A noise-robust acoustic method for recognizing foraging activities of grazing cattle

Luciano S. Martinez-Rau[a,b,1,*], José O. Chelotti[a,c], Mariano Ferrero[a], Julio R. Galli[d,e], Santiago A. Utsumi[f,g], Alejandra M. Planisich[e], H. Leonardo Rufiner[a,h], Leonardo L. Giovanini[a]

[a]*Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i), FICH-UNL/CONICET, Argentina*
[b]*Department of Computer and Electrical Engineering, Mid Sweden University, Sundsvall, Sweden*
[c]*TERRA Teaching and Research Center, University of Liège, Gembloux Agro-Bio Tech (ULiège-GxABT), 5030 Gembloux, Belgium*
[d]*Instituto de Investigaciones en Ciencias Agrarias de Rosario, IICAR, UNR-CONICET, Argentina*
[e]*Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Argentina*
[f]*W.K. Kellogg Biological Station and Department of Animal Science, Michigan State University, United States*
[g]*Department of Animal and Range Science, New Mexico State University, United States*
[h]*Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Argentina*

Abstract

Farmers must continuously improve their livestock production systems to remain competitive in the growing dairy market. Precision livestock farming technologies provide individualized monitoring of animals on commercial farms, optimizing livestock production. Continuous acoustic monitoring is a widely accepted sensing technique used to estimate the daily rumination and

[*]Corresponding author

Email addresses: luciano.martinezrau@miun.se (Luciano S. Martinez-Rau), jchelotti@sinc.unl.edu.ar (José O. Chelotti), mferrero@sinc.unl.edu.ar (Mariano Ferrero), jgalli@lidernet.com.ar (Julio R. Galli), sutsumi@nmsu.edu (Santiago A. Utsumi), aplanisich@gmail.com (Alejandra M. Planisich), lrufiner@sinc.unl.edu.ar (H. Leonardo Rufiner), lgiovanini@sinc.unl.edu.ar (Leonardo L. Giovanini)

[1]Present address: Department of Computer and Electrical Engineering, Mid Sweden University, Holmgatan, PO Box 85230 Sundsvall, Västernorrland, Sweden. (+46) 10-172 65 08

grazing time budget of free-ranging cattle. However, typical environmental and natural noises on pastures noticeably affect the performance limiting the practical application of current acoustic methods. In this study, we present the operating principle and generalization capability of an acoustic method called Noise-Robust Foraging Activity Recognizer (NRFAR). The proposed method determines foraging activity bouts by analyzing fixed-length segments of identified jaw movement events produced during grazing and rumination. The additive noise robustness of the NRFAR was evaluated for several signal-to-noise ratios using stationary Gaussian white noise and four different nonstationary natural noise sources. In noiseless conditions, NRFAR reached an average balanced accuracy of 86.4%, outperforming two previous acoustic methods by more than 7.5%. Furthermore, NRFAR performed better than previous acoustic methods in 77 of 80 evaluated noisy scenarios (53 cases with $p<0.05$). NRFAR has been shown to be effective in harsh free-ranging environments and could be used as a reliable solution to improve pasture management and monitor the health and welfare of dairy cows. The instrumentation and computational algorithms presented in this publication are protected by a pending patent application: AR P20220100910.
Web demo available at: `https://sinc.unl.edu.ar/web-demo/nrfar`

Keywords: Acoustic monitoring, foraging behavior, machine learning, noise robustness, pattern recognition, precision livestock farming.

---

1. Introduction

The new and diverse precision livestock farming tools and applications significantly reduce farm labor (Lovarelli et al., 2020; Tzanidakis et al., 2023). Precision livestock farming solutions allow individualized monitoring of animals to optimize herd management in most production systems (Michie et al., 2020). Monitoring the feeding behavior of livestock can provide valuable insights into animal welfare, including their nutrition, health, and performance (Banhazi et al., 2012; Garcia et al., 2020). Changes in grazing patterns, periodicity, and duration can be used to inform the management of pasture allocation (Connor, 2015), while changes in ruminant diets signal anxiety (Bristow and Holmes, 2007) or stress (Abeni and Galli, 2017; Schirmann et al., 2009), as well as an early indicator of diseases (Osei-Amponsah et al., 2020; Paudyal et al., 2018), rumen health (Beauchemin, 2018, 1991), and the onset of parturition (Kovács et al., 2017; Pahl et al., 2014) and estrus (Dolecheck et al., 2015; Pahl et al., 2015).

Free-ranging cattle spend 40-80% of their daily time budget on grazing and rumination activities (Kilgour, 2012; Phillips, 2002). A grazing bout involves the process of searching, apprehending, chewing, and swallowing herbage and is characterized by a sequence of ingestive jaw movement (JM) events associated with chews, bites, and composite chew-bites, without a fixed or predefined order. A bite event involves the apprehending and severing of the herbage, a chew event involves crushing, grinding, and processing previously gathered herbage, and a chew-bite event occurs when herbage is apprehended, severed, and comminuted in the same JM (Ungar and Rutter, 2006). Rumination is defined as the period of time during which an animal repeatedly regurgitates previously ingested food (cud) from its rumen, then chews the cud for 40-60 s, and re-swallows it. Rumination bouts begin with the first regurgitation and end with the last swallow (Beauchemin, 2018; Galli et al., 2020). Grazing and rumination involve JM-events taken at rates of 0.75-1.20 JM per second. Changes in the type and sequence of distinctive JM-events can be aggregated over time to determine the sequence and duration of foraging activities (Andriamandroso et al., 2016).

Feeding activity monitoring of cattle has primarily been approached through the use of noninvasive wearable sensors, including nose-band pressure, inertial measurement units, and microphone systems (Benos et al., 2021; Stygar et al., 2021). Each sensing technique has its advantages and disadvantages depending on the environment and application. Current nose-band pressure sensors are combined with accelerometers to log data from JMs. Raw data are analyzed by software to determine foraging behaviors and provide specific information associated with them (Steinmetz et al., 2020; Werner et al., 2018). Human intervention is required to process the data recorded on a computer, making it not scalable for use on commercial farms (Riaboff et al., 2022). Sensors based on inertial measurement units are widely used to recognize multiple behaviors such as grazing, rumination, posture, and locomotion (Aquilani et al., 2022; Chapa et al., 2020). Although accelerometer-based sensors are typically used in indoor environments (Balasso et al., 2021; Lovarelli et al., 2022; Wu et al., 2022), their use in outdoor environments has increased in the last years (Arablouei et al., 2023; Cabezas et al., 2022; Wang et al., 2023). One major drawback of inertial measurement units is their limited ability to estimate herbage intake in grazing (Wilkinson et al., 2020). Furthermore, the reliability of these sensors is highly dependent on their precise location, orientation, and secure clamping, making reproducing results difficult (Kamminga et al., 2018; Li et al., 2021a). For this reason, acoustic sensors are preferred over former sensors for monitoring the foraging and rumination behaviors of cattle outdoors. Head-placed microphones allow to collect detailed information on ingestive behaviors (Laca et al., 1992). Acoustic sensors are used to automatically recognize JM-events (Ferrero et al., 2023; Li et al., 2021b), estimate rumination and grazing bouts (Vanrell et al., 2018), distinguish between plants and feedstuffs eaten (Galli et al., 2020; Milone et al., 2012), and estimate differences in dry matter intake (Galli et al., 2018). Despite progress, the evaluation of the generalization capabilities of motion- and acoustic-based monitoring solutions are limited due to the scarcity of public and standardized datasets (Martinez-Rau et al., 2023b; Vanrell et al., 2020). As a result, there is room for improving the confidence in the acoustic monitoring of free-grazing cattle.

In recent years, acoustic methods have been developed for recognizing foraging activities. Vanrell et al. (2018) developed a method based on the analysis of the autocorrelation of the acoustic signal for the recognition of foraging activities. This method operates offline because it requires storing several hours of acoustic recording to discover the regularity patterns in the signal. Offline operation introduces considerable delays in making inferences about foraging activities, which could be critical for the early detection of estrus (Allrich, 1993; Reith and Hoy, 2012). The Bottom-Up Foraging Activity Recognizer (BUFAR) developed by Chelotti et al. (2020) operates online, meaning that the incoming digital acoustic signal is processed as it is generated. BUFAR analyzes 5-min segments of identified JM-events to determine grazing and rumination bouts, outperforming the method of Vanrell et al. (2018) with significantly lower computational costs. More recently, Chelotti et al. (2023) proposed an online Jaw Movement segment-based Foraging Activity Recognizer (JMFAR) that outperforms BUFAR. This is achieved by analyzing information from JMs that have been detected but not yet classified, enabling the recognition of grazing and rumination bouts. However, BUFAR and JMFAR exhibited an average confusion of approximately 10% between grazing and rumination, indicating a need for improvement in the recognition of these activities. Another significant drawback of these methods is their limited capability to recognize foraging activities in diverse operational conditions or in the presence of noise (Chelotti et al., 2023). To be useful in practical applications, acoustic foraging recognizers must work properly even under adverse noise and mismatch conditions, where variations in recording settings and environmental conditions are common. Additionally, low computation demands make them feasible for embedding in an acoustic monitoring sensor (Rehman et al., 2014). Motivated by this need, this paper describes in detail the operation, noise robustness and generalization capability of an alternative acoustic method for the recognition of grazing and rumination activities in free-range cattle. The proposed method involves a noise-robust methodology for detecting and classifying JM-events used to recognize foraging activities. In a recent proof-of-concept study, the implementation of the proposed method was assessed for real-time operation on a low-power microcontroller (Martinez-Rau et al., 2023a). The main contributions of this work are: (*i*) present an online acoustic method for estimating grazing and rumination bouts in cattle, characterized by a low computational cost. It classifies four classes of JM-events, which are analyzed in fixed-length segments to delimit activity bouts. (*ii*) The proposed method recognizes foraging activities in free-range environments under different and adverse acoustic conditions, using a robust JM event recognizer that is capable of identifying JM events under quiet and noisy operating conditions. (*iii*) Artificial noise sounds of different natures are used to simulate multiple adverse acoustic scenarios in controlled experiments (Skowronski and Harris, 2004).

The rest of this paper is organized as follows: Section 2 briefly describes a system for recognizing foraging activities and analyzes the operation and limitations of BUFAR and JMFAR. Section 3 introduces the proposed algorithm. This section also outlines the acquisition of the datasets, the experimental setup, and the performance metric used to validate the algorithms. The comparative results for the proposed and former algorithms are shown in Section 4. Section 5 explains and discusses the results of this work. Finally, the main conclusions follow in Section 6.

2. Current acoustic method analysis

In this section, a brief description and limitations of two current acoustic foraging activity recognizers, called BUFAR and JMFAR, are presented. Both methods follow the general structure of a typical pattern recognition system (Bishop, 2006; Martínez Rau et al., 2020) and can be represented by the common block diagram shown in Figure 1. A foraging activity recognizer can be analyzed at three temporal levels: bottom, middle, and top. These levels operate on the millisecond, second, and minute scales, respectively. A JM-event recognizer operates at both the bottom and middle levels to detect and classify different types of JM-events. The input digitized sound is conditioned, processed, and down-sampled using signal processing techniques to reduce the computational cost of the middle and top levels. The processed signals are used at the middle level for a JM detector based on adaptive thresholds. When a JM is detected, a set of distinctive JM features are computed over a time window centered on the JM. Finally, a machine learning model uses the extracted set of JM features to classify the JM-event with a corresponding timestamp. The middle level provides JM information to the top level. The top level buffers the JM information in nonoverlapping segments of 5-min duration. For each segment, a set of activity features is computed to serve as input to a classifier that determines the activity performed by the animal. Segments of 5-min duration store sufficient JM information data in the buffer to generate a confidence set of activity features, without significantly affecting the correct delimitation of foraging activity. Five-min duration agrees with the optimal segment duration value found in two previous studies (Chelotti et al., 2020; Rook and Huckle, 1997).

As previously mentioned, the type and sequence of distinctive JM-events can be analyzed to recognize foraging activities. Inspired by this, the BUFAR method uses a real-time JM-event recognizer developed by Chelotti et al. (2018) to detect and classify JM-events into three different classes: chews, bites, and chew-bites. The JM information comprises the timestamps and classes of the JM-events (see the top level of Figure 1). The JM information is analyzed in nonoverlapping 5-min segments. For each segment, a set of four statistical activity features is extracted, including (i) the rate of JM-events, and the proportion of the JM-events corresponding to the classes (ii) chew, (iii) bite, and (iv) chew-bite. These features are then used for a multilayer perceptron (MLP) classifier (Bishop, 2006) to determine the activities performed. However, inherent detection and classification errors of JM-events may cause misclassification of foraging activities. A more detailed description of BUFAR is provided by Chelotti et al. (2020).

The JMFAR method partially overcomes the limitation of BUFAR because it does not compute information about the JM-events classes. Instead, JMFAR analyses nonoverlapping 5-min segments from the detected JM. The same JM-event recognizer used in BUFAR is also used in JMFAR to compute the JM information. JM information consists of the signal used to detect the JM, the timestamps of the detected JM, and the extracted set of JM features. JM information, analyzed in segments, is employed to compute a set of activity features. The set of twenty-one statistical, temporal, and spectral features serves as input to an MLP classifier that determines the corresponding activity performed. A more detailed description of JMFAR is provided by Chelotti et al. (2023).

3. Material and Methods

3.1. Proposed foraging activity recognizer

The high sensitivity to noise of the JM-event recognizer used in BUFAR and JMFAR could lead to the misclassification of foraging activities. When the input audio signal is contaminated by noise, the accurate detection of JM, the computation of JM features, and the classification of JM-events are significantly impacted (Martinez-Rau et al., 2022). As a result, the noise directly impacts the JM information and consequently affects the computation of the set of activity features, leading to possible misclassification of activity. The activity recognition in quiet and noise conditions can be improved by using a better JM-event recognizer. This work proposes an online method called *Noise-Robust Foraging Activity Recognizer* (NRFAR). NRFAR introduces the use of the Chew-Bite Energy Based Algorithm (CBEBA) for the recognition of JM-events in diverse acoustic environments (Martinez-Rau et al., 2022). Similar to BUFAR, NRFAR analyses nonoverlapping segments of 5-min duration of recognized JM-events classes for the subsequent classification of foraging activities.

The CBEBA is a real-time pattern recognition method, able to distinguish individualized JM-events in terms of four different classes: *rumination-chews*, *grazing-chews*, *bites*, and *chew-bites*. It outperforms previously published methods in both the detection and classification of JM-events in both noiseless and noisy environments. Briefly, the implementation of CBEBA can be divided into four successive stages (Figure 1):

- Signal processor: the digitized input audio signal undergoes a second-order Butterworth band-pass filter to isolate the JM frequency range. The filtered signal is then squared to obtain the instantaneous power signal. To reduce computation, the former signal is used to compute two additional down-sampled signals: a decimated envelope signal and an energy signal calculated by frames.

- JM detector: the presence of a peak in the envelope signal above a time-varying threshold indicates the detection of a candidate JM-event.

When this indication occurs, the energy signal is compared with another adaptive threshold to delimit the boundaries of the candidate JM-event. The time-varying threshold considers short-timescale anatomical and behavioral characteristics of the animal, as well as, long-timescale variable feeding patterns. The adaptive threshold changes according to the background noise floor level on the acoustic signals.

- JM feature extractor: both delimited signals are used to extract a set of five robust JM features. These heuristic features are the duration, energy, symmetry of the envelope, zero-cross derivative of the envelope, and accumulated absolute value of the derivative of the envelope. To avoid the detection of a false-positive JM-event, it is classified only if the duration and energy are in a predefined range.

- JM classifier: A multilayer perceptron (MLP) classifier determines the class of the JM-event. The structure of the MLP classifier is 5-6-4 neurons in the input, hidden, and output layers. Furthermore, the adaptive thresholds are tuned based on the signal-to-noise ratio (SNR) estimated over the envelope and energy signals.

A more detailed description of CBEBA is provided by Martinez-Rau et al. (2022).

The top level of the proposed NRFAR processes the JM information provided by the JM-event recognizer CBEBA in nonoverlapping 5-min segments to establish the corresponding foraging activity. The JM information is the recognized JM-events, along with their respective timestamps. Each segment of JM information is used to generate a set of five activity features: (*i*) the rate of JM-events, and the proportion of the JM-events corresponding to the classes (*ii*) rumination-chew, (*iii*) grazing-chew, (*iv*) bite, and (*v*) chew-bite). The set of extracted activity features feeds an MLP activity classifier to label the foraging activity in terms of *grazing*, *rumination* and *other*. The classified label outputs are smoothed using a third-order median filter to reduce the possible misclassifications of the recognized activity along consecutive segments. Figure 2 shows an example of the proper operation of the smoothing filter.

3.2. Datasets description

This study uses two datasets to evaluate the algorithms under matched and mismatched conditions. The first one (referred to as DS1) is a public

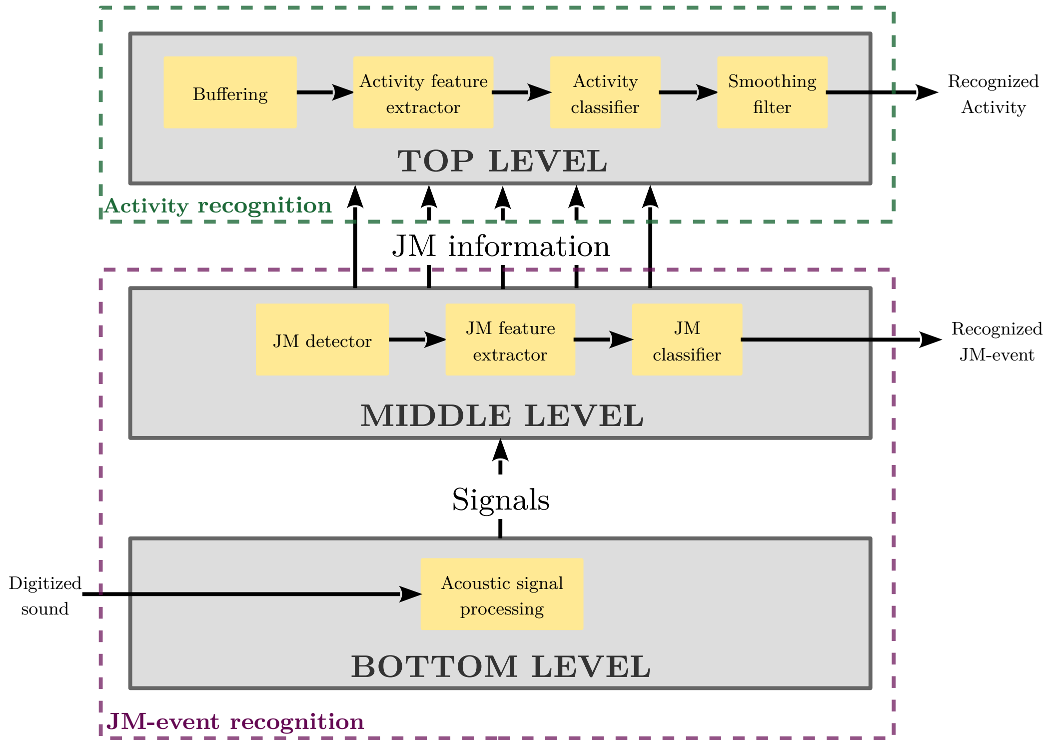

Figure 1: General block diagram of the BUFAR, JMFAR, and the proposed NRFAR methods divided into temporal scales. The JM information transferred to the top level is different in each method.

dataset collected at the Michigan State University's Pasture Dairy Research Center (W.K. Kellogg Biological Station, Hickory Corners, MI, USA) from July 31 to August 19, 2014 (Martinez-Rau et al., 2023b). In this dataset, the cows were handled using a pasture-based robotic milking system with unrestricted cow traffic, as described by Watt et al. (2015). Cows were voluntarily milked 3.0 ± 1.0 times per day using two Lely A3-Robotic milking units (Lely Industries NV, Maassluis, The Netherlands). Inside the dairy barn, the dairy cows were fed a grain-based concentrate. Cows had 24-h access to grazing paddocks with a predominance of either tall fescue (Lolium arundinacea), orchardgrass (Dactylis glomerata) and white clover (Trifolium repens), or perennial ryegrass (Lolium perenne) and white clover. From a herd of 146 lactating high-producing multiparous Holstein cows, 5 animals were selected to record acoustic signals and to monitor their foraging behavior in a noninvasive manner continuously. Specific information on grain-based

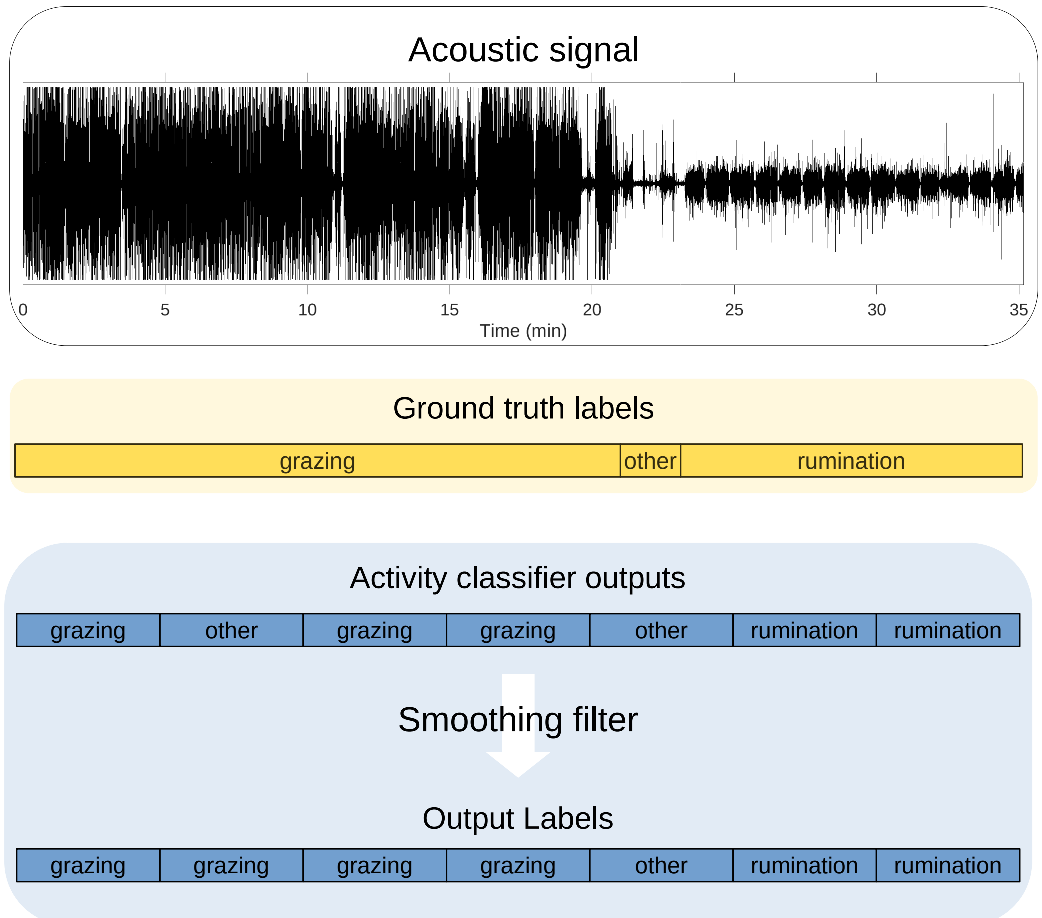

Figure 2: Example of recognized 5-min segments (blue color) compared to the ground truth reference labels (yellow color). The classified activity label assigned to every segment enters the smoothing filter to generate the output label of NRFAR.

concentrate, pasture on paddocks, and individualized characteristics of the 5 dairy cows are given in Martinez-Rau et al. (2023b).

Individualized 24-h of continuous acoustic recordings were obtained on 6 nonconsecutive days. The foraging behavior of the 5 dairy cows was recorded by 5 independent recording systems that were rotated daily, according to a 5 x 5 Latin-square design. This setup was allowed to verify differences in sound signals associated with a particular recording system, cow, or experimental day. The recording systems were randomly assigned to the cows on the first day. On the sixth day, the same order was used to reassign the recording systems to the cows. No prior training was considered necessary for the use of the recording systems before the start of the study.

Each recording system comprised two directional electret microphones connected to a digital recorder (Sony Digital ICD-PX312, Sony, San Diego, CA, USA). The digital recorder was protected in a weatherproof case (1015 Micron Case Series, Pelican Products, Torrance, CA, USA) and mounted on the top side of a halter neck strap (Figure 3). One microphone was positioned facing outwards in a noninvasive manner and pressed against the forehead of the cow to collect the sounds produced by the animal. The other microphone was placed facing inwards to capture the vibrations transmitted through the bones. The microphones kept the intended location using rubber foam and an elastic headband attached to the halter. This design prevents microphone movements, reduces wind noise, and protects microphones from friction and scratches (Milone et al., 2012). The digital recorders saved the audio recordings in MP3 format (Brandenburg and Stoll, 1994) with a 16-bit resolution at a sampling rate of 44.1 kHz. Each channel of the stereo MP3 files corresponds to the microphone facing inwards and outwards. In this study, the stereo MP3 files were converted to mono WAV files, and only those mono WAV files corresponding to the microphones facing inwards were used because they provide a better sound quality of the foraging activities with less presence of external noise sounds.

The second dataset (referred to as DS2) was collected at the Campo Experimental J.F. Villarino (Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Zavalla, Argentina) on August 1, 2022. The protocol used for the experiment has been evaluated and approved by the Committee on Ethical Use of Animals for Research of the Universidad Nacional de Rosario. This intensified pastoral-based dairy farm has a herd of 140-165 milking cows, with an individual production of 24-27 l of milk daily. Three 4-year-old lactating Holstein cows weighing 570-600 kg were selected for this experiment. The experimental cows were allowed to graze freely within a fully enclosed paddock measuring approximately 60 by 20 m, and they had continuous access to a watering trough. The paddock area was covered with naturalized perennial grasses (with a dominance of Cynodon sp., Lolium sp., and Festuca sp.). All cows were tamed and trained in the experimental routine before the experiment. Each animal was equipped with an acquisition data device consisting of an external microphone (IP57 100 mm, -42 $\pm$ 3 dB, SNR 57 dB) plugged via a 3.5 mm jack to a Moto G6 smartphone (Moto G6 smartphone specification, 2018). The smartphones were fixed inside plastic

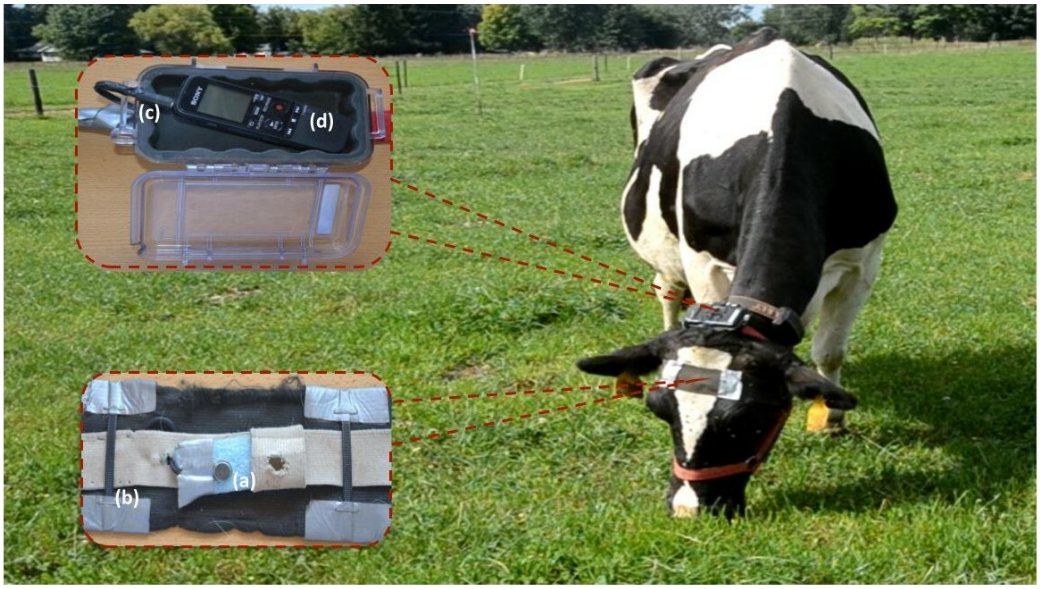

Figure 3: Recording system used to record the acoustic signals composed of microphones (a) that are covered by rubber foam and an elastic headband (b), which are wired and plugged (c) to a digital recorder placed inside a waterproof case (d) attached to a neck halter. Figure extracted from Martinez-Rau et al. (2023b)

boxes secured to prevent unintended internal movements. As in DS1, microphones were located on the cow's forehead and boxes were mounted to the top sides of halter neck straps (Figure 3). Audio recordings were stored in the Moto G6 using high-efficiency advanced audio coding (Bosi et al., 1997) with a bit rate of 128 kbps and a sampling rate of 44.1 kHz, single channel (mono).

Each fieldwork employed an experienced animal handler who had extensive knowledge of data collection on animal behavior. The handler observed the animals for blocks of approximately 5 min per h during daylight hours to ensure the proper placement and positioning of recording systems on the cows. The observations were conducted from a distance to minimize potential disruptions in animal behavior. The handler registered the observed foraging activities and other relevant parameters in a logbook. The ground truth identification of foraging activities was carried out by two experts with long experience in foraging behavior scouting and in the digital analysis of acoustic signals. An expert listened to the audio recordings to identify, delimit, and label the activities guided by the logbook. The results were doubleinspected and verified by the other expert. Although the experts agreed on all label assignments, there were some small differences in the start or end times of certain labels. In these cases, the experts collaborated to reach a mutual agreement on the labels. Activity blocks were labeled as *grazing*, *rumination*, or *other* (see Figure 2).

Additionally, this study uses audio clips from two open acoustic datasets to evaluate the algorithms under adverse conditions. The selection process for the useful audio clips is shown in Figure 4. The first dataset is a labeled collection of 2000 environmental audio clips of 5 s duration, organized into 50 categories with 40 audio clips per category (Piczak, 2015). The second dataset is a multilabeled collection of 51,197 audio clips, with a mean duration of 7.6 s, unequally distributed into 200 categories (Fonseca et al., 2022). To represent environmental and natural noises commonly found in field pastures, the categories *"aeroplane"*, *"chirping birds"*, *"cow"*, *"crickets"*, *"engine"*, *"insects"*, *"rain"*, *"thunderstorm"*, and *"wind"* from the first dataset and *"aircraft"*, *"animal"*, *"bird vocalisation and birds call and bird song"*, *"car passing by"*, *"cowbell"*, *"cricket"*, *"engine"*, *"fixed-wing aircraft and aeroplane"*, *"frog"*, *"insect"*, *"livestock and farm animals and working animals"*, *"rain"*, *"raindrop"*, *"thunder"*, and *"wind"* from the second dataset were selected. These categories were grouped into four exclusive sets according to their nature as follows:

1. Animals = {*animal*, *bird vocalisation and birds call and bird song*, *chirping birds*, *cow*, *cowbell*, *cricket*, *crickets*, *frog*, *insect*, *insects*, *livestock and farm animals and working animals*}
2. Vehicles = {*aeroplane*, *aircraft*, *car passing by*, *engine*, *fixed-wing aircraft and aeroplane*}
3. Weather = {*rain*, *raindrop*, *thunder*, *thunderstorm*, *wind*}
4. Mixture = {*Animals*, *Vehicles*, *Weather*}

The audio clips of the sets were listened to by the experts, and those that did not correspond with possible field pasture conditions were discarded. Overall, 3042 useful audio clips lasting 13.1 h were identified. For reproducibility, a list of selected audio clips is available as Supplementary Material.

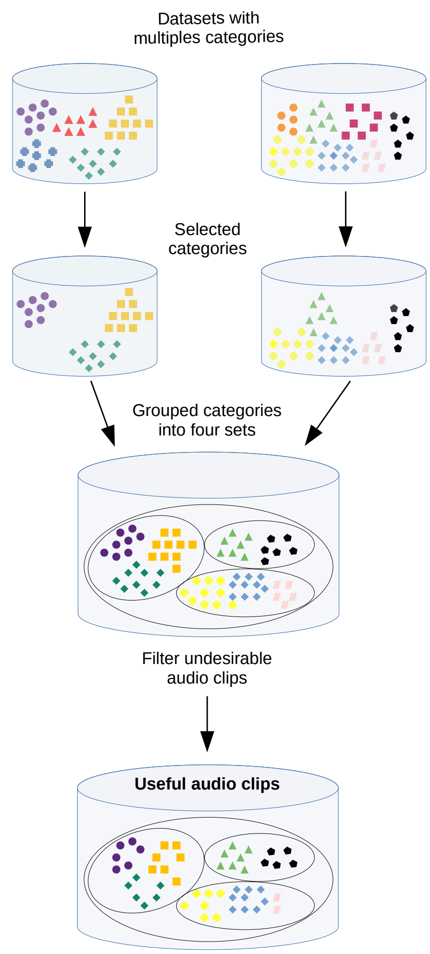

Figure 4: Top-down scheme for selecting useful audio clips.

3.3. *Numerical experiments setup*

3.3.1. *Experiment 1: performance evaluation under matched conditions*

In the initial experiment, the NRFAR performance was evaluated using DS1. This experiment assessed NRFAR effectiveness under consistent conditions, including the same animals, recording devices, and field conditions. NRFAR was coded, trained, and tested in Matlab R2019b (MathWorks, Natick, MA, USA), following a stratified 5-fold cross-validation scheme. A set of 349.4 h of outdoor audio recordings of DS1, composed of 50.5% *grazing*, 34.9% *rumination*, and 14.6% of *other* activities was used. The imbalanced distribution of classes is consistent with typical cattle behavior (Kilgour, 2012). Therefore, the test data were not balanced by class. From all available training data in each fold, 30% of the majority class (*grazing*) was randomly undersampled and 100% of the minority class (*other*) was synthetically oversampled (He et al., 2008), to generate a balanced dataset for training (35.6% *grazing*, 35.1% *rumination*, and 29.3% of *other* activities). The activity classifier is an MLP neural network formed by five input neurons (number of input features), one hidden layer, and three output neurons (number of output labels corresponding to the activity class). The activation functions used by the hidden and output layers are the hyperbolic tangent sigmoid and softmax transfer functions, respectively. During the MLP training phase, the scaled conjugate gradient backpropagation algorithm was used to find the optimal weight and bias of the network and optimize the MLP classifier's hyperparameters. The two hyper-parameters' learning rate and number of neurons in the hidden layer were fitted using a grid-search method. The learning rate was evaluated at values of 0.1, 0.01, 0.001, and 0.0001, whereas the number of neurons was evaluated within a range of 4 to 10.

3.3.2. Experiment 2: Generalization capability under clean mismatched conditions

The NRFAR generalization capability was evaluated by processing acoustic signals from different animals located in another field and recorded with different devices. NRFAR was trained on DS1 and tested on DS2. The training set was balanced using the same under- and over-sampling techniques applied in the first experiment. DS2 is composed of 13.2 h of audio recordings, corresponding to 51.8% *grazing*, 24.6% *rumination*, and 23.6% of *other* activities.

3.3.3. Experiment 3: Noise robustness evaluation

External noise may reduce the operability of acoustic foraging activity recognizers operating under free-range conditions. The particular properties of these noise sources, including their finite duration and limited bandwidth, make them difficult to distinguish and quantify in the context of this study, which analyzed almost 350 h of audio recordings. Although audio recordings captured in DS1 might occasionally contain some noise, the signals were assumed to be free of noise; that is, they had an infinite SNR. In this experiment, the robustness of the NRFAR to noise was evaluated in five trials for various levels of contamination with noise and measured in terms of the SNR in a range from 20 to -15 dB in steps of 5 dB. In each trial, NRFAR was trained in the same way as in the first experiment but a different noise source was artificially added to the audio recording of DS1 used for testing and then normalized. A stationary Gaussian white noise source was used in a trial, which is one of the most accepted methods for testing the algorithm noise robustness (Sáez et al., 2016). White noise is an *"infinite"* bandwidth signal with constant power spectral density across all frequencies. Furthermore, the previously mentioned set of audio clips (*Animals*, *Vehicles*, *Weather*, and *Mixture*) was used in four trials to represent nonstationary environmental and natural noises present on the pasture. In each trial, the audio clips were randomly selected without replacement and concatenated to represent the artificial noise source that was used to contaminate the original audio recordings. Some examples of waveforms and spectrograms at several SNRs produced during grazing and rumination are shown in the Supplementary Material.

3.4. Metrics

State-of-the-art BUFAR and JMFAR methods were evaluated under the same conditions as NRFAR to establish a comparison between different methods. Each audio recording has an associated ground-truth text file, specifying the start and end of the bouts, and the corresponding activity labels. The activity bouts, which last from several minutes to hours, were divided into nonoverlapping 1-s frames, following the approach described by Chelotti et al. (2023). This allowed a high-resolution activity recognition analysis to evaluate the performance of the methods. This action was performed on both the algorithm output and the ground truth for a direct comparison. In total, 1,257,759 frames and 47,606 frames were generated from the 349.4 h and 13.2 h of audio recordings of DS1 and DS2, respectively. For each audio signal, the balanced accuracy metric was calculated using the scikit-learn 1.2.2 library in Python[2] (Pedregosa et al., 2011). This metric provides a good indicator of the performance of multiclass imbalance problems (Mosley, 2013).

[2] `https://scikit-learn.org/stable/modules/generated/sklearn.metrics.balanced_accuracy_score.html`

4. Results

4.1. Experiment 1

The recognition performance of the different methods under matched conditions (i.e. trained and tested on DS1) reveals that NRFAR properly classifies $\geq 88.2\%$ of the frames into *grazing* or *rumination* classes, thus showing a significant improvement compared with the average of 79.5% for BUFAR and 84.3% for JMFAR (Figure 5). BUFAR exhibits the lowest recognition rate for the activities of interest but the highest recognition for *other* activities (88.1%). Moreover, confusion between *grazing* and *rumination* is lower for NRFAR ($\leq 1.2\%$), than for BUFAR ($\geq 11.2\%$) and JMFAR ($\geq 5.1\%$).

The computational cost of NRFAR, expressed in terms of operations per second (ops/s), is 13.4% higher than that of BUFAR (43,060 ops/s vs. 37,966 ops/s) and 14.6% lower than that of JMFAR (43,060 ops/s vs. 50,445 ops/s), with marginal variations presented between them. A detailed analysis and assumption of the operations involved are available in Appendix A.

4.2. Experiment 2

The generalization capability of the different methods to recognize foraging activities is evaluated in the independent DS2 dataset. Figure 6 shows the confusion matrices for the three methods. Qualitative previous results on DS1 are extended to those on DS2: NRFAR achieves a higher recognition rate for both *grazing* and *rumination* classes than JMFAR and BUFAR, with lower confusion between these classes.

The comparison of each method's performance in each dataset shows that NRFAR presents similar average balanced accuracies, being 86.4% in DS1 and 87.4% in DS2. Comparing Figure 6c versus Figure 5c, *grazing* is 5.9% higher in DS1 than in DS2, while *rumination* is 4.1% lower. On the other hand, JMFAR exhibits a 7.7% higher classification of *grazing* but 12.7% lower classification of *rumination* in DS1 than in DS2 (Figure 6b versus Figure 5b). The classification of *other* activity is similar in DS1 and DS2 for both NRFAR and JMFAR. BUFAR presents a similar capability for classifying *rumination* in DS1 and DS2. However, the classification of *grazing* decreases 26.1% from DS1 to DS2 (Figure 6a versus Figure 5a).

4.3. Experiment 3

The robustness to adverse conditions of the NRFAR method is evaluated and compared against the BUFAR and JMFAR methods using different noise

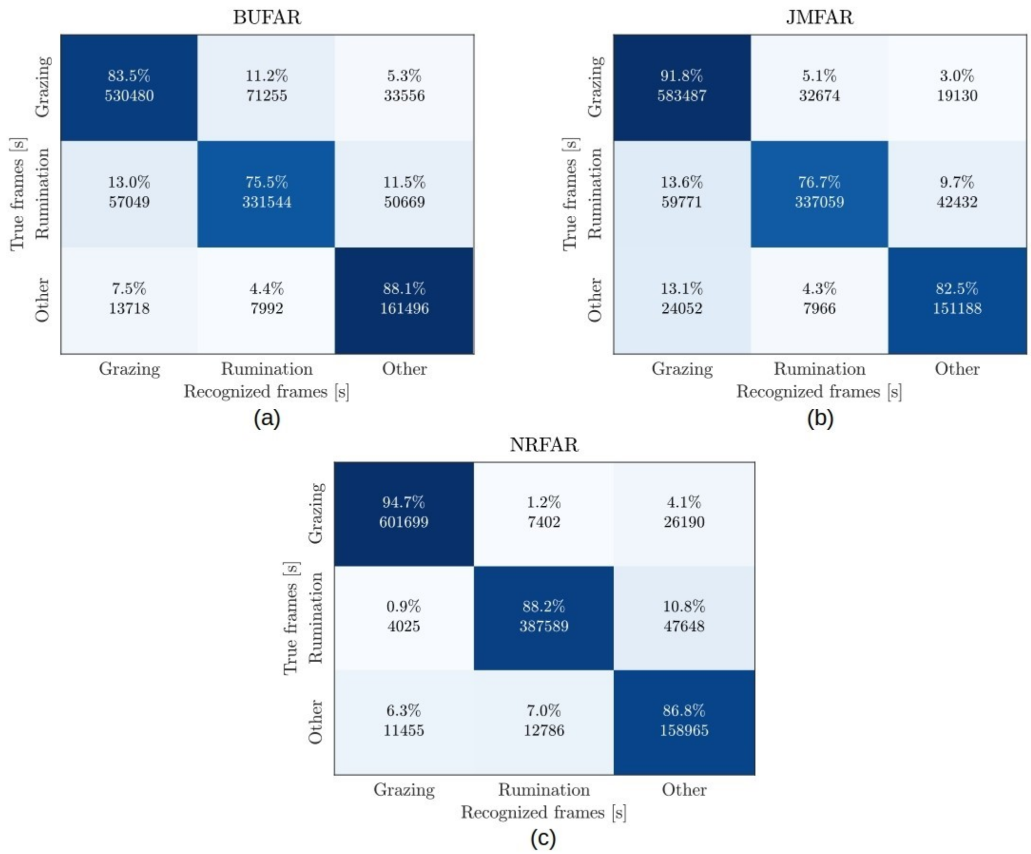

Figure 5: Confusion matrices for different foraging activities for the (a) BUFAR, (b) JMFAR, and (c) NRFAR methods when evaluating on DS1.

sources at multiple SNR levels. Gaussian white noise is added to the audio signals of DS1 in appropriate proportions, to achieve the desired SNR. Figure 7 shows the balanced accuracy, averaged over the audio signals, obtained with each method under different SNR conditions. NRFAR outperforms JMFAR and BUFAR in all cases ($p < 0.05$; Wilcoxon signed-rank test computed over the balanced accuracy of each signal (Wilcoxon, 1945)). The overall performance (average $\pm$ standard deviation) of NRFAR remains approximately constant, ranging from 0.86 ± 0.10 to 0.83 ± 0.13 for SNR ≥ 5 dB. Furthermore, the performance of JMFAR is higher (ranging from 0.79 ± 0.16 to 0.71 ± 0.16) than that of BUFAR (ranging from 0.76 ± 0.17 to 0.69 ± 0.17) under low noise conditions (SNR ≥ 10 dB). For moderate and high noise

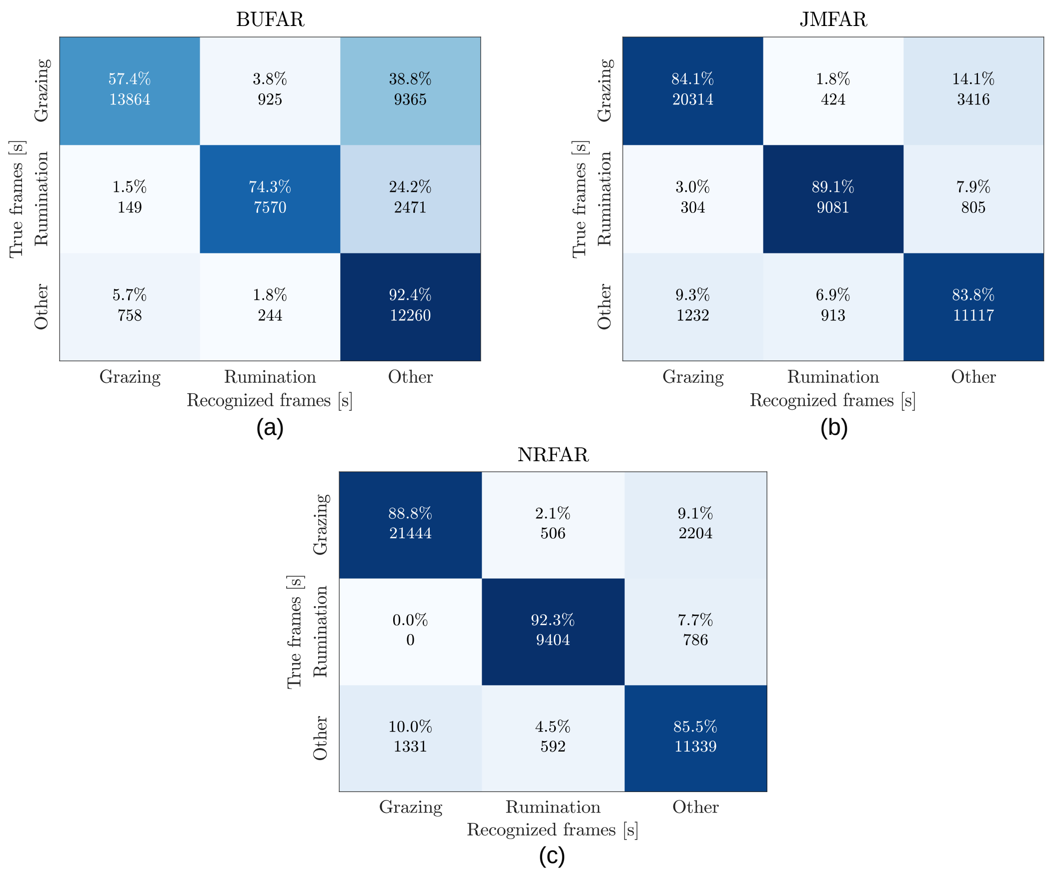

Figure 6: Confusion matrices for different foraging activities for the (a) BUFAR, (b) JMFAR, and (c) NRFAR methods when evaluating on DS2.

conditions (SNR $\leq$ 5 dB), BUFAR (ranging from 0.66 ± 0.17 to 0.39 ± 0.06) outperformed JMFAR (ranging from 0.65 ± 0.16 to 0.32 ± 0.10).

In a more challenging and realistic scenario, the original audio signals of DS1 are mixed with a nonstationary noise source in four independent trials. The noise source contains exclusively sounds of animals, vehicles, weather, or a mixture of these sounds. The balanced accuracy metrics reported by the methods using the four noise sources are shown in Figure 8. The performance of NRFAR decreases as the SNR decreases. However, the performance of BUFAR and JMFAR increases in general for SNR between 20 dB and 10 dB. In general, NRFAR outperforms BUFAR and JMFAR, particularly for SNR $\geq$ 15 dB and for SNR $\leq$ 0 dB. NRFAR has a higher

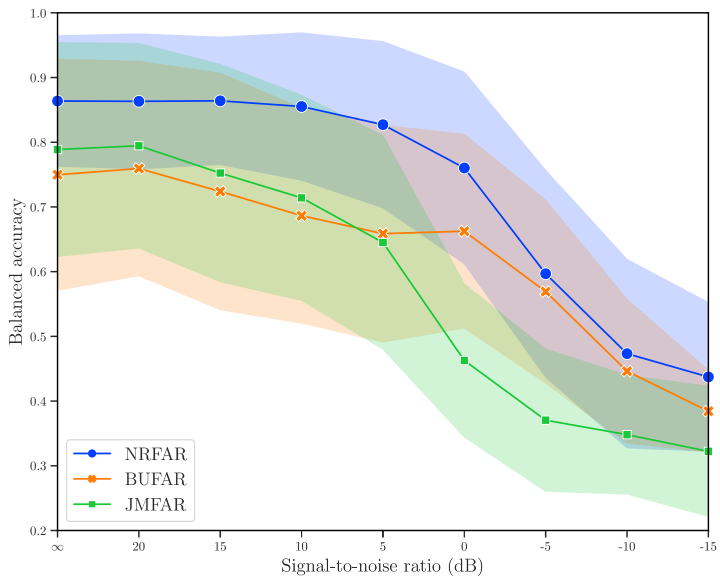

Figure 7: Performance rates (average ± standard deviation) for the NRFAR, BUFAR, and JMFAR methods using additive Gaussian white noise at several SNR levels.

balanced accuracy than BUFAR in the 32 evaluated cases ($p < 0.05$ in 25 cases). Additionally, NRFAR outperforms JMFAR for SNR $\geq$ 20 dB and SNR $\leq$ 0 dB ($p < 0.05$ in 14 of 16 cases). The results of comparing NRFAR with JMFAR for SNR between 15 dB and 5 dB are not always statistically significant, although NRFAR presents higher performances than JMFAR in most cases (Figure 8). On the other hand, JMFAR presents higher average balanced accuracy than BUFAR for SNR $\geq$ 0 dB for the four noise sources, particularly for 10 $\geq$ SNR $\geq$ 0 dB (with $p < 0.05$ in 19 of 20 cases). Reported statistical significance test values obtained in the experiments are available in Appendix B.

The previously reported results have been rearranged to provide a different interpretation. Figure 9 shows the performance degradation of the NRFAR, JMFAR, and BUFAR methods for the different noise sources. In Fig 9.a, the average balanced accuracy of NRFAR ranges from [0.86 - 0.85] for SNR = 20 dB to [0.44 - 0.33] for -15 dB. NRFAR reaches higher performance when Gaussian white noise is used. For a particular SNR value, NRFAR performs similarly between the noise sources representing more realistic acoustic pasture conditions. This is also true for JMFAR (Figure 9.b) but not for BUFAR (Figure 9.c).

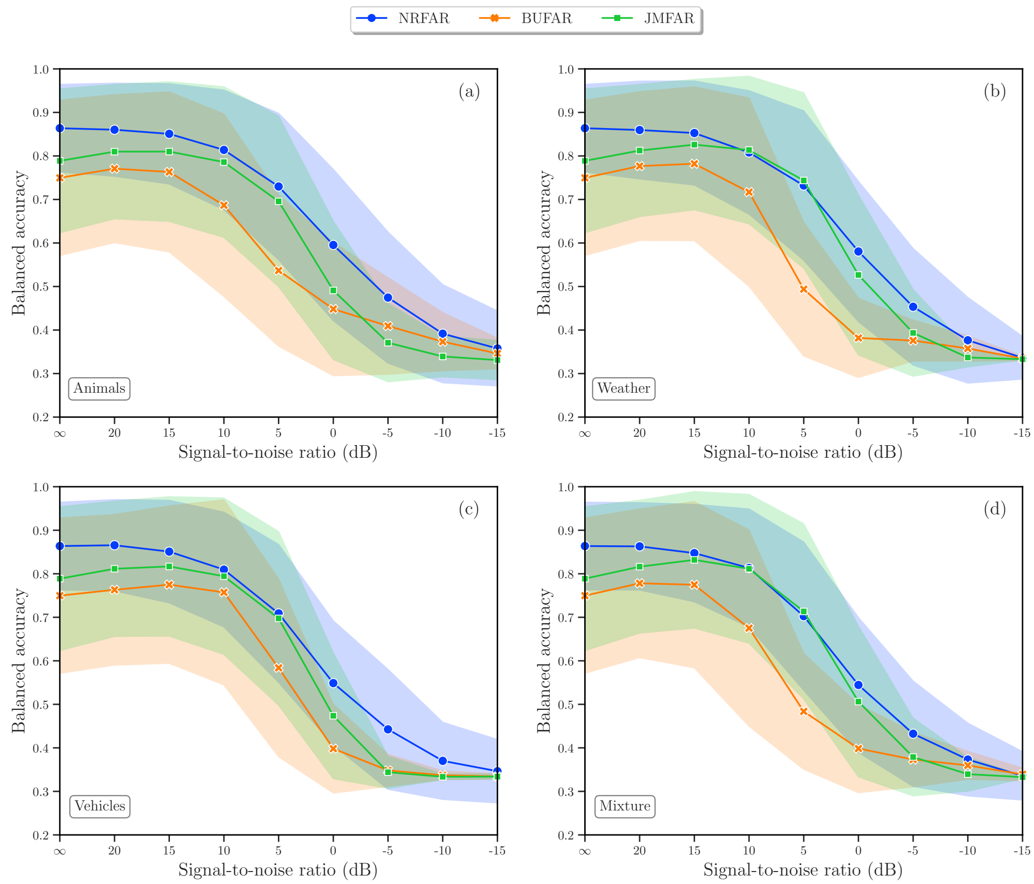

Figure 8: Performance rates (average ± standard deviation) for the NRFAR, BUFAR, and JMFAR methods using noises commonly present on pasture at several SNR levels.

By comparing stationary and nonstationary noise sources, BUFAR and NRFAR exhibit higher performance when Gaussian white noise is added to the audio signals in moderate and high levels (SNR ≤ 5 dB). However, for low noise conditions, the recognition performance of JMFAR is more affected when Gaussian white noise is used.

5. Discussion

Accurately classifying the most important ruminant foraging behavior provides useful information to monitor their welfare and health, and to gain insight into their pasture dry matter intake and utilization (Liakos et al.,

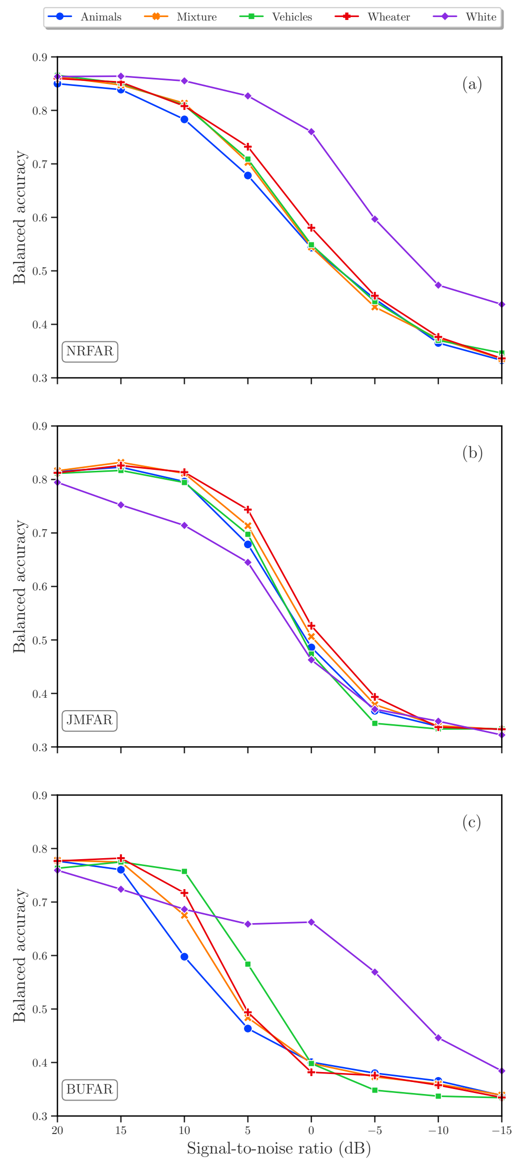

Figure 9: Variation of the performance metric across different noise sources for (a) NRFAR, (b) JMFAR, and (c) BUFAR. Marked points are the balanced accuracy, averaged over signals at a particular SNR level.

2018). This is typically achieved using accelerometers, pressure, or acoustic sensors. Commercial nose-band pressure sensors require handlers to analyze raw data recorded on a computer, which are not suitable for use in big rodeos (Riaboff et al., 2022). Ensuring the proper location, orientation, and attachment of accelerometer sensors mounted on a collar can become a laborious task for handlers to prevent their motion. Meeting these requirements is even more challenging under free-ranging conditions. Therefore, acoustic sensors are preferable for practical use under such conditions (Shen et al., 2020). Existing state-of-the-art acoustic methods for estimating the foraging activities of cattle, called BUFAR and JMFAR, are based on the analysis of fixed-length segments of sound signals. However, the misclassification of foraging activities remains a challenge. This study proposes an improved online acoustic foraging activity recognizer (NRFAR) that analyzes identified JM-event classes in nonoverlapping segments of 5-min duration. Like BUFAR, NRFAR computes statistical features of JM-events to identify foraging activities. NRFAR uses the CBEBA method to recognize JM-events into four classes: *rumination-chews*, *grazing-chews*, *bites*, and *chew-bites*. The NRFAR method represents a significant improvement over the previous BUFAR method, which only distinguished between *bites*, *chew-bites*, and *chews*, without discriminating between *rumination-chews* and *grazing-chews* events. The JMFAR method uses a different approach that does not require the identification of JM-events to delimit grazing and rumination bouts. Instead, it extracts information from the detected JM in the segment.

The results showed that the average correct recognition rate of the activities of interest (*grazing* and *rumination*) for NRFAR was 91.5% when evaluating in DS1, exceeding BUFAR by 12.0% and JMFAR by 7.2% (Figure 5). Importantly, this improvement in activity recognition was achieved without incurring substantial changes in computational cost. The remarkable performance improvement of NRFAR was due to the improved discrimination of JM-events produced during rumination and grazing by CBEBA. The good classification rate of JM-events allowed the computation of a confidence set of activity features with more specific discriminatory information than BUFAR and JMFAR to enhance activity classifications. NRFAR presented a minimal confusion of $\leq 1.2\%$ between *grazing* and *rumination*, which was lower than the confusion reported by BUFAR ($\geq 11.2\%$) and JMFAR ($\geq 5.1\%$). The authors hypothesized that the misclassification of foraging activities was reduced because it depends mainly on the misrecognition of JM-events associated with rumination (*rumination-chew*) and grazing (*grazing-chew*, *bite*, and *chew-bite*), and not between all possible JM-event classes. Therefore, NRFAR was less sensitive to JM-events misclassification than BUFAR. Likewise, discrimination between foraging activities and other activities presented a greater error in the NRFAR ($\geq 4.1\%$). This confusion was also observed in BUFAR and JMFAR and could be related to the great diversity of behavior represented by the *other* class. From a productivity standpoint, confusion of 5% or more between *grazing* and *rumination* can significantly affect the diagnoses of feeding performance (e.g. low dry matter intake) (Watt et al., 2015) or metabolic imbalances of nutritional origin in ruminants (e.g., subacute ruminal acidosis) (Beauchemin, 2018).

An acoustic method must be able to work effectively in different setups to have practical utility. NRFAR, JMFAR, and BUFAR, initially trained using DS1 signals, were tested on DS2 signals. Again, NRFAR exceeded the average recognition rate of *grazing* and *rumination* of JMFAR and BUFAR by 4.0% and 24.7%, respectively, with higher average balanced accuracy (87.4% for NRFAR, 84.4% for JMFAR, and 73.2% for BUFAR). Moreover, the average balanced accuracy of NRFAR in DS2 was 1.0% higher than in DS1, with similar recognition rates of the three classes in both datasets (Figure 5c and Figure 6c), demonstrating good generalization capability. JMFAR also exhibited good generalization performance (average balanced accuracy of 78.9% in DS1 and 84.4% in DS2) but an improvement in the recognition of *rumination* was compensated with a decrease in *grazing* (Figure 5b and Figure 6b). Noteworthy was the limited generalization ability of BUFAR to identify *grazing*, decreasing from 83.5% in DS1 to 57.4% in DS2 (Figure 5a and Figure 6a).

Acoustic methods often have lower performance in confined environments such as barns because of the high levels and varying types of noise present there. Acoustic reverberation existing in confined environments is the cause that noise has to be considered convolutional. In free-ranging conditions, noise is still present but is less intense and frequent, and can be considered additive. To reduce the unwanted effects of acoustic noise, an appropriate microphone setup (as shown in Figure 3) can be used. Hence, the proper operation of acoustic methods in free-ranging is not necessarily compromised. The effectiveness of an acoustic foraging activity recognizer depends on its ability to work well in adverse field conditions, making it a useful and effective tool for farmers and handlers. In this study, the noise robustness of NRFAR was evaluated and compared with previous methods by adding artificial noises to the original audio signals of DS1 at different levels ($20 \leq \text{SNR} \leq -15$ dB), which were even higher than those produced by real noises in classical pasture environments (Bishop et al., 2019). The noise robustness of the methods using a stationary noise source with different properties was evaluated (Figure 7). Artificial random Gaussian white noise was used to contaminate the audio signals. The white noise signal has a theoretical *"infinite"* bandwidth and a constant power spectral density across all frequencies, which can degrade important acoustic cues over the entire frequency range. NRFAR had great robustness to noise for SNR $\geq$ 10 dB, keeping their balanced accuracy almost constant. However, the performances of the JMFAR and BUFAR methods decreased with decreasing SNR. JMFAR performed better than BUFAR at low levels of noise (SNR $\geq$ 10 dB) since the noise had a similar impact on both methods in this SNR range. BUFAR outperformed JMFAR for moderate and high noise levels (SNR ≤ -5 dB) due to the higher robustness to noise of the JM information from recognized JM-events used by BUFAR. Furthermore, JMFAR exhibited the largest drop in performance in this experiment. The decreasing performance of JMFAR was due to the limited robustness to noise of the JM information, computed from detected JM-events, analyzed to recognize foraging activities (Figure 4). Additionally, NRFAR outperformed the other methods for the entire range considered in these numerical experiments (SNR $\geq$ -15 dB) (14 of 16 evaluated scenarios).

The effects of different nonstationary noise sources commonly present on pastures, such as sounds produced by animals, vehicles, weather, and a mixture of these sounds, were also evaluated. Figure 8 showed that JMFAR outperformed BUFAR, which is consistent with the results of Chelotti et al. (2023). In addition, NRFAR outperformed the previous methods in 61 of 64 evaluated scenarios, with 39 of those cases showing statistical significance ($p < 0.05$), as in the evaluations using Gaussian white noise (Figure 7). It should be noted that the largest differences in favor of NRFAR were observed for SNR $\geq$ 15 dB and SNR $\leq$ 0 dB, but NRFAR performed similarly to JMFAR for $10 \leq \text{SNR} \leq 5$ dB. Under high noise conditions, the performance of NRFAR was due to the high noise robustness and discriminative power of the JM features used to classify the JM-events by CBEBA (middle level of Figure 1) (Martinez-Rau et al., 2022).

The robustness of each method to different noise sources was analyzed. The performance of NRFAR using the four nonstationary noise sources was similar to each other for a particular SNR level (Figure 9.a), even though these noise sources have different spectral energy distributions (Özmen et al., 2022). A similar situation was observed for JMFAR (Figure 9b), but not for BUFAR (Figure 9c). It was noteworthy that NRFAR performed better when evaluated with stationary Gaussian white noise compared to the nonstationary noise sources (Figure 9a), particularly for moderate and high noise conditions. This particular situation was also observed in BUFAR (Figure 9c). nonstationary noise sources have uncertain onset, offset, and duration, which can lead to false detection of JM, classifying noises as JM-events (middle level of Figure 1). Figure 9b showed that JMFAR performed similarly with all nonstationary noise sources for SNR $\geq$ -5 dB because it did not depend on the identification of JM-events. Remarkably, JMFAR was less robust to stationary Gaussian white noise than to stationary noise sources at low noise levels (SNR $\geq$ 5 dB).

NRFAR has a low computational cost of 43,060 ops/s, which is of the same order of magnitude as BUFAR and JMFAR. It is important to note that most of the computational cost required by NRFAR (43,121 ops/s) comes from the computation of CBEBA (43,118 ops/s) (see Appendix A). This suggests that NRFAR could potentially be implemented in an application-specific ultra-low-power microprocessor, similar to the implementation of CBEBA (Martinez-Rau et al., 2023c). This computational cost value is theoretical and considers only the arithmetic and logic operations required to execute NRFAR. It is useful to compare the computational requirements of different methods independently on the platform. However, the total processing time of a constrained electronic device depends on available hardware resources (Manor and Greenberg, 2022). The recent deployment of NRFAR in a low-power microcontroller (Martinez-Rau et al., 2023a), combined with its strong noise robustness, positions NRFAR as a reliable tool to be embedded in an acoustic sensor for recognizing grazing and rumination activities.

6. Conclusion

This study proposes an improvement over former acoustic methods to recognize and delimit foraging activity bouts of grazing cattle. Inspired by the former BUFAR method, the proposed NRFAR method analyzes fixed-length segments of recognized JM-events. NRFAR uses a robust JM recognizer that discriminates JM-events produced during grazing and rumination under different operating conditions. This allows NRFAR to recognize foraging activities in free-range scenarios, even under adverse acoustic conditions. The method has shown a significant performance improvement over state-of-the-art acoustic methods in quiet and noisy conditions, and in different settings. The evaluation of noise robustness was performed by adding artificially different amounts of stationary Gaussian white noise, and nonstationary natural noise commonly present in free-range. Future work must include changes in the analysis of fixed-length segments to variable-length segments using dynamic segmentation to facilitate more accurate estimation of the foraging bouts of interest. Likewise, NRFAR could be used as a reference for developing new methods based on multi-modal data sensors to recognize feeding activities in more adverse environments, such as barns.

Acknowledgment

The authors wish to express their gratitude to the staff of the KBS Robotic Dairy Farm, who participated in the investigation. Additionally, we acknowledge the direct support from AgBioResearch-MSU. The authors would like to thank Constanza Quaglia (technical staff, CONICET) and J. Tomás Molas G. (technical staff, UNER-UNL) for their technical support in achieving the web demo. This work was supported by the Universidad Nacional del Litoral [CAID 50620190100080LI and 50620190100151LI]; Universidad Nacional de Rosario [AGR216, 2013 - AGR266, 2016 - and 80020180300053UR, 2019]; Agencia Santafesina de Ciencia, Tecnología e Innovación [IO-2018–00082], CONICET [PUE sinc(I), 2017]; and USDA-NIFA [MICL0222 and MICL0406].

CRediT authorship contribution statement

LSMR, JOC, MF, HLR, and LLG participated in conceptualization; LSMR participated in software stage; LSMR, JOC, JRG, and AMP participated in the data curation; LSMR, JOC, MF, HLR and LLG participated in the formal analysis; LSMR, JOC, MF, and HLR participated in the investigation stage; LSMR, JOC, MF, HLR, and LLG participated in methodology, validation and visualization stages; JRG, LLG, SAU, and HLR participated in the funding acquisition; JRG, LLG, and HLR participated in project administration; LSMR, JOC, MF, JRG, SAU, AMP, HLR and LLG contributed to the writing and reviewing of the original draft; All the authors reviewed and approved the manuscript.

Data availability

Data will be available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Computational cost

The computational cost of NRFAR depends on the input audio sampling frequency, the sub-sampling frequency used internally in CBEBA (fixed at $f_s = 150\ Hz$ in this analysis, according to its optimal value), the configuration of the two MLP neural networks used to classify the JM-events and foraging activities, and the duration of the segment lengths (fixed at 5 min). To obtain a valid comparison with other methods, an input sampling frequency of $f_i = 2\ kHz$ and 2 JM-events per second was chosen. Furthermore, the worst-case computational cost scenario was selected for both MLP classifiers. In addition, any arithmetic operation, arithmetic shift, logic comparison, or activation function is counted as one operation. The required number of operations per second for the computation stages of each level of NRFAR is:

Bottom level:

1. Audio pre-processing: limiting the bandwidth with a second-order band-pass filter and computing the instantaneous power signal requires $7 * f_i$ and f_i ops/s per sample, respectively. Then, 16,000 ops/s are required.
2. Signal computation: computing and decimating the envelope signal requires $11 * f_i + 150$ ops/s. Computing the energy signal by frames requires $f_i + 300$ ops/s. Altogether, this stage requires 24,450 ops/s.

Middle level:

1. JM-event detection: $4 + 0.925 * f_s$ and $12 + f_s$ operations per JM-event are necessary to detect and delimit the boundaries of JM-events. Then, this stage takes 610 ops/s.
2. Feature extraction: $3.5 * f_s$ operations per JM-event are necessary to compute the set of JM features. In total, 1050 ops/s are required.
3. JM-event classification: deciding whether an event should be classified requires $f_s + 3$ operations per JM event, whereas the MLP with 5-6-4 neurons requires 131 operations per JM-event, thus, 568 ops/s are required.

4. Tuning parameters: $f_s + 39$ operations per JM-event are necessary to update the thresholds. Then, 378 ops/s are required.

Middle level:

1. Segment buffering: this stage requires 2 operations per JM-event equivalent to 4 ops/s.
2. Feature extraction: computing the set of activity features requires 608 ops/segment.
3. Activity classification: considering the maximum number of neurons (10) in the hidden layer, the MLP requires 185 ops/segment.
4. Smoothing process: this filtering stage takes 2 ops/segment.

Finally, the total computational cost of NRFAR is 43,060 ops/s + 795 ops/segment $\approx$ 43,063 ops/s. Similar to BUFAR, the overall computational cost almost exclusively depends on the bottom and middle levels of Figure 1 (i.e., the JM event recognizer) because the top level is only executed once every 5 min (segment length). Hence, the total computational cost of NRFAR can be expressed as 12,918,795 ops/segment.

Appendix B. Statistical hypothesis test

The statistically significant discrepancies in the balanced accuracy between NRFAR and BUFAR, NRFAR and JMFAR, and JMFAR and BUFAR were evaluated using the Wilcoxon signed-rank test (Wilcoxon, 1945). Tables B.1, B.2, and B.3 show the p-values obtained from the comparison of these methods. P-values with a green background indicate a significant difference in performance with a confidence level of 5% ($p = 0.05$), and p-values with a pink background indicate a nonsignificant difference.

Table B.1: Statistically significant p-values were obtained by comparing the performance of the NRFAR and BUFAR methods with different noise sources at several noise levels.

SNR [dB]	NRFAR vs BUFAR				
	Animals	Vehicles	Weather	Mixture	White
20	3.88e-05	1.69e-08	8.75e-06	5.36e-06	1.02e-08
15	1.21e-04	7.79e-04	5.38e-04	8.33e-04	3.30e-11
10	1.58e-10	3.78e-01	9.34e-04	1.93e-06	7.36e-14
5	1.04e-15	1.92e-06	9.88e-15	1.34e-15	4.36e-13
0	1.43e-09	1.57e-09	1.71e-15	4.59e-10	1.16e-05
-5	7.39e-04	8.82e-06	5.20e-05	6.53e-04	1.98e-01
-10	6.23e-01	1.19e-02	9.68e-01	9.04e-01	2.16e-01
-15	5.63e-01	1.85e-01	9.44e-01	4.19e-01	6.01e-04

Table B.2: Statistically significant p-values were obtained by comparing the performance of the NRFAR and JMFAR methods with different noise sources at several noise levels.

SNR [dB]	NRFAR vs JMFAR				
	Animals	Vehicles	Weather	Mixture	White
20	8.45e-02	6.52e-04	1.80e-03	6.95e-03	5.45e-05
15	5.55e-01	2.30e-01	1.61e-01	9.76e-01	6.11e-10
10	3.66e-01	7.02e-01	3.28e-01	9.02e-01	2.61e-13
5	6.48e-01	5.98e-01	3.36e-01	2.69e-01	4.80e-15
0	3.12e-02	4.20e-04	3.77e-02	2.14e-01	8.13e-20
-5	3.29e-06	6.08e-07	8.82e-03	6.31e-03	2.83e-13
-10	4.04e-02	2.96e-03	1.20e-02	4.94e-03	6.17e-08
-15	5.95e-01	1.71e-01	7.00e-01	4.54e-01	3.15e-09

Table B.3: Statistically significant p-values were obtained by comparing the performance of the JMFAR and BUFAR methods with different noise sources at several noise levels.

SNR [dB]	JMFAR vs BUFAR				
	Animals	Vehicles	Weather	Mixture	White
20	4.67e-02	2.95e-03	2.33e-02	2.09e-02	4.39e-02
15	1.79e-04	6.66e-03	3.74e-03	2.36e-03	1.73e-01
10	2.01e-14	7.01e-02	4.646e-09	1.49e-10	1.58e-01
5	6.94e-17	1.04e-12	8.32e-18	3.47e-17	6.68e-01
0	1.25e-06	5.57e-10	2.58e-11	1.50e-10	1.07e-14
-5	6.81e-02	1.38e-01	5.61e-01	8.14e-01	4.71e-16
-10	9.58e-09	1.53e-04	7.81e-06	4.03e-08	3.89e-09
-15	4.20e-04	5.00e-01	2.73e-02	1.05e-04	5.31e-06

References

Abeni, F. and Galli, A. (2017). Monitoring cow activity and rumination time for an early detection of heat stress in dairy cow. *International Journal of Biometeorology*, 61(3):417–425.

Allrich, R. D. (1993). Estrous behavior and detection in cattle. *Veterinary Clinics of North America: Food Animal Practice*, 9(2):249–262.

Andriamandroso, A., Bindelle, J., Mercatoris, B., and Lebeau, F. (2016). A review on the use of sensors to monitor cattle jaw movements and behavior when grazing. *Biotechnologie, Agronomie, Société et Environnement*, 20.

Aquilani, C., Confessore, A., Bozzi, R., Sirtori, F., and Pugliese, C. (2022). Review: Precision livestock farming technologies in pasture-based livestock systems. *Animal*, 16(1):100429.

Arablouei, R., Wang, Z., Bishop-Hurley, G. J., and Liu, J. (2023). Multimodal sensor data fusion for in-situ classification of animal behavior using accelerometry and gnss data. *Smart Agricultural Technology*, 4:100163.

Balasso, P., Marchesini, G., Ughelini, N., Serva, L., and Andrighetto, I. (2021). Machine learning to detect posture and behavior in dairy cows: Information from an accelerometer on the animal's left flank. *Animals*, 11(10).

Banhazi, T. M., Lehr, H., Black, J., Crabtree, H., Schofield, P., Tscharke, M., and Berckmans, D. (2012). Precision livestock farming: an international review of scientific and commercial aspects. *International Journal of Agricultural and Biological Engineering*, 5(3):1–9.

Beauchemin, K. (2018). Invited review: Current perspectives on eating and rumination activity in dairy cows. *Journal of Dairy Science*, 101(6):4762–4784.

Beauchemin, K. A. (1991). Ingestion and mastication of feed by dairy cattle. *Veterinary Clinics of North America: Food Animal Practice*, 7(2):439–463.

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., and Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors*, 21(11).

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer Verlag.

Bishop, J. C., Falzon, G., Trotter, M., Kwan, P., and Meek, P. D. (2019). Livestock vocalisation classification in farm soundscapes. *Computers and Electronics in Agriculture*, 162:531–542.

Bosi, M., Brandenburg, K., Quackenbush, S., Fielder, L., Akagiri, K., Fuchs, H., and Dietz, M. (1997). ISO/IEC MPEG-2 advanced audio coding. *Journal of the Audio Engineering Society*, 45(10):789–814.

Brandenburg, K. and Stoll, G. (1994). Iso/mpeg-1 audio: A generic standard for coding of high-quality digital audio. *Journal of the Audio Engineering Society*, 42(10):780–792.

Bristow, D. J. and Holmes, D. S. (2007). Cortisol levels and anxiety-related behaviors in cattle. *Physiology & Behavior*, 90(4):626–628.

Cabezas, J., Yubero, R., Visitación, B., Navarro-García, J., Algar, M. J., Cano, E. L., and Ortega, F. (2022). Analysis of accelerometer and gps data for cattle behaviour identification and anomalous events detection. *Entropy*, 24(3).

Chapa, J. M., Maschat, K., Iwersen, M., Baumgartner, J., and Drillich, M. (2020). Accelerometer systems as tools for health and welfare assessment in cattle and pigs – a review. *Behavioural Processes*, 181:104262.

Chelotti, J. O., Vanrell, S. R., Galli, J. R., Giovanini, L. L., and Rufiner, H. L. (2018). A pattern recognition approach for detecting and classifying jaw movements in grazing cattle. *Computers and Electronics in Agriculture*, 145:83–91.

Chelotti, J. O., Vanrell, S. R., Martinez-Rau, L. S., Galli, J. R., Utsumi, S. A., Planisich, A. M., Almirón, S. A., Milone, D. H., Giovanini, L. L., and Rufiner, H. L. (2023). Using segment-based features of jaw movements to recognise foraging activities in grazing cattle. *Biosystems Engineering*, 229:69–84.

Chelotti, J. O., Vanrell, S. R., Rau, L. S. M., Galli, J. R., Planisich, A. M., Utsumi, S. A., Milone, D. H., Giovanini, L. L., and Rufiner, H. L. (2020). An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle. *Computers and Electronics in Agriculture*, 173:105443.

Connor, E. E. (2015). Invited review: Improving feed efficiency in dairy production: challenges and possibilities. *Animal*, 9(3):395–408.

Dolecheck, K. A., Silvia, W. J., Heersche Jr, G., Chang, Y. M., Ray, D., Stone, A., Wadsworth, B., and Bewley, J. (2015). Behavioral and physiological changes around estrus events identified using multiple automated monitoring technologies. *Journal of Dairy Science*, 98(12):8723–8731.

Ferrero, M., Vignolo, L. D., Vanrell, S. R., Martinez-Rau, L. S., Chelotti, J. O., Galli, J. R., Giovanini, L. L., and Rufiner, H. L. (2023). A full end-to-end deep approach for detecting and classifying jaw movements from acoustic signals in grazing cattle. *Engineering Applications of Artificial Intelligence*, 121:106016.

Fonseca, E., Favory, X., Pons, J., Font, F., and Serra, X. (2022). Fsd50k: An open dataset of human-labeled sound events. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30:829–852.

Galli, J., Cangiano, C., Pece, M., Larripa, M., Milone, D., Utsumi, S., and Laca, E. (2018). Monitoring and assessment of ingestive chewing sounds for prediction of herbage intake rate in grazing cattle. *Animal*, 12(5):973–982.

Galli, J. R., Milone, D. H., Cangiano, C. A., Martínez, C. E., Laca, E. A., Chelotti, J. O., and Rufiner, H. L. (2020). Discriminative power of acoustic features for jaw movement classification in cattle and sheep. *Bioacoustics*, 29(5):602–616.

Garcia, R., Aguilar, J., Toro, M., Pinto, A., and Rodriguez, P. (2020). A systematic literature review on the use of machine learning in precision livestock farming. *Computers and Electronics in Agriculture*, 179:105826.

He, H., Bai, Y., Garcia, E. A., and Li, S. (2008). Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, pages 1322–1328, Hong Kong.

Kamminga, J. W., Le, D. V., Meijers, J. P., Bisby, H., Meratnia, N., and Havinga, P. J. (2018). Robust sensor-orientation-independent feature selection for animal activity recognition on collar tags. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1).

Kilgour, R. J. (2012). In pursuit of "normal": A review of the behaviour of cattle at pasture. *Applied Animal Behaviour Science*, 138(1):1–11.

Kovács, L., Kézér, F., Ruff, F., and Szenci, O. (2017). Rumination time and reticuloruminal temperature as possible predictors of dystocia in dairy cows. *Journal of Dairy Science*, 100(2):1568–1579.

Laca, E. A., Ungar, E. D., Seligman, N. G., Ramey, M. R., and Demment, M. W. (1992). An integrated methodology for studying short-term grazing behaviour of cattle. *Grass and Forage Science*, 47(1):81–90.

Li, C., Tokgoz, K. K., Fukawa, M., Bartels, J., Ohashi, T., Takeda, K.-i., and Ito, H. (2021a). Data augmentation for inertial sensor data in cnns for cattle behavior classification. *IEEE Sensors Letters*, 5(11):1–4.

Li, G., Xiong, Y., Du, Q., Shi, Z., and Gates, R. S. (2021b). Classifying ingestive behavior of dairy cows via automatic sound recognition. *Sensors*, 21(15):5231.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8):2674.

Lovarelli, D., Bacenetti, J., and Guarino, M. (2020). A review on dairy cattle farming: Is precision livestock farming the compromise for an environmental, economic and social sustainable production? *Journal of Cleaner Production*, 262:121409.

Lovarelli, D., Brandolese, C., Leliveld, L., Finzi, A., Riva, E., Grotto, M., and Provolo, G. (2022). Development of a new wearable 3d sensor node and innovative open classification system for dairy cows' behavior. *Animals*, 12(11):1447.

Manor, E. and Greenberg, S. (2022). Custom hardware inference accelerator for tensorflow lite for microcontrollers. *IEEE Access*, 10:73484–73493.

Martinez-Rau, L. S., Adın, V., Giovanini, L. L., Oelmann, B., and Bader, S. (2023a). Real-time acoustic monitoring of foraging behavior of grazing cattle using low-power embedded devices. In *2023 IEEE Sensors Applications Symposium (SAS)*, pages 01–06.

Martinez-Rau, L. S., Chelotti, J. O., Ferrero, M., Galli, J., Utsumi, S., Planisich, A., Vignolo, L., Giovanini, L., and Rufiner, H. L. (2023b). Day-long acoustic recordings of grazing and rumination activities in dairy cows. *Scientific Data*, 10:782.

Martinez-Rau, L. S., Chelotti, J. O., Vanrell, S. R., Galli, J. R., Utsumi, S. A., Planisich, A. M., Rufiner, H. L., and Giovanini, L. L. (2022). A robust computational approach for jaw movement detection and classification in grazing cattle using acoustic signals. *Computers and Electronics in Agriculture*, 192:106569.

Martinez-Rau, L. S., Weißbrich, M., and Payá-Vayá, G. (2023c). A 4μw low-power audio processor system for real-time jaw movements recognition in grazing cattle. *Journal of Signal Processing Systems*, 95(4):407–424.

Martínez Rau, L., Chelotti, J. O., Vanrell, S. R., and Giovanini, L. L. (2020). Developments on real-time monitoring of grazing cattle feeding behavior using sound. In *2020 IEEE International Conference on Industrial Technology (ICIT)*, pages 771–776, Buenos Aires, Argentina.

Michie, C., Andonovic, I., Davison, C., Hamilton, A., Tachtatzis, C., Jonsson, N., Duthie, C.-A., Bowen, J., and Gilroy, M. (2020). The internet of things enhancing animal welfare and farm operational efficiency. *Journal of Dairy Research*, 87(S1):20–27.

Milone, D. H., Galli, J. R., Cangiano, C. A., Rufiner, H. L., and Laca, E. A. (2012). Automatic recognition of ingestive sounds of cattle based on hidden markov models. *Computers and Electronics in Agriculture*, 87:51–55.

Mosley, L. (2013). *A balanced approach to the multi-class imbalance problem.* PhD thesis, Industrial and Manufacturing Systems Engineering, Iowa State University, Ames.

Moto G6 smartphone specification (2018). Motorola Moto G6. `https://www.gsmarena.com/motorola_moto_g6-9000.php`. Accessed: June 28, 2024.

Osei-Amponsah, R., Dunshea, F. R., Leury, B. J., Cheng, L., Cullen, B., Joy, A., Abhijith, A., Zhang, M. H., and Chauhan, S. S. (2020). Heat stress impacts on lactating cows grazing australian summer pastures on an automatic robotic dairy. *Animals*, 10(5):869.

Özmen, G., Ozkan, İ. A., Seref, I., Tademir, S., Mustafa, Ç., and Arslan, E. (2022). Sound analysis to recognize cattle vocalization in a semi-open barn. *Gazi Journal of Engineering Sciences*, 8(1):158–167.

Pahl, C., Hartung, E., Grothmann, A., Mahlkow-Nerge, K., and Haeussermann, A. (2014). Rumination activity of dairy cows in the 24 hours before and after calving. *Journal of Dairy Science*, 97(11):6935–6941.

Pahl, C., Hartung, E., Mahlkow-Nerge, K., and Haeussermann, A. (2015). Feeding characteristics and rumination time of dairy cows around estrus. *Journal of Dairy Science*, 98(1):148–154.

Paudyal, S., Maunsell, F., Richeson, J., Risco, C., Donovan, D., and Pinedo, P. (2018). Rumination time and monitoring of health disorders during early lactation. *Animal*, 12(7):1484–1492.

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. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(85):2825–2830.

Phillips, C. (2002). *Nutritional Behaviour*. John Wiley & Sons, Ltd.

Piczak, K. J. (2015). ESC: Dataset for Environmental Sound Classification. In *Proceedings of the 23rd ACM International Conference on Multimedia*, page 1015–1018, Brisbane, Queensland, Australia.

Rehman, A. U., Abbasi, A. Z., Islam, N., and Shaikh, Z. A. (2014). A review of wireless sensors and networks' applications in agriculture. *Computer Standards & Interfaces*, 36(2):263–270.

Reith, S. and Hoy, S. (2012). Relationship between daily rumination time and estrus of dairy cows. *Journal of Dairy Science*, 95(11):6416–6420.

Riaboff, L., Shalloo, L., Smeaton, A., Couvreur, S., Madouasse, A., and Keane, M. (2022). Predicting livestock behaviour using accelerometers: A systematic review of processing techniques for ruminant behaviour prediction from raw accelerometer data. *Computers and Electronics in Agriculture*, 192:106610.

Rook, A. and Huckle, C. (1997). Activity bout criteria for grazing dairy cows. *Applied Animal Behaviour Science*, 54(2):89–96.

Sáez, J. A., Luengo, J., and Herrera, F. (2016). Evaluating the classifier behavior with noisy data considering performance and robustness: The equalized loss of accuracy measure. *Neurocomputing*, 176:26–35.

Schirmann, K., von Keyserlingk, M., Weary, D., Veira, D., and Heuwieser, W. (2009). Technical note: Validation of a system for monitoring rumination in dairy cows. *Journal of Dairy Science*, 92(12):6052–6055.

Shen, W., Cheng, F., Zhang, Y., Wei, X., Fu, Q., and Zhang, Y. (2020). Automatic recognition of ingestive-related behaviors of dairy cows based on triaxial acceleration. *Information Processing in Agriculture*, 7(3):427–443.

Skowronski, M. D. and Harris, J. G. (2004). Exploiting independent filter bandwidth of human factor cepstral coefficients in automatic speech recognition. *The Journal of the Acoustical Society of America*, 116(3):1774–1780.

Steinmetz, M., von Soosten, D., Hummel, J., Meyer, U., and Dänicke, S. (2020). Validation of the rumiwatch converter v0.7.4.5 classification accuracy for the automatic monitoring of behavioural characteristics in dairy cows. *Archives of Animal Nutrition*, 74(2):164–172.

Stygar, A. H., Gómez, Y., Berteselli, G. V., Dalla Costa, E., Canali, E., Niemi, J. K., Llonch, P., and Pastell, M. (2021). A systematic review on commercially available and validated sensor technologies for welfare assessment of dairy cattle. *Frontiers in Veterinary Science*, 8.

Tzanidakis, C., Tzamaloukas, O., Simitzis, P., and Panagakis, P. (2023). Precision livestock farming applications (plf) for grazing animals. *Agriculture*, 13(2).

Ungar, E. D. and Rutter, S. M. (2006). Classifying cattle jaw movements: comparing iger behaviour recorder and acoustic techniques. *Applied Animal Behaviour Science*, 98(1-2):11–27.

Vanrell, S. R., Chelotti, J. O., Bugnon, L. A., Rufiner, H. L., Milone, D. H., Laca, E. A., and Galli, J. R. (2020). Audio recordings dataset of grazing jaw movements in dairy cattle. *Data in Brief*, 30:105623.

Vanrell, S. R., Chelotti, J. O., Galli, J. R., Utsumi, S. A., Giovanini, L. L., Rufiner, H. L., and Milone, D. H. (2018). A regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle. *Computers and Electronics in Agriculture*, 151:392–402.

Wang, L., Arablouei, R., Alvarenga, F. A., and Bishop-Hurley, G. J. (2023). Classifying animal behavior from accelerometry data via recurrent neural networks. *Computers and Electronics in Agriculture*, 206:107647.

Watt, L., Clark, C., Krebs, G., Petzel, C., Nielsen, S., and Utsumi, S. (2015). Differential rumination, intake, and enteric methane production of dairy cows in a pasture-based automatic milking system. *Journal of Dairy Science*, 98(10):7248–7263.

Werner, J., Leso, L., Umstatter, C., Niederhauser, J., Kennedy, E., Geoghegan, A., Shalloo, L., Schick, M., and O'Brien, B. (2018). Evaluation of the rumiwatchsystem for measuring grazing behaviour of cows. *Journal of Neuroscience Methods*, 300:138–146.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83.

Wilkinson, J. M., Lee, M. R. F., Rivero, M. J., and Chamberlain, A. T. (2020). Some challenges and opportunities for grazing dairy cows on temperate pastures. *Grass and Forage Science*, 75(1):1–17.

Wu, Y., Liu, M., Peng, Z., Liu, M., Wang, M., and Peng, Y. (2022). Recognising cattle behaviour with deep residual bidirectional lstm model using a wearable movement monitoring collar. *Agriculture*, 12(8).