xView3-SAR: Detecting Dark Fishing Activity Using Synthetic Aperture Radar Imagery

Supplementary Material

A Label Confidence Levels

As we combine automated and manual annotations to create our ground truth labels, we use the following criteria to assign label confidence levels:

- If both the GFW algorithm and professional labelers agreed that an object is a `vessel`, then the object is assigned a label of `vessel` with `HIGH` confidence.
- If the GFW algorithm determined an object is a `vessel` and the professional labelers did not, then the object is `vessel` of `MEDIUM` confidence.
- If the professional labelers determined an object is a `vessel` with `HIGH` or `MEDIUM` confidence and and GFW did not, then the object is `vessel` of `MEDIUM` confidence.
- If the professional labelers determined an object is a `vessel` with `LOW` confidence and and GFW did not, then the object is `vessel` of `LOW` confidence.
- If the GFW algorithm determines an object to be `non-vessel` and the object is more than 2 kilometers away from shore, then the object is assigned a label of `non-vessel` with `HIGH` confidence.
- If the professional labelers determined an object is a `non-vessel` with `HIGH` confidence and and GFW did not, then the object is `non-vessel` of `MEDIUM` confidence.

Please see Appendix G for the confidence levels employed by human annotators and Table 6 for the distribution of confidence levels in the data broken out by all objects, vessels, and non-vessels.

B Reference Detection Approach

The reference implementation leverages the Faster-RCNN architecture [32]. The model backbone is initialized using pretrained weights from the ImageNet database. The first layer is replaced by a newly initialized convolutional layer with the appropriate number of input channels. The classification head is adjusted to handle the number of classes used by the model—three in the reference implementation case. Data is preprocessed by dividing each scene into 800 x 800 pixel chips, normalizing pixel values between 0 and 1 at the chip level, and creating a bounding box of 10 pixels on each side and a detection at its centroid for each ground-truth annotation. Each bounding box was provided a label of "non-vessel," "non-fishing vessel," or "fishing vessel." The normalized image is provided as input to the Faster-RCNN backbone, which can be trained as usual given the combination of bounding box and multiclass information provided for each detection.

The reference model is trained for five epochs on an NVIDIA DGX-1 using a single V100 GPU. Stochastic gradient descent with momentum is used to optimize the model with $\alpha = 5e^{-3}$, momentum of 0.9, and weight decay of $5e^{-4}$. A step learning rate scheduler is used with step size = 3 and $\gamma = 0.1$.

When evaluated against the *holdout* split, model scores for each of the tasks (described above) were as follows: $F1_D = 0.4302$ (object detection), $F1_S = 0.1293$ (close-to-shore detection), $F1_V = 0.6891$ (vessel classification), $F1_F = 0.3946$ (fishing classification), and $PE_L = 0.0000$ (length regression) for an aggregate score of 0.1904.

Note that the reference model makes no attempt to accurately predict lengths. This was implemented solely to demonstrate the required prediction output and test the computation of the metrics. For every detection centroid a square bounding box of fixed width was created, the diagonal of which was used as the predicted length of the vessel.

C Winning Model Performance Overview

We present a brief analysis of winning model detection performance along axes of geography and vessel size. In Figure 13, we see that—as expected—detection models tend to improve as the true vessel size becomes larger. Importantly, in this figure we present metrics computed only vessels for which have a label of either medium or high confidence. While many of the regions behave similarly, it is worth noting that smaller vessels appear harder to find on SAR in the Adriatic than in other areas; this may be due to more common non-metal construction in this region. High recall in the Gulf of Guinea across vessel sizes is also compelling, as this is a high-priority region for IUU activity. Finally, Figure 14 presents the distribution of ground-truth vessels in by size; while there are subtle differences between the different regions, overall the distributions skew strongly right, as one would expect.

D Geographical and Annotation Distribution

Data partition and percent of labels with vessel length:

- Train - 70.66%
- Validation - 36.42%
- Public - 34.06%
- Holdout - 32.74%

Data partition	Region name	Num. of scenes
Train	Adriatic	66
	Bay of Biscay	315
	Gulf of Guinea	62
	Iceland	13
	Norway	98
Validation	Adriatic	11
	Gulf of Guinea	16
	Iceland	2
	Norway	21
Public	Adriatic	42
	Gulf of Guinea	36
	Iceland	8
	Norway	64
Holdout	Adriatic	57
	Gulf of Guinea	65
	Iceland	18
	Norway	97

Table 2: Breakdown of the geographic distribution of xView3-SAR.

Data partition	Label source	Percent
Train	AIS	100.00
Validation	AIS	0.91
	AIS/Manual	57.36
	Manual	41.73
Public	AIS	1.73
	AIS/Manual	51.00
	Manual	47.27
Holdout	AIS	1.67
	AIS/Manual	53.48
	Manual	44.84

Table 3: xView3-SAR label source.

	Train	Validation	Public	Holdout
True	56.74	62.20	64.52	63.02
False	26.04	37.18	34.33	35.98
Not available	17.23	0.62	1.15	1.00

Table 4: Distribution of `is_vessel` labels, percentage of total detections.

	Train	Validation	Public	Holdout
True	19.51	5.00	5.75	5.40
False	37.22	15.24	11.44	11.47
Not available	43.26	79.76	82.81	83.13

Table 5: Distribution of `is_fishing` labels, percentage of total detections.

Confidence	All objects	Vessels	Non-vessels
High	38.9	20.0	80.3
Medium	41.2	56.3	19.7
Low	19.9	23.7	NA

Table 6: Confidence level for each object type, percentage of total detections. Note that "low" confidence level is not in use for non-vessels. See Appendix A for more details.

E Summary of Related Datasets

Name	Maritime Instances	Classes	Gigapixels	Task Type	Sensor Type
xView1 (training split) [16]	5,141	9	—	OD	Optical
DOTA 1.0 [45]	37,028	1	—	OD	Optical
OpenSARShip 2.0 [18][1]	34,528	16	0.14	OD	SAR
AIR-SARShip-1.0 [36]	461	1	0.28	OD	SAR
SSDD [48][2]	2,456	1	1.16	OD	SAR
FUSAR-Ship [13][3]	1851	15	1.31	OD	SAR
HRSC2016 [21]	2,976	**22**	1.91	IS, OD	Optical
Zhang et. al. [50]	5,175	1	2.40	OD	Optical
SAR-Ship-Dataset [42]	43,819	1	5.74	OD	SAR
HRSID [43]	16,951	1	10.76	IS, OD	SAR
LS-SSDD-v1.0 [47][4]	6,015	1	11.52	OD	SAR
SeaShips [35]	40,007	6	195.68	OD	Optical
Airbus Ship Detection [1]	192,556	1	368.35	IS, OD	Optical
xView3-SAR	**243,018**	2	**1,421.81**	OD	SAR

[1]Comprised of GRD and SLC images from Sentinel-1; AIS used to create labels. [2]RadarSat-2, TerraSAR-X, Sentinel-1. [3]Labeled using AIS automatically via a spatio-temporal Hungarian matching scheme. [4]AIS used in labeling but manual correlation only; Sentinel-1.

Table 7: Summary of related work. The xView1 and DOTA 1.0 datasets contains tasks unrelated to vessel detection, and the pixel numbers are not directly comparable to xView3-SAR. Under the Task Type column, "OD" stands for object detection; "IS" stands for instance segmentation. Bold cells indicate the largest number in each numerical category.

F Supplementary Figures

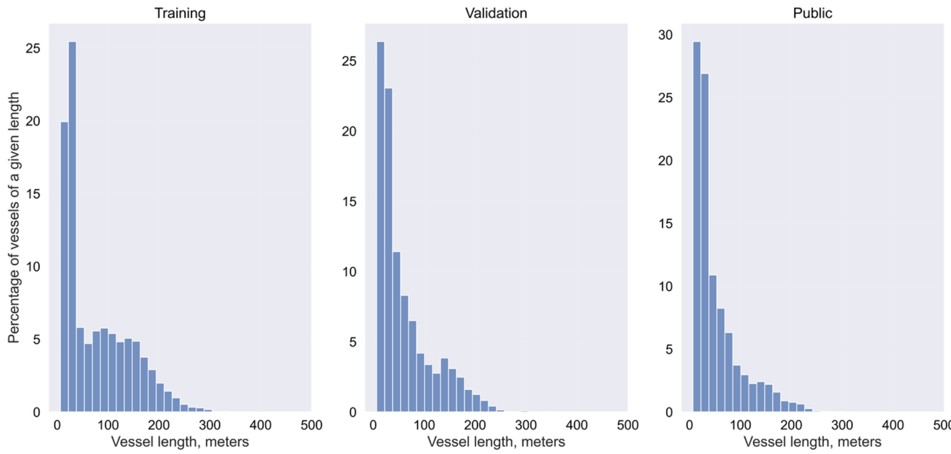


Figure 7: Distribution of vessel lengths.

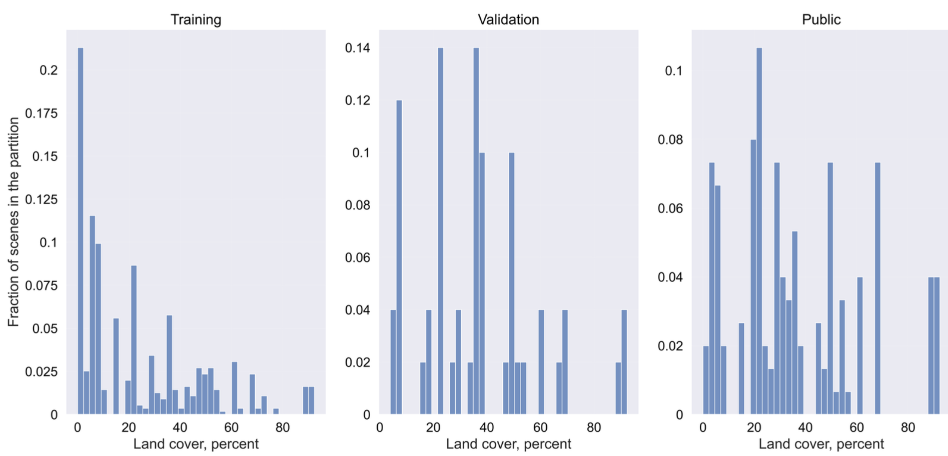


Figure 8: Distribution of land cover. From left to right: train, validation, public test, holdout.

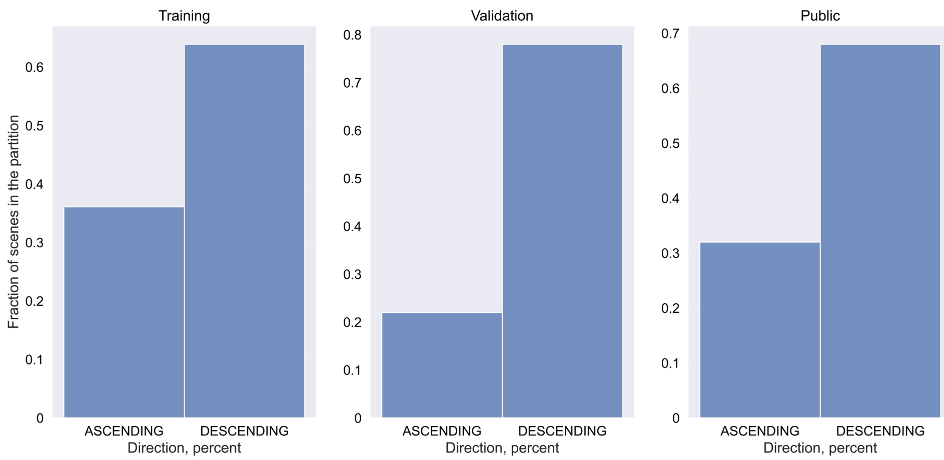


Figure 9: Distribution of satellite direction. From left to right: train, validation, public test, holdout.

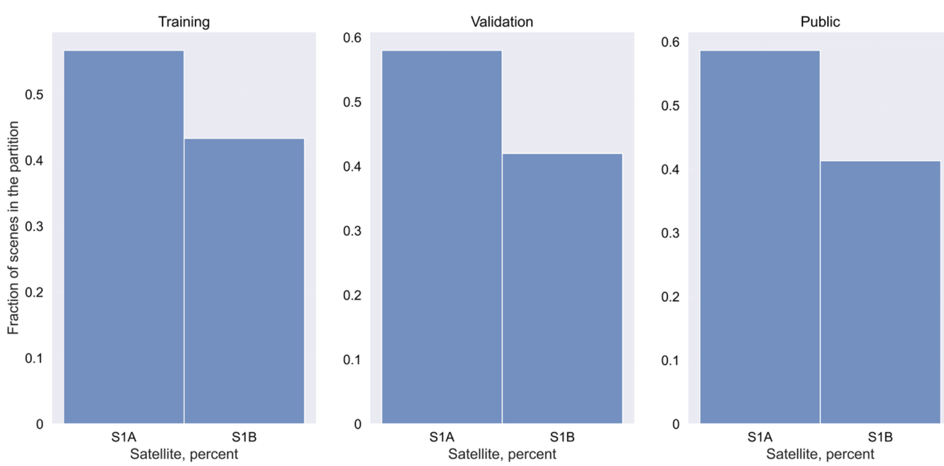


Figure 10: Distribution of satellite instrument source. From left to right: train, validation, public test, holdout.

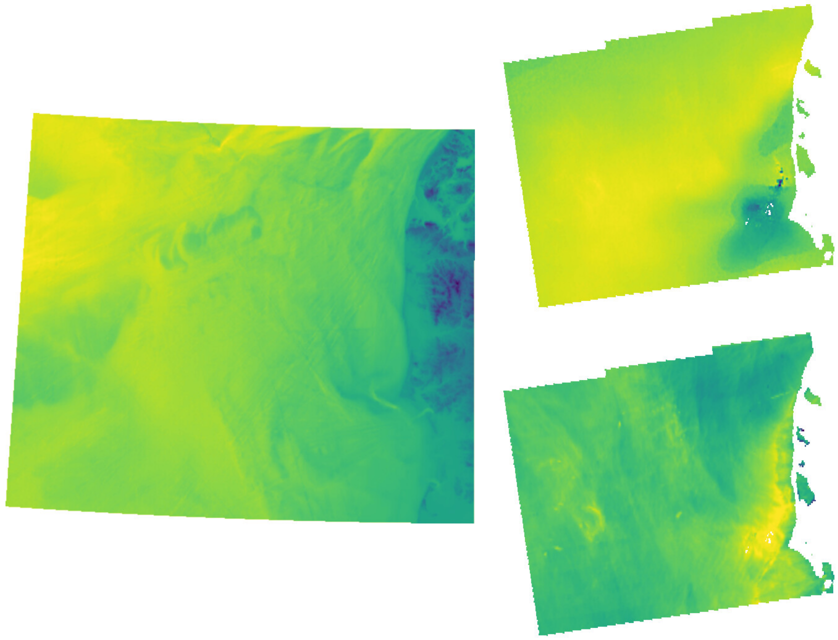


Figure 11: Visualization of ancillary data provided for a fixed area of interest; left: bathymetry, right top: wind direction, and right bottom: wind speed.

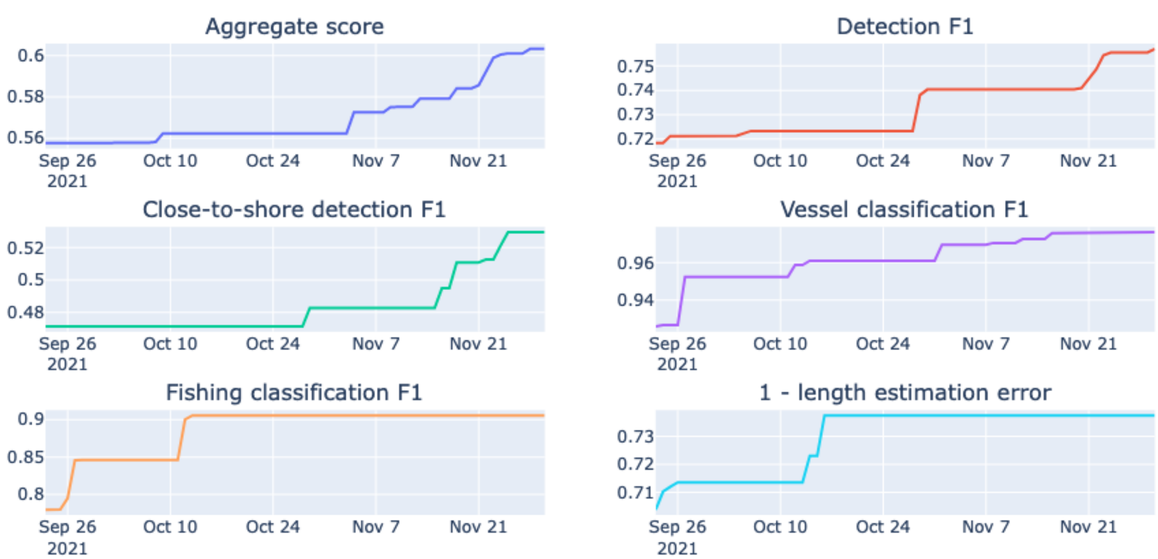
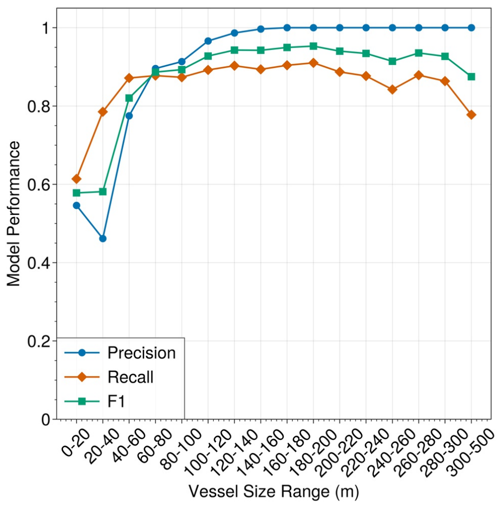
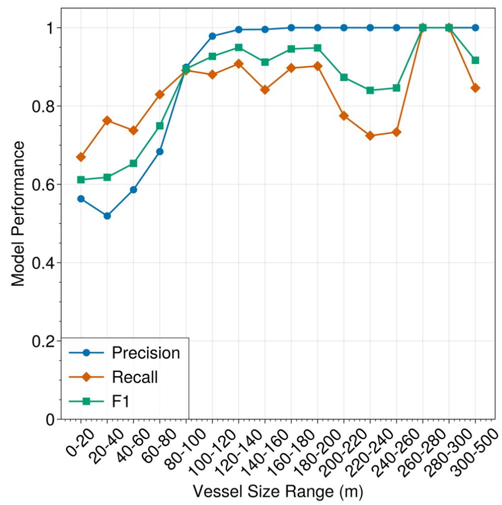


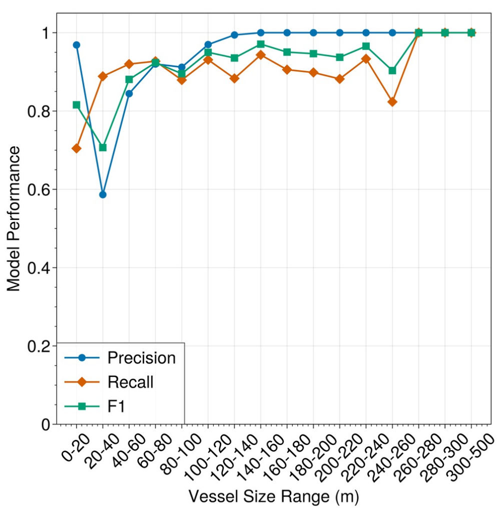
Figure 12: Improvement in aggregate and sub-scores from September–November.



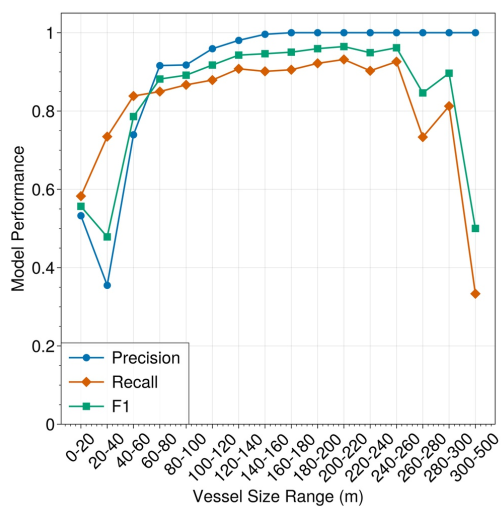
(a) All Regions



(b) Adriatic

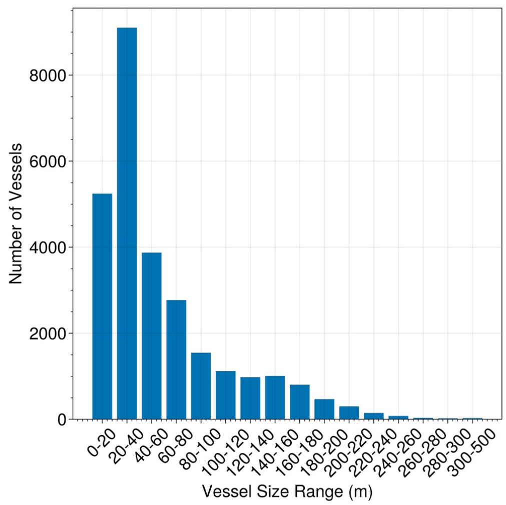


(c) Gulf of Guinea

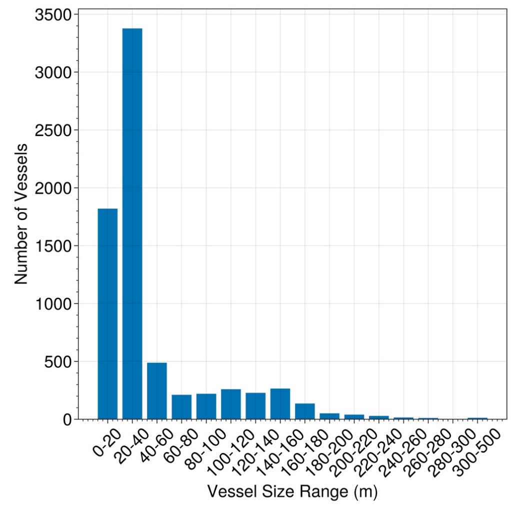


(d) Iceland and Norway

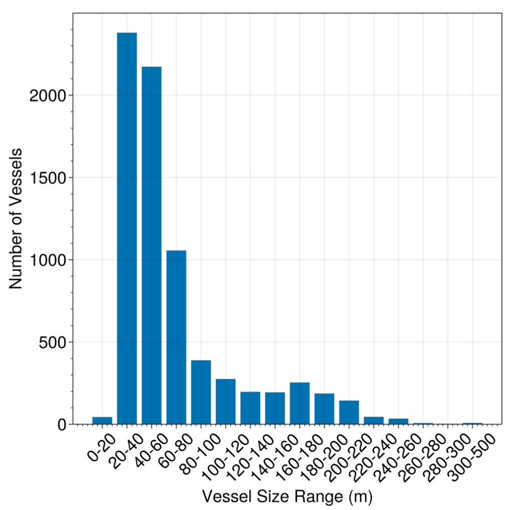
Figure 13: Winning model performance as a function of vessel size on high and medium confidence ground truth data by geographic region.



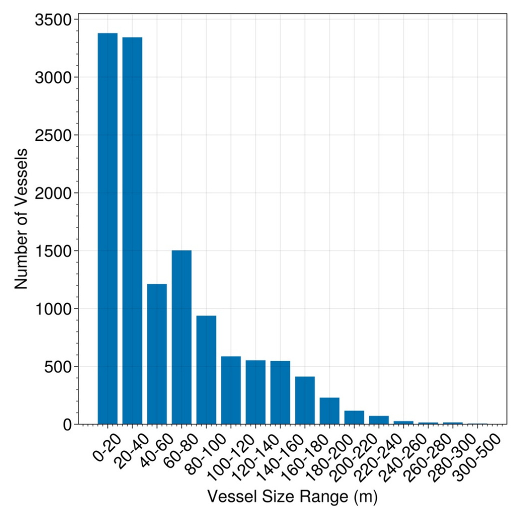
(a) All Regions



(b) Adriatic



(c) Gulf of Guinea



(d) Iceland and Norway

Figure 14: Distribution of vessel size on high and medium confidence ground truth data by geographic region.

G Instructions to Expert Annotators

In these tasks, you will see a Satellite Synthetic Aperture Radar image. The ground sample distance (GSD) is 20m, meaning that a single pixel is 20 meters over the ground.

The Task is to (1) identify vessels and other vessel-like objects in the imagery and label them as "non-vessels" when one is highly confident that they are not vessels, and label them as "vessels" otherwise; (2) provide a confidence rating.

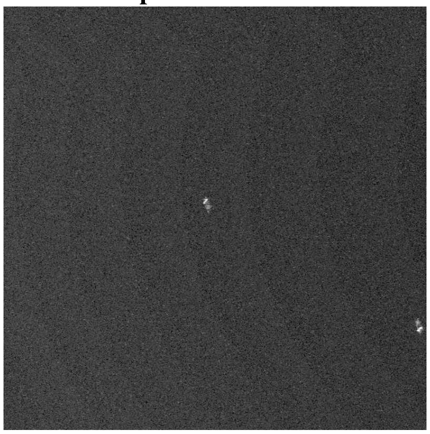
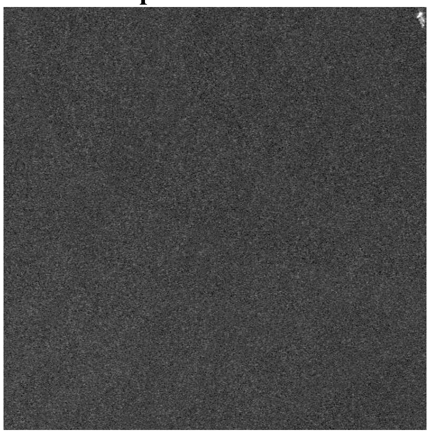
Contents

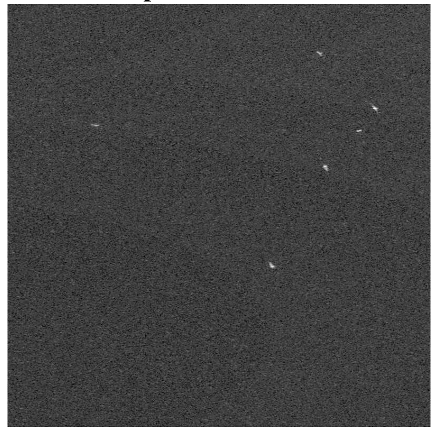
Categories

1. Vessel
2. Non-Vessel

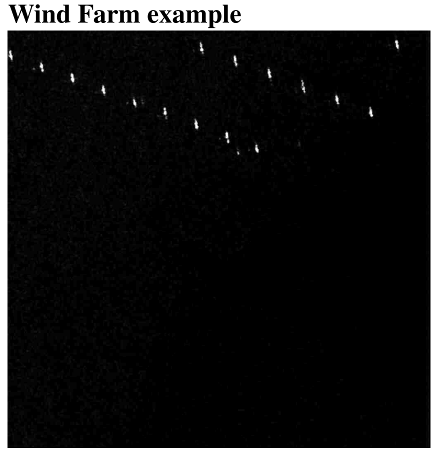
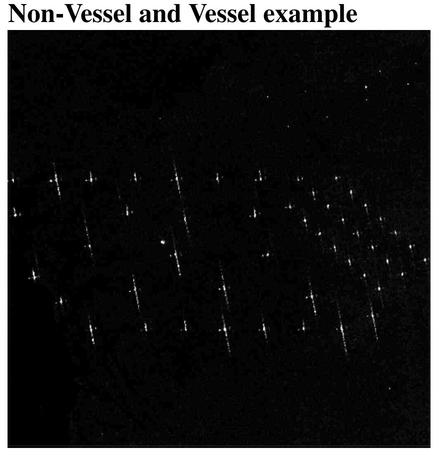
Annotation Rules

Vessels

Description	Example
Example: Place a bounding box around the area of the vessel the box should be tight around the object. *-In this example there should be two bounding boxes with the annotation for Vessel*	**Vessel example** 
Example: In this example there is a single object in the top right of the image chip; annotate with a bounding box. *-This example would have a single box annotation for vessel*	**Vessel example** 

Example: There are six objects in the image; annotate with a bounding box for each object. Note these objects do not form a clear pattern. *-This example would have six boxes with the annotation vessel*	**Vessel example** 

Non-Vessels

Description	Example
Example: This is an example of a wind farm, notable for the orderly pattern and the reflections formed by the objects. Please annotate these objects individually with a bounding box Non-Vessel. **Note that vessels may in some instances anchor in grids and they may anchor inside of wind farms. What separates a windmill from a vessel is likely going to be the backscatter from windmill blades.**	**Wind Farm example** 
Example: There is a cluster of objects in the center of the image, notable for the orderly pattern and the reflections formed by the objects. The objects in this cluster are windmills. Place a bounding box as Non-Vessel around each object in the cluster. **Note that vessels may in some instances anchor in grids and they may anchor inside of wind farms. What separates a windmill from a vessel is likely going to be the backscatter from windmill blades.** Note that there are also some objects in the top right hand corner; these are likely vessels. Place a bounding box as Vessel around each object in the top right hand corner.	**Non-Vessel and Vessel example** 

Negative example

Description	Example
Example: This image is an example of a bridge. The bridge is identifiable as a result of a long narrow line between land masses. ***Do Not*** annotate bridges or similar structures. Annotate all other objects that are likely Vessels and Non-Vessels with a bounding box	**Negative example** 

Confidence Rating

Description: For each annotation, what is your level of confidence for the detection of the object and its vessel/non-vessel label?

Confidence Attribute	Description
High confidence	I am absolutely sure it is a vessel.
Medium confidence	I am not so sure, but it looks like a vessel.
Low confidence	I am mostly guessing it is a vessel.