2025 ASEE Northeast Section Conference, March 22, 2025, University of Bridgeport, Bridgeport, CT, USA.

Automated Call Detection in Big Cats: A Non-Invasive Bioacoustic Tool for Conservation

Habibul Huq
School Of Engineering &
Computing Fairfield University
Fairfield, USA

Douglas Muratori
School Of Engineering &
Computing Fairfield University
Fairfield, USA

Chayne Howard
School Of Engineering &
Computing Fairfield University
Fairfield, USA

Mohammed Feroz Shaik
School Of Engineering &
Computing Fairfield University
Fairfield, USA

Luke Brown
School Of Engineering &
Computing Fairfield University
Fairfield, USA

Danushka Bandara
School Of Engineering &
Computing Fairfield University
Fairfield, USA

Raqibul Haque
School Of Engineering &
Computing Fairfield University
Fairfield, USA

Soyong Byun
College of Arts & Sciences Fairfield
University
Fairfield, USA

Abstract - **Tracking reproductive cycles in endangered big cats like the Amur leopard is critical for conservation but hindered by costly, invasive methods such as hormonal analysis, which can take months to yield results. This study introduces a non-invasive, automated bioacoustic tool to detect and classify saw calls—key vocalizations linked to mating behavior—using a web-based platform accessible to non-technical users. A "saw" is a single vocalization a big cat emits that resembles the sound of sawing through wood. A "saw call" refers to a sequence(3 or more) of these sawing sounds. Two algorithms were developed to detect these calls: an energy threshold method for manual verification and a Fourier Transform-based approach for rapid, automated detection. The Fourier method achieved 90% ± 2% accuracy, identifying 239 of 242 biologist-annotated saw calls, with false positive and negative rates of 8.4% ± 1.7% and 9.5%, respectively, and a processing speed of 7 seconds per file. Unlike semi-automated bioacoustic systems, our fully automated solution eliminates technical expertise barriers, offering real-time visualizations and Excel outputs. This tool streamlines estrus monitoring, potentially reducing detection time from months to days, and lays the groundwork for scalable conservation strategies in zoos and in wild settings. Our approach advances bioacoustic monitoring by integrating user-friendly design with high-performance detection for endangered species management.**

I. Introduction

This paper aims to study and analyze vocal patterns of endangered species, particularly Amur leopards and Amur tigers, to gain a better understanding of their reproductive patterns in a digital and non-invasive approach to further conservation efforts. This method entails a web application where researchers can access different audio recordings of big cats and analyze the data and create visualizations from them. This analysis allows the understanding of patterns related to mating calls, like saw calls, which can potentially be used to enhance reproduction rates among these endangered species. Research on endangered species, especially Amur leopards and Amur tigers, is crucial in order to develop concrete solutions for conservation. Studying the reproductive systems and cycles of these species is very important and is usually one of the most difficult aspects of the research.

In recent years, bioacoustic monitoring has emerged as a promising non-invasive alternative for studying animal behavior and reproductive patterns. Vocalizations, such as the "saw call," play a crucial role in communication among big cats, particularly for territorial marking and mate attraction. Research by [1] and [2] demonstrates the feasibility of classifying big cat vocalizations based on frequency, harmonics, and behavioral contexts. These studies underscore the potential of vocalization analysis to provide insights into reproductive cycles, stress levels, and social dynamics without the need for physical interaction.

Bioacoustic monitoring is heavily used in conservation. [3] and [4] highlight the effectiveness of bioacoustic analysis in tracking endangered species. [5] demonstrated the impact of environmental conditions on calling activity in the Mountaintop Frog, showing how animal vocalization might vary under different conditions [4], [5]. Manual analysis of vocalizations is time-consuming and requires expertise, as highlighted by [4], [5]. [5] explored multiple acoustic analysis methods and concluded that a semi-automated approach combining human and machine learning techniques yields the most accurate results [4].

Recognizing these challenges, this study proposes an integrated, user-friendly platform that harnesses the power of acoustic analysis to automate big cat call detection and analysis. By focusing on the vocalizations of Amur leopards and Amur tigers, our approach seeks to streamline the process for researchers, conservationists, and zoo personnel, ultimately aiding in more effective conservation strategies.

Some of the research mentioned earlier often focuses on a specific captive animal, which raises concerns about generalizability to other captive or wild animals. While papers like [5] provide insight into individual behavioral anomalies in captive Amur leopards, they lack an acoustic analysis component [5]. This presents an opportunity to combine behavioral observations with vocalization data in future research. Furthermore, the challenges of distinguishing vocalizations from background noise, as seen in [3], are relevant, as many noise detection algorithms face issues with false positives [3]. These gaps suggest the need for refining detection algorithms and incorporating environmental variables into analysis.

By leveraging vocalization patterns to track reproductive cycles and monitor animal communication and health, this tool will play a critical role in advancing conservation strategies for endangered big cat species.

As part of this study, a tool was developed that enhances research efficiency and accelerates conservation efforts for endangered species. By utilizing vocalization patterns to track reproductive cycles and monitor behavior, the system empowers researchers to make informed decisions about conservation strategies.

The user interface enables real-time monitoring and analysis, making the system accessible to both technical and non-technical users. By streamlining the analysis process and eliminating the need for complex manual processing and analysis of audio files, as well as prolonged and expensive hormonal and fecal analysis, this system equips researchers and conservationists with a powerful tool to track reproductive cycles of endangered species and develop conservation methods. Increasing the reach of this research through a digital platform can be a major step in the conservation efforts of endangered species [6].

II. Methodology

The audio data for this project is collected from zoo enclosures like Beardsley Zoo housing big cats like Amur leopards and tigers, where 16 one-hour audio recordings are captured daily between 4 PM and 8 AM, the period when vocalization activity is most prominent.

The recording setup includes:

- Recording Devices: Song Meter SM4 recorders placed in strategic locations within enclosures to minimize background noise interference.
- Storage and Retrieval: Audio files are stored locally on Hard Drives and retrieved weekly by zoo staff, then uploaded to Google Drive for processing.
- File Organization: Each recording is labeled with date, time, and recorder ID for easy tracking and retrieval.

Two core algorithms were developed in this study:

- Energy threshold-based detection
- Fourier Transform-based automatic detection, using frequency characteristics of saw calls.

The purpose of implementing two algorithms is to benefit from both as the energy threshold method is great for higher accuracy and manual verification while Fourier transform based method is faster and automatically detects/counts number of impulses in the saw call. Ultimately, these will be combined into an ensemble approach.

Once recordings are uploaded to the cloud environment, the files are processed by the detection algorithms. The performance of these methods were evaluated using manual annotations from Biologists.

Energy threshold based analysis

Audio Segmentation: Extracting audio snippets exceeding a set energy threshold. Audio that exceeds an energy threshold of around 60 are recognized as possible saws or saw calls and are isolated and extracted as separate audio files of 3.5 to 26 seconds.

User Validation: Researchers verify detections via an interface displaying spectrograms and waveforms and hearing the audio. Detected segments are categorized into saws, Neither, or saw call once verified.

Saw call Logging: Detected saws (each sawing noise) and calls (3 or more saws), along with timestamps, categorization, and more are recorded in an Excel database for further analysis.

The final output generates a detailed excel report per detected segment and a summary of detections per file. The energy threshold based algorithm processes each 1-hour audio file in approximately 45 seconds.

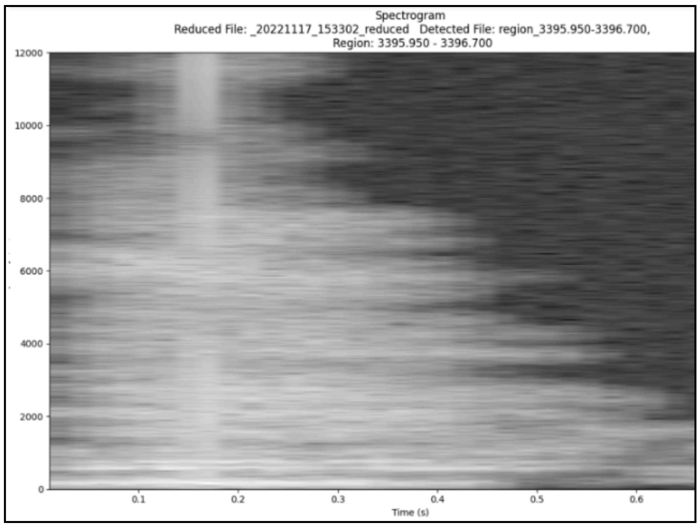

Fig. 1. Single Saw Spectrogram Visualization [1]

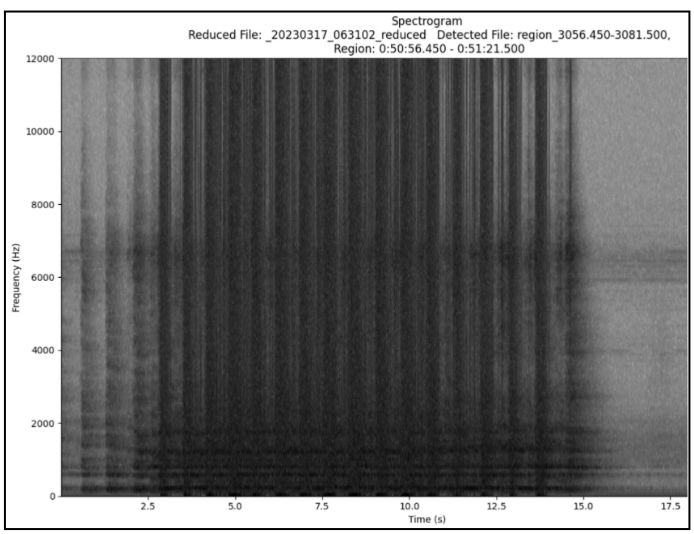

Fig. 2. Saw Call Spectrogram Visualization [2]

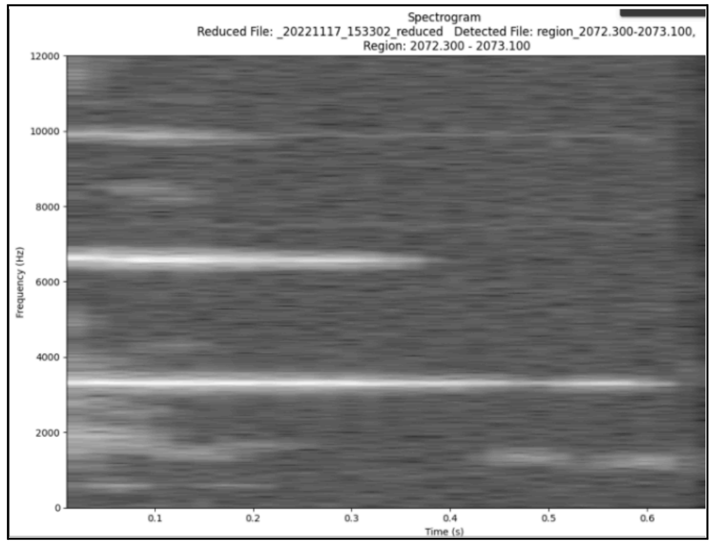

Fig. 3. Random(floor squeak) Noise Spectrogram Visualization [3]

File	Detected File	Region	Start Time	End Time	Duration	Processing Timestamp	Sound Type	Total SAWS	Total CALL
7823_20240309_060402	region_89.650-95.400	1	00:01:29	00:01:35	5.75 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_315.350-326.150	2	00:05:15	00:05:26	10.80 seconds	2024-11-20 03:37:04	CALLS	10	1
7823_20240309_060402	region_358.100-366.000	3	00:05:58	00:06:06	7.90 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_403.500-411.600	4	00:06:43	00:06:51	8.10 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_576.250-582.400	5	00:09:36	00:09:42	6.15 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_636.850-642.400	6	00:10:36	00:10:42	5.55 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_653.150-659.050	7	00:10:53	00:10:59	5.90 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_879.900-893.600	8	00:14:39	00:14:53	13.70 seconds	2024-11-20 03:37:04	CALLS	12	1
7823_20240309_060402	region_910.350-916.350	9	00:15:10	00:15:16	6.00 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_1058.650-1070.950	10	00:17:38	00:17:50	12.30 seconds	2024-11-20 03:37:04	CALLS	10	1
7823_20240309_060402	region_1117.450-1134.650	11	00:18:37	00:18:54	17.20 seconds	2024-11-20 03:37:04	CALLS	16	1
7823_20240309_060402	region_1212.050-1227.750	12	00:20:12	00:20:27	15.70 seconds	2024-11-20 03:37:04	CALLS	14	1
7823_20240309_060402	region_1918.450-1924.100	13	00:31:58	00:32:04	5.65 seconds	2024-11-20 03:37:04	Neither	0	0
7823_20240309_060402	region_2083.700-2096.300	14	00:34:43	00:34:56	12.60 seconds	2024-11-20 03:37:04	CALLS	10	1
7823_20240309_060402	region_3342.700-3348.350	15	00:55:42	00:55:48	5.65 seconds	2024-11-20 03:37:04	Neither	0	0
							Total SAWS/CALLS	72	6

Fig. 4. Excel Sheet with Detection Information [4]

Fourier transform based analysis

Fourier Transformation (FT) is a mathematical technique utilized to analyze and process waveforms by showing magnitude and frequency correlation. FT decomposes a function into frequency components. This method converts waveforms within audio recordings from the time domain, signals as a function of time, to the frequency domain, sinusoidal oscillations representing frequency and magnitude.

Our Fourier Transform implementation involves:

- Short-Time Fourier Transform (STFT) applied to audio frames.
- Spike Detection: Automatically identifying three consecutive spectral spikes as saw call indicators.
- Saw Merging: Grouping detections within a 5-second window to reduce duplication.
- Data Logging: Detections stored in an Excel database for trend analysis.

If there are more than three saws (impulses) within 5 seconds, the total audio segment is considered a call. The algorithm returns two files, one that counts all saws and saw calls, and another that prints out the time of the specific calls in the entire fileset.

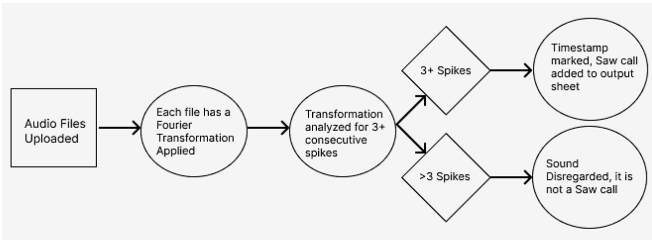

Fig. 5. Flow of Fourier Algorithm [5]

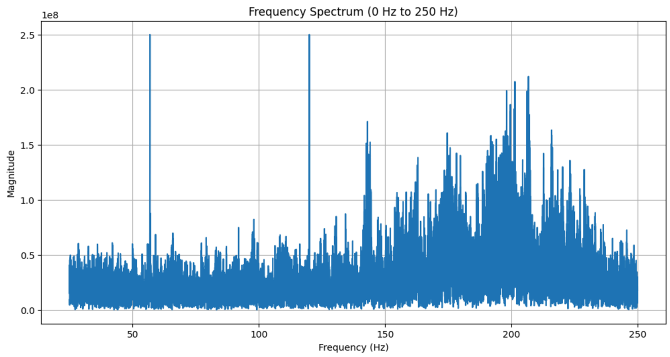

Fig. 6. Saw Call Fourier Visualization [6]

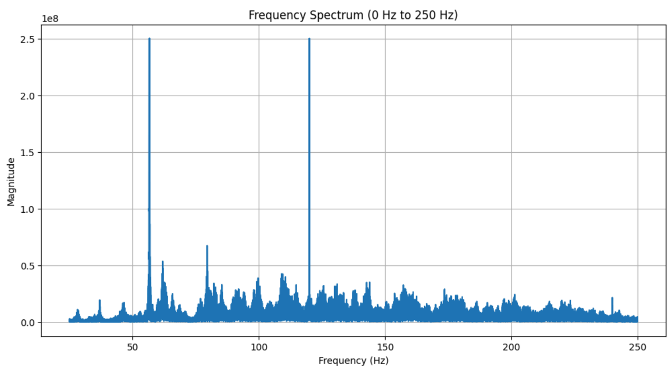

Fig. 7. Normal Vocalization Fourier Visualization [7]

Evaluation

The results were obtained for,

- Accuracy: Percentage of correctly detected saw calls.
- False Positive Rate (FPR): Percentage of incorrect detections.
- False Negative Rate (FNR): Missed saw calls.
- Processing Speed: Time required to analyze a 1-hour audio file.

User Interface Development

To enhance accessibility and usability, we're developing a web-based interface for researchers to interact with the detection system. Previously, users ran the detection algorithms manually using Google Colab, which required technical expertise and had session limitations. The new interface would allow for automated audio file uploading, processing, and visualization of results in an intuitive web environment.

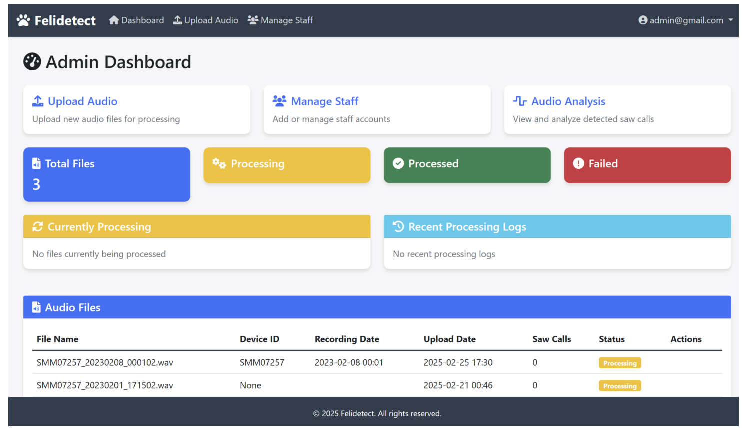

Fig. 8. Admin Page Wire Frame [8]

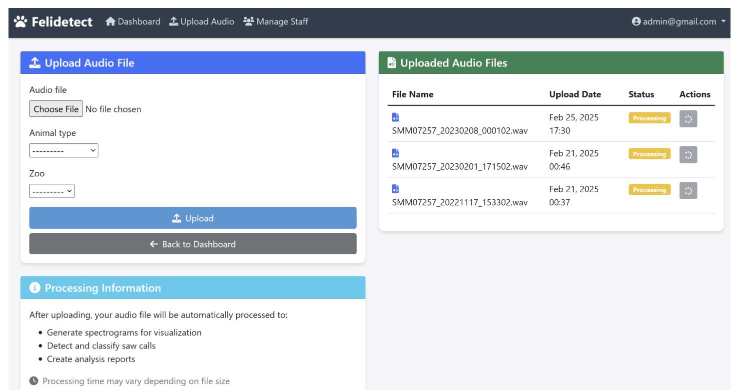

Fig. 9. Uploading Audio File Wire Frame [9]

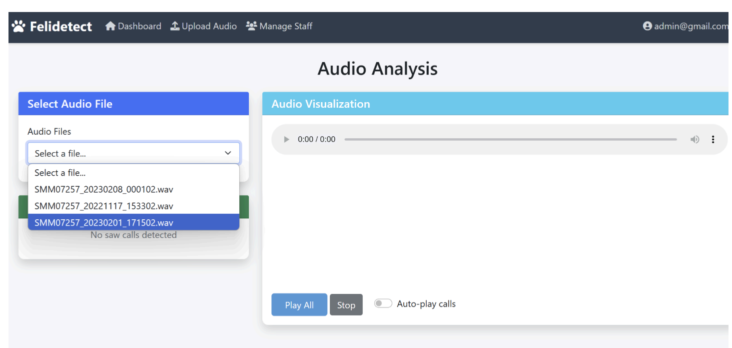

Fig. 10. Audio Analysis Page Wire Frame [10]

III. Results

With the energy threshold-based algorithm, the initial accuracy was low, prompting multiple performance testing iterations to refine the parameters and improve results. We compared our results with annotations from Fairfield University's Biology Department, which independently analyzed the same audio files.

- Performance Test #1 (Before Optimization, Using Reduced Noise File):
  - Detected Noises: 106
  - Verified Saw Calls: 7
  - Verified Saws: 56
  - Expected Saws (Bio Dept. Annotations): 77
- Performance Test #24 (After Optimization, Using Original Audio File):
  - Detected Noises: 11
  - Saw Calls: 6
  - Verified Saws: 78
  - Expected Saws (Bio Dept. Annotations): 78

Through these optimizations, the energy threshold-based algorithm achieved:

- 80%+ accuracy when using reduced noise audio files
- 90%+ accuracy when using the original audio file

In the testing of the Fourier transform-based method, 16 audio files were run through the code simultaneously. Within these files, 242 saw calls were confirmed on the annotated files by the biologists working alongside. Of these 242 daw calls, 239 were detected by the system. Of the 239 calls detected, it was discovered that 20 ± 4 false positives were identified by the model. Most files were assigned 1 false positive, though sometimes 2 or 3 were received, hence the tolerance established. This results in a false positive rate of 8.4% ± 1.7% being calculated. It also indicates that some calls were left undetected, termed false negatives. Approximately 23 false negatives were recorded, leading to a false negative rate of 9.5% being determined. This leaves the model with an overall accuracy rate of 90% ± 2% being achieved. A final average computation time of 7 seconds per file was recorded for the algorithm.

TABLE I

RESULTS OF METHODS

Method	Accuracy	False Positives	False Negatives	Processing Time
Energy Threshold	92%- 100%	45.5% - 70%	0 - 5%	45 sec/file
Fourier Transform	90+%	8.4% ± 1.7%	9.5%	7 sec/file

IV. Discussion

In zoo environments, where Amur leopards are often housed in controlled spaces, the virtual monitoring system could be an essential tool in ensuring consistent, non-invasive monitoring of estrus cycles. The virtual detection would create less stress for the animals, but could also be less effective depending on the technology level of the zoo. Additionally, the ease of data interpretation is crucial—caretakers need the system to be reliable and straightforward, avoiding any possible errors in detecting estrus.

The use of virtual monitoring systems in wild settings presents more complex challenges. Unlike in zoos, the inherent uncontrolled nature of the wild environment poses other problems, mainly with consistent and accurate data collection. The system would require greater flexibility for the unpredictability of wild leopards, which could possibly be mitigated by additional tools. The implementation in the wild could also face logistical and ethical hurdles. Deploying technology in remote areas may raise concerns about human impact on the environment or the animals' natural behaviors. Moreover, the model's reliance on specific data inputs—like motion and behavioral patterns—would require real-time updates and remote monitoring to ensure its success. Background noise from the wild environment can pose issues not discovered in controlled tasting. Despite these challenges. If adapted appropriately, this tool could be used to track not only reproductive health and improve population management, but also territorial behavior and stress monitoring. This could also lead to benefits in other conservation efforts, like anti-poaching activity.

V. Conclusion

This research demonstrates that bioacoustic monitoring can accurately track big cat reproductive cycles. Our two algorithms achieve over 90% detection accuracy, providing a scalable and non-invasive alternative to traditional monitoring. Utilizing this model, other researchers in the future will be able to detect saw calls accurately and at a low false negative rate. In addition to that, our approach is non-invasive and more efficient than existing invasive methods of studying reproductive patterns of big cats due to not needing physical contact with cats.

VI. References

[1] Rose, S. J., Allen, D., Noble, D., & Clarke, J. A. (2017). Quantitative analysis of vocalizations of captive Sumatran tigers (Panthera tigris sumatrae). Bioacoustics, 27(1), 13–26. https://doi.org/10.1080/09524622.2016.1272003

[2] X. Kong, D. Liu, A. Kathait, et al., "Behavioral-psychological motivations encoded in the vocal repertoire of captive Amur tiger (Panthera tigris altaica) cubs," BMC Zool., vol. 7, no. 2, 2022. [Online]. Available: https://doi.org/10.1186/s40850-021-00102-9.

[3] L. Bolitho, D. Newell, and H. Hines, "Bioacoustic Monitoring Reveals the Calling Activity of an Endangered Mountaintop Frog (Philoria kundagungan) in Response to Environmental Conditions," Diversity, vol. 15, no. 8, 2023. [Online]. Available: http://dx.doi.org.fairfield.idm.oclc.org/10.3390/d15080931.

[4] T. Dema, L. Zhang, M. Towsey, A. Truskinger, S. Sherub, J. Zhang, ... and P. Roe, "An investigation into acoustic analysis methods for endangered species monitoring: A case of monitoring the critically endangered White-bellied heron in Bhutan," in 2017 IEEE 13th International Conference on e-Science (e-Science), 2017, pp. 177-186. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8109135/authors#authors.

[5] Schwartz, R. (2015). The Curious Case of the Amur Leopard Sofiya. Undergraduate Journal of Service Learning & Community-Based Research, 4, 1–9. https://doi.org/10.56421/ujslcbr.v4i0.219

[6] A. V. Vitkalova, L. Feng, A. N. Rybin, B. D. Gerber, D. G. Miquelle, T. Wang, et al., "Transboundary cooperation improves endangered species monitoring and conservation actions: A case study of the global population of

Amur leopards," Conservation Letters, vol. 11, no. 5, p. e12574, 2018, doi: 10.1111/conl.12574