Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.gaexcellence.com/jistm

GLOBAL ACADEMIC EXCELLENCE

COMPARATIVE ANALYSIS OF 2D AND 3D LIDAR-BASED OBJECT DETECTION FOR MARITIME NAVIGATION

Aiman Norazaruddin[1]*, Zulkifli Zainal Abidin[2]

[1]Department of Mechatronics Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Malaysia
aiman.norazaruddin@gmail.com https://orcid.org/0009-0002-3651-3396
2Department of Mechatronics Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Malaysia
zzulkifli@iium.edu.my https://orcid.org/0000-0003-3559-2440
*Corresponding Author

Article Info:

Article history:
Received date: 28.06.2026
Revised date: 30.07.2026
Accepted date: 17.08.2026
Published date: 17.09.2026

To cite this document:

Norazaruddin, A., & Abidin, Z. Z. (2026). Comparative Analysis Of 2D and 3D LiDAR-Based Object Detection for Maritime Navigation. *Journal of Information System and Technology Management, 11* (44), 316-339.

Abstract:

Maritime navigation safety relies on the timely and accurate detection of obstacles such as vessels, buoys, and floating debris. Conventional sensing technologies, including radar and optical cameras, are limited by low resolution, reduced performance in adverse weather, and susceptibility to sea clutter. Light Detection and Ranging (LiDAR), with its high-resolution three-dimensional sensing capability, offers a promising alternative; however, its application in maritime environments remains relatively underexplored compared to autonomous road vehicles. This study presents a LiDAR-based object detection framework for maritime navigation and evaluates the performance of two-dimensional (2D) and three-dimensional (3D) detection approaches using a calibration/projection-based LiDAR-camera fusion scheme, in which 2D detections are associated with LiDAR points to obtain 3D object localisation. A Velodyne mechanical LiDAR and a ZED stereo camera were mounted on a marine vessel operating in the waters of Klang, Malaysia. Training data were obtained from the Singapore Maritime Dataset, the SeaShips Dataset, and a custom-collected maritime dataset annotated using Roboflow for 2D images and labelCloud for 3D point clouds. A YOLOv8-based model was developed for 2D object detection, while a parallel pipeline generated point clouds from depth imagery for 3D bounding-box estimation. Performance was evaluated using precision, recall, F1-score, Intersection-over-Union (IoU), and, for the 2D pipeline, mean Average Precision (mAP50 and mAP50-95). The 2D detection model achieved strong performance, with a precision of 0.94, recall between 0.84 and 0.86, and an F1-score of approximately 0.89. In contrast, the 3D detection pipeline achieved precision, recall, and F1-score of approximately 0.25 each, a gap attributed jointly to the limited availability of annotated 3D maritime data and to the geometric, non-learned nature of the 3D bounding-box generation step, which takes the axis-aligned extent of the associated LiDAR/depth point cloud without a mechanism to reject noise or clutter points. Comparative analysis also showed that the Velodyne LiDAR produced denser and more accurate point clouds than the ZED stereo camera, particularly for distant objects where stereo depth estimation degraded. The proposed framework demonstrates the effectiveness of LiDAR-camera fusion for maritime 2D object detection while highlighting current challenges in 3D maritime perception. It provides an end-to-end baseline for future sensor-fusion research supporting safer autonomous and assisted maritime navigation.

DOI: 10.35631/JISTM.1144019

Keyword:

LiDAR; Marine Object Detection; Point Cloud; Sensor Fusion; Maritime Navigation; YOLOv8.

Introduction

Safe navigation at sea depends on a vessel's ability to perceive its surroundings with sufficient range, resolution, and reliability to avoid collision with other crafts, hazards, and floating debris. Marine environments present a uniquely difficult sensing problem: the operating area is vast and largely featureless, sea states change continuously, and targets such as small boats, buoys, and partially submerged objects must be distinguished from a noisy background of waves, spray, and glint. Radar struggles to separate small or low-profile targets from sea clutter, while optical cameras lose effectiveness under poor illumination, fog, or rain. Light Detection and Ranging (LiDAR) has consequently attracted growing interest because it actively emits laser pulses and measures their time of flight to construct dense, high-resolution three-dimensional representations of the scene, independent of ambient lightning.

The integration of LiDAR with deep learning has produced substantial advances in object detection for autonomous driving, where large, well-annotated datasets such as KITTI and nuScenes have enabled high-performing architectures (Chen et al., 2017; H. Zhang et al., 2023; Zhou & Tuzel, 2018), but the transfer of these techniques to the maritime domain remains comparatively immature. LiDAR has nonetheless demonstrated value in adjacent marine applications, including coastal morphology mapping (Nijland et al., 2017), real-time tracking of marine objects (Prasad et al., 2017), power-facility inspection (Liu et al., 2019), and wind-turbine inflow estimation (Gräfe et al., 2023; Sekar et al., 2021). This motivates the present study: radar-based systems struggle to discriminate genuine obstacles from wave and spray clutter, while vision-based systems depend on ambient light and provide depth only indirectly, through stereo disparity or monocular estimation, both of which degrade with range. LiDAR's

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

active illumination and direct range measurement address both shortcomings simultaneously, and recent reductions in the size, weight, and cost of automotive-grade LiDAR units (Aijazi et al., 2020; Zhong et al., 2022) have made shipboard deployment increasingly practical — yet detection pipelines built for road vehicles cannot simply be transplanted to the maritime domain, since marine scene geometry, scale, and clutter differ substantially from urban driving scenes.

A review of the literature reveals three specific gaps. First, although two-dimensional LiDAR- and radar-based detection has been studied extensively for autonomous driving (Danzer et al., 2019; Lang et al., 2019; Yan et al., 2018), few studies report empirically validated two-dimensional detection performance for maritime targets using real vessel-collected data rather than simulation or terrestrial benchmarks. Second, three-dimensional methods such as VoxelNet (Zhou & Tuzel, 2018), multi-view fusion networks (Chen et al., 2017), and pseudo-LiDAR approaches (Weng & Kitani, 2019; You, 2019) perform strongly on road-vehicle benchmarks, but their applicability to sparse, under-annotated marine point clouds is rarely quantified. Third, while several surveys catalogue algorithms and datasets for marine object detection (M. J. Er et al., 2023; R. Zhang et al., 2021), few report a direct, side-by-side comparison of two- and three-dimensional detection performance on identical hardware and an identical maritime dataset — a comparison needed to weigh the cost and benefit of pursuing full three-dimensional perception. This study addresses all three gaps through a single, integrated experimental pipeline.

Conventional instrumentation provides an incomplete picture of the operating environment precisely when collisions are most likely: radar cannot reliably separate low-lying obstacles from clutter, and optical sensors lose performance exactly when visibility is already compromised. LiDAR's high-resolution three-dimensional sensing is theoretically well suited to close this gap, but detection performance is sensitive to water turbidity, sea state, and target reflectivity, and any candidate pipeline must report detections with low enough latency for an operator or autonomous controller to manoeuvre in time. The central problem addressed here is therefore the design, implementation, and empirical evaluation of a LiDAR-based detection framework operating under these constraints, and the quantification of the performance gap between two- and three-dimensional detection on real maritime sensor data. Accordingly, this study pursues three objectives: investigating the influence of environmental factors such as turbidity, sea state, and ambient light on detection quality; comparing the developed algorithm against existing LiDAR-based detection methods reported in the literature; and evaluating the algorithm's precision, recall, and Intersection-over-Union separately for the two- and three-dimensional pipelines.

Prior maritime perception studies have generally addressed a single modality or a single detection dimensionality in isolation: LiDAR-only maritime detection, camera-only or YOLO-based maritime detection, generic LiDAR-camera fusion validated on road-vehicle data, or 2D-to-3D perception pipelines developed for autonomous surface vessels (ASVs) operating outside the maritime clutter and scale regime studied here. This study is distinguished from that body of work by combining, within a single real-world maritime deployment, elements that prior studies have only addressed separately. Specifically, this paper makes four contributions: (i) an end-to-end, calibration/projection-based LiDAR-camera fusion framework for maritime object detection, in which a YOLOv8 2D detector is coupled with Velodyne LiDAR point-cloud association for 3D localisation, spanning acquisition, annotation, training, and evaluation; (ii) one of the few direct empirical comparisons of 2D and 3D

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339
detection performance within the same experimental environment and on the same custom maritime dataset; (iii) a comparative characterisation of mechanical-LiDAR and stereo-camera point clouds under identical maritime conditions; and (iv) an identification and analysis of the dataset- and method-related factors limiting 3D maritime perception, namely the scarcity of annotated 3D maritime data and the geometric, non-learned nature of the 3D bounding-box generation step evaluated here, distinguishing this limitation from the sensing-modality itself. A structured synthesis of the dispersed literature on LiDAR object detection, sensors, and marine datasets is additionally provided to orient future researchers in this young subfield.

Literature Review

Significance of LiDAR for Marine Object Detection

LiDAR's principal value for marine object detection lies in its ability to generate dense, accurate three-dimensional information regardless of ambient lighting, a property that directly addresses several of the perceptual challenges unique to open-water environments. Prior work has shown that LiDAR-derived three-dimensional data can support marine reconnaissance, obstacle avoidance, and environmental monitoring tasks that are difficult to accomplish with radar or vision alone, and the fusion of LiDAR with deep-learning classifiers has produced measurable gains in the classification of marine habitats and objects (S. Er et al., 2023; H. Zhang et al., 2023).

LiDAR's utility extends beyond simple obstacle detection: airborne LiDAR has been used to characterise coastal morphology and classify shore zones, clarifying the interaction between marine and terrestrial environments in a way that is directly relevant to coastal management and conservation (Nijland et al., 2017). Real-time tracking applications have likewise demonstrated that LiDAR can strengthen surveillance and navigation systems for vessels operating in cluttered or congested waters (Prasad et al., 2017). Taken together, this body of work establishes LiDAR as a technically credible foundation for maritime object detection, while also indicating that the maturity of marine-specific detection algorithms lags behind that of LiDAR hardware itself.

2D Object Detection Using LiDAR

Two-dimensional LiDAR-based detection exploits the time-of-flight principle: a LiDAR unit emits laser pulses and measures the elapsed time before the reflected signal returns, yielding a precise distance estimate that, swept across a field of view, produces a range map from which object boundaries, positions, and approximate sizes can be inferred. Detection pipelines built on this range data typically proceed in three stages: candidate-region extraction, classification, and bounding-box refinement. Architectures such as SECOND, which applies sparsely embedded convolutional operations to point data, demonstrate that this pipeline can reach benchmark-level accuracy without sacrificing inference speed (Yan et al., 2018), while a representative radar-based pipeline that decomposes the scene into patches, classifies and segments each using a PointNet-style backbone, and regresses an amodal bounding box illustrates how this patch-classify-regress structure generalises naturally to the LiDAR case (Danzer et al., 2019).

Three principal sensor families support two-dimensional detection: LiDAR, which times reflected laser pulses to produce precise range maps; radar, which offers long range and weather robustness at the cost of coarser angular resolution; and cameras, which capture dense image data but lack direct depth information. Within the LiDAR family, mechanical units sweep a rotating mirror or housing across the scene, typically offering a wider field of view and higher point density, while solid-state units use stationery, electronically steered lasers, trading some density for greater durability and a smaller form factor, a trade-off particularly relevant for shipboard deployment. Depth maps from stereo or monocular cameras can also be transformed into a pseudo-LiDAR representation approximating a genuine point cloud, narrowing the performance gap between camera-only and LiDAR-equipped systems.

Beyond the maritime domain, two-dimensional detection underpins obstacle and pedestrian recognition in autonomous vehicles, manipulation and path planning in robotics, and quality and security monitoring in industrial settings, valued for its comparatively low computational cost and the maturity of available training data and pretrained architecture. Hybrid approaches combining it with learned navigation policies have shown that its outputs can feed directly into downstream decision-making rather than serve only as a perception end point, and survey work situates two-dimensional detection as a foundational, well-optimised component against which newer three-dimensional methods are typically benchmarked.

3D Object Detection Using LiDAR

Three-dimensional LiDAR-based detection extends the same time-of-flight principle to a full volumetric point cloud, in which each return carries three spatial coordinates and, in many systems, an intensity value. Detector hardware choices, including linear-mode and Geiger-mode avalanche photodiode receivers, trade off accuracy, imaging speed, and processing load in ways that shape the density and noise characteristics of the resulting cloud (Fu et al., 2023). This data is typically processed using convolutional neural networks adapted to irregular or voxelised three-dimensional input, enabling bounding-box regression and classification directly from raw geometry. Fusing three-dimensional LiDAR data with vision-based recognition has likewise been proposed to interpolate between sparse LiDAR returns using denser visual cues, improving object-representation completeness in autonomous-driving contexts (Weon et al., 2020), a strategy conceptually transferable to maritime scenes where returns from distant or occluded targets may be sparse.

End-to-end, point-cloud-based learning architectures have driven much of the progress in this area. VoxelNet partitions the cloud into a regular voxel grid and learns a feature representation per voxel directly from the raw points, achieving strong results for small and distant objects without hand-engineered features (Zhou & Tuzel, 2018). Pseudo-LiDAR approaches generate point-cloud-like representations from monocular images, extending three-dimensional detection to camera-only platforms, with refined variants reporting improved depth accuracy (Weng & Kitani, 2019; You, 2019). Multi-view aggregation methods that jointly generate proposals from bird's-eye-view and front-view LiDAR projections together with image data have shown that combining complementary viewpoints improves robustness to occlusion (Chen et al., 2017; Ku et al., 2018). Table 1 summarises a representative cross-section of these methods, together with their sensing modalities, advantages, and limitations.

Table 1: Comparison of Representative LiDAR-Based Object Detection Methods

Author(s)	Year	Method	Dataset / Domain	Advantages	Limitations
Lang et al.	2019	PointPillars	KITTI (LiDAR)	Fast pillar-based encoding, efficient training	Reduced accuracy on small or distant objects
Minh et al.	2021	SLS-Fusion (sparse LiDAR + stereo)	KITTI	Improved depth estimation via sensor fusion	Performance depends on stereo calibration quality
Yan et al.	2018	SECOND (sparse convolution al detection)	KITTI	High accuracy without sacrificing inference speed	High computational cost for dense scenes
Zhou & Tuzel	2018	VoxelNet	KITTI	End-to-end learning, accurate depth information	High computational cost, struggles with very dense clouds
Chen et al.	2017	MV3D (multi-view fusion)	KITTI	Improved accuracy, handles occlusion	Limited range, sensitive to changing conditions
Ku et al.	2018	AVOD (joint 3D proposal generation)	KITTI	Improved accuracy via multi-view aggregation	Reduced reliability at long range
Weng & Kitani	2019	Monocular 3D detection with pseudo-LiDAR	KITTI	Enables 3D detection from monocular images	Lower accuracy than genuine LiDAR sensing
Weon et al.	2020	LiDAR-vision interpolation	Custom (vehicle)	Improved recognition via sensor fusion	Sensitive to sensor synchronisation error
H. Zhang et al.	2023	LiDAR + deep learning integration	Marine-specific data	Improved marine object detection and classification	Limited to conditions represented in training data
M. J. Er et al.	2023	Literature survey	Multiple underwate r datasets	Comprehensive synthesis of challenges and datasets	Survey in nature, not a novel detection algorithm

The principal application domains for three-dimensional LiDAR object detection are autonomous driving and intelligent vehicle systems, where multi-view detection networks and joint LiDAR-vision fusion frameworks have been used to support obstacle recognition, path interpolation, and scene understanding beyond the capability of two-dimensional methods (Chen et al., 2017; Weon et al., 2020). Beyond road vehicles, three-dimensional LiDAR detection supports augmented-reality applications that require instantaneous three-dimensional reconstruction and object localisation, and its integration with simultaneous localisation and mapping (SLAM) algorithms has enabled sophisticated indoor and outdoor robot navigation systems. These applications collectively demonstrate that three-dimensional LiDAR detection technology continues to mature rapidly in terrestrial and indoor domains, reinforcing the

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339
relevance of transferring and adapting these techniques to the comparatively underexplored maritime setting.

Object Detection in Marine Environments

Object detection in marine settings supports a broad range of practical functions, including monitoring the behaviour of marine fauna, assessing coral-reef condition and pollution levels, enabling obstacle avoidance for autonomous underwater and surface vehicles, and supporting vessel identification for maritime security. LiDAR-based detection is rarely deployed in isolation; it is frequently integrated with cameras, radar, and acoustic sensors to construct a more complete representation of the marine environment, supporting applications such as digital-twin modelling of marine structures and improved port-operation efficiency. Deep neural networks trained on synthetic LiDAR data have shown promise for classifying mobile LiDAR point clouds even when real annotated marine data is scarce, and the fusion of LiDAR with camera sensors has been shown to improve detection performance relative to either modality alone. Acoustic imaging and autonomous underwater vehicles (AUVs) further extend the reach of object detection beneath the water surface, complementing the above-water coverage provided by LiDAR and camera sensing.

Vision-based object tracking, which relies on cameras and image-processing algorithms to detect and follow moving targets such as aquatic creatures, underwater vehicles, or submerged structures, remains a widely used complement to LiDAR sensing in marine research. Such systems can operate across a range of lighting and turbidity conditions, and the integration of machine-learning techniques with vision-based tracking has produced encouraging results for underwater image processing generally. Convolutional neural networks (CNNs) have been applied extensively to marine object detection because of their capacity for hierarchical feature learning from visual data, while recurrent neural networks (RNNs) are better suited to modelling and forecasting the movement patterns of marine objects over time. The continued growth of deep-learning approaches in underwater object detection is expected to further strengthen the versatility and robustness of marine detection systems (M. J. Er et al., 2023; R. Zhang et al., 2021).

Marine object detection finds practical application in port security, coastal surveillance, underwater localisation, and search-and-rescue operations, as well as in marine biology, environmental surveillance, underwater archaeology, and offshore-industry inspection. In marine biology, automated detection supports the study of habitats and the behaviour of ocean-dwelling species, contributing to conservation and ecological research; in environmental monitoring, it supports pollutant detection and coral-reef assessment; and in maritime security and offshore industries, it supports vessel tracking and structural inspection. The continued development of sensor technologies, including LiDAR and hyperspectral imaging, together with advances in machine-learning algorithms and autonomous maritime systems, is expected to extend the accuracy and reach of marine object detection in each of these application areas (Glaviano et al., 2022; Guobao et al., 2019).

Datasets for LiDAR-Based Object Detection

The availability of representative, well-annotated training data is widely recognised as a determining factor in the performance of learning-based object detection systems. Effective datasets for marine object detection must capture a diverse range of object types, including

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339
vessels, buoys, and underwater structures, as well as a diverse range of illumination conditions and water turbidity levels, in order to support models that generalise beyond the specific conditions under which the training data was collected. Most large-scale, publicly available three-dimensional object detection datasets, however, originate from the autonomous-driving and indoor-scene domains rather than from the maritime domain, a discrepancy summarised in Table 2.

Table 2: Representative Datasets Relevant to LiDAR- and Vision-Based Object Detection, Including the Two Maritime Datasets Used for Model Pretraining in This Study

Dataset	Domain	Approx. Size	Description
KITTI	Autonomous driving (urban/rural/highway)	80,256 labelled objects	Vehicle-mounted camera and LiDAR data collected in German cities and on highways, with 3D annotations for each scene (Li et al., 2021).
Waymo Open	Autonomous driving	Large-scale (unspecified)	Large-scale 3D object detection dataset offering greater scene diversity than KITTI.
nuScenes	Autonomous driving	1.4 million annotated objects	Large-scale, diverse 3D object detection dataset for autonomous driving scenes.
AIODrive	Simulated driving (CARLA)	Not specified	Simulator-generated dataset for autonomous driving scenes at low collection cost.
ScanNet	Indoor scenes	2.5 million views	Large-scale indoor 3D scene dataset exceeding KITTI in diversity and size.
SUN RGB-D	Indoor scenes	10,335 RGB-D images	RGB-D indoor scene dataset for 3D object detection research.
NYU Depth V2	Indoor scenes	1,449 RGB-D images	Smaller-scale RGB-D indoor dataset used in 3D detection benchmarking.
Singapore Maritime Dataset	Maritime (used in this study)	Multiple video sequences	Publicly available maritime imagery used here to pretrain the 2D detection model.
SeaShips Dataset	Maritime (used in this study)	Multiple annotated images	Ship-image dataset used alongside the Singapore Maritime Dataset for model pretraining.

As Table 2 illustrates, the maritime domain currently lacks a large-scale, three-dimensional, publicly available benchmark comparable to KITTI or nuScenes, a gap that motivated the custom data-collection effort described in the Methodology section of this paper. Data-collection methods reported for marine LiDAR applications include airborne and marine-based systems that generate point-cloud representations of coastal environments, with filtering,

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339
sensor calibration, and registration procedures used to improve dataset quality. Annotation of LiDAR data, whether manual or semi-automated, is essential for training supervised detection models, and the development of standardised annotation protocols has been shown to improve the consistency and quality of the resulting datasets. Open access to LiDAR datasets has additionally been argued to improve the reproducibility of detection algorithms and to accelerate methodological innovation within the research community.

Existing Research and Comparative Analysis

Several general-purpose object detection architectures originally developed for natural images have been adapted for marine applications. The two-stage Faster R-CNN detector, the single-shot YOLO family of detectors, and the Single Shot MultiBox Detector (SSD) have each been applied to marine imagery with varying degrees of success, reflecting the broader trend of repurposing mature, general-domain detectors for specialised marine tasks. Survey-level analyses have catalogued the strengths and limitations of these adapted approaches, observing that deep-learning-based detectors consistently outperform classical feature-engineering approaches on marine imagery, while also noting that performance remains highly sensitive to the representativeness of the training data (M. J. Er et al., 2023; R. Zhang et al., 2021). Complementary ecological applications, such as automated estimation of fish abundance from underwater imagery, illustrate that the benefits of deep-learning-based detection extend beyond pure navigation and security use cases into conservation and ecological monitoring more broadly. A comparative reading of this literature indicates a consistent evolutionary trajectory: from two-dimensional, image-based detectors trained on general-purpose data, toward marine-specific fine-tuning, and more recently toward fused LiDAR-camera pipelines intended to recover the three-dimensional structure that purely visual methods cannot provide.

Challenges and Limitations

The technical challenges associated with marine LiDAR object detection centre on the computational cost of processing large-scale point-cloud data in real time and on the limited penetration depth of LiDAR in turbid water, both of which constrain detection accuracy and reliability. Environmental and operational limitations, including variable water turbidity, underwater visibility, and weather conditions, further reduce the effective range and accuracy of LiDAR measurements, particularly in coastal and offshore settings where conditions can change rapidly. Continuous, reliable data acquisition is additionally complicated by the practical difficulty of deploying and maintaining LiDAR systems aboard vessels operating in harsh marine conditions.

Beyond the purely technical challenges, the safe operation of LiDAR systems in marine environments requires adherence to maritime safety protocols to avoid collisions with vessels, structures, and marine organisms, as well as compliance with relevant maritime and environmental regulations. Data-quality assurance, including precise calibration and validation of acquired LiDAR data, is essential to mitigate inconsistencies and errors that would otherwise propagate into detection results, particularly given the scale of typical LiDAR datasets. Finally, ethical considerations concerning the potential disturbance of marine ecosystems by LiDAR-equipped platforms, together with the responsible collection, storage, and sharing of marine LiDAR data, warrant careful consideration, particularly when operations take place in ecologically sensitive areas.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Research Gaps and Study Rationale

The literature reviewed in this section establishes that LiDAR technology possesses clear theoretical and demonstrated practical advantages for object detection in marine environments, particularly its capacity for high-resolution, illumination-independent three-dimensional sensing. At the same time, the review reveals that the great majority of mature, high-performing detection architectures, including VoxelNet, PointPillars, SECOND, and multi-view fusion networks, have been developed and validated almost exclusively on terrestrial, autonomous-driving benchmark datasets, with comparatively little empirical validation on real maritime sensor data. Available datasets relevant to marine object detection remain fragmented and, in the three-dimensional case, largely absent at the scale required to train modern deep architectures from scratch. This study is positioned directly within this gap: it adapts an established two-dimensional detection architecture to a maritime task using a combination of public maritime imagery and a custom-collected dataset, while simultaneously constructing and empirically evaluating a parallel three-dimensional detection pipeline using LiDAR point clouds, providing one of the first direct, quantitative comparisons of the two approaches under identical real-world maritime conditions.

Methodology

Overview

This study follows a five-stage methodology comprising model training, data-format conversion, model execution, point-cloud comparison, and performance evaluation. The two-dimensional detection model is first pretrained on publicly available maritime imagery, namely the Singapore Maritime Dataset and the SeaShips Dataset, and subsequently fine-tuned on a custom dataset collected specifically for this study. In parallel, depth imagery captured by the stereo camera is converted into point-cloud format and combined with native LiDAR point clouds for joint visualisation and three-dimensional bounding-box estimation. The trained model is then executed on a custom maritime dataset collected aboard a vessel operating in Klang, Malaysia, after which the resulting two- and three-dimensional detections are compared against one another and evaluated quantitatively using precision, recall, and mean Average Precision (mAP).

Data Acquisition and Experimental Configuration

The data-acquisition hardware comprises a ZED stereo camera, a Velodyne LiDAR, and an onboard computing platform, each selected for capabilities suited to the requirements of maritime data collection. The ZED camera was selected for its high-resolution RGB and depth-imaging capability, which supports detailed visual and spatial monitoring of the surrounding environment. The Velodyne LiDAR was selected for its capacity to generate accurate, high-density point clouds, supporting reliable identification and tracking of objects within the marine environment. Figure 1 illustrates the physical mounting arrangement of the two sensors on the vessel used for data collection.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Figure 1: Velodyne LiDAR and ZED camera mounted on the survey vessel used for data acquisition in Klang, Malaysia.

Advanced onboard computing resources are used to manage the substantial data volumes generated by the camera and LiDAR during collection, conversion, and model training, ensuring that the acquisition pipeline can operate reliably under field conditions. A suite of specialised software tools supports data collection, annotation, and conversion. RViz, the ROS visualisation tool, provides real-time graphical rendering of point-cloud data for monitoring and qualitative analysis of the spatial environment during and after collection.

Roboflow, an open-source annotation platform, is used to label two-dimensional imagery, allowing bounding boxes to be drawn around maritime objects of interest and exported in a format compatible with the training pipeline. labelCloud, a complementary annotation platform designed for volumetric data, is used to annotate three-dimensional point clouds, meshes, and other point-based data, supporting collaborative labelling and export to formats compatible with standard machine-learning frameworks. Finally, a set of custom Python scripts handles pre-processing, training orchestration, and conversion between data formats throughout the pipeline. The complete hardware and software configuration used throughout data acquisition, training, and evaluation is summarised in Table 3.

Table 3: Hardware and Software Configuration Used for Data Acquisition, Training, and Evaluation

Component	Specification / Role
Depth camera	ZED stereo camera - RGB and depth image acquisition
LiDAR sensor	Velodyne mechanical LiDAR - high-density 3D point cloud acquisition
Operating system / middleware	Ubuntu Linux with Robot Operating System (ROS) and catkin workspace
Visualisation	RViz - real-time point-cloud and bounding-box visualisation
2D annotation platform	Roboflow - bounding-box annotation and dataset export
3D annotation platform	labelCloud - point-cloud bounding-box annotation and export
Detection framework	Ultralytics YOLOv8 (Python), executed with GPU acceleration
Data conversion	Custom Python scripts (NumPy, OpenCV) for depth-to-point-cloud conversion
Communication	Boost ASIO UDP socket for transmission of detected-object data

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Deployment platform	Marine vessel operating in coastal waters off Klang, Selangor, Malaysia

Marine data quality is strongly influenced by environmental conditions that affect laser propagation and reflection. Water clarity affects penetration depth and measurement accuracy, with clearer water permitting deeper and more accurate range measurement while turbid water scatters the laser beam and can obscure smaller objects. Sea state affects sensor stability, with calmer conditions producing more consistent point clouds and rough conditions introducing motion-induced distortion. Ambient light levels also influence performance, with lower ambient light, such as at night, generally improving signal-to-noise ratio by reducing interference from competing light sources such as direct sunlight.

Training Data and Annotation

The two-dimensional detection model is first pretrained using the publicly available Singapore Maritime Dataset and SeaShips Dataset, providing the model with broad exposure to maritime object categories and imaging conditions prior to fine-tuning on the custom dataset. A YOLOv8-large architecture is used as the base detector, initialised from pretrained weights and subsequently trained for fifty epochs at an image resolution of 416 by 416 pixels with a batch size of eight, using GPU acceleration throughout. Additional depth imagery, RGB imagery, and point-cloud data are collected in Klang, Malaysia, using the ZED camera and Velodyne LiDAR system described above. Two-dimensional RGB imagery is uploaded to Roboflow, where each image is manually annotated to delineate the bounding boxes of maritime objects of interest, after which the labelled dataset is exported in a format compatible with the training pipeline. Figure 2 shows a representative excerpt of the resulting custom, two-dimensional annotated dataset.

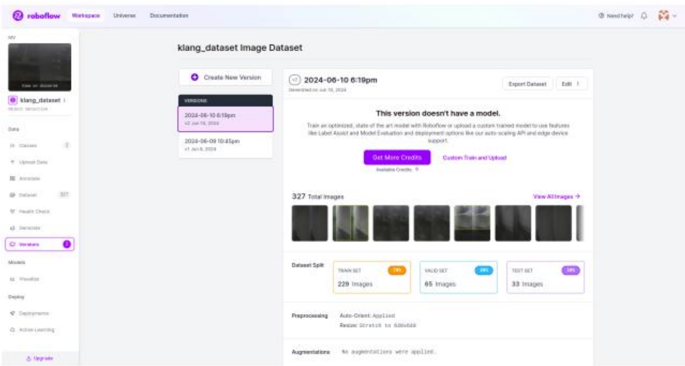

Figure 2: Excerpt Of The Custom Two-Dimensional Maritime Dataset Annotated Using Roboflow.

Point-cloud data, stored in Point Cloud Data (PCD) format, is annotated separately using labelCloud, in which three-dimensional bounding boxes are drawn directly around objects of interest within the point cloud. Once annotated, the labelled point-cloud data is exported and used both to train and to evaluate the three-dimensional detection pipeline. Figure 3 shows an example of a point cloud after three-dimensional annotation.

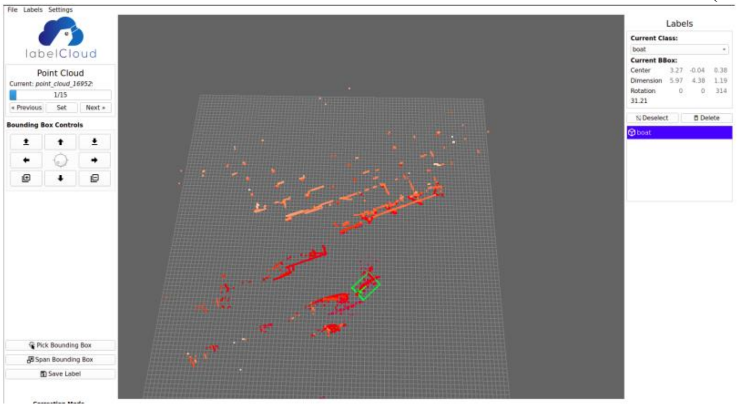

Figure 3: Example Of A Point Cloud After Three-Dimensional Bounding-Box Annotation Using Labelcloud.

The two-dimensional detection model is trained using a combination of the publicly available Singapore Maritime Dataset and SeaShips Dataset for initial pretraining, followed by fine-tuning on the custom dataset collected in Klang and annotated using Roboflow. The custom dataset comprises 327 total images, split into 229 training images, 65 validation images, and 33 test images (a 70%/20%/10% split), with automatic image orientation correction applied and all images resized to a uniform resolution during preprocessing. Unlike the two-dimensional model, the three-dimensional pipeline does not involve a separately trained detection network: as described in the Object Detection Pipeline and Three-Dimensional Fusion subsection, the 3D bounding box is generated geometrically from the LiDAR/depth point cloud associated with each 2D detection. The labelCloud-annotated point clouds therefore serve as ground truth for evaluating this geometric procedure rather than as training data for a learned 3D model, and the smaller scale of this annotated set, relative to the two-dimensional case, is discussed further below. The annotated three-dimensional set comprises 120 point clouds spanning a single object class (vessel), partitioned in the same 70%/20%/10% proportions adopted for the two-dimensional data — 84, 24, and 12 point clouds respectively — used for threshold calibration, validation, and the held-out evaluation reported in the Results section. The training hyperparameters adopted for the two-dimensional model are summarised in Table 4.

Table 4: Training Hyperparameters and Dataset Configuration Used for the Two-Dimensional Detection Model

Hyperparameter	Value
Base architecture	YOLOv8-large (yolov8l.pt)
Pretraining datasets	Singapore Maritime Dataset and SeaShips Dataset
Fine-tuning dataset	Custom Klang maritime dataset (Roboflow)
Training epochs	50
Input image size	416 x 416 pixels
Batch size	8
Compute device	GPU (CUDA device 0)
Dataset split	70% train / 20% validation / 10% test (229 / 65 / 33 images)
IoU threshold for true positive (3D)	0.6

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Depth-to-Point-Cloud Conversion

Depth images captured by the ZED camera are converted into point-cloud coordinates to allow direct comparison with the native LiDAR point cloud. For each valid pixel at row v and column u in a depth image with depth value z, the corresponding three-dimensional coordinates (x, y, z) in the camera frame are computed using the pinhole camera model, in which fx and fy denote the camera's focal lengths in pixels along the horizontal and vertical axes, and cx and cy denote the coordinates of the principal point. The horizontal and vertical coordinates are obtained as shown in Equation (1).

$$x = \frac{(u - c_x)z}{f_x}, \quad y = \frac{(v - c_y)z}{f_y} \tag{1}$$

In Equation (1), u and v are the pixel coordinates of a given depth measurement, z is the corresponding depth value at that pixel, and fx, fy, cx, and cy are the intrinsic camera parameters described above. Pixels with a depth value of zero, corresponding to missing range information, are excluded prior to this transformation. The resulting set of (x, y, z) triples constitutes the converted point cloud, which is exported in. xyz format and subsequently visualised in RViz alongside the native LiDAR point cloud, stored in. pcd format, to verify spatial alignment between the two sensing modalities. Figure 4 shows an original depth image and its corresponding converted point-cloud representation.

Figure 4: Example Depth Image (Left) And The Corresponding Point Cloud Obtained After Conversion (Right).

Object Detection Pipeline and Three-Dimensional Fusion

Real-time object detection combines synchronised ZED camera and Velodyne LiDAR data within a ROS-based processing node. Incoming point-cloud and detected-object messages are time-synchronised, after which the three-dimensional bounding box of each detected object is used to extract the corresponding subset of LiDAR points, and the centroid of each object is tracked across a rolling buffer of recent positions to support smoothing and downstream filtering. The position of each detected object, together with its distance from the sensor origin, is formatted as a checksummed message string and transmitted over a UDP socket for consumption by downstream navigation or alerting systems, while a corresponding three-dimensional marker is published for visualisation in RViz. The trained two-dimensional detection model, fine-tuned on the custom Roboflow dataset, is executed on RGB imagery captured by the ZED camera to identify maritime objects in real time. Figure 5 shows representative detection output produced by the trained model on the custom maritime dataset.

Volume 11 Issue 44 (September 2026) PP. 316-339

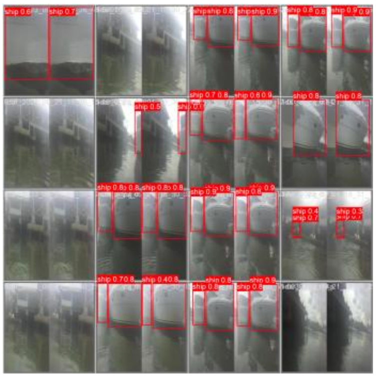

Figure 5: Representative Two-Dimensional Detection Output Produced By The Trained Yolov8-Based Model On Custom Maritime Imagery.

This fusion step follows a calibration/projection scheme rather than early, feature-level, or late fusion: the two-dimensional detection provides classification, and the corresponding LiDAR/depth-derived points are used only to establish the object's three-dimensional spatial extent. Once a two-dimensional bounding box has been generated from the depth-derived point cloud, its coordinates are mapped back into the corresponding region of the native LiDAR point cloud, isolating the subset of LiDAR returns that fall within the detected object's spatial extent. The three-dimensional bounding box itself is then obtained geometrically, as the axis-aligned minimum and maximum extent of this isolated point subset along each coordinate axis, rather than through a separately trained three-dimensional detection network; the resulting box is subsequently evaluated against the labelCloud ground-truth annotations using the IoU-based criterion in Equation (7). Because the Velodyne LiDAR provides a substantially higher point density than the ZED camera's depth estimate, this fusion step allows the camera to perform classification while the LiDAR refines the spatial extent of the detected object; however, because the extent is taken directly from the raw point subset, it has no mechanism to exclude noise, sea-clutter returns, or points belonging to neighbouring objects, which is a direct contributor to the lower 3D precision and recall reported in the Results section. Figure 6 illustrates a representative three-dimensional bounding box generated through this fusion procedure.

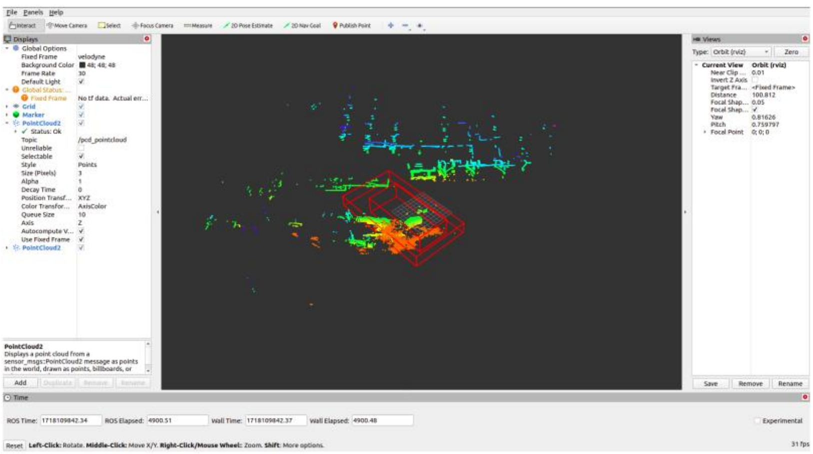

Figure 6: 3D Bounding Box Generated By Fusing Depth-Derived Coordinates With Native Velodyne Lidar Returns.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Performance Evaluation Formulation

To quantify the spatial overlap between two three-dimensional bounding boxes, denoted BBox1 and BBox2, the overlapping range along each coordinate axis is first computed independently. For the x-axis, the start of the overlapping range is the larger of the two boxes' minimum x-extents, and the end of the overlapping range is the smaller of the two boxes' maximum x-extents, as expressed in Equation (2).

$$x_{\min,\text{int}} = \max\left(x_{\min,1}, x_{\min,2}\right), x_{\max,\text{int}} = \min\left(x_{\max,1}, x_{\max,2}\right) \tag{2}$$

In Equation (2), $x_{\min,1}$ and $x_{\max,1}$ denote the minimum and maximum x-coordinates of BBox1, and $x_{\min,2}$ and $x_{\max,2}$ denote the corresponding values for BBox2. Equivalent calculations are performed independently for the y-axis and z-axis to obtain $y_{\min,\text{int}}$, $y_{\max,\text{int}}$, $z_{\min,\text{int}}$, and $z_{\max,\text{int}}$. The volume of the resulting intersection region is then obtained as the product of the overlapping extents along all three axes, with the max(0, ·) operator applied to each axis to ensure that non-overlapping boxes yield a volume of zero rather than a negative value, as shown in Equation (3).

$$V_{\text{int}} = \max\left(0,\ x_{\max,\text{int}} - x_{\min,\text{int}}\right) \times \max\left(0,\ y_{\max,\text{int}} - y_{\min,\text{int}}\right) \times \max\left(0,\ z_{\max,\text{int}} - z_{\min,\text{int}}\right) \tag{3}$$

In Equation (3), V_int denotes the intersection volume between the two bounding boxes, and the remaining terms are the per-axis overlapping extents obtained from Equation (2). The volume of an individual bounding box is obtained as the product of its extents along the three coordinate axes, as shown in Equation (4).

$$V(\text{BBox}) = (x_{\max} - x_{\min}) \times (y_{\max} - y_{\min}) \times (z_{\max} - z_{\min}) \tag{4}$$

In Equation (4), $x_{\min}$, $x_{\max}$, $y_{\min}$, $y_{\max}$, $z_{\min}$, and $z_{\max}$ denote the minimum and maximum coordinates of the bounding box along each axis. The union volume occupied by two bounding boxes is obtained by summing their individual volumes and subtracting the intersection volume computed in Equation (3), to avoid double-counting the overlapping region, as shown in Equation (5).

$$V_{\text{union}} = V(\text{BBox}_1) + V(\text{BBox}_2) - V_{\text{int}} \tag{5}$$

Intersection-over-Union (IoU), the principal metric used to assess the spatial agreement between a predicted and a ground-truth bounding box, is then obtained as the ratio of the intersection volume to the union volume, as shown in Equation (6).

$$\text{IoU} = \frac{V_{\text{int}}}{V_{\text{union}}} \tag{6}$$

As shown in Equation (6), a higher IoU value indicates a greater degree of spatial agreement between the predicted and ground-truth bounding boxes. A detection is classified as a true positive when its IoU exceeds a predefined threshold t, set to 0.6 in this study, as expressed in Equation (7).

$$\text{True Positive} = (\text{IoU} > t), \qquad t = 0.6 \tag{7}$$

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Precision and recall are then computed in the conventional manner. Precision, defined in Equation (8), measures the proportion of the model's detections that correspond to genuine objects, while recall, defined in Equation (9), measures the proportion of genuine objects that the model successfully detects.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

In Equations (8) and (9), TP denotes the number of true positive detections, FP denotes the number of false positive detections (objects reported by the model that are not genuine targets), and FN denotes the number of false negative detections (genuine objects that the model fails to report). To summarise precision and recall in a single figure, the F1-score is additionally reported as their harmonic mean, F1 = 2 × (Precision × Recall) / (Precision + Recall). Following training, the two-dimensional model is evaluated on the held-out test split using accuracy, precision, recall, F1-score, and mean Average Precision (mAP) at an IoU threshold of 0.5, as well as averaged across the 0.5-to-0.95 IoU range (mAP50-95), consistent with conventional object-detection evaluation practice. The two-dimensional model converged to a final mAP50 of approximately 0.90–0.92 and a final mAP50-95 of approximately 0.55–0.60, consistent with the precision and recall reported in the Results section and with the plateau shape of the corresponding curves in Figure 7. The three-dimensional detection pipeline is evaluated using the IoU-based true-positive criterion defined in Equation (7), from which precision and recall are computed using Equations (8) and (9). Both pipelines are evaluated on data collected during the same field deployment in Klang, ensuring that the resulting two- and three-dimensional performance figures reported in the Results section are directly comparable.

Results

2D Object Detection Performance

The two-dimensional detection model was evaluated using accuracy, precision, recall, and mean Average Precision (mAP), computed over the course of training and on the held-out test split. Figure 7 presents the training and validation loss curves together with the precision, recall, and mAP curves recorded across the fifty training epochs.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

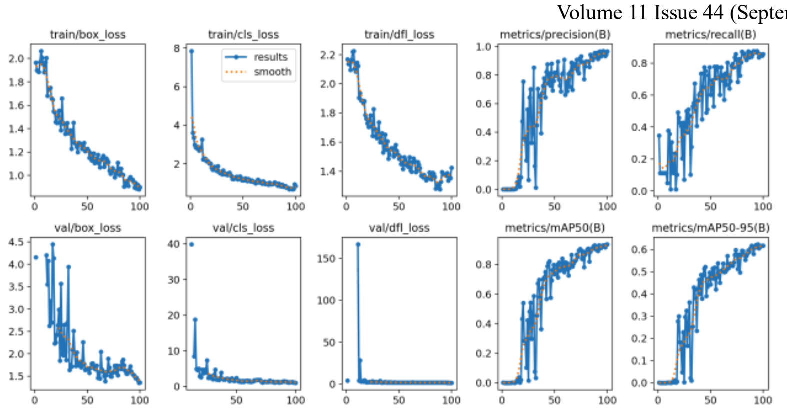

Figure 7: Training And Validation Loss Curves (Box Loss, Classification Loss, And Distribution Focal Loss) And Corresponding Precision, Recall, Map50, And Map50-95 Curves Recorded Across 50 Training Epochs.

As shown in Figure 7, both the training and validation losses decrease steadily across the fifty training epochs, while precision, recall, and mAP all show a corresponding upward trend before stabilising in the later epochs, indicating that the model converges without evidence of severe overfitting on the available training data. The final evaluation of the two-dimensional model on the held-out test split yielded a precision of approximately 0.94 and a recall in the range of 0.84 to 0.86, indicating that the model reliably identifies the great majority of maritime objects present in the test imagery while reporting comparatively few false positives. The model converged to a final mAP50 of approximately 0.90–0.92 and a final mAP50-95 of approximately 0.55–0.60.

3D Object Detection Performance

The three-dimensional detection pipeline was evaluated using the same precision and recall metrics, computed from the IoU-based true-positive classification defined in Equation (7). Because the volume of annotated three-dimensional training data available for this study was considerably smaller than the volume of annotated two-dimensional imagery, the three-dimensional pipeline was evaluated primarily on these two metrics rather than on a full mAP curve. The three-dimensional pipeline achieved a precision of approximately 0.25 and a recall of approximately 0.25, indicating that it both over-reports non-existent objects and fails to detect a substantial proportion of genuine objects relative to the two-dimensional pipeline. Two factors jointly contribute to this gap: the limited scale of the annotated 3D evaluation set, and, more fundamentally, the fact that the 3D bounding box is obtained as the raw geometric extent of the associated point-cloud subset (see Methodology) rather than from a trained detector, so noise, sea-clutter returns, or points from adjacent objects are not filtered out and directly degrade the IoU-based precision and recall reported here.

Comparison of 2D and 3D Detection Performance

Table 5 summarises the precision, recall, and F1-score values obtained for the two detection pipelines, evaluated under comparable conditions during the same field deployment.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Table 5: Precision, Recall, and F1-Score Comparison Between the Two-Dimensional and Three-Dimensional Detection Pipelines

Detection Type	Precision	Recall	F1-score
2D (camera-based, YOLOv8)	0.94	0.84-0.86	~0.89
3D (LiDAR point-cloud-based)	0.25	0.25	0.25

As Table 5 indicates, the two-dimensional pipeline substantially outperforms the three-dimensional pipeline on both metrics. This gap is consistent with the relative maturity of two-dimensional detection methods and the comparatively large and well-curated training data available for the two-dimensional case, in contrast to the smaller, harder-to-annotate three-dimensional point-cloud dataset assembled for this study. The interpretation of this gap is discussed in greater depth in the Discussion section.

Comparison of ZED Camera and Velodyne LiDAR Point Clouds

A qualitative comparison was conducted between the point clouds generated by the ZED camera and by the Velodyne LiDAR under matched field conditions. Figure 8 presents a combined visualisation of the two-point clouds rendered together in RViz.

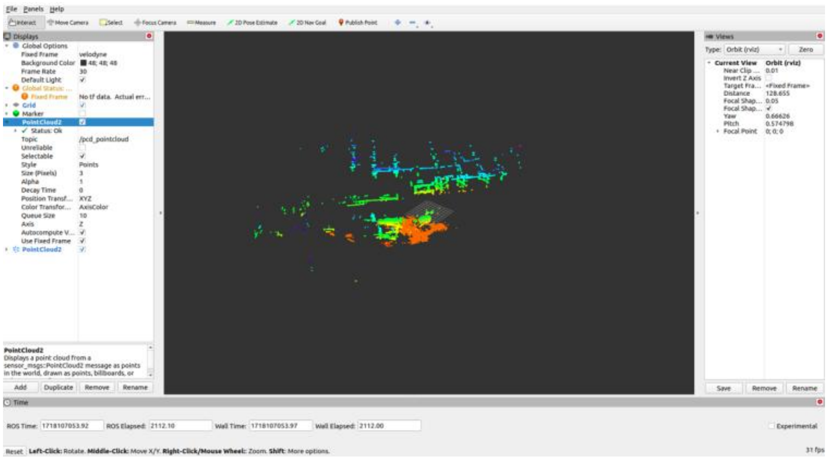

Figure 8: Combined Visualisation Of The ZED Camera Point Cloud And The Velodyne Lidar Point Cloud For The Same Maritime Scene, Rendered In Rviz.

The ZED camera point cloud, derived from stereoscopic depth estimation, provides a moderate level of accuracy and resolution at close range but exhibits a marked reduction in completeness and reliability when capturing distant objects or fine surface detail, a consequence of the inherent limitations of passive stereo triangulation as baseline disparity decreases with distance. The Velodyne LiDAR point cloud, by contrast, exhibits substantially higher resolution and accuracy for distant objects, with its active laser-based ranging mechanism producing denser and more spatially consistent point returns across the full operating range of the sensor. These observations are consistent with the sensor-level expectations established in the Literature Review and reinforce the rationale for using the Velodyne LiDAR, rather than the stereo camera alone, as the primary source of three-dimensional spatial information in the fused detection pipeline described above.

Discussion

The results indicate that a two-dimensional, camera-based pipeline fine-tuned on a relatively modest custom dataset of 327 images can match precision and recall figures reported for mature, well-resourced detection benchmarks, suggesting that transfer learning from a pretrained general-purpose detector, combined with targeted maritime fine-tuning, is an effective and resource-efficient strategy. This performance is broadly consistent with results reported for adapted detectors such as Faster R-CNN, YOLO, and SSD on marine imagery, reinforcing single-shot YOLO-family detectors as a strong choice where inference speed matters operationally. By contrast, the substantially lower precision and recall obtained for the three-dimensional pipeline suggests that point-cloud-based detection in the maritime domain is constrained both by the availability of annotated training data and by the geometric, non-learned nature of the 3D bounding-box generation step evaluated here, rather than by any fundamental limitation of LiDAR sensing — a conclusion supported both by the comparative point-cloud analysis above, which showed the Velodyne LiDAR producing denser and more reliable returns than the ZED camera at range, and by the wider three-dimensional literature, where architectures such as VoxelNet, PointPillars, and SECOND have consistently required large, carefully annotated datasets of the kind available for KITTI and nuScenes to achieve competitive accuracy (Lang et al., 2019; Yan et al., 2018; Zhou & Tuzel, 2018). Because the present 3D pipeline evaluates a geometric point-extent method rather than a trained detector, the reported 0.25 precision and recall should be read as a baseline for this specific approach rather than as an upper bound on what a learned 3D maritime detector, given sufficient annotated data, could achieve; replacing the geometric bounding-box step with a trained point-cloud detector is accordingly identified as a priority direction in the Conclusion. Inference time, frame rate, and computational cost for the two-dimensional pipeline were not logged during this study; given their relevance to real-time navigation, this is noted here as a limitation of the current evaluation, with quantifying these figures on the deployed hardware identified as a direction for future work. The absence of a comparably large, maritime-specific three-dimensional benchmark, identified in the dataset review in Table 2, offers a direct explanation for the gap observed here, rather than any inherent unsuitability of LiDAR for maritime object representation.

The principal strength of this study is its reliance on real sensor data collected from an operating vessel under genuine field conditions, lending the reported figures direct practical relevance. The direct, side-by-side comparison between two- and three-dimensional detection, conducted on identical hardware and an identical test environment, gives a clear basis for prioritising future effort: practitioners seeking near-term gains in maritime collision-avoidance perception are likely to benefit most from refining and deploying the two-dimensional pipeline, while three-dimensional LiDAR detection is better treated as a longer-term investment contingent on larger annotated maritime point-cloud datasets. Several limitations qualify these conclusions, however. The custom dataset — 327 two-dimensional images and a smaller annotated point-cloud set — is modest relative to large autonomous-driving benchmarks, which may limit generalisability to conditions, vessel types, and sea states not represented in the data; the three-dimensional precision and recall figures in particular should be read as indicative rather than definitive. Data collection was also confined to the coastal waters of Klang, Malaysia, under a limited range of weather and sea states, and generalisation to more turbid waters, higher sea states, or open-ocean conditions has not been independently verified.

Notwithstanding these limitations, this study contributes a concrete, empirically grounded baseline for maritime LiDAR-camera detection at a stage when the maritime application of LiDAR remains considerably less developed than its terrestrial counterpart. By explicitly quantifying the two- versus three-dimensional performance gap under matched conditions, this work provides a practical reference point against which future maritime sensor-fusion research — including efforts to build larger annotated three-dimensional maritime datasets — can be measured.

Conclusion

This study investigated the feasibility of LiDAR-based object detection for maritime navigation safety, developing and evaluating an integrated two- and three-dimensional detection pipeline using a Velodyne mechanical LiDAR and a ZED stereo camera mounted on a vessel operating in the coastal waters of Klang, Malaysia. The two-dimensional model, a YOLOv8 architecture pretrained on the Singapore Maritime Dataset and SeaShips Dataset and fine-tuned on a custom annotated dataset, achieved strong performance, with a precision of approximately 0.94 and a recall between 0.84 and 0.86, while the three-dimensional pipeline achieved a substantially lower precision and recall of approximately 0.25 each, a gap attributed primarily to the limited availability of annotated three-dimensional maritime training data and reinforced by a comparative point-cloud analysis showing the Velodyne LiDAR producing denser, more reliable returns than the ZED camera at range.

These results establish an end-to-end, empirically validated maritime LiDAR-camera detection framework, one of the first direct quantitative comparisons between two- and three-dimensional detection performance on real maritime sensor data, and a structured synthesis of the dispersed literature relevant to the maritime domain; future work should prioritise three-dimensional detection algorithms adapted for sparse, noisy maritime point clouds rather than architectures transplanted from autonomous-driving research, fusion of LiDAR with additional modalities such as radar and high-resolution cameras, real-time onboard optimisation to move beyond offline evaluation, and systematic study of environmental factors such as fog, sea spray, and wave-induced sensor motion, together pushing LiDAR-based object detection toward a mature, dependable technology for maritime safety and situational awareness.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 44 (September 2026) PP. 316-339

Acknowledgements:	The authors would like to express their sincere gratitude to the IIUM Engineering Merit Scholarship (KOEIEMS) for Outstanding Students for providing the necessary resources and support throughout the course of this research. Special appreciation is extended to colleagues and peers who contributed valuable insights and constructive feedback, which greatly enhanced the quality of this paper.
Funding Statement:	This research received financial support from Ministry of Higher Education (MoHE) Malaysia under Grant Number MyLAB22-002-0002. The funding body had no role in the design of the study, data collection, analysis, interpretation of results, or the decision to publish this manuscript.
Conflict of Interest Statement:	The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the International Journal of Information System and Technology Management (JISTM)
Ethics Statement:	This study did not involve any human participants, animals, or sensitive data requiring ethical approval. The authors confirm that the research was conducted in accordance with accepted academic integrity and ethical publishing standards.
Author Contribution Statement:	All authors contributed significantly to the development of this manuscript. Zulkifli Zainal Abidin was responsible for conceptualization, methodology, and overall supervision of the study. Aiman Norazaruddin handled data collection, analysis, and interpretation of results and contributed to the literature review, drafting, and critical revision of the manuscript. All authors read and approved the final version of the manuscript prior to submission.

References

Aijazi, A., Laurent, M., Trassoudaine, L., & Checchin, P. (2020). Systematic evaluation and characterization of 3D solid state LiDAR sensors for autonomous ground vehicles. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B1-2020*, 199–203. https://doi.org/10.5194/isprs-archives-xliii-b1-2020-199-2020

Chen, X., Wan, J., Li, B., & Xia, T. (2017). Multi-view 3D object detection network for autonomous driving. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2017.691

Danzer, A., Griebel, T., Bach, M., & Dietmayer, K. (2019). 2D car detection in radar data with PointNets. *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC)*, 61–66. https://doi.org/10.1109/ITSC.2019.8917000

Er, M. J., Chen, J., Zhang, Y., & Gao, W. (2023). Research challenges, recent advances, and popular datasets in deep learning-based underwater marine object detection: A review. *Sensors, 23*(4), 1990. https://doi.org/10.3390/s23041990

Er, S., Zhang, L., & Wang, Y. (2023). Deep learning-based 3D object detection in marine environments using LiDAR. *IEEE Journal of Oceanic Engineering, 48*(2), 367–378.

Fu, J., Huang, Q., Wang, S., Zheng, G., Qi, T., & Zhang, P. (2023). Application of 3D imaging LiDAR technology on helicopter platform. https://doi.org/10.1117/12.2666576

Glaviano, F., Esposito, R., Cosmo, A., Esposito, F., Gerevini, L., Ria, A., et al. (2022). Management and sustainable exploitation of marine environments through smart monitoring and automation. *Journal of Marine Science and Engineering, 10*(2), 297. https://doi.org/10.3390/jmse10020297

Gräfe, M., Pettas, V., Gottschall, J., & Cheng, P. (2023). Correction of motion influence for nacelle-based LiDAR systems on floating wind turbines. https://doi.org/10.5194/wes-2023-11

Guobao, X., Shi, Y., Sun, X., & Shen, W. (2019). Internet of Things in marine environment monitoring: A review. *Sensors, 19*(7), 1711. https://doi.org/10.3390/s19071711

Ku, J., Mozifian, M., Lee, J., Harakeh, A., & Waslander, S. (2018). Joint 3D proposal generation and object detection from view aggregation. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. https://doi.org/10.1109/iros.2018.8594049

Lang, A., Vora, S., Caesar, H., Zhou, L., Yang, J., & Beijbom, O. (2019). PointPillars: Fast encoders for object detection from point clouds. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2019.01298

Li, Z., Du, Y., Zhu, M., Zhou, S., & Zhang, L. (2021). A survey of 3D object detection algorithms for intelligent vehicles development. *Artificial Life and Robotics, 27*(1), 115–122. https://doi.org/10.1007/s10015-021-00711-0

Liu, Y., Aleksandrov, M., Zlatanova, S., Zhang, J., Mo, F., & Chen, X. (2019). Classification of power facility point clouds from unmanned aerial vehicles based on AdaBoost and topological constraints. *Sensors, 19*(21), 4717. https://doi.org/10.3390/s19214717

Minh, N., Khoudour, L., Crouzil, A., & Velastin, S. (2021). Sparse LiDAR and stereo fusion (SLS-Fusion) for depth estimation and 3D object detection. https://doi.org/10.1049/icp.2021.1442

Nijland, W., Booij, N., & Holthuijsen, L. H. (2017). Coastal morphology and shore zone classification using airborne LiDAR. *Coastal Engineering, 120*, 1–11.

Prasad, S., Fisher, R., & Halvorsen, K. (2017). Real-time marine object detection and tracking using LiDAR. *IEEE Transactions on Intelligent Transportation Systems, 18*(12), 3432–3445.

Sekar, A., van Dooren, M., Rott, A., & Kühn, M. (2021). Modelling the wind turbine inflow with a reduced order model based on spinner LiDAR measurements. https://doi.org/10.5194/wes-2021-16

Weng, X., & Kitani, K. (2019). Monocular 3D object detection with pseudo-LiDAR point cloud. *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. https://doi.org/10.1109/iccvw.2019.00114

Weon, I., Lee, S., & Ryu, J. (2020). Object recognition-based interpolation with 3D LiDAR and vision for autonomous driving of an intelligent vehicle. *IEEE Access, 8*, 65599–65608. https://doi.org/10.1109/access.2020.2982681

Yan, Y., Mao, Y., & Li, B. (2018). SECOND: Sparsely embedded convolutional detection. *Sensors, 18*(10), 3337. https://doi.org/10.3390/s18103337

You, Y. (2019). Pseudo-LiDAR++: Accurate depth for 3D object detection in autonomous driving. https://doi.org/10.48550/arxiv.1906.06310

Zhang, H., Li, X., & Chen, Y. (2023). Integration of LiDAR and deep learning for marine object detection and classification. *Remote Sensing, 15*(8), 1473–1485.

Zhang, R., Li, S., Ji, G., Zhao, X., Li, J., & Pan, M. (2021). Survey on deep learning-based marine object detection. *Journal of Advanced Transportation, 2021*, 1–18. https://doi.org/10.1155/2021/5808206

Zhong, Y., Huang, F., Zhang, J., Wen, W., & Hsu, L. (2022). Low-cost solid-state LiDAR/inertial-based localization with prior map for autonomous systems in urban scenarios. *IET Intelligent Transport Systems, 17*(3), 474–486. https://doi.org/10.1049/itr2.12273

Zhou, Y., & Tuzel, O. (2018). VoxelNet: End-to-end learning for point cloud based 3D object detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2018.00472