xView3 Ship Detection on AzureML

Phillip Hale, Microsoft, SMU

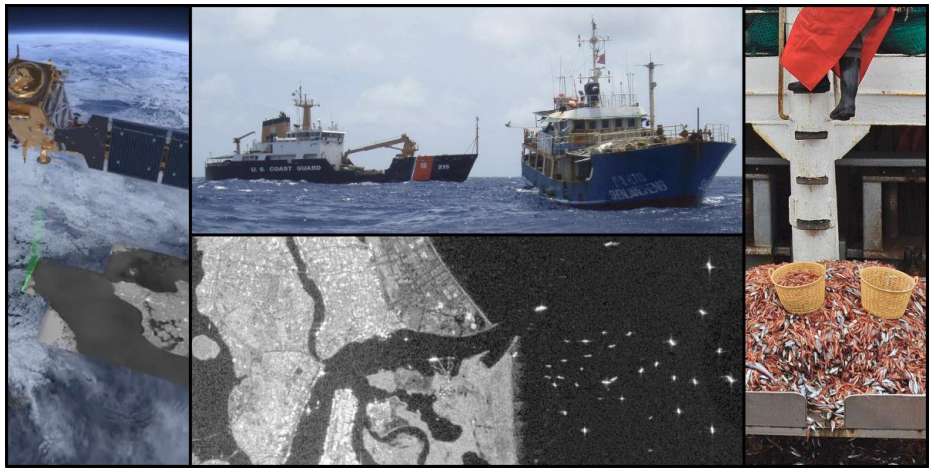
xView3 Challenge to Improve Illegal Ship Detection

Abstract short –

This report describes the techniques and experiments for improving automatic ship detection from synthetic aperture radar (SAR) satellite imagery carried out during the xView3 Dark Vessels Challenge in the Fall 2021.

Computer vision methods and Azure Machine Learning services utilized in this challenge seeks to advance research contributions in enhancements in extracting accurate ships mask and dimensions to enable positive improvements in ship detection, ship classification and estimating the length of detected ships.

Abstract Problem Statement –
The xView3 Challenge provides a large multi-dimensional dataset of SAR satellite views to benchmark new approaches to detect illegal fishing activities requiring object detection and classification techniques at a global scale.



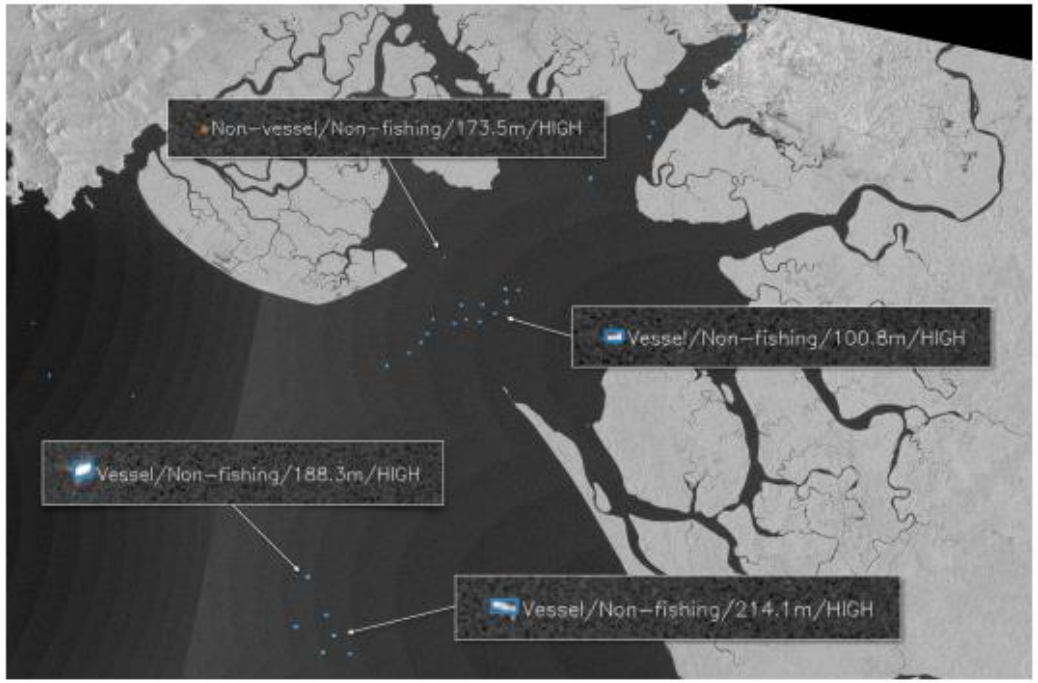
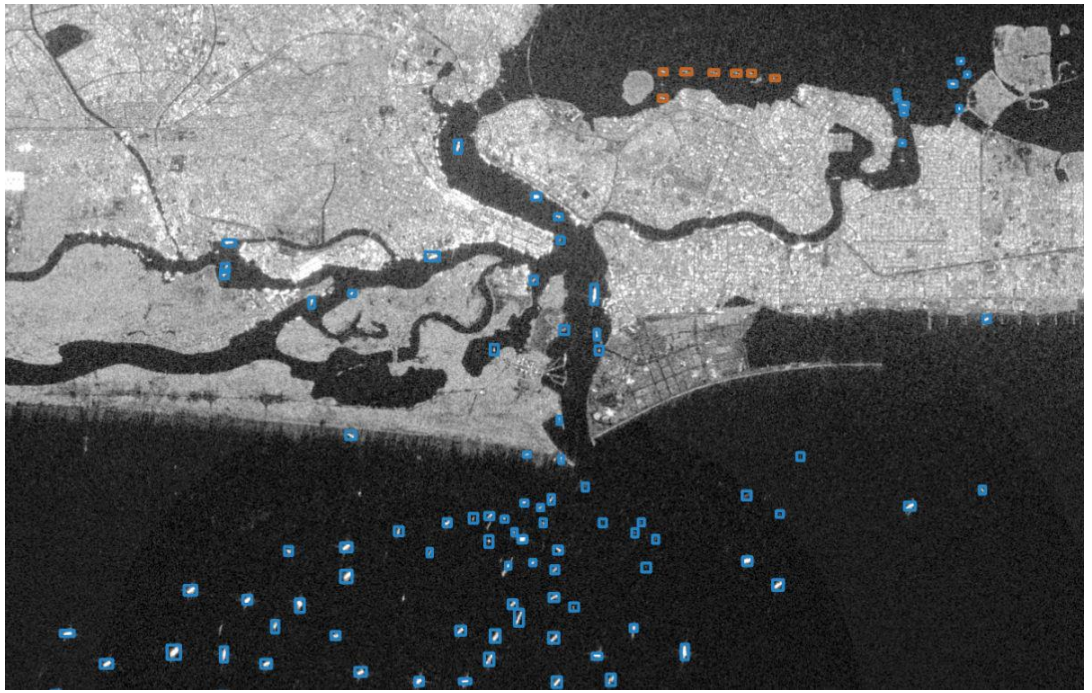


Figure 2: Sample annotations over an xView3 scene. Labels are showing vessel/non-vessel, fishing/non-fishing, l and confidence.



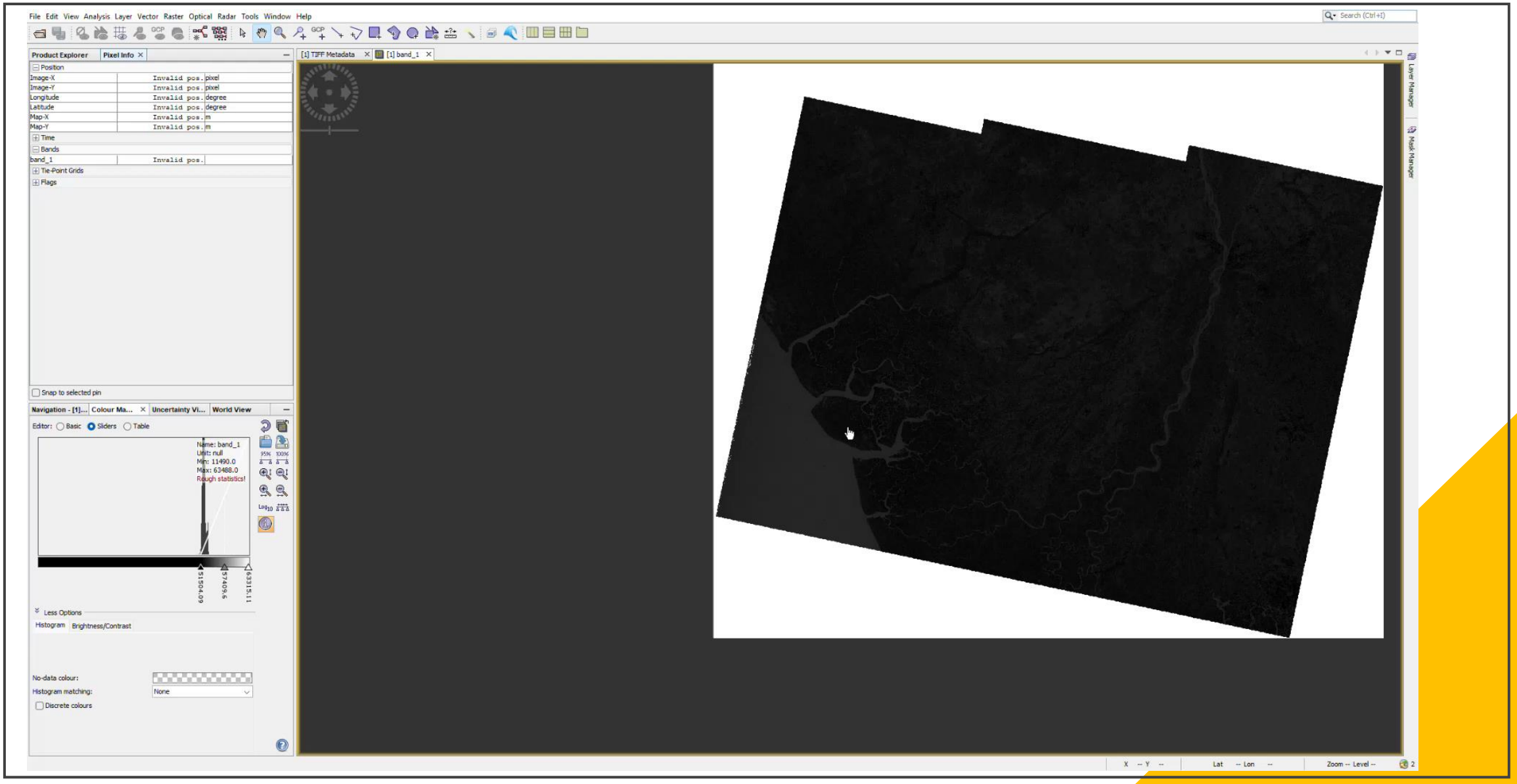
Source: xView3 - IUU

xView3 Prize Challenge: Developing Machine Learning Algorithms to Combat Illegal, Unreported, and Unregulated Fishing

For the xView3 Challenge, the prediction task is:

- Identify the maritime objects in each scene
- For each object, estimate its length, and classify it as vessel or non-vessel
- For each vessel, classify it as fishing or non-fishing

Synthetic Aperture Radar (SAR) imagery Visualization



Showing validation image from scene:
\0157baf3866b2cf9v\VH_dB.tif (28,760 x 24,644) 1.31 GB

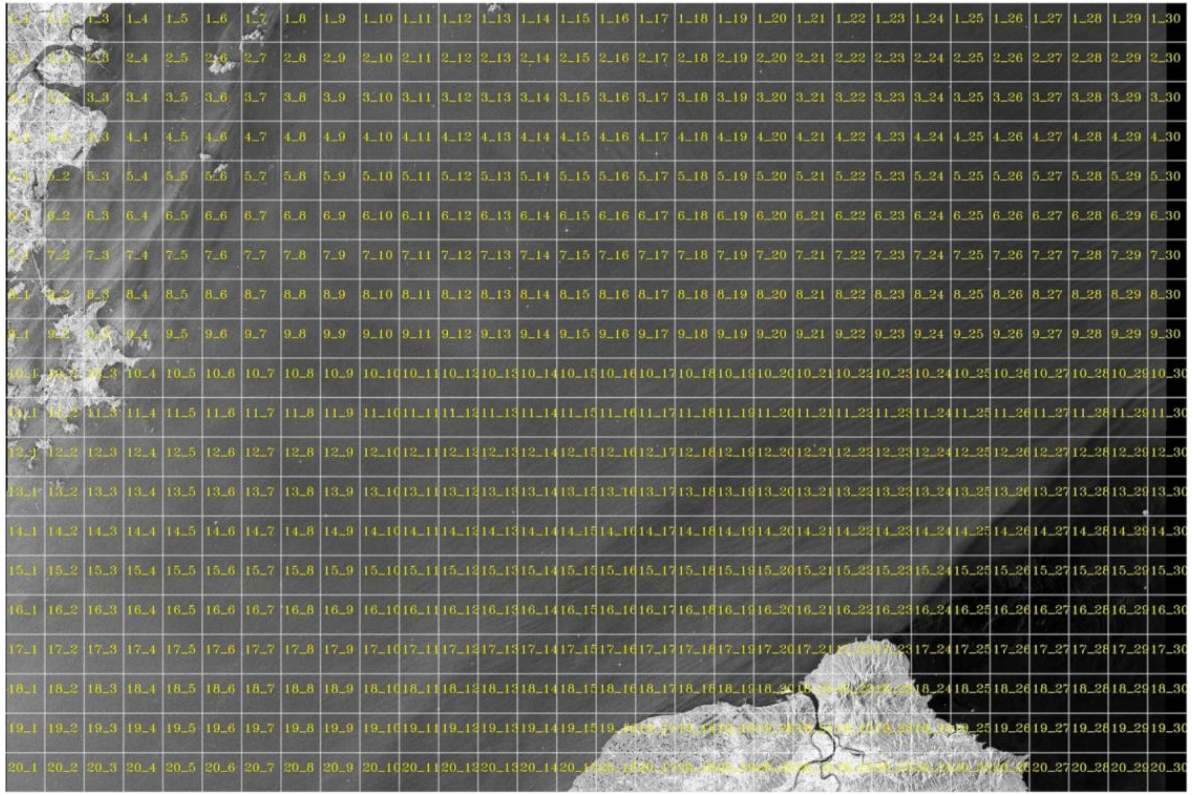


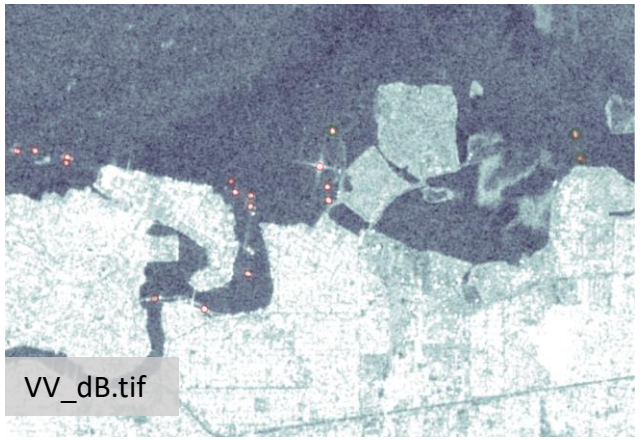
Image Cutting Process

- Crop large scale images to smaller scale (200x200 pixels)
- Most crops do not contain valid target
- Stored training images in a different blob storage due to size, potentially 2TB
- Initial cropping process took 14 days for the 1,000 scenes

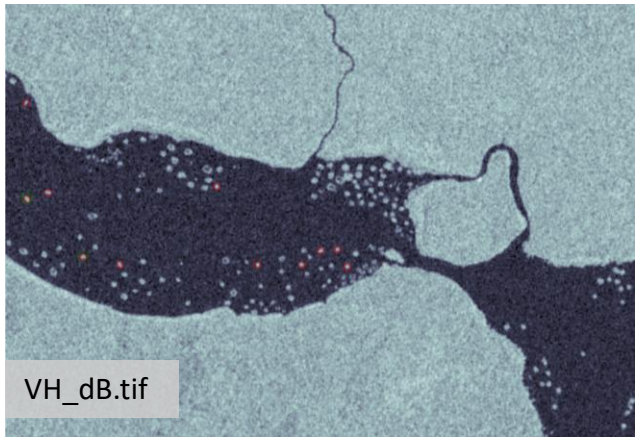
Xview3 Challenge Dataset Scene Files

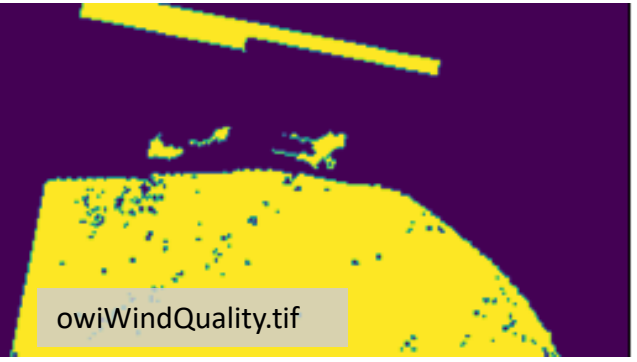

owiWindSpeed.tif

owiWindDirection.tif


VV_dB.tif

owiMask.tif


VH_dB.tif


owiWindQuality.tif


bathymetry.tif

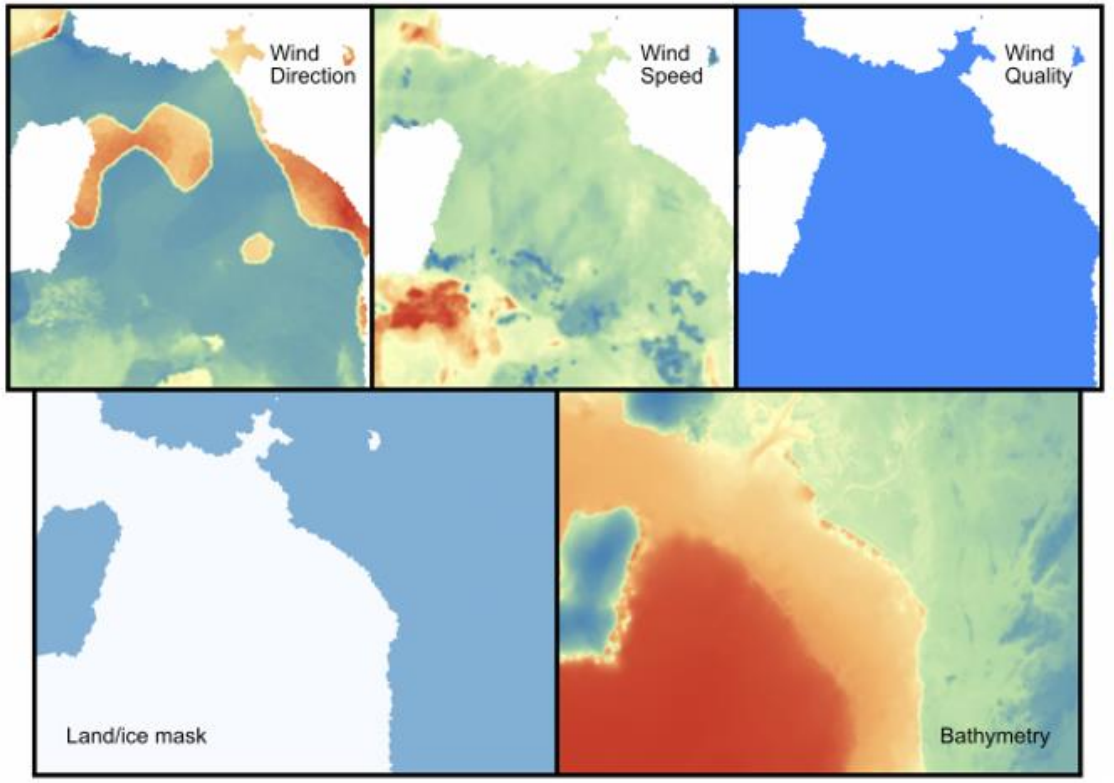
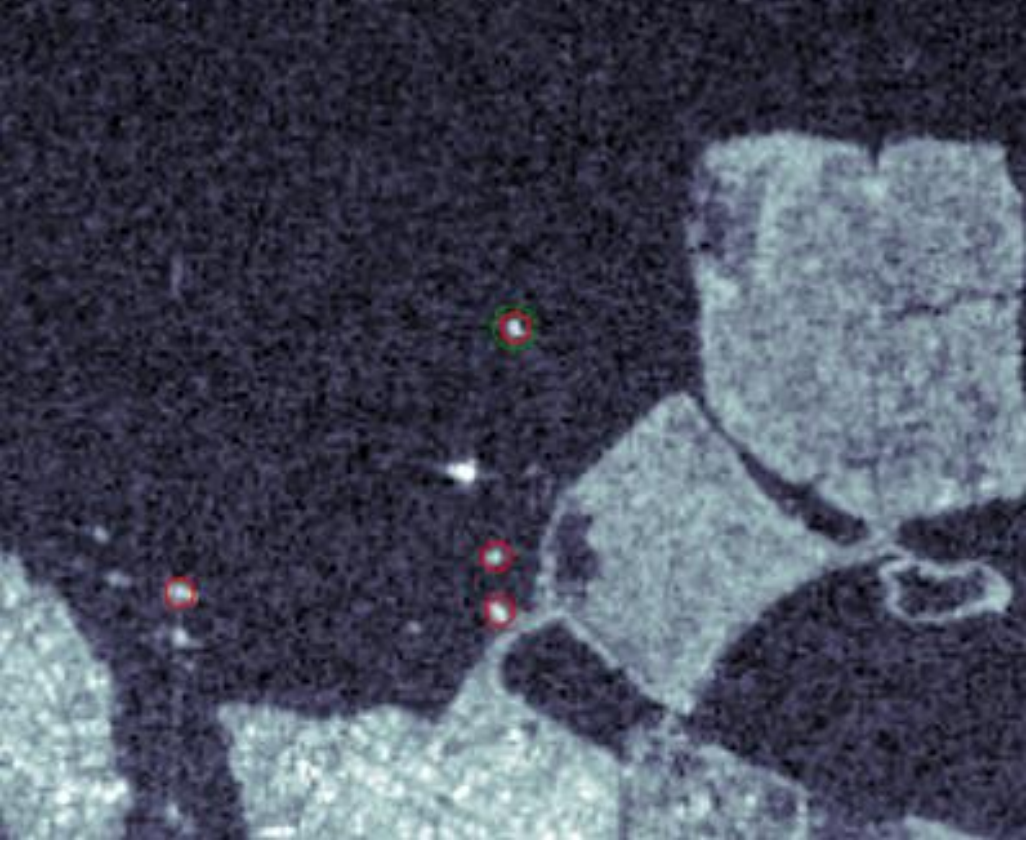


Figure 3: The xView3 dataset provides various wind data and bathymetry for each scene.

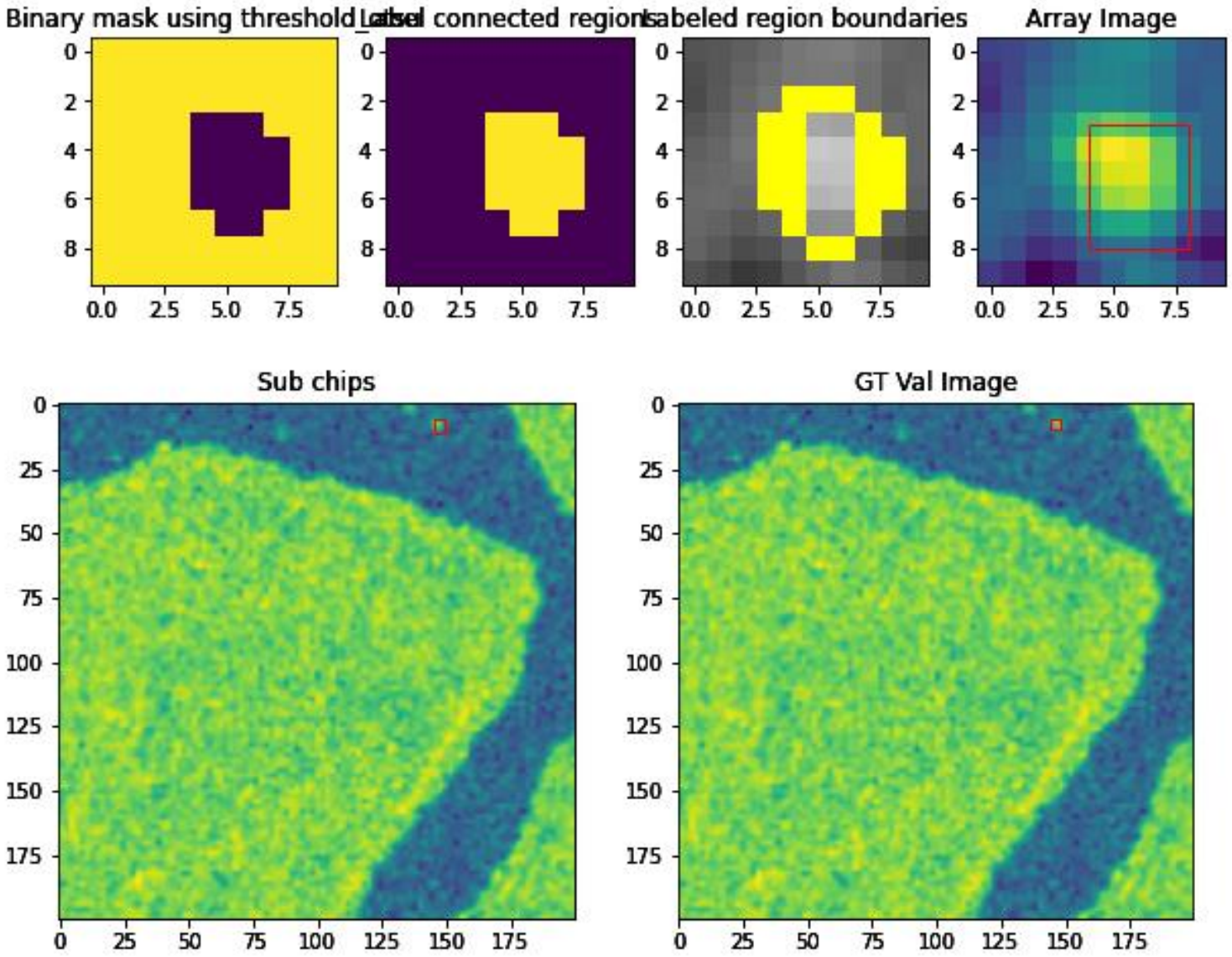
Source: xView3 - IUU



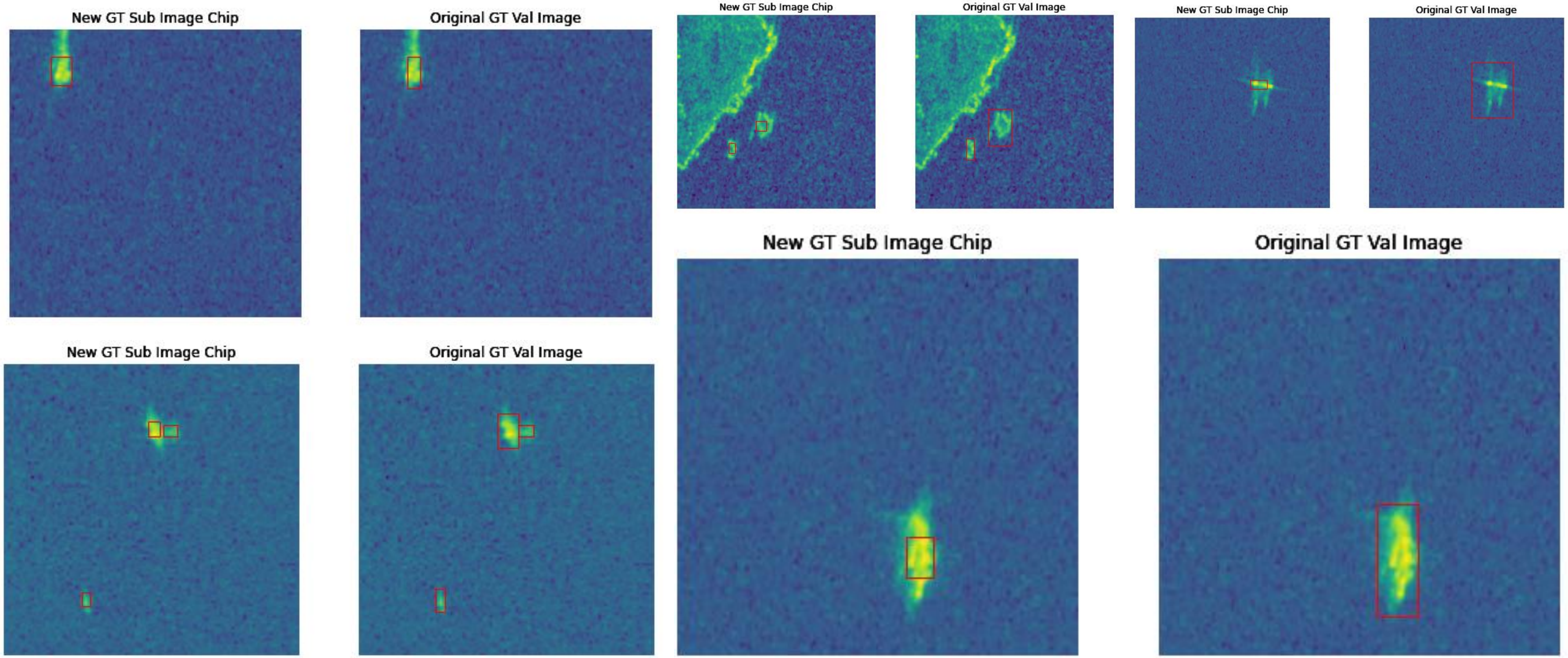
xView Mask dataloader v1 miss annotations

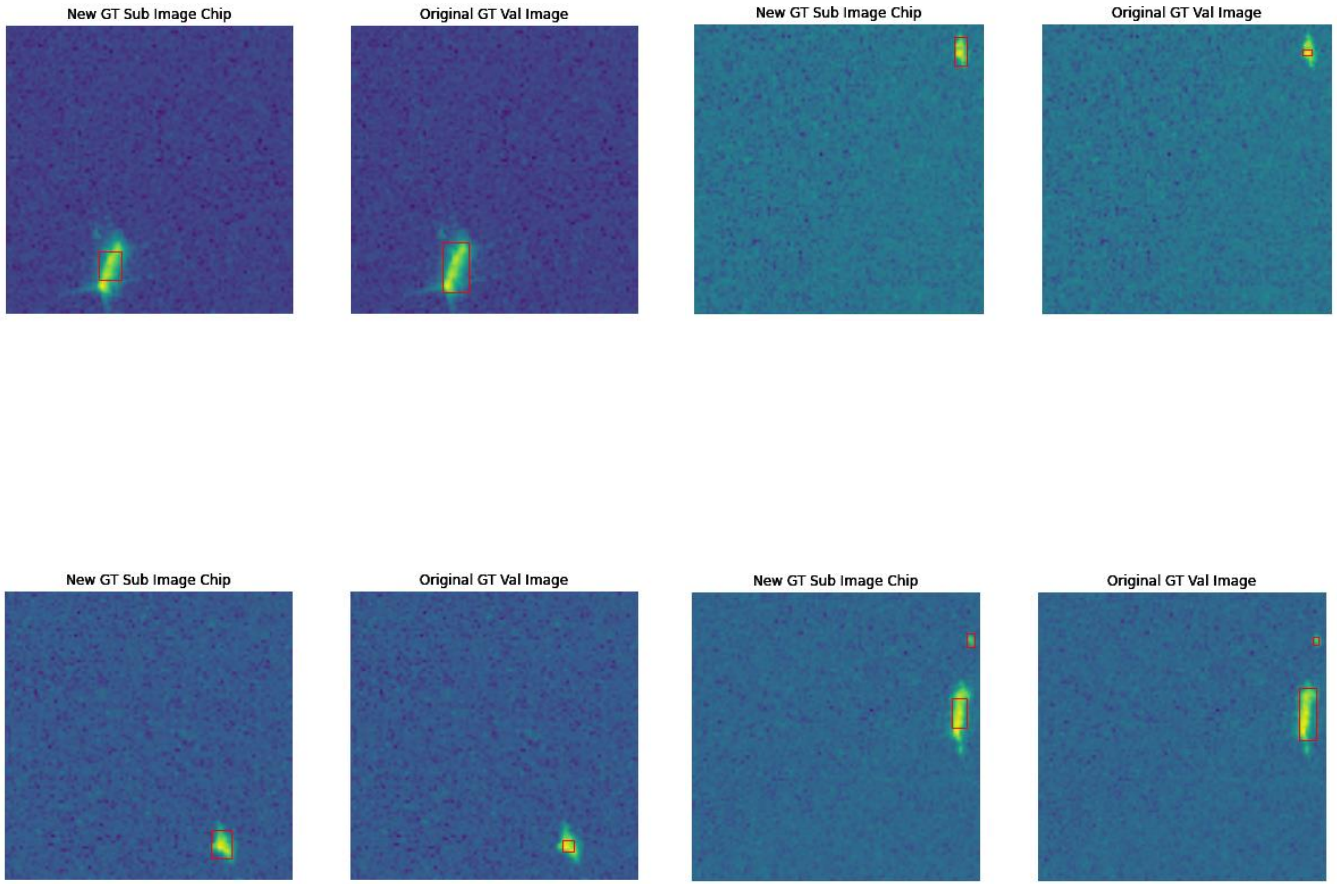
- Early version of the xview_dataloader did a poor job at creating annotations
- Many miss annotations that would error the training of the model

Sub chip zoom to correct bounding boxes and accurate mask for segmentaion



Enhancements on Original GT for Training

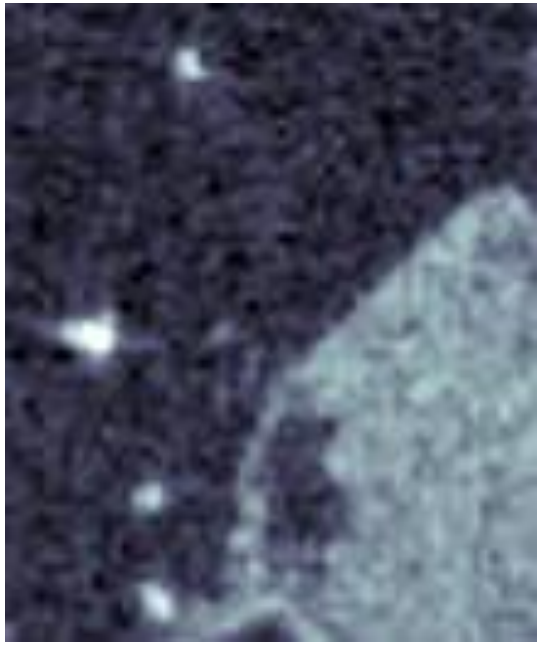




Errors on new GT

Overview of training xView3 Challenge in Azure ML

Raw File



Synthetics Mask Rendering

SAR raw .tif files are consumes raw and image and annotations
Create coco annotations files

Detectron2 in AzureML

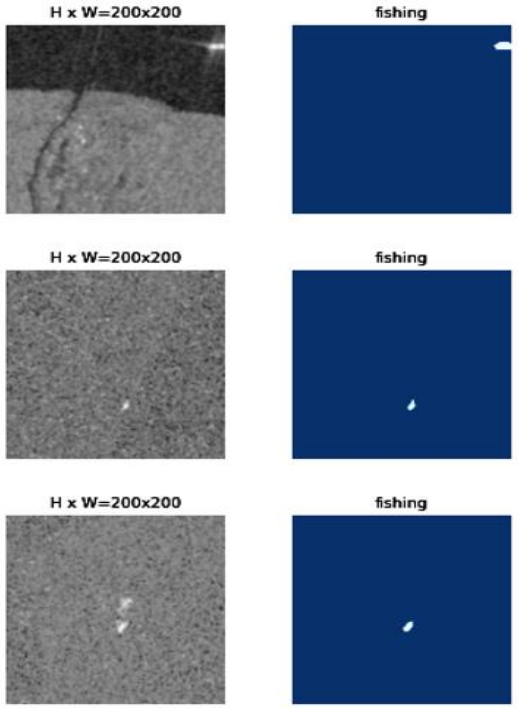
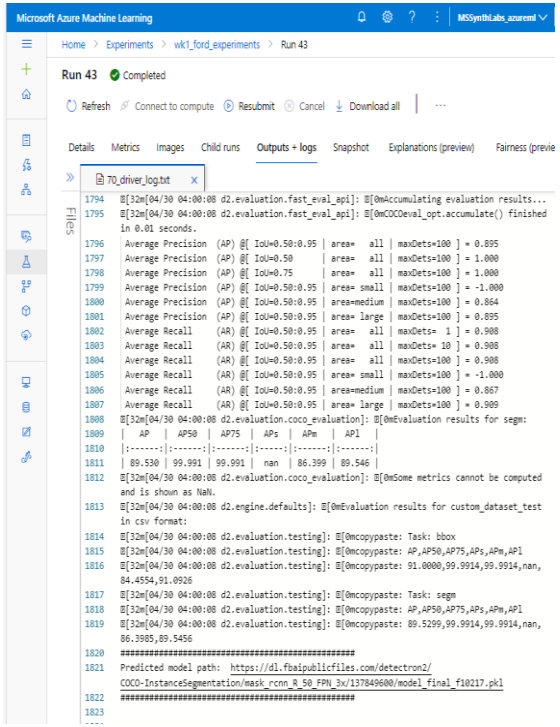


Image Picker at Train Time

Create new images for training instance segmentation models with configurable settings



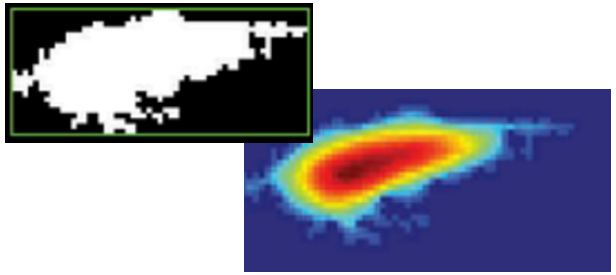
Integrated into AzureML

Utilizing AzureML enables structures and easy methods to kickoff new experiments

Model Explained

TBD

Gradient-weighted Class Activation Mapping



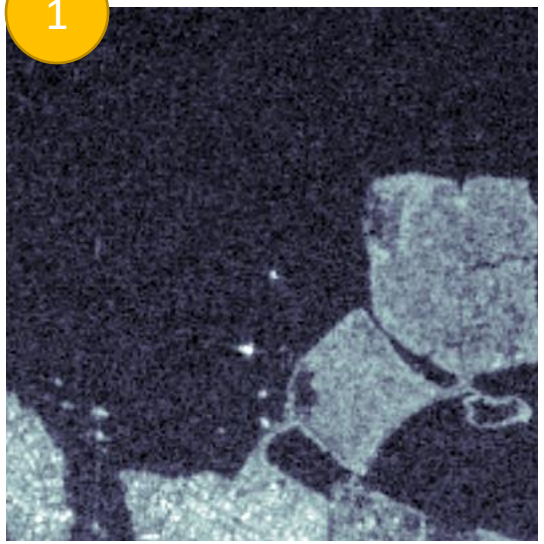
Visualizing experiments hyperparameters

Synthetics Explanation (Vis)

Hyperparameters and augmentations settings saved and utilized to measure the correctness of model

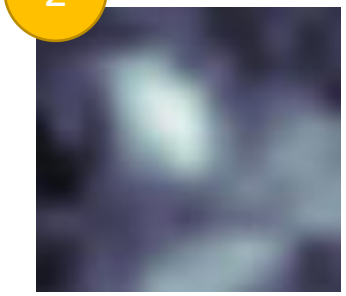
Estimating Ship Length using mask and ellipse

1



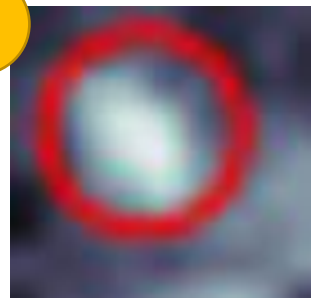
Input Test Image
Size 200 x 200

2



For each prediction,
filter mask crop

3

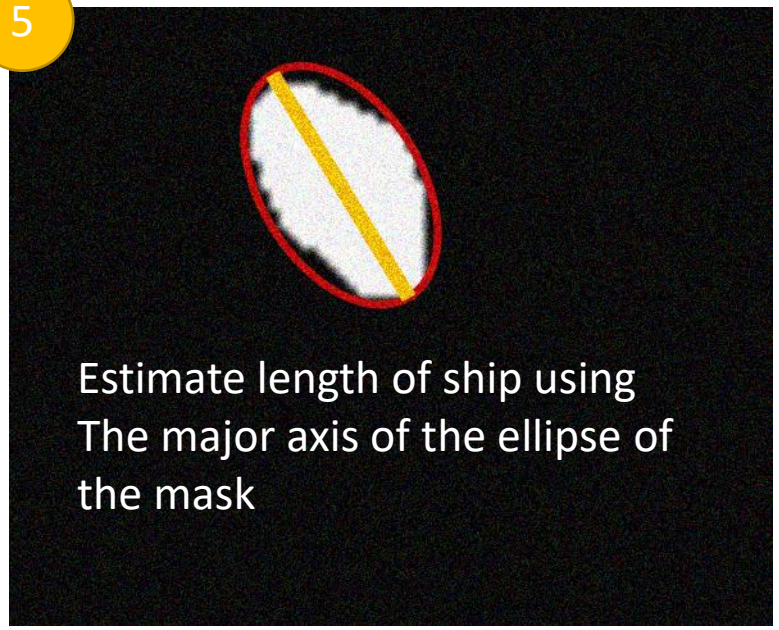


Class probability map

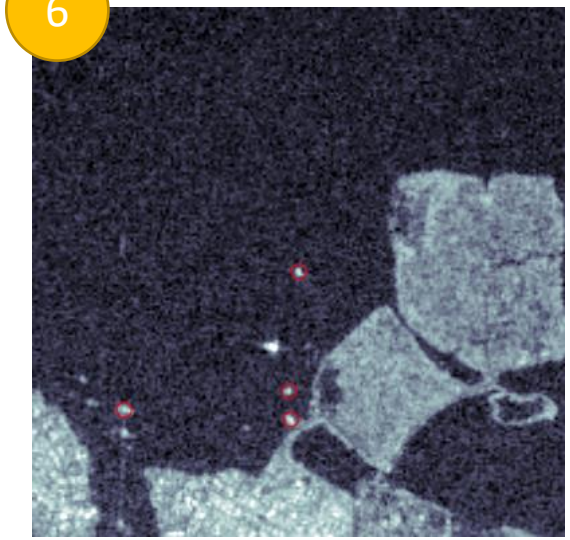
4

Bounding box + score

5



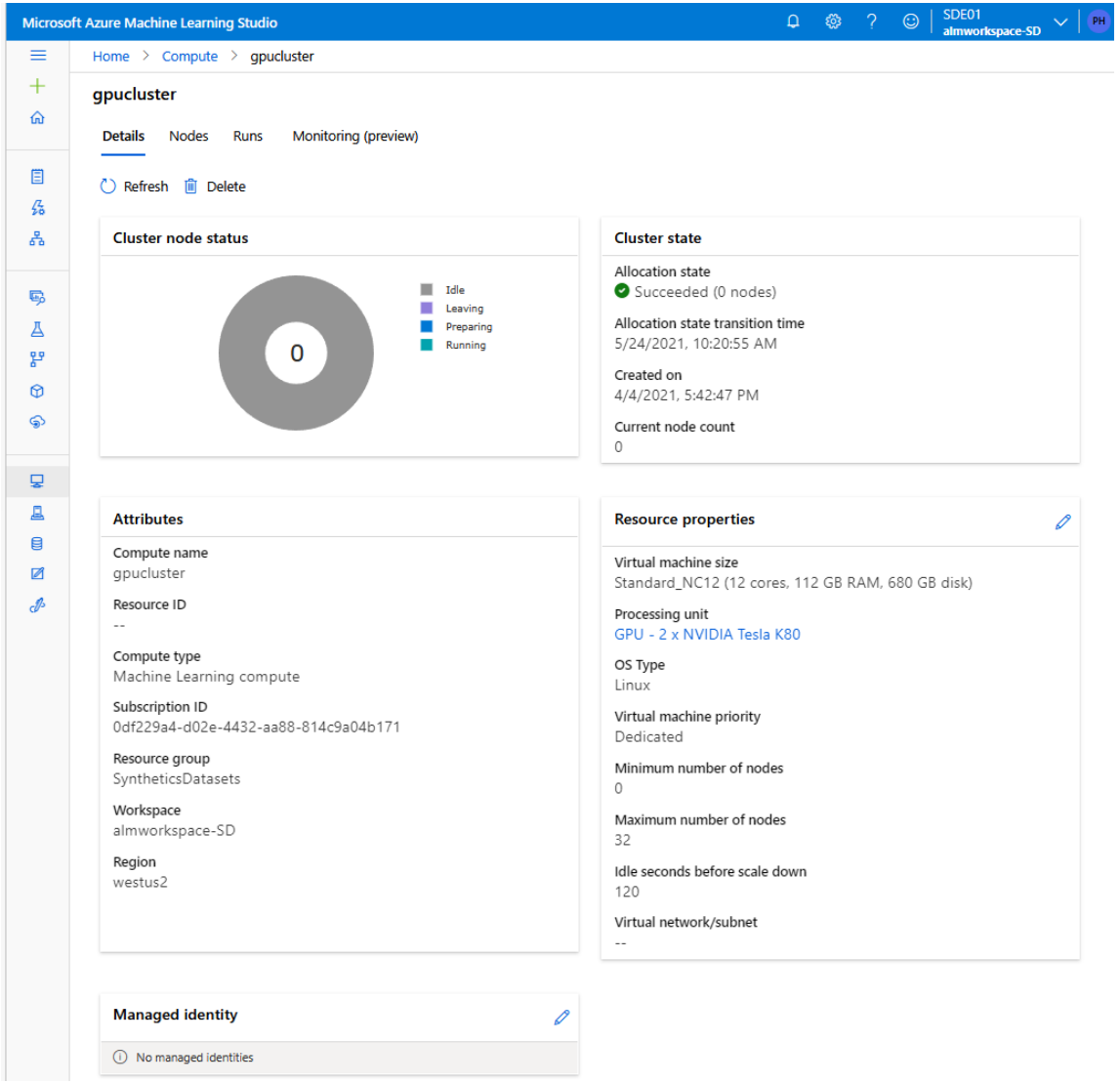
6

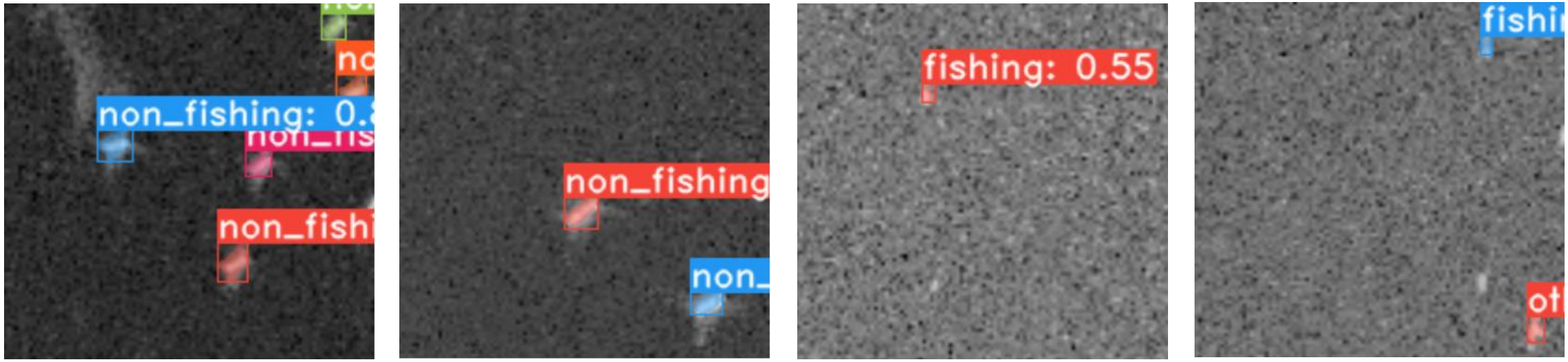


YOLACT prediction result
Size 200 x 200

Training with Detectron2

- Model Training Parms
- Model: mask_rcnn
- Backbone: Resnet50
- Neck: fpn
- Batch Normalization: Group Normalization (gn)
- GPUs: 2 NC12 based on K80
- Images per batch: 4
- Note: all models were tested on 1 GPU with





Yolact initial training 200 chip

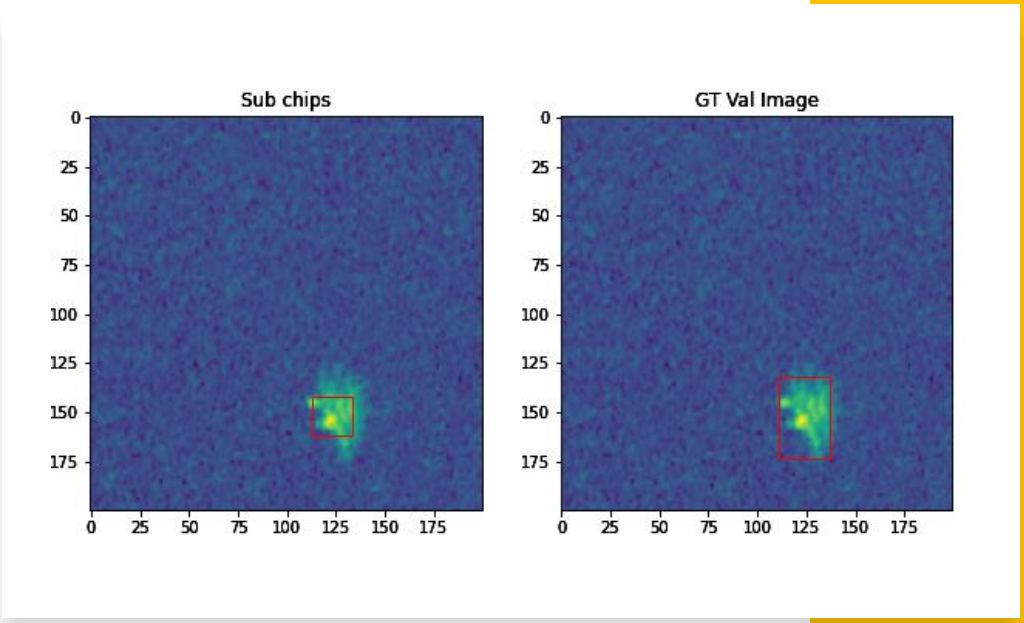
Current Ranking on xView Leaderboard

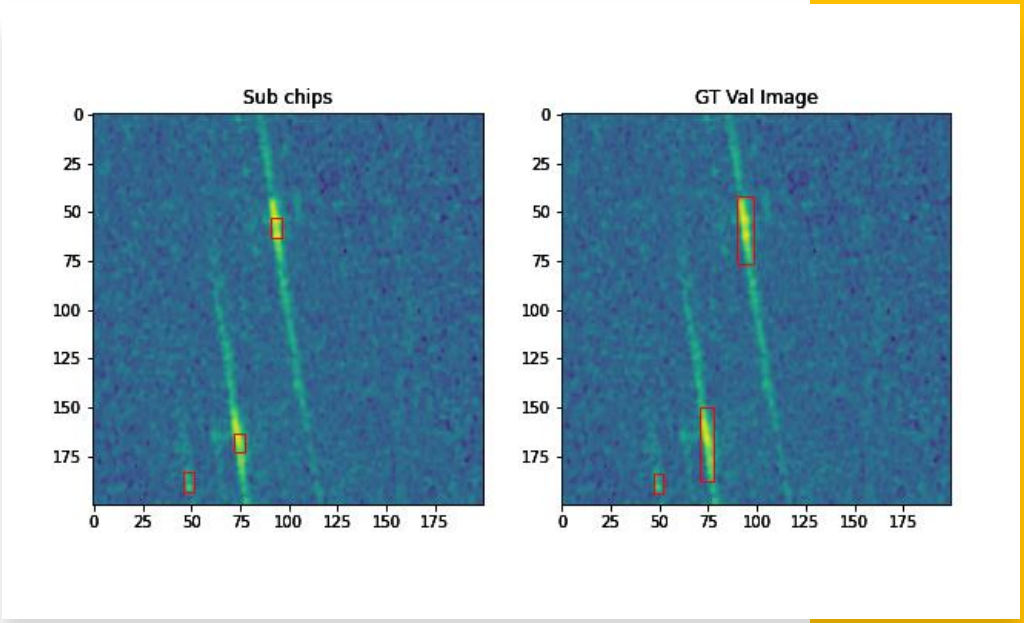
38	zolo	0.34255	0.59794	0.10651	0.86260	0.78818	0.10708	✗
39	Anonymous	0.33809	0.58400	0.08904	0.82452	0.49295	0.48811	✗
40	Anonymous	0.31778	0.56064	0.04148	0.83093	0.47406	0.48765	✗
41	Stanford 325B	0.22097	0.45577	0.12773	0.57027	0.18704	0.53909	✗
42	xView3 reference model	0.19003	0.42585	0.12092	0.71216	0.39804	0.00000	✓
43	Anonymous	0.18909	0.27082	0.05557	0.92336	0.83351	0.67856	✗
44	Anonymous	0.16636	0.27082	0.05557	0.80975	0.72614	0.47986	✗
45	naivelogic	0.16304	0.24289	0.16639	0.92127	0.75224	0.51647	✗
46	david	0.12163	0.23512	0.08977	0.81432	0.68243	0.00000	✗
47	Anonymous	0.04991	0.14223	0.02254	0.73197	0.00000	0.00000	✗

Challenges and Setbacks

Azure Cloud Challenges

- High Azure storage cost on uploading initial dataset
- Azure credits remaining only allowed for K80 GPU (V100 strongly preferred)
- K80 memory was insufficient to process the large SAR images requiring alternative chipping methods





Challenges and Setbacks

Project Execution Learnings

- Inconsistence with Ship ground truths in challenge datasets
- Image Splitting - Should have created dataset 1024x1024 rather than 200x200
- Spent too much time on dataloader and not enough time iterating on model selection and hyperparameters
- High knowledge barrier to utilize SAR images and ancillary images