Hawaiian monk seal vocalization detection and classification using Deep Learning

Kenneth Bremnes*[1], Lucas Hobrak*[1], Rebecca Moen*[1], Colleen Reichmuth[2], Jillian M. Sills[2], Jeppe H. Rasmussen[3]

* These authors contributed equally to the work
[1] Center for Artificial Intelligence Research, University of Agder, 4604 Kristiansand, Norway
[2] Institute of Marine Sciences, Long Marine Laboratory, University of California Santa Cruz, Santa Cruz, California 95060, USA
[3] Center for Coastal Research, University of Agder, 4604 Kristiansand, Norway

The Hawaiian monk seal, *Neomonachus schauinslandi*, is a highly endangered species with an urgent conservation priority. Applying passive acoustic methods to detect and monitor this species would inform research and management efforts. However, manually processing large acoustic datasets is slow, labour intensive and prone to human errors. Therefore, machine learning models were developed to automate the process and increase accuracy. This approach is divided into two classification problems with two Convolutional Neural Networks. The first model, based on ResNet, detects and isolates data where vocalizations are present. The second model, based on VGG net, classifies calls into 6 previously defined perceptual types. Both approaches convert acoustic data into spectrograms and utilize image classification. Approximately 100 hours of recordings from a captive, adult male Hawaiian monk seal were manually annotated, creating a dataset of 1200 labelled vocalizations used for model training and testing. After implementing data augmentation to balance the dataset and prevent overfitting, our models achieved an accuracy of 92% detection and 84% classification of different call types. Future work includes expanding the dataset by increasing the number of calls and including more individuals and environmental settings.