Research Paper

Street-Level Perception as Planning Intelligence

Reimagining Healthy Communities in Singapore through Multimodal Graph Learning

YingshanLIU, China

Abstract

This study takes Singapore as a case and, by building a new multimodal graph neural network analysis framework, examines the all-day accessibility of health-supporting facilities and the fairness of spatial experience in the urban environment. It further reveals the perception gap between planning data, AI algorithms and actual everyday life. The aim of the study is to explore how street-level perception can serve as a necessary intelligent medium for planning all-day healthy communities, and how multimodal graph learning can provide interpretable evidence for planning departments.

Methodologically, the study proposes the StreetCLIP-GNN framework. It selects streets within 800 metres of 52 MRT stations in Singapore and constructs 97,000 graph nodes, which are linked to health-related POIs and more than 10,000 street-view images downloaded from Mapillary. An undirected graph is built according to street connectivity, and a two-layer GraphSAGE model is used to aggregate neighbourhood information and to predict two accessibility indices, for daytime and night-time. The POIs include public hospitals, polyclinics, family clinics, community parks, fitness trails and pharmacies, and are grouped into two broad categories: medical facilities and health-supporting facilities. The division between day and night is based on typical opening hours and daily activity rhythms, and is further specified in the study through identification and classification of streetscape elements.

Node features include the number of POIs, distance to the nearest MRT station, and semantic features of street-view images extracted with CLIP. The accessibility indices are calculated using a distance-decay function: within the same time period, facilities that are open contribute more to a street node when they are closer and more numerous. Compared with a global-mean predictor and a feature-only multilayer perceptron without graph structure, the proposed model reduces RMSE on the daytime task, which indicates that it can effectively capture differences between streets. On the night-time task, it mainly increases spatial detail without sharply reducing the error, but this precisely reveals differences in accessibility between streets. The comparison of results shows that digital urbanisation needs finer spatial and temporal resolution and greater attention to service blind spots and gaps in perceived experience in the distribution of urban facilities.

Keywords

Street-level perception; Health accessibility; Graph neural networks; CLIP; 24-hour city; Singapore

1. Introduction

In the past ten years, many cities have turned to a 24-hour rhythm: logistics centers, takeaway platforms, airports, service work and nightlife economy have kept some cities awake for a long time after office hours (Seijas et al., 2024; World Economic Forum, 2025). Singapore has a population of about 6.04 million on about 735.7 square kilometers of land (Singapore Department of Statistics, 2024). These activities are concentrated in a very compact structure, with intensive housing, transportation and services coexisting, so late-night delivery workers, port and hospital shift workers, security guards, cleaners, and young people socializing in the business district are the most In the end, the unconventional time depends on this city. At the same time, the infrastructure for maintaining health and well-being still follows more traditional daytime practices. Most primary health care services are provided by government-funded general clinics and private general practitioner clinics: general clinics are usually open from about 8 a.m. to late afternoon on weekdays, only open for half a day on Saturdays, and closed on Sundays and public holidays (Ministry of Health, 2024; National University Polyclinics, 2025). Some general practitioner clinics extend to the evening, and a small number of clinics provide 24-hour service, but these late-night services are unevenly distributed throughout the island, so daily health needs are "not urgent enough" for the emergency department - to deal with chronic diseases, minor injuries, mental distress, or just need a safe place to rest and recover - it is difficult to solve outside the standard time (Ministry of Health, 2024).

This mismatch between the 24-hour urban economy and the time-limited health infrastructure has resulted in what I call a temporary service desert: the time and place where people actually exist and are potentially vulnerable, but the actual access to care is extremely limited. Relevant ideas appear in the study of health care "desert" and time-varying access. The study shows that even if there are many facilities on a static map, the accessibility will fluctuate significantly in a day (Kim and Kwon, 2022; Mao, 2025; Viana et al., 2025; Zazo-Moratalla et al., 2024). Here, I use this word to emphasize that these deserts come not only from distance, but also from opening hours and mobility. Interaction with street conditions.

Traditional urban health access research provides useful tools, but it also has obvious limitations. For a long time, distance-based, gravity-based and online travel time models have been used to assess who can reach clinics or hospitals within a certain threshold (McGrail, 2012; Jamtsho et al., 2014). They have generated important insights on a regional and urban scale, including Singapore's work on the geographic information system-based work on the service gap in comprehensive clinics and underserved groups (Wong et al., 2012; Deborah et al., 2018). But they usually assume that the demand is quite stable during the day and the population can be regarded as homogeneous users. Streets are neutral pipes, not places where experience is important.

At the same time, the "24-hour city" and the night economic agenda promote all-weather urban vitality, usually focusing on leisure and consumption in the central area (Seijas et al., 2024; World Economic Forum, 2024). Critical research on night urbanism shows that these projects often ignore workers, residents and vulnerable groups, who live at night in very different conditions (Fosuhene, 2025). Even if "24 hours" is used as a brand slogan, exclusion and risk can be incorporated into the night urban space, and health services are rarely debated. Heart.

Recently, the "Smart City" and Urban Artificial Intelligence Program promised to overcome these blind spots by using large-scale real-time data sets (from mobile phone tracks and platform logs to sensor networks), and claimed to provide constantly updated city views that can provide information for transportation planning, security strategies and sometimes medical service planning (Kitchin, 2014; Kandt and Batty, 2021). But in practice, they often reproduce another set of biases: a large number of regular patterns such as commuting traffic in the main corridor, trips to large shopping malls and activities in officially designated centers are the easiest to capture and model, while low-frequency behaviors, informal practices and activities through back streets or gaps leave The traces are weaker and more difficult to classify (Shelton et al., 2015; Leszczynski, 2016).

From the perspective of health equity, this is disturbing, because people who rely on off-peak time, irregular work or marginal space are precisely the most easily distorted or ignored; if the attention of the analysis is focused on things that are easy to calculate and predict, the needs of those who are no longer obvious in policy debates will be One step was pushed to the edge (Roubin et al., 2025; Viana et al., 2025).

All of these point to what I call the perception gap: the pattern between the perception and simulation of urban health needs by planning and artificial intelligence systems does not match the pattern with the actual street lifestyle. This gap is temporary, because data and models are usually calibrated during the day or under average conditions, and vulnerability often peaks at night, in the morning or during shift changes; it is spatial, because the main roads, MRT intersections and planning centers are well reflected in the official data set, while secondary streets, residences Access, informal shortcuts and transition spaces between Housing and Development Bureau (HDB) blocks are not well captured or generalized; this is empirical because there are few indicators available.

The rest of this article is structured as follows. Section 2 places this study in the discussion of healthy communities, time accessibility and street-level urban perception. Section 3 describes Singapore cases, multi-model data sets, and graph learning frameworks used to transform street-level perceptions into planning intelligence. Section 4 shows the results of empirical analysis. Section 5 discusses how these findings can help rethink healthy community planning in 24-hour urban conditions. Section 6 summarizes the main contributions and future direction of this study.

2. Methodology

2.1 From city to dataset: MRT catchments and street nodes

The starting point is a simple observation, that is, in Singapore, the MRT station acts as an "anchor" of daily life. If we want to know where we can get health support on foot, it is natural to analyze the tissue around these anchor points, rather than around any grid. Therefore, we have selected 52 MRT stations in mainland Singapore that have enough health-related points of interest (poi) and street scene images. This set covers the central business district and the night economic zone, mature and new group house towns, and industrial or logistics corridors on the edge of the city.

Around each station, we have defined an 800-meter walking water gathering area, which is roughly equivalent to a 10-minute walk for most adults. When the buffer zone overlaps, each street node is then assigned to the nearest site; therefore, the site collection area becomes the basic area for descriptive mapping and dividing data into training, verification and test folding. In the union of all the water collection areas, we narrowed it down to a scale suitable for walking. Using the pedestrian street network, excluding highways and non-walking lines, we regularly sample points along all pedestrian sections. Each sample is a street node: a potential location where a person may walk, wait or decide whether to continue. In the final data set, there are about 97,000 such nodes.

These street nodes are the basic spatial units of subsequent modeling. Each node carries attributes from nearby facilities, street view images and simple geometric measurements, and each node becomes a vertex in the diagram of StreetCLIP-GNN training.

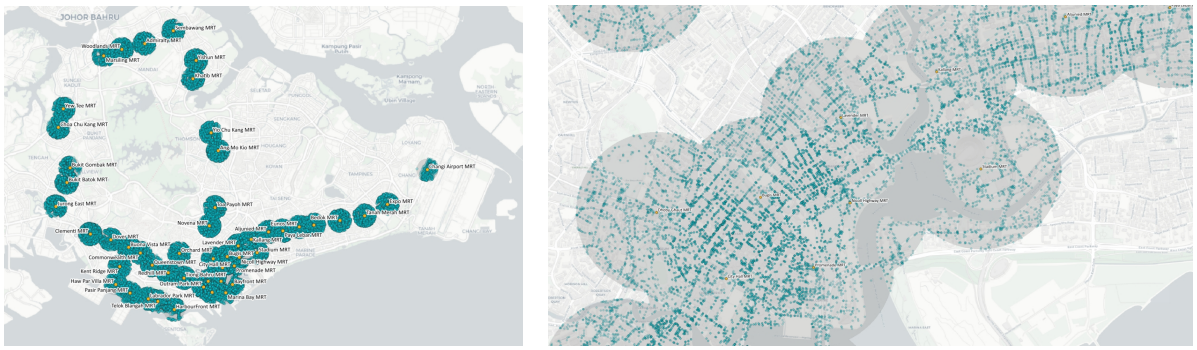


Fig. 1 Study area around 52 MRT stations in Singapore and sampled street segments within 800 m catchments.

2.2 Data layer: facilities, images, time

To understand the "all-weather healthy street", three elements are needed: what kind of health-related facilities exist, what the street environment looks like, and when different facilities are available.

Health-related action plans are based on official and public sources of data. They include, on the one hand, core health facilities such as hospitals, general clinics, 24-hour clinics and emergency departments, and on the other hand, health support facilities such as general practitioner clinics, pharmacies, sports centres and gyms, community clubs, parks and green corridors. For each POI, we keep unique identifiers, functional categories, geographical coordinates and the nearest MRT station identifier. This allows us to derive how many different types of facilities can be reached within a given distance for each street node and which parts of the pedestrian network are along.

If the facility tells us the location of the care, the street view image tells us how the care is embedded in the daily environment. We obtain images from Mapillary, a street-level image crowdsourcing platform covering the world, which has been widely used in urban mapping and computer vision missions (Mapillary, 2024). For each street node, we search for images within a fixed radius and keep a small number of recent images. Each image is processed using the open implementation of CLIP, which learns the joint embedding of images and short text prompts by predicting which title matches which image (Radford et al., 2021). Instead of embedding the original as a black box, we used a compact set of health-related text descriptions to query the model, such as "small community clinic on the ground floor", "a vibrant community park where people exercise", "well-lit commercial street at night" or "a dark, empty, no Industrial street with pedestrians". For each prompt image pair, CLIP returns a similarity score; a low-dimensional semantic vector can be obtained by averaging the available images of each node, which describes the type of street "seen" by the model. In the final feature set, each node is associated with this CLIP-based 16-dimensional semantic vector, and with a scalar perception index that summarises greener, better-lit and more pedestrian-friendly scenes into a single value between 0 and 1.

The third data layer is time. Instead of trying to model the full hours, we divided the day into two rough but policy-related parts: a day period from 10:00 to 18:00, including typical clinic opening hours, school and work routines, and most daytime errands; the evening period from 22:00 to 06:00 0, broadcast late-night work, emergency events and early morning activities. For each POI category, we specify which slots are considered to be available. Hospitals, emergency departments and 24-hour clinics contribute to both; most general practitioner clinics and general clinics offer daytime hours only; and many community and sports facilities are considered day or evening use rather than all-night support. These rules allow us to define a simple but transparent day and night accessibility index for each street node in the next step.

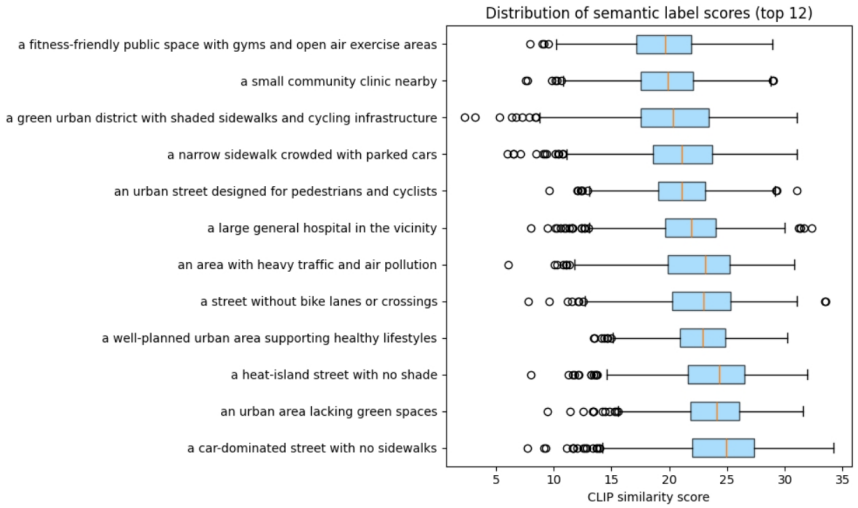


Fig. 2 Distribution of CLIP-based semantic label scores across all sampled street segments (top twelve labels).

2.3 Accessibility indices and StreetCLIP-GNN

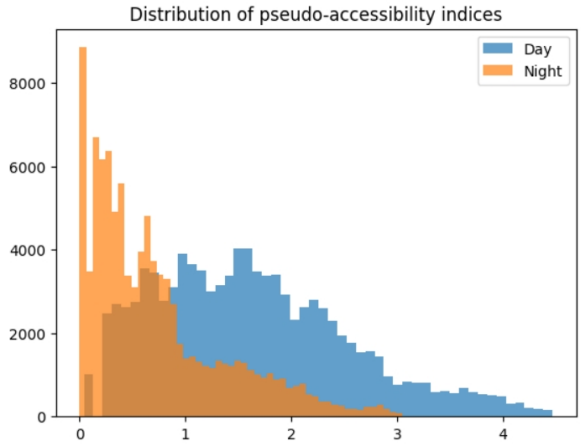
The core modelling task is to learn how day- and night-time health-supporting accessibility vary across the street network and how they relate to facilities, street semantics and spatial context. We approach this in two steps: first by defining pseudo-accessibility labels from the POI configuration, and then by training a graph neural network to reconstruct and refine these labels from multimodal node features.

2.3.1 Node-level accessibility pseudo-labels

For each street node i, we define a continuous accessibility index $A_i^{(t)}$ for each time slot $t \in \{\text{day,night}\}$. Intuitively, $A_i^{(t)}$ should be higher when there are many relevant facilities nearby that are open in slot t, and lower when such facilities are sparse or distant. Following the classic gravity-style formulations of Hansen and later accessibility work, we use an exponential distance-decay kernel (Hansen, 1959; Geurs & van Wee, 2004). A simple form is

$$A_i^{(t)} = \sum_{j \in \mathcal{F}^{(t)}} w_j^{(t)} \exp\left(-\frac{d_{ij}}{\beta}\right),$$

where $\mathcal{F}^{(t)}$ is the set of health-related POIs considered available in time slot t, d_{ij} is the network distance between street node i and facility j, β is a distance-decay parameter that controls how quickly influence fades with distance, and $w_j^{(t)}$ is a time-dependent service weight that reflects both facility type and opening hours (for example, a 24-hour clinic may carry a higher weight at night than a daytime-only GP clinic). The indices $A_i^{(\text{day})}$ and $A_i^{(\text{night})}$ thus serve as pseudo-labels: they do not come from observed patient flows, but from a transparent rule that captures the structural configuration of health support for each street node.



	access_day	access_night
access_day	1.000000	0.726905
access_night	0.726905	1.000000
a narrow sidewalk crowded with parked cars	0.235019	0.178914
a car-dominated street with no sidewalks	0.216235	0.195624
a heat-island street with no shade	0.145545	0.064855
a street without bike lanes or crossings	0.140926	0.127444
a run-down neighborhood with decaying buildings	0.075420	-0.027622
an urban street designed for pedestrians and cy...	0.073469	0.108696
an area with heavy traffic and air pollution	0.068474	0.048494
a poorly lit pedestrian walkway at night	0.015952	0.005036

Fig. 3 Distribution of pseudo-accessibility indices for day and night time slots across all street nodes.

2.3.2 Input features from facilities, geometry and semantics

For each street node i, the model uses a feature vector that combines three sources of information: health-related facilities, basic geometry and street-level semantics extracted from images.

$$\mathbf{x}_i = [e_i^{\text{POI}}, \mathbf{g}_i, \mathbf{c}_i, p_i]^{\text{T}}.$$

The feature vector x_i concatenates: (1) counts of nearby POIs by category e_i^{POI} (hospitals, polyclinics, GP clinics, pharmacies, parks, fitness trails, etc.); (2) simple geometric attributes g_i such as distance to the nearest MRT station and a local density indicator; and (3) semantic features c_i derived from CLIP similarities between street-view images and curated textual labels, reduced to a low-dimensional representation.

In addition, we construct a scalar perception index $p_i \in [0,1]$ that summarises how favourable the street environment is, balancing "supportive" labels (green, shaded, walkable streets, parks, community clinics) against "hostile" ones (heat-island streets, heavy-traffic roads, poorly lit walkways at night). In the extended model used in this paper, p_i is appended to the feature vector in Eq. (2), allowing StreetCLIP-GNN to learn how perceived street quality interacts with the spatial pattern of facilities.

2.3.3 Graph structure and neural architecture

To capture how conditions at one street segment influence its neighbours, we connect all street nodes into a single undirected graph. Two nodes are joined by an edge if they are within a specified network distance of each other or among each other's nearest neighbours, subject to a distance threshold. After removing self-loops and duplicate edges, the resulting graph has approximately 97,000 nodes and close to one million edges. This structure allows information about facilities and street semantics to propagate along the pedestrian network rather than remaining locked to isolated points.

On top of this graph, we build a GraphSAGE-style neural network with mean aggregation(Hamilton, Ying & Leskovec, 2017). Each node i starts with an initial hidden representation $\mathbf{h}_i^{(0)} = \mathbf{x}_i$. At each layer l, we update the representation by combining the node's current state with the mean of its neighbours' states:

$$\mathbf{h}_i^{(l+1)} = \sigma\left(W^{(l)}[\mathbf{h}_i^{(l)} \parallel \text{mean}_{j\in\mathcal{N}(i)}\mathbf{h}_j^{(l)}]\right),$$

where $\mathcal{N}(i)$ denotes the one-hop neighbourhood of node i, $W^{(l)}$ is a learnable weight matrix at layer l, $\sigma(\cdot)$ is a non-linear activation function such as ReLU, and $[\cdot\|\cdot]$ indicates vector concatenation. In the implementation used here, we employ two such graph layers, mapping the input features to 64-dimensional hidden vectors, with dropout applied between layers to reduce overfitting.

The final node embeddings $\mathbf{h}_i^{(L)}$ at the output of the second graph layer are then passed through a linear readout to obtain accessibility predictions. In a multi-task configuration, the readout produces a two-dimensional vector $(\widehat{A}_i^{(\text{day})}, \widehat{A}_i^{(\text{night})})$ for each node. In a single-task configuration, separate models are trained for day and night, each with a scalar output. In practice, the multi-task model captures the richer variance of day-time accessibility well, whereas a dedicated night-only model yields more stable and spatially coherent patterns for the night index, whose target distribution is more compressed.

2.4 Training, evaluation and scenario exercises

Training StreetCLIP-GNN means fitting the model parameters so that predicted accessibility values match the pseudo-labels as closely as possible on street nodes in station catchments that the model has seen, while still generalising to unseen catchments.

To assess this generalisation, we split the data by MRT station. All 52 stations are randomly ordered and assigned to three disjoint subsets: roughly 60% of stations form the training set, 20% the validation set and 20% the test set. Every street node inherits the split of its parent station catchment, so that entire station areas are either fully seen during training or fully held out. For each time slot $t \in \{\text{day,night}\}$, the model is trained to minimise a mean-squared error loss over the corresponding training nodes:

$$\mathcal{L}^{(t)} = \frac{1}{|\Omega_{\text{train}}^{(t)}|} \sum_{i\in\Omega_{\text{train}}^{(t)}} (\widehat{A}_i^{(t)} - A_i^{(t)})^2,$$

where $\Omega_{\text{train}}^{(t)}$ is the set of training nodes for time slot t, $A_i^{(t)}$ is the pseudo-accessibility index defined in Section 3.3.1, and $\widehat{A}_i^{(t)}$ is the corresponding model prediction. Optimisation is carried out with Adam, using early stopping based on validation loss. Performance is evaluated on the held-out test nodes using root mean squared error (RMSE) and mean absolute error (MAE), reported separately for the day and night tasks.

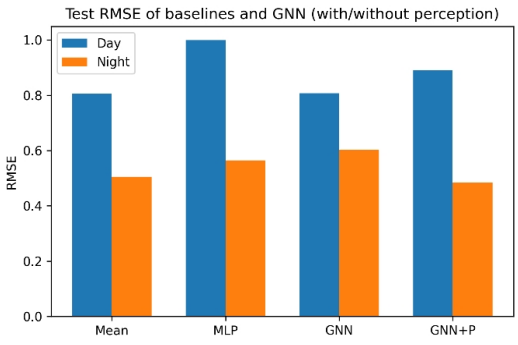
In order to understand whether the graph structure and street semantics really add value, we compared StreetCLIP-GNN with two non-spatial baselines. First of all, there is a global mean predictor, which assigns the average training set reachable value to each node; this is deliberately simplified, but it provides a useful lower bound. The second is a multi-layer sensor that only trains features, which is trained on the same node attributes as StreetCLIP-GNN, but there is no graph structure, and all nodes are regarded as independent observation values. Similar map-based comparisons have proven to be effective in recent work, which combines street-level images and gnn of urban scales (Lei et al., 2024). If GNN always matches or is better than these baselines on unseen sites, we can attribute the benefits to the combination of neighborhood clustering and multimodal features, rather than to pseudo-labeling alone.

Finally, once the models are trained, we will use them to explore simple scene exercises. The model parameters remain fixed, while the POI layer and time period availability rules are interfered with in a controllable way: for example, by adding a hypothetical 24-hour clinic near the MRT exit of the industrial corridor, or by extending the business hours of the existing GP clinic cluster in residential towns to make them useful for the night time. For each scenario, we recalculate the pseudo-reachability index, apply the

trained model, and map the predicted changes in accessibility on the street network. These scenarios will not be presented as optimization results; they are organized thought experiments that reveal where small-scale, targeted interventions will make the weak health pillars of 24-hour cities thicker, and where these interventions will hardly change the status quo.

3. Results and discussion

The first question is a modest but necessary question: Can StreetCLIP-GNN really learn more than a simple average? For the daytime task, the answer is obviously yes. The global average baseline captures some pseudo-label structures, but still produces large errors; multi-layer sensors with only characteristics regard each node as independent, which will over-fit the training data and perform poorly in idle sites. In contrast, the graph model combines editing-based semantics, POI indicators and neighborhood aggregation, which greatly reduces the RMSE and MAE in the test set. In other words, once the model is allowed to "listen" to the appearance of neighboring streets and the distribution of nearby facilities, it can reconstruct the area that can be reached during the day with a precision that a simple non-spatial model cannot achieve. This is encouraging because it shows that the model not only remembers the average of the station level, but also learns a meaningful street scale structure, which is consistent with recent work, showing the advantages of graph neural networks in spatial prediction tasks (Hamilton, Ying & Leskov Ec, 2017; Kipf & Welling, 2017; Lei et al., 2024).



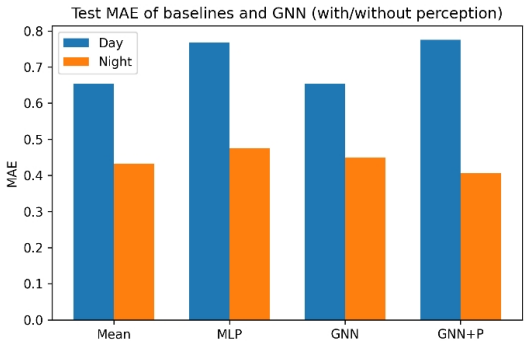


Fig. 4 Test RMSE of baselines and StreetCLIP-GNN variants for day and night accessibility prediction.

Fig. 5 Test MAE of baselines and StreetCLIP-GNN variants for day and night accessibility prediction.

The picture at night is more subtle. The global average baseline here is already very high, because the pseudo-tag itself is compressed: most networks have poor services, and only a few corridors have higher scores. In this case, the difference that no model can explain is relatively small. The performance of the first pure night GNN without a perception index is roughly equivalent to the average baseline, which confirms that there is little room for improvement in the label distribution. When the perception index is explicitly added as an input feature and a dedicated night model is trained, the test error is slightly lower than the average and the characteristic-only baseline, and the model produces a more differentiated space field. GNN no longer extensively cleans up "also low" values, but begins to choose main streets with relatively good traffic conditions at night, and instead chooses blocks with particularly lack of night facilities. From a numerical point of view, this is a moderate increase; in essence, this is important, because planners are often less interested in minor improvements in global errors, but more interested in identifying the shortcomings of the system in practical problems, especially in terms of health gaps in the city (Liu et al ., 2024).

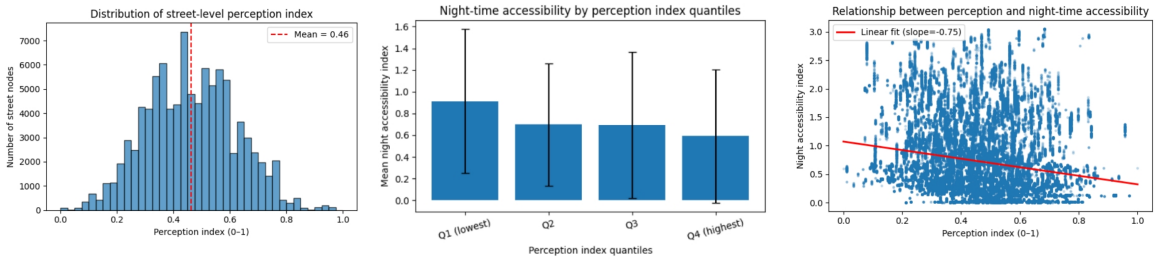


Fig. 6 Street-level perception index and its relationship to night-time accessibility

From numbers to maps, the reachable area during the day shows a pattern, which is generally intuitive, but more textured than the simple distance to the hospital. High-value values are concentrated around hospitals, general clinics, general clinics and parks-dense stations; the extended moderately accessible "ridges" extend along the MRT corridor through mature public housing towns, and community clubs, sports facilities and community parks are intertwined with residential areas. In these places, a 10-minute walk from the station may pass by several potential health support points. At the other end of the spectrum, the model reveals industrial and logistics areas and low-lying basins in some peripheral residential areas, where residential areas and workplaces are spatially separated from sanitary facilities. Here, even at noon, many street nodes can only see a small part of the accessible services in the water collection area. These models echo the known socio-spatial division in Singapore, but they are refined to the level of a single street segment, which is consistent with the recent chart-based analysis of urban health access (Lei et al., 2024; Sideris et al., 2024).

The field at night is not just a dull version of the day. It shrinks and splits. Most of the network, including many streets that look pleasing to the eye and relatively safe in the CLIP semantics, collapse their accessibility index once day clinics, pharmacies and community facilities exit the available set. The rest is a fragile night pillar, consisting of several areas adjacent to hospitals, clusters of 24-hour clinics and pharmacies, and some mixed-purpose corridors in which late-night commercial activities are carried out at the same time as health support functions. Between these main roads is a "healthy desert after dark": continuous industrial roads, waterfront promenades or purely residential lanes, where the nearest night support is beyond a comfortable walking distance. Because there are clinics and community facilities nearby, the contrast between day and night is particularly obvious in the neighborhood with a good score during the day, but at night, almost all these supports are lost. From the perspective of 24-hour cities, the urban structure of these places is only partially able to stand the test of time, and the preventive health commitment is structurally weak. In the experiment, a simple global correlation calculation also shows that the composite perception index and the night-time accessibility index are only weakly and slightly negatively correlated, which further confirms that visually “better” streets do not automatically enjoy stronger night-time support.

One way to interpret these maps is to regard them as a silent comment on the existing policy narrative. The agenda for car life and walking is assumed that under the vision of "20 minutes in a city in a 45-minute city", more people will rely on walking, cycling and public transportation for daily travel (Land Transport Administration, 2019). Preventive health initiatives are based on the premise that residents will have access to primary health care and a supportive environment close to their homes, as described in programmes such as healthier SG and community-level health zones (Ministry of Health, 2022). Both of these imaginations implicitly depend on the streets that act as continuous sanitary corridors. The results here show that this continuity is fragile. For some monitoring stations, especially those embedded in large medical campuses or mixed-use areas, the day and night areas remain relatively consistent; residents, workers and night shift workers can access care on foot most of the time. In other areas, especially in the industrial belt and newer marginal areas, the night monitoring system has gradually separated from the daytime monitoring system, leaving a long gap in the prevention and emergency safety net. This is not the failure of any single policy, but the failure of various measures such as

transportation, health, land use and night economic strategies that are currently encountered on the street (World Economic Forum, 2024; Seijas, 2024).

The model also helps to redefine the role of street perception in these debates. In several station areas, the accessibility during the day is average, but it is mainly concentrated on the streets. CLIP interprets it as green, shady and pedestrian-friendly, with parks, covered sidewalks or ground facilities. In other cases, accessibility is numerically similar, but concentrated on wider roads or visually harsh environments dominated by parking lots, blank walls or busy traffic. Traditional accessibility analysis regards these situations as equivalent; StreetCLIP-GNN implicitly attracts people's attention to differences by linking accessibility with the semantics of the street. For preventive health, arriving at the clinic along a shady and social street is different from arriving at the same clinic along an open industrial road, which echoes the results of perceived safety, greening and mental health research based on street scenes (Naik et al., 2014; Naik et al., 2016; Lu, 2019; Bardhan et al., 2024; Gebru et al., 2017). Although the model cannot directly measure the perception of safety or comfort, it shows how to incorporate these qualities into the same analytical framework as the distribution of facilities.



Fig. 7 Spatial patterns of predicted street-level accessibility: (a) day-time index and (b) night-time index.

Finally, even a simple scenario exercise using a training model illustrates how this analytical framework supports planned dialogue. When a hypothetical 24-hour clinic is added near an industrial-oriented

station, the night site responds in a highly localized way: within a comfortable walking distance, along the limited street corridor, accessibility rises sharply, with little impact on other places. When the business hours of an existing clinic group in a residential town are extended to the night hours, rather than a single new facility, the benefits are more dispersed, and the existing daytime ridge is thickened to something closer to a 24-hour sanitary corridor. In both cases, the global error index has hardly changed, but the pattern of street size will change in a way that planners and residents experience directly. This emphasizes the value of urban artificial intelligence as an oracre that provides the best solution, but as a diagnostic partner, it can reveal that small, targeted interventions may stitch together the weak health pillars of 24-hour cities, which is in line with the recent graphics-based decision-making support system in urban planning. Appeal (Sideris et al., 2024).

4. Conclusion

This paper regards the street as a sensor and the subject to raise a seemingly simple question: in a city that yearns for car life, is health-oriented and increasingly 24-hour, who can really get medical support on foot and when? Based on Singapore's MRT network, we have built a multi-mode street map, combining health-related facilities, clip-based street semantics and simple time rules for opening hours, and training StreetCLIP-GNN to learn day and night accessibility at the level of a single street node. Count.

The result depicts a clear picture. During the day, many MRT corridors form reasonably continuous "healthy ridges", especially around mature residential towns and large hospitals; at night, these ridges shrink into a long and fragmented backbone, leaving long industrial areas, waterfronts and marginal residential areas, which are actually healthy after dark. The desert. Digitally, the graphic model performs better than the non-spatial baseline in daytime tasks and generates more information for night access to the spatial field, even in the case of similar errors in simple averages. The introduction of a composite perception index helps the night model to sharpen this backbone and makes the mismatch between visually friendly streets and well-supported streets empirically visible. In essence, these maps show how existing transportation, sanitation and land use policies work together on the ground - sometimes successful, sometimes unbalanced.

In terms of methods, the study shows that perception at the street level can be woven into a graphic-based accessibility analysis without falling into a black box. The semantics derived from clip cannot replace traditional indicators; they add a missing layer, for example, distinguishing between receiving care through shady, socially active streets and receiving the same nominal care through hostile environments. Politically speaking, the results of the study show that the current narrative of 24-hour cities and preventive health is based on a fragile space: a few stations enjoy 24-hour walking support, while others rely on corridors during the day. Once the clinic is closed, the corridors will disappear.

For planners and public health institutions, this means using Urban AI as a diagnostic companion rather than installing another dashboard. The same model that highlights the ridges and deserts can be used to practice some simple hypothetical questions - adding a 24-hour clinic here and extending the opening hours there - and seeing how far these changes will actually go on the street network. Future work can link these patterns to real moving tracks or health outcomes, expand the time resolution beyond simple day and night divisions, and explore how similar methods perform in other cities, where the promise of all-weather healthy streets is even further.

5. References

Land Transport Authority (LTA) 2019, *Land Transport Master Plan 2040*, Land Transport Authority, Singapore.

Ministry of Health (MOH) 2022, *Healthier SG: White Paper*, Ministry of Health, Singapore.

World Economic Forum 2024, 'Rethinking 24-hour cities: night-time strategies to address urban challenges and thrive', *World Economic Forum*, 9 January.

Naik, N, Philipoom, J, Raskar, R & Hidalgo, CA 2014, 'Streetscore – Predicting the perceived safety of one million streetscapes', *CVPR Workshops*, pp. 793–799.

Naik, N, Raskar, R & Hidalgo, CA 2016, 'Cities are physical too: using computer vision to measure the quality and impact of urban appearance', *American Economic Review*, vol. 106, no. 5, pp. 128–132.

Gebru, T et al. 2017, 'Using deep learning and Google Street View to estimate the demographic makeup of the US', *PNAS*, vol. 114, no. 50, pp. 13108–13113.

Lu, Y 2019, 'Using Google Street View to investigate the association between street greenery and physical activity', *Landscape and Urban Planning*, vol. 191, 103435.

Radford, A et al. 2021, 'Learning transferable visual models from natural language supervision', in *Proceedings of ICML 2021*.

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

Kipf, TN & Welling, M 2017, 'Semi-supervised classification with graph convolutional networks', *International Conference on Learning Representations (ICLR)*.

Lei, B et al. 2024, 'Predicting building characteristics at urban scale using street-view imagery and graph neural networks', *Computers, Environment and Urban Systems*, vol. 107, 101990.

Liu, C et al. 2024, 'Graph attention networks unveil determinants of intra-urban health disparities', *Health & Place* (online first).

Bardhan, M et al. 2024, 'From space to street: a systematic review of the associations between visible greenery and bluespace in street view imagery and mental health', *Landscape and Urban Planning* (online first).

Sideris, N et al. 2024, 'Enhancing urban data analysis: leveraging graph-based decision support systems', *International Journal of Environmental Research and Public Health*, vol. 21, no. 2, 192.