Deep Learning

-- Project proposal --

A Comparison of Model Performances on Noisy SAR Data

Group 2

Germann Jakob
Dzhabarov Jafar
Maayan Janow
Isiaka Afolabi
Jared Heidt

DAAN 570 – Deep Learning
Spring Semester, 2025

TABLE OF CONTENTS

1.0 - Introduction

Synthetic Aperture Radar (SAR) is a fundamental technology for maritime surveillance with the advantage of acquiring high-resolution images independent of weather and daylight conditions. Such a characteristic renders SAR indispensable for time-critical applications like search-and-rescue operations, detection of illicit activities, and environmental monitoring. SAR images, nevertheless, are ubiquitously contaminated by speckle, clutter, and environmental interference, consequently rendering the accurate detection of maritime vessels challenging. Conventional computer vision techniques are overwhelmed by these difficulties, calling for sophisticated deep learning techniques to automate and improve the accuracy of detection.

The project aims at a systematic evaluation of the performance of state-of-the-art object detection and segmentation models under noisy SAR conditions. By experimenting with models like Faster R-CNN, YOLO variants, Graph Neural Networks (GNNs), in combination with various image masking methods, we will identify the most robust framework for detecting vessels from false targets and clutter. The findings will provide actionable insights into improving maritime safety, security, and operational effectiveness, while advancing methodologies for SAR data analysis under resource-scarce, high-noise settings.

2.0 - Problem Statement

Our objective for this study is to conduct a comprehensive analysis and comparison of the performances of various object detection models. Given a random Maritime Synthetic Aperture Radar Image, each model will be able to perform the following tasks:

A. Identify one or more maritime vessels within an image regardless of shape or size.
B. Distinguish between targets of interest and false targets and radar clutter.
C. Be capable of identifying targets of interest in images with high noise.

For this study, there are several issues that must be considered when building each model. When dealing with radar technology, it is expected to encounter issues both in collecting data and in processing it. Collecting radar data requires expensive technology and resources that few organizations can afford. Thus, we are limited in the amount of data and variability to what is given to us in our chosen dataset.

Processing radar data will bring additional issues that can negatively impact a model's performance. SAR image data often contains high amounts of noise surrounding targets of interest, be it from natural or man-made phenomena. These phenomena will reduce the quality of our data and the models' performance. Lastly, our objective is to identify more than one vessel in a single image. Due to the computation costs, SAR images often have their resolutions scaled down significantly, especially on systems with limited memory and computation power. This in turn reduces the detail of targets and makes distinguishing them far more difficult, particularly when in close proximity to each other.

Regardless of the aforementioned difficulties, using deep learning to identify targets in radar images is incredibly useful. The advantage of radar is that it can operate and collect target information in nearly any environment, regardless of the environment state. Having artificial intelligence models to detect anomalies in radar images can assist various tasks such as:

1. Search and rescue operations for lost ships.
2. Identifying ships engaging in illegal activity such as smuggling, piracy, etc.
3. Helping vessels plan safe routes through unpredictable seas.
4. Providing details on environmental impacts due to maritime incidents.

3.0 - Challenges

---

1. **Limited and Expensive SAR Data**: Acquiring data by SAR requires specialized equipment and extensive processing, yielding compact, non-diverse data samples. Limited data restricts generalization capacity, enhancing overfitting threats.
   Solution: Apply data augmentation strategies (e.g., rotation, and noise addition) and transfer learning from pre-trained models over analogous domains (e.g., optical imagery from satellites). Apply synthetic data generation models such as GANs to enhance variability across the data.
2. **High Noise and Background Clutter in SAR Imagery**: Speckles (Speckle noise) and ground clutter (e.g., waves, shore buildings) hide vessel signs, lowering the accuracy of vessel identification.
   Solution: Incorporate steps for data denoising (e.g., denoising by the discrete wavelet transform, Canny edge detection) beforehand. Apply attention models in CNN architectures (e.g., CBAM) that pay close attention while discarding unrelated pixels.
3. **Distinguishing Overlapping and Adjacent Targets**: SAR resolution limits adjacent vessels from appearing as single entities, hindering vessel instance segmentation.
   Solution: Use sophisticated segmentation models (Mask R-CNN, Hybrid Task Cascade network) that combine the detection and per-pixel masking. Apply GNNs' spatial-temporal analysis capabilities to track vessel patterns and resolve ambiguities.
4. **Identifying False Targets**: It still proves tricky to tell vessels apart from deceptions (decoys, e.g., buoys, floatables).
   Solution: Train models over multi-mode data (fusion of SAR data and AIS messages) so that context accompanies the identification. Apply ensemble models so that classification and detection models' outputs are mixed, enhancing decision stability.

4.0 - Related works

Several studies have explored object detection and segmentation in maritime environments using Synthetic Aperture Radar (SAR) images, given their importance in surveillance, navigation, and defense applications. Some of the important work in this topic include the following:

1. **Deep Learning for Maritime Vessel Detection**: Researchers have employed convolutional neural networks (CNNs) for vessel detection in SAR images. A study that was published in *Nature* developed modified versions of YOLOv4 and YOLOv5, object detection models that use CNNs as their backbone in order to detect ships from SAR imagery [1], even though the study demonstrated success in detecting maritime vessels it also highlighted challenges with small or obscured targets in noisy environments.
2. **Multi-Task Learning in Maritime Applications**: Some studies have developed unified models that combine object detection and segmentation, using shared features to boost performance, but these models still struggle with challenges like complex or highly variable vessel shapes and sizes. One study in *Frontiers in Marine Science* [2] explored instance segmentation in SAR images, blending semantic segmentation with object detection to tackle these complexities more effectively.
3. **Faster Region Based Convolutional Neural Networks** - Faster RCNN is aimed at solving the challenged traditional CNN models present when trying to detect multiple objects in an image, and their inefficiencies. RCNN region proposals through selective search algorithm which helps with much faster detection and a much higher precision in multi object detection. In our case, since we will be working with images that contain multiple objects and quite a noisy background, RCNN might be superior to the aforementioned Yolo based object detection.
4. **Graph Neural Networks**: One research group at the U.S. Army Research Laboratory found that using GNNs yielded comparable performance to Convolutional Neural Network models on low complexity SAR image data. [3] Figures in this paper show that little to no noise was within the SAR image data, not accurately reflecting the common problems encountered with such systems. Additionally, since Convolutional Neural Networks process the entire image for classification, this can take significant time and even skew the model's performance on images that contain few relevant pixels. The Graph Neural Network they used allowed them to bypass pixels with weak relationships to the target of interest and reduce computation time. Additionally, GNN model performance averaged 99.1% across both datasets, with only the Tai-SARnet achieving marginally better accuracy.
5. **GSDT**: Another group demonstrated state-of-the-art performance for multiple-object detection and tracking using Graph Neural Networks, a model called GSDT. [4] They proposed a multi-input-single-output approach that allowed the model to utilize data across the spatial and temporal domains. To reinforce relationships between individual objects in an image and predictions, the model will intake a current image and a prior variant that it made predictions on, if applicable. It will then use those previous predictions to generate feature vectors and use a sub-module called CenterNet to detect each object within the image. After detection, associations will be made at the pixel level so the model may learn an "identity embedding" for each detection at the current timestep. The group demonstrated significantly better performance in accuracy compared to several pre-proposed models, achieving a 60.7% on the 2DMOT2015, 74.5% on MOT16, 73.2% on MOT17, and 67.1% on MOT20.

Gaps in Existing Works and Research

- Many models are optimized for general object detection but lack specialization for maritime environments.

- There is limited exploration of approaches that can reliably differentiate between vessels and false targets in cluttered, noisy SAR images.
- Few studies systematically compare object detection and segmentation models on a common maritime SAR dataset, making it challenging to identify the best-performing techniques for specific scenarios.

5.0 - Data description

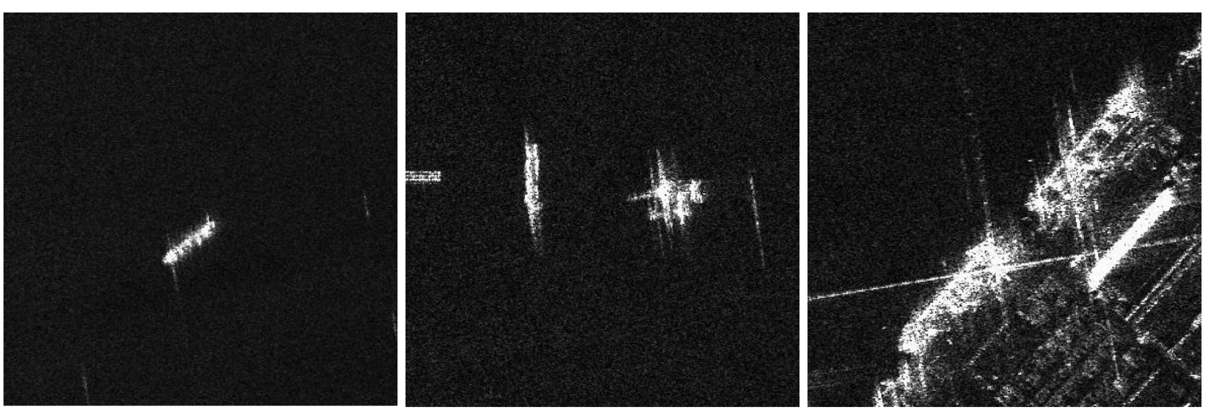
5.1 - Data Statistics

Our dataset consists of 6735 Synthetic Aperture Radar images across training, validation, and testing. Each image contains information representative of a typical airborne radar system, the dataset containing 17708 instances of various ship types and sizes. Due to the difficult nature inherent in processing SAR Image data, each image was resized to 640x640 for faster processing and training.

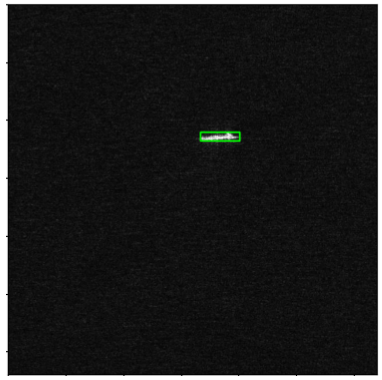
5.2 - Image Types and Annotations

Each image within the dataset contains one or more instances of an unknown type of maritime vessel. These images are set on a grayscale color system, with bright pixels representing an area of high reflectivity, implying that area contains a target. Dark pixels represent areas of low reflectivity, or areas with no targets or noise caught by the radar system.

The bottom left image is an example of a single ship target, represented by the bright pixels, out at sea, represented by the dark pixels. The middle image is an example of two targets within the image. The rightmost image demonstrates ship targets adjacent to irrelevant targets such as buildings, vehicles, and loading dock personnel. This will be one of our greatest challenges to help the model avoid detecting these confusing objects.



The data is also annotated to distinguish individual targets. Each image corresponds to a single annotation in its respective set's JSON file. An example of an annotated target is shown below.



These annotations will help the model also distinguish between targets of interest and confuser targets such as buildings, collections of noise, and other similarly sized irrelevant targets.

6.0 - Deep Learning Methodology

6.1 - Baseline

Since this is an image classification and a multi-object detection problem, we will be using supervised learning, as we will need to use labeled data to train our model.

The deep learning methodology we are suggesting is called Faster RCNN (Region based convolutional neural network) which is well known for its precision and performance, specifically with SAR based images. The idea is to use this model and optimize it with different numbers of hidden layers and experiment with different optimizers and learning rates until we get the desired results.

RCNN utilizes a selective search algorithm to approximately generate 2000 different suggestions for regions where objects can be found. This approach is especially useful in our scenario since our images will include multiple objects. Once the objects are detected, the extracted features are classified and boxes are formed around the suggested object or categorized as background or "noise". The boxes (multiples around each object) around the objects then become progressively more accurate until all the segments get combined into one, so that each object has an accurate box around it. This segmentation and localization process makes for a very accurate, multi-object detection model.

To use RCNN we will first need to select a Pre-trained model. AlexNet model is commonly used with RCNN. We will also need to adjust our max pooling and number of layers until we find the most accurate solution.

Once we have selected the pre-trained model we can begin with training the data. it is possible that due to the noise in the images, we will need to further process the images prior to training and annotate each object in the image for better, more accurate training.

Due to the images containing multiple objects in addition to SAR images being very noisy, we may have to add additional preprocessing steps to prepare the images before the training phase. In our research, one option would be to implement a filter called Canny mask detection (see related work section), which helps with reducing the noise around the images, making for a more accurate object detection.

After preprocessing and pre-training is concluded, we can proceed with the training phase on our test data, then use the model on test and validation datasets and evaluate the accuracy of the results.

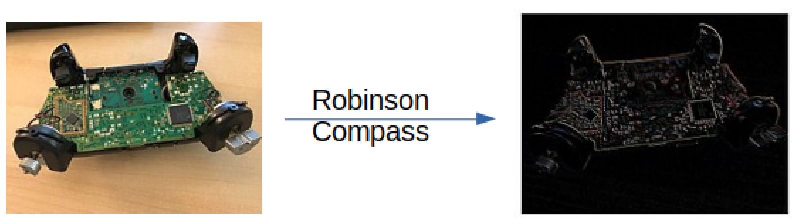
If, after we go through all the deep learning phases with RCNN, including tuning and optimizing, we find that the model does not perform well with SAR images, we may have to pivot and use other models such as Yolo for object detection, in which case our final submission will include the two models, and an in depth comparison between the models and their performance.

6.2 - Image Masking and Edge Detection Methods

Using Faster-RCNN and YOLO frameworks aims to serve as a baseline comparison when expanding on existing methods. Our second method involves inserting an additional step into our classification pipeline: Image Masking and Edge Detection.

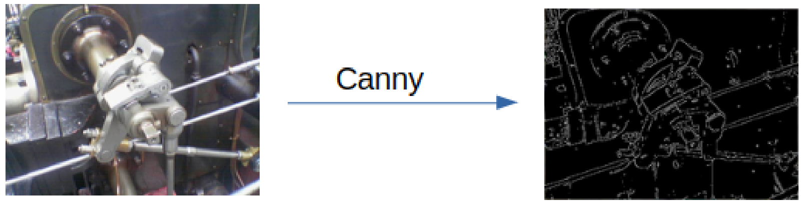
6.2.1 - Image Masking

One of the most popular image processing techniques is Image Masking. This involves selectively hiding or removing parts of an image based on a predetermined filter. The filter convolves over the image, identifying which pixels to retain, and which to remove. The result is an image where only a specific component of the original remains. In short, the mask isolates and highlights specific regions in the image based on the mask configuration. There are dozens of image masking techniques in use today, but the most popular tend to be the Robinson Compass, Kirsch Compass, and Sobel operators.



6.2.2 - Edge Detection

A similar but second image processing technique we intend to use is edge detection. As the name suggests, Edge Detection involves using a filtering method to detect edges of different objects within an image. This usually involves detecting extreme changes in pixel intensity, indicating a transition to another object in the image. The key difference between these methods and masking methods is that masking methods highlight specific areas whereas edge detection highlights the boundaries between objects. One of the most popular techniques is the Canny method.



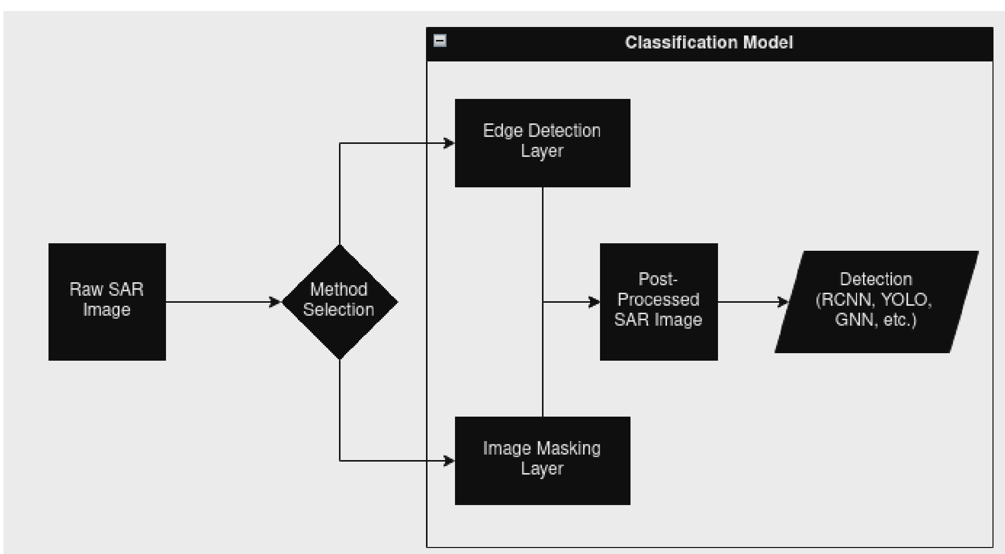
6.2.3 - Model Insertion

We propose that applying masking or edge detection techniques prior to training can assist in removing much of the unwanted noise found in SAR image data. As these techniques remove said noise, we hypothesize that this will isolate ship targets and make them significantly easier to detect compared to using the standard image format. The rationale behind each method is as follows: image masking will highlight the entire ship target and remove much of the unwanted clutter while edge detection will highlight the outline of the ship target.

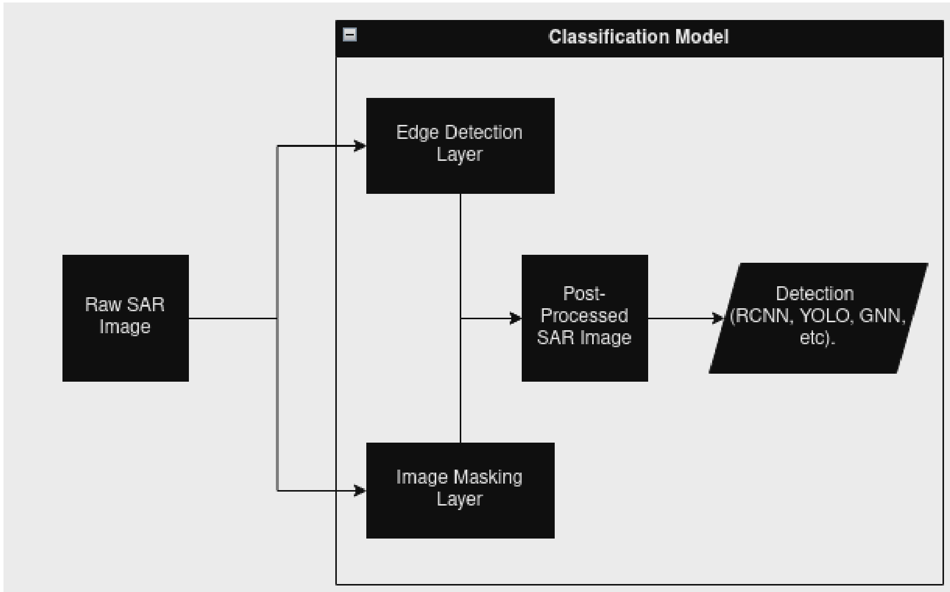
There are two method we intend to utilize our proposed techniques:

1. Using edge detection or image masking independently.
2. Using edge detection and image masking concurrently with the functional Keras API, then adjusting the model weights based on the multi-output result.

Below is an overview of our first model. Selecting one or the other avoids the possibility of removing too much of the target characteristics that the model cannot extract important features, at the cost of additional noise.



Using the Keras Functional API, we can also implement our second methodology to combine both images into a single layer for the detection step. With both image processing techniques removing different noise types from the image, the model may perform better by identifying unique characteristics from the output of both resulting operations.



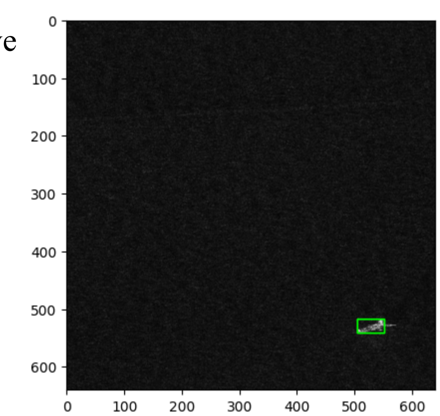
We hypothesize that using both layers in a Sequential model would be detrimental to performance as cascading the techniques may remove too many target details for the model to identify them.

For tools, we will be using GoogleColab for development, Tensorflow or PyTorch as framework and Github as a repo. We will also be using a Jira board to track progress and distribute work between team members.

7.0 - Expected Results

7.1 - Expected Image Results

We expect the model to both detect and annotate each respective sample within the testing set. Annotations, such as bounding boxes, must surround the desired target within an image. An example of the expected output is shown in the figure to the right. Each image will retain its 640x640 size with the annotation and the annotations should include as little background information as possible.



7.2 - Metrics

7.2.1 - Accuracy

Accuracy is a common metric used to measure CNN model performance used on similar tasks, as we've seen in related works. In order to benchmark the performance of the models we use, we intend to use similar metrics.

7.2.2 - Intersection Over Union

Intersection Over Union (IOU), or the Jaccard distance, is a common metric used for evaluating object detection models. It involves the use of set theory to determine the area of intersection of a model's predicted bounding box and the actual bounding box of the target. This metric will help us determine what noise or non-target entities are confusing the model and what improvements we can make to it. IOU is calculated via the following equation:

$$J(A \cap B) = \frac{|A \cap B|}{|A \cup B|}$$

8.0 - References to data sources

This dataset is designed for researchers interested in SAR ship detection and instance segmentation. It provides a variety of maritime SAR images suitable for developing and testing detection models.

Source: Sudheer, Kailas P. (2023). SARSCOPE: Synthetic Aperture Radar Maritime Images. Kaggle.
https://www.kaggle.com/datasets/kailaspsudheer/sarscope-unveiling-the-maritime-landscape:

MulTags-SARwv comprises 2,100 SAR images acquired by the Sentinel-1 satellite's wave mode. Each image covers a 20 km by 20 km area with a 5 m pixel resolution, providing detailed data for maritime analysis.

Source: Wang, C., Stopa, J., Vandemark, D., Foster, R., Ayet, A., Mouche, A., & Chapron, B. (2023). MulTags-SARwv Dataset: Synthetic Aperture Radar for Maritime Analysis. School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing, China.
seanoe.org

Gupta, H., Verma, O.P., Sharma, T.K. et al. (2024). Ship detection using ensemble deep learning techniques from synthetic aperture radar imagery. Scientific Reports 14, 29397.
https://doi.org/10.1038/s41598-024-80239-y

Muhammad, Y., Lili, Z., Shanwei, L., et al. (2023). Instance segmentation ship detection based on improved Yolov7 using complex background SAR images. Frontiers in Marine Science 10.
https://doi.org/10.3389/fmars.2023.1113669

Chen, X., Liu, Y., & Zhou, M. (2023). SSS-YOLO: Mitigating False Alarms in SAR Imagery via Hybrid Noise Classification Layers. Remote Sensing 15(8), 2100.

https://doi.org/10.3390/rs15082100

9.0 - Justification for Using Existing Code

Because SAR images are much more difficult to detect objects in than other, non SAR images, it would be helpful to try existing implementations first (we need still need to work on pre-processing, pre-training and annotating the images, as well as choosing the right pre-training model and optimization techniques) so that we can quickly pivot if the results are not successful. We will mainly be using the code as a skeleton/boilerplate to build on to the model further and optimize it to our unique needs. There aren't as many documentations for RCNN and therefore, using a starter code would allow us to concentrate on the intricacies of SAR analysis. This code we found in the following repository will not solve our problem as it is meant to be executed on regular images, and the model is not optimal even for regular images, let alone for SAR images which are much more vague and difficult to detect objects in, but it will give us a starting point so that we can come up with optimizations.

10.0 - References to Images, Figures, and Related Works

10.1 - Images

- **Canny Edge Detection Images**: Wikipedia Contributors. (2019, April 14). *Canny edge detector*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Canny_edge_detector
- **Robinson Compass Images**: Wikipedia Contributors. (2024, June 14). *Robinson compass mask*. Wikipedia; Wikimedia Foundation.
- **SAR Image Data**: Sudheer, K. P. (2025). *SARscope: Synthetic Aperture Radar Maritime Images*. Kaggle.com. https://www.kaggle.com/datasets/kailaspsudheer/sarscope-unveiling-the-maritime-landscape/data

10.2 - Related Works

1. Zhang, B., Wijeratne, S., Kannan, R., Prasanna, V., & Busart, C. (2023). Graph neural network for accurate and low-complexity sar atr. *arXiv preprint arXiv:2305.07119*.
2. Gupta, H., Verma, O.P., Sharma, T.K. *et al.* Ship detection using ensemble deep learning techniques from synthetic aperture radar imagery. *Sci Rep* **14**, 29397 (2024). https://doi.org/10.1038/s41598-024-80239-y
3. Wang, Y., Kitani, K., & Weng, X. (2021, May). Joint object detection and multi-object tracking with graph neural networks. In *2021 IEEE international conference on robotics and automation (ICRA)* (pp. 13708-13715). IEEE.
4. Muhammad, Y., Lili, Z., Shanwei, L., Jianhua, W., Sakaouth, H. M., Tugsan, C. A. I., Mengge, L., Qamar, U. I., Mehdi, S. R., & Qian, Y. (2023). Instance segmentation ship

detection based on improved Yolov7 using complex background SAR images. *Frontiers in Marine Science, 10*. Retrieved from https://www.frontiersin.org/articles/10.3389/fmars.2023.1113669/full.

All remaining images used were furnished by AI-570 Group 2 individually.