Detecting Marine Species in Echograms via Traditional, Hybrid, and Deep Learning Frameworks

Tunai Porto Marques[1]*, Alireza Rezvanifar[1]*, Melissa Cote[1], Alexandra Branzan Albu[1], Kaan Ersahin[2], Todd Mudge[2], Stéphane Gauthier[3]

[1]University of Victoria, BC, Canada [2]ASL Environmental Sciences, Victoria, BC, Canada [3]Fisheries and Oceans Canada, Sidney, BC, Canada

*Equal contribution

1. Motivation

Context:

- Marine biologists use acoustic images to monitor underwater sites in a thorough and non-invasive manner.
- Acoustic surveys are typically carried out using multifrequency echosounders, such as the Acoustic Zooplankton Fish Profiler (AZFP) [1].
- Acoustic data are displayed in 2D (vertical axis: depth or range, horizontal axis: time), producing an image called echogram. Echograms use colour codes representing the volumetric backscatter strength of fish, plankton, etc.

✗ Echograms are typically interpreted with manual or semiautomatic methods, which are time-consuming and prone to inter-expert disagreements.

Solution:

✓ Machine learning (ML) methods can provide solutions for the automatic detection of marine species in echograms.

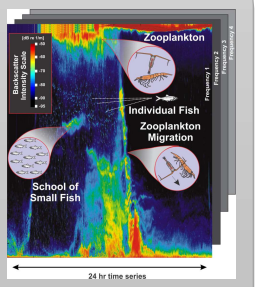


Illustration of an AZFP-generated echogram [1].

[1] D. Lemon, P. Johnston, J. Buermans, E. Loos, G. Borstad, and L. Brown, "Multiple-frequency moored sonar for continuous observations of zooplankton and fish," in 2012 Oceans. IEEE, 2012, pp. 1–6.

2. Contributions

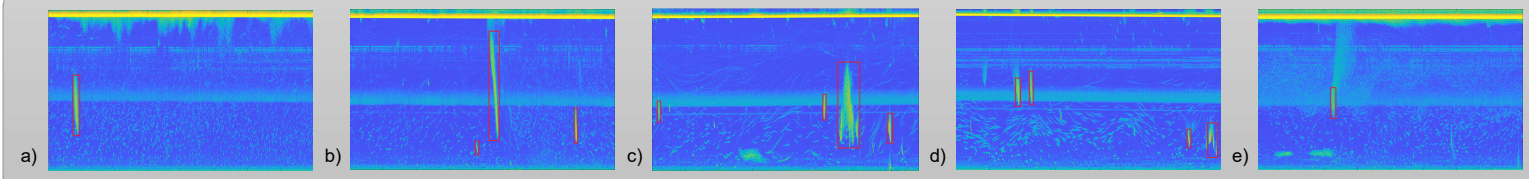
Theoretical: We provide a comprehensive comparative study that encompasses the whole spectrum of ML approaches.

- ➢ First of its kind for the application of marine species detection in echograms.
- ➢ We focus on schools of herring for a well-crafted and bounded study.

Practical: We address the problem of detecting biological targets in echograms by proposing a functional and high-performing framework that can be readily scaled to accommodate new species.

Our goal is to take advantage of decades of contributions and expertise from traditional approaches in fisheries acoustics and to assess what parts are still relevant in the DL era.

3. Challenges



Schools of herring (red boxes): typically elongated shape (a) that can vary in size (b) and shape (c), sometimes found in close proximity (d), with boundaries that may appear fuzzy due to bubbles (e).

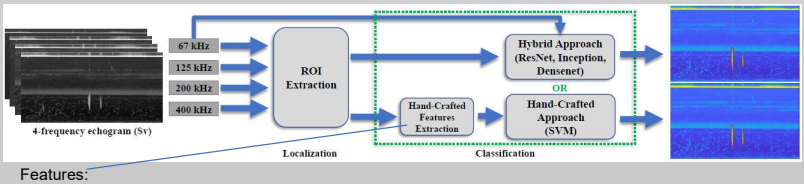
4. Proposed Method

Three compared approaches:

Approach	Localization	Classification
1) Hand-crafted	Custom ROI extractor	ML image classifier
2) Hybrid	Custom ROI extractor	DL image classifier
3) End-to-end	DL object detector	DL object detector

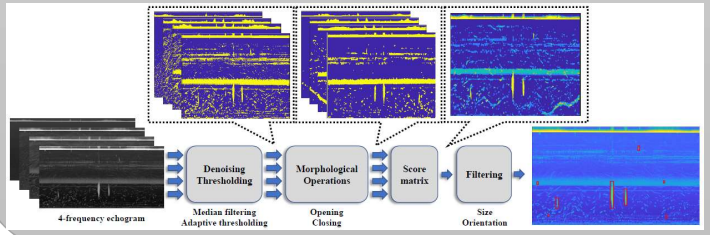
ROI = region of interest, ML = machine learning, DL = deep learning

Block diagram of hand-crafted and hybrid approaches:

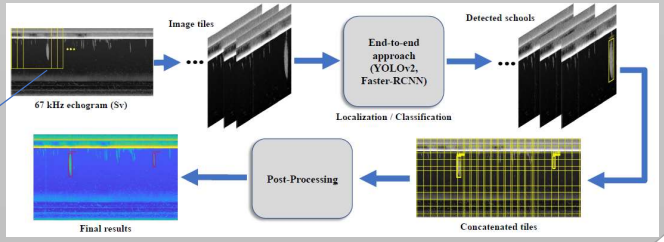


Features:
energetic (mean intensity), morphometric (ratio between major and minor axis, eccentricity, circularity)

ROI extraction:



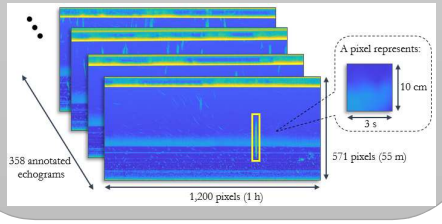
Block diagram of end-to-end approach:



5. Dataset

Dataset:

- 358 annotated echograms.
- Collected by Fisheries and Oceans Canada with ASL's AZFPs in the Okisollo channel, BC, Canada between May and October (2015-16).
- Training/valiation set: 80% (286 echograms with 617 instances of schools of herring).
- Test set: 20% (72 echograms with 128 instances of schools of herring).



6. Results

Quantitative evaluation:

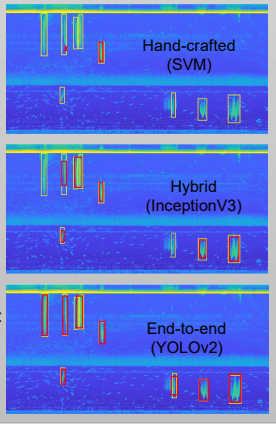
Method	IoU	P	R	F1
1) Hand-crafted (SVM)	0.3	45.41	65.63	53.67
1) Hand-crafted (SVM)	0.5	41.62	60.16	49.20
2) Hybrid (ResNet-50	0.3	72.37	85.94	78.57
2) Hybrid (ResNet-50)	0.5	62.50	74.20	67.80
2) Hybrid (DenseNet-201)	0.3	64.37	87.50	74.17
2) Hybrid (DenseNet-201)	0.5	55.75	75.78	64.24
2) Hybrid (InceptionV3)	0.3	**78.99**	85.16	**81.95**
2) Hybrid (InceptionV3)	0.5	**68.12**	73.44	70.68
3) End-to-end (YOLOv2)	0.3	73.08	**89.06**	80.28
3) End-to-end (YOLOv2)	0.5	67.31	**82.03**	**73.94**
3) End-to-end (Faster R-CNN)	0.3	66.90	74.22	70.37
3) End-to-end (Faster R-CNN)	0.5	45.07	50	47.41

IoU = intersection over union, P = precision, R = recall, F1 = F1-score

The best detection depends on the user's priorities:

- Best precision: hybrid (InceptionV3).
- Best recall: end-to-end (YOLOv2).
- Best high-quality: end-to-end (YOLOv2).

Qualitative evaluation:



- Generally, the hybrid and end-to-end approaches perform better than the hand-crafted one.
- The performance of the hand-crafted and hybrid approaches is limited by the ROI extractor's performance.
- When two schools are very close: very difficult case for all approaches.

7. Conclusion

An end-to-end DL-based framework is preferable to other learning approaches for marine species detection in echograms: performs equally or better than time-consuming hand-crafted methods even with limited training samples, and is readily scalable to new species.

Acknowledgements

This work was supported by NSERC Canada and ASL Environmental Sciences through the Engage and Engage Plus Grants programs.