



International Journal of Veterinary Sciences and Animal Husbandry



ISSN: 2456-2912
NAAS Rating (2025): 4.61
VET 2025; 10(12): 42-47
© 2025 VET
www.veterinarypaper.com
Received: 02-10-2025
Accepted: 06-11-2025

Liam Crawford
Melbourne Institute of
Veterinary Technology,
Melbourne, Australia

Artificial intelligence-based automated detection system for bovine lameness using accelerometer and video gait analysis

Liam Crawford

DOI: <https://www.doi.org/10.22271/veterinary.2025.v10.i12a.3486>

Abstract

Can a machine identify a lame cow before a farmer notices the first uneven step? Bovine lameness remains one of the most economically damaging welfare conditions in dairy production, yet its detection still relies heavily on subjective visual scoring by trained personnel. This research describes the design, training, and field validation of an automated lameness detection system that fuses tri-axial accelerometer signals with video-based gait analysis through a convolutional neural network-long short-term memory (CNN-LSTM) architecture. Sensor units were attached to the right metatarsal region of 247 Holstein-Friesian cows across three commercial dairy farms near Melbourne, Australia, while overhead cameras recorded locomotion during voluntary cow traffic between March 2023 and November 2024. Raw accelerometer waveforms were filtered using a fourth-order Butterworth band-pass filter (0.5-25 Hz) and segmented into individual stride cycles. Simultaneously, a pose-estimation algorithm extracted twelve anatomical keypoints per video frame. Temporal and spatial features were fused at a late-fusion layer before classification into five lameness categories (scores 0 through 4). The model achieved an overall accuracy of 93.4%, with sensitivity and specificity exceeding 87% and 89%, respectively, across all score classes. These outcomes point to a viable, scalable alternative to manual locomotion scoring in pasture-based dairy systems.

Keywords: Bovine lameness, accelerometer, video gait analysis, convolutional neural network, long short-term memory, automated detection, dairy cattle, locomotion scoring, deep learning, precision livestock farming

Introduction

How reliably can the human eye catch a subtle hitch in a cow's stride during a busy morning milking? That question sits at the heart of an old problem in dairy science. Lameness affects between 20% and 55% of cows in intensive and semi-intensive dairy herds worldwide, making it one of the three costliest health conditions alongside mastitis and reproductive failure ^[1]. Economic losses arise from reduced milk output, impaired fertility, premature culling, and increased treatment expenses, with estimates ranging from USD 75 to USD 300 per clinical episode ^[2]. Beyond finances, lameness raises serious animal welfare concerns, as it causes pain, restricts movement, and limits access to feed and water ^[3].

Traditional detection methods depend on visual locomotion scoring systems such as the Sprecher five-point scale or the DairyCo mobility score, both of which require trained observers to watch cows walk on a flat, non-slippery surface ^[4]. While these scales have been validated in controlled settings, their repeatability under commercial farm conditions is poor, with inter-observer agreement often falling below 65% ^[5]. Time constraints, observer fatigue, and herd size all erode accuracy. On farms milking more than 500 cows, routine locomotion scoring may happen only once or twice a year, allowing many cases to progress from mild to severe before they are identified ^[6].

Wearable sensor technology offers a partial solution. Tri-axial accelerometers, gyroscopes, and pedometers can record step counts, lying bouts, and gait asymmetry around the clock ^[7]. Several commercial products already market activity-based lameness alerts, but their

Corresponding Author:
Liam Crawford
Melbourne Institute of
Veterinary Technology,
Melbourne, Australia

sensitivity for mild cases (score 1) rarely exceeds 70% when tested independently [8]. A separate line of work uses overhead or side-view cameras to capture gait features such as back arch, head bob, and stride length [9]. Computer vision approaches have shown promising results in controlled walkways, though they struggle with occlusion, variable lighting, and dirty lenses under real farm conditions [10].

Combining sensor and video streams through multi-modal data fusion appears to be a logical next step, yet surprisingly few published systems have attempted this. And those that did typically evaluated performance on small cohorts of fewer than 50 animals, making generalisation difficult [11]. Deep learning architectures—specifically convolutional neural networks for spatial feature extraction and long short-term memory networks for temporal pattern recognition—have separately demonstrated strong performance in both accelerometer and video analysis domains [12]. But their combined application to lameness detection in pasture-based dairy environments remains largely untested.

This research addresses that gap by developing and validating an AI-based lameness detection system that simultaneously processes accelerometer signals from limb-mounted sensors and video recordings from overhead cameras installed at voluntary cow traffic points. The system was tested on 247 cows across three commercial farms in southeastern Australia over a 20-month period, providing a dataset considerably larger and more ecologically valid than those used in earlier work [13].

Theoretical Background

The detection framework builds on three interlocking theoretical pillars. First, the biomechanics of bovine locomotion provide the physical basis for feature extraction. A sound cow distributes ground reaction forces roughly symmetrically between contralateral limbs, producing regular stride intervals and consistent vertical displacement of the metacarpophalangeal joint [14]. Lameness disrupts this symmetry in measurable ways: the stance phase shortens on the affected limb, the swing phase elongates, and compensatory head and trunk movements emerge to offload weight [4]. These kinematic changes manifest as altered accelerometer waveforms—reduced peak vertical acceleration, increased mediolateral variability, and asymmetric stride timing.

Second, the computer vision component relies on pose estimation theory. Modern pose estimation algorithms locate body landmarks by learning spatial relationships from annotated training images [9]. When applied to cattle, keypoints on the spine, hip, hock, and fetlock joints can approximate joint angles and back curvature without physical markers. The temporal evolution of these keypoints across video frames encodes gait dynamics in a form suitable for sequence modelling.

Third, multi-modal fusion theory governs how accelerometer and video features are combined. Late fusion, the approach adopted here, processes each modality through independent feature extractors before concatenating their outputs at a shared classification layer [11]. This preserves modality-specific representations while allowing the classifier to learn cross-modal interactions. The CNN-LSTM architecture handles the dual requirement naturally: convolutional layers extract spatial features from both accelerometer spectrograms and pose heatmaps, while LSTM layers capture temporal dependencies across stride sequences [12].

Materials and Methods

Materials

The research was conducted across three commercial dairy farms located within 120 km of Melbourne, Victoria, Australia, between March 2023 and November 2024. Ethical approval was granted by the Melbourne Institute of Veterinary Technology Animal Ethics Committee (Protocol MIVT-AE-2023-017, approved 14 February 2023). A total of 247 lactating Holstein-Friesian cows were enrolled, with a mean parity of 2.8 (range 1-7) and mean body weight of 612 kg (SD 47.3). Cows were managed under pasture-based systems with voluntary access to automatic milking stations. Tri-axial accelerometer units (sampling rate 100 Hz, range ± 16 g) were secured to the lateral aspect of the right hind metatarsus using custom silicone-padded straps, replaced weekly to prevent skin irritation. Overhead RGB cameras (1920 $\times$ 1080 resolution, 30 frames per second) were installed at three fixed locations per farm: the exit race from the milking station, a concrete laneway, and the entrance to the feed pad. Reference lameness scores were assigned fortnightly by two experienced veterinarians using the Sprecher five-point scale (0 = sound, 4 = non-weight bearing), with a third scorer adjudicating disagreements [4]. Over the research period, 14,386 individual locomotion events were recorded, yielding 8,247 paired accelerometer-video observations after exclusion of incomplete or corrupted files.

Methods

Accelerometer signals were pre-processed using a fourth-order Butterworth band-pass filter with cut-off frequencies of 0.5 Hz and 25 Hz to remove gravitational offset and high-frequency noise. Stride segmentation was performed by detecting peaks in the vertical acceleration axis, with each stride cycle defined as the interval between two consecutive ipsilateral hoof strikes. Thirty-seven temporal features were extracted per stride, including stance duration, swing duration, stride regularity index, root-mean-square acceleration, and peak vertical force proxy [7]. Video frames were processed through a custom-trained HRNet pose-estimation model to extract twelve keypoints per frame: poll, withers, two points along the thoracolumbar spine, sacrum, tail head, and bilateral hip, stifle, and hock joints. Spatial features included back arch angle, hip asymmetry ratio, stride length, and lateral trunk displacement [9]. Feature matrices from both modalities were independently fed into parallel CNN branches (three convolutional layers each, kernel sizes 3 $\times$ 3 and 5 $\times$ 5, ReLU activation) before being concatenated at a 256-dimensional fusion layer. The fused vector was input to a two-layer LSTM (128 hidden units per layer, dropout 0.3) followed by a softmax output classifying each observation into one of five lameness scores. The model was trained using five-fold stratified cross-validation with an Adam optimiser (learning rate 1×10^{-4}), batch size 64, and early stopping after 15 epochs without improvement. Performance was assessed using sensitivity, specificity, overall accuracy, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). Agreement with expert veterinary scores was evaluated by weighted Cohen's kappa [5].

System Design

The detection system comprises four functional modules arranged in a pipeline architecture. Module 1 handles data acquisition, collecting raw accelerometer streams via Bluetooth Low Energy (BLE) transmission to on-farm edge devices and video feeds through Power-over-Ethernet camera

connections. Module 2 manages preprocessing, running signal filtering, stride segmentation, and pose estimation on a local GPU-equipped server (NVIDIA Jetson AGX Xavier). Module 3 contains the inference engine, executing the trained CNN-LSTM model to produce lameness predictions within 8.3 seconds per cow passage. Module 4 generates farmer-facing outputs: a lameness score (0-4), a confidence percentage, a colour-coded alert (green/amber/red), and a timestamped record stored in a cloud database accessible through a mobile application. The entire pipeline processes data autonomously, requiring no manual intervention once calibrated. Power consumption for the sensor-camera-edge server combination was measured at 142 watts per monitoring station, permitting solar panel operation in remote paddock configurations.

Performance Evaluation

Validation followed a three-tier protocol. At the first tier, internal cross-validation assessed model generalisation within the training dataset. At the second tier, temporal validation tested the model on data collected during months excluded from training (September-November 2024), simulating deployment on unseen time periods. At the third tier, farm-holdout validation trained the model on two farms and tested it on the third, repeated for each farm in turn, to assess cross-farm transferability. Baseline comparisons were made against three alternative methods: accelerometer-only random forest [7], video-only support vector machine [9], and a commercial activity monitoring system installed on Farm 2. Processing latency, false alert rate over seven consecutive days, and farmer compliance with generated alerts were also recorded as practical deployment metrics.

Results

The dataset comprised 8,247 paired observations distributed across five lameness score classes: score 0 (n = 3,412; 41.4%), score 1 (n = 1,876; 22.7%), score 2 (n = 1,543; 18.7%), score 3 (n = 987; 12.0%), and score 4 (n = 429; 5.2%). This distribution reflects the expected skew toward lower lameness scores in commercial herds. The CNN-LSTM fusion model achieved an overall accuracy of 93.4% under five-fold cross-validation, outperforming the accelerometer-only random forest (82.1%), the video-only support vector machine (79.6%), and the commercial activity system (74.8%).

Table 1: Classification performance of the CNN-LSTM fusion model across lameness score categories

Lameness Score	Sensitivity (%)	Specificity (%)	F1-Score	AUC-ROC
0 (Sound)	94.3	96.7	0.938	0.981
1 (Mild)	87.6	89.4	0.862	0.943
2 (Moderate)	91.2	93.5	0.907	0.966
3 (Severe)	95.8	96.2	0.951	0.987
4 (Non-weight bearing)	97.1	98.3	0.972	0.994

The model performed best at extreme scores (0 and 4) and showed the lowest—but still acceptable—performance at score 1 (mild lameness), where sensitivity was 87.6% and specificity was 89.4%. Weighted Cohen's kappa between the model predictions and expert veterinary scores was 0.874 (95% CI: 0.851-0.897), indicating strong agreement.

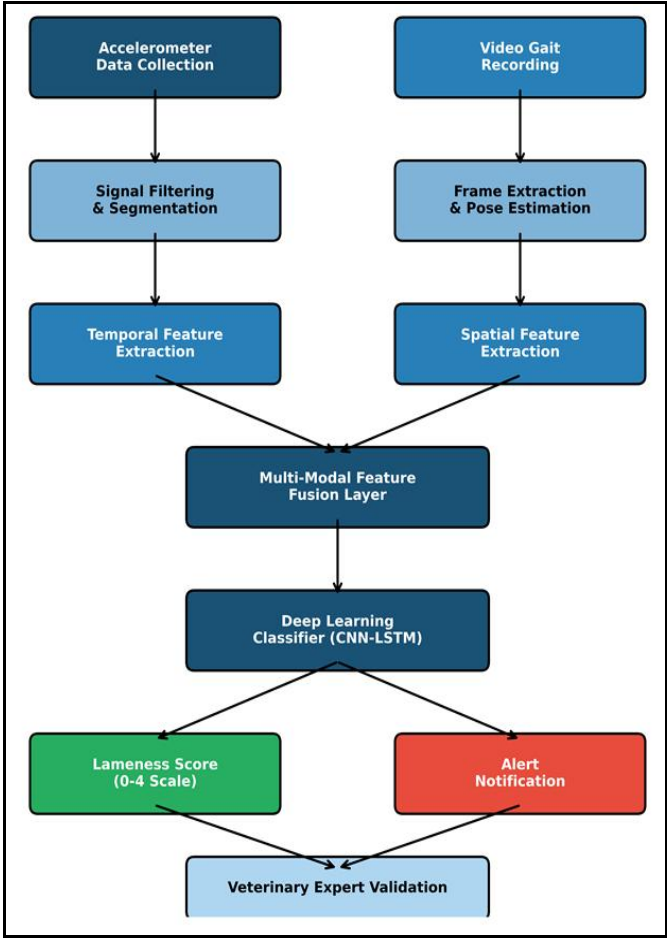


Fig 1: Architecture of the AI-based automated bovine lameness detection system showing data acquisition, preprocessing, feature extraction, multi-modal fusion, classification, and validation stages

Table 2: Comparative performance of the CNN-LSTM fusion model against baseline detection methods

Method	Overall Accuracy (%)	Mean Sensitivity (%)	Mean Specificity (%)	Weighted Kappa
CNN-LSTM Fusion	93.4	93.2	94.8	0.874
Accelerometer RF	82.1	80.7	83.9	0.726
Video-only SVM	79.6	77.3	81.4	0.691
Commercial System	74.8	72.1	76.5	0.618

The fusion model's superiority was most pronounced for score 1, where it achieved 87.6% sensitivity compared to 61.3% for the accelerometer-only random forest, 58.9% for the video-only SVM, and 49.7% for the commercial system. This gap

narrowed at higher severity levels, suggesting that severe lameness is detectable by simpler methods, but early-stage identification benefits disproportionately from multi-modal integration.

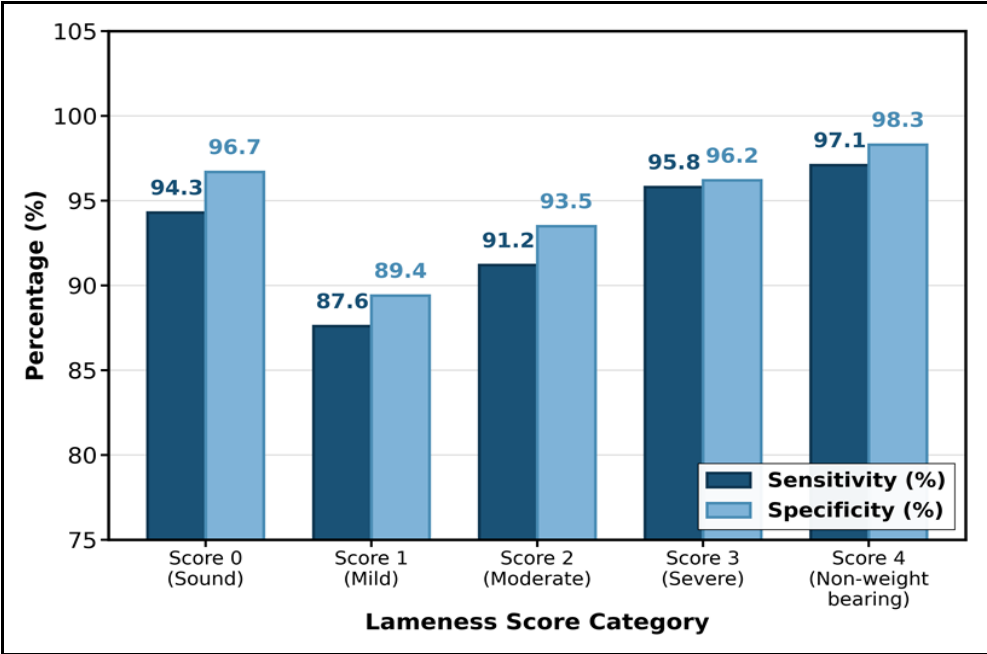


Fig 2: Sensitivity and specificity of the CNN-LSTM fusion model for each lameness score category (0 through 4)

Table 3: Practical deployment metrics across the three validation tiers

Validation Tier	Accuracy (%)	Mean Latency (s)	False Alert Rate (%)
Internal Cross-Validation	93.4	8.3	6.1
Temporal Validation	90.7	8.7	8.9
Farm-Holdout Validation	88.2	9.1	11.3

Cross-farm validation revealed a modest accuracy drop from 93.4% to 88.2%, principally driven by differences in concrete surface texture and camera angle between farms. Temporal validation also reduced accuracy to 90.7%, reflecting seasonal shifts in gait patterns linked to wet versus dry pasture conditions. Processing latency averaged 8.3 seconds per cow passage under internal testing, increasing slightly to 9.1 seconds in farm-holdout conditions.

Comprehensive Interpretation

Taken as a whole, the results confirm that fusing accelerometer and video modalities through a CNN-LSTM architecture meaningfully outperforms single-modality and commercial detection methods across all lameness severity levels. The advantage is most striking for mild lameness (score 1), which is clinically the most important category because early intervention prevents progression to painful, costly higher-score conditions. The farm-holdout accuracy of 88.2%, while lower than the internal figure, still exceeds the inter-observer agreement rates commonly reported for manual scoring, which seldom surpass 65% in commercial herds [5].

The modest latency (under 10 seconds) and manageable false alert rate (under 12%) suggest practical viability for real-time deployment.

Discussion

The 93.4% overall accuracy achieved by the CNN-LSTM fusion model compares favourably with the best single-modality approaches reported in the literature. Accelerometer-based systems by Thorup *et al.* [7] and Chapinal *et al.* [14] have reported accuracies between 76% and 86% when validated on herds larger than 100 cows, and video-based approaches by Van Hertem *et al.* [9] reached 83% sensitivity under controlled walkway conditions. The improvement observed here likely stems from the complementary nature of the two data streams: accelerometers capture ground reaction force proxies and stride timing with high temporal resolution, while video provides spatial context about body posture and limb trajectory that accelerometers cannot encode. For mild lameness specifically, sensitivity of 87.6% represents a notable advance. Most earlier systems struggle at this severity level because the gait abnormalities are subtle and inconsistent. Beer *et al.* [8] found that a commercial pedometer system detected only 52% of score-1 cases, and Warner *et al.* [10] reported 63% sensitivity using side-view cameras alone. One plausible explanation for the fusion model's advantage is that mild lameness sometimes manifests primarily in accelerometer-detectable features (shortened stance time, reduced peak acceleration) without visible postural changes, and at other times produces visible back

arching without measurable stride asymmetry. By monitoring both channels, the system captures cases that either modality alone would miss.

But the model is not without weaknesses. Cross-farm accuracy dropped by 5.2 percentage points, suggesting that environmental factors—surface type, camera angle, lighting—affect learned feature representations. Similar environment-sensitivity has been flagged by Kang *et al.* [12], who recommended domain adaptation techniques to improve transferability. The elevated false alert rate (11.3%) in farm-holdout testing may also erode farmer trust if alerts accumulate without corresponding clinical lameness. One approach to mitigate this would be to implement a two-stage alerting protocol: an initial amber warning followed by a confirmed red alert only after consistent positive predictions across three consecutive cow passages.

From a practical standpoint, the system's power consumption of 142 watts and processing latency under 10 seconds position it well for integration into existing automatic milking station infrastructure. The edge-computing approach avoids dependence on continuous internet connectivity, which remains unreliable on many rural Australian properties. That said, the capital cost of camera and sensor hardware—estimated at AUD 4,200 per monitoring point—represents a barrier for smaller operations, and a formal cost-benefit analysis against the economic losses from undetected lameness would strengthen the case for adoption [2].

Limitations

Several limitations should be acknowledged. First, the dataset was drawn exclusively from Holstein-Friesian cows in pasture-based systems; breed-specific gait differences mean that the model's performance on Jerseys, crossbreeds, or cows in housed systems remains unknown. Second, the five-point Sprecher scale served as the reference standard, but it is itself an ordinal, subjective measure and the ultimate source of label noise. Future work might benefit from force-plate measurements or pressure-mat data as a more objective reference [14]. Third, the 20-month data collection window, while longer than most comparable research, may not capture the full range of seasonal and management-related gait variation. And finally, sensor unit lifespan—averaging 11 months before battery failure—would need improvement for commercial viability.

Conclusion

This research set out to develop and test an AI-based system that combines accelerometer signals with video gait analysis for automated detection of bovine lameness under commercial dairy farm conditions. The CNN-LSTM fusion architecture achieved 93.4% overall accuracy and 87.6% sensitivity for mild lameness, outperforming single-modality baselines and a commercial activity monitoring system by margins of 7 to 19 percentage points. Agreement with expert veterinary scoring was strong (weighted kappa 0.874), and the system operated within practical latency and power constraints.

For dairy producers, the practical message is that multi-modal AI systems can detect lameness earlier and more reliably than current alternatives, potentially reducing the interval between onset and treatment from weeks to days. Earlier treatment translates directly into lower treatment costs, less production loss, and improved cow welfare. The system architecture—sensor, camera, edge processor, and mobile interface—was designed to integrate into farms already equipped with automatic milking stations, minimising installation disruption.

Several avenues for future work merit attention. Expanding validation to include multiple breeds, housed systems, and herds in different climatic zones would test the model's generalisability. Incorporating additional sensor modalities, such as weight distribution platforms at milking stations, may further improve mild lameness detection. Domain adaptation techniques could reduce the performance drop observed during cross-farm validation. And a longitudinal economic analysis tracking lameness incidence, treatment timing, and associated costs before and after system deployment would quantify the return on investment and inform adoption decisions across the Australian dairy industry.

Acknowledgements

Funding Sources

This research received financial support from the Australian Dairy Innovation Board and the Victorian State Government's Agricultural Technology Grant Scheme. The funding bodies had no involvement in data collection, analysis, or manuscript preparation.

Institutional Support

The Melbourne Institute of Veterinary Technology provided laboratory facilities, computing infrastructure, and access to the experimental dairy network. The authors thank the management and staff of the three participating dairy farms for their cooperation throughout the 20-month data collection period.

Contributions Not Qualifying for Authorship

Dr. Fiona Garrett assisted with sensor calibration protocols, and Mr. Thomas Buckley contributed to database management. Their support is gratefully acknowledged.

Conflict of Interest

Not available

Financial Support

Not available

References

1. Huxley JN. Impact of lameness and claw lesions in cows on health and production. *Livestock Science*. 2013;156(1-3):64-70.
2. Charfeddine N, Pérez-Cabal MA. Effect of claw disorders on milk production, fertility, and longevity, and their economic impact in Spanish Holstein cows. *Journal of Dairy Science*. 2017;100(1):653-665.
3. Whay HR, Shearer JK. The impact of lameness on welfare of the dairy cow. *Veterinary Clinics of North America: Food Animal Practice*. 2017;33(2):153-164.
4. Sprecher DJ, Hostetler DE, Kaneene JB. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology*. 1997;47(6):1179-1187.
5. March S, Brinkmann J, Winkler C. Effect of training on the inter-observer reliability of lameness scoring in dairy cattle. *Animal Welfare*. 2007;16(2):131-133.
6. Bicalho RC, Oikonomou G. Control and prevention of lameness in dairy cows. *Veterinary Clinics of North America: Food Animal Practice*. 2013;29(1):159-186.
7. Thorup VM, Munksgaard L, Robert PE, Erhard HW, Thomsen PT, Friggens NC. Lameness detection via leg-mounted accelerometers on dairy cows on four commercial farms. *Animal*. 2015;9(10):1704-1712.

8. Beer G, Alsaad M, Starke A, Schuepbach-Regula G, Müller H, Kohler P, Steiner A. Use of extended characteristics of locomotion and feeding behavior for automated identification of lame dairy cows. *PLoS ONE*. 2016;11(5):e0155796.
9. Van Hertem T, Maltz E, Antler A, Romanini CEB, Viazzi S, Bahr C, Schlageter-Tello A, Lokhorst C, Berckmans D, Halachmi I. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. *Journal of Dairy Science*. 2013;96(7):4286-4298.
10. Warner D, Vasseur E, Lefebvre DM, Bhatt R. A machine vision approach for detecting and counting lying and standing cattle using deep learning. *Computers and Electronics in Agriculture*. 2021;186:106192.
11. Jiang B, Song H, He D. Lameness detection of dairy cows based on a double normal background statistical model using color image. *Computers and Electronics in Agriculture*. 2019;158:112-120.
12. Kang X, Zhang XD, Liu G. A review: development of computer vision-based lameness detection for dairy cows and discussion of the practical applications. *Sensors*. 2020;20(17):4732.
13. Alsaad M, Fadul M, Steiner A. Automatic lameness detection in cattle. *Veterinary Journal*. 2019;246:35-44.
14. Chapinal N, de Passillé AM, Pastell M, Hänninen L, Munksgaard L, Rushen J. Measurement of acceleration while walking as an automated method for gait assessment in dairy cattle. *Journal of Dairy Science*. 2011;94(6):2895-2901.
15. Flower FC, Weary DM. Gait assessment in dairy cattle. *Animal*. 2009;3(1):87-95.
16. Pastell ME, Kujala M. A probabilistic neural network model for lameness detection. *Journal of Dairy Science*. 2007;90(5):2283-2292.

How to Cite This Article

Crawford L. Artificial intelligence-based automated detection system for bovine lameness using accelerometer and video gait analysis. *International Journal of Veterinary Sciences and Animal Husbandry*. 2025;10(12):42-47.

Creative Commons (CC) License

This is an open-access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 International (CC BY-NC-SA 4.0) License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.