Automated methods for early prediction of lameness in cattle

Vinoth Vincent Raj

April 22, 2019

Abstract

Lameness in cattle has been identified as one of the major concerns commercial farms are facing. Early identification of lameness not only saves economical costs in milk production for farmers but also helps the animal to get treated early and saves from severe pain. Using automated methods with the aid of sensor technology can be used to continuously monitor the behaviour of cows within a barn which gets rid of traditional time-consuming lameness detection methods.This research study focuses on the prediction of lameness in cows by analyzing the behaviour of cows within the barn. Also, it presents various techniques and approaches that has been done in the past and what techniques could be done in future in finding the lameness in cows.Significant patterns differentiating lame and non lame cows have been presented in this paper. The evaluation metrics that will be used and some of the feature engineering techniques that would be promising in helping the algorithms in detecting the lameness.This paper also discusses two sample approaches to predict lameness based on behaviours of cows within the barn.One is logistic regression which has been done by a previous study but with less variables in the sample 5 day dataset. Also, a simple Convolutional neural network to classify lameness from the heatmaps generated over the 5 days is also implemented with limited images and the paper also discusses the limitations of the approaches and and the future work to carry out in 5 months dataset.

1 Introduction

Lameness is one of the most critical issues concerning domestic animal health. Lameness is termed the third most expensive health issue for dairy cows following mastitis and reduced fertility [7].Issues caused by lameness-based production are disastrous for the farmer. There was a reduction in profit [1] due to a drop in milk and meat production and an increase in costs due to cattle's health care. The lameness-related diseases cost 66 per cattle, with 32 % of that for health care [1]. Detecting lameness in a timely and reliable manner [2][3] is important in order to reduce costs but also to ensure the animal's health. Reducing the health incidents and impact of lameness is therefore one of the biggest hurdles currently facing the dairy industry[8]. Needless to say, several dairy farmers are simply unaware of the number of lame cows in their flock and often do not have sufficient time to treat them if they are noticed[9]. Visual locomotion monitoring has the benefit of being carried out on any farm industry at any given time, taking into account pragmatic considerations such as labor resources. Besides the incoherence and subjectivity of visual monitoringsystems, it can be practically challenging to evaluate individual animals.

In the United Kingdom alone, the predicted cost of treatment to the dairy production, lost milk yield, and lost fertility exceed 128 million annually[4]. Timely lameness treatment can decrease severity and the percentage of treatments required[5], thus lowering economic costs, duration and impact of each animal's pain. Recognizing how disease or welfare status can affect cow's behaviour, barn utilizing patterns and local landscape interactions could provide hugely valuable insights and predictors to monitor and manage a decent range of animal species[6].

The herdsman or veterinarian usually detects lame cows on the basis of adjustments in cow gait, body posture or actions or the presence of hoof lesions during routine trimming. This analysis is the first of cow features that are used in practice to visually detect lame cows. These features are grouped into cow gait, behaviour and posture. Lameness is termed to be an unusual locomotion due to locomotive system pain. Though adjustments in the overall behavior of cattle have been correlated with lameness, such as lying, standing or feeding behaviour, adjustments in locomotion are the most widely used and accurate ways of monitoring lameness.

Lameness can be characterized as the medical symptom of painful illnesses, primarily associated with the locomotive system, leading to seriously impaired motion or divergence from posture or normal gait. The exact nature of lameness may vary from stiffness or reduced limb movement symmetry to an inability to weigh on a limb, or even total recumbency[10,11]. Illnesses can be found in the animal's legs, including painful lesions and abnormalities resulting in physical inability.Furthermore, lameness can also be correlated with nervous system injury (obturator paralysis) and musculoskeletal system (fractures, arthritis, tendonitis).

In horses[17], beef cattle[17] and dairy cows[18,19], accelerometers were previously used to detect lameness. In cattle, lameness was widely observed through differences in daily exercise or differences in lying position[18,19]. In the acceleration signal itself, some studies have tried to identify lameness through a difference. Pastell, Tiusanen[20] distinguished lame and healthy cows using accelerometers mounted on each of the four legs and successive data analysis using variance for each axis together with total acceleration of the legs.

Even though lameness is prevalent and it can not be completely eradicated, early lameness detection will prohibit the state from spreading rapidly within the group. The benefits of early lameness identification can therefore maximize the farm's income, enhance dairy welfare to improve the overall flock quality and reduce the veterinary, medicinal and labor costs[21].

Automated lameness detection approaches include computer-based vision[23] and pressure sensing[22] approaches. These strategies require costly equipment and are restricted to assessing just few moves per cow, which may not be sufficient to detect lameness accurately. Motion sensors were also researched for lameness detection. Many of these strategies monitor the physical activity of a cow (looking down, standing up and walking)[24]. At more advanced levels of the disorder, however, adjustments in physical exercise due to lameness occur. A change in the usual walking pattern (i.e. gait) of a cow is the first measurable symptom of lameness.

In recent years, there has been a massive involvement in deep learning[25]. Among various deep learning models, the most known algorithm is convolutional neural network (CNN), a group of artificial neural networks that has been a powerful method in machine vision tasks

Deep learning only operates well with lots of labeled data, significant computer resources, and modern architectures of neural networks. Here we merge the marked data generated from the architectures of data, present supercomputing, and state-of - the-art deep neural network (DNN) to test how deep learning can optimize information extraction fromx-y coordinate heatmap images. Like any technique, deep learning has biases that need to be taken into consideration, corrected and/or accounted for when using this technology .

Also with the successful implementation of the Convolution Neural Network in picture classification[27] as well as object detection[28], comprehensive features extracted from CNN layers are considered and applied in problem tracking, which solves challenges experienced in deep learning tracking approaches[26] because of the basic heterogeneity between tracking. This problem stems from CNN's robustness in objects classification resulting in inefficient representation of spatial details such as object location. In addition, CNN needs huge amounts of data to learn for classification and detection of objects as opposed to issues with visual object tracking that require only a small amount of samples during processing.

This paper is structured as follows: Section 2 presents the background researches that has been made in lameness and state of the art methodologies used in artificial intelligence domain. Section 3 present the data collection and pre-processing methods that has been implemented on the 5 day data which could also be used on the 5 months dataset . Section 4 describes the analysis and methodologies implemented for lameness prediction. A brief explanation of the Logistic regression and CNN is given. A description of the convolutional neural network used for lameness prediction and its working procedures are presented. Section 5 discusses the results of approaches of the networks and algorithms used. A comparison of the performances of the algorithms used in the 5 day dataset and Section 5 is the conclusion part of the paper.

2 Background

Significant studies on Precision Livestock have been formed in recent times depending on an automated method of machine learning that solves real-life problems. Machine learning was applied primarily to questions related to Precision Agriculture science. For instance, for the exact measurement of soil temperature, the machine learning applied [29]. Another implementation of machine learning deals with soil drying calculation [30] or with daily correct dew point prediction [31]. Even for the exact prediction in the product, machine learning is applied

Automatic behavior measurement linked to lame and non lame locomotion characteristics would allow measurements of everyday activities and might be a better option than traditional approaches[17]. Past studies verified the ability of accelerometers to describe patterns of animal behavior consistently and reliably[32,33,34,35]. Such devices ability to detect changes in lameness - symptoms of movement is yet to be assessed.The use of accelerometers deployed on animals for lameness prediction has been proposed by several authors, with commercial units already available in the dairy industry. .

Domestic and other commercial technology can be used to improve the welfare and performance of dairy cattle. Pedometers, for example, are efficient in sensing increased activity during estrus, thereby helping manage the fertility of dairy cattle [32]. There is, however, limited scope for the classification of cattle behavior using data from the pedometer. On the other hand, accelerometers have been used to identify lying and standing behavior in grazed dairy cattle [33,34]. Nielsen [34] uses a 3-dimensional head-mounted activity logger to determine weeding behavior.

A decision-tree algorithm was developed by [16] that uses tri-axial accelerometer data to detect transition events between lying and standing from a neck-mounted sensor to classify biologically important behaviour in dairy cows such as lying, standing and feeding . Subsequently, [16] they also analyzed barn use patterns to highlight differences in behaviors related to parity, lameness, and days in milk in barn-housed dairy cows where non-lame cows spent a lot of time and had larger site fidelity in the feeding area[16]. By contrast, lame cows spent longer and had higher site fidelity in the barn resting areas than non-lame cows.These findings suggest that differences in individual movement and space-use behavior can be used as indicators of state of health for computerized monitoring of Precision Livestock Farming, which could allow for better diagnosis and treatment and improve welfare of animals for livestock and other controlled species of animals.

Animals that are diseased often have an unusual or decreased activity. Changes in lying behavior in dairy cattle have therefore been used as potential health indicators and predictors, including dystocytes, postpartum disorders and lameness [35,36,37]. In order to mitigate its negative effects, early recognition and treatment of lameness is essential. Changes in lying behavior measures were therefore identified as a serious lameness behavior predictor based on differences in lame and non - lame cow lying responses [35].Lameness causes pain and both milk yield and reproductive performance are reduced [38] making it extremely expensive [39])

Findings on the lying behavior of lame cattle vary between studies. For instance, some scholars reported that lying bout length and variability were higher in lame cows compared to nonlame cows [40,35], whereas others did not report any difference in duration between lame and non-lame cows [41]. Detection of lameness is a struggle for dairy farmers and is therefore sometimes underestimated in its prevalence [42]. Computer controlled monitoring systems based on lying behavior changes could notify farmers to the onset of lameness or a high likelihood of lameness and would greatly benefit farm profitability and cow well-being [43]. Though lying behavior has ability as an predictor of lameness, computerized systems that delivers real-time lameness detection based on lying behavior changes have not been shown to be hugely accurate [37,43].

Conventional statistics often underperform in large datasets with forecast and classification [44], such as farm-based records which usually contain noisy and imbalanced data and are also incomplete. Machinelearning algorithms are generally able to better manage non-linear relationships and input variables interactions. Machine learning algorithms are usually able to better manage non-linear relations and input variables interactions. In a practical setting within a central dairy recording system, such an approach can also be implemented and automated to create alerts. In addition, as the volume and complexity of reported farm data is increasing, particularly with the use of existing data loggers to monitor cow behavior and physiology, machine-learning methods are becoming increasingly interesting in analyzing routine herd data and automatically predicting results [45].

The methods of predicting lameness in cattle vary from research to research. Most experiments use optical technologies to approach lameness. For example, (Song, 2008) attempted to examine the lameness by using high-resolution images and videos or in research [46] trying to identify the lameness by using 2-dimensional and3-dimensional cameras. A further way of dealing with lameness is by using sensors as reported in [47], in which authors tried to record and distinguish the cattle with lameness through the use of force sensors. Applying these methods to behavioural data may offer opportunities to learn more about subtle differences that may occur over time during disease onset. Positional data could also be extended to provide data on energy expenditure and pasture choice for grazing management as well as over time monitoring of health status

Animal Welfare research is no exception, as CNN has attained quality at the expert level in different areas[48,49].Many researchers proved the capacity of deep learning for diabetic retinopathy screening, detection of lymph node metastases and classification of skin lesions, respectively.

Analysis of animal behavior could enhance all biology tasks requiring identification of behaviors and state of health, such as animal tracking and management, habitat examination, and demographic estimation. Throughout this work, we use profound learning, a state-of - the-art machine learning technology that in recent years has led to dramatic improvements in artificial intelligence (AI), especially in computer vision.

3 Materials and Methods

The research was conducted in accordance with the United Kingdom Animal Welfare Act (2006)[16]. Lame cows were managed in accordance with the animal health plan of the farm and all of them.The data are collected from a local commercial cattle farm in Essex, UK. The data have an attribute of data time to capture the position of each cow. The cows are equipped with wireless sensors to monitor the location of the space. The sensors were installed on cows using a neck collar with a counter weight to try and keep the sensor at the top of the neck in a stable position.

3.1 Data Collection

The study was carried out over 5 months on a commercial cattle farmin Essex, UK. But in a period of 5 days from [16], the data mentioned in this report were recorded. Research data were collected for 5 consecutive days in January 2014 on a cattle farm in Essex, UK. A total of 210 Holstein pedigree cows were placed in a 30 m by 60 m rectangular free-stall barn. The cows were divided into groups of high yield of 120 cows and low yield of 90 cows, divided by a central feed alley. The high yield group had exposure to 120 free-stands and a sequential feed space of 0.43 m per cow in the upper part of the barn. The milking parlour and collection yard were located at the bottom of the barn, with a linking return passage on the right hand side of the barn.A total mixed ration was fed to all cows and milking was done at 05:00hrs, 13:00hrs and 21:00hrs)[16].

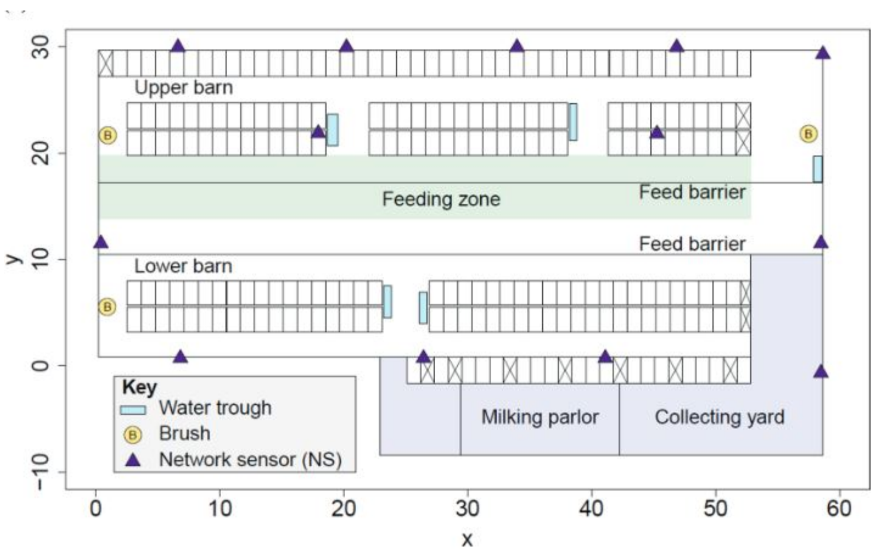

Figure 1: Barn structure

Prior to the research, all cattle in the high return group were scored using ZB's 4-point AHDB Dairy Mobility Score[50] where 3= severe lame and0= sound andat the departure to the milking parlour, and re-scoredby HH in the primary barn the following day[16]. During the prior three months, cows considered to have had a health event were excluded, including foot lesions and mastitis treatments[16].

Two study groups (10 lame cows and 10 non-lame cows) were chosen to compare yield and parity whenever possible in the 5-day data set relying on their mobility scores[16]. Cows with a mobility score were included in the lame group; score 3 cows's (severely lame') not included in the flock because of a handful of cows with this score and for ethical grounds (the selected cows would not be treated until the end of the study)[16]. Without any prior knowledge of the barn-use behavior, the cattle were selected. Individual parity ranged from 1 to 6 years with a mean of3,25 and standard deviation= 1,44and 44 to 220 withmean of125, andstandard deviation= 51,3and average daily milk yield ranged from 28,7 to 58,4 litres with a mean of42,5, and standard deviationof6,88 At the end of the study, all cows are medically tracked for lameness and, if necessary, foot trimming was performed[16].

To monitor the upper portion of the barn, the selected cows were fitted with portable sensors[28,35,47,48]. A RTLS wireless network is formed with a series of 500 sensors capable of measuring comparative spatial locations in (x, y, z) coordinates of each individual detector during the system using arrival time to pinpoint distances of continuous messages sent from each node to its neighbours. Twelve sensors were attached to understand fixed positions around the barn and a further eight are placed within the neighboring collection yard and milking parlour to improve network coverage and triangulation observations[16]. The detectors were discovered to operate well for individual cattle spatial tracking, even though performance was slightly worse than industrially commercialized configuration, which is likely due to metal characteristics in the barn environment interrupting the sensor signals. Using a neck collar, the sensors are placed on the cows.

Repeated location information are acquired for 24 hours a day over the 5 days of the study use a survey rate of 0.125Hz, leading in a theoretical limit of 54,000 location data points collected per cow over the duration of the study. However, location information were excluded during the three daily milking activities, each lasting about 90 minutes when the cattle left the upper barn area, as human intervention at this time restricted cow movement and space-use behavior[16]. In addition, some slight data loss occurred when sensors often encountered battery failure prior to replacement, or when sensor flaw appeared to position a cow outside the barn[16]. In the upper barn area (81 percent of the theoretical maximum) a total of 876,621 location information points are acquired and used in previous data analysis. For all cows, the average amount of location data data points received a day was 8767 (median= 8930), as well as the minimum average number of data points acquired for a single cowover 5 days of study was 8175 data points per day[16].

3.3 Data Preprocessing

The 5 days data from [16] has been used for this research study to analyze and setup up the experimental approaches for 5 months dataset. It could be estimated that a similar kind of data pre-processing steps and approaches would be involved in the 5 months dataset. A simple descriptive statistics has been done followed by some of the most common data pre-processing steps which was implemented in the 5 days dataset have been discussed in the upcoming paragraphs.

The dataset consists of 1080000 instances with 7 columns. The cowids, days, time, x ,y coordinates location of the barn to track cow movement and accelerometer mean and peak to track the activeness of the cow. Firstly the time is epoch format, which has been converted to human readable time format. The data contains 64248 null instances for x and y coordinates which was due to the lag for the sensors to actually start. The null values have been removed from the dataset resulting in 1015752 instances. The instances recorded were for every 8 seconds contains their x, y coordinates in the barn and their accelerometer readings. All the instances were aggregated by a 2 minute window by taking the mean of the x, y coordinates and their accelerometer readings.

The structure of the barn coordinates has been given in [16]. Using this coordinates, new features have been created to track the activity of the individual cows inside the barn. Some of the common activities include feeding, milking and resting. The location of these structures were discussed in the above sections and with the figures. Two new binary features have been generated if the x,y coordinates of the cow fall inside the feeding structure (0.25,10.5,58.65,17.2) and milking structure (29.4,-8.42,42.25,-1.62). These features help us to calculate the time spent by cows in feeding and milking and resting over the 2 minute window. However, these features just track whether the cows are present in the feeding area or in the milking parlour and it does not actually confirm if the cow was actually feeding and milking at that particular time. So, in this study we assume that the cows were actually feeding if it was in the feeding area and vice versa for milking parlour.

Distance moved between each instances can be a potential feature in analysing the behaviour between lameness and non-lame cows. Euclidian distance has been calculated over the 2 minute window data frame. For that old features of x,y coordinate has been created by shifted the x,y coordinate values to one step down where the instance would now have the current x,y coordinates and the old x,y coordinates in the same row. This setup would be helpful in calculating the Euclidean distance travelled by the individual cow between each 2 minute window.The formula used in shown below where x1 - x-coordinate of previous instance, x2- x-coordinate of current instance,y1 - y-coordinate of previous instance,y2- y-coordinate of current instance

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Finally, lame and non-lame cowids have been extracted from [16]. New binary feature has been populated called lame to distinguish between the lame cows and non-lame cows based on the cowid. Since the data has been already labelled in the 5 days data, it is not possible to track the behavioural difference of lame cows before it has been actually identified as lame, however, this could be done in the 5 months data to track the differences in the various metrics for each cows before it has been identified as lame.

These preliminary features will be useful for analysing the statistical methods and machine learning models. However, behaviour analysis can also be carried out using heatmap showing densities according to [16]. So the data set has been aggregated into two types of windows 1 minute window and 10 minute window to add more images for CNN training. Some of the sample heatmaps for lame and non lam cows have been shown in the figure for 1 minute and 10 minute window. Heatmaps would be helpful in getting a overall picture of the behaviour of the cow per day in the barn to find out where it has spent of the time. These heatmaps have been generated for all the 20 cows (10 lame and 10 non lame) per day for 5 days in 10 minute window and 1 minute window. This lead resulting in more heatmap images to train our Convolutional neural network model. Since the data is only for 5 days we had to aggregate it based on days over different time frame windows. However, for the 5 months dataset, the same approaches can be applied for each individual cow by aggregating all days and still can result in many images to train our network.

3.2 Data Exploration

The data has been smoothed by aggregating all the features on 2minute average to reduce the sensor error. The heat map for lame and non lame have plotted which can be shown in the below figure. Randomly lame and non lame cows have been selected. Feeding area and milking area has been plotted using the x,y coordinates for a perticular day to distinguish the area between feeding and resting in cubicles barn. The intensity shows the time spent on that particular area. More intensity generally means that the cow has spent more time.

Columns	Description
Cow-id	Unique identifier of the cow
Day	Day representing the count of days
Time	Represents the specific time when the sensor readings are captured
X-coordinate	X-coordinate position of the sensor in the barn
Y-coordinate	Y-coordinate position of the sensor in the barn
actm	Accelerometer mean
actp	Accelerometer Peak

Table 1: Data description for 5days dataset

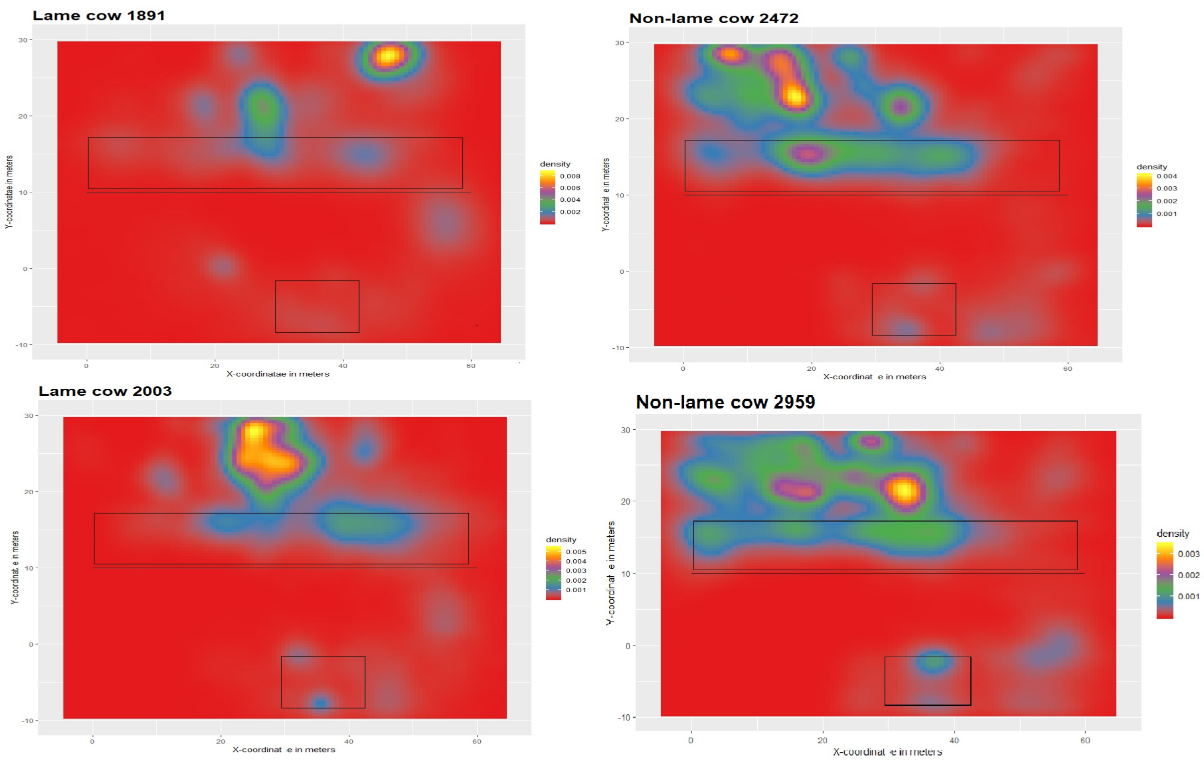

Figure 2: Lame and non-Lame cows heatmap

The rectangle in the middle area of the plot represents the feeding area and the small square at the bottom of the barn area represents the milking parlour in the barn. Clearly some patterns could be observed between lame and non lame cows. A wide spread inside the barn area is observed in non lame cows whereas intensity is higher at particular points for lame cows. Also, non lame cows spend significant amount of time in feeding area and less densities could be observed for lame cows in feeding area.

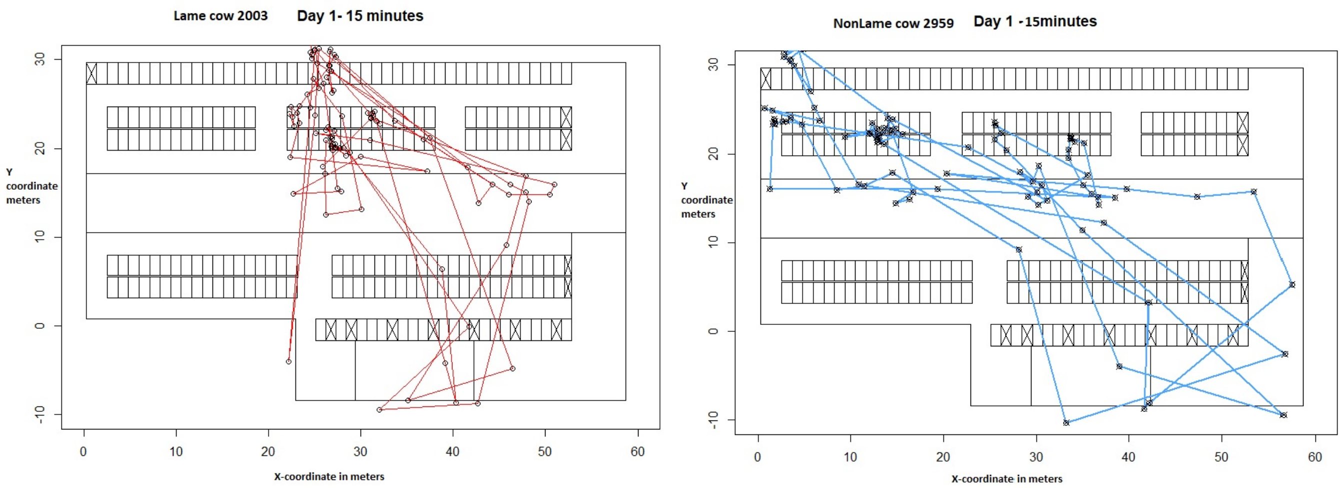

Figure 3: Space use plot inside the barn (Lame and Non-Lame)

The data has still been aggregated by every 15 minutes in order to have an overview of the spread of the cow which could be observed in Figure 3. A plot has been plotted for a particular day for lame and non lame cows.The same pattern which was observed int he above plot could also be observed in the plot. A wide spread of movement could be observed for non lame cows and the movement of lame cows was less spread inside the barn. Non lame cows has spent time in diff rent cubicles in the upper barn for resting whereas lame cows spend time only in selected cows which could be observed in the above plot.

Before the model approach some simple visualization on the data is made to identify key differences between lame and non lame cows. The data has been smoothed over days taking mean for x and y coordinates, actm and the sum of the duration of time spent in feeding and milking. Another data set has been created by aggregating all 5 days for individual cows by taking the mean of x and y coordinates, actm , feeding and milking duration in minutes. Firstly, The variable accelerometer reading mean has been analyzed. This variable denotes the activity or movement of cow at a particular time frame. A boxplot has been plotted to identify the key differences between lame and non lame against the actm feature. The plot could be seen in the figure 4.

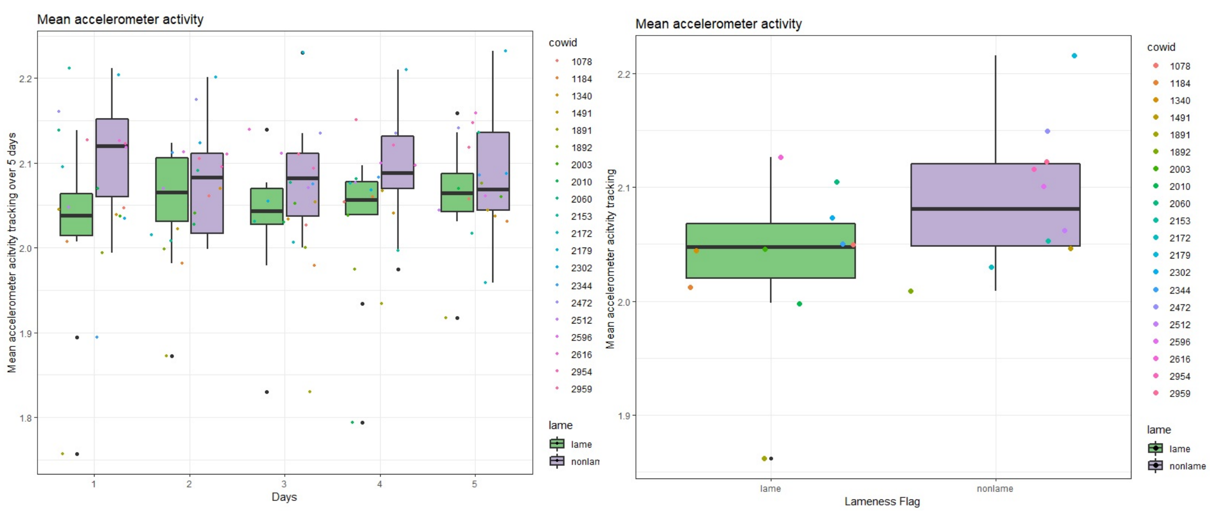

Figure 4: Accelerometer mean value differences over days (Lame and Non-Lame)

Clearly the mean of actm value shows significant difference between lame and non lame cows. The second plot shows that median actm value for non lame cows is higher than lame cows which makes sense as non lame cows has higher activity than lame cows. Some outliers have been observed in the plots, however, the spread of the points for lame cows is below 2.1 and spread is generally greater than 2.05 for non lame cows. Also first plot show the actm values for individual days for all 5 days for all cows distinguishing lame and non lame cows. The findings discussed above was supported by first plot as well. Higher activity could be observed for non lame cows in all five days than lame cows.

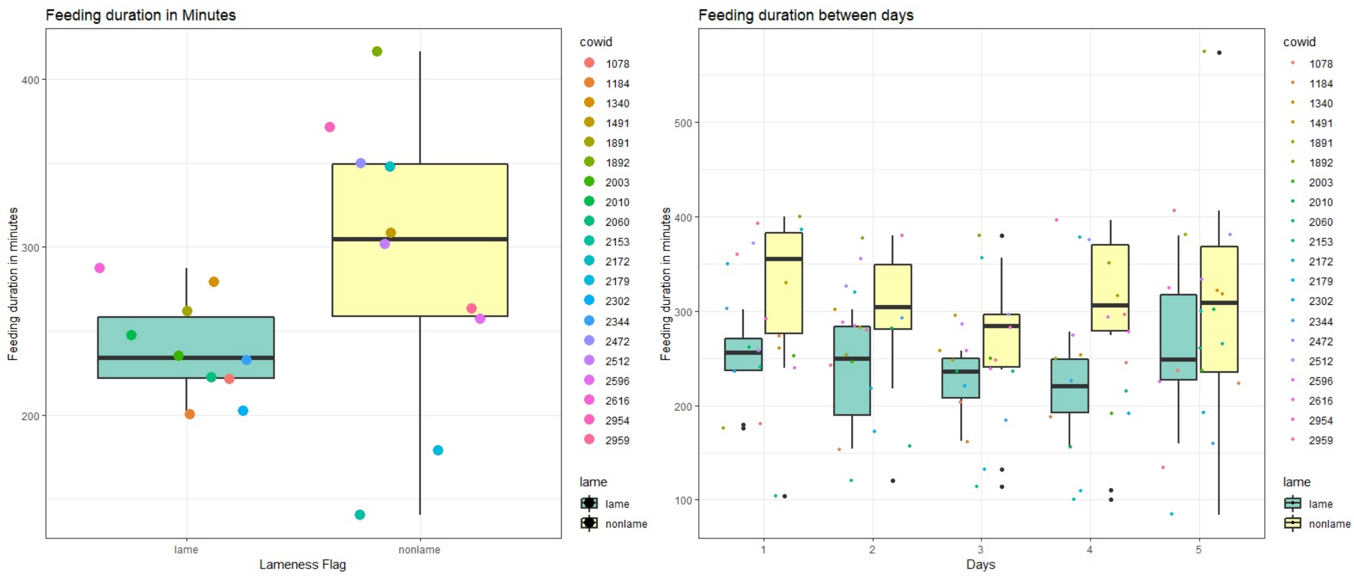

Figure 5: Feeding differences between Lame and Non-Lame cows

Analysis of feeding behaviour could be observed in the above figure. Plot 1 shows the mean feeding duration in minutes per day and it could be observed that for lame cows 25-75 percentage quartile lies between 225 minutes to 260 minutes with median of 235 minutes per day (approximately 4 hours per day)For non lame cows the median value is higher with 305 minutes (approximately 5 hours per day) and the range between 25-75 quartile is between 260 minutes to 350 minutes. Second plot shows feeding duration in minutes for all the 5 days. Clearly, the feeding duration for non lame cows is higher for all 5 five days. Huge difference in feeding duration could be observed in day 4. These plot support that lame cows feed less than non lame cows.

As mentioned earlier, flags populated in feeding feature was completely based on the feeding location of the barn and was not derived based on the activity. Hence, it is possible for cow to be in feeding area and was not actually eating. Hence, seperate dataset has been extracted for the feeding feature which has been reported as feeding. Analysis has been done on this dataset to find out the actm value which could be seen in the below figure.

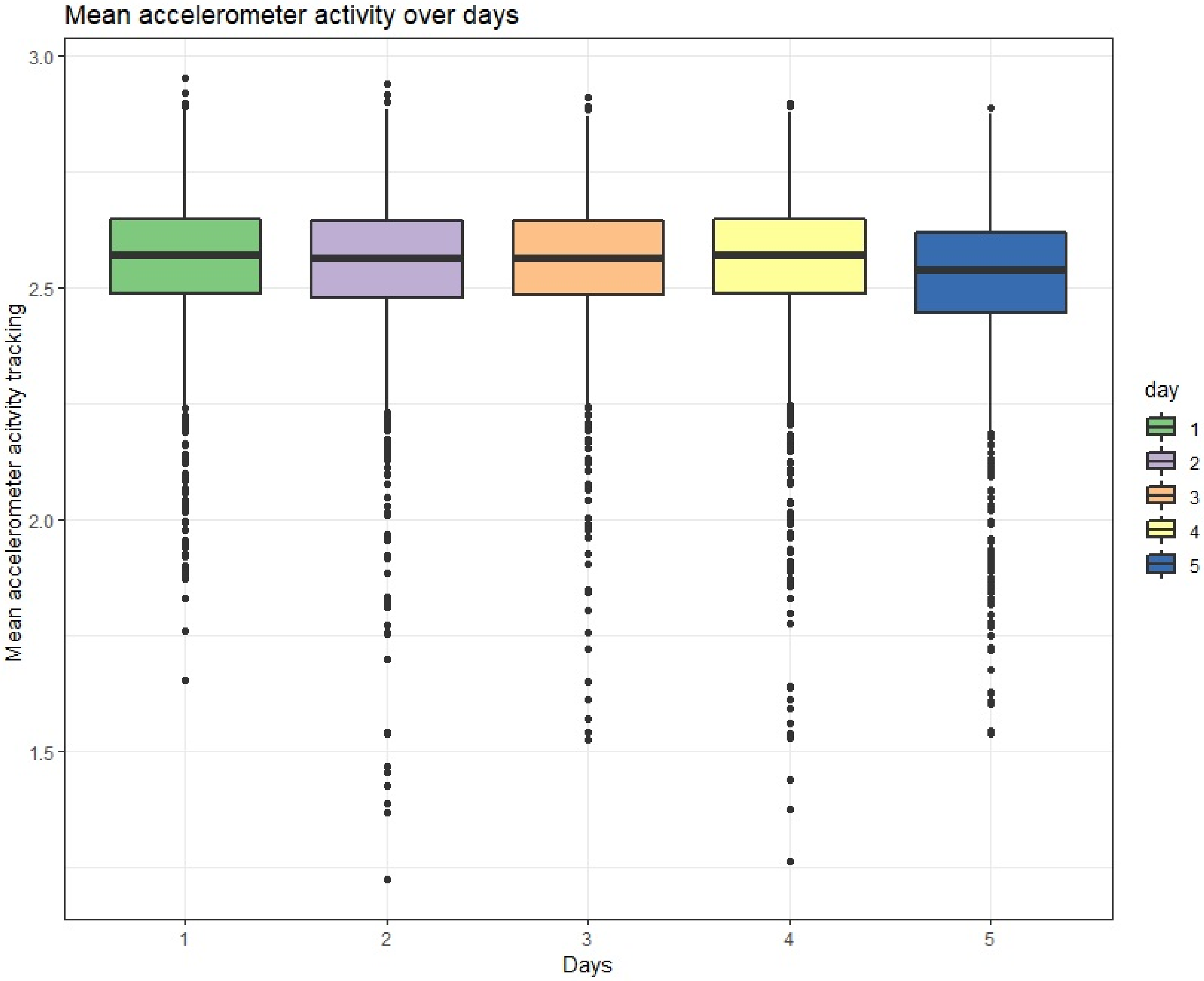

Figure 6: Feeding data with Accelerometer mean value over days to track actual feeding activity

The box-plot has been made showing the days in x axis and the mean actm value for the instances in feeding area. Low actm values corresponds to low activity of cows and hence we could observe many outliers are present in the lower quartile ranges especially below 2.3. It could be regarded that cows having actm values below 2.3 were actually not feeding in the feeding area. However, since we do not have data to support this argument, this could be taken as an assumption.

4 Analysis

This section describes two kinds of approaches in classification of lameness. One is through standard statistical techniques using logistic regression models which was implemented by [16] but without using parity, milk yield details as these data could be hard to get for 5 months dataset and the other approach using non linear method such as Convolutional neural network model as it was proven successfull in detecting patterns in images and heatmaps[27] iand evaluate the performance of both the models in terms of its accuracy. The key aim of this research project is to propose ways in identifying the lameness. So the behavioural differences before lameness and after lameness factors will be taken into consideration only for the 5 months dataset.

5a. Model Selection:

Some of the earlier research papers have proved that linear models works well in identifying lameness over a short period (5 days) of data[16]. However, for 5 months data the same linear models can be tested in terms of their performance and their results can be studied.

Firstly, equal number of lame, non lame and about to become lame in next few weeks cow data will be selected.Since we are dealing with the sensor data, it becomes necessary to aggregate the date based on certain time. Since we have 5 month data, it will be reasonable to aggregate data per day and calculate their behaviours. The following features will be extracted from the data for each of the cows and it has been derived fro the 5 days dataset as discussed earlier.

- Total distance the cow moved per day
- Feeding duration of the cow
- Milking duration of the cow

The respective labels will be our dependent variable 0 for Lameness and 1 for healthy cow.For the sake of analysis, labelling cows which were about to be lame in next x number of days is our primary focus in the 5 months dataset.So, finding the appropriate x number of days before identifying lameness where the observations significantly shows the behavioural differences between healthy and lameness will be identified. However, this could not be carried in 5 days dataset.

One approach in determining the x days before lameness would be to predict probabilities for each class and fitting an ROC curve to find the maximum area under curve for the predicted probability value. Once the model has been trained the coefficients can be tested for different months data.

$$\mathrm{y}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_i + .. + \beta_m x_n + \varepsilon_i,\ i = 1, \ldots, n.$$

where x_1 to x_m are the features derived such as mentioned above and β_0 is the intercept and β_1 to β_m are the coefficients for the features.

The following assumptions must hold in order to validate the multivariate linear regression model. First, a linear relationship must be established between the independent variables and the dependent variables. Second, the independent variables must not be multicollinear. To evaluate this, Variance Inflation Factor (VIF) scores were calculated for each predictor variable, and no scores should behigher than the VIF greater than 10 threshold . Thirdly, the residual model should be distributed normally ; and finally, the data must not contain heteroscedasticity. In each linear model, we use the Q-Q plot and the Breush-Pagan test to test the residuals for normality and heteroscedasticity.

Logistic Regression:
Here we consider a logit link mechanism from a generalized linear regression model of the form to discover the potential predictive capacity of the examined dependent variables to properly identify lameness in individual cows:

$$\frac{p(x)}{1 - p(x)} = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_i + .. + \beta_m x_i + \varepsilon_i}.$$

Where 0 is the intercept , Xi is the corresponding values of the dependent variables in the previous analysis , is the regression coefficients to be determined and,and p is the estimated probability of a cow being classified as lame. To prevent over-fitting the prediction model, we limit the option of model selection to those dependent variables, Xi, in which one or both of the independent variables in the analysis are found to also be significant. Akaike Information Criterion score is also usedto determine the best relative fitting model

5b.CNN Model over heatmaps:

Heatmaps can capture behaviour patterns over any given point of time. Hence it would interesting to develop predictions over the generated heatmap images. In this approach, heatmaps will be generated for each day for lame, non-lame and about to be lame cows and the corresponding labels will be noted. The first 3 months data will be used as the training set and next two months data will be used for validation set. The 3 months data heatmap images will be generated for each day and will be passed into a simple convolutional neural network to identify the patterns in the heatmap images to distinguish lame and non lame cows.

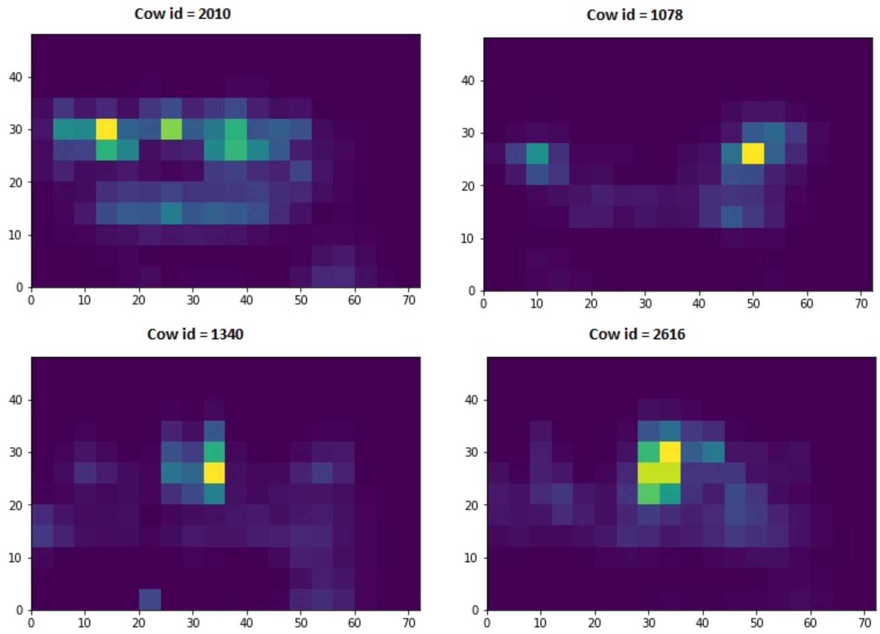

Figure 7: Heatmaps capturing cow behaviours over time

Convolutional Neural Networks have an design different from regular Neural Networks. By placing it through a sequence of hidden layers, regular neural networks transform an input. Each layer consists of a set of neurons in which each layer in the layer is fully connected to all neurons. Eventually, there is a finalfully-connected layer the layer of output representing the forecasts. Convolutional networks are multi-stage trainable architectures with multiple layers for each stage. Each stage's input and output are array combinations called as feature maps. In the event of a colored image, each feature map would also be a 2D array with an input image color channel, a video 3D array, and an audio input 1D array. The output phase reflects features that are derived from all input locations. Generally, each stage comprises of a convolution layer, non-linearity and a layer of pooling. After many convolution and pooling layers, there are a single or multiple fully connected layers.

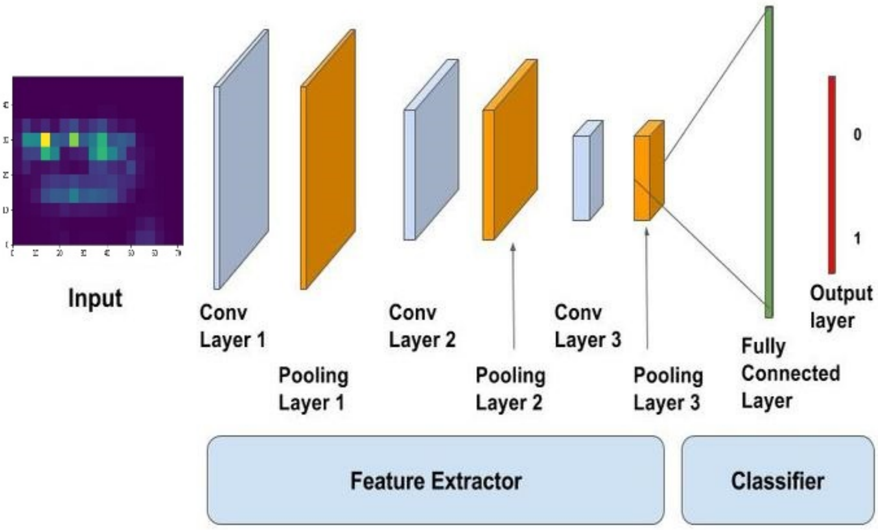

Figure 8: Architecture of CNN

A. Convolution layer

This layer is the core building block of a CNN. The characteristics of the layer are learnable kernels or filters that extend through the input's complete depth. Each neuronin this layer obtains input from a set of neurons in the prior layer located in the small neighborhood. Such a neighborhood in the prior layer is called the receptive field of the neuron. Each filter is converted with input that produces a map during the forward pass. When stacking numerous such feature maps produced from multiple filters, they form the convolution layer output. Share the weight vector generating the feature map that reduces the complexity of the model.Convolution is the first layer in an input image to obtain features. Convolution maintains the relationship between pixels through the use of squares of input data to learn image features. It is a mathematical operation that requires two inputs like the image matrix and a filter or kernel.

A. Strides

Stride is the amount of pixels moving over the matrix of the input. If the step is 1 we shift the filters at a time to 1 pixel. When the step is 2 we move the filters at once and so on to 2 pixels. The figure below shows that convolution would work with a 2 step.

B. Non-linearity Layer

This is a neuron layer that applies different activation functions. These processes implement nonlinearities for multi-layer networks that are desirable. Typically, the activation functions are sigmoid, tanh, and ReLU. Rectified linear units (ReLU) are preferable compared to other functionalities because neural networks train several times faster.

C. Pooling Layer

Pool Layer conducts a function to reduce the input's spatial dimensions and our model's computational complexity. And it controls overfitting as well. It runs on each input depth slice independently. There are various features like pooling Max, average pooling, or pooling L2-norm. Max pooling, however, is the most common type of pooling that takes the most essential partof the volume of the input. The pooling layer that takes small rectangular blocks from the convolution layer and subsamples it to produce a single maximum output from the block[19-21] can be followed by the Convolution layer. Pooling layer gradually reduces the representation's spatial size, thereby reducing the metrics to be calculated. It also supervises overfitting. Besides the maximum function, pooling units can also conduct other functions such as L2-norm pooling.

D. Fully Connected Layer

Perhaps one or more fully integrated layers execute high - level reasoning by taking all the neurons in the prior layer and linking them to each single neuron in the present layer to produce global semanticinformation.

5 Results

This research study explored the background and cow characteristics that were most influential in determining whether behavioural differences could help distinguish lame and non-lame cows. This research also explored the effect of different time on cows activity and the effect of activity determining the health status of the individual cows. Logistic regression was one of the primary analysis used for this research. After the new features have been derived our new data has the following structure

Descriptive statistics of the features are given in the below table. It could be observed that mean of x and y coordinate for all the 20 cows are 30.24 and 19.71 with a standard deviation of 5.65 and 1.91. The mean of the distance is 1514 meters per day for all 20 cows with a standard deviation of 140.608 meters per day. The feeding and milking were also summed up while aggregating for each day. The mean for feeding is 133 minutes per day and the mean for milking is 14.57 minutes per day.

Features	Mean	Standard deviation	Min	Max
X-coordinate	30.24	5.65	15.66	41.23
Y-coordinate	19.71	1.91	13.66	23.88
actm	2.063	0.082	1.757	2.232
Distance m/day	1514	140.608	1212	1894
Feeding min/day	133.1	41.65	42.0	287
Milking min/day	14.57	7.14	0	30

Table 2: Descriptive statistics for continuous variables

Analysis of Logistic Regression Assumptions Linearity of the Logit: The outcomes of the Box-Tidwell check for logit linearity. The regression results showed that linearity assumptions are violated for feeding and y coordinate variable. Tabachnick Fidell, suggest that features that lack linearity with the logit of the target feature can be converted to categorical variables. Since the experiment is used for research purpose ROC curves can be used to estimate the performance of the model.

Problems with multicollinearity consist of including various variables in the model that have a similar predictive relationship with the outcome. This can be evaluated by calculating the VIF value for each predictor.It is necessary to remove any variable with a high VIF value (above 5 or 10) from the model. This results in a simpler model without compromising the accuracy of the model, which is good.All the predictor variables for all cows had low VIF levels ranging from 1.24 to 1.42. The feature that was most alarming and had the highest VIF value for all cows was the actm variable.For

Data	Split	Lame	Non-lame	Total Instances	Total %
LR-Model1	Train	38	37	75	75
aggregated per day	Validation	12	13	25	25
LR-Model2	Train	7	8	15	75
aggregated per cow	Validation	3	1	5	25
Heatmap images	Train	14	14	28	73
from 1 min/10 min	Validation	6	4	10	27

Table 3: Number of observations and model aggregations, training and test data subset.

the heteroscedasticity, Breush-Pagan Test has been tested on the model and got an p value of 0.009 at 5% significance level. therefore we can accept the null hypothesis that the variance of the residuals is constant and infer that heteroscedasticity is not present.

The first thing to do when proceeding with a Logistic Regression model is to pick those factors (predictors) that claim to affect the test results based on intuition or experience. Table 1 presents the contingency table, i.e. the results of the Pearson correlation coefficient (R) between all the indicators assumed that affect the results.

The correlation plot below does not shows a very significant correlation between any of the features to eliminate one of them as a potential model predictor. The possible explanation is that if we accept strongly correlated predictors, they will be in continuous collision with the model; that is, they will share the variance and cause errors in the results of the output

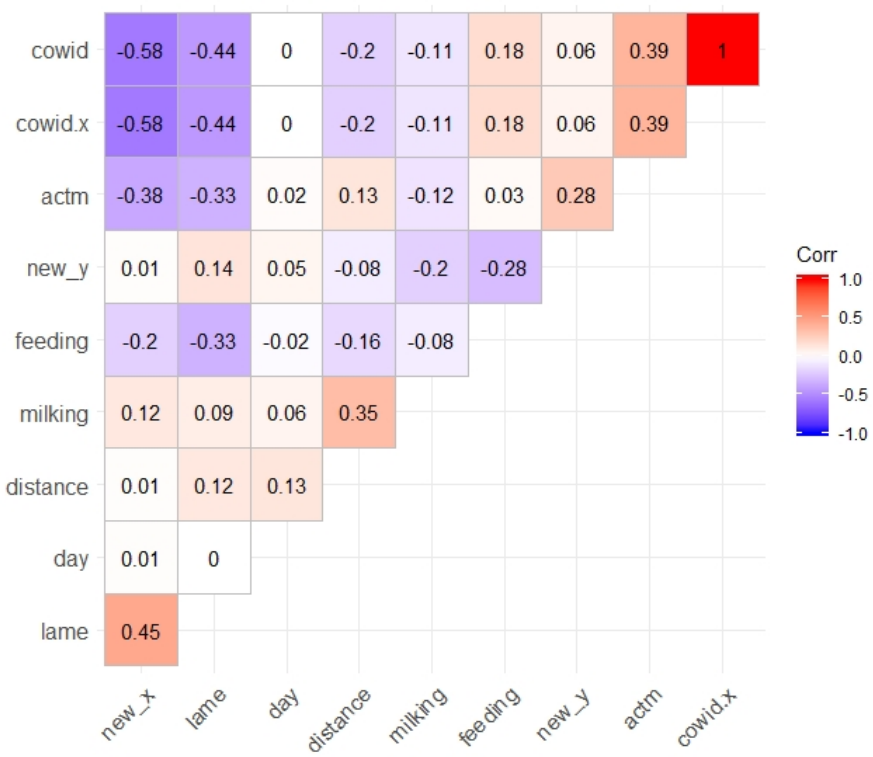

Figure 9: Architecture of CNN

The data has been split over 1 day and 5 day window size by taking mean of the x,y coordinates , distance moved, actm, and sum of milking and feeding duration. The 1 day window has generated over 100 samples for each days for all 20 cows and for 5 day instances the data has been grouped by cowed for all days and the mean of actm, distance, feeding and milking were averaged for all the rows. It results in 20 instances. For training and validation the data has been split into 75% as training and 25 % for testing for both the datasets. A comparative study will be carried out after fitting a logistic regression model as which dataset compression was working better in terms of predicting lameness. The split conversions has been shown in the table below. Proper care has been taken to split the dataset which has almost equal number of lame and non-lame cases. Similarly for CNN model the heatmaps has been generated for 1 minute window to capture the raw densities in heatmaps. Since deep learning requires lot of data to train a further 10 minute window images heatmaps has also been added. The training data has 28 images out of which 14 lame and 14 non lame cases were present and for testing 10 images were generated for which 6 lame and 4 non lame cases were selected.

Data	Test accuracy	Cross validation accuracy	Specificity	Sensitivity	AUC	AIC
1 day	72	77%	0.75	0.6923077	0.83	112.37
5day	0.8	85%	0.8	0.9	0.8535	26.629

Table 4: Results of Logistic Regression models

The results of logistic regression has been shown in the below table for 1 day and 5 days window. It could be observed that the 5 day window where only one instance for each cow was derived by aggregating the mean of actm , distance, x,y coordinates and taking the sum of feeding and milking seconds produced comparatively good results of 80 % test accuracy compared with that 1 day window size of 72% test accuracy. A 10 fold cross validation also has been performed on both the models and data over 5 day window achieved an accuracy of 85% compared with 77% for 1 day window size. The sensitivity and specificity is 0.69 and 0.75 for 1 day window and 0.9 and 0.8 for 5 day window size. To compare the model AIC metric has been calculated for each of the models. Generally, a very low AIC value corresponds to a better model. In this study, the model with 5 day window achieved a low AIC score of 26.629 and the model aggregated by one day has a AIC score of 112.37.

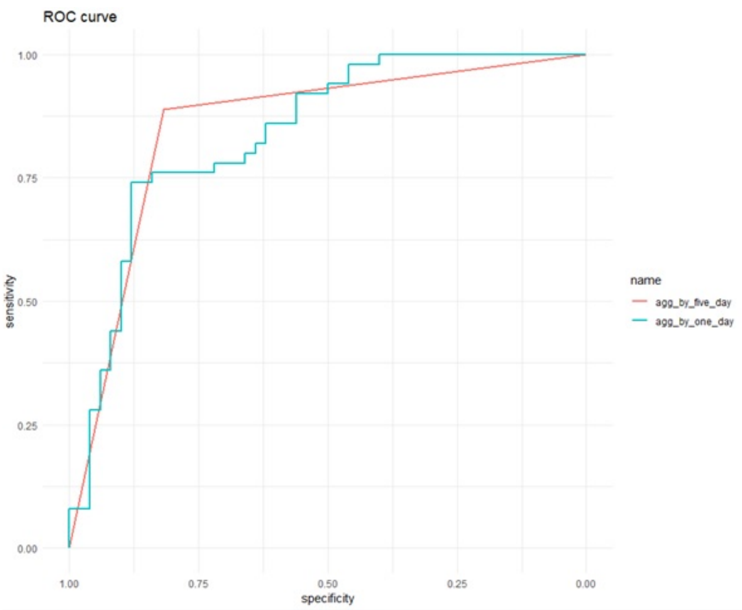

Figure 10: ROC Curves for both models

The design of the convolutional neural network used to detect lameness in heat map images is presented in this section. Experiments determine the appropriate values of the network's learning parameters. Note that of the 38 images obtained, 73% are used for training and 27% are used for network testing.

The network's input images are 50 size and 50 size. There are 2 classes in the outputs. There are 4 hidden layers in the proposed CNN. Table 3 shows the CNN's structure and parameters of learning. "Conv" here represents a layer of convolution, and "FC" is a fully connected layer. Note that size 3x 3 filters are used in all padding conversion operations, whereas all pooling operations are carried out using size 2x 2 max pooling windows.

Layer	Model 1	Model 2	Model 3	Model 4	Model 5
(input)		1 map	1 map	1 map	1 map
(convolutional)	32 maps	32 maps	32 maps	32 maps	32 maps
(max pooling)	32 maps	32 maps	32 maps	32 maps	32 maps
(convolutional)	64 maps	64 maps	64 maps	64 maps	64 maps
(max pooling)	64 maps	64 maps	64 maps	64 maps	64 maps
(convolutional)	128 maps		128 maps	128 maps	128 maps
(max pooling)	128 maps		128 maps	128 maps	128 maps
(convolutional)			256 maps	256 maps	128 maps
(max pooling)			256 maps	256 maps	128 maps
(convolutional)				256 maps	256 maps
(max pooling)				256 maps	256 maps
(convolutional)					256 maps
(max pooling)					256 maps
(fully connected)		100 neurons	100 neurons	100 neurons	100 neurons
(fully connected)	100 neurons	50 neurons			
(fully connected)	2 neurons	2 neurons	2 neurons	2 neurons	2 neurons

Table 5: Different CNN models and their layer structures

Different architecture of CNN has been tried in this research which can be seen above. In each CNN model many layers have been added to test the performance.
The first was the layer of convolution (Conv2D). For the firstconv2D layers, I chose to set 32 filters and the next 64 filters. Each filter uses the kernel filter to transform a portion of the picture (defined by the kernel size). The matrix of the kernel filter is applied to the entire image. Filters can be viewed as an image transformation .

From these converted images orfeature maps, the CNN can detach features that are helpful everywhere. The second major layer in CNN is the pooling layer. It is usedacts as a filter for downsampling. This filteractuallycomparesat the 2 adjacent pixels and selects the maximum value. It is used to reduce computational costs and also reduce overfitting to some extent.We need to select the pooling size ,the size of areapooled each time) the higher the pooling size, the more important the downsampling.

CNN can combine local characteristics and discover further global image features by combining convolutional and pooling layers. Dropout is a regularization method where a proportion of nodes in the layer are randomly disregarded for each training sample bysetting their wieghts to zero. This declines a networkproportionrandomly and forces the network in a distributed manner to learn features. This method also enhances widespreadness and decreases overfitting. Relu (activation functionmax(0,x)) is the rectifier. The activation function of the rectifier is used to add non-linearity to the network. It is used to convert the final feature maps into a single 1D vector. Flattening is a step thatcombines all the local characteristics found in the previous convolution layers. This flattening step is required so after some convolutional andmaxpool layers itmay make use fully connected layers.

A loss function and an optimization algorithm are added once our layers are added to the model. The costfunction measures the poor performance of our model with known labels on images. We use a specific form called binary crossentropy for categorical classifications (2 classes). The optimizer is the most important function. All models wereimplementedwith binary crossentropy for categorical classifications (lame and non lame). The optimizer is the most important function. This mechanism will incrementally improve variables that filter neuronal values, weights and bias to minimize the loss.To evaluate the performance of our model, the metric "accuracy" function is used. This indicator function is comparable to the loss function, except that metric evaluation outcomes are not use when the model is trained.

The Learning rate is the step by which the optimizer walks through the' loss landscape' to make the optimizer converge faster and closer to the global minimum of the loss function. The greater the LR, the greater the steps and the faster the convergence. But, with a high Learning rate, the sampling is very poor and the optimizer could likely fall into a local minima. Hence,a decreasing learning rate during the training can be used to effectively reach the global minimum of the loss function.

The overall performances of the models are tested using 38 images. Table 4 shows the recognition rates obtained for the back propagation networks using 50 50 pixels as the input image size.

Networks	Model-1	Model 2	Model 3	Model 4	Model 5
Training samples	28	28	28	28	28
Hidden Layers	3	2	4	5	6
Learning rate	0.01	0.01	0.01	0.01	0.01
Activation function	RELU	RELU	RELU	RELU	RELU
Epochs	10	10	10	10	10
Accuracy	0.8	0.8	0.8	0.8	0.6

Table 6: Results of the CNN models

The results of all the CNN models can be seen in the table. Different number of hidden layers are used for each model increasing the number of filters and structure. A default learning rate of 0.01 has been used.All the layers have Relu activation function and the final layer is activated by sigmoid function.

As shown in Tables, the networks behave differently during training and testing, and this is obviously due to the difference in the working principles, structures, and training algorithms of the five employed networks. Also in Table, Except model 5 all the other model achieved a test accuracy of 80%. It is understood that the test data has been tested with 10 heatmap images. Generally, CNN requires large image to train to attain a generalization, since training with only 38 images may overfit the data and might be applicable for other barn and cattle. For the 5months dataset, many heatmaps for lame and non lame cows can be generated which would open doors to more generalized algorithm training for CNN.

6 Discussion

Similar approaches can be explored on the 5 months dataset. Firstly, a more accurate way of measuring the feeding and milking activity can be carried out using the barn coordinates along with the actm data. Having threshold of actm values greater than the threshold value would be helpful in measuring the actual feeding activity. Similarly, actm value can also be used to track the milking activity inside the barn. These data can be calculated for each day having mean of actm. Also, cows are not active all the time and during night time resting time will be higher. Hence, segregating the data into morning, afternoon and night can also be done.Also before proceeding with the actual analysis, outliers and extreme values have to be handled and properly investigated. One approach in reducing outliers could be using Inter-Quartile range method.also, more sophisticated algorithms like isolation forest could also be applied in identifying and studying outliers.

Euclidean distance which has been calculated in this 5 days dataset can also be used to calculate the average speed of the cow per day. This feature can be used to differentiate lame and non lameness. However, this feature has to be excluded when cow is resting in the cubicle. Actm data could be used along with distance and x,y coordinate data to check if the cow is actually moving for the last 2 minutes. if it was moving, average speed can be calculated.

One drawback of this methods is that resting duration of cows has not been identified. Coordinate values for each cubicle can be calculated and as many boolean features can be added for each cubicle to track site fidelity and to track the amount of time spent by a cow at a particular cubicle. Another variable which would actually help us to determine if the cow was actually at rest is actm. Keeping a low activity threshold to update the resting feature would be helpful in determining lameness. Once the cubicle no visit feature has been created it would be easier to track the site fidelity to check if the cows have returned to the same cubicle every time.

Similar features which were extracted on daily basis like feeding time duration per day, milking duration per day, distance moved per day can also be analyzed over weeks and months from which a more generalized metrics and key information could be extracted. Also, some of these extracted information on analysis over different time frames could be transferred to other barns and farms, however, environmental factors and other barn structures should be taken into consideration. A more proper time frame for automated detection of lameness have to be selected and the process will be completely based on the trail and error methods.For example,in the above logistic regression method,the result for 5 days data showed that if the data has been aggregated over 5 days the algorithm was quite successfull in learning patterns and detecting lameness rather than aggregating on one day. However, since the algorithm was trained on small sample of data, there is a very high probability that the model has been overfitting. By trying out different time frames and different algorithms on the 5 months data and based on the results, an accurate time frame can be selected to confirm if cow is lame. In splitting of train and test data time series should be taken into consideration as well, random shuffling of data should be avoided.

Major drawback in the CNN model is the lack of images to train different patterns, since the above CNN models were trained only using 5 days data, it generally overfits and its accuracy should be taken with caution. The same model might not generalize well for other cows as it requires more images to train on to read the patterns. Generating heatmaps for each cow on each day over 3 months might be used for training and testing can be done on the next 2months dataset.

Early prediction of lameness is necessary, as the above mentioned approaches only differentiates between lameness and non lameness. The 5 months dataset would be helpful to track the changes in behaviour. The lame labelled cow on a particular date can be observed on the previous days. A study can be made on the cow, few days before, when the lameness have been reported. The behavioural changes in feeding, milking, distance moved, resting and site fidelity could be observed and similar approaches which has been demonstrated above can be done to identify and confirm threshold for lameness.

In this study, a simple demonstration has been made on how a wireless sensor system can be used to monitor movement and behaviors tracking the use of barn at high recording frequencies continuously, it provides additional sources of behavioral information that can not be easily collected using other accelerometer-based systems.

Lameness is one of the most critical issues of health and wellbeing affecting dairy cows all over the world. Early lameness identification can decrease animal distress and suffering[5] and reduce probable farmers' costs[4]. Existing lameness detection methods, based on specialist survivability observations, can be time-consuming and hence automated detection techniques are needed. Increasing the supply of dairy products and growing agricultural practices means that there was a need for automated behavioral monitoring systems which can act as a' advance warning' to detect and predict the health status.

In this 5-day cross-sectional study, by collecting high-resolution spatial location data, a sample demonstration have been shown how groups of 10 lame and 10 non-lame cows display a number of statistically relevant differences in their movement and space-use behaviors , time is spent at specific places in the barn. Within a Precision Livestock Farming approach[6], this form of space - use surveillance system could possibly be extended to allow automated on - farm lameness forecast in individual cows based on space - use and other behavioral differences.

Basic analysis on the 5 days data shows that non-lame cows spend more time in the feeding area , and the comparable result that lame cows spend more time in the cubicles area , is reliable with existing studies on feeding behavior in dairy cows[52,53,54]. As mentioned earlier,the study did not distinguish between cows that are actually feeding in the feeding area and those who are are not feeding. This distinction, however, can be made by combining basic spatial location data with additional activity data from the accelerometer[52]. Also, the study does not have measure of feed intake directly, earlier studies showed that lame cows may eat the same amount but at a faster rate than non-lame cows[55]. It may represent a reduced amount of time spent on the feed face to avoid confrontation and competition from other cows as it is known that lame cows are less likely to initiate an aggressive interaction[56]. Lame cattle may also boost their lying time[36] to reduce distress and pain, which may also explain the results we observed.

Even though the data have high-resolution spatial location data for each actual cow, the data is relatively small sample sizes (10 lame and 10 non-lame cows) and only 5 days of the cross-sectional study were running. Therefore, while our findings have exciting potential, we are careful about over-generalization and the study cannot confirm anything until it has been tested with 5 months data. Precise to this research group and barn environment are the parameter values of the model found during the statistical analysis and would be different for other barn locations and cows .

Basic analysis shows how barn use methods are linked to state of health in individual cows, but barn use activity is also likely to be affected by management actions, barn environment and design, milking rateand milking structure used, and individual cow genderageand breed. Research with large group sizes over an prolonged period would allow us to evaluate the consistency of any observed differences in space-use and what changes in space-use behavior might be detectable at the onset of lameness. More comprehensive social interactions and spatial structure that can affect individual barn-use activity could also be identified by tracking a complete flock over a longer period. In this research, the cattle being tracked created a subgroup of a much wider flock, and due to the complexity in differentiating between social interactions, we did not attempt to explore social relationships when many individuals within the whole flock were not aspect of the data set observed[16].

There were several limitations to our study. One is the training and test size. There were only limited images we could generate for the 5 day period for 20 cows. This implies that it can be difficult, as expected, to identify the absence of lameness with the limited data. Additional architecture and hyper parameter searches, threshold tuning, model assembly, or data increase could potentially improve absolute performance. In our study, we did not highlight these well-known techniques; instead, we focused on demonstrating viability and evaluating trends rather than optimizing end performance. However, our results suggest that only a large number of images (around tens of thousands) are needed to train the models with better classification images for useful performance.

7 Conclusion

In this study,a sample demonstration is shown on how location tracking data obtained from animal-mounted wireless sensors can be processed and analyzed using a Real Time Location System to provide a set of behavioral measures for space-use using 5 days data which could be transferred to 5 months data as well. Also,different models were used in a cross-sectional research design to discover changes in barn-use behavior between the lame and non lame barn-housed cows and found notable changes between lame and non-lame cows. Significant findings like lame cattle spent less time throughout the feeding area and more time in the barn's cubicle areas endorses with previous researches. The data size used in this study for 10 non lame and 10 lame cows was trivial which was aggregated over each day and aggregated over each cows, so care should be taken to extrapolate our findings and conclusions directly to other researches and proposals. Also, various methods ,steps and approaches were discussed shortly to experiment with the 5 months dataset. However,a more generalized results and related methods could possibly be created in upcoming dissertation study to create a new range of models for computerized tracking of barn cows, identifying and predicting health status in cattle by using the 5 months data.

References

[1] "Bruijnis, M. R., Hogeveen, H. and Stassen, E. N. (2010) Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model, Journal of Dairy Science, 93(6), pp. 24192432."

[2] "Booth, C. J., Warnick, L. D., Grohn, Y. T., Maizon, D. O., Guard, C. L. and Janssen, D. (2004) Effect of lameness on culling in dairy cows, Journal of Dairy Science, 87(12), pp. 41154122. doi: 10.3168/jds.S0022-0302(04)73554-7"

[3] "Holzhauer, M., Middelesch, H., Bartels, C. and Frankena, K. (2004) Evaluation of a Dutch claw health scoring system in dairy cattle, in Proceedings of the 13th International Symposium and 5th Conference on Lameness in Ruminants. Available at: email: m.holzhauer@gdvdieren.nl://search.ebscohost.com/login.aspx?direct=true db=lahAN=20043084821site=ehost-live"

[4] Willshire JA, Bell NJ. 2009 An economic review of cattle lameness. Cattle Pract. 17, 136141.

[5] Leach KA, Tisdall DA, Bell NJ, Main DCJ, Green LE. 2012 The effects of early treatment for hindlimb lameness in dairy cows on four commercial UK farms. Vet. J. 193, 626632. () pmid:22884565

[6] Berckmans D. 2014 Precision livestock farming technologies for welfare management in intensive livestock systems.Rev.Sci.Tech.Int.Off.Epizoot. 33, 189196.

[7] Enting H., Kooij D., Dijkhuizen A.A., Huirne R.B.M., Noordhuizen-Stassen E.N. Economic losses due to clinical lameness in dairy cattle. Livest. Prod. Sci. 1997;49:259267. doi: 10.1016/S0301-6226(97)00051-1

[8] Huxley J.N. Impact of lameness and claw lesions in cows on health and production.Livest. Sci.2013;156:6470. doi: 10.1016/j.livsci.2013.06.012.

[9] Leach K.A., Paul E.S., Whay H.R., Barker Z.E., Maggs C.M., Sedgwick A.K., Main D.C.J. Reducing lameness in dairy herdsOvercoming some barriers. Res. Vet. Sci. 2013;94:820825. doi: 10.1016/j.rvsc.2012.10.005.

[10] Straw B.E., Dewe C.E., Wilson M.R. Differential diagnosis of swine diseases. In: Straw B.E., DAllaire S., Mengeling W.L., Taylor D.J., editors.Diseases of Swine.Iowa State University Press; Ames, IA, USA: 1999. pp. 4186.

[11] Maas J. Musculoskeletal abnormalities. In: Smith B.P., editor. Large Animal Internal Medicine. Mosby, Inc.; ST. Louis, MS, USA: 2009. pp. 217231

[12] "Scheibe, K.; Gromann, C. Application testing of a new three-dimensional acceleration measuring system with wireless data transfer (WAS) for behavior analysis. Behav. Res. Methods 2006, 38, 427433"

[13] "Keegan, K.; Yonezawa, Y.; Pai, F.; Wilson, D.; Kramer, J. Evaluation of a sensor-based system of motion analysis for detection and quantification of forelimb and hind limb lameness in horses. Am. J. Vet. Res. 2004, 65, 665670"

[14] "Kokin, E.; Praks, J.; Veerme, I.; Poikalainen, V.; Vallas, M. IceTag3D accelerometric device in cattle lameness detection. Agron. Res. 2014, 12, 223230."

[15] "Chapinal, N.; de Passill, A.; Rushen, J.; Wagner, S. Automated methods for detecting lameness and measuring analgesia in dairy cattle. J. Dairy Sci. 2010, 93, 20072013.

[16] Space-use patterns highlight behavioural differences linked to lameness, parity, and days in milk in barn-housed dairy cows

[17] Scheibe K., Gromann C. Application testing of a new three-dimensional acceleration measuring system with wireless data transfer (WAS) for behavior analysis.Behav. Res. Methods.2006;38:427433. doi: 10.3758/BF03192796

[18] Higginson J., Millman S., Leslie K., Kelton D. Validation of a new pedometry system for use in behavioural research and lameness detection in dairy cattle; Proceedings of the First North America Conference of Precision Dairy Management; Toronto, ON, Canada. 25 March 2010; Elora, ON, Canada: Progressive Dairy Operators; 2010

[19] Kokin E., Praks J., Veerme I., Poikalainen V., Vallas M. IceTag3D accelerometric device in cattle lameness detection.Agron. Res.2014;12:223230.

[20] Pastell M., Tiusanen J., Hakojrvi M., Hnninen L. A wireless accelerometer system with wavelet analysis for assessing lameness in cattle.Biosyst. Eng.2009;104:545551. doi: 10.1016/j.biosystemseng.2009.09.007

[21] Van Nuffel A., Zwertvaegher I., Pluym L., Van Weyenberg S., Thorup V., Pastell M., Sonck B., Saeys W. Lameness detection in dairy cows: Part 1. How to distinguish between non-lame and lame cows based on differences in locomotion or behavior.Animals.2015;5:838860. doi: 10.3390/ani5030387.

[22] Maertens, W., J. Vangeyte, J. Baert, A. Jantuan, K. C. Mertens, et al. 2011. Development of a 492 real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive 493 walkway: The GAITWISE system. Biosystems Engineering 110:29-39

[23] Poursaberi, A., C. Bahr, A. Pluk, A. Van Nuffel D. Berckmans. 2010. Real-time automatic 517 lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow 518 with image processing techniques. Computers and Electronics in Agriculture 74(1):110-119

[24] A. Kiddy. Variation in physical activity as an indication of oestrus in dairy cows. Journal of Dairy Science, 60:235243, 1977

[25] Specht, D. F. 1990. Probabilistic neural networks. Neural Netw. 3:109118.

[26] X.-X. Niu, C. Y. Suen, A novel hybrid cnnsvm classifier for recognizing handwritten digits, Pattern Recognition 45 (4) (2012) 13181325.

[27] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, no. 11, pp. 22782324, 1998

[28] 1. KrizhevskyA,2. SutskeverI,3. HintonGE (2012)ImageNet classification with deep convolutional neural networks.Advances in Neural Information Processing Systems, edsPereiraF,BurgesCJC,BottouL,WeinbergerKQ(Curran Assoc,Red Hook, NY),Vol 25:10971105.

[29] "Nahvi, B., Habibi, J., Mohammadi, K., Shamshirband, S. and Al Razgan, O. S. (2016) Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature, Computers and Electronics in Agriculture. Elsevier B.V., 124, pp. 150160. doi: 10.1016/j.compag.2016.03.025"

[30] "Coopersmith, E. J., Minsker, B. S., Wenzel, C. E. and Gilmore, B. J. (2014) Machine learning assessments of soil drying for agricultural planning, Computers and Electronics in Agriculture. Elsevier B.V., 104, pp. 93104. doi: 10.1016/j.compag.2014.04.004"

[31] "Mohammadi, K., Shamshirband, S., Motamedi, S., Petkovi, D., Hashim, R. and Gocic, M. (2015) Extreme learning machine based prediction of daily dew pointtemperature, Computers and Electronics in Agriculture, 117, pp. 214225. doi: 10.1016/j.compag.2015.08.008"

[32] Roelofs J, LopezGatius F, Hunter RH, van Eerdenburg FJ, Hanzen C, 2010: When is a cow in estrus? Clinical and practical aspects Theriogenology 74, 327344.

[33] Munksgaard L, Reenen C and Boyce R 2006. Automatic monitoring of lying,standing and walking behaviour in dairy cattle. Journal of Dairy ScienceSuppl. 89, 304

[34] Nielsen LR, Pedersen AR, Herskin MS and Munksgaard L 2010. Quantifyingwalking and standing behaviour of dairy cows using a moving average basedon output from an accelerometer. Applied Animal Behaviour Science 127,1219

[35] "K. Ito, M.A.G. von Keyserlingk, S.J. LeBlanc, D.M. Weary Lying behavior as an indicator of lameness in dairy cows J. Dairy Sci., 93 (2010), pp. 3553-3560"

[36] "N. Blackie, J. Amory, E. Bleach, J. Scaife The effect of lameness on lying behaviour of zero grazed Holstein dairy cattle Appl. Anim. Behav. Sci., 134 (2011), pp. 85-91"

[37] "M. Alsaaod, C. Rmer, J. Kleinmanns, K. Hendriksen, S. Rose-Meierhfer, L. Plmer, W. Bscher Electronic detection of lameness in dairy cows through measuring pedometric activity and lying behavior Appl. Anim. Behav. Sci., 142 (2012), pp. 134-141"

[38] "O. Hernandez-Mendo, M.A.G. von Keyserlingk, D.M. Veira, D.M. Weary Effects of pasture on lameness in dairy cows J. Dairy Sci., 90 (2007), pp. 1209-1214"

[39] "J.F. Ettema, S. Ostergaard Economic decision making on prevention and control of clinical lameness in Danish dairy herds Livest. Sci., 102 (2006), pp. 92-106"

[40] "N. Chapinal, A.M. de Passill, D.M. Weary, M.A.G. von Keyserlingk, J. Rushen Using gait score, walking speed, and lying behavior to detect hoof lesions in dairy cows J. Dairy Sci., 92 (2009), pp. 4365-4374"

[41] "A. Gomez, N.B. Cook Time budgets of lactating dairy cattle in commercial freestall herds J. Dairy Sci., 93 (2010), pp. 5772-5781"

[42] "L.A. Espejo, M.I. Endres, J.A. Salfer Prevalence of lameness in high-producing Holstein cows housed in freestall barns in Minnesota J. Dairy Sci., 89 (2006), pp. 3052-3058"

[43] R.M.de Mol,G.Andr,E.J.B.Bleumer,J.T.N.van der Werf,Y.de Haas,C.G.van ReenenApplicability of day-to-day variation in behavior for the automated detection of lameness in dairy cows J. Dairy Sci.,96(2013), pp.3703-3712

[44] Heald, R., R. Tournebize, T. Blank, R. Sandaltzopoulos, P. Becker, A. Hyman, and E. Karsenti. 1996. Self-organization of MTs into bipolar spindles around artificial chromosomes in Xenopus egg extracts. Nature. 382:420425

[45] Morota G., Peagaricano F., Petersen J.L., Ciobanu D.C., Tsuyuzaki K. Nikaido I. (2015) An application of MeSH enrichment analysis in livestock. Animal Genetics 46, 3817.

[46] Viazzi, S., Bahr, C., Van Hertem, T., Schlageter Tello, A., Romanini, C.E.B., Halachmi, I., et al. (2014). Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. Computers and Electronics in Agricul-ture, 100(1), 139-147

[47] Pastell, M., Hautala, M., Poikalainen, V., Praks, J., Veerm ae, I.,Kujala, M., Ahokas, J., 2008. Automatic observation of cow leghealth using load sensors. Comput. Electron. Agric. 62

[48] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316(22):2402-2410. [PubMed] [Google Scholar]

[49] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-118

[50] Bell N, Huxley J. 2009 Locomotion, lameness and mobility in dairy cows. Vet. Rec. 164, 726726. () pmid:19502631

[51] Barker ZE, Vzquez Diosdado JA, Codling EA, Bell NJ, Hodges HR, Croft DP, et al. 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, 63106321. () pmid:29705427

[52] Barker ZE, Vzquez Diosdado JA, Codling EA, Bell NJ, Hodges HR, Croft DP, et al. 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, 63106321. () pmid:29705427

[53] Weigele HC, Gygax L, Steiner A, Wechsler B, Burla J-B. 2018 Moderate lameness leads to marked behavioral changes in dairy cows.J.Dairy Sci. 101, 23702382. () pmid:29290435

[54] .Norring M, Hggman J, Simojoki H, Tamminen P, Winckler C, Pastell M. 2014 Short communication: Lameness impairs feeding behavior of dairy cows.J.Dairy Sci. 97, 43174321. () pmid:24792807

[55] Gonzlez LA, Tolkamp BJ, Coffey MP, Ferret A, Kyriazakis I. 2008 Changes in feeding behavior as possible indicators for the automatic monitoring of health disorders in dairy cows.J.Dairy Sci. 91, 10171028. () pmid:18292258

[56] Galindo F, Broom DM. 2002 Effects of lameness of dairy cows.J.Appl.Anim.Welf.Sci.JAAWS5, 193201. () pmid:12578740