

Evaluation Report

Orca AI Computer Vision Assessment

Project no. : 2409-0018

Revision no.: 1

July 2025

ORCA AI

Contents

1. Introduction
2. Enhanced Situational Awareness (ESA) – Computer Vision
3. Evaluation Vessel and Voyage
4. Orca AI Onboard Placement
5. Orca AI System’s Screen
6. Evaluation Methodology
7. Evaluation Criteria – Precision and Recall Score
8. Results
    1. Condition
    2. Detection Sample
    3. Precision and Recall
    4. Missed Radar Detections
9. Situational Awareness Impact - General Crew Interaction

Introduction

Lloyd's Register Pte. Ltd., hereafter referred to as LR, carried out an assessment of Orca AI Computer Vision system to evaluate the object detection performance of the technology.

The scope of work encompasses evaluating the precision and recall metrics of the Orca AI computer vision detection algorithm by conducting field tests during an actual voyage. In this assessment, Orca AI classification capability is not evaluated.

This evaluation report provides an overview of the assessment conducted, including the testing methodology, prevailing conditions during the evaluation, and the specific technical criteria and definitions applied.

The report will first present the technical evaluation results, followed by a summary of the vessel crew's general observations and feedback regarding the computer vision system being tested. This structured approach aims to offer both quantitative performance insights and practical, user-based perspectives.

For this evaluation, LR and Orca AI jointly developed a methodology to assess AI-based Enhanced Situational Awareness solutions. The approach combined technical evaluation with crew behaviour and insights to assess system performance holistically.

Enhanced Situational Awareness (ESA) – Computer Vision

Navigational challenges at sea are growing more complex with rising traffic density, unpredictable weather, and tighter environmental standards. Enhanced Situational Awareness (ESA) systems are emerging in aiding crews to identify potential hazards sooner, supporting better decision-making, and ultimately achieve safer and more efficient navigation.

Watchkeeping, a critical component of situational awareness (SA) in maritime operations, stands to gain from advancements in automation and Artificial Intelligence (AI). In particular, the integration of computer vision technology into ESA offers real-time monitoring and analysis capabilities that support decision-making.

Orca AI SeaPod

Orca AI, established in 2018, developed the Orca AI SeaPod, a computer vision system designed for ESA in maritime operations.





The SeaPod features a fixed sensor head positioned near the ship's centre line, equipped with five optical cameras (FOV 225^0) and three thermal cameras (FOV 100^0). It employs AI-driven technology for real-time detection, classification, and distance estimation of surrounding relevant maritime objects.

Potential Use Cases & Operational Relevance

Night-time Lookout

In open waters, automation can relieve the Lookout, allowing for improved rest and overall wellbeing.

Congested Water

Assist in detection of small, non-AIS target, especially with minimal radar signatures. Improve accuracy in perception and FOV coverage. Reduce workload and fatigue for maintaining SA.

Poor Visibility

In poor weather like fog or rain, advanced systems improve visibility and target detection for safer navigation

Evaluation Vessel and Voyage

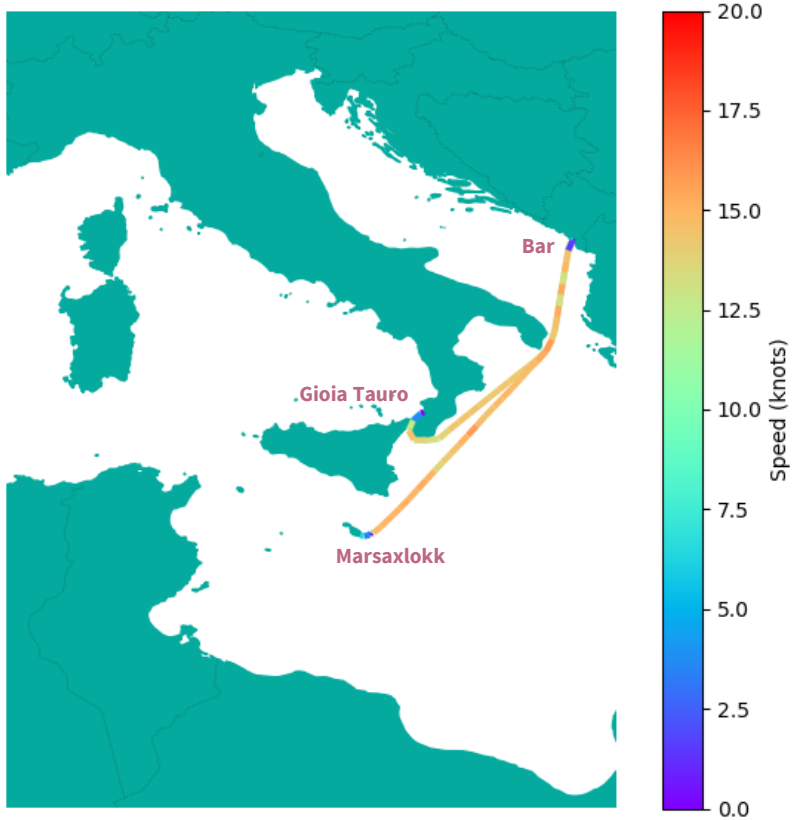
The evaluation was conducted on the MSC Adriana II (IMO 9169055) over a period of five days. The vessel began its voyage in Gioia Tauro, Italy, proceeded to Bar, Montenegro, and concluded in Marsaxlokk, Malta, where the evaluation was completed. The voyage took place from 30 June to 4 July 2025, with clear weather observed throughout. The total distance covered during the evaluation is 828 NM.

MSC Adriana II (IMO 9169055)



Parameter	Value
Length Overall (LOA)	216.00 m
Length between perpendiculars (LBP)	198.70 m
Breadth Moulded	26.66 m
Draught	10.40 m
Depth	18.95 m
Gross Tonnage (GT)	25,219
Deadweight (DWT	18,779

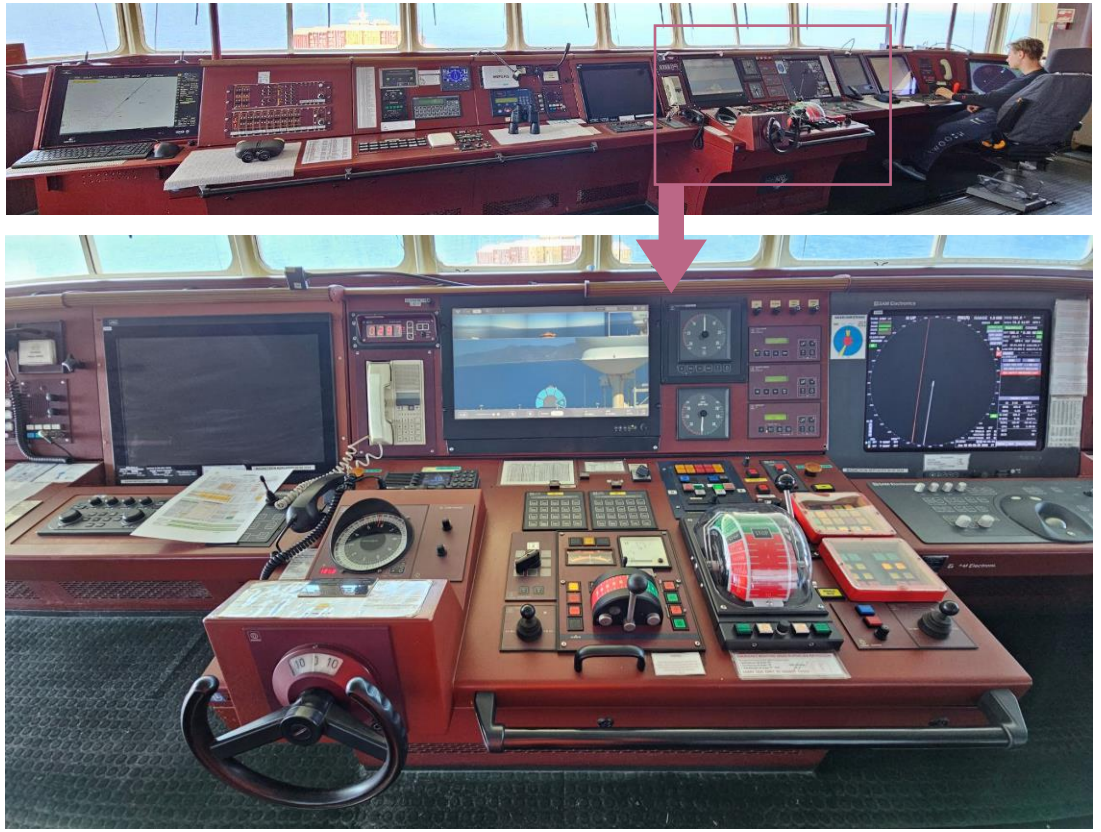
Vessel route during the evaluation



Orca AI Onboard Placement

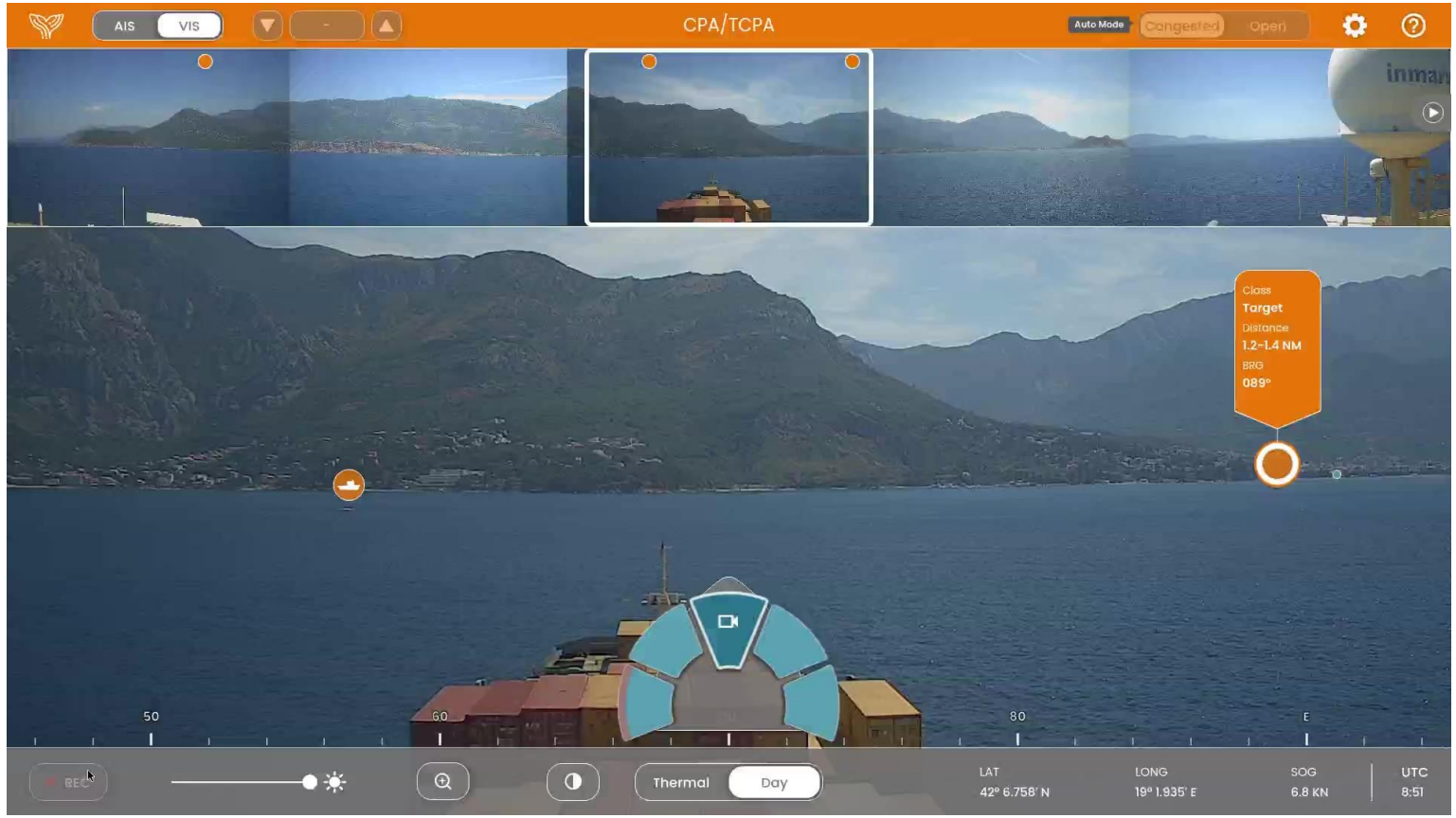


The Orca AI SeaPod camera is mounted on top of the vessel’s bridge, providing clear visibility of the vessel’s surroundings



The Orca AI system’s screen is centrally positioned on the bridge console, directly in front of the piloting officer.

Orca AI System's Screen



The Orca AI System's interface displayed on the vessel's bridge console shows the day camera in use, with the centre camera selected as the primary focus. Both Visual (VIS) and Congested modes are activated in this instance.

Classification & Indications



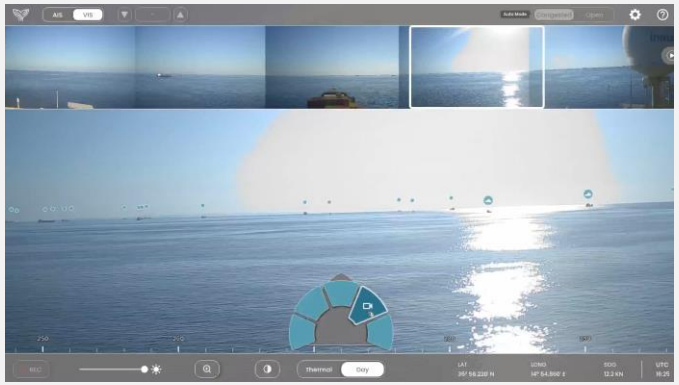
Symbols shown in Orca AI System's screen.
Source: Orca AI – SeaPod Training Booklet

Evaluation Methodology

During the voyage, observations were taken approximately every 30 minutes in open sea conditions. This interval was reduced to every 5 minutes when the vessel navigated congested areas, such as near ports, the Strait of Messina, and the anchorage area near Marsaxlokk.

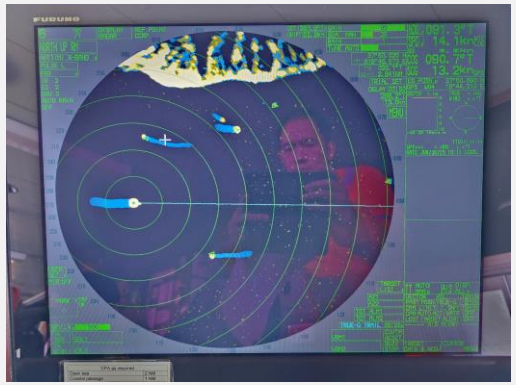
Ground truth data was generated and used as a benchmark to evaluate the computer vision detections. This was achieved through a combination of sources, including the Orca AI system's screen and its recordings, radar, AIS data via ECDIS, and visual observations. Ground truth is generated by LR and Orca AI representatives with the assistance of the officer of the watch.

Orca AI System's Screen



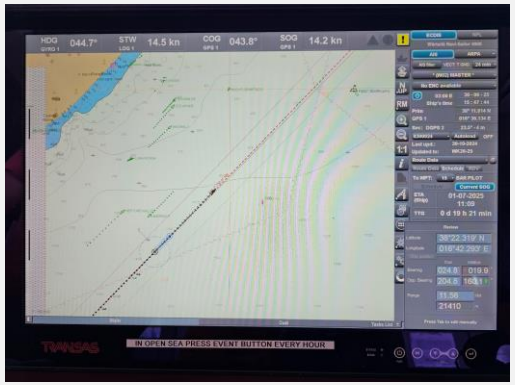
Detections captured by the Orca AI SeaPod in visual mode have been recorded, with evaluations conducted across all cameras (five in day mode and three in thermal mode).

Radar (via ARPA)



Targets detected by radar, as displayed on the ARPA, are used along with their corresponding distances.

AIS (via ECDIS)



AIS data displayed on the ECDIS, including available information such as object type and reported distance, is utilised.

Visual



Visual observations are used, with the aid of binoculars at times.

In each observation, a recording of the Orca AI system's screen is captured, along with photos of the ARPA and ECDIS screens, and visual observations are documented. The period that is assumed to be relevant in each of the observations is assumed to be approximately 1 minute.

Evaluation Criteria – Precision and Recall Score

The evaluation emphasises detection accuracy, with particular attention to assessing **Precision and Recall (P&R)** performance.

- Precision (P): Measures the system's accuracy in detecting and classifying targets, ensuring minimal false positives.
- Recall (R): Evaluates the completeness of detection, emphasising the system's ability to identify all relevant objects or events within the maritime environment

To calculate the P&R score of the computer vision technology, the following parameters are determined:

True Positive (TP)

Orca AI detections that match the Ground Truth (GT)-identified targets. For detection to be considered as a TP, the following criteria are used:

- **Target Detection.** The system must consistently detect maritime objects that are visibly present on the screen on the Orca AI system's screen, with the following operational ranges:
  - Large Vessels: up to 5 NM
  - Small Vessels: up to 3 NM
  - Buoys: up to 1 NM
- **Tracking Consistency.** Once detected, objects should remain tracked without significant loss errors.
- **Environmental Adaptability.** The system must operate reliably and detect all visible objects on the UI across varying conditions, including daylight and nighttime, as well as fog, rain, and clear weather

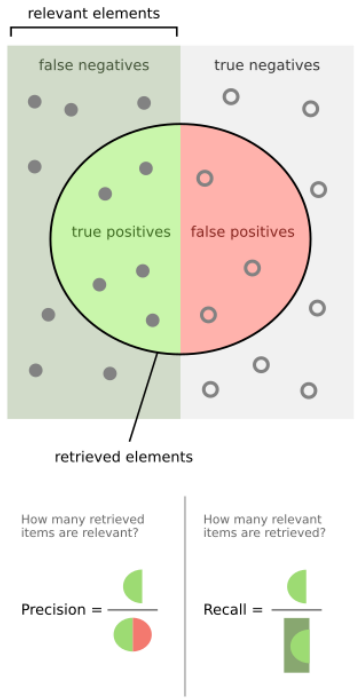
Good detection outside operational ranges is considered in this evaluation.

False Positive (FP)

Orca AI detections that do not match any ground truth (GT) target.
A detection is counted as an FP if it appears consistently, and this also covers cases where the system repeatedly detects the same single GT target as multiple separate objects.

False Negative (FN)

Ground truth (GT) targets that were not detected by Orca AI.
For a target to be counted as an FN, it must be within the camera's field of view. Targets that are completely blocked by other objects, even if they fall within the system's operational range, are not visible to the camera and therefore are excluded from the FN count.



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

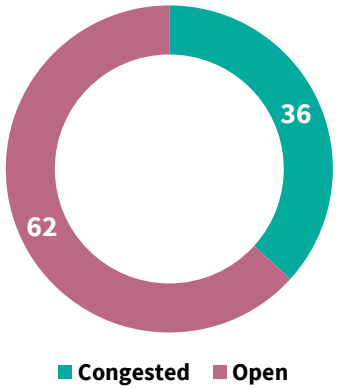
$$\text{Recall} = \frac{TP}{TP+FN}$$

Results

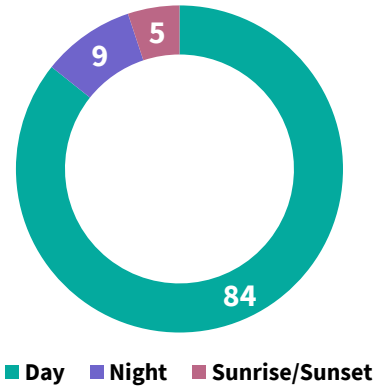
Results - Condition

- A total of 98 observations were collected during the evaluation, covering 739 relevant targets.
- In 72% of the observations, fewer than 7 targets were observed, while 13% recorded more than 15 relevant targets.
- The highest number of targets detected by the computer vision system in a single observation was 45.
- There were also instances with no targets detected, typically when the vessel was sailing alone in open sea within a 10 NM radius.

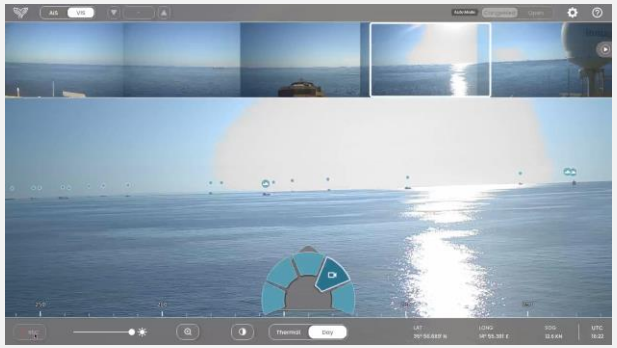
Mode Distribution



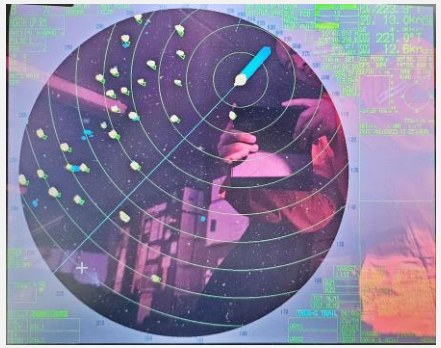
Time of Day Distribution



Example of Congested Condition



Orca AI System's Screen



Radar + AIS (via ARPA)

- The majority of the observations (63%) were conducted under congested conditions, providing an optimal environment to evaluate the system's detection performance.
- The majority of observations were conducted during daylight hours, as the vessel primarily sailed in open sea conditions at night when no relevant targets were present.

Results – Detection Sample

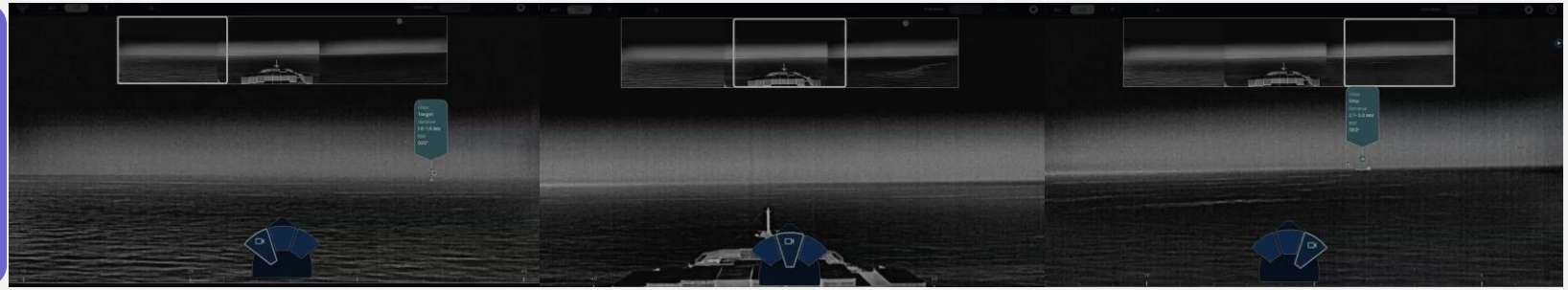
Sample Detection at Night

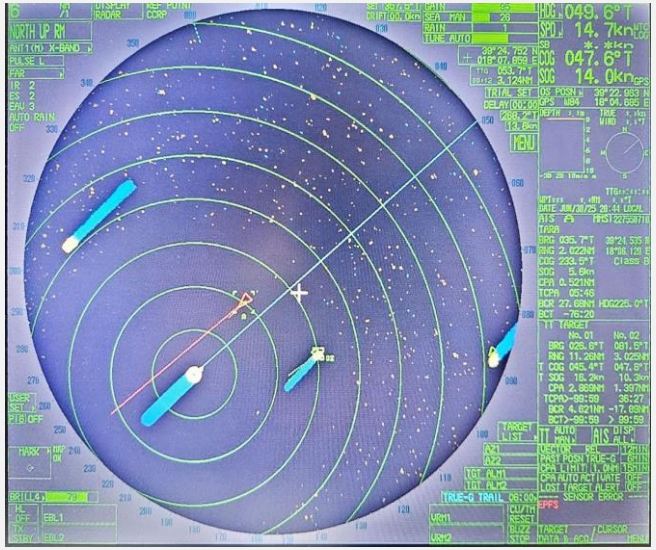
Day Camera



Left Camera | Center Camera | Right Camera

Thermal Camera





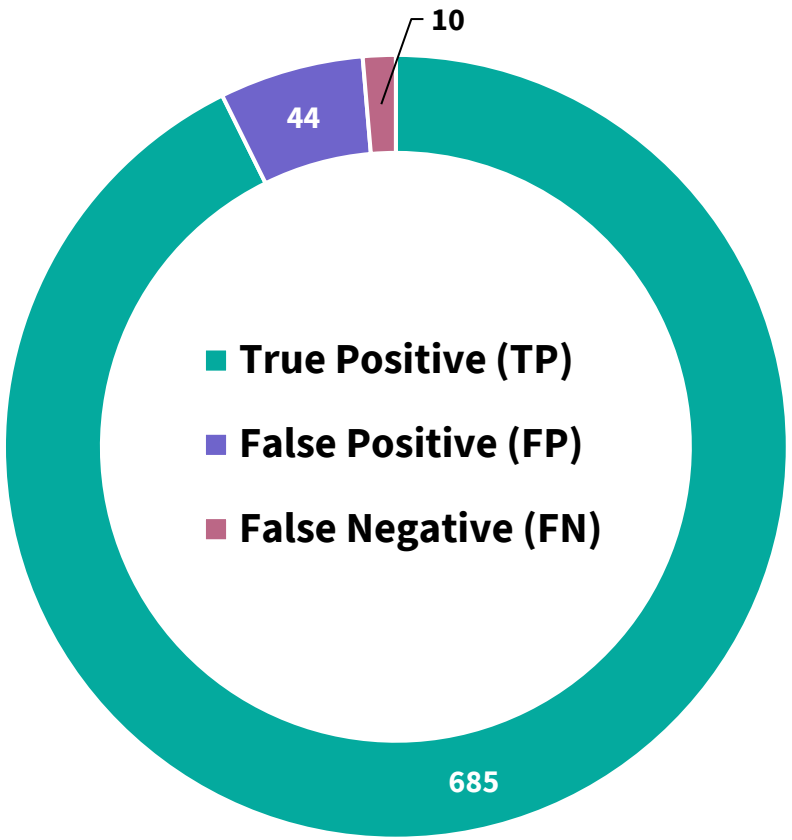
Radar + AIS (via ARPA)

Note: Field of View (FOV) of each day and thermal camera are different. The images shown above are not to scale and do not accurately represent the relative sizes or perspectives of the two camera views..

- Detection sample at night condition taken from Observation no. 24 (UTC Time: 30 June 2025, 19:46).
- Orca AI detected two nearby targets. First, a small vessel on the port side with a small radar signature and a large vessel (SOPHIA I – IMO 9546007) on the starboard side.
- Distance readings from the Orca AI system were periodically cross-checked with radar during the evaluation and showed good alignment, supporting confidence in the system's range estimation.

Results – Precision and Recall

Detection Performance Distribution



The number of relevant targets during the evaluation is 739

Precision	Recall
94.0%	98.6%

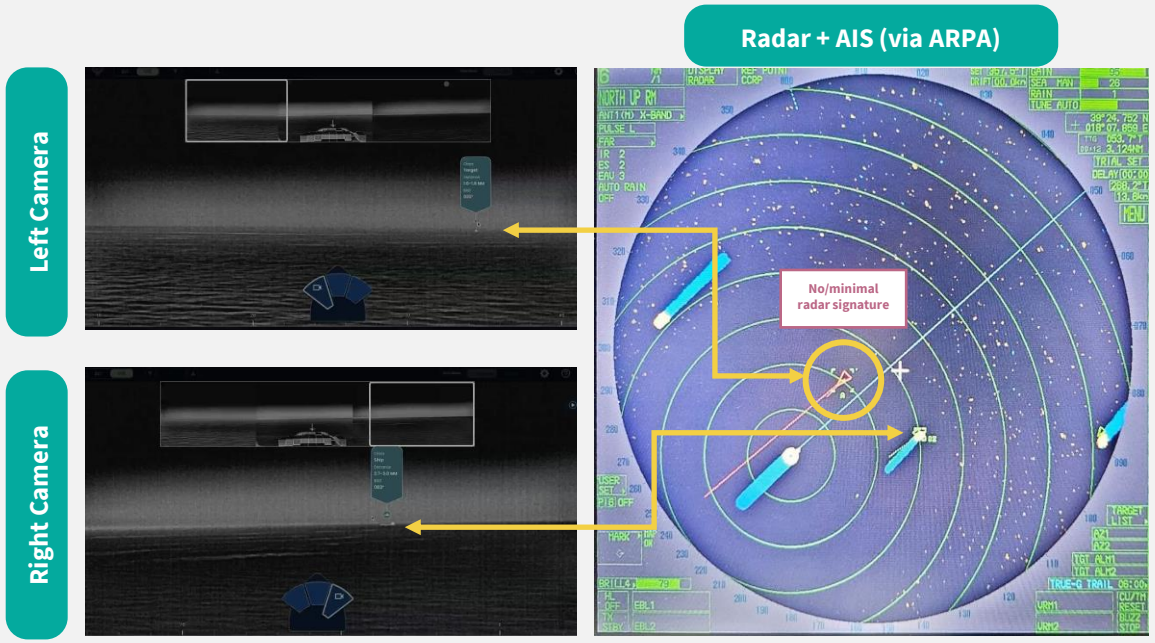
- The evaluation of Orca AI's Computer Vision system showed high Precision and Recall scores.
- The Recall score indicates that the system detected most relevant maritime objects during the evaluation, with an estimated missing rate of 1.4%.
- False Positives (FP) were observed primarily in two categories:
  - Duplicate detections of the same object.
  - Detections affected by visual conditions such as glare or reflections.

  Both types were included in the total FP count. Duplicate detections, which comprised approximately 73% of all FPs during the voyage, reflect repeated identification of existing targets rather than detections of non-existent objects.
- No system downtime was recorded throughout the evaluation period.

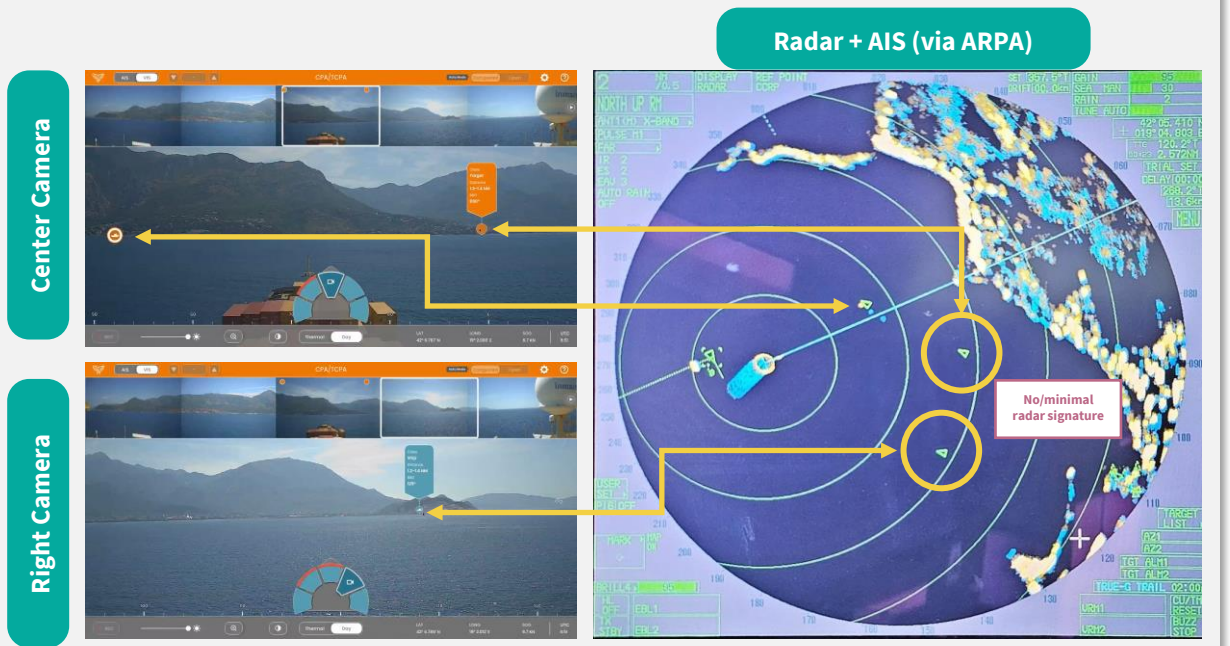
Results – Missed Radar Detections

- During the evaluation, several instances were observed in which the Orca AI system successfully detected close-range or low-signature targets that were not visible on the vessel's radar at the time.
- These detections typically occurred in visually challenging scenarios such as Low Radar Cross-Section objects or near clutter.
- Observation no. 24 illustrates a case where the Orca AI system detected a small wooden boat that was not reflected on radar.
- Observation no. 40 illustrates a case where the Orca AI system detected two small vessels that was not reflected on radar.
- These cases show the value of complementary visual sensing alongside traditional radar-based navigation.

Observation no. 24 (UTC Time: 30 June 2025, 19:46)



Observation no. 40 (UTC Time: 1 July 2025, 8:51)

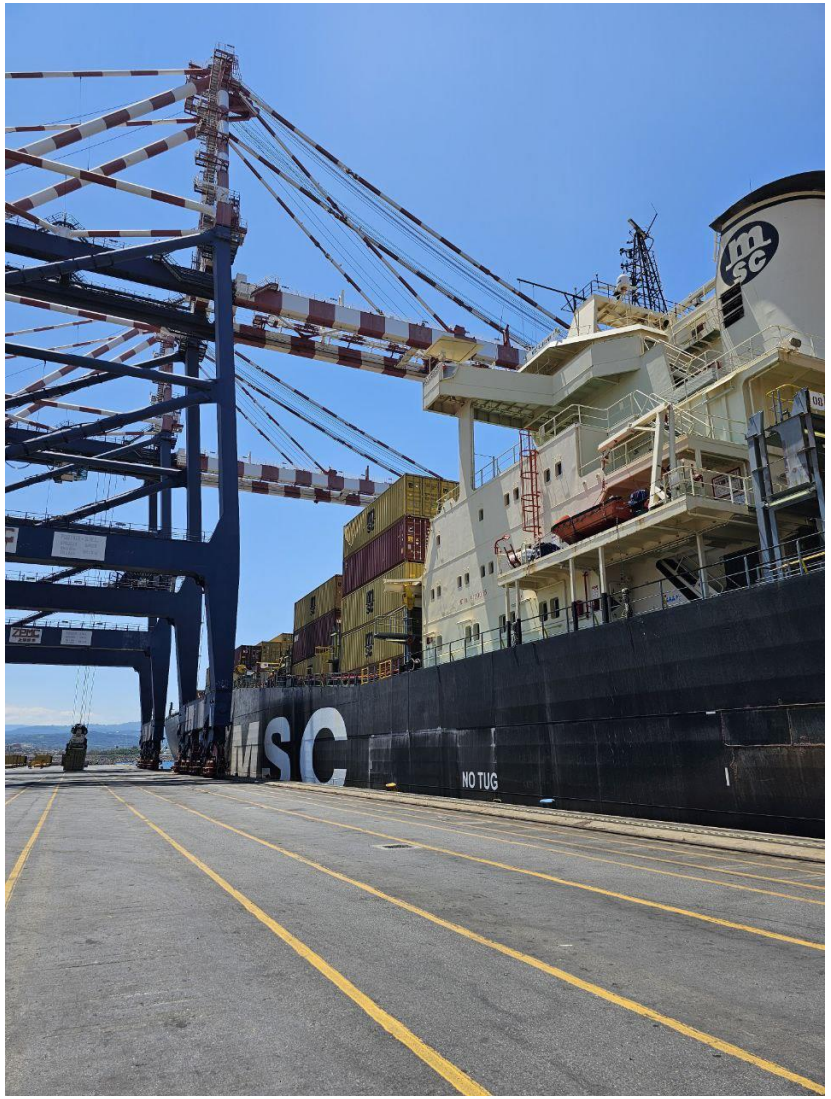


Note: The triangle symbol displayed on the ARPA represents an AIS target transmitted to the system and does not originate from radar detection.

Situational Awareness Impact – General Crew Interaction

Although this report focuses on technical performance, general observations noted during the evaluation are summarised below:

- Throughout the evaluation and across all operating conditions, including congested port areas, the Strait of Messina, the anchorage near Marsaxlokk, and open sea passages, **the crew consistently demonstrated strong situational awareness**.
- Overall, **the crew demonstrated an effective use of the system and showed confidence in the computer vision performance**. While adverse weather conditions were not encountered during the evaluation period, the crew noted that the system would be particularly valuable in such scenarios. They also **highlighted the usefulness of the thermal camera at night, especially for detecting small or non-AIS vessels that may not display lights** for visual identification.
- Notably, the Officer of the Watch (OOW) on MSC Adriana II reported that **a small, unlit wooden boat, which had drifted from its mooring at night and was not visible on radar, was first detected by the Orca AI system**. Under challenging visibility and sea conditions, **the initial alert was generated by Orca AI's** thermal and computer vision systems. This prompted the bridge team to initiate an early course adjustment and report the contact to local authorities, using the precise location (Lat/Long) provided by the Orca AI interface.
- During the evaluation, the crew appeared to have no difficulty navigating the system's user interface, likely aided by its **intuitive design** and possibly by prior training provided to them.
- A slight offset was noted between the user's touch point and the on-screen cursor. This behaviour appears to stem from the Marine Grade Rugged Touchscreen hardware and **does not reflect any limitation in the accuracy or performance of the Orca AI computer vision system.**



Thank you

Llyod’s Register Pte. Ltd