Zhou, C., Liu, P., & Zhang, D. (2018). Iterative reweighted Tikhonov-regularized multihypothesis prediction scheme for distributed compressive video sensing. *IEEE Transactions on Circuits and Systems for Video Technology*, *30*(1), 1–10.
Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., et al. (2016). The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 3213–3223).
Cottrill, C., Gaglione, F., Gargiulo, C., & Zucaro, F. (2020). Defining the characteristics of walking paths to promote an active ageing. In *Pedestrians, urban spaces and health* (pp. 209–213). CRC Press.
CSIRO's Data61 (2018). StellarGraph machine learning library. GitHub Repository. https://github.com/stellargraph/stellargraph.
Cureau, R. J., Pigliautile, I., Kousis, I., & Pisello, A. L. (2022). Multi-domain human-oriented approach to evaluate human comfort in outdoor environments. *International Journal of Biometeorology*, 1–13.
Deng, M., Yang, W., Chen, C., Wu, Z., Liu, Y., & Xiang, C. (2021). Street-level solar radiation mapping and patterns profiling using baidu street view images. *Sustainable Cities and Society*, Article 103289. http://dx.doi.org/10.1016/j.scs.2021.103289.
deSouza, P., Anjomshoaa, A., Duarte, F., Kahn, R., Kumar, P., & Ratti, C. (2020). Air quality monitoring using mobile low-cost sensors mounted on trash-trucks: Methods development and lessons learned. *Sustainable Cities and Society*, *60*, Article 102239. http://dx.doi.org/10.1016/j.scs.2020.102239.
Du, X., Sun, C., Zheng, Y., Feng, X., & Li, N. (2021). Evaluation of vehicle vibration comfort using deep learning. *Measurement*, *173*, Article 108634.
Efron, B. (1994). Missing data, imputation, and the bootstrap. *Journal of the American Statistical Association*, *89*(426), 463–475.
Feng, L., Liu, F., & Shi, Y. (2018). City brain, a new architecture of smart city based on the internet brain. In *2018 IEEE 22nd international conference on computer supported cooperative work in design (CSCWD)* (pp. 624–629). IEEE.
Florio, P., Peronato, G., Perera, A., Blasi, A. D., Poon, K. H., & Kämpf, J. H. (2021). Designing and assessing solar energy neighborhoods from visual impact. *Sustainable Cities and Society*, *71*, Article 102959. http://dx.doi.org/10.1016/j.scs.2021.102959.
Fu, Z., Song, M., Ma, C., Nasti, J., Tyagi, V., Lloyd, G., et al. (2022). An efficient hybrid model for low-light image enhancement in mobile devices. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 3057–3066).
Gao, W., Qian, Y., Chen, H., Zhong, Z., Zhou, M., & Aminpour, F. (2022). Assessment of sidewalk walkability: Integrating objective and subjective measures of identical context-based sidewalk features. *Sustainable Cities and Society*, *87*, Article 104142.
Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, *69*(4), 211–221.
Gozalo, G. R., Morillas, J. M. B., González, D. M., & Moraga, P. A. (2018). Relationships among satisfaction, noise perception, and use of urban green spaces. *Science of the Total Environment*, *624*, 438–450.
Grira, J., Bédard, Y., & Roche, S. (2010). Spatial data uncertainty in the VGI world: Going from consumer to producer. *Geomatica*, *64*(1), 61–71.
Grover, A., & Leskovec, J. (2016). Node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 855–864).

Gu, Z., Chen, C., & Zhang, D. (2018). A low-light image enhancement method based on image degradation model and pure pixel ratio prior. *Mathematical Problems in Engineering*, *2018*.

Guan, F., Fang, Z., Wang, L., Zhang, X., Zhong, H., & Huang, H. (2022). Modelling people's perceived scene complexity of real-world environments using street-view panoramas and open geodata. *ISPRS Journal of Photogrammetry and Remote Sensing*, *186*, 315–331.

Hamilton, W., Ying, Z., & Leskovec, J. (2017). Inductive representation learning on large graphs. *Advances in Neural Information Processing Systems*, *30*.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Hosseini, M., Miranda, F., Lin, J., & Silva, C. T. (2022). CitySurfaces: City-scale semantic segmentation of sidewalk materials. *Sustainable Cities and Society*, *79*, Article 103630. http://dx.doi.org/10.1016/j.scs.2021.103630.

Ito, K., & Biljecki, F. (2021). Assessing bikeability with street view imagery and computer vision. *Transportation Research Part C (Emerging Technologies)*, *132*, Article 103371.

Jacobs, C., Singh, T., Gorti, G., Iftikhar, U., Saeed, S., Syed, A., et al. (2019). Patterns of outdoor exposure to heat in three South Asian cities. *Science of the Total Environment*, *674*, 264–278.

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2020). GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, *34*(4), 625–636.

Jenks, G. F. (1967). The data model concept in statistical mapping. *International Yearbook of Cartography*, *7*, 186–190.

Ji, H., Qing, L., Han, L., Wang, Z., Cheng, Y., & Peng, Y. (2021). A new data-enabled intelligence framework for evaluating urban space perception. *ISPRS International Journal of Geo-Information*, *10*(6), 400.

Jo, H. I., & Jeon, J. Y. (2020). The influence of human behavioral characteristics on soundscape perception in urban parks: Subjective and observational approaches. *Landscape and Urban Planning*, *203*, Article 103890.

Kelly, C. M., Schootman, M., Baker, E. A., Barnidge, E. K., & Lemes, A. (2007). The association of sidewalk walkability and physical disorder with area-level race and poverty. *Journal of Epidemiology & Community Health*, *61*(11), 978–983.

Kim, J., Ahn, C. R., & Nam, Y. (2019). The influence of built environment features on crowdsourced physiological responses of pedestrians in neighborhoods. *Computers, Environment and Urban Systems*, *75*, 161–169.

Kim, J., Schiavon, S., & Brager, G. (2018). Personal comfort models–A new paradigm in thermal comfort for occupant-centric environmental control. *Building and Environment*, *132*, 114–124.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Lam, C. K. C., Yang, H., Yang, X., Liu, J., Ou, C., Cui, S., et al. (2020). Cross-modal effects of thermal and visual conditions on outdoor thermal and visual comfort perception. *Building and Environment*, *186*, Article 107297.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Lee, J., Kim, D., & Park, J. (2022). A machine learning and computer vision study of the environmental characteristics of streetscapes that affect pedestrian satisfaction. *Sustainability*, *14*(9), 5730.

Lei, B., Janssen, P., Stoter, J., & Biljecki, F. (2023). Challenges of urban digital twins: A systematic review and a delphi expert survey. *Automation in Construction*, *147*, Article 104716. http://dx.doi.org/10.1016/j.autcon.2022.104716.

Lei, B., Stouffs, R., & Biljecki, F. (2022). Assessing and benchmarking 3D city models. *International Journal of Geographical Information Science*, http://dx.doi.org/10.1080/13658816.2022.2140808.

Lepri, B., Oliver, N., & Pentland, A. (2021). Ethical machines: The human-centric use of artificial intelligence. *IScience*, *24*(3), Article 102249.

Li, W. (2020). GeoAI: Where machine learning and big data converge in GIScience. *Journal of Spatial Information Science*, (20), 71–77.

Li, P., Dai, P., Cao, D., Liu, B., & Lu, Y. (2022). Non-intrusive comfort sensing: Detecting age and gender from infrared images for personal thermal comfort. *Building and Environment*, Article 109256.

Li, L., Kun, Y., Shimokawa, T., Oyama, I., & Kitamura, S. (2013). Investigation of factors affecting the evaluation of streetscapes in Japan and China-An evaluation of streetscape based on the random forests method. *International Journal of Affective Engineering*, *12*(1), 1–10.

Li, H., Liu, Y., Zhang, H., Xue, B., & Li, W. (2021). Urban morphology in China: dataset development and spatial pattern characterization. *Sustainable Cities and Society*, Article 102981. http://dx.doi.org/10.1016/j.scs.2021.102981.

Li, Y., Yabuki, N., & Fukuda, T. (2022). Measuring visual walkability perception using panoramic street view images, virtual reality, and deep learning. *Sustainable Cities and Society*, *86*, Article 104140.

Li, D., Yu, W., & Shao, Z. (2021). Smart city based on digital twins. *Computational Urban Science*, *1*(1), 1–11.

Liu, P., & Biljecki, F. (2022). A review of spatially-explicit GeoAI applications in urban geography. *International Journal of Applied Earth Observation and Geoinformation*, *112*, Article 102936.

Liu, Y., Chu, L., Chen, G., Wu, Z., Chen, Z., Lai, B., et al. (2021). PaddleSeg: A high-efficient development toolkit for image segmentation. arXiv:2101.06175.

Liu, P., & De Sabbata, S. (2021). A graph-based semi-supervised approach to classification learning in digital geographies. *Computers, Environment and Urban Systems*, *86*, Article 101583.

Liu, P., Koivisto, S., Hiippala, T., van der Lijn, C., Vaisanen, T., Nurmi, M., et al. (2022). Extracting locations from sport and exercise-related social media messages using a neural network-based bilingual toponym recognition model. *Journal of Spatial Information Science*, (24), 31–61.

Liu, Y., Zhang, Y., Jin, S. T., & Liu, Y. (2020). Spatial pattern of leisure activities among residents in Beijing, China: Exploring the impacts of urban environment. *Sustainable Cities and Society*, *52*, Article 101806.

Lo, R. H. (2009). Walkability: what is it? *Journal of Urbanism*, *2*(2), 145–166.

Luo, J., Liu, P., & Cao, L. (2022). Coupling a physical replica with a digital twin: A comparison of participatory decision-making methods in an urban park environment. *ISPRS International Journal of Geo-Information*, *11*(8), 452.

Mai, G., Janowicz, K., Hu, Y., Gao, S., Yan, B., Zhu, R., et al. (2022). A review of location encoding for GeoAI: methods and applications. *International Journal of Geographical Information Science*, *36*(4), 639–673.

Mak, H. W. L., & Lam, Y. F. (2021). Comparative assessments and insights of data openness of 50 smart cities in air quality aspects. *Sustainable Cities and Society*, *69*, Article 102868. http://dx.doi.org/10.1016/j.scs.2021.102868.

Maxwell, A., Li, R., Yang, B., Weng, H., Ou, A., Hong, H., et al. (2017). Deep learning architectures for multi-label classification of intelligent health risk prediction. *BMC Bioinformatics*, *18*(14), 121–131.

Meng, Q., & Kang, J. (2016). Effect of sound-related activities on human behaviours and acoustic comfort in urban open spaces. *Science of the Total Environment*, *573*, 481–493.

Miranda, A. S., Fan, Z., Duarte, F., & Ratti, C. (2021). Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment. *Computers, Environment and Urban Systems*, *86*, Article 101563.

Motieyan, H., Kaviari, F., & Mostofi, N. (2022). Quantifying walking capability: a novel aggregated index based on spatial perspective and analyses. *Papers in Regional Science*, *101*(2), 483–503.

Nahavandi, S. (2019). Industry 5.0—A human-centric solution. *Sustainability*, *11*(16), 4371.

Natapov, A., & Fisher-Gewirtzman, D. (2016). Visibility of urban activities and pedestrian routes: An experiment in a virtual environment. *Computers, Environment and Urban Systems*, *58*, 60–70.

National University of Singapore (2022). Campus infrastructure. URL https://sustainability.nus.edu.sg/campus/campus-infrastructure/. (Accessed at 05.09.2022).

Nazarian, N., & Lee, J. K. (2021). Personal assessment of urban heat exposure: a systematic review. *Environmental Research Letters*, *16*(3), Article 033005. http://dx.doi.org/10.1088/1748-9326/abd350.

Nice, K. A., Nazarian, N., Lipson, M. J., Hart, M. A., Seneviratne, S., Thompson, J., et al. (2022). Isolating the impacts of urban form and fabric from geography on urban heat and human thermal comfort. *Building and Environment*, *224*, Article 109502. http://dx.doi.org/10.1016/j.buildenv.2022.109502.

Ning, H., Ye, X., Chen, Z., Liu, T., & Cao, T. (2022). Sidewalk extraction using aerial and street view images. *Environment and Planning B: Urban Analytics and City Science*, [ISSN: 2399-8083] *49*(1), 7–22. http://dx.doi.org/10.1177/2399808321995817.

Onan Demirel, H., Irshad, L., Ahmed, S., & Tumer, I. Y. (2021). Digital twin-driven human-centered design frameworks for meeting sustainability objectives. *Journal of Computing and Information Science in Engineering*, *21*(3).

PaddlePaddle Authors (2019). PaddleSeg, end-to-end image segmentation kit based on PaddlePaddle. https://github.com/PaddlePaddle/PaddleSeg.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, R. Garnett (Eds.), *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc..

Patterson, P. K., & Chapman, N. J. (2004). Urban form and older residents' service use, walking, driving, quality of life, and neighborhood satisfaction. *American Journal of Health Promotion*, *19*(1), 45–52.

Peng, Z., Bardhan, R., Ellard, C., & Steemers, K. (2022). Urban climate walk: A stop-and-go assessment of the dynamic thermal sensation and perception in two waterfront districts in Rome, Italy. *Building and Environment*, *221*, Article 109267.

Potvin, A. (2000). Assessing the microclimate of urban transitional spaces. *Proceedings of Passive Low Energy Architecture*, 581–586.

Rebuffi, S.-A., Gowal, S., Calian, D. A., Stimberg, F., Wiles, O., & Mann, T. A. (2021). Data augmentation can improve robustness. *Advances in Neural Information Processing Systems*, *34*, 29935–29948.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 779–788).

Ruffolo, M. (2022). The role of ethical AI in fostering harmonic innovations that support a human-centric digital transformation of economy and society. In *Harmonic innovation* (pp. 139–143). Springer.

Schroeder, H. W., & Anderson, L. M. (1984). Perception of personal safety in urban recreation sites. *Journal of Leisure Research*, *16*(2), 178–194.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 1–48.

Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). Text data augmentation for deep learning. *Journal of Big Data*, *8*(1), 1–34.

Siqi, J., Yuhong, W., Hien, W. N., Wu, C., & Xiaoli, D. (2022). Influences of the thermal environment on pedestrians' thermal perception and travel behavior in hot weather. *Building and Environment*, Article 109687.

Song, W., & Sun, G. (2010). The role of mobile volunteered geographic information in urban management. In *2010 18th international conference on geoinformatics* (pp. 1–5). IEEE.

Syed Ahmad, S. S., Yung, S. M., Kausar, N., Karaca, Y., Pamucar, D., & Al Din Ide, N. (2022). Nonlinear integrated fuzzy modeling to predict dynamic occupant environment comfort for optimized sustainability. *Scientific Programming*, *2022*.

Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, B., et al. (2020). AI for social good: unlocking the opportunity for positive impact. *Nature Communications*, *11*(1), 1–6.

Tribby, C. P., Miller, H. J., Brown, B. B., Werner, C. M., & Smith, K. R. (2017). Analyzing walking route choice through built environments using random forests and discrete choice techniques. *Environment and Planning B: Urban Analytics and City Science*, *44*(6), 1145–1167.

Van Dyk, D. A., & Meng, X.-L. (2001). The art of data augmentation. *Journal of Computational and Graphical Statistics*, *10*(1), 1–50.

Vasilikou, C., & Nikolopoulou, M. (2020). Outdoor thermal comfort for pedestrians in movement: thermal walks in complex urban morphology. *International Journal of Biometeorology*, *64*(2), 277–291.

Verma, D., Jana, A., & Ramamritham, K. (2020). Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. *Building and Environment*, *186*, Article 107340.

Wang, J., Hu, Y., & Joseph, K. (2020). NeuroTPR: A neuro-net toponym recognition model for extracting locations from social media messages. *Transactions in GIS*, *24*(3), 719–735.

Wang, M., Zheng, D., Ye, Z., Gan, Q., Li, M., Song, X., et al. (2019). Deep graph library: A graph-centric, highly-performant package for graph neural networks. arXiv preprint arXiv:1909.01315.

Wei, R., Ye, C., Sui, T., Ge, Y., Li, Y., & Li, J. (2022). Combining spatial response features and machine learning classifiers for landslide susceptibility mapping. *International Journal of Applied Earth Observation and Geoinformation*, *107*, Article 102681.

Wu, A. N., Stouffs, R., & Biljecki, F. (2022). Generative adversarial networks in the built environment: A comprehensive review of the application of GANs across data types and scales. *Building and Environment*, *223*, Article 109477. http://dx.doi.org/10.1016/j.buildenv.2022.109477.

Xia, Q., Zhang, X., Hu, Y., Tian, W., Miao, W., Wu, B., et al. (2022). The superposition effects of air pollution on government health expenditure in China—spatial evidence from GeoDetector. *BMC Public Health*, *22*(1), 1–15.

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295–316.

Yap, W., Chang, J.-H., & Biljecki, F. (2023). Incorporating networks in semantic understanding of streetscapes: Contextualising active mobility decisions. *Environment and Planning B: Urban Analytics and City Science*, http://dx.doi.org/10.1177/23998083221138832.

Ye, X., Du, J., Han, Y., Newman, G., Retchless, D., Zou, L., et al. (2022). Developing human-centered urban digital twins for community infrastructure resilience: A research agenda. *Journal of Planning Literature*, Article 08854122221137861.

Yoon, H.-Y., Kim, J.-H., & Jeong, J.-W. (2022). Classification of the sidewalk condition using self-supervised transfer learning for wheelchair safety driving. *Sensors*, *22*(1), 380.

Zaballos, A., Briones, A., Massa, A., Centelles, P., & Caballero, V. (2020). A smart campus' digital twin for sustainable comfort monitoring. *Sustainability*, *12*(21), 9196.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., et al. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, *180*, 148–160.

Zhao, T., Huang, Z., Tu, W., He, B., Cao, R., Cao, J., et al. (2022). Coupling graph deep learning and spatial-temporal influence of built environment for short-term bus travel demand prediction. *Computers, Environment and Urban Systems*, *94*, Article 101776.

Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., et al. (2019). T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*, *21*(9), 3848–3858.

Zhao, D., Wu, Z. S., et al. (2022). Towards a smart, safe, and sustainable sidewalk: A quantitative analysis on how sidewalk infrastructure affect personal delivery devices. Mobility21, Carnegie Mellon University.

Zhou, H., Liu, L., Lan, M., Zhu, W., Song, G., Jing, F., et al. (2021). Using Google street view imagery to capture micro built environment characteristics in drug places, compared with street robbery. *Computers, Environment and Urban Systems*, *88*, Article 101631.

Zhu, D., Liu, Y., Yao, X., & Fischer, M. M. (2021). Spatial regression graph convolutional neural networks: A deep learning paradigm for spatial multivariate distributions. *GeoInformatica*, 1–32.

Zhu, D., Zhang, F., Wang, S., Wang, Y., Cheng, X., Huang, Z., et al. (2020). Understanding place characteristics in geographic contexts through graph convolutional neural networks. *Annals of the American Association of Geographers*, *110*(2), 408–420.