Introduction

- HALLO = Humans and ALgorithms Listening for Orcas
- Two-year R&D project
- $550k in support from DFO
- Multi-institutional & multi-disciplinary team
- Goal: Develop **deep learning** software for detecting and classifying the vocalisations of killer whales (esp. Southern Residents) to support researchers and conservationists in BC.
Deep learning can ...

- help us create better acoustic detection and classification (DC) models
- change the way we develop, use, and interact with these models
Deep Learning
What is deep learning?

Deep learning is an approach to machine learning that utilizes deep neural networks.

Everyday applications:

- Speech recognition & synthesis
- Face Recognition
- Translation
- Object detection
Deep learning in marine bioacoustics

Deep learning works!

It's time to build software tools that makes deep learning accessible to researchers and conservationists.
Deep learning aims to be end-to-end

Machine Learning

Feature extraction → Classification

Deep Learning

Feature extraction + Classification
Deep learning aims to be end-to-end.

Machine Learning

[Feature extraction] → [Classification] → [Orcas]

Deep Learning

[Feature extraction] + [Classification] → [Orcas]
Transfer learning (model adaptation)

- Transfer learning can ...
  - drastically reduce amount of training data and training time
  - make models more adaptable and reusable
The deep learning workflow

Neural network adaptation to new data

1. Pre-trained network
   The Deep Neural Network detects sounds and proposes classifications.

2. Validation
   A human analyst validates the network’s detections and classifications.

3. Improved training data
   The new validated examples are added to the network’s pool of training data.

4. Enhanced performance
   The human analyst can now retrain the neural network to improve its performance.

p. 144-145:
"Towards Versatile and Adaptive Detection Algorithms in Underwater Acoustics"
The HALLO Project
HALLO - at a glance

- HALLO = Humans and ALgorithms Listening for Orcas
- Two-year R&D project
- $550k in support from DFO
- Multi-institutional & multi-disciplinary team

**Goal:** Develop deep learning software for detecting and classifying the vocalisations of killer whales (esp. Southern Residents) to support researchers and conservationists in BC.

- Marine bioacousticians
- Data scientists
- Deep learning experts
- Data managers
- Software developers
- System administrators
Say hello to the HALLO team!

Ruth Joy
Steven Bergner
Alex Harris
Emma Cummings
Kaitlin Palmer
Jennifer Wladichuk

Amalis Riera
Fabio Frazao
Bruno Padovese
Oliver Kirsebom

Dave Campbell
Paul Nguyen Hong Duc

Scott Veirs
Val Veirs
1. Academic studies

ORCA-SPOT: An Automatic Killer Whale Sound Detection Toolkit Using Deep Learning
Christian Bergler, Hendrik Schröter, Rachael Xi Cheng, Volker Barth, Michael Weber, Elmar Nöth, Heribert Hofer, and Andreas Maier

Whistle detection and classification for whales based on convolutional neural networks
Jia-jia Jiang, Ling-ran Bu, Fa-jie Duan, Xian-quan Wang, Wei Liu, Zhong-bo Sun, Chun-yue Li

Deep Learning and Domain Transfer for Orca Vocalization Detection
Paul Best, Maxence Ferrari, Marion Posnard, Sébastien Paris, Ricard Martor, Helena Symonds, and Paul Spong, Hervé Glotin

*Univ. Toulon, Aix Marseille Univ., CNRS, L3S, DFN Marseille, France
1LAMFA, CNRS Amiens France
2BOSONG SARL, France
3OrcaLab Alert Bay
Email: paul.best@univ-tln.fr
2. Collaborations between big tech and NGO/government
Related studies and projects

Results:
- The "CNN binary classifier" approach has produced impressive results (on restricted data sets)
- But models generalize poorly to new acoustic environments

Open questions:
- Can we make models generalize better? (e.g. by enlarging the training dataset)
- And/or can we develop methods (e.g. transfer learning) that allow us to readily adapt models to new environments?
- Can we teach models to distinguish between KW ecotypes, and pods?

Still missing:
- Ready-to-use tools
HALLO contributions

1. High-quality annotated KW acoustic data
2. Improved deep learning models; ecotype- and pod-level classification
3. Ready-to-use tools
4. Open-source software, developed with reusability and extensibility in mind
5. Collaborative approach; data and code publicly available
Also, ongoing conversations about data sharing with

- Cetacean Research Program, DFO Science (T. Doniol-Valcroze)
- Harald Yurk

among others ...

**Multiple data sources ➔ large variance**
Active learning approach

Active Learning

- Labeled Dataset
- Unlabeled Pool
- ML Model
- Annotator (Oracle)
- Add Labels
- Query Selection
- Retrain
- Predict

Credits: Kunal Mehta
HALLO products

1. Training and test datasets (accessible via web portal)

2. Trained deep learning models

3. Ready-to-use tools
   - Plug-in for PAMGuard
   - Active learning & transfer learning application

4. Workshops

5. Scientific papers
Thank you!