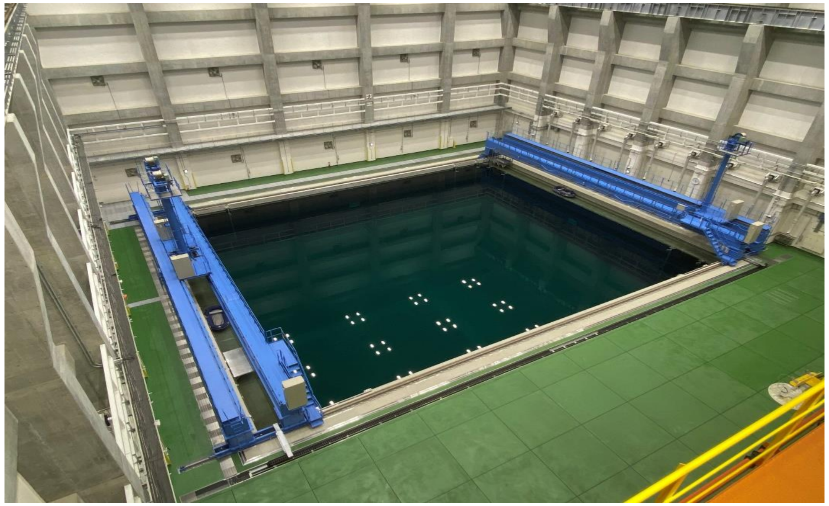
AI-based fishing net detection technology using optical sensors

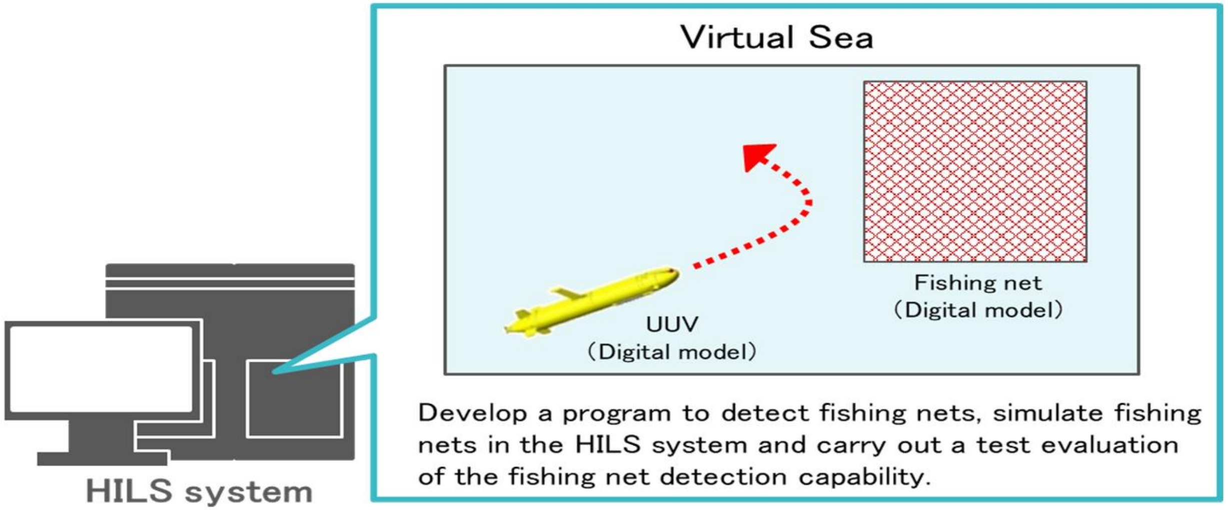
IMETS
IWAKUNI Maritime Environment Test & Evaluation Satellite
NSRC , ATLA

Uncrewed Maritime Vehicles Evaluation Research Section IMETS,ATLA
Defence Science Technology, Australia

- Obstacle detection and avoidance is a necessary for UUV. Fishing nets are difficult to detect and knowing the location of fishing nets is important for avoidance action.
- Therefore, research is being conducted to study the improvement of the accuracy of the fishing net detection program, as well as to conduct tests on the Hardware in the simulator(HILS) at the Iwakuni Maritime Environment Test & Evaluation Satellite(IMETS).

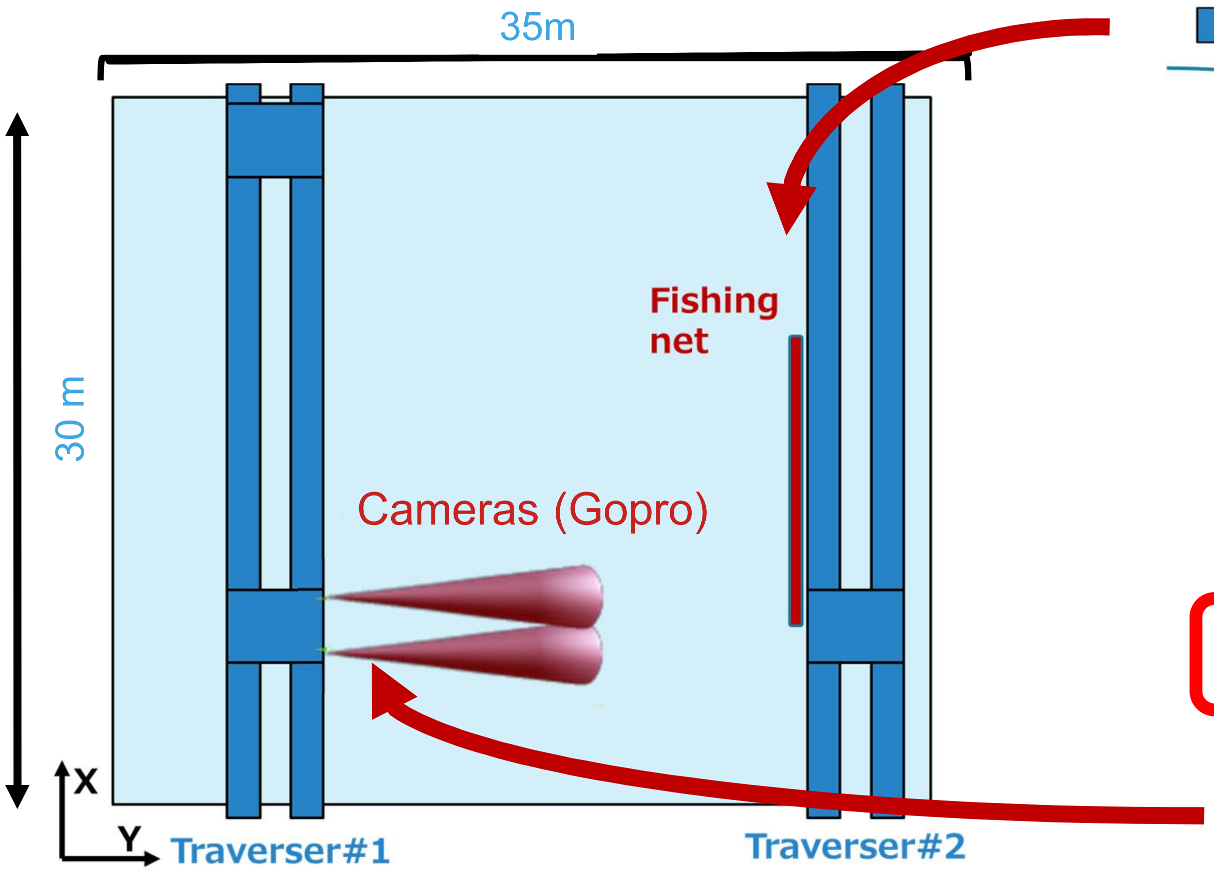
Large Tank in IMETS – 30m width X 35 m long X 11m depth (Plan view)



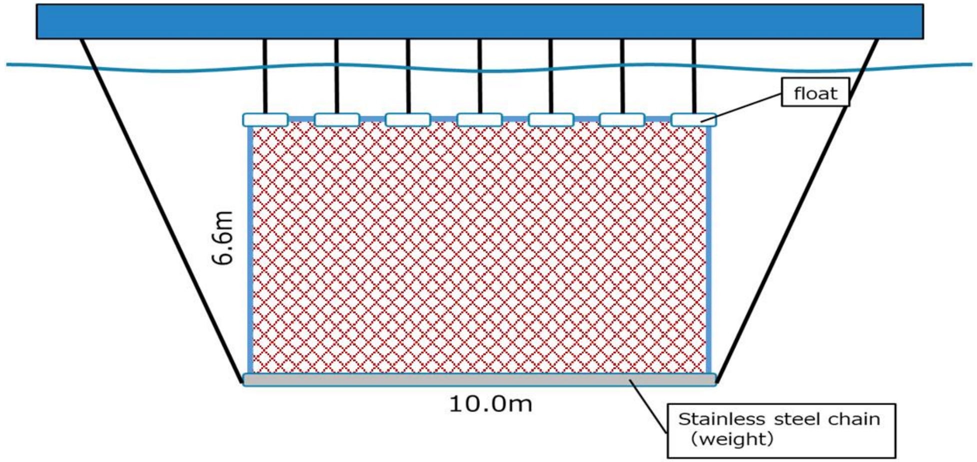


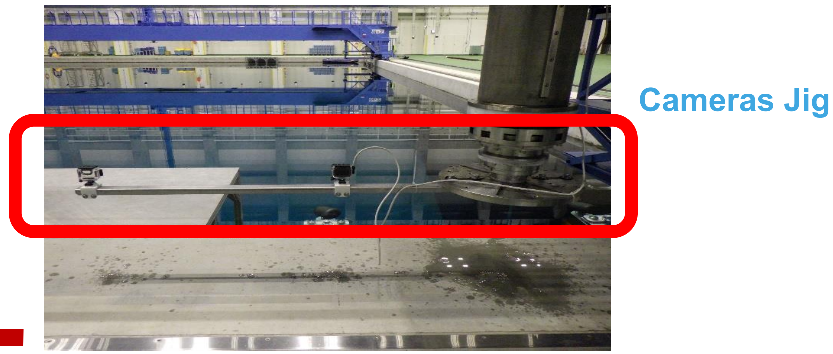
Research Questions
1. How to detect fishing nets using Machine Learning (ML) models object detection models?
2. How to integrate developed ML models into the HILS system?

Data Acquisition



Plan view

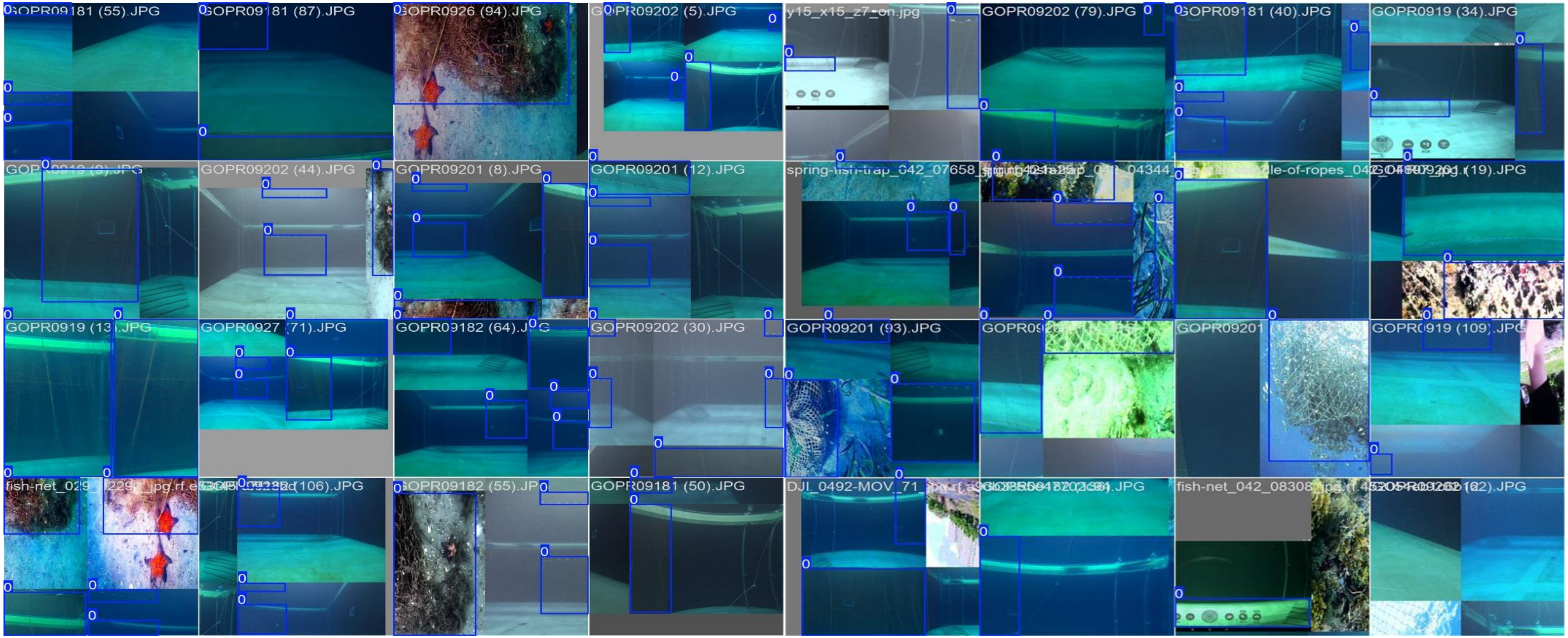




Machine learning models – Dataset

- 1678 images from IMETS Water Tank and various Fishing nets including 478 public fishing net images

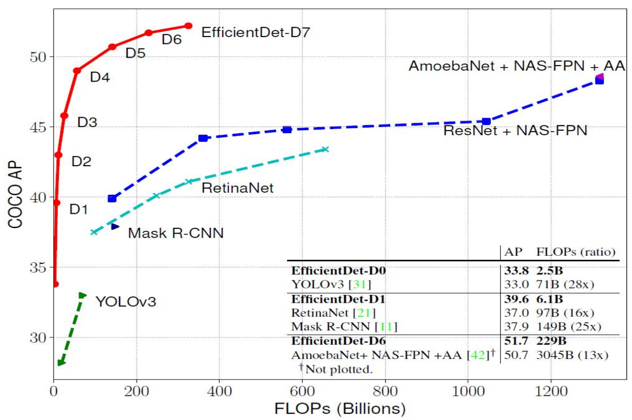
Training Set (Overlapping images but they are treated as separate images)



Overview of ML models used in this study

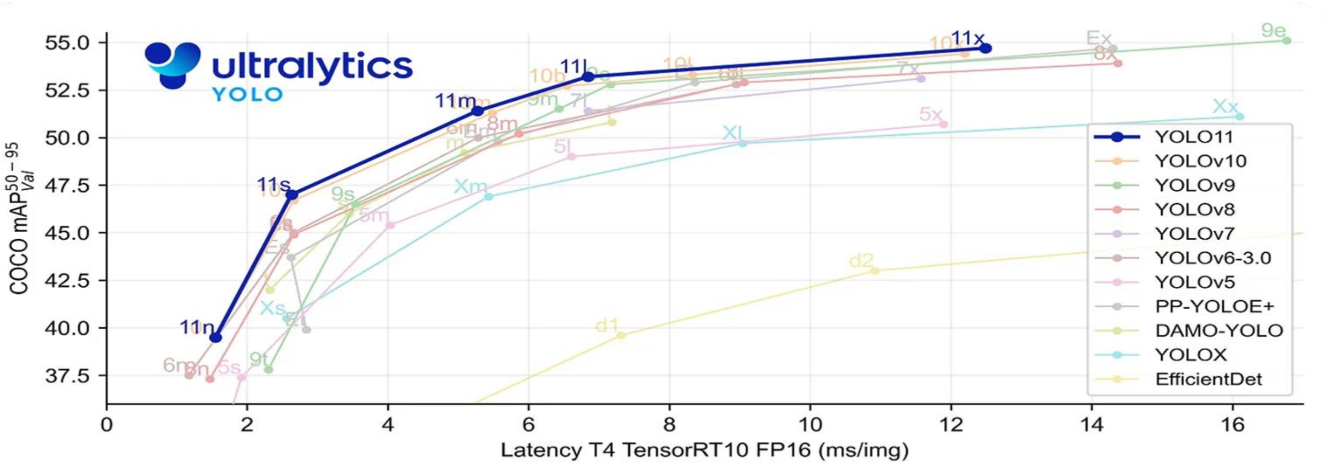
Model	YOLO (You Look Only Once model)	Faster R-CNN (Region-based Convolutional Neural Networks) RESNET (Residual Networks)	Faster R-CNN Swin Transformer (Shifted Window Transformer)	Efficientdet
Type	Single Stage	Two stage	Two stage	Single Stage
Speed	Very good	good	good	good
Accuracy	Very good	good	good	Very good
Feature extraction	CSP (Cross-Stage Partial)	RESNET34,50,101	SWIN Transformer	Efficient Net

Accuracy



Speed

Accuracy



Benchmarking YOLO and other models

Speed

Results – Faster R-CNN with a RESNET feature extractor

Faster R-CNN has a two-stage architecture. In the first stage, a Region Proposal Network (RPN) scans the image and proposes regions that might contain objects. Then, in the second stage, each proposed region is classified and its bounding box is refined.
RESNET – 34, RESNET-50, RESNET-101, RESNET – 152 are tested to extract features.

RESNET – 101

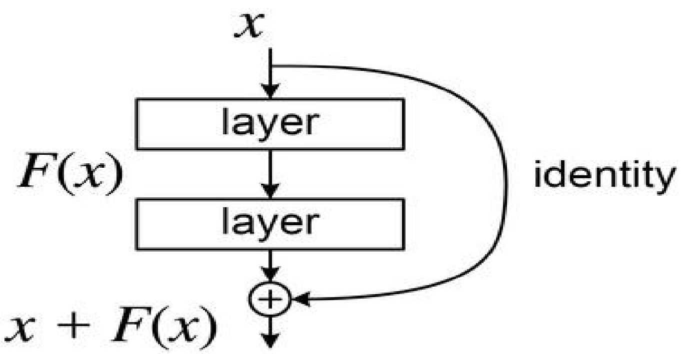
epoch	train_loss	valid_loss	mAP@IoU >0.5	mAP@IoU 0.5:0.95
0	2.270631	1.961415	0.493934	0.274304
1	1.213979	1.220275	0.733017	0.507363
2	0.8829	1.039559	0.828376	0.626742
3	0.910959	1.003706	0.841134	0.640351

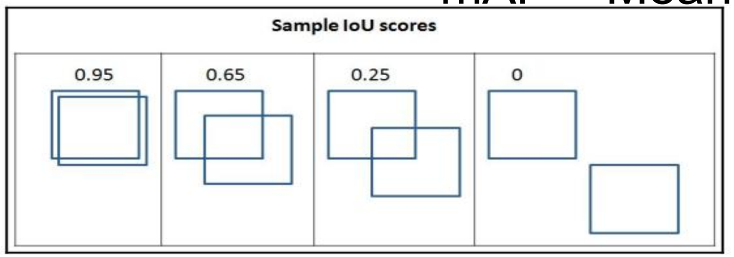
RESNET – 152

epoch	train_loss	valid_loss	mAP@IoU >0.5	mAP@IoU 0.5:0.95
0	1.635582	1.256892	0.631945	0.42013
1	0.829538	0.82414	0.836847	0.630583
2	0.655427	0.735213	0.877903	0.693057
3	0.612222	0.712022	0.872079	0.683245

Increasing Accuracy (mAP)

mAP – Mean Average Precision





Intersection over Union (IoU) measures the **overlap** between predicted and ground truth boxes

Results – Faster R-CNN with a Shifted Window Transformer

Using the Swin transformer has shown lower accuracy compared to RESNET. However, this may depend on the dataset and application. The Swin Transformer's architecture is built on a combination of hierarchical design and window-based self-attention for efficient working and feature extraction.

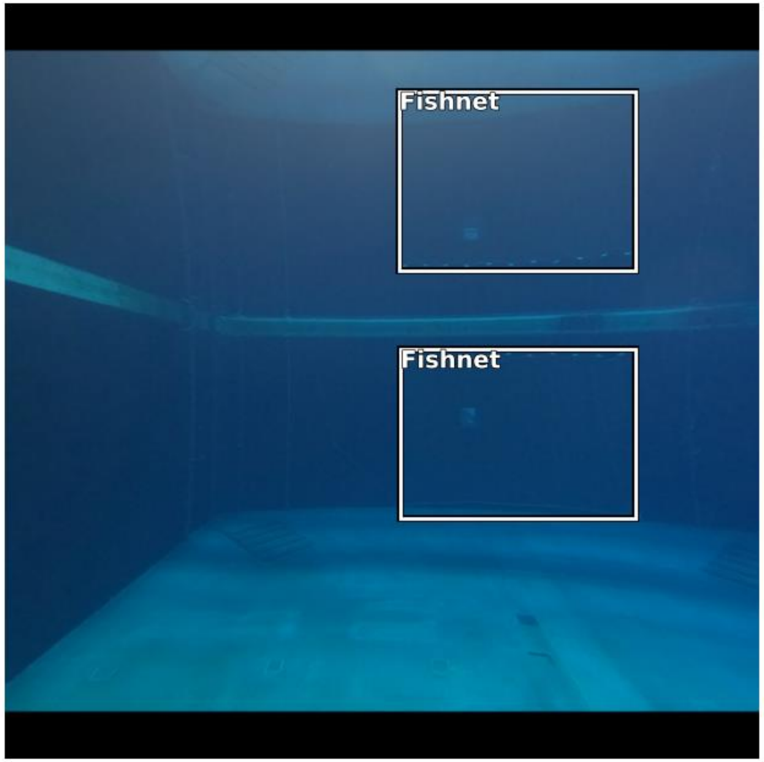
Fasterrcnn_swin_base

epoch	train_loss	valid_loss	mAP@IoU >0.5	mAP@IoU 0.5:0.95
0	0.215663	0.211138	0.740475	0.314077
1	0.12426	0.122734	0.790212	0.392322
2	0.16846	0.162612	0.865228	0.479912

Fasterrcnn_swin_large

epoch	train_loss	valid_loss	mAP@IoU >0.5	mAP@IoU 0.5:0.95
0	0.252893	0.225944	0.559516	0.211596
1	0.193553	0.164562	0.782309	0.387796
2	0.169299	0.151057	0.903895	0.526077
3	0.146452	0.143081	0.922239	0.563626

Increasing Accuracy (mAP)



Results – Efficientdet models

EfficientDets are developed based on the advanced backbone and a new scaling technique Implementing an EfficientNet backbone, it is possible to achieve much better accuracy.

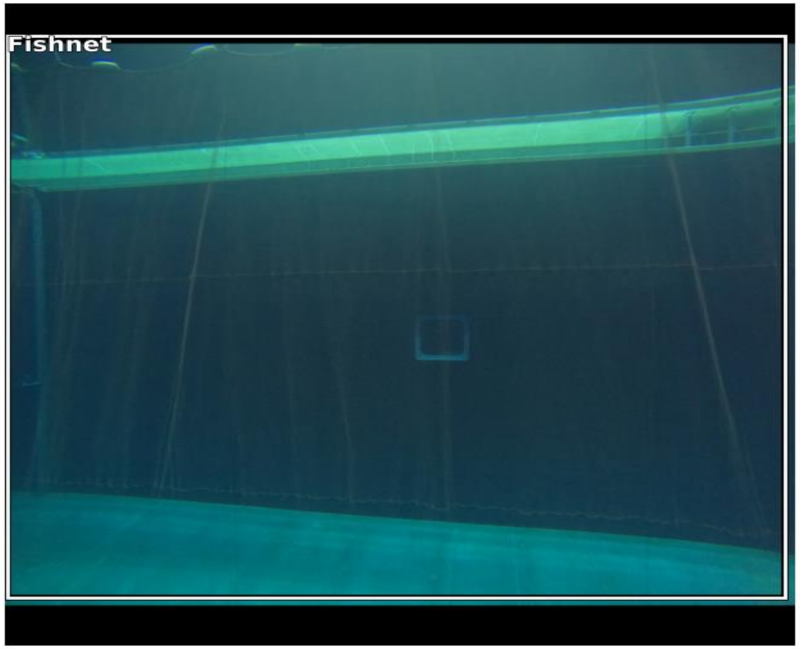
Efficientdet_d1

epoch	train_loss	valid_loss	mAP@IoU >0.5	mAP@IoU 0.5:0.95
0	0.759083	0.689068	0.835724	0.646752
1	0.793325	0.712323	0.853231	0.673431
2	0.832396	0.769463	0.871815	0.694135

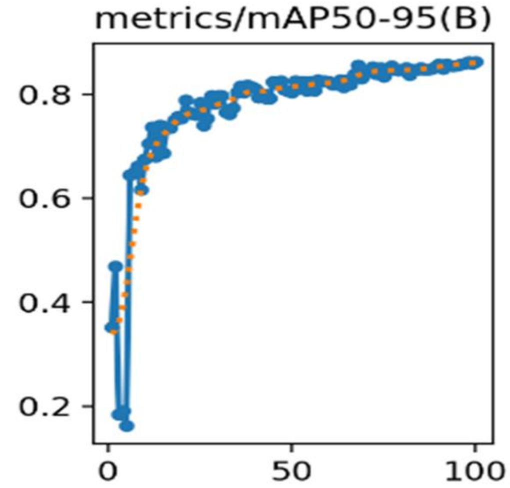
Efficientdet_d4

epoch	train_loss	valid_loss	mAP@IoU >0.5	mAP@IoU 0.5:0.95
0	0.800551	0.887823	0.885345	0.690316
1	0.782212	0.871231	0.903112	0.703145
2	0.754045	0.867921	0.912585	0.722585

Increasing Accuracy (mAP)



Results – Yolov11 model predictions



- Overall Accuracy metric after 0 -100 epochs

  (mAP0.50-0.95) ~ 0.862

Results – Summary

Machine Learning Models	Mean Average Precision@IoU 0.5:0.95
Fastrcnn_resnet101	0.640
Fastrcnn_resnet152	0.683
Fasterrcnn_swin_base	0.480
Fasterrcnn_swin_large	0.564
Efficientdet_d1	0.694
Efficientdet_d4	0.722
Yolo8n	0.838
Yolo11n	0.862

- YOLO11 model shown the best result
- Results reflects the ML detection potential in varying test positions and fishing net colours
- It could deploy within HILS potentially.

Next steps

- Incorporating the impact of angles and lighting conditions
- Further Hyper parametric optimisation

iMETS

IWAKUNI Maritime Environment Test & Evaluation Satellite

NSRC , ATLA