



## OPEN ACCESS

## EDITED BY

Peng Zhenming,  
University of Electronic Science and  
Technology of China, China

## REVIEWED BY

Tianlei Ma,  
Zhengzhou University, China  
Yun Zhang,  
Harbin Institute of Technology, China  
Zhengzhou Li,  
Chongqing University, China

## \*CORRESPONDENCE

Vinay Mohan

✉ [vinay.mohan@colostate.edu](mailto:vinay.mohan@colostate.edu)

Steven J. Simske

✉ [steve.simske@colostate.edu](mailto:steve.simske@colostate.edu)

## SPECIALTY SECTION

This article was submitted to  
Ocean Observation,  
a section of the journal  
Frontiers in Marine Science

RECEIVED 06 December 2022

ACCEPTED 13 February 2023

PUBLISHED 09 March 2023

## CITATION

Mohan V and Simske SJ (2023)  
Cross-sensor vision system for  
maritime object detection.  
*Front. Mar. Sci.* 10:1112955.  
doi: 10.3389/fmars.2023.1112955

## COPYRIGHT

© 2023 Mohan and Simske. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Cross-sensor vision system for maritime object detection

Vinay Mohan\* and Steven J. Simske\*

Department of Systems Engineering, Colorado State University, Fort Collins, CO, United States

Accurate and automated detection of maritime vessels present in aerial images is a considerable challenge. While significant progress has been made in recent years by adopting neural network architectures in detection and classification systems, these systems are usually designed specific to a sensor, dataset or location. In this paper, we present a system which uses multiple sensors and a convolutional neural network (CNN) architecture to test cross-sensor object detection resiliency. The system is composed of five main subsystems: Image Capture, Image Processing, Model Creation, Object-of-Interest Detection and System Evaluation. We show that the system has a high degree of cross-sensor vessel detection accuracy, paving the way for the design of similar systems which could prove robust across applications, sensors, ship types and ship sizes.

## KEYWORDS

deep learning, vessel detection system, maritime vessel, optical satellite system, object detection, convolutional neural network, synthetic aperture radar

## Introduction

From the advent of passenger ships in the late 19<sup>th</sup> century to container-revolutionized maritime transport in the 1970s, there has been increasing interest in monitoring, tracking and identifying vessels at sea. Before the first artificial earth satellite was placed into orbit in the mid 1950s, vessels were primarily tracked using either primitive cooperative systems such as inter-ship radio transmission or rudimentary non-cooperative systems such as coastal or on-board RADAR. Human interests at this time – revolving around safety & rescue, fishing and passenger transport - were largely satisfied by these systems.

More recently, effective understanding of the global maritime domain – or Maritime Domain Awareness (MDA) – has exploded in importance around the world with a significant number of commercial, defense and other government applications. There has been increasing attention given to exclusive economic zones (EEZ) and governance of a country's natural resources with state interests including maritime security, monitoring of marine traffic, illegal fishing, smuggling and maritime search & rescue. Commercial interests have expanded to include drilling and exploration of ocean floors, the management of fisheries, maritime piracy and cargo transportation. Private entities and NGOs have interests ranging from forecasting weather to the protection of ecology and sea health.

A number of these applications use knowledge of position and behavior of vessels as their cornerstone with MDA being enabled by information from land, sea, air and/or space systems and in some cases, vessel information repositories (Dekker et al., 2013). These systems can broadly be classified into one of two types – Cooperative and Non Cooperative Systems - based on whether the system is employed by vessels to communicate information about themselves or whether they are observation systems which function independently of vessel cooperation (Table 1). Information captured usually includes the vessel type, cargo, position, velocity, route as well as other identifying and tracking data.

Cooperative systems are rarely used for comprehensive MDA. Most small (<300) ton vessels are not required to carry either an Automatic Identification System (AIS) or a Long-Range Identification and Tracking System (LRIT) while fishing vessels – regardless of size – are not required to carry a Vessel Monitoring System (VMS). Additionally, illegally operating vessels rarely carry or operate their systems accurately. Some vessels turn off their systems while others spoof their mandatory position reports. Operations such as search and rescue can't be carried out effectively if one is to rely solely on cooperative reports either. These reasons make non-cooperative systems among the most beneficial sources of information for a number of the MDA applications outlined above. In particular, Synthetic Aperture Radar (SAR) and Optical Imaging Satellite Systems have several advantages such as their remote access, global reach, frequency of information updates and the high amount of data they can collect and process.

While the last century saw incremental progress made in computer-generated detection and classification of objects in images, the creation of the first convolutional neural network in the 1980s and GPU-accelerated training in the 2000s enabled significant strides in machine learning approaches to detection, segmentation and localization of objects in images.

However, a number of distinct challenges exist which prevent the robust detection of vessels at sea. Sea surfaces can be complex, and variations in weather and vessel reflectivity can lead to a loss of system precision. Small, densely packed and blurry vessels – as well as vessels very close to land – have all proven challenging for detection systems. While traditional systems are inefficient and generally have lower accuracy, modern systems have been time-consuming to build and require large amounts of labeled data.

Lastly, no single technique has proved robust across sensors, leading to piecemeal solutions for various sensors, datasets and locations.

This paper proposes a vision system which can provide robust target detection across disparate sensor types. The system is comprised of the following subsystems – Image Capture Subsystem, Image Processing Subsystem, Model Creation Subsystem, Object-of-Interest Detection Subsystem and System Evaluation Subsystem - and provides functionality for object detection using distinct independent data sources for model creation and object detection.

## Related work

LandSat-1, launched in 1972, was the first civil optical satellite. Since then, hundreds of optical satellites with varying resolutions have been launched with many continuing to orbit our planet. Recent VHR additions like the WorldView and GeoEye series have expanded spectral and spatial resolutions while others like QuickBird and IKONOS have a higher radiometric resolution as well. An increasing number of optical satellite sensors now also provide more frequent coverage of Earth. At the turn of the century, there was a significant increase in the availability of commercial VHR sensor data and with it an explosion in the number of publications exploring the viability of maritime vessel detection using satellite systems.

Some of the earliest systems for maritime vessel detection used a number of pre-processing steps prior to target detection. Sea-land separation was considered crucial for accurate detection of vessels in harbors (Willhauck et al., 2005) as well as reducing the high number of false positives generated when vessel detection systems were applied to land (Corbane et al., 2008). Consequently, coastline data was either incorporated from existing GIS data (Lavalle et al., 2011) or land masks were created from the images themselves (Dong et al., 2013). Similarly, key environmental effects – cloud coverage, waves and sunlight – were usually minimized using cloud masks (ESA, 2015), texture discrimination (Yang et al., 2014) or Fourier transform algorithms (Buck et al., 2007; Jin and Zhang, 2015).

Vessel Detection and Classification methods ranged from simple geometrical feature detection (Lin et al., 2012; Heiselberg, 2016) to machine learning techniques. Prior to the advancements made in object detection systems which used neural networks, support vector machines (SVM) – a supervised classification

TABLE 1 Cooperative vs Non Cooperative Systems.

Cooperative Systems	Non-cooperative Systems		
Examples	Type	Examples	Spectrum
Automatic Identification System (AIS)	Visual Sighting	Shore or Ship-based Sighting	Varies
Long-Range Identification & Tracking System (LRIT)	Optical Camera	Optical Imaging Satellites	Panchromatic, Multispectral
	Infrared Camera	FLIR, Weather Satellites	Thermal Infrared
Vessel Monitoring System (VMS)	Radar	Real Radar, SLAR, SAR	Electromagnetic

model – dominated publications (Bi et al., 2010; Bi et al., 2012; Kumar and Selvi, 2011; Li and Itti, 2011; Xia et al., 2011; Guo and Zhu, 2012; Satyanarayana and Aparna, 2012; Song et al., 2014). Other classifiers for vessel detection include the Bayesian classifier (Antelo et al., 2009), random forest models (Johansson, 2011) and Fisher classification (Zhang et al., 2012).

More recently, neural network-based systems have taken the world of image recognition and object recognition by storm. AlexNet, a convolutional neural network architecture designed by Alex Krizhevsky in 2012 achieved a top-5 error of 15.3%, a full 10 percentage points lower than that of the runner up and paved the way for significant strides in image classification, segmentation and object detection.

In Ramani et al., 2019 the authors build a vessel detection system for real-time maritime applications. The system employs a Mask R-CNN architecture to segment and classify 30 images every 30 seconds.

In Gallego et al. (2018), results from a Convolutional Neural Network are passed to a k-NN model to improve detection performance.

In Zhang et al. (2019), pre-processing of satellite images is performed using a support vector machine framework following which variations of the Faster R-CNN neural network architecture are applied to measure each system's performance on different sizes and types of vessels. The authors are able to identify a framework which performs reasonably well for both offshore and inland vessel detection.

Chen et al., 2020 also used CNNs to create an end-to-end detection system capable of detecting both inshore and offshore ships with an accuracy >90%. Their detection speed was 72 fps and their system intentionally balanced accuracy against speed of detection on the SAR Ship Detection Dataset (SSDD).

Li et al. (2017) used a CNN architecture-based detection system on a custom dataset consisting of ships of various sizes as well as a variety of environmental and sea conditions. Their paper established a higher precision with the custom framework than an equivalent Faster R-CNN system applied on the same dataset.

In contrast to common SAR and Optical Satellite Systems used in other publications, Yang et al. (2018) uses a remote sensing system which captured and segmented Google Earth images which were then used for vessel detection. The authors also used a custom neural network framework with a Feature Pyramid Network (FPN) to minimize false positives in images consisting of densely packed ships.

When we examine a collection of approaches used to build vessel detection systems, we observe a number of underlying trends:

- Neural networks have gained popularity in recent publications due to the largely scripted/automated approach to building highly accurate detection systems.
- Classification of vessels by vessel type has proven very challenging regardless of the type and resolution of the sensor(s) used.
- Most publications have built and tested their systems using a homogenous dataset of images collected from either a single sensor or a set of sensors, thereby failing to establish robustness of their system across sensor types.

In this paper, we tackle the challenge of building a system robust enough to collect and use images from one sensor to detect objects in images collected from a second, disparate sensor. Such a system would be tunable, adaptable and re-purposable across applications. In the Object-of-Interest Detection subsystem, we evaluate several state-of-the-art algorithms as well as create a custom model architecture from scratch.

While we do not intend to recommend a winning algorithm to solve cross-sensor vessel detection, we show that the designed system has a high degree of cross-sensor vessel detection accuracy, paving the way for future research in tunable, adaptable and re-purposable systems which could prove robust across applications, sensors, ship types and ship sizes.

## System overview

### Image capture subsystem

The Image Capture Subsystem (Figure 1) uses two satellite sensor feeds along with two XML file feeds to obtain and provide data to the consequent subsystems. The XML files contain annotated image information for the corresponding image feeds.

The first input is an optical aerial image feed of maritime scenes on the visible spectrum. The images are sourced in the RGB color scheme, and can contain zero, one or multiple maritime vessels in varying weather and lighting conditions. The images contain scenes from different regions of the world including Africa, Europe and Asia and different water bodies including the Mediterranean Sea as well as the Atlantic and Pacific Oceans. While the images are of different sizes, the average image has a spatial resolution of 512 x 512 pixels.

The second input is a synthetic aperture radar (SAR) generated feed of maritime scenes (sea waves, shallow sea topography, coastal zones, maritime vessels etc.) with a spatial resolution between 1-500m. This feed provides images of 256 pixels in both range and azimuth, and the vessels in these images have distinct scales and backgrounds. A given image can contain a single vessel, multiple vessels or none.

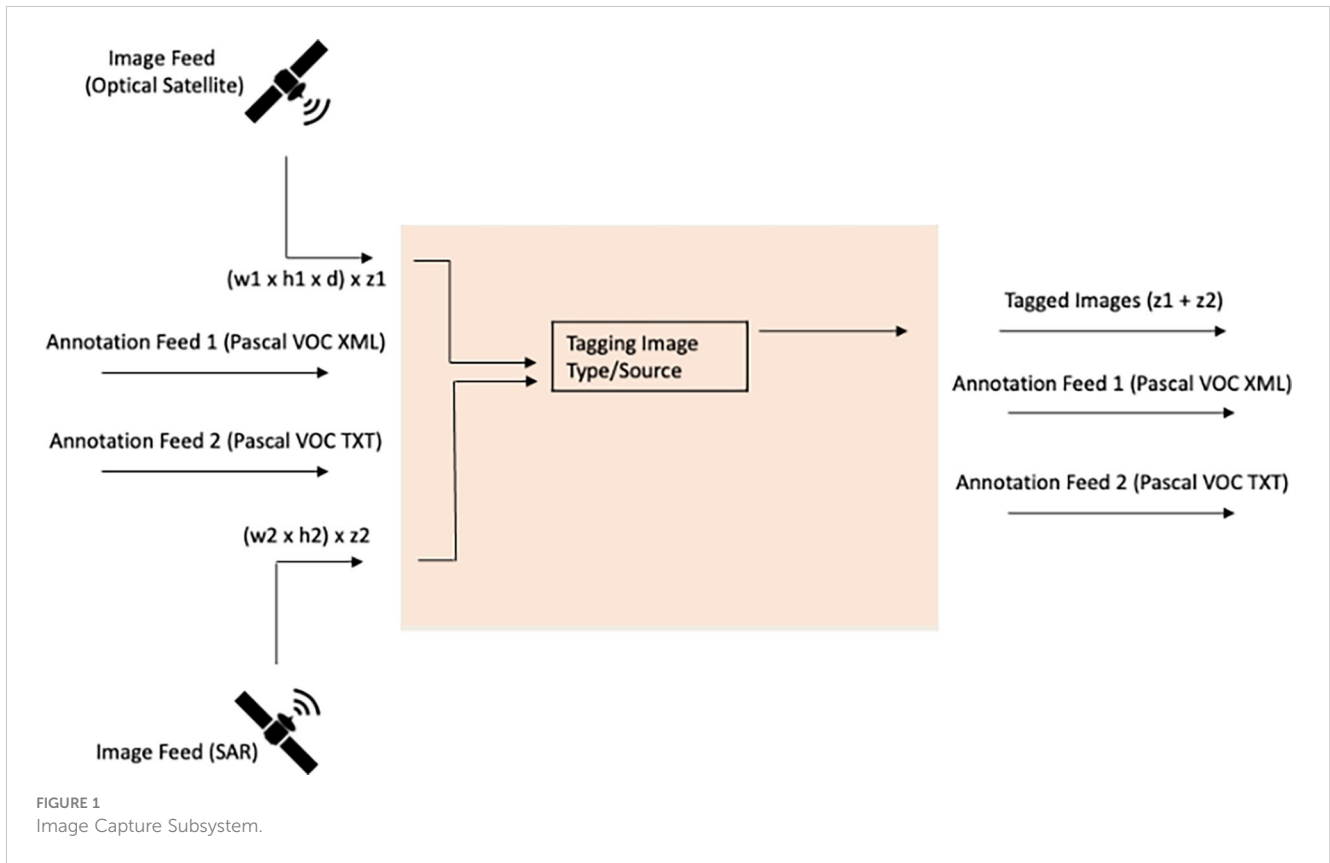
For each image feed, annotations are provided in the Pascal VOC format. The Pascal VOC format is a common annotation format for images which stores annotations in the XML file format with a separate XML annotation file for each image. Optionally, bounding box information is included in the [x-top-left, y-top-left, x-bottom-right, y-bottom-right] format.

The image feeds are tagged with their source before being merged together into a single stream and sent to the Image Processing Subsystem. The two streams of XML annotations comprise the other outputs of this system.

### Image processing subsystem

The inputs to the Image Processing Subsystem (Figure 2) are

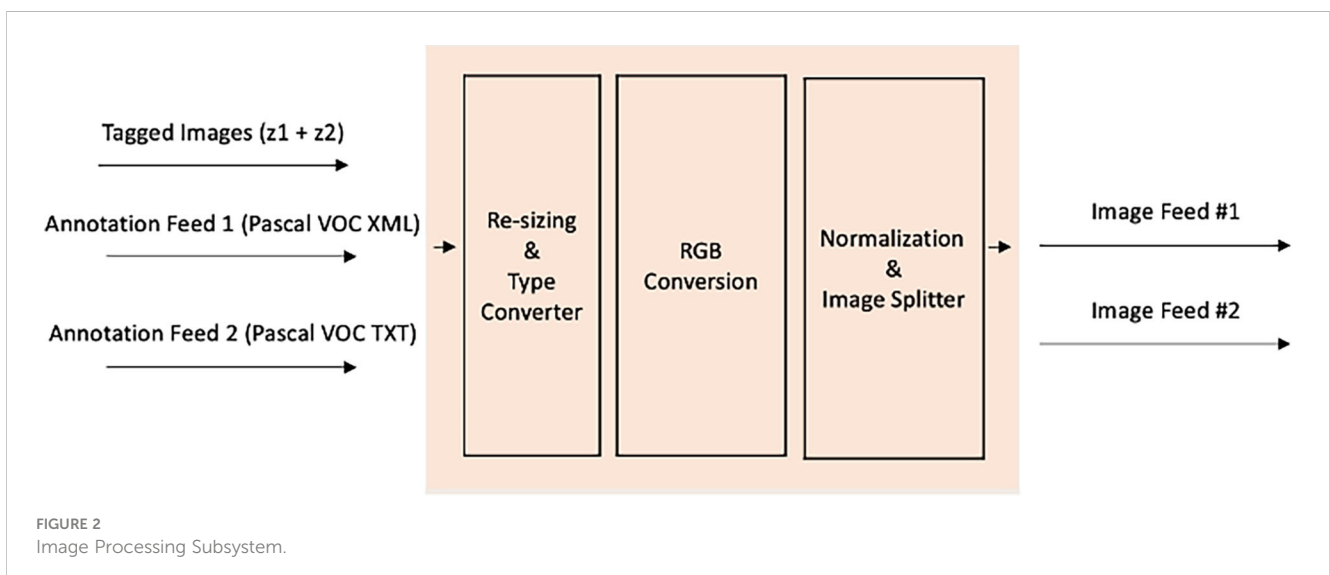
- a single stream of images tagged with their source, &

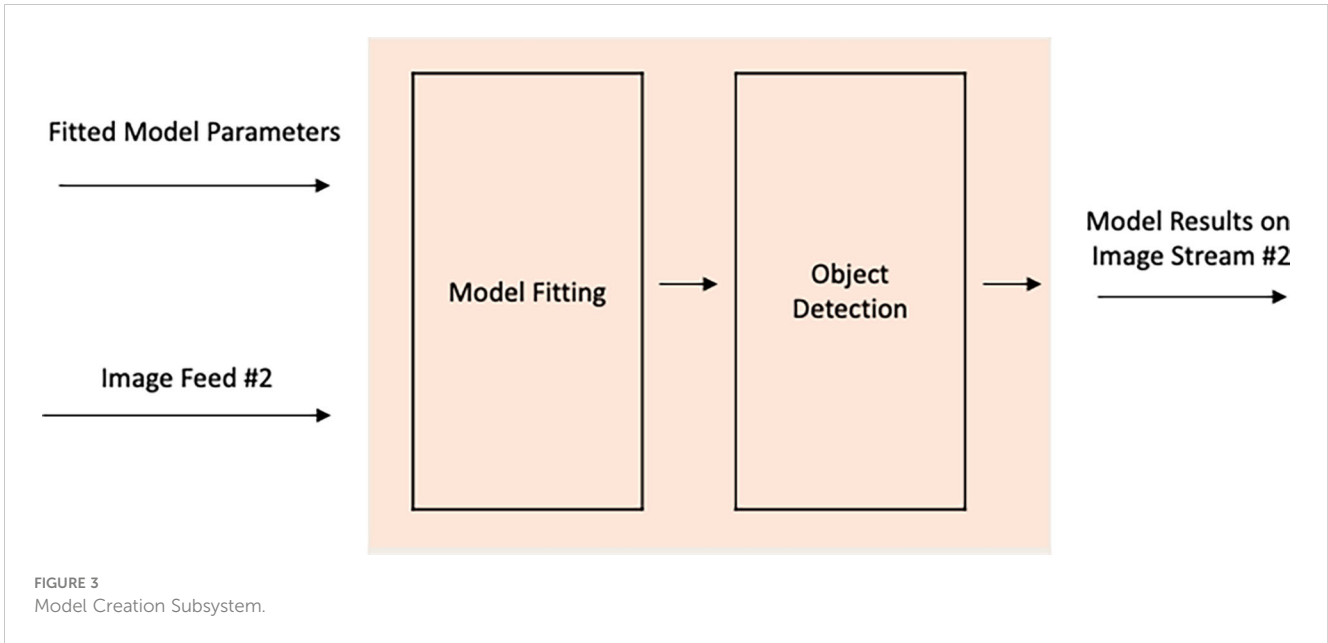


-two annotation feeds corresponding to the respective image streams.

First, the images are re-sized for uniformity across input streams and to match the dimensions of the input layer in the Model Creation Subsystem. The pixels in the image stream are then converted to the float datatype following which each image is normalized. Normalization scales the pixel values down from a

range of (0,255) to a range of (0,1). Lastly, the image streams are split based on their source, annotations are appended and the output of the subsystem consists of two tagged and annotated image streams. Convolutional Neural Network Architectures like AlexNet and GoogleNet perform various image chopping and feature extraction steps which, in conjunction with pooling layers make them translation (and to a large degree, rotation) invariant model architectures.



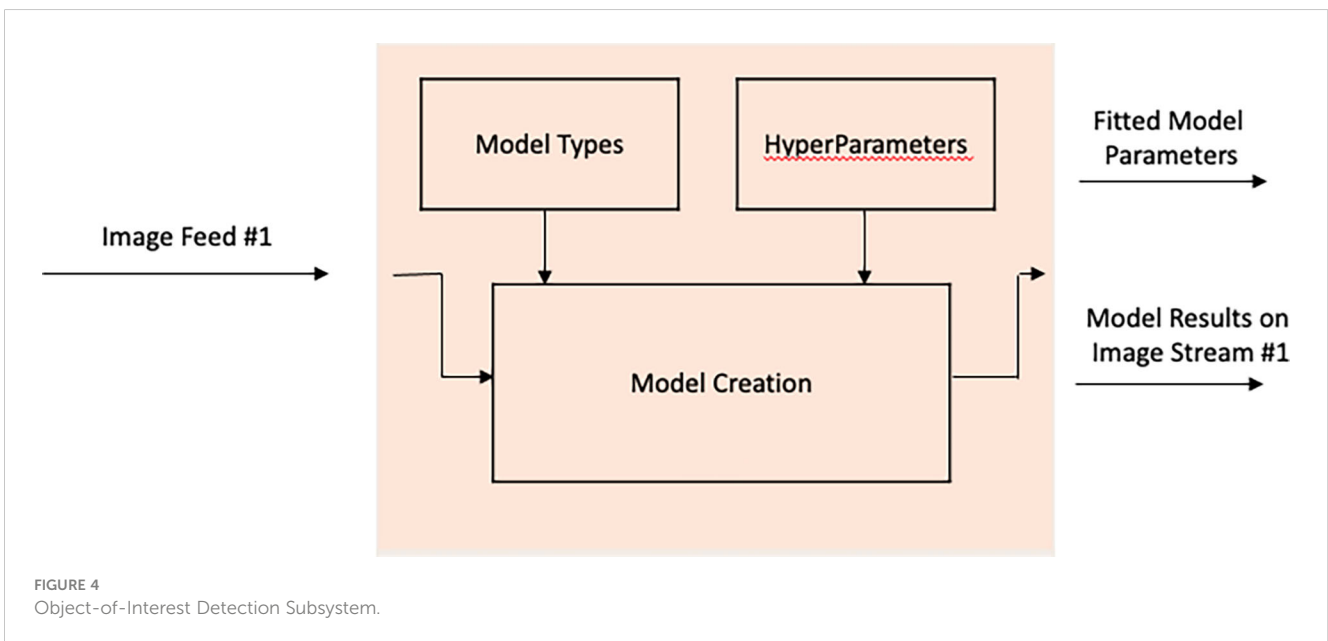


### Model creation subsystem

The input to the Model Creation Subsystem (Figure 3) is a single annotated image feed. This feed is used to train a binary classification model to detect the presence of a vessel in an image using a combination of pre-defined model frameworks and hyperparameters. The models employed are (a) a custom convolutional neural network architecture, defined and trained from scratch, and (b) transfer learning and benchmarking using four common computer vision model architectures. For the latter, we re-define and fine-tune the last layers for our specific task while leaving the architecture and weights of other layers as is. Model parameters for each fitted model comprise the output of the Model Creation subsystem as well as each model’s predictions on the input image feed.

### Object-of-interest detection subsystem

The inputs to the Object-of-Interest Detection Subsystem (OOIDS) (Figure 4) are the fitted model parameters and the second image feed on which OOI Detection is to be performed. The fitted model parameters can either be the hyperparameters of the model – in which case the model will need to be re-fit on the original dataset – or a fit model, as we have assumed here. The model is applied (‘scored’) on the second image stream producing predictions indicating the presence or absence of maritime vessels. The output of this subsystem are the model results on the second image stream. As a reminder, this is an image feed the model itself has not been exposed to, and is an attempt to measure the model’s power on a disparate and independent data source.



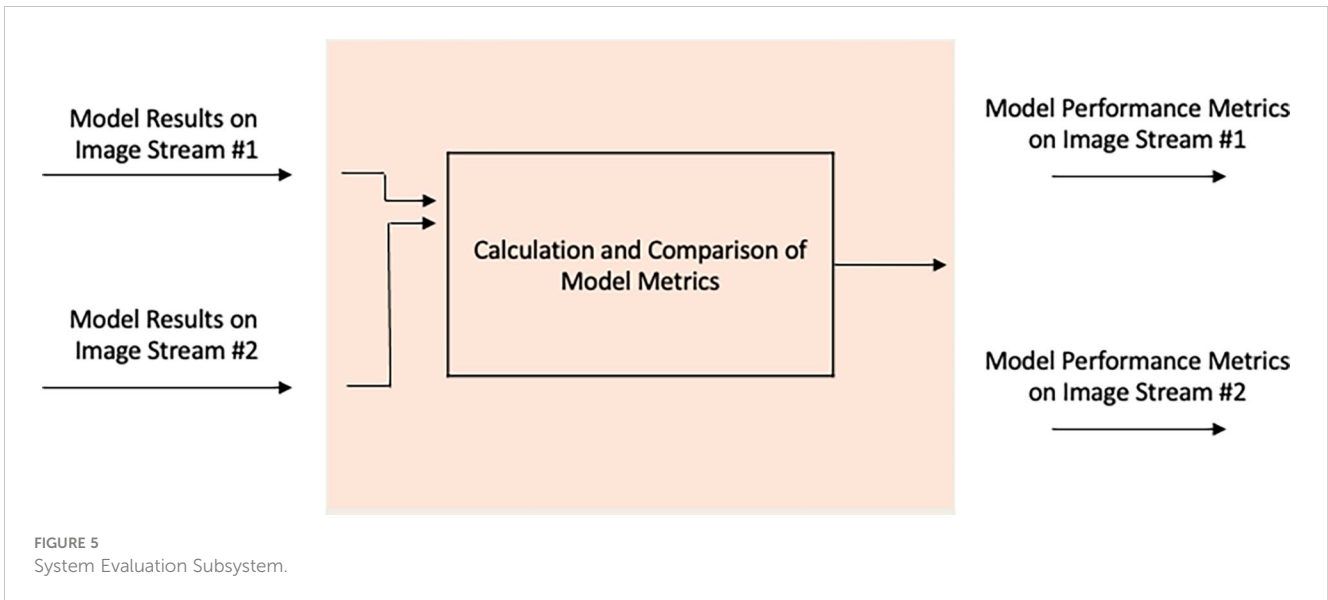


FIGURE 5 System Evaluation Subsystem.

### System evaluation subsystem

The System Evaluation Subsystem (Figure 5) calculates and produce model metrics which measure the performance of the model on the dependent (‘training’) and independent (‘test’) data sources. The inputs to this subsystem are the model results on both the training (image feed #1) and test (image feed #2) datasets. Using these model results, model metrics such as accuracy, precision and recall can be calculated, dependent on the number and frequency of classes in each dataset. The metrics indicate the overall performance of the system at performing object detection using different sensors and sensor types, i.e SAR and Optical Satellite sensors.,,,,

### Methodology

The System Block Diagram is shown in Figure 6.

To test system functionality and gauge performance, we use the MASATI (Maritime Satellite Imagery Dataset) and Sentinel datasets as

inputs to the Image Capture Subsystem. The images contained in these datasets contain maritime scenes in the visible spectrum using optical aerial cameras and SAR-based radio waves, respectively. These datasets mimic and satisfy the earlier outlined assumptions regarding the two satellite image feeds (III.A) and are accompanied by annotations indicating the presence/absence of maritime vessels which are treated as ground truth in the subsequent model design and evaluation subsystems.

The datasets are tagged with their source name, re-sized to standardized dimensions, normalized and the bits converted to the float datatype. The Model Creation Subsystem uses the Keras Deep Learning API with the Tensorflow backend to fit four pre-defined convolutional neural network architectures and one custom architecture on the MASATI dataset. The MASATI dataset consists of 1027 (48%) images containing one or more maritime vessels and 1132 (52%) images with none. In addition to a custom model trained from scratch, the 4 pre-defined architectures include

-VGG-16, proposed by Karen Simonyan and Andrew Zisserman of Oxford University in 2014, the ‘16’ in the

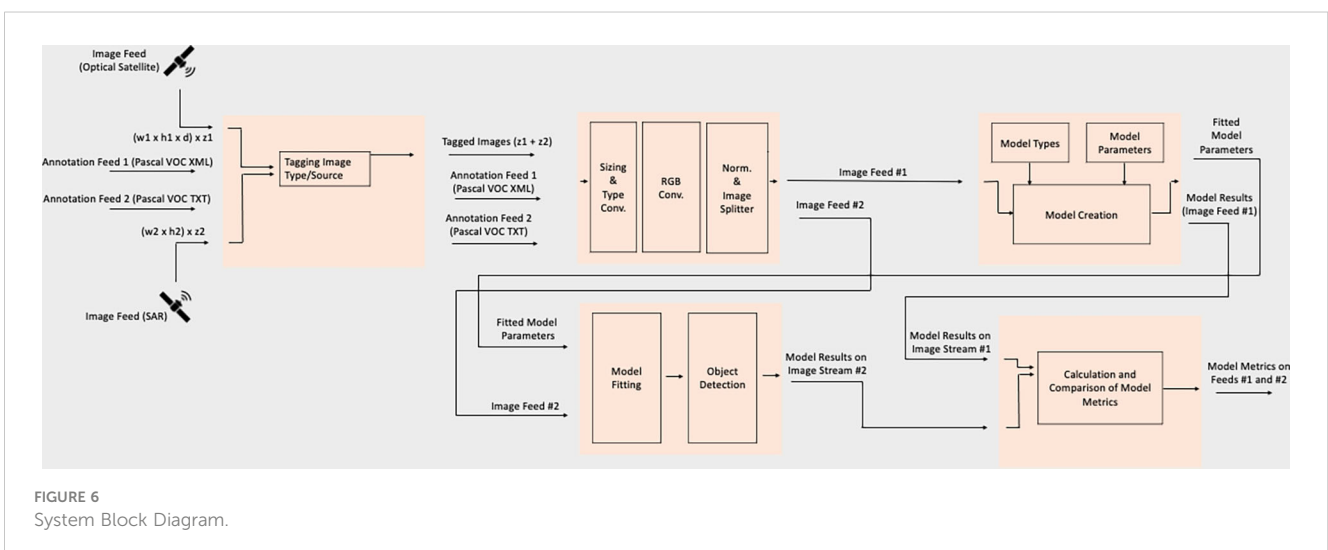


FIGURE 6 System Block Diagram.

name indicating the number of layers with weights (Simonyan and Zisserman, 2014).

- InceptionV3, originally a module for GoogleNet in 2015 (Szegedy et al., 2016)
- ResNet50, a variant of the ResNet Model consisting of 48 convolution layers and residual blocks, introduced in 2015, and
- Xception, a deep convolutional neural network architecture involving Depthwise Seperable Convolutions, introduced by Francois Chollet in 2016

Since (a) one of the primary goals behind most maritime object detection systems is real-time processing, and (b) our primary goal is to develop a system capable of using data from one sensor to detect objects in incoming data from a second sensor, each of the 5 models is trained for  $\leq 5$  epochs. There is no minimum stopping threshold or other optimization criteria since we want flexible models which aren't overfit or optimized on the MASATI dataset alone.

While the custom model is trained from scratch, each pre-defined architecture has the following changes:

- The input layer is altered to match the dimensions of the incoming data stream,
- The output layer is altered to a softmax function with two classes of interest, and
- while the fitted weights of most layers stay the same, the last five layers are re-trained for the purpose of optimizing detection of our classes of interest

Each model is trained using specific values of hyperparameters following which fitted parameter values are saved and transferred as outputs to the Object-of-interest Detection Subsystem. The OOI Detection Subsystem re-fits models and uses the fit models to predict the presence or absence of a maritime vessel in the second

(SAR) input stream. The results of the 5 models on the SAR image stream is an input to the System Evaluation Subsystem which calculates, compares and displays metrics for the system's user(s).

Model Configuration and hyperparameters for each model are shown in Table 2.

## Results and discussion

The System Evaluation Subsystem calculates each model's accuracy in detecting ships on the two datasets – MASATI and Sentinel, referred to as the training data and test data, respectively. The number of images in each dataset as well as the time to train and score each model on the respective datasets is also calculated. These results are shown in Table 3.

As we can see, the results are interesting and varied.

- While the custom model – trained from scratch – has a low accuracy, recall and F-Score, it has high precision and beats larger architectures like ResNet50 across the board when trained for only a few epochs.
- ResNet50, as we can see in Table 2 also has the highest number of parameters of all the architectures indicating that the extra learning potential of this network likely requires additional parameter tuning and in its current form results in overfitting on the training data.
- Most pre-trained models performed better than the custom model indicating that the extra layers, learning capacity and learned features in these models aided in our binary classification task, despite being designed for larger and more complex image classification and object detection tasks. In addition to the higher F-Scores, InceptionV3 and Xception have much lower training times than the custom architecture.
- Despite the datasets being collected from different sensors and sensor types, many of the models are successfully able to

TABLE 2 Parameters for Models 1-5.

Architecture	Custom	VGG-16	InceptionV3	ResNet50	Xception
Train-Val Ratio	N/A (Train = 1.0)	N/A (Train = 1.0)	N/A (Train = 1.0)	N/A (Train = 1.0)	N/A (Train = 1.0)
Image Width	200	200	200	200	200
Image Height	200	200	200	200	200
Number of Images	2,159	2,159	2,159	2,159	2,159
Number of Classes	2	2	2	2	2
Number of Epochs	5	5	1	1	1
Learning Rate	0.001	0.001	0.001	0.001	0.001
Decay Factor	0.9	0.9	0.9	0.9	0.9
Optimizer	RMSProp	RMSProp	RMSProp	RMSProp	RMSProp
Total Parameters	1,483,010	14,751,554	21,868,322	23,735,298	21,062,186
Trainable Parameters	1,483,010	2,396,674	65,538	200,706	204,802

TABLE 3 System Evaluation Subsystem Results.

		<i>Custom Model</i>	<i>VGG-16</i>	<i>InceptionV3</i>	<i>ResNet50</i>	<i>Xception</i>
Training Data	<i>Number of Images</i>	2,159	2,159	2,159	2,159	2,159
	<i>Time to Train</i>	88 seconds	536 seconds	31 seconds	65 seconds	57 seconds
	<i>Training Data - Overall Accuracy</i>	0.92	0.99	0.91	0.56	0.96
Test Data	<i>Number of Images</i>	19,500	19,500	19,500	19,500	19,500
	<i>Time to Score</i>	49 seconds	918 seconds	208 seconds	576 seconds	482 seconds
	<i>Precision (Ship)</i>	0.73	0.83	0.97	0.46	0.72

identify maritime vessels in one using data solely from the other with both high precision and recall despite the unbalanced Sentinel dataset.

- While most modern systems built on underlying neural network architectures require sufficiently large (a) computing power, (b) time, and/or (c) data to perform well, a dataset of ~2K images was used to sufficiently capture between 62% - 94% of vessels in a dataset 10x as large (21,682 images) with training and testing times of <=10 minutes.
- While many systems require significant tuning and selection of hyperparameters to optimize object detection, limited fine-tuning resulted in respectable vessel detection results.

While we have not examined incorrect classification results further to discover potential underlying trends, future research in cross-sensor vessel detection could prove robustness across ship and sensor types with longer training times, other model architectures and/or further hyperparameter tuning.

We propose the following guidelines which similar studies could consider:

- Using multiple sensors for both system training and testing
- Verification of algorithms on varied maritime scenes
- Validation of accuracy and false detection rates across different ship sizes and difficult conditions
- Introduction of ship classification algorithms for specific applications

Given that earth observation is a rapidly growing field with an increasing availability of open data and new satellite technology, cross-sensor vessel detection systems would be more adept than traditional systems and could prove less cost-sensitive for new applications. Future research using small datasets and low system

processing times may also lead to rapid detection rates, thereby aiding real-time maritime applications including safety, logistics and transportation.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

## Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

Antelo, J., Ambrosio, G., Gonzalez, J., and Galindo, C. (2009). "Ship detection and recognition in high-resolution satellite images. In 2009 IEEE International Geoscience and Remote Sensing Symposium (IEEE) 4, 514–517. doi: 10.1109/IGARSS.2009.5417426

Copernicus Open access hub. Available at: <https://scihub.copernicus.eu/>.

*Sentinel data access overview*. Available at: <https://sentinel.esa.int/web/sentinel/sentinel-data-access>.

*U.I for computer research - the MASATI dataset*. Available at: <https://www.iuii.ua.es/datasets/masati/>.



- Bi, F., Liu, F., and Gao, L. (2010). "A hierarchical salient-region based algorithm for ship detection in remote sensing images," in *Advances in neural network research and applications*. Eds. Z. Zeng and J. Wang (Berlin Heidelberg: Springer), 729–738.
- Bi, F., Zhu, B., Gao, L., and Bian, M. (2012). A visual search inspired computational model for ship detection in optical satellite images. *IEEE Geosci. Remote Sens. Lett.* 9, 749–753. doi: 10.1109/LGRS.2011.2180695
- Buck, H., Sharghi, E., Bromley, K., Guilas, C., and Cheng, T. (2007). "Ship detection and classification from overhead imagery," in *Applications of Digital Image Processing XXX (SPIE)* 6696, 522–536. doi: 10.1117/12.754019
- Chen, Y., Duan, T., Wang, C., Zhang, Y., and Huang, Mo (2020). End-to-End ship detection in SAR images for complex scenes based on deep CNNs. *J. Sensors* 2021, 8893182, 19. doi: 10.1155/2021/8893182
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1251–1258.
- Corbane, C., Pecoul, E., Demagistri, L., and Petit, M. (2008). "Fully automated procedure for ship detection using optical satellite imagery," in R.J. Frouin, S. Andrefouet, H. Kawamura, M.J. Lynch, D. Pan, T. Platt, et al *Remote Sensing of Inland, Coastal, and Oceanic Waters (SPIE)* 7150, 13. doi: 10.1117/12.805097
- Dekker, R. J., Bouma, H., Breejen, E., den, Broek, A. C., van den, Hanckmann, P., Hogervorst, M. A., et al. (2013). "Maritime situation awareness capabilities from satellite and terrestrial sensor systems," in *Proc. Maritime Systems and Technologies MAST Europe*.
- Dong, L., Yali, L., Fei, H., and Shengjin, W. (2013). "Object detection in image with complex background," in *3rd International Conference on Multimedia Technology (ICMT-13)* (Atlantis Press.), 471–478. doi: 10.2991/icmt-13.2013.58
- ESA. (2015). Sentinel-2 User Handbook. 2015. Available at: [https://scholar.google.com/scholar\\_lookup?title=Sentinel-2%20User%20Handbook&author=ESA&publication\\_year=2015](https://scholar.google.com/scholar_lookup?title=Sentinel-2%20User%20Handbook&author=ESA&publication_year=2015).
- Galleo, A.-J., Pertusa, A., and Gil, P. (2018). Automatic ship classification from optical aerial images with convolutional neural network. *Remote Sens.* 10 (4). doi: 10.3390/rs10040511
- Guo, J., and Zhu, C. R. (2012). A novel method of ship detection from spaceborne optical image based on spatial pyramid matching. *Appl. Mech. Mater.* 190–191, 1099–1103. doi: 10.4028/www.scientific.net/AMM.190-191.1099
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778. doi: 10.1109/CVPR.2016.90
- Heiselberg, H. (2016). A direct and fast methodology for ship recognition in sentinel-2 multispectral imagery. *Remote Sens* 8. doi: 10.3390/rs8121033
- Jin, T., and Zhang, J. (2015). "Ship detection from high-resolution imagery based on land masking and cloud filtering," in *Seventh International Conference on Graphic and Image Processing (ICGIP 2015)* (SPIE) 9817, pp. 268–273. doi: 10.1117/12.2228219
- Johansson, P. (2011). "Small vessel detection in high quality optical satellite imagery," in *Tech. report chalmers university of technology sweden. JR* (European Commission - Joint Research Centre, Ispra, Italy: Integrated Maritime Policy for the EU, Working Document III on Maritime Surveillance Systems).
- Kumar, S. S., and Selvi, M. U. (2011). Sea Object detection using colour and texture classification. *Int. J. Comput. Appl. Eng.* 1, 59–63.
- Lavalle, C., Rocha Gomes, C., Baranzelli, C., and Batista e Silva, F. (2011). *Coastal zones: Policy alternatives impacts on European coastal zones 2000–2050* (Joint Research Centre, Institute for Environment and Sustainability: European Commission).
- Li, Z., and Itti, L. (2011). Saliency and gist features for target detection in satellite images. *IEEE Trans. Image Process.* 20, 2017–2029. doi: 10.1109/TIP.2010.2099128
- Li, J., Qu, C., and Shao, J. (2017). Ship detection in SAR images based on an improved faster R-CNN. In *2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA)* (IEEE), 1–6. doi: 10.1109/BIGSAR DATA.2017.8124934
- Lin, J., Yang, X., Xiao, S., Yu, Y., and Jia, C. (2012). "A line segment based inshore ship detection method," in *Future control and automation*. Ed. W. Deng (Berlin Heidelberg: Springer-Verlag), 261–269.
- Ramani, S., Prabakaran, N., Kannadasan, R., and Rajkumar, S. (2019). Real time detection and segmentation of ships in satellite images. *Int. J. Sci. Technol. Res.* 8 (12).
- Satyanarayana, M. S., and Aparna, G. (2012). A method of ship detection from spaceborne optical image. *Int. J. Adv. Comput. Math. Sci.* 3, 535–540.
- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv 1409.1556*.
- Song, Z., Sui, H., and Wang, Y. (2014). "Automatic ship detection for optical satellite images based on visual attention model and LBP," in *2014 IEEE Workshop on Electronics, Computer and Applications*. 722–725. doi: 10.1109/IWECA.2014.6845723
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. B. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2818–2826. doi: 10.1109/CVPR.2016.308
- Willhauck, G., Caliz, J. J., Hoffmann, C., Lingenfelder, I., and Heynen, M. (2005) Object oriented ship detection from VHR satellite images (Accessed 6th Geomatic Week Conference).
- Xia, Y., Wan, S., and Yue, L. (2011). "A novel algorithm for ship detection based on dynamic fusion model of multi-feature and support vector machine," in *6th International Conference on Image and Graphics (ICIG)*. 521–526. doi: 10.1109/ICIG.2011.147
- Yang, G., Li, B., Ji, S., Gao, F., and Xu, Q. (2014). Ship detection from optical satellite images based on sea surface analysis. *IEEE Geosci. Remote Sens. Lett.* 11, 641–645. doi: 10.1109/LGRS.2013.2273552
- Yang, X., Sun, H., Fu, K., Yang, J., Sun, X., Yan, M., et al. (2018). Automatic ship detection in remote sensing images from Google earth of complex scenes based on multiscale rotation dense feature pyramid networks. *Remote Sens.* 10, 132. doi: 10.3390/rs10010132
- Yang, X., Sun, H., Fu, K., Yang, J., Sun, X., Yan, M., et al. (2011). A sea-land segmentation scheme based on statistical model of sea. *4th International Congress on Image and Signal Processing. CISP 2011*, 1155–1159. doi: 10.1109/CISP.2011.6100503
- Zhang, W., Bian, C., Zhao, X., and Hou, Q. (2012). "Ship target segmentation and detection in complex optical remote sensing image based on component tree characteristics discrimination," in *Proc. SPIE 8558, Optoelectronic Imaging and Multimedia Technology II*, 85582F, pp. 9. doi: 10.1117/12.2000688
- Zhang, S., Wu, R., Xu, K., Wang, J., and Sun, W. (2019). R-CNN-Based ship detection from high resolution remote sensing imagery. *Remote Sens.* 11 (6). doi: 10.3390/rs11060631