Automated detection of wildlife in proximity to marine renewable energy infrastructure using machine learning of underwater imagery

Mckenzie Love, Aiswarya Vellappally, Pierre Roy, Kate Smith, Gavin McPherson, and David P. Gold

Abstract— **The use of automated image analysis for monitoring wildlife interactions with marine renewable energy infrastructure can vastly reduce the time required to extract usable data from underwater imagery compared to manual expert processing. We present a novel industry-ready image processing workflow for automated wildlife detection developed using 1000+ hours of underwater video footage obtained by Nova Innovation Ltd. from their operational tidal stream turbine array at Bluemull Sound in Shetland, Scotland. The workflow includes object detection through advanced image analysis, image classification using machine learning, statistical analyses, and automated production of a summary report. Blind tests were undertaken on a subset of videos to quantify and iteratively improve the accuracy of the results. The final iteration of the workflow delivered an accuracy of 80% for the identification of marine mammals, diving birds and fish when a three-category (wildlife, algae, and background) classification system was used. The accuracy rose to 94.1 % when a two-category system was used, and objects were classified simply as 'target' or 'non-target'. The accuracy and speed of the workflow can be improved through expanding the initial training dataset of images with different species and water conditions. Application of this workflow significantly reduces manual processing and interpretation time, which can be a significant burden on project developers. Automated processing provides a subset for more focused manual scrutiny and analysis, while reducing the overall size of dataset requiring storage. Auto-reporting can be used to provide outputs for marine regulators to meet monitoring reporting conditions of project licences.**

Keywords—**automated reporting, environmental monitoring, machine learning, marine mammals.**

I. Introduction

ENVIRONMENTAL interactions of marine renewable energy (MRE) projects are challenging and costly to monitor, and questions remain about their potential effects on the physical and biological environment [1]. This uncertainty and the paucity of monitoring data from MRE projects increases the perception of risk about potential impacts which can confound regulatory decision-making and hamper efficient consenting (permitting) of projects [2]. A key concern and consenting risk for tidal energy technologies is the potential for injury to marine wildlife through collisions with moving parts of turbines [3]. To understand the real nature of this risk, information on animal presence and behaviour around operational tidal turbines is required [4, 5].

Large volumes of underwater imagery can be collected in a short time using submarine cameras, but there is a bottleneck at the processing stage required to extract usable biological and environmental data from imagery [6]. Manual image processing is also subject to observer bias, with inconsistency generated among and within observers [7]. Therefore, the automated processing of large volumes of environmental data acquired from submarine monitoring and the use of machine learning algorithms like convolutional neural networks to identify the presence of marine wildlife with MRE infrastructure are powerful tools for assessing the environmental response to MRE infrastructure [8].

A dataset of underwater video footage obtained by Nova Innovation Ltd. from their operational tidal stream turbine array at Bluemull Sound in Shetland, Scotland was

Sponsor and financial support acknowledgement: This work was internally funded by CGG Services (UK) Ltd. and Nova Innovation Ltd.

M. Love is with CGG Services (Australia) Pty Ltd., 1 Ord Street West Perth, WA 6005. Australia (e-mail: mckenzie.love@cgg.com).

A. Vellappally is with CGG Services (UK) Ltd., Crompton Way Manor Royal Estate Crawley, West Sussex RH10 9QN United Kingdom (e-mail: aiswarya.vellappally@cgg.com).

P. Roy is with CGG Services (UK) Ltd., Crompton Way Manor Royal Estate Crawley, West Sussex RH10 9QN United Kingdom (e-mail: pierre.roy@cgg.com).

K. Smith is with Nova Innovation Ltd., 45 Timber Bush, Leith, Edinburgh EH6 6QH United Kingdom (e-mail: kate.smith@novainnovation.com).

G. McPherson is with Nova Innovation Ltd., 45 Timber Bush, Leith, Edinburgh EH6 6QH United Kingdom (e-mail: gavin.mcpherson@novainnovation.com).

D. P. Gold is with CGG Services (UK) Ltd., Tyn Y Coed, Pentywyn Rd, Llanrhos LL30 1SA United Kingdom (e-mail: david.gold@cgg.com).

Digital Object Identifier: https://doi.org/10.36688/ewtec-2023-623

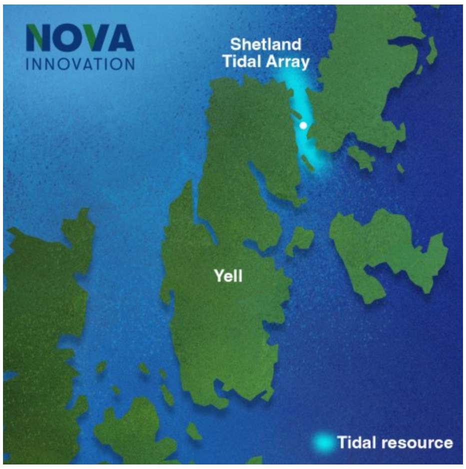

Fig. 1. Map of the location of Nova Innovation's Shetland Tidal Array site at Bluemull Sound between the islands of Yell and Unst. (Image provided by Nova Innovation Ltd.)

Fig. 2. Image showing size and scale of Nova's M100-D tidal turbine (Image provided by Nova Innovation Ltd.)

used to develop a workflow and associated algorithms to automatically filter many hours of underwater video, remove unwanted footage, and extract only video containing marine mammals, diving birds or fish.

Nova Innovation's Shetland Tidal Array at Bluemull Sound (Fig. 1) was the world's first offshore tidal energy array. The project has been operational since 2015, powering the Shetland grid and delivering reliable, predictable power to homes and businesses. Following the success of the first three turbines and the resulting positive impact on the local economy and Shetland's carbon emissions, Nova doubled the capacity of the Shetland Tidal Array to six turbines.

Nova's M100-D tidal turbine is a 2-bladed, horizontal axis device which uses gravity-based foundation that sits on the seabed. It is approximately 14m in height, with a rotor diameter of 8.5m (Fig. 2). Nova has a regulatory obligation to monitor and report on the presence of wildlife in proximity to its turbines to Marine Scotland under conditions of the project licences. To do this, it uses motion-detection triggered video cameras mounted to the turbines. To date, Nova has recorded fewer than 30 occurrences of diving birds or marine mammals close to the turbines in seven years of operation, while fish of the genus *Pollachius* are relatively common. However, 10s of 1000s of hours of submarine video footage have been recorded, which is often triggered by the motion of marine algae and other superfluous marine detritus in the tidal current.

This presents a processing and analysis, and data storage challenge whereby the majority of video footage contains unnecessary information, requiring manual review to identify and record the presence of marine wildlife of interest (mammals, diving birds and fish). Time spent reviewing hours of underwater footage has an associated monetary and resource cost to Nova which could be better distributed elsewhere. It also restricts the proportion of the total dataset that can be sampled and analysed. Therefore, the development of an automated process to automatically scan and filter video files, identify frames which contain wildlife of interest and produce statistical analyses in an auto-generated report is extremely valuable to both MRE developers and regulators.

II. METHODS

A. *Nova Innovation dataset*

Two sets of videos were received from Nova Innovation from its turbine-mounted cameras at the Shetland Tidal Array. Dataset A comprised videos recorded in November 2015, whilst videos from dataset B were videos from March to April 2016. Videos were obtained from motion-triggered cameras that were positioned on the side and top of the first installed turbine (the camera position will hereinafter be referred to as T1SIDE and T1TOP, respectively). Each video comprises a few seconds of footage before the initial motion was detected to ensure the entire interaction was captured. The lengths of the videos ranged from 10 seconds up to 15 approximately minutes. Dataset A exhibited a large amount of wildlife occurrences consisting of seals, diving birds and shoals of fish, whereas dataset B had relatively limited wildlife occurrences with just a few occurrences of diving birds and fish. The videos were of varying quality based on luminosity and hue of background (e.g. blue or green), turbidity and clarity of the water column, and degree of biofouling on the lens or close to the camera. There were a total of 263 and 668 videos from the T1SIDE and T1TOP positions, respectively, from dataset A, whereas dataset B contained 102 and 90 videos from the T1SIDE and T1TOP positions, respectively (Table 1).

B. *Machine learning model*

The machine learning model used here is a pretrained convolutional neural network called EfficientNet [9], which is widely used for image classification tasks as it has learned representations for a variety of real-life objects. The model was trained further on our dataset to enable it to correctly identify videos containing wildlife occurrences [10].

C. *Model training, validation and testing*

Videos from dataset A were used in model training, so were manually labelled and sorted into three categories of 'wildlife', 'detritus' and 'background' depending on the content. The 'wildlife' category consisted of videos containing seals, diving birds and fish, whereas the 'detritus' category contained objects such as large kelp or small pieces of unidentifiable plant-like detritus. The 'background' category was for videos that did not contain any objects of interest (i.e., contained no wildlife or detritus).

Videos from dataset A were sorted into further groups to use for model training or validation, respectively. Some videos from dataset B containing diving birds were used to supplement the training and validation process given the limited number of videos containing wildlife occurrences.

Model testing was conducted on videos from dataset B as it contained videos from a different temporal period and was deemed different enough to combat bias to the training datasets. Prior to testing the model, videos from dataset B were manually labelled to categorize videos into 'wildlife', 'background' or 'detritus' to determine the success of the model (Fig 3). However, due to many videos being 15 minutes long in dataset B, a sampling method of watching 5 seconds of video every 30 seconds was used in the labelling process.

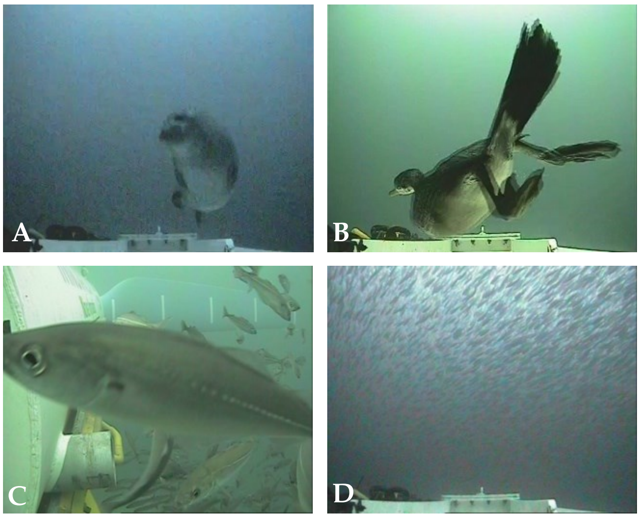

Fig 3. Examples of marine wildlife captured by motion-detection turbine-mounted cameras at Nova Innovation's Shetland Tidal Array, Bluemull Sound, Shetland, which were used in model training. A) Grey Seal, B) European Shag, C) Individual fish close to camera D) Large shoal of fish in the distance.

D. *Statistical analysis and auto-report generation*

To improve the efficiency of reviewing model results and to provide a tool to assist in Nova Innovation's reporting obligations under project licences, an auto-generated results report was created. The report contained the videos identified by the model to contain wildlife interactions and some statistics. Six frames per video were generated to include in the report, which was found to be a sufficient number of frames to capture a snapshot of the interaction with the turbine. The number of frames can be adjusted to suit the purpose of the report. The number of videos identified to contain wildlife interactions, as well as the amount of storage space potentially saved by using the model, is also included in the report (Fig 4). There are several additional statistics that could be calculated within the automated workflow and included the report. For example, a time analysis of the number of interactions recorded per month or season, or the frequency of bird detections in relation to fish presence.

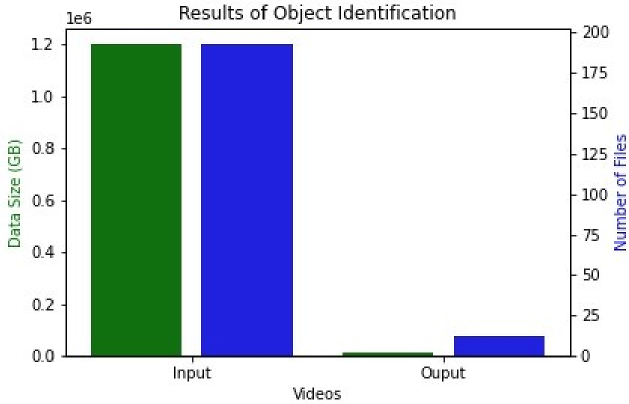

Fig 4. Example of statistics included in the final report. The data size (amount of storage) in green and the number of files in blue extracted from dataset B through applying image classification.

III. RESULTS

TABLE I
VIDEO DATASETS OF WILDLIFE INTERACTIONS WITH TIDAL TURBINES OBTAINED FROM NOVA INNOVATION LTD.

Dataset	Camera Position	Number of Videos
A	T1SIDE	263
	T1TOP	668
B	T1SIDE	102
	T1TOP	90

E. *Three categories (wildlife, algae, background)*

The model had good results when tested on unseen videos (dataset B). Out of 15 videos containing wildlife interaction, 12 (80.0 % accuracy) were correctly identified to contain wildlife (Fig. 5), one video was wrongly classed as 'background' and two videos wrongly classed as 'detritus'. The model struggled to differentiate between the 'background' and 'detritus' categories, with 71 out of 74 videos wrongly identified as containing detritus and the remaining 3 videos wrongly identified as containing wildlife. The model therefore had a 0 % accuracy for the 'background' category. The model performed well in detecting detritus, correctly identifying 59 out of 61 videos with detritus present (96.7 % accuracy). The remaining 2 videos that contained occurrences of detritus were incorrectly identified as containing wildlife. Overall, 5 out of 135 videos were wrongly identified to contain wildlife where they belonged to the 'background' or 'detritus' categories. The ability of the model to differentiate between the 'background' and 'detritus' categories is discussed further below.

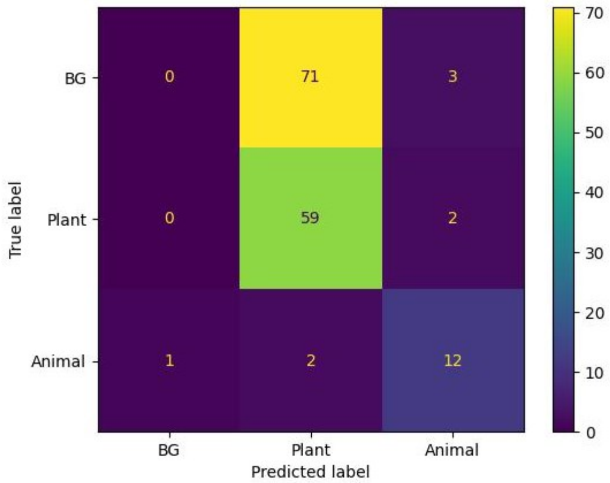

Fig 5: Model results for detecting 'wildlife' (Animals), 'detritus' (Plant) and 'background' (BG) categories in unseen videos.

F. *Two categories (target and non-target)*

As it was not a priority for Nova Innovation to differentiate between types of videos that contained false positives (i.e., no wildlife interactions), the model results on dataset B were re-evaluated against two new categories: 'target' and 'non-target'. Videos previously labelled as 'background' and 'detritus' were grouped into the 'non-target' category and videos previously labelled as containing 'wildlife' were attributed to the 'target' category. Against the new category conditions, 94.1 % of videos were accurately identified as either 'target' or 'non-target' by the model.

IV. Discussion

G. *Model performance*

With the rapid development of MRE infrastructure to combat the threat of climate change, and meet increasing energy security demands, automated processing of large volumes of environmental data and the use of machine learning algorithms could become essential tools in monitoring marine wildlife responses to MRE infrastructures, such as tidal turbines.

The use of convolutional neural networks proved to be effective in identifying seals, diving birds and fish from an operational tidal stream turbine at Bluemull Sound in Shetland. With 80% of targets correctly identified as wildlife, the model provides an operationally ready solution for Shetland (or similar ecosystems) to aid in understanding the impacts of tidal turbines on local marine life. The model also provides a solid foundation to improve detection accuracy, through increased training on videos from the same temperate environment, and to extend model capabilities through including training data from other tidal stream environments with different fauna present.

Two out of three videos containing wildlife that were incorrectly labelled by the model as 'background' or 'detritus' had fish present. Kelp fronds that had become entangled around the turbine structure would regularly trigger the motion-detection camera and were common in videos in both the training and testing datasets. It is possible that dense shoals of fish around the periphery of the camera, which were less common in training data, were mistaken by the model for entangled kelp. Through increasing the amount of training footage of fish in a variety of environmental parameters and shoal structures, model accuracy can easily be improved. The remaining video that was incorrectly labelled by the model to not contain any wildlife in fact contained a diving bird. It is likely that the video with the diving bird was not detected by the model due to the green colouration of the water column. Due to the limited occurrences of birds in the training data, birds often appeared in water column with blue colouration in the training dataset. Therefore, training the model on more instances of birds in a variety of environmental conditions would improve model results in this case.

There were a total of five videos out of 135 that belonged to the 'detritus or 'background' categories that were mislabelled as 'wildlife' by the model. Two videos were occurrences of macroalgal detritus, three videos belonged to the 'background' category and one video contained a small remotely operated vehicle (ROV). It is likely that the model misidentified these videos due to the unique shape of the macroalgal detritus and the ROV appearing as a small, dark object similar to that of a diving bird.

H. *Future work*

Although the success rates for three and two category identification are good, model accuracy can be improved through increased training data with a variety of shapes and sizes of detritus and wildlife. More images of the key wildlife groups of interest such as diving birds, marine mammals and fish are required to expand the training dataset and increase the model accuracy. In addition, the model can be extended or trained on different ecosystems and water quality conditions to increase versatility of the tool.

Currently, the model is applied as a post-processing workflow on batches of videos. Further work could include integrating the developed method into the turbines themselves so that data is processed *in situ* to aid real-time reporting and automated detection of wildlife. Integration of this workflow with automated passive acoustic monitoring systems linked by underwater and satellite communication systems can provide a holistic and complimentary approach using both visual and acoustic data for environmental monitoring.

V. Conclusion

This study successfully demonstrated the effectiveness of using machine learning model to automate the detection of wildlife in proximity to MRE infrastructure delivering significant efficiencies in the analysis and reporting for MRE monitoring programs. The model developed in this case study achieved an accuracy of 94.1 % in identifying videos within the 'target' ('wildlife') category and 'non-target' ('background' and 'detritus') category. The model has been integrated into a novel, industry-ready workflow that can ingest approximately 200 videos or 20 hours of footage and produce an automated detection report of the results in approximately 30 minutes. When using a manual approach, it takes approximately 320 person-hours of analysis for 1600 hours of video. By comparison, this automated workflow could analyse 1600 hours of video in 40 hours resulting in an 87.5% reduction in interpretation time.

The use of machine learning for automated processing provides a subset of data for more focused manual scrutiny and analysis, while reducing the overall size of the dataset requiring storage. This facilitates analysis of a much greater proportion of data and addresses the growing challenges of marine operators' data storage requirements.

References

[1] G. Boehlert, A. Gill, "Environmental and Ecological Effects of Ocean Renewable Energy Development – A Current Synthesis". Oceanography 23, 68–81, 2010. https://doi.org/10.5670/oceanog.2010.46

[2] A.E. Copping, N. Sather, L. Hannah, J. Whiting, G. Zydlewski, G. Staines, A. Gill, I.A.O. Hutchinson, T. Simas, J. Bald, C. Sparling, J. Wood, E. Masden "OES-Environmental 2020 State of the Science Report: Environmental Effects of Marine Renewable Energy Development Around the World". Report for Ocean Energy Systems (OES), 2020. https://doi.org/10.2172/1632878

[3] A.E. Copping, M.E. Grear "Applying a simple model for estimating the likelihood of collision of marine mammals with tidal turbines". International Marine Energy Journal, 1(1), 27–33, 2018. https://doi.org/10.36688/imej.1.27-33

[4] G.D. Hastie, D.J.F. Russell, P. Lepper P, "Harbour seals avoid tidal turbine noise: Implications for collision risk". Journal of Applied Ecology, 55, 684–693, 2018. https://doi.org/10.1111/1365-2664.12981

[5] C. Sparling, M. Lonergan, B. McConnell, "Harbour seals (*Phoca vitulina*) around an operational tidal turbine in Strangford Narrows: No barrier effect but small changes in transit behaviour". Aquatic Conservation: Marine and Freshwater Ecosystems, 28(1), 194-204, 2018. https://doi:10.1002/aqc.2790

[6] T.A. Wilding, A.B. Gill, A. Boon, E. Sheehan, J. Dauvin, J.-P. Pezy, F. O'Beirn, U. Janas, L. Rostin, I. De Mesel, "Turning off the DRIP ('Data-rich, information-poor') – rationalising monitoring with a focus on marine renewable energy developments and the benthos". Renew. Sustain. Energy Rev. 74, 848–859, 2017. https://doi.org/10.1016/j.rser.2017.03.013

[7] N. Piechaud, C. Hunt, P.F. Culverhouse, N.L. Foster, K.L Howell, "Automated identification of benthic epifauna with computer vision". Mar. Ecol. Prog. Ser. 615, 15–30, 2019. https://doi.org/10.3354/meps12925

[8] J.M Durden, B.J. Bett, T. Schoening, K.J. Morris, T.W Nattkemper, H.A. Ruhl, "Comparison of image annotation data generated by multiple investigators for benthic ecology." Mar. Ecol. Prog. Ser. 552, 61–70, 2016. https://doi.org/10.3354/meps11775

[9] N. MacLeod, M. Benfield, P. Culverhouse, "Time to automate identification". Nat. 2010 4677312 467, 154–155, 2010. https://doi.org/10.1038/467154a

[10] Tan, Mingxing, and Quoc Le. "EfficientNet: Rethinking model scaling for convolutional neural networks." In *International conference on machine learning*, pp. 6105-6114. PMLR, 2019.

[11] A. Vellappally, S. Hou, M. Love, F. Watkins, T. Jackson-Bue, "Machine learning for benthic taxon Identification". In: Proceedings of the International Conference on Learning Representations, April 25-29, 2022, London.

Acknowledgements

The authors would like to thank Nova Innovation Ltd., CGG Services (UK) Ltd. and Marine Scotland for making data from Bluemull Sound available to study and publish in these conference proceedings.