UNIVERSITY OF CAMBRIDGE
DEPARTMENT OF ENGINEERING

Agri FoRwArdS
EPSRC Centre for Doctoral Training in Agri-Food Robotics

Deep Learning for Antarctic Krill staging and morphology analysis from high resolution image pairs

Mazvydas Gudelis, Michal Mackiewicz

Introduction

British Antarctic Survey researchers conduct expeditions to assess the effects of global warming and overfishing. Antarctic Krill are vital to the food chain as whales, seals, penguins, albatrosses, petrels, squid, and many more feed on this prawn-like creature [1]. Given the potential for human error during these rigorous missions, automation offers a promising solution. The advent of deep learning has notably advanced Computer Vision, enabling nuanced image differentiation. Employing such sophisticated models could expedite research, facilitating precise animal identification and parameter assignment, thereby optimizing fieldwork efficiency and accuracy.

Methods

Our computer vision pipeline, comprised of modular components, facilitates the labelling, detection, preprocessing, and classification of Antarctic Krill imagery. Initial processing via a web application normalizes and sequences bounding box labels, accommodating variations in krill and box sizes. The resultant images, as shown in Figures 1 and 2, retain true size representation and offer composite views of each specimen.

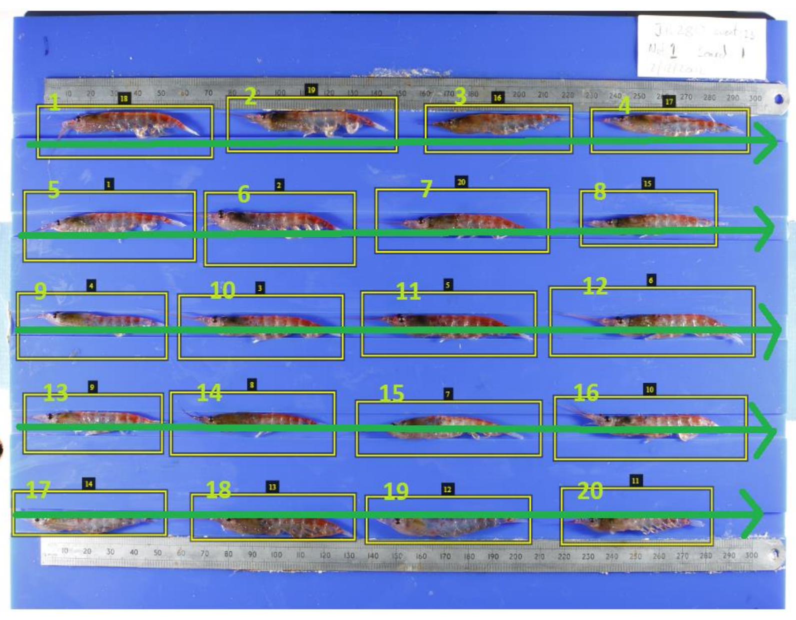
Fig 1: Krill detection from the "Krill tool" web application.

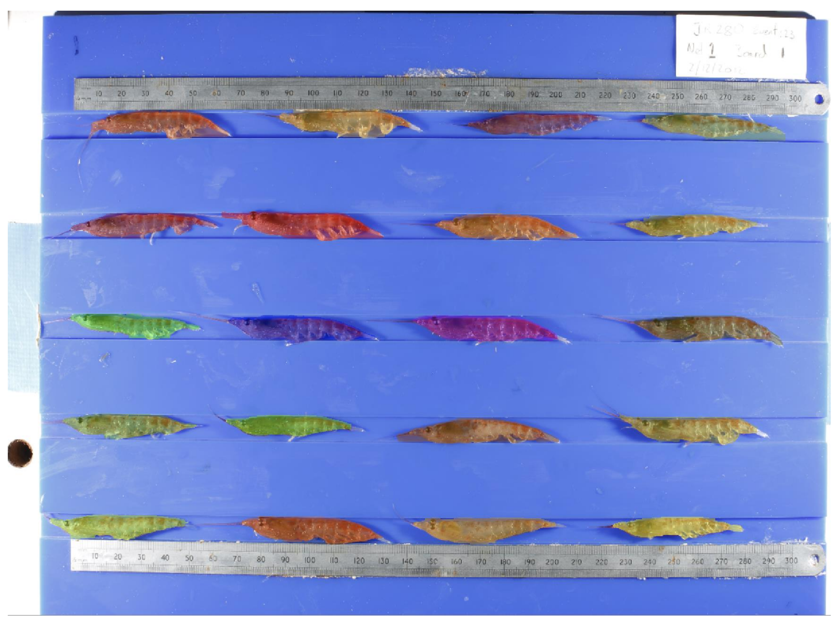
Fig 2: Accurate krill instance mask generation.

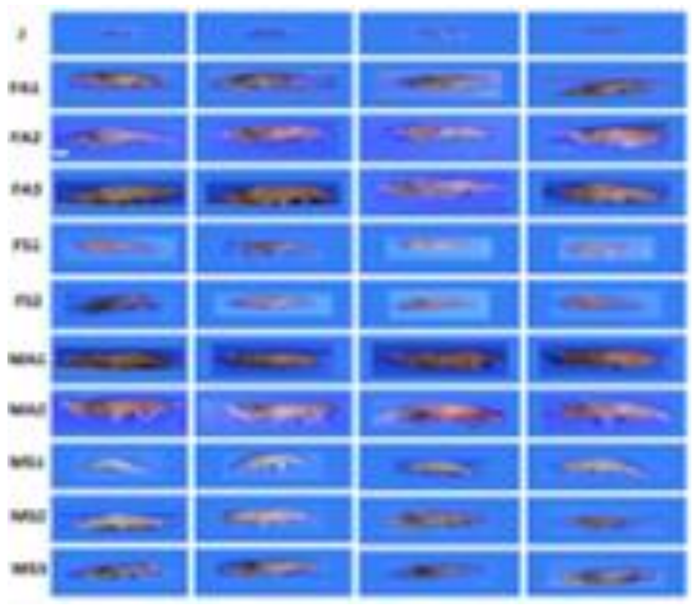
Fig 3: Krill sample visualization of the lateral view.

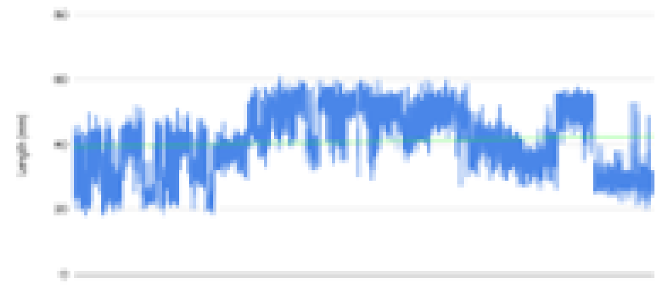
Fig 4: Maturity sample distribution

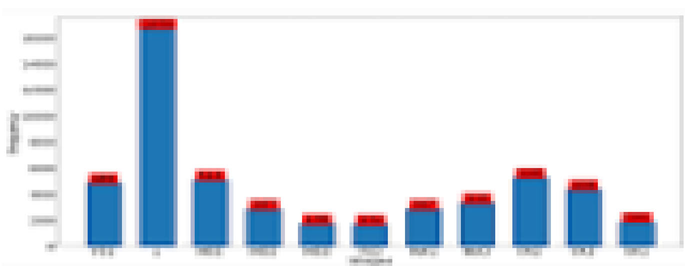
Fig 5: Length sample distribution.

With approximately 10,500 instances, each krill carries encoded maturity stage and length parameters, as depicted in Figures 3, 4, 5. Maturity stage is encoded as F/M - female/male, A/S - adult/subadult, 1/2/3 - advanced matureness indicator. Length is measured in millimeters with each specimen having a body length of 20-60mm.

Bounding box coordinates and instance masks constitute four-dimensional input vectors across various dataset views . This framework incorporates additional contextual information via box coordinates, while ensuring precise object representation through instance masks, thereby optimally balancing the visual information and object outline bounds.

$$\mathcal{L}_{CCE}(\hat{y}, y) = -\sum_{k}^{K} y^{(k)} log(\hat{y_i})^{(k)}$$

$$L_{RMSE}(\hat{y}, y) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2}$$

$$\mathcal{L}_{new} = 1(\mathcal{L}_{RMSE}) + 2(\mathcal{L}_{CCE})$$

The system leverages dual losses—categorical cross-entropy and root mean square error—to concurrently generalize two tasks from a single input image. Owing to the tasks' inherent correlation, an aggregate loss is employed via a hard parameter sharing approach [3], thereby promoting the evolution of task-specific backbone weights in a balanced manner.

Results


Fig 6: Multi-task learning result matrix showing 73% accuracy.

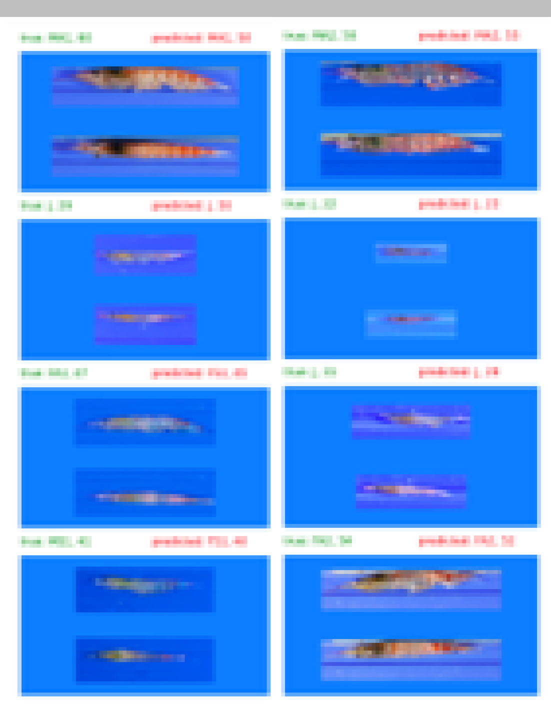
Fig 7: Prediction visualisation.

Resolution	Concatenated	Fused
340x200	63.69/2.40	62.86/2.31
680x400	65.45/2.14	65.61/1.97
1020x600	65.23/2.26	66.17/1.97
1360x800	69.78/1.65[2]	68.66/1.90
1700x1000	69.19/1.80[3]	73.28/1.50[1]

Table 1: Multi-task learning results. Best 3 models highlighted 1 , 2 , 3 . Values display the highest achieved balanced accuracy (higher is better) and root mean square error (lower is better) for the test set.

The proposed computer vision pipeline achieves a highest accuracy of 73% for maturity classification while maintaining a 1.5 mm error for the length estimation task. Misclassification mostly occurs between neighbouring maturity stages such as MS2-MS3, MA1-MA2, FA2-FA3, etc. These stages are difficult to distinguish even by an experienced researcher and often involves the use of a microscope. The experiment results indicate that the produced models can be integrated with an application which would drastically improve the speed of Antarctic Krill analysis.

References

[1] Flores, H., Atkinson, A., Kawaguchi, S., Krafft, B. A., Milinevsky, G., Nicol, S., ... & Werner, T. (2012). Impact of climate change on Antarctic krill. Marine Ecology Progress Series, 458, 1-19.
[2] N. Shazeer, K. Fatahalian, W. R. Mark and R. T. Mullapudi, "HydraNets: Specialized Dynamic Architectures for Efficient Inference," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 8080-8089, doi: 10.1109/CVPR.2018.00843.
[3] Caruana, R. "Multitask learning: A knowledge-based source of inductive bias." Proceedings of the Tenth International Conference on Machine Learning. 1993.
[4] Huang, Gao & Liu, Zhuang & van der Maaten, Laurens & Weinberger, Kilian. (2017). Densely Connected Convolutional Networks. 10.1109/CVPR.2017.243.

UKRI Engineering and Physical Sciences Research Council
UNIVERSITY OF CAMBRIDGE DEPARTMENT OF ENGINEERING
UEA University of East Anglia
UNIVERSITY OF LINCOLN