Machine Learning for Marine Dock Detection Using LiDAR Intensity and Detectron2

Harald Steiner[1], Louise Ethier[1], Olena Pylypenko[1]

[1]GeoBC, Ministry of Water, Land and Resource Stewardship, Government of British Columbia, Canada – Harald.Steiner@gov.bc.ca

Keywords: Machine learning; LiDAR intensity; Marine docks; Radiometric normalization; Detectron2; Ablation study

Abstract

The Province of British Columbia is undertaking a multi-year LiDAR mapping program to deliver high-quality elevation data and support open-access geospatial products. Although dock detection was not an initial priority, this study demonstrates the value of leveraging LiDAR acquisition programs to support scalable, automated coastal infrastructure mapping. We evaluate the impact of LiDAR intensity normalization on CNN-based marine dock detection using Detectron2. Through a controlled ablation, we compare three strategies—raw intensity, scan-angle correction, and range-based correction—to isolate radiometric effects on detection performance. Detection metrics (precision, recall, F-score, IoU) and shape-fidelity measures are reported separately to clarify trade-offs. Results show scan-angle correction delivers balanced precision and recall, while range-based correction improves precision at the cost of recall. Beyond technical findings, this study offers comparative evidence to guide preprocessing choices and provides evidence-based guidance for AI-derived coastal mapping products in British Columbia.

1. Introduction

British Columbia’s multi-year LiDAR mapping program (Ministry of Water, Land and Resource Stewardship, 2023) is accelerating provincial-scale remapping and expanding access to high-quality elevation data. As new acquisitions come online, coastal infrastructure layers—including marine docks—are updated rapidly under variable flight conditions. Current workflows lack automated dock extraction, and manual mapping is not well-suited to provincial-scale throughput. This paper addresses that gap by evaluating CNN-based dock detection from LiDAR intensity imagery using Detectron2, with novelty centered on an ablation study of three normalization strategies: raw intensity, scan-angle correction, and range-based correction. We further improve evaluation clarity by separating detection metrics (precision, recall, F1, IoU) from geometric fidelity metrics (Horizontal Root Mean Square Error ($\mathrm{RMSE_H}$), centroid displacement, Hausdorff distance, $F1_{Bbnd}(\tau)$, aligning with ASPRS Edition 2 (2024) guidance (ASPRS Positional Accuracy Standards Working Group, 2024; USGS NGP Standards and Specifications, 2024).

This paper contributes in three complementary ways. First, it provides a systematic comparison of LiDAR radiometric normalization strategies (Kashani et al., 2015) for CNN-based marine dock detection under realistic acquisition geometry. Second, it introduces a fully controlled Detectron2 workflow in which all processing steps are held constant so that normalization remains the sole experimental variable (Wu et al., 2019). Third, it offers practical guidance for large-scale coastal mapping by benchmarking detection, spatial accuracy, and boundary fidelity, thereby supporting consistent and defensible AI-derived marine infrastructure products and informing emerging specifications for operational AI mapping standards.

Together, these results offer comparative evidence for preprocessing decisions and contribute to future provincial guidance on AI-derived coastal mapping products.

2. Data and Study Area

2.1. Study Area

Figure 1 illustrates the study area located on the Sunshine Coast of British Columbia, approximately 40 km northwest of Vancouver. It is comprised of five tiles at a 1:2500 scale in the Geographic BC map tile system, each covering approximately 2.5 km^2 of coastal terrain. The area was selected for its diverse dock geometries—varying in size, orientation, and structural complexity—and its coverage by multiple LiDAR flight lines exhibiting the full range of scan angles. These characteristics provide a representative distribution of factors influencing point density, return structure, and intensity variability.

2.2. LiDAR Data Acquisition

Aerial LiDAR data were collected using a calibrated Riegl VQ-1560ii sensor flown at approximately 1400 m above ground level with a maximum scan angle of ±30° and 30% flight line overlap. The dataset achieved an average ground return density of 8pts/m^2, with up to six discrete returns and 16-bit raw intensity values. Acquisition occurred under clear weather conditions, minimizing atmospheric and water-surface interference. The coordinate reference system used was *NAD83(CSRS)V8.0*.

2.3. Validation Data

Spatial accuracy metrics quantify the horizontal positional agreement between predicted and reference dock geometries. Following the ASPRS Positional Accuracy Standards (Edition 2), horizontal positional accuracy is reported using RMSE_H computed from independently surveyed Real-Time Kinematic Global Positioning System (RTK–GPS) checkpoints.

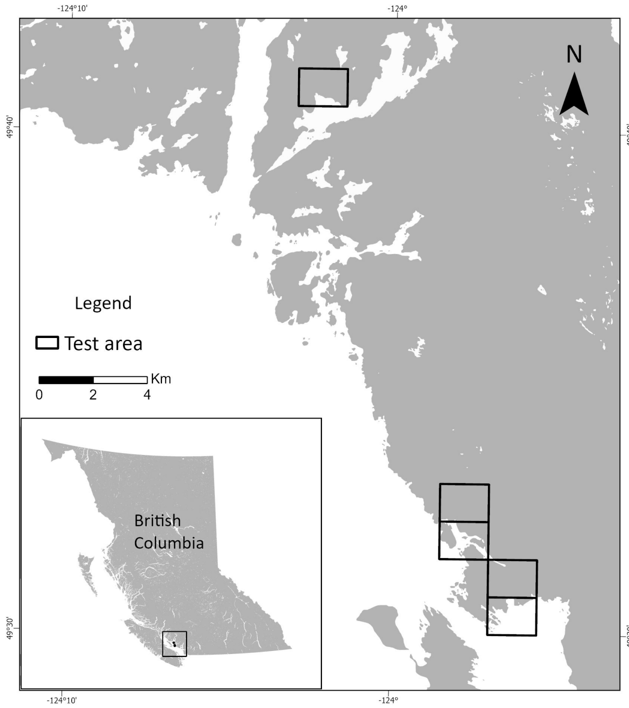

Figure 1. Map of the study area on the Sunshine Coast of British Columbia, Canada, showing BC Map Tile boundaries.

For each checkpoint, the horizontal residual was computed as the shortest Euclidean distance to any point along the corresponding dock polygon boundary and RMSE_H was then calculated from these residuals.

In addition to RMSE_H, centroid displacement was computed as a planimetric measure of positional offset between the centroids of predicted and reference dock polygons derived from the orthomosaics. Together, these measures characterize absolute checkpoint-based accuracy and relative footprint-level positional consistency.

2.4. Dataset Composition

Within the study area, a total of 543 marine docks were inventoried. Of these, 156 docks were reserved for detection evaluation; 59 of the 156 form the common true positive (TP) subset used for cross-domain shape-fidelity comparisons and 387 docks were manually delineated for training, ensuring a diverse mix of public and private structures spanning floating platforms, narrow gangways, multi-slip configurations, and irregular geometries. To prevent spatial leakage, training and test regions are geographically distinct and held constant across all experiments.

All dock labels were manually digitized as polygons in the provincial reference frame (NAD83(CSRS) V8.0; UTM Zone 10N). Details of rasterization, normalization domains, label rasterization, and prediction post-processing are described in Section 3.

In summary, the dataset combines a sufficiently large and diverse collection of dock structures with high-quality LiDAR and high-resolution imagery–derived reference data, ensuring that the ablation study evaluates normalization effects under realistic and representative coastal mapping conditions.

3. Methods

This study is designed as a controlled ablation experiment to isolate the impact of LiDAR intensity normalization on CNN-based marine dock detection. All components of the pipeline—tiling, rasterization, Detectron2 configuration, and evaluation—are held constant across experiments, with only the radiometric normalization module varying among raw, scan-angle-corrected, and range-corrected intensity imagery. The sections below describe the data preparation, normalization strategies, model training, inference workflow, and evaluation metrics.

3.1. Data Preparation

3.1.1. Radiometric normalization. We evaluate three radiometric domains: raw intensity, scan angle corrected intensity, and range-corrected intensity (Kashani, 2015).

(1) Raw intensity

$$I_{\mathrm{raw}}.$$

(2) Scan-angle correction

$$I_\theta = \frac{I_{\mathrm{raw}}}{\cos\theta + \epsilon}. \tag{1}$$

(3) Range correction

$$I_{corr} = I_{\mathrm{raw}} \left(\frac{R_i}{R_{ref}} \right)^2. \tag{2}$$

Here, R_i is the sensor-to-target range, defined below in Section 3.1.2 and R_{ref} is a user-defined reference range. After correction, intensities are normalized to 8-bit before export as GeoTIFFs.

3.1.2. Trajectory-free range estimation via forward intersection. When an external GNSS/IMU trajectory is unavailable, we estimate range entirely within a *local scan-plane coordinate system* without reconstructing the platform pose (Hartzell, 2020). For each scan line, we define a 2D local frame (X, Z) aligned with the scan plane: X is the across-scan axis, Z is the local vertical in the scan plane (scanner nadir aligned with the LiDAR Z-axis). All quantities below are expressed in this local frame, so no global trajectory is needed.

Forward-intersection in the local frame. Given two points from the same scan line with encoder angles and sufficient angular separation, it is possible to estimate their respective range values from the scanner. This approach assumes near-level flight and locally straight scan lines. The calculation is as follows: let $P_1(x_1, z_1)$ and $P_2(x_2, z_2)$ satisfy the previous requirements with encoder angles θ_1 and θ_2, respectively. The (unknown) scanner position $S(x_s, z_s)$ is obtained by intersecting the two rays in the *local* scan-plane frame:

$$Z_1 = z_1 + (x_s - x_1)\tan\theta_1 \quad \text{and} \tag{3}$$

$$Z_2 = z_2 + (x_s - x_2)\tan\theta_2, \tag{4}$$

Setting equations Z_1 and Z_2 equal yields

$$x_s = \frac{z_2 - z_1 + x_1\tan\theta_1 - x_2\tan\theta_2}{\tan\theta_1 - \tan\theta_2}, \tag{5}$$

$$z_s = z_1 + (x_s - x_1)\tan\theta_1. \tag{6}$$

Therefore the *range* from scanner to $P_1(x_1, z_1)$ in the local frame is

$$R_1 = \sqrt{(x_s - x_1)^2 + (z_s - z_1)^2}. \tag{7}$$

This process is performed on each point in the point cloud, obtaining a unique R_i value for every LiDAR point i in the given point cloud. Equations (5)–(7) operate solely in the local scan-plane coordinates (X, Z); they do not reconstruct or require a global trajectory.

3.1.3. Gridding and radiometric preparation. Point–level LiDAR intensities were gridded into 2D raster images using the Inverse Distance to a Power interpolation method implemented (GDAL Development Team, 2026). The gridding resolution was derived from the average LiDAR point density of approximately 8 ptsm^2 and set to a uniform output cell size of 15cm to ensure adequate sampling and spatial coherence of the intensity surface. For each raster cell, the interpolated intensity value $I(x, y)$ is computed as a weighted average of nearby returns :

$$I(x, y) = \frac{\sum_{i=1}^{n} \frac{I_i}{r_i^p}}{\sum_{i=1}^{n} \frac{1}{r_i^p}}, \tag{8}$$

where I_i denotes the intensity of the i^{th} return, n is the number of points falling within the search ellipse, and p is the inverse–distance power parameter. The point–to–grid distance is:

$$r_i = \sqrt{r_{ix}^2 + r_{iy}^2 + s^2}, \tag{9}$$

with r_{ix} and r_{iy} the offsets between the grid node and point i, and s an optional smoothing parameter. The resulting weight factor is $w_i = 1/r_i^p$.

Raw sensor intensities I_{raw} (16–bit) were subsequently scaled to 8–bit after normalization:

$$I_{\text{8-bit}} = \text{round}\left(255 \times \frac{I_{\text{raw}}}{4095}\right). \tag{10}$$

Following radiometric normalization (raw, scan–angle corrected, or range–corrected), interpolation was applied, and the final intensity rasters were exported as single–band GeoTIFFs for downstream CNN ingestion. These GeoTIFFs serve as the direct input to the Mask R–CNN model during both training and inference.

3.2. Workflow Overview

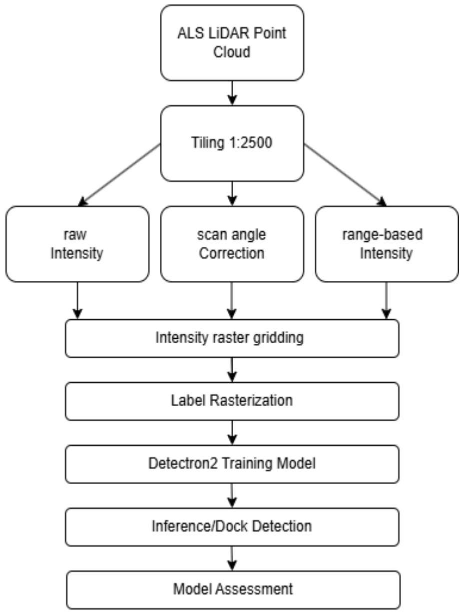

Figure 2. Workflow and ablation scope: only the intensity normalization module changes (raw, scan angle, range).

As Figure 2 highlights, the marine dock detection workflow follows a fixed processing pipeline in which only the radiometric normalization module varies across experiments. The pipeline begins with the tiling of the LiDAR point cloud. Then intensity is normalized by scan angle, range, or left uncorrected (raw), where these three options define the independent variable of the ablation study. Once an intensity normalization is chosen, gridding of intensity values is applied to the tiles to transform them into 2D rasters at 15 cm resolution.

Following labeling and normalization, dock reference polygons are rasterized into training masks using identical rules for buffer size, minimum object area, and label resolution. Each normalized intensity tile is paired with its corresponding mask to form the input for Detectron2’s Mask R–CNN architecture. All downstream components, tiling strategy, and hyperparameter settings, are held constant across experiments to ensure that radiometric normalization remains the sole varying factor.

During model training, batches of normalized tiles and their associated masks are used to optimize feature extraction for object detection and instance segmentation. Inference applies the trained model to new intensity imagery using the same preprocessing and normalization pipeline to avoid domain shift. Predicted dock masks are post–processed through non–maximum suppression, probability thresholding, and polygon vectorization. The resulting dock geometries are evaluated against the reference dataset using the detection and geometric accuracy metrics described in Section 3.6.

3.2.1. Ablation scope. The pipeline modules for tiling, rasterization, Detectron2 configuration, and evaluation remain identical across experiments; the only module that changes is the *intensity normalization* (raw, scan angle, range). To prevent domain–mismatch, each experiment uses training and testing imagery produced with the *same* normalization.

3.3. Model Training

3.3.1. Data split. We use a fixed geographic split of 27.4 km^2 for training and 12.4 km^2 for testing. While general machine learning guidance often motivates 70/30 or 80/20 splits from first principles (Gholamy et al., 2018), spatial imagery requires geographically independent partitions to avoid optimistic accuracy due to spatial autocorrelation (Karasiak et al., 2021). Empirical evaluations further show that the *spatial separation and block size* governing the split matter more than the exact percentage (Stock, 2025); hence we select a representative training area capturing variability in dock morphology, shoreline context, and acquisition geometry, and hold out a disjoint test area. This also reflects provincial-scale production constraints, where preparing a strict 70/30 geographic split over ~1,000,000 km^2 would be prohibitively expensive; instead, spatially disjoint and representative sampling is a recommended mitigation for leakage in remote sensing deep learning (Feng et al., 2023).

3.3.2. Training data curation and dock annotation. Dock reference polygons were manually delineated on the 15 cm LiDAR-intensity rasters. A dock was defined as any fixed or floating berthing platform, inclusive of attached gangways; vessels, isolated pilings, and breakwaters were excluded. Polygons followed the visible waterline footprint. Vertex placement adhered to a planimetric tolerance of ≤0.3 m or two pixels at resolution.

3.3.3. Training configuration. All models are implemented in Detectron2 with a Mask R–CNN architecture. Models are trained for 700 iterations with a learning rate of 0.00025, a batch size of 16, and input images rescaled to 1024 px on the short edge. Training is conducted on a single GPU with 128 GB of system RAM. Rasterization settings remain identical across experiments.

3.3.4. Training data per normalization. For each normalization type (raw, scan angle corrected, range–corrected), the *training set is generated from greyscale imagery–like intensity rasters produced with that same correction*, and evaluation is performed on test imagery using the identical correction. This pairing avoids domain–mismatch effects introduced by radiometric changes and isolates normalization as the only independent variable.

3.3.5. Inference. Inference refers to applying the trained model to new intensity imagery to produce dock predictions, using no further weight updates.

3.3.6. Ablation Design and Control Conditions. To ensure a *controlled* ablation:

- **Fixed components across experiments:** train–test split, network architecture, optimizer and hyperparameters, rasterization settings, data augmentation, and evaluation protocol.
- **Single varying factor:** the intensity normalization strategy (raw, scan angle, range).
- **Consistency rule:** training and inference use imagery processed with the same normalization to maintain domain consistency.

3.4. Evaluation Framework

Model performance is evaluated using three complementary classes of metrics: (1) *Detection metrics*, which quantify the correctness of dock identification, (2) *Spatial accuracy*, and (3) *Shape–fidelity metrics*, which assess positional agreement and boundary alignment between predicted and reference dock footprints. Detection metrics include Precision, Recall, F_1–score, and Jaccard Index (equivalent to IoU). Spatial accuracy metrics comprise RMSE_H and centroid displacement, while symmetric Hausdorff distance and $F1_{bnd}$ assess Shape–fidelity.

3.5. Detection Metrics

Detection metrics quantify whether docks are correctly identified; shape fidelity is evaluated separately in Section 3.7. We compare predicted dock polygons against reference polygons using one-to-one assignment based on the Jaccard criterion. A prediction is a true positive (TP) if its IoU with a unique reference polygon is positive; unmatched predictions are false positives (FP), and unmatched references are false negatives (FN). One-to-one matching is enforced via a maximum-IoU assignment.
From the counts, we compute,

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{11}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{12}$$

$$\text{F1} = \frac{2TP}{2TP + FP + FN}. \tag{13}$$

We also report the Jaccard Index J,

$$J = \frac{|P_{\text{pred}} \cap P_{\text{ref}}|}{|P_{\text{pred}} \cup P_{\text{ref}}|}, \tag{14}$$

where $P_{\text{pred}}, P_{\text{ref}}$ are the areas of the predicted and reference docks. Unless noted, metrics are aggregated per-tile followed by unweighted averaging.

3.6. Spatial Accuracy

Spatial accuracy metrics quantify the horizontal positional agreement between predicted and reference dock geometries. Following the *ASPRS Positional Accuracy Standards*

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-273-2026 | © Author(s) 2026. CC BY 4.0 License.

for Digital Geospatial Data, Edition 2 (ASPRS, 2024), horizontal positional accuracy is reported using RMSE_H, which measures the root-mean-square difference between predicted and reference positions at independently surveyed control points.

In addition to RMSE_H, centroid displacement was computed as a planimetric measure of positional offset between the centroids of predicted dock polygons and the corresponding reference polygons derived from high-resolution orthomosaics. Together, these two metrics characterize the spatial accuracy of the predicted dock outlines.

3.7. Shape Fidelity Metrics

Shape fidelity is evaluated using two complementary measures: (i) the symmetric Hausdorff distance, (ii) a $F1_{bnd}$ score computed within a fixed tolerance τ.

Hausdorff distance. For predicted polygon boundaries ∂P and reference boundaries ∂G, the symmetric Hausdorff distance is

$$H(P,G) = \max\left\{ \max_{x\in\partial P} \min_{y\in\partial G} \|x-y\|,\ \max_{y\in\partial G} \min_{x\in\partial P} \|y-x\| \right\}. \tag{15}$$

***F*1*bnd*.** Boundary matches are defined as

$$M_{P\to G}(\tau) = \left\{ p \in \partial P \ :\ \min_{g\in\partial G} \|p-g\| \le \tau \right\}, \tag{16}$$

$$M_{G\to P}(\tau) = \left\{ g \in \partial G \ :\ \min_{p\in\partial P} \|g-p\| \le \tau \right\}, \tag{17}$$

with

$$\mathrm{Prec}_{\mathrm{bnd}} = \frac{|M_{P\to G}|}{|\partial P|}, \qquad \mathrm{Rec}_{\mathrm{bnd}} = \frac{|M_{G\to P}|}{|\partial G|}, \tag{18}$$

and

$$F1_{\mathrm{bnd}}(\tau) = \frac{2\,\mathrm{Prec}_{\mathrm{bnd}}\,\mathrm{Rec}_{\mathrm{bnd}}}{\mathrm{Prec}_{\mathrm{bnd}} + \mathrm{Rec}_{\mathrm{bnd}}}. \tag{19}$$

We use a fixed τ matched to the annotation tolerance (2-pixel grid width ≈ 0.30 m).

Polygon inclusion. All shape-fidelity metrics are computed only for prediction–reference pairs classified as True Positives under the one-to-one IoU$\geq$ 0.50 rule; unmatched polygons are excluded. For cross-domain comparability, Shape-Fidelity metrics are computed on the same set of 59 reference docks that are True Positives in all three normalization domains.

4. Results

This section presents the experimental results for the three radiometric normalization strategies: raw intensity, scan angle–corrected intensity, and range–corrected intensity. All experiments use identical training/test regions and identical model configurations; only the intensity normalization differs across ablation runs.

4.1. Detection results

Tables 1, 2 and 3 summarize Precision, Recall, F_1, and Jaccard index J for each tile and each normalization strategy. Scan–angle correction produces the most balanced detection performance, while raw intensity yields higher recall at the cost of increased false positives. Range–based correction improves precision but results in reduced recall. Because each normalization domain yields its own set of TP, FP, and FN assignments, the number of matched and unmatched polygons therefore differs between normalization domains.

Tile	Ref	Pred	TP	FP	FN	Prec	Rec	F1	J
1	14	30	12	18	2	0.40	0.86	0.55	0.31
2	25	45	19	26	6	0.42	0.76	0.54	0.23
3	5	8	5	3	0	0.63	1.00	0.77	0.52
4	103	149	96	53	7	0.64	0.93	0.76	0.50
5	9	12	5	7	4	0.42	0.56	0.48	0.28
Avg	31.2	48.8	27.40	21.40	3.80	0.50	0.82	0.62	0.37

Table 1. Raw intensity detection metrics.

Tile	Ref	Pred	TP	FP	FN	Prec	Rec	F1	J
1	14	14	11	3	3	0.79	0.79	0.79	0.45
2	25	20	12	8	13	0.60	0.48	0.53	0.35
3	5	5	5	0	0	1.00	1.00	1.00	0.54
4	103	143	88	55	15	0.62	0.85	0.72	0.48
5	9	10	3	7	6	0.30	0.33	0.32	0.19
Avg	31.20	38.40	23.80	14.60	7.40	0.66	0.69	0.67	0.40

Table 2. Performance metrics for scan angle corrected intensity normalization.

Tile	Ref	Pred	TP	FP	FN	Prec	Rec	F1	J
1	14	10	10	0	4	1.00	0.71	0.83	0.52
2	25	16	14	2	11	0.88	0.56	0.68	0.40
3	5	7	5	2	0	0.71	1.00	0.83	0.61
4	103	123	86	37	17	0.70	0.84	0.76	0.50
5	9	10	2	8	7	0.20	0.22	0.21	0.15
Avg	31.20	33.20	23.40	9.80	7.80	0.70	0.67	0.66	0.43

Table 3. Performance metrics for range corrected intensity normalization.

4.2. Spatial Accuracy results

Spatial accuracy results, summarized in Table 4, include RMSE_H (38 RTK–GPS control points) and centroid displacement (156 reference docks). RMSE_H reflects absolute horizontal positional accuracy of predicted dock edges; centroid displacement captures planimetric offset between predicted and reference dock footprints.

Normalization	$RMSE_H$ (m)	Centroid Disp. (m)
Raw Intensity	0.99	1.662
Scan–Angle Corrected	2.01	1.773
Range Corrected	1.66	1.659

Table 4. Spatial Accuracy metrics for the three normalization strategies.

4.3. Shape Fidelity results

Shape fidelity was evaluated for 59 reference docks using symmetric Hausdorff distance and the $F1_{bnd}$ score. To control for selection effects, Table 5 reports boundary metrics only on the common TP subset (n=59) detected in all domains. Figure 3 illustrates how the three different intensity normalization methods affect shape fidelity.

Figure 3. Example dock outlines for raw (green), scan-angle (blue), and range-corrected (orange) intensity domains.

Normalization	Hausdorff (m)	$F1_{\mathrm{bnd}}(\tau)$
Raw Intensity	4.384	0.626
Scan–Angle Corrected	4.589	0.662
Range Corrected	4.427	0.634

Table 5. Shape fidelity metrics computed on the common TP subset (n=59) shared across all normalization domains.

4.4. Comparison of Normalization Strategies

Scan–angle correction produced the most balanced behavior across tiles, with improved consistency in both precision and recall. Raw intensity exhibited high recall but substantially lower precision, producing large numbers of false positives. Range–based correction improved precision but tended toward under–prediction, reducing recall. Consistent with this precision-leaning behavior, the mean Jaccard index (J) is highest for range-corrected imagery (0.43) compared with scan-angle (0.40) and raw intensity (0.37).

5. Discussion

5.1. Normalization as the primary driver of detection trade-offs

With tiling, rasterization, model configuration, and evaluation held constant, the ablation isolates radiometric normalization as the key factor shaping performance. As shown in Section 4.4, the three normalization strategies move the detector along the precision–recall operating spectrum; notably, scan–angle normalization provides the most stable operating regime across heterogeneous scan geometries.

This suggests that normalization choice is the primary lever for tuning precision–recall behavior: practitioners can favor sensitivity or conservatism via normalization before thresholding or post-processing. Future work should examine (i) robustness under more extreme incidence angles and ranges, (ii) calibration effects (e.g., reliability/ECE), and (iii) whether combining geometry-aware normalization with learned score calibration further stabilizes performance across tiles.

5.2. Detection versus geometric quality are distinct objectives

Separating *detection metrics* (Precision, Recall, F_1, IoU) from *spatial accuracy* ($RMSE_H$, centroid displacement) and *shape fidelity* (Hausdorff, $F1_{bnd}$) clarifies that “finding” a dock and “drawing” it accurately are related but distinct tasks.

5.3. Spatial accuracy interpretation $RMSE_H$

Table 4 reports horizontal $RMSE_H$ across normalization domains as 0.99–2.01 m. While this 1.02m spread indicates some sensitivity of absolute planimetric accuracy to normalization, its magnitude remains modest relative to stronger normalization-driven trade-offs observed in the detection metrics. Centroid displacements are tightly clustered (1.659–1.773 m), suggesting footprint level offsets are dominated by instance-mask geometry (15 cm grid, raster-to-vector steps) rather than radiometry.

5.4. Boundary quality and shape fidelity

On the common true-positive subset (n=59), $F1_{bnd}$ is highest for Scan–Angle correction (0.662), with range (0.634) and raw (0.626) close behind; symmetric Hausdorff distances are comparable across domains (4.38–4.59 m). Figure 3 shows Scan–Angle correction delivers acceptable outlines, while range correction can tighten boundaries in some cases but misses marginal structures more often, consistent with its precision/recall trade-off.

5.5. Operational considerations for provincial scale mapping

For workflows that prioritize high recall (e.g., discovery, inventory), raw intensity is acceptable provided QA/QC can handle more false positives. Where balanced performance is required, scan–angle correction offers a defensible default. When analyst time is limited and map cleanliness is the priority, range-based correction reduces false positives but increases false negatives, particularly for omission-prone targets.

5.6. Limitations and threats to validity

We use a fixed geographic split selected to preserve spatial independence; conclusions are derived from a single coastal region and sensor configuration; polygon vectorization was not regularized with cartographic priors; and annotated training data volume may exceed some production regimes.

5.7. Recommendations

Based on the ablation, we recommend:

1. Use *Scan–angle correction* as the default normalization to balance precision and recall across heterogeneous flight conditions.
2. When maximizing *recall* for inventory/triage, consider *raw intensity* and prune commission errors via lightweight post-filters or focused analyst review.
3. When minimizing *false positives* and QC effort, consider *range correction* and mitigate omissions with targeted checks in omission-prone settings.
4. Report detection, spatial accuracy, and boundary metrics separately in production to make trade-offs explicit to stakeholders.

5.8. Outlook

The trajectory-free range estimation (Sec. 3.1.2) extends applicability to legacy or low-metadata datasets by enabling radiometric correction without external trajectory information. Future work should evaluate reduced-label and cross-region generalization scenarios and investigate kernel point convolution-based deep learning methods (Thomas et al., 2019) capable of leveraging full 3-D structure.

6. Conclusion

This controlled ablation isolates radiometric normalization as the primary factor shaping CNN-based dock detection from LiDAR intensity. Across tiles, *scan–angle correction* provides the most stable balance between precision and recall, *raw intensity* increases recall at the cost of commission errors, and *range correction* reduces commission errors but can under-predict in challenging cases.

7. Additional Considerations and Future Directions

Governance for AI-derived mapping (data lineage, normalization provenance, QC thresholds, and change-management) will be essential as products scale beyond pilot regions. Label scarcity and annotation costs motivate workflows that combine targeted labeling, active sampling, and transfer learning to reduce dependence on large in-domain training sets.

Acknowledgements

This research was conducted within the **shíshálh** Nation’s *swiya* (territory) on the Sunshine Coast, British Columbia. We gratefully acknowledge the **shíshálh** People.

References

ASPRS Positional Accuracy Standards Working Group, 2024. ASPRS Positional Accuracy Standards for Digital Geospatial Data, Edition 2. Technical report, American Society for Photogrammetry and Remote Sensing, Baton Rouge, LA.

Feng, H., Wang, Y., Li, Z., Zhang, N., Zhang, Y., Gao, Y., 2023. Information Leakage in Deep Learning-Based Hyperspectral Image Classification: A Survey. *Remote Sensing*, 15(15), 3793.

GDAL Development Team, 2026. GDAL Grid Tutorial — GDAL Documentation. https://gdal.org/en/stable/tutorials/gdal_grid_tut.html.

Gholamy, A., Kreinovich, V., Kosheleva, O., 2018. Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. Technical Report UTEP-CS-18-09, University of Texas at El Paso, El Paso, TX, USA.

Hartzell, P. J., 2020. Airborne lidar sensor trajectory estimation from LAS scan angle field. https://github.com/pjhartzell/scan-angle.

Karasiak, N., Dejoux, J.-F., Monteil, C., Sheeren, D., 2021. Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. *Machine Learning*, 111(8), 2715–2740.

Kashani, A. G., Olsen, M. J., Parrish, C. E., Wilson, N., 2015. A Review of LiDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. *Sensors*, 15(11), 28099–28128.

Ministry of Water, Land and Resource Stewardship, 2023. Free, accessible data will help communities manage, adapt, plan. https://news.gov.bc.ca/releases/2023WLRS0010-000512.

Stock, A., 2025. Choosing blocks for spatial cross-validation: lessons from a marine remote sensing case study. *Frontiers in Remote Sensing*, 6, 1531097.

Thomas, H., Qi, C. R., Deschaud, J.-E., Marcotegui, B., Goulette, F., Guibas, L. J., 2019. KPConv: Flexible and Deformable Convolution for Point Clouds. arXiv preprint, arXiv:1904.08889.

USGS NGP Standards and Specifications, 2024. Adopt updated accuracy standards (ASPRS Edition 2, 2024 updates). https://www.usgs.gov/ngp-standards-and-specifications/adopt-updated-accuracy-standards.

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., Girshick, R., 2019. Detectron2. https://github.com/facebookresearch/detectron2.