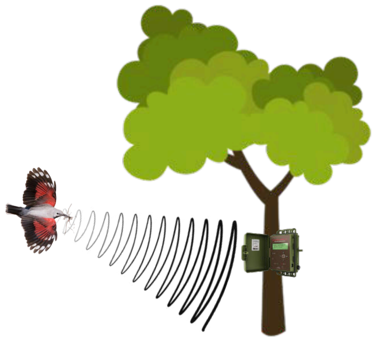
Enabling Multi-Species Bird Classification On Low-Power Bioacoustic Loggers

Stefano Ciapponi, Leonardo Mannini, Jarek Scanferla, Matteo Anderle, Elisabetta Farella

Problem



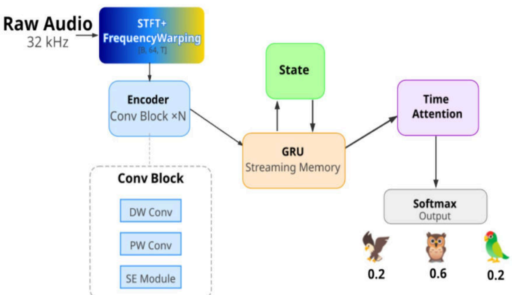
Goal:

- **Enable multi-species avian Classification** on field Microcontrollers
- **Deploy the model on a standard Passive Acoustic Monitoring device** within practical, real world constraints.

CHALLENGE: *Developing Efficient algorithms at the edge requires both **systems knowledge and domain expertise**.*

Insight into device specifications and acoustic data characteristics facilitates **careful tuning and optimization of algorithms tailored to specific applications**.

Architecture



- **Backbone:**
  - Efficient DWPW **1-d convolutional backbone** w/ squeeze-and-excitation on frequencies and a **MatchboxNet** inspired structure
- **GRU:**
  - Good compromise between standard RNN and LSTM, allows for **chunk based processing** of spectrogram in a streaming fashion
- **Time Attention:**
  - For feature **aggregation**

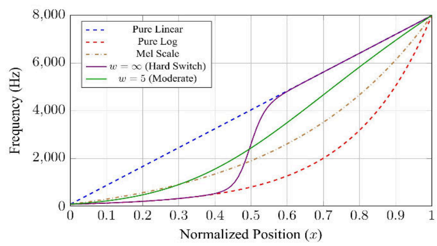
Deployment

- Tests on Audiomoth Dev (MCU 32 +256kb of RAM) and RPi 3B+ (SBC 1GB of RAM).

Highlights

Device	Energy/Inf. [J]	Time/Inf. [s]	Power [W]
Audiomoth	0.077	1.69	**0.046**
RPi 3 B+ (**ours**)	**0.172**	0.061	2.80
RPi 3 B+ (BirdNET)	2.79	0.978	2.84

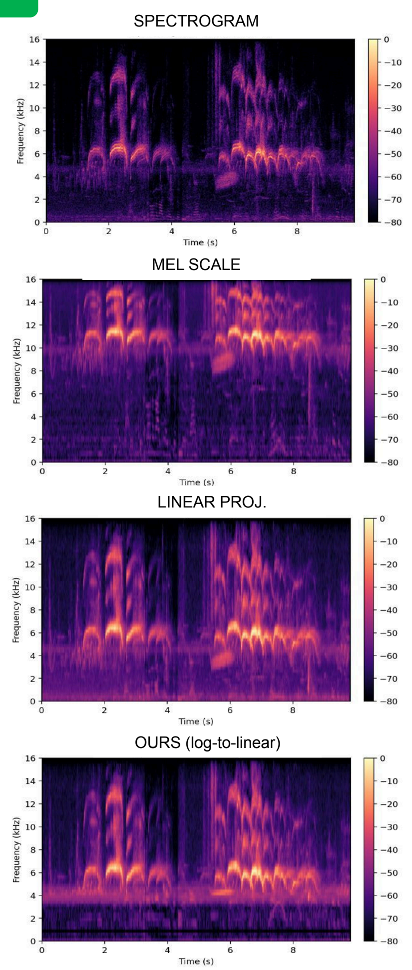
Semi-Learnable Filterbanks



Rationale

- Many approaches to bird classification use **mel-scale because of legacy Deep Learning approaches**.
- Most information lies in **higher frequencies**, so blending them together is decremental.
- We hypothesize that the right approach is **log-to-linear but we don't know where the switch happens** in the frequency domain, for this reason we decided to make it **semi-learnable**.

Filterbank Type	Learning Mode	Best Val Acc (%)	Test Acc (%)
Mel	Fixed	82.49	79.61
Linear Triangular	Fixed	82.10	81.45
Combined Log Linear	**Semi-learnable**	**85.74**	**87.22**
Fully Learnable	Fully learnable	84.26	83.83



Species Benchmarks

Configuration	#Species	Test Acc (%)	F1 (%)	Breakpoint (Hz)	Trans. Width
Monoclass (Corvus Corax)	1	**92.37**	92.62	1955	51.79
Easy species (semi-learnable)	8	90.76	90.90	1.5	6
Hard species	13	77.47	77.97	1224	27.42
Full dataset (57k params)	70	66.51	67.49	851	19
Full dataset (136k params)	70	70.14	**70.81**	1390	33.86

Future Works

A Distributed approach:

- Working on a whole network of devices spread across the Alps.
- Reframe the task as late stage fusion and confidence error correction in a Bayesian framework.
- Use ecological information between recording sites to refine predictions.