DOI: 10.1111/tgis.13005

RESEARCH ARTICLE

Transactions in GIS WILEY

Street-level traffic flow and context sensing analysis through semantic integration of multisource geospatial data

Yatao Zhang[1,2] | Martin Raubal[1,2]

[1]ETH Zurich, Future Resilient Systems, Singapore-ETH Centre, Singapore, Singapore

[2]Institute of Cartography and Geoinformation, ETH Zurich, Zurich, Switzerland

Correspondence
Yatao Zhang, ETH Zurich, Future Resilient Systems, Singapore-ETH Centre, 1 CREATE Way, #06-01 CREATE Tower, Singapore 138602, Singapore.
Email: yatzhang@ethz.ch

Funding information
National Research Foundation Singapore

Abstract

Sensing urban spaces from multisource geospatial data is vital to understanding the transportation system in the urban context. However, the complexity of urban context and its indirect interaction with traffic flow deepen the difficulty of exploring their relationship. This study proposes a geo-semantic framework first to generate semantic representations of multi-hierarchical urban context and street-level traffic flow, and then investigate their mutual correlation and predictability using a novel semantic matching method. The results demonstrate that each street is associated with its multi-hierarchical spatial signatures of urban context and street-level temporal signatures of traffic flow. The correlation between urban context and traffic flow displays higher values after semantic matching than those in multi-hierarchies. Moreover, we found that utilizing traffic flow to predict urban context results in better accuracy than the reversed prediction. The results of signature analysis and relationship exploration can contribute to a deeper understanding of context-aware transportation research.

1 | INTRODUCTION

Context awareness refers to linking entities with their contextual information to handle the interpretation task of these entities (Perera et al., 2014). Recently, context-aware techniques have attracted much attention in the fields of ubiquitous computing, smart cities, and traffic management (Bibri, 2018; Liu, Wan, et al., 2017). These techniques are particularly important for transportation research, as the movement of entities that constitutes traffic flow is embedded in the urban context (Buchin et al., 2012). Thus, exploring the relationship between urban context and traffic flow is of great significance for a better understanding of the transportation system.

The effects of urban context on human movement are varied, which can both enable and limit traffic flow in the transportation network (Sharif & Alesheikh, 2017). From a meteorological perspective, urban context can help detect human mobility patterns under different weather conditions (Lana et al., 2018), and boost traffic analysis when encountering extreme weather situations that are non-recurrent (Yu et al., 2017). Many previous studies have explored the importance of urban weather context in traffic flow and their relationship (Theofilatos & Yannis, 2014). From a geographic perspective, urban context is closely associated with most factors in human–environment research spanning from population density to road structures to land use (Tedjopurnomo et al., 2020). The richness of geographic context leads to the diversity of data sources, such as satellite images, point of interest (POI) data, and GPS trajectories (Li et al., 2016). Buchin et al. (2012) also enumerated some examples of geographic context that are relevant for analyzing human movement data, including the transportation network, land cover, and locations. Despite the importance of geographic context in traffic flow, its richness deepens the difficulty to measure urban context from the geographic view and explores the relationship between urban context and traffic flow.

Essentially, traffic flow in the transportation network connects two essential aspects of urban context geographically, that is, social and natural aspects (Tu et al., 2020; Wang et al., 2019). The social aspect depicts the thematic characteristics of urban entities based on multisource urban data, such as POI and social media data (Liu, He, et al., 2017; Yuan et al., 2014). The natural aspect can be described using remote sensing images to uncover the physical characteristics of urban entities, including spectral, textural, and local features (Shahriari & Bergevin, 2017; Zhong et al., 2015). By combining the social and natural aspects, the capability of multisource geospatial data to sense the urban environment can be greatly improved (Liu, He, et al., 2017; Zhang, Li, et al., 2019). For example, Zhang, Li, et al. (2019) proposed a cross-correlated framework to integrate the social and natural aspects of urban context to recognize functional urban land use, which presented a better performance than using either aspect separately. However, there is still a lack of studies to incorporate the two aspects of urban context into traffic flow analysis.

In different situations, the compositions of urban context that may influence human movement are diverse due to the spatial variance of urban environments. Given a pick-up or drop-off GPS point, the context refers to nearby POIs of the GPS point that passengers may visit, that is, their origin or destination places (Huang & Gartner, 2014). Given the traffic flow in a street, the context refers to the distribution of nearby land parcels that may influence route planning and traffic congestion (Zhang, Sun, et al., 2017). Unlike a pick-up or drop-off point that has a direct interaction with its context, for example, passengers' visit and stay (Huang & Gartner, 2014; Liu et al., 2012), the interaction between traffic flow and its context is indirect and challenging to measure because of their near-untouchable attribute. This indirect interaction can present complicated outcomes in different situations. For example, the common situation is that traffic flow in the downtown area has a higher probability of being congested than in the suburban area during peaking hours (Zhao & Hu, 2019). But there are also some exceptions, such as dense residential areas away from the city center in China. Meanwhile, not every traffic flow has a meaningful relationship with its nearby urban context, such as traffic flow on an overpass. Thus, an in-depth study is needed to disclose the indirect and complicated interaction between urban context and traffic flow.

To integrate the social and natural aspects of urban context and reveal its relationship with traffic flow, this study proposes a novel geo-semantic framework to represent urban context and traffic flow at the street level and explore their correlation and predictability. Due to the scale-dependent characteristics of geographic objects (Tu et al., 2020), the framework first establishes a multi-hierarchical structure of urban context to cover various influential regions for each street. Following that, we utilize the probabilistic topic modeling method to extract semantic representations of multi-hierarchical urban context by combining its social and natural aspects and traffic flow by capturing its mobility patterns (Miao et al., 2021; Roller & Walde, 2013). A semantic matching method is then designed to produce correlated versions of all semantic representations borrowed from the cross-modal retrieval system (Pereira et al., 2013). Finally, we can identify the most suitable hierarchy of each street by comparing the distance between correlated representations of urban context and traffic flow, and explore their relationships through correlation and predictability.

The remainder of this article is organized as follows. Section 2 introduces previous research on urban context sensing, especially land use recognition. Section 3 provides a detailed description of datasets and the proposed framework. The experimental results and discussion are reported in Sections 4 and 5, respectively, followed by a conclusion in Section 6.

2 | URBAN CONTEXT SENSING

Accurately sensing urban spaces from multisource geospatial data can benefit the understanding of urban context and human activities in cities worldwide (Raubal et al., 2021; Zhang, Li, et al., 2021). Several data-driven frameworks have been proposed to perceive urban spaces. For example, Zheng et al. (2014) introduced the framework of urban computing, which connects urban sensing, data management, data analytics, and service providing into a recurrent process, to handle urban problems using urban big data. In addition, social sensing, urban informatics, and crowd-sourcing studies are also promising frameworks for understanding urban spaces from a data-driven perspective (Foth et al., 2011; Li, 2017; Liu et al., 2015).

In urban sensing, utilizing geospatial data to sense the natural and social aspects of urban context has attracted much attention, especially in recognizing urban land use (Yuan et al., 2014; Zhang et al., 2018). From the natural aspect of urban context, the object-based classification method plays a critical role in detecting urban land use from high-spatial resolution (HSR) satellite images (Zhang, Du, et al., 2017). This method extracts feature descriptions within land parcels (i.e., objects) to classify their land use (Gamanya et al., 2007). Initially, some low-level features from HSR satellite images were used to describe land parcels, such as spectral, texture, and local features. These features can effectively provide the physical characteristics of urban land parcels but fail to reflect the scene information of geographic objects, for example, high-level semantic information (Tu et al., 2020). In other words, it is difficult for these features to overcome the semantic gap between low-level physical characteristics and high-level scene information within land parcels (Liu, He, et al., 2017).

To handle this semantic gap, probabilistic topic models were proposed to extract semantic features from urban land parcels to represent the scene information (Tu et al., 2020). The models contain a variety of methods to discover hidden semantic representations from urban context, such as probabilistic latent semantic analysis (pLSA) and latent Dirichlet allocation (LDA) (Blei et al., 2003). Zhong et al. (2015) evaluated the performance of using pLSA and LDA to recognize urban land use through extracted semantic features. The results suggested that the LDA model performs better than the pLSA model in different datasets. Although the LDA model can produce qualified representations of high-level scene information from satellite images, it is challenging to discern land parcels with similar physical features but different thematic attributes, for example, public service buildings and commercial buildings located on the same street (Liu, He, et al., 2017).

This challenge leads to the social aspect of urban context sensing, which can be solved by taking multi-source geospatial data with thematic attributes into account, such as POI data (Yao et al., 2016), social media data (Zhang, Li, et al., 2019), and mobile phone data (Tu et al., 2017). The solution is highly motivated by the data-driven frameworks of urban computing, social sensing, and urban informatics. Generally, fusing HSR satellite images with the above mentioned geospatial data can integrate the physical and thematic characteristics to simultaneously reveal the natural and social aspects of urban context (Tu et al., 2020; Zhang, Li, et al., 2019). For example, Liu, He, et al. (2017) presented a framework for combining HSR satellite images with social media data to classify urban functions using the LDA model. Zhang, Li, et al. (2019) revealed the advantages of correlated semantic representations obtained from HSR satellite images, POI data, and social media data in land use recognition by extending the LDA model. Compared to studies purely using satellite images, multisource geospatial data fusion can significantly improve the performance of the LDA model in urban context sensing. Following the fusion method, this study will investigate a multi-hierarchical structure of urban context sensing and its application to transportation research.

3 | RESEARCH DESIGN

3.1 | Study area and datasets

This study was conducted in Singapore with a population of 5.45 million and an area of 728.6 km^2 (Figure 1a). As a high-density city, the landscape in Singapore is dominated by manufactured structures and holds many residential and commercial facilities (Sidhu et al., 2018). Urban streets in high-density cities can promote massive traffic flow among land parcels and improve urban vitality (Sulis et al., 2018; Zhang, Li, et al., 2021), making Singapore an ideal case to explore the relationship between urban context and traffic flow.

The fused HSR satellite image, POI data, and Grab-Posisi GPS dataset constitute the primary data to extract semantic representations from urban context and traffic flow. Also, transportation network and land use data from the Urban Redevelopment Authority provide auxiliary information to conduct this study. The information of datasets and their basic preprocessing are as follows.

- The HSR satellite image in Figure 1a holds a spatial resolution of 2.39 m after fusing the Sentinel-2 and Google Earth images. Two Sentinel-2 images with four channels (NIR, R, G, and B) in 2020 and 2021 were first fused to remove clouds. The result was then fused into the Google Earth image collected in 2021 using the high-pass filter resolution merge method (Gangkofner et al., 2007). Finally, we obtain the fused HSR satellite image with four bands and consistent spectral information across Singapore.

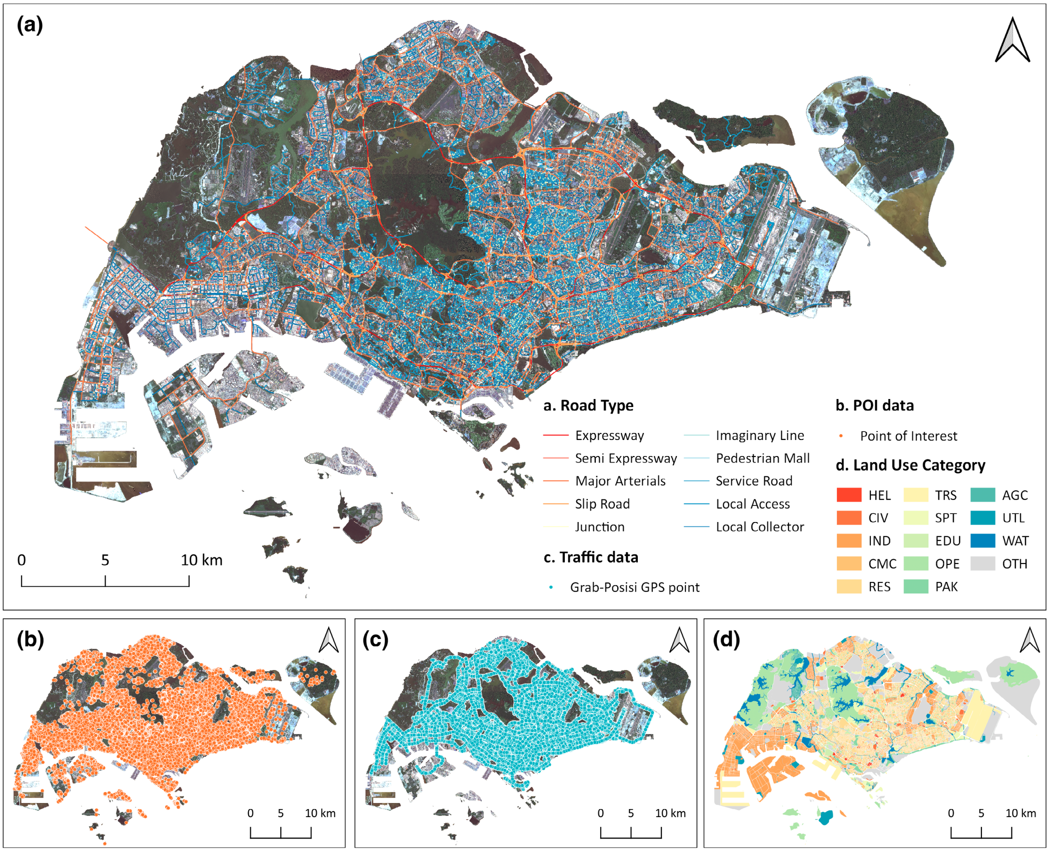

FIGURE 1 Study area and the used datasets in this study. (a) The fused satellite image of the Google and Sentinel-2 images and the transportation network. (b) POI data. (c) Grab-Posisi dataset. (d) Land use data.

- The POI data in Figure 1b is from Open Street Map (OSM) in 2021, including the original POIs and building polygons. Given a building polygon, its centroid point was extracted as a POI to replenish the number of POIs in Singapore. In addition, their categories were reclassified to align with land use information in Figure 1d based on their locations. In total, we have 116,097 POIs in Singapore.
- The Grab-Posisi dataset in Figure 1c is an open-source GPS trajectory dataset in Singapore provided by Grab,[1] a Southeast Asia's ride-sharing company (Huang et al., 2019). The Grab-Posisi dataset was collected from Grab driver's phones while in transit during April 2019 with a 1-s sampling rate, that is, from 08/04/2019 to 21/04/2019. Each GPS point has the trajectory ID, time, location, speed, etc. The dataset also provides contextual information, including bearing, the accuracy level, driving modes, and data acquisition modes, which can be used in map inference, mode detection, traffic detection and forecast, and next location prediction (Huang et al., 2019). In Singapore, there are 28,000 trajectories in total, and we employ this dataset to extract traffic flow by aggregating the GPS trajectories in each street segment.
- The transportation road network and land use data from 2019, shown in Figure 1a,d, can be obtained from Singapore's Urban Redevelopment Authority. The transportation road network has the following types: expressway, semi-expressway, major arterials, local collector, local access, service road, etc. After reclassifying the category of land use data, it contains the following 14 types: health and medical care (HEL), civic and community institution (CIV), industry (IND), commercial (CMC), residential (RES), transport facilities (TRS), sports and recreation (SPT), educational institution (EDU), open space (OPE), park (PRK), agriculture (AGC), utility (UTL), waterbody (WAT), and other lands (OTH).

3.2 | Research framework

The proposed geo-semantic framework aims to extract the semantic representations of multi-hierarchical urban context and traffic flow, and then explore their relationship according to correlation and predictability, shown in Figure 2. First, the framework adopts street segmentation to partition street segments for computing traffic flow and implements hierarchical region segmentation to establish a multi-hierarchical structure for computing urban context. Then, we utilize the probabilistic topic modeling method to generate semantic representations of urban context in multi-hierarchies and semantic representations of traffic flow in each street (Miao et al., 2021; Roller & Walde, 2013). Finally, the relationship between urban context and traffic flow is explored by selecting the most suitable hierarchy of urban context for per-street traffic flow through a designed semantic matching method (Pereira et al., 2013; Rasiwasia et al., 2010).

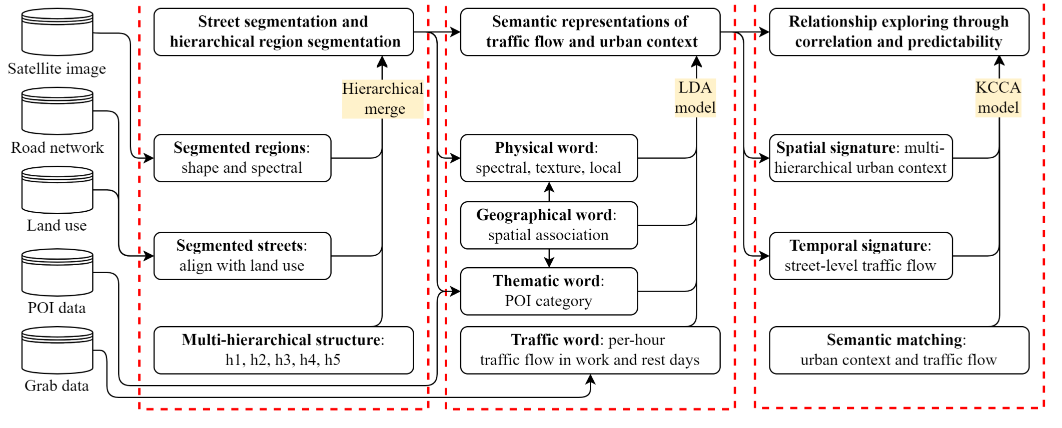

FIGURE 2 The proposed geo-semantic framework to explore the relationship between multi-hierarchical urban context and traffic flow. In the multi-hierarchical structure, h_1–h_5 refer to the five hierarchies of urban context (hierarchy 1–hierarchy 5), respectively.

3.2.1 | Street segmentation and hierarchical region segmentation

A geographic object can exhibit various attributes and geographic phenomena at different spatial scales (Tu et al., 2020). This scale-dependent property also applies to traffic flow and urban context. For example, the value of traffic flow would vary if streets were partitioned by different nodes; the content of urban context nearby a street would also be varied if setting different neighboring coverage regions. Thus, the primary issue is to segment streets and regions properly for computing traffic flow and urban context.

As traffic flow calculated in each street segment is potentially affected by land use information (Zhang, Sun, et al., 2017), each segmented street should generally accord with the distribution of its nearby land parcels. This accordance is also helpful to explore the relationship between traffic flow and urban context since they are spatially associated. In detail, the procedure of street segmentation follows these steps, which finally results in 5581 available street segments.

1. We selected main roads from the transportation road network to accommodate traffic flow, that is, expressways, major arterials, minor arterials, local collectors, and junction roads.
2. The above-selected roads simultaneously satisfying three criteria were merged to generate new street segments. (a) The roads to be merged should spatially touch each other. (b) The roads possess the same road name. (c) Land parcels nearby the roads have the same land use category. The third criterion was evaluated by matching the category of the nearest land parcel to the road.
3. For all merged street segments, the over-long street segments were iteratively split by their middle points until their length was <2 km. The over-short street segments <50 m were filtered out. Details are shown in Figure 3a.

For region segmentation, we used satellite image segmentation to partition a given image into a series of non-overlapping homogeneous regions (Hu et al., 2016). The internal heterogeneity of these segmented regions is controlled by a scale parameter according to shape and spectral criteria in Equation (1) (Draguct et al., 2010; Tu et al., 2020). The weight (w_{shape}) of the shape heterogeneity was set as 0.5 to maintain the balance between shape (c_{shape}) and spectral ($c_{spectral}$) heterogeneity, and the compactness of the shape was also set as 0.5. In this study, we employed the estimation of scale parameter 2 (ESP 2) to identify the best scale parameter (Drăguţ et al., 2014), which was 116.0 for the given satellite image. Figure 3b illustrates the example of the satellite image segmentation result.

$$S_{criteria} = w_{shape} \bullet c_{shape} + \left(1 - w_{shape}\right) \bullet c_{spectral} \tag{1}$$

After identifying the spatial range of each segmented street and region, the second issue is to associate each segmented street with nearby segmented regions. To solve this problem, we designed a multi-hierarchical structure to attach nearby segmented regions in different ranges to the segmented street. Given a street in Figure 3c, its buffer zone was drawn based on a predefined buffer size, and all segmented regions touching this zone were merged to

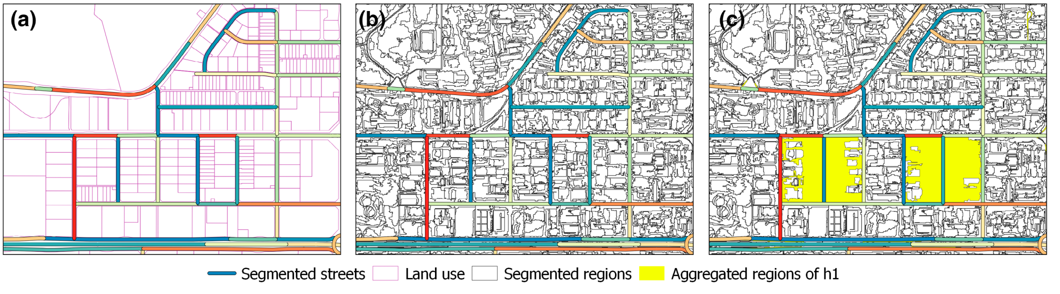

FIGURE 3 The diagram of street segmentation and hierarchical region segmentation. (a) Street segmentation. (b) Region segmentation. (c) Region merging based on multi-hierarchical buffer zones.

generate a new aggregated region. The merging process was implemented based on the pre-segmented regions from the satellite image. Compared to purely using the buffer zone to identify the region of urban context, the merging method takes the shape and spectral homogeneity of each pre-segmented region into account from bottom to top (Hu et al., 2016). This method guarantees that the boundary of the aggregated region, to a large extent, accords with the actual boundary of geographic objects. In addition, each pre-segmented region holds a complete function from a land-use perspective, making the aggregated region consistent with the edge of actual land patches to describe urban context information. In addition, we provided five buffer zone sizes to correspond with five hierarchies, that is, 100, 200, 300, 400, and 500 m, denoted as hierarchy 1 (h_1), hierarchy 2 (h_2), hierarchy 3 (h_3), hierarchy 4 (h_4), and hierarchy 5 (h_5), respectively. From h_1 to h_5, the area of aggregated region increases step by step so that the relationship between traffic flow and its urban context could be minutely evaluated in the further stage.

3.2.2 | Probability topic modeling for traffic flow and urban context

The probabilistic topic modeling method has been widely used to extract high-level semantic representations in urban spaces, which consists of the bag-of-word (BOW) and LDA models (Liu, He, et al., 2017; Zhong et al., 2015). The method regards the set of all research units as the corpus, each research unit as a document, and each basic feature in the document as a word. The BOW model quantizes basic features extracted from raw data into words and utilizes the histogram of words to represent a document (Shahriari & Bergevin, 2017). Then, the LDA model generates semantic representations of each document through a mixture over an underlying set of topics, and each topic is characterized by a distribution of words (Blei et al., 2003; Zhang, Li, et al., 2019). In addition, we adopted the indicator of coherence (C_V) to measure topic coherence with various topic numbers and assist in the selection of topic numbers (Röder et al., 2015). This part introduces the construction of words for traffic flow and urban context using the BOW model and the generation of their semantic representations using the LDA model.

Traffic flow

Its word refers to the amount of traffic flow within a time slot in each street, denoted as a traffic word $w_{tf} = \{s,t,m\}$. s, t, and m represent street ID, time slot ID, and the amount of traffic flow, respectively. The time interval to count traffic flow in each street was set to 1 h, indicating that there are 24 time slots within one day. However, due to the difference of mobility patterns on work and rest days (Yao et al., 2019; Zheng et al., 2014), we separately counted their traffic flow to generate the traffic word for each time slot, that is, 24 time slots on a work day and 24 time slots on a rest day. Then, the range of t contains 48 different time slots, for example, w00:00–01:00 and w23:00–00:00 referring to the first and last time slot on the work day, respectively; r00:00–01:00 and r23:00–00:00 referring to the first and last time slot on the rest day, respectively. For the Grab-Posisi dataset on work and rest days, m_{st} was separately calculated to average the amount of traffic flow in the street s for each time slot t. Given a segmented street s, its document regarding the traffic word can be expressed as:

$$D_s = \left\{ w_{tf}^1 : m_{s1}, w_{tf}^2 : m_{s2}, \ldots, w_{tf}^t : m_{st} \right\} \tag{2}$$

where w_{tf}^t is the t_{th} word of traffic flow and m_{st} is the count of this word. Then, the corpus of traffic flow is the set of all documents $\mathcal{D}_{tf} = \{D_1, D_2, \ldots, D_s\}$, which is input into the LDA model to generate semantic representations of traffic flow, denoted as $\mathcal{S}_{tf}$.

Urban context

POI data and satellite images represent the social and natural aspects of urban context (Zhang, Li, et al., 2019), leading to multimodal LDA topic modeling that involves two data modalities. To capture multimodal components of urban context, we develop three kinds of context words as basic words, that is, thematic, physical, and geographic words. The thematic and physical words are constructed based on POI data and satellite images to characterize the social and natural aspects of urban context, respectively. The geographic word is first devised to associate each thematic word with its surrounding physical words by characterizing their spatial relations. As there are five hierarchies containing urban context in different ranges, semantic representations are separately extracted to model urban context for each hierarchy. Given all aggregated regions in a hierarchy h_i, each aggregated region can be regarded as a document D_r.

(1) Thematic word: the category information of POI data is directly treated as the word due to its discrete properties satisfying the BOW model (Zhang, Li, et al., 2019). Assuming there were K kinds of POIs, the thematic word refers to the amount of POI data for different categories in each aggregated region, denoted as $w_{ct} = \{l, c, m\}$. l, c, and m represent the location, category, and the amount of category c in the aggregated region, respectively. Then, the thematic part of a document in the hierarchy h_i can be expressed as:

$$D_{ri}^{ct} = \left\{ w_{ct}^{1} : m_1, w_{ct}^{2} : m_2, \ldots, w_{ct}^{k} : m_k \right\} \tag{3}$$

where w_{ct}^{k} is the k_{th} thematic word and m_k is the count of this word.

(2) Physical word: previous studies have verified the effectiveness of three kinds of features from satellite images to display multifaceted physical attributes of urban spaces, including spectral, texture, and local features (Liu, He, et al., 2017; Zhang, Li, et al., 2019; Zhong et al., 2015). When collecting these three physical features, the size of a sliding window is set to 150 × 150 pixels with 25 overlapping pixels (Zhong et al., 2015). Given a small region in the sliding window, its spectral feature describes the average value and standard deviation in four bands, denoted as f_{spectral}; its texture feature is measured by four Haralick's statistical features, that is, contract, energy, correlation, and homogeneity, from the gray-level co-occurrence matrix in four bands (Mohanaiah et al., 2013), denoted as f_{texture}; its local feature applies the dense scale-invariant feature transform algorithm in the near-infrared band to provide a 128-dimension vector (Zhang, Li, et al., 2019; Zhong et al., 2015), denoted as f_{local}.

After extracting the spectral, texture, and local features from all sliding windows in the satellite image, we used the K-Means method to cluster these features separately into various clustering numbers (Tu et al., 2020). The Davies–Bouldin index (DBI) was then employed to estimate the clustering performance of each clustering number (Davies & Bouldin, 1979). As shown in Equation (4),

$$\text{DBI} = \frac{1}{N} \sum_{u=1}^{N} \max_{v \neq u} \frac{S_u + S_v}{M_{u,v}} \tag{4}$$

the DBI measures the clustering quality through the separation between cluster u and v, denoted as $M_{u,v}$, and the within-cluster scatter for cluster u and v, denoted as S_u and S_v (Davies & Bouldin, 1979). The lower DBI means a better clustering quality. The clustering results with the lowest DBIs can be regarded as their corresponding feature descriptors, that is, three sub-words of the physical word (Zhang, Li, et al., 2019). Assuming the optimal clustering numbers of f_{spectral}, f_{texture}, and f_{local} are x, y, and z respectively, their descriptors constitute a tuple as (w_{spe}, w_{tex}, w_{loc}). Then, the physical part of a document in the hierarchy h_i can be expressed as:

$$D_{ri}^{cp} = \left\{ \begin{array}{c} w_{\text{spe}}^{1} : m_{\text{spe}}^{1}, \ldots, w_{\text{spe}}^{x} : m_{\text{spe}}^{x}, \\ w_{\text{tex}}^{1} : m_{\text{tex}}^{1}, \ldots, w_{\text{tex}}^{y} : m_{\text{tex}}^{y}, \\ w_{\text{loc}}^{1} : m_{\text{loc}}^{1}, \ldots, w_{\text{loc}}^{z} : m_{\text{loc}}^{z} \end{array} \right\} \tag{5}$$

where w_{spe}^{x}, w_{tex}^{y}, and w_{loc}^{z} are the xth spectral word, yth texture word, and zth local word; m_{spe}^{x}, m_{tex}^{y}, and m_{loc}^{z} are the counts of these three words, respectively.

(3) Geographic word: we first devise a novel method to build connections between thematic and physical words through their relative spatial locations, ensuring that multimodal components of urban context are spatially related. Given a thematic word, $w_{ct} = \{l, c, m\}$, its location l falls into one particular sliding window that can be used to extract

a tuple of the physical word, (w_{spe}, w_{tex}, w_{loc}). The geographic feature is then constructed by combining the thematic and physical words, denoted as $f_{geo} = (w_{ct}, w_{spe}, w_{tex}, w_{loc})$. Similar to the construction of physical words, the DBI-based *K*-Means method is utilized to produce the discrete word of this geographic feature through clustering. Assuming the optimal clustering number is q, the geographic part of a document in the hierarchy h_i can be expressed as:

$$D_{ri}^{cg} = \left\{ w_{cg}^{1} : m_1, w_{cg}^{2} : m_2, \ldots, w_{cg}^{q} : m_q \right\} \tag{6}$$

where the w_{cg}^{q} is the qth geographic word and m_q is the count of this word.

After integrating the thematic, physical, and geographic words in Equations (3, 5, 6) within an aggregated region, its document in the hierarchy h_i can be expressed as $D_{ri} = \left\{D_{ri}^{ct}, D_{ri}^{cp}, D_{ri}^{cg}\right\}$. Then, the corpus of urban context in this hierarchy is the set of all documents $\mathcal{D}_{uc}^{i} = \{D_{1i}, D_{2i}, \ldots, D_{ri}\}$, which is input into the LDA model to generate semantic representations of urban context in h_i, denoted as $\mathcal{S}_{uc}^{i}$. Similarly, semantic representations of urban context at multi-hierarchies can be obtained using the same method, denoted as $\mathcal{S}_{uc} = \left\{\mathcal{S}_{uc}^{1}, \mathcal{S}_{uc}^{2}, \mathcal{S}_{uc}^{3}, \mathcal{S}_{uc}^{4}, \mathcal{S}_{uc}^{5}\right\}$.

3.2.3 | Semantic matching and relationship exploration

The semantic representations of multi-hierarchical urban context and traffic flow allow us to analyze the spatial signatures of urban context and temporal signatures of traffic flow for each street. After that, we design a semantic matching method to project urban context and traffic flow into their common feature spaces and explore their relationship through correlation and predictability.

We began by analyzing the signatures of multi-hierarchical urban context and street-level traffic flow through the DBI-based K-Means clustering method. For the spatial signature of urban context, K-Means clustered its semantic representations $\mathcal{S}_{uc}^{i}$ in each hierarchy into various clustering numbers. The DBI evaluated the clustering performance to select the optimal clustering result (Davies & Bouldin, 1979). Then, land use information within each aggregated region was attached to the optimal clusters by averaging the areas of different kinds of land use, that is, the spatial signature specified by the spatial distribution of land use information. For the temporal signature of traffic flow, we analyzed its signature pattern using the same method to cluster semantic representations $\mathcal{S}_{tf}$. Then, the temporal amount of traffic flow was attached to the optimal clusters by averaging its per-hour amount, that is, the temporal signature specified by the temporal variations of traffic flow amount. Finally, each street can be represented by its multi-hierarchical spatial signatures of urban context and street-level temporal signature of traffic flow.

The next question arises of which hierarchy of urban context best matches traffic flow in a street, as there are multi-hierarchical semantic representations of urban context but only one semantic representation of traffic flow. To handle this problem, we devise a semantic matching method borrowed from the cross-modal retrieval system to retrieve the most suitable semantic representation of urban context from multi-hierarchies for each street (Pereira et al., 2013; Rasiwasia et al., 2010). The core of this method is kernel canonical correlation analysis (KCCA), which nonlinearly projects the semantic representations of two modalities, that is, traffic flow and urban context, into their maximally correlated common spaces (Lai & Fyfe, 2000; Zhang, Li, et al., 2019). Assuming the semantic representations of traffic flow and urban context are $\mathcal{S}_{tf}$ and $\mathcal{S}_{uc} = \left\{\mathcal{S}_{uc}^{1}, \mathcal{S}_{uc}^{2}, \mathcal{S}_{uc}^{3}, \mathcal{S}_{uc}^{4}, \mathcal{S}_{uc}^{5}\right\}$, their correlated representations obtained from their common spaces through KCCA can be expressed as $\mathcal{S}_{tf}'$ and $\mathcal{S}_{uc}' = \left\{\mathcal{S}_{uc}^{1\prime}, \mathcal{S}_{uc}^{2\prime}, \mathcal{S}_{uc}^{3\prime}, \mathcal{S}_{uc}^{4\prime}, \mathcal{S}_{uc}^{5\prime}\right\}$. For each street, we measured the distance between $\mathcal{S}_{tf}'$ and each hierarchy in $\mathcal{S}_{uc}'$, and the hierarchy with the minimal distance was selected as the most suitable hierarchy. For all segmented streets, the set of correlated representations of urban context and traffic flow in their most suitable hierarchies was regarded as the matched version.

After semantic matching, we explored the mutual relationship between traffic flow and urban context through two indicators, that is, correlation and predictability. For the correlation, the Pearson's coefficients between traffic flow $\mathcal{S}_{tf}'$ and urban context $\mathcal{S}_{uc}'$ were calculated to compare the difference between the multi-hierarchical and matched versions. For the predictability, it can be divided into two parts, that is, using urban context to predict traffic flow and using traffic flow to predict urban context. When using urban context $\mathcal{S}_{uc}$ in multiple hierarchies to predict traffic flow $\mathcal{S}_{tf}$, we generated the predicted labels through clustering $\mathcal{S}_{tf}$ and then utilized semantic representations in $\mathcal{S}_{uc}$ to predict these labels. During this process, this study developed multi-group labels by clustering semantic representations of traffic flow into various clustering numbers, that is, 2–10. This multi-group method can guarantee that the measure of the predictability is comprehensive to consider various scenarios. Similarly, we can measure the performance of using traffic flow to predict urban context in multiple hierarchies. Regarding the matched version, we applied $\mathcal{S}'_{uc}$ and $\mathcal{S}'_{tf}$ instead of $\mathcal{S}_{uc}$ and $\mathcal{S}_{tf}$ to predict each other in their most suitable hierarchies through the same method. We used the random forest algorithm to implement the prediction task and adapted the 10-fold cross-validation method to evaluate the performance of predictability. As there are five hierarchies and a matched version, the correlation and predictability in each hierarchy and the matched version are separately computed and then compared their difference.

4 | RESULTS

The three main outcomes from the proposed framework are urban context signatures by clustering their multi-hierarchical semantic representations, traffic flow signatures by clustering their street-level semantic representations, and the relationship between urban context and traffic flow. In the following subsections, we discuss our experiments on these three outcomes and investigate the sensitivity analysis of several critical parameters of the proposed framework.

4.1 | Multi-hierarchical spatial signatures of urban context

We developed a multi-hierarchical structure to cluster aggregated regions through the semantic representations of urban context, and employed land use distribution as the spatial signature for each cluster. Such a structure can uncover the hierarchical characteristic of the complex connotations of spatial urban context. This characteristic holds significant future potential in urban planning since mixed land use is becoming increasingly common in the current urban environments (Dovey & Pafka, 2017).

We used a coherence value (C_V) to measure the topic coherence of the LDA model, assisting to identify the topic number (Röder et al., 2015). Figure 4a shows the values of C_V under various topic numbers in multi-hierarchies. We observe that the C_V values for low hierarchies (e.g., h_1 and h_2) are consistently higher than those for high hierarchies (e.g., h_4 and h_5). This can be explained by delving into the meaning of high hierarchies. With the increase of hierarchies, more geographic objects would be included to model hidden topics, deepening the complexity of urban context. The relatively smaller C_V values in high hierarchies reveal the influence of mixed land use on this spatial complexity. Furthermore, the change trends of C_V values in all five hierarchies are similar, that is, beginning with a sudden rise in the early part, achieving the peak value in the middle, and ending with a slow downtrend. We selected the topic number with the highest C_V value as the parameter of the LDA model, marked by gray dots in Figure 4a, to generate semantic representations of aggregated regions for each hierarchy.

Based on the generated semantic representations, we applied the DBI value to measure the clustering quality by computing the separation between clusters and the scatter within clusters (Davies & Bouldin, 1979). Figure 4b shows the DBI values under various clustering numbers using the *K*-Means clustering method. In five hierarchies, the trends of DBI values are all descending, indicating that a larger clustering number refers to a better clustering quality. We selected the clustering results with the lowest DBI values, marked by gray dots in Figure 4b, to analyze the spatial signatures of aggregated regions through land use information. For each cluster, the area ratio of each kind of land use to this cluster was computed. Then, this area ratio was ranked among clusters to emphasize the difference in land use distribution across clusters.

From five hierarchies, we selected two middle hierarchies to reveal the spatial patterns of urban context signatures, that is, h_2 and h_4 in Figures 4c,d. We find that the spatial signatures of h_2 and h_4 present complex compositions

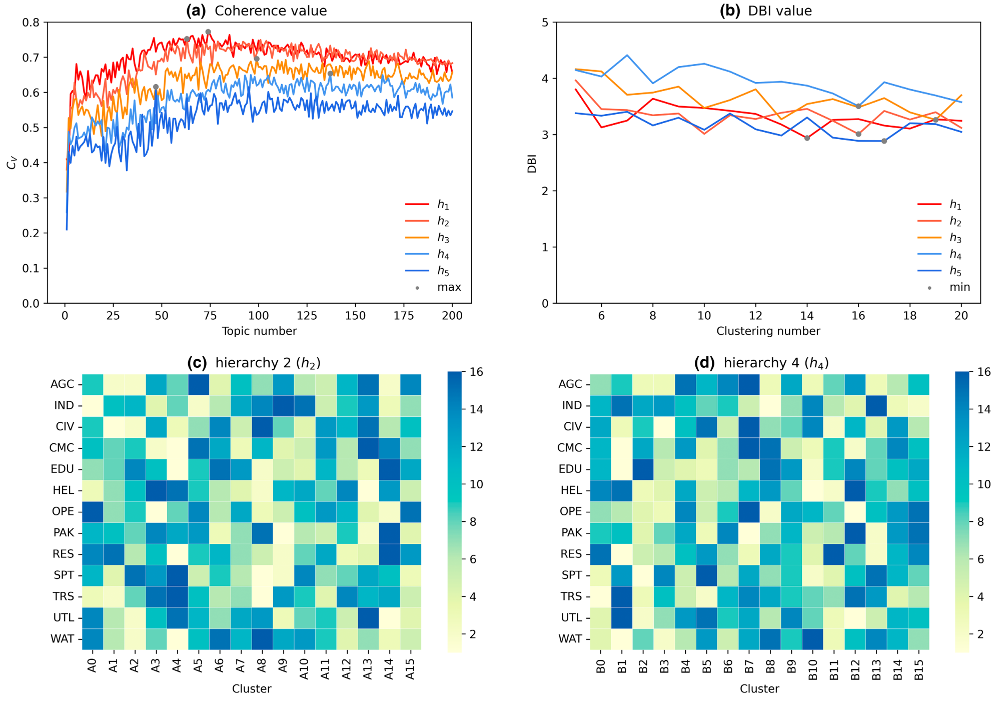

FIGURE 4 Multi-hierarchical urban context sensing analysis and the spatial signatures. (a) The coherence (C_V) values of LDA by setting various topic numbers. (b) The DBI values of *K*-Means by setting various clustering numbers. (c) The area ratio rank heatmap of the spatial signature in h_2. (d) The area ratio rank heatmap of the spatial signature in h_4. In (a, b), h_1–h_5 represent the five hierarchies of urban context (hierarchy 1–hierarchy 5), respectively. In (c, d), the area ratio was first computed in each column, and then this ratio was ranked by each line. The higher number in the color bar means that the area ratio in the cluster has priority over other clusters.

due to the involvement of mixed land use. For example, the highest rank of RES in A14 of h_2 holds a close connection with the CMC, EDU, HEL, and PAK land, but the highest rank of RES in B11 of h_4 is isolated from other land use. Furthermore, we discover the co-occurrence relationship of some kinds of land use despite different hierarchies. For example, the spatial distribution of RSE is positively related to the CMC, EDU, and PAK land but negatively associated with the IND land. Despite land use distribution, the multi-hierarchical semantic representations can be associated with other urban variables, for example, building function and population distribution (Lin et al., 2021; Yao et al., 2017), to help urban planners understand contextual information from a hierarchical perspective.

4.2 | Street-level temporal signatures of traffic flow

In line with St 4.1, we utilized street-level semantic representations of traffic flow to cluster segmented streets, and employed traffic flow variations as the temporal signature for each cluster. This part aims to disclose the mobility patterns of temporal traffic flow in work and rest days. The result can provide references for understanding human activities and locating the distribution of travel demand as the data is from a ride-sharing company.

We also used the C_V and DBI values to evaluate the performance of the LDA and *K*-Means models under various topic numbers and clustering numbers, respectively. Figure 5a presents the change trend of C_V, achieving the marked peak value as the topic number is 6. We selected this topic number to generate street-level semantic representations of traffic flow using the LDA model. The interesting phenomenon appears that the DBI also achieves the marked valley value as the clustering number is 6 in Figure 5b. The same topic and clustering numbers imply the potential

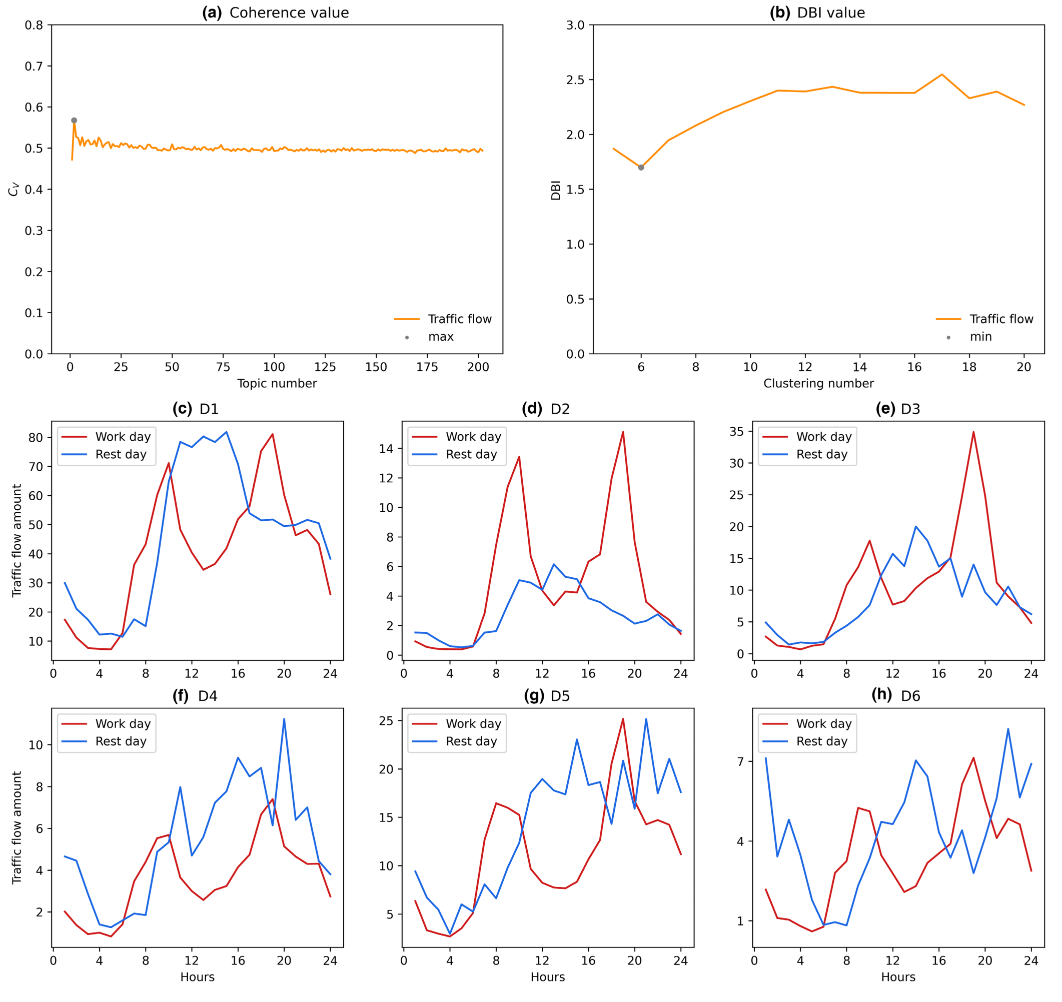

FIGURE 5 Street-level traffic flow analysis and the temporal signatures. (a) The coherence (C_V) values of LDA by setting various topic numbers. (b) The DBI values of *K*-Means by setting various clustering numbers. (c–h) Temporal signatures: The work-day and rest-day traffic flow signatures of six clusters, and each cluster represents its mobility pattern specified by traffic flow. In (c–h), the *x*-axis refers to the *i*th time slot (hour) of the work day and rest day.

consistency of semantic meanings between hidden topics and clustering results. Essentially, the LDA model aims to discover hidden topics to represent traffic flow in each street (Blei et al., 2003), and the streets with similar topics were gathered into corresponding clusters during the clustering process. An evidence is that the values of C_V and DBI at the point of 6 both have an obvious advantage over any other measure points in Figures 5a,b. Moreover, we observe that the change trends of C_V and DBI of urban context in Figure 4 fluctuates and is hard to discern the extreme values, compared to those of traffic flow in Figure 5. This can be explained by the spatial complexity of urban context due to the rich connotations of geographic objects.

Given six clusters of different mobility patterns, Figures 5c–h illustrate their temporal signatures in work and rest days by averaging traffic flow amount in each cluster. In Figure 5c, we observe that D1's signature presents an enormous traffic flow amount, indicating the most common mobility pattern of people taking the Grab. It contains two peak periods in the work day displaying commuting activities and a long peak period in the rest day displaying entertainment behavior. D2 in Figure 5d has a similar workday mobility pattern to D1, but its rest-day signatures differ. The rest-day amount of traffic flow in D2 reduces substantially compared to workday's, which reveals that the vibrancy of urban streets within D2 is mainly activated during the work day. D3 in Figure 5e can be regarded as the enhanced version of D2. A divergent part is that D3's traffic flow amount in the morning is less than that in the afternoon for the rest day. This phenomenon can also be observed in D4, D5, and D6, implying that consumers prefer to take the Grab in the afternoon rather than in the morning. Furthermore, we find that the workday mobility patterns in D4, D5, and D6 can find clues from other clusters, but their rest-day signatures are in different situations. For example, in Figure 5g, the traffic flow amount of D5 in the rest day achieves the peak period until the night, reflecting that there may be many night activities. The rest-day signatures of D4 and D6 in Figures 5f,h show large fluctuations and relatively low traffic amount, which are closely related to the characteristics of origin and destination points of traffic flow.

4.3 | Relationship between urban context and traffic flow

For each street, matching multi-hierarchical urban context representations to one traffic flow representation can be converted into a retrieval problem across multi-modalities. Given two modalities describing the same street, one is a determined modality of its traffic flow representation, and the other is an undetermined modality of its multi-hierarchical context representations. After converting, the problem becomes how to retrieve the most suitable urban context representation from multi-hierarchies through the traffic flow representation. It leads to the third outcome of our framework, that is, semantic matching, which aims to solve the converted problem by projecting the two modalities into their common spaces. Then, we can explore the relationship between urban context and traffic flow through correlation and predictability regarding the multi-hierarchical and the matched versions.

In terms of correlation, Figure 6a demonstrates the Pearson's coefficient values between urban context and traffic flow for the multi-hierarchical and the matched versions. We find that the box plot of the matched version presents overall higher correlation values than the multi-hierarchical version. The 50th percentile in the box plot of the matched version achieves 0.76, higher than 0.58 shown in any other hierarchies. It reveals the effectiveness of semantic matching to capture correlated information between urban context and traffic flow. Also, we find that the correlation difference among the five hierarchies is not apparent, maintaining a similar range in their box plots. This is because urban context in either hierarchy cannot entirely interpret its correlation with traffic flow using a fixed spatial range across the whole of Singapore. Furthermore, we noticed that there also exist some negative correlation values in the matched version. This suggests that not every street would keep a strong positive correlation between traffic flow and its nearby urban context. To illustrate this phenomenon, the correlation value map of Singapore and its two enlarged areas are shown in Figures 6c–e, where the color indicates the correlation value and the width indicates the matched hierarchy.

In Figure 6c, the spatial distribution of correlation implies that most streets retain high correlation values according to observing a large portion of dark red streets in Singapore. This can be verified by the box plot of the matched version in Figure 6a, with its 50th percentile over 0.75 and its 75th percentile over 0.60. Figure 6d presents the correlation distribution in the central area of Singapore, where traffic flow is positively related to its urban context. However, in the southwest area shown in Figure 6e, the situation differs a lot. Traffic flow in partial streets is weakly related to its urban context, especially streets on Jurong Island. Regarding this phenomenon, we infer there are two potential explanations. First, the region is mainly covered by industrial land and distant from the central area, leading to the lack of travel needs by taking the Grab cars. The access of social vehicles to Jurong Island is also limited. Thus, traffic flow in this region could be largely underestimated because of ignoring other traffic vehicles. Second, not every street is closely related to its nearby urban context mentioned above, such as the street on the elevated highway and the context of the complex mixed land use, resulting in low correlation.

In terms of predictability, Figure 6b visualizes two reversed parts of prediction accuracy, that is, "traffic flow predicts context" and "context predicts traffic flow." Similar to correlation, we find that the matched version holds better predictability than the multi-hierarchical version for both two parts. This reveals the advantages of the semantic matching method to better capture the cross-modality information between traffic flow and urban context. In addition, we discover that "traffic flow predicts context" performs much better than "context predicts traffic flow" in all box plots

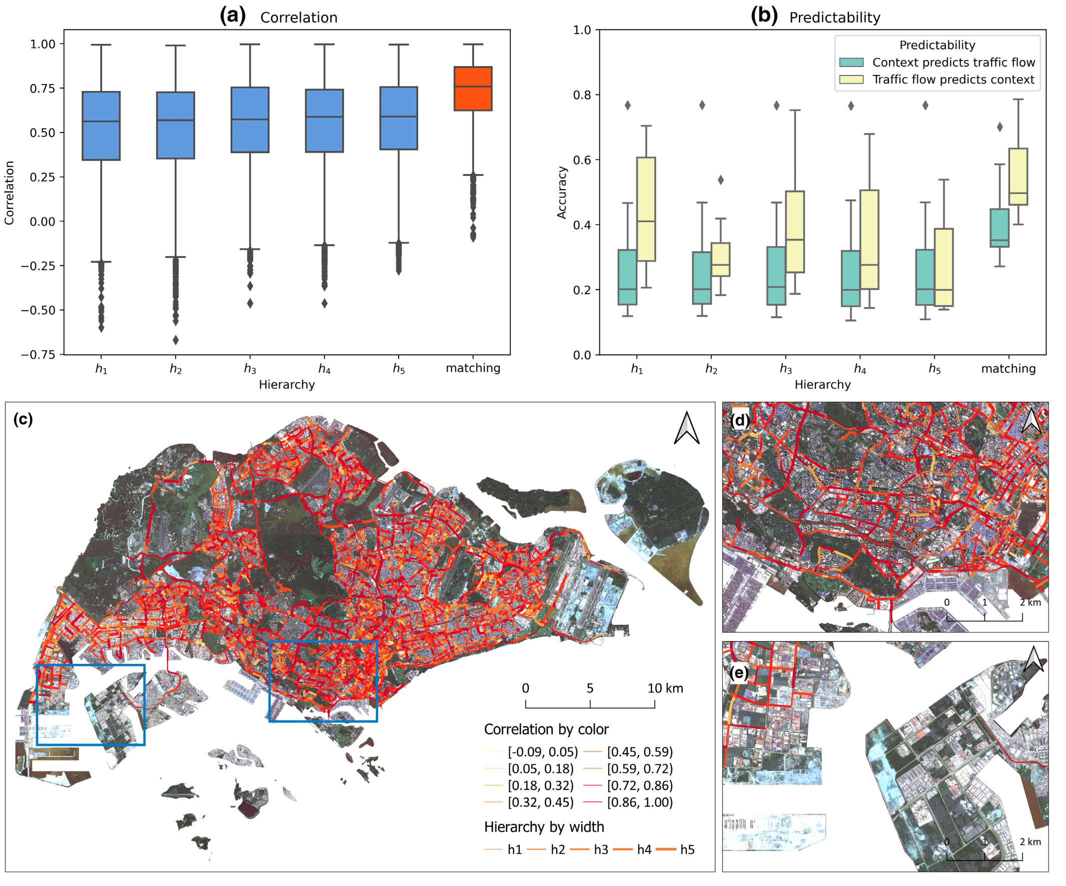

FIGURE 6 The relationship between urban context and traffic flow and their correlation map in Singapore. (a) The correlation box plot for the multi-hierarchical and the matched versions. (b) The predictability box plot for the multi-hierarchical and the matched versions. (c) The correlation map in each street for the matched version in Singapore. (d) The enlarged correlation map in the central area (right rectangle in c). (e) The enlarged correlation map in the southwest area (left rectangle in c). In (a–c), h_1–h_5 represents the five hierarchies of urban context (hierarchy 1–hierarchy 5), respectively. For each street in (c–e), its color and width indicate the correlation value and the matched hierarchy, respectively.

in Figure 6b. Although urban context holds the potential to unveil spatiotemporal urban mobility patterns (Zhang, Wu, et al., 2019), it is still a challenging task to predict dynamic traffic flow accurately by purely utilizing static urban context. In contrast, mining human activities hidden in mobility data can significantly benefit recognizing urban land use and environmental attributes (Liu et al., 2012; Yao et al., 2019; Yuan et al., 2012). Meanwhile, the biased sampling of the Grab-Posisi dataset can also lead to the unbalanced performance of two reversed parts. This dataset only provides the driving records of partial Grab's consumers in Singapore that cannot represent all commuting activities of residents.

5 | DISCUSSION

This study proposed a geo-semantic framework to produce representations of urban context and traffic flow using the BOW and LDA models. In this framework, how to determine the range of documents, the category of words, and the number of topics are critical (Tu et al., 2020; Zhang, Li, et al., 2019). For traffic flow, its documents and words, that is, street segments and traffic words w_{tf}, are fixed due to its discrete properties. For urban context, its documents and

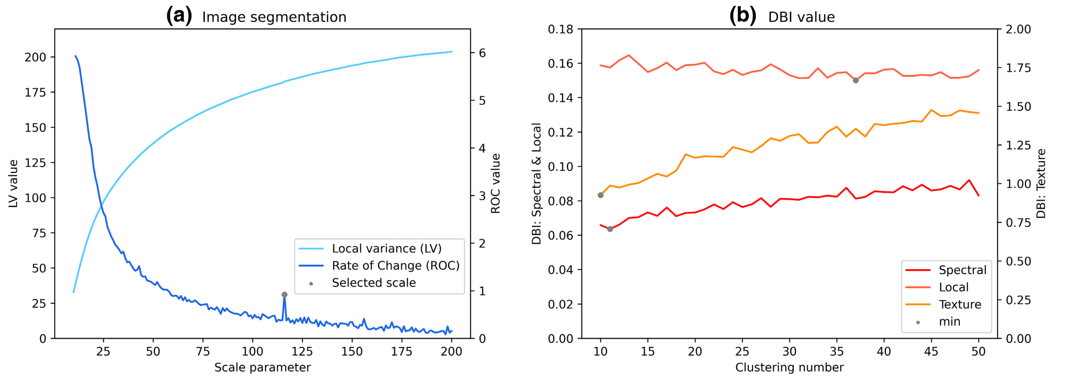

FIGURE 7 The selection of the scale parameter and physical word numbers. (a) Local variance (LV) and its rate of change (ROC) with different scale parameters for satellite image segmentation. (b) DBI values of the physical words along with various clustering numbers for the spectral, texture, and local words.

words are decided by a scale parameter of satellite image segmentation that influences the size of pre-segmented regions and the numbers of physical words, including the spectral (w_{spe}), texture (w_{tex}), and local (w_{loc}) words. As to the selection of topic numbers, they have been discussed in Section 4.1 and Figure 4a for urban context and in Section 4.2 and Figure 5a for traffic flow. Thus, this part discusses the scale parameter selection in satellite image segmentation and the selection of physical word numbers in the BOW model.

In satellite image segmentation, ESP 2 detects scale transition by measuring local variance (LV), which can reflect the segmentation performance of the satellite image (Drăguţ et al., 2014). The rate of change (ROC) between LVs of two neighboring scales can help to select a suitable scale parameter. In Figure 7a, LV increases along with the increase in the scale parameter, implying that the homogeneity of segmented regions increases. Since satellite image segmentation aims to ensure homogeneity within segmented regions and heterogeneity among segmented objects, the higher LV value contributes to the better segmentation result for interpretation. After that, the segmented results need to be merged to generate multi-hierarchies. Thus, the size of each segmented region cannot be too large in case the merged region contains an oversized area. Considering all these factors, we selected the peak point in the curve of ROC as the targeted scale with a scale parameter of 116.0, where LV has a sudden increase in Figure 7a. In general, the LV at this point is high to guarantee the segmentation quality, and the average size of segmentation results is also suitable to aggregate multi-hierarchies shown in Figure 3.

In the BOW model, the solution to identifying the word numbers of non-discrete features is to discretize these features by clustering them into multiple clusters (Tu et al., 2020; Zhang, Li, et al., 2019). Figure 7b visualizes the DBI values of three physical words with various clustering numbers, including the spectral (w_{spe}), texture (w_{tex}), and local (w_{loc}) words. The lowest value of DBI means the best clustering quality, whose clustering number can be regarded as the word number. For the spectral and texture features, their overall trends of DBIs increase along the *x*-axis, which achieve the lowest DBIs when the clustering numbers are 11 and 10, respectively. For the local feature, its overall trend presents a decreasing shift along with the clustering number, and we can identify its lowest DBI value when the clustering number is 37. The obtained clusters of these three features are the three kinds of sub-words under the physical word.

6 | CONCLUSION

To fully understand the relationship between urban context and traffic flow, this study proposed a novel geo-semantic framework to analyze multi-hierarchical urban context and traffic flow at the street level. In each street, we provided multi-hierarchical urban context signatures specified by land use distribution from a spatial perspective and the traffic flow signature specified by mobility patterns from a temporal perspective. Furthermore, we found that urban context and traffic flow present a close relationship in terms of correlation and predictability after semantic matching.

The contribution of this study is threefold. First, we devised a novel framework to associate urban context with traffic flow by projecting the cross-modal semantic representations into their common spaces. This framework can be further extended to related research to explore the relationship across multiple modalities. Second, we demonstrated the use of a multi-hierarchical structure in comprehensively depicting street-level contextual information, which can provide personalized contextual representations depending on the targeted research. Also, the geographic word enables the generated semantic representations to better connect the social and natural aspects of urban context. Third, the proposed framework can provide support for other urban studies, such as utilizing semantic representations of urban context and traffic flow in POI recommendation and next place prediction.

However, this study fails to integrate dynamic contextual information into urban context sensing from the temporal dimension, which is crucial to improve the ability to predict traffic flow. In the future, this research will be expanded to incorporate temporal information from urban environments into urban context sensing, which can help urban planners make more evidence-based transportation planning strategies.

ACKNOWLEDGMENTS

The research was conducted at the Future Resilient Systems program at the Singapore-ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation Singapore. This research is supported by the National Research Foundation Singapore (NRF) under its Campus for Research Excellence and Technological Enterprise (CREATE) program. The authors would like to acknowledge the valuable feedback received from Nishant Kumar.

CONFLICT OF INTEREST

The authors report there are no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data and codes supporting main findings of this study are available in figshare.com with the identifier https://figshare.com/s/7cdb6266493e95498546.

ORCID

Yatao Zhang https://orcid.org/0000-0001-5701-2836
Martin Raubal https://orcid.org/0000-0001-5951-6835

ENDNOTE

[1] Data details: https://engineering.grab.com/grab-posisi.

REFERENCES

Bibri, S. E. (2018). Smart sustainable cities of the future: The untapped potential of big data analytics and context–aware computing for advancing sustainability. Springer. https://doi.org/10.1007/978-3-319-73981-6

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022. https://dl.acm.org/doi/10.5555/944919.944937

Buchin, M., Dodge, S., & Speckmann, B. (2012). Context-aware similarity of trajectories. International Conference on Geographic Information Science, 7478(612), 43–56. https://doi.org/10.1007/978-3-642-33024-7_4

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2, 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

Dovey, K., & Pafka, E. (2017). What is functional mix? An assemblage approach. Planning Theory & Practice, 18(2), 249–267. https://doi.org/10.1080/14649357.2017.1281996

Draguct, L., Tiede, D., & Levick, S. R. (2010). ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. International Journal of Geographical Information Science, 24(6), 859–871. https://doi.org/10.1080/13658810903174803

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. ISPRS Journal of Photogrammetry and Remote Sensing, 88, 119–127. https://doi.org/10.1016/j.isprsjprs.2013.11.018

Foth, M., Choi, J. H., & Satchell, C. (2011). Urban informatics. 2011 ACM Conference on Computer Supported Cooperative Work, Hangzhou, China (pp. 1–8). ACM. https://doi.org/10.1145/1958824.1958826

Gamanya, R., De Maeyer, P., & De Dapper, M. (2007). An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization. Expert Systems with Applications, 32(2), 616–624. https://doi.org/10.1016/j.eswa.2006.01.055

Gangkofner, U. G., Pradhan, P. S., & Holcomb, D. W. (2007). Optimizing the high-pass filter addition technique for image fusion. Photogrammetric Engineering & Remote Sensing, 73(9), 1107–1118. https://doi.org/10.14358/PERS.73.9.1107

Hu, Z., Li, Q., Zou, Q., Zhang, Q., & Wu, G. (2016). A bilevel scale-sets model for hierarchical representation of large remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 54(12), 7366–7377. https://doi.org/10.1109/TGRS.2016.2600636

Huang, H., & Gartner, G. (2014). Using trajectories for collaborative filtering-based POI recommendation. International Journal of Data Mining, Modelling and Management, 6(4), 333–346. https://doi.org/10.1504/IJDMMM.2014.066762

Huang, X., Yin, Y., Lim, S., Wang, G., Hu, B., Varadarajan, J., Zheng, S., Bulusu, A., & Zimmermann, R. (2019). Grab-posisi: An extensive real-life GPS trajectory dataset in Southeast Asia. 3rd ACM SIGSPATIAL International Workshop on Prediction of Human Mobility, Chicago, IL (pp. 1–10). ACM. https://doi.org/10.1145/3356995.3364536

Lai, P. L., & Fyfe, C. (2000). Kernel and nonlinear canonical correlation analysis. International Journal of Neural Systems, 10(5), 365–377. https://doi.org/10.1142/S012906570000034X

Lana, I., Del Ser, J., Velez, M., & Vlahogianni, E. I. (2018). Road traffic forecasting: Recent advances and new challenges. IEEE Intelligent Transportation Systems Magazine, 10(2), 93–109. https://doi.org/10.1109/MITS.2018.2806634

Li, Q. (2017). From geomatics to urban informatics. Geomatics and Information Science of Wuhan University, 42(1), 1–6. https://doi.org/10.13203/j.whugis20160200

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A., & Cheng, T. (2016). Geospatial big data handling theory and methods: A review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing, 115, 119–133. https://doi.org/10.1016/j.isprsjprs.2015.10.012

Lin, A., Sun, X., Wu, H., Luo, W., Wang, D., Zhong, D., Wang, Z., Zhao, L., & Zhu, J. (2021). Identifying urban building function by integrating remote sensing imagery and POI data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 8864–8875. https://doi.org/10.1109/JSTARS.2021.3107543

Liu, J., Wan, J., Jia, D., Zeng, B., Li, D., Hsu, C. H., & Chen, H. (2017). High-efficiency urban traffic management in context-aware computing and 5G communication. IEEE Communications Magazine, 55(1), 34–40. https://doi.org/10.1109/MCOM.2017.1600371CM

Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., & Hong, Y. (2017). Classifying urban land use by integrating remote sensing and social media data. International Journal of Geographical Information Science, 31(8), 1675–1696. https://doi.org/10.1080/13658816.2017.1324976

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., & Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. Annals of the Association of American Geographers, 105(3), 512–530. https://doi.org/10.1080/00045608.2015.1018773

Liu, Y., Wang, F., Xiao, Y., & Gao, S. (2012). Urban land uses and traffic 'source-sink areas': Evidence from GPS-enabled taxi data in Shanghai. Landscape and Urban Planning, 106(1), 73–87. https://doi.org/10.1016/j.landurbplan.2012.02.012

Miao, N., Xue, F., & Hong, R. (2021). Multi-modal semantics-based supervised latent dirichlet allocation for event classification. IEEE Multimedia, 28(4), 8–17. https://doi.org/10.1109/MMUL.2021.3077915

Mohanaiah, P., Sathyanarayana, P., & GuruKumar, L. (2013). Image texture feature extraction using GLCM approach. International Journal of Scientific and Research Publications, 3(5), 1.http://www.ijsrp.org/research-paper-0513.php?rp=P171166

Pereira, J. C., Coviello, E., Doyle, G., Rasiwasia, N., Lanckriet, G. R., Levy, R., & Vasconcelos, N. (2013). On the role of correlation and abstraction in cross-modal multimedia retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(3), 521–535. https://doi.org/10.1109/TPAMI.2013.142

Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. (2014). Context aware computing for the internet of things: A survey. IEEE Communications Surveys and Tutorials, 16(1), 414–454. https://doi.org/10.1109/SURV.2013.042313.00197

Rasiwasia, N., Costa Pereira, J., Coviello, E., Doyle, G., Lanckriet, G. R., Levy, R., & Vasconcelos, N. (2010). A new approach to cross-modal multimedia retrieval. 2010 ACM Multimedia International Conference, Firenze, Italy (pp. 251–260). ACM. https://doi.org/10.1145/1873951.1873987

Raubal, M., Bucher, D., & Martin, H. (2021). Geosmartness for personalized and sustainable future urban mobility. In W. Shi, M. F. Goodchild, M. Batty, M. P. Kwan, & A. Zhange (Eds.), Urban informatics (pp. 59–83). Springer. https://doi.org/10.1007/978-981-15-8983-6_6

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China (pp. 399–408). ACM. https://doi.org/10.1145/2684822.2685324

Roller, S., & Walde, S. S. I. (2013). A multimodal LDA model integrating textual, cognitive and visual modalities. 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, WA (pp. 1146–1157). ACL. https://aclanthology.org/D13-1115.pdf

Shahriari, M., & Bergevin, R. (2017). Land-use scene classification: A comparative study on bag of visual word framework. Multimedia Tools & Applications, 76(21), 1–17. https://doi.org/10.1007/s11042-016-4316-z

Sharif, M., & Alesheikh, A. A. (2017). Context-awareness in similarity measures and pattern discoveries of trajectories: A context-based dynamic time warping method. GIScience & Remote Sensing, 54(3), 426–452. https://doi.org/10.1080/15481603.2017.1278644

Sidhu, N., Pebesma, E., & Câmara, G. (2018). Using Google earth engine to detect land cover change: Singapore as a use case. European Journal of Remote Sensing, 51(1), 486–500. https://doi.org/10.1080/22797254.2018.1451782

Sulis, P., Manley, E., Zhong, C., & Batty, M. (2018). Using mobility data as proxy for measuring urban vitality. Journal of Spatial Information Science, 16, 137–162. https://doi.org/10.5311/JOSIS.2018.16.384

Tedjopurnomo, D. A., Bao, Z., Zheng, B., Choudhury, F., & Qin, A. K. (2020). A survey on modern deep neural network for traffic prediction: Trends, methods and challenges. IEEE Transactions on Knowledge and Data Engineering, 34(4), 1544–1561. https://doi.org/10.1109/TKDE.2020.3001195

Theofilatos, A., & Yannis, G. (2014). A review of the effect of traffic and weather characteristics on road safety. Accident Analysis & Prevention, 72, 244–256. https://doi.org/10.1016/j.aap.2014.06.017

Tu, W., Cao, J., Yue, Y., Shaw, S. L., Zhou, M., Wang, Z., Chang, X., Xu, Y., & Li, Q. (2017). Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. International Journal of Geographical Information Science, 31(12), 2331–2358. https://doi.org/10.1080/13658816.2017.1356464

Tu, W., Zhang, Y., Li, Q., Mai, K., & Cao, J. (2020). Scale effect on fusing remote sensing and human sensing to portray urban functions. IEEE Geoscience and Remote Sensing Letters, 18, 38–42. https://doi.org/10.1109/LGRS.2020.2965247

Wang, J., Wu, J., Wang, Z., Gao, F., & Xiong, Z. (2019). Understanding urban dynamics via context-aware tensor factorization with neighboring regularization. IEEE Transactions on Knowledge and Data Engineering, 32(11), 2269–2283. https://doi.org/10.1109/TKDE.2019.2915231

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., & Mai, K. (2016). Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. International Journal of Geographical Information Science, 31(4), 825–848. https://doi.org/10.1080/13658816.2016.1244608

Yao, Y., Liu, X., Li, X., Zhang, J., Liang, Z., Mai, K., & Zhang, Y. (2017). Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data. International Journal of Geographical Information Science, 31(6), 1220–1244. https://doi.org/10.1080/13658816.2017.1290252

Yao, Y., Zhang, Y., Guan, Q., Mai, K., & Zhang, J. (2019). Sensing multi-level urban functional structures by using time series taxi trajectory data. Geomatics and Information Science of Wuhan University, 44(6), 875–884. https://doi.org/10.13203/j.whugis20170111

Yu, R., Li, Y., Shahabi, C., Demiryurek, U., & Liu, Y. (2017). Deep learning: A generic approach for extreme condition traffic forecasting. 2017 SIAM International Conference on Data Mining, Houston, TX (pp. 777–785). SIAM. https://doi.org/10.1137/1.9781611974973.87

Yuan, N. J., Zheng, Y., Xie, X., Wang, Y., Zheng, K., & Xiong, H. (2014). Discovering urban functional zones using latent activity trajectories. IEEE Transactions on Knowledge and Data Engineering, 27(3), 712–725. https://doi.org/10.1109/TKDE.2014.2345405

Yuan, Y., Raubal, M., & Liu, Y. (2012). Correlating mobile phone usage and travel behavior—A case study of Harbin, China. Computers, Environment and Urban Systems, 36(2), 118–130. https://doi.org/10.1016/j.compenvurbsys.2011.07.003

Zhang, A., Li, W., Wu, J., Lin, J., Chu, J., & Xia, C. (2021). How can the urban landscape affect urban vitality at the street block level? A case study of 15 metropolises in China. Environment and Planning B: Urban Analytics and City Science, 48(5), 1245–1262. https://doi.org/10.1177/2399808320924425

Zhang, F., Wu, L., Zhu, D., & Liu, Y. (2019). Social sensing from street-level imagery: A case study in learning spatio-temporal urban mobility patterns. ISPRS Journal of Photogrammetry and Remote Sensing, 153, 48–58. https://doi.org/10.1016/j.isprsjprs.2019.04.017

Zhang, T., Sun, L., Yao, L., & Rong, J. (2017). Impact analysis of land use on traffic congestion using real-time traffic and POI. Journal of Advanced Transportation, 2017(1985), 1–8. https://doi.org/10.1155/2017/7164790

Zhang, X., Du, S., & Wang, Q. (2017). Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data. ISPRS Journal of Photogrammetry and Remote Sensing, 132, 170–184. https://doi.org/10.1016/j.isprsjprs.2017.09.007

Zhang, X., Li, W., Zhang, F., Liu, R., & Du, Z. (2018). Identifying urban functional zones using public bicycle rental records and point-of-interest data. ISPRS International Journal of Geo-Information, 7(12), 459. https://doi.org/10.3390/ijgi7120459

Zhang, Y., Li, Q., Tu, W., Mai, K., Yao, Y., & Chen, Y. (2019). Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. Computers, Environment and Urban Systems, 78(3688), 101374. https://doi.org/10.1016/j.compenvurbsys.2019.101374

Zhang, Y., Li, X., Wang, S., Yao, Y., Li, Q., Tu, W., Zhao, H., Zhao, H., Feng, K., Sun, L., & Hubacek, K. (2021). A global north-south division line for portraying urban development. iScience, 24(7), 102729. https://doi.org/10.1016/j.isci.2021.102729

Zhao, P., & Hu, H. (2019). Geographical patterns of traffic congestion in growing megacities: Big data analytics from Beijing. Cities, 92, 164–174. https://doi.org/10.1016/j.cities.2019.03.022

Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing: Concepts, methodologies, and applications. ACM Transactions on Intelligent Systems & Technology, 5(3), 38. https://doi.org/10.1145/2629592

Zhong, Y., Zhu, Q., & Zhang, L. (2015). Scene classification based on the multifeature fusion probabilistic topic model for high spatial resolution remote sensing imagery. IEEE Transactions on Geoscience and Remote Sensing, 53(11), 6207–6222. https://doi.org/10.1109/TGRS.2015.2435801

How to cite this article: Zhang, Y., & Raubal, M. (2022). Street-level traffic flow and context sensing analysis through semantic integration of multisource geospatial data. *Transactions in GIS*, *26*, 3330–3348. https://doi.org/10.1111/tgis.13005