Applications of Deep Learning to Sound Detection and Classification in Underwater Acoustics

Fabio Frazao
Acoustic Data Analyst

Ocean Networks Canada, Victoria, BC. 2019.04.29

Outline

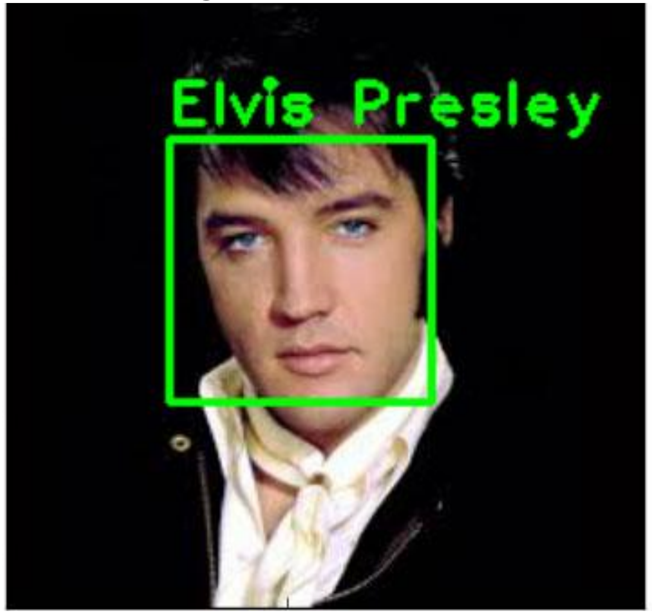
- Brief introduction to Deep Neural Networks
  - Examples of applications (image recognition, natural speech processing, etc)
  - Why is deep learning successful now?
  - Feature extraction, transfer learning
  - Basic network architectures (Multilayer perceptron , CNN, RNN)
  - Acoustic data representations
- Projects
  - Detecting arctic cod grunts (UVic data)
  - Classifying killer and pilot whales (WHOI)
  - Matching individual Killer whale calls (WHOI data)
  - Detecting Baleen whales (ONC Barclay Canyon data)
- Training Datasets
  - HF5 standard
- Data augmentation strategies
  - Image manipulation
  - Sound propagation modelling
  - Deep Generative models
- Ketos library
  - HDF5 database
- Workflow visions
  - Interactive app
  - Data augmentation tool
- Your inputs

Brief introduction to neural networks

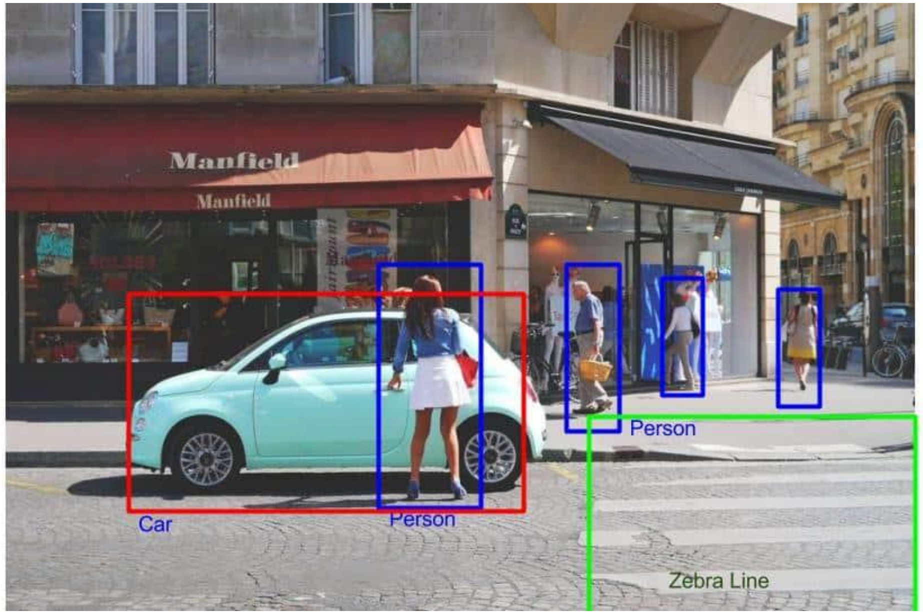
Ocean Networks Canada, Victoria, BC. 2019.04.29

Examples of applications- Computer vision

Facial Recognition



Object detection



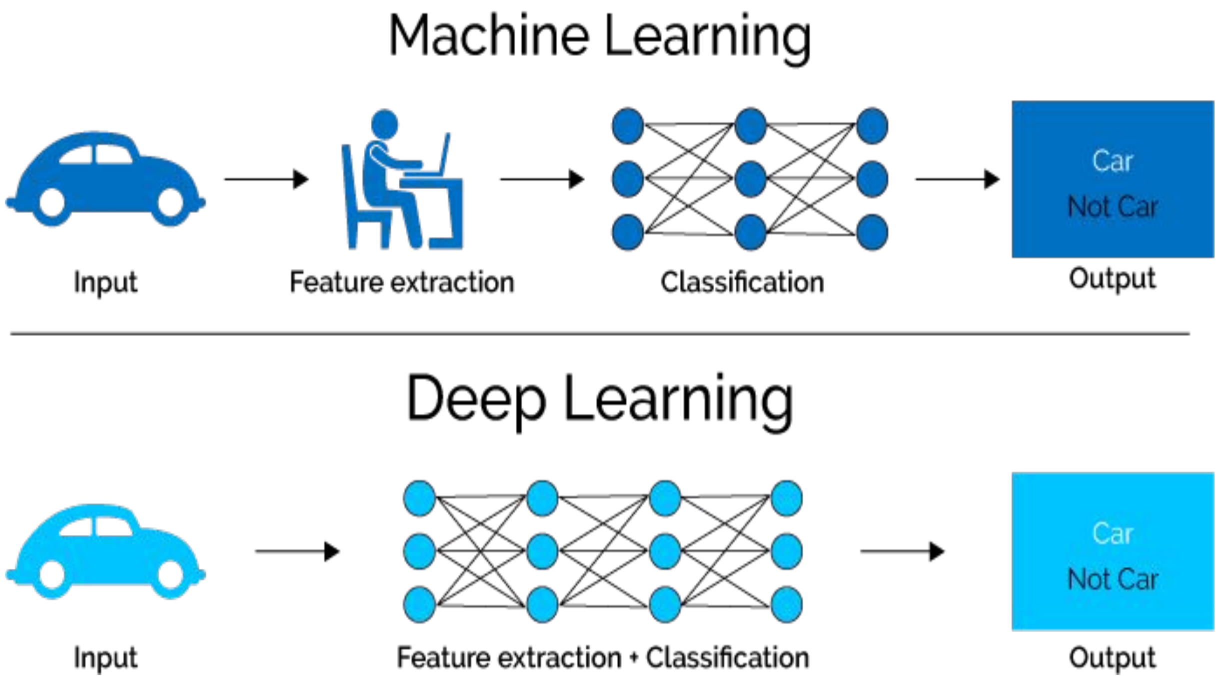
Examples of applications- Natural language processing

Speech recognition/ synthesis

Translation

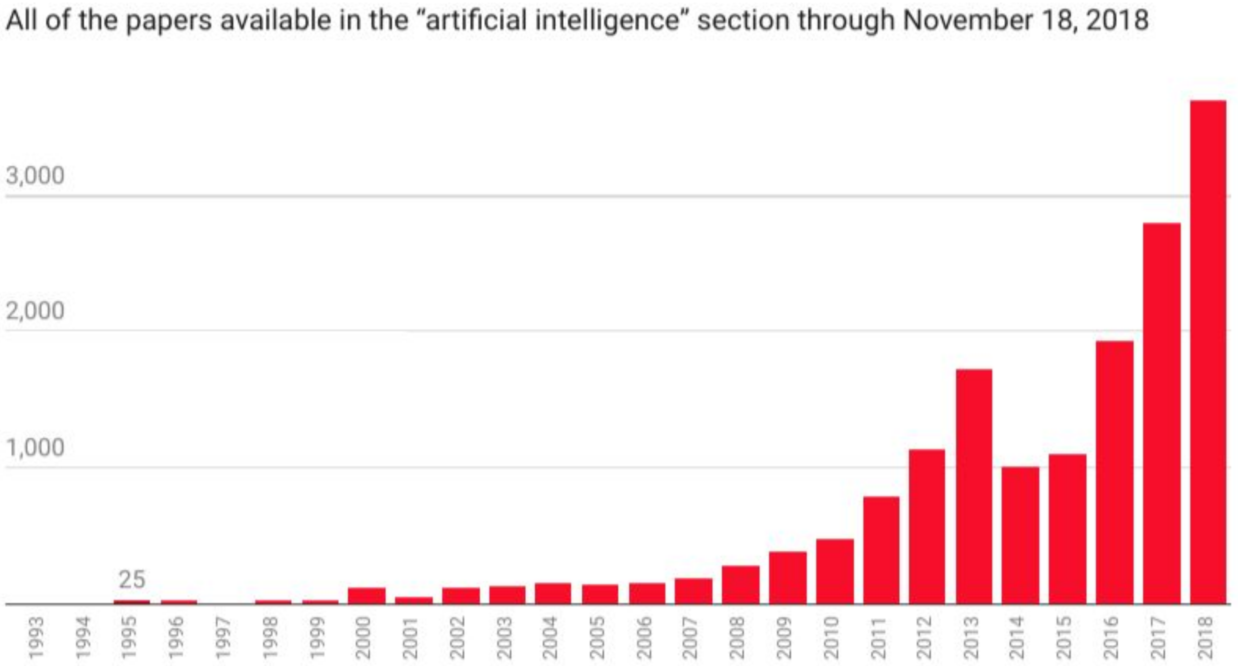
Feature extraction

Deep Learning systems aim to be end-to-end, although in practice there's usually some level of input (signal) processing left



Why is deep learning successful now?

- Software improvements (algorithms + frameworks)
- Hardware Improvements (Larger storage. Faster processors:cpus, gpus, tpus, etc.)
- Data improvements (More and better data)



16,000 + papers on ArXiv.org

MIT Technology review
(https://www.technologyreview.com/s/612768/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/)

Why is deep learning successful now?

- Software improvements (algorithms + frameworks)
- Hardware Improvements (Larger storage. Faster processors:cpus, gpus, tpus, etc.)
- Data improvements (More and better data)

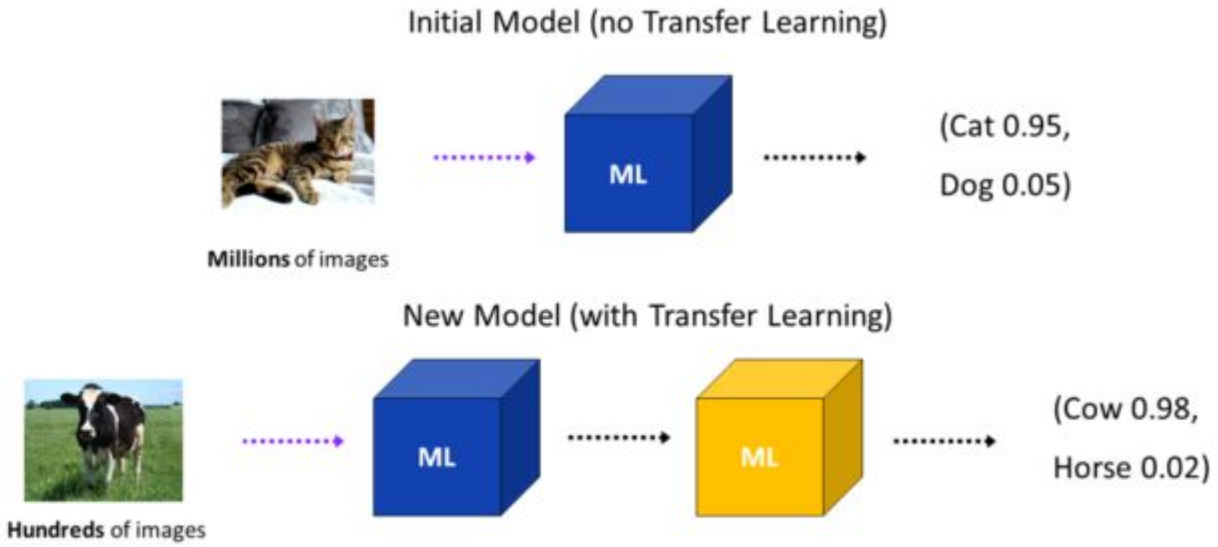
3 million images, 9 thousand individuals

IMAGENET

14 million images, 921 thousand synsets (categories)

Transfer learning

- Transfer learning can drastically reduce the training time and amount of data required
- Make models more adaptable and reusable



Transfer learning

- Transfer learning can drastically reduce the training time and amount of data required
- Make models more adaptable and reusable

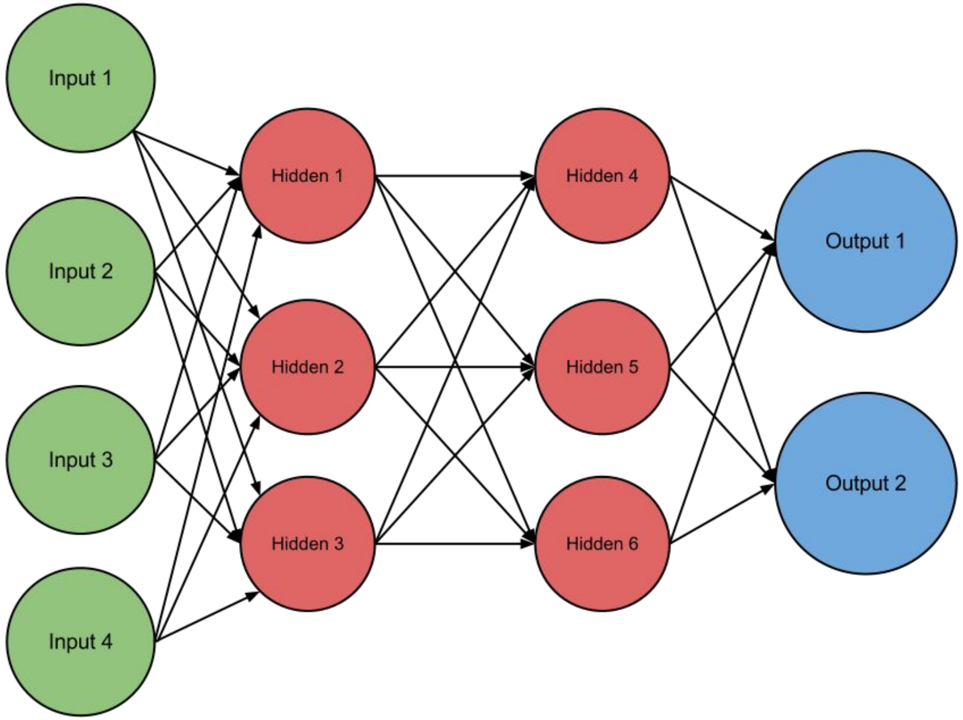
FaceNet: A Unified Embedding for Face Recognition and Clustering

Florian Schroff, Dmitry Kalenichenko, James Philbin

(Submitted on 12 Mar 2015 (v1), last revised 17 Jun 2015 (this version, v3))

Basic network architectures

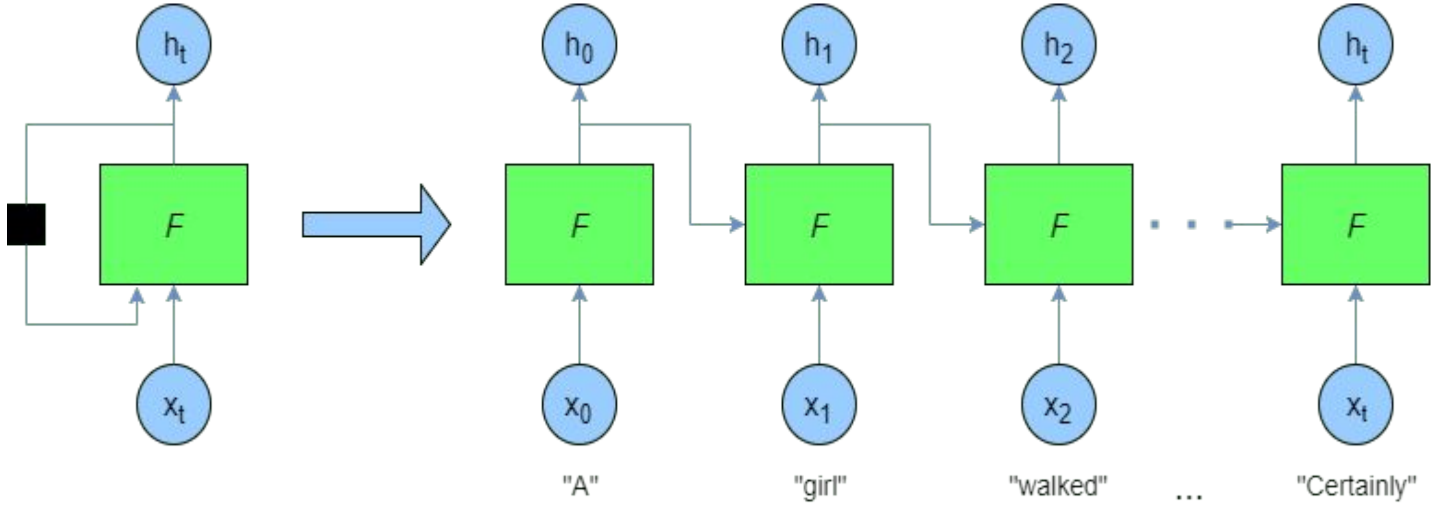
Multilayer perceptron



- One of the foundational architectures for modern Deep Learning
- Rarely used on its on nowadays

Basic network architectures

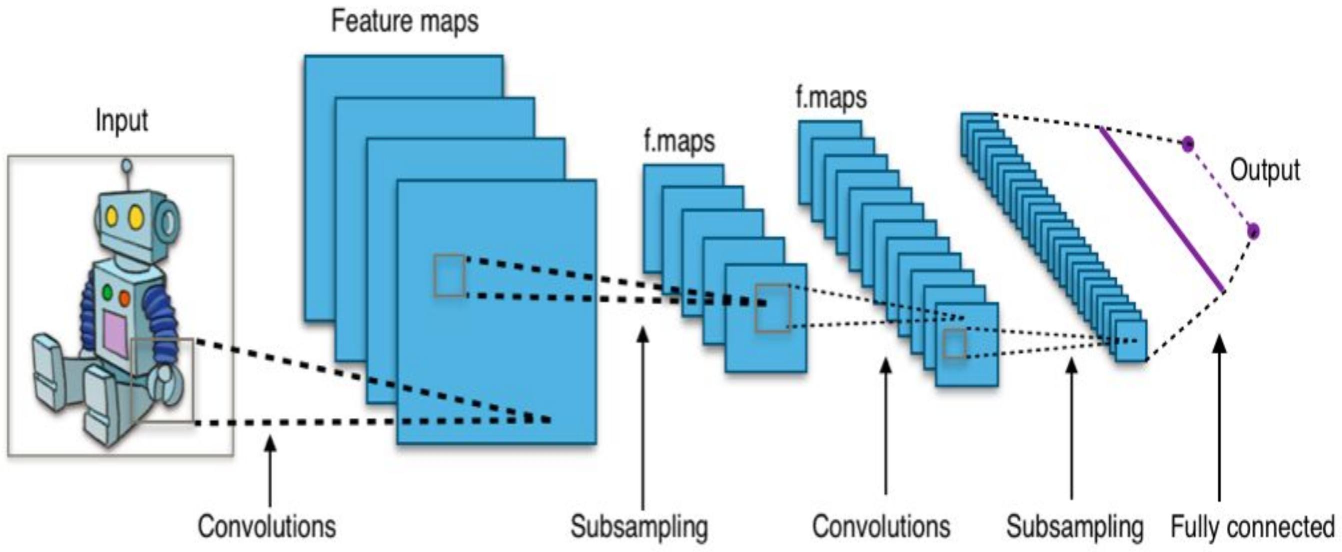
Recurrent Neural Networks (RNNs)



- The output from previous cells is used by the next cells in addition to the inputs ("memory")
- Works very well for sequences/ time-series
- Very useful for language modelling

Basic network architectures

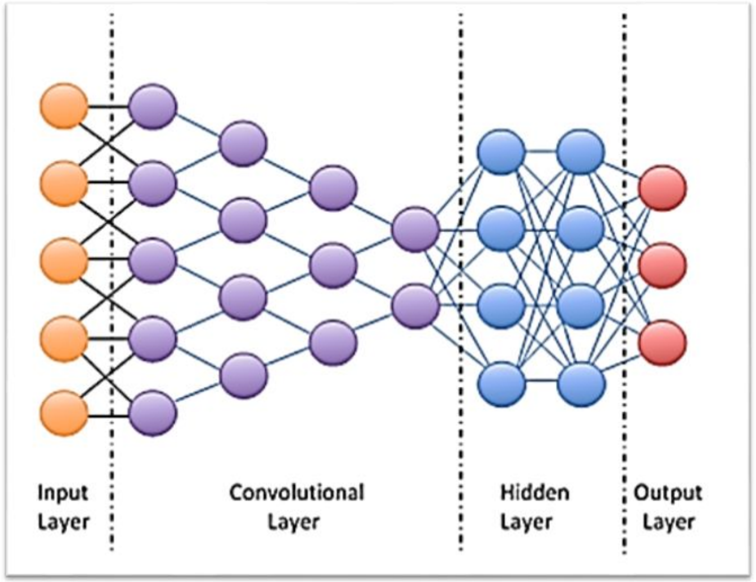
Convolutional neural network (CNN)



- Creates feature maps that summarize information from one layer to the next
- Works very well for images/spatially related features
- Very useful for dealing with images; the building block of modern computer vision

Hybrid architectures

Neural Networks are usually built by combining basic architectures

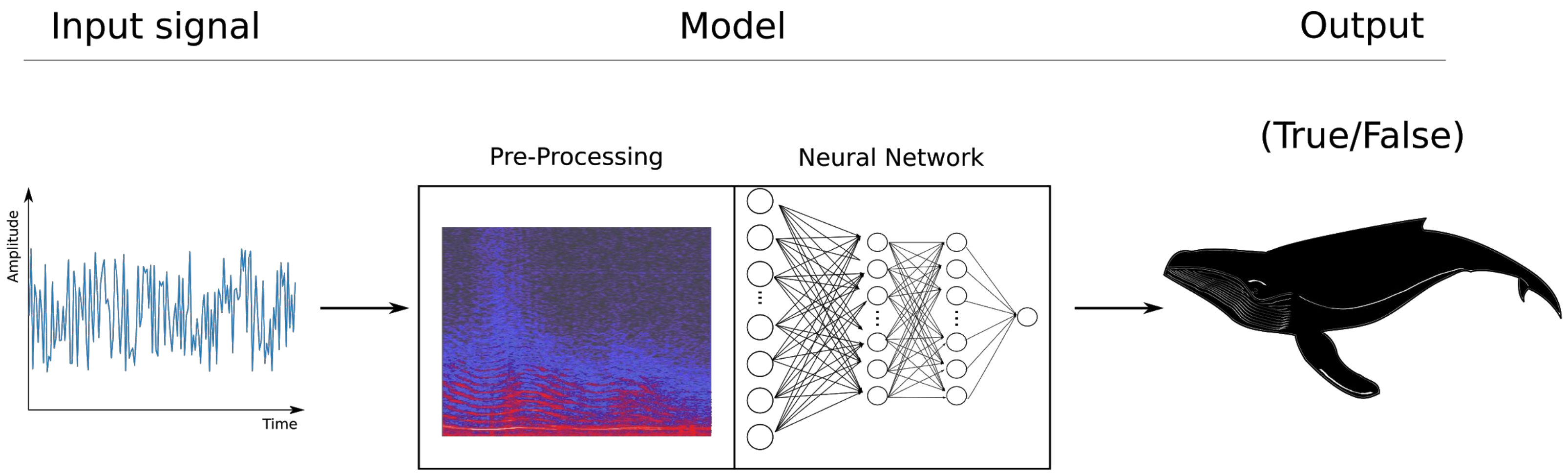


Acoustic data representations

Ocean Networks Canada, Victoria, BC. 2019.04.29

Transfer learning

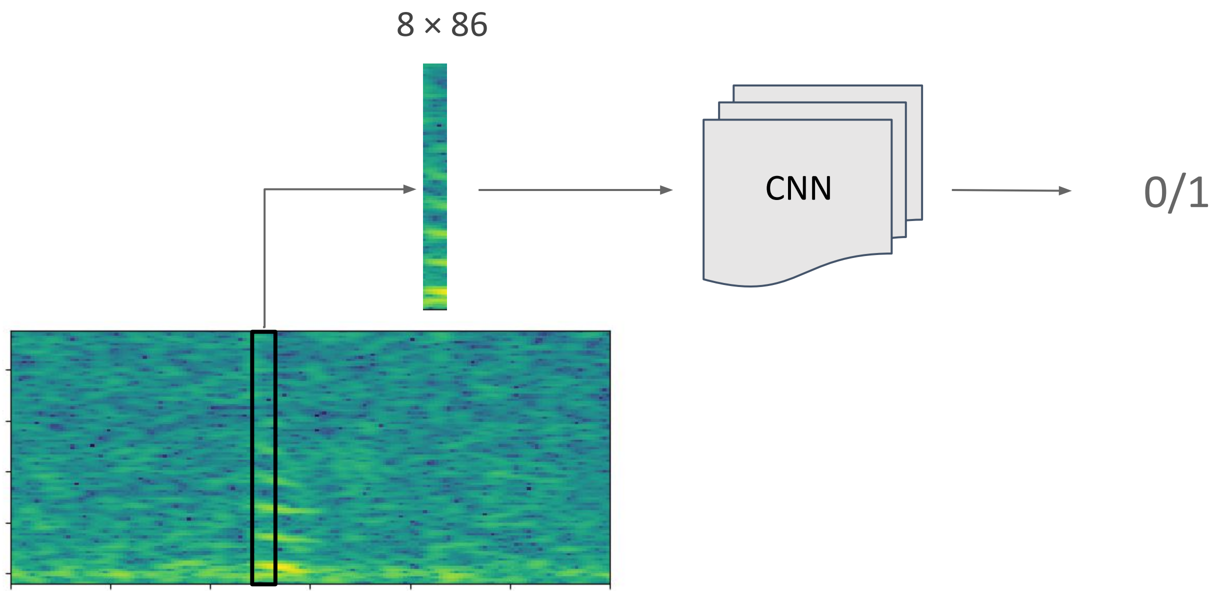
- Waveforms
- Spectral representations
  - Constant Q Transform
  - Fourier transform
  - Morlet wavelet transform
- Scalar features



Projects

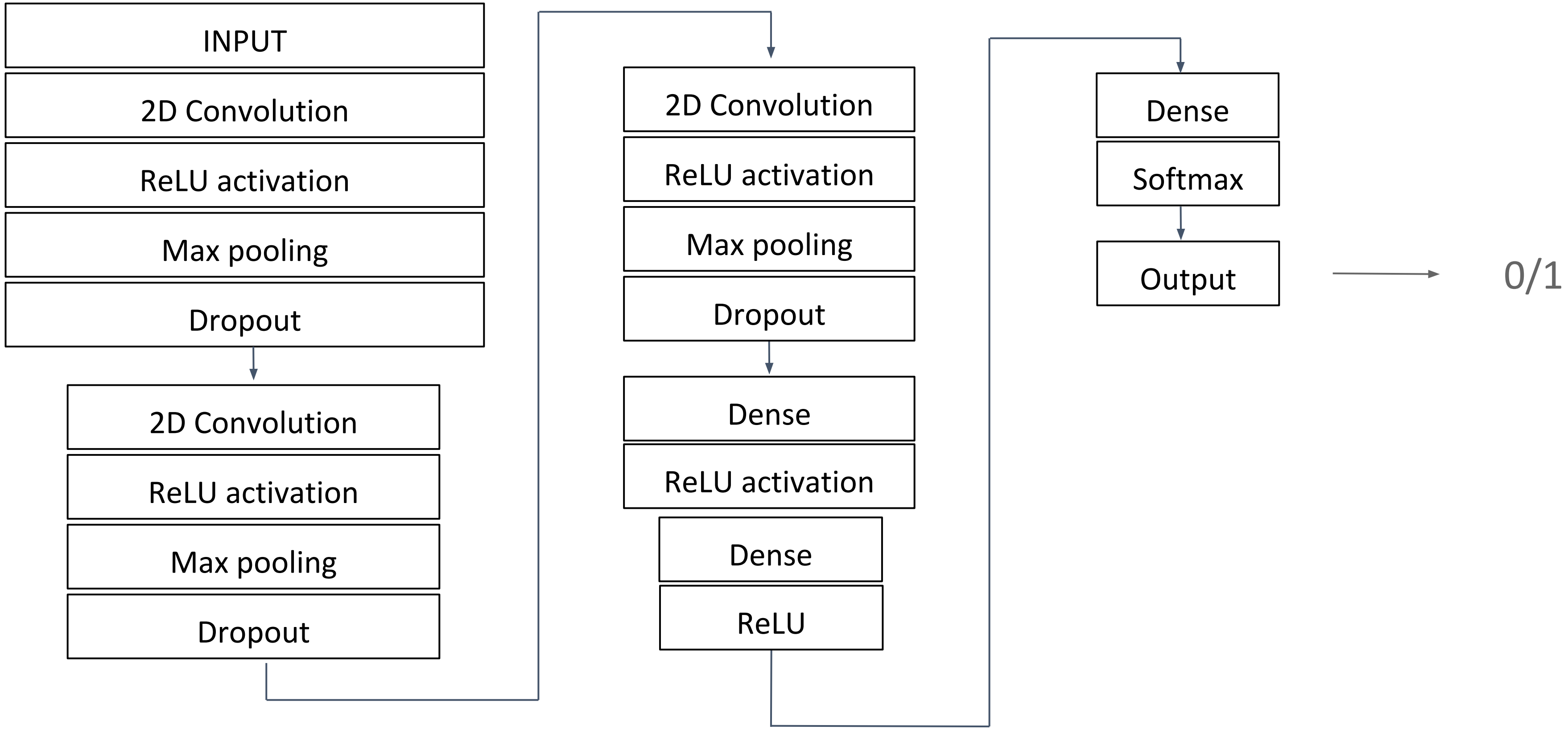
Arctic cod grunts

Using a CNN (convolutional neural network) to detect fish sounds



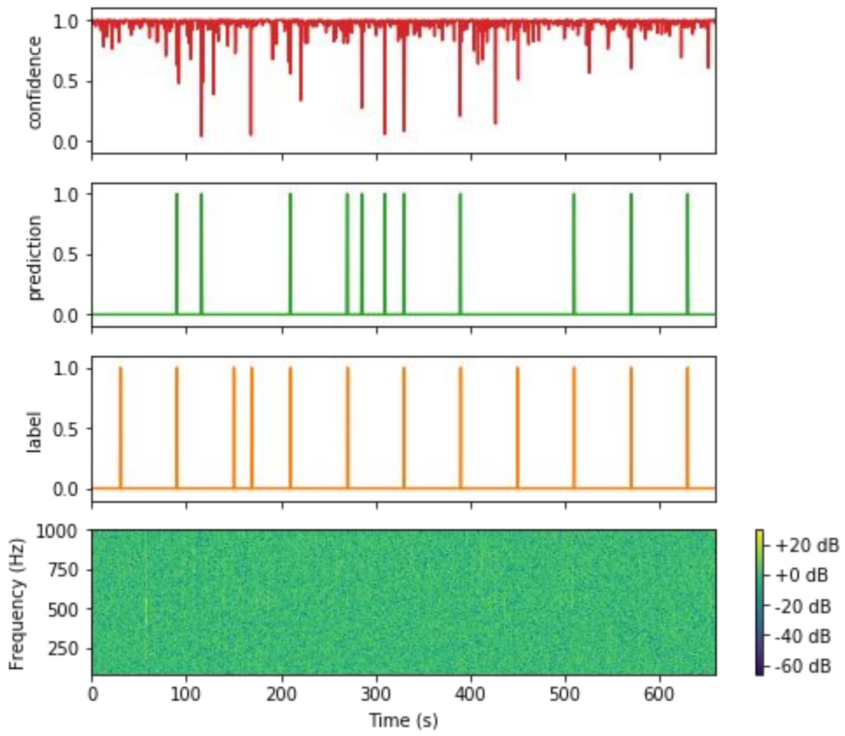
Arctic cod grunts

Using a CNN (convolutional neural network) to detect fish sounds



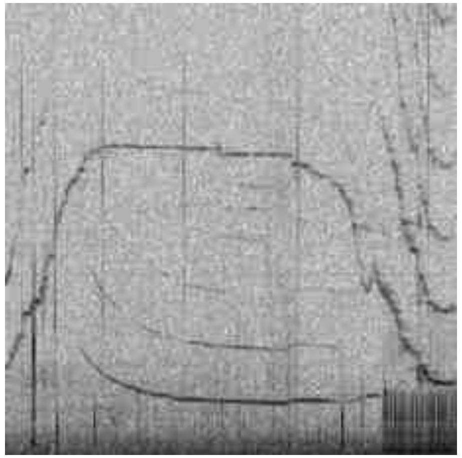
Arctic cod grunts

Using a CNN (convolutional neural network) and a to detect fish sounds



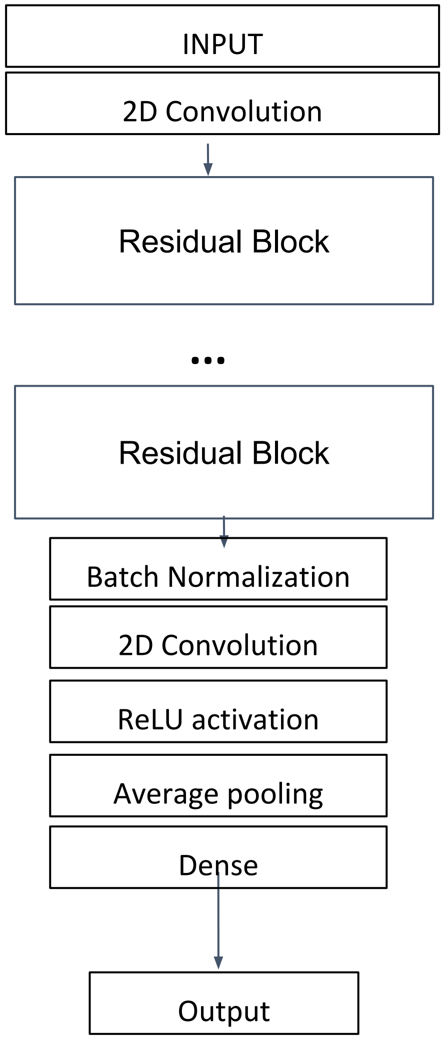
Accuracy: 98.4%
Precision: 72.7%
Recall: 66.7%

Differentiating between killer and pilot whales using ResNETs

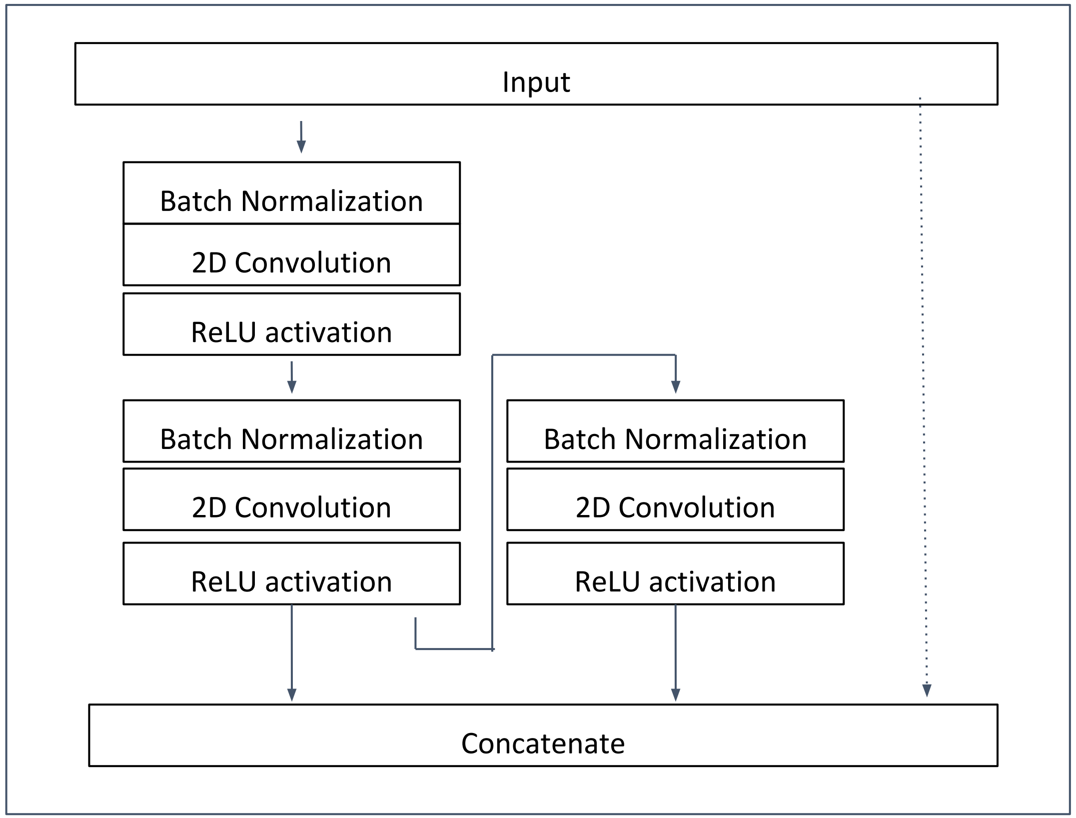


Accuracy: 98.44%

Differentiating between killer and pilot whales using ResNETs



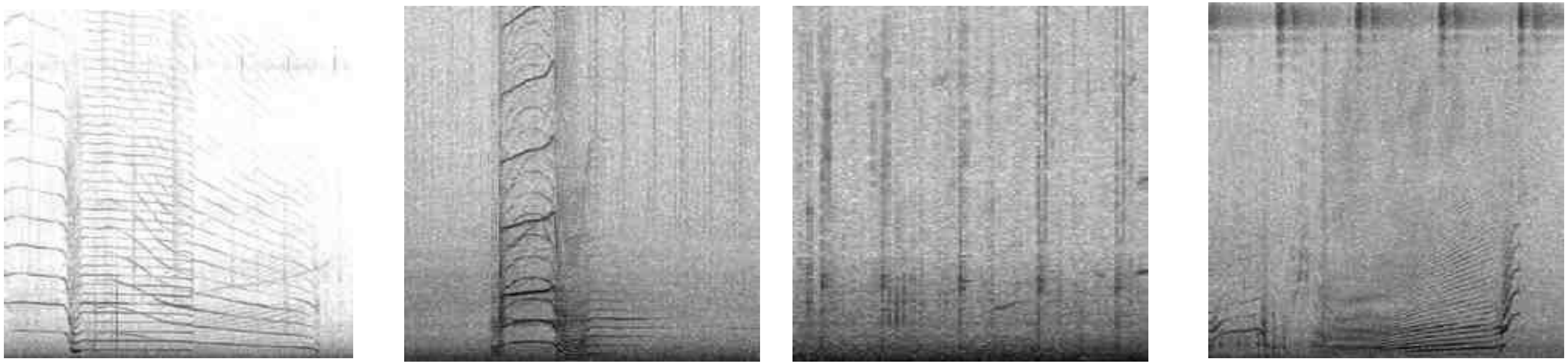
Residual Block



Matching individual killer whale calls with Siamese Networks

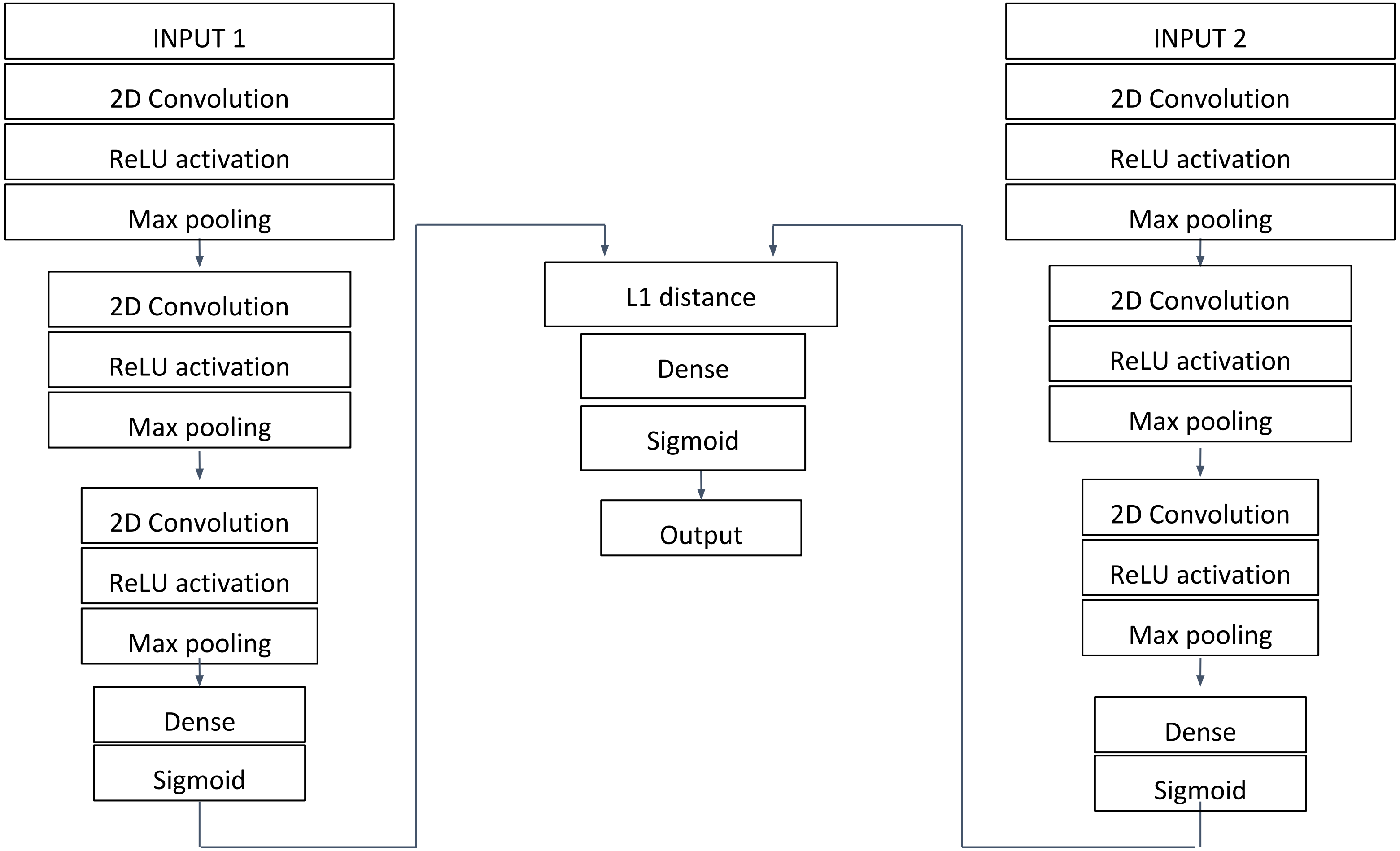
15 individuals

Accuracy: 94.6%



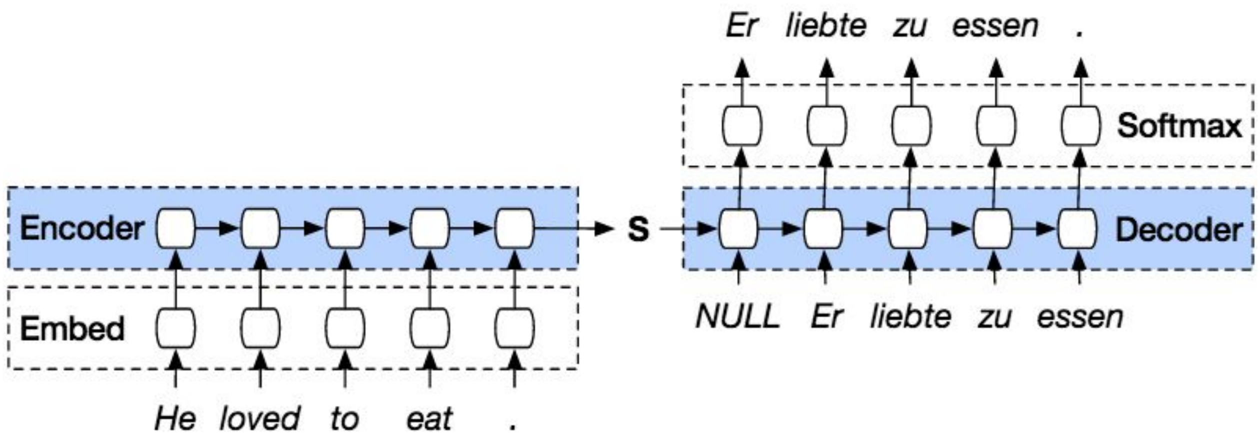
Matchingindividual killer whale calls with Siamese Networks

Accuracy: 94.6%



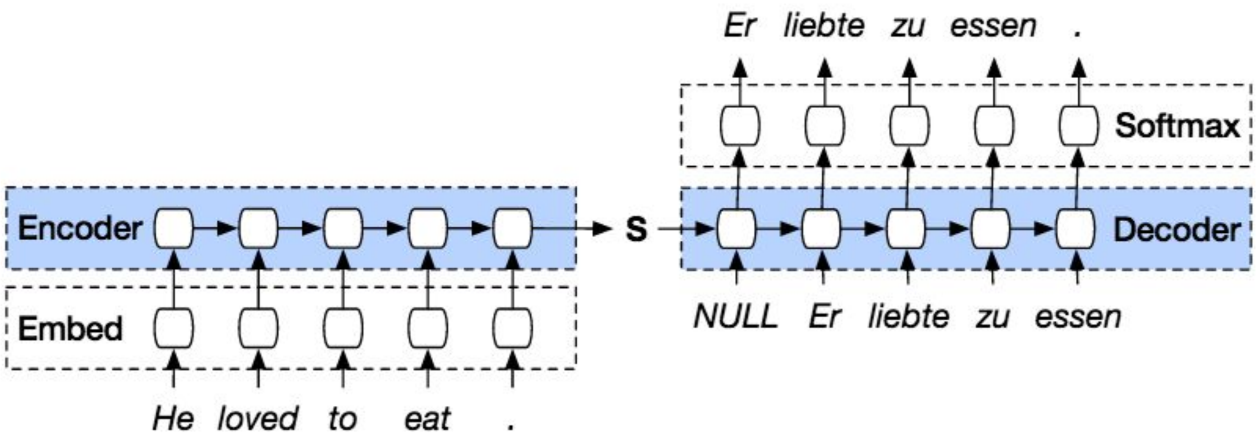
ONC Barclay Canyon

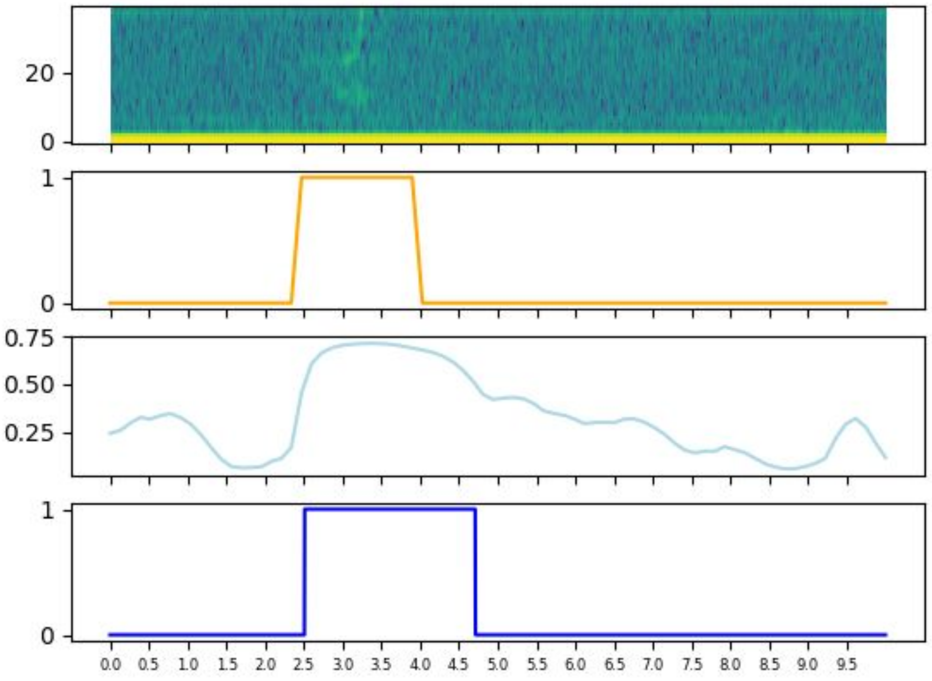
Using a sequence to sequence model to detect humpback whales



ONC Barclay Canyon

Using a sequence to sequence model to detect humpback whales

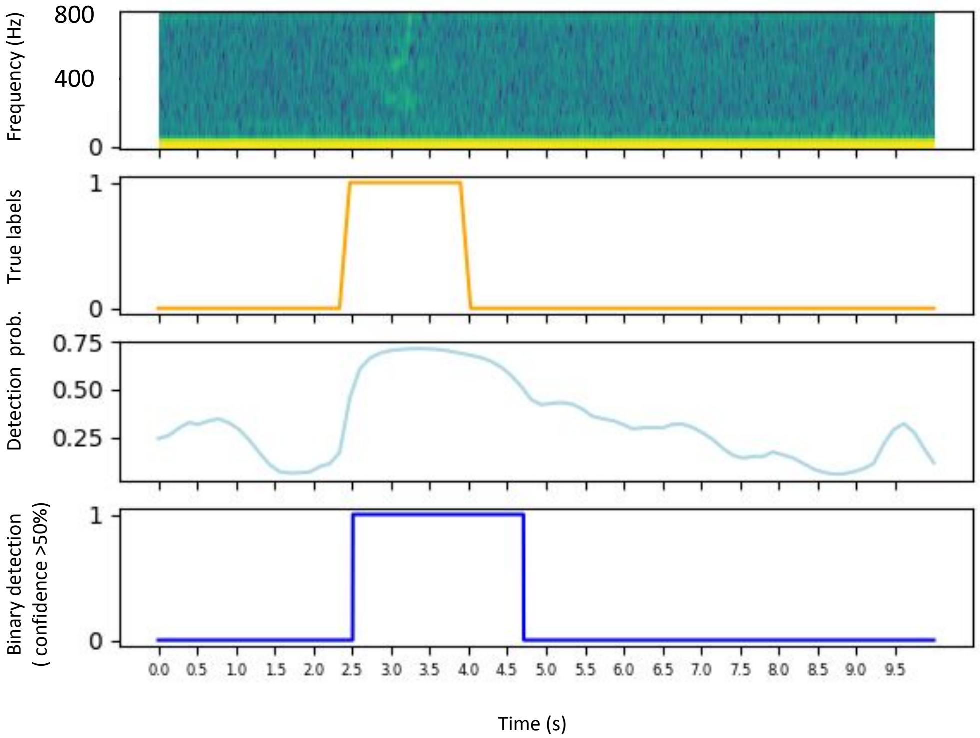




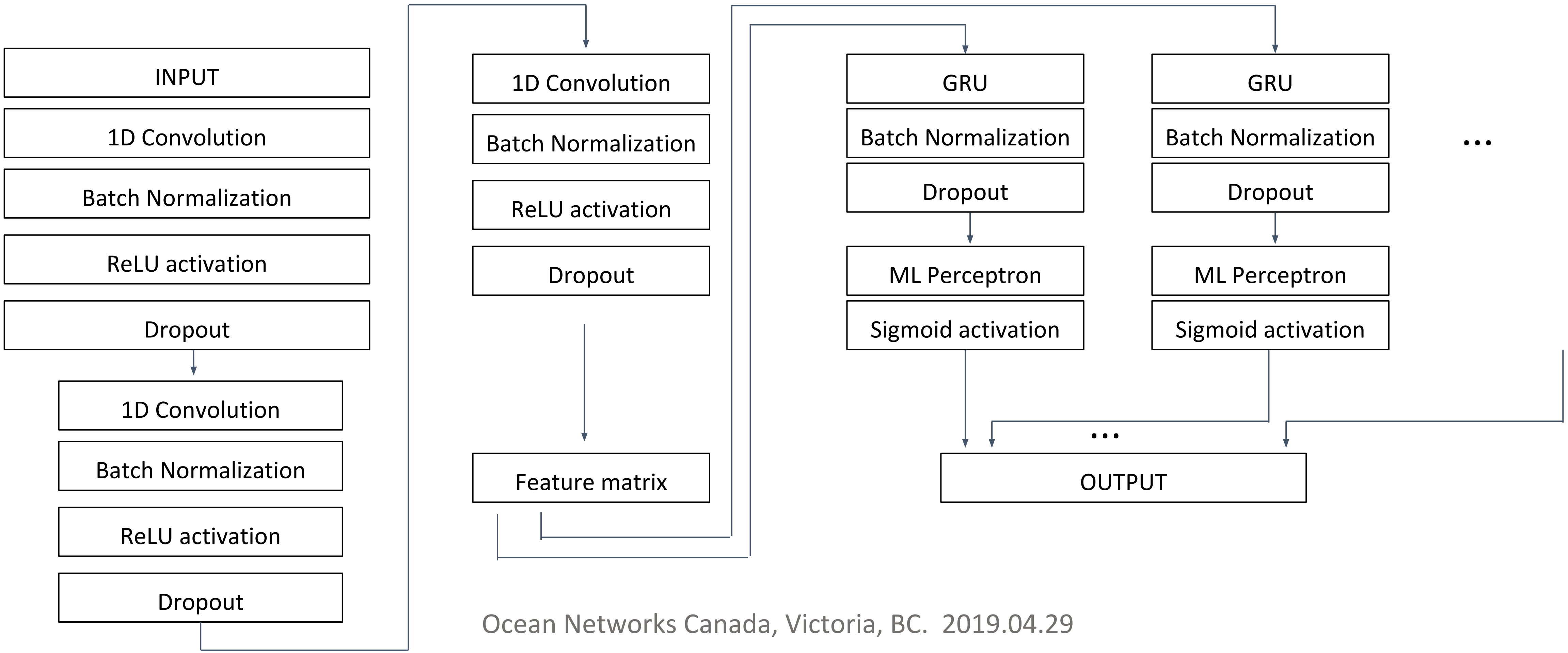
ONC Barclay Canyon

Using a sequence to sequence model to detect humpback whales

Accuracy: 87%
Precision: 72%
Recall: 36%



ONC Barclay Canyon

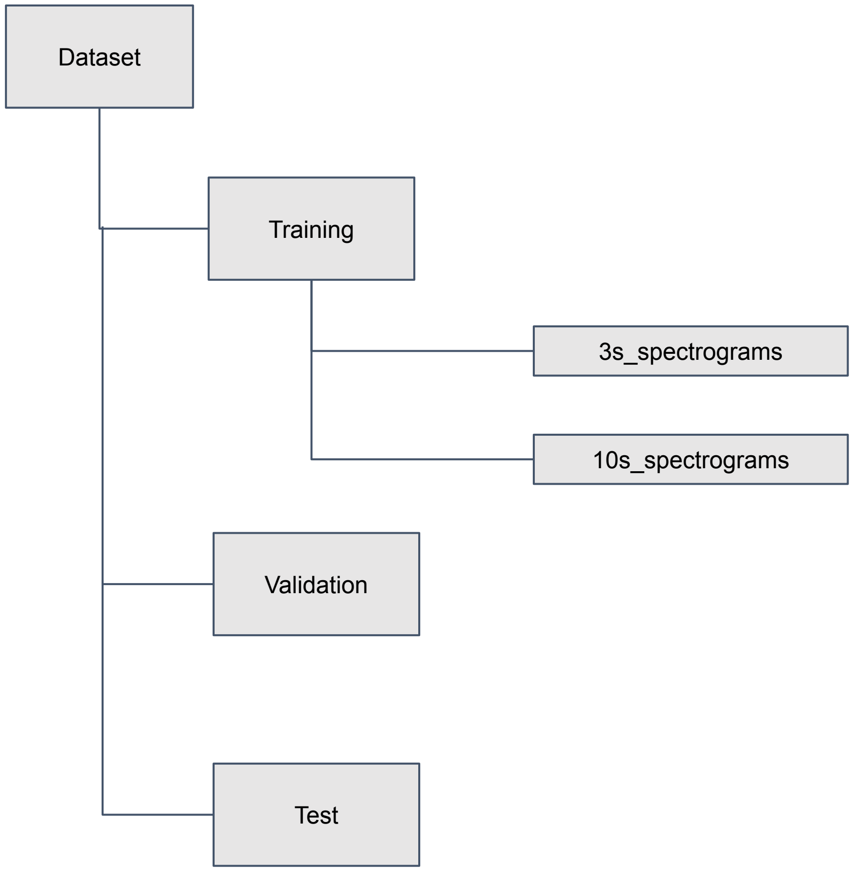


Ocean Networks Canada, Victoria, BC. 2019.04.29

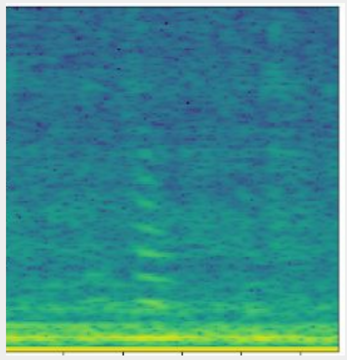
Training datasets

Dataset standard

Datasets are organized in hierarchical format (using HDF5)



3s_spectrograms

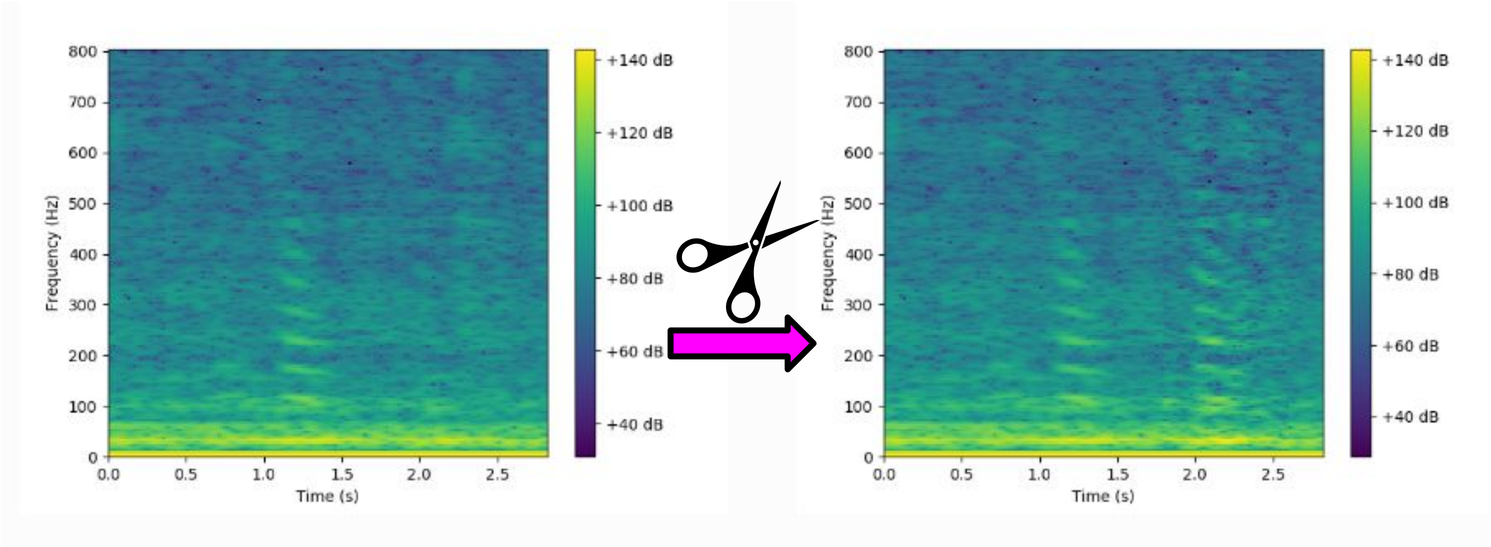
id	data	labels	boxes
uvic_23		1	(1.1,1.6,80,600)

Data Augmentation

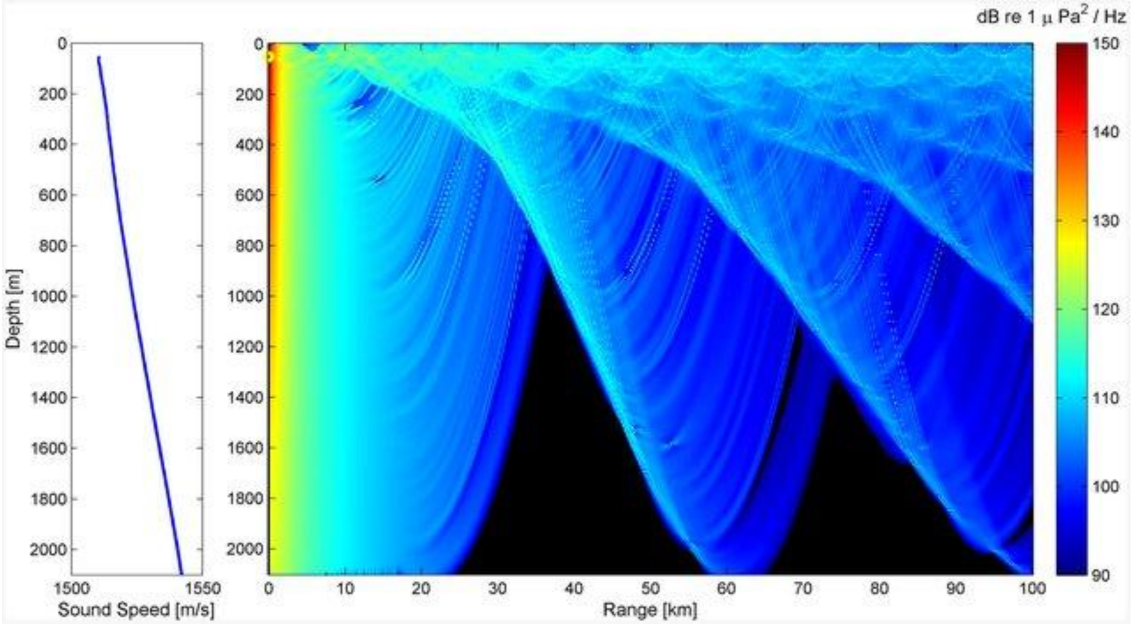
Ocean Networks Canada, Victoria, BC. 2019.04.29

Simple signal manipulation tricks

- Extract the signal of interest and place it in a different background
- Change amplitude



Sound propagation modeling



- Pass a signal through a sound propagation model, to simulate what it would sound like in different environments

Deep generative models

Generative Adversarial Networks

- WaveNETs
- Variational Autoencoders
- Recurrent Neural Networks
- etc

Generative Adversarial Networks

Sageev Oore

Canada CIFAR AI Chair,
Associate Professor, Faculty of Computer Science, Dalhousie University
Research Faculty Member, The Vector Institute

Previously:

Visiting Research Scientist, Google Brain
Associate Professor & Chairperson (on leave)
Department of Mathematics & Computer Science, SMU

Ketos library

Ocean Networks Canada, Victoria, BC. 2019.04.29

Ketos library

- Open-source python package (GPL3 license)

- Available on PyPI:

  pip install ketos

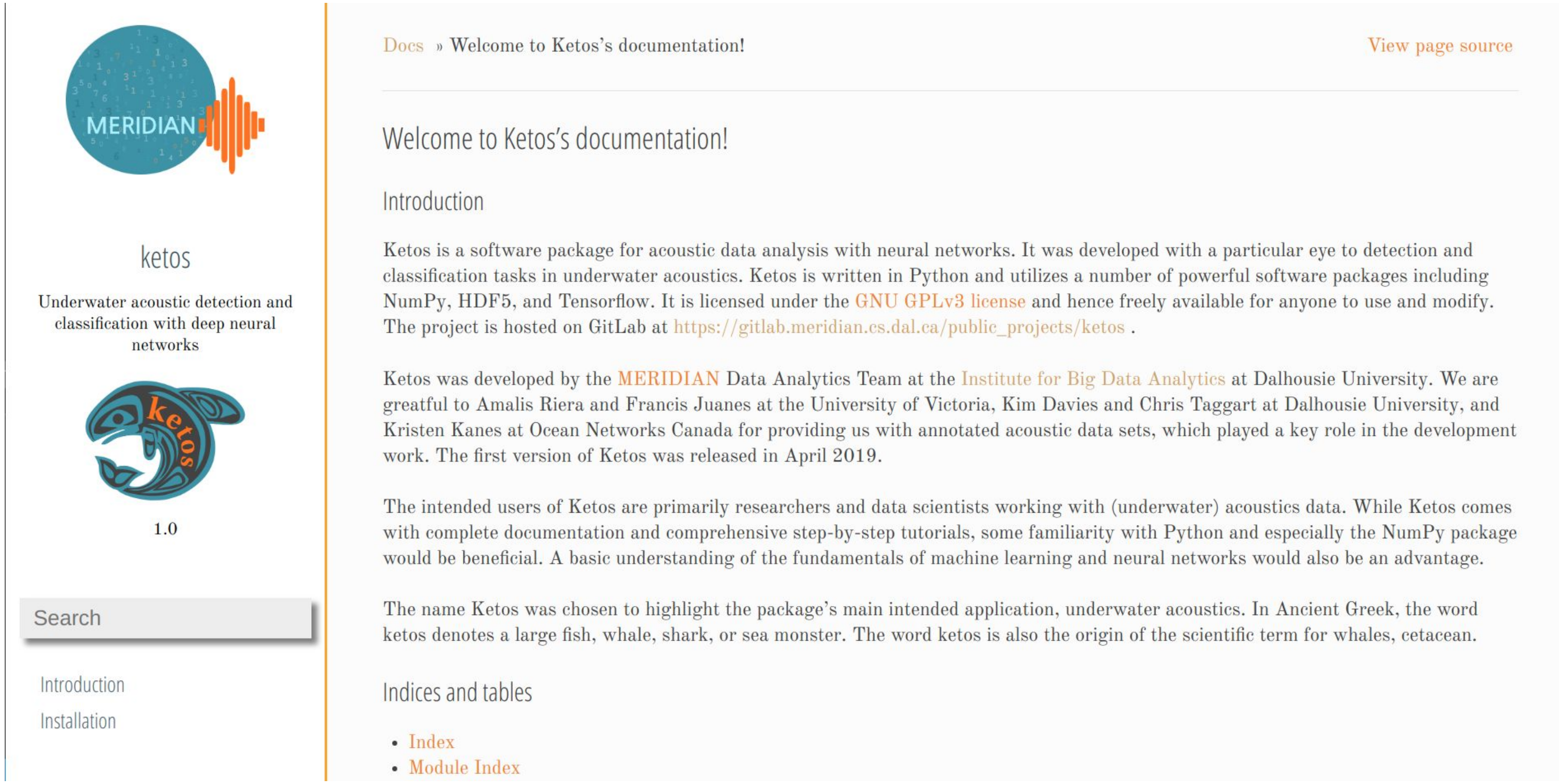
- Built on top of Numpy, Tensorflow and HDF5/PyTables

- Provides:
  - Data handling tools (including for larger than memory datasets)
  - Signal processing methods
  - Useful network architectures with a common interface

Ketos library

- Documentation
  docs.meridian.cs.dal.ca/ketos

- Source code
  gitlab.meridian.cs.dal.ca/public_projects/ketos



MERIDIAN

ketos

Underwater acoustic detection and classification with deep neural networks

ketos

1.0

Search

Introduction
Installation

Docs » Welcome to Ketos's documentation!

View page source

Welcome to Ketos's documentation!

Introduction

Ketos is a software package for acoustic data analysis with neural networks. It was developed with a particular eye to detection and classification tasks in underwater acoustics. Ketos is written in Python and utilizes a number of powerful software packages including NumPy, HDF5, and Tensorflow. It is licensed under the GNU GPLv3 license and hence freely available for anyone to use and modify. The project is hosted on GitLab at https://gitlab.meridian.cs.dal.ca/public_projects/ketos .

Ketos was developed by the MERIDIAN Data Analytics Team at the Institute for Big Data Analytics at Dalhousie University. We are greatful to Amalis Riera and Francis Juanes at the University of Victoria, Kim Davies and Chris Taggart at Dalhousie University, and Kristen Kanes at Ocean Networks Canada for providing us with annotated acoustic data sets, which played a key role in the development work. The first version of Ketos was released in April 2019.

The intended users of Ketos are primarily researchers and data scientists working with (underwater) acoustics data. While Ketos comes with complete documentation and comprehensive step-by-step tutorials, some familiarity with Python and especially the NumPy package would be beneficial. A basic understanding of the fundamentals of machine learning and neural networks would also be an advantage.

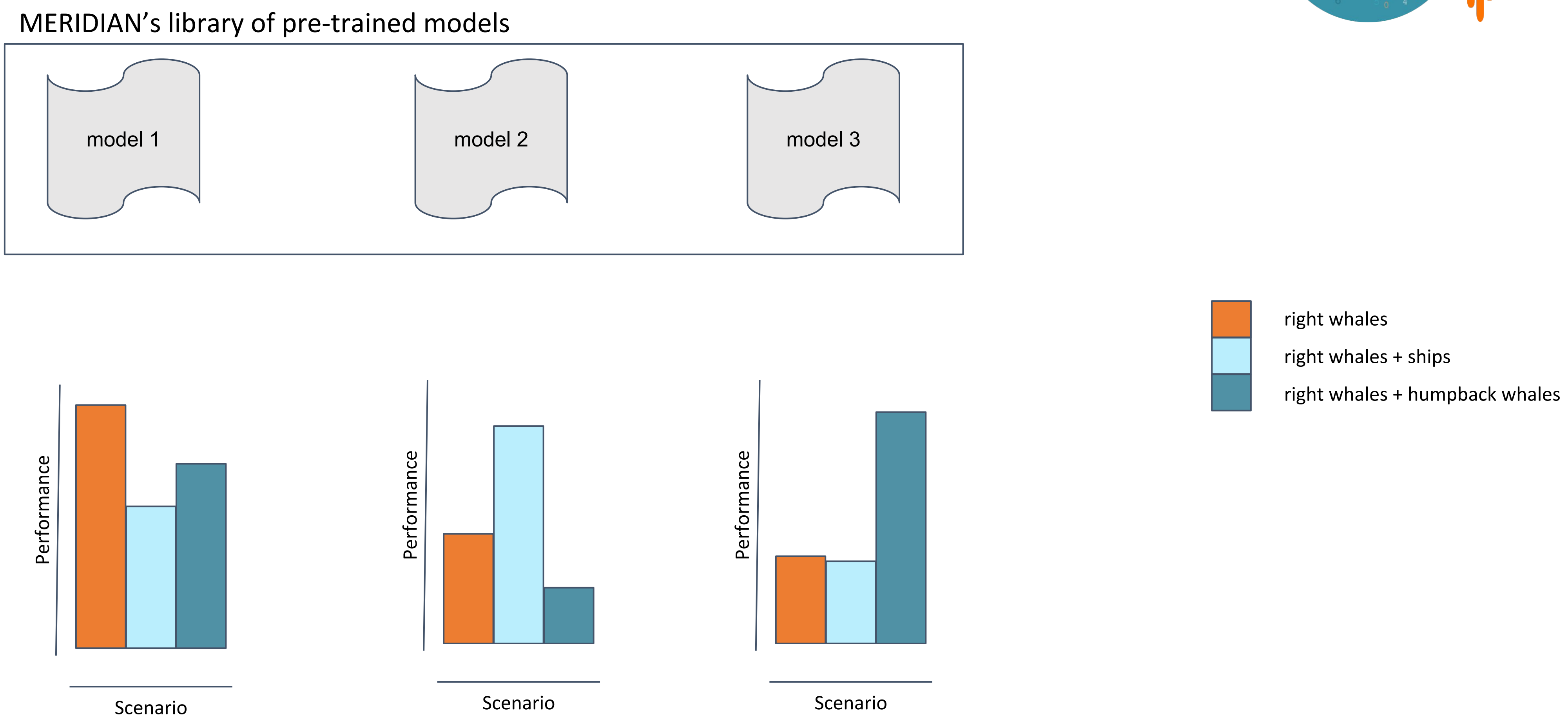
The name Ketos was chosen to highlight the package's main intended application, underwater acoustics. In Ancient Greek, the word ketos denotes a large fish, whale, shark, or sea monster. The word ketos is also the origin of the scientific term for whales, cetacean.

Indices and tables

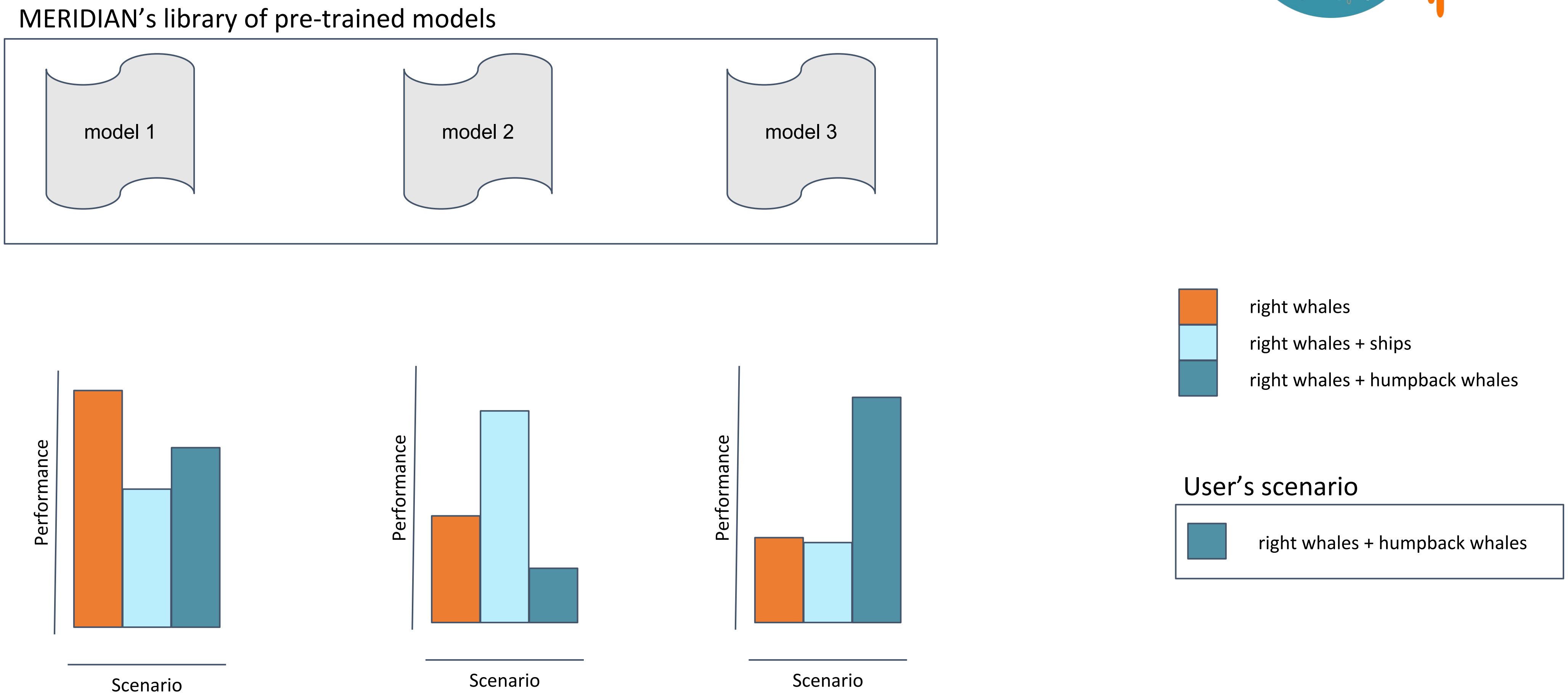
- Index
- Module Index

Workflow vision

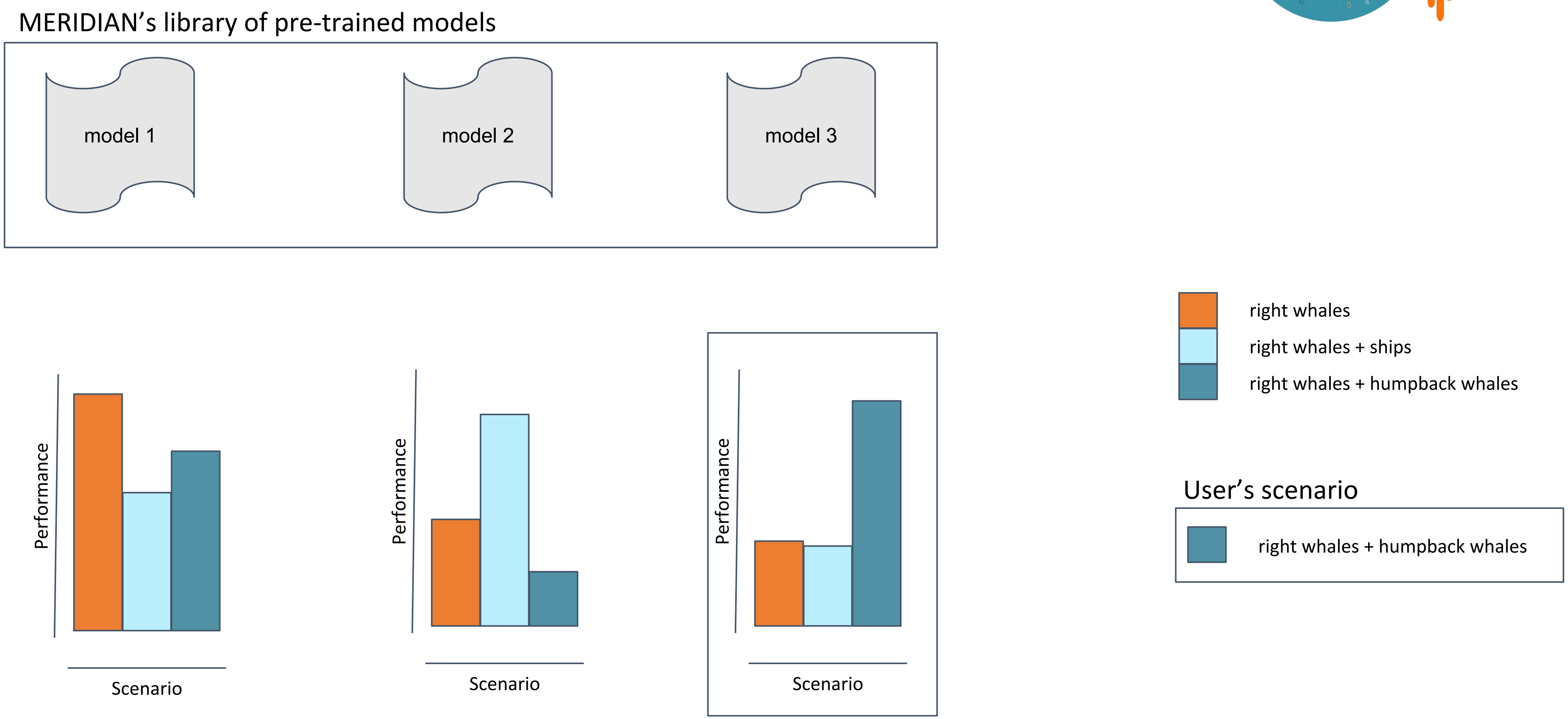
Workflow visions



Workflow visions

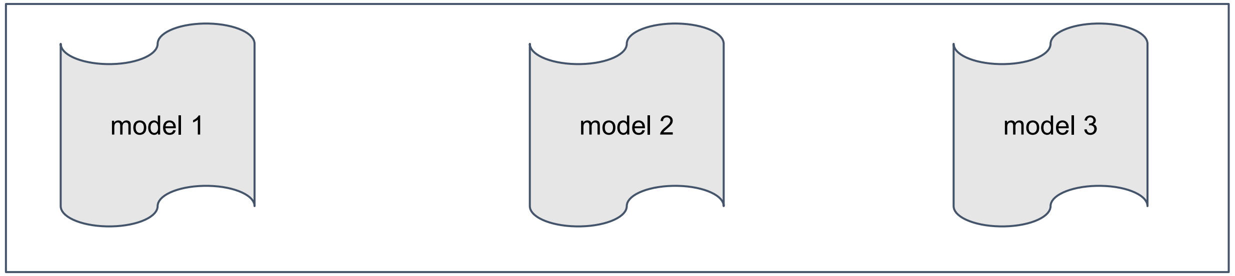


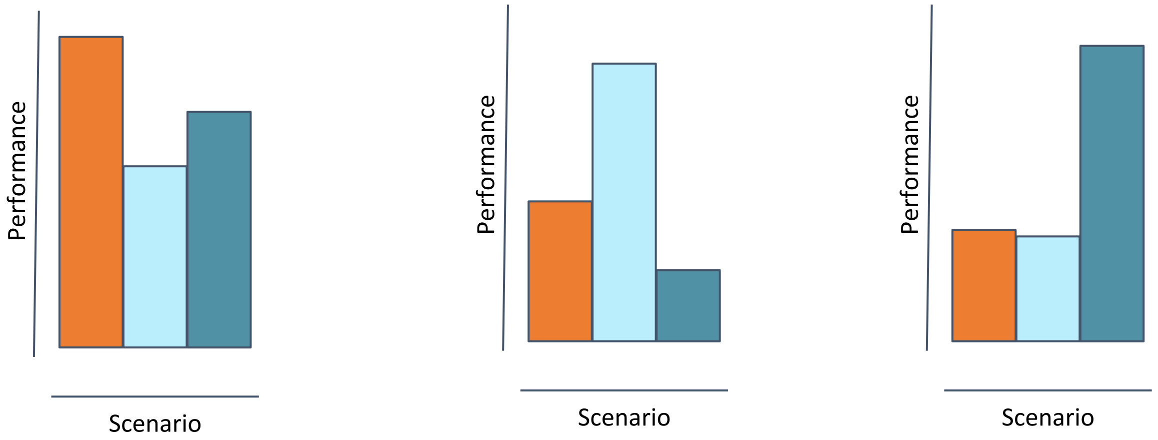
Workflow visions

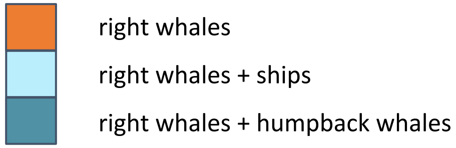


Workflow visions

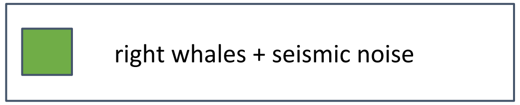
MERIDIAN's library of pre-trained models





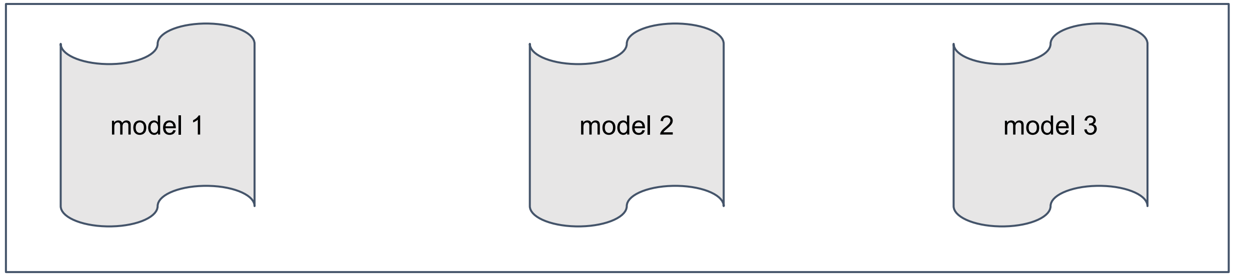


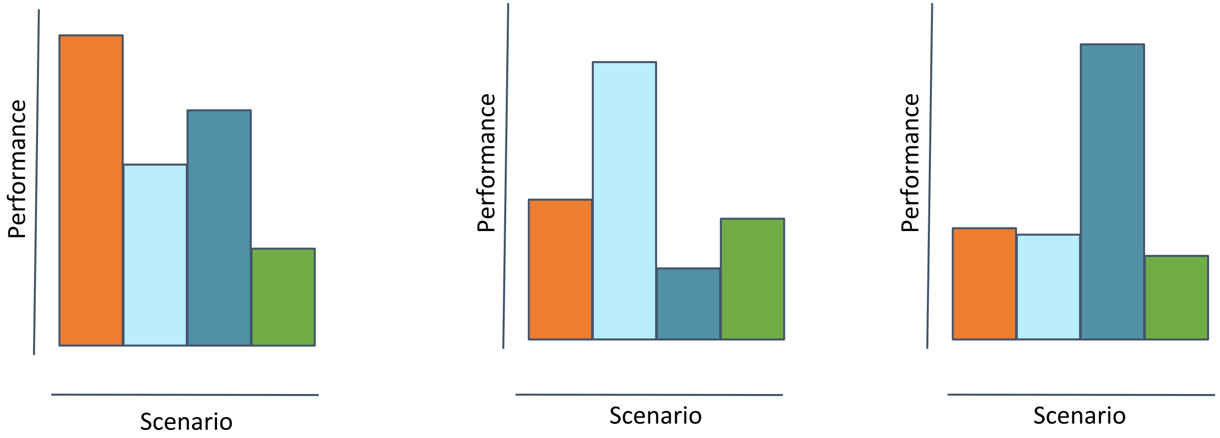
User's scenario



Workflow visions

MERIDIAN's library of pre-trained models





right whales

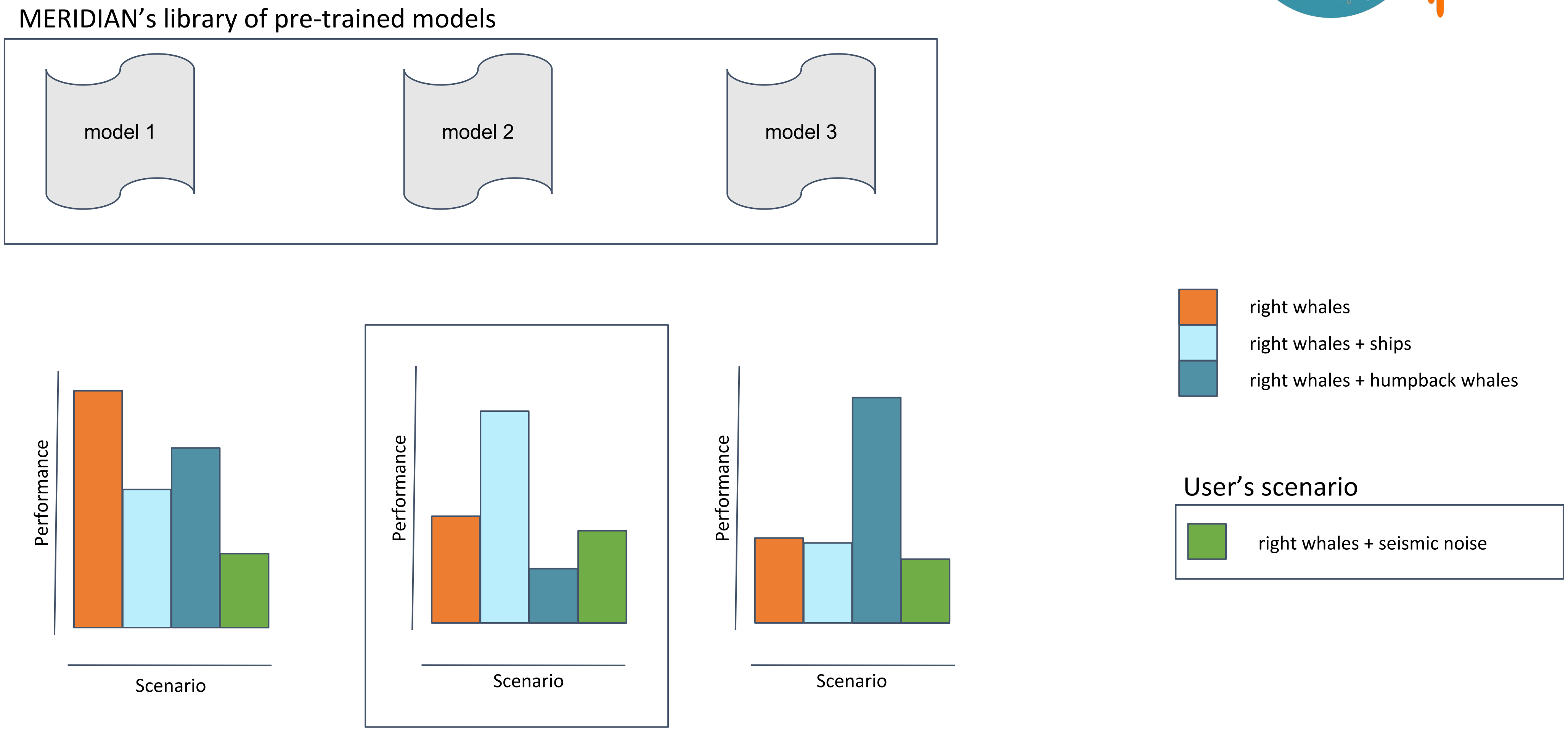
right whales + ships

right whales + humpback whales

User's scenario

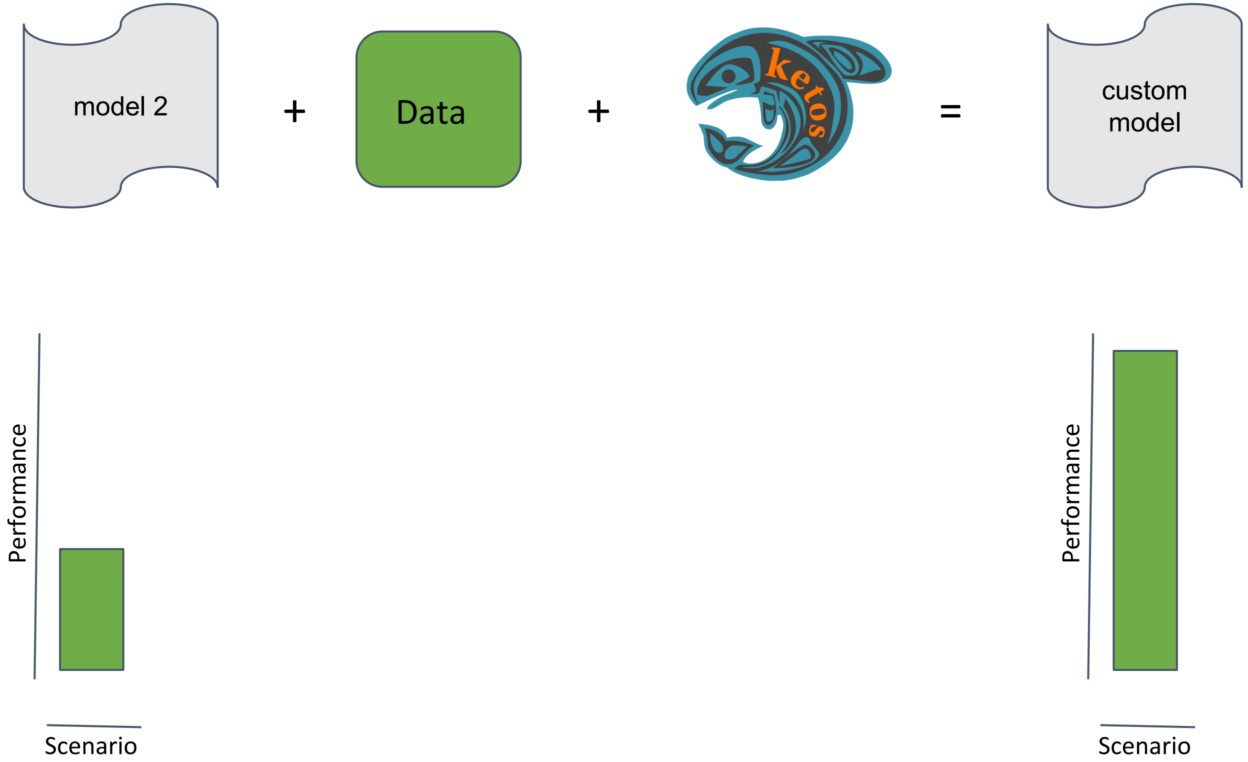
right whales + seismic noise

Workflow visions



Workflow visions

Transfer learning

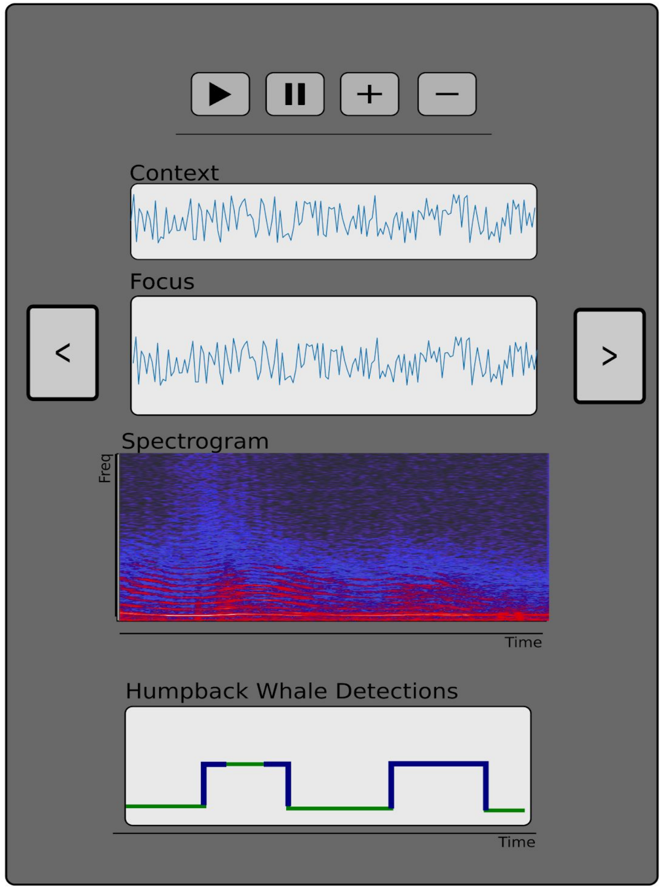


User's scenario

right whales + seismic noise

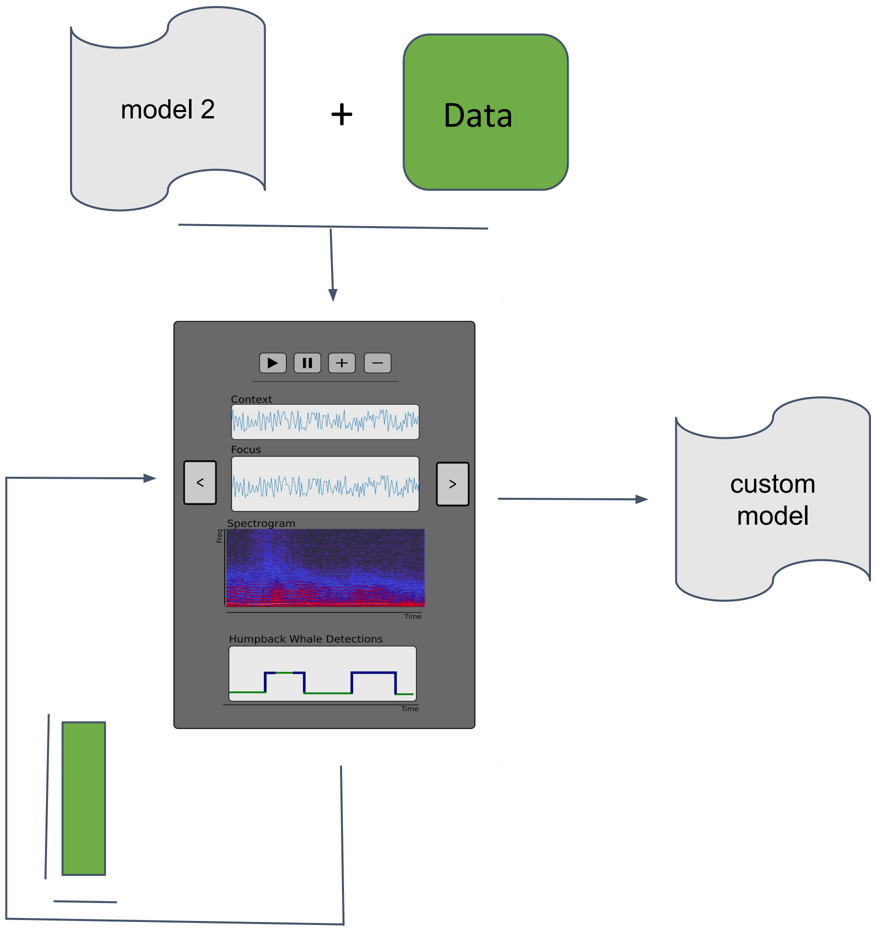
Interactive training tool

- visual interactive learning application
- human analyst and neural network work together
- human analyst inspects (and corrects) classifications proposed by neural network.
- neural network improves its performance in response to feedback from analyst



Workflow visions

(Inter)active learning



Workflow visions

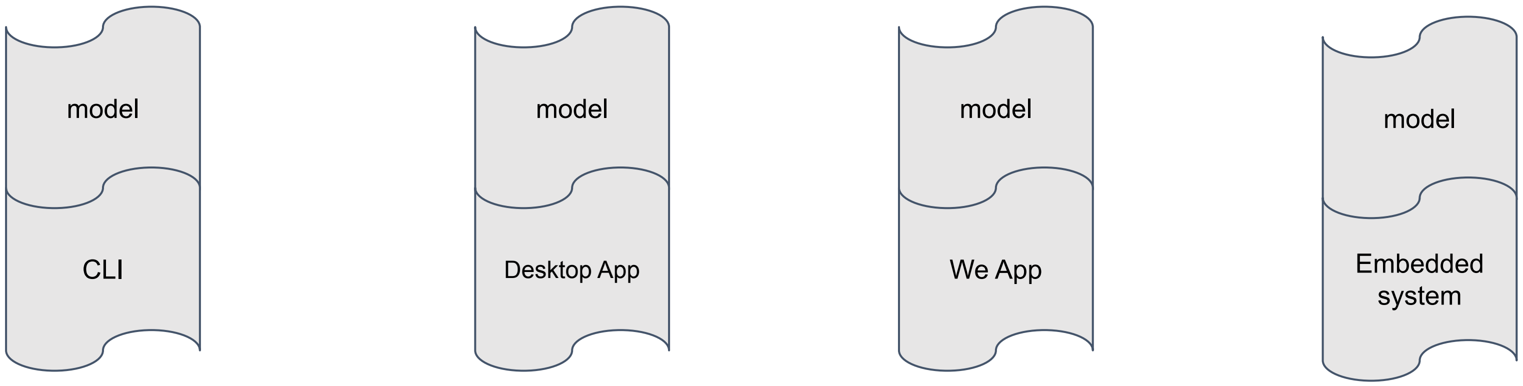
Data augmentation tool

- A large database of annotated underwater acoustic samples (multiple species/ call types)
- A variety of background noises
- Users can mix and match
- Data augmentation, signal processing and other methods also available

Your inputs!

How can we best serve the community?

What kind of interface to build for trained detectors/classifiers?



Thank you!