Passive Acoustic Monitoring Of Animal Populations With Compressed Sensing

Milanto Rasolofohery[1,2,3] – Emmanuel Dufourq[1,2,3]

AIMS African Institute for Mathematical Sciences SOUTH AFRICA (1)

AIMS African Institute for Mathematical Sciences RESEARCH & INNOVATION (2)

Stellenbosch UNIVERSITY IYUNIVESITHI UNIVERSITEIT (3)

1. INTRODUCTION

Passive Acoustic Monitoring (PAM) allows researchers to use microphones to record the natural world. Machine learning algorithms can be trained to detect vocalisations of animals as a means of assisting in conservation efforts. Digital recorders typically save audio files in uncompressed wave (.wav) format, leading to expensive storage requirements and difficult to transmit over a network in remote locations.

We hypothesized that the use of Compressed Sensing (CS) would enable bioacoustic data to be compressed rapidly, thereby facilitating the transmission of large files across constrained networks. This advancement aims to enable near real-time wildlife monitoring.

2. OBJECTIVES

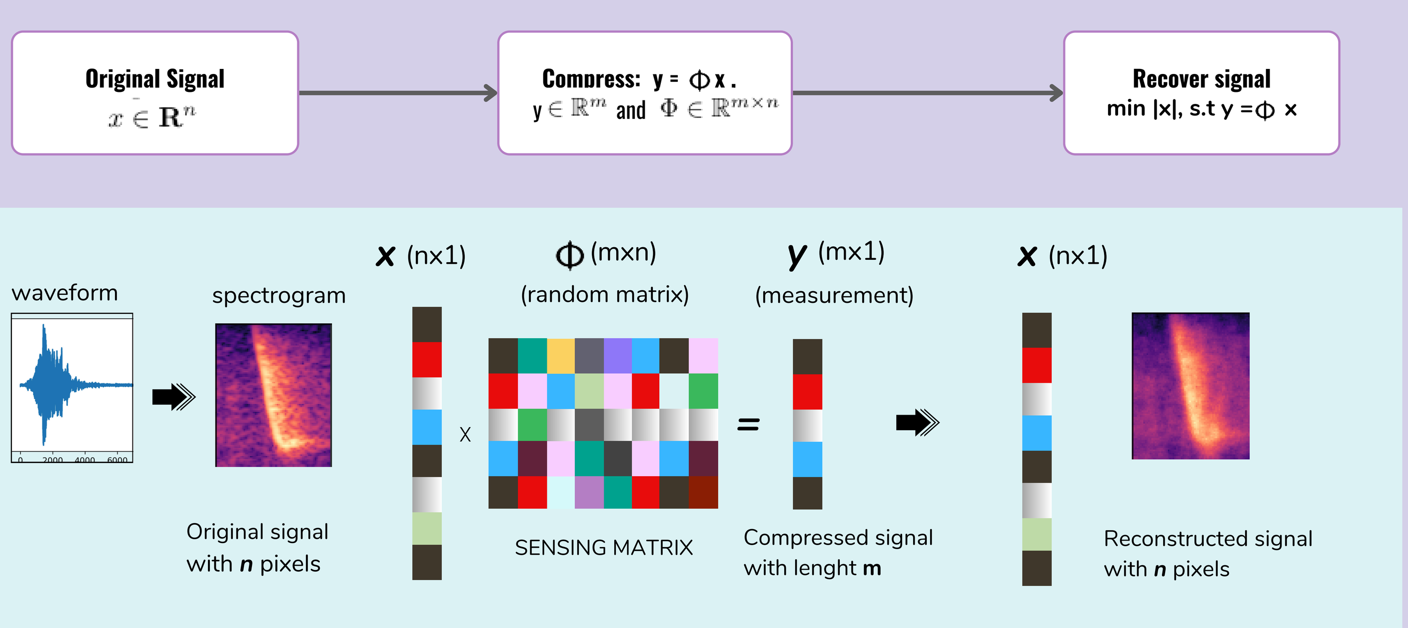
- Develop a method for detecting specific animal vocalizations in audio data to help passively monitor animal species in remote locations.
- Reduce the power consumption and data usage associated with data collection and transmission over limited networks.

3. METHODOLOGY

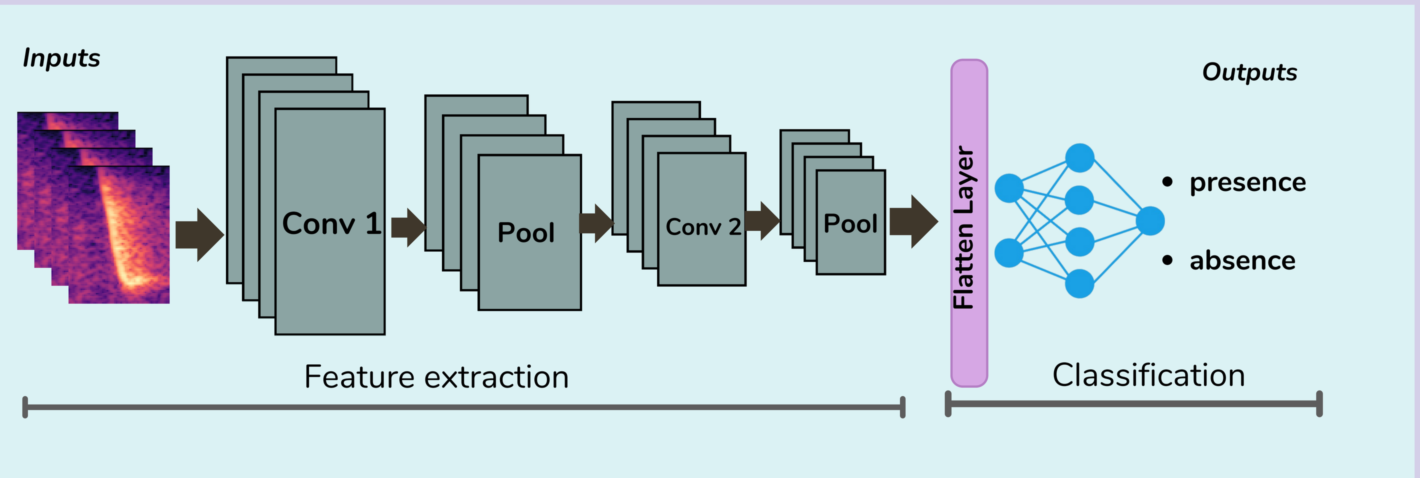
We extracted the spectrograms from the waveform, then compressed and reconstructed according to the CS process.

3.1. CS scheme for compression and reconstruction

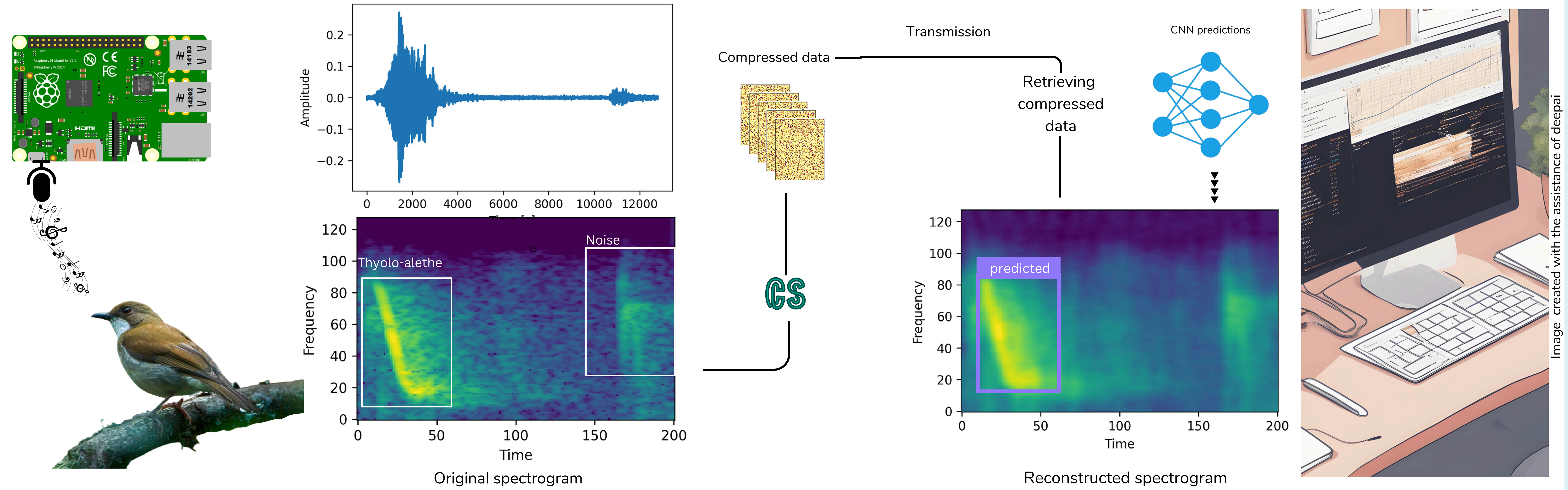
- ***x*** is an ***n***-dimensional vector representing the original signal (spectrogram).
- ***x*** is compressed into signal ***y*** using the measurement matrix Φ.
- Φ a random matrix of ***(m×n)*** dimensions, and **y** an ***m***-dimensional vector.
- ***n<<m***, where ***m*** is the number of measurement (small samples)
- Reconstruct ***x*** from ***y*** using optimization algorithms (BP,LASSO,OMP,...)



3.2. CNN for audio classification as spectrogram inputs



4. IMPLEMENTATION



Automatic data recording and compression on site

- We developed an **ARU** system made up of Raspberry Pi and microphones attached to trees to record audio data. The data is immediately compressed on the Raspberry Pi and transferred to a remote server.
- The compression is computationally inexpensive and fast, taking less than 1 second to complete.

Reconstruction and Deep Learning on Remote Computer

- We retrieve the compressed data from the remote server, and reconstruct the signals.
- We apply a convolutional neural network (CNN) to the reconstructed data and predict animal presence.

5. RESULTS

5.1. Hainan Gibbon

Table 1: Evaluation across 8 hours h of test recordings containing 7405 segments (324 gibbon calls, 7081 background noise). The original size of audio in wav format is 1914.48MB (1.87GB).

METHODS	PARAMETERS	F1-SCORE	FILE SIZE (MB)	COMPRESSION RATE
RAW	SPECTROGRAMS	92.29%	485.48	74.64%
CS	10% SAMPLES	84.69%	48.55	97.46%
	15% SAMPLES	87.15%	72.82	96.20%
	20% SAMPLES	88.30%	97.11	94.93%
JPEG	QUALITY 25	90.35%	59.3	96.90%
	QUALITY 50	91.19%	59.3	96,90%
MP3	BITRATE 16	79.77%	58.83	96.87%
	BITRATE 32	89.19%	119.66	93.75%
	BITRATE 64	82.18%	239.33	87.50%
AAC	BITRATE 16	84.17%	61.71	96.78%
	BITRATE 32	91.72%	121.58	93.65%
	BITARE 64	92.70	241.28	87.40%
OGG	QLEVEL 0	90.04%	78.58	95.90%
	QLEVEL 2	91.01%	114.80	90.00%
	QLEVEL 6	91.07%	183.17	94.43%
FLAC	LEVEL 0	91.25%	1144.11	40.42%
	LEVEL 2	90.41%	1099.88	42.55%

5.2. Thyolo Alethe

Table 2: Evaluation across 5 hours h of test recordings containing 17739 segments (2553 alethe calls, 15186 background noise). The total size of the original file is 1161.81 MB (1.13 Gb).

METHODS	PARAMETERS	F1-SCORE	FILE SIZE (MB)	COMPRESSION RATE
RAW	SPECTROGRAMS	87.70%	213.93	81.59%
CS	10% OF SAMPLES	85.20%	21.40	98.16%
	15% OF SAMPLES	85.98%	32.09	97.27%
	20% OF SAMPLES	86.66%	42.79	96.32%
JPEG	QUALITY 25	85,68%	50.9	95.61%
	QUALITY 50	86.16%	50.9	95.61%
MP3	BITRATE 32	86.37%	67.12	94.22%
	BITRATE 64	86.93%	135,25	88.44%
	BITRATE 128	87.13%	268.50	76.89%
AAC	BITRATE 16	86.75%	37.74	96.75%
	BITRATE 32	84.93%	71.35	93.86%
	BITARE 64	87.39%	138.65	88.07%
OGG	QLEVEL 0	87.93%	94.95	91.83%
	QLEVEL 2	87.55%	123.00	89.41%
	QLEVEL 6	87.74%	166.86	85.64%
FLAC	LEVEL 0	88.89%	718.17	61.29%
	LEVEL 6	87.11%	693.64	40.30%

CS can achieve high compression ratios (only keep 15% of samples) with an acceptable drop in classification performance, resulting in a data size reduction of 96.20% for gibbon and 97.27% for thyolo compared to the original waveform.

Vulnerable Thyolo alethe (Malawi)
Sample rate: 32000Hz

- **Thyolo alethe** : reduction of the data from 1.13 GB to 32.09 MB as an array while experiencing only a decrease of 1.72% points in the detection F1-score
- **Hainan gibbon:** reduction from 1.87 GB to 72.82MB and decrease of 5.15% points in F1-score.

Critically Endangered Hainan gibbon (China)
Sample rate: 8000Hz

SCAN ME

Acknowledgments

Milanto Rasolofohery is supported by the financial assistance of the UK-South Africa Newton Fund through the Development in Africa with Radio Astronomy (DARA Big Data) in collaboration with the National Research Foundation (NRF), and supported financially with the AIMS Research & Innovation Center (AIMS RIC) Rwanda. Emmanuel Dufourq was funded by a grant from the Carnegie Corporation of New York (provided through the AIMS Research and Innovation Centre). The statements made and views expressed are solely the responsibility of the authors.