Urban on-street parking monitoring with drone flights

Yura Tak*[1], Maksymiliann Schoeffel[1], and Nikolas Geroliminis[1]

[1]Urban Transport Systems Laboratory (LUTS), EPFL, Switzerland

*Corresponding Author

Short summary

Efficient on-street parking management is a critical challenge for modern cities due to its direct impact on traffic congestion, energy consumption, and driver satisfaction. The dispersed nature of on-street parking across large regions makes monitoring difficult using traditional sensors, severely limiting their scalability in dynamic environments. This paper presents a novel, fully automated drone-based framework to extract on-street parking locations and monitor occupancy directly from raw aerial video. Our approach eliminates the need for predefined maps or inaccurate crowd-sourced data. Instead, it infers parking masks from observed vehicle trajectories. The framework detects vehicles and tracks their trajectories to isolate stationary parking behavior from active traffic. We then use a spatial-temporal dwell-time scoring technique combined with a clustering algorithm to generate precise parking masks. A real-world case study in Galatsi, Greece, demonstrates the effectiveness of this method, capturing two-day parking occupancy dynamics with high autonomy. Results highlight the potential of drone-based aerial systems to deliver scalable, cost-effective, and adaptive solutions for studying urban mobility and monitoring on-street parking without reliance on any prior information.

Keywords: On-street parking, Drone monitoring, Computer vision

1 Introduction

Effective parking management lies at the foundation of sustainable urban mobility, directly influencing traffic congestion, energy consumption, and driver satisfaction (Mantouka et al., 2021; Shoup, 2006; Zhu et al., 2020). Unlike traditional off-street parking facilities, which feature explicitly defined physical boundaries and controlled access points that are easily monitored by centralized barrier sensors—on-street parking operates in a highly unstructured environment. On-street parking spaces are continuously distributed along active roadways, often lack clear geometric delineations, and are heavily intertwined with standard traffic flow. This decentralized and open nature, combined with the highly dynamic turnover of vehicles throughout the day, makes on-street parking monitoring a uniquely challenging task that traditional static infrastructure struggles to scale with.

Traditional monitoring systems, such as in-ground sensors, fixed surveillance cameras, or manual surveys, are often limited in spatial coverage, expensive to deploy, and inflexible to urban changes (Georgopoulou & Papantoniou, 2025). These limitations motivate the exploration of aerial and mobile sensing platforms, particularly unmanned aerial vehicles (UAVs), as a new paradigm for large-scale urban observation. UAVs, combined with advanced computer vision, offer a unique capability to survey wide urban areas rapidly and flexibly, capturing high-resolution imagery that can be processed to extract vehicle positions and infer parking occupancy (Fonod et al., 2025).

Parking is a low-dynamic phenomenon, where vehicles remain stationary for extended periods. Consequently, low-frequency drone observations are sufficient to provide reliable information about parking availability and usage patterns. This characteristic makes on-street parking monitoring highly compatible with drone-based data acquisition, even when drones operate intermittently or across large coverage zones. Furthermore, Lu et al. (2025) provide compelling evidence that drone-based enforcement systems can reduce operational costs by over 90% compared to traditional manual patrolling.

Despite these advantages, a critical gap remains in the literature. Existing UAV-based parking frameworks are often constrained by their reliance on prior infrastructural knowledge (Wang et al., 2020). Many current methodologies depend on pre-mapped ShapeFiles or highly precise satellite

orthoimages to define parking boundaries before monitoring can begin (Kim et al., 2024). Moreover, crowd-sourced data, such as OpenStreetMap (OSM), is often inconsistent, outdated or non-existent for on-street parking locations (Hellekes et al., 2023). This reliance on predefined maps severely limits the scalability of drone deployments in dynamic urban zones.
To bridge this gap, this paper introduces a fully automated framework capable of extracting on-street parking locations and monitoring occupancy directly from raw aerial video. We utilize robust multi-object tracking (Zhang et al., 2022) to continuously monitor vehicle positions, distinguishing locations where vehicles consistently remain stationary from areas of active traffic flow. By mapping these accumulated stationary periods and applying unsupervised clustering, the framework constructs precise on-street parking zone without any prior infrastructural data.
The main contributions of this paper are:

- Automatic inference of on-street parking locations strictly from raw drone data.
- A scalable methodology that eliminates the dependency on previous knowledge of parking locations or inconsistent crowd-sourced maps.
- Analysis of real-world occupancy dynamics and usage patterns based on a case study in Galatsi, Greece.

The remainder of the paper is structured as follows: Section 2 presents the proposed methodology, Section 3 discusses the results, and Section 4 concludes the study.

2 Methodology

The framework consists of three main components: drone video processing, on-street parking location estimation, and parking occupancy extraction. An overview of the proposed framework is illustrated in Fig. 1.

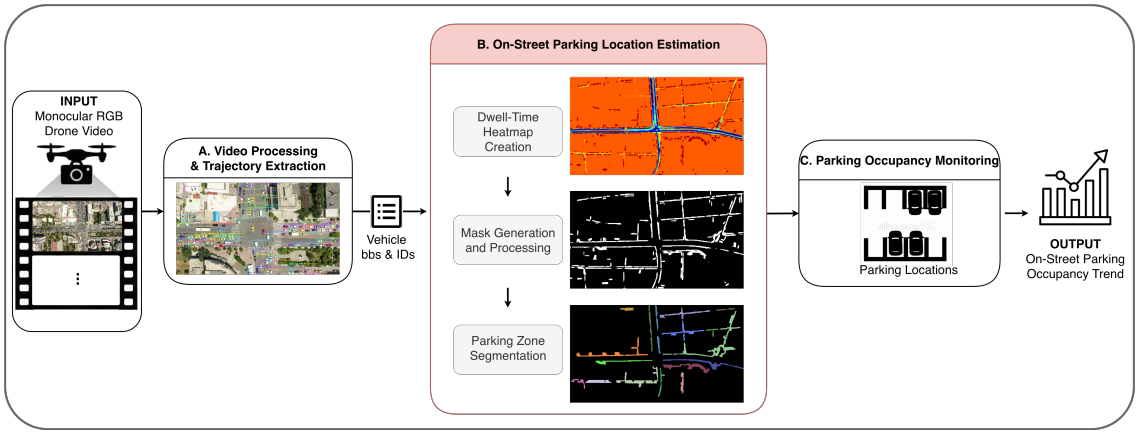


Figure 1: Overview of the proposed drone-based on-street parking monitoring framework.

A. Video Processing and Trajectory Extraction

The initial phase converts raw aerial video into stable vehicle trajectories. First, we stabilize the video using Stabilo, an open-source framework designed to extract trajectories from drone imagery (Fonod et al., 2025). This process aligns consecutive frames into a coherent reference, ensuring that vehicle coordinates are consistent across time regardless of drone movement.
Vehicle detection is performed using YOLO11-m-OBB (He et al., 2025; Jocher & Qiu, 2024). This one-stage detector predicts orientation angles and outputs Oriented Bounding Boxes (OBB). The OBB approach is essential for top-down aerial views to provide precise localization of vehicles at arbitrary orientations. Unlike standard axis-aligned boxes, OBB minimizes the inclusion of non-object background pixels and improves the separation of closely packed vehicles in dense parking rows. We use a model fine-tuned on DOTAv1 (Xia et al., 2018), a large-scale dataset specifically designed for object detection in aerial and satellite imagery. DOTA features Bird's Eye View (BEV) images with vehicle categories. However, because vehicle pixel scales in raw drone footage often

differ from those in the training set, an image patching strategy is applied. Frames are divided into overlapping patches to maintain high-resolution features. A global non-maximum suppression (NMS) step is then applied to merge detections and remove duplicates across patch boundaries.
To maintain vehicle identities across frames, we apply the ByteTrack algorithm (Zhang et al., 2022). This algorithm matches detections based on spatial proximity, OBB overlap, and motion consistency. ByteTrack allows the system to maintain consistent tracks even during partial occlusions, such as trees or shadows. The output of this stage is a set of continuous vehicle trajectories, which are then classified as stopped or moving based on the displacement of each tracked OBB over time. To improve robustness, motion estimates are smoothed using exponential moving average (EMA) and a hysteresis-based thresholding mechanism is applied to prevent unstable switching between states due to tracking noise or slow movement.

B. On-Street Parking Location Estimation

Our framework infers parking locations directly from vehicle behavior without prior maps or annotations. The core heuristic is straightforward: areas where tracked vehicles remain stationary for extended periods are highly likely to be parking spaces. When these stationary zones are directly adjacent to active traffic lanes, we classify them as on-street parking.
Because this approach is entirely data-driven, it is naturally limited by the observation timeframe. Many valid parking spots may simply remain empty during the drone flights. As a result, the initial detection process often produces fragmented, scattered parking footprints rather than continuous curbside lanes.
To overcome this data sparsity, the framework is designed to merge nearby, isolated detections. This step transforms fragmented spots into continuous, lane-like masks that better reflect actual parking infrastructure, which is essential for accurate downstream occupancy analysis. Furthermore, because the raw tracking data inevitably captures off-street parking (such as private lots) and includes tracking errors, the system must also filter out this unwanted noise.
To systematically address these challenges, the on-street parking zone extraction process is divided into three consecutive stages: mask generation, mask processing, and parking zone segmentation (see Fig. 2).

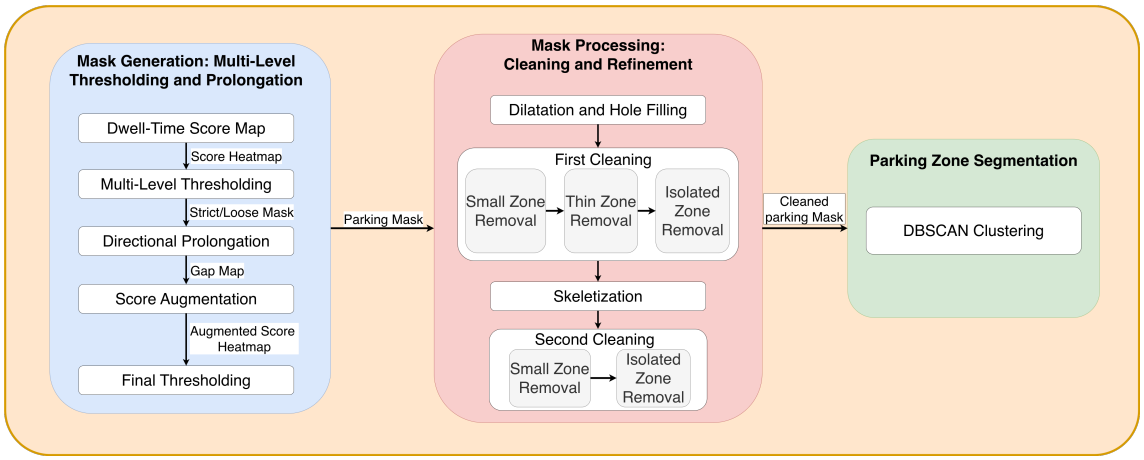


Figure 2: Parking zone extraction pipeline.

Mask Generation: Multi-Level Thresholding and Prolongation

The first stage of the extraction pipeline bridges the gap between discrete vehicle trajectories and static parking geometry. Intuitively, a genuine parking space is characterized by high cumulative stationary time (dwell time) and a distinct absence of active traffic flow. Therefore, the process begins by translating the temporal tracking data into a continuous spatial probability surface—a heatmap. By mathematically "rewarding" stationary behavior and "penalizing" movement, this heatmap provides the foundational evidence needed to extract the initial parking boundaries.
To construct this heatmap, the region of interest is first discretized into a regular spatial grid with a resolution of 10×10 pixels. For each frame, the OBB of every tracked vehicle is projected onto this grid. We generate a spatio-temporal dwell-time map by accumulating time increments in grid

cells occupied by stationary vehicles D_{stop}. Simultaneously, we penalize cells occupied by moving vehicles M_{move} to filter out active traffic flow. This accumulation yields a raw evidence score S_{raw}, which is normalized to produce a probabilistic spatial heatmap of parking activity.

$$S_{\text{raw}} = w_{\text{stop}}\, D_{\text{stop}} - w_{\text{move}}\, M_{\text{move}}$$

Where, w_{stop} and w_{move} are scalar weighting coefficients that control the relative importance of stationary evidence versus moving evidence in the final score. The normalized grid is smoothed with a Gaussian filter to produce a spatially coherent heatmap.

The score heatmap is binarized using a two-level thresholding scheme to ensure robust detection while enabling the recovery of fragmented information. First, a strict threshold (τ_{high}) is applied to generate a mask of high-confidence seed regions. Second, a permissive (loose) threshold (τ_{low}) captures a broader candidate band of weak but potentially valid parking areas. This established strict/loose approach ensures that subsequent refinement focuses on plausible extensions of confident detections. For each connected seed component, the dominant orientation is estimated from its minimum-area bounding rectangle, and a constrained morphological expansion is performed along this principal axis using a line-shaped structuring element. This expansion is bounded by the inferred road mask, obtained from the inverse parking mask while preserving only roadway-connected regions, thereby preventing the mask from extending into roads. Potential gaps are identified as pixels belonging simultaneously to the candidate band and the oriented expansion, while remaining outside the original seed mask.

These gap regions are reinforced through a distance-weighted boosting mechanism. A distance transform is computed from the seed mask such that pixels closer to existing parking evidence receive greater reinforcement than more distant candidates. The underlying score map is selectively increased within these regions, enabling fragmented but spatially consistent parking segments to reconnect without uniformly enlarging all detections. The final, refined parking mask is generated by thresholding the updated, boosted score map.

Mask Processing: Cleaning and Refinement

The raw parking mask is refined to retain only coherent on-street parking zones while removing noise, isolated artifacts, and likely off-street parking detections. The key idea is to progressively enforce the geometric properties expected of on-street parking, namely elongated, continuous, and spatially connected structures, while suppressing detections that do not match these characteristics. This refinement also regularizes the mask into smoother and more uniform parking regions for subsequent analysis.

The generated mask undergoes a comprehensive post-processing pipeline to produce a spatially coherent and morphologically regularized representation of on-street parking regions, as detailed in Algorithm 1.

This process consists of three main phases:

1. Morphological and Geometric Cleanup: The mask is initially processed with morphological closing and dilation to bridge internal gaps and establish continuous parking bands. Subsequently, connected-component analysis systematically eliminates geometrically implausible detections. Candidate components are discarded if they fall below a minimum area (τ_{area}), lack sufficient thickness based on their distance transform (τ_{thick}), or exceed a spatial isolation distance from the main curbside network (τ_{iso_1}).

2. Iterative Skeleton-based Refinement: To bridge remaining discontinuities between adjacent vehicles without distorting global geometry, the mask is iteratively refined. In each iteration, the medial topological structure is extracted via skeletonization. Nearby disconnected segments are reconnected through minor isotropic dilation and immediately re-skeletonized. This continuous skeletal network is then dilated by a fixed radius (R) to reconstruct a smoother parking mask.

3. Final Cleaning: A concluding pass of area and isolation filtering removes any residual small artifacts introduced during the topological reconstruction operations. The result is the final binary on-street parking mask, cleanly shaped for subsequent zone-level segmentation.

Parking Zone Segmentation

The global parking occupancy mask is subsequently partitioned into independent functional parking zones using DBSCAN (Density-Based Spatial Clustering of Applications with Noise) (Ester et al., 1996). DBSCAN groups spatially dense parking pixels into coherent clusters while remaining robust to minor discontinuities and residual noise, thereby separating the global mask into distinct parking zones corresponding to individual parking areas or street segments.

Algorithm 1 Parking Mask Cleaning and Refinement

Require: Raw binary mask M_{raw}, Iterations N, Reconstruction radius R
Require: Thresholds: $\tau_{area}, \tau_{thick}, \tau_{iso_1}, \tau_{final_area}, \tau_{iso_2}$
Phase 1: Morphological and Geometric Cleanup
$M \leftarrow \text{Dilate}(\text{Close}(M_{raw}))$ ▷ Bridge gaps and establish bands
for each connected component $C \in M$ **do**
  if $\text{Area}(C) < \tau_{area}$ **then**
    Remove C from M ▷ Suppress small noise
    continue
  end if
  $T_C \leftarrow \text{EstimateThickness}(\text{DistanceTransform}(C))$
  if $T_C < \tau_{thick}$ **then**
    Remove C from M ▷ Remove spurious narrow structures
    continue
  end if
  $D_{neighbor} \leftarrow \text{MinDistanceToNearest}(C, M \setminus \{C\})$
  if $D_{neighbor} > \tau_{iso_1}$ **then**
    Remove C from M ▷ Reject isolated off-street noise
  end if
end for
Phase 2: Iterative Skeleton-based Refinement
for $i = 1$ **to** N **do**
  $S \leftarrow \text{Skeletonize}(M)$ ▷ Extract medial structure
  $S_{recon} \leftarrow \text{Skeletonize}(\text{SmallDilation}(S))$ ▷ Reconnect segments
  $M \leftarrow \text{Dilate}(S_{recon}, \text{radius} = R)$ ▷ Reconstruct smooth mask
end for
Phase 3: Final Cleaning
$M \leftarrow \text{FilterByArea}(M, \tau_{final_area})$
$M_{final} \leftarrow \text{FilterByIsolation}(M, \tau_{iso_2})$
Ensure: Final globally refined parking mask M_{final}

C. On-Street Parking Monitoring

The final component of the framework quantifies parking usage by spatially correlating vehicle detections with the established parking masks. For every processed frame, the system evaluates the spatial intersection between each vehicle's OBB and the binary masks of the estimated parking zones. A vehicle is recorded as occupying a spot if its OBB is located within a parking zone boundary, with IoU above a threshold τ. By performing this check frame-by-frame, the system calculates the occupancy count per zone at each timestamp. This automated extraction provides the parking dynamics, enabling the study of urban mobility patterns without manual intervention.

3 Results and discussion

The framework has been developed and evaluated on drone videos collected in Galatsi Urban District, Greece. A swarm of 6 drones has been deployed. Among them, 4 were moving to fly through a predefined flight path, 2 were static drones, hovering over the departure regions of the moving drones. The static drones flew at a height of 250 meters in compliance with greek drone regulations. The experiments have been conducted over 2 consecutive days in the urban area of Galatsi, generating a large stream of visual traffic data. The framework has been deployed on the static drones.



Figure 3: Manually annotated ground-truth map of visible on-street parking locations.

On-Street Parking Estimation Evaluation

To evaluate the proposed framework, ground truth on-street parking locations were manually annotated from the drone video frames, as shown in Fig. 3. Only spots visible from the BEV were included in the ground truth. This assumes that if a human annotator cannot identify a parking location due to extreme occlusion, the algorithm should not be penalized for failing to recover it. Based on these manual annotations, a total of 77 distinct on-street parking segments were identified within the observation zone.

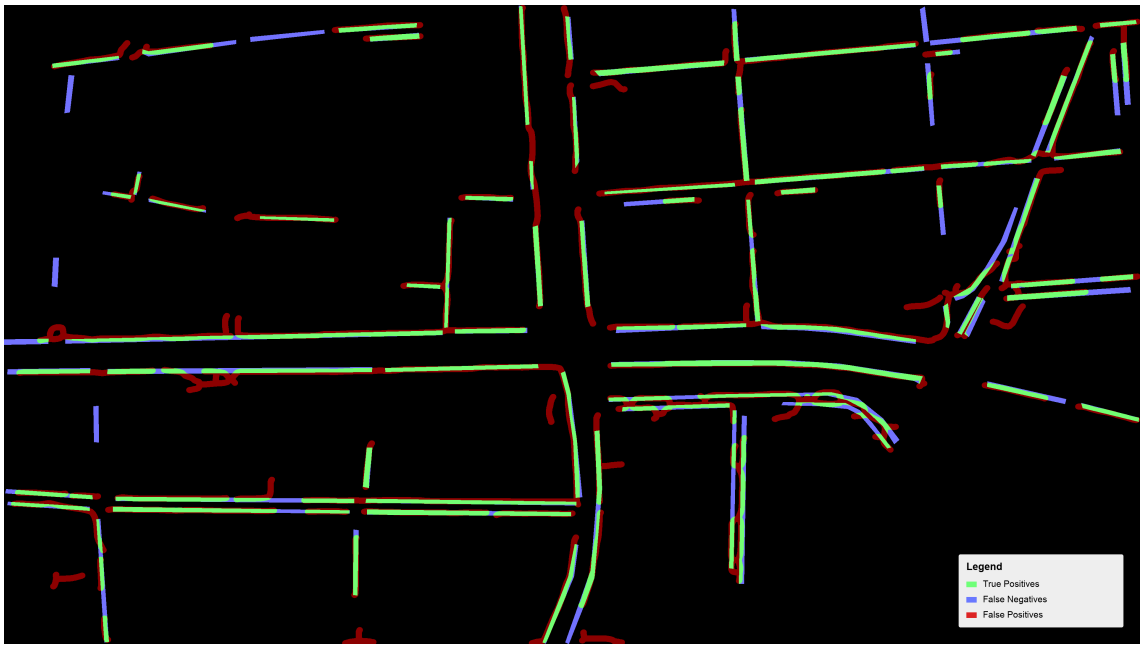


Figure 4: Qualitative evaluation of the mask estimation.

The qualitative performance of the mask estimation is illustrated in Fig. 4, which overlays the generated prediction mask onto the annotated ground truth. In this visualization, green regions represent successful detections (True Positives), indicating accurate overlap between the estimation

and the true parking lanes. Blue regions correspond to False Negatives (FN), highlighting valid parking areas that the framework failed to recover due to a lack of observed parking events. Conversely, red regions denote False Positives (FP). These erroneous detections are primarily caused by vehicles stopped at traffic lights or illegal parking events that the algorithm misclassified as valid parking zones.

Table 1: Quantitative Evaluation of Parking Mask Estimation

Precision	Recall	F1-Score	Accuracy
0.57	0.79	0.66	0.96

Quantitative results are summarized in Table 1. We utilize a standard segmentation metric: precision, recall, F1 score and accuracy. While the F1 score is 0.66, the overall accuracy reaches 0.96. The lower F1 score and precision values are attributed to the fact that our data-driven masks are generated from behavioral dwell-times, which produce shapes that do not perfectly match the rigid, geometric polygons of the manual annotations, which is heavily penalized by the metrics.

Table 2: Lane-level Parking Mask Evaluation

Mean Coverage	Median Coverage	Zone Found at 40% Coverage
0.75	0.85	69/77 (89.61%)

To provide a more practically meaningful evaluation, we additionally assess detection performance at the lane level, as summarized in Table 2. Under a 40% coverage criterion, the method successfully identifies 69 out of 77 annotated parking lanes. Furthermore, the mean and median coverage of detected lanes reach 0.75 and 0.85 respectively, indicating that although geometric boundaries may differ from the ground truth, the majority of parking structures are well localized and substantially covered by the predicted mask.

Parking Occupancy Analysis

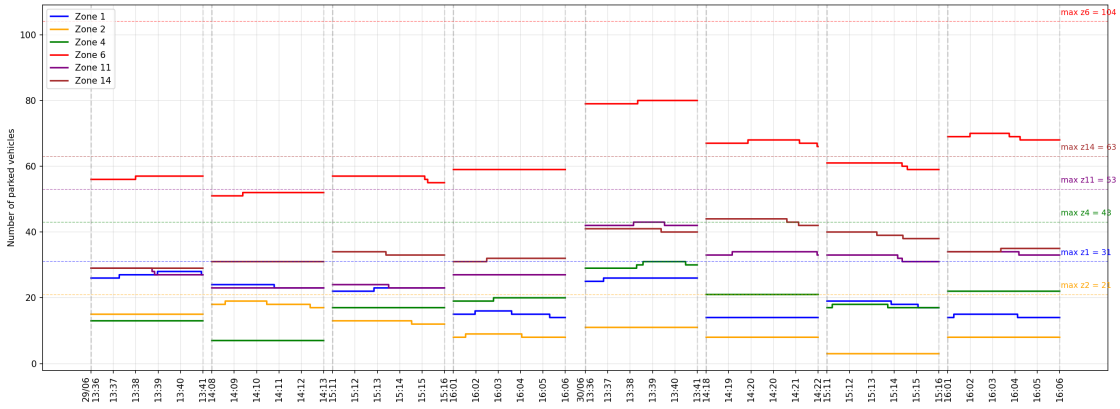


Figure 5: Temporal analysis of on-street parking occupancy across 6 representative zones during 4 distinct sessions over 2 consecutive days.

To establish a baseline for the parking occupancy, the maximum capacity for each zone was calculated. These capacities were determined by dividing the area of the estimated parking zone divided by the average area of a vehicle augmented by offset δ, providing a localized ceiling for vehicle density. As illustrated in Fig. 5, Zone 6 is the largest identified on-street parking zone, with a maximum capacity of 104 vehicles, while the remaining five representative zones provide smaller scales of available supply.

The occupancy data reveals distinct intra-day variances, particularly during the 13:30 to 16:00 intervals. A notable observation is the significant peak in occupancy occurring on the second day (June 30) around 13:30 across nearly all zones. For instance, Zone 6 reaches approximately 80% of its total capacity during this period. This localized surge in demand can likely be attributed

to the presence of a morning market on the second day, which draws higher traffic volumes to the area. This correlation is further supported by the observed drop in occupancy across these zones after 14:00 as the market concludes.
In contrast, Zone 2 exhibits minimal fluctuation during this same period. This relative stability is linked to its geographical displacement from the market site, rendering it less susceptible to event-driven demand spikes. Finally, the consistency of the zone occupancy rankings across both June 29 and June 30 confirms that the proposed methodology is stable, providing a robust framework for assessing parking demand and optimizing urban mobility.

4 Conclusions

This paper presented a fully automated framework for monitoring on-street parking using only raw aerial video data. Our approach eliminates the dependency on auxiliary geospatial data or prior maps by inferring parking boundaries directly from vehicle trajectories. The system provides a comprehensive pipeline spanning from boundary detection to the analysis of occupancy dynamics. The proposed data-driven methodology utilises a pixel-wise scoring function based on the contrast between stationary dwell times and transit frequency. Experimental results demonstrate that this approach accurately detects 89.61% of the parking spots. While the pixel-level metrics are relatively low due to the discrepancy between the data-driven masks and manual geometric annotations, the framework offers high functional utility for parking monitoring.
The primary limitation of this data-driven approach is the requirement for at least one parking event to occur within a spot for it to be identified. Additionally, the current framework is sensitive to outliers, meaning illegal parking events may be incorporated into the estimated boundaries. Future work will focus on implementing outlier detection and more conservative classification strategies to refine the parking zones. Extending the temporal observation window will also enhance the robustness and completeness of the inferred parking inventories.
In conclusion, this study highlights the potential of drone-based aerial systems as scalable, cost-effective, and adaptive solutions for urban mobility. By operating without reliance on prior infrastructural data, the framework provides a viable path for monitoring dynamic urban environments.

References

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd* (Vol. 96, pp. 226–231).

Fonod, R., Cho, H., Yeo, H., & Geroliminis, N. (2025). Advanced computer vision for extracting georeferenced vehicle trajectories from drone imagery. *Transportation Research Part C: Emerging Technologies*, *178*, 105205.

Georgopoulou, C., & Papantoniou, P. (2025). A literature review of automated roadside parking monitoring using artificial intelligence algorithms. *Electronics*, *14*(20), 4119.

He, L.-h., Zhou, Y.-z., Liu, L., Cao, W., & Ma, J.-h. (2025). Research on object detection and recognition in remote sensing images based on yolov11. *Scientific Reports*, *15*(1), 14032.

Hellekes, J., Kehlbacher, A., Díaz, M. L., Merkle, N., Henry, C., Kurz, F., & Heinrichs, M. (2023). Parking space inventory from above: Detection on aerial images and estimation for unobserved regions. *IET Intelligent Transport Systems*, *17*(5), 1009–1021.

Jocher, G., & Qiu, J. (2024). *Ultralytics yolo11.* GitHub. Retrieved from `https://github.com/ultralytics/ultralytics`

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*.

Lu, S.-H., Chang, C.-C., Raj, R., & Lai, W.-T. (2025). Application of drones in issuing parking tickets: an innovative solution for route optimization and workforce replacement. *International Journal of Intelligent Computing and Cybernetics*, *18*(4), 812–855.

Mantouka, E. G., Fafoutellis, P., & Vlahogianni, E. I. (2021). Deep survival analysis of searching for on-street parking in urban areas. *Transportation Research Part C: Emerging Technologies*, *128*, 103173.

Shoup, D. C. (2006). Cruising for parking. *Transport policy*, *13*(6), 479–486.

Wang, H., Liu, Y., Huang, H., Pan, Y., Yu, W., Jiang, J., ... others (2020). Atg-pvd: ticketing parking violations on a drone. In *European conference on computer vision* (pp. 541–557).

Xia, G.-S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J., ... Zhang, L. (2018). Dota: A large-scale dataset for object detection in aerial images. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 3974–3983).

Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., ... Wang, X. (2022). Bytetrack: Multi-object tracking by associating every detection box. In *European conference on computer vision* (pp. 1–21).

Zhu, Y., Ye, X., Chen, J., Yan, X., & Wang, T. (2020). Impact of cruising for parking on travel time of traffic flow. *Sustainability*, *12*(8), 3079.