

Article

Identification of Elephant Rumbles in Seismic Infrasonic Signals Using Spectrogram-based Machine Learning

A. M. J. V. Costa, C. S. Pallikkonda, H. H. R. Hiroshan, G. R. U. Y. Gamlath, C. U. S. Edussooriya, and S. R. Munasinghe  0000-0003-3445-4767

Department of Electronic and Telecommunication Engineering, University of Moratuwa, Moratuwa 10400, Sri Lanka

Dept. of Global Development, CALS, Cornell University, NY 14853, USA

* Correspondence: srm278@cornell.edu

Abstract: This paper presents several machine-learning methods and highlights the most effective one for detecting elephant rumbles in infrasonic seismic signals. The design and implementation of electronic circuitry to amplify, filter, and digitize the seismic signals captured through geophones are presented. The process converts seismic rumbles to the spectrogram and the existing methods of spectrogram feature extraction and appropriate machine learning algorithms are compared on their merit of automatic seismic rumble identification. A novel method of denoising the spectrum that leads to enhanced accuracy in identifying seismic rumbles is presented. It was experimentally found that the combination of the Mel frequency cepstral coefficient (MFCC) feature extraction method and the Ridge classifier machine learning algorithm gives the highest accuracy of 97% in detecting infrasonic elephant rumbles hidden in seismic signals. The trained machine learning algorithm can run quite efficiently on general-purpose embedded hardware such as Raspberry Pi, hence the method provides a cost-effective and scalable platform to develop a tool to remotely localize elephants, which would help mitigate the human-elephant conflict.



Citation: Costa A. M. J. V.; Pallikkonda C. S.; Hiroshana H. H. R.; Gamlath G. R. U. Y.; Edussooriya C. U. S.; Munasinghe S.R. Identification of Elephant Rumbles in Seismic Infrasonic Signals Using Spectrogram-based Machine Learning. *Appl. Sci.* **202y**, *0*, 0. <https://doi.org/>

Academic Editor: Silvio Cocuzza

Received: dd mm yyyy

Accepted: dd mm yyyy

Published:

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Elephants use different vocalization patterns to communicate between herds and each other for food, water, mating, and warning. Among these patterns, the infrasonic rumbles have a very specific significance in long-range communication [6]. Elephant rumbles have acoustic and seismic components. The acoustic component propagates through the air as a three-dimensional wave, and gets attenuated through the foliage fairly quickly. The seismic component on the other hand propagates through the ground as a two-dimensional Rayleigh wave and therefore, travels a longer distance compared to the acoustic component. Elephant's foot is believed to have the capability of a seismic transponder to facilitate long-range seismic infrasonic communication. These seismic signals can be a powerful component of an elephant's communication system, serving crucial functions such as mate finding, prey detection, and interspecific and intraspecific warnings [1]. The attenuation of seismic waves during transmission increases monotonically with frequency, hence the range of 10 Hz to 40 Hz is considered the "sweet zone" for seismic signal propagation. Elephant rumbles, which have a fundamental frequency of around 20 Hz [2], fall within this sweet zone, allowing them to communicate through the ground. The elephant rumbles that propagate as Rayleigh waves through the ground can be captured using geophones [13].

Detecting and localizing wild elephants through airborne rumbles encounter certain challenges that impede its effectiveness. Airborne signals, though easy to detect, attenuate rapidly limiting the range over which elephants can be identified. The amplitude of a

Rayleigh wave during ground surface transmission is inversely proportional to the square root of the distance, with a loss of 3 dB for every doubling of the distance. In airborne signals, amplitude is inversely proportional to the distance, with a loss of 6 dB for every doubling of the distance. Factors such as weather, time of day, scattering, and reflections due to objects, and trees also attenuate the airborne signal [1]. The seismic wave undergoes a lower attenuation compared to the airborne wave, hence the detectable range of the seismic component of a rumble is greater than that of the airborne acoustic component [3]. In [22], a method is proposed to detect elephant rumbles using the shape of the formant frequency tracks of a rumble. However, the assumption in this research that the first two formant frequencies remain stationary is controversial [4]. In [24], a template matching method is used for rumble identification in which a reference feature template is used to clarify whether or not a test sample contains a rumble. This method delivers subpar performance with a 78.6% detection rate in which 78.2% is false positives as reported in [23]. In [23], seismic elephant rumbles are identified using support vector machines (SVMs) and signal enhancement, which has achieved an 88.2% detection rate with 13.7% false positives. In [5], elephant localization using an array of infrasonic seismic rumble detectors is demonstrated. Among the limited research conducted in detecting seismic elephant rumbles signal acquisition and signal processing have not yet been fully investigated. There is a clear research gap and a need to develop the right signal acquisition and processing methods to improve seismic elephant rumble detection.

This research intends to explore this gap and propose the most effective signal acquisition and signal processing methods for infrasonic seismic elephant rumble detection. After experimentally evaluating various feature extraction techniques, we observe that the three machine learning (ML) algorithms: Ridge Classifier, Decision Tree Classifier, and Light Gradient Boosting Machine (LGBM) classifier provide the best results when Mel frequency cepstral coefficients (MFCC), Hjorth parameters, and spectral energy distributions, respectively, are used as features in identifying infrasonic seismic elephant rumbles. We further find that the Ridge classifier and MFCC feature extraction method together achieves the highest accuracy of 97%, which is the best performance in elephant rumble detection to date. This outcome will help devise a tool to detect wild elephants at a distance and timely alert the villagers in the neighborhood to mitigate the human-elephant conflict [21].

2. Spectrogram of Infrasonic Seismic Signals

2.1. Overview of Signal Acquisition and Processing

The seismic elephant rumbles are in the micro-volt range. They are picked up by geophones, and amplified electronically to a level suitable for signal processing. The amplified seismic signal is then bandpass filtered to within the 5Hz to 150Hz range, which is the typical bandwidth of elephant infrasonic rumble. This filtering ensures that noise and other unwanted signals are removed, enhancing the signal quality. To achieve the desired voltage gain, the signal is guided through a variable gain amplifier in which the gain is optimally adjusted. Then, the signal is sent to an analog-to-digital converter (ADC) for digitization. As an additional precaution, a voltage clipping stage is employed to safeguard against any potential voltage spikes that could adversely impact the data integrity. The digital signal is transmitted to a Raspberry Pi single-board computer through the I2C communication protocol. In the Raspberry Pi, the spectrogram features are extracted and fed to a trained machine-learned algorithm to determine if the spectrogram contains an elephant rumble. The overview of the system is shown in Figure 1.

2.2. Electronic Circuitry

Figure 2 shows the seismic signal acquisition workflow which involves the geophone, amplifier I, bandpass filter, amplifier II, clamping and clipping circuitry, and the 16-bit ADC.

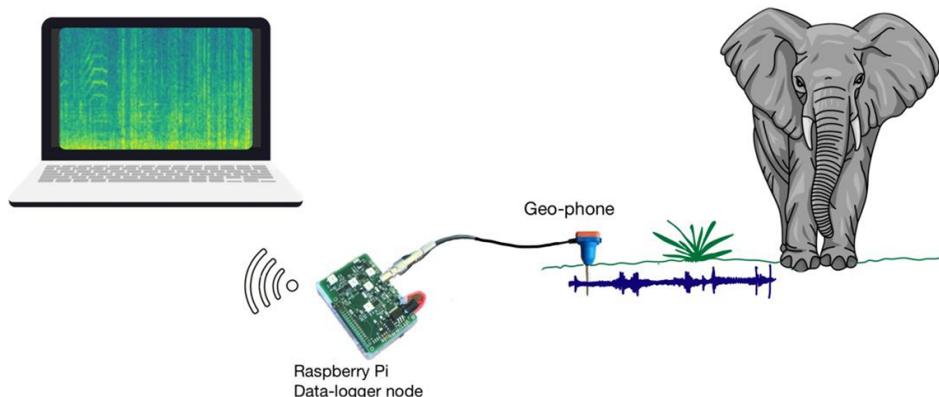


Figure 1. An overview of the system. The system captures infrasonic seismic signals using geophones. The captured signals are conditioned and transmitted to a computer for signal processing and implement machine learning algorithm to detect elephant rumbles.

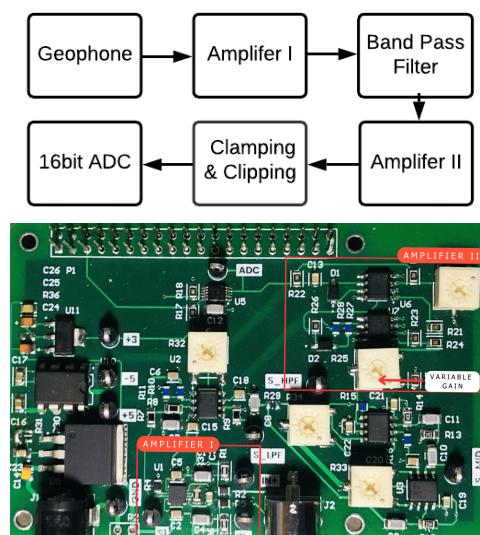


Figure 2. Seismic signal processing workflow (above), and the electronic circuit (below).

2.2.1. The Geophone

The geophone consists of a coil of wire suspended between magnets with a mass attached to it. When seismic waves pass through the ground, the mass vibrates, generating electrical signals proportional to the vertical component of the ground vibration. This research uses SG-05 geophone of which the technical specifications are listed in Table 1.

2.2.2. Amplifier I

The micro-volt signal generated by the geophone is amplified by Amplifier I, which is an instrumental amplifier. It accepts a differential input and eliminates the common mode noise significantly with a high common mode rejection ratio (CMRR). Being an integrated circuit, it has a lower instrumental noise and offset. The passive components of the circuit were carefully selected considering the tolerances, as they can affect the signal quality significantly. The amplifier has a gain of 500 in this stage.

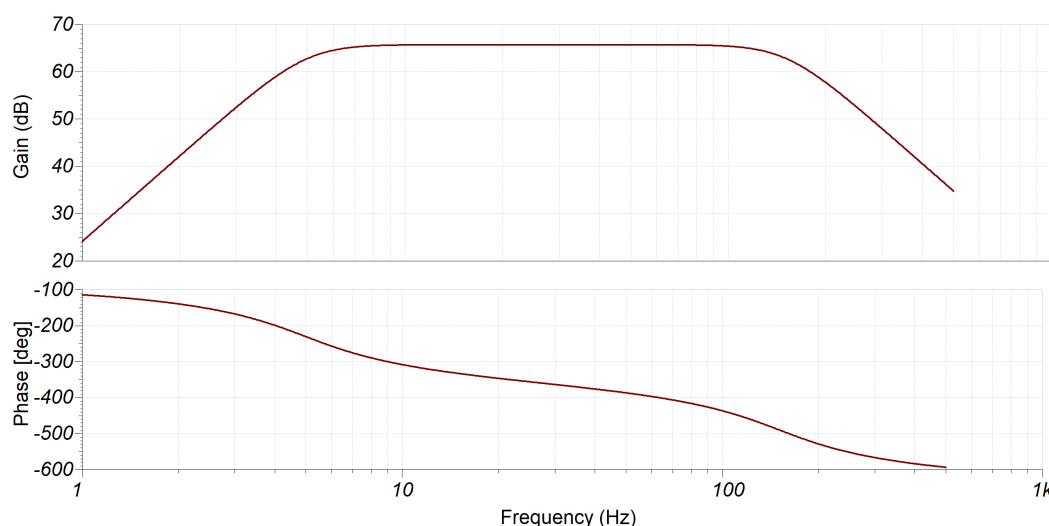
2.2.3. Band Pass Filter and Anti-Aliasing

The seismic elephant rumble spread over infrasonic frequencies, hence other frequencies the geophone picks up are to be removed. For this, a third-order Butterworth filter is used. Since the Butterworth filter has a flat passband gain, the filtering will uniformly affect

Table 1. Geophone SG-05 Technical Specifications

Property	Value
Natural frequency	5 Hz
Sensitivity	80 V/ms ⁻¹
Spurious frequency	150 Hz
Weight	22.7 g
Operating temperature	-40°C to +80°C

the amplitudes of the passband frequencies in the geophone signal. This filter further acts as an anti-aliasing filter before the ADC. The filter implementation employs a Sallen-key topology and uses a low-noise operational amplifier (OPAmp) [17]. Figure 3 shows the gain and phase response of the filter.

**Figure 3.** Frequency response of the Bandpass filter: magnitude(above) and phase(below) responses

2.2.4. Amplifier II

After the band-pass filtering, the signal is mostly infrasonic, and it may or may not contain an elephant rumble. The signal is still weak and therefore it is amplified again using the Amplifier II, which is an inverting amplifier with a variable gain that is adjusted using a variable resistor shown in Figure 2 (below). Amplifier II elevates the signal to the full voltage range of 0 V–3.3 V. The total amplification therefore is in the range of 3000–6000.

2.2.5. Clamping and Clipping Circuit

The voltage clipping circuit limits the analog signal to a safe range of 0 V to 3.3 V before feeding it into the ADC. This protective measure serves two primary purposes: limiting the analog signal range; and protecting against voltage spikes. The voltage clamp circuit adds a DC offset to the signal to make it a positive voltage before feeding it to the ADC.

2.2.6. Analog to Digital Converter

The 16-bit ADC converts the amplified band-limited analog signal to digital format. Considering the geophone signal bandwidth of 150 Hz, the sampling rate is selected as 475 samples per second so that aliasing does not happen. One elephant rumble acquired is illustrated in Figure 4. The signal spectrum of the rumble after filtering is shown in Fig.?? The digitized signal in a 10 Hz–150 Hz band, is then transferred to the Raspberry Pi single-board computer for further processing.

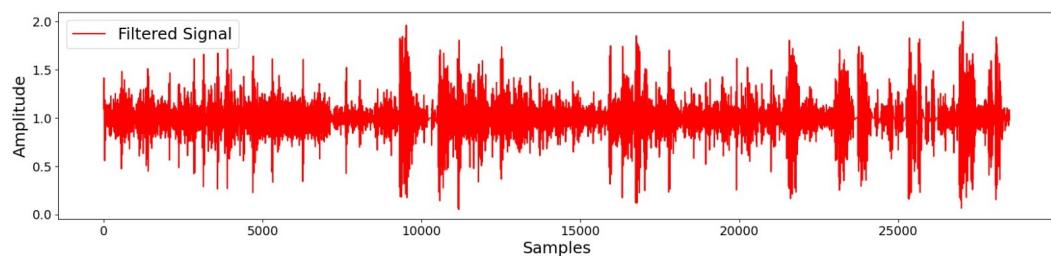


Figure 4. An elephant rumble after filtering

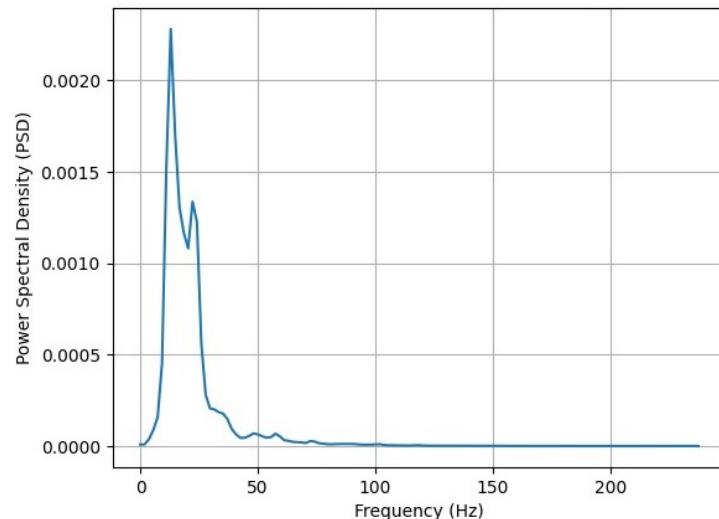


Figure 5. The spectrum of an elephant rumble

2.2.7. Sensitivity

The sensitivity of the 16-bit ADC with a range of 3.3 V can be obtained as

$$S_{ADC} = \frac{3.3 \text{ V}}{2^{16} - 1} = 50.354 \mu\text{V}. \quad (1)$$

Given the geophone sensitivity S_G and the amplification factor g , the sensitivity of the system S_s is given by

$$S_s = \frac{S_{ADC}}{g S_G}. \quad (2)$$

The sensitivity of the geophone is $S_G = 80 \text{ V/ms}^{-1}$ (see Table 1), and with $g=3000$, $S_s = 0.3147 \text{ ms}^{-1}$.

2.3. Signal Processing

2.3.1. Wavelet Decomposition and Reconstruction

The wavelet transform is used for the time-frequency decomposition of the seismic signal while also denoising the signal [20]. The wavelet transform decomposes the signal into a sum of wavelets with varying positions and scales, enabling to capture of low-frequency and high-frequency details. Here we employ Daubechies 3 wavelet for the wavelet decomposition and denoising.

2.3.2. Identification of Frequency Ridges

A ridge filter is first applied to identify the ridge structures of the spectrogram. Then, the filtered spectrogram is applied to the original spectrogram as a weighting function to enhance it further. At each spatial position (t, f) in the spectrogram $S(t, f)$, where t represents time and f denotes frequency, the partial derivatives (gradients) of the spectrogram

were calculated concerning both time and frequency. Then, the gradients are subjected to a Gaussian smoothing process. The structure tensor, which captures the local intensity distribution structure in the spectrogram's neighborhood is then constructed from the smoothed gradients as follows [4]:

$$T(t, f) = \begin{bmatrix} \nabla_t^2 & \nabla_t \cdot \nabla_f \\ \nabla_t \cdot \nabla_f & \nabla_f^2 \end{bmatrix}, \quad (3)$$

where $T(t, f)$ is the structure tensor at (t, f) , ∇_t is the partial derivative with respect to time, and ∇_f is the partial derivative with respect to frequency. From the structure tensor, we derive the eigenvalues λ_1 and λ_2 using the following expression

$$\lambda_{1,2} = \frac{1}{2} \left[\nabla_t^2 + \nabla_f^2 \pm \sqrt{(\nabla_t^2 - \nabla_f^2)^2 + 4(\nabla_t \cdot \nabla_f)^2} \right]. \quad (4)$$

When there is an edge-like structure such as a frequency contour in the local neighborhood of the spectrogram, the condition $\lambda_1 > \lambda_2$ is fulfilled. In the case of a perfect edge, λ_2 becomes zero while the condition $\lambda_1 > \lambda_2$ is still fulfilled. If $\lambda_1 = \lambda_2$, it indicates that the underlying structure is rotationally symmetric, such as an isolated spectral peak. If both eigenvalues become zero, it suggests that the underlying structure is homogeneous throughout the local neighborhood. From the eigenvalues of the structure tensor at a given position (t, f) in the spectrogram, the coherence is computed as follows:

$$c = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}. \quad (5)$$

2.3.3. Spectrogram Enhancement

To enhance the visibility of frequency contours while simultaneously reducing noise, the coherence is used as a weighting function. The enhanced spectrogram, $\hat{S}(t, f)$, is computed as shown in **Algorithm 1**.

$$\hat{S}(t, f) = S(t, f) \cdot (c(t, f) + 1). \quad (6)$$

Subsequently, as shown in **Algorithm 2** three distinct percentiles (25^{th} , 50^{th} , and 75^{th}) are computed based on the input spectrogram, and used as threshold values to partition the spectrogram's pixel intensity values into three categories: low, medium, and high. The algorithm iterates through each pixel within the spectrogram and applies adjustments contingent on the intensity of the pixel. Pixels with values surpassing the 75^{th} percentile undergo enhancement by adding 5, thus accentuating the regions with the highest intensity. Pixels falling within the range of the 50^{th} and 75^{th} percentiles experience a moderate enhancement of 2, ensuring that mid-level features remain discernible. Pixels below the 50^{th} percentile have their intensity reduced by subtracting 2, consequently amplifying the contrast between the background and the features of interest. Finally, for pixels with intensities below 25^{th} percentile, a reduction of 5 is applied, maintaining a clear distinction between the lowest-intensity areas and other regions of the spectrogram. Additionally, following the enhancement process, a Gaussian blur operation is applied. This operation serves to further refine the spectrogram, smoothing out noise and improving its overall quality for more precise feature extraction and interpretation.

2.3.4. Assessment of Spectrogram Enhancement

To assess the effectiveness of the spectrogram enhancement algorithm, the structural similarity index (SSIM), which is a widely employed metric in image analysis and quality assessment [10], is used. It serves as a valuable tool for comparing two images by evaluating their structural similarity, taking into account luminance, contrast, and structure. SSIM quantifies how well the structural elements of an image, such as edges, textures, and patterns, are preserved when compared to a reference image. It produces a value between

Algorithm 1: Enhanced Spectrogram Calculation

Data: s - Input spectrogram;

Result: Enhanced spectrogram;

Apply ridge filtering ($Ridge_s$) for s ;

Calculate time gradient ($time_gradient$) along axis 0;

Calculate frequency gradient ($frequency_gradient$) along axis 1;

Calculate tensor elements:

$$tensor_{xx} = time_gradient^2$$

$$tensor_{yy} = frequency_gradient^2$$

$$tensor_{xy} = time_gradient * frequency_gradient$$

Apply Gaussian filtering to $tensor_{xx}$, $tensor_{yy}$, and $tensor_{xy}$ to obtain smoothed versions;

Calculate eigenvalues λ_1 and λ_2 ;

Calculate coherence (c) as:

$$c = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}$$

Calculate enhanced spectrogram ($enhanced_spectrogram$) as:

$$enhanced_spectrogram = (10 \cdot \log_{10}(Ridge_s)) \cdot (1 + c)$$

return $enhanced_spectrogram$;

-1 and 1, where a score of 1 indicates a perfect similarity, meaning the two images are identical, while lower scores signify increasing dissimilarity. The SSIM index is calculated as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (7)$$

where μ_x is mean of x , μ_y is mean of y , σ_x is standard deviation of x , σ_y is standard deviation of y , and σ_{xy} is covariance between x and y . The constance c_1 and c_2 are defined by $c_i = (Lk_i)^2; i = 1, 2$, where L is the dynamic range of pixel values, $k_1 = 0.01$ and $k_2 = 0.03$ are default constants.

3. Spectral Features for Seismic Elephant Rumble Detection

The three feature sets: MFCC, Hjorth parameters, and spectral energy distribution, that represent the spectral information of seismic data, are tested in order to determine the effectiveness of extracting spectral features of seismic elephant rumbles.

3.1. Mel-frequency Cepstral Coefficients

The MFCCs of a signal provide a characteristic representation of the signal [15]. When extracting MFCCs, the signal is divided into overlapping frames of fixed 25 ms. The frames are shifted by 50% to create an overlap to ensure continuity. On each frame, a Hamming window is applied to taper the signal towards the frame boundaries. This is done to enhance the harmonics, smooth the edges, and reduce the edge effect while taking the discrete Fourier transform (DFT) [14] of the signal. Each windowed frame is converted into a magnitude spectrum by applying DFT as follows:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j\frac{2\pi}{N}nk}, \quad k = 0, 1, \dots, N-1, \quad (8)$$

Algorithm 2: Threshold-Based Spectrogram Enhancement

Data: $s1$ - Input spectrogram;
Result: $final_spectrogram$ - Further Enhanced spectrogram;
Calculate Threshold values from $s1$ - $threshold1, threshold2, threshold3$;
for i in range($s1.shape[0]$) **do**
 for j in range($s1.shape[1]$) **do**
 if $s1[i, j] > threshold3$ **then**
 $final_spectrogram[i, j] = s1[i, j] + 5$;
 else if $s1[i, j] > threshold2$ **then**
 $final_spectrogram[i, j] = s1[i, j] + 2$;
 else if $s1[i, j] > threshold1$ **then**
 $final_spectrogram[i, j] = s1[i, j] - 2$;
 else
 $final_spectrogram[i, j] = s1[i, j] - 5$;
 return $final_spectrogram$;

where $X(k)$ is DFT of the input sequence $x(n)$, N is number of samples in the sequence, and k is frequency index. The Mel spectrum is obtained by applying a Mel-filter bank to the Fourier-transformed signal. The approximation of the Mel frequency scale from the actual frequency is expressed as follows.

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right), \quad (9)$$

where m is Mel-frequency scale of the actual frequency f in Hz. The Mel spectrum of the magnitude spectrum $X(k)$ is computed by multiplying the magnitude spectrum by each of the triangular Mel weighting filter as follows.

$$s(m) = \sum_{k=0}^{N-1} |X(k)|^2 H_m(k); \quad 0 \leq m \leq M-1, \quad (10)$$

where $s(m)$ is filter bank output at index m , $X(k)$ is the DFT of the signal, $H_m(k)$ is the Mel filter bank coefficient at k and m , N is the number of DFT points, and M is the number of Mel frequency bins. The triangular Mel weighting filters are defined as

$$H_m(k) = \begin{cases} 0 & \text{if } k < f(m-1) \\ \frac{2(k-f(m-1))}{f(m)-f(m-1)} & \text{if } f(m-1) \leq k \leq f(m) \\ \frac{2(f(m+1)-k)}{f(m+1)-f(m)} & \text{if } f(m) < k \leq f(m+1) \\ 0 & \text{if } k > f(m+1) \end{cases}, \quad (11)$$

where $H_m(k)$ is the Mel filter bank coefficient at k and m , m is the Mel index, and $f(m)$ is the Mel frequency corresponding to index m .

The discrete cosine transform is applied to the log-compressed filterbank coefficients. Since most of the signal information is represented by the first few MFCC coefficients and the dataset is small, only the first few coefficients are extracted. Finally, the MFCC is calculated as

$$c(n) = \sum_{m=0}^{M-1} \log_{10}(s(m)) \cos \left(\frac{\pi n(m-0.5)}{M} \right), \quad (12)$$

where $c(n)$ is cepstral coefficient at index n , $s(m)$ is filterbank output at index m , M is the number of Mel frequency bins, and C is the number of cepstral coefficients.

3.2. Hjorth Parameters

Hjorth parameters form a set of statistical measures used to characterize the time-domain properties of a signal [9]. From those parameters we have used Hjorth activity (H_a), Hjorth mobility (H_m), and Hjorth complexity (H_c) which are defined as

$$H_a = \frac{1}{N-1} \sum_{n=1}^N [x(n) - \mu]^2 \quad (13)$$

$$H_m = \sqrt{\frac{\text{Var}(x[n] - x[n-1])}{\text{Var}(x[n])}} \quad (14)$$

$$H_c = \frac{H_m(x[n] - x[n-1])}{H_m(x[n])}, \quad (15)$$

where μ is the mean of $x(n)$, and $\text{Var}()$ is the variance operator.

3.3. Spectral Energy Distribution

To compute the spectral energy distribution, the spectrogram is split into frequency bands. In this research, twenty-five frequency bands are used. Let's denote the frequency band by $B_k; k = 1, 2, \dots, 25$. The energy within each frequency band E_k is calculated by summing the squared magnitudes of the spectrogram's elements within B_k as follows:

$$E_k = \sum_{i \in B_k} \sum_{j=0}^{N-1} |\text{spectrogram}[i, j]|^2, \quad (16)$$

where i is index of frequency bin within k , and j is index of time frame $(0, N-1)$.

4. Machine Learning Algorithms for Seismic Elephant Rumble Detection

A number of ML algorithms together with three spectral feature sets described in Section 3 are tested for seismic elephant rumble detection. To this end, we employ the open-source *LazyPredict* machine learning platform using sci-kit learn 1.3.2 [19] Python library. This ML platform allows the training and testing of various ML algorithms quite efficiently. The MFCC feature set performed well with the following machine learning algorithms:

- Ridge Classifier
- Support vector machine (SVM) - Linear
- Logistic Regression.

Furthermore, the Hjorth parameters performed well with the following machine learning algorithms:

- Decision tree classifier
- AdaBoost classifier
- Random Forest classifier,

and the spectral energy distribution performed well with the following machine learning algorithms:

- Light gradient boosting machine classifier (LGBM)
- Gradient boosting classifier
- AdaBoost classifier.

In our study, we use a dataset consisting of 7 rumbles and 60 non-rumble instances collected at Pinnawala elephant orphanage. Given the limited size of our dataset, we employ cross-validation to evaluate the performance of ML models, ensuring a more robust assessment of model generalization. As shown in Figure 6 the *LazyPredict* receives the feature set in MFCC, Hjorth, or spectral energy distribution format, and executes corresponding ML algorithms. For each ML algorithm and feature format combination, the performance is recorded in terms of accuracy, balance accuracy, and F1 score. It is

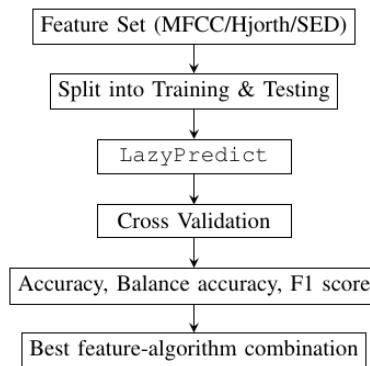


Figure 6. Determination of the best feature-algorithm combination for machine learning using `LazyPredict` platform. Here, SED means spectral energy distribution.



Figure 7. Free ranging area in Pinnawala elephant orphanage. The observation hall is the marked rectangle closer to the right-bottom corner

worthwhile to note that, here, we present the best performing ML algorithms for a given set of features, and the use of a set of features, e.g., MFCC, is not limited to the presented ML algorithms only. Furthermore, we may employ an early fusion of features (before the classification) or a late fusion of classification results (using an ensemble learning approach with a selected set of ML algorithms) to improve the classification performance. We intend to incorporate these approaches in future work.

5. Experimental Data Collection

The seismic signal collection was conducted at the elephant orphanage in Pinnawala, Sri Lanka (Figure 7), where there are about twenty-five elephants in a free-ranging area. The elephants' behavior was closely observed before the experiment. It was clear that in hot weather, elephants tend to be inactive and spend most of the day under trees. However, in normal weather, they become active and engaged. Seismic data was collected under average weather conditions using three equispaced geophones buried near the fence that is marked with straight lines in Figure 7. Elephant behavior was video recorded during the few-hour data capture. At the time the food truck arrived, one elephant rumbled quite noticeably as shown in Figure 8 where it shows that the elephant stands literally on two legs exerting maximum pressure on the ground that would enhance the seismic rumble propagate through the ground. The seismic data that is aligned with audio and video records of the rumble was extracted from the recorded seismic signal. The spectrogram clearly shows the frequency patterns of strong elephant rumbles as shown at the 10 s mark of the top two illustrations of Figure 9.



Figure 8. An elephant while rumbling - Pinnawala elephant orphange, Sri Lanka.

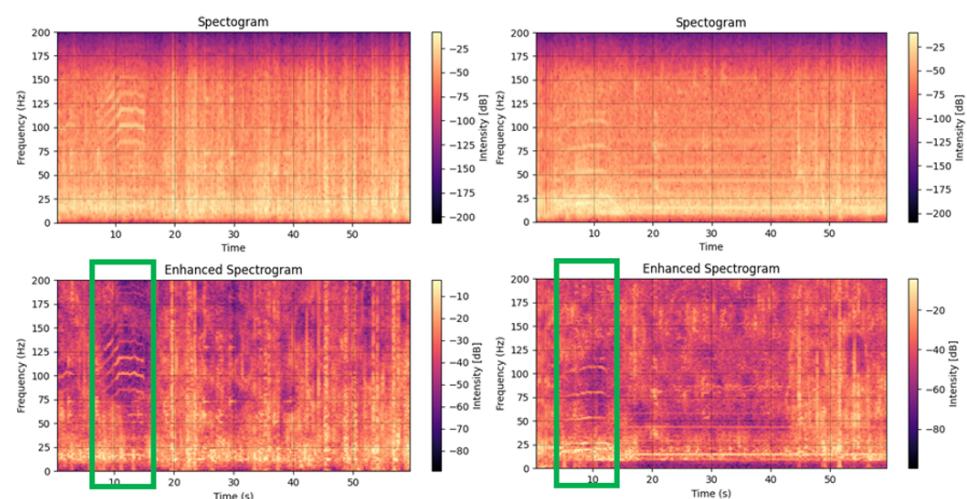


Figure 9. Spectrograms of two relatively strong rumbles verified by video-audio proof, before (top) and after (bottom) the enhancement using the proposed algorithm.

6. Results

The performance of each ML algorithm along with the relevant feature sets are compared using accuracy, balanced accuracy (BA) and F1 Score defines as follows:

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{TP + FP + TN + FN}; \\ BA &= \frac{\text{Sensitivity} + \text{Specificity}}{2}; \\ F1 \text{ Score} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}; \end{aligned}$$

in which

$$\begin{aligned} \text{Sensitivity} &= \frac{TP}{TP + FN} \\ \text{Specificity} &= \frac{TN}{TN + FP} \\ \text{Precision} &= \frac{TP}{TP + FP} \\ \text{Recall} &= \frac{TP}{TP + FN}, \end{aligned}$$

MFCC Feature	P-Value
mfcc_2	4.198520e-11
mfcc_12	7.750290e-06
mfcc_22	3.081888e-04
mfcc_32	4.038959e-04
mfcc_7	6.701393e-04
mfcc_11	1.539645e-03
mfcc_30	1.756065e-03
mfcc_40	2.829550e-03
mfcc_21	2.907236e-03
mfcc_19	3.459980e-03
mfcc_29	1.106669e-02
mfcc_13	1.471919e-02
mfcc_9	1.511950e-02
mfcc_25	1.820666e-02
mfcc_38	3.290245e-02
mfcc_36	3.836483e-02
mfcc_31	4.856187e-02

Table 2. Selected MFCC features based on P values

Feature	P-Value
Activity	0.001120
Complexity	0.043674
Mobility	0.044281

Table 3. Selected Hjorth features based on P values

Feature	P-Value
SpecraEnergy_14	0.007169
SpecraEnergy_15	0.009102
SpecraEnergy_13	0.038727
SpecraEnergy_16	0.039069

Table 4. Selected Spectral Energy features based on P values

where TP is number of true positives, FP is number of false positives, TN is number of true negatives, and FN is number of false negatives.

6.1. Performance Comparison

In this study, feature selection was based on p-values. Features with p-values below the threshold of 0.05 were considered significant to improve the model's interpretability and reduce the risk of overfitting. Table 2, Table 3, and Table 4 show the selected features from MFCC features, Hjorth features, and Spectral energy features respectively.

Table 5 shows the rumble identification performance when the MFCC feature extraction method was used with the ridge classifier, SVM-linear, and Logistic regression as machine learning algorithms.

Table 6 shows the rumble identification performance when Hjorth parameters were used as features and the Decision tree classifier, AdaBoost classifier, and RandomForest classifier were used as machine learning algorithms. Table 7 shows the rumble identification performance when spectral energy distribution was used as features and LGBM classifier, Gradient boosting classifier, and AdaBoost classifier were used as machine learning algorithms.

ML Algorithm	Accuracy	BA	F1 score
Ridge Classifier	$97.05 \pm 0.051\%$	95.83%	95.65%
SVM-linear	$94.04 \pm 0.067\%$	90%	88.89%
Logistic Regression	$93.85 \pm 0.069\%$	84.16%	77.78%

Table 5. Performance when MFCCs were used as features

ML Algorithm	Accuracy	BA	F1 score
Decision Tree Classifier	$93.84 \pm 0.096\%$	83.33%	73.68%
AdaBoost Classifier	$92.26 \pm 0.097\%$	78.33%	74.04%
RandomForest Classifier	$89.68 \pm 0.056\%$	70%	57.14%

Table 6. Hjorth parameters as features

Spectrogram	SSIM with the initial spectrum [4]	SSIM with the enhanced spectrum
Fig. 9(left)	0.86	0.65
Fig. 9(right)	0.89	0.66

Table 8. Denoising effect in the spectrogram enhancement process, measured in SSIM index.

ML Algorithm	Accuracy	BA	F1 score
LGBM Classifier	$86.67 \pm 0.119\%$	64.17%	42.86%
Gradient Boosting Classifier	$81.95 \pm 0.095\%$	63.33%	40%
AdaBoost Classifier	$80.24 \pm 0.136\%$	58.33%	28.57%

Table 7. Spectral energy distribution as features

Based on the results in Table 5 and Table 6, MFCC with the Ridge classifier is the best combination for identifying elephant rumbles in seismic signals.

The performance of the best model for each feature extraction method was assessed using ROC curves, shown in Figures 10 to 12. The Ridge Classifier performed exceptionally well with an AUC of 0.99, indicating near-perfect classification. The Decision Tree Classifier achieved an AUC of 0.88, while the LGBM Classifier reached a respectable 0.74. Overall, the combination of MFCC features and the Ridge Classifier is the most effective for the classification task.

6.2. Effectiveness of Spectrogram Enhancement

One of the main contributions of this work is spectrogram enhancement presented in 2.3.3. It is evident qualitatively from Figure 9 that the proposed approach enhance the features of the spectrogram. In order to compare the performance quantitatively, we employed the SSIM index. Here, we considered the spectrogram generated just after the ADC as the ground truth before any digital-domain processing and denoising. The achieved SSIM indices for the spectrograms enhanced using the proposed algorithm and that in [4] are presented in Table 8. We note that higher SSIM means a more similarity for the original noisy spectrogram and vice versa. Accordingly, the low SSIM index achieved with the proposed algorithm indicate that proposed algorithm provides a better enhancement compared to the algorithm proposed in [4].

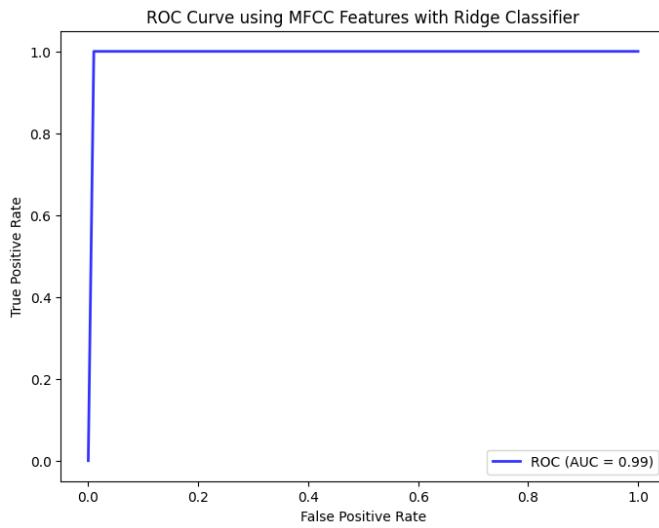


Figure 10. ROC Curve using MFCC features with Ridge Classifier.

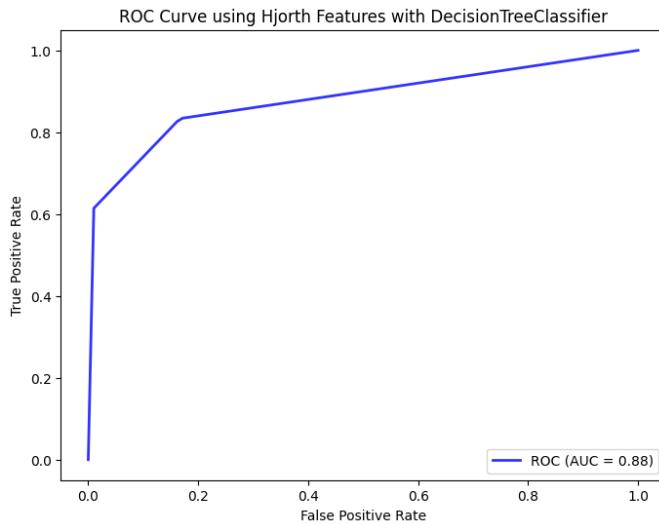


Figure 11. ROC Curve using Hjorth features with Decision Tree Classifier.

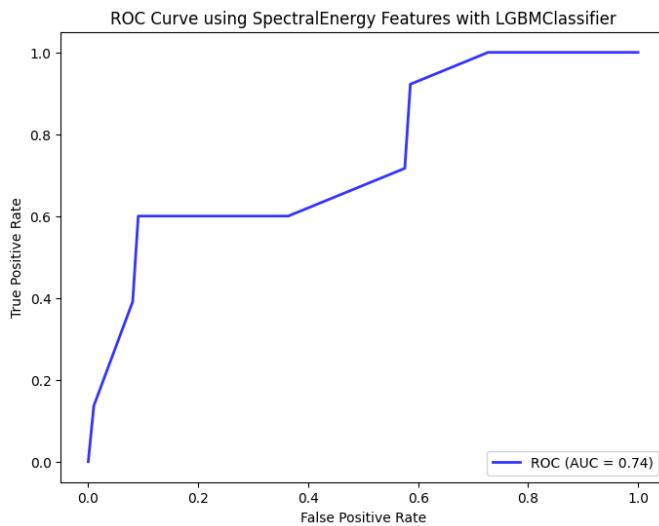


Figure 12. ROC Curve using Spectral Energy features with LGBM Classifier.

7. Conclusion

The electronic circuitry and signal conditioning techniques have been developed quite successfully leading to acquiring seismic elephant rumbles in good quality. A few elephant rumbles were visually and audibly noticed during the data collection. Their seismic infrasonic copies were extracted from the geophone data, and the spectrograms were determined. The spectrogram enhancement has reduced noise in the seismic data quite significantly leading to achieving higher accuracy in feature detection and machine learning. Using `lazy predict`, it was found that the MFCC feature extraction method and Ridge classifier ML algorithm give the best performance in automatic elephant rumble detection.

The proposed elephant rumble identification method could be used to devise an early warning system for the villages troubled by wild elephants. The circuits involved are low-power and low-cost. It could be made even more cost-effective and less power-consuming by replacing the Raspberry Pi board with a Raspberry Pi zero board, which extends the battery life in field deployment. For the practical deployment, the best ML algorithm, Ridge classifier can be implemented on an edge device using `scikit-learn` library in Python. It is important to mention that elephants do not rumble often, hence rumble detection will not lead to the development of a perfect early warning system. The elephant localization using seismic rumbles is yet to be investigated, though some early work has been reported.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Not available yet

The video of the traditional and bird takeoff can be found at, <https://www.youtube.com/watch>

Acknowledgment

The authors would like to thank Dr. Daniela Hedwig, Director of the Elephant Listening Project, Dr. Geoffrey Abers, Chair of the Earth and Atmospheric Sciences at Cornell University, USA, and Dr. Chase A. LaDue, at Oklahoma City Zoo and Botanical Garden, USA for their insightful guidance.

Dr. Lori Lenard, Dr. Ronnie Coffman, and Dr. Richard Cahoon are acknowledged for facilitation and guidance provided for this research during the Fulbright Fellowship of the principal investigator Dr. Rohan Munasinghe at the Dept of Global Development, CALS, Cornell University, Ithaca, NY 14853, USA.

The authors would also acknowledge the staff of the National Zoological Department, and the Elephant Orphanage, Pinnawala, Sri Lanka for their approval and invaluable support during data collection.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. O'Connell-Rodwell, C. E.; Arnason, B. T; and Hart, L. A.; Seismic properties of Asian elephant (*Elephas maximus*) vocalizations and locomotion. *The Journal of the Acoustical Society of America*, vol. 108, no. 6, pp. 3066-3072, Acoustical Society of America, 2000.
2. Payne, K. B; Langbauer, W. R; and Thomas, E. M; Infrasonic Calls of the Asian Elephant (*Elephas Maximus*). *The Journal of Behavioral Ecology and Sociobiology*, vol. 18, no. 4, pp. 297-301, Springer, 1986.
3. Sayakkara, A. P.; Jayasuriya, N.; Ranathunga, T.; Suduwella, C.; Vithanage, N.; Keppitiyagama, C.; De Zoysa, K.; Hewage, K.; and Voigt, T.; Eloc: locating wild elephants using low-cost infrasonic detectors. *13th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 44-52, 2017.
4. Zeppelzauer, M.; Hensman, S.; and Stoeger, A. S.; Towards an automated acoustic detection system for free-ranging elephants. *Bioacoustics*, vol. 24, no. 1, pp. 13-29, Taylor & Francis, 2015.
5. Reinwald, M.; Moseley, B.; Szenicer, A.; Nissen-Meyer, T.; Oduor, S.; Vollrath1, F.; Markham A.; and Mortimer, B.; Seismic localization of elephant rumbles as a monitoring approach. *Interface*, vol. 18, pp. 1-11, The Royal Society, 2021.

6. Nair, S.; Balakrishnan, R.; Seelamantula, C. S.; and Sukumar, R.; Vocalizations of wild Asian elephants (*Elephas maximus*): structural classification and social context. *The Journal of the Acoustical Society of America*, vol. 126, no. 5, pp. 2768-2778, 2009.
7. Phinyomark, A.; Phukpattaranont, P.; and Limsakul, C.; The usefulness of wavelet transform to reduce noise in the SEMG signal. *EMG methods for evaluating muscle and nerve function*, pp. 107-132, nTech Rijeka, Croatia, 2012.
8. Daubechies, I.; Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909-996, Wiley, 1988.
9. Wannawijit, I.; and Kaiwansil, S.; Ruthaisujaritkul, S.; and Yingthawornsuk, T.; ECG classification with modification of higher-order Hjorth descriptors. In Proc. 2019 15th IEEE International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), pp. 564-571, 2019.
10. Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P.; Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600-612, 2004.
11. Clemins, P. J.; Johnson, M. T.; Generalized perceptual linear prediction features for animal vocalization analysis. *The Journal of the Acoustical Society of America*, vol. 120, no. 1, pp. 527-534, AIP Publishing, 2006.
12. Kim, Y-E.; Su, D-H.; Jeon, C-H.; Lee, J-K.; Cho, K-J.; Chung, J-G.; Sound Source Localization Method Using Region Selection. *Advances in Sound Localization*, IntecOpen, 2011.
13. Günther, R. H.; O'Connell-Rodwell, C. E.; Caitlin, E.; and Klemperer. S. L.; Seismic waves from elephant vocalizations: A possible communication mode?. *Geophysical Research Letters*, vol. 31, no. 11, Wiley Online Library, 2004.
14. Oppenheim, A. V.; Discrete-time signal processing. Ed.3 Pearson Higher Education Inc., Upper Saddle River, NJ, 2010.
15. Jurafsky, D.; and Martin, J. H.; Speech and Language Processing, Ed. 3 Online available, 2003.
16. Keen, S. C.; Shiu, Y.; Wrege, P. H.; and Rowland, E. D.; Automated detection of low-frequency rumbles of forest elephants: A critical tool for their conservation *The Journal of the Acoustical Society of America*, vol. 141, no. 4, pp. 2715-2726, AIP Publishing, 2017.
17. Zumbahlen, H.; Phase Relations in Active Filters Analog Dialogue, vol. 41, no. 4, 2007.
18. Linghu, J.; and Dong, H.; and Cui, J.; Ensemble wavelet-learning approach for predicting the effective mechanical properties of concrete composite materials *Computational Mechanics*, vol. 70, no. 2, pp. 335–365, 2022.
19. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; and Duchesnay, E.; Ensemble wavelet-learning approach for predicting the effective mechanical properties of concrete composite materials *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
20. Donoho, D. L.; De-noising by soft-thresholding *IEEE transactions on information theory*, vol. 41, no. 3, pp. 613-627, 1995.
21. Prakash, T. G.; Supun L.; Wijeratne, A. W.; and Fernando, P.; Human-Elephant Conflict in Sri Lanka: Patterns and Extent Gajah, vol. 51, pp. 16-25, IUCN/SSC Asian Elephant Specialist Group, 2020.
22. Wijayakulasooriya, J. V.; Automatic recognition of elephant infrasound calls using formant analysis and hidden markov model *Proceedings of the international conference on industrial and information systems*, pp. 244-248, 2011.
23. Zeppelzauer, M.; Stoeger, A. S.; and Breiteneder, C.; Acoustic Detection of Elephant Presence in Noisy Environments *Proceedings of the 2nd ACM International Workshop on Multimedia Analysis for Ecological Data*, pp. 3-8, ACM Press, 2013.
24. Hao, Y.; Campana, B.; and Keogh, E.; Monitoring and mining animal sounds in visual space *Journal of Insect Behavior*, vol. 26, no. 4, pp. 466-493, 2012.