Computers and Electronics in Agriculture 216 (2024) 108500

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

CowScreeningDB: A public benchmark database for lameness detection in dairy cows

Shahid Ismail [a], Moises Diaz [b,*], Cristina Carmona-Duarte [b], Jose Manuel Vilar [c], Miguel A. Ferrer [b]

[a] Universidad del Atlantico Medio, Las Palmas de Gran Canaria, Spain
[b] iDeTIC Institute, Universidad de Las Palmas de Gran Canaria, 35017 Las Palmas, Spain
[c] Faculty of Veterinary Medicine, Universidad de Las Palmas de Gran Canaria, 35017 Las Palmas, Spain

A R T I C L E I N F O

Keywords:
Cow
Lameness
Machine learning
Public database
Dairy
Support vector machine

A B S T R A C T

Lameness is one of the costliest pathological problems affecting dairy animals. It is usually assessed by trained veterinary clinicians who observe features such as gait symmetry or gait parameters as step counts in real time. With the development of artificial intelligence, various modular systems have been proposed to minimize subjectivity in lameness assessment. However, the major limitation in their development is the unavailability of a public database, as most existing ones are either commercial or privately held. To tackle this limitation, we have introduced CowScreeningDB, a multi-sensor database which was built with data from 43 dairy cows. Cows were monitored using smart watches during their normal daily routine. The uniqueness of the database lies in its data collection environment, sampling methodology, detailed sensor information, and the applications used for data conversion and storage, which ensure transparency and replicability. This data transparency makes CowScreeningDB a valuable and objectively comparable resource for further development of techniques for lameness detection for dairy cows. In addition to publicly sharing the database, we present a machine learning technique which classifies cows as healthy or lame by using raw sensory data. To facilitate fair comparisons with state-of-the-art methods, we introduce a novel benchmark. Combining the database, the machine learning technique and the benchmark validate our major objective, which is to establish the relationship between sensor data and lameness. The developed technique reports an average accuracy of 77 % for the best case scenario and presents perspectives for further development. By introducing this framework which encompasses the database, the classification algorithm and the benchmark, we significantly reduce subjectively in lameness assessment. This contribution to lameness detection fosters innovation in the field and promotes transparent, reproducible research in the pursuit of more effective management of dairy cow lameness.

Implications: Lameness detection is one of the main tasks in dairy systems, given its importance in the production ambit. However, the data used during detection is generally either held privately or sold commercially. In this study, we create a multi-sensor database (CowScreeningDB), which can be used for lameness. Because we have made the database public[1] and free of charge for research purposes, it should act as a benchmark allowing to objectively compare techniques put forth to deal with lameness. We also provide details of the sampling system used, comprised of hardware and a baseline classification algorithm.

1. Introduction

Productivity in livestock farming is negatively affected by housekeeping costs and dairy diseases such as lameness. In recent years, sensor-based artificial intelligence systems have been used to assess the overall health of cows, including behavioral changes, body part detection (Jiang et al., 2019a), etc. Among the physiological behaviors, i.e., the physical motion of the cow under observation, lameness is the main

* Corresponding author at: F109 Basic Science Building. Campus de Tafira s/N. 35017. Las Palmas de Gran Canaria. Spain
E-mail addresses: shahiid.ismail@gmail.com (S. Ismail), moises.diaz@ulpgc.es (M. Diaz), cristina.carmona@ulpgc.es (C. Carmona-Duarte), jose.vilar@ulpgc.es (J.M. Vilar), miguelangel.ferrer@ulpgc.es (M.A. Ferrer).

[1] (CowScreeningDB can be downloaded from https://github.com/Shahid-Ismail/CowScreeningDB-A-public-database-for-lameness-detection or https://gpds.ulpgc.es/)

https://doi.org/10.1016/j.compag.2023.108500
Received 15 August 2023; Received in revised form 5 November 2023; Accepted 3 December 2023
Available online 8 December 2023

cause of the most critical change observed in the gait or stance. Lameness is a collective term for three types of abnormalities, namely, claw horn disruption lesions (CHDL), skin lesions, and non-foot lameness (Mason, 2007). CHDL is not a single disease, but rather, covers a number of non-infectious foot lesions, including sole ulcers, sole hemorrhage, and white line disease (Griffiths et al., 2020). Skin lesions are indicative of dysfunctional housing of dairy cows (Kielland et al. (2009), while non-foot generally includes all non-foot conditions, such as lameness due to bone injury and muscle and joint damage (Mason, 2007).

As lameness is associated with a broad spectrum of health disorders, it can serve as a good biomarker for detecting them. Veterinary clinicians are thus trained for this at individual and herd levels. However, lameness detection can be hard to realize, and additional factors such as rain need to be considered (Thompson et al., 2019). Generally, a veterinary clinician can detect lameness by focusing on lameness predictors, mainly gait and posture traits. Asymmetric gait, reluctance to bear weight, short steps, joint flexibility, arched back, head bobbing, walking speed, difficulty in turning, and difficulty in rising are the predictors which clinicians generally use during a visual inspection (Schlageter-Tello et al., 2014). Recent years have seen an increase in automation of lameness detection through the use of sensor-based systems that record force distribution and kinematics (linear, temporal and/or angular) measurements or any other combination related to the stance or motion of the animal.

Force platforms are generally used to perform kinetic assessments to detect lameness in dairy animals objectively. These devices have been used by researchers such as Chapinal and Tucker (2012), who have employed them along with video recordings to assess the weight distribution among limbs and step counts in lame cows. In the same vein, Hertem et al. (2013) detected lameness via a tri-variable model using logistic regression, which is based on milk yield, rumination, and neck activity. Another study, which used logistic regression, was conducted by Kamphuis et al. (2013). They used uni-variables such as weight, activity (average steps per hour) and features related to milk, and their combinations (weight, activity and features related to milk together) which converted analysis to multivariable to assess lameness. Bruijnis et al. (2010) related foot disorder with lameness using a dynamic stochastic simulation model in a related study. Similarly, Pastell et al. (2006) used piezoelectric force sensors for lameness detection.

Kinematics parameters such as stride length, support time, and articular range of motion are the characteristics used by veterinarians to assess lameness. In this context, Poursaberi et al. (2010) introduced a real-time system for lameness detection based on shape analysis using back posture. The authors concluded that posture-based classification can operate in real time. Piette et al. (2020) also employed back posture analysis to validate the performance of an automatic camera-based system. Hartem et al. (2016) compared the performance of a graphic video system for lameness detection with a multi-sensor system verifying milk production, activity, and postural changes. They concluded that sensor-based videography enhances lameness detection in systems with other kinematic sensors in place. In a related work, Zhaoa et al. (2018) used a multivariable analysis looking at gait asymmetry, speed, tracking up, stance time, stride length, and tenderness to classify a cow as sound or lame. Other studies using kinematics were conducted by Vázquez Diosdado et al. (2018), Van De Gucht et al., 2017, Maertens et al. (2011), and Pastell et al. (2008). However, the studies by Van De Gucht et al., 2017 and Pastell et al. (2008) are different from the others as they use force sensor-based kinematics.

The major modalities in the field of automated lameness detection systems are reflected in video and image-based methods. These systems can detect lameness both online and offline. Among studies considered herein are those conducted by Jiang et al. (2022), Kang et al. (2020), Jiang et al. (2020), Piette et al. (2020), Jiang et al., (2019b), Zhao et al. (2018) and Van Hertem et al. (2014). Unlike digital systems, which are used for lameness detection (Kang et al., (2020); Jiang et al. (2020); Piette et al. (2020); Jiang et al., (2019b); Zhao et al. (2018) and Van Hertem et al. (2014)), Jiang et al. (2022) used an analog system (phase alteration line (PAL)). Due to their associated high financial and computational costs, the latter systems generally have limited storage capacity. For example, media in the systems of Jiang et al. (2022), Kang et al. (2020), Jiang et al., (2019b) and Zhao et al. (2018) are limited respectively to 40 s, 1 k images, 40 s, 30 s and 7 min. The systems of Jiang et al. (2022) and Jiang et al., (2019b) are real-time systems which use the back position of cows and a double normal distribution statistical model, respectively. However, lameness classification by Kang et al. (2020), Piette et al. (2020), Zhao et al. (2018) and Van Hertem et al. (2014) is carried out offline, and the authors use supporting phase, back posture, leg swing analysis and consecutive night-time milking sessions, respectively.

To reduce associated costs, inertial sensors such as pedometers and accelerometers have gained in popularity in recent years. They are in widespread use, as reflected by studies such as Lemmens et al. (2023), Frondelius et al. (2022), O'Leary et al. (2020a), O'Leary et al. (2020b), Borghart et al. (2021), Jarchi et al. (2021), Antanaitis et al. (2021), Shahinfar et al. (2021), Taneja et al. (2020), Byabazaire et al. (2019), Weigele et al. (2018), Barker et al. (2018), Vázquez Diosdado et al. (2018), Beer et al. (2016), Thorup et al. (2016), Thorup et al. (2015), Garcia et al. (2014), De Mol et al. (2013), Van Hertem et al. (2013), Kamphuis et al. (2013), Van Hertem et al. (2013), Chapinal and Tucker (2012), Maertens et al. (2011) and Nielsen et al. (2010). The main reason for the popularity of these systems is that they involve a continuous sampling of the lameness predictor. Generally, lameness is detected using inertial measurements (Borghart et al., 2021; Jarchi et al., 2021; Weigele et al., 2018; Barker et al., 2018 and others), milk related measurements (Lemmens et al., 2023; Borghart et al., 2021; Van Hertem et al., 2016 and others), behavior-related predictors such as lying time, number of lying bouts, maximum length of the lying bout, roughage feeding time, etc. (Frondelius et al., 2022; Zhao et al., 2018; Thorup et al., 2016 and others) and a mixture of these mentioned predictors (Riaboff et al., 2021; Shahinfar et al., 2021 and others).

In addition to the preceding modalities, videography for sensor data is a composite approach in which the computational load is reduced by using sensor data-based classification. Studies that have looked at this approach include Beer et al. (2016), Van Hertem et al. (2016) and Kokin et al. (2014). Table A (Appendix A. Supplementary material) presents a summary overview of the above literature review.

In Table A, studies are compared via hardware (sensors) used to sample the prediction parameter (signal), data statistics, lameness levels and public sharing. Data statistics are given by Equations (1) and (2):

$$S_S = N_C\left(N_S{}^*N_{D/O}\right) \tag{1}$$

$$S_V = N_C\left(N_S{}^*I_R{}^*D\right) \tag{2}$$

where, S_S, S_V are sensor- and video/image-based statistics, respectively. $N_C, N_S, N_{D/O}, I_R andD$ are the number of cows, sensors/video (images) files, number of days/observations, image resolution in pixels and duration for videos, respectively. Here, S_S indicates two types of sensor-based results. In the first, results are reported by number of cows, number of sensors, and sensor active duration. Results are also reported in simplified form by number of cows, number of sensors and number of observations. They are differentiated by * which shows the second type. Similarly, S_V reflects both video- and image-based results represented by the † sign. Here, a distinction is made between the video and image results by adding two † signs for video based research works. Again, the results are given in terms of number of cows, number of images/video files and duration, respectively.

An objective comparison can be done by number of observations/files from a given number of cows. This comparison would be valid for image/video -based studies as well as for the sensor-based studies, in which the number of observations/files are given. Jiang et al. (2022), Kang et al. (2020), Zhao et al. (2018), Lemmens et al. (2023), Borghart et al. (2021) are example of such studies which hold information needed for comparison. However, this comparison is not valid for the Frondelius et al. (2022), Antanaitis et al. (2021) and similar studies. A further complication is added in videography-based studies when the duration of videos is considered in comparison with images. A similar difficulty is also encountered in sensor-based studies which are based on samples (Shahinfar et al., 2021; Borghart et al., 2021), versus studies containing sampling duration (Riaboff et al., 2021; Byabazaire et al., 2019). The preceding information highlights a critical challenge in objectively comparing studies for lameness detection. Another complication arises from the need for objective comparisons within similar studies. For example, it is challenging comparing Borghart et al. (2021) with Jarchi et al. (2021), although both studies are reported in terms of cows, number of sensors and observations/files. However, there are 3799 observations from 164 cows, using 6 sensors, versus 25,624 observations from 23 cows, using eight sensors. An extra layer of difficulty is added by the lameness scoring. Generally, scoring is reported according to the Sprecher et al. (1997) ranking, and is in the 1–5 range. However, Frondelius et al. (2022) and Jarchi et al. (2021) have used the paired scores for lameness levels. Similarly, certain scores are missing in some studies (Garcia et al., (2014) and Barker et al. (2018)). The hardware used in lameness detection is also very diverse, and includes a RumiWatch noseband halter, a SONY HDR-CX290E, a Bosch BMI160 inertial measurement unit, among others. Finally, all studies (with the exception of the present one) do not grant public access to their data, which limits objective comparisons of the developed techniques.

The above-mentioned complicating factors relating to objective comparisons all point to the need to establish a public database. In order to remove this major limitation, we have introduced CowScreeningDB[1], along with the necessary information. The critical importance of data in the machine learning domain cannot be overemphasized, as evidenced by many research works (Martens (2018); Zhao, et al. (2020); Celi et al. (2019); Artrith et al. (2021); Shimron, et al. (2022); Rodgers, et al. (2023); Aldoseri, et al. (2023); Jain, et al. (2020); Hettinga et al. (2023); Hu et al. (2020); Catillo et al. (2022); Dekker (2006); Paullada et al. (2021) and Trisovic et al. (2021)), to name but a few. Among these studies, certain research endeavors (Zhao,et al., (2020); Hettinga et al. (2023); Hu et al. (2020); Catillo et al. (2022) and Celi et al. (2019)) reflect the importance of public access to data for machine learning-based AI development. Aside from the database being publicly accessible, its quality must be sufficiently high to allow its reuse. In fact, data gatherers are expected to abide by the FAIR (Findability, Accessibility, Interoperability, Reusability) principle (Martin et al. (2017); Roche et al. (2022) and Llebot and Steven (2019)) for public data sharing. CowScreeningDB data is also collected following this principle, and is currently shared at multiple public repositories[1]. We would like to point out that there are indeed private databases using inertial equipment similar to that in our study. For example, Benaissa et al. (2019) collected data using an accelerometer, but during their study, the dairy cows stayed in the barn, and the sampling frequency was less than 1 Hz. It is thus not possible to compare their results with our database, in which the dairy cows move from the barn to the milking parlor, and the sampling frequency is 100 Hz. To the best of our knowledge, there therefore is no equivalent database against which ours can be compared. The present study is the only one that freely shares full raw inertial measurement unit data (accelerometer, gyroscope, and magnetometer) sampled at 100 Hz, making it the only database available for public research and development for lameness detection in dairy cows. Hence, public sharing represents the main contribution of our study.

Data collection is realized in a bid to study the relationship between sensory data and lameness scores. The present study establishes this relationship using a machine learning-based inference technique while employing classical classification measures, thereby, fulfilling the major objective. In short, we are not only offering a database, but we are also providing a baseline technique along with it, and this should help researchers as they extend the present work.

The following are the major highlights of our research:

- A labeled database, CowScreeningDB, is made available to public. The data for this database is collected from both healthy and lame cows; the lame cows are characterized by different lameness levels, which were assigned by trained veterinary clinicians. The reference provided along with the database allows an unbiased assessment of lameness.
- During CowScreeningDB sampling, each cow is observed for approximately 6.7 h during their routine life. While the database is created, continuous sampling is performed, ensuring that it is transparent, and can thus be used for continuous evaluation.
- Information regarding the sampling and classification system is also shared, allowing its reuse by other researchers or domain experts.
- A benchmark containing standard measures is introduced, allowing to objectively compare different techniques.

2. Material and methods

The material and methods section is divided into four major subsections: data collection, data distribution, benchmarking criterion, and methodology introduced, which includes features used as part of the methodology.

2.1. Data collection

Farm details and data sampling are included in the data collection. The first details of the farms which are used during the study are given next.

2.1.1. Farm details

The data for this study were obtained from an intensive dual-purpose Friesian dairy cattle farm in Gran Canaria, Spain. This farm, boasting a total of 1100 animals, focuses on milk production, generating an average of 5,750,000 L annually. Each cow contributes to this figure by producing an average of 12,700 L per year, equivalent to 35 L of milk per day. The herd consists of 580 lactating cows, 220 rebreeding cows, and 300 designated for meat consumption. The farm's facilities lack individual cubicles, and provide a minimum of 10 square meters of space per animal. The farm has implemented an artificial insemination program for over a decade, enhancing genetic diversity and overall livestock quality. The corral bedding, a mixture of manure and straw, is regularly oxygenated to prevent bacterial growth. Concrete surfaces in feeding and milking areas facilitate easy cleaning.

For the study, a sample size of 25–30 animals, both healthy and pathological, was determined. This number allowed for comprehensive data collection within a reasonable timeframe. Recordings, lasting 5–6 h, ensured thorough movement data availability. Although the cause of lameness in the cows may have had different origins, our focus was on differentiating between healthy and lame animals, enabling prompt detection of issues.

The animals' gait and pathology were assessed using a 5-point locomotion scoring system developed by Sprecher et al. (1997), which focuses on back arching, weight distribution on limbs when standing, and movement patterns. Classification and observation were carried out by a professional veterinarian with six years of extensive experience with dairy cattle. Animals were assessed once or twice, and rechecks were performed within 12 h, ensuring consistency in classifications. Previous studies (Eriksson et al. (2020); Sahar et al., 2022) corroborated the reliability of this approach, affirming the stability of classifications even after 24 h.

2.1.2. Data sampling

Data sampling and transfer were carried out using the network diagram shown in Figs. 1-3. From this diagram (Fig. 1), we can see that dairy cows are fitted with smart watches fitted with Wi-Fi connectivity.

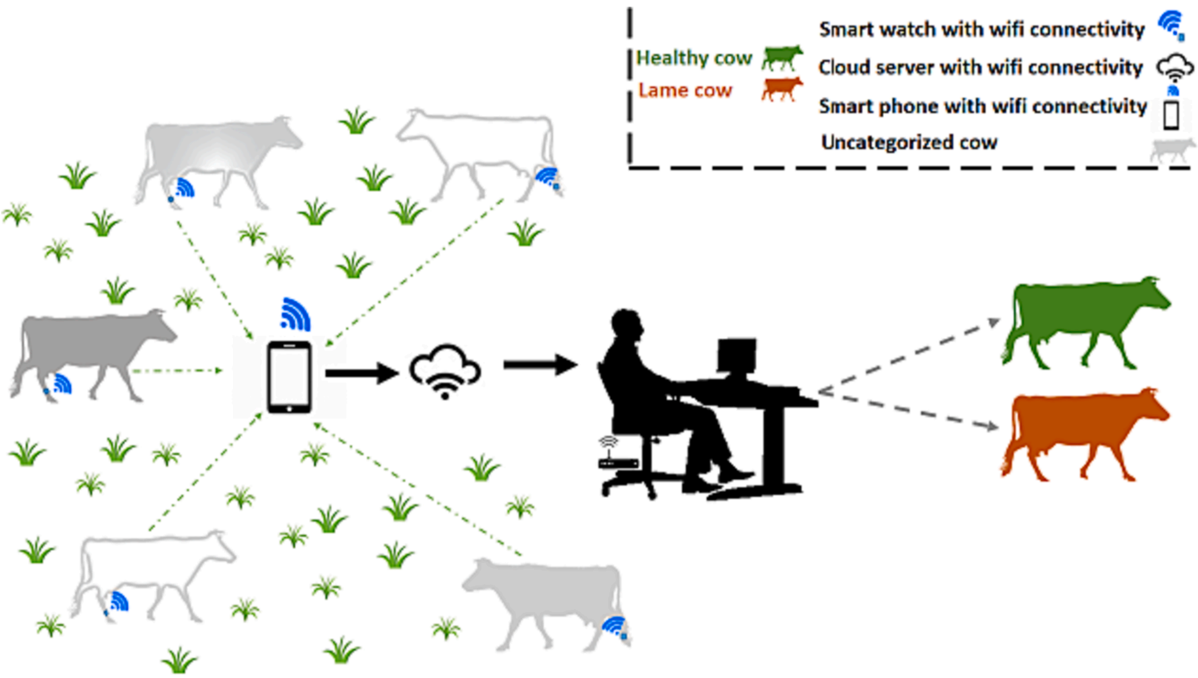

Fig. 1. Network Diagram. The figure shows the main components used in the sampling of data and classification. Smart watches are attached to the lower part of hind legs, which are connected to a smart phone. An application is installed in the smart phone to convert raw data collected to a format suitable for saving on cloud servers. The classifier, who is generally a person, downloads the data from cloud and feeds it to a machine learning based-classification algorithm which categorizes cows as healthy or lame.

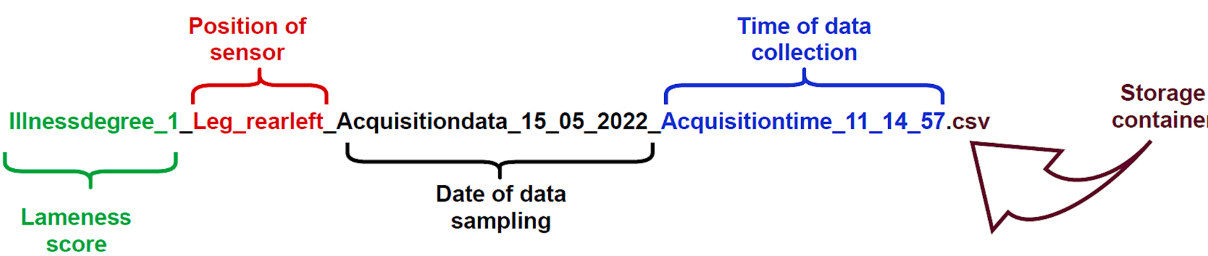

Fig. 2. The figure shows the name of a file which itself contains the sampling information. It can be seen that this specific file was sampled from a cow which was healthy as its lameness level was 1. Sensors were located on the rear left leg, and data was collected at 14 min and 57 s after 11o'clock on May 15, 2022. The data was saved using the comma-separated value file format.

These watches are connected to a smart phone, which in turn is connected to a cloud service. From the cloud service, data is given as an input to the classification algorithm, which classifies a cow as either healthy or lame. Fig. 2 presents details regarding the sampled file. From the Figure, we can see that the file was saved using comma-separated values (CSV) on May 15, 2022. The file name also includes the storage start time along with the position of the sensor. For recording the full database, a smart watch was placed on one of the cow's four legs, as can be seen in Fig. 3. From points, A and B, it can be seen that watches were firmly attached to the legs of the cows using straps attached to an Apple Watch 6, and thus, the orientations of all the sensors, including the accelerometer, the gyroscopes and the magnetometer remained the same. By keeping the sensors at a relatively fixed position on the leg, only inertial data from the former were recorded, and the noise contribution due to the motion of the sensors themselves was very small. Utilizing the apparatus enabled the recorded signals to maintain a high signal-to-noise ratio, eliminating the need to apply any signal conditioning technique before segmentation (as shown in Fig. 6).

The placement of the device on the limbs was done randomly, including on the affected limb. This random placement of the sensor helped eliminate bias and proved the sensor's capacity to detect lameness without targeting specific limbs. Animals adapted well to wearing the sensors, which were lightweight, and the initial minutes of recordings made during placement manipulation were excluded to ensure accurate data collection.

Our teams at the University of Las Palmas de Gran Canaria created an app specifically designed to capture data in this study. This app allows the veterinarian to select the paw where the watch will be located. As such, the registered leg, recording date, and animal identification number are included in each stored file name. The watch records continuously, storing recorded data in consecutive 90-second files. Once the data is recorded on the smart watch, it is synchronized with a smart phone and uploaded to the cloud for storage. Each file contains 13 columns with temporal information data, cow-generated acceleration (without gravity) for all three axes of the device, gravity (three axes), gyroscope data (radial velocity) of the three axes, and the attitude (yaw, pitch, and roll). An Apple Watch 6, an Apple iPhone and iCloud were used, but any other smart watch, smart phone and cloud service combination could also be utilized.

Three sensors, namely, a gyroscope, a magnetometer and accelerometers, were used in the present study. Information regarding acceleration and radial velocity was collected directly using the gyroscope and accelerometer. Similarly, the Apple watch also provided information on gravity after post-processing. However, for information on the

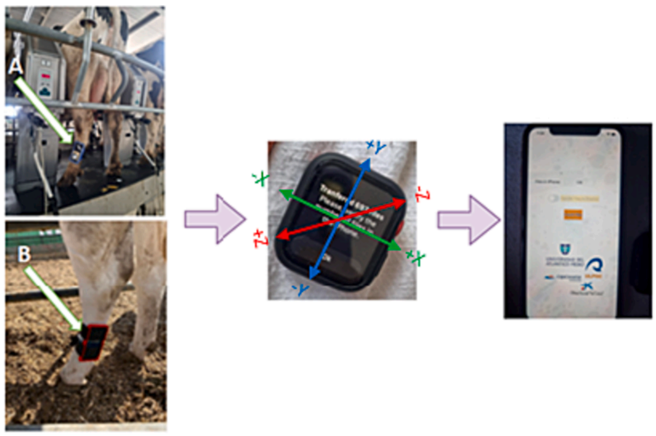

Fig. 3. Data transfer system composed of sensor attached to the cow, an Apple Watch 6 and an iPhone App. An Apple watch was recording data during the routine life of a cow. Arrows 'A' and 'B' show the watches attached to the paw of the cow during two dairy activities. during milking in milking parlor and during a normal walk. The firm grip of the straps attached to the watch (Apple Watch 6) kept the Apple Watch intact, ensuring the fixed orientation of sensors inside. By keeping the Watch and embedded sensors in place, signal noise due to the watch's motion was minimized. The figure also shows the reference orientation of the Apple Watch 6 with the respect to the x-, y- and z-axes. Once data was collected, they were transmitted to a local system as well as to a cloud service.

attitude (radial position), the direction of the reference pole using a magnetometer was also initialized. The sampling rate, i.e., the sampling frequency used, was 100 Hz.

Afterwards, the data was transmitted to the mobile phone and to the cloud service. Data can be appended in serial or parallel, giving 43 cows and 11,518 samples, respectively. Section 2.2 presents statistics for CowScreeningDB. An analytical overview of the multi-sensor data is given in Table 1, where it can be seen that the data is composed of different physical signals, i.e., acceleration, gravity, radial velocity and attitude, meaning that the data is separable when different signals are considered separately.

2.2. Data distribution

Figs. 4 and 5 show the data collected for CowScreeningDB along with the necessary associated statistics. From Fig. 4, it can be seen that there are five categories of cows in terms of lameness. A cow with lameness score of 1 represents a sound (healthy) cow, followed by cows with lameness scores of 2–5, which are with various levels of lameness. Fig. 4-a also shows a binary categorization of cows as healthy or lame. There are 19 healthy cows, and 7, 6, 6 and 5 lame cows respectively with lameness scores of 2–5. Fig. 4-b presents the number of samples per lameness score. There are 4787 samples (a file of 90 s) taken from healthy cows, 19 in number, as shown in Fig. 4-a. Similarly, 1038, 1630, 2432 and 1631 samples are respectively collected from the lame cows.

Table 1
Details of the signals in the database. Here, ST is the signal type for which units are also given, respectively. Acceleration, gravity, angular rotation and roll, pitch, and yaw all have three components, which are directed along x-, y- and z-axes.

Channel No./ ST(Unit)	Time	Acceleration	Gravity	Angular Rotation	Roll, Pitch and Yaw
1	S	–	–	–	–
2:4	–	m/s^2	–	–	–
5:7	–	–	m/s^2	–	–
8:10	–	–	–	Radian/s	–
11:13	–	–	–	–	Radian

As can be seen, there are 19 versus 24 cows for the healthy and lame categories, with 4787 and 6731 samples, respectively.

Fig. 5 presents data in the form of average number of samples and average time duration for the different lameness scores using box plots. From the box plots shown, it can be inferred that the data varies between different lameness scores both with respect to samples as well as time duration.

2.3. Benchmarking criteria

For CowScreeningDB, the benchmarking criteria for the evaluation of technique are as follows.

Classification benchmark: Table 2 indicates that data could be used for binary as well as multi-class classification using the standard measures of sensitivity, specificity, precision and accuracy. If the binary class points towards an abnormal animal, the multi-class classification screens the particular level of lameness for that case.

Ablation study: An ablation study is generally conducted to select the most discriminating features or signals.

Since four different types of signals were used in this study, any subsequent technique introduced will therefore have to report results regarding the most discriminating signals and features used. Prospective studies could be further incorporate to include the fusion of different signals such as combining acceleration with gravity and another similar combination. As mentioned above, the minimum criterion for benchmarking is the binary classification, along with an ablation study regarding the selection of the most discriminating signals and features. However, an in-depth study could be extended to include multi-class classification. In the next section (Section 2.4), the details of a machine learning-based classification system are given, which could be used as a baseline for future studies.

2.4. Methodology

Once data is converted into serial format, it is then classified using the machine learning-based classification system. However, data conversion to the serial format is challenging due to length variation among the samples. To address this issue, a length normalization could be done. In the present case, length variation between signals makes them unsuitable for length equalization using the normalization method as the variations are too great. Therefore, features which could reflect the global nature of signals are extracted. For example, the power distribution in the frequency spectra could reflect the frequency behavior related to the inertial data. For the present study, feature extraction is applied for segmented as well as non-segmented signals, as provided in the classification technique in Fig. 6. The system can be divided into three main modules, namely, segmentation algorithm, feature calculation, and classification network. The modules are introduced next, starting with the segmentation algorithm.

Segmentation algorithm: This algorithm is comprised of a 3rd-order moving median filter and an 8th-order zero phase filter homomorphic filter with a cutoff of 3 Hz, and normalization. A moving median filter removes the spike noise from the signal, which is then given as an input to the homomorphic filter. Homomorphic filtering uses the Hilbert transform to enhance the average components.

The details of operations applied during homomorphic filtering are given in Equations 3–7 and Fig. 7. From the equations, we can infer that during homomorphic filtering, the signal $x[n]$ is considered to be comprised of low frequency ($f[n]$) and high frequency ($h[n]$) components, and the latter component refers to the lameness contents in the present case. To filter these contents, a Hilbert transformation along with filtering and log-based manipulations are applied. Fig. 7 shows the sequence in which these operation are applied, with the Hilbert transformation being the starting operation. The cutoff of the low pass filter was at 3 Hz, which was set empirically.

Finally, normalization ([0,1]) is applied to the filtered signal. For the

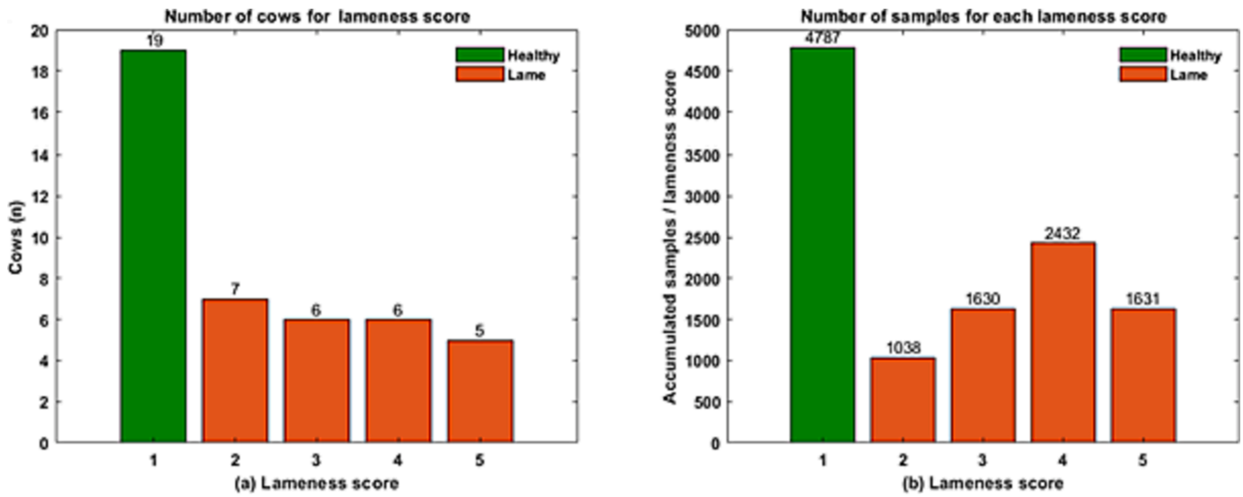

Fig. 4. Data statistics for CowScreeningDB using lameness scores. (a) The figure shows the data division by designated lameness score (1–5), with level 1 indicating a healthy cow. Lame cows are represented by lameness scores of 2–5. (b) The figure shows the data by number of samples/lameness score.

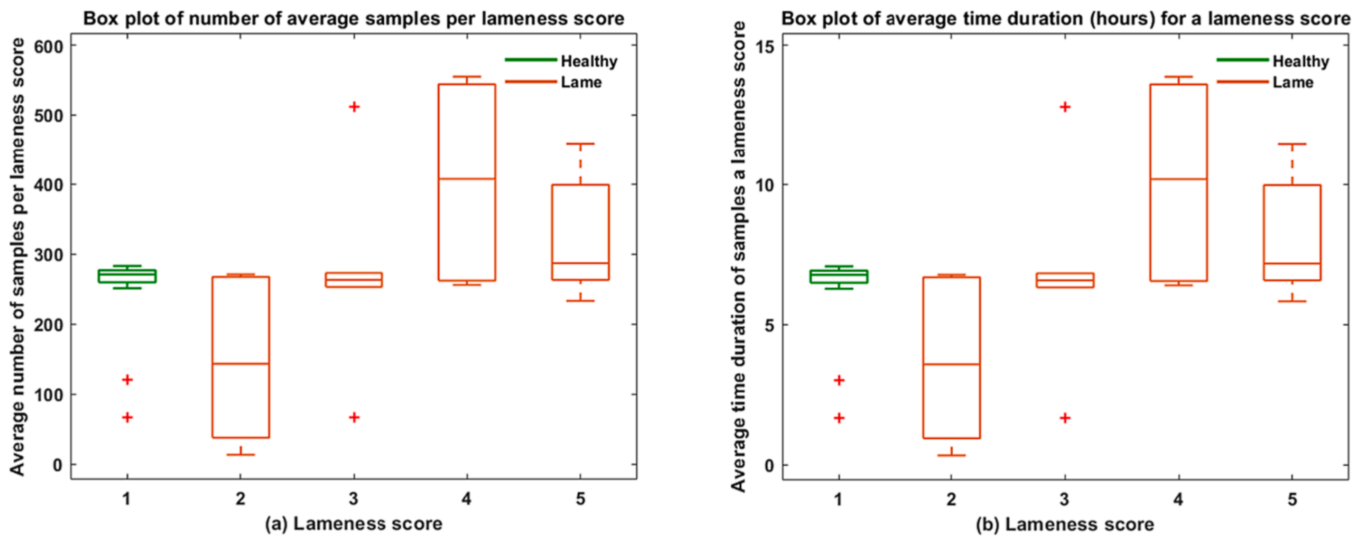

Fig. 5. Data distribution by average number of samples and average time duration related to samples. (a) Box plot of average number of samples for different lameness scores; (b) Box plot of average time duration of samples collected for a given lameness score. The average time duration is in hours.

segmentation algorithm, every sample with a power greater than a given threshold (10 %) is considered as motion, with the remaining samples representing resting positions. Lameness predictors, such as gait symmetry, steps per unit time and ratio of rest to motion are calculated during motion. Hence, segmenting a signal into motion and rest can aid in lameness detection. However, in this study, segmentation is used to increase the number of features (Section 2.4) from 184 to 370 for lameness detection.

Features: Features are characteristics or attributes which can be used to discriminate a healthy cow from a lame one. Two such features, namely, the cumulative distribution frequency (CDF) and the Chirplet Z transform (CZT), are shown in Fig. 8. From Fig. 8-a, it can be seen that the CDF slope for a lame cow is close to unity, but it is changing continuously for a healthy cow.

$$x[n] = h[n]f[n] \tag{3}$$

$$log(x[n]) = log(h[n]f[n]) \tag{4}$$

$$log(x[n]) = log(h[n]) + log(f[n]) \tag{5}$$

$$e^{log(x[n])} = e^{log(f[n])} \tag{6}$$

$$x[n] = f[n] \tag{7}$$

A change in slope refers to an abrupt change in power when the cow is in motion. Hence, a healthy cow moves with more power and agility in comparison with a lame one. Similarly, in Fig. 8-b, the distribution of frequency contents is shown for an acceleration signal. From the figure, we can see that the morphology (shape) of spectral contents can be used to differentiate between a lame and a healthy cow. Features used during the study are primarily based on the mentioned features, along with the power-based features. Details on features are given below, starting with transform-based features.

Chirplet Z Transform features: These are spectral features that are calculated using the Chirplet transform. The Chirplet Z Transform (Equation (9) is a modified version of the Fast Fourier Transform, which is given here in Equation (8):

$$X(K) = \sum_{n=0}^{N-1} x[n]e^{-\frac{j\pi nk}{N}} \qquad \text{n, K} = 0, \cdots\cdots \text{N} - 1) \tag{8}$$

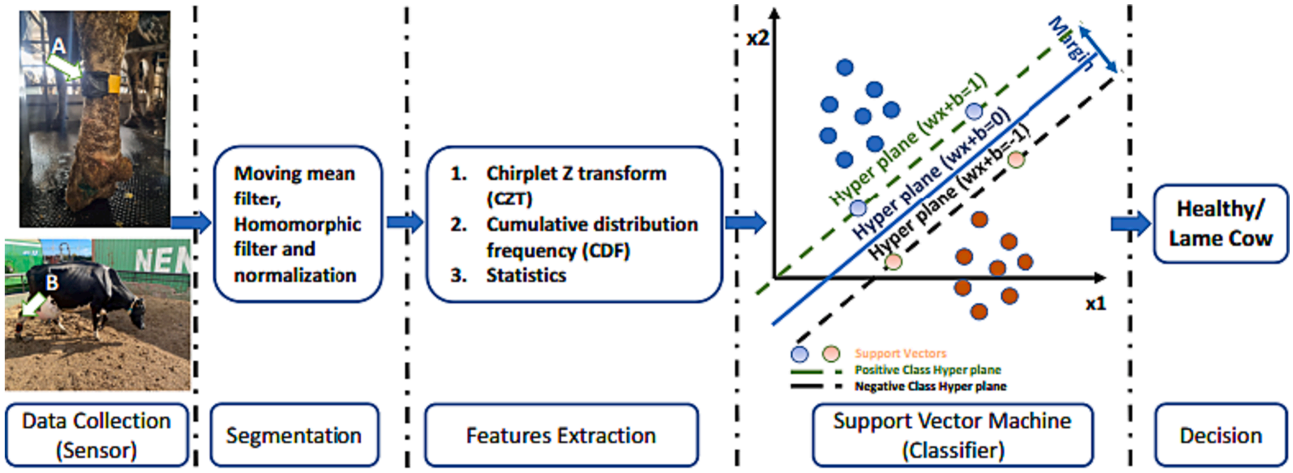

Fig. 6. Classification system consisting in sensor-based data acquisition followed by its conversion into format suitable for input to a machine learning algorithm. An SVM (support vector machine) classifies the data in terms of healthy and lame cows. Data is sampled during the routine life of a cow, as shown in the figure. Under "Data Collection. (Sensor)", arrows 'A' and 'B' highlight the sensor, i.e., the Apple Watch 6.

Table 2
Dimensionality and parameters of features used during the study. The total number of features is 370, and out of these, 362 are either CZT- or CDF-based features, where N_M, N_H, $\delta_{Avg.}$ and DF represent the filter orders of moving median, filter orders of homomorphic filters, average threshold, and decimation factor, respectively.

N_M	N_H	Normalization	$\delta_{Avg.}$	DF	CZT	Statistical	Power	CDF	Total
3	8	[0,1]	90	100	2 × 90	2 × 3	2 × 2	2 × 90	370

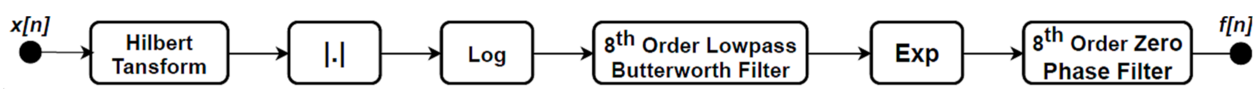

Fig. 7. Overview of homomorphic filtering process. Basic building blocks are the Hilbert transform, absolute value calculation, log and antilog transformations and filtering. Here, low pass and zero phase filters are applied to retain the low frequency along with the mentioned processes.

$$CZT[x(n)] = X(Z_k) = \sum_{n=0}^{N-1} x[n] z_k^{-n}$$

$$z_k = AW_k = A_0 e^{j\theta_0} W_0 e^{-j\varphi_0} \text{ n, K } = 0, 1 \cdots\cdots \text{N} - 1 \quad (9)$$

In the above equations (8),9), x[n] is the input signal and X(K) and $X(Z_k)$ are the Fast Fourier Transform and the Chirplet Z Transform, respectively (Hu and Zhu (2011)). From Equations (8),9), it can be inferred that the Fourier and Chirplet Transforms are calculated on a unit circle ($e^{\frac{-j\pi nk}{N}}$) and an arc, which is parameterized ($A_0 e^{j\theta_0}$, $W_0 e^{-j\varphi_0}$). The Chirplet Z Transforms are calculated for original as well as segmented signals. The resultant transforms are both normalized ([0,1]) and greatly decimated with a decimation factor. Three statistical features, namely, the mean, standard deviation and inverse coefficient of variation, are also calculated from the decimated and normalized Chirplet Z transforms.

Power-based features: The motion percentage, the power percentage and the power crossing are global power-based features. The motion percentage is a heuristic feature that is based on the consecutivity of the samples, and is zero for consecutive samples. The power percentage is the ratio of the absolute power of segmented to unsegmented signals (Equation (10), and the power crossing indicates the point where the Chirplet spectrum crosses its mean.

$$PP = \frac{P_s}{P_{us}} \quad (10)$$

In Equation (10), P_s, P_{us} represent the power in the segmented and unsegmented signal, respectively.

Cumulative distribution frequency (CDF)-based features: These are features based on the power profile of the signals. The major steps used to calculate these features are shown in Fig. 8-c. Initially, a signal amplitude is converted into an absolute amplitude, which is then converted into a power profile using Equation (11) given below. For this study, the power profile is decimated such that N is equal to 90, as shown in Fig. 8-a.

$$X(K) = \sum_{n=1}^{K} x[n] \qquad \text{n, K } = 1, 2, \cdots\cdots \text{N} \quad (11)$$

The total number of features used in the study is 370, as shown in Table 2.

• **Classification network:** The support vector machine is a classifier which uses a decision boundary, support vectors and hyper-planes to distinguish between classes. Support vectors are placed on the positive (Class 1) as well as negative hyper-planes (Class 2). The support vector machine tries to increase the distance between both hyper-planes in order to distinguish classes [20, 21]. The following is the model used for classification:

$$y = \begin{cases} +1, if\ \overrightarrow{X} \bullet \overrightarrow{\omega} + \overrightarrow{b} \geq 0. \\ -1, if\ \overrightarrow{X} \bullet \overrightarrow{\omega} + \overrightarrow{b} < 0. \end{cases} \quad (12)$$

In Equation (12), X, w, b and '•' are the input (features), weights, biases and dot operator. As the model in the equation is linear model, w and b represent the slope and y-intercept, respectively. Moreover, the dot product is conducted between the features and weights, and the result is added to the biases. From Equation (12), it is apparent that Equation $\overrightarrow{X} \bullet \overrightarrow{\omega} + \overrightarrow{b} > 0$ points to one class, i.e., a healthy cow, in the

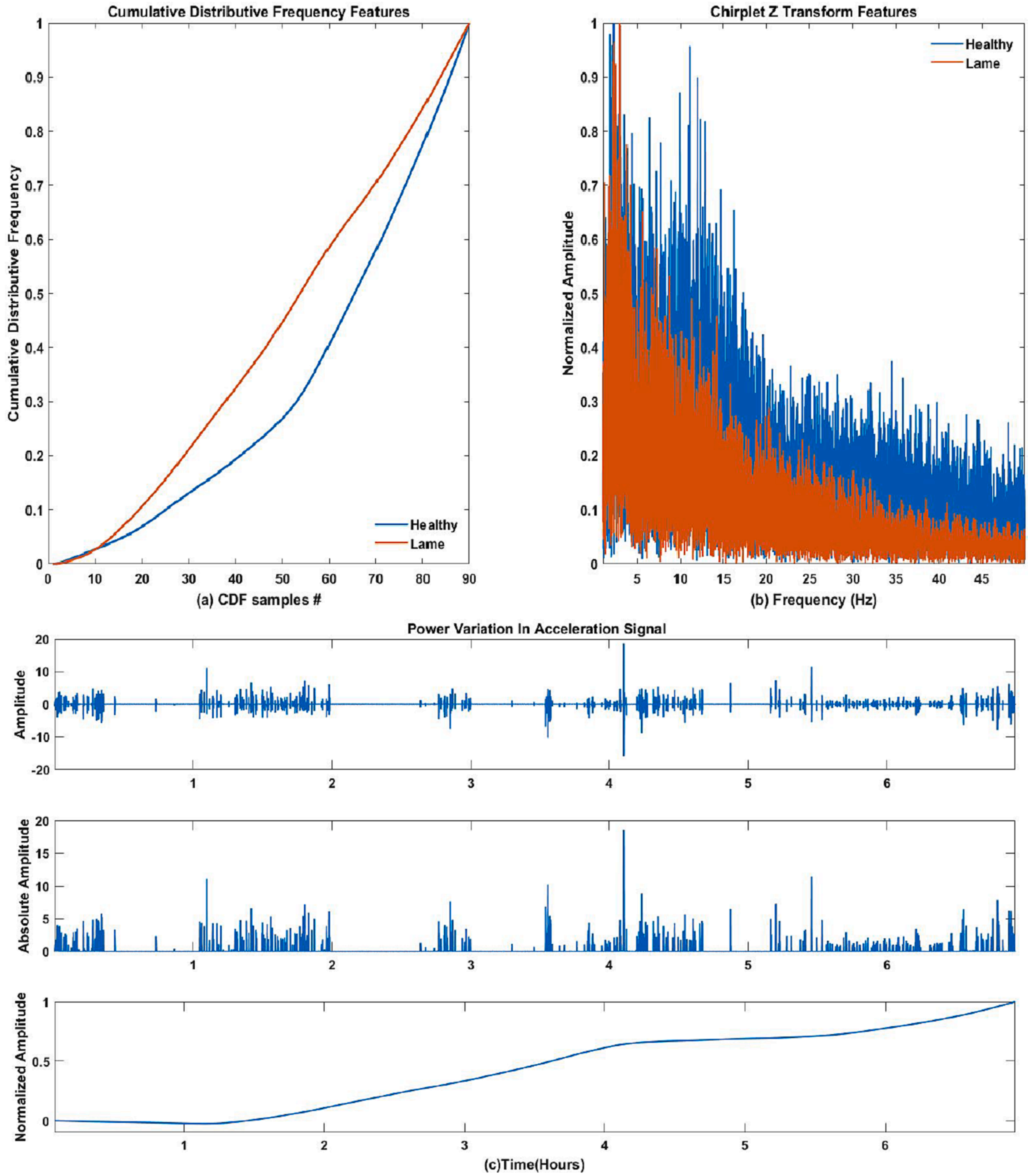

Fig. 8. Features used during the study for classification and procedure to calculate cumulative distribution feature. (a, b) show the cumulative distribution frequency and Chirplet Z transform-based features. From the figures, healthy and lame cows can be discriminated using cumulative distribution feature as the slope of healthy and lame profiles can easily be differentiated. Similarly, the distribution of the spectral contents of the lame cow is different from healthy cow's. (c) First, the signal amplitude is converted into an absolute amplitude. Then, Equation (11) is used to convert the absolute amplitude into cumulative features, during which decimation is also applied.

present study. A failure of the condition mentioned in Equation (12) will result in the assignment of label of another class in the case of binary classification. A pictorial representation of a support vector machine is shown in Fig. 6.

For a binary SVM, a linear SVM with a third-order polynomial kernel function was used, along with an iterative single data algorithm solver. Default values were used for box constraints (Cost ([0 1]), empirical prior probabilities, and nonstandard predictor values, and initial weights were set to [1 1]), with zero probability for outliers. In the present study, K-fold validation was also applied, giving binary classification in terms of healthy versus lame cows. The classification system results are given in Section 3. The technique presented above was implemented using MATLAB 2022a on a Dell Inspiron 15 7000 Gaming series computer with 16 Gb RAM, Ci7, 7th Generation, and 4 Gb GPU RAM.

3. Results

The performance of the presented technique was validated using the minimum criterion given above, i.e., binary classification (healthy/lame cow), along with ablation studies. The criteria applied were named Protocol-I and Protocol-II hereafter, and evaluation metrics used in the study are given below in Equation 13.

In Equation (5), True positive (TP) and True negative (TN) refer to the situations when a 'healthy cow' is detected as a 'healthy cow' and a 'lame cow' is not detected as a 'healthy cow'. False positive (FP) and false negative (FN) convert a 'lame cow' into a 'healthy cow' and a 'healthy cow' is detected as a 'lame cow', respectively. Using these definitions, sensitivity, specificity, precision and accuracy evaluation measures are defined in Equation 13.

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{13}$$

Along with the above measures, an area under curve (AUC) using a receiver operator characteristics (ROC) curve is also used. An ROC curve is a graphical tool to show the performance of a classification model at all classification thresholds. It is plot of a false positive rate (1-specificity) along the x-axis and of a true positive rate, i.e., sensitivity, along the y-axis. As mentioned above, the AUC can be calculated using the ROC curve. The larger the area, the greater the classification accuracy of the classifier used.

Protocol I. During Protocol I, all features from all signals, i.e., acceleration, gravity, angular position and angular velocity, are used. Hence, all 12 signals and features calculated from these signals are used. Table 3 and Fig. 9 present the performance of the system. The table shows that the best and worst accuracies are 77 % and 46 %, respectively. The areas under the curves are 0.69 and 0.38, calculated using ROC curves for the biggest and the smallest areas, and these are shown in Fig. 9 (a, c).

Protocol II. Protocol II assesses the most discriminating signals and features (Table 4). From the results, it can be inferred that the angular velocity represents the most discriminating signals (Table 4-a). Similarly, the performance of CZT and CDF features is similar as they represent the majority of features used for classification (Table 4-b).

The data distribution used during the study was 70/30 %, where 70 % represents the training data and 30 % represents the test data.

Table 3
System performance for Protocol I. Precision (Pre.), Sensitivity (Sen.), Specificity (Spe.), and Accuracy (Acc.) vary between 46 and 77 % for worst to best case scenarios, where H and L represent healthy and lame cows, respectively.

	H	L	Avg.	H	L	Avg.	H	L	Avg.
TP	4	2	3	3	4	3.5	5	5	5
FP	4	3	3.5	2	4	3	1	2	1.5
FN	3	4	3.5	4	2	3	2	1	1.5
TN	2	4	3	4	3	3.5	5	5	5
Pre.	50	40	46	60	50	54	83	71	77
Sen.	57	33	46	43	67	54	71	83	77
Spe.	33	57	46	67	43	54	83	71	77
Acc.	46	46	46	54	54	54	77	77	77

Moreover, for all protocols, 10-fold cross-validation was used along with the area under the curve to compare the system performance. The performance is given in Tables 3 and 4 and in Fig. 9. The tables and figure show that the selected features can accurately classify the given signal.

4. Discussion

This section covers general information about major contributions, future recommendations and comparisons with the state of the art. The study's main contributions encompass public sharing of the database, signal analysis, and the impact of machine learning algorithms. Additionally, the research recommends the development of handheld devices for lameness detection.. Finally, a brief comparison with the state of the art is carried out.

4.1. Public sharing of data

From the studies mentioned in the introduction (Section 1) and in Table A (Appendix A, Supplementary material), we can see that various sensor-based techniques are being used to predict lameness. However, none of the studies mentioned share data in the public domain, and any referenced data is either commercially available or private in nature (Chapinal and Tucker, 2012). To the best of our knowledge, the present study is the only one that shares a database for public research and development. Another unique feature of the study is that it shares complete information on the hardware and firmware utilized. In Section 2, we presented the information on the hardware used, composed of an Apple iWatch 6, an Apple iPhone, and iCloud. Although we used this specific equipment, any other smart watch, smart phone and cloud service combination could also be used. The only proprietary component of our study is the App, which is used to convert data from smart watches to a smart phone. The presented detailed information can be used for future development of similar lameness detection projects using a sensor-based methodology.

4.2. Signal analysis

The data we shared consists of four different types of physical signals, which can be processed individually, as shown in Table 1. To further stress the individual characteristics of these signals, a data example is plotted in Fig. 10. From the figure, we can see that acceleration and gravity signals have more dispersion in their values than does the gravity signal. Similarly, roll, yaw, and pitch are more reflective of the digital behavior of the magnetometer. Moreover, the morphologies of the signals during very low amplitudes, as indicated by points 'A', are very different. With the preceding information, the lameness assessment can be enhanced.

Furthermore, information from similar properties could be used for the binary, as well as multi-class, assessment of lameness. Another aspect of the database is its average duration of observation, which is approximately 6.7 h; this is helpful when it comes to assessing the individual motion of a cow. Using our database alongside an observation study in a case-control scenario of healthy versus lame cows, we can thus extend the analysis to identify a cow.

4.3. Performance enhancement of machine learning algorithm

In Section 1, it was highlighted that lameness is caused by many factors. Therefore, to improve the average performance, new features must be added. Moreover, the information extracted from these features can be enhanced by calculating them at high resolution. That means that instead of using features composed of 370 points, they could be calculated at 1 k or even 2 k resolution. Features from multiple domains could be used to enrich this information. For example, employing the Hilbert transform enables to extract the lower and upper envelopes of a signal, thereby representing a time-based analysis. Moreover, time–frequency

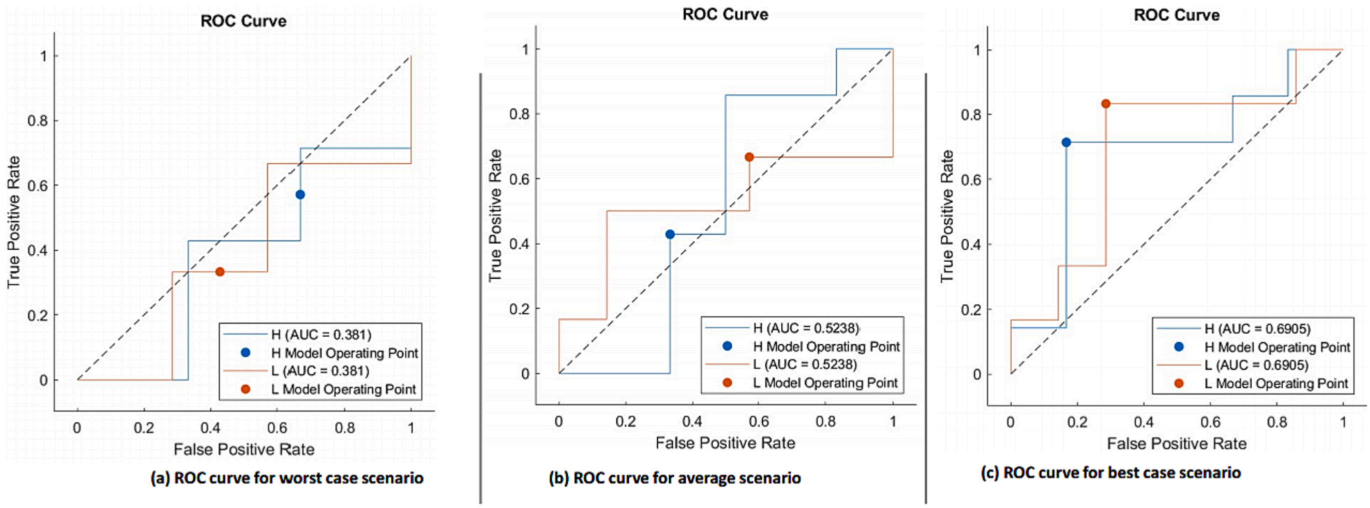

Fig. 9. Performance of the system using Protocol I. During this protocol, the system is tested using all features from all signals. (a, b) Worst and average case scenario, where the area under the curve (AUC) varies between 0.38 and 0.69. (c) A significant increase can be observed in AUC for the best case scenario, where H and L represent healthy and lame cows, respectively.

Table 4
Results of ablation studies with respect to the selection of the most discriminating signals and features. Results from Table 4-a show that angular velocity signals are the most discriminating. H, L, and Avg. represent healthy, lame, and average cows, respectively.

(a)Ablation study for selection of the most discriminating signals

	Acceleration signals			Gravity signals			Angular position signals			Angular velocity signals		
	H	L	Avg.	H	L	Avg.	H	L	Avg.	H	L	Avg.
TP	1	6	3.5	0	5	2.5	3	3	3	6	2	4
FP	0	6	3	1	7	4	3	4	3.5	4	1	2.5
FN	6	0	3	7	1	4	4	3	3.5	1	4	2.5
TN	6	1	3.5	5	0	2.5	3	3	3	2	6	4
Pre.	100	50	54	0	42	38	50	43	46	60	67	62
Sen.	14	100	54	0	83	38	43	50	46	86	33	62
Spe.	100	14	54	83	0	38	50	43	46	33	86	62
Acc.	54	54	54	38	38	38	46	46	46	62	62	62

(b) Ablation studies for CZT and CDF features

	CZT Features									CDF Features								
	H	L	Avg.	H	L	Avg.	H	L	Avg.	H	L	Avg.	H	L	Avg.	H	L	Avg.
TP	3	2	2.5	3	3	3	4	5	4.5	3	2	2.5	4	3	3.5	6	2	4
FP	4	4	4	3	4	3.5	1	3	2	4	4	4	3	3	3	4	1	2.5
FN	4	4	4	4	3	3.5	3	1	2	4	4	4	3	3	3	1	4	2.5
TN	2	3	2.5	3	3	3	5	4	4.5	2	3	2.5	3	4	3.5	2	6	4
Pre.	43	33	38	50	43	46	80	63	69	43	33	38	57	50	54	60	67	62
Sen.	43	33	38	43	50	46	57	83	69	40	33	38	57	50	54	86	33	62
Spe.	33	43	38	50	43	46	83	57	69	33	43	38	50	57	54	33	86	62
Acc	38	38	38	46	46	46	69	69	69	38	38	38	54	54	54	62	62	62

.

decomposition could be used to get information at multiple levels. In this regard, techniques such as Wavelet, empirical mode decomposition, and others are useful.

The above-detailed techniques can be further divided into goal-oriented and data-driven techniques. We can also use numerical techniques such as principal component analysis, non-negative factorization, and singular vector decomposition (SVD). All the techniques can be used individually or can be combined. For example, SVD can be applied after decomposing the signal using Wavelets. However, the main purpose is to extract and enrich information, which is useful for

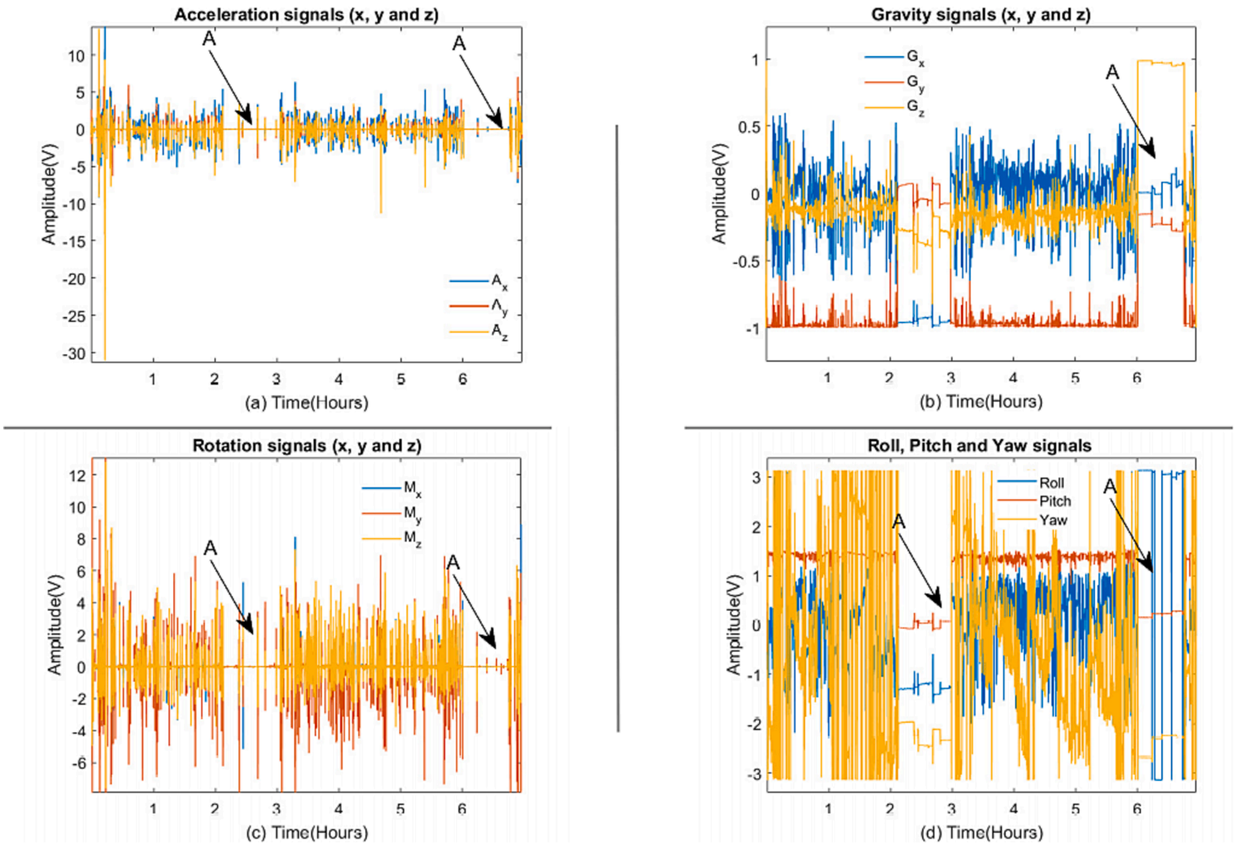

Fig. 10. Details of the signals within multi-sensor data for a cow. (a) Acceleration signals oriented along the x-, y- and z-coordinates. (b) Gravity signals are also oriented along the x-, y- and z-coordinates. (c-d) Two types of rotational signals, i.e., angular rotation along the x-, y-, and z-coordinates and roll, pitch, and yaw signals. The arrow 'A' points to the duration of the signals, noting where activity is at a minimum.

detecting lameness. Nevertheless, the machine learning technique could be further enhanced as indicated. However, these techniques lack the sequential information that is present in the signal. As such, deep learning techniques could be explored to use this sequential information.

We have presented the specifics of binary classification in the results section. Additionally, the database provides an opportunity for conducting multi-class classification experiments, as highlighted in the benchmark section (Section 2.3) for the benefit of the research community. Building upon the framework employed for binary classification, our multi-class classification achieved a precision of 0.21, a sensitivity of 0.20, a specificity of 0.80, and an accuracy of 0.26. These results indicate that our data exhibits variance that can be harnessed, albeit requiring the incorporation of high-resolution features to validate the aforementioned proposals. This underscores the positive potential for further enhancements and advancements in our research.

4.4. Development of hand-held devices for lameness detection

During this study, we also shared our base system for binary classification to allow objective comparisons with future studies. This base system uses the support vector machine (SVM) as a machine learning technique, which is a supervision-based strategy. However, other classification techniques such as clustering and Monte Carlo can also be explored. Clustering is an example of an unsupervised- based technique and Monte Carlo represents reinforcement learning. The major difference between the supervisory and non-supervisory approaches is the availability of labeled data for classification as supervision-based techniques need initial training before being used during evaluation. Similarly, semi-supervision-based methods can also be used. Another perspective for exploration could be hardware implementation in real-time processing. It represents the transfer of such algorithms to controller-based hardware such as the Raspberry pi (Raspberry pi, 2023) and Atmega and Alf and Vergart's Risc (AVR Microchip Technology, 2023).

4.5. Comparison with existing studies

As highlighted in the introduction (Section 1), an objective comparison with existing studies is not possible due to variations in data (number of files, sampling frequency, duration, etc.). Moreover, in the Data Sampling section (Section 2.1.4), we also highlight that different hardware could be used. This approach maintains a fair comparison by keeping the underlying framework consistent. For such a comparison, we also provide a base algorithm which could be used to that end in future studies. It should also be pointed out that the data provided ($43 \times 11{,}518 = 495{,}274$) can be used in developing both machine learning- and deep learning-based lameness detection systems as similar (and even smaller) data volumes have already been used in other studies (Borghart et al., 2021($164 \times 3799 = 623{,}036$); Jarchi et al., 2021($23 \times 25624 = 589{,}352$); Van De Gucht et al., 2017 ($45 \times 1240 = 55{,}800$); Van Hertem et al., 2016 ($242 \times 3629 = 878{,}218$); Thorup et al., 2015 ($348 \times 959 = 333{,}732$); Kokin et al., 2014 ($33 \times 481 = 15{,}873$) and Van Hertem et al., 2014 ($186 \times 744 = 138{,}384$)). Hence, our study represents a breakthrough in the field of lameness detection using inertial data.

The raw data from sensors could be used in multiple ways; for example, it could predict lameness based on the activity of a cow. As shown in Fig. 10, Point A shows the part of a signal where the cow's activity level is at a minimum, and such an area could represent a resting/zero-activity position. The relationship between activity and resting position could be explored. Moreover, in our system, the data is being recorded as the cow goes about its daily activities, as the sensor is attached in the morning while the cow is in the milking parlor, and is removed during the second session, which is generally after 5–7 h. As such, the daily life of a cow is recorded, which can also be used for a transition of the health level from healthy to lame.

5. Conclusion

Our study proves that different sensors in close proximity to one another can be embedded to create a multi-sensor which is lightweight and capable of doing recordings for long periods. For dairy farms, this data is useful for lameness detection, as we have successfully correlated lameness with recorded inertial data. We also have hypothesized that the type of sampling covered herein can be useful for other quadrupeds, such as goats, pigs, camels, horses, etc., because of their similarity with the cows' structure.

By introducing our database, CowScreeningDB, and making it publicly available, we therefore contribute significantly to removing limitations faced in obtaining lameness data. Moreover, we have presented a base algorithm along with a benchmark allowing objective comparisons. A number of analysis techniques are proposed, with deep learning, which considers the sequential nature of data, coming out on top.

CRediT authorship contribution statement

Shahid Ismail: Investigation, Methodology, Resources, Software, Validation, Visualization, Formal analysis, Writing – original draft, Writing – review & editing. **Moises Diaz:** Conceptualization, Investigation, Data curation, Formal analysis, Funding acquisition, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. **Cristina Carmona-Duarte:** Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Jose Manuel Vilar:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Miguel A. Ferrer:** Conceptualization, Investigation, Formal analysis, Funding acquisition, Project administration, Resources, Software, Supervision, Validation.

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.

Data availability

The authors do not have permission to share data.

Acknowledgments

This research was supported by the Spanish MINECO (PID2019-109099RB-C41 project), the CajaCanaria and the la Caixa bank grant, 2019SP19. The authors would like to thank Eva Soler, a veterinarian from the Gran Canaria agricultural products company, Spain, for her assistance during the data collection from cows.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2023.108500.

References

Aldoseri, A., et al. (2023). Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges. Applied Sciences 13.12. 7082.

Antanaitis, R., Juozaitienė, V., Urbonavičius, G., Malašauskienė, D., Televičius, M., Urbutis, M., Baumgartner, W., 2021. Impact of Lameness on Attributes of Feeding Registered with Noseband Sensor in Fresh Dairy Cows. Agriculture 11 (9), 851.

Artrith, N., Butler, K.T., Coudert, F.-X., Han, S., Isayev, O., Jain, A., Walsh, A., 2021. Best practices in machine learning for chemistry. Nat. Chem. 13 (6), 505–508.

Barker, Z.E., Vázquez Diosdado, J.A., Codling, E.A., Bell, N.J., Hodges, H.R., Croft, D.P., Amory, J.R., 2018. Use of novel sensors combining local positioning and acceleration to measure feeding behavior differences associated with lameness in dairy cattle. J. Dairy Sci. 101 (7), 6310–6321.

Beer, G., Alsaaod, M., Starke, A., Schuepbach-Regula, G., Müller, H., Kohler, P., Steiner, A., Weary, D., 2016. Use of extended characteristics of locomotion and feeding behavior for automated identification of lame dairy cows. PLoS One 11 (5).

Borghart, G.M., O'Grady, L.E., Somers, J.R., 2021. Prediction of lameness using automatically recorded activity, behavior and production data in post-parturient Irish dairy cows. Ir. Vet. J. 74 (1).

Bruijnis, M.R.N., Hogeveen, H., Stassen, E.N., 2010. Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model. J. Dairy Sci. 93 (6), 2419–2432.

Byabazaire, J., et al. (2019). Lameness detection as a service: application of machine learning to an internet of cattle. In 2019 16th IEEE annual consumer communications & networking conference (CCNC) (pp. 1-6). IEEE.

Catillo, M., Del Vecchio, A., Pecchia, A., Villano, U., 2022. Transferability of machine learning models learned from public intrusion detection datasets: the cicids2017 case study. Softw. Qual. J. 30 (4), 955–981.

Celi, L.A., Citi, L., Ghassemi, M., Pollard, T.J., Mueck, L.A., 2019. The PLOS ONE collection on machine learning in health and biomedicine: Towards open code and open data. PLoS One 14 (1).

Chapinal, N., Tucker, C., 2012. Validation of an automated method to count steps while cows stand on a weighing platform and its application as a measure to detect lameness. J. Dairy Sci. 95 (11), 6523–6528.

de Mol, R.M., André, G., Bleumer, E.J.B., van der Werf, J.T.N., de Haas, Y., van Reenen, C.G., 2013. Applicability of day-to-day variation in behavior for the automated detection of lameness in dairy cows. J. Dairy Sci. 96 (6), 3703–3712.

Dekker, Ronald. (2006). The importance of having data-sets.

Eriksson, H.K., Daros, R.R., von Keyserlingk, M.A.G., Weary, D.M., 2020. Effects of case definition and assessment frequency on lameness incidence estimates. J. Dairy Sci. 103 (1), 638–648.

Frondelius, L., Lindeberg, H., Pastell, M., 2022. Lameness changes the behavior of dairy cows: daily rank order of lying and feeding behavior decreases with increasing number of lameness indicators present in cow locomotion. J. Vete. Behav. 54, 1–11.

Garcia, E., Klaas, I., Amigo, J.M., Bro, R., Enevoldsen, C., 2014. Lameness detection challenges in automated milking systems addressed with partial least squares discriminant analysis. J. Dairy Sci. 97 (12), 7476–7486.

Griffiths, B.E., Mahen, P.J., Hall, R., Kakatsidis, N., Britten, N., Long, K., Robinson, L., Tatham, H., Jenkin, R., Oikonomou, G., 2020. A prospective cohort study on the development of claw horn disruption lesions in dairy cattle; furthering our understanding of the role of the digital cushion. Front. Vet. Sci. 7.

Hettinga, S., van 't Veer, R., Boter, J., 2023. Large scale energy labelling with models: The EU TABULA model versus machine learning with open data. Energy 264.

Hu, W., et al., 2020. Open graph benchmark: Datasets for machine learning on graphs.". Adv. Neural Inf. Proces. Syst. 33, 22118–22133.

Hu, G.-S., Zhu, F.-F., 2011. An improved chirplet transform and its application for harmonics detection. Circuits Syst 2 (3), 107–111.

Jain, A., et al. (2020). Overview and importance of data quality for machine learning tasks. Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining.

Jarchi, D., Kaler, J., Sanei, S., 2021. Lameness detection in cows using hierarchical deep learning and synchrosqueezed wavelet transform. IEEE Sens. J. 21 (7), 9349–9358.

Jiang, B., et al., 2022. Dairy cow lameness detection using a back curvature feature. Comput. Electron. Agric. 2022 (194), 106729.

Jiang, B.o., Wu, Q., Yin, X., Wu, D., Song, H., He, D., 2019a. FLYOLOv3 deep learning for key parts of dairy cow body detection. Comput. Electron. Agric. 166.

Jiang, B.o., Song, H., He, D., 2019b. Lameness detection of dairy cows based on a double normal background statistical model. Comput. Electron. Agric. 158, 140–149.

Jiang, B.o., Yin, X., Song, H., 2020. Single-stream long-term optical flow convolution network for action recognition of lameness dairy cow. Comput. Electron. Agric. 175.

Kamphuis, C., Frank, E., Burke, J.K., Verkerk, G.A., Jago, J.G., 2013. Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological characteristics of dairy cows to detect lameness. J. Dairy Sci. 96 (11), 7043–7053.

Kang, X., Zhang, X.D., Liu, G., 2020. Accurate detection of lameness in dairy cattle with computer vision: A new and individualized detection strategy based on the analysis of the supporting phase. J. Dairy Sci. 103 (11), 10628–10638.

Kielland, C., Ruud, L.E., Zanella, A.J., Østerås, O., 2009. Prevalence and risk factors for skin lesions on legs of dairy cattle housed in freestalls in Norway. J. Dairy Sci. 92 (11), 5487–5496.

Kokin, E., et al., 2014. Icetag3d™ accelerometric device in cattle lameness detection. Agron. Res. 12 (1), 223–230.

Lemmens, L., et al. (2023). The Combined Use of Automated Milking System and Sensor Data to Improve Detection of Mild Lameness in Dairy Cattle. Animals, 13(7), 1180.

Llebot, C., Steven, V.T., 2019. Peer Review of Research Data Submissions to ScholarsArchive@ OSU: How can we improve the curation of research datasets to enhance reusability? J. eSci. Librariansh. 8, 2.

Maertens, W., Vangeyte, J., Baert, J., Jantuan, A., Mertens, K.C., De Campeneere, S., Pluk, A., Opsomer, G., Van Weyenberg, S., Van Nuffel, A., 2011. Development of a real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: The GAITWISE system. Biosyst. Eng. 110 (1), 29–39.

Martens, B. (2018). The importance of data access regimes for artificial intelligence and machine learning. JRC Digital Economy Working Paper 2018-09.

Martin, E.G., Law, J., Ran, W., Helbig, N., Birkhead, G.S., 2017. Evaluating the quality and usability of open data for public health research: a systematic review of data offerings on 3 open data platforms. J. Public Health Manag. Pract. 23 (4), e5–e13.

Mason, Colin. (2007). Preventing Lameness in Dairy Cows; Hoof Lesions; Their Identification, Treatment, Management and Prevention.

Microchip Technology Inc. https://www.microchip.com/en-us/about/contact-us. May 15, 2023.

Nielsen, L.R., Pedersen, A.R., Herskin, M.S., Munksgaard, L., 2010. Quantifying walking and standing behaviour of dairy cows using a moving average based on output from an accelerometer. Appl. Anim. Behav. Sci. 127 (1-2), 12–19.

O'Leary, Niall W., et al. (2020)b. Grazing cow behavior's association with mild and moderate lameness. Animals 10.4. 661.

O'Leary, N. W., et al. (2020)a. Invited review: Cattle lameness detection with accelerometers. Journal of dairy science 103.5. 3895-3911.

Pastell, M., et al. (2006). Measuring lameness in dairy cattle using force sensors. In CIGR World Congress Proceedings.

Pastell, M., Kujala, M., Aisla, A.-M., Hautala, M., Poikalainen, V., Praks, J., Veermäe, I., Ahokas, J., 2008. Detecting cow's lameness using force sensors. Comput. Electron. Agric. 64 (1), 34–38.

Paullada, A., Raji, I.D., Bender, E.M., Denton, E., Hanna, A., 2021. Data and its (dis) contents: A survey of dataset development and use in machine learning research. Patterns 2 (11), 100336.

Piette, D., Norton, T., Exadaktylos, V., Berckmans, D., 2020. Individualised automated lameness detection in dairy cows and the impact of historical window length on algorithm performance. Animal 14 (2), 409–417.

Poursaberi, A., Bahr, C., Pluk, A., Van Nuffel, A., Berckmans, D., 2010. Real-time automatic lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques. Comput. Electron. Agric. 74 (1), 110–119.

Raspberry pi. https://www.raspberrypi.com/. May 15, 2023.

Riaboff, L., Relun, A., Petiot, C.-E., Feuilloy, M., Couvreur, S., Madouasse, A., 2021. Identification of discriminating behavioural and movement variables in lameness scores of dairy cows at pasture from accelerometer and gps sensors using a partial least squares discriminant analysis. Prev. Vet. Med. 193.

Roche, D.G., Berberi, I., Dhane, F., Lauzon, F., Soeharjono, S., Dakin, R., Binning, S.A., 2022. Slow improvement to the archiving quality of open datasets shared by researchers in ecology and evolution. Proc. R. Soc. B 289 (1975).

Rodgers, C.M., Ellingson, S.R., Chatterjee, P., 2023. Open Data and transparency in artificial intelligence and machine learning: A new era of research. F1000Res 12, 387.

Sahar, M.W., Beaver, A., Daros, R.R., von Keyserlingk, M.A.G., Weary, D.M., 2022. Measuring lameness prevalence: Effects of case definition and assessment frequency. J. Dairy Sci. 105 (9), 7728–7737.

Schlageter-Tello, A., et al., 2014. Manual and automatic locomotion scoring systems in dairy cows: A review. Prev. Vet. Med. 116 (1-2), 12–25.

Shahinfar, S., Khansefid, M., Haile-Mariam, M., Pryce, J.E., 2021. Machine learning approaches for the prediction of lameness in dairy cows. Animal 15 (11). https://doi.org/10.1016/j.animal.2021.100391.

Shimron, E., Tamir, J.I., Wang, K.e., Lustig, M., 2022. Implicit data crimes: Machine learning bias arising from misuse of public data. Proc. Natl. Acad. Sci. USA 119 (13).

Sprecher, D.J., Hostetler, D.E., Kaneene, J.B., 1997. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. Theriogenology 47 (6), 1179–1187.

Taneja, M., Byabazaire, J., Jalodia, N., Davy, A., Olariu, C., Malone, P., 2020. Machine learning based fog computing assisted data-driven approach for early lameness detection in dairy cattle. Comput. Electron. Agric. 171.

Thompson, A.J., Weary, D.M., Bran, J.A., Daros, R.R., Hötzel, M.J., von Keyserlingk, M. A.G., 2019. Lameness and lying behavior in grazing dairy cows. J. Dairy Sci. 102 (7), 6373–6382.

Thorup, V.M., Munksgaard, L., Robert, P.-E., Erhard, H.W., Thomsen, P.T., Friggens, N. C., 2015. Lameness detection via leg-mounted accelerometers on dairy cows on four commercial farms. Animal 9 (10), 1704–1712.

Thorup, V.M., Nielsen, B.L., Robert, P.-E., Giger-Reverdin, S., Konka, J., Michie, C., Friggens, N.C., 2016. Lameness affects cow feeding but not rumination behavior as characterized from sensor data. Front. Vet. Science 3.

Trisovic, A., Mika, K., Boyd, C., Feger, S., Crosas, M., 2021. Repository approaches to improving the quality of shared data and code. Data 6 (2), 15.

Van De Gucht, T., Saeys, W., Van Weyenberg, S., Lauwers, L., Mertens, K., Vandaele, L., Vangeyte, J., Van Nuffel, A., 2017a. Automatic cow lameness detection with a pressure mat: Effects of mat length and sensor resolution. Comput. Electron. Agric. 134, 172–180.

Van De Gucht, T., Saeys, W., Van Nuffel, A., Pluym, L., Piccart, K., Lauwers, L., Vangeyte, J., Van Weyenberg, S., 2017b. Farmers' preferences for automatic lameness-detection systems in dairy cattle. J. Dairy Sci. 100 (7), 5746–5757.

Van Hertem, T., et al., 2016. Lameness detection in dairy cattle: single predictor v. multivariate analysis of image-based posture processing and behaviour and performance sensing. Animal 10 (9), 1525–1532.

Van Hertem, T., Maltz, E., Antler, A., Romanini, C.E.B., Viazzi, S., Bahr, C., Schlageter-Tello, A., Lokhorst, C., Berckmans, D., Halachmi, I., 2013. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. J. Dairy Sci. 96 (7), 4286–4298.

Van Hertem, T., Viazzi, S., Steensels, M., Maltz, E., Antler, A., Alchanatis, V., Schlageter-Tello, A.A., Lokhorst, K., Romanini, E.C.B., Bahr, C., Berckmans, D., Halachmi, I., 2014. Automatic lameness detection based on consecutive 3D-video recordings. Biosyst. Eng. 119, 108–116.

Vázquez Diosdado, J.A., Barker, Z.E., Hodges, H.R., Amory, J.R., Croft, D.P., Bell, N.J., Codling, E.A., Loor, J.J., 2018. Space-use patterns highlight behavioural differences linked to lameness, parity, and days in milk in barn-housed dairy cows. PLoS One 13 (12).

Weigele, H.C., Gygax, L., Steiner, A., Wechsler, B., Burla, J.-B., 2018. Moderate lameness leads to marked behavioral changes in dairy cows. J. Dairy Sci. 101 (3), 2370–2382.

Zhao, K., Bewley, J.M., He, D., Jin, X., 2018. Automatic lameness detection in dairy cattle based on leg swing analysis with an image processing technique. Comput. Electron. Agric. 148, 226–236.

Zhao, L., Ciallella, H.L., Aleksunes, L.M., Zhu, H., 2020. Advancing computer-aided drug discovery (CADD) by big data and data-driven machine learning modeling. Drug Discov. Today 25 (9), 1624–1638.