Smart Agricultural Technology 9 (2024) 100682

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Acoustic-based models to assess herd-level calves' emotional state: A machine learning approach

Maíra Martins da Silva [a], Robson Mateus Freitas Silveira [b,c,*], Gean Gobo da Cruz [a], Karen Airosa Machado de Azevedo [c], Carla Maris Machado Bittar [b], Iran José Oliveira da Silva [c]

[a] São Carlos School of Engineering, University of São Paulo (USP), Av. Trab. Sancarlense 400, 13566-590 São Carlos, SP, Brazil

[b] Department of Animal Science, "Luiz de Queiroz" College of Agriculture (ESALQ), University of São Paulo (USP), Av. Pádua Dias, 11, Piracicaba, SP, Brazil

[c] Livestock Environment Research Group (NUPEA), Department of Biosystems Engineering, "Luiz de Queiroz" College of Agriculture (ESALQ), University of São Paulo (USP), Av. Pádua Dias, 11, Piracicaba, SP, Brazil

ARTICLE INFO

Keywords:
Calves' vocalization
Machine-learning classifiers
Frequency-domain-based metrics
Sound quality domain metrics

ABSTRACT

Animal bioacoustics is an important tool for monitoring different aspects of the physiology, behavior and well-being of animals remotely, non-invasively and continuously. Studies using this science are growing mainly due to the development of machine learning. This work aims to investigate the use of machine-learning classifiers to determine whether calves' vocalization audio data can be used to assess their welfare condition regarding feeding. Firstly, we collected several calves' vocalization audio data before and after feeding the animals at different day times and ages. Then, eleven time-domain, frequency-domain and sound quality-based metrics were extracted from these audio data and used as features for the classifiers. These features were used to determine whether vocalization audio data belonged to before or after feeding classes. Moreover, the most relevant ones were identified using the Random Forest algorithm. Finally, seven machine-learning classifiers were trained and tested, considering the entire set of features and a subset containing the most relevant features. The k-nearest neighbor classifier trained with the subset of the most relevant features obtained a 98.37% accuracy. Both frequency-domain and sound quality features played important roles in this classification. The main implications of this study are the development of a methodological proposal to study acoustics using machine learning and the fact that vocalization is a biomarker of animal welfare.

1. Introduction

Animal welfare encompasses theories concerning an animal's natural life, biological functioning, and emotional condition [1–4]. As Nielsen et al. [4] highlighted, understanding feeding behavior and nutritional requirements is essential for assessing animal welfare. Calves are usually fed milk (or milk replacer) twice daily until they are 8 or 9 weeks old. During this life stage, assessing their welfare at a herd level is of utmost importance since it can be valuable in assessing the health and growth of calves [5–6]. Many strategies based on physiology and/or animal behavior on have been explored for assessing animal welfare using feeding behavior. On the one hand, de Passillé et al. [7] investigated the correlation between several physiological behaviors as open-field conditions with sniffing, licking, running, walking, vocalizing, and jumping with welfare conditions, including feeding behavior. In addition, Macmillan et al. [8] and Gaillard et al. [9] propose behavior-monitoring techniques. Macmillan et al. [8] monitored rumination and activity behavior using a collar-mounted automated activity monitoring system. Due to the importance of estimating animals' nutrient requirements, Gaillard et al. [9] claimed that new technologies, such as real-time sensing animal condition techniques, should be extensively used.

The relationship between vocal parameters and animal welfare has been widely investigated due to the potential of vocalizations as noninvasive indicators of the emotional and physiological state of animals. Vocal parameters, such as frequency, amplitude, duration and tonal variation, can reflect responses to the environment, handling conditions and social interactions. Animals in situations of stress, pain or discomfort tend to emit vocalizations with distinct characteristics, such as higher frequencies and greater intensity, compared to states of relaxation or contentment, which generate softer and more harmonic sounds. Furthermore, analysis of vocalizations can identify signs of distress associated with isolation, hunger, thirst or physical restraint,

* Corresponding author.
E-mail address: robsonmateusfs1994@gmail.com (R.M.F. Silveira).

https://doi.org/10.1016/j.atech.2024.100682
Received 30 October 2024; Received in revised form 27 November 2024; Accepted 29 November 2024
Available online 30 November 2024

allowing adjustments in handling practices [10–13]. Schnaider et al. [10] realized that high vocal frequency could happen during the arrival of the calf in the corral and the reunion of the calf and its mother. This claim was found by assessing the fundamental frequency of vocalization audio data. Similarly, de la Torre et al. [14] aimed to identify vocal parameters that can characterize cow and calf contact calls. Shorten and Hunter [15] used acoustic sensors to automate the detection of cow vocalization duration and type (open mouth, closed mouth, and mixed mouth).

Several authors have been addressing animal welfare conditions using machine learning strategies. In fact, Siegford et al. [16] indicated that automated behavior analysis tools had been rapidly developed for use with animals. Regarding the calves' welfare condition, Gavojdian et al. [17] proposed the use of machine learning techniques to perform this evaluation. Jung et al. [18] exploited a deep learning-based cattle vocal classification model for real-time livestock monitoring, while Peng et al. [19] extracted and combined multiple acoustic features for cattle call pattern classification. Our proposal also exploits machine-learning classifiers to assess calves' emotional state using vocal audio data. This emotional state can be further correlated to animal welfare metrics. The novelty of this work is the inclusion of sound quality-based metrics in the set of investigated features for the derivation of machine learning classifiers related to calves' feeding condition at herd level. Sound quality refers to how individuals perceive and react to a sound. Perception is a mental capability shaped by sensory input and influenced by objective and subjective factors. Sound quality metrics reflect the extent to which the sound is deemed suitable for that specific aim. Several metrics have been proposed for evaluating the human perception of a sound, such as loudness, sharpness, and roughness, among others [20].

In summary, we investigate the importance of some features extracted from audio data acquired during calves' vocalization to identify their feeding condition at a herd level. The primary objective of this work is to investigate the use of machine-learning classifiers to identify if calves' vocalization audio data belongs to the before-feeding or after-feeding classes. However, the efficiency of a classifier can be significantly improved by exploring proper features extracted from the audio data. Therefore, a secondary objective of this work, which is equally important, is to evaluate some time and frequency-domain and sound quality-based features by assessing their importance for the classification. This evaluation of the features' importance, denoted as feature engineering, is a crucial step in our research and could potentially significantly impact the classifiers' performance. In summary, this work presents a tool that distinguishes between the vocals of a group of animals, which can be a manner to assess welfare.

2. Materials and methods

2.1. Ethical note

The study was carried out at São Paulo University - Luiz de Queiroz College of Agriculture (Piracicaba City, São Paulo State, Brazil). The region where the research was conducted is located at the geographical coordinates 22° 42′ 30″ S and 47° 38′ 00″ W and at an altitude of 546 m. It was approved by the Animal Ethics Committee (CEUA) of the same university, under protocol n. 582,210,722. The study was carried out in this same higher education institution and was in compliance with the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines.

2.2. Animals

A total of 75 calves aged between 3 and 8 weeks and clinically healthy were studied. After birth, calves were immediately separated from their mothers and fed the equivalent of 5% of birth weight with high-quality colostrum and a dose of powdered colostrum (SSCL, 100 g IgG) within 2 h of birth.

Calves were housed in individual hanging cages until 14 days of age and then in individual shelters until 8 weeks of age (Fig. 1). The animals had access to drinking water and starter concentrate (20% Crude Protein). The calves were fed with commercial milk replacer (Agroceres, Brazil), diluted to 14% solids, in a volume of 6 L/d until 42 days of age, 4 L/d until 49 days and then 2 L/d until weaning at 56 days of age. The total volume was supplied in teat-buckets for two meals (7am and 5pm).

2.3. Data organization

We acquired this audio data before and after feeding on different days and calves' ages (Fig. 2). Subsequently, the data were organized as before and after feeding the milk replacer (Table 1). The main objective of collecting audio data sets at different ages (July/2023 and November/2023) and daytime (morning and afternoon) is to investigate if the same classifier can be used, considering different animals' ages and different daytimes. This is important since a simpler strategy could be used if a single classifier achieves good accuracy with different data sets.

Some individuals may have produced sound during the acquisition, while others remained silent. We extracted the samples from the audio data sets considering the individuals' variety. This approach corroborates achieving results at the herd level rather than the individual level.

2.4. Data acquisition

Audio data have been recorded before and after feeding to obtain the distinct calves' vocalizations. Calf vocalizations were recorded using a Sennheiser MKE 200 microphone (frequency response 40–20,000 Hz, max SPL 120 dB at 1000 Hz) and recorded on a Zoom Hn1 digital recorder with stereo input (sampling rate 44.1 kHz). Each was stored in .wav format with 16-bit amplitude resolution. The microphone was hung near the rain houses (Figs. 1 and 2) at a distance of approximately 1 m above the ground for calf vocal recordings. An audio data sample of a calf vocalization in the morning before feeding is illustrated in Fig. 3, and the absolute values of this audio data set's fast fourier transform (FFT) are depicted in Fig. 4. The audio data are proportional to the sound pressure level, a measure that can be acquired in *Pa* and has been scaled so that its maximum value is below 15,050. This maximum value has been arbitrarily selected.

Fig. 1. Experimental location.

Fig. 2. Calves fed with milk replacer (The organization of the calves during feeding).

Table 1
Amount of acquired data during a specific month and day time.

Amount of audio data	Before feeding	After feeding
July/morning	144	36
July/afternoon	100	36
November/morning	72	10
November/afternoon	48	44
July	244	72
November	120	54
All data	364	126

2.5. Feature extraction

Feature engineering aims to identify pertinent details within raw data and convert them into a format that a machine-learning based model can readily explore, improving its performance. In this work, the raw data consists of audio data sets. Each audio data set contains an entire vocalization of a single calf. Data containing two vocalizations are not considered in this work but may investigated in the future. Aiming to identify features that may play an important role in the detection of the feeding status, this work investigates three classes of features: frequency-domain, time-domain and sound quality metrics. Table 2 describes the eleven features explored in this work.

The features *F*1 and *F*2 were derived from the absolute values of each audio data set's FFT (depicted in Fig. 4). *F*1 is the average value of the FFT's amplitudes in the frequency range from 0 up to 2600 Hz, while *F*2 is the frequency value of the FFT's peak value. The features *F*4, *F*5, *F*6, *F*7 and *F*8 are the bin histogram values. This 5-bin histogram is found by evaluating the actual envelope of the time-domain signal. The actual envelope is derived by the mean of the upper and lower envelopes. While Fig. 5(a) illustrates a single time-domain audio data, the upper and the lower envelopes and the actual envelope, Fig. 5(b) shows the histogram of the actual envelope with 5 bins. Finally, *F*11 is the abscissa of the highest value obtained by deriving a spectrogram. Fig. 6 depicts the spectrogram of the audio illustrated in Fig. 3. A cross signal highlights the highest value.

Acoustic engineers and researchers develop sound quality metrics to assess sound qualities objectively. This work investigates the use of such metrics for identifying calves' vocalization condition. The features *F*3, *F*9, and *F*10 are sound quality metrics [21], also denoted as psycho-acoustic models [22]. The following sound quality metrics were investigated:

Loudness (*F*3): According to American National Standards Institute [23], loudness can be defined as the attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud. In other words, it is the subjective perception of sound pressure [20]. This metric was derived for each dataset using the MOSQITO Python toolbox [21]. MOSQITO offers the possibility of computing the acoustic loudness of a signal according to the Zwicker method for time-varying signals according to ISO 532–1:2017 [24].

Roughness (*F*9): Roughness is understood as the texture perception of the sound. It depends on the distance between the partials measured in critical bandwidths [20]. Despite several models have been developed to compute the acoustic roughness [20], the most commonly used is the one proposed by Daniel [25]. This calculation is available in MOSQITO and is used in this work.

Sharpness (*F*10): Sharpness is a hearing perception associated with frequency, differently from loudness. Its derivation is carried out by a specific loudness distribution of the sound using weighting functions. At MOSQITO, the default weighting functions are defined in DIN 45,692:2009–08 [26].

2.6. Feature engineering

In this study, the objective of the classifiers is to distinguish between before-feeding and after-feeding classes based on audio data sets. In the preceding section, we discussed the features under investigation. The classifiers were trained and tested using the complete feature set and a

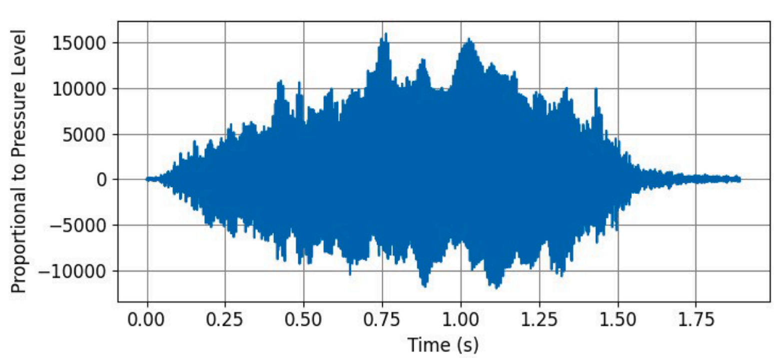

Fig. 3. Audio data of a calf mooing in the morning before feeding.

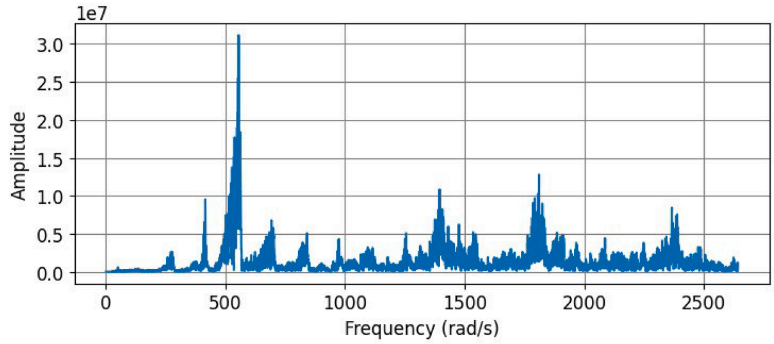

Fig. 4. The absolute value of sound signal's FFT of a calf mooing in the morning before feeding.

Table 2
Frequency-domain, and Psycho-acoustic metrics.

Feature (*Fj*)	Description	Metric's Nature
*F*1	Frequencies' Mean	Frequency-domain
*F*2	Max. Amplit.'s Freq.	Frequency-domain
*F*3	Loudness	Sound quality
*F*4-*F*8	Histogram	Time-domain
*F*9	Roughness	Sound quality
*F*10	Sharpness	Sound quality
*F*11	Spectrum	Frequency-domain

carefully selected subset of the most significant features using feature selection.

Feature selection plays a crucial role in improving the performance of classification models while reducing computational complexity. As highlighted in previous research by Khaire and Dhanalakshmi [27], evaluating the importance of features allows for identifying those most relevant to the task at hand, enhancing model accuracy. By using a more compact yet pertinent subset of features, we can streamline the model and improve its efficiency.

This study focus on applying the Random Forest (RF) algorithm to select the most relevant features for classification from a set of 11 features, as described in Table 2. The RF algorithm has been widely used in previous works [28–29] for classification, regression, and feature selection tasks. It constructs multiple decision trees, as shown in Fig. 7, and uses the average of these trees to determine the final classification output [30]. In this work, we generate 100 decision trees to derive this average and assess their effectiveness in identifying the most significant features for classification.

In Fig. 7, *target*=0 means that the data set was acquired before feeding the animals (before-feeding class), and *target*=1 means that the data set was acquired after feeding the animals (categorized as the after-feeding class). The features are denoted as *F*?, and each decision tree consists of a root node, intermediate nodes, and leaf nodes. The figure illustrates a decision tree with three nodes, i.e. a decision tree with a single intermediate node. An optimal decision is made for each node using the Gini impurity measure to identify the most relevant features

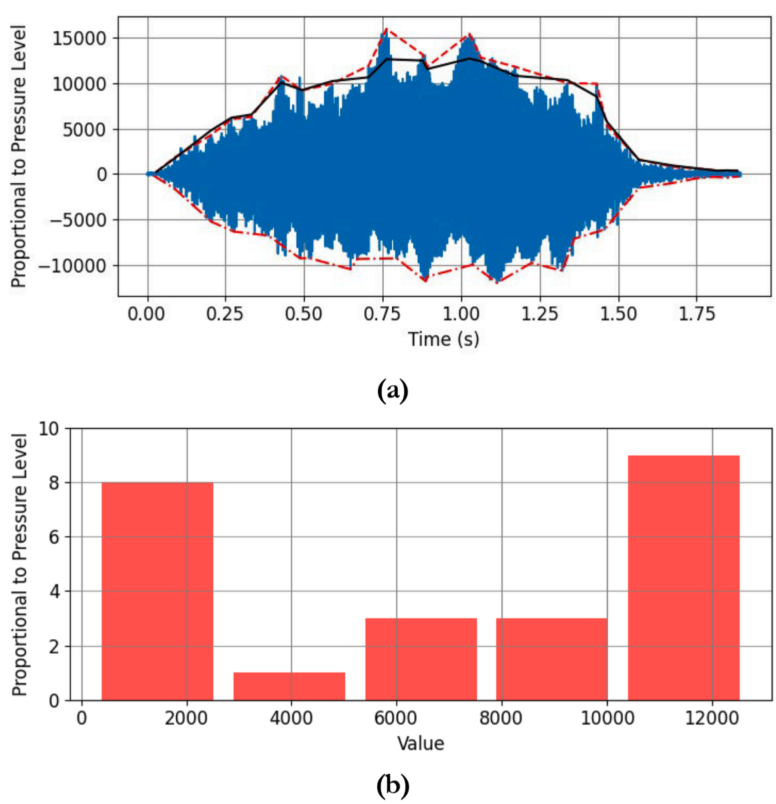

Fig. 5. (a) Time-domain audio data and envelopes and (b) Histogram with 5 bins.

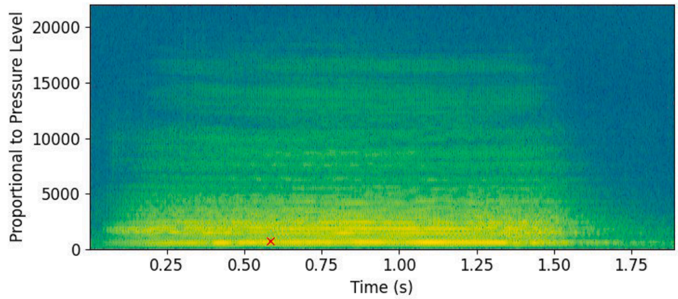

Fig. 6. Spectrogram and its highest value.

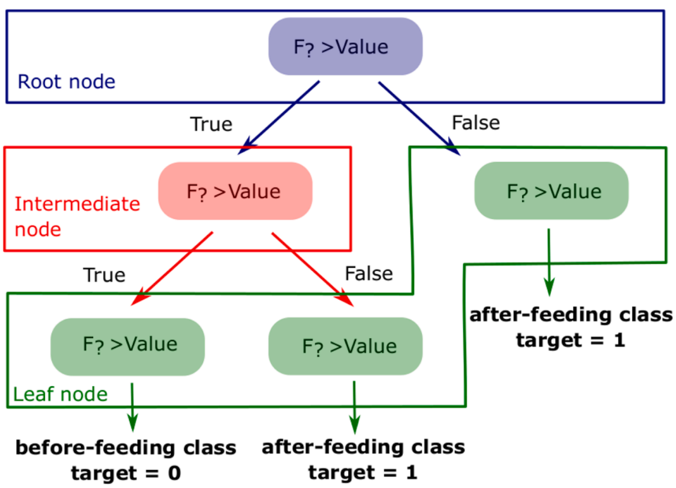

Fig. 7. Illustration of a Decision Tree for Classification of before-feeding and after-feeding classes.

for the classifications.

The decision values can be either "true" or "false" based on the condition *F?>Value*, as depicted in Fig. 7. It is worth noting that a decision tree may have fewer nodes than the total number of features, highlighting the possibility of classification with a reduced feature set consisting of the most critical features. This reduced set was adopted in our study.

2.7. Machine-learning based classifiers

This work exploits two sets of features: the entire set containing 11 features and a reduced subset containing the most relevant ones. This investigation compares seven well-known machine-learning based classifiers: k-nearest neighbors (k-NN) [31], support vector machines considering a linear Kernel function (SVM-Linear) [32–33], support vector machines considering radio basis Kernel function (SVM-RBF) [32–33], decision tree (DT) [34], RF [28–29], multi-layer perceptron (MLP) [35] and AdaBoost [36]. For completeness, these classifiers and their hyperparameters are briefly explained. In this work, we used the Python library scikit-learn to implement the classifiers.

2.7.1. k-nearest neighbor (k-NN)

The k-NN classifier, introduced by Cover and Hart [31] in their seminal work, operates by assigning the class label of the most frequently occurring pattern among the k nearest training samples. For example, when $k = 1$, the assigned pattern corresponds to the nearest neighbor.

This assignment can be carried out uniformly or weighted according to a distance metric between the samples. In this work, the distance metrics investigated are Euclidean and Manhattan distance metrics, according to Table 3. The former metric corresponds to the Euclidean distance, while the latter corresponds to the sum of the absolute differences between the samples' Cartesian coordinates.

One notable advantage of this approach is its ability to create a decision surface that readily adapts to the distribution shape of the training data. However, a significant drawback lies in the complexity of the training phase, which involves the derivation of multiple distance metrics and the sorting and verification of several neighboring samples.

2.7.2. Support vector machine (SVM) - Linear and RBF

Support Vector Machine (SVM) is a supervised learning method designed to derive a hyperplane, often referred to as a 'hard margin', that effectively separates data patterns. In Fig. 8, one can perceive two distinct patterns categorized as 'target $= 0'$ and 'target $= 1$,' visually represented by the hyperplane equation $\boldsymbol{wT{\cdot}xi}+b = 0$. The data points closest to this hyperplane are referred to as 'support vectors'. The coefficients $\mathbf{w}$ and b are determined during the training phase.

The primary objective is to maximize the distances between the hyperplane and these support vectors, as outlined by [32–33]. However, there are scenarios where a hyperplane cannot perfectly separate datasets. In such cases, a dimensionless transformation can be applied to the input sets using Kernel functions, which reshape the data to be linearly separable. Various Kernel functions have been proposed in the literature, and our study investigates two specific ones: the linear and the Radial Basis Function (RBF) Kernel functions, as discussed by [33].

Additionally, SVM allows for some margin violations by imposing a penalty factor on the optimization problem. This penalty factor is known as the box constraint, represented by hyperparameter 'C'. This approach is referred to as 'soft-margin' SVM. The RBF Kernel function introduces a hyperparameter, denoted as 'Y', which controls the degree of flexibility in this transformation. In Fig. 8, one can observe the concept of soft margins applied to our case study.

2.7.3. Decision tree (DT) and random forest (RF)

A Decision Tree (DT), as depicted in Fig. 7, is a series of if-else statements created during the training stage and can be used as a classification tool [34].

The tree begins at the root node, progresses through intermediate nodes, and terminates at the leaves, as shown in Fig. 7. It's important to note that as the tree depth increases, the number of samples required for tree expansion doubles at each level. Users can control the tree's depth using the hyperparameter 'max_depth' to mitigate overfitting. Similarly, the user can regulate the number of features considered at each

Table 3
Classifiers' grid parameters.

Classifier	Parameters	Range
k-NN	k	1,3,5,7,13
k-NN	weights	uniform and distance
k-NN	metric	Euclidean and Manhattan
SVM-Linear	C	0.01, 0.02, 0.05, 0.1, 0.2, 0.25, 0.3, 0.33, 0.4, 0.5, 0.75, 0.9, 1, 2, 3
SVM-RBF	C	0.01, 0.02, 0.05, 0.1, 0.2, 0.25, 0.3, 0.33, 0.4, 0.5, 0.75, 0.9, 1, 2, 3, 5, 10, 20, 50, 100, 150
SVM-RBF		0.01, 0.02, 0.05, 0.1, 0.2, 0.25, 0.3, 0.33, 0.4, 0.5, 0.75, 0.9, 1, 2, 3
DT	max_depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20
DT	max_features	auto, sqrt, log2
RF	max_depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20
RF	max_features	auto, sqrt, log2
RF	n_estimators	50, 70, 80, 90, 100, 110, 120, 130, 150
MLP	alpha	0.0001, 0.001, 0.01, 0.1, 1, 10
MLP	solver	lbfgs, sgd, adam
MLP	activation	identity, logistic, tanh, relu
Adaboost	n_estimators	50, 70, 80, 90, 100, 110, 120, 130, 150
Adaboost	learning_rage	0.0001, 0.001, 0.01, 0.1, 1, 9

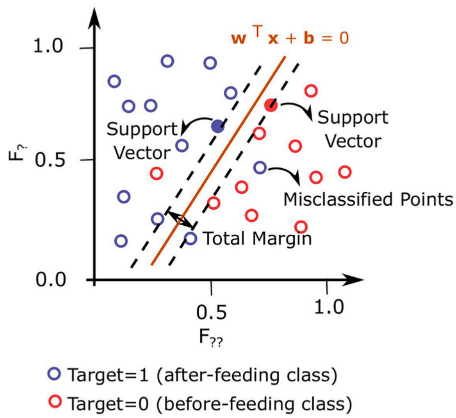

Fig. 8. Illustration of a soft-margins SVM-based classifier.

branching decision using the 'max_features' hyperparameter. This hyperparameter can be automatically determined or manually set based on factors such as the square root or logarithm of the total number of features. For example, the algorithm uses 5 randomly selected features in the splitting process if the 'max_features' is set to 5.

As mentioned, the Random Forest (RF) algorithm calculates several decision trees and their average results in the output's classifier [30]. Several decision trees have been evaluated during the training stage. This hyperparameter is denoted as the number of estimators ('n_estimators').

2.7.4. Multi layer perceptron (MLP)

A Multi-Layer Perceptron (MLP) classifier maps an input dataset, $x \in \mathbb{R}$, into appropriate targets (outputs), $\mathbf{o}\cdot \in \mathbb{R}O$ (Rumelhart and McClelland, 1987). In this work, we evaluated the performance of this classifier considering two input subsets $I = 11$ or I=the number of relevant features, and two possible outputs (target $=0$ or target $=1$) yield $O = 1$. An MLP, also defined as a feed-forward artificial neural network model, is composed of input, hidden and output layers. The user can specify the number of hidden layers since it is a hyperparameter. In this work, we used 2 hidden layers and 100 neurons in each layer.

A single layer of an MLP can be mathematically described by

$$f(\mathbf{x}) = G\{(2) + (2)[s((1) + (1)\mathbf{x})]\} \tag{1}$$

with bias vectors $\mathbf{b}(1)$, $\mathbf{b}(2)$; weight matrices $\mathbf{W}(1)$, $\mathbf{W}(2)$ and activation functions G and s.

Weight matrices are determined through an optimization process in the training phase. A variety of optimization solvers can be applied, including the Limited-Memory BFGS, the stochastic gradient descent and Adam methods [37]. Our study considers these optimization solvers as hyperparameters denoted as 'lbfgs,' 'sgd,' and 'adam,' respectively.

Given that a is the input for the activation functions, we can define the typical activation functions, G: the identity (which returns a), the logistic sigmoid function (which $1/(1 + exp(-a))$, the hyperbolic tan function (which returns tanh(a)) and the rectified linear unit function (which returns max(0, a)). In this work, these activation functions are denoted as 'identity', 'logistic', 'tanh' and 'relu', respectively.

Aiming to mitigate overfitting, we also used a regularization term labeled 'alpha'. This term controls the magnitude of the weights. By increasing 'alpha', simpler decision boundary curves are produced, while more complex decision boundaries can be derived by reducing' alpha'. More details about the MLP method can be found at [38].

2.7.5. AdaBoost

The AdaBoost algorithm is known for its effectiveness in binary classification problems, as discussed by Breiman (2001). This algorithm leverages boosting iterations of weak learners to create a strong learner. Typically, a weak learner comprises a simple tree structure, denoted as a decision stump, consisting of a root node and two leaves (for binary classification).

AdaBoost constructs a decision stump during each iteration based on a decision metric, such as the Gini impurity measure. In the training phase, the importance of features in creating an effective classifier influences the likelihood of a feature being included in the next evaluation. The algorithm goes through numerous boosting iterations, resulting in multiple decision stumps. The final classification is determined by the most common output across these decision stumps.

The 'n_estimators' hyperparameter allows users to specify the maximum number of evaluations, and the algorithm may terminate earlier upon convergence. Moreover, there is a 'learning_rate' hyperparameter, which can adjust the contribution of each decision stump. However, a high learning rate can lead to premature convergence.

2.8. Statistical analysis

We investigated seven different classifiers to derive machine-learning classifiers and identify whether calves' vocalization audio data belongs to the before-feeding or after-feeding classes. The data set was divided into training and testing using the stratified cross-validation method with k-fold=5. During the training stage, a grid search was carried out considering several hyperparameters, as detailed in Table 3. We used the set of best hyperparameters for each classifier to derive the results discussed in this section.

We split up the audio data sets in different manners to investigate the role of the daytime and the calves' age in the vocalization. For instance, we investigated the audios acquired in July/morning and July/afternoon to verify the impact of the daytime on the results. The same reasoning has been exploited considering the audios acquired in November/morning and November/Afternoon. We also considered all the data obtained in July and November to understand the role of age in the classifiers. Finally, we derived classifiers including all data acquired in the morning/afternoon and July/November. This final investigation is the most important since it can indicate if a single classifier can distinguish the before-feeding and the after-feeding classes.

The performance of these classifiers is compared using accuracy (ACC), F1-score [39], training, and testing times. The computational routine was developed in Python 3 using the scikit-learn library, among others.

3. Results and discussion

Table 4 presents the performance of the classifiers based on all features listed in Table 3, along with their optimal hyperparameters. When analyzing data from distinct times of day (July morning, July afternoon, November morning, and November afternoon), both the k-NN and SVM-RBF classifiers achieved perfect accuracy and F1-scores of 100%. These classifiers also exhibited fast training and testing times, making them efficient for practical applications.

When considering the age of the calves (July and November), the k-NN classifier performed exceptionally well, achieving 98.73% accuracy in July and 100% in November. These findings suggest that the k-NN classifier can effectively handle variations in calves' ages, highlighting its robustness in differentiating between the before-feeding and after-feeding classes. Notably, the k-NN classifier also achieved 100% accuracy across the entire dataset, reinforcing its versatility. The results demonstrate that the k-NN classifier can reliably distinguish between the before-feeding and after-feeding states, independent of the time of day or the age of the calves when all features are considered. However, it is important to acknowledge that deriving the 11 features can be computationally demanding, which may limit real-time application in certain conditions.

From an animal welfare perspective, these findings underscore the

Table 4
Classifiers' performance considering all features - before-feeding and after-feeding classes.

Classifier	Month/day time	ACC	Performance F1-Score	Indicators *Ttrain*(s)	*Ttest*(s)	Best hyperparameters		
	–	–	–	–	–	metric	*k*	weights
	July/morning	100	100	0.0015	0.0080	euclidean	1	uniform
	July/afternoon	100	100	0.0053	0.0036	manhattan	1	uniform
k-NN	November/morning	100	100	0.0016	0.0010	manhattan	13	distance
	November/afternoon	100	100	0.0049	0.0120	euclidean	13	distance
	July	98.73	97.99	0.090	0.016	euclidean	1	uniform
	November	100	100	0.0094	0.0018	euclidean	5	distance
	All data	100	100	0.0130	0.0021	manhattan	1	uniform
	–	–	–	–	–		C	
	July/morning	100	100	0.1221	0.0005		0.20	
	July/afternoon	91.18	89.04	0.1645	0.0005		0.20	
SVM-linear	November/morning	100	100	0.2894	0.0003		0.25	
	November/afternoon	86.96	86.00	2.1309	0.0004		0.50	
	July	92.41	86.19	0.3971	0.0005		0.20	
	November	84.09	78.11	8.6495	0.0007		0.20	
	All data	80.49	76.04	2.9192	0.0010		0.20	
	–	–	–	–	–		C & gamma	
	July/morning	100	100	0.0060	0.0010	2 & 0.01		
	July/afternoon	100	100	0.0045	0.0005	0.9 & 0.02		
SVM-RBF	November/morning	100	100	0.0013	0.0003	0.9 & 0.01		
	November/afternoon	100	100	0.0019	0.0003	0.4 & 0.01		
	July	98.73	97.99	0.0067	0.013	0.9 & 0.01		
	November	68.18	40.54	0.0034	0.0012	2 & 0.01		
	All data	95.12	94.20	0.0115	0.026	2 & 0.01		
	–	–	–	–	–	max_depth & max_features		
	July/morning	98.67	98.54	0.0014	0.0002	7 & sqrt		
	July/afternoon	100	100	0.0010	0.0001	7 & log2		
DT	November/morning	100	100	0.0043	0.0002	8 & auto		
	November/afternoon	95.65	95.52	0.0036	0.0002	5 & log2		
	July	96.20	94.25	0.0012	0.0002	11 & sqrt		
	November	88.64	87.55	0.0014	0.0002	17 & log2		
	All data	98.37	98.15	0.0053	0.0002	19 & auto		
	–	–	–	–	–	max_depth max_features n_estimators		
	July/morning	100	100	0.2313	0.0073	7 7 90		
	July/afternoon	100	100	0.1196	0.0045	7 9 70		
RF	November/morning	95.2	82.1	0.2262	0.0082	7 7 150		
	November/afternoon	95.65	95.52	0.0709	0.0027	12 7 50		
	July	96.20	94.25	0.0067	0.0013	7 7 80		
	November	97.73	97.33	0.2804	0.0081	11 7 100		
	All data	95.93	95.20	0.1699	0.0050	10 7 70		
	–	–	–	–	–	activation alpha solver		
	July/morning	100	100	0.6887	0.0004	identity 0.0001 lbfgs		
	July/afternoon	91.18	85.73	0.4109	0.0012	logistic 0.001 lbfgs		
MLP	November/morning	100	100	0.2411	0.0005	tanh 0.01 adam		
	November/afternoon	91.30	91.15	0.4022	0.0010	tanh 0.0001 lbfgs		
	July	91.14	83.37	0.3470	0.0004	identity 0.01 lbfgs		
	November	88.64	86.03	0.8905	0.0005	logistic 0.001 lbfgs		
	All data	84.55	79.29	0.7549	0.0011	tanh 10 adam		
	–	–	–	–	–	learning_rate & n_estimators		
	July/morning	100	100	0.2313	0.0073	1 & 90		
	July/afternoon	100	100	0.1789	0.0112	0.1 & 80		
AdaBoost	November/morning	100	100	0.2148	0.0165	0.1 & 130		
	November/afternoon	95.65	95.52	0.1241	0.0102	0.1 & 70		
	July	98.73	97.89	0.3204	0.0222	1 & 130		
	November	97.73	97.33	0.3842	0.0230	1 & 110		
	All data	94.31	93.38	0.2128	0.0130	1 & 90		

potential of using animal acoustics as a phenotypic biomarker to assess the well-being of calves. Acoustic features, particularly those related to vocalizations, can be a non-invasive indicator of various aspects of animal health, including feeding behavior and stress levels. This biomarker could contribute to a more comprehensive assessment of calves' welfare at the herd level when combined with other welfare indicators, such as behavioral observations or physiological measures.

From a set of 11 features (Table 2), the RF algorithm derived the most relevant ones for identifying the before-feeding and after-feeding classes according to their importance. The RF algorithm operates by evaluating how each feature contributes to the prediction, helping to identify key features that can be used to differentiate between the two classes. For illustration, Fig. 9 shows the importance of the features according to the RF algorithm for the data acquired in November. In this case, the most important features are $F10$, $F9$, $F3$, and $F2$: three are sound quality-based metrics, while the other is a frequency-domain-based metric. This finding suggests that sound quality metrics significantly distinguish between the before-feeding and after-feeding classes, potentially reflecting important physiological or behavioral changes that occur before and after feeding events. Sound-based metrics are often sensitive to environmental and behavioral factors, and their dominance in the feature set may indicate that audio signals are strongly correlated with the feeding process.

Fig. 10 depicts the correlation between the selected features. One can observe that the diagonal of this plot presents values equal to one, as expected since each feature is perfectly correlated with itself. More importantly, the values out of the diagonal should be smaller than one, demonstrating that the features are not highly correlated with each

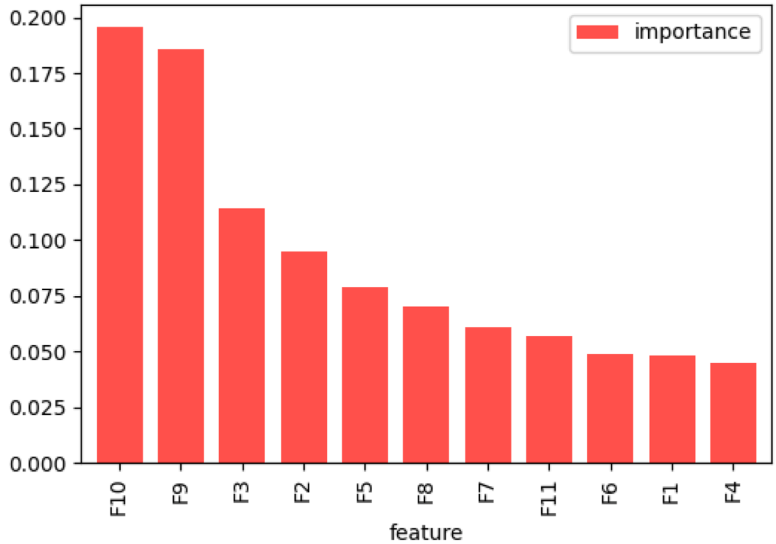

Fig. 9. Features' importance according to Random Forest algorithm for data acquired in November.

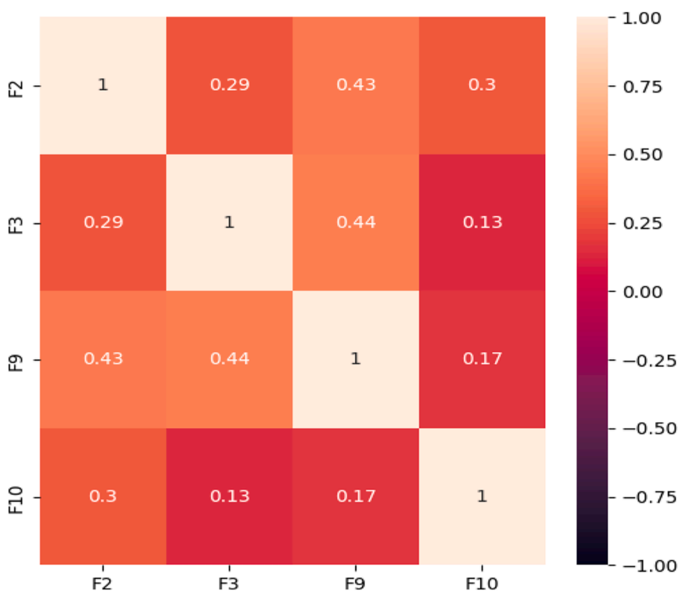

Fig. 10. Features' correlation data acquired in November.

other. The analysis of correlations is critical in feature selection, as highly correlated features can introduce redundancy, which might negatively affect the performance of machine learning models. In this case, the highest correlation values are 0.44 and 0.43 between *F*3/*F*9 and *F*2/*F*9, respectively. These moderate correlations indicate that while the features are not entirely independent, the redundancy between them is adequate. The lack of highly correlated features in the selected set is an important outcome, as it confirms that the RF algorithm has identified a diverse and complementary set of variables. This diversity enhances the robustness of the model, ensuring that it leverages different aspects of the data to improve predictive performance.

Table 5 describes the most important features, considering the different daytime and calves' ages. The RF algorithm selected two frequency-based metrics (*F*1 and *F*2), one time-domain metric (*F*4), and two sound quality-based metrics (*F*3 and *F*10) to serve as input for the classifiers when considering all data. The selection of both frequency-based and time-domain metrics indicates that the RF algorithm recognizes the relevance of the data's spectral and temporal characteristics in differentiating feeding conditions. Frequency-based metrics, such as *F*1 and *F*2, likely capture critical patterns related to the acoustic environment during feeding. The inclusion of the time-domain metric *F*4 further suggests that temporal features, such as the duration and intensity of sounds over time, are important for detecting shifts in feeding behavior. This metric may also capture variations in feeding activity that occur over time, such as changes in the pacing or frequency of feeding events. Notably, the two sound quality-based metrics (*F*3 and *F*10) are of particular interest. The fact that the RF algorithm highlights sound quality features implies that the quality of sound, rather than just the quantity or frequency, can be a sensitive indicator of feeding conditions. Sound quality-based features are often associated with the clarity, richness, or sharpness of sounds, which can vary based on the animal's feeding behavior, the environment, and even the animal's welfare. Therefore, sound quality metrics may provide a higher level of sensitivity to feeding conditions compared to traditional methods, potentially leading to more accurate and timely detection of feeding events.

Table 5
Features selected by random forest algorithm.

Month/ day time	Features
July/ morning	F_2, F_4, F_7
July/ afternoon	F_1, F_2, F_3, F_4, F_{10}
November/ morning	F_1, F_2, F_3, F_7, F_8, F_{11}
November/ afternoon	F_3, F_5, F_9, F_{10}
July	F_1, F_2, F_3, F_4
November	F_2, F_3, F_9, F_{10}
All data	F_1, F_2, F_3, F_4, F_{10}

Table 6 describes the performance of the investigated classifiers considering the set of selected features, as shown in Table 5, and the best hyperparameters. The results indicate that when considering distinct daytimes (July/morning, July/afternoon, November/morning, and November/afternoon), the best classifiers were the k-NN, the SVM-RBF, the DT, and the RF. These four classifiers achieved perfect scores of 100% in both accuracy and F1-score for the conditions of July/morning, July/afternoon, and November/morning. This performance clearly indicates that these classifiers are well-suited for detecting feeding conditions under varying daytime conditions.

The fact that these classifiers scored 100% accuracy and F1-score across different time periods suggests that the selected features—derived from sound quality, frequency, and time-domain metrics—are robust and effective for distinguishing between feeding conditions. The classifiers' consistency across different times of day highlights their generalizability and reliability in real-world

Table 6
Classifiers' performance considering the features selected by RF - Before and After Dairy.

Classifier Month/day time	ACC	Performance F1-Score	Indicators *Ttrain*(s)	Ttest(s)		Best parameters	
–	–	–	–	–	metric	*k*	weights
July/morning	100	100	0.0050	0.0103	euclidean	1	uniform
July/afternoon	100	100	0.0057	0.0084	manhattan	1	uniform
k-NN November/morning	100	100	0.0223	0.0121	euclidian	1	uniform
November/afternoon	95.65	95.52	0.0036	0.025	manhattan	13	distance
July	100	100	0.063	0.0222	euclidean	1	uniform
November	100	100	0.0133	0.0062	euclidean	5	distance
All data	98.37	98.12	0.0117	0.0037	manhattan	3	distance
–	–	–	–	–		C	
July/morning	88.89	80.00	0.0050	0.0103		0.05	
July/afternoon	88.24	83.65	0.4120	0.0018		0.20	
SVM-linear November/morning	95.24	82.05	0.4785	0.0017		0.25	
November/afternoon	82.61	82.58	0.138	0.0014		0.50	
July	88.61	79.90	0.2964	0.0034		0.20	
November	–	–	–	–		–	
All data	–	–	–	–		–	
–	–	–	–	–		C & gamma	
July/morning	100	100	0.0059	0.0017	0.75 & 0.02		
July/afternoon	100	100	0.0028	0.0014	0.75 & 0.01		
SVM-RBF November/morning	100	100	0.0034	0.0014	0.9 & 0.01		
November/afternoon	95.65	95.62	0.0040	0.0018	3 & 0.01		
July	100	100	0.0076	0.036	1 & 0.02		
November	100	100	0.0047	0.0023	0.75 & 0.01		
All data	67.48	40.29	0.0097	0.040	2 & 0.01		
–	–	–	–	–	max depth & max features		
July/morning	100	100	0.0038	0.0013		16 & sqrt	
July/afternoon	100	100	0.0032	0.0015		7 & auto	
DT November/morning	100	100	0.0037	0.0013		4 & sqrt	
November/afternoon	95.65	95.52	0.0044	0.0011		15 & log2	
July	100	100	0.4488	0.0164		16 & sqrt	
November	97.73	97.33	0.0048	0.0019		17 & auto	
All data	75.61	68.94	0.0040	0.0020		4 & auto	
–	–	–	–	–	max depth	max features	n estimators
July/morning	100	100	0.1133	0.0049	7	7	90
July/afternoon	100	100	0.1662	0.0102	7	7	110
RF November/morning	100	100	0.2112	0.0031	7	7	50
November/afternoon	95.65	95.52	0.0791	0.0039	7	13	50
July	100	100	0.4488	0.0164	10	10	90
November	100	100	0.3289	0.0065	7	14	90
All data	95.12	94.29	0.2076	0.0108	7	7	80
–	–	–	–	–	activation	alpha	solver
July/morning	88.89	80.00	0.1361	0.0019	identity	0.0001	lbfgs
July/afternoon	85.29	82.78	0.5790	0.0031	tanh	0.01	lbfgs
MLP November/morning	100	100	0.2112	0.0031	logistic	0.1	lbfgs
November/afternoon	100	100	0.1248	0.0022	identity	0.01	lbfgs
July	88.61	78.62	0.7726	0.0035	identity	1	lbfgs
November	79.55	74.86	0.2691	0.0026	logistic	1	adam
All data	68.29	42.93	0.1011	0.0032	identity	0.0001	adam
–	–	–	–	–	learning rate & estimators		
July/morning	100	100	0.0818	0.0083	1 & 30		
July/afternoon	100	100	0.1563	0.0119	0.1 & 70		
AdaBoost November/morning	85.71	65.95	0.1318	0.0138	0.1 & 50		
November/afternoon	95.65	95.52	0.1141	0.0106	0.01 & 70		
July	100	100	0.1532	0.0084	1 & 30		
November	93.18	91.99	0.2637	0.0219	1 & 130		

applications, where conditions might vary. Additionally, the high performance across different daytimes might reflect the fact that the feeding behaviors and corresponding acoustic signals do not undergo significant changes depending on the time of day, further suggesting that the classifier's decision-making process is based on strong, invariant patterns in the data.

Interestingly, while the testing times for the k-NN, SVM-RBF, DT, and RF classifiers are fast, the training times for the RF classifiers appear more demanding. This difference in computational demands is an important consideration for real-time applications. While RF classifiers may provide robust performance with high accuracy, the need for longer training times might limit their usability in systems that require quick model updates or real-time feedback. This tradeoff between accuracy and computational efficiency is a common challenge in machine learning applications, and in practice, the choice of classifier may depend on the specific requirements of the system. For example, if real-time predictions are a priority, k-NN or SVM-RBF might be preferred due to their faster training and testing times, even though RF offers slightly superior predictive accuracy in certain cases.

When considering distinct calves' ages (July and November), the best classifiers were again the k-NN, SVM-RBF, and RF, all achieving 100% accuracy in both July and November. This suggests that the classifiers can handle variability in the data associated with age, further emphasizing the robustness of the selected features. The ability to maintain high accuracy across different age groups indicates that the model is relatively insensitive to age-related differences in feeding behavior or acoustic signals, which is important for developing a generalized system.

The k-NN classifier achieved 98.37% accuracy when considering all data and the selected features, which is very close to 100%, indicating that even with slight variations in performance across data subsets (e.g., time of day or age), the overall accuracy remains exceptionally high. This is a noteworthy result, as it demonstrates that the k-NN classifier, despite being relatively simple compared to more complex classifiers. This is particularly valuable in scenarios where simplicity, interpretability, and computational efficiency are prioritized over marginal gains in accuracy.

From these results, we can conclude that using a reduced, yet relevant set of features has minimal impact on the accuracy of the best classifiers. This is important because it suggests that simplifying the feature set does not necessarily compromise the model's predictive accuracy, which can be crucial for developing real-time or embedded systems where computational resources might be limited. By focusing on a more compact set of features, the complexity of the calculation can be reduced, allowing for faster inference times. This also facilitates the creation of real-time algorithms for detecting calves' feeding conditions, which could be deployed in field applications where quick decisions are needed, such as in automated livestock monitoring or precision farming. Moreover, the results underscore the importance of feature selection in improving classifier performance. By selecting only the most relevant features, the model is not burdened with redundant or irrelevant data, which can lead to overfitting or unnecessary complexity.

4. Conclusions

It is concluded that the k-NN classifier achieved 100% accuracy using all features and 98.37% with a relevant subset, demonstrating that it is possible to obtain high accuracy with a smaller set of metrics. The SVM-RBF classifier also achieved 100% accuracy with all features, but showed unsatisfactory performance with the reduced subset. The time, frequency and sound quality domain-based metrics used by k-NN proved to be effective in assessing the emotional condition of calves, suggesting that audio data and the proposed methodology can be explored in other animal welfare assessments. Although promising, these tools should be complemented with other indicators for a more comprehensive assessment, pointing out directions to improve the welfare and biological function of animals. Future studies can expand the use of these metrics to other contexts and species, refining their application in animal welfare.

Research regulation

The study was approved by the Ethics Committee on Animal Use (CEUA) of University of São Paulo – Luiz de Queiroz Agriculture School (USP/ESALQ), Piracicaba City, São Paulo State, Brazil, under protocol n. 582,210,722. The study was carried out in this same higher education institution and was in compliance with the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Code availability

Not applicable.

Funding

This work was supported by the São Paulo Research Foundation - FAPESP [grant number 2022/07,442–8] and by the National Council for Scientific and Technological Development - CNPq [grant number 303,884/2021–5].

CRediT authorship contribution statement

Maíra Martins da Silva: Data curation, Conceptualization, Writing – original draft, Validation, Methodology, Funding acquisition. **Robson Mateus Freitas Silveira:** Conceptualization, Writing – original draft. **Gean Gobo da Cruz:** Data curation, Conceptualization. **Karen Airosa Machado de Azevedo:** Conceptualization. **Carla Maris Machado Bittar:** Validation, Supervision, Resources, Writing – original draft. **Iran José Oliveira da Silva:** Formal analysis, Data curation, Conceptualization, Writing – original draft, Funding acquisition.

Declaration of competing interest

None.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References

[1] J. Rushen, Assessing the welfare of dairy cattle, J. Appl. Animal Welf. Sci 4 (2001) 223–234, https://doi.org/10.1207/S15327604JAWS0403_05. Available from.
[2] R. Uehleke, S. Seifert, S. Hüttel, Do animal welfare schemes promote better animal health? An empirical investigation of German pork production, Livest. Sci. 247 (2021) 104481, https://doi.org/10.1016/j.livsci.2021.104481. Available from.
[3] G.D. Rosa, R.D. Palo, R. Serafini, F. Grasso, A. Bragaglio, A. Braghieri, F. Napolitano, Different assessment systems fail to agree on the evaluation of dairy cattle welfare at farm level, Livest. Sci. 229 (2019) 145–149, https://doi.org/10.1016/j.livsci.2019.09.024. Available from.
[4] Nielsen B.L., de Jong I.C., De Vries T.J. The use of feeding behaviour in the assessment of animal welfare. In: Phillips, C. (eds) Nutrition and the Welfare of Farm Animals. Animal Welfare. 2016; 16. Springer, Cham. Available from: https://doi.org/10.1007/978-3-319-27356-3_4.
[5] S. Ivemeyer, J. Preußer, D. Haager, C. Simantke, P. Waldherr, K. Kull, G. Utz, U. Knierim, C. Winckler, Impact of enhanced compared to restricted milk feeding on the behaviour and health of organic dairy calves, Appl. Anim. Behav. Sci. 252 (2022) 105655, https://doi.org/10.1016/j.applanim.2022.105655. Available from.
[6] M. Machado, R.M.F. Silveira, C.M.M. Bittar, C.M.V. Lobos, I.J.O. da Silva, Can the emotional state of calves be noticed by their facial expression and heart rate? Appl. Anim. Behav. Sci. 260 (2023) 105874 https://doi.org/10.1016/j.applanim.2023.105874. Available from.
[7] A.M. de Passillé, J. Rushen, F. Martin, Interpreting the behaviour of calves in an open-field test: a factor analysis, Appl. Anim. Behav. Sci. 45 (3–4) (1995) 201–213, https://doi.org/10.1016/0168-1591(95)00622-Y. Available from.
[8] K. Macmillan, M. Gobikrushanth, M. Colazo, Activity and rumination changes as predictors of calving in primiparous and multiparous holstein cows, Livest. Sci. 260 (2022) 104944, https://doi.org/10.1016/j.livsci.2022.104944. Available from.
[9] C. Gaillard, M. Durand, C. Largouët, J.Y. Dourmad, C. Tallet, Effects of the environment and animal behavior on nutrient requirements for gestating sows: future improvements in precision feeding, Anim. Feed Sci. Technol. 279 (2021) 115034, https://doi.org/10.1016/j.anifeedsci.2021.115034. Available from.
[10] M. Schnaider, M. Heidemann, A. Silva, C. Taconeli, C. Molento, Vocalization and other behaviors as indicators of emotional valence: the case of cow-calf separation and reunion in beef cattle, J. Veterin. Behav 49 (2022) 28–35, https://doi.org/10.1016/j.jveb.2021.11.011. Available from.
[11] H. Pedersen, K. Malm, Cross-disciplinary method development for assessing dog welfare in canine-assisted pedagogical work: a pilot study, J. Appl. Animal Welfare Sci (2023) 1–14, https://doi.org/10.1080/10888705.2023.2211205. Available from.
[12] G. Smith-Vidaurre, V. Perez-Marrufo, T.F. Wright, Individual vocal signatures show reduced complexity following invasion, Anim. Behav. 179 (2021) 15–39, https://doi.org/10.1016/j.anbehav.2021.06.020. Available from.
[13] N. Sharma, V. Prakash, S. Kohshima, R. Sukumar, Asian elephants modulate their vocalizations when disturbed, Anim. Behav. 160 (2020) 99–111, https://doi.org/10.1016/j.anbehav.2019.12.004. Available from.
[14] M.P. de la Torre, E.F. Briefer, T. Reader, A.G. McElligott, Acoustic analysis of cattle (bos taurus) mother–offspring contact calls from a source–filter theory perspective, Appl. Anim. Behav. Sci. 163 (2015) 58–68, https://doi.org/10.1016/j.applanim.2014.11.017. Available from.
[15] Shorten P., Hunter L. Acoustic sensors for automated detection of cow vocalization duration and type. Comput. Electron. Agric. 2023; 208: 107760. Available from: https://doi.org/10.1016/j.compag.2023.107760.

[16] J.M. Siegford, J.P. Steibel, J. Han, M. Benjamin, T. Brown-Brandl, J.R. Dórea, D. Morris, T. Norton, E. Psota, G.J Rosa, The quest to develop automated systems for monitoring animal behavior, Appl. Anim. Behav. Sci. 265 (2023) 106000.
[17] D. Gavojdian, T. Lazebnik, M. Mincu, A. Oren, I. Nicolae, A. Zamansky, Bovinetalk: machine learning for vocalization analysis of dairy cattle under negative affective states, Front. Vet. Sci. 11 (2024), https://doi.org/10.3389/fvets.2024.1357109. Available from.
[18] D.H. Jung, N.Y. Kim, S.H. Moon, C. Jhin, H.J. Kim, J.S. Yang, H.S. Kim, T.S. Lee, J. Y. Lee, S.H. Park, Deep learning-based cattle vocal classification model and real-time livestock monitoring system with noise filtering, Animals 11 (2) (2021) 357, https://doi.org/10.3390/ani11020357. Available from.
[19] Y. Peng, Kondo N Wulandari, T. Fujiura, T. Suzuki, H. Yoshioka, E Itoyama, Japanese black cattle call patterns classification using multiple acoustic features and machine learning models, Comput. Electron. Agric. 204 (2023) 107568, https://doi.org/10.1016/j.compag.2022.107568. Available from:.
[20] H. Fastl, E. Zwicker, Psychoacoustics: Facts and Models, Springer, Berlin Heidelberg, 2007, https://doi.org/10.1007/978-3-540-68888-4. URL.
[21] Green Forge Coop, Mosqito, 2022. Available from: https://zenodo.org/record/5284054. doi:10.5281/ZENODO.5284054.
[22] J.A. Mosquera-Sánchez, W. Desmet, L.P.R. de Oliveira, A multichannel amplitude and relative-phase controller for active sound quality control, Mech. Syst. Signal. Process. 88 (2017) 145–165, https://doi.org/10.1016/j.ymssp.2016.10.036. Available from.
[23] American National Standards Institute, Committee On Bioacoustics. American National Standard Psychoacoustical Terminology, American National Standards Institute New York, ANSI, American National Standards Institute, NY, 1973. Available from, https://books.google.com.br/books?id=5Fo0GQAACAAJ.
[24] ISO 532-1:2017, Acoustics — Methods for Calculating Loudness — Part 1: Zwicker method, Standard, International Organization for Standardization, Geneva, CH, 2017. Available from, https://www.en-standard.eu/bs-iso-532-1-2017-acoustics-methods-for-calculating-loudness-zwicker-method/?msclkid=2a6450470d2c1eb2903a36c29968a972.
[25] R.W.P. Daniel, Psychoacoustical roughness: implementation of an optimized model, Acta Acustica 83 (1997) 113–123.
[26] DIN 45692:2009-08, Messtechnische Simulation der Hörempfindung Schärfe, Beuth Verlag GmbH, Standard, 2017, https://doi.org/10.31030/1521326. Available from:.
[27] U.M. Khaire, R. Dhanalakshmi, Stability of feature selection algorithm: a review, J. King Saud Uni. - Comput. Inform. Sci 34 (2022) 1060–1073, https://doi.org/10.1016/j.jksuci.2019.06.012. Available from.
[28] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324. Available from.
[29] R. Genuer, J.M. Poggi, C. Tuleau-Malot, Variable selection using random forests, Pattern. Recognit. Lett. 31 (2010) 2225–2236, https://doi.org/10.1016/j.patrec.2010.03.014. Available from.
[30] X. Li, W. Chen, Q. Zhang, L. Wu, Building auto-encoder intrusion detection system based on random forest feature selection, Comput. Secur. 95 (2020) 101851, https://doi.org/10.1016/j.cose.2020.101851. Available from.
[31] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEe Trans. Inf. Theory. 13 (1967) 21–27, https://doi.org/10.1109/TIT.1967.1053964. Available from.
[32] Vapnik V. The support vector method of function estimation, Springer U.S., Boston, M.A. 1998: 55–85. Available from: https://doi.org/10.1007/978-1-4615-5703-6.
[33] A. Christmann, I. Steinwart, Support Vector Machines, Information Science and Statistics, Springer, New York, NY, 2008. Available from, https://link.springer.com/book/10.1007/978-0-387-77242-4.
[34] S. Safavian, D. Landgrebe, A survey of decision tree classifier methodology, IEEE Trans. Syst. Man Cybern. 21 (1991) 660–674, https://doi.org/10.1109/21.97458. Available from.
[35] D.E. Rumelhart, J.L. McClelland, Learning internal representations by error propagation. Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations, MIT Press, 1987, pp. 318–362.
[36] R.E. Schapire, Explaining AdaBoost, in: B. Schölkopf, Z. Luo, V. Vovk (Eds.), Empirical Inference, Springer, Berlin, Heidelberg, 2013, https://doi.org/10.1007/978-3-642-41136-6_5. Available from.
[37] D.P. Kingma, Ba J. Adam, A method for stochastic optimization, in: International Conference on Learning Representations, 2014, https://doi.org/10.48550/arXiv.1412.6980. Available from:.
[38] T. Windeatt, Accuracy/diversity and ensemble MLP classifier design, IEEe Trans. Neural Netw. 17 (2016) 1194–1211, https://doi.org/10.1007/10.1109/TNN.2006.875979. Available from.
[39] A.A. Taha, A. Hanbury, Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool, BMC. Med. ImAging 15 (2015) 29, https://doi.org/10.1186/s12880-015-0068-x. Available from.