The Veterinary Journal 304 (2024) 106091

Contents lists available at ScienceDirect

The Veterinary Journal

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

Automated dairy cattle lameness detection utilizing the power of artificial intelligence; current status quo and future research opportunities

Nektarios Siachos *, Joseph M. Neary, Robert F. Smith, Georgios Oikonomou

Department of Livestock and One Health, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Leahurst Campus, Chester High Road, CH64 7TE, UK

ARTICLE INFO

Keywords:
Cattle
Locomotion
Machine learning
Mobility

ABSTRACT

Lameness represents a major welfare and health problem for the dairy industry across all farming systems. Visual mobility scoring, although very useful, is labour-intensive and physically demanding, especially in large dairies, often leading to inconsistencies and inadequate uptake of the practice. Technological and computational advancements of artificial intelligence (AI) have led to the development of numerous automated solutions for livestock monitoring. The objective of this study was to review the automated systems using AI algorithms for lameness detection developed to-date. These systems rely on gait analysis using accelerometers, weighing platforms, acoustic analysis, radar sensors and computer vision technology. The lameness features of interest, the AI techniques used to process the data as well as the ground truth of lameness selected in each case are described. Measures of accuracy regarding correct classification of cows as lame or non-lame varied with most systems being able to classify cows with adequate reliability. Most studies used visual mobility scoring as the ground truth for comparison with only a few studies using the presence of specific foot pathologies. Given the capabilities of AI, and the benefits of early treatment of lameness, longitudinal studies to identify gait abnormalities using automated scores related to the early developmental stages of different foot pathologies are required. Farm-specific optimal thresholds for early intervention should then be identified to ameliorate cow health and welfare but also minimise unnecessary inspections.

Introduction

Lameness represents a clinical demonstration of an underlying (most likely) painful foot, or, occasionally, musculoskeletal pathologies, and is a leading cause of reduced cow welfare, reduced milk production, impaired fertility, and increased culling risk, negatively affecting farm profitability (Collick et al., 1989; Melendez et al., 2003; Walker et al., 2008; Huxley, 2013; Puerto et al., 2021). According to an extended literature review (Thomsen et al., 2023), the average reported herd lameness prevalence has changed only slightly over the last 30 years and with minimal geographical differences. Average reported lameness prevalence, using various scoring methods, in European and North American dairy herds was 22.0% and 24.6%, respectively, ranging from 5.1% to 45% between studies (Thomsen et al., 2023). Furthermore, untrained farmers commonly underestimate the lameness problem in their herds across all farming systems. Australian producers in grazing dairy herds only identified as lame 25% of cows that were identified as lame by researchers (Beggs et al., 2019). On average, Canadian dairy farmers estimated lameness prevalence in their herds to be 9%; less than half of that observed by researchers (22%) (Cutler et al., 2017).

Impact of regular lameness monitoring

Frequent monitoring and early intervention have been described as key components of any effective lameness management approach. When researchers performed weekly or biweekly locomotion scoring without giving access to these records to farmers, the time elapsed from a cow being identified as lame by the researchers to treatment ranged from a median of three weeks (Alawneh et al., 2012) to as long as nine weeks, with some cows waiting up to 16 weeks for treatment (Leach et al., 2012). Leach et al. (2012) investigated a more stringent intervention approach to the farmer's recognition of lame cows, by performing fortnightly mobility scoring and treating any cow identified as lame within 48 h. This led to an approximately 30% ($P < 0.01$) difference in lameness prevalence within four weeks of enrolment and a reduction of recurrence rates, compared to the control farmers' approach. Although

* Corresponding author.
E-mail address: Nektarios.siachos@liverpool.ac.uk (N. Siachos).

https://doi.org/10.1016/j.tvjl.2024.106091
Received 20 December 2023; Received in revised form 25 February 2024; Accepted 27 February 2024
Available online 29 February 2024

this difference waned during the following weeks of the study, fewer cows managed under the strict protocol were identified as severely lame. Additionally, Groenevelt et al. (2014) have shown that fortnightly mobility scoring by a trained scorer and foot trimming of any cows identified as lame within 3–48 h had higher cure rates, decreased the odds of lameness by half and resulted in fewer cases of severe foot lesions after a period of 18 weeks, compared to cows relying on farmers' "normal procedures" for detection and treatment of lame cows.

Visual mobility scoring systems

Early lameness detection should rely on regular mobility/locomotion scoring performed by a trained scorer or by farm personnel (Groenevelt et al., 2014; Horseman et al., 2014). Visual scoring systems have been developed as inexpensive and non-intrusive methods to visually assess gait abnormalities in cows standing and/or walking, focusing on different gait and posture indicators to define lameness. Most systems distinguish between non-lame cows and cows with mild or severe lameness. Available visual systems have been summarized (Schlageter-Tello et al., 2014b; Van Nuffel et al., 2015a). Five point (1−5) lameness scoring methods (Sprecher et al., 1997; Flower and Weary, 2006; Thomsen et al., 2008), or modifications based on these methods, and the four scale (0−3) mobility scoring system (Whay et al., 2003) of the U.K. Agricultural and Horticultural Development Board (AHDB) are the most referred and frequently used systems in practice. The method of Sprecher et al. (1997) relies on assessing animals when standing and when moving while the AHDB method and several other methods only rely on assessing moving animals.

Frequency of scoring in dairy farms usually varies from weekly to quarterly, and is often driven by milk contracts as a requirement to improve animal welfare and meet acceptable industry standards. However, many farmers are still detecting lameness only during foot trimming or by *ad hoc* observation while performing other tasks like moving cows, or as cows walk through the milking parlour (Leach et al., 2010; Horseman et al., 2014; Dolecheck and Bewley, 2018; Sadiq et al., 2019). Visual mobility scoring, although a very helpful tool, is time-consuming and labour-intensive. Consistency is difficult to maintain especially in large herds (Van Nuffel et al., 2015a). Moreover, it is prone to subjectivity leading to significant variation between and within observers. The scorer's background and experience, the ground surface of the walking passageway, the cow-flow speed, and whether assessment is performed live or from video footage of walking cows are factors contributing to this variability (Schlageter-Tello et al., 2014b; Nejati et al., 2023). In addition, since impaired mobility is not expressed in the same way for all cows, the decision of each human scorer on which traits are more important to assign a specific score adds to the subjectivity of visual scoring systems (Schlageter-Tello et al., 2014b).

The simple percentage agreement (PA) and the unweighted or weighted Cohen's kappa coefficient are the most used measures of categorical agreement. The inter-rater agreement between different human scorers assessing mobility of cows on-farm varied remarkably within and across studies and reported measures were: weighted kappa = 0.24–0.68 for a five-level scale (Thomsen et al., 2008), while kappa = 0.00–0.57 (Linardopoulou et al., 2022), PA = 88%, kappa = 0.41, and Gwet's coefficient = 0.85 (Anagnostopoulos et al., 2023), have been reported for a two-level scale agreement.

Accordingly, measures of inter-rater agreement reported in studies assessing mobility from videos with walking cows were: PA = 23–82% and kappa = 0.28–0.84 for a five-level scale (Schlageter-Tello et al., 2014a), and kappa = 0.42–0.49 (Dahl-Pedersen et al., 2018), and PA = 79% and kappa = 0.57 (Gardenier et al., 2021), for a two-level scoring scale. The intra-rater agreement also varied within and across studies and reported measures of accuracy were: weighted kappa = 0.38–0.78 (Thomsen et al., 2008), PA = 60–83% and weighted kappa = 0.63–0.86 (Schlageter-Tello et al., 2014a), PA = 63–68%, kappa = 0.53–0.59 and weighted kappa = 0.72–0.78 (Garcia et al., 2015), for a five-level scale and PA = 82% and kappa = 0.63 (Gardenier et al., 2021), for a two-level scoring scale.

According to Landis and Koch (1977) kappa values should be interpreted as showing slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80) or almost perfect (0.81–1.00) agreement. A kappa value of ≥ 0.60 (Gibbons et al., 2012; Schlageter-Tello et al., 2014a) and a PA of ≥ 80% (McHugh, 2012) have been suggested as benchmarks of acceptable reliability for various welfare and health indices.

Automated lameness detection systems

A plethora of automated lameness detection systems have been developed in the last decades, to overcome the drawbacks of visual scoring and to provide an objective and accurate early lameness detection monitoring tool. These systems have been reviewed thoroughly before for their development, the features they relied on for lameness detection and their performance (Schlageter-Tello et al., 2014b; Van Nuffel et al., 2015b; Alsaaod et al., 2019; O'Leary et al., 2020; Nejati et al., 2023). Automated systems can be classified into three categories, based on the characteristics they rely on for lameness detection:

(1) kinematic methods, including vision-based techniques, accelerometers and pressure-sensitive mats;

(2) kinetic methods, including ground reaction force plates, and weight distribution platforms; and

(3) indirect methods, defined as those that do not rely on gait characteristics, including infrared thermography, and systems gathering data from milk production, lying time, eating behaviour and activity sensors.

Cost of lameness, willingness to pay for lameness control and cost-benefit of automated systems

The cost of lameness in dairy herds varies across studies and is highly dependent on the lesion type and severity. The estimated average cost per case of lameness regardless the cause ranged from USD $185 to $333 for primiparous and multiparous cows, respectively (Liang et al., 2017). A study using data from over 108,000 trimming sessions in 804 dairy farms in Spain (Charfeddine and Pérez-Cabal, 2017) was used to quantify the economic effects of specific foot lesions. The average annual cost per cow per case of digital dermatitis, sole ulcer and white line disease was USD $10.8, $50.9 and $43.3, respectively. Of interest is that mild cases could cost up to $232, whereas severe cases could cost up to $622 per affected cow annually, due to milk losses, more days open and premature culling expenses. Recently, Robcis et al. (2023) estimated that the average cost per lameness case and per active case of digital dermatitis is approximately €308 and €392 per year, respectively, based on bioeconomic simulation analysis examining different scenarios. Besides milk losses that were the main component of lameness costs, impaired fertility, treatment, culling and the dynamics of active and inactive states of digital dermatitis cases were also accounted for in this stochastic modelling approach.

The cost of an automated lameness detection system, either due to equipment purchase costs or to annual subscription fees, poses a concern affecting a farmer's perception of these systems. A large questionnaire survey conducted among UK dairy farms revealed that farmers were willing to pay a median cost of £50 per cow per year to achieve a 0% lameness prevalence in their herds and were also willing to pay a median of £250 per lame cow per year (Bennett et al., 2014). Farmers were also willing to pay a median amount of £16 per cow per year to avoid the inconvenience related to the employment of conventional lameness control measures.

Kaniyamattam et al. (2020) conducted a cost-benefit analysis of investing in automatic lameness detection systems under different scenarios of lameness incidence, herd size, efficiency, and cost of the system combinations. They concluded that 295 of the 351 tested scenarios produced a positive net return for an automated system over visual lameness detection performed by a human scorer, which ranged from USD $13 to $99 per cow per year, with the cost of investment ranging from $10,000 to $50,000 with a lifespan of 10 years. Comparably, Edwardes et al. (2022) assessed the economic value of using an automated wearable sensor-based system to manage lameness. A bioeconomic simulation model was used to examine 80 different scenarios of "sub-optimal" mobility prevalence, foot trimming protocols and sensor performance for a typical small-size Dutch dairy farm with cows having access to pasture. They accounted for the cost of the sensors, trimming costs, milk production losses and culling costs. A positive net return was obtained in 39 of the tested scenarios, with the best showing a maximum net gain of USD $51 per cow per year.

Application of artificial intelligence in livestock monitoring

Artificial intelligence (AI) in general describes the technological advances in hardware and software that enable programming a machine and especially a computer system to perform various tasks simulating the intelligence of a human being (Fatima and Pasha, 2017). Machine learning (ML) and deep learning (DL) are subfields of AI and include models built to resolve decision-making problems (Hossain et al., 2022). The intensity of dairy farming and the already extended utilization of sophisticated technological tools offers many opportunities for the emerging capabilities of AI in performing tasks previously difficult or even impossible for classical computer systems to carry out.

Aside from lameness detection, ML through video monitoring has been used for other livestock monitoring purposes and will likely become a core component in the delivery of future precision livestock farming technologies. A brief description of the opportunities of AI applications in modern dairy farms is provided by De Vries et al., (2023). Monitoring the health and welfare of livestock is time-consuming and often mundane task with results dependent on the human's experience and motivation (Garcia et al., 2020). The automation offered by machine learning approaches can surmount these problems but many of these new technologies, despite commercial availability, remain to be scientifically validated. A recent systematic review of sensor technologies for welfare assessment of dairy cattle reported that just 14% of commercially available sensors had been scientifically validated (Stygar et al., 2021).

Machine learning has recently been used to successfully categorise the behaviour of calves based on accelerometer and gyroscope data obtained from a collar-based sensor (Carslake et al., 2021). It may also detect subacute ruminal acidosis from spatio-temporal data, although with low specificity (Wagner et al., 2020), as well as determine dairy cattle posture and resting behaviour from flank-based accelerometer data (Balasso et al., 2021), and lying behaviour in cows on pasture or in a barn based on data from a collar-mounted sensor containing an accelerometer, a magnetometer and a gyroscope (Schmeling et al., 2021). The aforementioned studies were all based on wearable devices with some using video recordings as the source of ground truth data. Some studies, however, have used DL on video surveillance, primarily for the purpose of cattle tracking and behavioural evaluation (Zambelis et al., 2021; Han et al., 2023). We are likely to see considerable growth in the use of artificial intelligence across the precision livestock farming sector.

Peer-reviewed articles of automated lameness detection systems using artificial intelligence

We performed a literature review following the PRISMA guidelines (Page et al., 2021). Peer-reviewed articles and conference papers in Scopus and PubMed were searched using the following keywords in the title, abstract and/or keywords: cow*; cattle; locomot*; mobility; lame*; automat*; machine learning; machine vision; deep learning; neural network; artificial intelligence, in different combinations. We included only relevant original research studies written in English language. The list of references in those studies was also used. We did not filter according to year of publication. The last search was conducted on June 30, 2023. Main characteristics of the systems described in each publication and information regarding ground truth used for validation are summarized in Table 1.

Accelerometers

Haladjian et al. (2018) developed an electronic device attached in the hind leg sensing motion on a 3D orientation. Ten multiparous non-lame cows that were allowed to walk freely inside their barn for some time wearing the device, were used to extract data that served as the features of "normal" strides. Then, "abnormal" stride features were obtained by applying a plastic block in the lateral claw of any hind foot at the same cows, and data during walking was collected again. A support vector machine algorithm classified cows into lame or non-lame with an accuracy (ACC) of 91%, 74% sensitivity (SE) and 92% specificity (SP).

Barker et al. (2018) tested wearable 3D sensors monitoring position, activity and behaviour that would potentially detect early signs of lameness. After mobility scoring 47 cows using the 0–3 grade AHDB method, nine cows with a score 0 (non-lame) and 10 cows with a score 2 (lame), matched for parity, days-in-milk and milk yield, were selected. The following day, motion sensors mounted on neck collars were fitted to the cows and collected acceleration and position data for five days. A decision tree algorithm was chosen to classify feeding and non-feeding behaviour of cows, but not for lameness *per se*; however, several behavioural measures extracted differed between lame and non-lame cows, indicating a potential use of these measures to train algorithms for lameness detection.

In another study (Taneja et al., 2020), 146 cows were used wearing a commercial long-range accelerometer (ISM band, ENGS Systems, Israel) on one of the front legs. These accelerometers gathered data regarding the lying time and the count of steps and swaps from lying down to standing up. Visual locomotion scoring performed by an agricultural science student and by the farmer using a four-scale method, was used as the ground truth of lameness. A k-nearest neighbor algorithm produced the best classification accuracy, with 87% ACC, 90% SE and 73% SP.

Gertz et al. (2020) applied ML methodology to data collected by wearable activity sensors to accurately detect locomotor system-associated diseases in general. Data from a commercial system using both leg and neck sensors (CowScout, GEA Farm Technologies, Germany) measuring activity, behaviour, leg activity, lying time and walking time, was collected from 397 cows. Any locomotor system-related disease and disorder was recorded on the farm's management software by a trained trimmer and veterinarians. Despite the wide range of conditions pooled within this definition, these records were used to classify cows as "healthy" or with "impaired locomotion" and served as the ground truth. A scalable decision tree boosting system (XGBoost) was selected as the classifier algorithm producing an Area under the Receiver Operating Characteristic Curve (AUC) of 86%, 81% SE, 78% SP and 81% F-measure (the harmonic mean of precision and recall).

Borghart et al. (2021) evaluated predictive models gathering data from sensors and animal records to automatically detect lameness. Cows (n= 164) equipped with commercial neck-mounted 3D accelerometers (MooMonitor+, Dairymaster, Ireland) recording behavioural, rumination, resting and feeding time data, were used with data gathered over an 11-month period. Ground truth of lameness was established with weekly locomotion scores performed by a trained veterinarian using the 1–5-point scoring method (Sprecher et al., 1997). Cows with locomotion scores (LS) of 1 and ≥ 2 were considered as "sound" and "unsound", respectively. Among four evaluated models using an XGBoost algorithm, the model including metrics collected by the sensors, along with data regarding parity, bodyweight and milk production, achieved the best

Table 1
List with peer-reviewed articles and conference papers describing automated cattle lameness detection systems using artificial intelligence. Studies are classified according to the type of sensor applied to the system. Sensors, methodology, ground truth used for lameness definition and metrics of performance are presented.

Reference	Sample size	Sensor	Features for lameness detection	Statistical Model / Algorithm	Ground truth used for validation	Visual scoring system used	Validation with foot lesion records	Identification of lame foot	Performance / measures of correct classification
Accelerometers									
Haladjian et al., 2018	10	3D motion sensor (leg-worn)	Deviation from normal stride pattern	Support Vector Machine Classification algorithm	Strides of blocked non-lame cows as ground truth	NA	No	No	ACC: 91.1% SE: 74.2% SP: 91.6%
Barker et al., 2018	20	Neck-mounted 3D accelerometer	Position and acceleration data	Decision tree classification algorithm	Human scorers	Whay et al. (2003) (0–3), only cows with score 0 and 2 were used	No	No	No measures of accuracy for lameness detection. Only for feeding and non-feeding behaviour, which were found to be affected by lameness
Taneja et al., 2020	146	3D long-range Accelerometer (leg-worn)	Behavioural and activity profile	K-Nearest Neighbor	Human scorers (agricultural science student or farmer)	Four point grading method	No	No	ACC: 87% SE: 89.7% SP: 72.5%
Gertz et al., 2020	397	Neck- and leg-mounted activity sensors (CowScout, GEA Farm Technologies, Germany)	Behavioural and activity data	Gradient Boosted Decision Tree Learning algorithm (XGBoost)	Recorded data for locomotor-associated diseases by trained trimmer and veterinarian	NA Binary score “sick” or “healthy”	Yes	No	AUC: 0.86 F-measure: 0.81 SE: 81% SP: 78%
Borghart et al., 2021	164	3D Accelerometers (neck-mounted, Moomonitor+)	Behavioural metrics, BW and milk production data	Gradient Boosted Decision Tree Learning algorithm (XGBoost)	Human scorer (trained veterinarian)	Sprecher et al. (1997) (1–5)	No	No	Kappa: 0.58 AUC: 0.85 ACC: 78% SE: 78% SP: 78%
Jarchi et al., 2021	23	Accelerometers (leg-worn)	Instantaneous gait frequencies for each leg	Long Short-Term Memory (recurrent Neural Network) and Synchrosqueezed Wavelet Transform	Human scorers	Whay et al. (2003) (0–3)	No	Potentially	SE: 90.1–99% SP: 98.4–100% F-measure: 0.95–0.99
Force plates									
Pastell and Kujala, 2007	73	Four balance platform in robotic milking systems	Changes from dynamic leg load, steps and kicking behaviour during milking	Probabilistic Neural Network	Human scorers	Sprecher et al. (1997) 1–5 scores. 1 and 2: non-lame; ≥3: lame	No Hoof pathologies recorded but not used for training or validation	No	ROC curves for binary lameness classification: AUC: 0.86 CC: 96.2% DR: 100% SE: 100% SP: 57.5%
Ghotoorlar et al., 2012	105	Four force plate balance system	Ground reaction force data	Artificial Neural Network	Human scorer	Sprecher et al. (1997) (1–5)	No	No	SE: 50–100% across scores SP: 91–100% across scores PPV: 81–100% across scores NPV: 92–100% across scores
Radar sensing									
Shrestha et al., 2018	5	Radar sensor	Radar Micro-Doppler signature	K-Nearest Neighbor	Human scorer (qualified veterinarian)	Whay et al. (2003) (0–3)	No	No	ACC: 80% SE: 70% SP: 91.4%

(continued on next page)

Table 1 (*continued*)

Reference	Sample size	Sensor	Features for lameness detection	Statistical Model / Algorithm	Ground truth used for validation	Visual scoring system used	Validation with foot lesion records	Identification of lame foot	Performance measures of correct classification
Busin et al., 2019	54	Radar sensor	Radar Micro-Doppler signature	Naïve Bayesian and K-Nearest Neighbor	Human scorer (qualified veterinarian)	Whay et al. (2003) (0−3)	No	No	ACC: 83% SE: 81% SP: 83%
Acoustic analysis									
Volkmann et al., 2021	64	Walk-over panel equipped with a recording piezoelectric sensor	Acoustic analysis of cow impact sound records walking on a panel	Random Forest Algorithm	Detection of foot lesions during professional trimming	Binary score as bearing any foot lesion	Yes	No	Kappa: 0.80 SE: 81% SP: 97% Precision: 0.90 NPV: 0.93
Computer vision									
Zhao et al., 2018	98	Commercial web Camera – side view	Leg swing analysis (motion curve of moving leg, swinging speed and step time)	Decision tree classification algorithm	Human scorer -assessment of videos	NA 1–3 scores based on presence of various lameness indicators	No	No	ACC: 90.2% SE: 90.3% SP: 94.7%
Kang et al., 2020	100	Commercial 2D digital camera – side view	Deviations in movement speed during supporting phase of a cow's hoof	Receptive Field block Net single Shot Detector	Human scorers (average score of on-farm and on-video evaluation)	Three grade scoring method 1: no lameness; 2: mild lameness; 3: severe lameness)	No	Yes	Spearman's ρ: 0.864 Kappa: 0.93 ACC: 96% Precision: 87%
Jiang et al., 2020	90	Commercial digital 2D camera – side view	Behavioural pattern recognition	Convolutional Neural Network	NA	Four classes: normal, slight, moderate or severe lameness	No	No	NA
Wu et al., 2020	50	Commercial digital 2D camera – side view	Step size of front and rear legs	YOLOv3 (Deep Learning algorithm) and Long Short-Term Memory classification model	NA	NA	No	No	ACC: 98.6% SE: 97% SP: 97% Precision: 99.2% Recall: 97.5%
Kang et al., 2022	456	Commercial 2D digital camera – side view	Position of cow's hooves	YOLOv4 (object detection algorithm) and DenseNet classification algorithm	Two human scorers - assessment of videos	1–3 scale Engel et al. (2003) O'Callaghan (2003)	No	No	ACC: 98.5% SE: 98.5% SP: 99.3%
Russello et al., 2022	30	Camera – side view	Pose estimation based on annotated landmarks	Deep Learning based Pose Estimation model (convolutional neural network)	NA (Ground-truth confidence maps of anatomical landmarks)	NA	No	No	NA
Jiang et al., 2022	90	Camera, video – side view	Fitting curvature of cow's back	Filter Layer YOLOv3 (Deep Learning algorithm) and Long Short-Term Memory classification model	Three human scorers - assessment of videos	Three grade scoring method (not lame; early-lame; moderate or severely lame)	No	No	ACC: 96.6% Precision: 98.3% Recall: 97.5%
Anagnostopoulos et al., 2023	6040	2D Camera – overhead view	Changes in coordinates of reference anatomic points	Convolutional Neural Network	Human scorer	Whay et al. (2003) (0−3)	Yes	No	Binary classification: PA: 87% Kappa: 0.40 Gwet's: 0.83 Predicton of potentially painful lesions: ACC: 73.8%

(*continued on next page*)

Table 1 (*continued*)

Reference	Sample size	Sensor	Features for lameness detection	Statistical Model / Algorithm	Ground truth used for validation	Visual scoring system used	Validation with foot lesion records	Identification of lame foot	Performance / measures of correct classification
Barney et al., 2023	250	Camera (GoPro) – side view	Pose estimation based on specific anatomical points	Mask regions with Convolutional Neural Network features and CatBoost classification algorithm	Three human scorers	Whay et al. (2003) (0–3)	No	No	SE: 52% SP: 81% Kappa: 0.88 ACC: 92% Precision: 0.87 Recall: 0.92
Zheng et al., 2023	60	Commercial digital camera – side view	Step length of front and rear legs to detect limping	Convolutional Siamese Neural Network with attention mechanism	NA (step length of "average lame" cow to observe changes)	NA	No	No	ACC: 94.7% SE: 95.1% SP: 93.7% Precision: 96.2%
Li et al., 2023	222	Commercial digital camera – side view	Motion changes	Temporal Aggregation network using micromotion features	Two human scorers - assessment of videos	Scoring for specific lameness indicators	No	No	ACC: 98.9% SE: 98.9% SP: 99.4% Precision: 98.9 % F-measure: 98.9
Zhao et al., 2023	52	Commercial digital camera – side view	Head-hoof and back-hoof linkage features	DeepLabCut (v2.2b8) algorithm with ResNet50 as backbone network.	Three human scorers - assessment of videos	Flower and Weary (2006) 1–5 point scoring system	No	No	ACC: 87.3% Recall: 88.8% Precision: 86.4% NPV: 88.2% SP: 85.7%
Siachos et al., 2023	27,082	2D Camera – overhead view	Changes in coordinates of reference anatomic points	Convolutional Neural Network	Four human scorers	Whay et al. (2003) (0–3)	Yes	No	Binary classification: PA: 81.5 – 86.3% Kappa: 0.23 – 0.40 Gwet's: 0.0.76 – 0.83 Prediction of potentially painful lesions: ACC: 83% SE: 40% SP: 88%

NA, not available; ACC, accuracy; SE, sensitivity; SP, specificity; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the (receiver operating characteristic) curve; CC, correct classification; DR, detection rate; PA, percentage agreement.

classification performance with kappa = 0.58, AUC = 0.85, 78% SE and 78% SP.

Jarchi et al. (2021) used 23 cows wearing four 3D motion sensors, one on each leg, producing low frequency acceleration signals. Ground truth of lameness was based on visual mobility scoring using the 0–3-point AHDB method, with cows receiving a mobility score (MS) ≥ 2 being considered as lame. A deep neural network was developed that classified cows into lame and non-lame using instantaneous gait frequencies for each leg, with 90–99% SE and 98–99% SP using a 60-second window.

Weighing platforms and ground reaction force plates

A system consisting of four balance platforms placed on the floor of a robotic milking unit was developed by Pastell and Kujalaf (2007). Leg weights during 10,000 milkings from 73 cows over a 5-month period were obtained. Data were split into two parts for training and validation. Data included features for the weight of each leg, steps and kicking behaviour during milking. Weekly visual lameness scoring by experienced scorers using the 1–5-point method (Sprecher et al., 1997) was used as the ground truth of lameness. Cows with a lameness score (LS) ≤ 2 were considered sound and those with LS ≥ 3 were considered lame. Foot lesions were also systematically recorded during this study but were not used for training or validation. A probabilistic neural network was selected to classify cows into sound or lame. The system's performance for the binary classification produced an AUC = 0.86 and was highly sensitive (100%) but less specific (58%).

Ghotoorlar et al. (2012) developed a lameness detection system based on four separate ground reaction force plates for each foot installed on a foot-trimming box. From 105 cows were used, 60% were randomly selected for development of an artificial neural network, and the remaining 40% for validation. As the ground truth of lameness, visual lameness scoring was performed by an experienced scorer while cows were walking before and after standing on the plates, using the 1–5-point method (Sprecher et al., 1997). The repeatability of the system's continuous scores from 25 cows obtained twice on the same day was high (Pearson's $r = 0.95$). The system was highly sensitive for cows with LS = 1 and LS = 4 (94 and 100%, respectively), but less sensitive for LS = 5 (50%), while it was highly specific across all LS (from 91% to 100%). This system required cows standing still for about six minutes to obtain accurate data, making it more suitable to be installed on the floor of the milking parlour or the robotic unit, rather than on a passageway.

Acoustic analysis

The hypothesis that alterations in normal gait due to lameness would affect the impact sound produced by a cow's steps was tested by Volkmann et al. (2021). They constructed a panel installed on the slatted floor of a passageway exiting the milking parlour, that cows would walk over. The panel was equipped with a piezoelectric sensor that recorded the impact sound for each cow identified through an RFID sensor. Based on 640 audio files from 64 cows, a random forest algorithm was chosen for binary classification of cows as non-lame or lame. The presence of foot lesions recorded during a professional trimming session of 43 cows were used as the ground truth of lameness and preferred over visual lameness scoring. Binary transformation of a three-grade scoring method (Volkmann et al., 2019), produced a kappa = 0.64 for the agreement with the lesion findings. The model's predicted classification yielded a substantial agreement with lesion findings (kappa = 0.80), with 81% SE and 97% SP. It must be noted, though, that grade or severity of foot lesions were not considered. It has been demonstrated that several lesions such as inactive cases of digital dermatitis and mild cases of sole hemorrhage and white line separation may go undetected by human scorers assessing a cow's gait (Chapinal et al., 2009).

Radar sensing

Shrestha et al. (2018) tested the use of radar signal transmission that generated micro-doppler signatures in five dairy cows, among other species as well, while approaching and walking away from the radar to identify lameness. Cows were scored for lameness by two human scorers using the 0–3-grade AHDB method. A supportive vector machine and k-nearest neighbor were the models considered as classifiers, showing similar measures of accuracy. The posterior view was the best, meaning when cows were walking away from the radar, producing an ACC of 80%, with 70% SE and 91% SP.

The previous original small-scale proof-of-concept study was replicated by Busin et al. (2019) using 44 cows and 80 sheep. The micro-doppler signatures were obtained as cows were walking away from a radar sensor on an exit-race. A human scorer assessed the mobility of each cow using again the 0–3-grade AHDB method, at the same time. A naïve Bayesian algorithm at a three seconds segment duration produced better accuracies, over the k-nearest neighbor classifier, with ACC of 83%, 81% SE and 83% SP.

Computer vision

Most publications describing the development and the evaluation of automated systems for lameness detection using AI, rely on the use of computer vision technology. These systems can be grouped into those focusing on changes in the position or the movement of the feet or limbs of the cows, and those assessing changes in position and movement of multiple anatomical body regions.

Systems focusing exclusively on position or movement of feet/limbs

Zhao et al. (2018) used a commercial web camera to capture a side-view video from 98 cows walking through an alley to automatically detect lame cows using leg swing analysis. Six features extracted from the motion curves of the legs during a step were used to evaluate locomotion from a total of 621 acquired videos. Ground truth of lameness was established according to single trained farm personnel who assessed videos of the cows and classified them using a 3-grade method; non-lame cows were scored as 1, cows with one indicator of lameness received a score 2 and cows with two or more indicators received a score 3. The repeatability of the human scores was assessed and produced a kappa = 0.91. A decision tree learning algorithm was chosen for classification with an average ACC of 90%, 90% SE and 95% SP.

Kang et al. (2020) developed a system focusing on the movement of the hooves and thus being unaffected by individual cow gait characteristics. A commercial 2D digital camera was placed near the entrance of the milking parlour. Authors suggested that the added discomfort due to the weight of milk before milking would make the signs of lameness more obvious. Side-view footage of the whole body of 100 walking multiparous Holstein cows of a single farm was captured. Degree of lameness was established by measuring deviations in the movement speed of the hooves during the supporting phase. A receptive field block network (Receptive Field Block Net Single Shot Detector), which is a high-performing object detector algorithm based on deep convolutional neural networks, was used to track and detect the position of the hooves. Two trained scorers assessed the locomotion of these cows on-site, after milking though, and on the acquired videos, using a 3-point scoring method (score 1: no lameness; score 2: mild lameness; score 3: severe lameness). The average scores for each cow were used as the ground truth. Agreement between human and predicted scores was almost perfect (kappa = 0.93). The classification ACC was 96% and the positive predictive value (PPV) was 87% across all scores. The identification of the lame foot with an ACC of 93% was an additional important novelty of this study.

Kang et al. (2022), similarly to the work of Kang et al. (2020), developed a system to detect the position of the hooves. A commercial 2D digital camera was placed near the exit race of the milking parlour. Side-view footage from 456 multiparous Holstein cows from a single farm were obtained. Visual locomotion scoring was performed using the same 3-point scoring method as in Kang et al. (2020). Selected cows were equally distributed among three locomotion scores. Data was divided into training (60%), validation (20%) and test (20%) datasets. A YOLOv4 algorithm, a machine learning model able of real-time object detection, was used to track the hooves of the cows. A convolutional neural network architecture (DenseNet) was then used to classify the spatiotemporal gaitmaps of each hoof into lame or non-lame. The system produced a remarkably high classification ACC of 99%, with 99% SE and 99% SP.

Zheng et al. (2023) collected side-view footage from 60 Holstein cows. Videos were captured using a commercial digital 2D camera. Twenty-five video clips were recorded from lame cows and 35 from non-lame cows. However, the authors did not describe the scoring method they used to define lameness. Following leg location, the cow's relative step size was the selected feature to train a convolutional Siamese neural network architecture with an attention mechanism for lameness detection. A support vector machine algorithm produced the best classification accuracy with 95% ACC, 95% SE, 94% SP and 96% precision (the PPV).

Systems focusing on position or movement of multiple body parts

Jiang et al. (2020) developed a convolutional neural network architecture for lameness recognition using spatiotemporal features of cows' behaviour while they were walking, which were enriched with optical flow algorithms. From 90 cows on a single farm, 1080 side-view videos were captured using a digital 2D camera. A part of the available videos was used as the test dataset. Although lameness definition used for validation was not clearly described, video clips were classified into four classes that showed cows as "normal", slightly lame, moderately lame or severely lame. The system's performance was evaluated by calculating the correct classification ACC per clip (95%) and the "mean average precision" (98%), a metric used to evaluate object-detection models.

Wu et al. (2020) developed an algorithm for lameness detection by processing the relative step size of cows. From an initial pool of 750 side-view videos of 50 Holstein cows on a single farm, 30 videos of non-lame cows and 20 of lame cows were selected and used for this study. However, they did not provide information on the method used for definition of lameness. After manually labelling the head, back and leg regions in the captured frame images, an object detection algorithm (YOLOv3) was developed that tracked the position of the legs in each frame. This allowed for the calculation of the relative step size of front and hind legs. A long short-term memory model outperformed the other evaluated classifiers, producing an ACC of 99%, 97% SE and 97% SP.

Jiang et al. (2022) used 810 side-view videos from 90 Holstein-Friesian cows on a single farm walking in a passageway. They focused on extracting the curvature of the cow's back, using a real-time object detection algorithm (FLYOLOv3), considering this feature changes considerably in lame cows. Three human observers assessed the videos and classified cows into three categories as being "normal", "early" lame or moderately and severely lame, and served as the ground truth. Selected videos were equally distributed across the three classes. The predictive model they developed, a bi-directional long short-term memory neural network with noise layer (NOISE+BiLSTM), yielded an average classification ACC of 97%.

Russello et al. (2022) proposed a new deep learning framework for estimation of body parts position using 3D convolutions on automatic lameness detection using pose estimation of cows in videos. Side-view videos from 30 cows (20 black and white, 10 red and white) on a single farm from an original population of 70 filmed Holstein-Friesian cows, were captured using a stereo camera with sensors at the visible wavelength. Cows were filmed outdoors while walking in a passageway from the barn to pasture. Data was divided into training and test datasets, but the authors did not include validation of the system's performance against a ground truth method of lameness detection. The percentage of head-normalized key-points metric (PCKh) obtained from known and unknown cows showed promising results for generalization on different farms.

The work of Li et al. (2023) was based on that of Kang et al. (2020) and Kang et al. (2022), in terms of the camera setup and the ground truth scoring system applied. Holstein cows (n=222) on a single farm were randomly selected and four side-view videos per cow were captured using a commercial digital 2D camera. The ground truth reference for lameness was set by two human experts assessing videos using a three-grade scoring method and classified cows as non-lame (score 1), or those one indicator of lameness (score 2) or cows with two or more indicators (score 3). Videos were selected to be equally distributed within each score. A temporal aggregation network architecture was employed with the extraction of micromotion features of the cow's body parts to detect lameness. This method produced highly accurate classifications with 99% ACC, 99% SE, 99% SP, 99% precision and 99% F-measure.

Zhao et al. (2023) used a deep convolutional network software package (DeepLabCut) to estimate the pose of cows by tracking defined anatomical key point coordinates. Side-view videos from 52 walking cows were captured using a commercial 2D digital camera. The head, back and the four feet were the user-defined key points that allowed the extraction of locomotion features regarding the motion curve of the hooves linked to the swing curve of the head and the posture of the back. Ground truth of lameness was established by three experienced human scorers assessing a total of 216 videos using a 1–5-point scoring method (Flower and Weary, 2006) and classified cows into 3 classes as non-lame (scores 1 and 2), early lame (scores 3 and 4) or severely lame (score 5). Among several classifiers evaluated, a logistic regression model produced the best ACC (87%) with 86% precision, 88% negative predictive value (NPV), 89% recall and 86% SP.

Barney et al. (2023) developed a fully automated system for lameness detection of multiple cows at the same time. In Holstein-Friesian cows (n=250) from a 300-cow dairy herd at Newcastle University, 25 side-view videos containing 10 cows each were captured using a commercial 2D digital camera. Three experienced human scorers assessed the mobility of the cows appearing in these videos using the AHDB scoring method and the average score was set as the ground truth for lameness. A pose estimation machine learning algorithm (Mask R-CNN, masked regions with convolutional neural network features) was trained for object detection using 15 anatomical key-points on the back, the head, the neck, the legs, and the feet. These key-points were then used to estimate deviation from best fit for the posture of the back, the position of the head and the angle of the neck. The CatBoost algorithm, a gradient boosting decision tree algorithm working with categorical data, was used for classification. The algorithm classified cows within each mobility score with an ACC of 95 ± 0.5%, 92% recall and 87% precision. Accordingly, the inter-rater agreement between predicted and visual mobility scores for the four classes was substantial with a kappa = 0.88. Moreover, the binary classification (scores 0 and 1: non-lame; scores 2 and 3: lame) produced an ACC of 100%, 92% recall and 90% precision and kappa = 0.82 for the inter-rater agreement.

A commercially-available fully automated system for real-time lameness detection was recently evaluated by Anagnostopoulos et al. (2023). The system used a 2D surveillance camera mounted over an exit race capturing footage of walking cows from an overhead angle. An object-tracking algorithm identified coat patterns and the shape of the cow's body. Certain reference points on the body of the cow were marked and their coordinates across frames recorded. The stored information was then processed by a convolutional neural network architecture which produced a mobility score for each cow on a scale from 0 to 100, with higher values representing increased likelihood of lameness. Each increment of 25 points corresponded to one grade of the 0–3 four-grade AHDB scoring method.

This system was validated using the visual mobility scores of two experienced human scorers as the ground truth, who scored approximately 7,000 dairy cows during 19 whole-herd scoring sessions in three dairy farms equipped with a camera. Binary transformed visual scores (0 and 1: non-lame; 2 and 3: lame) produced an overall PA > 86%, kappa = 0.32–0.40 and Gwet's coefficient = 0.80–0.83. It is noteworthy that the inter-rater agreement between the two human scorers scoring the same 903 cows was similar to that between the system and the humans, producing 88% percentage agreement, kappa = 0.41 and Gwet's coefficient = 0.85. The accuracy of the system in detecting cows bearing potentially painful lesions considered to impair mobility was also assessed. The systematic recording of foot lesions graded for severity in all four feet of 84 cows during trimming sessions was used as the ground truth. The system was more sensitive than the human assessor in identifying cows with potentially painful lesions, with 52% SE and 81% SP compared to the assessor's 29% and 89% respectively.

A larger scale validation study for the same system was conducted with 29 visits for whole milking herd mobility scoring in seven dairy farms performed by four experienced veterinarians (Siachos et al., 2023). Approximately 27000 visual mobility scores were matched to the automatically generated scores. Additionally, data regarding the presence and the severity of any foot lesion from all four feet were collected from 991 cows during 17 trimming sessions in three farms. All cows were mobility scored by the system, while 340 were also scored by the same human scorer 1–3 days before trimming. Percentage agreement, kappa and Gwet's agreement coefficient for the binary (lame vs. non-lame) categorical agreement ranged from 82% to 86%, from 0.23 to 0.41 and from 0.76 to 0.83, respectively. Additionally, ACC, SE and SP of the system and the human scores in detecting cows bearing potentially painful lesions were 83%, 40% and 88%, and 80%, 53% and 83%, respectively.

Limitations of current studies and future prospects

Strengths and limitations of currently available automated systems have been thoroughly described (Schlageter-Tello et al., 2014b; Alsaaod et al., 2019; O'Leary et al., 2020; Nejati et al., 2023). The same could apply to automated systems using AI instead of conventional methods regarding the sensors, farm applicability and performance. A systematic literature review described thoroughly the characteristics, advantages and limitations of the different ML techniques that can apply as classification models to livestock health and welfare monitoring (Garcia et al., 2020).

The systems described in this review rely on a wide range of kinematic, kinetic and indirect methods for cattle lameness detection. Measures of accuracy reported for most of them are promising. However, with some exceptions, more independent validation studies are needed to investigate the application on-farm and justify the commercialization of the available AI capabilities. Furthermore, while most studies use visual mobility scoring as the ground truth for lameness detection, only a few studies use the inspection of the feet and recording of specific foot pathologies.

It has been suggested that automated systems for lameness detection should produce a SE > 90% and SP > 99% to justify their use by farmers (O'Leary et al., 2020). The rationale is that lower SP would result in many cows classified as false positive, unnecessarily increasing trimming labour and costs, especially in large herds. Although this is a fair argument, visual mobility scoring on-farm used as the ground truth to assess measures of accuracy of automated systems has been shown to produce notable disagreement within- and between-observers. Moreover, Logan et al. (2023) have recently shown that visual mobility scoring is highly specific (SP = 94%) but has poor sensitivity to correctly identify cows bearing moderate and severe lesions (SE = 35% and 43%, respectively), by scoring and recording foot lesions of approximately 600 cows with an overall lameness prevalence (mobility score 2 and 3) of 12%. Besides, there is lack of consensus on which lesions and at what grade of severity impair mobility of cows to a detectable degree. It is possible that a cow bearing mild lesions affecting mobility and identified by an automated system as moderately lame could be treated in the early stages and would not deteriorate if presented to the trimmer. Early intervention has been shown to effectively reduce lameness prevalence (Leach et al., 2012; Groenevelt et al., 2014). Whether a human observer could correctly classify this cow as lame is uncertain. This gap in our knowledge does not allow us to properly validate an automated system carrying out a reasoning task otherwise performed by a trained human observer.

Moreover, we believe that the cross-sectional observation of cows' mobility and of the presence of foot pathologies during a trimming session is imperfect when used as the ground truth of lameness. Foot lesions represent mostly chronic pathologies, and it is unclear at which stages of their development they become clinically apparent by impaired mobility. Therefore, given that AI can handle large and complex data, effort should be put into identifying cows at the early developmental stages of lesions with mild signs of lameness considering the cow's individuality and the farm's characteristics. We could obtain useful information and gain a better understanding of how each lesion type affects mobility by training an algorithm using longitudinal data from individual cows with known foot lesions history from a common starting time-point (i.e. at drying-off or early lactation routine foot trim) and repeatedly examining for development of new lesions. To effectively control lameness and at the same time, to minimise trimming costs and labour, an optimal farm-specific threshold must then be determined, which could be used as the decisive point for early intervention dependent to the farm's needs and goals.

CRediT authorship contribution statement

Georgios Oikonomou: Writing – review & editing, Methodology, Conceptualization. **Robert F. Smith:** Writing – review & editing. **Nektarios Siachos:** Writing – original draft, Methodology, Investigation, Conceptualization. **Joseph M. Neary:** Writing – review & editing, Investigation.

Declaration of Competing Interest

None of the authors has any financial or personal relationships that could inappropriately influence or bias the content of the paper.

References

Alawneh, J.I., Laven, R.A., Stevenson, M.A., 2012. Interval between detection of lameness by locomotion scoring and treatment for lameness: a survival analysis. Vet. J. 193, 622–625.
Alsaaod, M., Fadul, M., Steiner, A., 2019. Automatic lameness detection in cattle. Vet. J. 246, 35–44.
Anagnostopoulos, A., Griffiths, B.E., Siachos, N., Neary, J., Smith, R.F., Oikonomou, G., 2023. Initial validation of an intelligent video surveillance system for automatic detection of dairy cattle lameness. Front. Vet. Sci. 10, 1111057.
Balasso, P., Marchesini, G., Ughelini, N., Serva, L., Andrighetto, I., 2021. Machine learning to detect posture and behaviour in dairy cows: information from an accelerometer on the animal's left flank. Animals 11, 2792.
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 behaviour differences associated with lameness in dairy cattle. J. Dairy Sci. 101, 6310–6321.
Barney, S., Dlay, S., Crowe, A., Kyriazakis, I., Leach, M., 2023. Deep learning pose estimation for multi-cattle lameness detection. Sci. Rep. 13, 1–19.
Beggs, D.S., Jongman, E.C., Hemsworth, P.E., Fisher, A.D., 2019. Lame cows on Australian dairy farms: a comparison of farmer-identified lameness and formal lameness scoring, and the position of lame cows within the milking order. J. Dairy Sci. 102, 1522–1529.
Bennett, R.M., Barker, Z.E., Main, D.C.J., Whay, H.R., Leach, K.A., 2014. Investigating the value dairy farmers place on a reduction of lameness in their herds using a willingness to pay approach. Vet. J. 199, 72–75.
Borghart, G.M., O'Grady, L.E., Somers, J.R., 2021. Prediction of lameness using automatically recorded activity, behaviour and production data in post-parturient Irish dairy cows. Irish Vet. J. 74, 1–10.
Busin, V., Viora, L., King, G., Tomlinson, M., Lekernec, J., Jonsson, N., Fioranelli, F., 2019. Evaluation of lameness detection using radar sensing in ruminants. Vet. Rec. 185, 572.
Carslake, C., Vázquez-Diosdado, J.A., Kaler, J., 2021. Machine learning algorithms to classify and quantify multiple behaviours in dairy calves using a sensor–moving beyond classification in precision livestock. Sensors 21, 1–14.
Chapinal, N., de Passillé, A.M., Weary, D.M., von Keyserlingk, M.A.G., Rushen, J., 2009. Using gait score, walking speed, and lying behaviour to detect hoof lesions in dairy cows. J. Dairy Sci. 92, 4365–4374.
Charfeddine, N., Pérez-Cabal, M.A., 2017. Effect of claw disorders on milk production, fertility, and longevity, and their economic impact in Spanish Holstein cows. J. Dairy Sci. 100, 653–665.
Collick, D.W., Ward, W.R., Dobson, H., 1989. Associations between types of lameness and fertility. Vet. Rec. 125, 103–106.
Cutler, J.H.H., Rushen, J., de Passillé, A.M., Gibbons, J., Orsel, K., Pajor, E., Barkema, H. W., Solano, L., Pellerin, D., Haley, D., et al., 2017. Producer estimates of prevalence and perceived importance of lameness in dairy herds with tie-stalls, free-stalls, and automated milking systems. J. Dairy Sci. 100, 9871–9880.
Dahl-Pedersen, K., Foldager, L., Herskin, M.S., Houe, H., Thomsen, P.T., 2018. Lameness scoring and assessment of fitness for transport in dairy cows: agreement among and between farmers, veterinarians and livestock drivers. Res. Vet. Sci. 119, 162–166.
De Vries, A., Bliznyuk, N., Pinedo, P., 2023. Invited Review: Examples and opportunities for artificial intelligence (AI) in dairy farms. Appl. Anim. Sci. 39, 14–22.
Dolecheck, K., Bewley, J., 2018. Animal board invited review: Dairy cow lameness expenditures, losses and total cost. Animal 12, 1462–1474.
Edwardes, F., Van der Voort, M., Hogeveen, H., 2022. The economics of sensor-based management of dairy cow suboptimal mobility. J. Dairy Sci. 105, 9682–9701.
Fatima, M., Pasha, M., 2017. Survey of machine learning algorithms for disease diagnostic. J. Intell. Learn. Syst. Appl. 9, 1–16.
Flower, F.C., Weary, D.M., 2006. Effect of hoof pathologies on subjective assessments of dairy cow gait. J. Dairy Sci. 89, 139–146.
Garcia, E., König, K., Allesen-Holm, B.H., Klaas, I.C., Amigo, J.M., Bro, R., Enevoldsen, C., 2015. Experienced and inexperienced observers achieved relatively high within-observer agreement on video mobility scoring of dairy cows. J. Dairy Sci. 98, 4560–4571.
Garcia, R., Aguilar, J., Toro, M., Pinto, A., Rodriguez, P., 2020. A systematic literature review on the use of machine learning in precision livestock farming. Comput. Electron. Agric. 179, 105826.
Gardenier, J., Underwood, J., Weary, D.M., Clark, C.E.F., 2021. Pairwise comparison locomotion scoring for dairy cattle. J. Dairy Sci. 104, 6185–6193.
Gertz, M., Große-Butenuth, K., Junge, W., Maassen-Francke, B., Renner, C., Sparenberg, H., Krieter, J., 2020. Using the XGBoost algorithm to classify neck and leg activity sensor data using on-farm health recordings for locomotor-associated diseases. Comput. Electron. Agric. 173, 105404.
Ghotoorlar, S.M., Mehdi Ghamsari, S., Nowrouzian, I., Shiry Ghidary, S., 2012. Lameness scoring system for dairy cows using force plates and artificial intelligence. Vet. Rec. 170, 126.
Gibbons, J., Vasseur, E., Rushen, J., De Passillé, A.M., 2012. A training programme to ensure high repeatability of injury scoring of dairy cows. Anim. Welf. 21, 379–388.
Groenevelt, M., Main, D.C.J., Tisdall, D., Knowles, T.G., Bell, N.J., 2014. Measuring the response to therapeutic foot trimming in dairy cows with fortnightly lameness scoring. Vet. J. 201, 283–288.
Haladjian, J., Haug, J., Nüske, S., Bruegge, B., 2018. A wearable sensor system for lameness detection in dairy cattle. Multimodal Technol. Interact. 2, 27.
Han, S., Fuentes, A., Yoon, S., Jeong, Y., Kim, H., Sun Park, D., 2023. Deep learning-based multi-cattle tracking in crowded livestock farming using video. Comput. Electron. Agric. 212, 108044.
Horseman, S.V., Roe, E.J., Huxley, J.N., Bell, N.J., Mason, C.S., Whay, H.R., 2014. The use of in-depth interviews to understand the process of treating lame dairy cows from the farmers' perspective. Anim. Welfare 23, 157–165.

Hossain, M.E., Kabir, M.A., Zheng, L., Swain, D.L., McGrath, S., Medway, J., 2022. A systematic review of machine learning techniques for cattle identification: datasets, methods and future directions. Artif. Intell. Agric. 6, 138–155.

Huxley, J.N., 2013. Impact of lameness and claw lesions in cows on health and production. Livestock Sci. 256 64–70.

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

Jiang, B., Song, H., Wang, H., Li, C., 2022. Dairy cow lameness detection using a back curvature feature. Comput. Electron. Agric. 194, 106729.

Jiang, B., Yin, X., Song, H., 2020. Single-stream long-term optical flow convolution network for action recognition of lameness dairy cow. Comput. Electron. Agric. 175, 105536.

Kang, X., Li, S., Li, Q., Liu, G., 2022. Dimension-reduced spatiotemporal network for lameness detection in dairy cows. Comput. Electron. Agric. 197, 106922.

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, 10628–10638.

Kaniyamattam, K., Hertl, J., Lhermie, G., Tasch, U., Dyer, R., Gröhn, Y.T., 2020. Cost benefit analysis of automatic lameness detection systems in dairy herds: a dynamic programming approach. Prevent. Vet. Med. 178, 104993.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159.

Leach, K.A., Tisdall, D.A., Bell, N.J., Main, D.C.J., Green, L.E., 2012. The effects of early treatment for hindlimb lameness in dairy cows on four commercial UK farms. Vet. J. 193, 626–632.

Leach, K.A., Whay, H.R., Maggs, C.M., Barker, Z.E., Paul, E.S., Bell, A.K., Main, D.C.J., 2010. Working towards a reduction in cattle lameness: 1. Understanding barriers to lameness control on dairy farms. Res. Vet. Sci. 89, 311–317.

Li, Q., Chu, M., Kang, X., Liu, G., 2023. Temporal aggregation network using micromotion features for early lameness recognition in dairy cows. Comput. Electron. Agric. 204, 107562.

Liang, D., Arnold, L.M., Stowe, C.J., Harmon, R.J., Bewley, J.M., 2017. Estimating US dairy clinical disease costs with a stochastic simulation model. J. Dairy Sci. 100, 1472–1486.

Linardopoulou, K., Viora L., Fioranelli F., Kernec J., Abbasi Q., King G., Borelli E., Jonsson N., 2022. Time-series observations of cattle mobility: accurate label assignment from multiple assessors, and association with lesions detected in the feet. Proceeding of the 31st World Buiatrics Congress, Madrid, Spain, pp. 297.

Logan, F., McAloon, C.G., Ryan, E.G., O'Grady, L., Duane, M., Deane, B., McAloon, C.I., 2023. Sensitivity and specificity of mobility scoring for the detection of foot lesions in pasture based Irish dairy cows. J. Dairy Sci. (In press.).

McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochemia Medica 22, 276–282.

Melendez, P., Bartolome, J., Archbald, L.F., Donovan, A., 2003. The association between lameness, ovarian cysts and fertility in lactating dairy cows. Theriogenology 59, 927–937.

Nejati, A., Bradtmueller, A., Shepley, E., Vasseur, E., 2023. Technology applications in bovine gait analysis: a scoping review. PLoS One 18, e0266287.

O'Leary, N.W., Byrne, D.T., O'Connor, A.H., Shalloo, L., 2020. Invited review: cattle lameness detection with accelerometers. J. Dairy Sci. 103, 3895–3911.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Moher, D., 2021. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J. Clin. Epidemiol. 134, 103–112.

Pastell, M.E., Kujalaf, M., 2007. A probabilistic neural network model for lameness detection. J. Dairy Sci. 90, 2283–2292.

Puerto, M.A., Shepley, E., Cue, R.I., Warner, D., Dubuc, J., Vasseur, E., 2021. The hidden cost of disease: II. Impact of the first incidence of lameness on production and economic indicators of primiparous dairy cows. J. Dairy Sci. 104, 7944–7955.

Robcis, R., Ferchiou, A., Berrada, M., Ndiaye, Y., Herman, N., Lhermie, G., Raboisson, D., 2023. Cost of lameness in dairy herds: an integrated bioeconomic modeling approach. J. Dairy Sci. 106, 2519–2534.

Russello, H., van der Tol, R., Kootstra, G., 2022. T-LEAP: occlusion-robust pose estimation of walking cows using temporal information. Comput. Electron. Agric. 192, 106559.

Sadiq, M.B., Ramanoon, S., Mossadeq, W.S., Mansor, R., Hussain, S.S., 2019. Dairy farmers' perceptions of and actions in relation to lameness management. Animals 9, 270.

Schlageter-Tello, A., Bokkers, E.A.M., Groot Koerkamp, P.W.G., Van Hertem, T., Viazzi, S., Romanini, C.E.B., Halachmi, I., Bahr, C., Berckmans, D., Lokhorst, K., 2014a. Effect of merging levels of locomotion scores for dairy cows on intra- and interrater reliability and agreement. J. Dairy Sci. 97, 5533–5542.

Schlageter-Tello, A., Bokkers, E.A.M., Koerkamp, P.W.G.G., Van Hertem, T., Viazzi, S., Romanini, C.E.B., Halachmi, I., Bahr, C., Berckmans, D., Lokhorst, K., 2014b. Manual and automatic locomotion scoring systems in dairy cows: a review. Prevent. Vet. Med. 116, 12–25.

Schmeling, L., Elmamooz, G., Hoang, P.T., Kozar, A., Nicklas, D., Sünkel, M., Thurner, S., Rauch, E., 2021. Training and validating a machine learning model for the sensor-based monitoring of lying behaviour in dairy cows on pasture and in the barn. Animals 11, 2660.

Shrestha, A., Loukas, C., Le Kernec, J., Fioranelli, F., Busin, V., Jonsson, N., King, G., Tomlinson, M., Viora, L., Voute, L., 2018. Animal lameness detection with radar sensing. IEEE Geosci. Remote Sens. Lett. 15, 1189–1193.

Siachos N., Anagnostopoulos A., Griffiths B.E., Neary J.M., Smith R.F., Oikonomou G., 2023. Evaluation of an automated cattle lameness detection system. Proceedings of 74th Annual Meeting of the European Federation of Animal Science, Lyon, France, pp. 769.

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, 1179–1187.

Stygar, A.H., Gómez, Y., Berteselli, G.V., Dalla Costa, E., Canali, E., Niemi, J.K., Llonch, P., Pastell, M., 2021. A systematic review on commercially available and validated sensor technologies for welfare assessment of dairy cattle. Front. Vet. Sci. 8, 634338.

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, 105286.

Thomsen, P.T., Munksgaard, L., Togersen, F.A., 2008. Evaluation of a lameness scoring system for dairy cows. J. Dairy Sci. 91, 119–126.

Thomsen, P.T., Shearer, J.K., Houe, H., 2023. Prevalence of lameness in dairy cows. Vet. J. 295, 105975.

Van Nuffel, A., Zwertvaegher, I., Pluym, L., Van Weyenberg, S., Thorup, V.M., Pastell, M., Sonck, B., Saeys, W., 2015a. Lameness detection in dairy cows: part 1. How to distinguish between non-lame and lame cows based on differences in locomotion or behaviour. Animals 5, 838–860.

Van Nuffel, A., Zwertvaegher, I., Van Weyenberg, S., Pastell, M., Thorup, V.M., Bahr, C., Sonck, B., Saeys, W., 2015b. Lameness detection in dairy cows: part 2. Use of sensors to automatically register changes in locomotion or behaviour. Animals 5, 861–885.

Volkmann, N., Kulig, B., Hoppe, S., Stracke, J., Hensel, O., Kemper, N., 2021. On-farm detection of claw lesions in dairy cows based on acoustic analyses and machine learning. J. Dairy Sci. 104, 5921–5931.

Volkmann, N., Stracke, J., Kemper, N., 2019. Evaluation of a gait scoring system for cattle by using cluster analysis and Krippendorff's α reliability. Vet. Rec. 184, 220.

Wagner, N., Antoine, V., Mialon, M.M., Lardy, R., Silberberg, M., Koko, J., Veissier, I., 2020. Machine learning to detect behavioural anomalies in dairy cows under subacute ruminal acidosis. Comput. Electron. Agric. 170, 105233.

Walker, S.L., Smith, R.F., Routly, J.E., Jones, D.N., Morris, M.J., Dobson, H., 2008. Lameness, activity time-budgets, and estrus expression in dairy cattle. J. Dairy Sci. 91, 4552–4559.

Whay, H.R., Main, D.C.J., Green, L.E., Webster, A.J.F., 2003. Assessment of the welfare of dairy cattle using animal-based measurements: Direct observations and investigation of farm records. Vet. Rec. 153, 197–202.

Wu, D., Wu, Q., Yin, X., Jiang, B., Wang, H., He, D., Song, H., 2020. Lameness detection of dairy cows based on the YOLOv3 deep learning algorithm and a relative step size characteristic vector. Biosyst. Eng. 189, 150–163.

Zambelis, A., Saadati, M., Dallago, G.M., Stecko, P., Boyer, V., Parent, J.P., Pedersoli, M., Vasseur, E., 2021. Automation of video-based location tracking tool for dairy cows in their housing stalls using deep learning. Smart Agric. Technol. 1, 100015.

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, K., Zhang, M., Ji, J., Zhang, R., Bewley, J.M., 2023. Automatic lameness scoring of dairy cows based on the analysis of head- and back-hoof linkage features using machine learning methods. Biosyst. Eng. 230, 424–441.

Zheng, Z., Zhang, X., Qin, L., Yue, S., Zeng, P., 2023. Cows' legs tracking and lameness detection in dairy cattle using video analysis and Siamese neural networks. Comput. Electron. Agric. 205, 107618.