The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-G-2025
ISPRS Geospatial Week 2025 “Photogrammetry & Remote Sensing for a Better Tomorrow…”, 6–11 April 2025, Dubai, UAE

DroneVision: A Low-Code Solution for Urban Parking Occupancy Detection Using Vision-Language Models

Meng Jin, Melanie Handrich, Bernd Bienzeisler

Fraunhofer Institute for Industrial Engineering IAO, Heilbronn, Germany
(meng.jin, melanie.handrich, bernd.bienzeisler)@iao.fraunhofer.de

Keywords: Parking Occupancy Detection, Aerial Image Analysis, Computer Vision, Vision-Language Models (VLMs), Smart City Solutions.

Abstract

Accurate detection of on-street parking occupancy is crucial for urban traffic management. We present DroneVision, a low code solution that enables municipalities to analyze parking occupancy using drone-acquired aerial imagery and vision language models (VLM). This approach allows municipal authorities to upload commercial drone imagery and parking space information, automatically processes these inputs, and generates comprehensive analysis reports. We evaluated two state-of-the-art open-source VLMs against traditional pretrained convolutional neural networks (CNNs) for drone imagery analysis. Results demonstrate the effectiveness of our VLM-based approach, offering a scalable solution that can be integrated into smart city infrastructures while making advanced AI technology accessible to municipal authorities through a low-code interface.

1. INTRODUCTION

The challenges of urban parking management have grown increasingly complex with rapid urbanization and the surge in vehicle ownership across cities worldwide (Mavlutova et al., 2023)(Ceder, 2021). Efficient monitoring and management of on-street parking spaces have become crucial components of urban planning and traffic management systems (Arora et al., 2019). Current approaches to monitoring parking occupancy can be categorized into several main categories, each with its own limitations and challenges.

Static sensor-based systems have demonstrated their capability in precise parking occupancy detection. For instance, San Francisco implemented 8,200 in-ground geomagnetic sensors for monitoring on-street parking spots in pilot areas (SFpark, 2018). However, these solutions face significant scalability challenges due to their high installation and maintenance costs, limiting their widespread adoption across cities (Chen et al., 2020) (Lin et al., 2017) (Šolić et al., 2020). Static camera systems represent another traditional approach to parking monitoring, typically focusing on specific outdoor parking spaces at a small scale (Gören et al., 2019). While effective for individual locations, these systems require extensive infrastructure for comprehensive coverage, leading to substantial implementation costs (Amato et al., 2017). Moreover, the deployment of widespread camera networks faces varying legal restrictions across different countries, particularly concerning data protection and privacy regulations.(Pannerselvam, 2021)(Mahaarachchi et al., 2023) (Shivaprasad et al., 2024) Crowdsensing-based solutions have emerged as an alternative, utilizing smartphone sensors (Nawaz et al., 2013) (Bock et al., 2019) or on-vehicle sensors (Bazzaza et al., 2024) (Liao et al., 2022) to monitor parking availability. However, these approaches introduce additional costs, such as monetary incentives for participant recruitment, and create dependencies on existing parking systems. Similarly, parking occupancy inference based on payment transactions, while leveraging existing infrastructure and complementary data sources (traffic flow, population, POIs), proves inadequate for non-paid residential parking areas and fails to account for discrepancies between payment duration and actual parking behavior (Zhao et al., 2021) (Assemi et al., 2021).

Given these limitations, drone technology presents a promising alternative for urban parking analysis. Drones offer distinct advantages: they don’t occupy road infrastructure and can efficiently capture aerial imagery of large urban areas in relatively short timeframes (Telli et al., 2023) (Mohsan et al., 2023). While subject to weather conditions, government regulations, and privacy considerations, drones have proven valuable in various traffic-related studies, including road density estimation (Kim et al., 2023), and vehicle emissions analysis (Barmpounakis et al., 2021). Previous research has explored drone-based parking occupancy detection, particularly for high-density off-street parking scenarios (Kim et al., 2024). However, these approaches typically rely on extensive manual labeling for training traditional deep learning networks and require substantial AI expertise, limiting their practical adoption by municipal authorities. The technical complexity of implementing and maintaining such systems creates significant barriers for government agencies seeking to deploy these solutions.

To address these limitations, we present DroneVision, a low-code platform that leverages vision-language models (VLMs) (Zhu et al., 2023)(Bai et al., 2023) to analyze parking occupancy from drone-acquired aerial imagery. The platform enables municipalities to automatically process aerial images and generate comprehensive reports without requiring technical expertise. The integration of VLMs marks a significant advancement over conventional computer vision approaches, as traditional convolutional neural network (CNN)-based methods rely heavily on extensive training data and require fine-tuning of specific model architectures (Zhang et al., 2024). Our VLM-based approach offers greater flexibility and robustness in interpreting aerial parking scenarios, while providing valuable comparative insights against traditional CNN approaches. We evaluated our approach through a comprehensive case study in a residential area of Karlsruhe, Germany, demonstrating its practical applicability in real-world urban environments.

This research advances urban mobility management by making sophisticated aerial image analysis accessible to municipal authorities through a low-code approach. The remainder of this paper reviews related work in parking occupancy detection and vision-language models, describes DroneVision's architecture and implementation, presents experimental results, and discusses implications for future research.

2. System Architecture and Technical Implementation

Our research presents a user-friendly approach designed for municipal authorities to analyze parking occupancy using drone imagery. We demonstrate a complete end-to-end solution that includes an interactive web interface built with React and Django, allowing users to visualize and analyze detection results efficiently. Figure 1 illustrates our system workflow, which consists of four main components: parking space registration, drone image processing and reconstruction, occupancy detection, and automated report generation. This architecture enables seamless integration between the detection system and user interface, providing a practical tool for parking management authorities.

2.1 Parking space registration

We developed an interactive mapping interface that integrates geographic information systems (GIS) and custom polygon drawing tools for efficient parking space registration. As shown in Figure 2, the interface features vector-based drawing functionality that enables municipal staff to delineate street segments and annotate parking space information. Each segment is assigned a street-corresponding name and includes detailed parking capacity data, facilitating comprehensive analysis of parking pressure across different road sections. This digitization process transforms traditional paper-based parking records into a structured digital format through a streamlined, one-time registration workflow. The system maintains these geospatial annotations in a database, establishing persistent records that serve as the foundation for subsequent analyses.

2.2 Drone image processing and reconstruction

An automated image processing pipeline for handling drone-acquired aerial imagery is presented. This system integrates OpenDroneMap's (ODM, n.d.) photogrammetry capabilities to process overlapping aerial photographs, reconstructing orthophotos that ensure both geometric accuracy and spatial consistency. As illustrated in Figure 3, the reconstructed orthophoto demonstrates effective geometric correction and precise spatial alignment of the study area.

Challenges arise in processing drone-captured images, especially under suboptimal lighting conditions such as during twilight hours. To mitigate these low-light issues, multiple image enhancement techniques were implemented using OpenCV's Python library. The enhds: Histogram Equalization (HE) (Pizer et al., 1987) to improve global contrast by redistributing intensity values; Contrast Limited Adaptive Histogram Equalization (CLAHE) (Setiawan et al., 2013) for region-specific enhancement while minimizing noise amplification; and Gamma Correction (Guan et al., 2009) to adjust brightness and contrast via a non-linear transformation of pixel intensities.

Furthermore, the system utilizes registered parking space coordinates to automatically extract pertinent image segments from the orthophotos through geospatial data analysis, thereby distinguishing between parking spaces and driving lanes. A manual registration approach was deliberately utilized over advanced computer vision techniques for several strategic reasons. First, the absence of clear parking space demarcation lines in many urban contexts renders automated detection unreliable. Second, this approach allows for the integration of existing administrative data, which enhances the accuracy of public parking space inventories and streamlines data management. Overall, this methodology serves as an effective bridge between traditional administrative records and digital spatial data, enabling authorities to maintain precise control over parking space designations while capitalizing on modern digital analysis techniques.

2.3 Occupancy detection

For the detection of parking occupancy, we implemented two parallel approaches: a traditional CNN-based method and a VLMs based approach that leverages recent advances in generative AI.

The first approach utilizes a CNN-based vehicle detection system implemented with the YOLO (Redmon and Farhadi, 2018) architecture. We trained this model using a hybrid dataset that combines open-source satellite imagery with custom drone-acquired data collected from German cities (Aalen and Stuttgart-Vaihingen). This hybrid training strategy ensures the model's adaptability to aerial images captured at various altitudes, from satellite imagery to high-resolution drone footage. Using the LabelImg annotation tool, we manually labeled different vehicle types in our dataset, including private cars, commercial vans, and trucks, and the model was subsequently fine-tuned to optimize detection accuracy. The final implementation successfully identifies and counts vehicles from aerial imagery. In addition, a filtering step was applied to enhance the robustness of the results. For instance, during segmentation, some vehicles might be split between two road segments, or objects with low detection probabilities may be erroneously detected. To address these issues, detected vehicles with a confidence score of less than 70% were filtered out prior to conducting traditional statistical analyses.

For our second approach, we integrated two state-of-the-art VLMs, Qwen2.5-VL (Wang et al., 2024) and InternVL2.5 (Chen et al., 2024), from Hugging Face. Both models employ separate vision encoders to process image information, after which the encoded image features are combined with text and passed to the language model decoder for further processing. The VLM architecture accepts aerial images along with text-based queries as input. To analyze parking occupancy, we designed two specific prompts:

1. Prompt 1: Please first outline the position of each car and output all the coordinates in JSON format, and then answer how many cars are there in the image. This prompt generates bounding box detections that require subsequent counting.
2. Prompt 2: Please answer how many cars are there in the image. This prompt directly outputs the precise vehicle count.

We subsequently compared the accuracy of these two prompting strategies to evaluate their performance in analyzing parking occupancy.

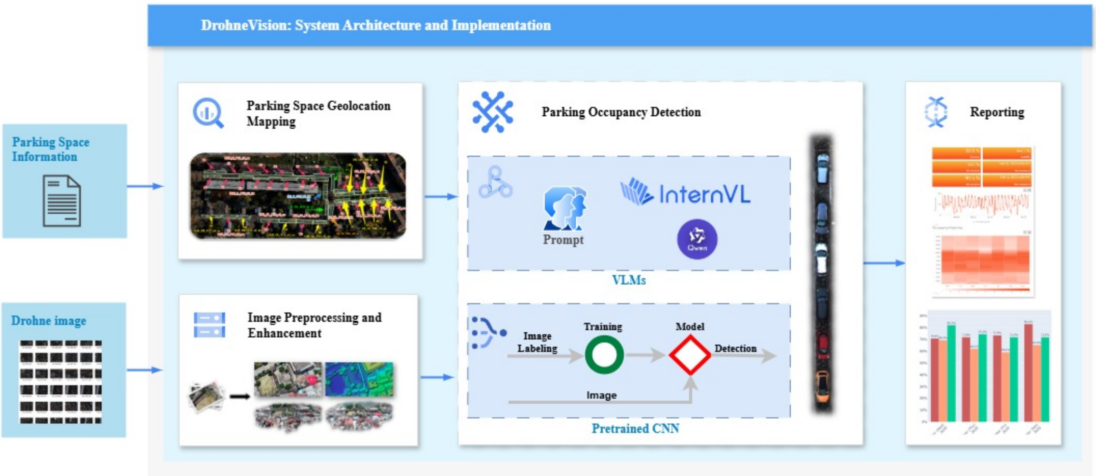

Figure 1. DrohneVision: System Architecture

Figure 2. Interface for Parking Space Registration

Figure 3. An example of a Processed Orthophoto in Test Area

2.4 Report generation

To facilitate practical application, an automated reporting module was developed using traditional statistical methods to generate comprehensive parking analytics. The system processes detection results to produce statistical visualizations of occupancy rates and temporal usage patterns, complemented by a spatial distribution analysis of parking utilization. Our reporting interface transforms complex detection data into accessible insights for municipal authorities.

3. Evaluation and Results

We evaluated our approach through a series of experiments conducted in real-world urban environments, focusing on both technical performance metrics and practical applicability for parking management. To thoroughly evaluate our approach's performance and practical applicability, we conducted a comprehensive case study in a residential district of Karlsruhe, Germany. Additionally, we validated our approach through supplementary testing at two public parking facilities in Heilbronn, enabling us to assess the system's adaptability across different urban environments and parking management contexts. In the following sections, we present a detailed analysis of our primary case study in Karlsruhe, which demonstrates the effectiveness of the approach in a residential parking scenario.

3.1 Experimental Setup

To evaluate our approach, we conducted an extensive data collection campaign using drone imagery in residential areas of Karlsruhe, Germany. The study area, spanning approximately 0.14 square kilometers, encompasses three residential streets characterized by high parking demand, where residents regularly face challenges in finding available parking spaces (Figure 3). Data collection was systematically carried out over a ten-day period, including both weekdays and weekends. To capture temporal variations in parking patterns, we performed three daily drone flights - during morning, noon, and evening hours. Operating at an altitude of 90 meters, in compliance with German aviation regulations that impose a 100-meter ceiling for drone operations, each flight captured 165 images. This resulted in a comprehensive dataset of 4,950 images (165 images × 3 flights × 10 days). The dataset provides diverse coverage of the 204 parking spaces within the study area, incorporating variations in both temporal conditions (different times of day) and environmental factors (varying weather conditions).

3.2 Detection Performance

The implementation of our parking analysis system began with a comprehensive mapping of the test area's parking infrastructure. We adopted a fine-grained segmentation strategy to achieve high analytical precision, systematically dividing the area into multiple small parking segments. Figure 4 illustrates this detailed segmentation approach, providing a complete overview of all delineated parking areas. To ensure data persistence and facilitate subsequent analyses, we established a permanent database record of all segmented parking spaces, creating a robust digital representation of the parking infrastructure.

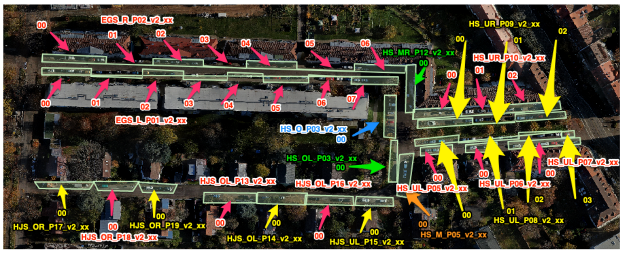

Figure 4. Parking Space Registration and Street Segments in Test Area

Subsequently, we processed and orthorectified drone imagery from 30 different acquisition times to create aerial orthophotos. From these composite images, we extracted the corresponding parking space segments based on our registered coordinates. These segments were then analyzed using both CNN-based and VLMs approaches to detect vehicles and calculate parking occupancy rates. Ground truth data was established through simultaneous manual vehicle counting during drone imagery acquisition.

Model	Results			
	Accuracy	precision	recall	F1-score
Pretrained Yolo	0.979	0.993	0.986	0.989
Qwen2.5 VL	0.932	0.932	1.0	0.965
InternVL2.5	0.878	0.878	1.0	0.935

Table 1. Performance Comparison of CNN-based and VLM-based Vehicle Detection Approaches.

Our detailed analysis of vehicle detection performance compared traditional CNN-based models with emerging VLMs using aerial imagery. In a representative sample, manual counting established a ground truth of 143 parked vehicles. The YOLO model demonstrated high accuracy by detecting an identical total number of vehicles, although closer examination revealed two minor discrepancies: one undetected vehicle and one false detection of a vehicle shadow. The model's performance was quantitatively impressive, achieving a precision of 99.3%, recall of 98.6%, and F1-Score of 98.9%, indicating robust and balanced detection capabilities. As shown in Figure 5, our analysis identified specific detection challenges, particularly in cases involving partially obscured vehicles and motion-induced blur from moving vehicles during image acquisition.

Figure 5. Representative Detection Error Cases from Pretrained YOLO

In our evaluation of VLMs, we first tested InternVL2.5 using two distinct approaches. The grounding+counting method identified 160 vehicles, though it frequently counted single vehicles multiple times. The counting-only method yielded similar results with 159 vehicles. Subsequently, Qwen2.5 VL's grounding+counting approach detected 146 vehicles, showing improved accuracy over InternVL2.5. While Qwen2.5 VL successfully identified all vehicles, it occasionally double-counted less common vehicle types, such as blue cars or those with sunroofs, though it performed reliably with common vehicle colors. Figure 6 illustrates both successful and erroneous detection cases by Qwen2.5 VL. Notably, while the model avoided false positives from vehicle shadows, its primary error pattern was overdetection, leading to multiple counts of single vehicles. The counting-only prompt produced 189 vehicles, significantly overestimating the actual count.

Figure 6. Representative Vehicle Detection Results from Qwen2.5 VL

The comparative analysis demonstrates the YOLO model's superior performance in terms of accuracy and reliability. While VLMs showed promise, they struggled with double-counting and misidentification issues. The notable difference between grounding+counting and counting-only results emphasizes the crucial role of localization in achieving accurate counting results. Currently, CNN-based models maintain their superiority due to extensive fine-tuning with large datasets. However, the rapid evolution of VLM technology presents encouraging progress, as evidenced by the significant improvements in InternVL2.5 (January 2025) compared to InternVL2.5 (December 2024).

VLM present expanded capabilities beyond the specialized tasks of CNN-based models, offering comprehensive identification of diverse urban features including rivers, roads, and micromobility infrastructure. The system architecture developed in this research integrates with open-source VLMs through the Hugging Face platform, enabling seamless incorporation of emerging models that leverage improvements in few-shot learning, zero-shot generalization, and contextual understanding. The inherent capability of VLMs to comprehend complex spatial relationships and natural language descriptions presents significant research opportunities. Through domain-specific fine-tuning and optimized prompting strategies, these models can address complex urban mobility challenges beyond conventional vehicle detection, including traffic flow analysis, pedestrian behavior monitoring, and infrastructure utilization assessment. The continuous advancement in large language model capabilities positions VLMs as a viable solution for comprehensive urban analysis applications, demonstrating their potential to complement and enhance traditional computer vision approaches in future urban monitoring systems.

Following the implementation of both detection models, we conducted statistical analyses to examine parking occupancy patterns across different days and times. Figure 7 presents a visualization of our analysis results, where three distinct columns, represented by different colors, indicate occupancy rates during morning, noon, and evening flight sessions. The data reveals clear temporal patterns in parking utilization. Notably, parking occupancy rates demonstrate significant variation throughout the day, with the lowest occupancy consistently observed during midday periods, while morning and evening hours exhibit substantially higher utilization rates. Furthermore, our analysis indicates a distinct difference between weekday and weekend patterns, with weekend periods showing notably higher parking pressure. These temporal patterns align with typical residential mobility behaviors, where residents tend to be present during early morning and evening hours on weekdays, and maintain higher presence during weekends.

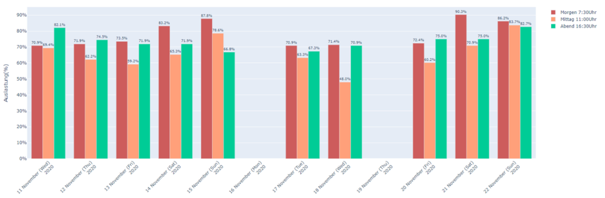

Figure 7. Temporal Analysis of Parking Space Occupancy Rates Across Different Time Periods

3.3 Efficiency

All processing tasks were performed on a workstation equipped with an NVIDIA RTX 4080 GPU. The processing pipeline, including image enhancement, vehicle detection, and report generation, requires approximately 20 minutes to process images from a single drone flight (165 images). The most time-consuming component is the aerial image reconstruction and orthorectification, which takes approximately 14 minutes. The subsequent processing times for different detection approaches are compared in Table 2, including both YOLO and VLM-based methods. The complete workflow demonstrates efficient performance that enables practical deployment for regular parking monitoring tasks.

Model	Processing Time (s)
Pretrained Yolo	2
Qwen2.5 VL (Grounding+Counting)	73
Qwen2.5 VL (Counting)	17
InternVL2.5 (Grounding+Counting)	164
InternVL2.5 (Counting)	23

Table 2. Processing Time Comparison.

We compared our dual-approach system against traditional manual counting methods and existing automated solutions. In our test area, the drone flight and image capture process required 273 seconds. In contrast, simulated manual counting using Google Maps showed that walking the 700-meter street segment would take approximately 9 minutes, even without recording license plate information. This represents a time savings of 49.4%. The time advantage becomes even more pronounced when scaled to larger areas. We conducted a simulation for Heilbronn's city center, covering an area of 340m × 560m with multiple street segments totaling 4.2 km. While manual inspection of this area would require at least one hour of walking, drone coverage could be completed in a maximum of 30 minutes.

To further validate the scalability of our approach, we extended our testing to two public parking lots in Heilbronn. The system demonstrated robust performance across these different environments, maintaining consistent accuracy levels despite the variation in parking space layouts and environmental conditions.

3.4 Challenges and Limitations

While our drone-based parking monitoring system demonstrates promising results, several challenges and limitations need to be addressed in future work:

Environmental Constraints: Our approach's performance is currently limited by weather conditions and lighting variations. Night-time operations and adverse weather conditions (such as rain) can significantly impact image quality and system reliability. To address these limitations, we propose integrating multiple sensing modalities, including infrared cameras for night-time operations and weather-resistant sensors. Additionally, advanced image enhancement techniques and temporal data integration could help maintain system performance under suboptimal conditions.

Regulatory Compliance: Drone operations in urban environments face strict regulatory requirements regarding flight frequency, altitude restrictions, and permitted areas of operation. These regulations can impact the system's deployment flexibility and monitoring frequency. To optimize within these constraints, we suggest developing adaptive flight planning algorithms that maximize coverage while ensuring compliance with local aviation regulations. Furthermore, coordination with municipal authorities to establish dedicated drone corridors could facilitate more regular monitoring operations.

Technical Challenges: Current technical limitations primarily involve partial vehicle occlusions and motion artifacts in captured images. To overcome these challenges, we envision implementing a comprehensive solution combining multiple approaches:

1. Multi-perspective imaging using coordinated drone flights at various altitudes and angles
2. Advanced image processing techniques, including motion deblurring and super-resolution enhancement
3. Integration of temporal data to reduce motion artifacts and resolve occlusions
4. Implementation of deep learning-based image restoration and 3D scene reconstruction

These challenges present opportunities for future research and development, particularly in integrating emerging technologies and optimizing system performance for large-scale urban deployments.

4. Conclusions and outlook

This paper presents DroneVision, a novel approach to urban parking monitoring that combines drone-based aerial imagery with advanced computer vision techniques. Our dual-method system, integrating both CNN-based and VLM-based approaches, achieved a high detection accuracy of 97.9% and 93.2% in real-world testing scenarios. The system demonstrates significant advantages over traditional manual counting methods, reducing survey time by approximately 50% in our test area while maintaining high reliability.

Our extensive evaluation across residential areas and public parking facilities has demonstrated the robust performance and practical viability of our proposed approach. The development of an intuitive interface tailored for municipal authorities ensures seamless system adoption by city planners and parking management teams, eliminating the need for specialized technical knowledge.

While our current implementation shows promising results, we have identified several opportunities for future enhancement. From a technical perspective, the integration of multi-perspective imaging techniques and advanced image processing algorithms could further improve detection accuracy and reliability. Additionally, the incorporation of state-of-the-art VLMs or their fine-tuning with domain-specific data presents promising avenues for further optimization. The scalability of our system also merits consideration, as deployment across larger urban areas would provide valuable insights into its performance at scale and its potential for integration with existing smart city infrastructure. This expansion would enable comprehensive urban parking management while validating the system's adaptability to diverse environmental conditions and parking scenarios.

References

Amato, G., Carrara, F., Falchi, F., Gennaro, C., Meghini, C., Vairo, C., 2017. Deep learning for decentralized parking lot occupancy detection. *Expert Systems with Applications*, 72, 327–334.

Arora, N., Cook, J., Kumar, R., Kuznetsov, I., Li, Y., Liang, H.-J., Miller, A., Tomkins, A., Tsogsuren, I., Wang, Y., 2019. Hard to park? estimating parking difficulty at scale. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2296–2304.

Assemi, B., Paz, A., Baker, D., 2021. On-street parking occupancy inference based on payment transactions. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 10680–10691.

Bai, J., Bai, S., Yang, S., Wang, S., Tan, S., Wang, P., Lin, J., Zhou, C., Zhou, J., 2023. Qwen-vl: A versatile vision-language model for understanding, localization, text reading, and beyond. *arXiv preprint arXiv:2308.12966*, 1(2), 3.

Barmpounakis, M., Montesinos-Ferrer, M., Gonzales, E., Geroliminis, N., 2021. Empirical investigation of the emission-macroscopic fundamental diagram. *Transportation Research Part D: Transport and Environment*, 101, 103090.

Bazzaza, T., Tohidypour, H. R., Wang, Y., Nasiopoulos, P., 2024. Accurate Detection and Localization of Individual Free Street Parking Spaces Using AI and Innovative Global Motion Estimation. *IEEE Transactions on Intelligent Vehicles*.

Bock, F., Di Martino, S., Origlia, A., 2019. Smart parking: Using a crowd of taxis to sense on-street parking space availability. *IEEE Transactions on Intelligent Transportation Systems*, 21(2), 496–508.

Ceder, A., 2021. Urban mobility and public transport: future perspectives and review. *International Journal of Urban Sciences*, 25(4), 455–479.

Chen, X., Wu, S., Shi, C., Huang, Y., Yang, Y., Ke, R., Zhao, J., 2020. Sensing data supported traffic flow prediction via denoising schemes and ANN: A comparison. *IEEE Sensors Journal*, 20(23), 14317–14328.

Chen, Z., Wang, W., Tian, H., Ye, S., Gao, Z., Cui, E., Tong, W., Hu, K., Luo, J., Ma, Z. et al., 2024. How Far Are We to GPT-4V? Closing the Gap to Commercial Multimodal Models with Open-Source Suites. *arXiv preprint arXiv:2404.16821*.

Gören, S., Öncevarlık, D. F., Yldz, K. D., Hakyemez, T. Z., 2019. On-street parking spot detection for smart cities. *2019 IEEE International smart cities conference (ISC2)*, IEEE, 292–295.

Guan, X., Jian, S., Hongda, P., Zhiguo, Z., Haibin, G., 2009. An image enhancement method based on gamma correction. *2009 Second International Symposium on Computational Intelligence and Design*, 1, 60–63.

Kim, S., Anagnostopoulos, G., Barmpounakis, E., Geroliminis, N., 2023. Visual extensions and anomaly detection in the pNEUMA experiment with a swarm of drones. *Transportation Research Part C: Emerging Technologies*, 147, 103966.

Kim, S., Tak, Y., Barmpounakis, E., Geroliminis, N., 2024. Monitoring Outdoor Parking in Urban Areas With Unmanned Aerial Vehicles. *IEEE Transactions on Intelligent Transportation Systems*.

Liao, F., Sun, Y., Wu, Y., Wang, J., 2022. Real-time occupancy detection of on-street parking spaces based on an edge device. *2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML)*, IEEE, 621–625.

Lin, T., Rivano, H., Le Mouël, F., 2017. A survey of smart parking solutions. *IEEE Transactions on Intelligent Transportation Systems*, 18(12), 3229–3253.

Mahaarachchi, B. K., Cohen, S., Bookhagen, B., Doskoč, V., Friedrich, T., 2023. Sustainable on-street parking mapping with deep learning and airborne imagery. *International Conference on Intelligent Data Engineering and Automated Learning*, Springer, 209–221.

Mavlutova, I., Atstaja, D., Grasis, J., Kuzmina, J., Uvarova, I., Roga, D., 2023. Urban transportation concept and sustainable urban mobility in smart cities: a review. *Energies*, 16(8), 3585.

Mohsan, S. A. H., Othman, N. Q. H., Li, Y., Alsharif, M. H., Khan, M. A., 2023. Unmanned aerial vehicles (UAVs): Practical aspects, applications, open challenges, security issues, and future trends. *Intelligent Service Robotics*, 16(1), 109–137.

Nawaz, S., Efstratiou, C., Mascolo, C., 2013. Parksense: A smartphone based sensing system for on-street parking. *Proceedings of the 19th annual international conference on Mobile computing & networking*, 75–86.

ODM, O. A., n.d. command line toolkit to generate maps, point clouds, 3D models and DEMs from drone, balloon or kite images. *GitHub Page 2020; https://github.com/OpenDroneMap/ODM*.

Pannerselvam, K., 2021. Adaptive parking slot occupancy detection using vision transformer and llie. *2021 IEEE International Smart Cities Conference (ISC2)*, IEEE, 1–7.

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. B., Zuiderveld, K., 1987. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3), 355–368.

Redmon, J., Farhadi, A., 2018. Yolov3: An incremental improvement.

Setiawan, A. W., Mengko, T. R., Santoso, O. S., Suksmono, A. B., 2013. Color retinal image enhancement using clahe. *International Conference on ICT for Smart Society*, 1–3.

SFpark, 2018. Parking Sensor Data Guide—SFMTA. *Available:*, https://www.sfmta.com/sites/default/files/reports-and-documents/2018/ 08/sfpark$_{d}ataguide_{p}arkingsensordata.pdf$.

Shivaprasad, S., Anand, M., Chilkunda, S. A., Kamalesh, A., Oruganti, R., Radhakrishna, S., Venugopal, N., 2024. Finding potential on-street parking spots: An object detection and segmentation approach. *International Conference on Smart Computing and Communication*, Springer, 433–443.

Šolić, P., Leoni, A., Colella, R., Perković, T., Catarinucci, L., Stornelli, V., 2020. IoT-ready energy-autonomous parking sensor device. *IEEE internet of things journal*, 8(6), 4830–4840.

Telli, K., Kraa, O., Himeur, Y., Ouamane, A., Boumehraz, M., Atalla, S., Mansoor, W., 2023. A comprehensive review of recent research trends on unmanned aerial vehicles (uavs). *Systems*, 11(8), 400.

Wang, P., Bai, S., Tan, S., Wang, S., Fan, Z., Bai, J., Chen, K., Liu, X., Wang, J., Ge, W., Fan, Y., Dang, K., Du, M., Ren, X., Men, R., Liu, D., Zhou, C., Zhou, J., Lin, J., 2024. Qwen2-VL: Enhancing Vision-Language Model’s Perception of the World at Any Resolution. *arXiv preprint arXiv:2409.12191*.

Zhang, J., Huang, J., Jin, S., Lu, S., 2024. Vision-language models for vision tasks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Zhao, D., Ju, C., Zhu, G., Ning, J., Luo, D., Zhang, D., Ma, H., 2021. MePark: Using meters as sensors for citywide on-street parking availability prediction. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), 7244–7257.

Zhu, D., Chen, J., Shen, X., Li, X., Elhoseiny, M., 2023. Minigpt-4: Enhancing vision-language understanding with advanced large language models. *arXiv preprint arXiv:2304.10592*.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-731-2025