International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 8, Number 2, Year 2026
DOI: https://doi.org/10.62051/ijcsit.v8n2.01
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Cattle Behavior Recognition and Modeling: A Review of Sensing Modalities, Intelligent Methods, and Engineering Applications

Xiangyi Zeng

Henan University of Science and Technology, Luoyang 471000, China

ABSTRACT

Cattle behavior recognition is a key component of precision livestock farming, enabling continuous monitoring of animal health, welfare, and management conditions. This review provides a structured overview of recent advances in cattle behavior recognition from an engineering perspective. Existing studies are systematically summarized according to three main sensing paradigms: contact-based sensing using wearable devices, non-contact sensing based on vision and other remote techniques, and multimodal fusion approaches. Wearable sensing methods are reviewed with respect to sensor types, data acquisition, feature representation, and temporal modeling, demonstrating strong capability for fine-grained behavior recognition. Non-contact approaches, particularly vision-based methods, are discussed in terms of detection, tracking, pose estimation, and spatiotemporal analysis under realistic farm environments. In addition, multimodal fusion and behavior modeling methods are highlighted for their ability to improve robustness and support long-term monitoring, including digital twin–oriented behavior perception. Finally, engineering system design and practical deployment issues, such as energy efficiency, communication, edge computing, and long-term maintainability, are discussed, together with key challenges and future research directions. This review aims to provide a concise reference for developing reliable and scalable cattle behavior monitoring systems.

KEYWORDS

Cattle behavior recognition; Wearable sensing; Vision-based monitoring; Multimodal fusion; Precision livestock farming

1. INTRODUCTION

The automatic recognition and modeling of cattle behavior is a key enabling technology for precision livestock farming, as it provides continuous and objective information for animal health monitoring, welfare assessment, and management decision support [1–3]. Traditional behavior assessment relies heavily on manual observation, which is labor-intensive, subjective, and difficult to scale in large commercial farms [4]. To overcome these limitations, data-driven behavior recognition methods based on sensing technologies and intelligent analysis have been widely investigated in recent years [1–6].

Contact-based sensing represents one of the most mature and widely adopted approaches for cattle behavior recognition. By deploying wearable devices equipped with accelerometers, gyroscopes, acoustic sensors, and positioning modules, fine-grained motion and activity information can be continuously collected from individual animals [2, 5–8]. Such approaches have been successfully applied to the recognition of feeding, rumination, locomotion, estrus, and lameness-related behaviors, demonstrating high sensitivity to subtle body movements and temporal activity patterns [5–10].

In parallel, non-contact sensing approaches have rapidly advanced with the development of perception and computer vision technologies. Video-based monitoring systems, vision-based pose estimation, multi-object tracking, as well as radar- and ultrasound-based techniques enable remote and non-invasive observation of cattle behavior without attaching devices to animals [11–16]. These methods are particularly suitable for large-scale farm environments and allow the extraction of spatial and spatiotemporal behavior characteristics under complex scenes [11, 12, 14].

Beyond single-modality sensing, multimodal fusion methods have attracted increasing attention in recent studies. By integrating heterogeneous data sources such as wearable sensor signals, visual information, acoustic cues, and location data, multimodal approaches provide more comprehensive and robust representations of cattle behavior [17–20]. Furthermore, behavior modeling and digital twin–oriented frameworks have been explored to support long-term monitoring and system-level behavior perception under incomplete or noisy sensing conditions [18–21].

Along with sensing and modeling advances, machine learning and deep learning techniques have become the dominant tools for cattle behavior analysis. These methods enable effective extraction of temporal, spatial, and spatiotemporal features from high-dimensional data, leading to improved recognition performance in realistic farm environments [3, 6, 12, 17]. Meanwhile, engineering considerations such as sensor deployment, energy consumption, communication infrastructure, and system scalability play an increasingly important role in translating research prototypes into practical livestock monitoring systems [7, 15, 19].

Although a large number of studies have been reported, existing reviews often focus on individual sensing technologies or isolated modeling techniques. A unified and structured overview that systematically summarizes contact-based sensing, non-contact sensing, and multimodal fusion approaches for cattle behavior recognition and modeling, together with their engineering implications, remains limited. To address this gap, this review provides a comprehensive synthesis of recent advances in cattle behavior recognition by organizing existing studies according to sensing modalities, intelligent methods, and engineering applications. Accordingly, this review first surveys contact-based sensing methods, followed by non-contact perception approaches and multimodal fusion strategies, and then discusses engineering system design and future research challenges.

2. CONTACT-BASED SENSING FOR CATTLE BEHAVIOR RECOGNITION

Contact-based sensing is one of the most mature and extensively investigated paradigms for cattle behavior recognition, primarily due to its ability to capture fine-grained motion and activity signals through direct measurement of animal movements [1–4]. By attaching sensors to the animal body, contact-based approaches enable continuous and high-resolution monitoring at the individual level, making them particularly suitable for recognizing behaviors closely related to posture, locomotion, and repetitive motion patterns [2, 5, 6].

Compared with non-contact approaches, wearable sensing systems are less affected by environmental factors such as illumination variation and visual occlusion, and can provide stable time-series measurements under diverse farm conditions [3, 7]. As a result, contact-based methods have been widely adopted for monitoring feeding, rumination, locomotion, estrus, and lameness-related behaviors in both controlled experiments and practical farm deployments [5–10].

2.1. Wearable Sensors and Data Acquisition

Wearable sensing systems for cattle behavior recognition typically rely on inertial, acoustic, and positioning sensors deployed on collars, nose rings, or other animal-mounted devices [2, 3, 4, 10]. Among these, accelerometers are the most commonly used sensors, owing to their low cost, low power consumption, and strong capability in capturing body movement intensity and posture variation [1, 4, 6]. Accelerometer signals have been extensively used as the primary data source for recognizing a wide range of cattle behaviors, including feeding, lying, walking, and grazing activities [4, 6, 8].

Gyroscopes are often integrated with accelerometers to provide complementary rotational motion information, which is beneficial for distinguishing complex locomotion patterns and detecting gait abnormalities associated with lameness [5, 6]. In addition to inertial sensing, acoustic sensors have been employed to capture jaw movement and chewing sounds, enabling effective recognition of feeding and rumination behaviors through sound-based or vibration-based cues [8, 17].

To incorporate spatial context, positioning technologies such as Ultra-Wideband (UWB) and Global Navigation Satellite System (GNSS) sensors have been combined with inertial measurements to provide location-aware behavior recognition [7, 19]. By jointly modeling activity intensity and spatial movement patterns, these systems enhance discrimination of behaviors that exhibit similar motion characteristics but occur in different spatial contexts, such as grazing and walking [7, 19].

From a data acquisition perspective, contact-based sensing systems generate high-frequency time-series data and are often designed for long-term continuous monitoring [6, 10]. Sampling rates are selected by balancing motion resolution and energy efficiency, which is a critical consideration for battery-powered wearable devices deployed in real farms [1, 10]. Furthermore, sensor placement, attachment stability, and animal comfort directly affect signal quality, as improper installation may introduce noise, drift, or motion artifacts that degrade recognition performance [1, 3, 4, 10].

2.2. Feature Representation and Temporal Modeling

Raw signals collected from wearable sensors are usually transformed into discriminative feature representations prior to behavior classification. Early studies mainly relied on handcrafted features extracted from time-domain and frequency-domain analyses, including statistical descriptors, signal magnitude areas, and spectral features, which effectively characterize periodic and repetitive motion patterns [4, 6, 8]. These features are commonly combined with conventional machine learning classifiers to perform behavior recognition tasks [3, 4, 8].

With the increasing availability of large-scale wearable datasets, deep learning–based temporal modeling methods have become the dominant approach for feature representation learning in contact-based cattle behavior recognition [1, 2, 6]. Recurrent neural networks and their variants have been widely adopted to capture temporal dependencies in accelerometer and gyroscope signals, demonstrating improved robustness to inter-animal variability and sensor placement differences [5, 6]. Semi-supervised and representation learning strategies have further been explored to reduce labeling costs and improve generalization under limited annotated data [5].

These approaches enhance representation capability under complex and noisy measurement conditions, and show improved performance for recognizing subtle behaviors with overlapping motion patterns [2, 6].

2.3. Representative Applications and Practical Limitations

Contact-based sensing methods have been extensively applied to a variety of cattle behavior monitoring tasks, including feeding and rumination detection, activity level assessment, estrus identification, and early lameness detection [5–7, 8, 19]. Experimental results reported in existing studies demonstrate that wearable sensing approaches can achieve high recognition accuracy and stable performance under both experimental and real-farm conditions [1, 4, 6].

Despite these advantages, several limitations remain. Wearable devices require physical attachment and periodic maintenance, which may increase labor cost and raise concerns related to animal comfort and long-term usability [3, 4, 10]. Sensor drift, device loss, and limited battery capacity also pose challenges for continuous long-term deployment in large-scale farming environments [1, 6, 10]. These limitations motivate the exploration of complementary non-contact sensing approaches and multimodal fusion strategies, which aim to enhance robustness and scalability and are discussed in the following sections.

3. NON-CONTACT SENSING FOR CATTLE BEHAVIOR RECOGNITION

Non-contact sensing enables remote and non-invasive monitoring of cattle behavior without the need to attach devices to animals, which makes it particularly attractive for large-scale farm deployment and long-term observation [22–24]. Compared with wearable sensing, non-contact approaches can directly capture spatial and spatiotemporal information such as posture, gait, inter-individual interaction, and group dynamics, thereby supporting a richer understanding of cattle behavior under practical farm conditions [22, 28].

3.1. Vision- and Video-Based Perception

Vision-based methods constitute the dominant category of non-contact cattle behavior recognition. Most existing studies adopt camera-based perception pipelines that include cow detection, individual identification, multi-object tracking, and behavior recognition [22–28]. Robust detection and tracking are essential in real farm environments, where frequent occlusion, background clutter, and dense animal distributions pose significant challenges [26–28].

To address these issues, a variety of deep learning–based detection and tracking frameworks have been proposed. Efficient object detectors and tracking algorithms have been designed to improve accuracy while reducing computational complexity, enabling continuous monitoring in crowded barns and outdoor pastures [22, 26, 27]. Multi-object tracking techniques further support the extraction of individual-level temporal trajectories, which are critical for behavior analysis and long-term monitoring [23, 24].

Beyond detection and tracking, pose estimation and keypoint-based representations have been widely explored to capture fine-grained motion characteristics associated with locomotion and lameness [11, 29–31]. By modeling spatial posture configurations together with temporal gait dynamics, these methods achieve higher sensitivity to abnormal motion patterns than coarse activity-level classification [29, 31]. In addition, spatiotemporal deep models, such as 3D convolutional networks and two-stream architectures, have been employed to recognize complex behaviors directly from video sequences, demonstrating improved representation capability for dynamic motion patterns [12, 32].

To enhance robustness across different farm layouts and camera viewpoints, multi-camera systems and cross-view fusion strategies have been investigated. Multi-camera tracking and bird's-eye-view mapping frameworks can alleviate occlusion and improve monitoring stability at the herd level [23, 24]. Furthermore, learning strategies that reduce dependence on large labeled datasets, such as contrastive learning, have been explored for individual identification and behavior recognition under limited data conditions [33].

3.2. Radar-Based and Ultrasound-Based Sensing

In addition to vision-based approaches, radar- and ultrasound-based sensing methods provide complementary non-contact solutions for cattle behavior monitoring, particularly in scenarios where lighting conditions are poor or camera installation is constrained [13, 16]. Frequency-modulated continuous-wave (FMCW) radar has been applied to monitor respiration-related signals in dairy cows, enabling continuous physiological behavior monitoring without physical contact [13]. Such methods are robust to illumination changes and can operate reliably in enclosed or low-visibility environments.

Ultrasound-based techniques have also been investigated for cattle positioning and behavior-related inference. By leveraging ultrasound signal propagation characteristics and machine learning classifiers, these systems can estimate cow position and detect specific behavior patterns under controlled deployment conditions [16]. Although radar and ultrasound methods typically provide lower spatial resolution than vision-based systems, they offer valuable complementary information and enhanced robustness under challenging environmental conditions [13, 16].

3.3. Practical Challenges and Engineering Considerations

Despite significant progress, non-contact sensing in real farm environments remains challenging. Vision-based systems are sensitive to occlusion, illumination variation, and background complexity, which can degrade detection, tracking, and recognition performance [26–28]. Differences in camera viewpoints, farm layouts, and animal appearance further limit model generalization across deployment sites [22, 33].

From an engineering perspective, continuous video processing and multi-camera deployment impose substantial computational and communication burdens on farm monitoring systems. Efficient model design, edge computing strategies, and stable data transmission infrastructures are therefore essential to ensure real-time performance and scalability [23, 24]. These challenges motivate the development of lightweight perception models and the integration of non-contact sensing with wearable sensing through multimodal fusion, which is discussed in the next section.

4. MULTIMODAL FUSION AND BEHAVIOR MODELING

Multimodal fusion aims to integrate heterogeneous sensing sources—such as wearable inertial signals, acoustic cues, location data, and visual observations—to achieve more comprehensive and robust cattle behavior recognition than single-modality approaches [17–21]. This direction is particularly important in practical farms, where any single sensing modality may become unreliable due to device loss, signal drift, occlusion, illumination variation, or infrastructure constraints [3, 12, 22, 24]. By exploiting complementary information across modalities, fusion-based methods can improve recognition accuracy, enhance robustness under real-world disturbances, and support higher-level behavior modeling for long-term monitoring [18–21].

4.1. Fusion Strategies and Feature Alignment

Existing multimodal studies typically adopt either feature-level fusion or decision-level fusion. Feature-level fusion combines representations from different modalities into a unified embedding before classification, enabling the model to learn cross-modal correlations directly [17, 19]. Decision-level fusion aggregates predictions from modality-specific models, which is often simpler to implement and can offer better fault tolerance when one modality becomes degraded [19, 20].

A critical challenge is cross-modal alignment, since different sensors operate at different sampling rates, measurement scales, and noise characteristics. Practical fusion systems therefore need to handle asynchronous signals and heterogeneous feature spaces in a principled manner [17, 19]. In addition, multimodal fusion has been used to enrich behavior context; for example, combining movement patterns with location trajectories enables location-aware behavior interpretation, which is beneficial for behaviors that exhibit similar motion patterns but occur in distinct spatial contexts [7, 19].

4.2. Robustness Under Incomplete or Degraded Modalities

Farm deployments commonly encounter missing or degraded modalities, such as temporary camera blind spots, partial sensor failures, or intermittent connectivity. Multimodal recognition under incomplete modalities therefore becomes a key requirement for stable monitoring systems [18]. Recent studies have explored strategies that explicitly address modality incompleteness by learning robust cross-modal representations, so that behavior recognition remains feasible even when one or more modalities are unavailable [18].

Beyond missing modalities, another practical issue is modality quality variation. Visual signals may degrade under occlusion or illumination changes, while wearable signals may be affected by attachment looseness or drift [3, 12, 22]. Fusion-based designs can mitigate these issues by leveraging redundancy across modalities and adapting the reliance on each modality according to its reliability in the current environment [18, 19]. Such robustness considerations are also closely linked to engineering constraints, since real farms require systems that continue functioning under imperfect sensing and limited maintenance conditions [21, 24].

4.3. Behavior Modeling and Digital Twin–Oriented Frameworks

In addition to recognizing short-term behavior categories, recent works have started to emphasize behavior modeling frameworks that support long-term monitoring and system-level perception [9, 21]. Digital twin–oriented approaches are particularly relevant in this context, as they aim to represent cattle behavior states in a structured manner that can be continuously updated by multimodal observations and used for downstream monitoring tasks [18, 21]. Compared with isolated recognition pipelines, such modeling-oriented frameworks provide a more coherent basis for integrating heterogeneous sensing inputs and for tracking behavior evolution over time [18, 21].

Moreover, multimodal behavior modeling can benefit from explicit mechanisms that cope with incomplete sensing, since long-term monitoring inevitably encounters missing data segments. Approaches that incorporate modality completion or modality mapping concepts provide a practical path toward maintaining stable behavior perception under incomplete modalities [18]. Overall, multimodal fusion and behavior modeling form the natural bridge between “behavior recognition as classification” and “behavior perception as an operational monitoring system” in modern precision livestock farming [18, 21].

5. ENGINEERING SYSTEMS AND PRACTICAL DEPLOYMENT

Engineering system design is the key step that transforms cattle behavior recognition methods from laboratory demonstrations into reliable on-farm monitoring services. Practical deployments must jointly consider sensing hardware, communication networks, edge/cloud computing pipelines, energy constraints, and long-term maintainability to ensure continuous operation under realistic farm conditions [1, 10, 15, 23, 24]. Importantly, these factors directly affect data quality, inference latency, and system scalability, and therefore determine whether behavior recognition outputs are actionable for management and health monitoring [1, 21, 23].

5.1. System Architecture: Sensing–Communication–Computing Pipeline

A typical deployment follows a layered pipeline: (i) sensing and data acquisition at the animal or barn level, (ii) local communication and aggregation, and (iii) computation for real-time or near-real-time inference [1, 10, 15]. For contact-based systems, wearable devices continuously generate high-frequency time-series data and require energy-aware sampling, buffering, and transmission strategies to support long-duration operation [1, 6, 10]. For non-contact systems, camera-based monitoring introduces substantially higher data volume, which intensifies requirements on bandwidth, edge computing capability, and storage management, especially in multi-camera settings [23, 24].

Edge computing is increasingly adopted to reduce transmission load and provide timely feedback. For example, edge-enabled behavior recognition designs have been explored for ruminating behavior recognition, illustrating the practicality of performing inference closer to data sources when continuous monitoring and low latency are required [34]. Similarly, IoT-oriented deployments that combine on-site perception modules with local computing can support individual-level monitoring tasks while reducing dependence on constant cloud connectivity [15].

5.2. Communication and Coverage for Farm-Scale Operation

Communication infrastructure is a frequent bottleneck in real farms, particularly for pasture-based scenarios and large barns with many animals. Low-power network designs and site-use monitoring strategies have been studied to improve coverage and energy efficiency for long-term grazing cattle monitoring, highlighting the need to match networking choices to farm geography and operational constraints [35]. For wearable systems, communication plans must balance update frequency, battery life, and robustness to intermittent connectivity, because unstable transmission can create missing data segments that degrade downstream recognition and modeling reliability [1, 6, 35].

For vision-based deployments, multi-camera monitoring further amplifies network and synchronization requirements. In practice, camera placement, coverage overlap, and data routing must be co-designed with tracking and mapping modules to maintain stable identity association over time and reduce the impact of occlusion [23, 24]. These engineering decisions strongly influence the effectiveness of herd-level monitoring in crowded farm environments [23, 24].

5.3. Reliability, Maintainability, and Long-Term Usability

Long-term usability depends on both hardware robustness and operational maintenance cost. Wearable systems require regular inspection, charging or battery replacement, and occasional re-attachment, which may increase labor cost and introduce data discontinuities if devices are lost or loosened [1, 6, 10]. Non-contact systems reduce the need for animal-mounted devices, but require stable installation, cleaning, and calibration of cameras or other sensors, and may still face performance degradation under complex scenes [22–24].

From a system reliability perspective, practical deployments should explicitly account for sensor failures, data dropouts, and modality degradation. This consideration naturally motivates integrating multiple sensing sources and adopting fusion or modeling strategies that tolerate missing or unreliable inputs [18–21]. Such robustness-oriented system design is particularly important for digital-twin–oriented monitoring, where long-term continuity of behavior perception is required to support farm management and health-related assessment workflows [21].

6. CHALLENGES AND FUTURE DIRECTIONS

Despite rapid progress in cattle behavior recognition, several technical and engineering challenges still limit reliability, scalability, and long-term usability in real farms [1, 21–24]. Addressing these challenges requires coordinated advances in sensing, modeling, data management, and system deployment, rather than isolated improvements to individual recognition models [1, 18–21, 23, 24].

6.1. Data Quality, Annotation Cost, and Evaluation Consistency

A persistent challenge is the variability of data quality across farms, sensor types, and deployment conditions, which can lead to distribution shifts and inconsistent model performance [3, 12, 22–24]. For wearable sensing, signal drift, attachment looseness, and intermittent transmission can introduce noise and missing segments that degrade recognition stability over time [1, 6, 10, 35]. For vision-based monitoring, occlusion, illumination variation, and crowded scenes remain major sources of recognition errors in practical environments [22–24, 26–28].

Another limiting factor is the high cost and inconsistency of behavior annotation, especially for fine-grained behaviors and long-term monitoring. Differences in behavior definitions, labeling protocols, and dataset composition make cross-study comparison difficult and hinder reproducibility [3, 12, 33]. Future work should emphasize consistent evaluation settings and transparent reporting of deployment conditions to enable fair comparison and trustworthy conclusions across farms and sensing modalities [3, 22, 24, 33].

6.2. Generalization Across Animals, Farms, and Devices

Model generalization remains a core barrier to real-world adoption. Wearable-based recognition models are sensitive to inter-animal variability and device placement differences, and cross-device transfer remains challenging under heterogeneous hardware configurations [2, 6]. Vision-based models also face cross-farm generalization issues due to differences in camera viewpoints, barn layouts, backgrounds, and animal appearance, which can reduce robustness when models are deployed outside the training environment [22–24, 33].

Future research should therefore prioritize learning strategies that improve cross-domain robustness, including training protocols that explicitly consider device diversity and cross-scene variation, and model designs that remain stable under realistic farm disturbances [6, 22, 24, 33]. In addition, leveraging multi-camera configurations and fusion-based mapping can reduce viewpoint dependency and partially mitigate occlusion-induced failure modes in crowded environments [23, 24].

6.3. Multimodal Fusion Under Missing and Degraded Modalities

Although multimodal fusion is a promising direction, practical fusion systems must handle missing modalities and quality degradation as first-class design constraints [18, 19, 21]. In realistic farms, wearable signals may degrade due to maintenance issues, while visual signals may become unreliable due to occlusion or lighting changes, and communication limitations can further exacerbate data incompleteness [1, 12, 22–24, 35]. Methods that explicitly address incomplete modalities and learn robust cross-modal representations offer a practical path toward stable behavior perception under imperfect sensing conditions [18, 19].

Future work should focus on reliability-aware fusion strategies that adapt modality weighting according to sensor availability and quality, and on deployment-oriented fusion architectures that remain functional under intermittent connectivity and limited maintenance [18, 19, 21, 35]. Such designs are particularly relevant for long-term monitoring objectives that extend beyond short-term classification and require continuous behavior perception [21].

6.4. From Behavior Classification to Behavior Modeling and Digital Twins

Many existing studies focus on recognizing short-term behavior categories, while practical management tasks often require long-term behavior modeling and continuous state representation [21]. Digital twin–oriented research highlights the need for structured behavior perception frameworks that integrate multimodal observations, support temporal continuity, and provide interpretable behavior states for downstream monitoring and management applications [18, 21]. This direction naturally connects recognition outputs with operational decision support, while also providing a unified basis for incorporating heterogeneous sensing sources over time [18, 21].

Future studies should further clarify the relationship between behavior recognition outputs and modeling objectives, such as how recognition uncertainty, missing data segments, and modality degradation should be represented and propagated in long-term monitoring systems [18, 21, 35].

Establishing system-level validation protocols for digital twin–oriented monitoring will also be essential to demonstrate practical usefulness beyond accuracy metrics reported on short-term datasets [18, 21].

6.5. Deployment Constraints: Energy, Computation, and Connectivity

Finally, deployment constraints remain decisive for practical adoption. Wearable systems must balance sampling rate, inference frequency, and battery lifetime, while ensuring stable communication and manageable maintenance workload [1, 6, 10, 35]. Vision-based systems, especially multi-camera monitoring, introduce substantial data volume and computational burden, making edge computing, efficient model design, and robust data pipelines essential for real-time operation [15, 23, 24]. Edge-enabled recognition, as demonstrated for specific behaviors, illustrates the practicality of performing inference closer to data sources under latency and bandwidth constraints [34].

Future work should continue to integrate algorithm design with system engineering, including energy-aware sensing strategies, lightweight perception models, and deployment-tested communication schemes suitable for diverse farm settings [1, 15, 23, 24, 34, 35]. Such end-to-end considerations are necessary to bridge the gap between research prototypes and reliable on-farm behavior monitoring services [1, 21, 23, 24].

7. CONCLUSION

This review has provided a systematic and structured overview of recent advances in cattle behavior recognition by organizing existing studies according to sensing modalities, intelligent methods, and engineering applications. Contact-based sensing approaches based on wearable devices have demonstrated strong capability in capturing fine-grained motion and activity patterns, enabling reliable recognition of a wide range of behaviors closely related to locomotion, feeding, and health conditions. In parallel, non-contact sensing approaches, particularly vision-based systems, have rapidly advanced and offer scalable, non-invasive solutions for monitoring individual and herd-level behavior under practical farm environments.

Beyond single-modality designs, multimodal fusion and behavior modeling have emerged as a key direction for improving robustness and supporting long-term monitoring. By integrating heterogeneous sensing sources, fusion-based approaches mitigate the limitations of individual modalities and provide richer behavior representations. Furthermore, behavior modeling and digital twin–oriented frameworks establish a natural bridge between short-term behavior recognition and system-level behavior perception, which is essential for continuous monitoring and management-oriented applications.

From an engineering perspective, practical deployment remains a decisive factor for real-world adoption. Issues related to sensing hardware, communication infrastructure, energy efficiency, edge computing, and long-term maintainability strongly influence system reliability and scalability. Addressing these constraints requires end-to-end system design that jointly considers sensing, computation, and deployment conditions, rather than isolated algorithmic improvements.

Overall, cattle behavior recognition is evolving from isolated recognition tasks toward integrated perception systems that combine multimodal sensing, intelligent modeling, and robust engineering design. Continued progress along this trajectory is expected to enhance the practicality, reliability, and impact of behavior monitoring technologies in precision livestock farming, ultimately supporting improved animal health, welfare, and management efficiency.

REFERENCES

[1] Ni G, Jia Y, Shi Z, et al. A Full end-to-end analytical framework for livestock behavior modeling and health assessment using wearable electronic recording system and machine learning [J]. Smart Agricultural Technology, 2026, 13: 101686.

[2] Yue Y, Yu L, Hu Y, et al. Class-Specific feature selection for enhanced cattle behavior recognition using wearable-based system [J]. Smart Agricultural Technology, 2025, 12: 101548.

[3] Essien D, Inyang S, Umoren I. Evaluating machine learning classifiers and explainability for monitoring cow behaviour with wearable nose rings [J]. Preventive Veterinary Medicine, 2025, 244: 106630.

[4] Arablouei R, Currie L, Kusy B, et al. In-situ classification of cattle behavior using accelerometry data [J]. Computers and Electronics in Agriculture, 2021, 183: 106045.

[5] Zhang K, Han S, Wu J, et al. Early lameness detection in dairy cattle based on wearable gait analysis using semi-supervised LSTM-Autoencoder [J]. Computers and Electronics in Agriculture, 2023, 213: 108252.

[6] Arablouei R, Bishop-Hurley G, Bagnall N, et al. Cattle behavior recognition from accelerometer data: Leveraging in-situ cross-device model learning [J]. Computers and Electronics in Agriculture, 2024, 227: 7.

[7] Benaissa S, Tuyttens F, Plets D, et al. Improved cattle behaviour monitoring by combining Ultra-Wideband location and accelerometer data [J]. Animal, 2023, 17(4): 100730.

[8] Chelotti J, Vanrell S, Martinez-Rau L, et al. Using segment-based features of jaw movements to recognise foraging activities in grazing cattle [J]. Biosystems Engineering, 2023, 229: 69-84.

[9] Zhang Y, Zhang Y, Gao M, et al. Digital twin perception and modeling method for feeding behavior of dairy cows [J]. Computers and Electronics in Agriculture, 2023, 214: 108181.

[10] Dutta D, Natta D, Mandal S, et al. MOOnitor: An IoT based multi-sensory intelligent device for cattle activity monitoring [J]. Sensors and Actuators A: Physical, 2022, 333: 113271.

[11] Wu S, Han S, Mo X, et al. A top-down deep neural network for multi-dairy cows pose estimation and lameness detection [J]. Computers and Electronics in Agriculture, 2025, 239: 110911.

[12] Yang J, Jia Q, Han S, et al. An Efficient Multi-Scale Attention two-stream inflated 3D ConvNet network for cattle behavior recognition [J]. Computers and Electronics in Agriculture, 2025, 239: 110911.

[13] Tuna S, Rustia D, Hsu J, et al. Frequency modulated continuous wave radar-based system for monitoring dairy cow respiration rate [J]. Computers and Electronics in Agriculture, 2022, 196: 106913.

[14] Geng H, Hou Z, Liang J, et al. Motion focus global–local network for cow behavior recognition [J]. Computers and Electronics in Agriculture, 2024, 226: 109399.

[15] Chen Y, Rustia D, Huang S, et al. IoT-based system for individual dairy cow feeding behavior monitoring using cow face recognition and edge computing [J]. Internet of Things, 2025, 33: 101674.

[16] Madola V, Vacho L, Cviklovic, et al. Detection of dairy cow position using ultrasound technique with support vector machine, Computers and Electronics in Agriculture, 2025, 235: 110385.

[17] Ferrero M, Chelotti J O, Martinez-Rau L S, et al. A multi-head deep fusion model for recognition of cattle foraging events using sound and movement signals [J]. Engineering Applications of Artificial Intelligence, 2025, 157: 111372.

[18] Zhang Y, Zhang Y, Gao M, et al. Multimodal behavior recognition for dairy cow digital twin construction under incomplete modalities: A modality mapping completion network approach [J]. Artificial Intelligence in Agriculture, 2025, 15(3): 459–469.

[19] Arablouei R, Wang Z, Bishop-Hurley G J, et al. Multimodal sensor data fusion for in-situ classification of animal behavior using accelerometry and GNSS data [J]. Smart Agricultural Technology, 2023, 4: 100163.

[20] Wang J, Zhao Y, Li X, et al. Multimodal fusion-based detection method of estrus cows using multisource data inspired by hidden Markov model algorithms [J]. Computers and Electronics in Agriculture, 2025, 235: 110391.

[21] Zhang Y, Zhang Y, Jiang H, et al. New method for modeling digital twin behavior perception of cows: Cow daily behavior recognition based on multimodal data [J]. Computers and Electronics in Agriculture, 2024, 226: 109426.

[22] Araújo V M, Rili I, Gisiger T, et al. AI-powered cow detection in complex farm environments [J]. Smart Agricultural Technology, 2025: 100770.

[23] Abbas K, Afzal Z, Raza A, et al. Vision transformer-based multi-camera multi-object tracking framework for dairy cow monitoring [J]. Smart Agricultural Technology, 2025: 101525.

[24] Nasir M F, Fuentes A, Han S, et al. Multi-camera fusion and bird-eye view location mapping for deep learning-based cattle behavior monitoring [J]. Artificial Intelligence in Agriculture, 2025.

[25] Giannone C, Sahraeibelverdy M, Lamanna M, et al. Automated dairy cow identification and feeding behaviour analysis using a computer vision model based on YOLOv8 [J]. Smart Agricultural Technology, 2025: 101304.

[26] Zheng Z, Li J, Qin L. YOLO-BYTE: An efficient multi-object tracking algorithm for automatic monitoring of dairy cows [J]. Computers and Electronics in Agriculture, 2023, 209: 107857.

[27] Zheng Z, Qin L. PrunedYOLO-Tracker: An efficient multi-cows basic behavior recognition and tracking technique [J]. Computers and Electronics in Agriculture, 2023, 213: 108172.

[28] Han S, Fuentes A, Yoon S, et al. Deep learning-based multi-cattle tracking in crowded livestock farming using video [J]. Computers and Electronics in Agriculture, 2023, 212: 108044.

[29] Li Q, Si Y, Chu M, et al. A novel lameness detection method for dairy cows based on temporal gait and spatial post features [J]. Computers and Electronics in Agriculture, 2024, 224: 109172.

[30] Wu S, Han S, Zhang J, et al. Multi-scale keypoints detection and motion features extraction in dairy cows using ResNet101-ASPP network [J]. Journal of Integrative Agriculture, 2024.

[31] Zhao K, Zhang M, Ji J, et al. Automatic lameness scoring of dairy cows based on the analysis of head-and back-hoof linkage features using machine learning methods [J]. Biosystems Engineering, 2023, 230: 424–441.

[32] Ma S, Zhang Q, Li T, et al. Basic motion behavior recognition of single dairy cow based on improved Rexnet 3D network [J]. Computers and Electronics in Agriculture, 2022, 194: 106772.

[33] Hooker J M, de Medeiros B B, Saha C, et al. Multitask contrastive learning for individual dairy cow recognition across different behavior classes based on small image sets [J]. Journal of Dairy Science, 2025.

[34] Shen W, Sun Y, Zhang Y, et al. Automatic recognition method of cow ruminating behaviour based on edge computing [J]. Computers and Electronics in Agriculture, 2021, 191: 106495.

[35] Mancuso D, Castagnolo G, Parlato M C M, et al. Low-power networks and GIS analyses for monitoring the site use of grazing cattle [J]. Computers and Electronics in Agriculture, 2023, 210: 107897.