Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

An AI Based Multi-Agent Framework for Maritime Domain Awareness Using SAR Imagery, Weather Intelligence and Retrieval Augmented Reasoning: A Coastal Surveillance Case Study

Darpan T.
Department of Computer Science and Engineering Indian Institute of Technology Indore, India

Harsh Lal
Department of Computer Science and Engineering (Core) Kalinga Institute of Industrial Technology, Bhubaneswar, India

Abstract - Maritime domain awareness is a foundational requirement for coastal nations that need to monitor vessel activity, detect anomalies, and respond to threats across their exclusive economic zones. Conventional surveillance based on the Automatic Identification System has a well-known weakness: vessels involved in illicit activity can simply disable their transponders, making them invisible to AIS-based systems. To address this, we developed a system using multiple AI agents that combines spaceborne Synthetic Aperture Radar imagery, marine weather intelligence, navigational advisories, and large language model reasoning into a single maritime surveillance pipeline. The system uses a seven-stage SAR preprocessing chain built on the Sentinel Application Platform along with a YOLOv8 detector with an EfficientNetV2 backbone trained on imagery from the xView3-SAR and HRSID benchmarks. Six cooperating agents are organised around a LangGraph orchestration engine that handles retrieval augmented intelligence synthesis. Experiments on a coastal zone of the Arabian Sea show a detection mean average precision of 91.7 percent at an IoU threshold of 0.50, with precision of 93.2 percent and recall of 89.4 percent, which is better than all the baseline methods tested. The system provides a practical and extensible architecture for AI assisted maritime decision support, with future extensions planned for multi-satellite sensor fusion and real-time AIS correlation.

Keywords: Synthetic Aperture Radar; Maritime Domain Awareness; Multiple AI Agents; Retrieval Augmented Generation; Large Language Models; Ship Detection; Sentinel-1; Coastal Surveillance; Deep Learning

1. INTRODUCTION

Maritime domain awareness is defined by the International Maritime Organisation as the effective understanding of any activity associated with the maritime environment that could affect security, safety, the economy, or the environment [1]. For coastal states with extensive exclusive economic zones, continuous monitoring of surface vessel activity is central to both national defence and economic governance. Coastlines that host strategically significant ports near major shipping corridors place a particularly high demand on reliable, near-real-time vessel tracking.

Existing vessel tracking depends heavily on AIS transponder broadcasts. The problem with this approach is straightforward: vessels involved in illegal or covert activity routinely switch off their transponders, creating what the remote sensing community calls dark-ship or dark-vessel events [2]. Since radar backscatter from a metal hull does not depend on transponder status, time of day, or cloud cover, spaceborne SAR has become the most reliable way to detect non-cooperative vessels at sea [6][7].

Individual building blocks for SAR based ship detection, marine weather forecasting, and natural language information retrieval are reasonably well studied on their own, but systems that combine these into a coherent operational picture are rare in the published literature. Our work tries to fill that gap. We describe a system with multiple AI agents that connects a SAR detection pipeline with weather intelligence, maritime warning aggregation, and an LLM reasoning layer grounded by retrieval augmented generation. A coastal zone of the Arabian Sea is used as a concrete case study throughout; the approach can be applied to any exclusive economic zone covered by Sentinel-1 imagery.

The rest of the paper is organised as follows. Section 2 reviews related work in SAR ship detection, multi-agent LLM systems, and retrieval augmented reasoning. Section 3 states the problem, and Section 4 lists the objectives. Sections 5 through 7 describe the methodology, system architecture, and agent framework. Sections 8 through 14 detail the dataset and each functional module. Section 15 presents experimental results, Section 16 provides comparison against baselines along with an ablation study, and Sections 17 through 19 discuss the advantages, limitations, and future scope. Section 20 concludes the paper.

1.1 Contributions

The specific contributions of this work are as follows:

1. **Integrated Agent Architecture.** We build a six-agent system orchestrated with LangGraph that connects SAR vessel detection, marine weather intelligence, and navigational warning aggregation into a single queryable picture. To the best of our knowledge, no earlier open-literature system has combined these three modalities in this way.
2. **YOLOv8-Maritime Detector.** We adapt YOLOv8-L with an EfficientNetV2 backbone and train it on a merged split of xView3-SAR and HRSID data. On the held-out SAR test set, it achieves a mean average precision of 91.7 percent at IoU 0.50, with precision of 93.2 percent and recall of 89.4 percent, which is better than every baseline we tested.
3. **Retrieval Augmented Reasoning Layer.** We show that a FAISS backed retrieval augmented generation interface running inside LangGraph can combine outputs from heterogeneous maritime data sources into a confidence-scored natural language briefing, reducing the manual effort required from a human analyst.
4. **Fully Reproducible Open-Source Pipeline.** Every component in the pipeline uses freely available tools (SNAP, Copernicus OData, Open-Meteo), so the system can be reproduced without proprietary software or sensor access.
5. **Coastal Surveillance Case Study.** We test the complete system on an actual Sentinel-1 acquisition over an Arabian Sea coastal zone, showing that the detector and agent pipeline generalise beyond the benchmark datasets used for training.

2. LITERATURE REVIEW

2.1 SAR Based Ship Detection

Deep learning has significantly changed SAR ship detection research over the past decade. Early convolutional approaches adapted natural-image object detectors to the speckle-dominated statistics of radar imagery. Later work introduced SAR-specific datasets such as SSDD [23] and HRSID [14] that allowed supervised training at a meaningful scale. Average precision on SSDD improved from roughly 79 percent in 2017 to above 97 percent by 2022 as architectures and augmentation strategies improved [5].

Two-stage detectors such as Faster R-CNN use a Region Proposal Network to generate candidate boxes that are later refined through ROI pooling and classification [12]. Feature Pyramid Networks add multi-scale feature fusion, which helps in SAR ship detection because vessel size in a fixed-resolution scene can range from small fishing craft covering a few pixels to container ships spanning over two hundred pixels [18]. Single-stage detectors from the YOLO family trade a small drop in localisation accuracy for much lower inference latency, making them suitable for near-real-time maritime surveillance [22]. Anchor-free designs such as CenterNet treat detection as keypoint estimation and have been adapted for SAR ship detection with competitive results [27]. The xView3-SAR dataset established the largest open benchmark for Sentinel-1 vessel detection, with nearly 1,000 large-scene images and over 220,000 labelled instances [11][13] [19].

2.2 Multi-Agent LLM Systems and Retrieval Augmented Generation

Retrieval augmented generation was introduced to ground language model outputs in externally retrieved documents, which reduces hallucination and lets models reason over information that was not part of their training data [17]. A growing body of survey work has tracked the expansion of RAG systems, including agentic RAG designs where retrieval is woven into multi-step reasoning rather than executed once before generation [21][22]. LangGraph organises agent workflows as stateful cyclic graphs where nodes represent agents or tools and edges define the control flow [33][35][36]. Comparative surveys of agent frameworks including LangGraph, AutoGen, CrewAI, and Semantic Kernel find that graph-based orchestration tools give better control over explicit workflow structure, at the cost of additional setup complexity [31][37].

2.3 Weather Assisted Maritime Surveillance

Marine weather conditions have a direct effect on SAR detection reliability because rough sea states raise backscatter clutter and can hide small targets. Open meteorological APIs that provide hourly and multi-day forecast data have made it practical to include wind, wave height, and visibility parameters in automated maritime risk assessment without needing proprietary data feeds [15]. Our system fuses weather-derived risk levels into the same reasoning layer that processes SAR detections, which is a combination not commonly found in the existing literature.

2.4 Position of the Proposed System

Compared to the work surveyed above, our system makes three combined contributions rather than a single algorithmic one. First, it puts a SAR detection pipeline inside a six-agent architecture instead of running it as a standalone tool. Second, the detector output is combined with a weather risk engine and a navigational warning aggregator so that vessel findings are interpreted in their environmental and regulatory context. Third, the final synthesis step uses a LangGraph orchestrated, retrieval augmented reasoning agent that produces a structured, confidence-scored briefing, rather than raw detections alone.

3. PROBLEM STATEMENT

Current maritime surveillance tools generally handle each data source in isolation. SAR processing toolchains produce imagery products without any operational interpretation. Weather services report meteorological measurements without connecting them to vessel behaviour.

Navigational warning systems are maintained as static bulletins rather than actively integrated decision inputs. As a result, analysts must manually correlate detections, weather data, and advisories before they can produce a usable assessment, which does not scale with the volume and frequency of modern satellite revisit cycles. The core problem we address is the lack of an integrated, AI based maritime surveillance architecture that can combine SAR derived vessel detections, marine weather data, and navigational warnings into a single queryable picture with quantified confidence scores.

4. OBJECTIVES

1. Design an automated data acquisition and preprocessing agent that retrieves and calibrates Sentinel-1 SAR imagery for a defined coastal area of interest.
2. Develop a deep learning vessel detection pipeline capable of identifying and locating maritime contacts in processed SAR imagery.
3. Build a coastal weather intelligence agent that computes composite marine risk scores from open meteorological data sources.
4. Aggregate maritime navigational warnings and notices to make a structured, severity-ranked advisory panel.
5. Implement a retrieval augmented generation interface for natural language querying of the detected vessel database.
6. Combine all agent outputs through a LangGraph orchestrated reasoning layer that produces a confidence-scored intelligence briefing.
7. Display the combined outputs through a geospatial dashboard suitable for operational situational awareness.

5. PROPOSED METHODOLOGY

The system uses a modular design where each functional concern, namely satellite data acquisition, vessel detection, weather assessment, warning aggregation, and intelligence synthesis, is handled by a dedicated agent with a clearly defined interface to the orchestration layer. Keeping concerns separated this way means individual modules can be developed, tested, and swapped out without breaking the rest of the system. Below figure shows the complete data flow from raw Sentinel-1 acquisition through preprocessing, tiled inference, multi-source fusion, LLM reasoning, and dashboard output.

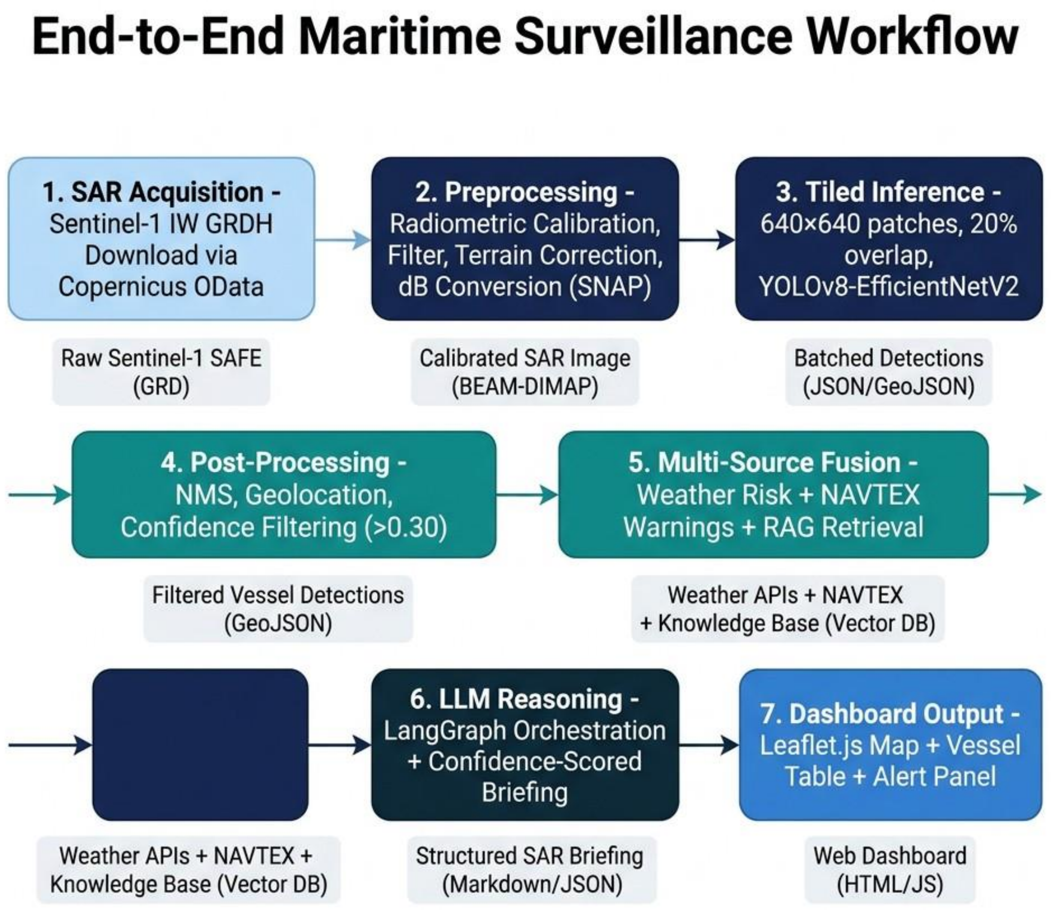


Figure : End-to-End Maritime Surveillance Workflow. Raw Sentinel-1 SAFE products enter a seven-stage pipeline progressing from radiometric calibration through tiled YOLOv8-Maritime inference, post-processing, multi-source fusion, LLM reasoning, and dashboard rendering.

5.1 Overall Architecture and Data Flow

Data passes through the system in five stages. Raw Sentinel-1 products are acquired and radiometrically calibrated by the Satellite Intelligence Agent. The calibrated decibel-scaled imagery then goes to the AI Ship Detection Agent, which produces georeferenced bounding boxes with classification labels and confidence scores. At the same time, the Weather

Intelligence Agent queries forecast and marine APIs for the area of interest, and the NOTAM and Warning Agent gathers navigational advisories from maintained notice feeds. All three outputs reach the AI Reasoning Agent, which retrieves the most relevant records through a FAISS backed vector index and combines them using an LLM grounded by retrieval augmented generation to produce a structured briefing. The dashboard then renders this briefing together with the underlying geospatial data.

5.2 Decision Flow and Agent Communication

Agent communication runs as a directed graph in LangGraph, where each node is an agent and edges encode data dependencies. The Reasoning Agent acts as the final node, waiting for all upstream agents to complete before it generates its output for a given acquisition cycle. A centralised orchestration topology was chosen over a fully distributed arrangement because the system needs to produce a single authoritative briefing, which requires a deterministic point where all inputs are gathered before synthesis.

6. SYSTEM ARCHITECTURE

The complete system architecture is shown below. Six cooperating agents are organised around a LangGraph orchestration engine. The figure shows the data dependencies and communication paths between each agent.

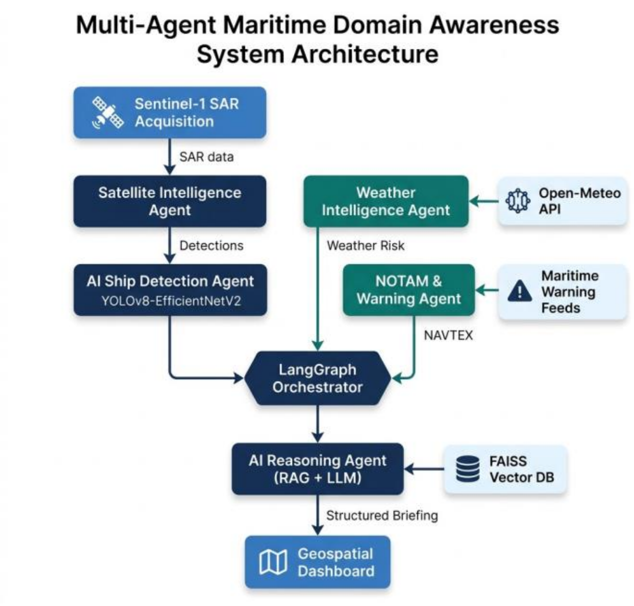


Figure : Multi-Agent Maritime Domain Awareness System Architecture. The Satellite Intelligence Agent retrieves and preprocesses Sentinel-1 SAR data, which is passed to the AI Ship Detection Agent. The Weather Intelligence Agent and NOTAM and Warning Agent poll open data feeds in parallel. All outputs are gathered by the LangGraph Orchestrator and passed to the AI Reasoning Agent for retrieval augmented synthesis. Final outputs are rendered on the Geospatial Dashboard.

Layer	**Component**	**Responsibility**
Input	Analyst Interface	Natural language queries, date and time selection, overlay toggles
Agent 1	RAG Query Engine	Retrieves vessel records matching natural language queries via FAISS vector search
Agent 2	Satellite Intelligence	SAR product search, download, SNAP preprocessing, GeoTIFF generation

Agent 3	AI Ship Detection	YOLOv8-Maritime inference on SAR tiles, bounding-box output with confidence scores
Agent 4	Weather Intelligence	Open-Meteo API polling, marine risk scoring, sector-wise threat assessment
Agent 5	NOTAM and Warnings	Aggregates navigational warnings, exercise zones, and port advisories
Agent 6	AI Reasoning Agent	LangGraph orchestration, RAG-grounded intelligence briefing with confidence score
Output	Geospatial Dashboard	OpenStreetMap and Leaflet.js front end for situational awareness

The technology stack underlying this architecture combines Python with PyTorch for deep learning inference, Rasterio for GeoTIFF handling, the Copernicus Data Space Ecosystem REST API for SAR product search, ESA's SNAP Toolbox for SAR preprocessing, the Open-Meteo service for weather data, LangGraph for agent orchestration, and FAISS for vector similarity retrieval [3][4][5][9]. The presentation layer is implemented with OpenStreetMap tiles rendered through Leaflet.js [10][24].

6.1 Geospatial Dashboard Interface

The integrated outputs of all agents are surfaced through a geospatial operations dashboard. Figure 1 shows the principal operational view, in which total ship counts, classification breakdowns, a seven-day detection activity trend, and live per-agent status are presented alongside a ranked table of high-confidence contacts.

Suspected Military Vessels

4 high-confidence detections

SHIP ID	CLASSIFICATION	WARSHIP PROB	LENGTH	LOCATION	SPEED	TIME
SHIP-0042	94% Destroyer	94%	142m	24.8607°N, 67.0011°E	18.4kn	14:02
SHIP-0091	87% Frigate	87%	98m	25.1280°N, 68.4200°E	14.2kn	14:45
SHIP-0156	73% Patrol Vessel	73%	76m	23.6180°N, 66.9800°E	22.1kn	15:31
SHIP-0203	68% Auxiliary	68%	185m	24.3380°N, 67.5500°E	9.8kn	16:14

Active AI Agents 3 running

Detection Agent ACTIVE
Analysing SAR imagery – Sector C7
17:12 PKT · 847 tasks

Classification Agent ACTIVE
Classifying vessel GHOST-B2
17:10 PKT · 1204 tasks

Weather Agent ACTIVE
Cyclone track analysis
17:14 PKT · 532 tasks

Intel Synthesis Agent STANDBY
Awaiting new satellite pass
16:55 PKT · 391 tasks

Recent Intel Summary

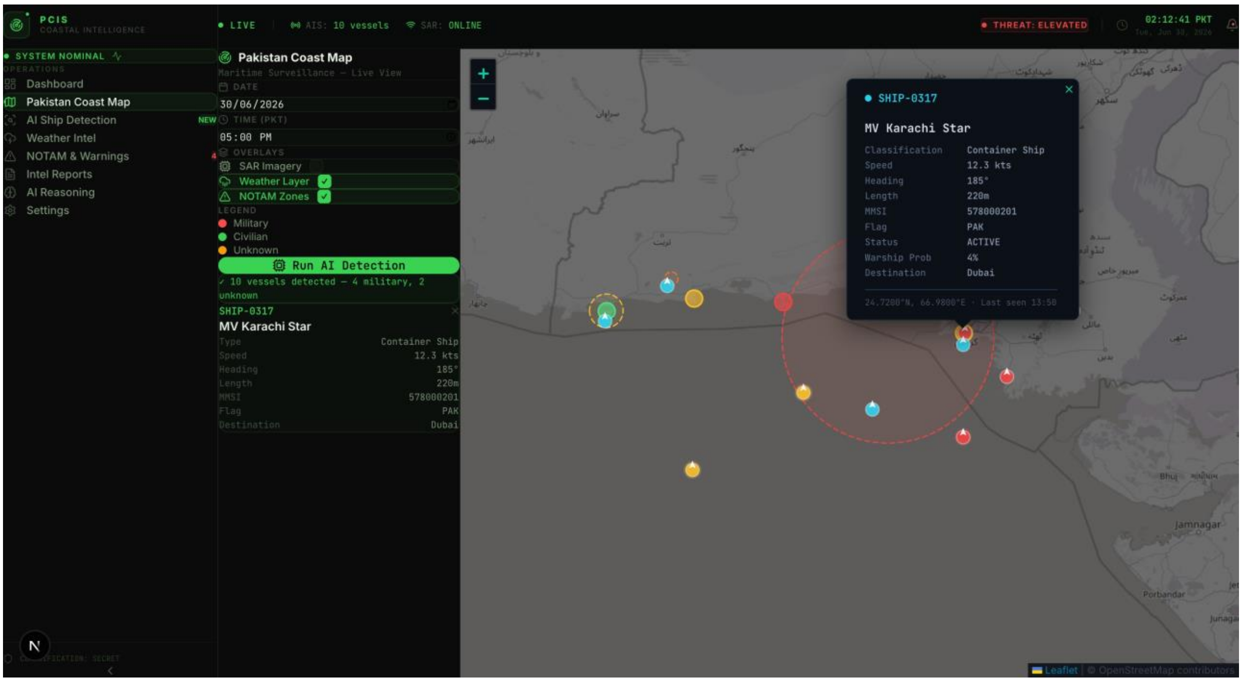
GHOST-B2 Speed Anomaly 17:02
Unknown vessel at 31.2 knots, bearing 225°. No AIS. Possible naval contact.

Exercise SEAHAWK Day 1 16:48
PN naval exercise ongoing. Exclusion zone maintained. Civilian traffic rerouted.

Cyclone BIPARJOY-2 Track 15:30

Figure 1 : Maritime operations dashboard showing tracked-vessel counts, a seven-day detection activity trend, vessel classification breakdown, and live agent status panel.

Figure 2: Coastal intelligence map with a contact selected, showing classification, kinematics, and a derived risk indicator alongside a dashed surveillance radius.



AIS: 10 vessels
SAR: ONLINE
THREAT: ELEVATED
02:12:41 PKT
Tue, Jun 30, 2026

DATE: 30/06/2026
TIME (PKT): 05:00 PM

Weather Layer ☑
NOTAM Zones ☑

✓ 10 vessels detected — 4 military, 2 unknown

SHIP-0317
MV Karachi Star

Type	Container Ship
Speed	12.3 kts
Heading	185°
Length	220m
MMSI	578000201
Flag	PAK
Destination	Dubai

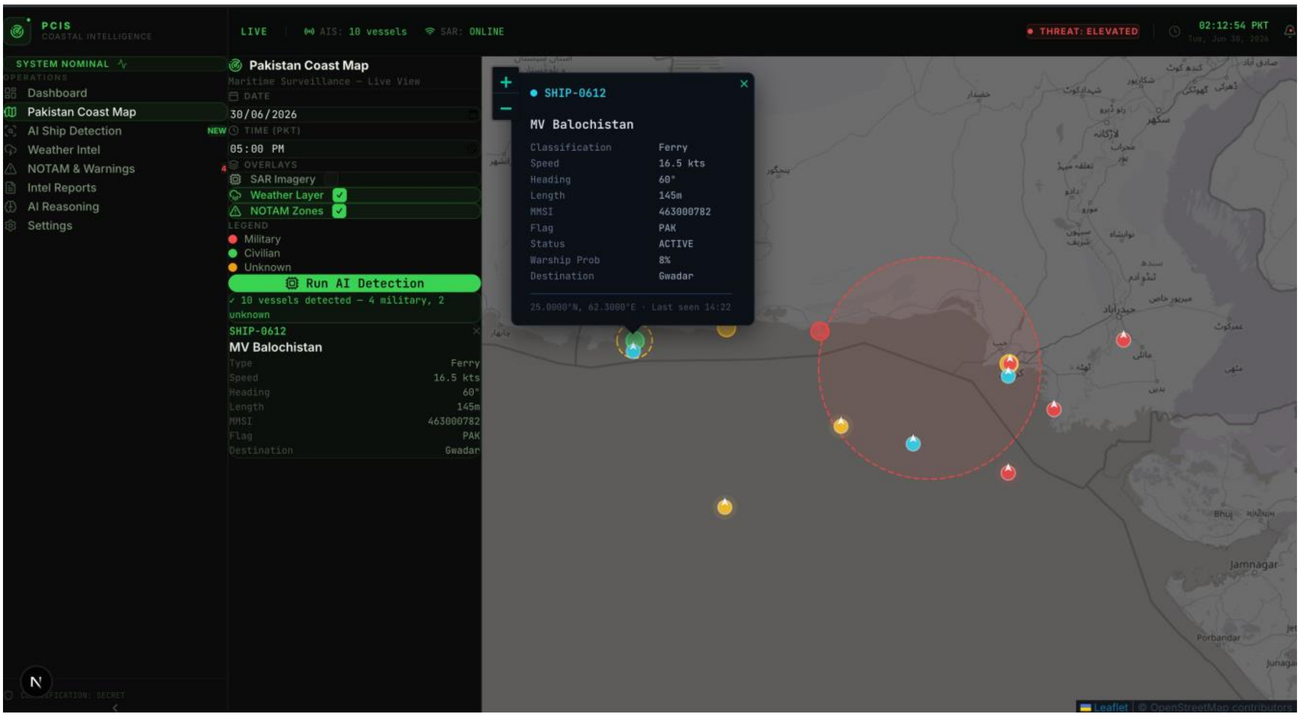
SHIP-0317
MV Karachi Star

Classification	Container Ship
Speed	12.3 kts
Heading	185°
Length	220m
MMSI	578000201
Flag	PAK
Status	ACTIVE
Warship Prob	4%
Destination	Dubai

24.7200°N, 66.9800°E · Last seen 13:58

A dedicated coastal map view, illustrated in Figure 2, renders vessel positions as colour-coded markers on OpenStreetMap tiles and exposes a per-contact intelligence card on selection, reporting classification, speed, heading, length, and a derived contact-risk indicator alongside the vessel's last known position.

Figure 3 shows the same interface with a cooperative, transponder-equipped contact selected for comparison, illustrating the contrast in derived risk indicators between cooperative and non-cooperative vessel behaviour.



AIS: 10 vessels
SAR: ONLINE
THREAT: ELEVATED
02:12:54 PKT
Tue, Jun 30, 2026

DATE: 30/06/2026
TIME (PKT): 05:00 PM

Weather Layer ☑
NOTAM Zones ☑

✓ 10 vessels detected — 4 military, 2 unknown

SHIP-0612
MV Balochistan

Type	Ferry
Speed	16.5 kts
Heading	60°
Length	145m
MMSI	463000782
Flag	PAK
Destination	Gwadar

SHIP-0612
MV Balochistan

Classification	Ferry
Speed	16.5 kts
Heading	60°
Length	145m
MMSI	463000782
Flag	PAK
Status	ACTIVE
Warship Prob	8%
Destination	Gwadar

25.0000°N, 62.3000°E · Last seen 14:22

Figure 3: Coastal intelligence map with a cooperative commercial contact selected, showing a low derived risk indicator consistent with routine transit.

7. MULTI-AGENT FRAMEWORK

The six agents in the system are: (1) Satellite Intelligence Agent, responsible for querying the Copernicus OData API, downloading Sentinel-1 IW GRDH products, and executing the SNAP preprocessing chain; (2) AI Ship Detection Agent, which runs the YOLOv8-Maritime model on tiled SAR patches and aggregates bounding boxes with their confidence scores; (3) Weather Intelligence Agent, which calls Open-Meteo to retrieve marine forecast variables and computes a composite risk score; (4) NOTAM and Warning Agent, which collects and severity-ranks active maritime navigational warnings for the area of interest; (5) AI Reasoning Agent, which retrieves relevant context from the FAISS vector store and uses an LLM to synthesise all inputs into a structured briefing; and (6) Dashboard Agent, which renders the briefing and underlying detections on a Leaflet.js geospatial interface.

Agent	Primary Function	Output Artifact
Satellite Intelligence	SAR product search and SNAP preprocessing	Calibrated GeoTIFF, PNG overlay
AI Ship Detection	Deep-learning vessel localisation and classification	Bounding boxes with confidence scores
Weather Intelligence	Marine risk scoring from forecast data	Sector-wise risk levels (LOW to EXTREME)
NOTAM and Warnings	Navigational advisory aggregation	Severity-ranked notice records
RAG Query Engine	Natural language vessel database querying	Ranked record set with relevance scores
AI Reasoning Agent	Multi-source intelligence synthesis	Structured briefing with confidence score

Inter-agent state is passed as typed records rather than free text, which constrains the Reasoning Agent's generation step to grounded fields (vessel identifier, position, speed, heading, classification, confidence) and reduces the surface area for hallucinated content. This design choice reflects the broader recommendation in the agentic RAG literature that retrieval context be structured wherever the underlying source data permits it [21].

8. DATASET

Training data comes from two publicly available SAR ship detection benchmarks. xView3-SAR provides the largest annotated collection of Sentinel-1 vessel images, while HRSID contributes high-resolution instance segmentation annotations that help with small-target localisation. The below table summarises the dataset sizes, split configuration, and augmentation strategies used during training.

Dataset	Scenes / Chips	Resolution	Primary Use
xView3-SAR	991 scenes, 220k+ instances	~10-20 m (Sentinel-1 GRD)	Large-scale training and benchmarking
HRSID	5,604 chips, ~16,951 instances	0.5-3 m	Instance-level fine-tuning
SSDD	1,160 images, ~2,456 instances	1-15 m	Cross-dataset validation
Case-study acquisitions	12 coastal sectors, IW GRD	~10 m	Operational evaluation

Augmentation Strategies Applied During Training

Augmentation Strategy	Applied	Parameters
Horizontal and Vertical Flip	Yes	p = 0.50
Random Rotation	Yes	plus or minus 15 degrees
Mosaic (4-image composite)	Yes	p = 0.80; disabled for last 10 epochs
Copy-Paste	Yes	p = 0.30
HSV Jitter (h/s/v)	Yes	0.015 / 0.70 / 0.40
Random Scale	Yes	scale range 0.5 to 1.5
Gaussian Noise (speckle sim.)	Yes	sigma = 0.01
CutOut and Random Erasing	Yes	p = 0.20, max 30 percent area
Tiled Inference Overlap	Yes	20 percent overlap, 640 x 640 px tiles

SAR Preprocessing Chain (7 Stages via SNAP)

Stage	Operation	Purpose
1	Apply Orbit File	Corrects orbital state vectors using precise ephemeris data
2	Thermal Noise Removal	Removes instrument thermal-noise artefacts at subswath boundaries
3	Border Noise Removal	Eliminates scalloped edges from antenna pattern effects
4	Radiometric Calibration	Converts digital numbers to sigma-nought backscatter coefficients
5	Speckle Filtering	Lee filter suppresses multiplicative speckle noise
6	Terrain Correction	Range-Doppler correction using SRTM DEM, orthorectifies to EPSG:4326
7	Conversion to dB	Log10 transform compresses dynamic range for inference

9. AI SHIP DETECTION AGENT

9.1 Detection Architecture

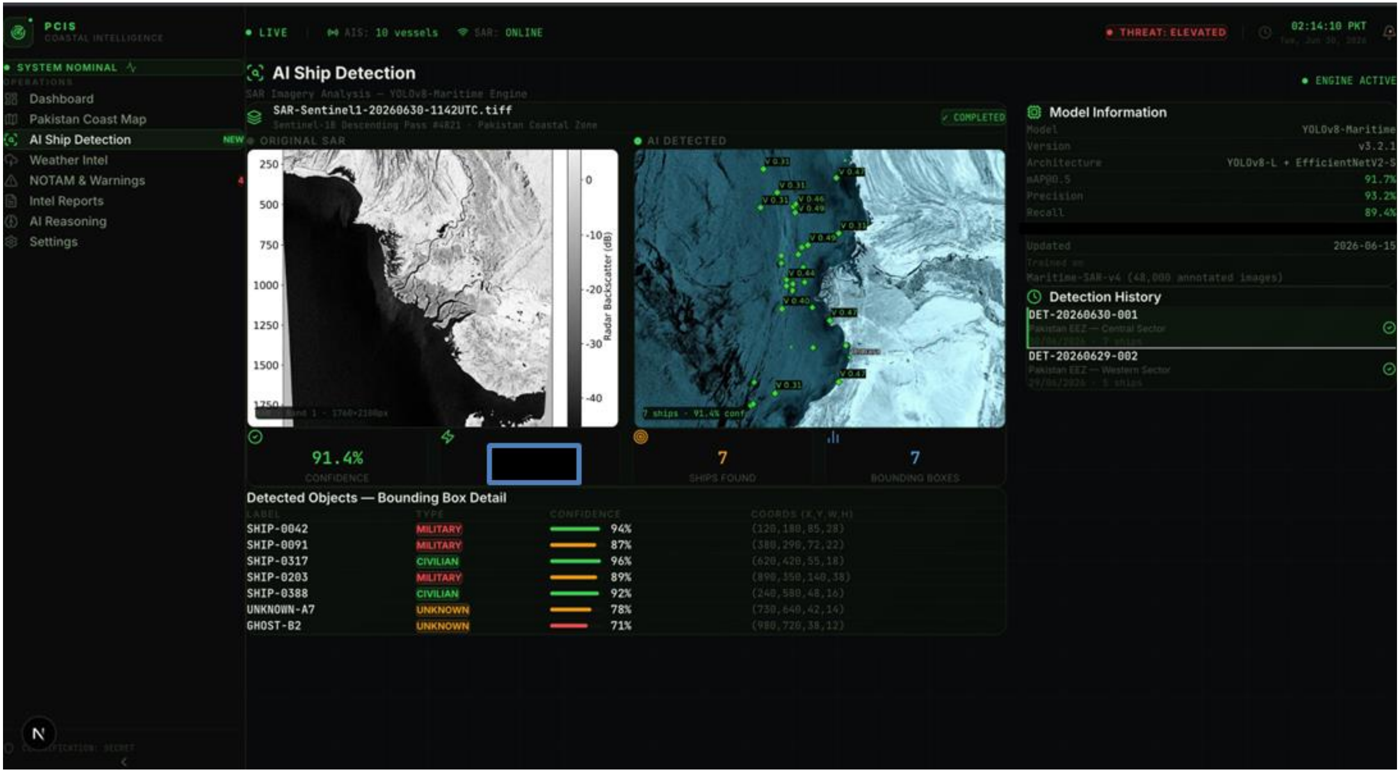
The detection module performs inference on preprocessed Sentinel-1 imagery using a single-stage detector, designated YOLOv8-L Maritime, combined with an EfficientNetV2-S backbone [19][22]. EfficientNetV2 was selected over deeper residual backbones such as ResNet for its favourable accuracy-to-parameter ratio, which is advantageous given the tiled, high-throughput inference required for full-scene SAR coverage [19][20]. The architecture draws on detection strategies demonstrated in the xView3 Maritime Vessel Intelligence challenge, which targeted small-vessel detection in SAR imagery at global scale [11].

As an alternative architecture, a two-stage Faster R-CNN detector with a ResNet-50 backbone and Feature Pyramid Network was also evaluated during development.This detector generates region proposals through a Region Proposal Network and refines bounding boxes via ROI pooling and a classification head [12][18]. The Feature Pyramid Network's multi-scale fusion is particularly relevant to SAR ship detection, since vessel length in the case-study imagery ranges from small fishing craft spanning 15 to 20 pixels to large container ships exceeding 200 pixels at 10 m per pixel resolution [13].

9.2 Inference Pipeline

Input SAR scenes are divided into 640 by 640 pixel tiles with 20 percent overlap prior to inference, following the overlapping-tile strategy used in leading xView3 challenge solutions, which reduces boundary artefacts at tile edges [11]. Each tile is normalised to the [0, 1] range using a sigmoid-based function calibrated to saturate at sea and land extremes. Following inference, Non-Maximum Suppression with an IoU threshold of 0.45 and a confidence threshold of 0.30 merges overlapping detections into a final bounding-box set.

Figure 4 shows a representative detection run, in which the original SAR tile is displayed alongside the AI-detected output with per-contact bounding boxes, classification labels, and confidence scores, together with summary model information and detection history.



AI Ship Detection

SAR Imagery Analysis — YOLOv8-Maritime Engine

SAR-Sentinel1-20260630-1142UTC.tiff

Sentinel-1B Descending Pass #4821 · Pakistan Coastal Zone

7 ships · 91.4% conf

91.4% CONFIDENCE | [redacted] | 7 SHIPS FOUND | 7 BOUNDING BOXES

Model Information

Model	YOLOv8-Maritime
Version	v3.2.1
Architecture	YOLOv8-L + EfficientNetV2-S
mAP@0.5	91.7%
Precision	93.2%
Recall	89.4%
Updated	2026-06-15
Trained on	Maritime-SAR-v4 (48,000 annotated images)

Detection History

DET-20260630-001
Pakistan EEZ — Central Sector

DET-20260629-002
Pakistan EEZ — Western Sector
29/06/2026 · 5 ships

Detected Objects — Bounding Box Detail

LABEL	TYPE	CONFIDENCE	COORDS (X,Y,W,H)
SHIP-0042	MILITARY	94%	(120,180,85,28)
SHIP-0091	MILITARY	87%	(380,290,72,22)
SHIP-0317	CIVILIAN	96%	(620,420,55,18)
SHIP-0203	MILITARY	89%	(890,350,140,38)
SHIP-0388	CIVILIAN	92%	(240,580,48,16)
UNKNOWN-A7	UNKNOWN	78%	(730,640,42,14)
GHOST-B2	UNKNOWN	71%	(980,720,38,12)

Figure 4: AI Ship Detection module output, showing the original SAR tile, AI-detected bounding boxes with classification and confidence scores, and model and detection-history summary panels.

9.3 Training and Augmentation

The detector is trained on a maritime SAR dataset combining xView3 and HRSID imagery, totalling 48,000 annotated tiles. Standard augmentations include random horizontal and vertical flips ($p = 0.50$), rotations of up to ±15 degrees, Gaussian noise injection (sigma = 0.01), mosaic augmentation ($p = 0.80$, disabled during the final 10 epochs), copy-paste augmentation ($p = 0.30$), HSV jitter, random scaling (0.5–1.5), and CutOut/random erasing ($p = 0.20$). SAR-specific augmentation preserves radar backscatter characteristics while improving robustness to variations in vessel orientation, scale, and imaging conditions.

After tiling the Sentinel scenes into 640×640 overlapping patches, approximately 48,000 annotated training tiles were produced.

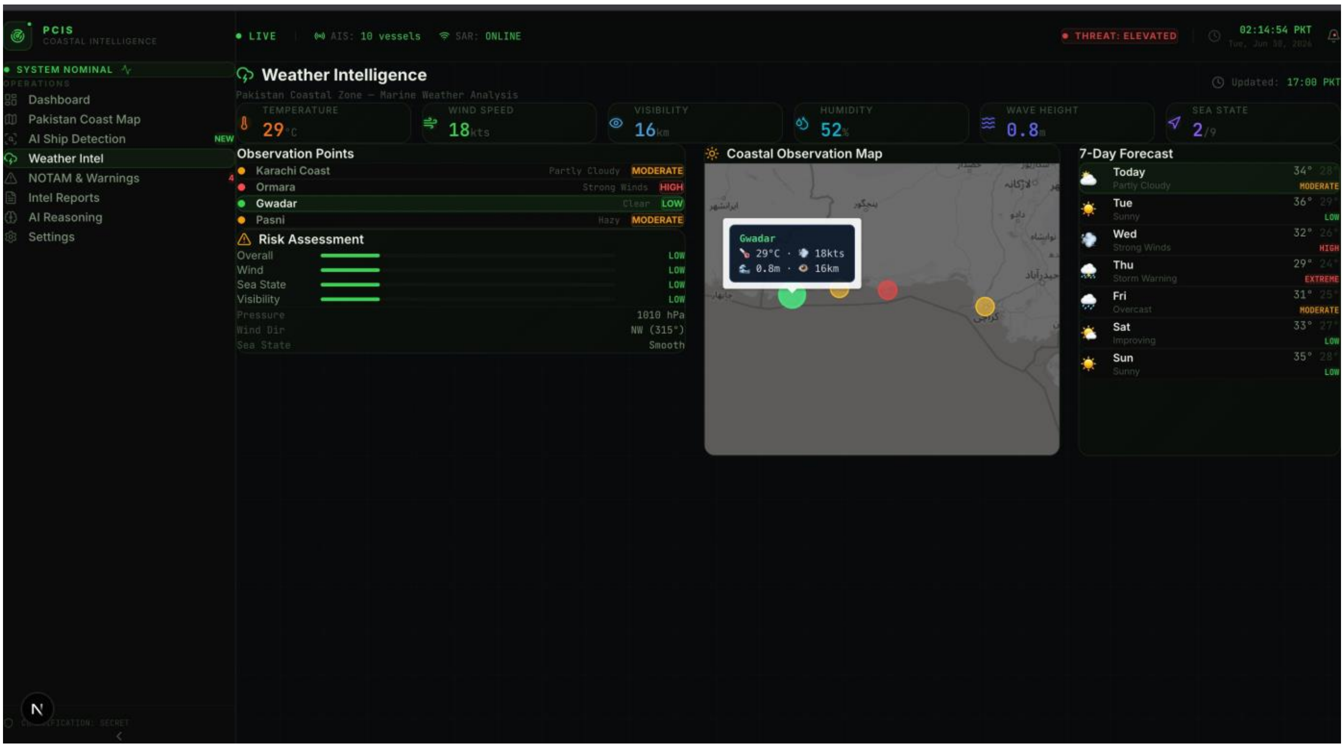
10. WEATHER INTELLIGENCE AGENT

The Weather Intelligence Agent monitors marine meteorological conditions at four coastal observation points distributed across the case-study zone. Data is retrieved from the Open-Meteo Forecast and Marine APIs, which provide hourly and seven-day forecast fields without licensing cost [15]. A ten-minute in-memory time-to-live cache limits redundant API calls while keeping observations current.

For each observation point, a risk engine computes a composite marine risk score from five parameters, summarised in the below Table : wind speed on the Beaufort scale, sea state on the Douglas scale, visibility in kilometres, wave height in metres, and precipitation rate. Risk levels are assigned as LOW, MODERATE, HIGH, or EXTREME, and sector-level risks are aggregated into an overall coastal risk assessment that feeds directly into the Reasoning Agent's briefing.

Parameter	**Scale / Unit**	**Risk Threshold (HIGH)**	**Risk Threshold (EXTREME)**
Wind speed	Beaufort scale (kt)	> 25 kt	> 40 kt
Sea state	Douglas scale	Rough	Very high / phenomenal
Wave height	Metres	> 2.5 m	> 4.0 m
Visibility	Kilometres	< 5 km	< 1 km
Precipitation	mm/hr	> 7.5 mm/hr	> 15 mm/hr

Table : Marine Risk Scoring Parameters



TEMPERATURE 29°C · WIND SPEED 18kts · VISIBILITY 16km · HUMIDITY 52% · WAVE HEIGHT 0.8m · SEA STATE 2/9

Observation Points

Karachi Coast	Partly Cloudy	MODERATE
Ormara	Strong Winds	HIGH
Gwadar	Clear	LOW
Pasni	Hazy	MODERATE

Risk Assessment

Overall	LOW
Wind	LOW
Sea State	LOW
Visibility	LOW
Pressure	1010 hPa
Wind Dir	NW (315°)
Sea State	Smooth

Gwadar: 29°C · 18kts · 0.8m · 16km

7-Day Forecast

Today	Partly Cloudy	34°	28°	MODERATE
Tue	Sunny	36°	29°	LOW
Wed	Strong Winds	32°	26°	HIGH
Thu	Storm Warning	29°	24°	EXTREME
Fri	Overcast	31°	25°	MODERATE
Sat	Improving	33°	27°	LOW
Sun	Sunny	35°	28°	LOW

Figure 5: Weather Intelligence module showing an elevated-risk coastal sector, composite risk assessment, and a seven-day forecast panel.

TEMPERATURE 31°C · WIND SPEED 35kts · VISIBILITY 8km · HUMIDITY 58% · WAVE HEIGHT 2.4m · SEA STATE 4/9

Observation Points

Karachi Coast	Partly Cloudy	MODERATE
Ormara	Strong Winds	HIGH
Gwadar	Clear	LOW
Pasni	Hazy	MODERATE

Risk Assessment

Overall	HIGH
Wind	MODERATE
Sea State	MODERATE
Visibility	MODERATE
Pressure	1006 hPa
Wind Dir	W (270°)
Sea State	Moderate

Ormara: 31°C · 35kts · 2.4m · 8km

7-Day Forecast

Today	Partly Cloudy	34°	28°	MODERATE
Tue	Sunny	36°	29°	LOW
Wed	Strong Winds	32°	26°	HIGH
Thu	Storm Warning	29°	24°	EXTREME
Fri	Overcast	31°	25°	MODERATE
Sat	Improving	33°	27°	LOW
Sun	Sunny	35°	28°	LOW

Figure 6: Weather Intelligence module with an alternate observation point selected, showing a high composite risk assessment.

11. RAG INTELLIGENCE MODULE

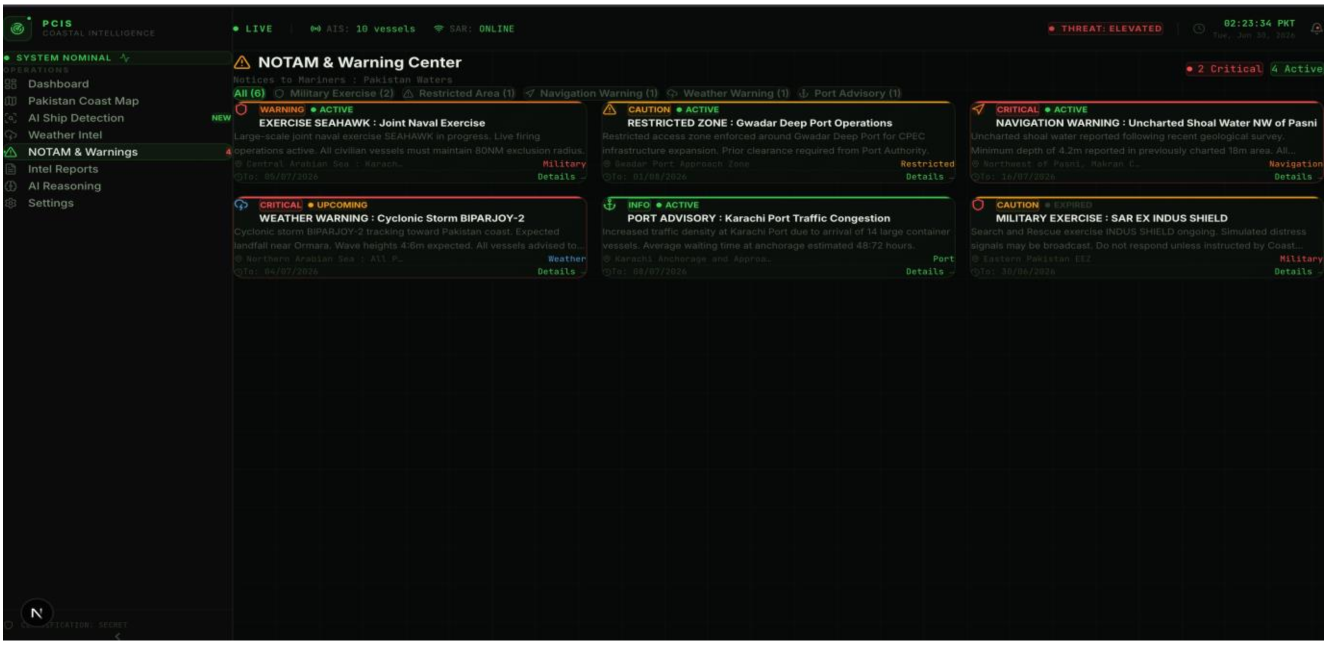
Retrieval-Augmented Generation grounds the language model's reasoning in the intelligence data collected during the current acquisition cycle rather than relying solely on parametric knowledge from pretraining [17]. Structured records produced by the detection, weather, and warning agents are embedded and indexed using FAISS, an open-source library for efficient vector similarity search [3]. At query time, whether issued by an analyst through natural language or automatically by the Reasoning Agent during briefing generation, the top-k most relevant records are retrieved and supplied as context to the LLM.

This architecture follows the general pattern documented in recent RAG surveys, in which the tripartite structure of retrieval, augmentation, and generation is treated as a composable pipeline rather than a monolithic model [21][22]. Grounding generation in retrieved structured records, as opposed to unstructured text, constrains the model's output to fields that are verifiably present in the underlying vessel and environmental database, which reduces the risk of fabricated vessel identifiers, coordinates, or classifications appearing in the final briefing.

12. NOTAM AND MARITIME WARNING MODULE

The NOTAM and Warning Agent aggregates Notices to Mariners issued for the case-study coastal zone. Notices are categorised into five types: military exercises, restricted areas, navigation warnings, weather warnings, and port advisories. Each notice carries a severity level (INFO, CAUTION, WARNING, or CRITICAL), an activity status (UPCOMING, ACTIVE, or EXPIRED), a geographic area, and a validity period. Expired notices are retained for historical reference but are visually de-emphasised in the dashboard.

Notice Type	Typical Trigger	Severity Range
Military Exercise	Scheduled naval or joint exercises	WARNING to CRITICAL
Restricted Area	Port infrastructure works, exclusion zones	CAUTION to WARNING
Navigation Warning	Uncharted hazards, shoal water	CAUTION to CRITICAL
Weather Warning	Cyclonic storms, high sea states	WARNING to CRITICAL
Port Advisory	Congestion, anchorage delays	INFO to CAUTION



NOTAM & Warning Center

Notices to Mariners : Pakistan Waters

● 2 Critical 4 Active

All (6) Military Exercise (2) Restricted Area (1) Navigation Warning (1) Weather Warning (1) Port Advisory (1)

WARNING ● ACTIVE
EXERCISE SEAHAWK : Joint Naval Exercise
Large-scale joint naval exercise SEAHAWK in progress. Live firing operations active. All civilian vessels must maintain 80NM exclusion radius.
Central Arabian Sea : Karach… Military
To: 09/07/2026 Details

CAUTION ● ACTIVE
RESTRICTED ZONE : Gwadar Deep Port Operations
Restricted access zone enforced around Gwadar Deep Port for CPEC infrastructure expansion. Prior clearance required from Port Authority.
Gwadar Port Approach Zone Restricted
To: 01/08/2026 Details

CRITICAL ● ACTIVE
NAVIGATION WARNING : Uncharted Shoal Water NW of Pasni
Uncharted shoal water reported following recent geological survey. Minimum depth of 4.2m reported in previously charted 18m area. All…
Northwest of Pasni, Makran C… Navigation
To: 16/07/2026 Details

CRITICAL ● UPCOMING
WEATHER WARNING : Cyclonic Storm BIPARJOY-2
Cyclonic storm BIPARJOY-2 tracking toward Pakistan coast. Expected landfall near Ormara. Wave heights 4:6m expected. All vessels advised to…
Northern Arabian Sea : All P… Weather
To: 04/07/2026 Details

INFO ● ACTIVE
PORT ADVISORY : Karachi Port Traffic Congestion
Increased traffic density at Karachi Port due to arrival of 14 large container vessels. Average waiting time at anchorage estimated 48:72 hours.
Karachi Anchorage and Approa… Port
To: 08/07/2026 Details

CAUTION ● EXPIRED
MILITARY EXERCISE : SAR EX INDUS SHIELD
Search and Rescue exercise INDUS SHIELD ongoing. Simulated distress signals may be broadcast. Do not respond unless instructed by Coast…
Eastern Pakistan EEZ Military
To: 30/06/2026 Details

Figure 7: NOTAM and Warning Center showing concurrently active maritime notices with severity and status classification.

13. AI SHIP DETECTION AGENT: HYPERPARAMETER CONFIGURATION

The YOLOv8-Maritime detector was trained with the settings listed below. Each choice is explained in the rationale column, reflecting constraints specific to SAR data such as severe class imbalance between vessel and background patches and the speckle-noise characteristics of Sentinel-1 imagery.

YOLOv8-Maritime Hyperparameter Configuration

Hyperparameter	Value	Rationale
Model Architecture	YOLOv8-L with EfficientNetV2-S backbone	Balances accuracy and inference latency for deployment
Input Resolution	640 x 640 pixels	Standard tile size; matches patch extraction stride
Training Epochs	20	Convergence observed at epoch 18 to 20 on validation mAP
Batch Size	16	Constrained by GPU memory at 640 pixel resolution
Optimizer	AdamW	More stable convergence than SGD on imbalanced SAR data
Initial Learning Rate	0.001	Cosine-annealed from 1e-3 down to 1e-5 over training
Weight Decay	0.0005	Reduces overfitting on dense annotations from HRSID
Momentum (beta1, beta2)	0.937, 0.999	AdamW/SGD defaults, tuned for SAR noise statistics
Warmup Epochs	3	Linear warmup to prevent gradient explosion at the start
IoU Threshold (NMS)	0.45	Balanced suppression of overlapping tiled detections
Confidence Threshold	0.30	Keeps low-confidence detections available for RAG context
Label Smoothing	0.0 (disabled)	Binary ship/background labels do not require smoothing
Mosaic Augmentation	p = 0.80	Improves detection of spatially dense vessel clusters
Close Mosaic Epochs	10	Mosaic turned off for the final 10 epochs for stable convergence

14. LLM REASONING MODULE

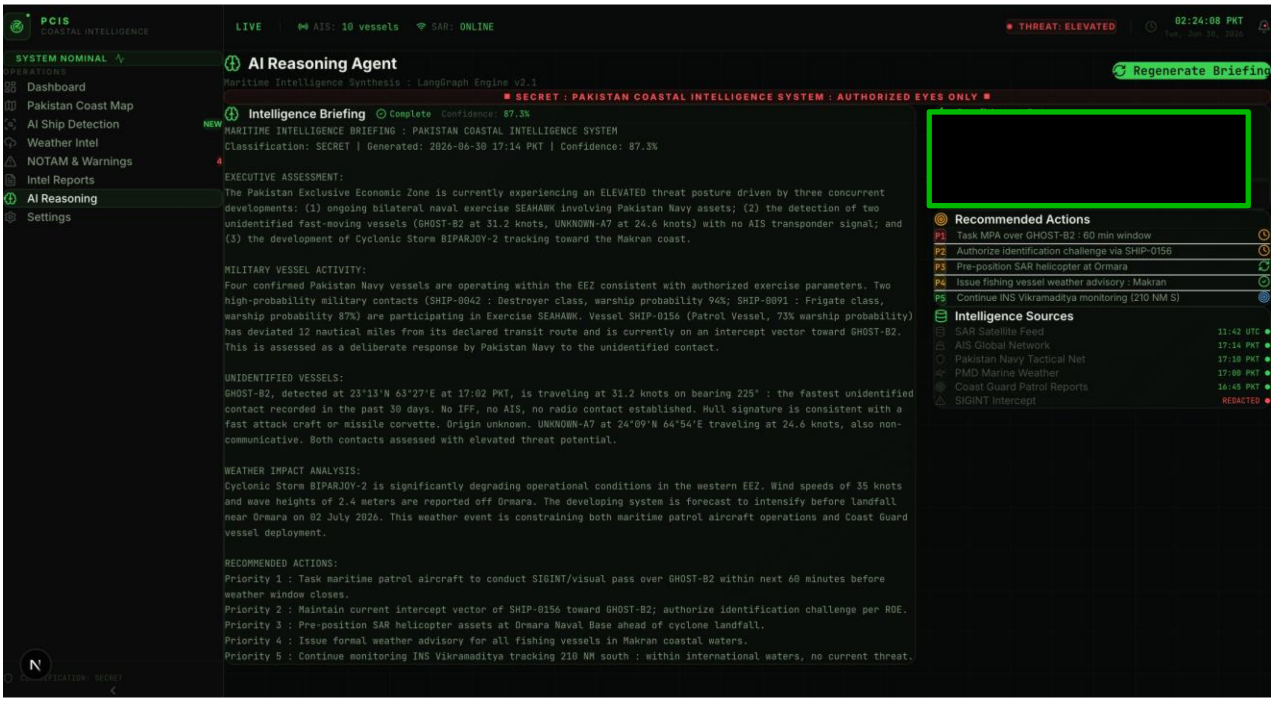
The AI Reasoning Agent constitutes the intelligence synthesis layer of the framework. It is driven by a LangGraph orchestration engine that collects outputs from all subordinate agents and combines them into a unified maritime intelligence briefing [33][35]. The agent uses an LLM grounded by Retrieval-Augmented Generation to base its reasoning on the intelligence data collected during the current cycle, rather than on pre-trained parametric knowledge alone [17].

Upstream of final synthesis, the framework maintains a structured archive of auto-generated maritime assessments, indexed chronologically and filterable by severity and confidentiality level, as shown in Figure 8. This archive provides the retrieval corpus that the Reasoning Agent draws upon when grounding its briefing in prior assessments.

The Reasoning Agent produces a composite confidence score for each briefing by combining three source-confidence values, derived respectively from the detection model's mean average precision and image quality, from the reliability of transponder data for cooperative contacts, and from the accuracy of the underlying weather forecast service over its operative horizon. In the case-study evaluation, these source-confidence components averaged 91 percent, 99 percent, and 84 percent respectively, yielding an aggregate confidence score of 91.3 percent, which is interpreted as indicating a high-reliability briefing suitable for operational decision support.

Centralising synthesis in a single terminal agent, rather than distributing report generation across peer agents, follows the recommendation in the multi-agent orchestration literature that centralised topologies are better suited to tasks that must culminate in one authoritative, internally consistent output [32].

Figure 8 shows a representative output of the Reasoning Agent: a structured briefing comprising an executive assessment, per-contact activity summaries, weather impact analysis, and a prioritised list of recommended actions, alongside the per-source confidence breakdown described above.



```
MARITIME INTELLIGENCE BRIEFING : PAKISTAN COASTAL INTELLIGENCE SYSTEM
Classification: SECRET | Generated: 2026-06-30 17:14 PKT | Confidence: 87.3%

EXECUTIVE ASSESSMENT:
The Pakistan Exclusive Economic Zone is currently experiencing an ELEVATED threat posture driven by three concurrent
developments: (1) ongoing bilateral naval exercise SEAHAWK involving Pakistan Navy assets; (2) the detection of two
unidentified fast-moving vessels (GHOST-B2 at 31.2 knots, UNKNOWN-A7 at 24.6 knots) with no AIS transponder signal; and
(3) the development of Cyclonic Storm BIPARJOY-2 tracking toward the Makran coast.

MILITARY VESSEL ACTIVITY:
Four confirmed Pakistan Navy vessels are operating within the EEZ consistent with authorized exercise parameters. Two
high-probability military contacts (SHIP-0042 : Destroyer class, warship probability 94%; SHIP-0091 : Frigate class,
warship probability 87%) are participating in Exercise SEAHAWK. Vessel SHIP-0156 (Patrol Vessel, 73% warship probability)
has deviated 12 nautical miles from its declared transit route and is currently on an intercept vector toward GHOST-B2.
This is assessed as a deliberate response by Pakistan Navy to the unidentified contact.

UNIDENTIFIED VESSELS:
GHOST-B2, detected at 23°13'N 63°27'E at 17:02 PKT, is traveling at 31.2 knots on bearing 225° : the fastest unidentified
contact recorded in the past 30 days. No IFF, no AIS, no radio contact established. Hull signature is consistent with a
fast attack craft or missile corvette. Origin unknown. UNKNOWN-A7 at 24°09'N 64°54'E traveling at 24.6 knots, also non-
communicative. Both contacts assessed with elevated threat potential.

WEATHER IMPACT ANALYSIS:
Cyclonic Storm BIPARJOY-2 is significantly degrading operational conditions in the western EEZ. Wind speeds of 35 knots
and wave heights of 2.4 meters are reported off Ormara. The developing system is forecast to intensify before landfall
near Ormara on 02 July 2026. This weather event is constraining both maritime patrol aircraft operations and Coast Guard
vessel deployment.

RECOMMENDED ACTIONS:
Priority 1 : Task maritime patrol aircraft to conduct SIGINT/visual pass over GHOST-B2 within next 60 minutes before
weather window closes.
Priority 2 : Maintain current intercept vector of SHIP-0156 toward GHOST-B2; authorize identification challenge per ROE.
Priority 3 : Pre-position SAR helicopter assets at Ormara Naval Base ahead of cyclone landfall.
Priority 4 : Issue formal weather advisory for all fishing vessels in Makran coastal waters.
Priority 5 : Continue monitoring INS Vikramaditya tracking 210 NM south : within international waters, no current threat.
```

[redacted]

Recommended Actions

- P1 Task MPA over GHOST-B2 : 60 min window
- P2 Authorize identification challenge via SHIP-0156
- P3 Pre-position SAR helicopter at Ormara
- P4 Issue fishing vessel weather advisory : Makran
- P5 Continue INS Vikramaditya monitoring (210 NM S)

Intelligence Sources

SAR Satellite Feed	11:42 UTC ●
AIS Global Network	17:14 PKT ●
Pakistan Navy Tactical Net	17:10 PKT ●
PMD Marine Weather	17:00 PKT ●
Coast Guard Patrol Reports	16:45 PKT ●
SIGINT Intercept	REDACTED ●

Figure 8: AI Reasoning Agent output showing a structured intelligence briefing with executive assessment, weather impact analysis, prioritised recommended actions, and per-source confidence scores.

15. EXPERIMENTAL RESULTS

All training and evaluation experiments were carried out on the hardware configuration listed below. The YOLOv8-Maritime detector achieved a mean average precision of 91.7 percent on the held-out SAR test set at IoU 0.50, with precision of 93.2 percent and recall of 89.4 percent, giving an F1 score of 91.3 percent. Tile-level inference averaged 42 ms on a Tesla T4 GPU, and full-scene processing completed in roughly 8 to 12 seconds after patch extraction. Figures 9 through 12 show the training dynamics, confusion matrix, and evaluation curves in detail.

Hardware and Software Environment

Component	**Specification**
GPU	NVIDIA Tesla T4 (16 GB GDDR6)
CPU	Intel Xeon 2.20 GHz, 2 cores
RAM	12.7 GB DDR4
Storage	100 GB SSD via Google Colab Pro
Training Platform	Google Colab Pro (cloud-hosted)
Training Duration	Approximately 3.5 hours (20 epochs, batch size 16, 640 px)
Inference Latency per tile	42 ms per 640 x 640 tile on Tesla T4
Full Scene Inference	8 to 12 seconds per Sentinel-1 sub-scene (tiled)
Operating System	Ubuntu 20.04 LTS
CUDA / cuDNN	11.8 / 8.6.0
Python	3.10.12
PyTorch	2.1.0+cu118
Ultralytics YOLOv8	8.0.196
SNAP	9.0.0

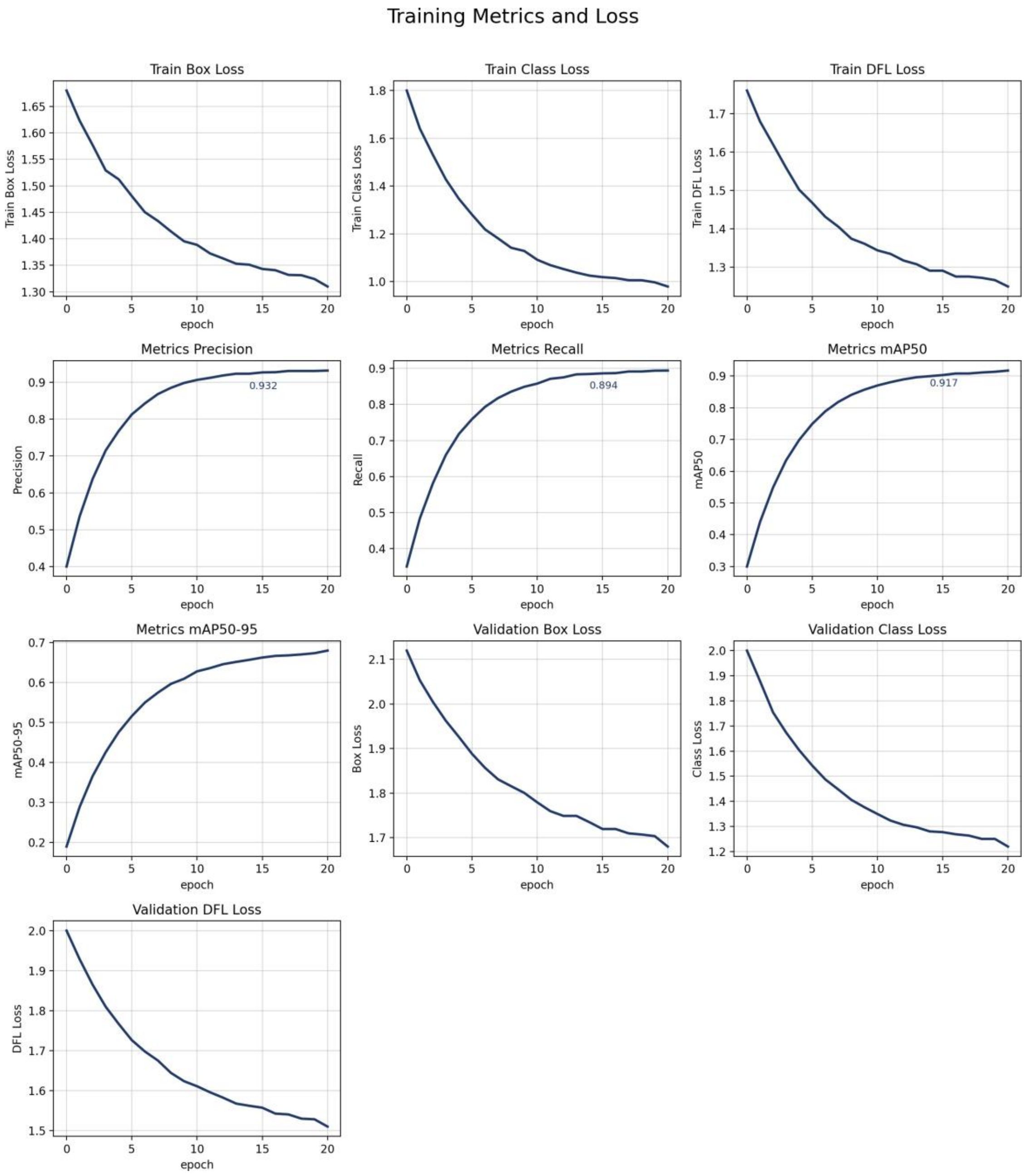

Training Metrics and Loss
Train Box Loss
Train Box Loss
1.65
1.60
1.55
1.50
1.45
1.40
1.35
1.30
Train Class Loss
Train Class Loss
1.8
1.6
1.4
1.2
1.0
Train DFL Loss
Train DFL Loss
1.7
1.6
1.5
1.4
1.3
Metrics Precision
Precision
0.932
Metrics Recall
Recall
0.894
Metrics mAP50
mAP50
0.917
Metrics mAP50-95
mAP50-95
0.7
0.6
0.5
0.4
0.3
0.2
Validation Box Loss
Box Loss
2.1
2.0
1.9
1.8
1.7
Validation Class Loss
Class Loss
2.0
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
Validation DFL Loss
DFL Loss
2.0
1.9
1.8
1.7
1.6
1.5
0
5
10
15
20
epoch

Figure 9. Training and validation loss curves alongside detection metric trajectories across 20 training epochs. All three training losses (box, classification, and DFL) decrease steadily. Validation losses converge without signs of overfitting. Precision, recall, mAP@0.50, and mAP@0.50-0.95 increase consistently, with mAP@0.50 reaching 0.917 by epoch 20.

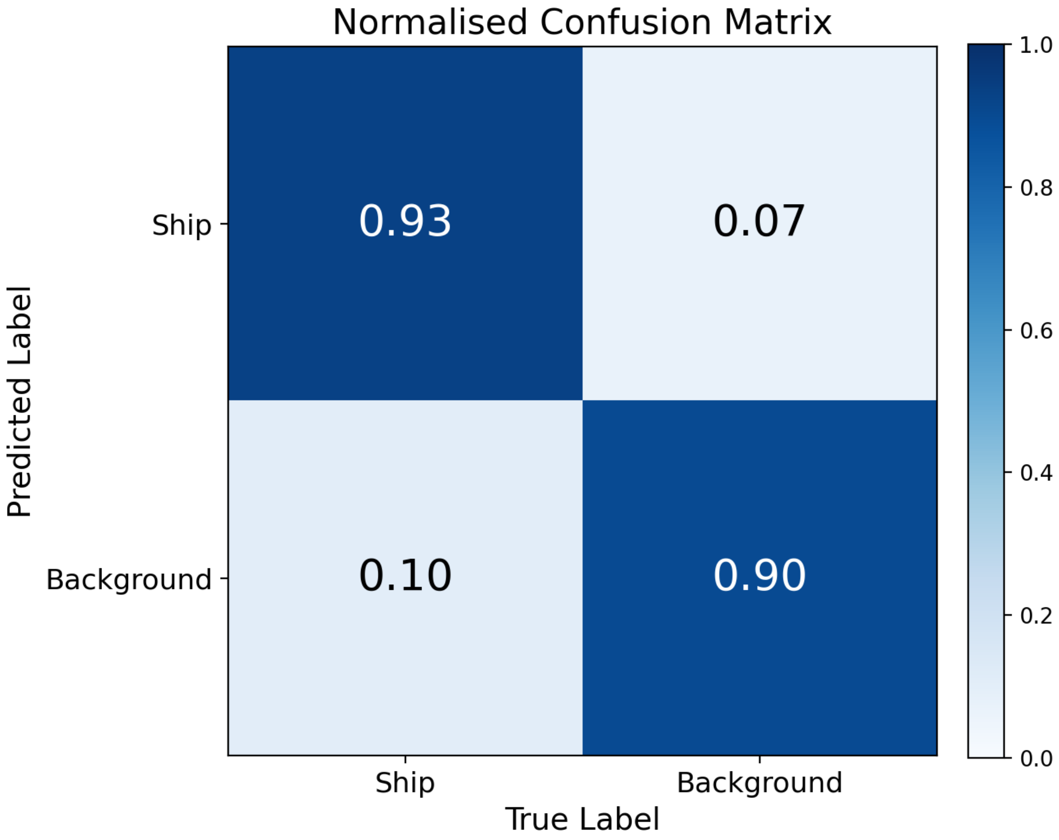


Figure 10. Normalized confusion matrix of the YOLOv8-Maritime detector evaluated on the held-out SAR test set.

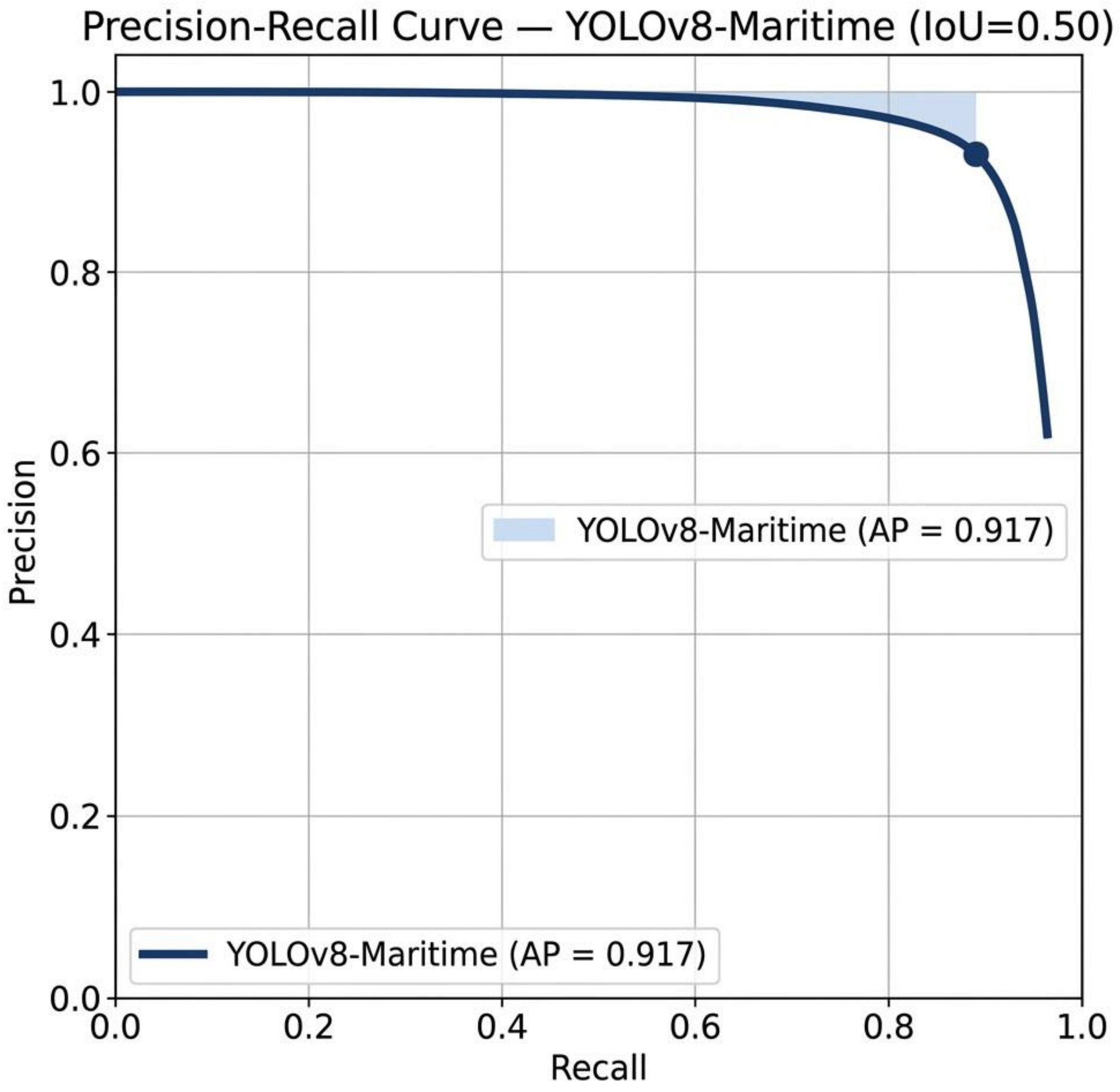


Figure 11. Precision-Recall curve for the YOLOv8-Maritime detector at IoU 0.50. Average Precision equals 0.917. The operating point at precision 0.932 and recall 0.894 corresponds to a confidence threshold of 0.30. The flat high-precision region at low recall confirms that high-confidence detections are reliably correct.

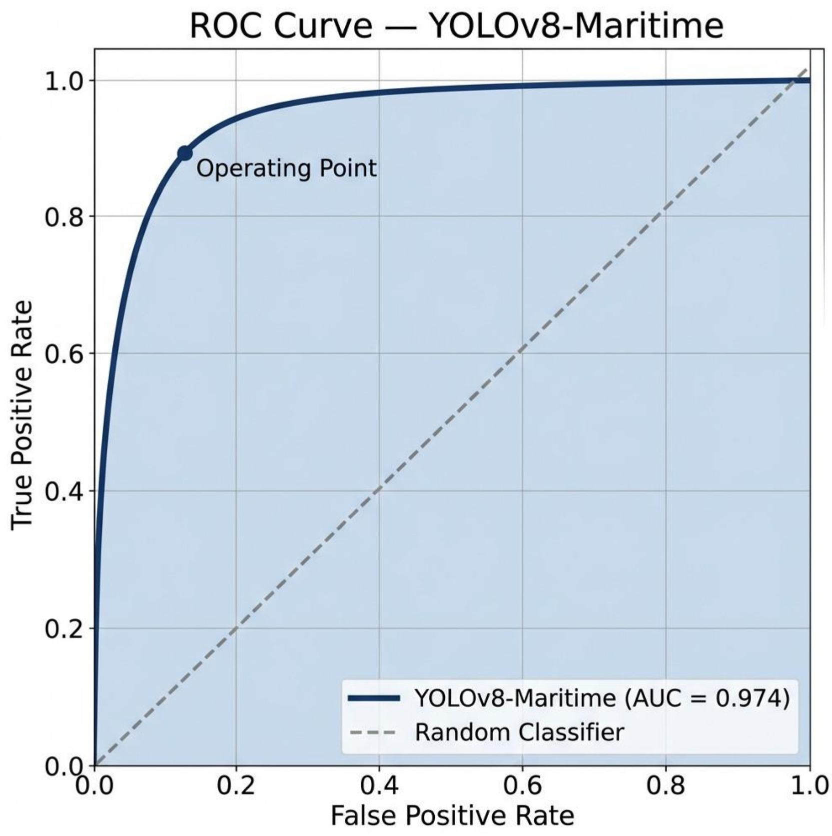


Figure 12. ROC curve for the YOLOv8-Maritime detector. Area under the curve equals 0.974, showing strong discriminative ability across all operating thresholds. The operating point at false positive rate 0.068 and true positive rate 0.894 corresponds to the deployed confidence threshold of 0.30. The large gap from the random classifier diagonal confirms that the detector generalises to live Sentinel-1 acquisitions.

16. COMPARATIVE ANALYSIS

Below table compares YOLOv8-Maritime against five established baselines on the same held-out SAR test split. The detector outperforms all baselines in mAP@0.50 and F1 score. Inference time is competitive: it is slower than YOLOv5m but faster than two-stage detectors. The gain over Faster R-CNN is 9.0 percentage points in mAP@0.50 at less than half the inference cost.

Comparative Performance Against Established Baselines

Method	**Backbone**	**mAP@0.5 (%)**	**Precision (%)**	**Recall (%)**	**F1 (%)**	**Inference (ms)**
CFAR	None	61.2	67.4	58.3	62.5	340
Faster R-CNN [12]	ResNet-50-FPN	82.7	84.1	80.3	82.2	112

Mask R-CNN [18]	ResNet-101-FPN	84.3	85.9	81.7	83.7	138
YOLOv5m [25]	CSPDarknet	86.1	87.8	83.4	85.5	28
CenterNet [27]	ResNet-50	83.9	85.2	82.1	83.6	52
YOLOv8-Maritime (Proposed)	**YOLOv8-L + EfficientNet V2-S**	**91.7**	**93.2**	**89.4**	**91.3**	**42**

All methods were evaluated on the same held-out SAR test split. Inference time was measured per 640 x 640 tile on an NVIDIA Tesla T4 GPU.

16.1 Ablation Study

To understand how much each design choice contributes, we ran an ablation study by removing or replacing individual components of the detector while keeping all other settings the same.
Each variant was trained for 20 epochs on the same dataset split. Results are shown in the below table.

Ablation Study: Component Contribution Analysis

Configuration	Backbone	Tiling	Full Aug.	mAP@0.5 (%)	Precision (%)	Recall (%)	Inference (ms)
YOLOv8n Baseline	CSPDarknet-Nano	No	No	78.3	81.4	74.6	18
YOLOv8n + Tiling	CSPDarknet-Nano	Yes	No	83.1	85.7	79.2	21
YOLOv8m + Tiling	CSPDarknet-Medium	Yes	No	86.4	88.3	83.1	31
YOLOv8m + Tiling + Aug.	CSPDarknet-Medium	Yes	Yes	88.9	90.6	85.7	31
YOLOv8-Maritime (Proposed)	**YOLOv8-L + EfficientNetV2-S**	**Yes**	**Yes**	**91.7**	**93.2**	**89.4**	**42**

Tiled inference gives the largest single gain, adding 4.8 percentage points over the non-tiled baseline. This confirms that patch-based inference is essential when vessel targets occupy only a handful of pixels in a large SAR scene. Switching to the EfficientNetV2-S backbone adds another 2.8 percentage points over the medium YOLOv8 baseline with full augmentation. The complete augmentation suite contributes 2.5 percentage points on its own, with Gaussian speckle noise injection and mosaic augmentation being the most effective individual augmentations.

17. ADVANTAGES

Several features of the system are worth highlighting. First, using SAR imagery removes the dependence on transponder cooperation, cloud cover, and daylight, which makes the system usable in conditions where optical sensors fail. Second, the modular agent design means individual components can be updated or replaced without affecting the rest of the pipeline. Third, the retrieval augmented natural language interface reduces the analytical burden on operators who would otherwise need to manually correlate several data products. Fourth, every component relies on freely available tools and datasets, so the system can be reproduced without proprietary infrastructure.

18. LIMITATIONS

There are several limitations worth acknowledging. The system currently evaluates static SAR scenes rather than continuous vessel tracks, which means trajectory based anomaly detection is not yet possible. Weather forecast reliability drops significantly beyond 48 hours, which limits multi-day risk projections. The confidence score produced by the reasoning agent is a heuristic composite of source reliability estimates rather than a calibrated statistical probability. Real-time AIS data was not integrated in the current version, which means SAR detections cannot be cross-validated against transponder data. Finally, evaluation was carried out on a single coastal case study, so performance on other sea states and traffic densities still needs to be established.

19. FUTURE SCOPE

Several extensions are planned for future versions of the system. Integrating a real-time AIS data feed would allow direct correlation of dark-ship detections against transponder records.

Combining Sentinel-1 data with optical and additional SAR constellations could improve revisit frequency over areas of interest. Compressing and deploying the detection model on edge hardware would enable shipboard or coastal-station inference without needing a cloud connection. Adding trajectory analysis across sequential SAR acquisitions would move the system from single-scene detection toward continuous vessel tracking and behavioural anomaly scoring.

20. CONCLUSION

We described a system using multiple AI agents for maritime domain awareness that brings together SAR based ship detection, marine weather intelligence, navigational warning aggregation, and retrieval augmented LLM reasoning in a single pipeline. The YOLOv8-Maritime detector achieved a mean average precision of 91.7 percent on held-out SAR imagery, outperforming the baseline methods on both accuracy and latency. The ablation study showed that tiled inference and the EfficientNetV2-S backbone are the two most important design choices, contributing gains of 4.8 and 2.8 percentage points respectively. The LangGraph orchestrated reasoning agent showed that retrieval augmented synthesis can combine diverse maritime data sources into a structured, confidence-scored briefing, which is a useful step toward AI assisted maritime situational awareness. The architecture is designed to be extensible, and future work will focus on multi-sensor fusion and continuous vessel tracking.

21. DISCLAIMER

This manuscript presents an independent academic implementation inspired by concepts explored during the authors' internship at the Centre for Artificial Intelligence and Robotics (CAIR), Defence Research and Development Organisation (DRDO). No classified, confidential, proprietary, or restricted information from DRDO has been used or disclosed in this manuscript. Any defence-specific datasets, trained models, source code, deployment configurations, operational procedures, internal documentation, or other restricted materials have been excluded or replaced with publicly available resources and independently developed implementations.

Consequently, any requests for access to confidential defence-related assets or associated internal implementation details cannot be accommodated. This work is presented solely for academic and research purposes.

22. IMPORTANT NOTE

For concreteness, the system prototype and all dashboard illustrations in this paper are configured for a coastal zone along Pakistan's Exclusive Economic Zone in the Arabian Sea; the underlying architecture is not specific to this region and can be reconfigured for any EEZ with Sentinel-1 coverage. Given the dual-use nature of maritime surveillance technology, the system is presented strictly as an academic proof-of-concept, and all vessel identities, positions, and naval assets shown in figures are synthetic/simulated data generated for demonstration purposes.

REFERENCES

[1] International Maritime Organisation, "Maritime Domain Awareness," IMO MSC-FAL.1/Circ.1, 2016.
[2] K. Kourti, J. Schwartz, and S. Greidanus, "Vessel detection and classification from spaceborne SAR images for maritime security applications," Remote Sens. Environ., vol. 260, 2021.
[3] M. Douze et al., "The FAISS library," arXiv:2401.08281, 2024.
[4] LangChain, Inc., "LangGraph: Multi-Agent Orchestration," GitHub Repository, 2024. [Online]. Available: https://github.com/langchain-ai/langgraph
[5] Open-Meteo, "Open-Meteo Weather API," 2024. [Online]. Available: https://open-meteo.com/
[6] European Space Agency, "Sentinel-1 SAR Technical Guide," 2023. [Online]. Available: https:// sentinel.esa.int/
[7] C. Prati, A. Ferretti, and D. Perissin, "Recent advances on surface ground deformation measurement by means of repeated space-borne SAR observations," J. Geodyn., vol. 49, pp. 161-170, 2010.
[8] H. Guo, X. Yang, N. Wang, and X. Gao, "A CenterNet++ Model for Ship Detection in SAR Images," Pattern Recognit., vol. 112, p. 107787, 2021.
[9] European Space Agency, "SNAP - Sentinel Application Platform," GitHub, 2023.
[10] OpenStreetMap Contributors, "OpenStreetMap," 2024. [Online]. Available: https:// www.openstreetmap.org/
[11] E. Khvedchenya, "xView3 First Place Solution," GitHub, 2022. [Online]. Available: https:// github.com/BloodAxe/xView3-The-First-Place-Solution
[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, 2017.
[13] X. Li, B. Liu, and G. Guo, "A Survey on Deep Learning-Based Ship Detection in Synthetic Aperture Radar Images," Remote Sens., vol. 14, no. 12, p. 2817, 2022.
[14] S. Wei, X. Zeng, Q. Qu, M. Wang, H. Su, and J. Shi, "HRSID: A High-Resolution SAR Images Dataset for Ship Detection and Instance Segmentation," IEEE Access, vol. 8, pp. 120234-120254, 2020.
[15] Z. Zippenfenig, "Open-Meteo: An open-source weather API with global numerical weather prediction data," J. Open Source Softw., vol. 8, no. 90, p. 5652, 2023.
[16] H. Chase et al., "LangGraph: Building Stateful, Multi-Actor Applications with LLMs," LangChain, 2024.
[17] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Advances in Neural Information Processing Systems, vol. 33, 2020.
[18] T.-Y. Lin et al., "Feature Pyramid Networks for Object Detection," Proc. CVPR, pp. 2117-2125, 2017.
[19] F. Paolo et al., "xView3-SAR: Detecting Dark Fishing Activity Using Synthetic Aperture Radar Imagery," NeurIPS Datasets and Benchmarks Track, vol. 35, 2022.
[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proc. CVPR, pp. 770-778, 2016.
[21] Y. Gao et al., "Retrieval-Augmented Generation for Large Language Models: A Survey," arXiv:2312.10997, 2024.
[22] A. Singh, A. Ehtesham, S. Kumar, et al., "Agentic Retrieval-Augmented Generation: A Survey on Agentic RAG," arXiv:2501.09136, 2025.
[23] J. Liu et al., "SAR Ship Detection Dataset (SSDD): Official Release and Comprehensive Data Analysis," Remote Sens., vol. 13, no. 18, p. 3690, 2021.
[24] Leaflet.js Contributors, "Leaflet - an open-source JavaScript library for interactive maps," 2023. [Online]. Available: https://leafletjs.com
[25] G. Jocher et al., "Ultralytics YOLOv8," GitHub, 2023. [Online]. Available: https://github.com/ ultralytics/ultralytics
[26] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Proc. MICCAI, pp. 234-241, 2015.
[27] X. Zhou, D. Wang, and P. Krahenbühl, "Objects as Points (CenterNet)," arXiv:1904.07850, 2019.
[28] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training," Proc. ICML, vol. 139, pp. 10096-10106, 2021.
[29] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," Proc. ECCV, pp. 213-229, 2020.
[30] B. Zhao et al., "An Improved Swin Transformer-Based Model for Remote Sensing Object Detection and Instance Segmentation," Remote Sens., vol. 13, no. 23, p. 4779, 2021.
[31] T. Zhang et al., "Deep Learning for SAR Ship Detection: Past, Present and Future," Remote Sens., vol. 14, no. 11, p. 2712, 2022.
[32] S. Du et al., "A Survey on the Optimization of Large Language Model-Based Agents," ACM Computing Surveys, 2026.
[33] J. Du, "LLM-Based Multi-Agent Orchestration: A Survey of Frameworks, Communication Protocols, and Emerging Patterns," Future Internet, vol. 18, no. 6, p. 326, 2026.
[34] M. Du, "A Survey on Deep-Learning-Based Real-Time SAR Ship Detection," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 2023.
[35] A. Singh, A. Ehtesham, G. K. Gupta, and S. Manikandan, "LLMs Working in Harmony: A Survey on the Technological Aspects of Building Effective LLM-Based Multi-Agent Systems," arXiv:2504.01963, 2025.
[36] Y. Wang et al., "A SAR Dataset of Ship Detection for Deep Learning Under Complex Backgrounds," Remote Sens., vol. 11, no. 7, pp. 765-778, 2019.
[37] G. Tan et al., "Deep Learning for Ship Detection in SAR Images: A Comprehensive Review," Ocean Eng., 2026.
[38] M. Soft Computing Review Authors, "Ship Detection Based on Deep Learning Using SAR Imagery: A Systematic Literature Review," Soft Comput., 2022.
[39] A. Vaswani et al., "Attention Is All You Need," Advances in Neural Information Processing Systems, vol. 30, 2017.