Decoding Asian Elephant Vocalisations: Unravelling Call Types, Context-Specific Behaviors, and Individual Identities

Seema Lokhandwala[1(✉)], Rohit Sinha[1], Sreeram Ganji[2], and Balakrishna Pailla[2]

[1] Indian Institute of Technology Guwahati, Guwahati, India
{seema171655001,rsinha}@iitg.ac.in
[2] Reliance Jio AICoE, Hyderabad, India
{ganji.sreeram,balakrishna.pailla}@ril.com

Abstract. This paper investigates the automatic classification of four types of Asian elephant vocalizations (rumble, roar, trumpet, and chirp) recorded in Kaziranga National Park. Apart from the call type classification, the study explores individual identification and contextual analysis. Various classifiers using openSMILE features are developed to facilitate the classification process. The results demonstrate accurate classification of elephant call types and successful classification of context-specific behavior and individual identity based on trumpet and chirp calls, respectively. This study highlights the potential of acoustic analysis for understanding elephant communication and well-being, offering insights into their context-specific behavior and individual identities.

Keywords: Bioacoustics · Animal behavior · Animal communication · Elephant communication

1 Introduction

Asian elephants (*Elephas maximus*) are social and widely distributed mammals, relying on acoustic communication to navigate their complex social dynamics and geographically dispersed locations [3,28]. Acoustic communication plays a crucial role in various aspects of their lives, including, maintaining group cohesiveness, fostering cooperation, mating, and facilitating mother-infant interactions [11,17,22].

The repertoire of vocalizations produced by Asian elephants is remarkably diverse, encompassing a wide range of sounds that facilitate their communication. These vocalizations include low-frequency rumbles and growls, which can propagate over long distances, allowing elephants to communicate effectively over large areas [9,15]. Additionally, they are capable of producing high-frequency trumpets, chirps, roars, and barks, showcasing their remarkable vocal versatility in acoustic communication [14,22,25].

A. Karpov et al. (Eds.): SPECOM 2023, LNAI 14339, pp. 367–379, 2023.
https://doi.org/10.1007/978-3-031-48312-7_30

However, despite the significance of elephant vocalizations, there is still limited research in the field of automatic classification of call types. Clemins *et al.* conducted a study utilizing a hidden Markov model (HMM) for classifying call types, which stands as one of the few studies in this domain. The authors utilized mel-frequency cepstral coefficients, log energy, and spectrally derived features to train HMM classifiers and reported an overall classification accuracy of 79.7% [4]. However, there is a need for further investigations to address the existing limitations and develop more refined and accurate call-type classifiers.

Recent research has found a connection between the acoustic structure of elephant vocalizations and their arousal or motivational states. Berg *et al.* classified elephant calls into ten types, with respect to their corresponding behaviors. High-frequency calls, like trumpets, were associated with emotionally charged situations, while low-frequency calls like rumbles were prevalent in relaxed social contexts [1]. The acoustic properties of elephant rumbles reflect individual emotional arousal [23]. For example, rumbles occurring during socializing and agitation exhibit increased fundamental frequencies and decreased duration [30]. Wesolek *et al.* found that post-nursing cessation rumbles had distinct acoustic characteristics [29].

In a recent study, Stoeger *et al.* investigated African elephants *(Loxodonta africana)* and their ability to produce various call types, including snorts, rumbles, and trumpets in response to verbal cues from trainers (mahouts). The study revealed that rumbles produced during social interactions with conspecifics had distinct acoustic characteristics compared to rumbles elicited by trainer cues [24]. Sharma *et al.* found that rumbles, but not trumpets, were modulated during disturbances among wild Asian elephants [21]. Fuchs *et al.* observed trumpet calls conveying individual identity information but no modulation between greeting and disturbance contexts [7].

For a complete understanding of a species' communication system, it is essential to comprehend the information conveyed by the various vocalizations in its vocal repertoire. The use of vocalizations by a number of non-human mammalian species to communicate sex, caller identity, emotional state, and context has been documented [13,19,20,26]. Caller identity is important in social species in particular, and this has been observed in the majority of mammal and bird species.

Individual anatomical and morphological variations in the sound-producing structures, as well as internal factors and the physiology of sound production, all affect vocal identity [26]. Soltis *et al.* [23] and Fuchs *et al.* [7] used discriminant functional analysis to study the individual identity of rumble and trumpet vocalizations respectively, revealing distinctive acoustic characteristics associated with each call type. In the assessment of individual identity for trumpet calls, the authors report a classification accuracy of 71.7% [7]. In the analysis carried out Soltis *et al.*, the accuracy of individual identity based on rumbles stood at 60.0% [23]. In speaker recognition experiments carried out on rumble calls, Clemins *et al.* report an individual identification accuracy of 82.5% [4]. Based on the findings from the aforementioned studies, we have been motivated to propose a framework and build models to classify call types followed by individual identity and context-specific behavior.

The remainder of the article is structured as follows. Information on the elephant vocalization data gathered for this study is described in Sect. 2. Section 3 describes the proposed framework for detecting call types, individual identities, and the broad nature of context-specific behaviors in the collected elephant vocalization database. Section 4 discusses the findings of the study. Section 5 summarizes the conclusions drawn.

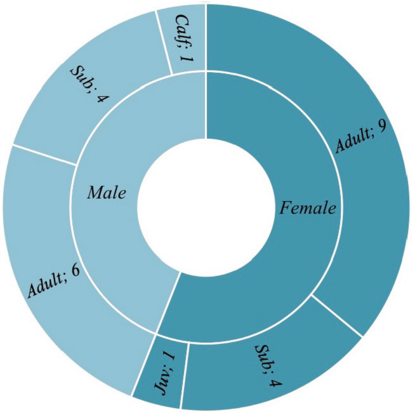


Fig. 1. Distribution of sexes and age groups among the studied subjects. (Sub refers to sub-adult and Juv refers to juvenile).

2 Database

2.1 Subjects and Study Site

The Kaziranga National Park and Tiger Reserve (hereinafter referred to as KNP), is a World Heritage Site in Assam (India), where the elephant vocalization data was collected. KNP houses around 60 semi-captive Asian elephants used for activities like patrolling, anti-poaching efforts, and tourism. Each elephant has a mahout to meet their daily needs. These elephants socialize, bathe together, play, freely browse in the forested area, and are accustomed to the presence of humans. For this study, based on their locations within KNP, a total of 25 elephants were selected with their ages spanning from less than 1 year to 60 years. They were categorized into four major age classes based on previous reports [27]: calf (below 1 year), juvenile (1–5 years), sub-adult (5–15 years), and adult (15 years and above). Figure 1 lists the sexes and age groups of the individuals under study.

2.2 Context-Specific Data Recording

The recording sessions were held throughout the field site, including the elephants' bathing areas, browsing areas, and places where they are tethered at night. No manipulative experiments to elicit responses from the elephants were carried out during these sessions. A round-robin approach was used during these sessions, with an average of 4 h spent monitoring each subject. A minimum of 15 min to an hour-long observation of their behavior per session was recorded at an interval of 30 s. We categorized elephant behavior into several broad categories, including locomotion, social interactions, handler interactions, self-directed actions, foraging, comfort, and other behaviors [12,18].

In the study, close observation of elephant behavior allowed for interpretation in the context of their social interactions. Based on this interpretation, three specific context-specific behaviors were assigned: positive, negative, and neutral.

The positive context-specific behavior was assigned when elephants were observed interacting with other non-dominant elephants. These interactions involved socializing, playing, and engaging in behaviors such as physical contact and moving toward each other.

The negative context-specific behavior was assigned when elephants interacted with their mahouts (human caretakers) or other dominant individuals. During these interactions, elephants exhibited specific behaviors such as head bobbing and body swaying, which are commonly associated with stress or agitation. Head bobbing refers to repetitive and rhythmic movements of the head, while body swaying refers to rhythmic side-to-side movements of the body. These interactions often occurred when elephants displayed signs of fear or distress, such as retreating or showing avoidance behaviors.

Neutral context-specific behavior was assigned when individuals engaged in contact calls without clearly displaying positive or negative context-specific behavior. Contact calls are vocalizations made by elephants to communicate with each other over long distances.

2.3 Collection and Categorization of Acoustic Data

The elephant vocalization data was recorded during the daytime from February to April 2021. The data was recorded for a total of 47 d, yielding 103 h of acoustic data. Behavior, caller's identity, approximate recording distance, and context were noted for each recorded vocalization. The vocalizations were collected with a Sound Devices MixPre-3 II recorder connected to an Earthworks QTC-40 omni condenser microphone with a frequency response of 3 Hz to 40 kHz, sampling at a rate of 48 kHz, and a range of 5–100 meters. Nikon D5100 and a Cannon 1200-D digital single-lens reflex camera were used for video recordings.

Through field notes, auditory observation, and spectrogram analysis, all vocalizations were identified. Based on the results of earlier studies, the vocalizations were divided into four main call types and combination calls [1,11,22]. A total of 401 elephant calls, encompassing individuals of all age groups and sexes were captured during our fieldwork. Detailed information about the dataset can

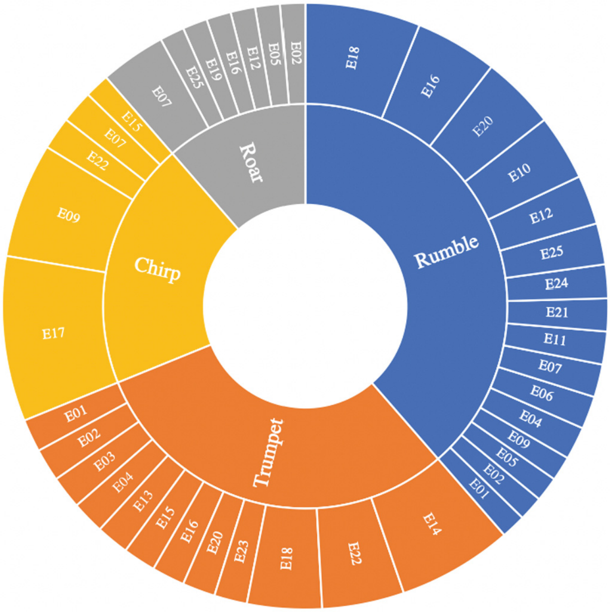


Fig. 2. Relative contribution of individual elephant in the data used for call type classification experiment.

be found in reference [10]. From this dataset, 226 calls, representing all call types, were selected for developing call type classifiers. The selection was based on the behavioral context and quality of the calls, as depicted in Fig. 2.

In order to conduct individual identification and context-specific behavior experiments, two types of calls: trumpet and chirp, were chosen. The rumble and roar call types were not selected due to the limited amount of data available per individual for rumble and roar calls. For the individual identification experiment, only the individuals who produced more than three calls were included. On average, we obtained 12.5 chirp calls from two individuals and 7.5 trumpet calls from eight individuals. Table 1 provides information regarding the number of calls used in chirp and trumpet call types for context-specific behavior experiments.

3 Proposed Framework

A comprehensive framework for analyzing elephant acoustic data is represented in Fig. 3. The process involves several key steps, starting with the segmentation of elephant calls. Once the calls are segmented, relevant features are extracted using advanced techniques. These features capture important acoustic characteristics of the elephant vocalizations. Next, the study utilizes five independent classifiers, each trained on the extracted features and associated with a specific label representing call type. Once the call type label is identified, we deployed five independent classifiers, each trained on the extracted features and associated with a specific label representing individual elephant and context-specific behavior for trumpet and chirp call types. Classification metrics are calculated to evaluate the performance of the classifiers and assess the quality of the classification results. Overall, this framework enables a systematic and effective analysis of elephant acoustics data, providing valuable insights into call type, individual identities, and behavioral patterns.

Table 1. Distribution of trumpet and chirp calls across three context-specific behaviors.

Context-specific behavior	Number of calls	
	Chirp	Trumpet
Positive	12	13
Neutral	5	22
Negative	8	25

3.1 Segmentation

The process of analyzing the acoustic recordings began with a visual examination using the PRAAT 6.2.03 software [2]. This involved opening the recordings and carefully observing the waveforms and spectrograms. We referred to our field notes and listened to the recordings to gather additional information about the calls. Once the calls were identified within the raw data, they were precisely located, marking the start and end times of each call. To extract relevant information, the calls were then trimmed, selecting the specific portions that contained the calls of interest.

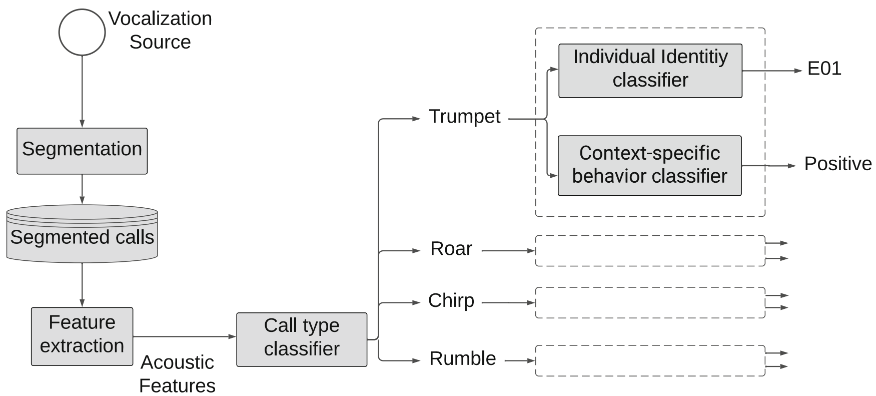
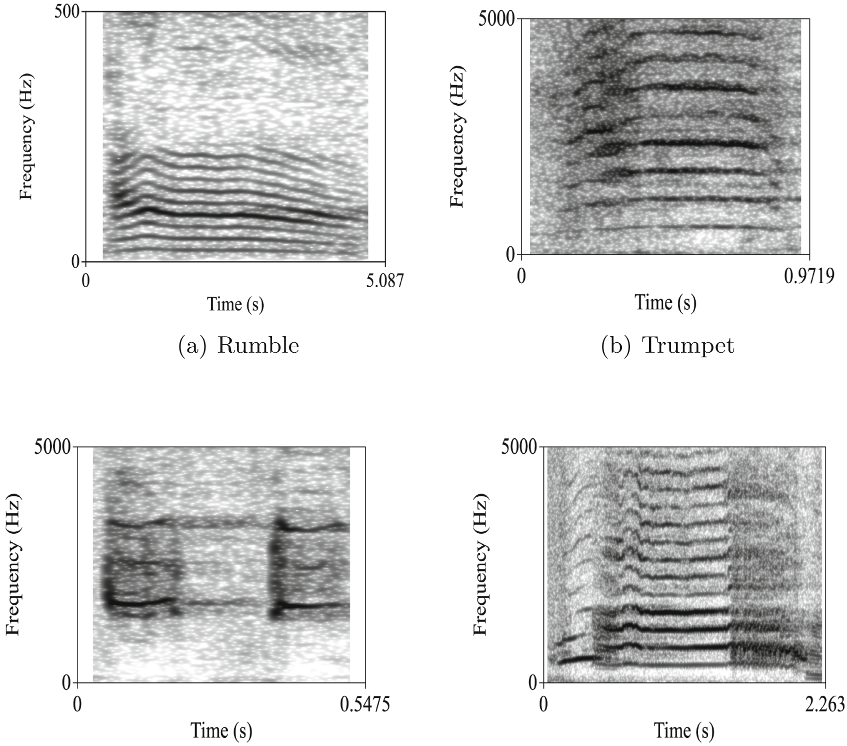


Fig. 3. A block diagram of the proposed framework. The modules within the dotted box are replicated for the rest of the call types, indicating their similar functionality.



(a) Rumble (b) Trumpet

(c) Chirp (d) Roar

Fig. 4. Spectrograms depicting four types of elephant calls: (a) rumble, (b) trumpet, (c) chirp, and (d) roar. These visual representations highlight the unique characteristics of the rumble calls, allowing for a clear classification.

3.2 Feature Extraction and Acoustic Analysis

Features were extracted using the Python-based open-source feature extraction openSMILE toolkit [5]. The feature set is an acoustic parameter set for various areas of automatic voice analysis. The feature set was extracted from the openSMILE toolkit using an extended Geneva Minimalistic Acoustic Parameter Set (eGeMAPS) which resulted in 25 low-level descriptors (LLDs) and 88 functionals were extracted. The 25 LLDs are made up of voicing features, spectral features, cepstral features, and energy features. There are 88 functionals produced after statistics like the variances, arithmetic mean, standard deviations, and percentiles of the LLDs are calculated. These LLDs were obtained from 25 ms frames and extracted every 10 ms. Only the 88 functionals served as inputs for training individual classifiers associated with specific labels, representing particular individuals, context-specific behaviors, or call types. The spectrograms of rumbles, trumpet, and roar calls of elephants exhibited distinct features, reflecting the unique acoustic characteristics of each call type. Rumbles displayed strong energy in the lower frequency range, spanning from infrasound frequencies to several hundred Hertz, often with harmonic or quasi-harmonic patterns and various modulation patterns. Trumpet calls were characterized by a wide frequency range, showcasing broadband energy across the entire spectrum, an initial transient or burst of energy, and potential harmonic structures. Roar calls demonstrated a broadband distribution of energy with an emphasis on the mid-frequency range, irregular or modulated patterns, and the presence of harmonic structures or non-harmonic components. Chirp calls were characterized by their unique temporal structure. In comparison to the other calls, they were significantly shorter in duration, making them stand out noticeably. It's important to note that these spectrogram features can vary among individual elephants and may be influenced by factors such as age, sex, and social context. Further analysis utilizing advanced signal processing techniques can extract quantitative features from spectrograms to facilitate classification and in-depth analysis.

3.3 Classifiers

To classify the elephant calls based on context-specific behaviors, call types, and individual identification, three distinct models were developed, each utilizing different criteria or feature sets. The models employed five different classification algorithms: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes, Multi-layer Perceptron (MLP), and Random Forest. SVM determined an optimal hyperplane to separate data points of different classes, while KNN classified new data based on the majority class of its nearest neighbors. Naive Bayes calculated the probability of a data point belonging to a certain class based on the assumption of feature independence. MLP, a type of artificial neural network, learned complex relationships between inputs and outputs. Random Forest combined multiple decision trees to make predictions.

For the small size of the database, a k-fold validation methodology, with k set to 3, is employed to evaluate the classification performances of a model. The subsets were created so that the sets of utterances within each of the three subsets were mutually exclusive. Data from test utterances made up subsets 1, 2, and 3 respectively, with each subset accounting for 30% of the testing set. This approach allowed for comprehensive evaluation and validation of the performance of each of the models.

3.4 Evalution Metrics

In this study, Accuracy is used to determine how well the classification model performed. It can be defined as,

$$\text{Accuracy}(\%) = 100 \times \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

where TP stands for "true positives", TN for "true negatives", FP for "false positives", and FN for "false negatives". The percentage accuracies for 3-fold of the data are calculated and reported in Tables 2 and 3.

4 Results and Discussion

4.1 Call Type Experiment

In the context of call type classification, the Random Forest model achieved the highest accuracy of 82.7%, followed by the Naive Bayes model. The confusion matrix for this experiment is presented in Fig. 5. Notably, rumbles were classified with the highest accuracy at 100% due to their low-frequency nature, whereas roars exhibited the lowest accuracy at 33%. One potential reason for this issue is the limited availability of training data for the "roar" call type. For the call type experiment, our findings are consistent with those reported by Clemins *et al.* [4].

Table 2. The following table showcases the average accuracy of the five models for call types.

Classification model	Average Accuracy (%)
Support Vector Machine	65.0
K-Nearest Neighbors	64.5
Naive Bayes	72.6
Multi-layer Perceptron	61.0
Random Forest	**82.7**

To determine which features play an important role in this classification, feature importance was analyzed. The top five features identified were alphaRatioV-sma3nz-stddevNorm, loudness-sma3-stddevNorm, loudnessPeaksPerSec, mfcc2-sma3-stddevNorm, and F1bandwidth-sma3nz-amean. The first feature, alphaRatioV, represents the variation of the alpha ratio in an audio signal. This ratio provides insights into the spectral balance of the signal. The second feature, loudness-stddevNorm, reflects the normalized standard deviation of the signal's loudness. The third feature, loudnessPeaksPerSec, denotes the number of loudness peaks detected per second in the audio signal. These peaks represent instances of significantly high amplitude, such as trumpet, chirp, and roar which have higher frequency calls compared to rumble. Thus, loudness serves as a distinguishing factor in classifying these sounds, as evident from its inclusion in the top features for such classification. The fourth feature, mfcc2-stddevNorm, characterizes the second mel-frequency cepstral coefficient (MFCC), and is widely employed in audio signal analysis for tasks such as speech recognition and speaker identification. Finally, the fifth feature, F1bandwidth-sma3nz-amean refers to the width or range of frequencies around the first formant peak, which is an important component in speech analysis. The first formant represents the primary resonance frequency of the vocal tract during speech production.

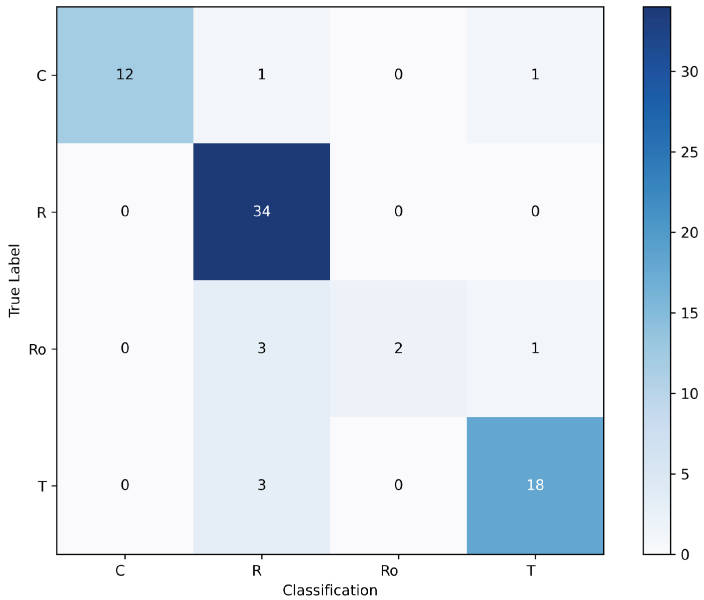


Fig. 5. Confusion matrix for call type experiment. (C-chirp, R-rumble, Ro-roar, T-trumpet).

4.2 Context-Specific Behavior and Individual Identification Experiment

For context-specific behavior and individual identification experiments, we utilized trumpet and chirp calls as the primary data. In the case of context-specific behavior classification, the Random Forest model achieved the highest accuracy of 75%, followed by the Naive Bayes model. Similarly, for chirp calls in the case of context-specific behavior classification, the Random Forest model achieved the highest accuracy of 72.6%, followed by the Naive Bayes model. Compared with earlier literature, Fuchs *et al.* observed trumpet calls in greeting and disturbance contexts, achieving a classification accuracy of 58.3%. In our own study, we observed trumpet calls in three different contexts and achieved a higher classification accuracy of 75%.

In terms of individual identification experiments, the Random Forest model is noted to outperform the others, achieving an accuracy of 91.6% for chirp calls compared to 71.6% for trumpet calls. Chirp calls had higher accuracy for individual identification because only data from two individuals were used, whereas eight individuals' data were used for trumpet calls. Compared with earlier literature, rumble, and trumpet calls were classified according to individual identity, indicating that acoustic characteristics varied based on the individual identity of the caller [7,23]. When specifically comparing trumpet calls, it is worth noting that Fuchs *et al.* [7] reported a classification accuracy of 71.7% for six individual

Table 3. The average accuracies of the five different models for context-specific behavior (Context) and individual identification (Identity) experiments for trumpet and chirp call types.

Classification model	Average Accuracy (%)			
	Trumpet		Chirp	
	Context	Identity	Context	Identity
Support Vector Machine	54.9	43.3	48.6	83.7
K-Nearest Neighbors	48.3	46.6	51.8	79.6
Naive Bayes	58.3	58.3	64.3	87.9
Multi-layer Perceptron	53.3	46.6	43.9	68.0
Random Forest	**75.0**	**71.6**	**72.6**	**91.6**

elephants, whereas our study achieves a closely similar classification accuracy of 71.6% for a slightly larger sample size of eight individual elephants.

Overall, these models employ different approaches and algorithms to classify elephant calls based on context-specific behaviors, call types, and individual identification. The Random Forest model demonstrated strong performance in classifying call types, context-specific behaviors, and individual identification. These results highlight the strengths and weaknesses of each model in capturing the underlying patterns in elephant calls for different classification tasks.

In the future, a detailed analysis needs to be conducted to determine which features play a significant role in the analysis. The Asian elephant is a social species which lives in matriarchal family groups [6,28]. They form social bonds (relationships) with unrelated individuals in captivity and even provide reassurance to distressed conspecifics [8,16]. Therefore, identifying the caller holds significant value as it fosters support and enhances social interactions among individuals. Playback experiments are recommended in the future to determine how well Asian elephants can identify and differentiate familiar conspecifics based on their vocalizations.

5 Conclusion

In conclusion, this study developed a framework to classify call types and then also demonstrated that the acoustic of elephant calls are context-specific, exhibiting distinct characteristics in relation to different context-specific behavioral states. Furthermore, the aim in the future is to develop an end-to-end architecture that can not only classifies context-specific behavior in all elephant calls but also recognize individual identity. This comprehensive understanding of elephant communication, encompassing both context-specific behavioral states and individual variations, contributes to a more nuanced comprehension of elephant behavior and communication. These findings have potential implications for conservation efforts, captive elephant welfare, and advancing our understanding of how elephants express themselves through vocalizations.

Acknowledgements. The first author extends sincere appreciation to the Assam Forest Department for graciously granting permission to gather the essential data required for this study. Their invaluable cooperation and consent have made it possible to conduct this research. Additionally, we would like to express heartfelt gratitude to the dedicated forest guards and diligent mahouts who wholeheartedly supported and actively participated in the fieldwork.

References

1. Berg, J.K.: Vocalizations and associated behaviors of the African elephant (Loxodonta africana) in captivity. Zeitschrift für Tierpsychologie **63**, 63–79 (1983). https://doi.org/10.1111/j.1439-0310.1983.tb00741.x
2. Boersma, P., Weenink, D.: Praat: doing phonetics by computer (version 5.1.13) (2009). http://www.praat.org
3. Calabrese, A., et al.: Conservation status of Asian elephants: the influence of habitat and governance. Biodivers. Conserv. **26**(9), 2067–2081 (2017). https://doi.org/10.1007/s10531-017-1345-5
4. Clemins, P.J., Johnson, M.T., Leong, K.M., Savage, A.: Automatic classification and speaker identification of African elephant (Loxodonta africana) vocalizations. J. Acoust. Soc. Am. **117**(2), 956–963 (2005)
5. Eyben, F., Wöllmer, M., Schuller, B.: openSMILE - the munich versatile and fast open-sourceaudio feature extractor. In: Proceedings of the 9th ACM International Conference on Multimedia, (MM), pp. 1459–1462 (2010)
6. Fernando, P., Lande, R.: Molecular genetic and behavioral analysis of social organization in the Asian elephant (Elephas maximus). Behav. Ecol. Sociobiol. **48**, 84–91 (2000)
7. Fuchs, E., Beeck, V.C., Baotic, A., Stoeger, A.S.: Acoustic structure and information content of trumpets in female Asian elephants (Elephas maximus). PLoS ONE **16**, 1–19 (2021). https://doi.org/10.1371/journal.pone.0260284
8. Gadgil, M., Nair, P.V.: Observations on the social behaviour of free ranging groups of tame Asiatic elephant (Elephas maximus Linn). Proc. Anim. Sci. **93**, 225–233 (1984)
9. Leighty, K.A., Soltis, J., Wesolek, C.M., Savage, A.: Rumble vocalizations mediate interpartner distance in African elephants Loxodonta africana. Anim. Behav. **76**(5), 1601–1608 (2008)
10. Lokhandwala, S., Sarmah, P., Sinha, R.: Classifying mahout and social interactions of Asian elephants based on trumpet calls. In: Prasanna, S.R.M., Karpov, A., Samudravijaya, K., Agrawal, S.S. (eds.) Speech and Computer. SPECOM 2022. Lecture Notes in Computer Science, vol. 13721, pp. 426–437. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-20980-2_37
11. Nair, S., Balakrishnan, R., Seelamantula, C.S., Sukumar, R.: Vocalizations of wild Asian elephants (Elephas maximus): structural classification and social context. J. Acoust. Soc. Am. **126**, 2768–2778 (2009). https://doi.org/10.1121/1.3224717
12. Olson, D.: The Elephant Husbandry Resource Guide. American Zoo and Aquarium Association, Silver Spring, Maryland (2004)
13. Palacios, V., Font, E., Márquez, R., Carazo, P.: Recognition of familiarity on the basis of howls: a playback experiment in a captive group of wolves. Behaviour **152**(5), 593–614 (2015)

14. Pardo, M.A., et al.: Differences in combinatorial calls among the 3 elephant species cannot be explained by phylogeny. Behav. Ecol. **30**, 809–820 (2019). https://doi.org/10.1093/beheco/arz018
15. Payne, K.B., Langbauer, W.R., Thomas, E.M.: Infrasonic calls of the Asian elephant (Elephas maximus). Behav. Ecol. Sociobiol. **18**, 297–301 (1986)
16. Plotnik, J.M., De Waal, F.B.: Asian elephants (Elephas maximus) reassure others in distress. PeerJ **2**, e278 (2014)
17. Poole, J.H., Payne, K., Langbauer, W., Moss, C.: The social contexts of some very low-frequency calls of African elephants. Behav. Ecol. Sociobiol. **22**, 385–392 (1988). https://doi.org/10.1007/BF00294975
18. Rees, P.A.: Activity budgets and the relationship between feeding and stereotypic behaviors in Asian elephants (Elephas maximus) in a zoo. Zoo Biol. **28**, 79–97 (2009)
19. Rendall, D.: Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. J. Acoust. Soc. Am. **113**(6), 3390–3402 (2003)
20. Rendall, D., Rodman, P.S., Emond, R.E.: Vocal recognition of individuals and kin in free-ranging rhesus monkeys. Anim. Behav. **51**(5), 1007–1015 (1996)
21. Sharma, N., S, V.P., Kohshima, S., Sukumar, R.: Asian elephants modulate their vocalizations when disturbed. Anim. Behav. **160**, 99–111 (2020). https://doi.org/10.1016/j.anbehav.2019.12.004
22. Silva, S.D.: Acoustic communication in the Asian elephant Elephas maximus maximus. Behaviour **147**, 825–852 (2010). https://doi.org/10.1163/000579510X495762
23. Soltis, J., Leong, K., Savage, A.: African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. Anim. Behav. **70**, 589–599 (2005). https://doi.org/10.1016/j.anbehav.2004.11.016
24. Stoeger, A.S., Baotic, A.: Operant control and call usage learning in African elephants. Philos. Trans. R. Soc. B **376**(1836), 20200254 (2021)
25. Stoeger, A.S.: An Asian elephant imitates human speech. Curr. Biol. **22**, 2144–2148 (2012). https://doi.org/10.1016/j.cub.2012.09.022
26. Taylor, A.M., Reby, D.: The contribution of source-filter theory to mammal vocal communication research. J. Zool. **280**(3), 221–236 (2010)
27. Varma, S., Baskaran, N., Sukumar, R.: Field key for elephant population estimation and age and sex classification. Resource material for synchronized elephant population count using block count, line transect dung count method and waterhole count. Asian Nature Conservation Foundation, Innovation Centre, Indian Institute of Science, Bangalore - 560 012 (2012)
28. Vidya, T.N., Sukumar, R.: Social organization of the Asian elephant (Elephas maximus) in Southern India inferred from microsatellite DNA. J. Ethol. **23**, 205–210 (2005). https://doi.org/10.1007/s10164-005-0144-8
29. Wesolek, C.M., Soltis, J., Leighty, K.A., Savage, A.: Infant African elephant rumble vocalizations vary according to social interactions with adult females. Bioacoustics **18**, 227–239 (2009). https://doi.org/10.1080/09524622.2009.9753603
30. Wood, J.D., McCowan, B., Langbauer, W.R., Viljoen, J.J., Hart, L.A.: Classification of African elephant loxodonta Africana rumbles using acoustic parameters and cluster analysis. Bioacoustics **15**, 143–161 (2005). https://doi.org/10.1080/09524622.2005.9753544