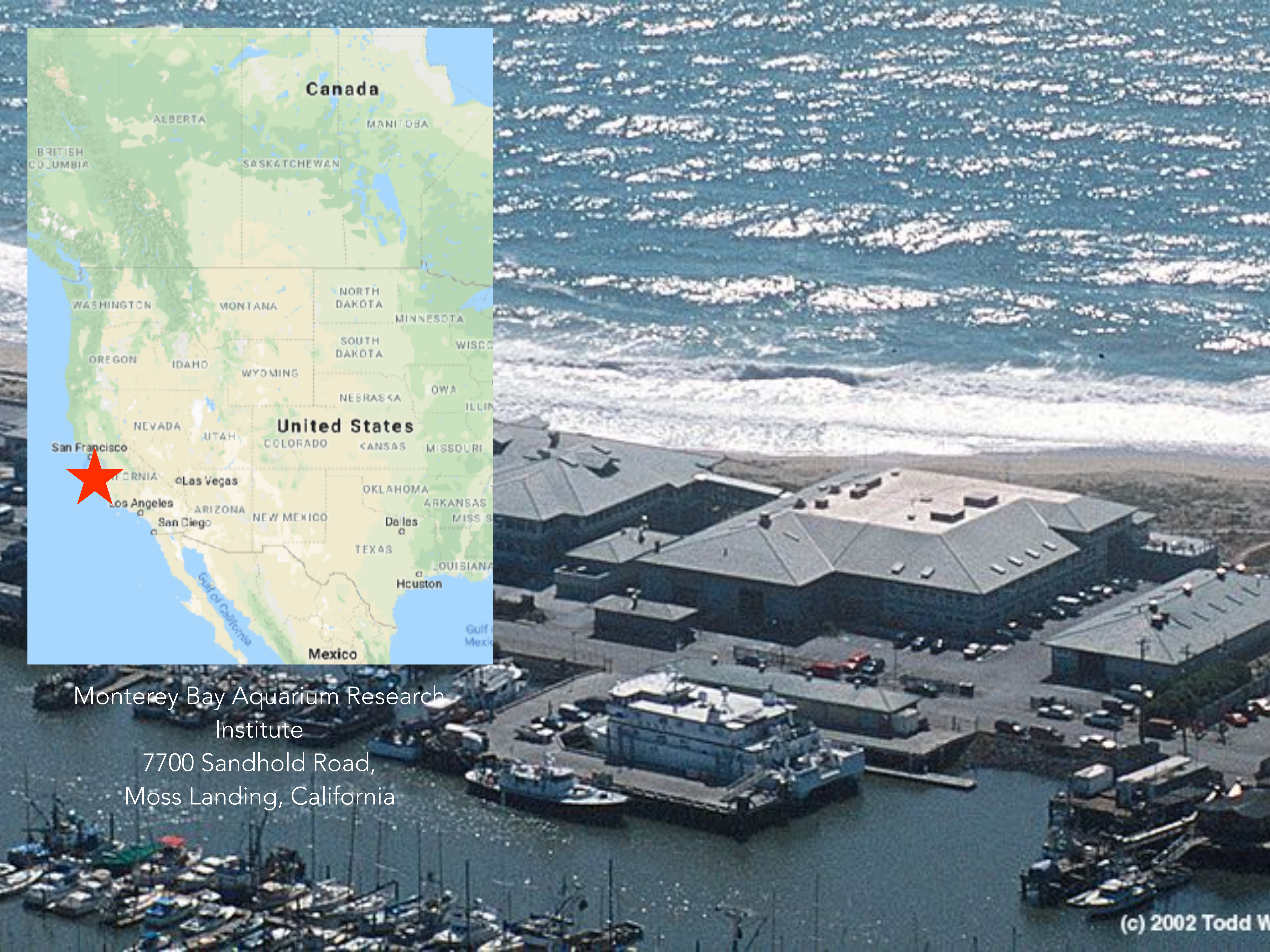
DANELLE CLINE

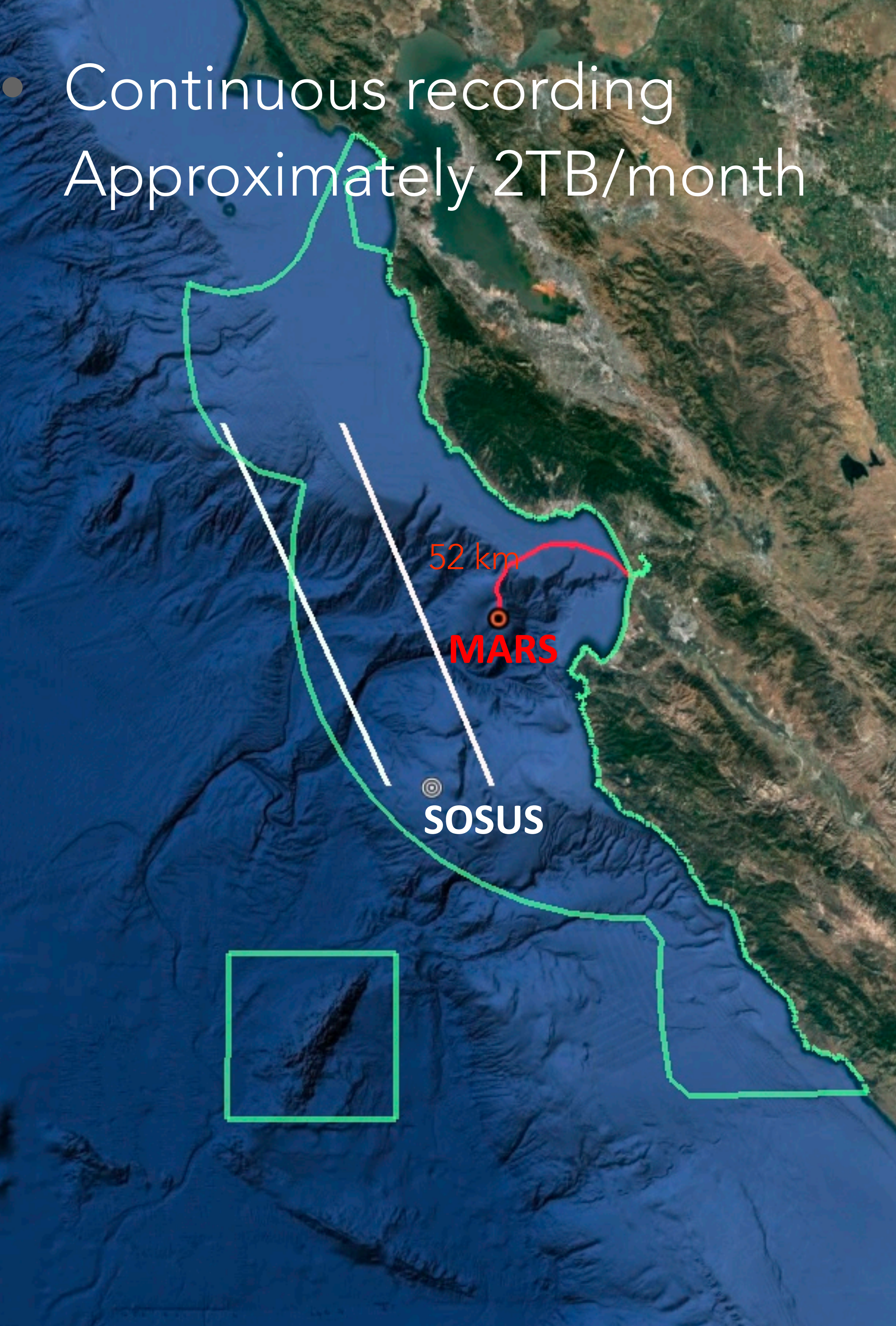
Detection and classification of whales calls using band-limited energy detection and transfer learning

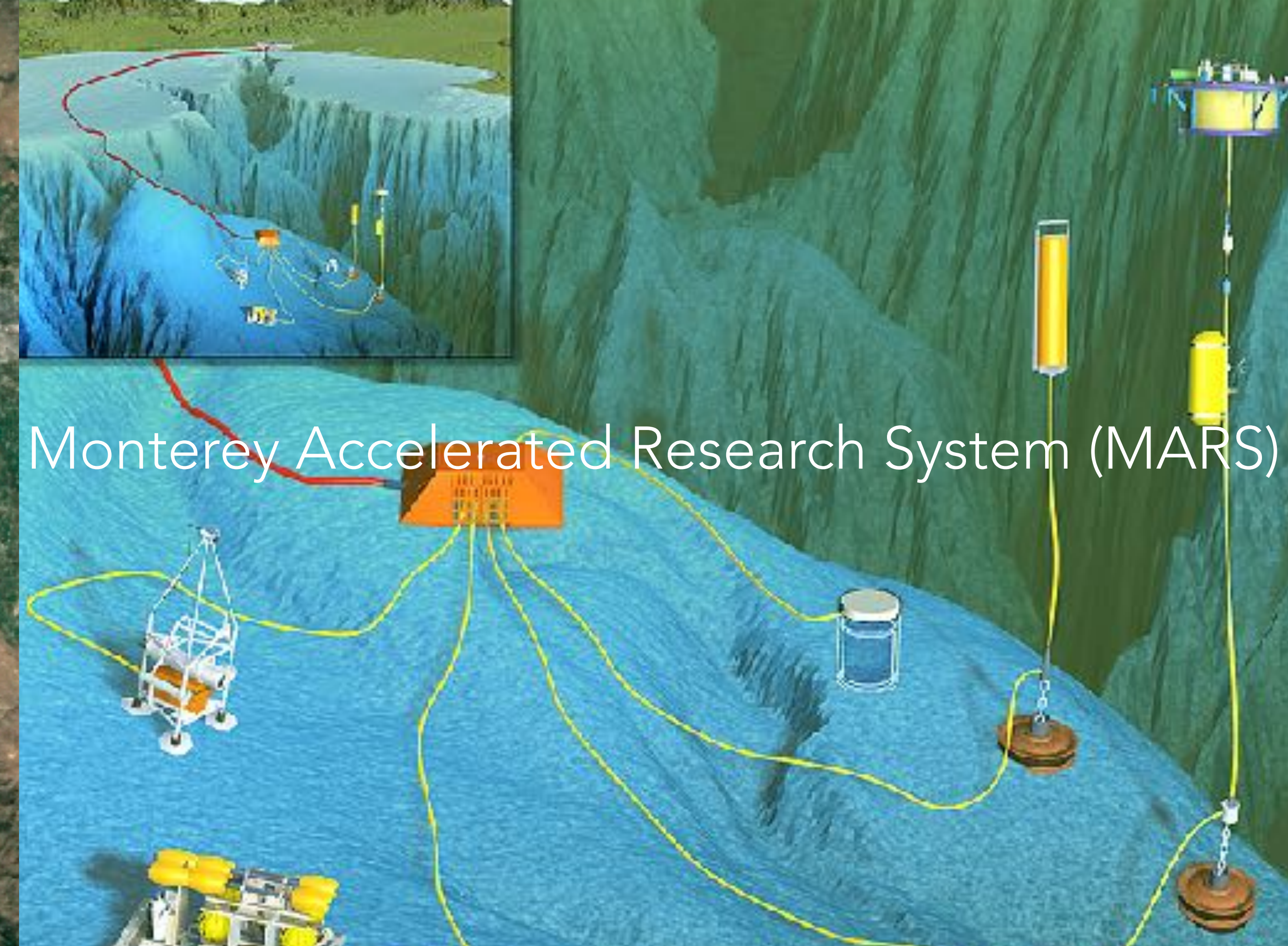
Monterey Bay Aquarium Research Institute

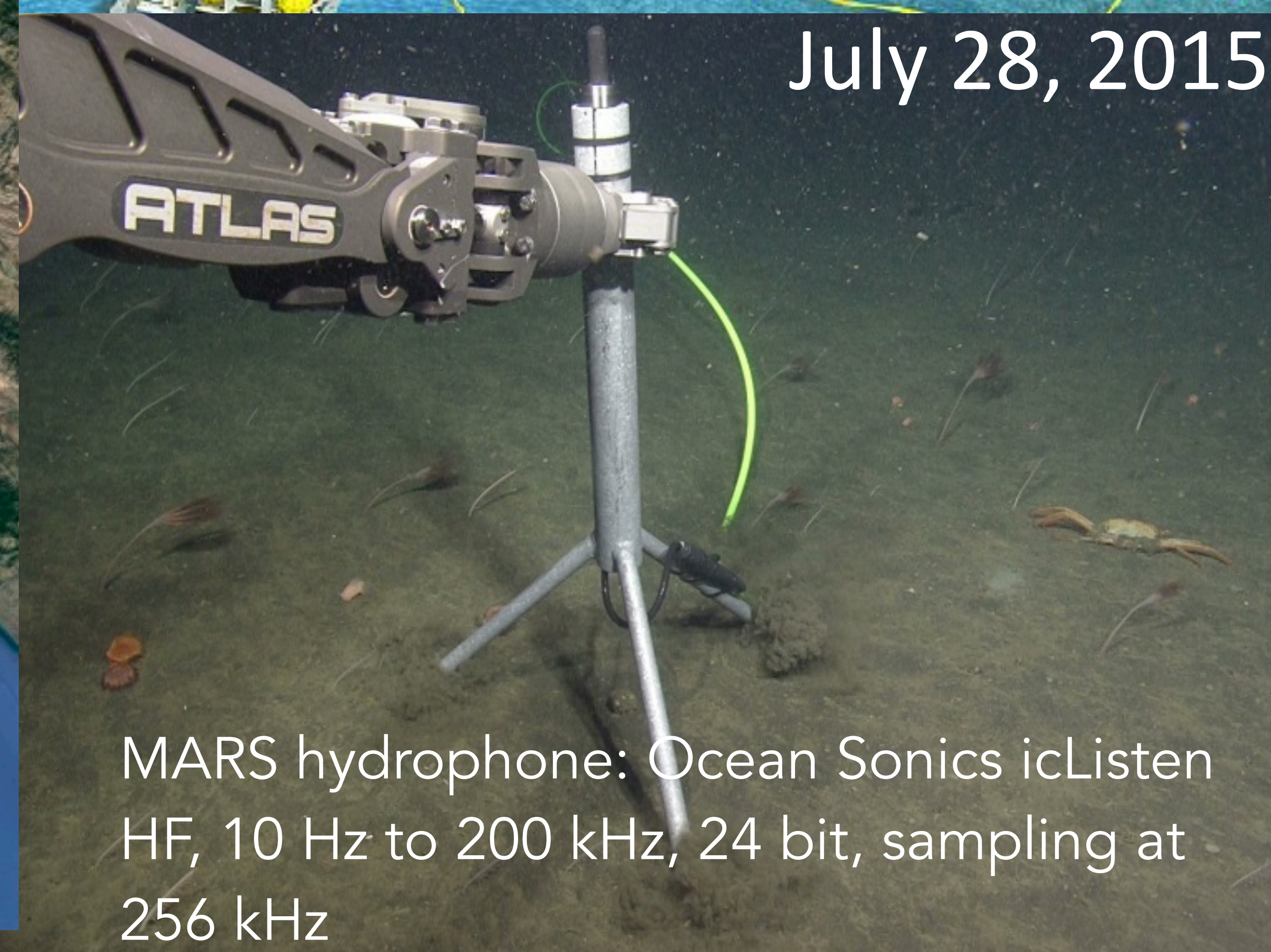


Monterey Bay Aquarium Research Institute
7700 Sandhold Road,
Moss Landing, California

- Continuous recording
  Approximately 2TB/month

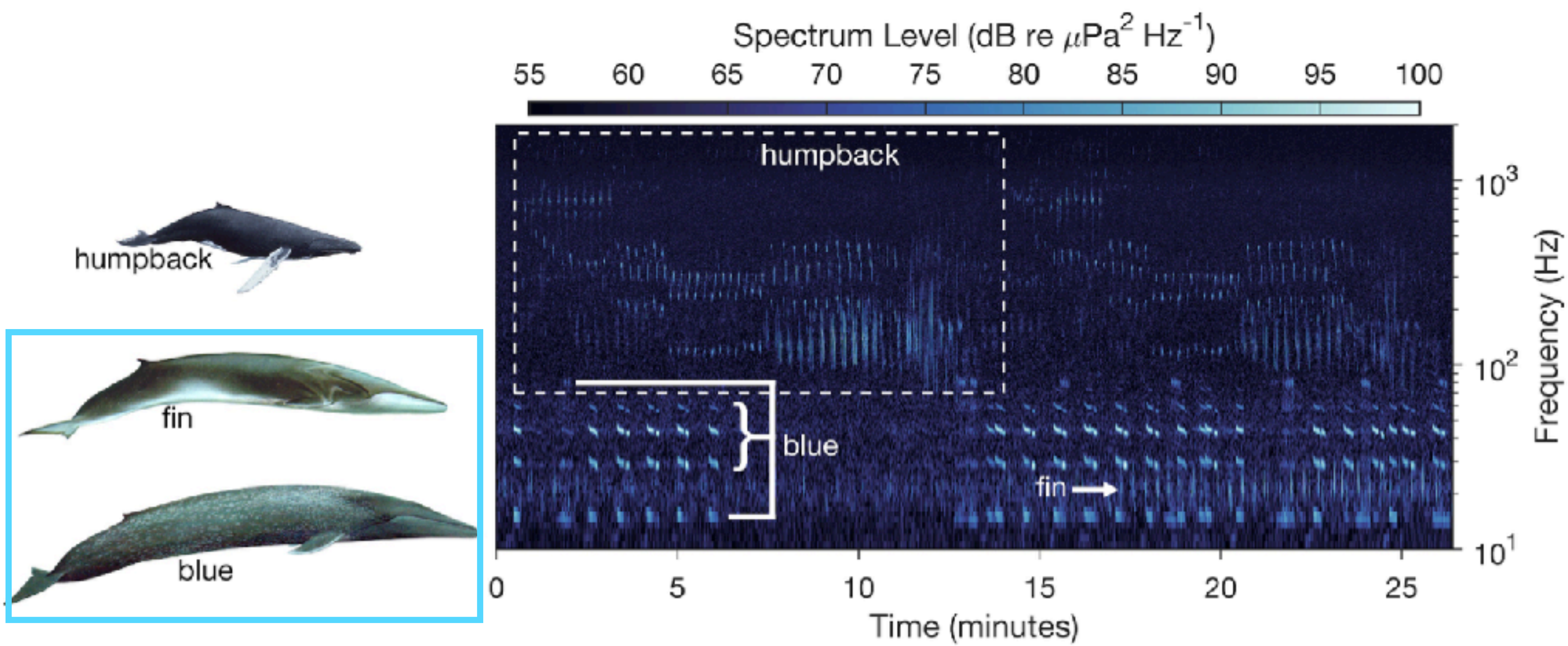






GOAL

DETECTION AND CLASSIFY BLUE AND FIN WHALE CALLS



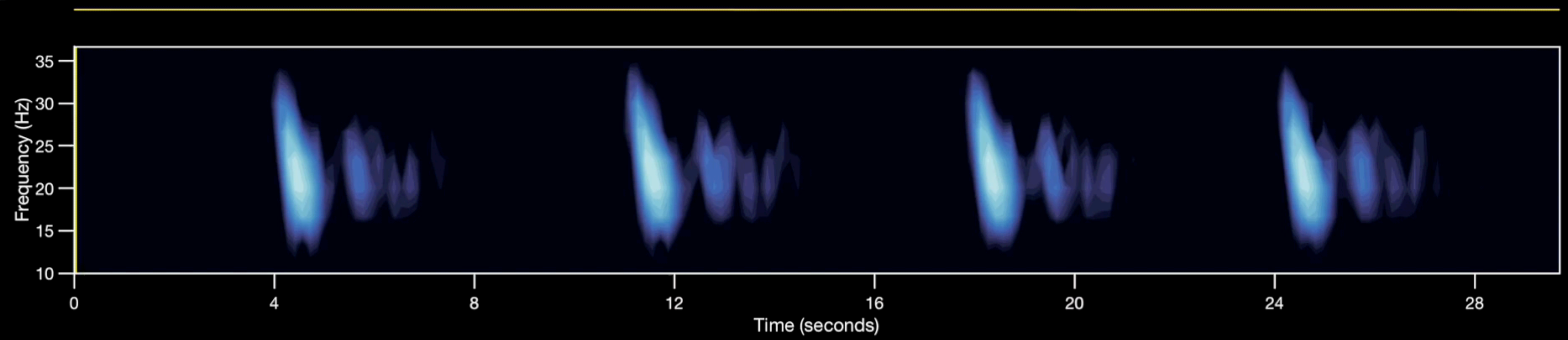
whale images by Larry Foster   spectrogram image by John Ryan

"Greyhound of the sea"

FIN WHALE

endangered

forages on schooling fish, squid, crustaceans

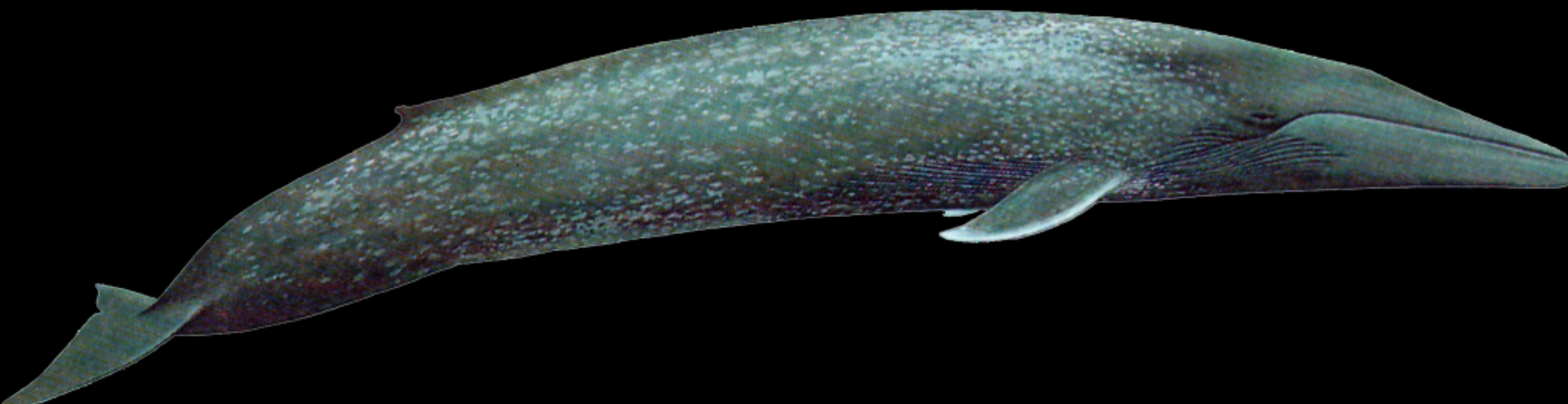


20 Hz pulses

"Largest animal ever known to have ever existed"
- Wikipedia

BLUE WHALE

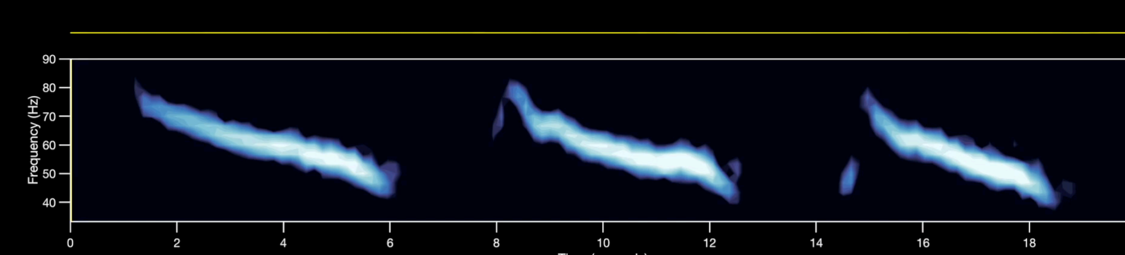
Eastern North Pacific hosts the largest population
forages on krill - up to 3,600 kg/day!



A call

B call

D call



MBARI

OVERVIEW

- **Data Decimation and Filtering**
- **Detection and Classification**

DECIMATION AND FILTERING

- Goal is to decimate the 256 kHz hydrophone data down to 2 kHz, a factor of 128, avoiding higher frequency aliasing

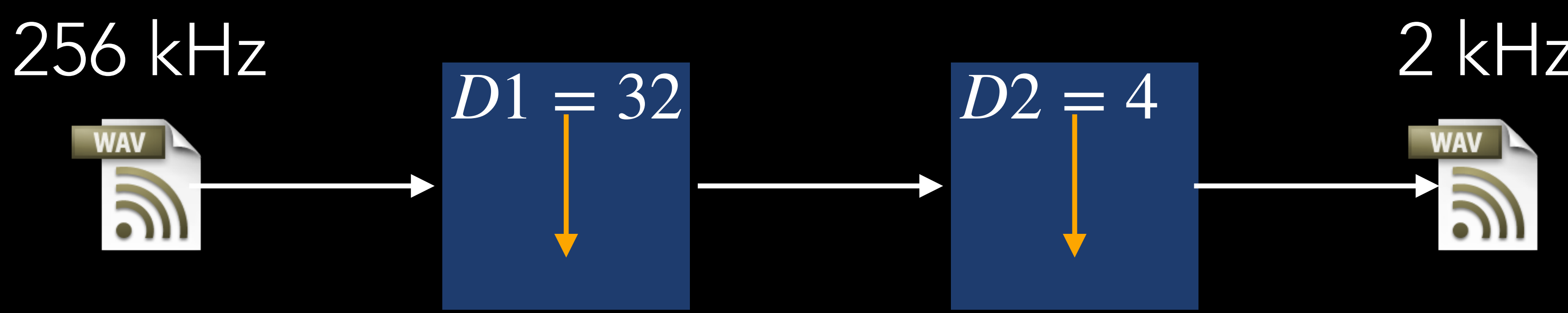
- Requires low-pass filtering the signal to limit its bandwidth to less than 1 kHz, and downsampling (i.e. throwing out 127 out of every 128 samples)

- Why not use available routines such as Matlab's decimate or Scipy's scipy.signal.decimate?

- We want control of the details, e.g. filter type (IIR or FIR), filter length, window type, phase linearity, etc.

MBARI

DECIMATION FILTER IMPLEMENTATION DETAILS



Filtering in multiple stages is faster than a single stage

$$D1 = 2 * D * (1 - \sqrt{D * F/(2 - F)})/(2 - F * (D + 1))$$

$$D = 128$$

$$F = f/fstop$$

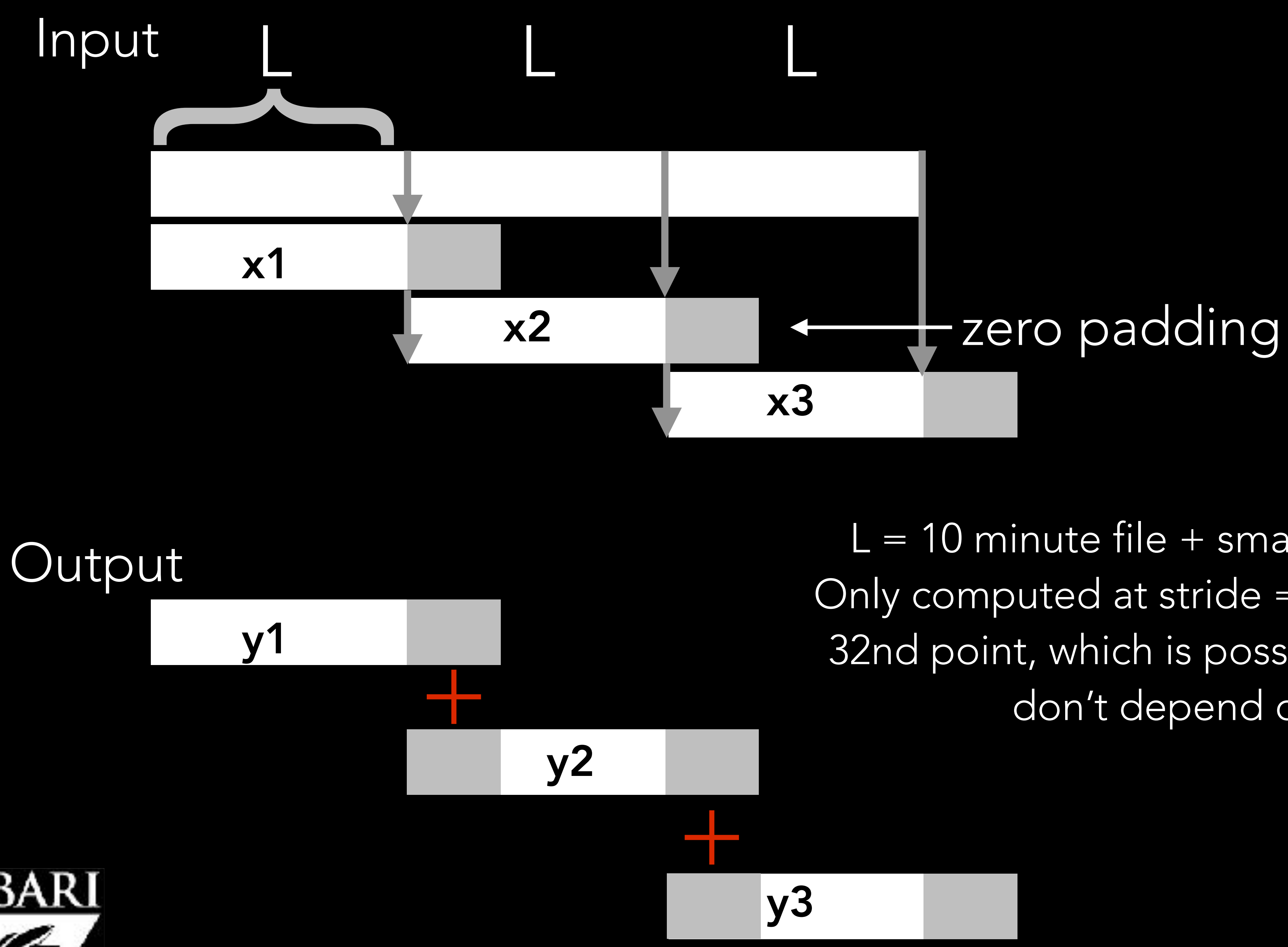
$$f = 100(10\,\%\,offinalstopband)$$

$$F = 100/1000$$

Lyons, R. G. (2004). Understanding digital signal processing (2nd ed.). Upper Saddle River, NJ: Prentice Hall

https://bitbucket.org/mbari/pam-decimate-notebook

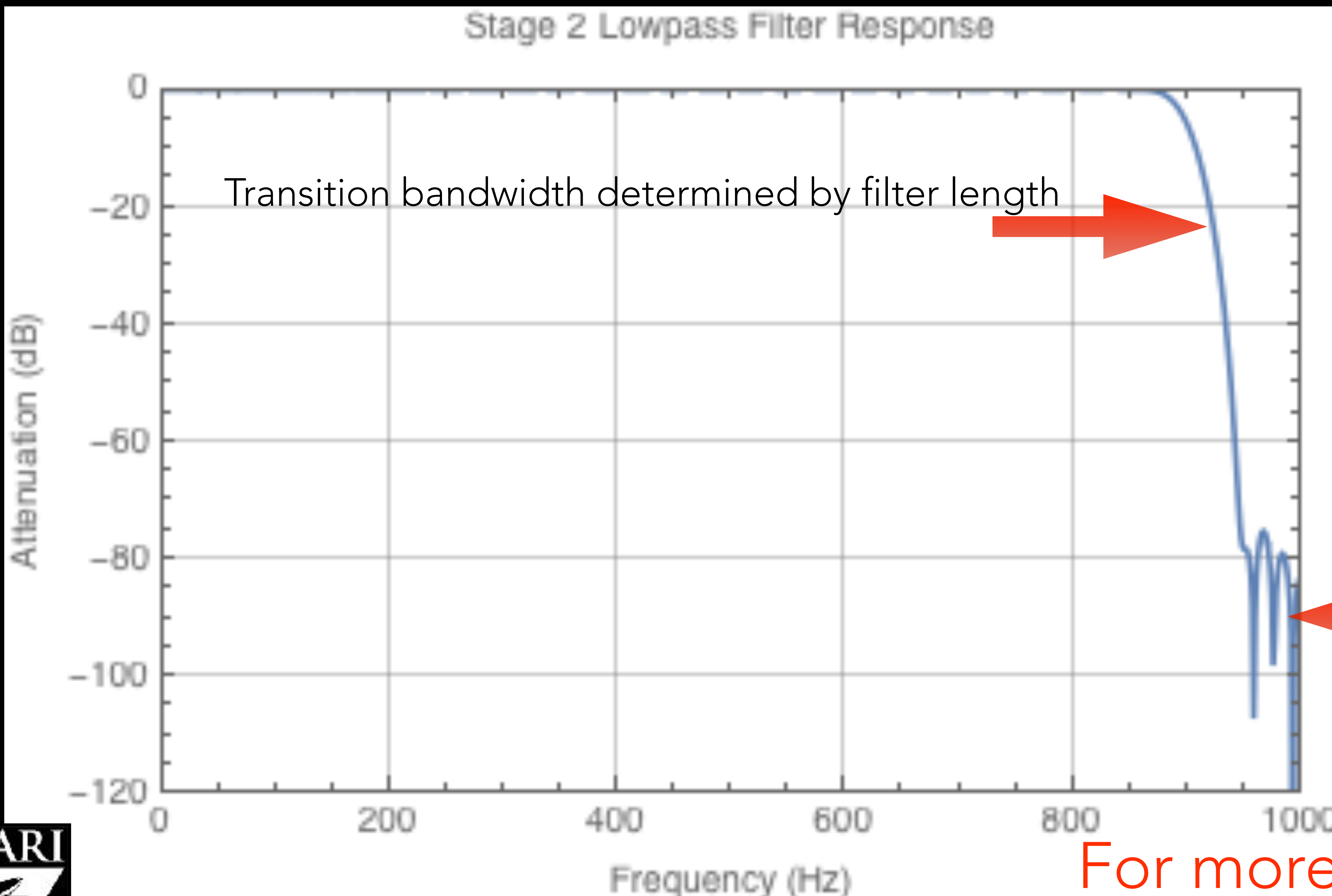
DECIMATION FILTER IMPLEMENTATION DETAILS



L = 10 minute file + small padding either side
Only computed at stride = decimation, e.g. every 32nd point, which is possible because FIR filters don't depend on past input

MBARI

DECIMATION FILTER



Attenuation in stop band determined by window type

For more details

https://bitbucket.org/mbari/pam-decimate-notebook

DETECTION AND CLASSIFICATION

- **Detect** candidate calls using Band-Limited-Energy-Detectors *(simple, fast to compute, not very intelligent)*

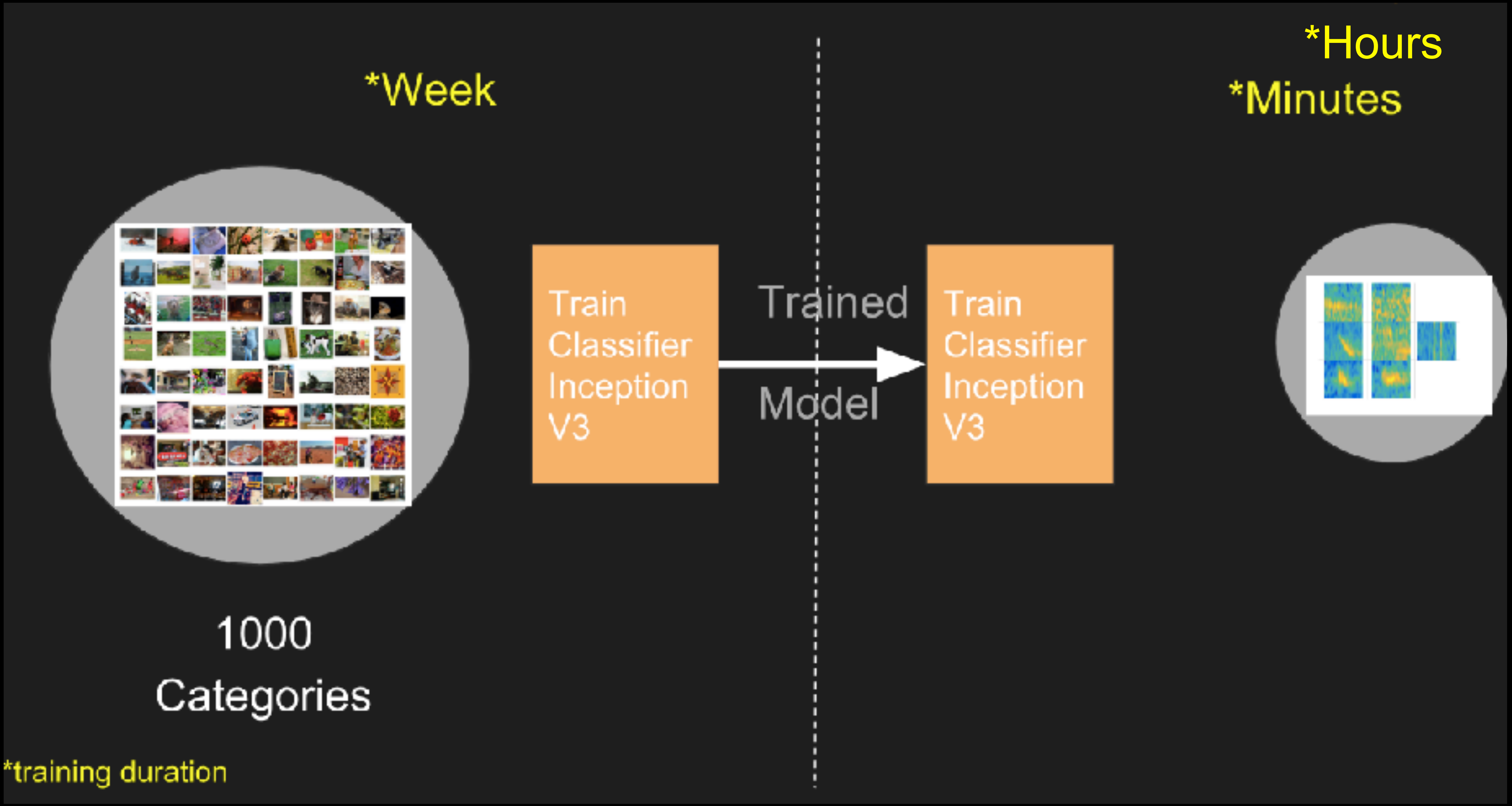
- **Classify** candidate calls using transfer learning (more intelligent)

MBARI

BAND-LIMITED ENERGY DETECTOR - TWO-STAGE APPROACH

- Stage 1: find every sequence of consecutive spectral slices where the in-band power exceeds a specified threshold
- Stage 2: candidates that are too short, too long, or too sparse (occupancy is below the specified threshold) are discarded
- Don't expect perfection with this method. This requires iteration. Raven simplifies this process with an interactive tool.

DETECTION AND CLASSIFICATION



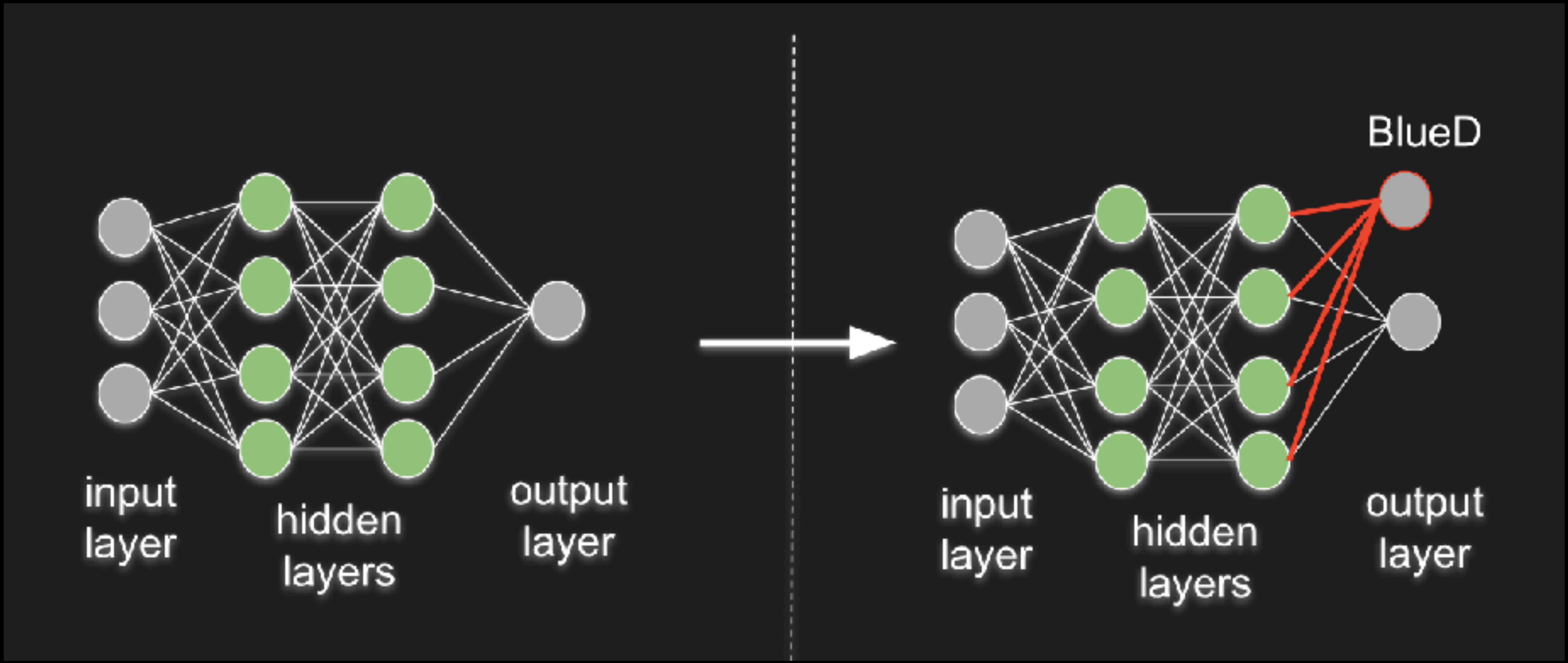
Transfer Learning Concept

Start with pre-trained model (this may have taken a week to train, but you don't need to worry about that)

train, or "fine-tune" for the problem of whale classification

MBARI

DETECTION AND CLASSIFICATION



Retrain the output dense layers only

MBARI

SPECTROGRAM ENHANCEMENT

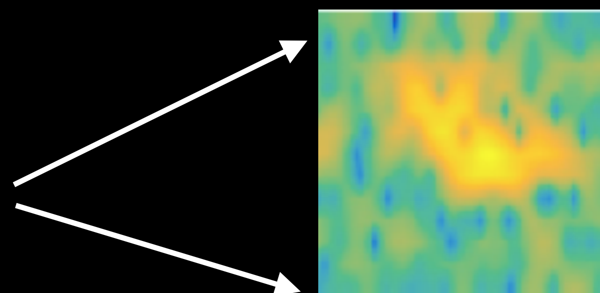
1000 Hz

0 Hz

Raw

Cropped and normalized

36 Hz



10 Hz

3x3 blur convolving in in frequency/time

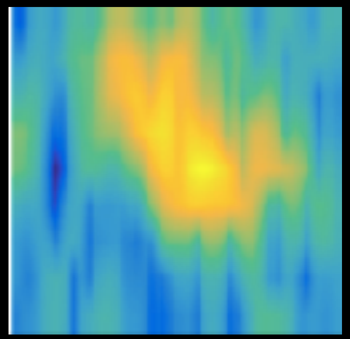
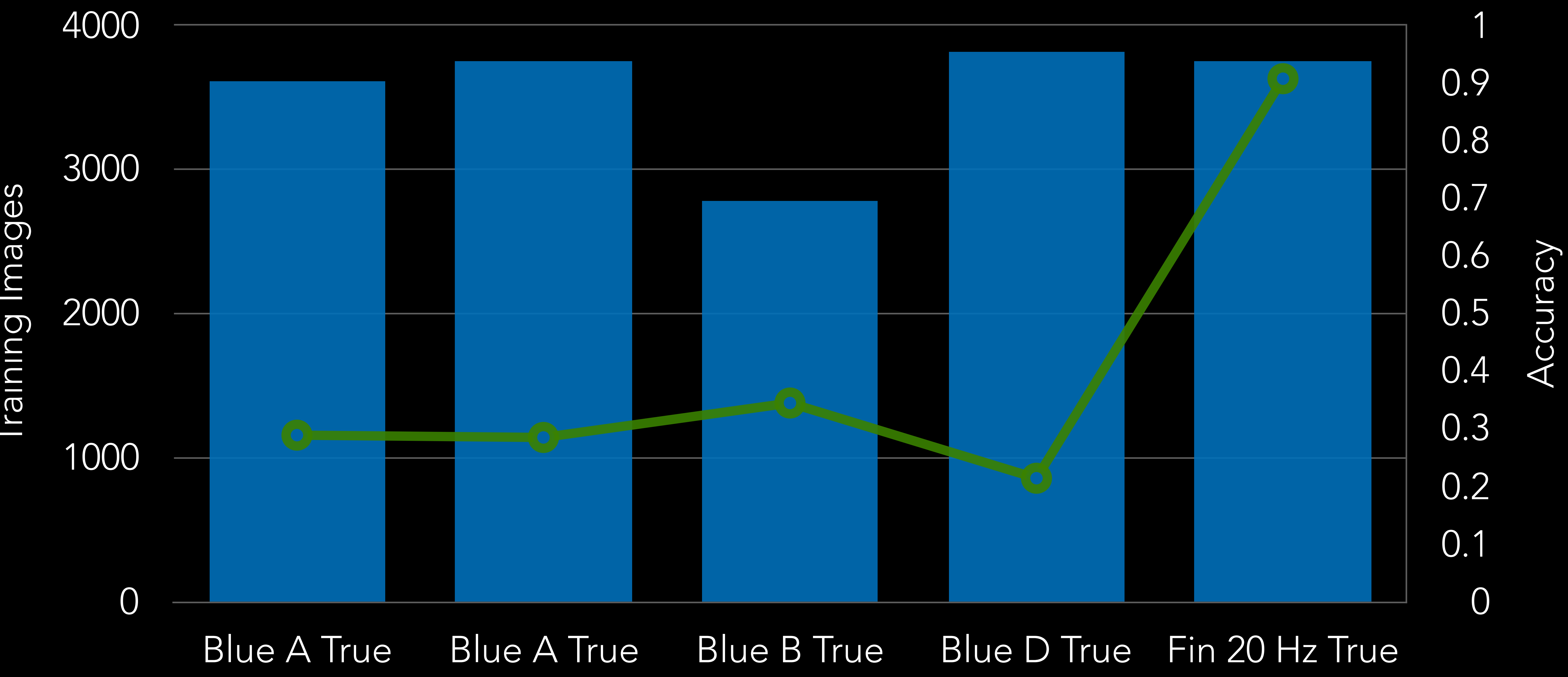


Image resized 299x299 pixels as required for CNN

https://bitbucket.org/mbari/pam-spectrogram-notebook

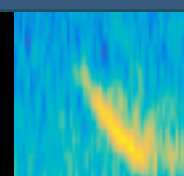
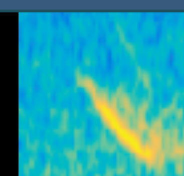
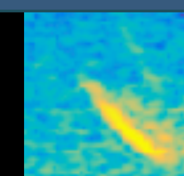
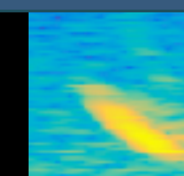
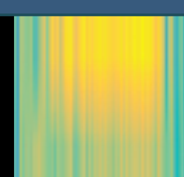
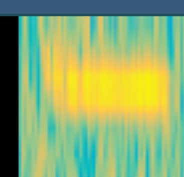
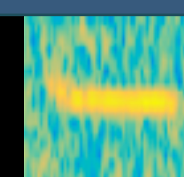
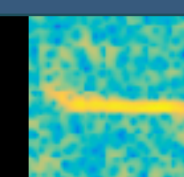
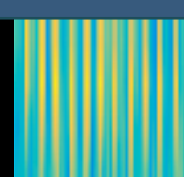
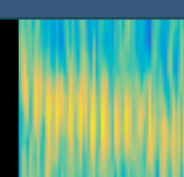
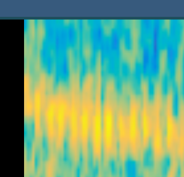
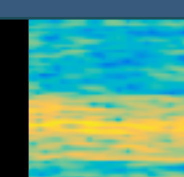
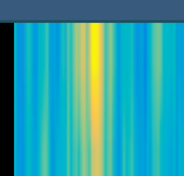
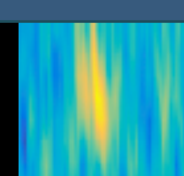
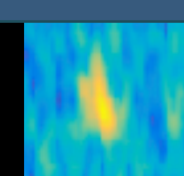
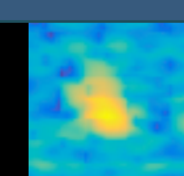
CLASSIFICATION ACCURACY

ACCURACY 80/20 TRAIN/TEST SPLIT, 30K STEPS, RANDOM CROP



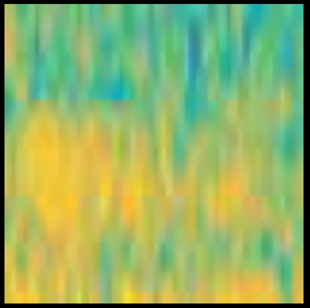
CLASSIFICATION VARYING FFT LENGTHS

AVERAGE ACCURACY 80/20 TRAIN/TEST SPLIT, 30K STEPS, RANDOM CROP

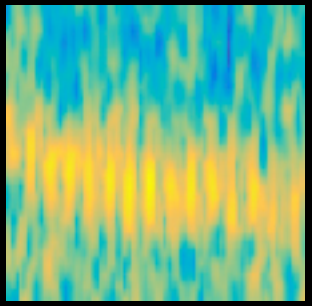
	FFT LENGTH, 0.95% OVERLAP, HANNING			
	512	1024	2048	4096
BLUE D	**0.967**	0.952	0.953	0.919
				
BLUE B	0.637	0.643	0.631	**0.669**
				
BLUE A	0.953	0.938	0.95	**0.96**
				
FIN 20 HZ	0.893	0.907	0.895	0.909
				

MBARI

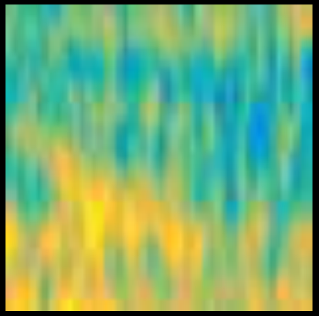
MISCLASSIFICATIONS



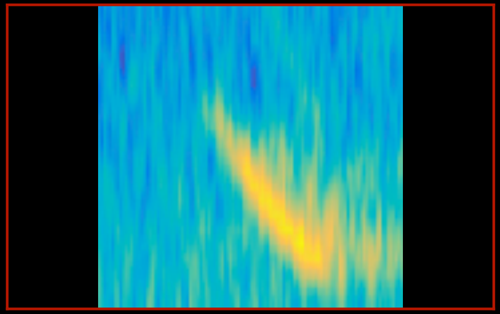
Blue A True
(predicted as false)



Example

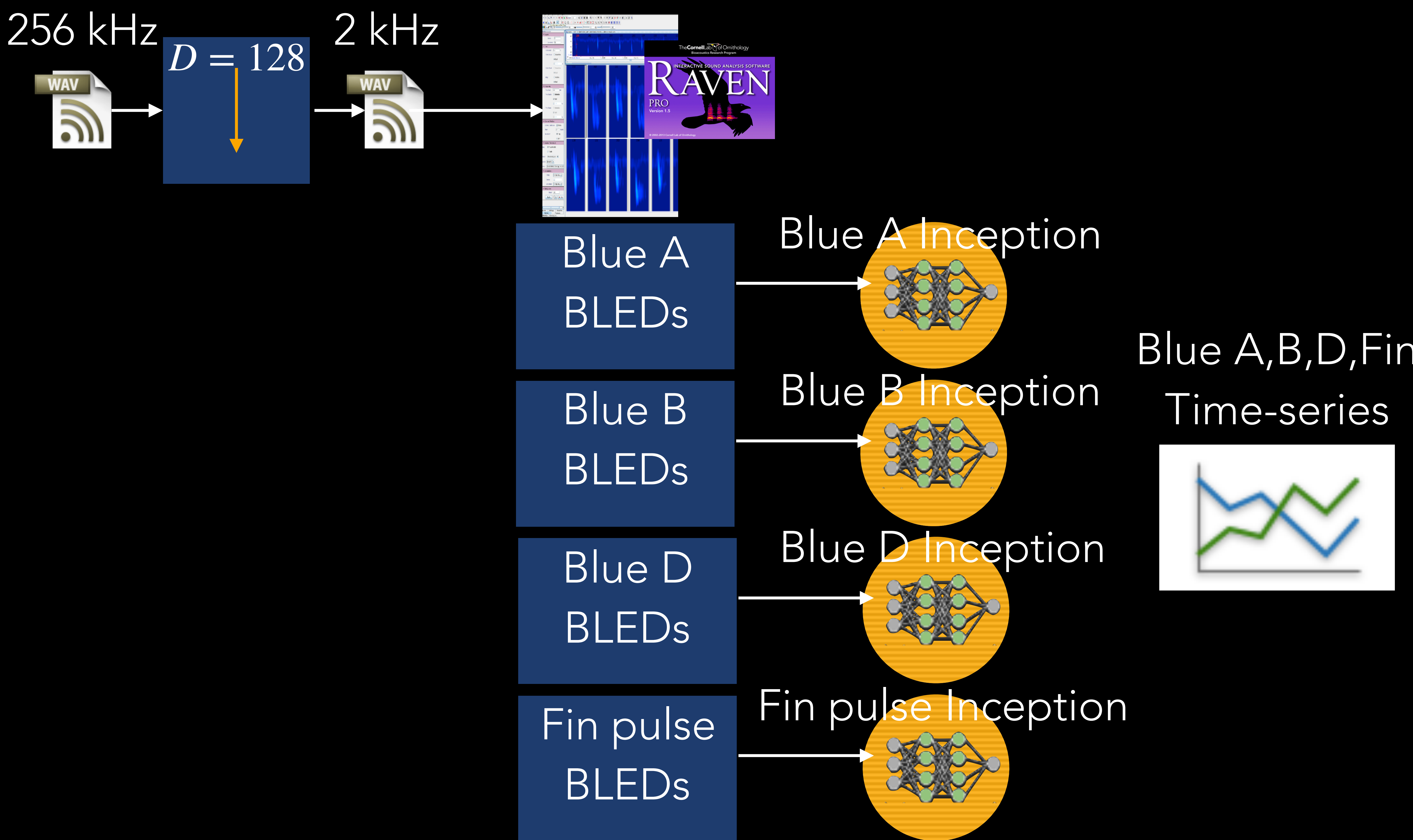


Blue D True
(predicted as False)



Example

DECIMATION, DETECTION AND CLASSIFICATION WORKFLOW



CONCLUSIONS

- BLED can difficult to fine-tune. Verbose detections required for accurate classification. Sensitive to noise.

- High classification accuracy for most classes. More training data and curation should improve accuracy.

MBARI

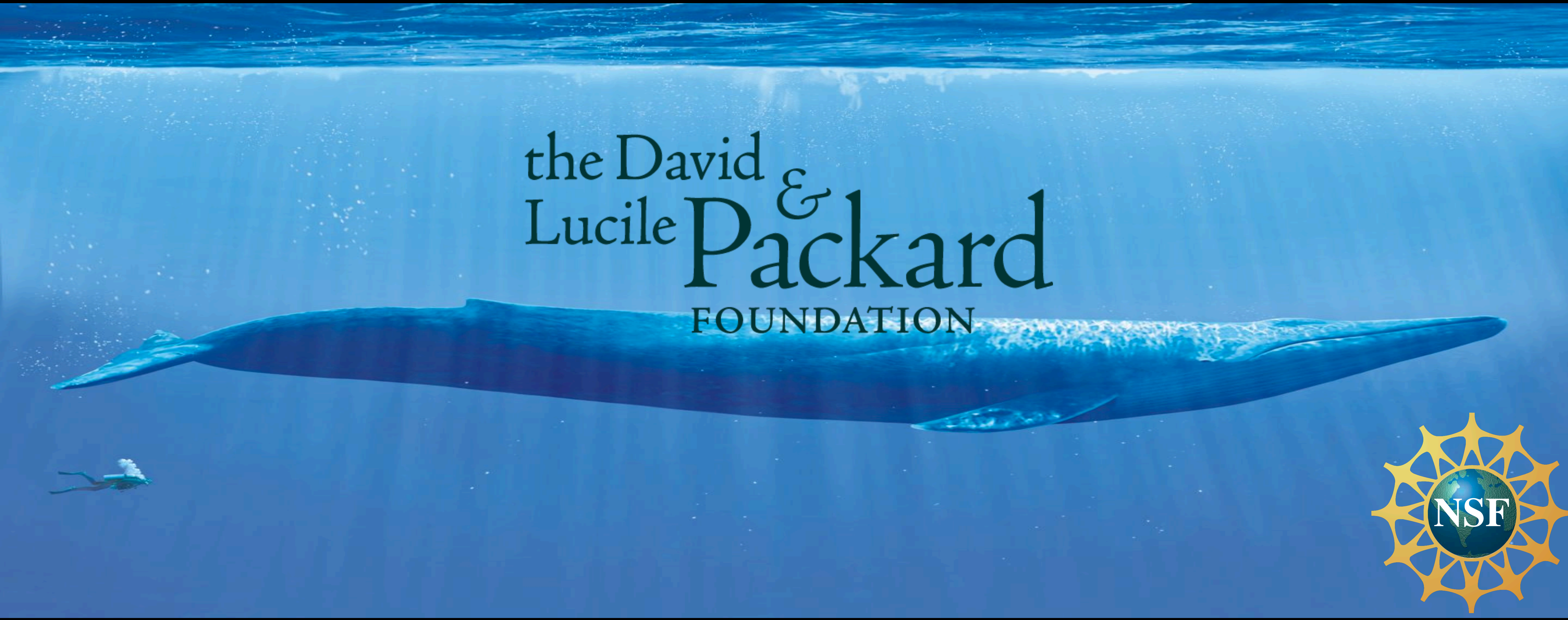
FUTURE WORK

- Refine and augment training libraries (using existing models)
- Run entire 3 year time-series
- Explore replacing BLED with ResNet-50
- Evaluate per-channel energy normalization (PCEN) as method to suppress noise

MBARI

THANK YOU

ACKNOWLEDGMENTS






John Ryan **Paul McGill** **Yanwu Zhang** **Danelle Cline**

MBARI


Daniel De Leon
Cal Poly Student

Will Oestreich
PhD Student
Stanford University
Hopkins Marine Station