

RESEARCH ARTICLE



A Privacy-Aware UAV Parking Framework with Polygon-Based Slot Association and Full-Frame Anonymization

Aisha Alahmed¹ , Saed Alqaraleh² , Qanout Alsharida¹, Mariam Alahmed¹, Leen Boodai¹ and A. H. Abdul Hafez^{1,*}

¹Department of Computer Science, College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa, Saudi Arabia

²Department of Data Science and Artificial Intelligence, Mutah University, Jordan

Abstract: Large outdoor parking facilities are difficult to monitor with fixed sensors, whereas unmanned aerial vehicles (UAVs) provide wide-area coverage without per-space hardware. Converting aerial imagery into operational records, however, requires more than vehicle detection: detections must be assigned to specific parking spaces, linked to stored locations, and anonymized before retaining frames. We present a modular privacy-aware UAV parking framework combining YOLOv8s detection, box-normalized polygon association, geo-referenced slot records, and an independent full-frame anonymization branch. Because anonymization operates on the complete frame, sensitive regions remain eligible for processing regardless of occupancy-detector output. The four-stage occupancy detector achieved held-out test precision of 0.984, recall of 0.986, and mAP@0.5 of 0.963. At the predefined association threshold ($\tau = 0.50$), the mapping module correctly resolved 1,697 of 1,712 annotated slots, achieving 99.12% slot-state accuracy (95% image-cluster bootstrap CI: 98.34–99.66%). Under calibration-separated evaluation, end-to-end anonymization recall reached 91.37% for 6,015 WIDER FACE instances and 92.24% for 902 license plates, compared with 58.27% and 37.47% for same-data baselines. These results establish a traceable image-to-record pipeline in which occupancy detection, slot association, and anonymization are evaluated as distinct components.

Keywords: smart parking, UAV imagery, parking-occupancy detection, polygon-based slot association, full-frame anonymization

1. Introduction

Parking search can increase cruising time and associated vehicle emissions in dense urban areas [1]. As smart-city services become more data-driven, parking operators increasingly require slot-level availability information rather than only aggregate occupancy estimates. Recent reviews demonstrate that smart-parking systems span sensing, communication, computation, and user-service layers [2].

Existing smart-parking deployments commonly use magnetic or ultrasonic sensors, Radio-Frequency Identification (RFID)-based access systems, fixed surveillance cameras, or hybrid Internet-of-Things (IoT) architectures. These approaches can be effective in controlled sites, but large outdoor facilities may require substantial installation, calibration, maintenance, or camera-coverage planning. Unmanned aerial vehicle (UAV)-view monitoring provides a complementary sensing perspective because elevated imagery can cover broad spatial regions, although the resulting images introduce small-object scale, dense perspective,

occlusion, illumination variation, geometric alignment, and privacy challenges [2–4].

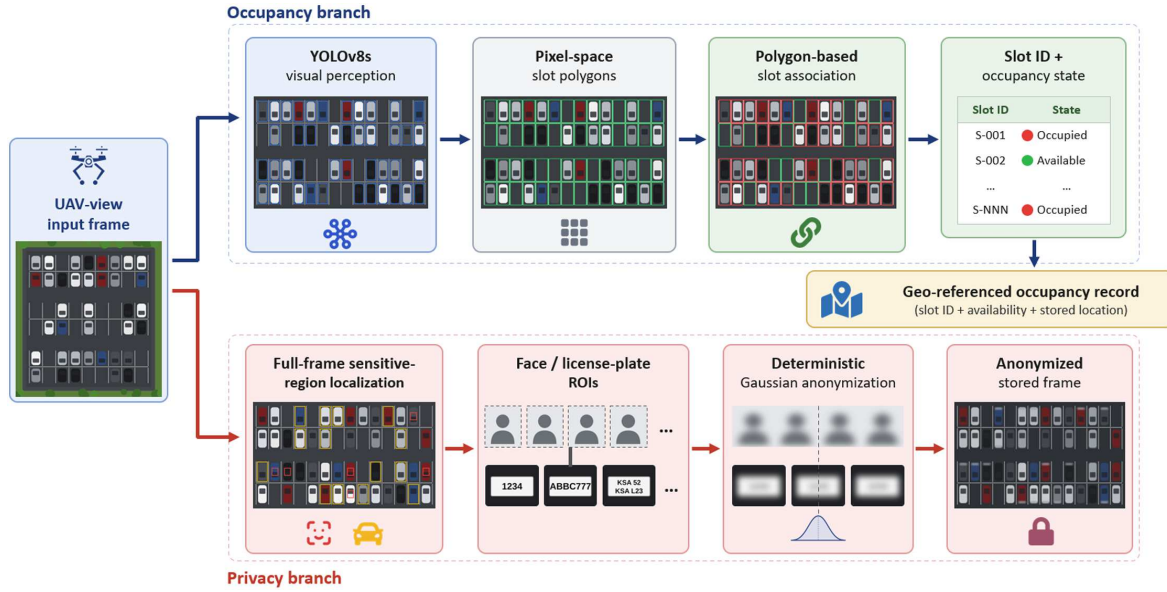
Deep neural detectors, particularly YOLO-family one-stage models, are widely used in traffic and parking analytics because they offer practical real-time performance. However, detector accuracy alone does not provide slot-level parking information. A deployable research pipeline must also determine which fixed parking slot a visual prediction belongs to, maintain a traceable connection between that slot and its stored location record, and handle sensitive visual regions independently of occupancy detection. These downstream functions create distinct error sources and therefore require their own evaluation rather than being inferred from detector mAP.

This study addresses the system-level gap between object detection and operational parking information through a modular image-to-record architecture (Figure 1).

The UAV-view input frame is processed through two parallel branches. The occupancy branch performs YOLOv8s-based visual perception, pixel-space polygon association, and slot-state generation linked to the corresponding geo-referenced occupancy record. The privacy branch independently processes the full input frame to localize face and license-plate regions and applies deterministic Gaussian anonymization before storage. The two branches operate independently so that occupancy estimation is

*Corresponding author: A. H. Abdul Hafez, Department of Computer Science, College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa, Saudi Arabia. Email: aabdulhafez@kfu.edu.sa

Figure 1
System architecture and evaluation boundaries of the proposed privacy-aware UAV parking framework



unaffected by anonymization and sensitive-region processing is not conditioned on successful parking detection.

The framework separates UAV-view perception, polygon-based slot association, geo-referenced record generation, and full-frame anonymization into independently measurable stages. Its contribution is the combined formulation and validation of these stages around a traceable slot-level output, supported by dedicated ground truth, predefined decision rules, calibration-separated privacy testing, uncertainty estimates, and diagnostic analysis.

1.1. Main contributions

The main contributions of this study are summarized as follows:

- An integrated dual-branch UAV parking architecture that converts full-frame visual input into traceable slot-level occupancy records while localizing and anonymizing faces and license plates independently of occupancy detection.
- An auditable box-normalized polygon-association method that links vehicle detections to fixed parking-space geometry and geo-referenced slot identifiers, evaluated on 1,712 PKLot slots with class-wise metrics, image-cluster uncertainty, and threshold sensitivity.
- A calibration-separated full-frame anonymization protocol that freezes model, threshold, inference resolution, and padding before final face and plate scoring, with end-to-end coverage, blur-area cost, same-data baselines, and size-stratified Wilson intervals.
- A multi-layer validation strategy that keeps detector, association, and anonymization evidence distinct, combining held-out detector testing, deterministic real-data slot diagnostics, and quantitative privacy analysis without reproducing sensitive benchmark imagery.

2. Related Work

2.1. Smart parking and UAV-view monitoring

Smart parking has been studied through sensor-based, camera-based, crowdsourced, and hybrid IoT approaches.

Fixed-camera vision systems reduce per-slot hardware but depend on stable imaging geometry and robust visual recognition. UAV monitoring has also been investigated for road and parking applications because it offers flexible viewpoints and broad coverage [3, 4]. In the parking setting, however, UAV imagery introduces additional challenges related to object scale, perspective, occlusion, illumination, and the relationship between image-space detections and fixed slot geometry.

VisDrone-DET2019 [5] provides large-scale visible-spectrum drone imagery for dense aerial object detection, while High-altitude Infrared Thermal dataset for Unmanned Aerial Vehicle-based object detection (HIT-UAV) [6] contributes high-altitude infrared UAV imagery across varied scenes. These datasets broaden the aerial operating conditions available for training, but neither provides the scene-specific parking-slot geometry required for deterministic slot-level occupancy assignment. The parking-specific dataset [7] is therefore used separately for the final available/occupied task.

Aerial detection surveys converge on scale as the dominant difficulty. Leng et al. [8] organize the field around a set of imbalance problems, of which scale and spatial imbalance bear most directly on parking scenes; Hua and Chen [9] treat small-object size as the defining constraint; and Tang et al. [10] document the accompanying sensitivity to altitude, viewpoint, and illumination. Sun et al. [11] add a modality axis, learning under the uncertainty that arises when RGB and infrared views of the same scene disagree. These observations shaped the training data used here rather than the architecture. Stages 1–3 progressively broaden the aerial training data, beginning with the VisDrone baseline (Stage 1), followed by parking-domain adaptation (Stage 2) and multi-source aerial training using VisDrone, HIT-UAV, and parking imagery (Stage 3). The best Stage 3 checkpoint then initializes Stage 4, the parking-specific occupancy fine-tuning stage.

2.2. Detector context and adopted backbone

YOLOv8 is an anchor-free YOLO-family detector with mature training, inference, export, and multi-task support in the Ultralytics ecosystem [12, 13]. YOLOv10 introduces an end-to-end Non-Maximum Suppression (NMS)-free design [14], while Real-Time Detection Transformer (RT-DETR) provides

a real-time transformer-based detection framework [15]. Recent comprehensive reviews trace the continuing development of the YOLO family from earlier versions through YOLO11 and subsequent architectures [13, 16]. Broader object-detection surveys contextualize these developments within the evolution of modern Convolutional Neural Network (CNN)- and transformer-based methods [17], while transformer-focused reviews examine DETR-style detection architectures, their design variants, comparative performance, and remaining challenges [18]. These works provide architectural context, although their published metrics are not directly comparable with the parking-specific task evaluated here.

YOLOv8s was retained as the implemented backbone because it supported the complete development chain, offered mature training and export tooling, and provided a strong final precision–recall operating point. Section 4.4 reports a retrospective descriptive benchmark against YOLOv5s, YOLOv10s, YOLO11s, and RT-DETR; this comparison was assembled to contextualize recorded detector behavior and to provide additional evidence regarding the suitability of the adopted YOLOv8s backbone, rather than to conduct prospective architecture selection. The benchmark records that YOLOv8s reports a precision of 0.984, recall of 0.986, and mAP@0.5 of 0.963. CMCA-YOLO [19] provides additional parking-specific context from the literature.

More broadly, detector selection is only one layer of the parking pipeline. Slot association and privacy-aware visual processing are evaluated separately because errors in those modules cannot be inferred from detector mAP, precision, or recall alone. Accordingly, the study examines whether a modular UAV-view pipeline can convert visual predictions into accurate, traceable slot records while independently quantifying association behavior and privacy-aware anonymization coverage.

2.3. Evaluation gap and positioning of the proposed study

Computer-vision-based parking management encompasses several related tasks, including parking-space classification, automatic parking-space detection, and vehicle detection and counting [20]. These tasks have generally been investigated in different experimental settings. PKLot established a large fixed-camera parking-space classification benchmark [21]. Rafique et al. [22]

developed a deep-learning-based parking-management framework that uses YOLOv5s vehicle detection to determine parking availability and evaluated the approach on PKLot and real-world parking scenarios. Grbić and Koch [23] explicitly evaluated automatic parking-slot detection followed by occupancy classification in fixed-camera imagery.

UAV-based parking surveillance has also been demonstrated [4], establishing the relevance of aerial sensing but not, in the cited study, the module-level polygon-association and privacy evaluation considered here.

Recent parking-navigation work demonstrates that visual perception can be combined with homography, GeoJSON site models, and mobile Global Navigation Satellite System (GNSS) for real-world localization [24]. In the present framework, association is performed directly in image coordinates using scene-specific pixel-space polygons; the resulting slot identifier then links to a stored geo-referenced record.

Privacy-aware visual processing must be evaluated separately from the primary vision task because privacy transformations and sensitive-region localization can introduce their own performance trade-offs. Hukkelås and Lindseth [25] showed that image anonymization can affect the utility of data used for computer-vision training, highlighting the need to evaluate anonymization rather than treating it as a purely cosmetic post-processing operation. De Coninck et al. [26] further demonstrated privacy-preserving visual analysis through learned video obfuscation designed to retain task-relevant information while suppressing sensitive visual content. In the vehicle domain, Zou et al. [27] combined YOLOv8-based localization of privacy-sensitive license-plate information with a dedicated protection mechanism. These studies motivate the present framework’s full-frame, occupancy-independent sensitive-region branch and the dedicated privacy metrics defined in Section 4.6.

Together, these studies establish the need for an integrated evaluation perspective. The present work advances this perspective by combining aerial perception, polygon-based slot association, geo-referenced record generation, and privacy-aware anonymization within one modular framework and by evaluating the two operational downstream stages independently.

Table 1 positions representative prior work by sensing perspective, system function, and downstream-evaluation scope, highlighting the distinctive combination addressed in the present study.

Table 1
Positioning of representative related work by system component and evaluation scope

Study/reference	View/setting	Primary function	Slot-association evaluation	Privacy evaluation	Scope relative to present study
Sarkar et al. [4]	UAV parking surveillance	Aerial parking monitoring/detection	Not reported	Not reported	UAV setting; downstream modules not evaluated in cited study
Grbić & Koch [23]	Fixed-camera parking	Automatic slot detection + occupancy classification	Yes	Not reported	Strong fixed-view slot-level baseline
CMCA-YOLO [19]	Parking surveillance imagery	Vehicle/object detection	Not reported	Not reported	Parking-specific detector baseline
Laptiev et al. [24]	Fixed multi-camera + mobile GNSS	Segmentation + homography + GeoJSON navigation	Yes, localization error	Not reported	Geo-referenced localization; not UAV occupancy/privacy

(Continued)

Table 1
(Continued)

Study/reference	View/setting	Primary function	Slot-association evaluation	Privacy evaluation	Scope relative to present study
Hukkelås and Lindseth [25]	General computer-vision imagery	Image anonymization and downstream vision utility	N/A	Yes	Privacy/utility evaluation outside parking
Present study	UAV framework; fixed-camera module audit	Detection + polygon association + geo-referenced records + anonymization	Yes; 1,712-slot module audit	Yes; calibration-separated	Integrated image-to-record framework with independently evaluated downstream modules

3. Proposed Methodology

3.1. Framework overview and evaluation decomposition

The proposed framework converts UAV-view imagery into slot-level parking records while keeping perception, slot association, geographic record linkage, and privacy-aware image processing conceptually separate. This decomposition is deliberate because each stage can fail independently and therefore requires its own evidence rather than being subsumed under detector mAP.

The occupancy branch first applies YOLOv8s to obtain vehicle or available/occupied visual predictions, depending on the experimental stage. These visual outputs are then compared with scene-specific parking-slot polygons defined in image coordinates. The association module produces a slot identifier and occupancy state, and that identifier links to the corresponding geo-referenced record used for downstream display or storage.

The privacy branch starts from the full input frame rather than from the output of the occupancy detector. Dedicated sensitive-region localization is therefore logically independent of successful parking or vehicle detection. Candidate face and license-plate regions are anonymized before storage or secondary visual use, allowing localization performance and anonymization coverage to be measured independently of occupancy accuracy. Because both branches consume the same unmodified frame in

parallel, anonymization is applied to the retained or exported frame and does not alter the input to the occupancy detector; the occupancy metrics reported below are therefore unaffected by the blur-area cost quantified in Section 4.6.

The training program comprises four progressively specialized stages with distinct dataset scopes and label spaces. Stages 1–3 are initialized independently from pretrained YOLOv8s weights, whereas Stage 4 is initialized from the best Stage 3 checkpoint for parking-specific binary available/occupied fine-tuning.

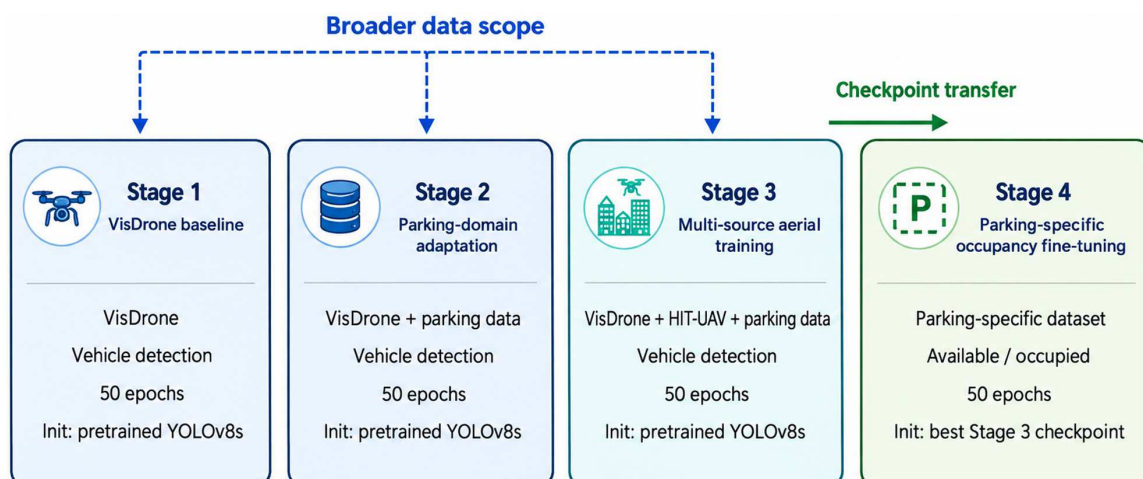
The evaluation is correspondingly decomposed into detector performance (Sections 4.2–4.4), polygon-based slot-association performance (Section 4.5), privacy-localization and anonymization performance (Section 4.6), and qualitative success/failure analysis (Section 4.7). This structure directly separates the system-level claims from the individual algorithms used to implement them.

Figure 1 summarizes the proposed system architecture. Figure 2 summarizes the progressive training stages and distinguishes the broadening of the training-data scope across Stages 1–3 from the checkpoint transfer used to initialize Stage 4.

3.2. Polygon-based slot association and geo-referenced mapping

Parking slots are digitized as polygons using QGIS (formerly Quantum GIS), following established Geographic Information System (GIS)-based parking-site representation practices [28].

Figure 2
Progressive training stages and model initialization strategy



The geometries are exported as GeoJSON records [24], and each slot is assigned a unique identifier together with its stored geographic coordinates. Parking-slot association itself is performed in image coordinates using scene-specific pixel-space slot polygons. The geographic coordinates are not used to geometrically transform YOLO detections during association; instead, the slot identifier provides a fixed link from the image-space association result to the corresponding geo-referenced record.

For vehicle-detection stages, let B_i denote the bounding box of detected vehicle i and S_j the pixel-space polygon of parking slot j . For the association protocol, the Intersection-over-Area (IoA) score is prospectively defined as the fraction of the detected vehicle box that lies inside the candidate slot. Normalizing by the detection area rather than by the slot area makes the score invariant to how tightly each space was digitized, so that a correctly parked vehicle activates its slot whether its polygon is drawn tightly or generously, and a small vehicle is not penalized for lying inside a large polygon. The association score is therefore defined as:

$$\text{IoA}(B_i, S_j) = \frac{\text{Area}(B_i \cap S_j)}{\text{Area}(B_i)} \quad (1)$$

Each detection is assigned to the slot with the maximum IoA when that score meets or exceeds the predefined association threshold τ . If no slot satisfies the threshold, the detection is marked unassigned. If multiple detections compete for the same slot, the highest-IoA detection is retained for the slot, and the event is recorded as a conflict for evaluation rather than being manually corrected.

The resulting slot state is defined as follows.

$$O_j = \begin{cases} 1, & \max_i \text{IoA}(B_i, S_j) \geq \tau, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

The primary operating point is $\tau = 0.50$: at least half of a detected vehicle box must lie within the assigned slot polygon. Section 4.5 additionally evaluates $\tau = 0.30, 0.40, 0.50, 0.60$, and 0.70 to characterize threshold sensitivity.

Parking-specific available/occupied predictions can be linked to fixed slot identifiers through the same scene geometry. The independent PKLot audit in Section 4.5 uses authoritative polygons and occupancy labels together with a fixed domain-matched vehicle-proposal stream to evaluate the box-to-slot mapping rule.

3.3. Privacy-aware visual processing

Sensitive-region localization operates on the complete input frame independently of the parking-occupancy detector. Localization is performed at an enlarged 1600-pixel input rather than at the 640-pixel occupancy resolution because sensitive regions in wide-area frames are frequently only a few tens of pixels across; whether the residual resolution is sufficient for anonymization to be meaningful is treated here as an empirical question and is answered directly by the size-stratified analysis in Section 4.6. The face configuration uses a YOLO-based face localizer at 1600-pixel inference, confidence 0.05, and 40% horizontal/vertical padding; the exact implementation resource is identified in the Data Availability Statement. Recent YOLO-family face-detection research demonstrates the effectiveness of scale- and occlusion-aware detection on challenging face benchmarks such as WIDER FACE [29]. The plate configuration uses a YOLOv8s-based license-plate localizer at 1600-pixel inference, confidence 0.30, 20% horizontal padding, and 40% vertical padding; the exact checkpoint resource is identified in the Data Availability Statement. Recent

YOLOv8-based license-plate detection and recognition research provides additional methodological context for this choice [30].

This full-frame branch covers sensitive regions throughout the scene, including areas outside mapped parking slots. Let Ω denote the union of the padded sensitive-region masks.

$$I'(x, y) = \begin{cases} (G_{\sigma, k} \times I)(x, y), & (x, y) \in \Omega, \\ I(x, y), & (x, y) \notin \Omega. \end{cases} \quad (3)$$

We selected the detector, threshold, inference strategy, and padding using calibration sets of 120 face images and 250 plate images. The selection objective first maximized calibration F2, then chose the smallest blur-area policy that retained at least 95% of the best calibration end-to-end coverage. Final scoring used separate sets of 500 WIDER FACE images [31] and the complete 882-image test split of the Keremkerke License Plate Object Detection benchmark [32]. Primary one-to-one localization matching used Intersection over Union (IoU) ≥ 0.30 , with IoU ≥ 0.50 reported as a secondary diagnostic. Predicted regions were expanded using the frozen padding rule and anonymized with a 31×31 Gaussian kernel ($\sigma_x = \sigma_y = 7.0$) [33]. End-to-end anonymization recall measures the proportion of ground-truth instances for which an automatically generated blur region covers at least 95% of the sensitive region.

4. Experimental Evaluation and Results

4.1. Datasets, label harmonization, and training protocol

The detector-training program uses VisDrone [5], HIT-UAV [6], and a parking-specific dataset [7]. Downstream validation uses PKLot parking-space polygons and occupancy labels [21], WIDER FACE annotations [31], and the Keremkerke License Plate Object Detection benchmark [32], a COCO-format derivative of the Augmented Startups Vehicle Registration Plates dataset. The plate benchmark contains 6,176 training, 1,765 validation, and 882 test images and is distributed under CC BY 4.0. We used a prospectively selected 250-image validation subset only for calibration; we reserved the full 882-image test split, containing 902 annotated plates, for final scoring. These sources have different roles and are not pooled into a single headline benchmark. Table 2 separates detector training from the independent slot and privacy audits.

Parking-relevant vehicle categories in Stages 1–3 were harmonized into a single vehicle class. Stage 4 reformulated the task as binary available/occupied prediction and was initialized from the best Stage 3 checkpoint. Stages 1–3 each began independently from pretrained YOLOv8s weights, whereas the Stage 3-to-Stage 4 transition used checkpoint transfer, as shown in Figure 2. Stages 1–3 are independently initialized from pretrained YOLOv8s weights while progressively broadening the training data from the VisDrone baseline to parking-domain adaptation and multi-source aerial training. Stage 4 performs parking-specific available/occupied fine-tuning using the best checkpoint obtained from Stage 3, representing the only sequential checkpoint transfer in the training pipeline.

Training was executed offline in a CUDA-enabled cloud environment using 640×640 inputs and a batch size of 16. Stages 1–4 each used a maximum training schedule of 50 epochs with an early-stopping patience of 15, as summarized in Figure 2. For these runs, we left the optimizer, learning-rate schedule, and augmentation settings at the Ultralytics release defaults and did not

Table 2
Training-stage and downstream-evaluation data sources, tasks, dataset splits, and roles

Stage	Dataset source(s)	Task/classes	Split/size	Purpose
Stage 1 – VisDrone baseline	VisDrone	Vehicle	Train: 6,471; validation: 548; test: 3,190	Generic aerial-vehicle baseline
Stage 2 – Parking-domain adaptation	VisDrone + parking-specific dataset	Vehicle	Source splits retained/merged	Parking-domain adaptation
Stage 3 – Multi-source aerial training	VisDrone + HIT-UAV + parking-specific dataset	Vehicle	Train: 8,260; validation: 1,083; remaining records: test	Broader multi-source aerial training
Stage 4 – Parking-specific occupancy fine-tuning	Parking-specific occupancy dataset	Available/occupied	1,216 images: train 851; validation 245; test 120	Binary parking-occupancy fine-tuning
Slot-association audit	PKLot	Slot state	29 registered images; 1,712 slots	Independent polygon-association audit
Face-privacy audit	WIDER FACE	Face region of interest (ROI)	120 calibration; 500 final images; 6,015 ground-truth (GT) instances	Independent anonymization audit
License-plate privacy audit	Keremberke plate benchmark	Plate ROI	250 calibration; 882-image test; 902 GT instances	Independent anonymization audit

tune them per stage; no stage-specific hyperparameter search was performed.

The parking-specific dataset contains 1,216 images partitioned into 851 training, 245 validation, and 120 test images. We fixed the Stage 4 training schedule and operating configuration before scoring the held-out 120-image test partition. Test AP@0.5 was 0.954 for available and 0.973 for occupied; overall precision was 0.984, recall was 0.986, and mAP@0.5 was 0.963. Throughout the paper, detector metrics are reported on the 0–1 scale to three decimal places, association and privacy rates are reported as percentages to two decimal places, and mean matched IoU is reported to four decimal places.

4.2. Results across training stages

Stage 1 (VisDrone baseline) established the generic aerial-vehicle baseline, with precision of 0.893, recall of 0.747, and mAP@0.5 of 0.825. Stage 2 (parking-domain adaptation) incorporated parking-specific imagery and recorded precision of 0.890, recall of 0.750, and mAP@0.5 of 0.830. Stage 3 (multi-source aerial training) broadened the training scope to VisDrone,

HIT-UAV, and parking imagery, yielding precision of 0.940, recall of 0.830, and mAP@0.5 of 0.900. Stage 4 (parking-specific occupancy fine-tuning) used the best Stage 3 checkpoint for available/occupied prediction and achieved held-out test precision of 0.984, recall of 0.986, and mAP@0.5 of 0.963. Stages 1–3 report recorded development/validation outputs, whereas Stage 4 provides the final result on the 120-image held-out test partition.

The verified training configuration and evaluation protocol are summarized in Table 3.

4.3. Final binary occupancy detector

For the final binary detector, the held-out 120-image test partition yielded AP@0.5 values of 0.954 for available and 0.973 for occupied. Overall precision was 0.984, recall was 0.986, and mAP@0.5 was 0.963. This final test result provides detector-level evidence for the implemented backbone; Sections 4.5 and 4.6 independently evaluate slot association and anonymization.

We fixed the Stage 4 checkpoint, stopping schedule, and operating configuration before scoring the test partition. Test images were not used for training, early stopping, hyperparameter

Table 3
Training configuration and evaluation protocol for the four-stage YOLOv8s training pipeline

Component	Configuration
Backbone	YOLOv8s
Input/batch	640 × 640; batch size 16
Compute environment	CUDA-enabled cloud GPU; NVIDIA T4 reported in project records
Initialization	Stages 1–3: pretrained YOLOv8s; Stage 4: best Stage 3 checkpoint
Epochs	Stages 1–4: maximum 50 epochs each; early-stopping patience 15
Stage 4 split	Train 851; validation 245; test 120
Reported Stage 4 metrics	Held-out test: P 0.984; R 0.986; mAP@0.5 0.963
Optimizer/augmentation	Ultralytics release defaults; no per-stage tuning
Downstream evaluation	Slot association and privacy-aware anonymization evaluated separately in Sections 4.5–4.7

selection, or the retrospective architecture benchmark in Section 4.4, preserving a clear separation between development and final evaluation.

Table 4 summarizes the precision, recall, and mAP@0.5 values recorded across the four training stages.

4.4. Retrospective detector benchmark

Table 5 places the implemented YOLOv8s backbone alongside YOLOv5s, YOLOv10s, YOLO11s, and RT-DETR on the parking-occupancy task at 640×640 input resolution. It is a retrospective descriptive benchmark: comparator values are recorded validation outputs, whereas the YOLOv8s row is the final Stage 4 held-out test record. Note that this experiment was not used to select the backbone or tune the reported test result.

Among the comparator validation records, YOLO11s achieved the highest mAP@0.5 (0.975). The final YOLOv8s held-out test record shows precision 0.984, recall 0.986, and mAP@0.5 0.963. YOLOv8s remains the operational backbone because it underpins the complete evaluated workflow and provides mature training, integration, and export support.

RT-DETR recorded mAP@0.5 = 0.944 with recall 0.838; YOLOv5s and YOLOv10s recorded mAP@0.5 values of 0.905 and 0.915, respectively. These results contextualize detector behavior across architectures but are not interpreted as a controlled superiority test because the evaluation records serve different roles. The paper's primary detector claim therefore rests on the held-out Stage 4 result, while the system contribution is supported by the independent downstream audits.

4.5. Polygon-based slot-association evaluation

The slot-association module was evaluated on 29 fixed-camera PKLot images registered to local validation frames through feature matching and verified homographies. Original PKLot parking-space contours and occupancy labels served as ground truth. PKLot was selected because it is, to our knowledge, the only large public parking benchmark that distributes authoritative per-space polygon geometry along with per-space occupancy labels; no public UAV parking dataset currently

provides the slot-level geometric ground truth required for a direct association audit. To isolate association behavior under a defined, domain-matched proposal stream, vehicle boxes were generated with a VisDrone YOLO detector using sliced inference. Accordingly, the audit quantifies the box-to-slot mapping stage under registered geometry; it does not conflate association accuracy with final binary-detector accuracy or live multi-site UAV performance.

At the prospectively fixed threshold $\tau = 0.50$, the audit covered 1,712 slots: 1,143 occupied and 569 available. It correctly activated 1,134 occupied slots, missed 9 occupied slots, falsely activated 6 available slots, and correctly retained 563 available slots. Slot-level accuracy was 99.12% (95% image-cluster bootstrap CI: 98.34%–99.66%). Occupied precision, recall, and F1 were 99.47%, 99.21%, and 99.34%; available precision, recall, and F1 were 98.43%, 98.95%, and 98.69%, respectively. The interval was estimated by nonparametric image-cluster bootstrap, resampling the 29 audit images with replacement for 2,000 replicates using the prospectively fixed random seed 10,637. Table 6 summarizes the primary module-level result.

The predefined sensitivity sweep yielded slot accuracies of 99.47%, 99.59%, 99.12%, 96.09%, and 88.49% at $\tau = 0.30, 0.40, 0.50, 0.60$, and 0.70 , respectively. Accuracy remained above 99% across $\tau = 0.30$ – 0.50 ; stricter overlap requirements increased missed occupied slots. The primary result remains the prospectively fixed $\tau = 0.50$. The detector generated 3,608 boxes: 1,140 canonical assignments, 2,020 unassigned detections, and 448 duplicate-slot conflicts. Unassigned detections include vehicles beyond the annotated PKLot spaces, while duplicate conflicts were resolved reproducibly by retaining the highest-IoA assignment. Inspection of the 15 misclassified slots identifies two distinct mechanisms. The six false activations occur when a vehicle parked in a neighboring space produces a box whose majority area falls inside the target polygon, as in Figure 3(c), where the winning score reaches IoA = 0.91. The nine missed occupied slots occur when an oblique viewing angle, partial occlusion, or an enlarged proposal box reduces the fraction of the box lying inside the correct polygon to just below τ , as in Figure 3(d), where IoA = 0.46. Both mechanisms are therefore near-threshold rather than gross mapping failures, which is consistent with the sweep: relaxing τ to 0.40 recovers several missed slots (99.59%), relaxing further to 0.30

Table 4
YOLOv8s performance across the four progressive training stages. Stages 1–3 report recorded development/validation metrics, whereas Stage 4 reports the final result on the held-out 120-image test partition

Exp.	Training focus	Output class(es)	Precision	Recall	mAP@0.5
Stage 1	VisDrone baseline	Vehicle	0.893	0.747	0.825
Stage 2	Parking-domain adaptation	Vehicle	0.890	0.750	0.830
Stage 3	Multi-source aerial training	Vehicle	0.940	0.830	0.900
Stage 4	Parking-specific occupancy fine-tuning	Available/occupied	0.984	0.986	0.963

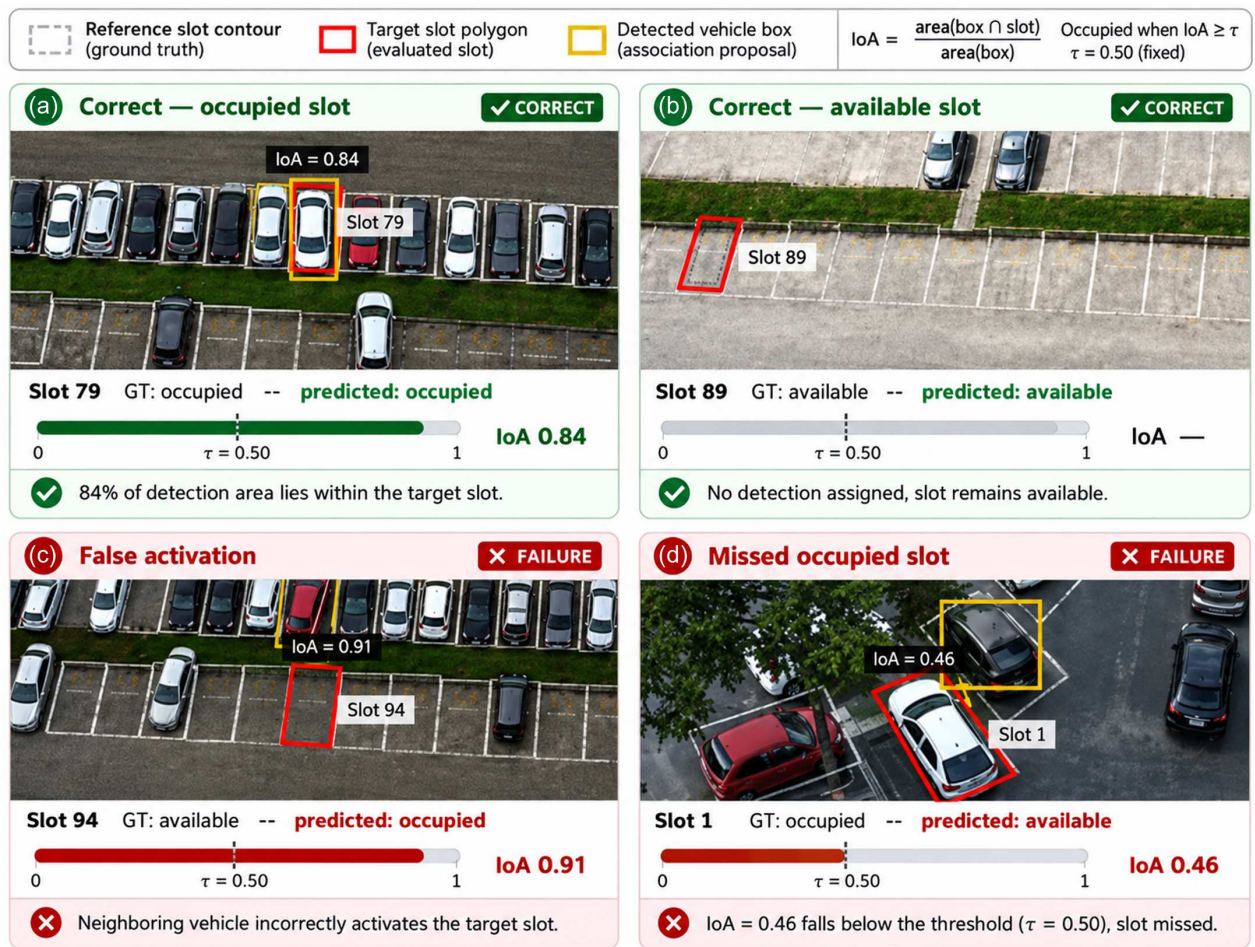
Table 5
Retrospective descriptive benchmark on the parking-occupancy task at 640×640 . Comparator rows report validation records, whereas YOLOv8s reports the final Stage 4 held-out test result

Model	Architecture type	Precision	Recall	mAP@0.5
YOLOv5s	Legacy CNN	0.909	0.921	0.905
YOLOv10s	NMS-free CNN	0.926	0.958	0.915
YOLO11s	Recent CNN	0.977	0.982	0.975
RT-DETR	Transformer	0.931	0.838	0.944
YOLOv8s (adopted backbone)	CNN with C2f	0.984	0.986	0.963

Table 6
Primary PKLot slot-state audit at the predefined IoA threshold $\tau = 0.50$

Outcome	Result	95% CI	Count/note
Slot accuracy	99.12%	98.34%–99.66%	1,697/1,712
Occupied precision	99.47%	–	1,134/(1,134 + 6)
Occupied recall	99.21%	98.39%–99.81%	1,134/1,143
Occupied F1	99.34%	–	–
Available precision	98.43%	–	563/(563 + 9)
Available recall	98.95%	–	563/569
Available F1	98.69%	–	–

Figure 3
Representative slot-association outcomes from the frozen PKLot audit at the predefined threshold ($\tau = 0.50$)



trades part of that gain for additional false activations (99.47%), and tightening τ to 0.70 converts many correct activations into misses (88.49%).

4.6. Privacy-aware anonymization evaluation

Privacy performance was evaluated using established public benchmarks with independently provided annotations, thereby ensuring an objective and reproducible assessment without project-specific manual ROI labeling. WIDER FACE and the public license-plate benchmark were selected because currently available UAV parking datasets do not provide face or

license-plate ground truth. Because these benchmarks predominantly contain ground-level imagery, their results should not be interpreted as direct estimates of privacy performance in UAV parking footage. They provide independently annotated test sets for evaluating the sensitive-region localization and anonymization mechanism under controlled conditions. The size-stratified analysis is therefore reported to expose the expected degradation for smaller targets and to clarify the remaining domain-transfer limitation. Accordingly, the blur-area measurements reported below can be interpreted as an upper-bound estimate relative to UAV parking imagery, where faces and license plates generally occupy a substantially smaller proportion of the frame.

Detector/model choice, confidence threshold, inference strategy, and padding were selected on calibration data only and frozen before final scoring. The face final set contained 6,015 ground-truth instances in 500 WIDER FACE images; the plate final set contained 902 ground-truth instances in 882 public benchmark images. No final-set example was used for model, threshold, resolution, ensemble, or padding selection. Binomial 95% confidence intervals were calculated using the Wilson score method. Because multiple sensitive targets can occur within one image, a post hoc robustness analysis additionally resampled images as clusters for 2,000 bootstrap replicates using seed 10,637, retaining all targets in every sampled image; the target-level Wilson intervals remain the primary uncertainty estimates. Ground-truth regions were stratified by equivalent bounding-box side length, $s = \sqrt{(w_{GT} \times h_{GT})}$: small ($s < 32$ pixels), medium ($32 \leq s < 96$ pixels), and large ($s \geq 96$ pixels). For faces, YOLO produced 8,576 predictions: 5,486 true positives (TP), 3,090 false positives (FP), and 529 false negatives (FN). Precision was 63.97%, recall 91.21% (95% CI: 90.46–91.90%), F1 75.20%, F2 84.05%, and mean matched IoU 0.7288. End-to-end anonymization recall was 91.37% (5,496/6,015; 95% CI: 90.64–92.06%), with a mean blurred image area of 16.03%. End-to-end successes (5,496) slightly exceed one-to-one localization true positives (5,486) because coverage is evaluated against the union of all padded blur regions: a small number of ground-truth faces that failed the one-to-one IoU ≥ 0.30 assignment were nevertheless fully covered by the padded region of an adjacent or overlapping prediction. Relative to the same-data baseline, the selected operating point increased localization recall by 16.48 percentage points and end-to-end recall by 33.10 percentage points. This configuration intentionally prioritizes sensitive-region coverage while quantifying the associated blur-area cost.

For license plates, the selected YOLOv8s-based license-plate localizer produced 954 predictions, with 858 TP, 96 FP, and 44 FN. Precision was 89.94% (95% CI: 87.87%–91.69%), recall 95.12% (95% CI: 93.51%–96.35%), F1 92.46%, F2 94.04%, and mean matched IoU 0.7292. End-to-end anonymization recall was 92.24% (832/902; 95% CI: 90.31%–93.81%), with a mean blurred area of 5.90%. The same-data baseline reached 69.40% localization recall and 37.47% end-to-end recall. Both the frozen face and plate baselines use 1280-pixel inference, a confidence threshold of 0.20, and 10% horizontal/vertical padding.

Size-stratified end-to-end anonymization succeeded for 322/337 large faces (95.55%, 95% Wilson CI: 92.79–97.28%), 1,115/1,167 medium faces (95.54%, 94.20–96.59%), and 4,059/4,511 small faces (89.98%, 89.07–90.82%). Plate results were 36/44 for large instances (81.82%, 68.04–90.49%), 679/707 for medium instances (96.04%, 94.34–97.25%), and 117/151 for small instances (77.48%, 70.19–83.42%). These strata identify small-target coverage as a priority for the next development stage; the large-plate estimate will benefit from a larger evaluation sample.

In the image-cluster robustness analysis, the 95% bootstrap intervals for end-to-end anonymization recall were 88.73%–93.76% for faces and 90.54%–93.95% for license plates. These sensitivity intervals account for within-image dependence and do not replace the prespecified target-level Wilson intervals. Table 7 summarizes the final privacy-preservation results and the corresponding same-dataset baseline comparisons.

4.7. Qualitative and quantitative diagnostic analysis

Slot panels were selected deterministically from frozen association outputs after all operating settings had been fixed. The selection includes two correct slot states and two representative

Table 7
Final privacy-preservation results and same-dataset baseline comparisons. E2E denotes end-to-end anonymization recall, defined as the proportion of ground-truth (GT) regions achieving at least 95% coverage

Modality/ setting	P/R/F2 (%)	E2E recall (%)	Blur area (%)
Face, selected	63.97/91.21/84.05	91.37	16.03
Face, baseline	90.66/74.73/77.45	58.27	6.30
Plate, selected	89.94/95.12/94.04	92.24	5.90
Plate, baseline	74.70/69.40/not recorded	37.47	3.02

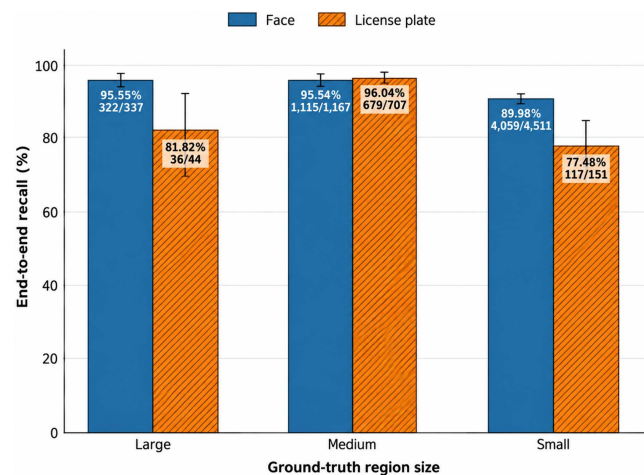
error states. The privacy diagnostic focuses on size-stratified end-to-end recall with exact denominators and Wilson uncertainty, complementing rather than duplicating the aggregate results in Table 7.

Figure 3 shows correct occupied and available decisions together with a false activation and a missed occupied slot. Because the displayed result is a complete slot output under the fixed proposal stream, an error may reflect either the incoming box or the polygon assignment. Figure 4 visualizes face and license-plate end-to-end recall by object size, with exact successful/total counts and Wilson intervals that make scale-dependent behavior explicit.

5. Discussion

Taken together, the results support the framework’s central image-to-record contribution: visual predictions are converted into traceable slot states through an explicit association rule, while sensitive regions are processed through an independent branch with measurable coverage. At the predefined $\tau = 0.50$, the registered PKLot module audit achieved 99.12% slot-level accuracy, with occupied and available F1 scores of 99.34% and 98.69%, respectively. Under the fixed domain-matched proposal stream, this result demonstrates reliable box-to-slot mapping and is appropriately interpreted as module-level evidence rather than a claim about unconstrained live UAV deployment.

Figure 4
Size-stratified end-to-end anonymization recall from the frozen final evaluation



The association diagnostics clarify why the method remains auditable in crowded scenes. Detections outside annotated spaces remain unassigned, multiple detections competing for one slot are resolved by maximum IoA, and both events are retained in the evaluation record rather than manually corrected. Accuracy remained above 99% across $\tau = 0.30\text{--}0.50$, indicating a stable operating region around the predefined threshold while preserving deterministic error analysis.

The privacy audit demonstrates the value of optimizing sensitive-region coverage while reporting its visual cost. Relative to the same-data baselines, the selected face configuration increased end-to-end recall from 58.27% to 91.37%, with mean blurred area increasing from 6.30% to 16.03%; the selected plate configuration increased end-to-end recall from 37.47% to 92.24%, with mean blurred area of 5.90%. Image-cluster robustness intervals supported the interpretation of the prespecified Wilson intervals. Size stratification localizes the remaining challenge: small-face recall was 89.98%, small-plate recall was 77.48%, and the large-plate estimate of 81.82% was based on only 44 instances, whereas both modalities exceeded 95% in the best-supported medium or large strata.

Detector evidence has a separate, clearly defined role. The final Stage 4 YOLOv8s occupancy detector achieved held-out test precision of 0.984, recall of 0.986, and mAP@0.5 of 0.963. The retrospective benchmark records YOLO11s as the highest-validation-mAP comparator, but it did not drive model selection or test-time tuning. This interpretation supports the implemented backbone without extending the evidence to an unsupported universal architecture ranking.

Panels in Figure 3 show (a) a correctly identified occupied slot, (b) a correctly retained available slot, (c) a false activation caused by assignment from a neighboring vehicle, and (d) a missed occupied slot caused by insufficient overlap. White outlines denote reference parking-space contours, red outlines indicate the evaluated target slot, and yellow boxes denote detected vehicles. The reported IoA is the fraction of the detected vehicle-box area contained within the target-slot polygon. The examples illustrate both successful associations and the principal failure modes observed in the quantitative analysis of Section 4.5.

Ground-truth regions are grouped by equivalent box side length $s = \sqrt{(w_{GT} \times h_{GT})}$: large ≥ 96 pixels, medium 32–<96 pixels, and small <32 pixels. Bars report observed recall and successful/total counts; error bars show 95% Wilson intervals. Face counts are 322/337, 1,115/1,167, and 4,059/4,511; license-plate counts are 36/44, 679/707, and 117/151.

5.1. Scope and future work

The present study intentionally prioritizes controlled, auditable module-level validation. Registered PKLot geometry isolates the box-to-slot association mechanism, public privacy benchmarks provide independently reproducible ground truth, and the held-out Stage 4 evaluation establishes the implemented detector's operating performance. This design enables assessment of each component on its own evidence without conflating detector, association, and anonymization errors.

Building on this validated foundation, future work will extend the framework to multi-site UAV video with temporal tracking, aerial privacy evaluation, resolution-adaptive processing for small targets, controlled same-split architecture comparisons, and latency and energy profiling on embedded hardware. These extensions will broaden the framework's deployment-scale evidence while preserving the calibration-separated and module-specific evaluation discipline established in this study.

6. Conclusion

This study establishes a modular privacy-aware UAV parking framework that combines detector-based perception, box-normalized polygon association, geo-referenced slot records, and independent full-frame anonymization. The final Stage 4 YOLOv8s occupancy detector achieved held-out test precision of 0.984, recall of 0.986, and mAP@0.5 of 0.963. Under registered fixed-camera geometry and a fixed proposal stream, the 1,712-slot association audit achieved 99.12% accuracy. Calibration-separated final evaluation achieved end-to-end anonymization recall of 91.37% for faces and 92.24% for license plates. Together, these results provide a traceable, independently evaluated path from visual input to slot-level records and privacy-aware output without conflating detector, association, and anonymization evidence.

Future research will build on this validated foundation by extending the framework toward continuous, multi-site UAV deployment while preserving the rigorous, calibration-separated evaluation protocol established in this study.

Funding Support

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia, under Grant No. KFU263592.

Ethical Statement

This study used only pre-existing, publicly available, and previously published benchmark datasets for academic computer-vision research. The authors recruited no participants, collected no new human-subject data or facial images, and had no direct interaction with individuals. The datasets were accessed and used in accordance with their respective terms of use, licensing conditions, and established research practices. No identifiable facial images are reproduced in the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in VisDrone-DET2019 [GitHub] at <https://github.com/VisDrone/VisDrone-Dataset>, in HIT-UAV [GitHub] at <https://github.com/suojiashun/HIT-UAV-Infrared-Thermal-Dataset>, in the parking-specific occupancy dataset [GitHub] at <https://github.com/Dineshmanchala19/parking-space-detector>, in PKLot at <https://web.inf.ufpr.br/vri/databases/parking-lot-database/>, in WIDER FACE at <https://shuoyang1213.me/WIDERFACE/>, in the License Plate Object Detection benchmark [Hugging Face] at <https://huggingface.co/datasets/keremberke/license-plate-object-detection>, in yolov8-face [GitHub] at <https://github.com/lindevs/yolov8-face>, and in Hugging Face at <https://huggingface.co/Babblu2821/alpr-plate-detector>.

Author Contribution Statement

Aisha Alahmed: Software, Investigation, Data curation, Writing – original draft, Visualization. **Saed Alqaraleh:** Conceptualization, Writing – review & editing, Supervision. **Qanout Alsharida:** Methodology, Software, Validation, Data curation, Writing –

original draft. **Mariam Alahmed:** Software, Formal analysis, Writing – original draft. **Leen Boodai:** Software, Validation, Writing – original draft. **A. H. Abdul Hafiz:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition.

References

- [1] Xiao, R. I., & Jaller, M. (2025). Prediction framework for parking search cruising time and emissions in dense urban areas. *Transportation*, 52(4), 1289–1317. <https://doi.org/10.1007/s11116-023-10455-4>
- [2] Channamallu, S. S., Kermanshachi, S., Rosenberger, J. M., & Pamidimukkala, A. (2026). Smart parking systems: A comprehensive review of digitalization of parking services. *Green Energy and Intelligent Transportation*, 5(1), 100293. <https://doi.org/10.1016/j.geits.2025.100293>
- [3] Kujawski, A., & Nürnberg, M. (2023). Analysis of the potential use of unmanned aerial vehicles and image processing methods to support road and parking space management in urban transport. *Sustainability*, 15(4), 3285. <https://doi.org/10.3390/su15043285>
- [4] Sarkar, S., Totaro, M. W., & Elgazzar, K. (2019). Intelligent drone-based surveillance: Application to parking lot monitoring and detection. In *Unmanned Systems Technology XXI*, 11021, 13–19. <https://doi.org/10.1117/12.2518320>
- [5] Du, D., Zhu, P., Wen, L., Bian, X., Lin, H., Hu, Q., ..., & Liu, Z. (2019). VisDrone-DET2019: The vision meets drone object detection in image challenge results. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 213–226. <https://doi.org/10.1109/ICCVW.2019.00030>
- [6] Suo, J., Wang, T., Zhang, X., Chen, H., Zhou, W., & Shi, W. (2023). HIT-UAV: A high-altitude infrared thermal dataset for Unmanned Aerial Vehicle-based object detection. *Scientific Data*, 10(1), 227. <https://doi.org/10.1038/s41597-023-02066-6>
- [7] Manchala, D. (2024). Parking Space Detection using YOLOv8 [Data set] GitHub. Retrieved from: <https://github.com/Dineshmanchala19/parking-space-detector>
- [8] Leng, J., Ye, Y., Mo, M., Gao, C., Gan, J., Xiao, B., & Gao, X. (2024). Recent advances for aerial object detection: A survey. *ACM Computing Surveys*, 56(12), 296, 1–36. <https://doi.org/10.1145/3664598>
- [9] Hua, W., & Chen, Q. (2025). A survey of small object detection based on deep learning in aerial images. *Artificial Intelligence Review*, 58(6), 162. <https://doi.org/10.1007/s10462-025-11150-9>
- [10] Tang, G., Ni, J., Zhao, Y., Gu, Y., & Cao, W. (2024). A survey of object detection for UAVs based on deep learning. *Remote Sensing*, 16(1), 149. <https://doi.org/10.3390/rs16010149>
- [11] Sun, Y., Cao, B., Zhu, P., & Hu, Q. (2022). Drone-based RGB-infrared cross-modality vehicle detection via uncertainty-aware learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(10), 6700–6713. <https://doi.org/10.1109/TCSVT.2022.3168279>
- [12] Jocher, G., Chaurasia, A., & Qiu, J. (2023). *Ultralytics YOLOv8* (Version 8.0.0) [Computer software]. *Ultralytics YOLOv8*. Retrieved from: <https://github.com/ultralytics/ultralytics>
- [13] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, 5(4), 1680–1716. <https://doi.org/10.3390/make5040083>
- [14] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-time end-to-end object detection. *Advances in Neural Information Processing Systems*, 37, 107984–108011. <https://doi.org/10.52202/079017-3429>
- [15] Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G., Dang, Q., ..., & Chen, J. (2020). DETRs beat YOLOs on real-time object detection. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 16965–16974. <https://doi.org/10.1109/CVPR52733.2024.01605>
- [16] Sapkota, R., Flores-Calero, M., Qureshi, R., Badgujar, C., Nepal, U., Poulouse, A., ..., & Karkee, M. (2025). YOLO advances to its genesis: A decadal and comprehensive review of the You Only Look Once (YOLO) series. *Artificial Intelligence Review*, 58(9), 274. <https://doi.org/10.1007/s10462-025-11253-3>
- [17] Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE*, 111(3), 257–276. <https://doi.org/10.1109/JPROC.2023.3238524>
- [18] Li, Y., Miao, N., Ma, L., Shuang, F., & Huang, X. (2023). Transformer for object detection: Review and benchmark. *Engineering Applications of Artificial Intelligence*, 126, 107021. <https://doi.org/10.1016/j.engappai.2023.107021>
- [19] Zhao, N., Wang, K., Yang, J., Luan, F., Yuan, L., & Zhang, H. (2024). Cmca-yolo: A study on a real-time object detection model for parking lot surveillance imagery. *Electronics*, 13(8), 1557. <https://doi.org/10.3390/electronics13081557>
- [20] De Almeida, P. R. L., Alves, J. H., Parpinelli, R. S., & Barddal, J. P. (2022). A systematic review on computer vision-based parking lot management applied on public datasets. *Expert Systems with Applications*, 198, 116731. <https://doi.org/10.1016/j.eswa.2022.116731>
- [21] De Almeida, P. R., Oliveira, L. S., Britto, A. S., Jr., Silva, E. J., Jr., & Koerich, A. L. (2015). PKLot–A robust dataset for parking lot classification. *Expert Systems with Applications*, 42(11), 4937–4949. <https://doi.org/10.1016/j.eswa.2015.02.009>
- [22] Rafique, S., Gul, S., Jan, K., & Khan, G. M. (2023). Optimized real-time parking management framework using deep learning. *Expert Systems with Applications*, 220, 119686. <https://doi.org/10.1016/j.eswa.2023.119686>
- [23] Grbić, R., & Koch, B. (2023). Automatic vision-based parking slot detection and occupancy classification. *Expert Systems with Applications*, 225, 120147. <https://doi.org/10.1016/j.eswa.2023.120147>
- [24] Laptiev, O., Thuruthel Murali, A., Saab, N., Soltanov, N., & Paulauskaitė-Tarasevičienė, A. (2026). An integrated vision–mobile fusion framework for real-time smart parking navigation. *Logistics*, 10(4), 84. <https://doi.org/10.3390/logistics10040084>
- [25] Hukkelås, H., & Lindseth, F. (2023). Does image anonymization impact computer vision training? In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 140–150. <https://doi.org/10.1109/CVPRW59228.2023.00019>
- [26] De Coninck, S., Wang, W. C., Leroux, S., & Simoens, P. (2024). Privacy-preserving visual analysis: Training video obfuscation models without sensitive labels. *Applied Intelligence*, 54(8), 6041–6052. <https://doi.org/10.1007/s10489-024-05489-9>
- [27] Zou, C., Liu, Y., Yang, Y., Zhou, C., Yu, Y., & Shang, Y. (2025). A privacy-preserving license plate encryption scheme

- based on an improved YOLOv8 image recognition algorithm. *Signal Processing*, 230, 109811. <https://doi.org/10.1016/j.sigpro.2024.109811>
- [28] Kaplan, T., & Can, E. (2025). Mitigating urban traffic: Identifying potential parking lots using GIS and AHP methods in Yalova, Türkiye. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 14(4), 2577–2609. <https://doi.org/10.17798/bitlisfen.1759305>
- [29] Yu, Z., Huang, H., Chen, W., Su, Y., Liu, Y., & Wang, X. (2024). YOLO-FaceV2: A scale and occlusion aware face detector. *Pattern Recognition*, 155, 110714. <https://doi.org/10.1016/j.patcog.2024.110714>
- [30] Moussaoui, H., Akkad, N. E., Benslimane, M., El-Shafai, W., Baihan, A., Hewage, C., & Rathore, R. S. (2024). Enhancing automated vehicle identification by integrating YOLO v8 and OCR techniques for high-precision license plate detection and recognition. *Scientific Reports*, 14(1), 14389. <https://doi.org/10.1038/s41598-024-65272-1>
- [31] Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). WIDER FACE: A face detection benchmark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5525–5533. <https://doi.org/10.1109/CVPR.2016.596>
- [32] Keremberke. (2022). *License Plate Object Detection [Data set]*. *Hugging Face*. Retrieved from: <https://huggingface.co/datasets/keremberke/license-plate-object-detection>
- [33] Park, S., Ma, J., Oh, S. W., & Kim, W. S. (2025). Blur image representation in defocus using generalized Gaussian kernel. *Journal of Information Display*, 26(3), 321–328. <https://doi.org/10.1080/15980316.2025.2483195>

How to Cite: Alahmed, A., Alqaraleh, S., Alsharida, Q., Alahmed, M., Boodai, L., & Hafez, A. H. A. (2026). A Privacy-Aware UAV Parking Framework with Polygon-Based Slot Association and Full-Frame Anonymization. *Artificial Intelligence and Applications*. <https://doi.org/10.47852/bonviewAIA620210637>