AI based satellite system architecture to detected and classify ship activities

Renan Menezes
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
renang@ita.br

Gleisson Bezerra
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
gleisson@ita.br

Lidia Sato
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
shibuya@ita.br

Luis Loures
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
loures@ita.br

Abstract—**Maritime Domain Awareness (MDA) is essential for ensuring global maritime security and combating threats such as piracy, smuggling, and illegal, unreported, and unregulated (IUU) fishing. Traditional surveillance systems, heavily reliant on the Automatic Identification System (AIS), face limitations due to AIS deactivation or manipulation by vessels engaged in illicit activities. These blind spots highlight the need for alternative methods to monitor maritime operations more effectively. This study proposes a satellite-based system architecture that enhances MDA capabilities by employing high-resolution imagery and machine learning (ML) algorithms to detect and classify vessels in near real-time. The system begins with the acquisition of satellite images from targeted Regions of Interest (ROIs). These images are processed using supervised ML models trained on annotated maritime datasets, enabling accurate ship identification based on characteristics such as hull shape, dimensions, and structural features. A key innovation is the consideration of onboard processing, where a machine learning model could be deployed directly on the satellite. This approach significantly reduces latency in detection and reporting, addresses bandwidth limitations in downlink communications, and increases the system's operational responsiveness—crucial for time-sensitive maritime interventions. While AIS data is not fused in the current implementation, its integration is planned for future work. Incorporating AIS signals can provide valuable metadata for cross-verification with satellite detections, improving the identification of non-compliant or suspicious vessels. However, in this study, AIS is acknowledged only as a contextual reference. The proposed system is evaluated through performance simulations, focusing on detection accuracy, classification precision, and onboard processing efficiency. This work demonstrates the potential of combining satellite imagery and machine learning for autonomous maritime monitoring. It establishes a scalable foundation for future MDA systems, with applications including enforcement against IUU fishing, enhancing port security, and aiding in search and rescue operations. Future developments will include AIS fusion and multispectral image analysis to broaden system functionality and reliability.**

Keywords—Maritime Domain Awareness (MDA), Satellite Imagery, Machine Learning, Ship Detection

I. INTRODUCTION

Maritime Domain Awareness (MDA) refers to the effective understanding of anything associated with the maritime domain that could impact national security, commerce, or the environment. As global maritime traffic increases and geopolitical tensions persist, MDA becomes a strategic imperative for securing territorial waters, combating illicit activities, and ensuring the sustainable use of marine resources. Applications of MDA range from tracking commercial shipping routes and monitoring fishing zones to supporting search and rescue operations and enforcing maritime law. As such, governments and agencies increasingly invest in technologies and systems that enhance situational awareness across vast oceanic areas.

One of the primary tools in achieving MDA is the Automatic Identification System (AIS), which is mandated for most large vessels under international regulations. AIS transmits real-time information such as vessel identity, position, heading, and speed, allowing authorities to track maritime traffic efficiently. However, AIS has significant limitations. Vessels involved in illicit activities, such as illegal fishing, smuggling, or unauthorized maritime incursions, often deactivate their AIS transponders or falsify their location data to avoid detection. Furthermore, AIS signals can be spoofed or jammed, creating vulnerabilities in maritime situational awareness. These limitations reveal critical blind spots in surveillance systems that rely solely on AIS, thus reinforcing the need for complementary and resilient monitoring technologies.

To address these gaps, vision-based computational monitoring using satellite imagery has emerged as a promising solution. High-resolution optical and synthetic aperture radar (SAR) imagery provides wide-area surveillance capabilities that are passive and independent of AIS transmissions. By applying machine learning algorithms, particularly convolutional neural networks (CNNs), to satellite images, it is possible to autonomously detect, classify, and track vessels, even in the absence of AIS signals. This study focuses specifically on the use of high-resolution optical imagery combined with supervised learning models for ship detection and classification.

This work is organized to present a comprehensive overview of the proposed satellite-based maritime surveillance system and its potential for enhancing Maritime Domain Awareness (MDA). Initially, the study introduces the challenges of traditional vessel monitoring approaches, emphasizing the limitations of the Automatic Identification

System (AIS) in detecting non-compliant or illicit maritime activity. Following this, the system architecture is detailed, including image acquisition strategies, machine learning model development for vessel detection and classification, and the rationale for onboard processing implementation. Subsequently, a performance evaluation of the proposed approach is conducted through simulations, highlighting detection accuracy, classification precision, and computational efficiency in a satellite-constrained environment. The work concludes with a discussion on future developments, including the integration of AIS data for cross-validation and the potential for real-time operational deployment.

II. Literature Survey

A. Limitations of AIS and the Need for Complementary Systems

The Automatic Identification System (AIS) is a standardized maritime communication system designed to enhance the safety and efficiency of vessel navigation. AIS operates primarily in the Very High Frequency (VHF) maritime band, broadcasting real-time navigational and identification data, including vessel name, Maritime Mobile Service Identity (MMSI), position, course over ground (COG), speed over ground (SOG), navigational status, and voyage-related information. This data is transmitted autonomously at regular intervals and can be received by other ships, coastal stations, and increasingly, by satellite-based AIS (S-AIS) receivers, enabling a global reach.

The AIS plays a critical role in Maritime Domain Awareness (MDA), but its effectiveness is challenged by several technical, operational, and policy-related issues. Technically, the vast volume of AIS data comprising real-time vessel positions and voyage information, requires robust systems for validation, fusion, storage, and integration, while the risk of overloading the AIS VHF Data Link (VDL) due to increased usage and binary applications may necessitate updates to the communication standard or additional infrastructure. Policy challenges include the management of AIS binary applications, regulation of data sharing among governmental and private entities, and the unintended commercial use of AIS data, which may conflict with international laws and discourage vessel compliance. Frequency allocation within the congested VHF maritime band further complicates long-range AIS reception, especially by satellite. Additionally, the reliability of AIS data depends on correct onboard equipment operation and data entry, necessitating improved enforcement mechanisms to ensure system integrity and prevent issues such as interference and jamming. Addressing these challenges is essential for AIS to fully support global maritime security and surveillance objectives. [1]

B. Satellite Imagery for Vessel Detection

The integration of deep learning techniques into maritime surveillance has significantly advanced the capabilities of vessel detection and classification using satellite imagery. Reference [2] introduced a deep learning framework that leverages convolutional neural networks (CNNs) to detect vessels in satellite images, demonstrating high accuracy in various maritime scenarios. Similarly, reference [3] employed CNN-based models for vessel detection, emphasizing the importance of robust preprocessing and data augmentation to enhance detection performance. A hybrid approach combining traditional image processing with deep learning was proposed in a study presented at the IEEE International Geoscience and Remote Sensing Symposium (IGARSS) [4], which improved detection accuracy in high-resolution satellite imagery by integrating multiple feature extraction methods. In the realm of vessel classification, reference [5] explored the use of capsule networks, which capture spatial hierarchies more effectively than traditional CNNs, resulting in improved classification of maritime vessels. Collectively, these studies underscore the transformative impact of deep learning on maritime domain awareness, offering enhanced accuracy and efficiency in vessel detection and classification tasks.

C. AIS and Optical Vessel Classification Data Fusion

The fusion of Automatic Identification System (AIS) data with satellite imagery has emerged as a pivotal technique in enhancing maritime surveillance and vessel detection. A method was developed that integrates AIS data with visual detections from deep learning models, such as YOLOv5, to enrich datasets with vessel-specific information like type, size, speed, and direction [6]. This approach utilizes homography-based matching to associate detected ships with corresponding AIS messages, achieving an association accuracy of up to 85.06% for fixed cameras. Another study proposed a fusion method combining AIS data with high-resolution satellite imagery to improve the identification and positioning of maritime targets [7]. By employing a point-set matching algorithm and a fuzzy comprehensive decision method, this approach significantly reduced positioning errors, with over a 70% reduction in root mean square error and positioning errors controlled within 4 pixels. A fully automated framework was also introduced that fuses AIS data with Synthetic Aperture Radar (SAR) satellite images for vessel detection [8]. This method automatically annotates satellite images by correlating them with AIS data, enabling the training of convolutional neural networks without the need for manually labeled datasets. The trained model achieved an accuracy of 88% in ship detection, demonstrating the effectiveness of the approach in identifying vessels, including those that may have turned off their AIS transponders.

III. Methodology

This work proposes a structured and systematic methodology aimed at simulating, constructing, training, and evaluating a maritime vessel detection model using Google Earth image. The overall approach was designed to reflect realistic operational conditions by considering the use of a Commercial Off-The-Shelf (COTS) optical payload onboard a small satellite operating in low Earth orbit (LEO) as parameter to build the dataset. The methodology integrates orbital mechanics, sensor modeling, region-specific data collection, synthetic data generation, and machine learning model development. The methodology follows a structured pipeline, as illustrated in Figure 1, which ensures consistency between the system-level assumptions and the conditions under which the model is trained. The primary objective of this pipeline is to produce a dataset and a detection algorithm that accurately reflect the constraints and capabilities of a real satellite mission. The pipeline is composed of the following sequential steps:

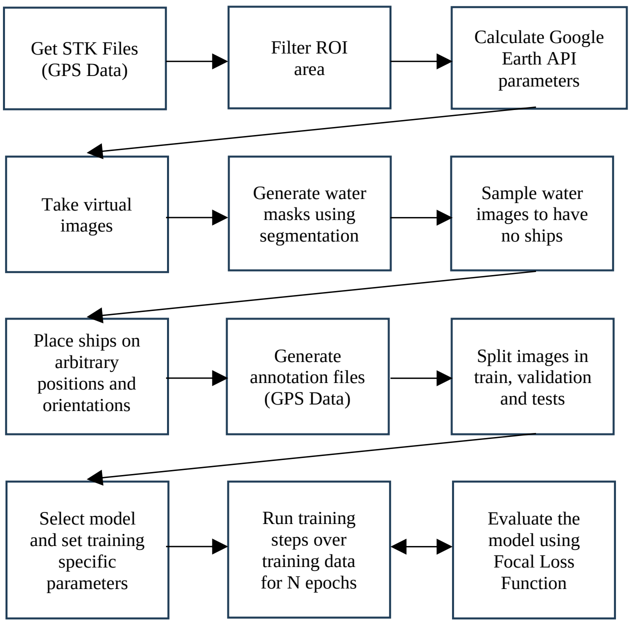

Figure 1 - End-to-End Methodology Pipeline for Mission-Constrained Dataset Generation and Model Training

A. *Camera Consideration*

The dataset selection and interpretation were guided by the assumption of using a Commercial Off-The-Shelf (COTS) optical camera module suitable for small satellite platforms. The assumed payload characteristics include a Ground Sampling Distance (GSD) of 4.75 meters at an orbital altitude of 500 km and a swath width of 19.4 km at the same altitude. This assumption is critical for the system design as it defines the spatial resolution and coverage area that influence the detectable ship sizes and the number of vessels per image file.

B. *Orbit Consideration*

Orbital Altitude Influence on GSD, for this study, an orbital altitude of approximately 370 km is considered. This lower altitude provides an improved Ground Sampling Distance (GSD) when compared to higher orbits. Assuming a linear relationship between altitude and GSD, and using the baseline GSD of 4.75 meters at 500 km, the GSD at 370 km can be estimated by the proportional formula:

$$GSD370 = GSD500 \times \left(\frac{370}{500}\right) \quad (1)$$

This higher resolution enhances the ability to detect and classify smaller vessels and fine features. The table 1 illustrates the comparative effect of orbital altitude on GSD.

Table 1 - Estimated Ground Sample Distance (GSD) at Different Orbital Altitudes and Corresponding Impact on Image Resolution

Altitude (km)	Estimated GSD (m)	Resolution Impact
600	5.70	Coarse – suitable for large ships
500	4.75	Baseline – moderate detail
400	3.80	Improved – small vessel detection
370	3.52	High – enhanced structural detail

These values assume a nadir-pointing camera with consistent focal length and sensor properties. This reduction in altitude directly impacts revisit time, swath width, and atmospheric drag considerations, which must be accounted for in mission planning.

C. *Region of Interest Consideration*

The selected Region of Interest (ROI) encompasses the South Atlantic region, specifically the Brazilian Exclusive Economic Zone (EEZ), Figure 2 highlights in red the selected Region of Interest (ROI). This region comprises extensive maritime zones that are subject to significant economic activities and recurrent incidents of Illegal, Unreported, and Unregulated (IUU) fishing. The selection of this area is justified by its strategic relevance to national interests, the challenges associated with effective surveillance coverage, and its critical role within Brazil's maritime security policy.

Figure 2 - Brazilian Exclusive Economic Zone (EEZ), Region of Interest (ROI)

D. *Dataset Description*

To assess the potential imaging opportunities and construct a representative dataset, an orbital analysis was conducted using Systems Tool Kit (STK) software. The analysis simulated a scenario over a 30-day period with a time step of 1 second, allowing precise estimation of access durations and satellite visibility windows over targeted maritime regions. Based on the visibility events obtained from this analysis, image acquisition opportunities were derived. The resulting synthetic dataset comprises a total of 2,592,000 satellite image captures across multiple oceanic and coastal locations globally. From this global dataset, a subset of 17,344 images was specifically selected from the Region of Interest (ROI) encompassing the Brazilian Exclusive Economic Zone (EEZ) Also known as Blue Amazon. This selection was guided by the orbital passes that yielded favorable imaging geometries over the ROI. The extracted subset formed the basis for training and evaluating the detection models under operationally relevant conditions.

E. Dataset Preparation

To simulate the conditions of a COTS camera onboard a satellite, images were captured using the Google Earth API. The capture parameters of the API were chosen to emulate a Ground Sampling Distance (GSD) of 3.52 meters and a swath width of 14.356 km. Each image was rendered with a resolution of 1280 × 1280 pixels, replicating the characteristics of the assumed onboard optical sensor at an altitude of 370 km. This step was fundamental in ensuring consistency between dataset generation and the expected spatial resolution of an real system.

After acquisition, a segmentation step was applied to identify and retain only images containing any portion of water, as illustrated in Figure 3. This process was essential to filter out frames dominated by island regions, port infrastructure, or areas with minimal ocean coverage, common occurrences due to the proximity to coastlines. A water binary segmentation algorithm was employed to exclude land-dominated images, ensuring that the retained dataset accurately reflects typical maritime observation conditions

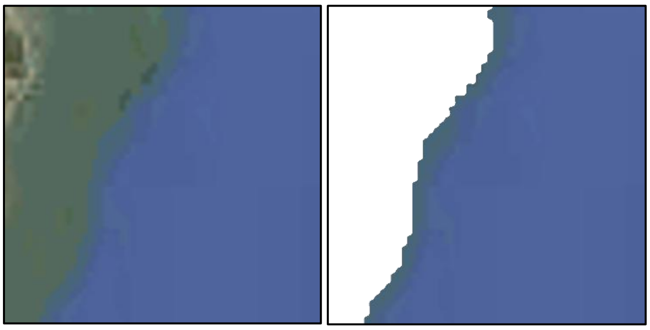

Figure 3 - Water Mask Segmentation Process for Filtering Maritime Observation Imagery

To enrich the dataset, augmentation procedure was implemented. Notably, this augmentation involved the insertion of synthetic ship images rather than real ship cutouts. Computer-generated models of ships, designed to match real-world vessel types and dimensions, were superimposed onto oceanic backgrounds according to the segmentation. The synthetic ships were varied in orientation, scale, illumination, and sea conditions to simulate operational diversity. This method allowed for controlled training data generation while preserving the generalization ability of the model.

F. Ship Classification

The synthetic ships added to the dataset were modeled to represent a range of vessel categories commonly encountered in maritime monitoring operations. These vessel types were classified into three primary size categories based on their length: small (10–20 meters), medium (21–50 meters), and large (51–100 meters). The pixel dimensions were estimated based on a Ground Sampling Distance (GSD) of 3.52 meters at 370 km altitude and an image resolution of 1280 × 1280 pixels as referenced in table 2.

Table 2 - Vessel Size Categories and Image Resolution Details

Size Category	**Estimated Length (m)**	**Estimated Width (m)**	**Pixel Length (px)**	**Pixel Width (px)**
Large	51–100	10–20	15–28	3–6
Medium	21–50	5–10	6–14	1–3
Small	10–20	2–4	3–6	1–2

This classification scheme ensured diversity in the types and dimensions of maritime targets present in the training dataset, which is essential for generalizing the detection model across a wide spectrum of operational scenarios, see Figure 4 for the types of ships included.

Figure 4 - Representative Types of Ships Used in the Training Dataset

G. Model's Benchmark

In this study, four distinct convolutional neural network (CNN) architectures were implemented and evaluated for the task of ship classification using satellite imagery. All models were selected based on their proven effectiveness in object detection tasks and were configured with Feature Pyramid Networks (FPN) to enhance multi-scale feature extraction, a critical capability for detecting ships of varying sizes and orientations. The first model employed was Faster R-CNN with MobileNetV3 Small and FPN, designed to offer a favorable trade-off between inference speed and accuracy, particularly suitable for deployment in resource-constrained environments such as onboard satellite processing units. The second architecture, Faster R-CNN with MobileNetV3 Large and FPN, builds upon this by incorporating a larger backbone, thereby increasing the model's representational capacity while maintaining relatively efficient computation. The third model utilized was RetinaNet with ResNet50 and FPN. RetinaNet is a one-stage detector that incorporates focal loss to address the class imbalance problem inherent in object detection. The ResNet50 backbone provides deep feature extraction capabilities, making it a robust choice for complex scenes involving cluttered backgrounds or partially occluded vessels. Finally, Faster R-CNN with ResNet50 and FPN was also implemented to leverage the high detection accuracy of the two-stage Faster R-CNN framework, combined with the depth and generalization ability of the ResNet50 backbone. Table 3 summarizes the comparative characteristics of each model in terms of parameter size and response time in frames per second (FPS):

Table 3 - Comparison of model sizes and response times (FPS)

Model Name	**Size (Parameters)**	**Response Time (FPS)**
FasterRCNN Mobilenet v3 Small FPN	3.9M	70-80
FasterRCNN Mobilenet v3 Large FPN	6.1M	50-60
RetinaNet ResNet50 FPN	34.0M	20-30
FasterRCNN ResNet50 FPN	41.0M	14-20

The results in Table 3 highlight a clear trade-off between model complexity and processing speed. The MobileNetV3-based models, with significantly fewer parameters (3.9M and 6.1M), demonstrate superior real-time performance with response times reaching up to 80 FPS, making them suitable for real-time applications and hardware-limited platforms. In contrast, the ResNet50-based models offer greater detection accuracy due to deeper feature extraction, but this comes at the cost of reduced inference speed, with the heaviest model (Faster R-CNN with ResNet50 FPN) reaching only 14–20 FPS. This performance comparison supports a balanced selection of CNN architectures depending on the constraints and goals of the deployment scenario, particularly for onboard processing in space systems where computational efficiency is critical.

The reported performance values were obtained by averaging results during inference using the model.eval() mode in PyTorch, with the torch.no_grad() context enabled to prevent gradient computation. A batch size of 2 was utilized during testing, and all experiments were conducted on a system equipped with an NVIDIA RTX 4060 GPU. The integration of the Feature Pyramid Network (FPN) in all models contributed to improved handling of multi-scale object detection, which is particularly relevant for detecting ships of varying dimensions. MobileNet backbones were selected for their optimization toward near real-time performance on resource-constrained devices. In contrast, the Faster R-CNN framework is generally associated with higher detection accuracy, while RetinaNet is recognized for offering a balanced compromise between accuracy and computational efficiency. An initial evaluation was performed to identify which models could effectively utilize a fourth image channel, an alpha channel, alongside the standard RGB configuration. This additional channel encodes spatial information derived from the image segmentation process, indicating precise ship locations within the dataset. To ensure a fair and consistent comparison across architectures, all models were modified to accept both three-channel (RGB) and four-channel (RGB+A) inputs.

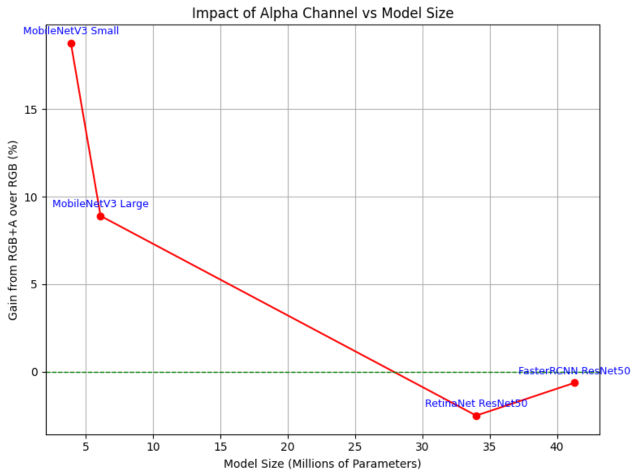

Figure 5 - Model Performance Evaluation and Comparison of Multi-Scale Ship Detection with Multi-Channel Input Configurations

The Total Loss function used in the training of these models was as follows:

Classification Errors + Bounding Box Regression Error + Detection Error of Any Object + FPN Regression Error

The Figure 5 presents the impact of model complexity on the performance gains obtained from incorporating an additional alpha channel in the input images, in comparison to the standard RGB configuration. The vertical axis indicates the relative gain (in percentage) of using four-channel inputs (RGB + Alpha), while the horizontal axis represents the model size in millions of parameters. To produce these results, each model was trained on a subset of 2,000 labeled samples from the training dataset. Evaluation was conducted over 1,000 validation samples by computing the average of the total accumulated losses, allowing for a consistent and comparative performance assessment across all models. The alpha channel, introduced as a fourth input channel, encodes spatial priors derived from the segmentation stage, highlighting areas of probable ship presence. This feature was included to assess whether such spatial guidance could enhance detection capabilities, particularly in low-complexity models. The experimental results clearly demonstrate that lightweight architectures, notably MobileNetV3 Small, significantly benefit from the inclusion of the alpha channel, achieving over

15% improvement in loss reduction compared to their RGB-only counterparts. MobileNetV3 Large also exhibits a notable gain, albeit lower, around 8%. This behavior suggests that lightweight models, which are constrained in depth and feature extraction capacity, are able to leverage the additional spatial information to better localize and classify ship targets. In contrast, larger models such as RetinaNet ResNet50 and Faster R-CNN ResNet50 show minimal to negative performance gains when incorporating the alpha channel. In the specific case of RetinaNet, performance degradation was observed, indicating that the additional channel may be treated as noise. This can be attributed to the high representational capacity of such models, which are equipped with numerous convolutional filters designed to autonomously extract complex spatial and contextual features. Consequently, these models may already learn to localize regions of interest effectively using RGB inputs alone, rendering the alpha channel either superfluous or counterproductive. In summary, Figure 5 underscores the relevance of model capacity in determining the effectiveness of data augmentation strategies such as multi-channel input expansion. While lightweight architectures benefit markedly from the addition of a spatially informative channel, larger and more complex models exhibit limited or adverse responses, emphasizing the importance of tailoring input design to model characteristics.

H. Training Setup

After preliminary analysis, two models were chosen for full experimentation: FasterRCNN ResNet50 configured with RGB only and The MobileNet V3 Small configured for RGB+A.

The dataset was partioned into three subsets as follows:

- **Training Set**: 9,621 images
- **Validation Set**: 4,124 images
- **Test Set**: 4,576 images (unseen during training)

Of the total annotated images containing ships, the distribution was as follows:

- **Training Images with Ships**: 8,664
- **Validation Images with Ships**: 3,713

The synthetic ships used for training comprised a total of 90 distinct vessel models, distributed across different size categories as defined below:

Key training settings included:

- Input image resolution: 1280x1280 pixels
- Batch size: 4 for training and 2 for validation
- Learning rate: 0.01 with scheduled decay at 0.8 according to validation monitoring
- Optimizer: Stochastic Gradient Descent (SGD)
- Epochs: 30

To enhance model robustness and generalization, standard data augmentation techniques were employed, including horizontal flipping, rotation, color jitter, and mosaic augmentation. Additionally, label smoothing and Intersection over Union (IoU)-based loss functions were applied to improve the precision of bounding box predictions.

I. Evaluation Framework

The evaluation framework for the vessel detection and classification model was structured around widely adopted object detection performance metrics to ensure comprehensive assessment and comparability. These metrics provide a multi-faceted view of the model's capabilities in identifying vessels under varied maritime scenarios.

1. Mean Average Precision (mAP@0.5)

The principal metric for evaluating detection performance is the mean Average Precision at an Intersection over Union (IoU) threshold of 0.5, denoted as mAP@0.5. This metric reflects the model's ability to correctly localize objects across different classes.

2. Focal Loss with Weighted Class (FLWC)

Custom loss function that extends the traditional Focal Loss by integrating class-specific recall weights that emphasizes hard-to-classify examples by reducing the loss contribution of well-classified ones, the addition of weighted recall further biases the optimization process toward improving the recall of underrepresented or critical classes. This approach is particularly useful in object detection or classification tasks where missing (or less present) certain classes (low recall) is more detrimental than false positives. By dynamically scaling the loss with recall-based weights, the model is guided to pay more attention to classes where recall performance is lacking, leading to improved sensitivity and robustness in challenging scenarios.

3. F1 Score

The F1 Score is the harmonic mean of precision and recall, providing a single performance metric when an equilibrium between the two is desired. It is particularly useful in maritime contexts where both under-detection (missed vessels) and over-detection (false alarms) can carry operational costs.

4. *Training Results*

1. FastrRCNN ResNet50 (RGB)

The FastrRCNN ResNet50 network used a validation loss function as follows:

$$ValLoss = (1 - mAP@0.5) + FLWC(F1) \quad (2)$$

The table 4 present the weight given to the classes for the FLWC:

Table 4 - Weight given to the classes for the FLWC

Class	**Weight**
LARGE	0.5
MEDIUM	1.0
SMALL	1.5

Since there is no expectation of running a network of this size onboard, the choice of these parameters was made to pursue more precision in small objects without losing classes using F1 to balance. The detailed comparison of training and validation loss is shown in Figure 6.

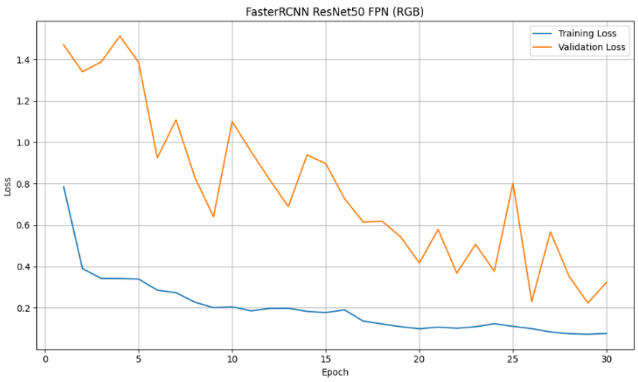

Figure 6 - Training and Validation Loss Over Epochs for Object Detection Precision – FasterRCNN ResNet50 FPN (RGB)

It is interesting to observe validation oscillation due to the combination of Validation Loss parameters. However, the training loss follows a smooth curve demonstrating the high power of this network in detecting small objects.

Figure 7 - Trust Percentage on an Unseen Example Not Included in Training or Validation - FasterRCNN ResNet50 FPN (RGB)

As shown in Figure 7, the very high percentage of trust obtained by this network can be observed in an example that was not in either the training or validation datasets.

2. Mobilenet V3 Small (RGB+A)

The Mobinet network used a validation loss function as follows:

$$ValLoss = 0.5 \times (1 - mAP@0.5) + 0.1 \times FLWC(Recall) \quad (3)$$

The table 5 present the weight given to the classes for the FLWC:

Table 5 - Weight given to the classes for the FLWC

Class	Weight
LARGE	1.5
MEDIUM	1.0
SMALL	0.5

This combination of parameters allowed the network to converge in its learning and, generalizing enough, not being so penalized by the deficiency of detecting smaller ships, since it has less capacity for small objects.

This network has more possibility of embedded execution. So, the weights between the terms of the equation were chosen to favor more the hit of bounding boxes than the class and leave an eventual disambiguation to be done on the ground. For this reason, Recall was also chosen as a measure to weigh against the classes. The detailed comparison of training and validation loss is shown in Figure 8.

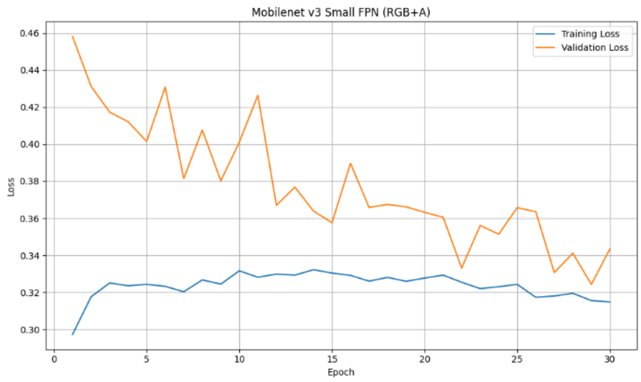

Figure 8 - Training and Validation Loss Over Epochs for Object Detection Precision – Mobilenet v3 Small FPN (RGB +A)

It is interesting to observe the same oscillation of the validation due to the combination of parameters of the Validation Loss. However, in this case it tends to take much longer to stabilize due to the size of the network and the difficulty of the network alone detecting smaller ships correctly without the help of Validation to correct its learning rate.

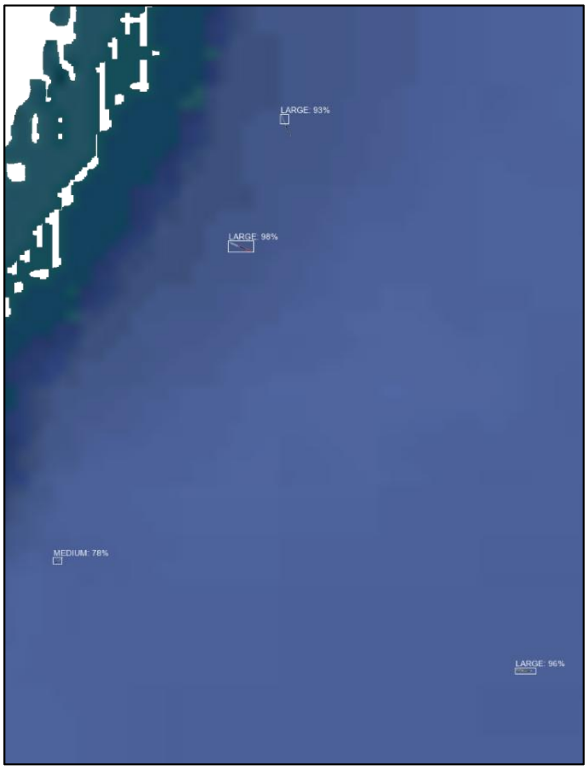

Figure 9 - Trust Percentage on an Unseen Example Not Included in Training or Validation - Mobilenet v3 Small FPN (RGB +A)

As shown in Figure 9, Comparatively, we can see the same image where a detection below the threshold of 0.5 was missed and small divergences in the boxes with a lower degree of confidence in the smaller classes as well.

IV. SYSTEM ARCHICTUCTURE AND ONBOARD CONSTRAINTS

Designing a satellite-based system capable of detecting and classifying maritime vessels onboard requires the consideration of various architectural and environmental constraints. Based on performance results of two selected networks, it is shown in the Figure 10 a high-level diagram of potential architecture to be developed as next steps.

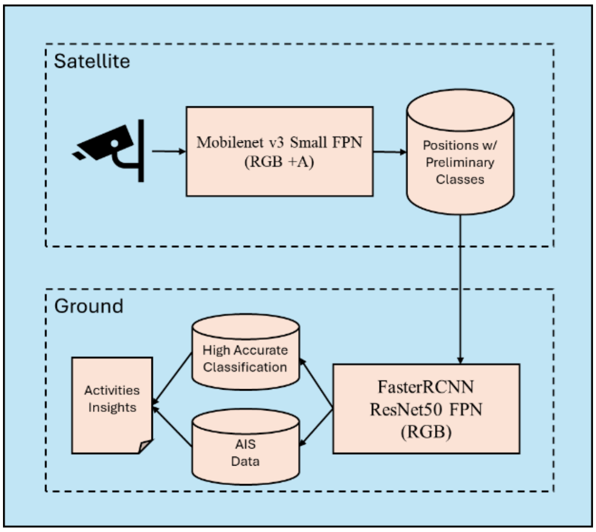

Figure 10 – Proposed AI based system hybrid architecture to detected and classify ship activities

A. System Pipeline Considerations

The system architecture must be capable of supporting a sequence of tasks: image acquisition, preprocessing, inference using a machine learning model, event handling, and data downlink. Each stage must be optimized to operate efficiently within the temporal and spatial constraints imposed by the satellite's orbit. The pipeline should be designed with deterministic timing in mind, ensuring that operations complete within the available visibility windows.

B. Processing Constraints

Given the limited computational resources onboard small satellites, particularly CubeSats, the inference algorithm must be lightweight and optimized for constrained environments. The system should account for memory limitations, processing delays, and energy availability. The software should also be able to handle variable workloads depending on image content and orbital coverage, making it necessary to incorporate adaptive task scheduling.

C. Thermal and Environmental Considerations

Thermal dissipation becomes a significant issue when operating high-performance computing tasks in space. Since radiative heat transfer is the only viable cooling mechanism, the architecture must minimize peak processing loads or include thermal cycling strategies. In addition, the system must be resilient to the radiation environment, which can affect processing accuracy and memory integrity.

D. Power Management

All onboard systems share a limited power budget. The design must consider the need to prioritize mission-critical tasks, ensuring that image processing does not interfere with core spacecraft operations. Power-aware scheduling and energy-efficient algorithmic design are therefore essential. The operational profile should be configured to activate processing only during specific time windows or conditions, such as when passing over regions of interest.

E. Data Throughput and Communication Limitations

Another critical consideration is the limitation in data downlink capacity. As high-resolution images are data-intensive, the onboard system must reduce the data volume by extracting relevant features (e.g., bounding boxes, classifications) and transmitting only essential metadata. This requires a careful balance between onboard processing complexity and ground-based post-processing requirements.

F. Autonomy and Fault Tolerance

Due to the restricted access and command opportunities, the system must operate autonomously for extended periods. The architecture must include mechanisms for fault detection, self-recovery, and robust error handling. Additionally, software updates or reconfigurations must be supportable without direct physical access, often via remote patching or scheduled command uploads.

G. Attitude Control System (ACS) Considerations

To ensure that the camera maintains appropriate orientation for continuous maritime observation, the attitude control system must provide fine-pointing accuracy and stability. Challenges include:

- **Precise Geolocation**: Accurate pointing enables the extraction of meaningful maritime scenes and reduces geolocation errors.
- **Stability During Imaging**: ACS must prevent jitter or drift that could blur images or compromise detection accuracy.
- **Sun-Synchronous Requirements**: For consistent lighting, orbit and attitude planning must synchronize with solar positioning.

The system design must incorporate attitude-aware imaging strategies to coordinate between ACS and the payload schedule, ensuring alignment with regions of interest.

H. TT&C Constraints

Telemetry, Tracking, and Command (TT&C) subsystems must handle status reporting, command execution, and the downlink of compressed results. Constraints include:

- **Limited Bandwidth**: Only reduced metadata and selected image crops should be transmitted.
- **Prioritization Logic**: The system must decide which detections are most critical to send, based on geolocation, class, or detection confidence.
- **Command Execution Delay**: As satellite visibility windows are limited, command responsiveness may be delayed, necessitating autonomous fallback behavior.

I. Modularity and Scalability

The system should be modular to facilitate future upgrades or adaptations to new missions or sensor types. Scalability is also important, as the same architectural principles should ideally apply whether the platform is a single CubeSat or part of a distributed constellation. This enables coordinated coverage of larger maritime regions and more persistent monitoring capabilities.

V. Conclusion

This study presents a comprehensive and operationally grounded methodology for the development of a maritime vessel detection system based on satellite imagery, with a particular focus on the use of Commercial Off-The-Shelf (COTS) optical payloads onboard small satellites operating in low Earth orbit. The proposed framework integrates multidisciplinary elements—ranging from orbital analysis, sensor modeling, dataset generation and machine learning techniques—within a cohesive pipeline that maintains consistency between real-world mission constraints and training data assumptions. By emulating satellite imaging parameters using Google Earth API and augmenting the dataset with synthetically generated ship models tailored to realistic maritime scenarios. The methodology ensures representativeness of the data and robustness of the detection models. The evaluation of four distinct CNN architectures, enhanced with Feature Pyramid Networks and tested across RGB and RGB+Alpha configurations, offers valuable insights into the trade-offs between inference speed, detection accuracy, and model complexity—particularly relevant for resource-constrained satellite platforms to run those models onboard. Experimental results underscore the effectiveness of lightweight models such as Faster R-CNN with MobileNetV3, which demonstrate significant gains in performance with the incorporation of spatial priors via an alpha channel. Conversely, deeper architectures such as ResNet50 show diminishing returns or even performance degradation with additional channel information, suggesting that model capacity must be carefully matched with input complexity. Furthermore, the study addresses critical system-level considerations required for onboard deployment, including thermal management, power constraints, limited computational resources, data reduction strategies, and the need for autonomy and fault tolerance. The proposed architecture supports a modular and adaptive inference pipeline capable of real-time operation, prioritizing the detection and classification of vessels over strategic maritime regions such as the Brazilian Exclusive Economic Zone.

VI. Future Works

The advancement of autonomous ship detection and classification onboard small satellites represents a promising step toward achieving persistent maritime domain awareness, particularly over remote or under-monitored oceanic regions. While significant progress has been made in developing an onboard image processing framework capable of identifying vessels from satellite imagery, further research is required to integrate this capability with AIS data in a coherent and operationally effective manner.

AIS is a cooperative system that transmits vessel identification, location, speed, and other navigation-related data via VHF radio signals. Although widely adopted, AIS presents limitations due to intentional disabling, spoofing, or coverage gaps in certain areas, especially where ground-based AIS receivers are unavailable. The integration of an data generated by an onboard vision-based detection system with AIS receivers on the same satellite platform offers the opportunity to cross-validate observed and reported vessel positions, thus enabling the detection of non-cooperative or suspicious maritime activity.

In this context, future work will focus on establishing a data fusion framework that allows for the correlation of visual detections obtained from optical or SAR imagery with AIS broadcasts acquired in real time. This could be done primarily using the proposed hybrid architecture before deploying onboard. By accurately geolocating vessels detected through image processing and associating them with AIS-transmitted positions, the system can not only confirm ship behavior and characteristics but also flag discrepancies in reported and observed information. Ships that are visually detected but not transmitting AIS can be marked as potentially non-compliant or illicit, warranting further analysis or reporting to maritime authorities.

The integration requires the development of a software architecture capable of harmonizing heterogeneous data sources onboard the satellite. This includes the ingestion and temporal alignment of AIS data streams with satellite image acquisition events. The challenge lies in real-time or near-real-time matching of AIS signals with detected ships, especially considering uncertainties in AIS message delays, image georeferencing errors, and variations in satellite attitude. Probabilistic data association techniques, such as Kalman filters or joint probabilistic data association (JPDA), may be employed to associate AIS targets with visual detections in a dynamic maritime environment.

Furthermore, the implementation of onboard decision-making logic will be necessary to handle discrepancies in AIS-image correlation. For example, in cases where a detected vessel lacks a corresponding AIS signal within a defined spatial-temporal window, the system should autonomously classify this event as an anomaly and prioritize the image or metadata for downlink. Conversely, when a match is found, the onboard system may only transmit a condensed metadata packet summarizing the detection and AIS match, thereby conserving bandwidth. Such event-driven data transmission strategies are essential for CubeSat missions operating under stringent power and telemetry constraints.

Operational integration also involves consideration of AIS receiver capabilities onboard small satellites. Given the limitations imposed by orbital dynamics, antenna gain, and signal processing complexity, AIS receivers must be designed to ensure high-fidelity reception in low Earth orbit, even in high-density maritime zones. Additionally, regulatory compliance must be observed in terms of the use of AIS frequencies and VHF antennas in space.

The fusion of satellite-based image analytics and AIS data presents a novel and powerful tool for maritime situational awareness. Future research will explore the co-optimization of hardware and software components to achieve robust, real-time integration of these systems onboard CubeSats. This includes testing the framework in relevant orbital environments, evaluating performance over real maritime traffic, and validating the end-to-end system against ground-truth data. Ultimately, the convergence of visual detection and AIS reception aboard a single satellite platform can significantly enhance the identification, classification, and behavioral analysis of maritime traffic on a global scale.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided by the project FINEP 01.20.0194.00 (2700/20) – ITASAT 2: a space weather mission to monitor the ionosphere with three satellites flying in formation. The support from FINEP and the ITASAT 2 mission team was essential for the development and simulation activities associated with the proposed onboard maritime surveillance system.

REFERENCES

[1] B. J. Tetreault, "Use of the Automatic Identification System (AIS) for maritime domain awareness (MDA)," *Proceedings of OCEANS 2005 MTS/IEEE*, Washington, DC, USA, 2005, pp. 1590-1594 Vol. 2, doi: 10.1109/OCEANS.2005.1639983.

[2] UOW-Vessel: A Benchmark Dataset of High-Resolution Optical Satellite Images for Vessel Detection and Segmentation. ***Ly Bui, Son Lam Phung, Yang Di, Thanh Le, Tran Thanh Phong Nguyen, Sandy Burden, Abdesselam Bouzerdoum***; Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024, pp. 4428-4436

[3] Sankhe, D. and Bhosale, S. 2024. Vessel Detection in Satellite Images using Deep Learning. Engineering, Technology & Applied Science Research. 14, 6 (Dec. 2024), 18357–18362. DOI:https://doi.org/10.48084/etasr.8755.

[4] Karantaidis, Ioannis & Bereta, Konstantina & Zissis, Dimitris. (2023). A Hybrid Method for Vessel Detection in High-Resolution Satellite Imagery. 5738-5741. 10.1109/IGARSS52108.2023.10282146.

[5] Cameron Hilton, Shibin Parameswaran, Marissa Dotter, Chris M. Ward, and Josh Harguess "Classification of maritime vessels using capsule networks", Proc. SPIE 10992, Geospatial Informatics IX, 109920E (18 June 2019); https://doi.org/10.1117/12.2518775.

[6] Gülsoylu, Emre & Koch, Paul & Yıldız, Mert & Constapel, Manfred & Kelm, Andre. (2024). Image and AIS Data Fusion Technique for Maritime Computer Vision Applications. 859-868. 10.1109/WACVW60836.2024.00098.

[7] Wang, X.; Song, X.; Zhao, Y. Identification and Positioning of Abnormal Maritime Targets Based on AIS and Remote-Sensing Image Fusion. *Sensors* 2024, *24*, 2443. https://doi.org/10.3390/s24082443.

[8] Milios, Aristides & Bereta, Konstantina & Chatzikokolakis, Konstantinos & Zissis, Dimitris. (2019). Automatic Fusion of Satellite Imagery and AIS data for Vessel Detection. 1-5. 10.23919/FUSION43075.2019.9011339.