City-scale roadside electric vehicle parking and charging capacity: A deep learning augmented street-view-image data mining and analytic framework

Yifan Pu [a], Rui Zhu [b,*], Shu Wang [a], Linlin You [c], Teng Zhong [d], Yanqing Xu [e], Zheng Qin [b]

[a] *Department of Geography, National University of Singapore, Singapore 117570, Republic of Singapore*
[b] *Institute of High-Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, Singapore 138632, Republic of Singapore*
[c] *School of Intelligent Systems Engineering, Sun Yat-Sen University, Shenzhen 510006, China*
[d] *Key Laboratory of Virtual Geographic Environment (Ministry of Education of PRC), Nanjing Normal University, Nanjing 210023, China*
[e] *School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China*

HIGHLIGHTS

- Developed a DL-augmented framework to estimate city-scale EV charging capacity.
- Enhanced street-view-image segmentation with transfer learning and refined labels.
- Calculated geo-object areas by developing 3D space geometric projection method.
- Combined GIS and DL to improve parking station planning efficiency and reliability.
- Integrated EV charging records with GIS to estimate roadside charging capacity.

ARTICLE INFO

ABSTRACT

In response to the escalating sales of electric vehicles (EVs), roadside parking and charging have been developed to facilitate EV penetration in many cities. However, its city-scale capacity is usually unknown, hindering effective planning of parking and charging infrastructures. To tackle this problem, we develop a deep learning augmented street-view-image (SVI) data mining and analytic framework, consisting of three hierarchical modules. The first module retrieves geo-locations along roads in the government authorized parking zones (APZs) and obtains SVIs that present both sides of roads centralized at these geo-locations, which is used to identify suitable roadside parking locations. The second module conducts transfer learning that determines a suitable SVI dataset with well-defined classes of interested street-view geo-objects and obtains the optimal DL model capable of refined segmentation of various types of roads. The third module identifies different urban functional zones to suggest locations suitable for roadside parking, develops a 3D space geometric projection method that estimates parking areas in each location, and unravel roadside charging capacity through geospatial statistics of existing EV charging records. As a case study using 55,724 SVIs in Singapore, the IoU of segmented avenues, paths, and sidewalks is as high as 92.51 % to 89.71 %, and we suggest 54,812 roadside parking stalls available from 6761 locations in the APZs, which can support up to 590,315 kWh/day and 5,685,923 kWh/day in the commercial zones and residential zones, respectively. Our study is significant in fundamental geospatial data construction and scaling roadside EV parking and charging in dense urban areas.

1. Introduction

1.1. Background

In recent years, Electric Vehicles (EVs) have met an irreversible trend in increasingly penetrating people's daily life. Three key factors driving this shift are their ability to reduce greenhouse gas emissions by 40 %–60 % over their lifecycle compared to traditional fuel vehicles [1], the significant increase in endurance mileage over the past five years [2], and the decrease of battery costs by 86 % in the last decade, significantly

lowering purchase prices [3]. Additionally, government incentives, such as Singapore's SGD 25,000 "Green Vehicle Rebate", stimulate EV sales [4]. Global EV sales reached 6.6 million vehicles in 2021 – double the sales in 2020 [5], and the expansion of EV charging infrastructure, including over 250,000 public EV Charging Stations (EVCSs) in the European Union as of 2023 with rising fuel prices, promotes the shift towards EVs [6].

As EV adoption accelerates worldwide, the demand for EVCSs has surged, particularly for flexible roadside charging. This flexibility is essential in densely populated urban areas where dedicated charging stations may be limited, addressing the needs of EV users and alleviating range anxiety. With the acceleration of modern life, the convenience offered by roadside EV parking and charging is crucial for enabling EVs to compete with gasoline vehicles in terms of refueling efficiency. The existing roadside parking and charging facilities are insufficient to meet the growing demand, and the distances between them are often inadequate to alleviate range anxiety. Accurately estimating roadside EV charging capacity at a city scale is crucial for promoting efficient roadside EV parking infrastructure. Without fundamental EVCS datasets, the current roadside charging infrastructure falls short by not fully accounting for actual streetscapes and local traffic regulations, leading to inefficient placement and potential regulatory conflicts.

The integration of Geographic Information Science (GIS), Deep Learning (DL), and Street View Images (SVIs) is indispensable in this context [7]. The DL model can utilize SVIs to detect suitable locations for roadside EV parking and charging based on factors such as parking availability and proximity to power sources. Recent research on EVCS site selection has employed various methodologies, including GIS models, Multi-Criteria Decision-Making (MCDM) models, and optimization algorithms [8,9]. MCDM models can introduce bias due to subjective weight assignments, impacting decision accuracy, and subjective judgments may also overlook local traffic legality. Unlike MCDM, this study aims to use cutting-edge technology to provide a comprehensive analysis of roadside EV parking and charging capacity while ensuring site selection complies with traffic regulations by using the data of APZs.

1.2. Geospatial analysis for site selection of EVCSs

Geospatial analysis is an effective approach for planning EVCSs, allowing evaluation of numerous factors crucial to the placement of charging stations, such as proximity to static power infrastructure, road information, traffic patterns, and land availability [10]. By identifying various parking locations based on these spatial indicators, geospatial analysis can strategically determine the optimal ones to effectively serve high-demand areas [11]. Furthermore, a more detailed investigation, which integrates dynamic population density and traffic flow with the distribution of existing charging stations, power network capacity, and land use, can further help identify optimal charging station locations. For instance, a study in Athens, Greece, employed GIS-based multi-criteria analysis to ascertain the best locations for charging stations [12].

In addition, another approach combines geospatial technologies with the Analytic Hierarchy Process (AHP), assigning weights to various factors to select the optimal charging station locations [13]. This method not only considers geospatial data but also incorporates expert opinions and socio-economic factors, thus enhancing the scientific basis for site selection. A relevant study successfully identified optimal locations for EVCSs using GIS and AHP methods [14]. In terms of methodology, the AHP method is effective for planning charging stations. By applying spatial modeling after fundamental GIS data integration using AHP, criteria such as environmental impact, cost-effectiveness, and user demand can be comprehensively evaluated to ensure a well-rounded and rational siting decision [15]. However, these studies face challenges due to a lack of fundamental GIS data regarding EV roadside parking. Specially, identifying roadside EV parking zones solely based on road types without considering the surrounding geo-environment is unreliable and unconvincing. Therefore, effectively tackling this challenge

through the analysis of SVIs and an understanding of the street-view environment is urgently needed.

1.3. Measurement of geo-objects in SVIs

SVIs can be used to obtain detailed information about streets and the associated urban landscape, including road width [16], roadside facilities [17], vehicle parking conditions [18], and pedestrian activities [19], all of which are useful for identifying suitable parking locations. Although these studies have extracted various urban information from SVIs, there remains a lack of research focused on the extraction of specific geo-objects that can affect roadside EV parking. This gap encourages us to integrate this demand into the research framework. Numerous studies have utilized SVIs to estimate the geometry of geo-objects in the real world [20]. These studies often combine depth estimation techniques with pixel counting, analyzing the distribution of pixels in images, along with camera parameters, to accurately infer the real size of geo-objects [21]. This approach has been applied in a wide range of fields, including urban planning [22] and traffic monitoring [23], demonstrating the feasibility and accuracy of geometric measurements based on SVIs.

While there are several robust studies on estimating the size of geo-objects, research that relies on SVIs to estimate the area of specific geo-objects remains limited. This limitation arises from the challenges associated with interpreting SVIs for area estimation. Additionally, the variability in image quality, occlusions, and perspectives in SVIs further complicates the process. In contrast, most existing studies focus on non-SVIs method, where pixel counting techniques, combined with precise camera calibration and known scales, are employed to convert pixel counts into real-world area measurements [24]. Therefore, to suggest roadside locations suitable for EV parking, our study will explore a method that integrates robust semantic segmentation of SVIs with a geometric measurement technique, projecting pixels from two-dimensional (2D) images into three-dimensional (3D) space to calculate the area of geo-objects.

1.4. DL-based semantic segmentation

In the context of semantic segmentation, DL has significantly enhanced the ability to understand and classify every pixel in an image, allowing for more precise interpretation of complex scenes [25]. Recent advancements in infrastructure planning have incorporated smart charging strategies, such as joint models for shared electric vehicles, which optimize charging demand and energy distribution. For example, a study on the electrification of vehicles in Turin highlights the importance of smart charging strategies. The research uses modeling and simulation of EV user behavior to forecast charging scenarios, identify management challenges, and improve EV infrastructure [26]. It applies discrete choice modeling based on socio-economic and transport data to describe user charging behavior geographically, estimating the number of EVs and user characteristics. These insights provide tools to evaluate modifications and indicate an adequate charging network to facilitate EV diffusion. Early models, such as Fully Convolutional Network (FCN), U-Net, and SegNet, enabled end-to-end learning and pixel-level predictions but faced disadvantages due to trade-offs between resolution and efficiency [27]. To address these limitations, advanced models like PSPNet and DeepLabV3+ addressed these issues through multi-scale context aggregation and improved geo-objects representation [28]. For model training, MXNet offers distinct advantages in scalability and deployment [29]. Unlike PyTorch [30], MXNet supports both dynamic and static computational graphs, and its built-in distributed training capabilities make it more efficient for large-scale production across multiple devices. Additionally, MXNet is more memory-efficient, which is beneficial for handling larger models in production environments [31]. Thus, MXNet will be used in our study for loading pre-trained models and tuning the environment for the well-trained model.

In the comparison of three datasets in heatmap segmentation experiments based on DeepLabV3+, the Cityscapes dataset achieved better segmentation results than the SCUT-SEG and SODA datasets [32]. These studies inspired us to conduct a similar approach to compare various semantic segmentation models with control variables to segment SVIs and determine suitable roadside EV parking zones. A GIS-based study conducted EVCSs site selection by analyzing the traffic accessibility of public commercial zone [33]. Although it does not utilize GIS data related to roadside parking, this approach provides valuable insights into actual parking demand and behavioral patterns, supporting the effective deployment of EVCSs in high-demand areas.

Cross-domain transfer learning addresses data scarcity and model generalization challenges in street view image segmentation by leveraging deep convolutional networks to extract features from source and target domains [34]. Models pre-trained on rich source data (e.g., Cityscapes) are fine-tuned on limited target data using domain adaptation techniques like MMD or DANN to reduce domain shift [35]. This approach enhances model robustness and mitigates annotation costs. Applications include transferring knowledge from remote sensing to street view, using synthetic data with domain adaptation, and cross-city street view transfer [36]. Challenges like class imbalance are addressed through weighted loss functions, while multi-scale feature fusion improves object detection across sizes. Practical recommendations involve selecting suitable pre-trained models (e.g., PSPNet, DeepLabV3+), integrating domain adaptation, and using incremental fine-tuning to prevent overfitting [37]. Data augmentation and multimodal fusion further improve accuracy in complex scenes. Future research could explore cross-modal transfer and lightweight deployment for practical applications such as autonomous driving [38].

1.5. Datasets for SVI-based semantic segmentation

Semantic segmentation classifies each pixel in an image into geo-objects like *avenues, buildings, pedestrians, vehicles,* and *trees,* enabling detailed SVI analysis critical for autonomous driving and smart city development [34]. Standard datasets support model training and evaluation. For example, the ADE20K dataset, with over 25,000 images covering 150 types, is widely used for tasks such as semantic segmentation and geo-objects detection [35,36]. However, for complex scenarios, it may be necessary to increase the representation of specific geo-objects in the training tasks to improve their segmentation accuracy. For the Cityscapes dataset, focusing on urban street scenes with 5000 high-resolution images annotated with 30 types, is well-suited for urban traffic tasks but may struggle in more diverse environments [37]. Similarly, the COCO dataset contains over 200,000 images covering 80 types but presents challenges with complex backgrounds and many small geo-objects [38]. Vision Transformers (ViT) have more parameters but similar mIoU and Dice scores to UNet CNN models [39]. ViT outperforms UNet in some metrics [40] but struggles with background categories, limiting its practical use in tasks requiring background-foreground distinction.

CNN-based models like DeepLabV3+ handle background categories better by explicitly classifying background pixels. DeepLabV3+ is chosen for its robust semantic segmentation and modular architecture, which adapts dynamically to road density and zoning regulations. Experiments show it achieves 88 % cross-city deployment accuracy in complex urban settings, far superior to non-adaptive models at 51 % [41]. This makes DeepLabV3+ ideal for urban EV charging infrastructure planning, where accurate background segmentation is crucial. Choosing DeepLabV3+ emphasizes the importance of selecting models not just for overall performance but also for their ability to address specific task challenges, such as background management in semantic segmentation.

CamVid, offering continuous video frames, is ideal for dynamic scene analysis but lacks scene diversity [42]. Instead, Mapillary Vistas, covering diverse street scenes globally, tests the robustness of models

like Panoptic-DeepLabV3+ but lacks pixel analysis [43]. Additionally, the BDD100K dataset, with rich metadata, supports multi-task learning [44]. Since the datasets introduced above demonstrated their effectiveness in urban planning, we will also utilize the publicly available datasets to facilitate the development of semantic segmentation models in transfer learning [45,46]. Meanwhile, this study aims to develop a study-area-specific dataset by comparing general SVI datasets to create a model that accurately segments and identifies parking area within the SVI.

1.6. Contributions

This study presents three key innovations. First, the combination of DL and SVIs enables a comprehensive estimation of roadside parking and charging capacity compared to site selection based solely on fundamental geospatial dataset construction for parking locations. Second, UFZs are identified by analyzing geo-objects from segmented SVIs, providing recommendations for suitable roadside parking. Thirdly, segmentation accuracy is improved by applying transfer learning on pre-trained models using datasets from the study area, enabling a more precise quantification of geo-objects in the segmentation results. Additionally, this study integrates government APZs data with road networks and validates the intersections of APZs with land use and road networks, respectively, to enhance the accuracy of EV charging station selection. It customizes a semantic segmentation model for urban settings and combines potential EVCS data with existing EVCS data to improve demand forecasting and enable the efficient deployment of EVCS infrastructure. Current EV parking and charging estimation methods have significant flaws. GIS-based models lack detailed real-world insights from dynamic urban settings, resulting in overgeneralized capacity assessments. Heuristic planning methods oversimplify spatial complexities, and existing segmentation approaches are inaccurate in quantifying geo-objects due to poor adaptation to local contexts, which limit precision in site selection and infrastructure planning.

2. Estimation of city-scale roadside parking capacity

This study proposed a research framework to suggest appropriate locations for EV roadside parking and charging at the city scale, consisting of three interconnected modules (Fig. 1). The first module develops geospatial analysis and identifies locations that can theoretically be used for roadside parking in the government APZs. The second module (i) performs effective transfer learning that pre-trains several DL models using a selected SVI dataset, which well-defines classes of interested street-view geo-objects, (ii) selects the optimal DL model by comparing segmentation performance of our custom SVIs with dedicatedly prepared labels, and (iii) conducts robust SVI segmentation in the whole study area. The third module (i) categories SVIs into different UFZs based on geo-objects to recommend different principles for calculating roadside parking areas, (ii) estimates roadside parking areas and the number of parking stalls by developing a 3D space geometric projection model, and (iii) quantifies roadside EV charging capacity through geospatial statistics of real EV charging records in each district of the city. Fig. 2 illustrates the overall research diagram.

2.1. Study area and datasets

Singapore is in southeast Asia, close to the equator. Recently, the Singapore government has initiated a plan to gradually phase out petrol and diesel vehicles, ceasing new diesel vehicles registrations by 2025 and permitting only EV registrations by 2030, with a target for full adoption of cleaner energy vehicles by 2040 [3]. In this context, the Singapore EV market was valued at 63.03 million USD in 2022 and is projected to reach 650.92 million USD by 2030, reflecting a compound annual growth rate of 34.4 % [4]. To meet the continuously rising EV charging demand, the government plans to install 60,000 EV charging
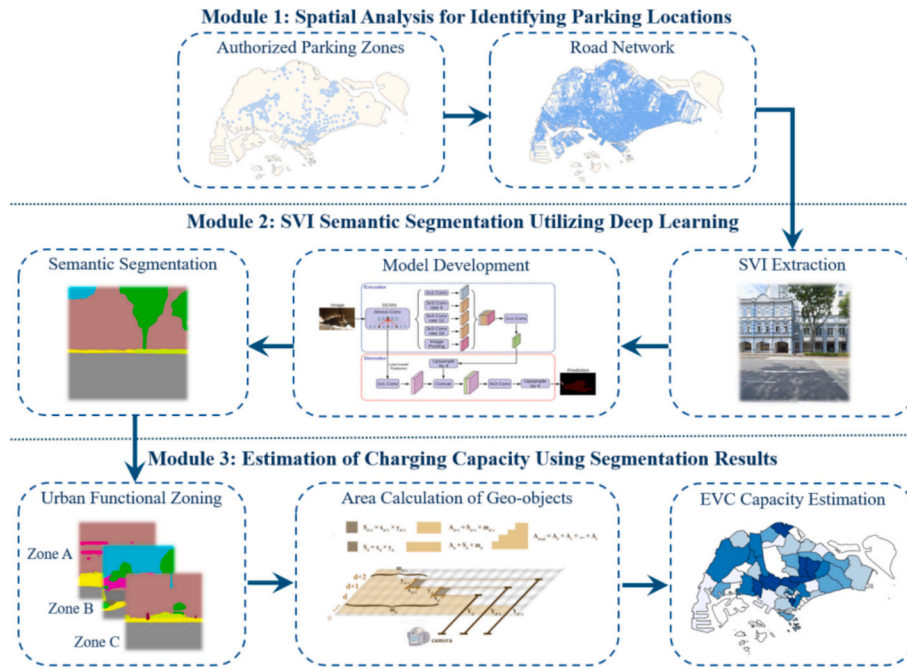
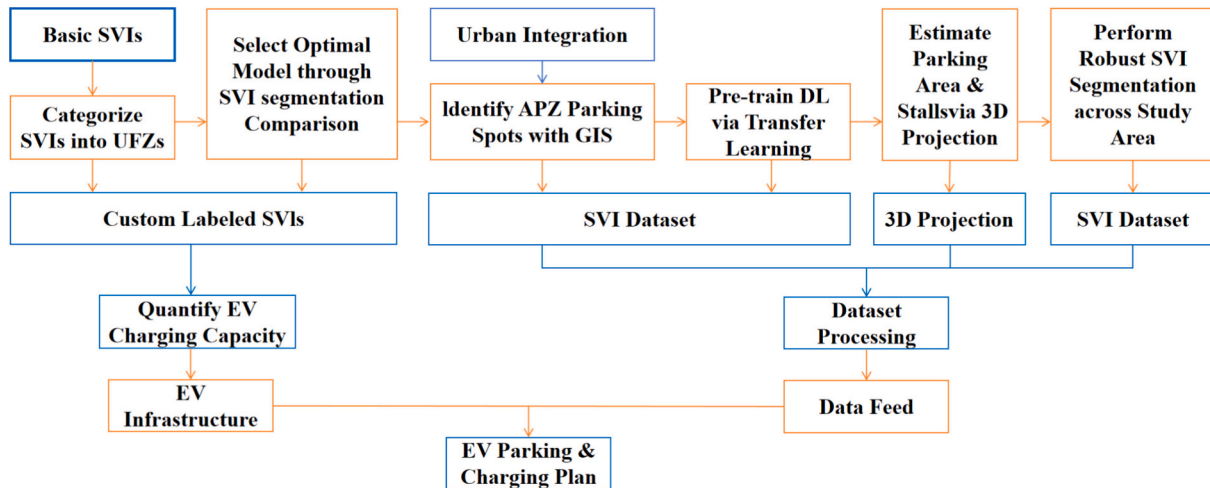**Fig. 1.** Research framework for estimating roadside parking potential.

Fig. 2. Workflow for estimating roadside parking potential.

stations by 2030, an increase from the current 3600 in 2022 [5]. As a high-density city-state with heavy traffic and limited land, Singapore presents a unique case for exploring roadside EV parking capacity [6], which makes Singapore is an ideal study area for this research, offering valuable insights for other cities.

We collected the latest road network data (polylines) and land use data (polygons) in Singapore from OpenStreetMap [9]. The government APZs (polygons) were obtained from the Urban Redevelopment Authority (URA) of Singapore [47]. We purchased Google Cloud API [48] and downloaded Google SVIs on both sides of the road at 90 degrees to the road directions. The dimensions were set to 400 pixels by 400 pixels, with a 90-degree field of view (FOV) [49]. The pitch angle was 0 degrees, and the camera was shot in a horizontal position without tilting up or down. This wide angle allowed the image to capture a large area and gather a broad range of environmental information across the horizontal plane. To evaluate the roadside EV charging capacity, we also collected existing EV charging records from the government data portal to explore the roadside EV charging capacity. The dataset comprised 193

charging stations, with records updated every 30 min from November 7, 2023, to April 17, 2024. Each record included information on the charging location, time, and power of electricity charged (Fig. 3).

2.2. Identification of suitable locations allowing roadside parking

We consider that roadside parking is only allowed in the government APZs. Thus, road networks (polylines) within the APZs (polygons) are maintained. Then, we extract all points along the maintained road networks at every 50 m interval to balance between details and redundancy [50,51]. Since some points are duplicated at the same location or too close to each other, we refine the set of points by removing these duplications. One this basis, SVIs centralized at the refined set of locations are collected to investigate suitable roadside for EV parking and charging. All the generated points that are along the road networks and fall within the government APZs represent the preliminary locations for further investigation. When points have the same coordinates, only one is retained, and when two points on the adjacent polylines of road
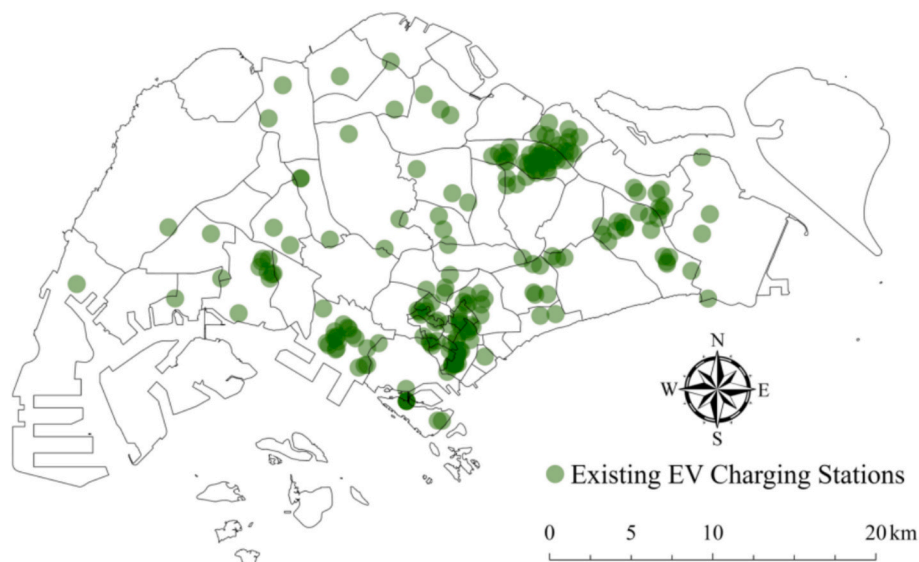
Fig. 3. Locations of existing EV charging stations.

networks are shorter than 50 m, one of the two is kept. Fig. 4 illustrates that locations (green dots) along road networks and in the APZs will be investigated to identify the parking suitability, while those (red dots) outside the APZs will not be filtered out for further consideration. As a result, we collected 55,724 SVIs in Singapore, which consists of 13,522 SVIs in the APZs and 42,202 SVIs out of the APZs.

2.3. Transfer learning

Transfer learning is an effective approach to facilitate an accurate segmentation of geo-objects from SVIs. To obtain a pre-trained DL model that is suitable for segmenting a variety of geo-objects particularly related to roads (e.g., *pedestrians*, *motorways*, and *footpaths*), four major steps were processed. First, 100 SVIs with a high presence of specific geo-objects were selected from those located within APZs. These SVIs were meticulously annotated at the pixel level using the VGG Image Annotator tool, labeling 19 types of SVI geo-objects in each image, such as avenues, buildings, pedestrians, vehicles, and trees. Second, rigorous quality checks were conducted to ensure annotation consistency. Third, to enhance dataset diversity and prevent model overfitting, data

augmentation techniques were applied to the SVIs, including adjustments to brightness and contrast. The images were then standardized and resized to 400 × 400 to ensure uniform input. Fourth, DeepLabV3+ was retained as the DL model, with pre-trained models from ADE20K [52] and Cityscapes [53] applied to segment the 100 SVIs (90 % for training and 10 % for validation). This approach allowed us to identify the optimal SVI dataset for transfer learning.

It is also imperative to select a suitable DL model for transfer learning and accurate segmentation of all collected SVIs in the study area. Here, we utilize three well-known DL models (i.e., DeepLabV3+ [25], PSPNet [26], and FCN [27]), which have been pre-trained on the optimal SVI dataset, to perform semantic segmentation of the 100 well-labeled SVIs in Singapore (90 % for training and 10 % for validation). When the optimal DL model is selected, all SVIs in the APZs are segmented.

Regarding model configurations, the Adam optimizer is used along with a learning rate scheduler to progressively reduce the learning rate and improve model stability [54], and Dropout is applied to reduce the risk of overfitting during training [55]. The Cross-Entropy Loss function, suitable for addressing class imbalance in semantic segmentation tasks, is used to measure the performance of the model, and the output was a
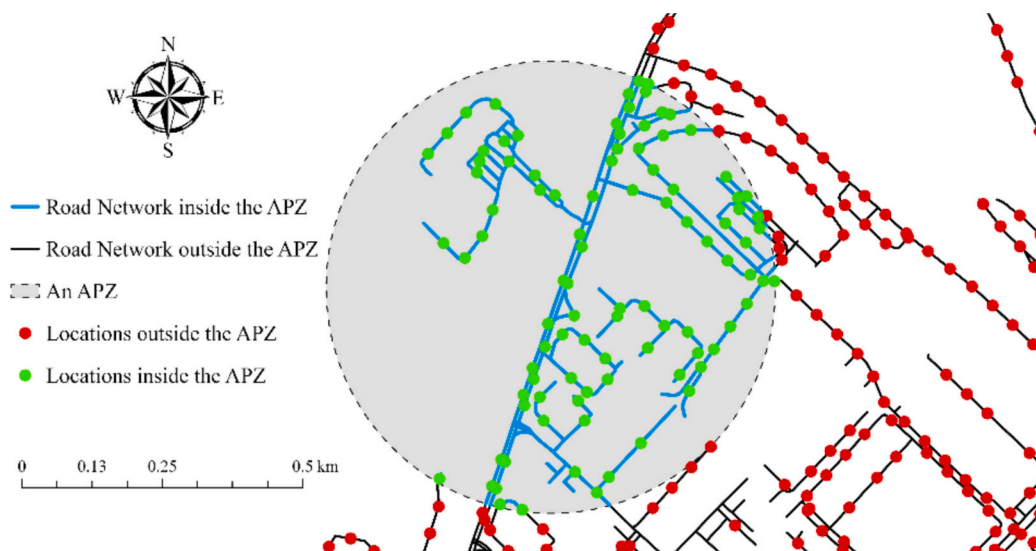

Fig. 4. Selection of locations on road networks in the APZs.

probability value between 0 and 1. The function is defined in Eq. (1), where N is the number of samples, C is the number of classes, $y_{i,c}$ is the ground truth label, and $\widehat{y}_{i,c}$ is the predicted probability for class c for sample i.

$$Loss = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c} log\left(\widehat{y}_{i,c}\right) \tag{1}$$

2.4. Accuracy assessment indicators

The metrics of IoU, Precision, Recall, Accuracy, and F1-score are used to systematically evaluate the segmentation accuracy. Specifically, the *IoU* is calculated in Eq. (2), where *TP* is the number of true positives, *FP* is the number of false positives, and *FN* is the number of false negatives, *TN* is the number of true negatives. *Precision* is defined in Eq. (3), *Recall* is defined in Eq. (4), and *Accuracy* is calculated in Eq. (5). The *F1-score* is defined in Eq. (6), combining *Precision* and *Recall* into a single metric that balances the trade-off between them.

$$IoU = \frac{TP}{TP + FP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

2.5. Identification of UFZs

Based on the semantic segmentation results of SVIs which have been identified allowing roadside parking, we estimated roadside parking capacity in all UFZs. To start with, we defined shared geo-objects which are commonly seen in different UFZs, including *buildings, avenues, vehicles,* and *trees.* Meanwhile, we defined unique geo-objects, which only appear in a specific UFZ, such as residential *mailboxes* which can only appear in the residential zone and *signboards* which are most probably built in the commercial zone. If an SVI contains all shared geo-objects, then it can be classified based on the unique geo-objects contained. Notably, if an SVI lacks any of these shared geo-objects, the corresponding urban areas are categorized as Unidentified Zones and excluded from further classification. This approach ensured that only SVIs containing all shared geo-objects were used to identify UFZs, thereby minimizing the risk of misclassification due to incomplete shared characteristics. Table 1 illustrates the information above.

In a study area characterized by dense and diverse urban functions, the same geo-objects across different UFZs may serve different purposes. For example, a *sidewalk* should not be used for parking in the commercial zones having heavy traffic flows, while it can be used for parking in the residential zones in certain periods. Therefore, the more diverse the dataset classification, the less likely the model was to overlook the

Table 1
Definition of UFZs.

Type of zones	Shared geo-objects	Unique geo-objects
Commercial zone		trade, signboard
Clearway zone	building, house, vehicle, avenue, sidewalk, path, tree, streetlight, person, bin, grass	traffic lights, skyscraper, bus path, fence, mailbox
Residential zone		

functional diversity of the same geo-object. We first calculated the shared road-related geo-objects for each UFZ. It was important to note that, regardless of whether the geo-objects are shared or unique, we must determine their parking availability within a specific UFZ before calculating the area.

2.6. Estimation of roadside parking areas

2.6.1. Regression of distance scales in a 3D space

A geometric measurement method is developed to estimate the available roadside parking areas from segmented SVIs. This is achieved by projecting 2D images into a 3D space, and using standardized geo-objects, such as avenues, pedestrians, zebra crossing, and camera distance to estimate the varying scales of investigated geo-objects. Specifically, a 3D space consists of three axes: x, y, and z. As shown in Fig. 5, it illustrates the parameters represented by x and z in reality and in the SVI, respectively. (The y-direction will be discussed separately after x and z due to its different computational principles.) If a geo-object touches the ground at the d pixel from the bottom of the SVI, its horizontal size in the x direction and vertical size in the z direction can be calculated using its real-world dimensions based on Eqs. (7) and (8):

$$x_{width} = \frac{x_{real}}{x_{pixel}} \tag{7}$$

$$z_{height} = \frac{z_{real}}{z_{pixel}} \tag{8}$$

where x_{width} is the width of one pixel, x_{real} represents the real width of the geo-object, and x_{pixel} represents the number of pixels that the geo-object occupies in the horizontal direction of the image. Similarly, z_{height} is the height of one pixel, z_{real} is the real height of the geo-object, and z_{pixel} calculation method is similar to x_{width}, we could fit them and derive Eqs. (9) and (10), where x_d and z_d represent x_{width} and z_{height} at the pixel position d respectively. Due to x increases with the increase of d, when $d = 0$, $x > 0$, the fitted equations require $a_x > 0$, $b_x > 0$, $c_x > 0$ and $a_z > 0$, $b_z > 0$, $c_z > 0$.

$$x_d = a_x \times d^{b_x} + c_x \tag{9}$$

$$z_d = a_z \times d^{b_z} + c_z \tag{10}$$

In addition to the calculation based on the above method, the real distance Y from a known building to the camera (the sampling point) can be measured directly. Based on the given the pixel position d, Y and d then be fitted in Eq. (11), where Y_d represents the real distance from the camera to the geo-object when it is at d. Note that the increasing rate of Y_d is getting faster nonlinearly with the increase of d.

$$Y_d = a_y \times d^{b_y} + c_y \tag{11}$$

Therefore, the fitting parameter here needs $a_y > 0$, $b_y > 1$, $c_y > 0$. To get the varying distance of a single pixel on the y, further calculation was required based on Eqs. (12) and (13):

$$Y_{d+1} = a_y \times (d+1)^{b_y} + c_y \tag{12}$$

$$y_d = Y_{d+1} - Y_d \tag{13}$$

During the formula fitting process, 80 % of the collected data are used to fit exponential regression and the remaining 20 % data are tested. After applying the fitting method described above, calculations in Eqs. (14), (15) and (16) get the real distance in the real world.

$$x_d = 7.6400 \times 10^{-11} \times d^{5.2955} + 14.95 \tag{14}$$

$$z_d = 2.4754 \times 10^{-11} \times d^{5.5840} + 13.07 \tag{15}$$

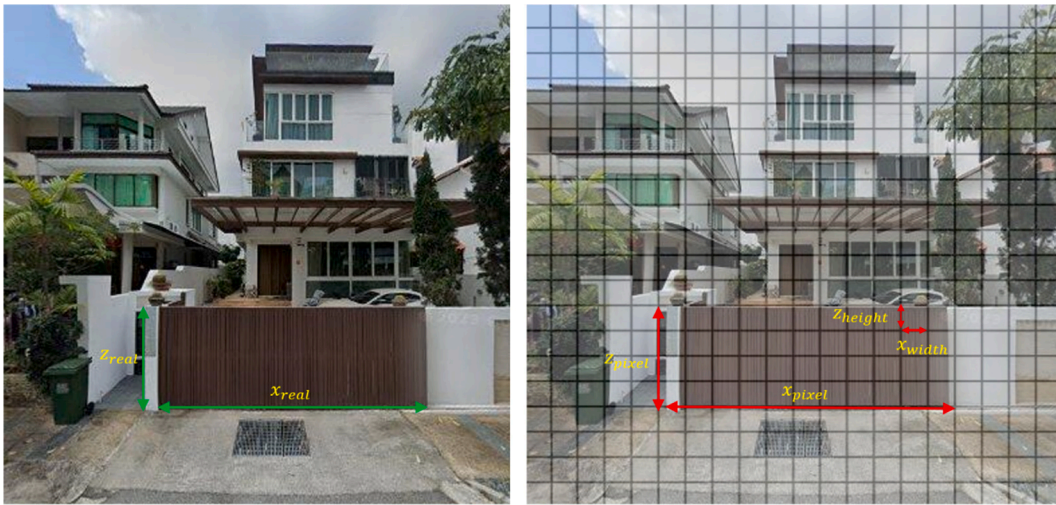$$Y_d = 1.6199 \times 10^{-12} \times d^{7.1093} + 3096.67 \tag{16}$$

Fig. 5. The principle of calculating the area of geo-objects based on pixel count. (a) Original SVI. (b) Calculation principle.

For the area calculation of Geo-objects in 3D space, we considered errors in the X, Z, and Y directions (Table 2). In the X direction, data filtering via residual screening and a standard train-test split (80 % training, random seed 42) ensured high model stability. The power function model achieved an MAE of 3.85 mm, MAPE of 6.47 %, and R^2 of 0.9886, with a confidence interval width of ± 2.5 mm (68 % confidence level). In the Z direction, using the complete dataset with an 80 %–20 % split (seed 80) resulted in an R^2 of 0.9409 and MAE of 8.17 mm, with a confidence interval width of ± 5.8 mm. In the Y direction, model parameters were preset based on theoretical assumptions, and two test datasets showed alignment with fitted curve trends. Future work includes exploring data screening strategies for the X direction, introducing stratified sampling or weighted fitting for the Z direction, and combining actual data with covariance analysis for the Y direction to enhance model adaptability and performance.

Overall, the X direction demonstrated efficient data screening and modeling capabilities, the Z direction showed tolerance for complex data, and the Y direction provided an initial validation basis for the theoretical framework. Targeted optimizations, such as adjusting sample selection thresholds for the X direction, enhancing noise robustness for the Z direction, and supplementing actual data fitting for the Y direction, could lead to higher precision in each direction.

2.6.2. Area calculation of geo-objects with polygonal coordinates

Since the segmented areas of *avenues* in the SVIs are confined to the ground plane with minimal vertical variation, indicating insignificant extension in the direction of z axis of 3D space, we define the height of the *avenues* as $z = 0$. As shown in the SVI segmentation example (Fig. 6), the left side represents a comparison between the original SVI and the segmentation result, while the right side illustrates the breakdown of the SVI into pixels. The yellow section represents the semantic segmentation result for the geo-object labeled *avenues* and every SVI used in the calculation is captured with the *avenues* as the starting point.

The area is calculated using an accumulation method. First, we take a row, d, within the segmentation result of geo-object labeled *avenue* as an example. The distance from the camera to the geo-object represented at that point is the real-world distance Y_d, and y_d is the difference between Y_{d+1} and Y_d, which represents the length of the one pixel at d [42–45]. Combined with the horizontal length of the one pixel x_d, we calculate the real area of one pixel for row d. x_d is unaffected by depth as it is not along the photographing direction of SVIs, while y_d is affected by depth. Next, if the row d contains m_d pixels, the area represented by that row is the sum of the areas of the m_d pixels. The total area of the *avenue* in the SVI, A_{total}, is the cumulative sum of the areas for each row. The calculation is presented in Eqs. (17), (18) and (19):

$$S_d = x_d \times y_d \tag{17}$$

$$A_d = S_d \times m_d \tag{18}$$

$$A_{total} = \sum_{n=0}^{i} A_n \tag{19}$$

3. Results

3.1. Statistics of potential roadside parking locations

Using overlapping analysis between land use types and road networks in the APZs, it is found that 39 types of land uses (Table 3) and 33 types of road networks (Table 4) may exhibit roadside parking. Among these, 36 land use types and 31 road network types can support roadside parking. Specifically, 99.56 % of *residential* land use and 85.36 % of *commercial* land use can be utilized, indicating its high capability in supporting roadside parking. By contrast, the value is much lower for *industrial*, at only 34.87 %, and land use such as *playground*, *proposed* and *utility* exhibit no parking capacity, suggesting significant constraints in roadside parking. In the other perspective, 98.87 % of *service* road networks can be used for roadside parking, and road networks such as

Table 2
Comparison of the 3D-axis.

| 3D-Axis	Data Strategy	Source of Model Parameters	MAE (mm)	MAPE (%)	Confidence Interval Width	Optimization Direction
X	Low residual screening + random split	Fitted from training set	3.85	6.47	± 2.5 mm	Balance screening strategy with generalization
Z	Full data retention + random split	Fitted from training set	8.17	12.11	± 5.8 mm	Explore noise suppression and stratified modeling
Y	Theoretical assumption guidance	Preset parameters + to be verified	To be supplemented	To be supplemented	Theoretical reference value	Combine measured data to refine parameters and covariance analysis
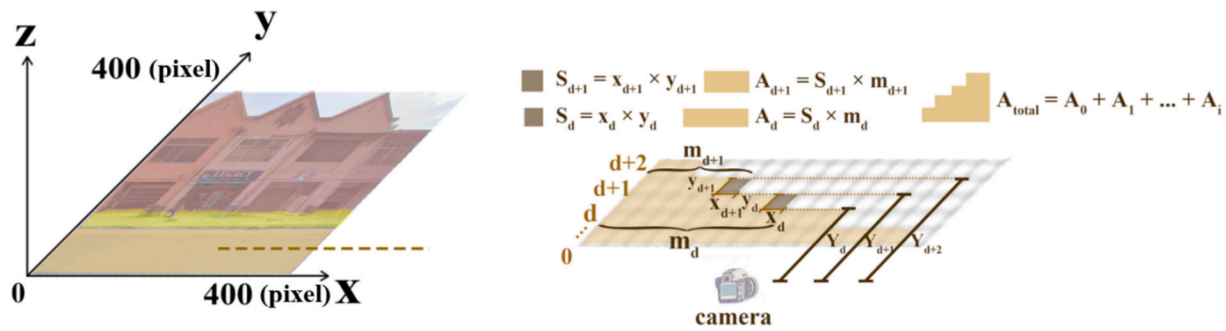
Fig. 6. Area calculation of geo-objects based on SVI.

Table 3
Statistics of land use areas in the APZs per km^2.

Type of land use	Area in APZs (m^2)	Share in all study areas (%)	Type of land use	Area in APZs (m^2)	Share in all study areas (%)
residential	86,135.00	99.56	cemetery	882.85	92.25
grass	56,357.00	75.52	basin	771.14	8.75
construction	32,076.00	98.36	village	675.56	90.00
industrial	12,996.00	34.87	brownfield	518.70	75.00
commercial	9070.00	85.36	railway	458.72	97.25
meadow	8797.00	90.05	healthcare	107.04	95.88
religious	6715.00	2.36	greenhouse	66.08	82.50
retail	5639.00	15.75	reservoir	35.27	0.70
recreation	2541.00	9.94	government	14.92	5.00
military	2257.00	2.20	institutional	8.27	98.00
garages	2093.00	94.52	mixed	5.29	1.25
plant	2049.00	2.25	public	5.24	67.50
allotments	1678.00	4.45	quarry	2.85	100
farmyard	1597.00	84.25	station	1.27	100
greenfield	1554.00	95.55	churchyard	1.20	100
farmland	1072.00	10.00	depot	1.12	100
flowerbed	921.50	10.23	garage	1.00	100
aquaculture	920.27	10.00	playground	0	0
education	920.28	79.69	proposed	0	0
			utility	0	0

Table 4
Statistics of road distance in the APZs per km.

Type of road network	Length in APZs (m)	Share in all study areas (%)	Type of road network	Length in APZs (m)	Share in all study areas (%)
service	3269.16	98.87	trunk	19.25	16.67
footway	1571.00	12.56	pedestrian	15.21	83.33
residential	596.27	8.13	secondary with link	15.89	12.75
primary	501.72	61.36	corridor	10.07	92.26
secondary	434.74	95.82	tertiary with link	7.04	90.15
tertiary	337.78	3.38	proposed	6.08	85.73
steps	192.75	7.75	living street	4.27	97.50
unclassified	130.13	88.76	elevator	4.28	75.00
construction	81.87	96.43	raceway	2.94	50.00
motorway with link	68.47	95.75	busway	2.72	75.00
restway	52.24	90.85	bridleway	2.68	50.00
cycleway	51.86	16.37	bus stop	2.52	97.50
motorway	46.57	84.78	rest area	2.46	90.15
primarily with link	45.75	3.38	lane	2.21	100
path	28.87	25.76	channel	1.97	8.25
track	23.42	10.15	crossing	0	0
			no stopping	0	0

primary demonstrate a modest parking potential of 61.36 %, suggesting a need for improved design to enhance usability. In contrast, types of road networks such as *crossing* and *no stopping* exhibit no parking capacity. As a result, we generated 27,862 locations along the entire road networks, then we obtained 7431 locations inside the APZs that can be theoretically used for roadside parking. After denoising (i.e., removing duplicate coordinates which are within 50 m), 6761 locations were retained as potential parking options for further investigation. It was found that a large number of these locations are concentrated in densely populated districts in the northern (Fig. 7(a)) and southern (Fig. 7(b)) parts of Singapore.

3.2. Comparison of transfer learning results

We successfully segmented *sidewalk*, using DeepLabV3+ pre-trained on ADE20K (Fig. 8a), which, however, was not achieved when it was pre-trained on Cityscapes (Fig. 8b). Thus, ADE20K was ultimately selected for training and evaluation. On this basis, the segmentation results demonstrate that DeepLabV3+ achieved the highest accuracy (Table 5), outperforming PSPNet and FCN regarding all accuracy indicators. Particularly, DeepLabV3+ outperforms PSPNet and FCN in 19 segmented geo-objects in terms of mIoU, Precision, Recall, Accuracy and F1-Score (Table 6). Although PSPNet performs better than FCN in mIoU and Precision, it only surpasses FCN by approximately 1 % in Recall.

We suggest that three geo-objects (*avenue*, *path*, and *sidewalk*) are important fundamentals to estimate roadside parking. Table 5 also represents satisfactory results for the three geo-objects, with DeepLabV3+ getting the largest IoU as high as 98.16 %, 97.84 %, and 97.75 % (The bolded parts in the table). respectively. DeepLabV3+ has a Recall 7.7 % higher than PSPNet and 8.6 % higher than FCN. It also has an Accuracy 6.76 % higher than PSPNet and 10.41 % higher than FCN. The F1-Score of DeepLabV3+ is 7.65 % higher than PSPNet and 10.28 % higher than FCN. This means that, after manually annotating these geo-objects in SVIs, the model can adapt to specific street-view environments in Singapore, highlighting the effectiveness of the proposed transfer learning. Furthermore, as shown in Fig. 9, the segmentation results of the DeepLabV3+ (Fig. 9 (b)) model for *path* closely aligned with the SVIs (Fig. 9 (a)), whereas PSPNet (Fig. 9 (c)) and FCN (Fig. 8 (d)) produce fragmented segmentation for *path*, exhibiting significant errors.

3.3. Segmentation and classification of SVIs in Singapore

The above results confirm that DeepLabV3+ pre-trained on the ADE20K is the optimal DL model for our SVI segmentation task, which is thus used for segmenting SVIs of the locations in APZs. The performance of the well-trained model on the test set is presented in Table 7. Based on the segmentation results and the 19 types of shared and unique geo-objects, geo-objects presented in SVIs are classified into different UFZs (Table 8). We also recorded the proportions of the SVIs in different UFZs (Table 8), which is approximately 72.1 % in the residential zone, 15.8 % in the commercial zone, 6.1 % in the clearway zone, 6.0 % in the
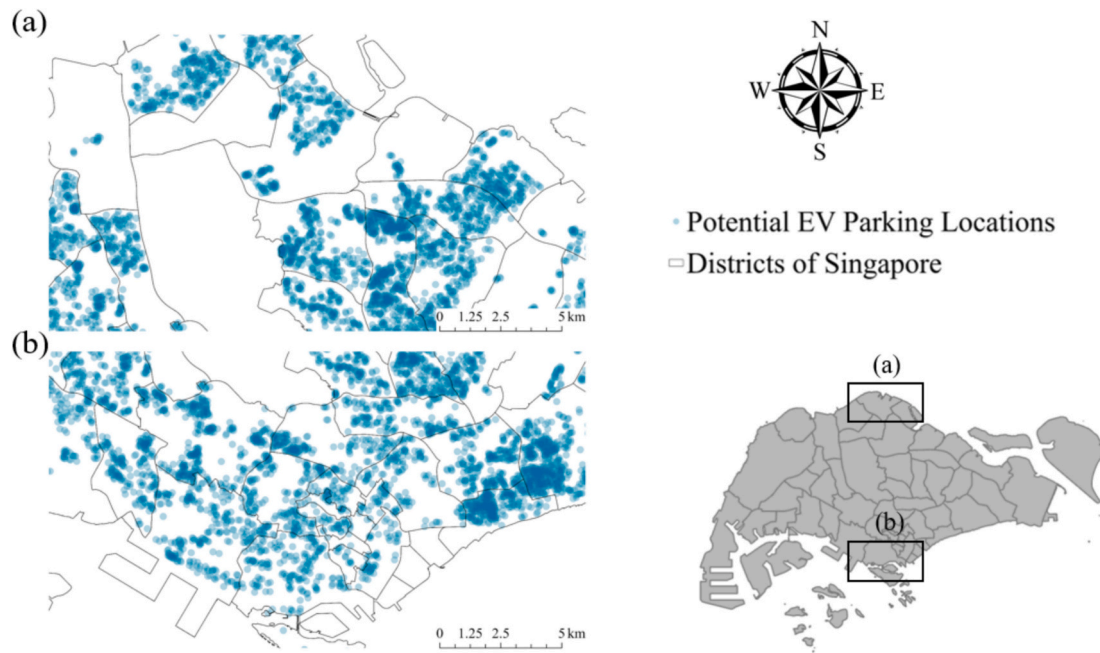
Fig. 7. Potential roadside parking locations. (a) Northern part of Singapore. (b) the Southern part of Singapore.
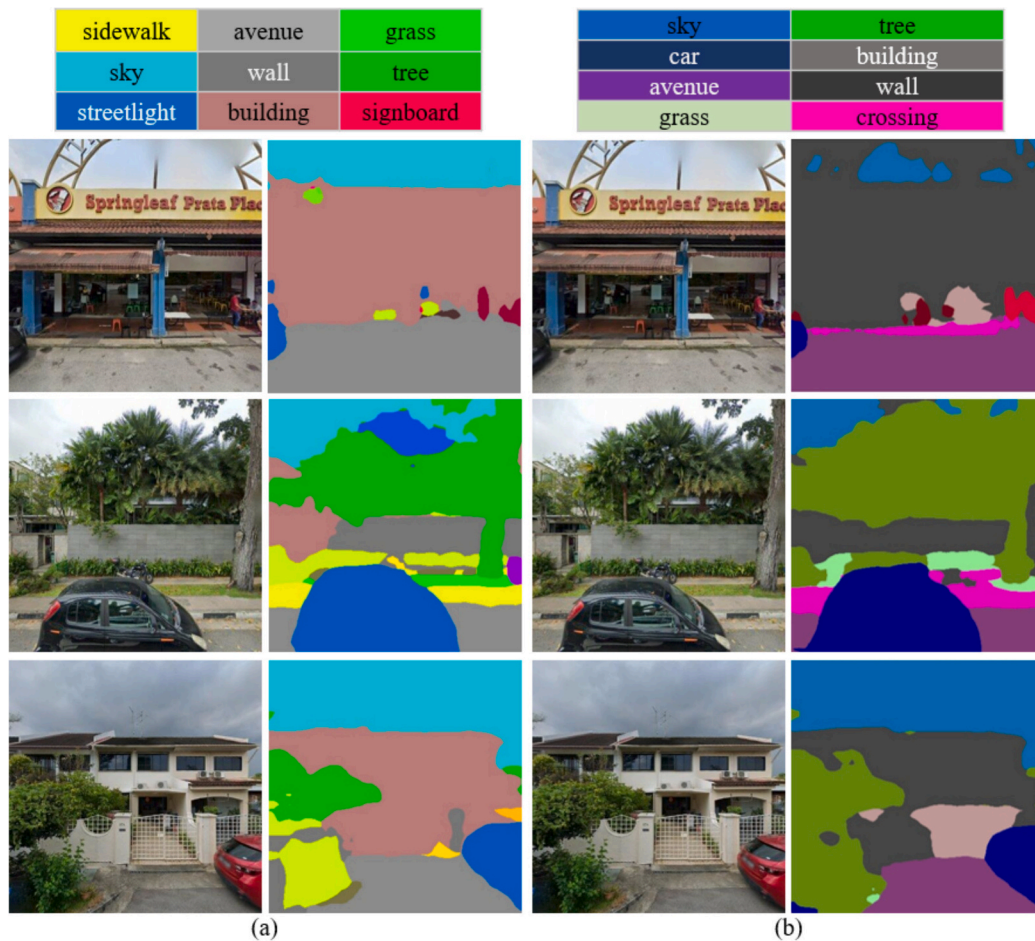

Fig. 8. Semantic segmentation using DeepLabV3+. (a) ADE20K. (b) Cityscapes.

Table 5
The overall segmentation accuracy of the 100 SVIs in Singapore, using DL models pre-trained on ADE20K.

DL models	IoU (%)	Precision (%)	Recall (%)	Accuracy (%)	F1-Score (%)
DeepLabV3+	62.80	82.10	86.20	73.07	84.10
PSPNet	48.10	74.50	78.50	66.31	76.45
FCN	39.60	70.40	77.60	62.66	73.82

unidentified zone that cannot be used for analyzing roadside parking such as mosaic and nighttime SVIs.

It is found that geo-objects in the commercial zone have a high concentration of trade-related elements, such as *signboards*, indicating active commerce (Fig. 10(a)). The clearway zone includes urban geo-objects like *streetlights*, *bus*, highlighting its role as a major transportation thoroughfare (Fig. 10(b)), whereas the residential zone comprises *fences*, representing primarily housing areas (Fig. 10(c)). The unidentified zone, due to poor visibility of nighttime and mosaic SVIs, means that it does not have any valid geo-objects. According to the district-based statistics, we found that the distribution of appropriate roadside parking locations exhibits different distribution characteristics across various UFZs. In the commercial zones (Fig. 11(a)), they are primarily distributed in the southern and eastern districts; in clearway zones (Fig. 11(b)), those roadside parking locations primarily distribute in the southwestern and southeastern districts; in residential zones (Fig. 11(c)), they are predominantly distributed in the eastern and western districts, with the eastern districts having significantly more than the western ones, and with a noticeably larger quantity compared to the other three UFZs. By contrast, unidentified zones (Fig. 11(d)) are primarily distributed in the middle eastern and middle western districts.

3.4. Estimation of the number of roadside parking stalls

According to the regulation issued by URA [44], any temporary or long-term parking on all types of roads is prohibited in the clearway zone, which is thus excluded from further estimation. In the commercial zone, parking on the *sidewalk* is also illegal; therefore, only the area of *avenue* and *path* is considered. Since the residential zone does not have any explicit constraints, *sidewalk* and *avenue*, and *path* areas are usable for roadside parking. Next, for each SVI, the total usable parking areas are summarized. According to the URA standard [45], a parking stall is 2.4 m by 4.8 m, occupying an area of 11.52 m^2. As a result, Fig. 12

presents the estimated number of roadside parking stalls in each district in Singapore from two types of UFZs: commercial zones (Fig. 12(a)) and residential zones (Fig. 12(b)), resulting in a total of 54,812 parking stalls across the entire study area. Overall, the estimated roadside parking stalls for both residential zones and commercial zones are mainly concentrated in the southern and eastern districts. However, the parking stalls in residential zones are significantly more than those in commercial zones.

3.5. Roadside EV charging capacity

To estimate the EV charging demand (kWh/day) across Singapore, we first assume that the number of roadside parking stalls in each district equals the number of EV charging piles in that district [56,57]. Using existing charging records, we calculate the EV charging demand per district. Subsequently, we divide the EV charging demand per district by the number of EV charging piles in each district to obtain the average EV charging demand per pile for each district. While calculating the existing charging records (kWh/day) for Singapore's total of 55 districts, 9 districts were excluded due to no roadside parking stalls, resulting in no charging demand for those areas.

As shown in Fig. 13, our results suggest that Singapore's charging demand primarily avoids two major reservoirs and several industrial islands (indicated by the white areas in Fig. 13). The estimated charging demand in commercial zones is mainly concentrated along the southern and southeastern coast, while avoiding areas near the Changi airport and offshore islands, while the residential charging demand extends from west to east. Both coastal residential zones and commercial zones exhibit high demand for EV charging. In comparison, the EV charging demand in commercial zones aligns with the distribution characteristics of the estimated number of EV charging piles (Fig. 13(a)), whereas the estimated roadside charging demand in residential zones shows an additional high demand in the northern districts (Fig. 13(b)). The total electricity demand in the commercial zone is 590,315 kWh/day, while the total electricity demand in the residential zone is 5,685,923 kWh/day.

Several additional constraints should be considered for practical implementation. Firstly, roadside EV parking and charging in high-density urban areas during peak hours may cause congestion [58]. Three strategies can be proposed to mitigate the potential adverse effects: (i) implement spatiotemporal pricing incentives based on real-time traffic data to discourage peak-hour charging, as validated in Singapore's ERP system [59]; (ii) integrate EV charging schedules with grid

Table 6
Accuracy of segmented geo-objects of the 100 SVIs in Singapore[#].

Geo-object	IoU (%)			Precision (%)			Recall (%)			Accuracy (%)			F1-Score (%)		
	DL-1	DL-2	DL-3	DL-1	DL-2	DL-3	DL-1	DL-2	DL-3	DL-1	DL-2	DL-3	DL-1	DL-2	DL-3
avenue	98.16	97.46	97.04	97.71	96.53	95.36	91.51	88.52	81.53	86.96	85.91	84.87	94.51	92.35	87.9
path	97.84	96.96	96.46	94.1	95.96	93.82	89.89	87.04	80.2	83.75	85.4	83.5	91.95	91.28	86.48
sidewalk	97.75	96.88	96.21	92.26	92.13	92.11	88.71	85.99	80.08	82.11	82	81.98	90.45	88.95	85.67
building	96.46	95.8	95.11	92.89	93.82	90.75	87.86	85.02	78.54	82.67	83.5	80.77	90.31	89.2	84.2
house	95.41	93.18	93.06	91.35	92.1	90.1	87.25	85.12	76.25	81.3	81.97	80.19	89.25	88.47	82.6
vehicle	93.43	92.76	91.76	88.94	91.37	89.96	85.32	84.89	74.81	79.16	81.32	80.06	87.09	88.01	81.69
stoplight	93.1	91.8	90.46	91.41	90.38	89.34	81.39	80.29	72.5	81.35	80.44	79.51	86.11	85.04	80.04
skyscraper	91.59	90.22	88.38	86.23	90.01	88.21	78.88	75.12	70.36	76.74	80.11	78.51	82.39	81.89	78.28
tree	91.24	88.36	87.89	85.33	86.19	86.39	76.29	75.76	68.24	75.94	76.71	76.89	80.56	80.64	76.25
streetlight	84.55	79.08	78.64	80.84	81.64	77.32	73.03	72.8	65.57	71.95	72.66	68.81	76.74	76.97	70.96
person	79.79	78.05	76.31	78.36	74.29	70.09	70.46	71.65	61.83	69.74	66.12	62.38	74.2	72.95	65.7
bus	76.73	75.06	73.39	73.05	67.97	62.74	68.77	69.11	62.25	65.01	60.49	55.84	70.85	68.54	62.49
bin	76.85	75.17	73.5	72.33	67.12	61.75	68.23	68.06	65.9	64.37	59.74	54.96	70.22	67.59	63.76
grass	75.25	73.61	71.97	72.05	66.79	61.36	67.36	65.63	62.9	64.12	59.44	54.61	69.63	66.2	62.12
trade	71.35	70.9	69.32	71.2	62.16	55.98	65.86	64.04	62.26	63.37	55.32	49.82	68.43	63.09	58.95
signboard	72.48	70.28	68.71	68.16	58.79	52.06	64.81	63.81	58.76	60.66	52.32	46.33	66.44	61.2	55.21
wall	71.84	70.05	68.45	65.32	56.6	51.74	62.36	58.29	52.08	58.13	50.37	46.05	63.81	57.43	51.91
fence	71.5	69.04	68.38	62.5	55.44	48.16	61.45	55.97	51.5	55.63	49.34	42.86	61.97	55.7	49.77
mailbox	64.74	63.33	61.92	59.61	52.00	44.16	57.88	54.73	50.58	53.05	46.28	39.3	58.73	53.33	47.15

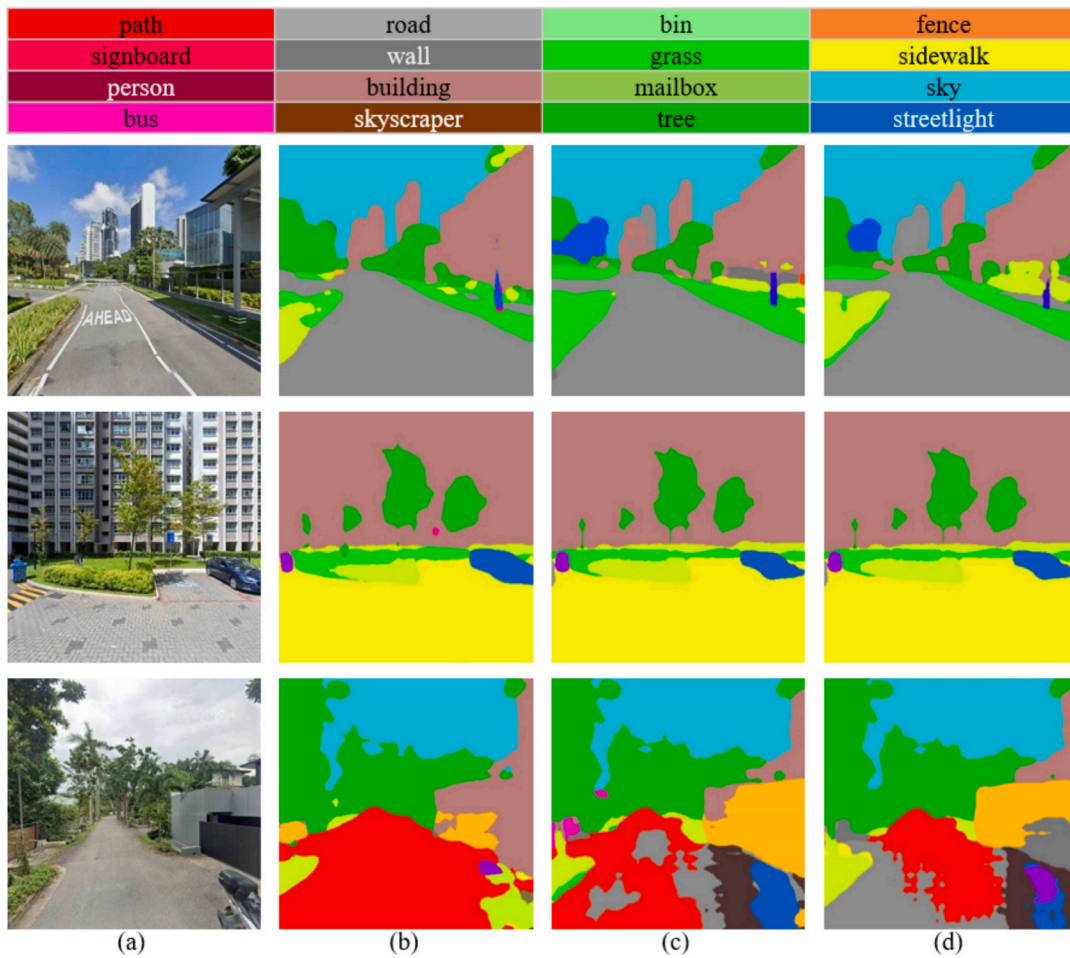[#] DL-1, DL-2, and DL-3 refer to DeepLabV3+, PSPNet, and FCN, respectively pre-trained on ADE20K.

Fig. 9. Semantic segmentation using the pre-trained ADE20K dataset. (a) Custom SVIs in Singapore. (b) DeepLabV3+. (c) PSPNet. (d) FCN.

Table 7
Segmentation accuracy of SVIs in Singapore using the pre-trained DeepLabV3 +
.

Geo-object	IoU (%)	Precision (%)	Recall (%)	Accuracy (%)	F1-Score (%)
avenue	92.51	98.71	99.16	87.85	98.93
path	90.89	95.10	98.84	84.64	96.93
sidewalk	89.71	93.26	98.75	83.00	95.93
building	88.86	93.89	97.46	83.56	95.64
house	88.25	90.35	96.41	80.41	93.28
vehicle	86.32	89.94	94.43	80.05	92.13
stoplight	82.39	90.41	94.10	80.46	92.22
skyscraper	79.88	87.23	92.59	77.63	89.83
tree	77.29	86.33	92.24	76.83	89.19
streetlight	72.03	81.84	85.55	72.84	83.65
person	69.46	79.36	80.79	70.63	80.07
bus	67.77	74.05	77.73	65.90	75.85
bin	66.23	73.33	77.85	65.26	75.52
grass	66.36	73.05	76.25	65.01	74.62
trade	64.86	72.20	72.35	64.26	72.27
signboard	63.81	69.16	73.48	61.55	71.25
wall	61.36	66.32	72.84	59.02	69.43
fence	60.45	63.50	72.50	56.52	67.70
mailbox	56.88	60.61	63.74	53.94	62.14

Table 8
Results of UFZ identification.

Type of zones	Number of SVIs	Proportion (%)
Residential zone	4876	72.10
Commercial zone	1067	15.80
Clearway zone	415	6.10
Unidentified zone	403	6.00

balancing algorithms can be developed to optimize energy usage and reduce grid stress, as demonstrated in Tokyo [62].

Secondly, to refine planning locations, we need to address safety, accessibility, and regulatory requirements, such as adequate space, compliance with safety standards, and proximity to urban infrastructure. These factors require detailed studies beyond initial estimations, and our results aim to guide urban planning with interval-based recommendations. Thirdly, balancing fast and slow chargers is crucial for different urban needs. Future work will collect data on charger power and usage patterns to develop more reliable allocation strategies and conduct cost-benefit analyses to optimize deployment. Additionally, applying the framework to other cities may face challenges from incomplete GIS datasets. We suggest using crowd-sourced data from platforms like Mapillary (https://www.mapillary.com/) or integrating real-time traffic data with GIS to identify high-demand areas. Machine learning algorithms can forecast demand for dynamic adjustments, and continuous monitoring will improve model accuracy and planning efficiency.

load forecasts through time of use charging coordination to shift peak demand to off-peak periods, as shown in California's SGIP program [60]; (iii) co-locate fast-charging stations with public transport nodes to reduce intra-urban charging trips, following the Berlin S-Bahn model [61]. Additionally, reinforcement learning-based dynamic load
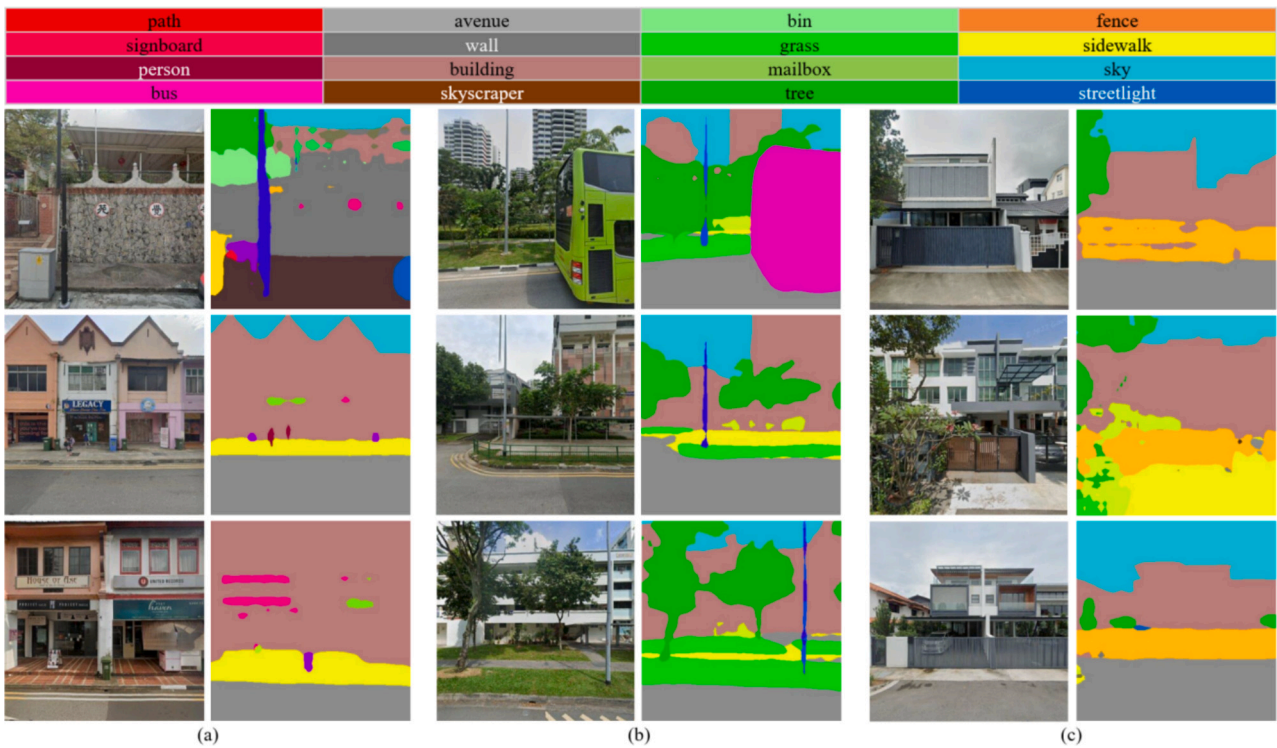
Fig. 10. Segmentation results of UFZs. (a) Commercial zones. (b) Clearway zones. (c) Residential zones.
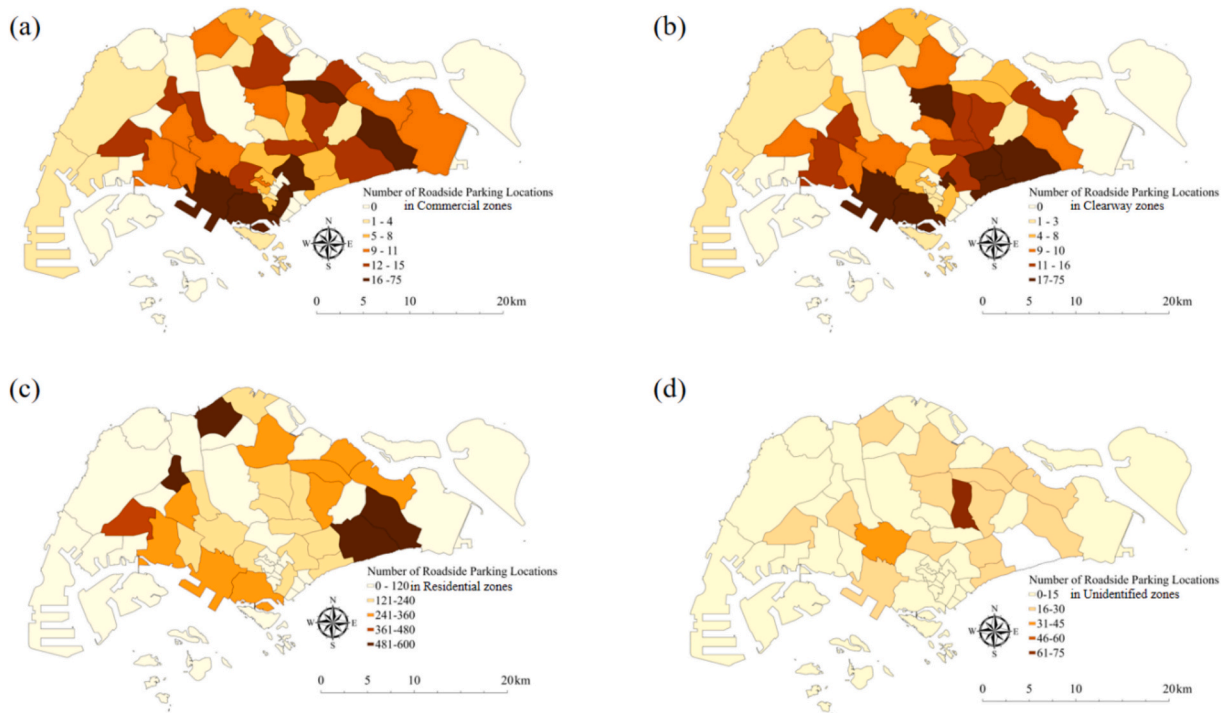

Fig. 11. The number of locations to be used for investigating roadside parking in four different UFZs. (a) Commercial zones. (b) Clearway zones. (c) Residential zones. (d) Unidentified zones.

4. Discussion

The developed framework consists of three interconnected modules, integrating GIS data, DL-enhanced SVI data, and 3D geometry projection. The framework provides a robust approach for identifying optimal EV charging locations. Our findings suggest that Singapore has a total of 54,812 suitable roadside parking stalls, which aligns closely with the government's goal of installing 60,000 EV charging piles by 2030, indicating great potential to plan city-scale EV roadside parking and charging infrastructures.

This study presents several key impacts. First, it introduces an innovative study framework that combines policy regulations (APZs)
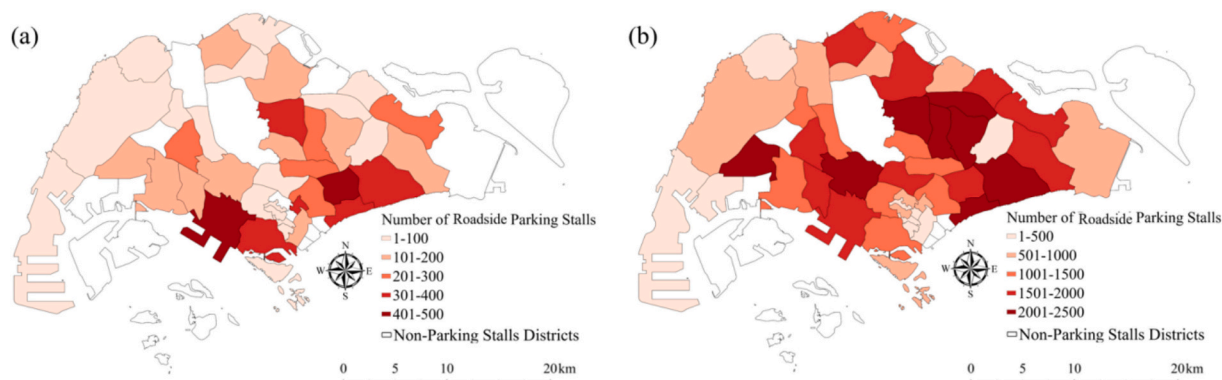
Fig. 12. The number of estimated roadside parking stalls in Singapore. (a) Commercial zones. (b) Residential zones.
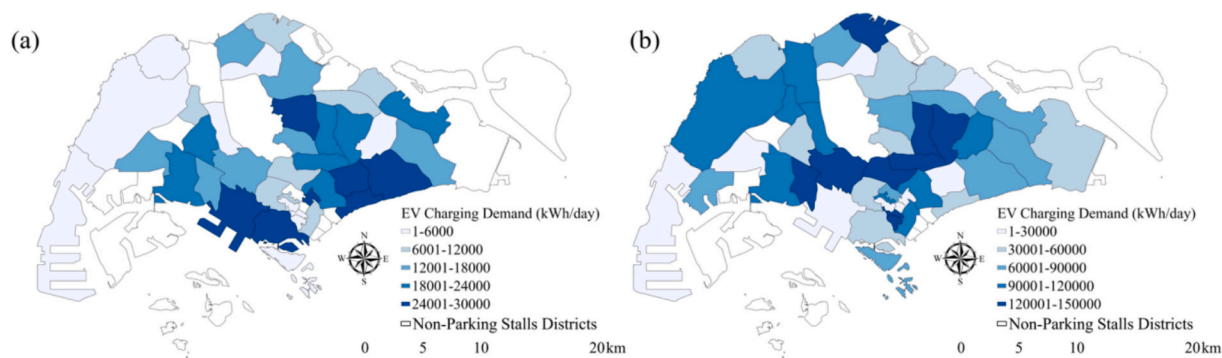

Fig. 13. The estimated roadside daily mean EV charging demand per district. (a) Commercial zones. (b) Residential zones.

with geospatial analysis to construct fundamental geospatial datasets. This approach not only aids in understanding socioeconomic phenomena on urban streets but also enhances data availability for various urban applications, such as traffic management, land use planning, and optimizing urban infrastructure layout. Moreover, the SVI-based method outperforms traditional resource-intensive surveys, significantly reducing costs and time, offering a practical, data-driven solution for resource-limited cities. Our approach can support flexible and convenient roadside parking, promote city-scale vehicle electrification, and provide urban planning insights for governments and policymakers. This enables policymakers to precisely identify high-demand EV charging areas, which is crucial for addressing urban challenges such as the mismatch between the number of parking stalls and the EV charging demand in a district, whether it is undersupply or oversupply, ensuring a smoother transition to an electrified transportation system.

Second, for compact study areas like Singapore, characterized by a complex combination of different UFZs, we achieved more accurate recognition of various geo-objects in SVIs compared to using general DL datasets by leveraging effective transfer learning and selecting the optimal semantic segmentation model. Additionally, by integrating existing charging records, we provided accurate estimations of roadside parking areas to ensure that our urban EV charging capacity estimations reflect real-world conditions. This directly contributes to creating actionable insights for cities with high EV density and complex urban layouts, ensuring that EV infrastructure deployment aligns with actual demand and operational feasibility.

Third, the framework presented in this study is scalable and can be applied to other cities. Since the tools used in this study, such as OpenStreetMap and Google Cloud API, have global coverage, and MXNET is open source, the framework can be extended to any city that can access GIS data and SVI images through these sources. By applying the 3D geometry projection method introduced in this study and adapting MXNET for DL as needed, it is also possible to determine EV

charging demand for other cities. Considering the generalization capability of the developed framework, the major concern is the availability of geospatial data and SVIs. Given the flexible nature of the methodology, the framework can be easily applied to other cities by using similar data infrastructure, making it a practical solution for promoting global vehicle electrification.

Nevertheless, our study contains three limitations. First, due to the manpower limitation, we carefully selected and manually labeled a small number of 100 representative SVIs in Singapore to train a custom DL model. The segmentation accuracy and UFZ identification can be further improved if more labeled SVIs are available. Second, considering Singapore has tiny terrain variation, we assume that all types of roads like avenue and sidewalk are flat. Yet, a more accurate parking area estimation through 3D geometry projection needs to model the deformation along the z-axis in SVIs. This is especially unignorable when applying the model to other cities presenting considerable terrain variation (e.g., Hong Kong and Bergen). Third, we do not consider dynamic roadside parking allocation (i.e., allowing parking during certain dates and periods), which has been implemented in many cities including Singapore. Future study can address this issue to meet the increasing parking and charging demand by incorporating detailed policy analysis.

To improve prediction reliability across diverse regions, we will develop localized models by fine-tuning pre-trained models with location-specific data, including traffic dynamics, seasonal demand shifts, and urban infrastructure traits. This approach enhances adaptability while maintaining computational efficiency. For rural or data-scarce areas, we will leverage alternative data streams such as satellite imagery, aggregated mobile data, and publicly contributed inputs to compensate for limited infrastructure records. We will redesign models to accommodate rural features like decentralized transportation networks and sparse energy grids. Additionally, we will develop lightweight pre-processing pipelines and transfer learning strategies for low-resource region settings. We will also explore synergies with emerging

technologies, including vehicle-to-grid (V2G) integration for smarter station placement, LiDAR-SVI fusion for geometric precision, and regulatory constraint modeling to enhance real-world feasibility.

5. Conclusion

To conclude, this study develops a scalable research framework that couples DL and GIS technologies to estimate city-scale roadside EV parking stalls and charging capacity. The results demonstrate that DL-based segmentation and 3D geometry projection can effectively simulate real-world measurements, significantly improving the efficiency of geospatial data construction. Our study highlights a promising opportunity, particularly in densely populated areas such as Singapore, for deploying roadside EV parking and charging infrastructures. The pre-trained model not only evaluates a city's overall energy efficiency based on EV charging demand but also provides valuable guidance for other cities in formulating sustainable strategies for EV charging station deployment. Future work could (i) apply the framework to rural or less-developed areas to address sparse infrastructure, (ii) collaborate with policymakers and government agencies to refine deployment strategies and enhance the framework's capability for sustainable mobility, and (iii) integrate dynamic data (e.g., near real-time traffic flows, charging demand, and PV generation) to provide adaptive charging scheduling services.

CRediT authorship contribution statement

Yifan Pu: Writing – original draft, Visualization, Software, Resources, Methodology, Formal analysis, Data curation. **Rui Zhu:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Shu Wang:** Conceptualization. **Linlin You:** Supervision. **Teng Zhong:** Writing – review & editing, Conceptualization. **Yanqing Xu:** Writing – review & editing, Supervision. **Zheng Qin:** Writing – original draft, Supervision.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

[1] Seddig K, Meyer-Hübner N, Jochem P, Fichtner W. Two-stage stochastic optimization for cost-minimal charging of electric vehicles at public charging stations with photovoltaics. Appl Energy 2019;240:49–61.
[2] Wang Y, Zhang C, Zhao T, Zhao H. Siting and sizing of fast charging stations in highway network with budget constraint. Appl Energy 2018;225:449–62.
[3] Luo L, Yang G, Hu Z, Song Y. Joint planning of distributed generation and electric vehicle charging stations considering real-time charging navigation. Appl Energy 2019;238:1–10.
[4] Luo L, Wu T, Zhang C. Optimal planning of electric vehicle charging stations comprising multi-types of charging facilities. Appl Energy 2018;229:407–18.
[5] Sadeghi-Barzani P, Rajabi-Ghahnavieh A, Kazemi A. Optimal fast charging station placing and sizing. Appl Energy 2014;125:289–99.
[6] Abbasi MH, Marzbani F, Rashidinejad M, Van Oort G. Coordinated operation of electric vehicle charging and wind power generation as a virtual power plant: a multi-stage risk constrained approach. Appl Energy 2019;238:214–28.
[7] Quddus MA, Yang Z, Ai X, Wu Q. A collaborative energy sharing optimization model among electric vehicle charging stations, commercial buildings, and power grid. Appl Energy 2018;229:1215–27.
[8] Wang W, Zhang Y, Zhang X, Li H. Linking energy-cyber-physical systems with occupancy prediction and interpretation through WiFi probe-based ensemble classification. Appl Energy 2019;239:102–15.
[9] Huang P, Li H, Zhou S, Zhang X. Uncertainty-based life-cycle analysis of near-zero energy buildings for performance improvements. Appl Energy 2018;229:468–78.
[10] Huang P, Li H, Zhou S, Zhang X. A clustering-based grouping method of nearly zero energy buildings for performance improvements. Appl Energy 2019;237:193–203.
[11] Wang C, Chen J, Zou J. CO2 mitigation scenarios in China's road transport sector. Energy Convers Manag 2007;48:2110–8.
[12] Weiss M, Dekker P, Moro A, Scholz H, Patel MK. On the electrification of road transportation–a review of the environmental, economic, and social performance of electric two-wheelers. Transp Res Part D: Transp Environ 2015;41:348–66.
[13] Borozan S, Neugebauer S, Psarras C. Strategic network expansion planning with electric vehicle smart charging concepts as investment options. Adv Appl Energy 2022;5:100087.
[14] Bartłomiejczyk M, Kalina T, Połom M. The reduction of auxiliaries' power demand: the challenge for electromobility in public transportation. J Clean Prod 2020;258:120879.
[15] Gallet M, Luethi G, Molinaro A, Wietlisbach M, Leuba M. Estimation of the energy demand of electric buses based on real-world data for large-scale public transport networks. Appl Energy 2018;230:344–56.
[16] Gao Z, Lin Z, LaClair TJ. Battery capacity and recharging needs for electric buses in city transit service. Energy 2017;122:588–600.
[17] Bi Z, Song L, De Kleine R, Mi CC, Keoleian GA. Plug-in vs. wireless charging: life cycle energy and greenhouse gas emissions for an electric bus system. Appl Energy 2015;146:11–9.
[18] Liu L, Wang Z, Zhang W, Xu G, He J. A review on electric vehicles interacting with renewable energy in smart grids. Renew Sust Energ Rev 2015;51:648–61.
[19] Oh M, Kim D, Kim S, Yi Y. Estimation of photovoltaic potential of solar bus in an urban area: case study in Gwanak, Seoul. Korea Renew Energy 2020;146:13–22.
[20] Vur Rehman N, Park H, Kim H, Han M, Alam MM. Solar potential assessment of public bus routes for solar buses. Renew Energy 2020;146:2159–71.
[21] Alam M, Wang JF, Guangpei C, Yunrong L, Chen Y. Convolutional neural network for the semantic segmentation of remote sensing images. Mobile Networks and Appl 2021;26:146–59.
[22] Kemker R, Salvaggio C, Kanan C. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS J Photogramm Remote Sens 2018;145:60–77.
[23] Liu T, Yang X. Monitoring land changes in an urban area using satellite imagery, GIS, and landscape metrics. Appl Geogr 2015;56:53–63.
[24] Longbotham N, Chaapel C, Bleiler L, Padwick C, Emery WJ, Pacifici F. Very high resolution multiangle urban classification analysis. IEEE Trans Geosci Remote Sens 2012;50(4):1155–70.
[25] Cui G, Zhang L. Improved faster region convolutional neural network algorithm for UAV target detection in complex environments. Results in Engineering 2024;67(5):155–70.
[26] Sica L, Deflorio F. Estimation of charging demand for electric vehicles by discrete choice models and numerical simulations: application to a case study in Turin. Green Energy and Intelligent Trans 2023;2(2):100069.
[27] Xie M, Ruan J, Bai W, Qiao Q, Bai L, Zhang J, et al. Pollutant payback time and environmental impact of Chinese multi-crystalline photovoltaic production based on life cycle assessment. J Clean Prod 2018;184:648–59.
[28] Energy Market Authority. Installed Capacity & Number of Installations. 2023.
[29] Hason MM, Abbood IS, Odaa SA. Land cover reflectance of Iraqi marshlands based on visible spectral multiband of satellite imagery. Results in Engineering 2020;12(4):130–5.
[30] Kamga GAF, Bitjoka L, Akram T, Mbom AM, Naqvi SR, Bouroubi Y. Advancements in satellite image classification: methodologies, techniques, approaches and applications. Int J Remote Sens 2021;42(20):7662–722.
[31] Zhou B, Zhao H, Puig X, Xiao T, Fidler S, Barriuso A, et al. Semantic understanding of scenes through the ADE20K dataset. Int J Comput Vis 2019;127(3):302–21.
[32] Zhang C, Sargent I, Pan X, Li H, Gardiner A, Hare J, et al. Joint deep learning for land cover and land use classification. Remote Sens Environ 2019;221:173–87.
[33] Neupane B, Horanont T, Aryal J. Deep learning-based semantic segmentation of urban features in satellite images: a review and meta-analysis. Remote Sens 2021;13(4):808–9.
[34] Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. Int J Comput Vis 2018;137(2):357–73.
[35] Malof JM, Hou R, Collins LM, Bradbury K, Newell R. Automatic solar photovoltaic panel detection in satellite imagery. In 2015 International Conference on Renewable Energy Research and Applications. 2015. p. 1428–31.
[36] Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. Int J Comput Vis 2017;128(3):288–303.
[37] Gibson GM, Johnson SD, Padgett MJ. Single-pixel imaging 12 years on: a review. Opt Express 2020;28:128–90.
[38] Edgar MP, Gibson GM, Padgett MJ. Principles and prospects for single-pixel imaging. Nat Photonics 2018;13(1):13–20.
[39] Stojek R, Pastuszczak A, Wrobel P, Kotynski R. Single pixel imaging at high pixel resolutions. Opt Express 2022;30:22730–45.
[40] Phillips DB, Sun MJ, Taylor JM, Edgar MP, Barnett SM, Gibson GM, et al. Adaptive foveated single-pixel imaging with dynamic supersampling. Sci Adv 2017;3(4):15–7.
[41] Weakly semi supervised extraction of rooftop photovoltaics from high-resolution images based on segment of anything model and class activation map. Appl Energy 2024;343:121170.
[42] ETHOS.PASSION: an open-source workflow for rooftop photovoltaic potential assessments from satellite imagery. Sol Energy 2023;260:115506.
[43] Mapping of decentralised photovoltaic and solar thermal systems by remote sensing aerial imagery and deep machine learning for statistic generation. Energy and AI 2023;11:100236.
[44] A transformer-based multimodal-learning framework using sky images for ultra-short-term solar irradiance forecasting. Appl Energy 2023;341:121033.

[45] PVNet: a novel semantic segmentation model for extracting high-quality photovoltaic panels in large-scale systems from high-resolution remote sensing imagery. Int J Appl Earth Obs Geoinf 2023;120:103621.

[46] Segmenter Rooftop PV. A size-aware network for segmenting rooftop photovoltaic systems from high-resolution imagery. Remote Sens 2023;15(3):738.

[47] Downscaling of nighttime light imagery with a spatially local estimation model using human activity-physical features. Int J Appl Earth Obs Geoinf 2024;123:104688.

[48] Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS. Deep learning for visual understanding: a review. Neurocomputing 2016;187:27–48.

[49] He X, Wu Y, Sun H, Zhang Z. Using open data and deep learning to explore walkability in Shenzhen, China. Transp Res Part D: Transp Environ 2023;124:103601.

[50] Wani MA, Mujtaba T. Segmentation of satellite images of solar panels using fast deep learning model. Int J Renew Energy Res 2021;11(1):31–45.

[51] He N, Yang J, Helbich M, Guan Q, Liu Y. Urban neighborhood environment assessment based on street view image processing: a review of research trends. Environmental Challenges 2021;3:100045.

[52] Helbich M, de Beurs D, Kwan MP, O'Connor D, Groenewegen PP. Outdoor light at night, air pollution, and depressive symptoms: a cross-sectional study in the Netherlands. Sci Total Environ 2020;706:135975.

[53] Eskew J, Ratledge M, Wallace M, Gheewala SH, Rakkwamsuk P. An environmental life cycle assessment of rooftop solar in Bangkok. Thailand Renewable Energy 2018;123:781–92.

[54] Todde G, Murgia L, Deligios PA, Hogan R, et al. Energy and environmental performances of hybrid photovoltaic irrigation systems in Mediterranean intensive and super-intensive olive orchards. Sci Total Environ 2019;651(2):2514–23.

[55] Zhu R, Wong MS, You L, Santi P, Nichol J, Ho HC, et al. The effect of urban morphology on the solar capacity of three-dimensional cities. Renew Energy 2020;153:1111–26.

[56] Park HS, Jeong K, Hong T, Ban C, Koo C, Kim J. The optimal photovoltaic system implementation strategy to achieve the national carbon emissions reduction target in 2030: focused on educational facilities. Energ Buildings 2016;119:101–10.

[57] de Wild-Scholten MM. Energy payback time and carbon footprint of commercial photovoltaic systems. Sol Energy Mater Sol Cells 2013;119:296–305.

[58] Li P, Zhang H, Guo Z, Lyu S, Chen J, Li W, et al. Understanding rooftop PV panel semantic segmentation of satellite and aerial images for better using machine learning. Adv Appl Energy 2021;4:100057.

[59] Malof JM, Collins LM, Bradbury K, Newell RG. A deep convolutional neural network and a random forest classifier for solar photovoltaic array detection in aerial imagery. In 2016 IEEE International conference on renewable energy research and applications. 2016. p. 650–4.

[60] Yuan J, Yang HHL, Omitaomu OA, Bhaduri BL. Large-scale solar panel mapping from aerial images using deep convolutional networks. In 2016 IEEE International Conference on Big Data. 2016. p. 2703–8.

[61] Marimuthu C, Kirubakaran V. Carbon payback period for solar and wind energy project installed in India: a critical review. Renew Sust Energ Rev 2013;23:80–90.

[62] Guo Z, Lu J, Chen Q, Liu Z, Song C, Tan H, et al. TransPV: refining photovoltaic panel detection accuracy through a vision transformer-based deep learning model. Appl Energy 2024;355:122282.