

DRGD: A Deterministic Color–Geometry–PnP Gate Detector for Reproducible CPU-Based Evaluation on Indoor and Outdoor Sequences

Víctor Alejandro Méndez-López¹, Carlos Soubervielle-Montalvo¹, José Martínez-Carranza^{2,*}

¹ Universidad Autónoma de San Luis Potosí (UASLP), Faculty of Engineering, Mexico

² Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Computer Science Department, Puebla, Mexico

alejandromdzlpz.uaslp@gmail.com, carlos.soubervielle@uaslp.mx, carranza@inaoep.mx

Abstract. This paper presents DRGD (Drone Racing Gate Detector), a modular classical computer-vision detector for single-gate localization in indoor and outdoor drone-gate sequences. Its contribution is system-level rather than operator-level: standard color, morphology, geometry, and PnP components are integrated into a deterministic color–geometry–PnP pipeline with frozen domain configurations, split-disciplined evaluation, staged ablation, and explicit reporting of negative evidence. DRGD follows a two-stage classical detection structure: Stage 1 converts full-frame color evidence into mask-based color–geometry proposals, while Stage 2 applies geometric constraints, PnP plausibility validation, and top- K candidate selection before emitting the final bounding box. The detector is evaluated on a custom four-sequence 640×480 dataset comprising SQ3, SQ4, SQ5, and SQ7, using a split-disciplined CPU protocol with calibration, validation, configuration freezing, A0–A3 ablation, and a single-run final test. After indoor calibration, the frozen 1d9 indoor configuration achieved 0.953 F1 and 45.91 FPS on validation, and 0.905 F1 and 48.46 FPS on the indoor final test. The frozen 1d11 outdoor configuration achieved 0.891 F1 and 42.08 FPS on validation, but reached 0.474 F1 globally on the outdoor final test, mainly due to the lower SQ7 result. To provide an external quantitative reference, a YOLO11n baseline was fine-tuned on the same splits and evaluated with the same frame-level test protocol; it reached 0.998 indoor F1 and 0.840 outdoor F1, improving SQ7 recall while still producing false positives on gate-absent SQ7 frames. The results show

that DRGD is accurate and real-time under controlled indoor conditions, remains usable on SQ5 outdoors, and identifies sequence-dependent outdoor robustness as the main direction for future embedded deployment.

Keywords. Classical object detection, gate detection, resource-constrained vision, CPU benchmarking, reproducible evaluation.

1 Introduction

Classical computer-vision pipelines remain relevant when deterministic behavior, transparency, low implementation complexity, and explicit stage-level profiling are important. These properties are valuable for embedded perception, where an algorithm may later need to be simplified, ported, or accelerated by embedded systems, including SoC–FPGA platforms. In this work, the CPU implementation is used as a reproducible reference for the same fixed detector stages.

Gate detection is a useful case for this design space. The object has a simple physical structure, but its apparent scale, contour completeness, color response, and background separation change with viewpoint, distance, illumination, and scene clutter. A modular detector can expose why a candidate was accepted or rejected, which stage dominates the decision, and which sequence conditions

remain unresolved. This makes the detector suitable for reproducible design analysis rather than only for final-score reporting.

This paper reports DRGD as a standalone classical detector focused on the detector architecture, the CPU-based experimental workflow, and the calibrated behavior of a color–geometry–PnP pipeline on indoor and outdoor sequences. DRGD is a system-level contribution rather than a new low-level image-processing operator: HSV masking, morphology, contour filtering, and PnP are standard components, while the novelty lies in their deterministic two-stage integration for gate detection. Stage 1 generates color–geometry gate proposals from full-frame evidence, and Stage 2 applies PnP validation and top- K candidate selection before emitting the final bounding box.

The main contributions of this work are: (i) a modular two-stage classical gate detector with manually configurable static parameters and frozen indoor/outdoor configurations for domain-specific execution; (ii) a split-disciplined evaluation protocol with calibration, validation, configuration freezing, validation-stage ablation, and a one-shot final test; (iii) an ordered A0–A3 ablation that isolates the color-mask, geometric-proposal, PnP-validated, and top- K early-accept operating variants; (iv) a final-test diagnostic analysis that documents strong indoor behavior, usable SQ5 outdoor behavior, and the SQ7 held-out limitation while retaining negative evidence; and (v) a YOLO11n reference baseline trained and tested on the same data partitions and evaluated using DRGD-compatible frame-level precision, recall, F1, and throughput metrics.

2 Related Work

Classical single-object detection pipelines are often structured as explicit stages for image evidence extraction, candidate generation, and interpretable verification. HOG-based detection [1] and deformable part models [2] are representative examples of this staged design, where features, scoring, and selection can be analyzed separately. Pose-aware detection for known objects, as illustrated by textureless-object pipelines such as LINEMOD, further shows that structural constraints

and geometric verification are useful when the target has a known physical form [5]. The proposed DRGD detector follows this philosophy for a single known object class: a colored square racing gate. Instead of using a learned dense detector, it combines color evidence, geometric filtering, and PnP-based plausibility verification.

Drone racing provides a useful motivating scenario because gates have constrained geometry, while their apparent scale, viewpoint, illumination, blur, and visibility can vary significantly [4]. Recent autonomous drone-racing perception increasingly relies on learned gate-detection or gate-perception modules. GateNet estimates gate center, distance, and orientation from a fish-eye RGB camera using a lightweight CNN architecture [6]; AlphaPilot integrates learned visual abstraction with state estimation, mapping, and trajectory planning in a complete autonomous racing system [3]; continual-learning gate perception has been studied for dynamic lighting conditions [7]; and recent nano-drone work combines a compact deep gate-detection front-end with monocular gate-to-gate visual navigation [8].

These works provide the learned-perception context for the YOLO11n reference baseline reported in this paper. In contrast, DRGD is not intended to replace learned gate detectors; rather, it is positioned as a deterministic classical reference pipeline for transparent color–geometry–PnP gate detection under frozen configurations, split-disciplined evaluation, and resource-constrained deployment scenarios.

3 Proposed DRGD Detector

3.1 Modular Pipeline

DRGD is implemented as a configuration-driven classical pipeline. Fig. 1 summarizes the logic as two sequential stages. Stage 1 receives an RGB frame, applies lightweight preprocessing, converts the image to HSV, and forms a binary color mask from two calibrated gate-color bands. The mask is cleaned using morphology and hole filling. Contours extracted from the cleaned mask define candidate ROIs, which are filtered by

area, aspect ratio, extent, solidity, border contact, and rectangular support. Rectangle and corner evidence then convert the remaining ROIs into gate candidates.

Stage 2 receives these candidates, applies PnP plausibility validation against the known gate model, ranks the surviving hypotheses, and emits at most one final bounding box using top- K candidate validation. This stage separation is relevant for CPU profiling and for later migration to resource-constrained embedded platforms, including SoC–FPGA implementations, because color masking, morphology, contour filtering, PnP validation, and top- K selection expose clear computational boundaries.

3.2 Candidate Scoring and Output

Candidate records in the frozen implementations include ROI location and coverage, border contact, geometry scores such as rectangularity, solidity, aspect ratio and corner evidence, PnP status and reprojection error, and a final selection score. These fields connect Stage 1 proposal quality with Stage 2 validation quality without turning DRGD into a learned detector. The gate model used by PnP is a fixed square-gate geometry with ordered corner correspondences. A candidate is physically consistent with this model when its image corners can be explained by a plausible camera-to-gate pose with acceptable reprojection error and validation status.

The frozen deployable configuration uses top- K candidate validation and final candidate selection. Candidates are first ranked by the available color–geometry and validation evidence. Then, at most the top- K candidates are considered for final acceptance; validation rules check the enabled geometric and PnP conditions, and output rules emit one bounding box only when an accepted candidate exists. If no candidate satisfies these rules, the frame is reported as no detection. DRGD uses two static, manually calibrated parameter sets, one indoor and one outdoor, while preserving the same algorithmic structure. The frozen 1d9 indoor and 1d11 outdoor values are reported in Appendix A.

Table 1. Custom dataset and contiguous split policy used for DRGD evaluation. Counts correspond to usable frames after excluding the 4-frame gaps

	SQ3	SQ4	SQ5	SQ7	Total
Domain	Indoor	Indoor	Outdoor	Outdoor	All
Usable	3310	2210	2050	1920	9490
Train (80%)	2648	1768	1640	1536	7592
Val. (10%)	331	221	205	192	949
Test (10%)	331	221	205	192	949

4 Methodology

The methodology defines the dataset partitions, scoring rules, execution platform, and staged evaluation order used to report DRGD.

Its purpose is to keep parameter calibration, validation, ablation analysis, and the final test separated, so that the reserved test split is used only for final reporting. All development and evaluation runs were executed in CPU-only mode on the same laptop-class platform: Intel® Core™ i7-8565U CPU at 1.80 GHz, 16 GB of system memory, and a 64-bit Windows environment. The platform includes Intel UHD Graphics 620, but the reported DRGD timing corresponds to CPU execution.

4.1 Custom Dataset and Split Policy

The experiments use a custom four-sequence drone-gate dataset with 640×480 RGB frames. SQ3 and SQ4 define the indoor domain, while SQ5 and SQ7 define the outdoor domain.

The frozen records store evaluated frames as RGB PNG images; camera make/model, lens, exposure, and shutter metadata are not encoded in the supplied manifests. The reproducible dataset description is therefore specified at the frame, split, annotation, and sequence-condition level. Indoor sequences were captured under controlled illumination, whereas outdoor sequences used natural daylight with stronger background, scale, and transition variability. To avoid cross-sequence leakage, each sequence is independently split into contiguous train, validation, and test partitions. Table 1 summarizes usable-frame counts and the split policy. Fig. 2 shows one representative frame from each sequence.

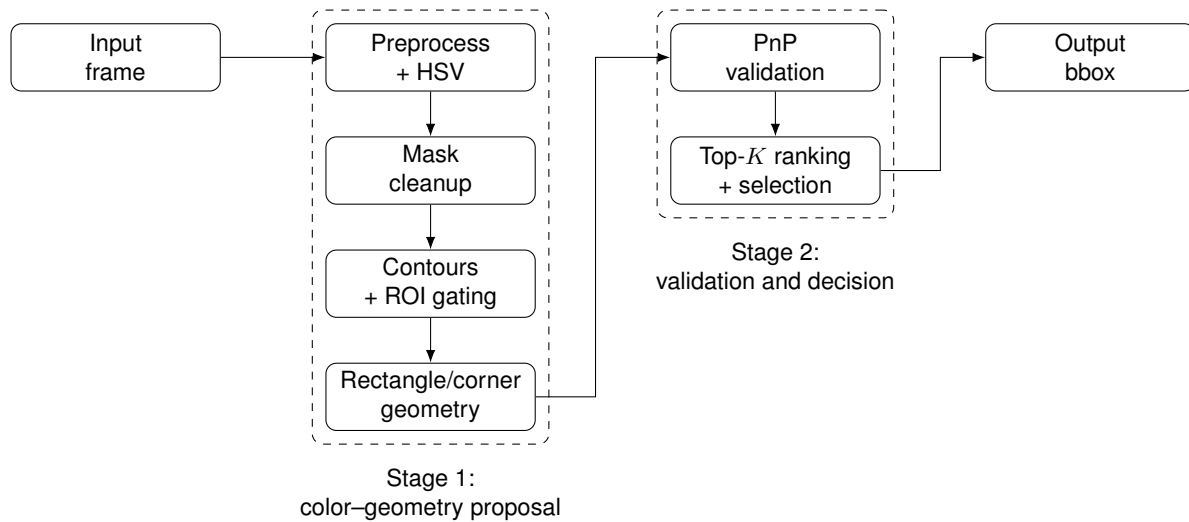


Fig. 1. Overview of the two-stage DRGD pipeline. Stage 1 generates bounded color-geometry gate proposals from full-frame evidence, and Stage 2 applies PnP validation plus top- K candidate validation before emitting the final bounding box

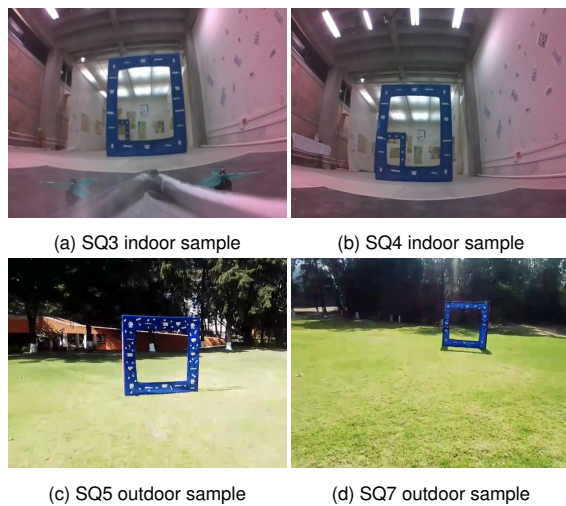


Fig. 2. Representative frames from the custom four-sequence dataset. The indoor pair (SQ3, SQ4) and outdoor pair (SQ5, SQ7) show the main domain contrast addressed by the detector

Ground-truth annotations use one manually verified bounding box for the visible gate when a gate is present. Frames without a visible gate are represented by empty label files and are used as negative evidence for false-positive and true-negative counting.

The annotation target is the visible gate region used for frame-level IoU scoring, and each sequence was frozen through the v5.1 split and manifest files before final testing.

4.2 Metrics

Metrics are computed frame by frame. For each frame, the detector may either emit one bounding box or emit no detection. If a gate is present and the emitted box overlaps the ground-truth box with $\text{IoU} \geq 0.50$, the frame is counted as a true positive (TP). If a gate is present but no valid box is emitted, or the emitted box has $\text{IoU} < 0.50$, the frame is counted as a false negative (FN). If no gate is present but the detector emits a box, the frame is counted as a false positive (FP). If no gate is present and no box is emitted, the frame is counted as a true negative (TN). The reported indicators are accepted frames, TP, FP, FN, TN, precision, recall, F1, accuracy, and end-to-end FPS. In splits where all frames contain a gate, TN is absent; therefore, recall and F1 are emphasized over accuracy alone.

Timing is computed over the same evaluated split. For each run, N denotes the number of processed frames, and t_{e2e} denotes the total elapsed time required to process those

frames. This elapsed time includes image input and orchestration overhead, so the reported end-to-end FPS reflects the complete CPU evaluation loop rather than only the detector core. Table 2 summarizes the formulas used in the experiment logs.

4.3 Experimental Protocol

The experiments followed a staged calibration–validation–freeze–ablation–test workflow. Indoor calibration used the training split, and 1d9 was selected for freezing after giving the best validation behavior among the tested indoor iterations. Outdoor calibration then used a separate configuration file while keeping the shared detector logic unchanged; 1d11 was frozen after outdoor validation.

Later outdoor attempts, 1d12–1d14, did not improve validation behavior and were not adopted. After freezing, the A0–A3 ablation was performed only on the validation split to analyze the ordered contribution of detector stages, without retuning the frozen configurations. Finally, each frozen configuration was evaluated once on its corresponding reserved test split, with no further tuning. The final test therefore uses the frozen parameter values listed in Appendix Appendix A, rather than parameters selected from test-set behavior.

Table 3 summarizes the stage/split policy and its relation to the main workflow. Calibration tunes the static indoor and outdoor configuration files. Validation checks calibrated configurations before freezing. Ablation uses validation frames to isolate A0–A3 behavior. The final test uses only the test split and is executed once after the corresponding freeze for reporting.

The A0–A3 ablation follows the natural execution order of DRGD. Each variant adds one design layer over the previous one to measure the contribution of color-mask evidence, geometric proposal generation, PnP-based physical plausibility, and the final top- K early-accept operating mode. Table 4 defines the four validation variants.

Table 2. Detection and timing metrics used in the evaluation

Metric	Formula
Precision	$TP/(TP + FP)$
Recall	$TP/(TP + FN)$
F1	$2PR/(P + R)$
Accuracy	$(TP + TN)/(TP + FP + FN + TN)$
End-to-end FPS	N/t_{e2e}

Table 3. Stage/split policy and main DRGD experimental workflow

Stage	Split	Main output
Calibration	Training	Indoor/outdoor parameter files
Validation	Validation	Frozen configurations
Ablation	Validation	A0–A3 stage contribution
Final test	Test	Reserved-test performance

Table 4. Ordered DRGD ablation variants used for validation-stage analysis

Variant	Configuration added
A0	HSV mask and binary-mask cleanup
A1	A0 plus contour and rectangle/corner geometry
A2	A1 plus PnP physical-plausibility validation
A3	A2 plus top- K candidate validation and early acceptance

4.4 YOLO11n Reference Baseline Protocol

A compact learned-detector reference baseline was added using YOLO11n [9]. The baseline used the same GateDataset v5.1 train, validation, and final-test image lists as DRGD, with a single *gate* class and empty label files for no-gate frames.

YOLO11n was initialized from generic pretrained weights and fine-tuned for 100 epochs with image size 640, batch size 16, and seed 0.

The validation split was used only to select the confidence threshold from 0.10, 0.15, 0.20, 0.25, 0.30, 0.40, 0.50; the selected value, 0.10, was then fixed for the one-shot final-test evaluation. For DRGD-compatible reporting, YOLO detections were reduced to at most one prediction per frame by retaining the highest-confidence gate box.

A true positive required the same IoU criterion as DRGD, $\text{IoU} \geq 0.50$; gate-present frames below this threshold were counted as non-localized,

Table 5. Consolidated DRGD results for the indoor and outdoor parameter sets. Emphasis marks the validation F1 values used to justify freezing and the final test F1/FPS values used for reporting. Header abbreviations: I/O = indoor/outdoor; Cal./Val./Test = calibration/validation/final test

	I-Cal.	I-Val.	I-Test	O-Cal.	O-Val.	O-Test
Frames	4416	552	552	3176	397	397
GT	4416	552	552	3062	397	173
Accepted	3512	517	489	2424	342	197
TP	3090	502	456	1831	319	74
FP	0	0	0	17	0	65
FN	1326	50	96	1231	78	99
F1	0.823	0.953	0.905	0.746	0.891	0.474
FPS	43.93	45.91	48.46	46.89	42.08	47.63

Table 6. A0–A3 validation ablation results. Emphasis marks the main F1 values used to interpret stage contribution and the early-accept top- K variant

Indoor validation				
Metric	A0	A1	A2	A3
Accepted	492	492	517	517
TP	477	477	502	502
FN	75	75	50	50
Recall	0.864	0.864	0.909	0.909
F1	0.927	0.927	0.953	0.953
FPS	50.76	50.65	48.76	45.58

Outdoor validation				
Metric	A0	A1	A2	A3
Accepted	224	255	342	314
TP	215	232	319	290
FN	182	165	78	107
Recall	0.542	0.584	0.804	0.730
F1	0.703	0.738	0.891	0.844
FPS	51.66	55.22	54.07	49.65

and unmatched detections were counted as false positives.

This preserves the same frame-level TP/FP/FN/TN structure used in the comparison table. YOLO11n was executed on the X470 workstation with an AMD Ryzen 7 1700 CPU and NVIDIA GeForce GTX 1660 Ti GPU. Thus, the reported YOLO FPS values correspond to end-to-end GPU-assisted prediction throughput, not to a hardware-matched CPU-speed comparison against DRGD; the main comparable quantity is final-test detection behavior under the same split and scoring rule.

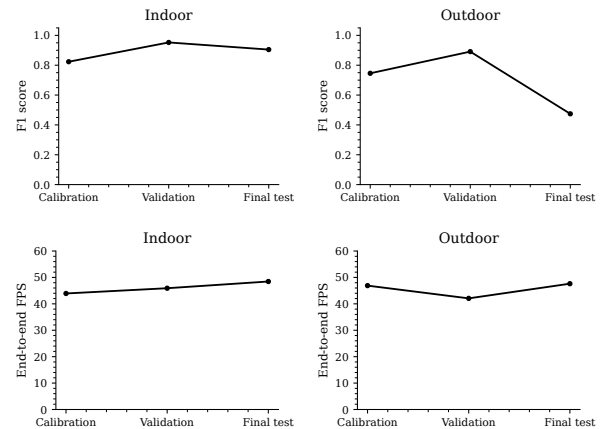


Fig. 3. Stage-level behavior for the indoor and outdoor parameter sets: F1 score (top) and end-to-end FPS (bottom)

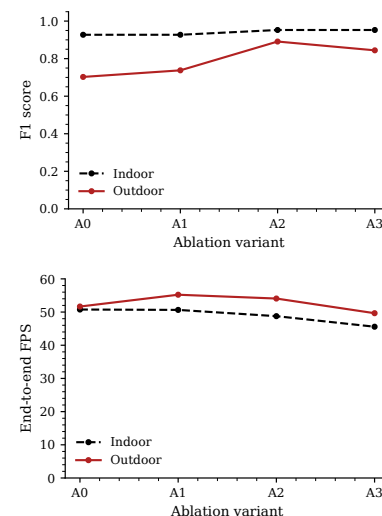


Fig. 4. Validation ablation behavior: F1 score (top) and end-to-end FPS (bottom) for the A0–A3 variants

5 Results

5.1 Calibration, Validation, and Freeze Procedure

Table 5 summarizes the frozen indoor and outdoor parameter sets across the official stages. Calibration established the static configuration files.

Table 7. Sequence-level final test results. Emphasis marks the sequence-level F1 values that drive the indoor/outdoor interpretation

	SQ3	SQ4	SQ5	SQ7
Domain	Indoor	Indoor	Outdoor	Outdoor
Frames	331	221	205	192
GT	331	221	121	52
Accepted	328	161	103	94
TP	325	131	72	2
FP	0	0	7	58
FN	6	90	49	50
Precision	1.000	1.000	0.911	0.033
Recall	0.982	0.593	0.595	0.038
F1	0.991	0.744	0.720	0.036

Table 8. True-positive (TP) IoU summary on the reserved final test split

Domain	Seq.	TP frames	Mean IoU	Median IoU
Indoor	SQ3	325	0.947	0.951
Indoor	SQ4	131	0.925	0.937
Outdoor	SQ5	72	0.798	0.837
Outdoor	SQ7	2	0.748	0.748

On validation, revision 1d9 reached 0.909 recall, 0.953 F1, and 45.91 FPS indoors, supporting the indoor freeze. Revision 1d11 reached 0.804 recall, 0.891 F1, and 42.08 FPS outdoors, supporting the outdoor freeze before the final test. The table also includes final test results for compact comparison.

Figs. 3 summarize the main experimental contrast. Indoor F1 decreases moderately from 0.953 on validation to 0.905 on the final test while remaining above 48 FPS. Outdoor F1 decreases from 0.891 on validation to 0.474 on the final test, while final-test throughput remains 47.63 FPS, indicating that the outdoor limitation is detection selectivity and localization robustness rather than runtime saturation.

5.2 Ablation Analysis

Table 6 reports the A0–A3 validation ablations. Indoors, A0 and A1 are already strong, while PnP validation raises recall from 0.864 to 0.909 and F1 from 0.927 to 0.953.

The gain comes mainly from SQ4, whose localized frames increase from 161 to 186, while SQ3 remains unchanged at 316 localized frames across A0–A3.

Table 9. Quantitative reference comparison on the reserved final test split. DRGD is reported with the CPU-only platform used throughout the paper, whereas YOLO11n is a learned-detector reference evaluated on a GTX 1660 Ti GPU

Indoor final test				
Metric	SQ3 DRGD	SQ3 YOLO	SQ4 DRGD	SQ4 YOLO
Platform	CPU	GPU	CPU	GPU
Precision	1.000	0.997	1.000	1.000
Recall	0.982	0.997	0.593	1.000
F1	0.991	0.997	0.744	1.000
FPS	50.62	27.35	45.54	22.70

Outdoor final test				
Metric	SQ5 DRGD	SQ5 YOLO	SQ7 DRGD	SQ7 YOLO
Platform	CPU	GPU	CPU	GPU
Precision	0.911	1.000	0.033	0.452
Recall	0.595	0.983	0.038	1.000
F1	0.720	0.992	0.036	0.623
FPS	46.72	21.15	48.63	20.88

Outdoors, A2 increases F1 from 0.703 in A0 to 0.891, showing that PnP plausibility and candidate ranking are important for the harder validation frames. A3 represents the early-accept top- K operating variant; it slightly lowers validation recall but remains real-time at 49.65 FPS.

Fig. 4 confirms that the added validation logic improves F1 without sacrificing real-time execution. The A2–A3 difference is an operating-mode trade-off: A2 isolates the highest validation accuracy, while A3 measures the top- K early-accept variant.

5.3 Final Held-Out Test

The sequence-level final test in Table 7 clarifies how the global indoor/outdoor results are distributed across sequences. Indoors, SQ3 reaches very high F1, while SQ4 remains the harder indoor case because recall decreases despite perfect precision. Outdoors, SQ5 remains usable with high precision and moderate recall. In contrast, SQ7 shows both very low recall and very low precision, indicating that the outdoor limitation is sequence-dependent and concentrated in that test sequence.

Table 10. SQ7 final-test frame-level localization and selectivity summary for DRGD

	Gate present	Gate absent
Total frames	52(100%)	140(100%)
Correct localization (TP)	2(3.8%)	–
Accepted, IoU < 0.50	34(65.4%)	–
No accepted output	16(30.8%)	–
False positive (FP)	–	58(41.4%)
No output (TN)	–	82(58.6%)

Table 8 reports a lightweight TP-IoU localization-quality summary for true-positive frames only. The indoor true positives have high median IoU, while SQ5 remains lower but stable. SQ7 has only two true positives, so its IoU values are not representative of the full sequence; they only show that the few accepted correct detections are well localized once the correct candidate is selected.

The lower outdoor final-test performance is concentrated in SQ7: SQ5 reaches 0.720 F1, whereas SQ7 reaches 0.036 F1 because only 2 of 52 gate-present frames are localized and 58 false positives occur on gate-absent frames.

5.4 YOLO11n Reference Comparison

Table 9 compares DRGD and YOLO11n under the same final-test split and frame-level scoring protocol. The YOLO11n baseline provides external quantitative context for the DRGD results, reaching 0.998 aggregate F1 indoors and 0.840 aggregate F1 outdoors. On SQ7, YOLO11n localizes all 52 gate-present frames but produces 63 false positives, corresponding to 32.8% of the 192 SQ7 final-test frames. This supports the interpretation of SQ7 as a combined localization–selectivity challenge under outdoor transition conditions rather than only a localization failure.

The FPS column should be interpreted cautiously because it reports the measured end-to-end throughput of each execution profile: CPU-only execution for DRGD and GPU-assisted prediction for YOLO11n. DRGD reaches higher throughput in Table 9, which is plausible for a lightweight task-specific pipeline, whereas YOLO11n includes neural-detector preprocessing, tensor conversion, GPU synchronization,

inference, and postprocessing. Therefore, the table provides a quantitative reference for detection quality and reported execution behavior, not a hardware-normalized CPU/GPU speed ranking.

6 Discussion

6.1 Main DRGD Results

The indoor results support DRGD as a two-stage real-time classical detector under controlled conditions. The frozen 1d9 indoor configuration reaches high validation and final test F1 while maintaining nearly 50 FPS end-to-end.

The ablation also shows that color and geometry provide a strong baseline, while PnP plausibility increases correctly localized frames in the harder indoor sequence, SQ4, without reducing performance on SQ3. This indicates that PnP is useful when the color–geometry stage already proposes plausible gate candidates: it improves the physical consistency of selected SQ4 candidates rather than compensating for absent or severely ambiguous proposals.

The outdoor results are more nuanced. The frozen 1d11 outdoor configuration is feasible and fast on validation, with 0.804 recall, 0.891 F1, and 42.08 FPS. However, the final test shows markedly lower SQ7 performance even though SQ5 remains usable. Because no tuning was performed after the final test, this result is retained as a valid limitation. SQ7 also includes gate-absent transition frames between the approach to the first and second gates, so its final-test result reflects both localization difficulty and false-positive control under transition conditions.

The SQ7 evidence suggests that the remaining outdoor instability is dominated by proposal and candidate-selection conditions—background confusion, gate scale, and transition frames—rather than by PnP alone; once an incorrect or weak candidate set reaches Stage 2, PnP validation cannot fully recover the intended gate hypothesis.

Table 11. Shared camera and gate model used by the frozen DRGD configurations

Camera model				
Parameter	f_x	f_y	c_x	
Value	289.6602	389.2576	317.9839	
Parameter	c_y	d_0	d_1	
Value	265.7634	-0.21973489	0.03835729	
Parameter	d_2	d_3	d_4	
Value	-0.00018976	-0.00010129	0	

Gate model and preprocessing				
Parameter	Width	Height	CLAHE clip	CLAHE tile
Value	2.10 m	2.10 m	2.0	8
Parameter	CLAHE on V		Pyramid levels	
Value	1		0	

Table 12. Frozen color-segmentation and mask-cleanup parameters. HSV values use the OpenCV scale $H \in [0, 179]$ and $S, V \in [0, 255]$

HSV range 1				
Config.	H min	H max	S min–max	V min–max
Indoor 1d9	88	142	40–255	25–255
Outdoor 1d11	92	136	55–255	35–255

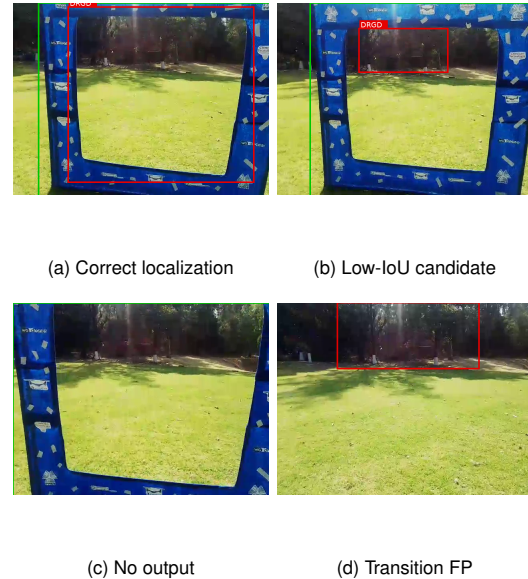
HSV range 2				
Config.	H min	H max	S min–max	V min–max
Indoor 1d9	140	179	35–255	25–255
Outdoor 1d11	150	179	45–255	30–255

Mask cleanup			
Config.	Opening	Closing	Fill holes
Indoor 1d9	3, 1	11, 2	1
Outdoor 1d11	3, 1	9, 1	1

6.2 SQ7 Localization–Selectivity Case

Table 10 and Fig. 5 summarize the SQ7 frame-level behavior. In the 52 gate-present frames, DRGD localizes 2 frames, accepts 34 low-IoU candidates, and produces no output in 16 frames, indicating that most gate-present errors are associated with candidate localization or scale selection rather than absence of a detected mask.

In the 140 gate-absent transition frames, 58 detections are emitted, showing that SQ7 also stresses selectivity between consecutive gate approaches.

**Fig. 5.** Representative SQ7 final-test cases for DRGD: correctly localized gate, low-IoU selected candidate, gate-present frame with no accepted output, and false positive during a gate-absent transition frame. Green boxes denote ground truth and red boxes denote the DRGD output

7 Conclusion

This paper presented DRGD as a deterministic color–geometry–PnP gate detector evaluated through a staged CPU workflow covering calibration, validation, ablation, configuration freezing, and a one-shot final test.

The proposal is a system-level classical detection pipeline rather than a new low-level image-processing operator: its main contribution is the reproducible integration of color segmentation, geometric proposal generation, PnP plausibility validation, and top- K candidate selection under fixed indoor and outdoor configurations.

The frozen 1d9 indoor configuration achieved strong final-test performance, with 0.905 F1 and 48.46 FPS end-to-end, showing that DRGD can provide accurate and real-time gate localization under controlled indoor conditions. The frozen 1d11 outdoor configuration remained real-time but reached 0.474 F1 on the final test, mainly due to the lower SQ7 result. This outcome does not negate the indoor behavior or the

Table 13. Frozen ROI, border, geometry, and corner-refinement parameters

ROI area and shape		
Config.	Parameter	Value
Indoor 1d9	Area min–max	0.0024–0.95
	Solidity min	0.46
	Extent min	0.21
	Aspect min–max	0.20–1.90
Outdoor 1d11	Area min–max	0.0012–0.95
	Solidity min	0.46
	Extent min	0.16
	Aspect min–max	0.16–2.20

ROI size, border rule, and geometry		
Config.	Parameter	Value
Indoor 1d9	BBox max area	0.58
	Border margin, penalty	6, 0.25
	Rectangularity min	0.28
	Solidity min	0.43
	Aspect min–max	0.23–1.35
Outdoor 1d11	BBox max area	0.70
	Border margin, penalty	10, 0.40
	Rectangularity min	0.26
	Solidity min	0.42
	Aspect min–max	0.20–1.45

Corner refinement		
Config.	Parameter	Value
Indoor 1d9	Enabled, window	1, 7
	Iterations, epsilon	30, 0.01
Outdoor 1d11	Enabled, window	1, 7
	Iterations, epsilon	30, 0.01

usable SQ5 outdoor result; instead, it identifies the robustness boundary of the current frozen color–geometry–PnP configuration. In particular, SQ7 provides useful negative evidence because outdoor transition frames affected both localization quality and false-positive rejection.

A YOLO11n reference baseline trained on the same split achieved higher final-test F1, especially outdoors, but under a different learned and GPU-evaluated execution profile. Therefore, DRGD's higher measured FPS in this comparison should be interpreted as the behavior of a lightweight task-specific CPU pipeline rather than as a hardware-normalized CPU/GPU speed ranking.

Overall, the results support DRGD as a fast, interpretable, and reproducible classical reference detector that performs strongly indoors, remains informative outdoors, and exposes the sequence-dependent limitations that future versions should address.

Future work will focus on stronger candidate ranking, negative-frame selectivity, and robustness to partial, off-scale, or transitional gate views, while preserving the transparency of the

Table 14. Frozen PnP validation, candidate-selection, output, and top- K parameters

PnP validation and candidate limit			
Group	Parameter	Indoor 1d9	Outdoor 1d11
PnP	Enabled, method	1, ITER.	1, ITER.
PnP	Iter., min inliers	200, 4	200, 4
PnP	Reproj. max px	9.0	9.0
PnP score	Inlier weight	0.15	0.15
PnP score	Reproj. weight	0.20	0.15
PnP score	Coverage weight	0.65	0.70
Candidate limit	ROI top- K	12	12

Runtime and selection score			
Group	Parameter	Indoor 1d9	Outdoor 1d11
Runtime policy	Stop first accept	0	0
A3 policy	Stop first accept	1	1
Selection score	hscore weight	0.62	0.45
Selection score	PnP weight	0.08	0.05
Selection score	bbox-area weight	0.25	0.42
Selection score	center weight	0.05	0.08

Selection references and output			
Group	Parameter	Indoor 1d9	Outdoor 1d11
Selection refs	bbox-area ref.	0.10	0.065
Selection refs	small-bbox floor	0.012	0.0045
Selection refs	center sigma	0.55	0.60
Selection refs	border penalty	0.08	0.12
Output	Emit score min	0.36	0.38
Output expansion	Enabled	–	1
Output expansion	x, y , min px	–	0.03, 0.03, 1

frozen-configuration protocol and progressing toward implementation on a resource-constrained SoC–FPGA platform.

Appendix A Frozen DRGD Parameters

This appendix reports the frozen numerical parameters used by the final DRGD experiments. The indoor values correspond to the frozen 1d9 configuration used for SQ3–SQ4, and the outdoor values correspond to the frozen 1d11 configuration used for SQ5–SQ7. The values were taken from the corresponding DRGD_i.cfg and DRGD_o.cfg files and were not changed after the final test. The tables focus only on parameters associated with the DRGD stages described in this paper: camera/gate calibration, color segmentation, mask cleanup, ROI and geometry filtering, PnP validation, candidate scoring, top- K selection, and final output generation.

References

1. Dalal, N., Triggs, B. (2005). Histograms of oriented gradients for human detection.

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 886–893. DOI: 10.1109/CVPR.2005.177.

2. **Felzenszwalb, P. F., Girshick, R. B., McAllester, D., Ramanan, D. (2010).** Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 9, pp. 1627–1645. DOI: 10.1109/TPAMI.2009.167.
3. **Foehn, P., Brescianini, D., Kaufmann, E., Cieslewski, T., Gehrig, M., Muglikar, M., Scaramuzza, D. (2022).** Alphapilot: Autonomous drone racing. *Autonomous Robots*, Vol. 46, pp. 307–320. DOI: 10.1007/s10514-021-10011-y.
4. **Hanover, D., Loquercio, A., Bauersfeld, L., Romero, A., Penicka, R., Song, Y., et al. (2024).** Autonomous drone racing: A survey. *IEEE Transactions on Robotics*, Vol. 40, pp. 3044–3067. DOI: 10.1109/TR0.2024.3400838.
5. **Hinterstoisser, S., Cagniard, C., Ilic, S., Sturm, P., Navab, N., Fua, P., Lepetit, V. (2012).** Gradient response maps for real-time detection of textureless objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 5, pp. 876–888. DOI: 10.1109/TPAMI.2011.206.
6. **Pham, X. H., Bozcan, I., Sarabakha, A., Haddadin, S., Kayacan, E. (2021).** Gatenet: An efficient deep neural network architecture for gate perception using fish-eye camera in autonomous drone racing. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. DOI: 10.1109/IROS51168.2021.9636207.
7. **Qiao, Z., Pham, X. H., Ramasamy, S., Jiang, X., Kayacan, E., Sarabakha, A. (2024).** Continual learning for robust gate detection under dynamic lighting in autonomous drone racing. *Proceedings of the International Joint Conference on Neural Networks*. DOI: 10.1109/IJCNN60899.2024.10649903.

8. **Scarciglia, L., Paolillo, A., Palossi, D. (2025).** A map-free deep learning-based framework for gate-to-gate monocular visual navigation aboard miniaturized aerial vehicles.

9. **Ultralytics (2024).** YOLO11 Models. Available at: <https://docs.ultralytics.com/models/yolo11/>. Accessed 2026-06-11.

Article received on 25/03/2026; accepted on 18/07/2026.

** Corresponding author is Jose Martinez-Carranza.*