Towards Versatile & Adaptive Detection Algorithms in Underwater Acoustics

by Oliver Kirsebom, Fabio Frazao, Ines Hessler, and Stan Matwin

In lakes and oceans, sounds produced by aquatic animals combined with the noise created by the environment (e.g., splashing waves) and human activities produce a rich underwater soundscape. This underwater soundscape holds a wealth of information about animal behaviour, geographic distribution, habitat use, social interaction, and more. By immersing specially designed microphones into the water, scientists are able to tap into the constant stream of acoustic information to study animals in their natural environment. This approach, known as passive acoustic monitoring, permits valuable data to be collected over long time periods with minimal effort and has the potential to inform decisions on environmental and conservation issues.

Unlocking the information contained in the acoustic recordings is, however, a formidable challenge as underwater soundscapes are highly complex and variable, and hence difficult to interpret. Therefore, the analysis is usually performed manually by an expert human analyst or automatically by a carefully engineered algorithm, which requires significant effort and resources to implement in the first place. Either way, the data analysis represents a bottleneck in the use of passive acoustic monitoring, restricting the large scale application of the method.

Deep neural networks are a type of machine-learning algorithm modelled loosely after the human brain, which can be trained to perform highly non-linear operations. They have proven highly successful in tasks related to image classification, video segmentation, and speech recognition, and shown great promise in applications to underwater acoustics. For example, a recent partnership between Google and the U.S. National Oceanic and Atmospheric Administration produced a neural network capable of detecting the songs of humpback whales (https://patternradio.withgoogle.com). Such success stories depend, however, on the availability of large training datasets with thousands of evaluated examples, a rarely afforded luxury in underwater acoustics. Yet, there are reasons for being optimistic: The difficulties imposed by the shortage of training data could be overcome by employing techniques such as transfer learning, where a neural network trained to perform a particular task is adapted to perform a different, but related task, using a modest number of training examples. Moreover, with the active help of human analysts, deep neural networks may be able to learn more efficiently, further reducing the data requirements.

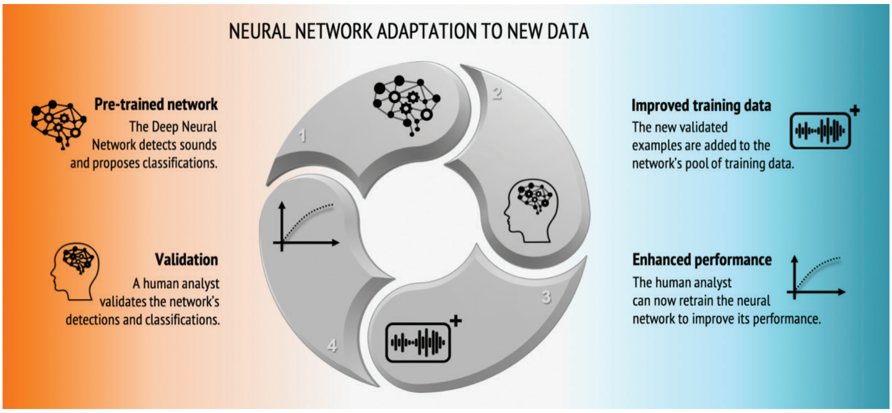

Figure 1: Neural network adaptation to new data. As the neural network becomes more confident in its predictions, the human analyst will be able to entrust more decisions to the network.

MERIDIAN is currently exploring this idea through the creation of an interactive learning application with a visual interface, which will allow the human analyst to inspect and correct detections and classifications proposed by the neural network. Adding newly validated data to the pool of training data, the neural network will gradually improve its performance (Figure 1). As the neural network becomes more confident in its predictions, the human analyst will be able to entrust more decisions to the network, greatly speeding up the overall analysis task. The positive feedback loop will allow neural networks to be trained more efficiently and save precious resources. The envisioned user interface will enable the data analyst to analyze the data, verify, and improve the performance of the detector in the same setting. As a critical step towards this goal, we are currently exploring various training schemes to identify the most efficient learning strategy for the neural network.

Where traditional approaches to sound detection and classification rely on handcrafted algorithms for extracting acoustic features from the raw audio data, deep neural networks can work directly on the raw audio data (or its spectral representation). This ability, more than anything, is what sets deep neural networks apart from traditional approaches, as it allows them to learn the features that are most useful for a given task. This makes deep neural networks versatile and adaptive – highly desirable properties for detecting and classifying underwater sounds against a variable spectrum of background noises, which differ substantially between locations and change with time.

One day soon sound detection algorithms based on Deep Neural Networks may become part of the standard toolkit available to scientists for analyzing large bodies of underwater acoustic data. At MERIDIAN, we are already working towards this goal. However, the success of Deep Neural Networks in underwater acoustics will also depend on the availability of sufficiently large training datasets that are shared openly within the community.

Dr. Oliver Kirsebom is the Lead Acoustic Data Analyst at MERIDIAN and an Adjunct Professor in the Faculty of Computer Science at Dalhousie University. Fabio Frazao is the Mighty Data Scientist at MERIDIAN. Dr. Ines Hessler is the Senior Data Manager at MERIDIAN and Adjunct Professor in the Faculty of Computer Science at Dalhousie University. Dr. Stan Matwin is the P.I. of the MERIDIAN project, Director of the Institute for Big Data Analytics, and Professor in the Faculty of Computer Science at Dalhousie University. https://meridian.cs.dal.ca