Computers and Electronics in Agriculture 238 (2025) 110804

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Robust piglet nursing behavior monitoring through multi-modal fusion of computer vision and ambient floor vibration

Yiwen Dong [a,*,1], Zihao Song [a,1], Jesse R. Codling [b], Gary Rohrer [c], Jeremy Miles [c], Sudhendu Sharma [d], Tami Brown-Brandl [d], Pei Zhang [b], Hae Young Noh [a]

[a] Department of Civil and Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, 94305, CA, USA
[b] Electrical Engineering and Computer Science, University of Michigan, 2260 Hayward Street, Ann Arbor, 48109, MI, USA
[c] U.S. Meat Animal Research Center, Clay Center, U.S. Department of Agriculture, 844 Rd 313, Clay Center, 68933, NE, USA
[d] Biological Systems Engineering, University of Nebraska-Lincoln, 200 L.W. Chase Hall, Lincoln, 68583, NE, USA

ARTICLE INFO

Keywords:
Nursing
Multi-modal
Computer vision
Structural vibration
Precision livestock farming

ABSTRACT

Nursing is a critical activity during the lactation period of swine farming. Continuous monitoring of piglet nursing behavior during the lactation period is essential to informing animal caretakers about the health status of piglets to reduce the mortality rate, maximize lactational growth, and improve animal welfare. Traditional approaches rely on manual observation and wearable devices, which are labor-intensive and can cause discomfort to the animals. Recent advancement in computer vision and ambient vibration sensing enables non-contact piglet nursing monitoring: The computer vision approach captures piglet location but has limited observation of their detailed movement due to lighting, resolution, and visual obstruction constraints; the ambient vibration sensing approach captures piglet movement patterns but has limited location information. In this study, a novel approach to integrate these two complementary sensing modalities is developed for robust piglet nursing behavior monitoring during the lactation period. This study leverages the state-of-the-art Segment Anything Model (SAM) to first convert images into sparse representations of piglet behaviors and then combine with ambient vibration to collaboratively infer piglet nursing pattern and intensity. This new approach enables piglet nursing monitoring with much lower computing and storage requirements than conventional computer vision methods, making it more practical for farm settings. Real-world experiments were conducted at a pig farm for continuous vision and vibration monitoring of 8 pens over 3 farrowing cycles. This study has a 97% accuracy in classifying 5 nursing stages, representing a significant 3× and 3.8× error reduction compared to the baseline method using only vision or vibration data, respectively. The multi-modal fusion approach leads to an efficient, robust, and accurate piglet nursing model that can immediately inform caretakers of issues that arise during this crucial time point of a piglet's life.

1. Introduction

The survival and well-being of piglets during the lactation period are critical factors that influence the overall success of modern swine farming (Baxter and Edwards, 2021; Lay et al., 2002). In particular, continuous monitoring of piglet nursing behavior plays an essential role in reducing mortality rates and improving animal welfare (Sadeghi et al., 2023; Kumar et al., 2025). First, nursing behavior monitoring ensures that piglets are receiving the necessary nutrients from the mother sow's milk (Blavi et al., 2021; Hojgaard et al., 2020; Valros et al., 2002), improving the overall breeding success and profitability of the farm. In addition, changes in nursing behavior can be an early indicator of health issues in both the sow and piglets such as mastitis, metritis, and agalactia (Martin et al., 1978), which can be timely detected by continuous monitoring to allow prompt intervention and disease management (Manteuffel et al., 2017), preventing the spread of illnesses within the herd. To this end, automated nursing behavior monitoring provides a resource-efficient solution for precision swine farming (Sadeghi et al., 2023), assisting farm staff in decision-making.

Traditional piglet nursing monitoring has primarily relied on labor-intensive methods such as direct observation and wearable devices (Sadeghi et al., 2023; Jensen et al., 1991; Pan et al., 2023). Direct observation is not only resource-intensive but also limited to non-continuous monitoring, and the presence of human observers can affect the natural behavior of the animals. On the other hand, wearable

* Corresponding author.
E-mail address: ywdong@stanford.edu (Y. Dong).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.compag.2025.110804
Received 24 September 2024; Received in revised form 20 June 2025; Accepted 24 July 2025
Available online 11 August 2025

Fig. 1. This figure illustrates a scenario where both nursing and sleeping behaviors of the piglets occurred in the same location near the sow's belly. In such cases, the vision approach has relatively low accuracy.

sensing requires all piglets to carry devices, which can cause discomfort to the animals and may not provide a holistic view of the complex nature of piglet nursing behaviors. Recent advancements in computer vision and ambient vibration sensing have created new possibilities for non-invasive and continuous piglet nursing monitoring (Yang et al., 2019; Li et al., 2023; Gan et al., 2021, 2023; Dong et al., 2023). Recent development in computer vision has enabled contactless monitoring of piglet locations and movements, but brings in new challenges of visual occlusion, especially when the piglets are packed and overlapping during nursing. Specifically, piglet activities such as nursing and sleeping near the sow's belly can happen at the same location and look similar from the videos, lowering the accuracy of nursing monitoring (see Fig. 1). Recent studies developed ambient floor vibration sensing in capturing piglet movements through time- and frequency-domain signal patterns, which shows promise in low-cost, low-maintenance pig activity monitoring (e.g., recognizing nursing stages) (Bonde et al., 2021; Dong et al., 2023). However, it has limitations in localizing individual piglets within a large group.

This paper introduces a novel multi-model approach for piglet nursing behavior monitoring through the fusion of computer vision and ambient floor vibration sensing, which overcomes the limitations in single sensors and to provide a robust prediction of piglet nursing stages. The main idea was to leverage the advantages of each sensing modality to compensate for their inherent limitations. A top-down surveillance camera and floor-mounted vibration sensors are combined to accurately capture both the location of individual piglets and the movements of occluded piglets. To achieve optimal performance, computer vision features were automatically extracted from the unsupervised Segment Anything Model (SAM) (Kirillov et al., 2023) and integrated with the vibration features that were effective in prior work (Dong et al., 2023). The integration of these features leads to non-intrusive and more reliable nursing monitoring during the lactation period with much lower computing and storage requirements than conventional computer vision methods, making it more practical for farm settings.

The primary research challenge lies in the mixture of nursing and non-nursing activities among a group of piglets, which were not clearly delineated in the existing studies. For example, some piglets may fall asleep during nursing and some may walk around the pen to explore the area, especially during the transition periods from sleeping to nursing (Puppe et al., 2008). These heterogeneous activities lead to a large number of corner cases, resulting in unsatisfactory performance in nursing stage prediction when applying the state-of-the-art computer vision or vibration signal processing models directly. To overcome this challenge, domain knowledge of piglet nursing patterns was incorporated to decompose the heterogeneous piglet behaviors. Instead of a binary classification of nursing, three additional transition stages were identified through the collaboration with animal scientists. A new computer vision algorithm was developed to identify potential nursing periods based on the piglet's postures and locations, and then verified by the vibration data through the intensity of their movements.

Therefore, the core objectives of this paper include:

- Developing a robust pig nursing monitoring system through a complementary multi-modal fusion of computer vision and ambient floor vibration sensing;
- Characterizing the heterogeneous nursing behavior among a group of piglets by integrating domain knowledge of the nursing stages, locations, and movement patterns in farrowing pens to enhance the performance of video segmentation and feature extraction;
- Evaluating the multi-modal monitoring system through real-world deployments with 12 lactating pigs, demonstrated its effectiveness under varying environmental and behavioral settings.

2. Materials

The materials used in this study involves sow farrowing management of eight pens over three farrowing cycles (24 sows and around 220 piglets), the collection of video and vibration data, and the labeling effort based on video and vibration data observation.

2.1. Sow farrowing management

The study was conducted on a research sow farm at the USDA-ARS U.S. Meat Animal Research Center (USMARC) located at Clay Center, Nebraska, U.S.A. The animals used for this study were from the USMARC commercial crossbred population consisting primarily of Yorkshire and Landrace genetics. This study utilized the standard breeding protocols for females, gestation ration, and lactation ration at the USMARC, which have been reported previously (Rempel et al., 2022).

The farrowing room consisted of two rows of 10 pens separated by a 1.2 m aisle. Each pen measured 1.8 m × 2.5 m and included a confinement area for the sow measuring 0.6 m × 2.1 m (Condous et al., 2016), as shown in Fig. 2. The piglets were allowed to roam freely around the entire pen. The lactation period in this facility typically refers to the first 30 days after farrowing, which was the monitoring period of this study. In this deployment, three types of pen layouts were used: (1) offset, (2) diagonal, and (3) centered. These layout types represent variations in sow and piglet positioning relative to the geophone sensors and will be discussed in more detail in Section 6.3. The flooring was constructed with metal slats 8.5 mm wide and gaps of 9.5 mm (TriDEK, Hog Slat, Inc.; Newton Grove, NC, USA).

The climate-controlled room began at 25 °C and was gradually lowered to 20 °C as the piglets grew older. Evaporative cooling systems were employed to reduce the room temperature, while forced-air heaters and 10 cm radiant heaters were used to increase the air temperature when needed. Additionally, each pen had a 175 W ceramic infrared lamp installed over a rubber mat (1.1 m × 0.3 m), controlled by a thermostat to maintain a nominal ambient temperature of 25 °C for the piglets.

For this study, eight farrowing crates from one row of pens in the farrowing room were utilized.

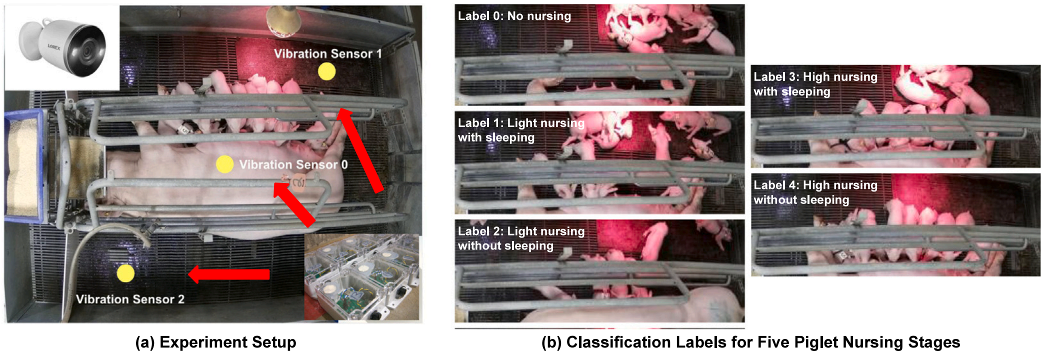

Fig. 2. Experiment setup and nursing stage labels: (a) Experiment setup for a sample pig pen with video and vibration data acquisition system. The video camera provides a top-down view of the entire pen. The three vibration sensors are installed underneath the floor to record pig-induced floor vibrations. (b) Classification labels for piglet nursing behavior are defined using the image frames from the videos. The labels are defined based on (1) the number of nursing piglets and (2) the number of sleeping piglets into five stages — no nursing (label 0), light nursing with sleeping (label 1), light nursing without sleeping (label 2), high nursing with sleeping (label 3), and high nursing without sleeping (label 4).

2.2. Video and vibration data acquisition

To collect the video data, an 8-channel Network Video Recording (NVR) camera system (Lorex N842A81) was installed on top of each farrowing pen approximately 2.6 m above the floor. The cameras continuously recorded video at a resolution of 1920×1080 pixels and a frame rate of 10 frames per second. The recorded footage was stored on a 10TB hard drive. The video data size amounted to approximately 880 megabytes per hour for each camera, equating to around 7 GB per hour collectively for all eight cameras. In addition, each video frame was overlaid with the corresponding crate number and the time, displayed in the upper corners of the image (see Fig. 2a).

The vibration data was collected from the three geophone sensor nodes attached underneath the pen floor, as described in Fig. 2. Each sensor node consists of an SM-24 geophone integrated with a self-developed PCB, along with a data transmission and storage unit. The sensor system features an extended spurious frequency response above 240 Hz, enabling full-bandwidth signal acquisition at a 2-ms sampling interval. The sensor layout consists of one sensor underneath the sow (in the middle of the pen), one sensor under the heat lamp where the piglets tend to sleep, and another sensor at the other side of the pen to cover the entire area. Sensors were mounted at three locations, including (1) center, (2) under the heat lamp, and (3) under the drinking nozzle, as shown in Fig. 2. All geophone sensors were concealed in plastic boxes to prevent water leakage during cleaning and excretion from the pigs and were fixed with multiple zip-ties to ensure firm coupling with the crate, as described in the prior work (Dong et al., 2023). The sampling rate was set as 500 Hz with a 2–10× amplification rate to achieve optimal signal-to-noise ratio without clipping the data. All data collection was performed in accordance with federal and institutional regulations regarding proper animal care practices and was approved by the U.S. Meat Animal Research Center's Institutional Animal Care and Use Committee as $EO\#171.0$.

2.3. Piglet nursing stage labels

The collected data were labeled based on variations in piglet nursing behaviors using video data. First, the overall piglet behavior was observed from the videos in order to develop meaningful labels that are representative of various nursing stages. We observed that piglets' nursing behaviors are usually mixed with sleeping behavior through observation. To overcome the mixed nursing and sleeping behavior challenge, combined behavior characterization was leveraged from a prior work (Kim et al., 2025), aiming to capture the complex piglet behavior during various nursing stages that involve sleeping. In addition, we found that piglet location and their behaviors were correlated — piglets typically nurse near the sow's belly and sleep near the heat lamp with their feet visible from the camera or near the sow's belly when stacked with the other piglets. In cases when distinguishing sleeping and nursing behavior was difficult, the vibration data was used as a reference to observe the signal intensity. Typically, piglet sleeping behaviors occur at the beginning and ending phases of nursing.

With a basic understanding of the piglet nursing behaviors and suggestions from the animal scientists, the labels were assigned to each data frame. Based on the nursing percentile (i.e., the percentage of nursing piglets among all piglets), discrete labels were defined to quantify the mixed nursing and sleeping behavior based on the observations. Specifically, nursing stages were divided into light nursing (piglets nursing percentile less than 50%) and high nursing (piglets nursing percentile greater or equal to 50%) and then combined with the behavior of no sleeping (piglets sleeping percentile equal to 0%) and sleeping (piglets sleeping percentile greater than 0%) to the label. With the nursing and sleeping combinations, the piglet nursing behaviors were defined into five stages — no nursing (label 0), light nursing with sleeping (label 1), light nursing without sleeping (label 2), high nursing with sleeping (label 3), and high nursing without sleeping (label 4). These five types of labels are collectively exhaustive (include all types of piglet nursing patterns that are commonly observed in the pen) and are mutually exclusive (no overlaps among various stages). A summary of each type of label and its representative images is presented in Fig. 2b.

3. Methods

In this section, our multi-modal pig nursing monitoring system framework (see Fig. 3) is introduced, with an emphasis on the cross-modality correction and multi-modal fusion as our key methodology innovation. The framework consists of 4 modules: (1) vibration data processing and feature extraction, (2) video data processing and feature extraction, (3) potential nursing period identification, and (4) multi-modal nursing stage classification. First, the vibration data were collected from three sensors implemented under the crate floor, and the video data from a camera mounted on the ceiling facing downwards. To extract features from the acquired vibration and video data, the vibration data were processed through sliding windows (Module 1), and the image data were processed through zone division and segmentation (Module 2). Then, vision features were used to identify the potential nursing period through active feature sampling at the nursing locations in Module 3. Lastly, piglet nursing stages were classified through multi-modal fusion and cross-modality correction in Module 4.

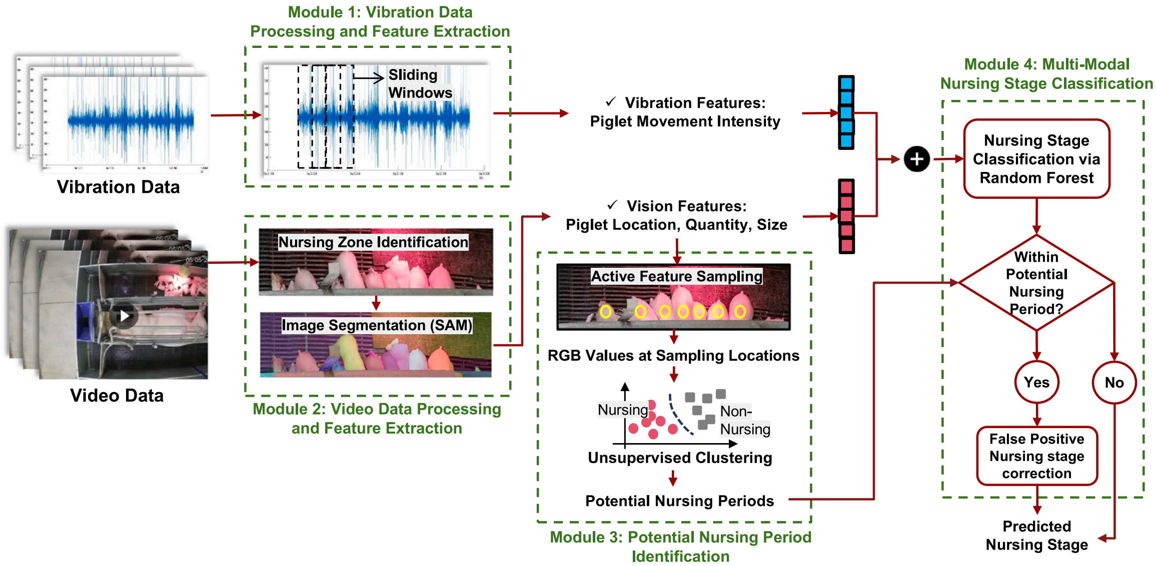

Fig. 3. Overview of our multi-modal piglet nursing behavior monitoring framework, which consists of four modules: (1) vibration data processing and feature extraction, (2) video data processing and feature extraction, (3) potential nursing period identification, and (4) multi-modal nursing stage classification.

3.1. Vibration data processing for piglet movement feature extraction

Vibration feature extraction plays a fundamental role in capturing subtle piglet movement patterns relevant to nursing behavior classification. Features representing the piglet movements during nursing and non-nursing periods were extracted by processing the vibration data, based on existing studies' findings, such that movement information can be reflected by ambient vibration patterns (Bonde et al., 2021; Dong et al., 2023). To pre-process the raw data for noise reduction, a low-pass filter was applied, and the signals were then segmented on a per-minute basis to match the time stamp with the vision features. To isolate the piglet-induced vibration from the sow-induced vibration, we excluded the frames when the vibration data exceeds 90 percentile among all signal amplitudes because those vibrations are typically induced by the sow's posture change or eating and not relevant to nursing, characterized by high signal amplitudes as suggested by prior work (Bonde et al., 2021). This 90% threshold choice was informed by clear peaks in vibration intensity observed during specific periods, such as May 05, 5:00–6:00 a.m. and 6:30–7:00 p.m. in Fig. 4, which coincide with the farm's daily food refill schedule. These events cause sharp spikes in vibration magnitudes due to sow excitation and unrelated piglet movement. Applying the 90% cutoff allowed us to mitigate the influence of these rare but disruptive peaks while preserving the majority of the signal variation that reflects typical piglet behavior patterns.

3.1.1. Characterizing nursing-induced floor vibrations

To understand how vibration data assist in nursing classification, data were first characterized to extract representative features that distinguish nursing from non-nursing behaviors. Nursing-induced floor vibration data were characterized by comparing the signals between nursing and non-nursing periods. Fig. 4 shows the sample vibration data with respect to the nursing behavior over one day. During the nursing period, piglets typically have less movement and also lie down on the ground, which induces less vibration. Therefore, the nursing period identified in Fig. 4 has lower signal magnitudes compared with the non-nursing period. In contrast, the vibration amplitudes during the non-nursing period are higher because the piglets may be walking around and may exhibit other movements that are more intense than nursing.

3.1.2. Vibration feature extraction for nursing

Features were extracted from the time domain and frequency domain of the floor vibration signals to represent collective piglet behaviors in a pen. These features included the maximum, minimum, median, and variance values of signal magnitudes within each 60-s window and each 5-Hz frequency range. The window size was chosen in order to align with the temporal resolution of the per-minute video frames to keep consistency between the modalities, as introduced in Section 3.2. The frequency resolution was selected based on the findings from prior work (Bonde et al., 2021), in which a 5-Hz interval was sufficient to reflect different types of piglet movement while effectively reducing the feature dimensions. The statistical values, including the maximum, minimum, median, and variance of signal magnitudes, were chosen for their ability to capture the range, intensity, and variability of the floor vibration signals, which are critical for accurately representing the collective behaviors of piglets in the pen. The time-domain signal features can represent how intense the piglets' behaviors are, as piglets' different behaviors will induce different amounts of energy. The mathematical formulas for time-domain features were shown in Eqs. (1) to (4). The frequency-domain features can represent what type of forces the piglets exert on the floor. For example, the 25-Hz frequency corresponds to the fundamental frequency of the floor structure. Frequencies in the range of 200-Hz to 250-Hz capture the dynamic behaviors of piglets, which involve quick movements, while frequencies in the range of 50-Hz to 100-Hz typically reflect static postures such as lying down. The mathematical formulas for frequency-domain features were shown in Eqs. (5) to (8). Prior work (Dong et al., 2023) has demonstrated the effectiveness of these features in classifying nursing stages.

$$\text{Maximum: } x_{\max} = \max(x_i), \quad i = 1, 2, \dots, N \tag{1}$$

$$\text{Minimum: } x_{\min} = \min(x_i), \quad i = 1, 2, \dots, N \tag{2}$$

$$\text{Median: } x_{\text{median}} = \text{median}(x_i), \quad i = 1, 2, \dots, N \tag{3}$$

$$\text{Variance: } x_{\text{var}} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2, \quad \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{4}$$

$$\text{Maximum: } X_{\max}(f) = \max(|X(f_k)|), \quad f_k \in [f, f+5\text{ Hz}] \tag{5}$$

$$\text{Minimum: } X_{\min}(f) = \min(|X(f_k)|), \quad f_k \in [f, f+5\text{ Hz}] \tag{6}$$

$$\text{Median: } X_{\text{median}}(f) = \text{median}(|X(f_k)|), \quad f_k \in [f, f+5\text{ Hz}] \tag{7}$$

$$\text{Variance: } X_{\text{var}}(f) = \frac{1}{M-1}\sum_{k=1}^{M}(|X(f_k)| - \bar{X}(f))^2,$$

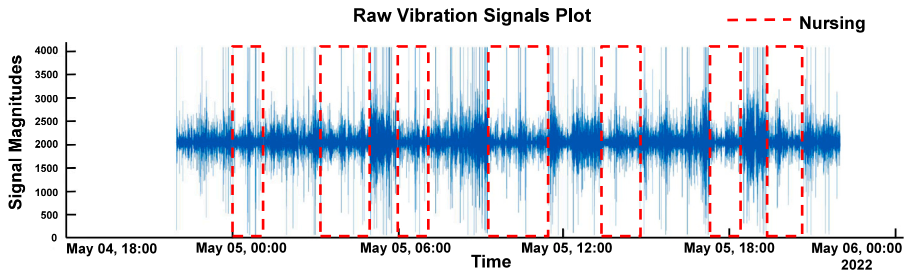

Fig. 4. Sample raw vibration data plot over one day. The signals within red boxes represent the nursing period, while the other period of signals shows the non-nursing period.

$$\bar{X}(f) = \frac{1}{M}\sum_{k=1}^{M} |X(f_k)| \tag{8}$$

where:

- x_i represents the ith time-domain sample within each 60-s window, and N is the number of samples in this window.
- $X(f_k)$ represents the frequency-domain magnitude at frequency f_k, and M is the number of frequency points within the 5-Hz frequency interval.

3.2. Video data processing for piglet nursing information extraction

The video data were processed to extract piglet nursing information such as location, quantity, and size through a data processing pipeline leveraging the state-of-the-art Segment Anything Model (SAM) (Kirillov et al., 2023). First, the representative image frames were captured per minute from the video based on the empirical observation that the nursing pattern had a gradual change that typically spans several minutes. This method also reduces the redundancy and dimensionality in data analysis while maintaining the effective information of piglet nursing. Then, features from each image frame produced by SAM to represent individual piglet behavior were extracted for future classification. The SAM was developed by Meta Research for image segmentation and has been trained with the largest segmentation dataset so far. SAM is promotable to perform segmentation with input points, bounding boxes, and so on, and it enables zero-shot generation of unfamiliar objects without additional training. By sampling the features of individual segments generated by SAM near the sow's belly, this active feature sampling method can differentiate potential nursing and non-nursing periods.

3.2.1. Nursing location characterization and nursing zone identification

The nursing locations were characterized, and the nursing zones were identified by observing the video data, in order to focus on the analysis of nursing-related behaviors only. First, based on empirical observation, the sow is limited to a fixed position due to the confinement in which she can only turn her belly to either the upper or lower side to allow nursing. As a result, the piglets will only nurse in two fixed areas, either in the upper or lower area near the sow's belly. Therefore, these two rectangular nursing areas were cropped from each image frame as the vision dataset to capture nursing-related behaviors only (see Fig. 5). Since this study focuses on nursing behavior monitoring, these two rectangular areas were defined as piglet "nursing zones" to exclude other unrelated objects, such as the sow itself, trough, steel pipes, etc. Compared with processing the complete image frame, focusing on the "nursing zones" not only reduces environmental noises but also improves the data processing efficiency.

3.2.2. Video frame segmentation and feature extraction

Image segmentation was performed on video frames using SAM to extract parse features to represent the quantity, size, and position information of piglets for future multi-modal fusion and nursing stages classification. First, SAM was leveraged to produce segments within the nursing zones that were identified from the prior processing. SAM then outputs a dictionary containing detailed properties of all segments, including segment boundaries, areas, and other relevant information about segments. Then, features to represent the characteristics of individual piglets were extracted, summarized as follows:

- **Sum and Median of Segment Sizes**: For each frame, let $A = \{a_i\}_{i=1}^{N}$ denote the set of areas for N segments identified by SAM. The sum and median of segment sizes are defined as:

$$A_{\text{sum}} = \sum_{i=1}^{N} a_i, \quad A_{\text{median}} = \text{median}(a_1, a_2, \ldots, a_N)$$

- **Number of Segments in a Nursing Zone**: Let N denote the total number of segments detected by SAM within each nursing zone of a given frame:

$$N_{\text{segments}} = |A|$$

- **Centroids of the Segments**: For each segment S_i, the centroid C_i is calculated based on pixel coordinates:

$$C_i = (x_i, y_i) = \left(\frac{1}{|S_i|} \sum_{(x,y)\in S_i} x, \ \frac{1}{|S_i|} \sum_{(x,y)\in S_i} y \right)$$

where (x, y) denotes pixel coordinates belonging to segment S_i.

3.3. Potential nursing period identification through active feature sampling

In face of the challenge due to mixed nursing and sleeping behaviors, identifying potential nursing periods is a crucial component of the proposed methodology, where an active feature sampling strategy was developed to filter out the irrelevant frames and narrow down to the potential nursing periods. To meet practical requirements on minimal labeling effort, this strategy was designed to be fully unsupervised, and thus no manual labeling was involved. The difference in color between piglets and crate ground was leveraged to identify potential nursing and non-nursing periods. Based on the physical intuition that the piglets are often pink in color while the crate color is often dark gray, the mean RGB values of the nursing zone segments were extracted to visualize the contrast between nursing and non-nursing periods. As shown in Fig. 6a, the high Red and Blue values in the nursing zone image align well with the nursing period, indicating that the color difference between piglets and other objects helps to identify the nursing piglets.

Based on the aforementioned intuition, this study develops an unsupervised active feature sampling strategy to determine nursing periods. According to the nursing zone defined in Section 3.2.1, features were sampled in the two nursing zones identified next to the sow's belly. For each video frame, the segment with the highest "quality score" generated by SAM was picked, and its RGB values were extracted. The "quality score" quantifies the robustness of the segmentation by computing the average Intersection-over-Union (IoU) between the mask obtained at the default threshold and masks generated at slightly perturbed thresholds. Specifically, if M denotes the original mask obtained at the default threshold (e.g., 0.0) and $\{M_i'\}_{i=1}^{N}$ denotes masks

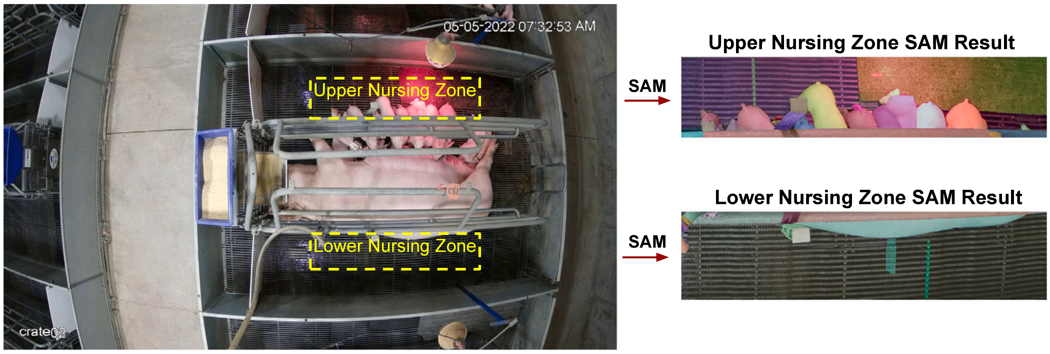

Fig. 5. A sample of the nursing zones and the image segmentation results. This figure shows the upper and lower nursing zones near the sow's belly and their corresponding SAM results.

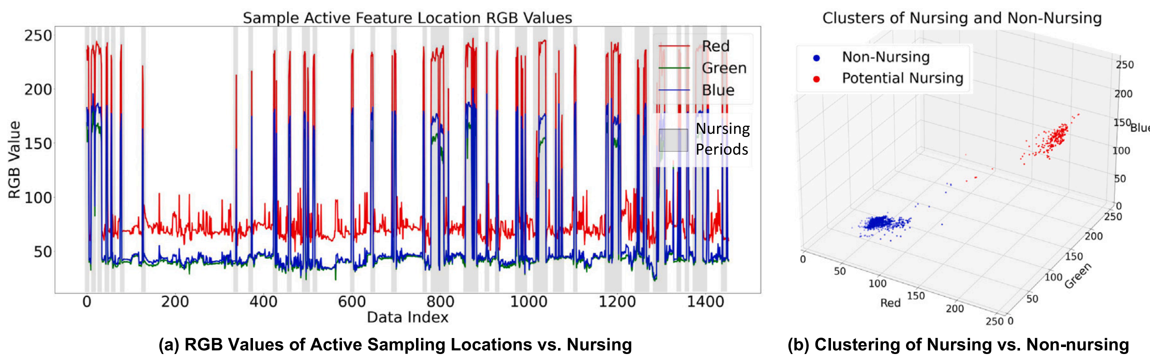

Fig. 6. Active feature sampling of RGB values for potential nursing period identification. (a) shows the association between high RGB values and nursing periods over one day. (b) shows a clear separation between nursing and non-nursing periods: the red cluster represents the nursing piglets with higher RGB values, and the blue cluster represents the crate background with lower RGB values.

generated at N perturbed thresholds, the quality score Q is computed as:

$$Q = \frac{1}{N}\sum_{i=1}^{N}\frac{|M \cap M_i'|}{|M \cup M_i'|}.$$

The "quality score" presents the model's confidence in producing segments, where lower scores indicate that unrelated objects may exist around the target object or only part of the target object was segmented at the input active feature location. The RGB values in this selected segment represented the presence of piglets in the nursing zone. Then, K-Means clustering was performed to separate the RGB values into two clusters for all frames. The choice of two clusters ($k = 2$) was based on the expectation that the frame contents would naturally form two visually distinct groups: frames showing only the crate background and frames containing piglets. Empirical observations confirmed this hypothesis, as piglets contrasted clearly against the crate background. The results in Fig. 6b showed that the frames with crate background only and frames with piglets were separated into two different clusters. Since piglets are typically pink in color, their segments have more red colors than the crate background, the cluster with a higher value in the red color component was picked as the potential nursing period. In the other cluster, the RGB values were approximately (50, 50, 50), matching the dark gray crate background color. This implies there were no piglets in the nursing zone, thus clustered as non-nursing periods. To this end, the frame indices of the nursing and non-nursing periods were saved for subsequent multi-modal fusion. The main benefit of the active feature sampling strategy is that it can identify nursing periods without the need for labeled training data.

Moreover, the active feature sampling method is also effective in reducing background noises that were originally misidentified as piglets by SAM. While SAM produces segments mainly based on the object contour, this active sampling method leveraged domain knowledge of pig and crate colors to further rule out the segments that are not piglets. As shown in Fig. 7, the green crosses representing the centroids of the noise segments are completely removed after this proposed active feature sampling method. This allows us to focus on the piglet segments for further analysis.

3.4. Multi-modal nursing stage classification

Through the final module of nursing stage classification, multi-modal fusion and cross-modality correction were performed to classify piglet nursing stages accurately. Results from the active feature sampling strategy (introduced in the previous module) were leveraged to improve prediction accuracy without requiring additional labeled data. The extracted vibration and vision features from prior modules were fused to represent piglet nursing behaviors from various aspects (e.g., location, activity type, intensity), aiming to improve the robustness and accuracy of nursing stage classification. More specifically, the vision and vibration features were concatenated into a single comprehensive feature vector. This fused feature set was then used as the input to the final classifier for nursing stage prediction. The vision and vibration features are highly complementary because the vibration features contain the piglet movement intensity and posture information, while the vision features provide additional piglet location and quantity information.

First, the vision and vibration features were normalized through `MinMaxScaler` of Scikit-Learn (Pedregosa et al., 2011) to ensure that one set of features does not dominate the other due to scale differences. This scaler transforms all features to a [0, 1] range prior to model training. Then, the normalized vision and vibration features were fused into a single feature vector by concatenating them as the designed features for the final prediction. The final dataset with the designed features was split into 70% for training and 30% for testing. The random forest model was chosen for nursing stage prediction because it is

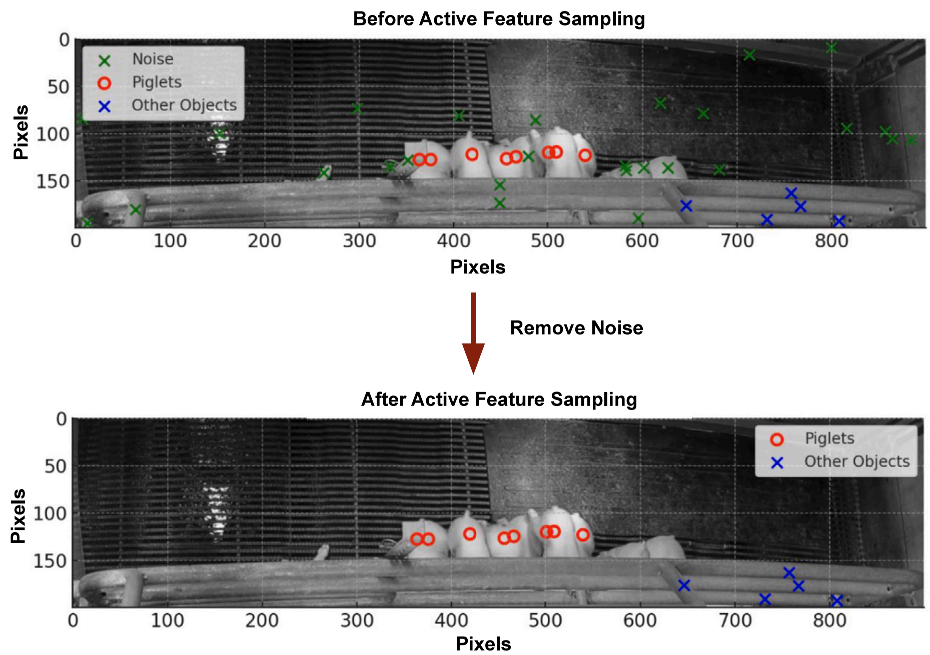

Fig. 7. This figure shows the effectiveness of the proposed method in reducing background noises produced by SAM with the pixel coordinates overlaid for precise reference. The noises are removed after feature sampling based on the unique locations and colors of the piglets that contrast with the crate background.

robust in dealing with multi-class classification tasks in many previous literature (Dong et al., 2023; Bonde et al., 2021), which aligns with the objective of five-stage nursing classification. The model consists of 150 trees in the forest, with all other parameters set to their default values as defined by the `RandomForestClassifier()` of Scikit-Learn (Pedregosa et al., 2011). Next, the potential nursing frame indices generated from Module 3 (potential nursing period identification) were used to correct the nursing stage prediction by checking if the potential nursing frame index list contains the frames through prediction. If it is within the potential nursing list and has been predicted to be non-nursing (label 0), then the largest probability of stage prediction other than non-nursing (label) is picked for this frame. The output of the method is the nursing stage (label 0–4 as defined in Fig. 2b). If it is not included in the potential nursing list, the current stage prediction will be retained as the final prediction. After correcting the nursing period information, the predicted nursing stage is produced as the output of our method. The performance of the model is assessed using prediction accuracy and the F1 score, as defined in Eqs. (9) and (10).

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{All Predictions}} \tag{9}$$

$$\begin{aligned}\text{F1 Score} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \\ &= \frac{2 \times (\text{True Positives})}{2 \times (\text{True Positives}) + (\text{False Positives}) + (\text{False Negatives})}\end{aligned} \tag{10}$$

4. Results

4.1. Nursing stage classification results

The effectiveness of multi-modal fusion in nursing stage classification is first evaluated by comparing the vision-feature-only and vibration-feature-only models with the proposed fusion model.

The single-modality baseline models uses either vision or vibration features only to classify piglet nursing stages, with 0.82 and 0.77 F1 scores using Eq. (10), respectively. The baseline model with vision-only features shows the effectiveness of piglets' location and quantity information. The baseline model with vibration-only features shows how piglets' movement information helps with piglets' nursing stage classification. Fig. 8 shows the confusion matrix for baseline classification results.

The proposed multi-modal fusion method has an average F1 score of 0.94, which is significantly higher than both of the baselines (0.82 for vision-only and 0.77 for vibration-only). Fig. 9 shows the confusion matrix for the overall piglets' nursing stage classification results.

Moreover, the proposed active feature sampling method that produces the potential nursing periods is effective in further improving the precision, where 100% of false positives of non-nursing predictions are corrected. Overall, the final model presents 97% accuracy in classifying five different piglet nursing stages using Eq. (9).

To optimize model performance, various types of classification models are compared to select the most suitable one. Table 1 presents the evaluation metrics comparison among widely adopted classification models in the existing literature (Breiman, 2001; Cortes and Vapnik, 1995; Murtagh, 1991; Friedman, 2001; Mucherino et al., 2009; Gorishniy et al., 2021). Upon examination of the table, the results determined that the random forest model outperforms others. Further analysis and interpretation of the nursing behavior classification results are provided in the subsequent Discussion section.

4.2. Comparison with vision-based image classification methods

To effectively assess the performance of the proposed multi-modal framework, the results were compared with the state-of-the-art computer vision models employed in prior research specifically focused on sow nursing behavior detection (Gan et al., 2021; Pan et al., 2023; Yin et al., 2023), as well as several additional advanced vision-only image classification models that have demonstrated competitive performance on ImageNet (Deng et al., 2009). Each vision-based model was pre-trained on extensive image datasets and fine-tuned to classify piglets' nursing levels from video frames. Table 2 summarizes the results, including model backbone, framework, accuracy, and brief model characteristics.

The results demonstrate that the proposed multi-modal framework achieves superior accuracy compared to these vision-only approaches, underscoring the advantage of integrating multiple data modalities in the classification of piglets' nursing behaviors. The following Discussion

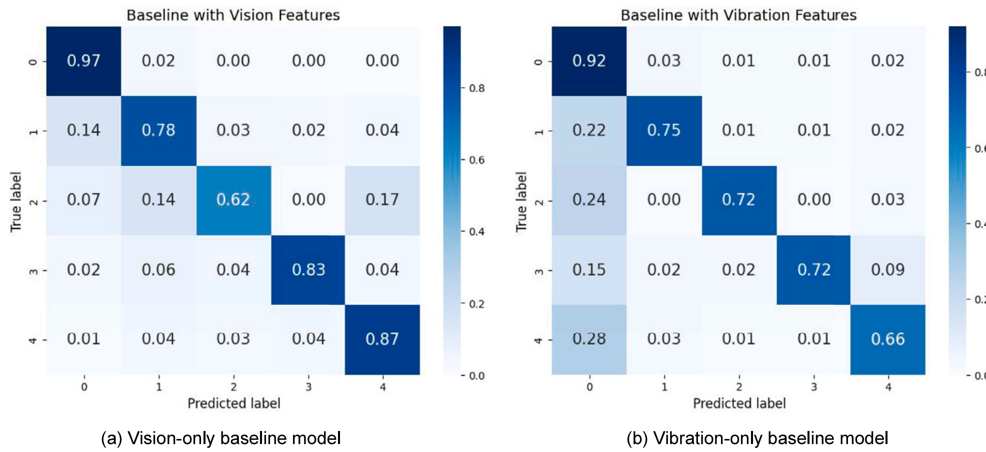

Fig. 8. Confusion matrices of single-modality nursing classification results. This confusion matrix is classified into five distinct nursing stages for analysis: 0 (no nursing), 1 (light nursing with sleeping), 2 (light nursing without sleeping), 3 (high nursing with sleeping), and 4 (high nursing without sleeping). (a) shows results with vision features only. (b) shows results with vibration features only.

Table 1
Comparison of performance metrics for various classification models.

Model	Model type	Library/Framework	F1 Score	Accuracy
Our model (random forest) (Breiman, 2001)	Ensemble trees	Scikit-Learn	**0.94**	0.97
Support vector machine (Cortes and Vapnik, 1995)	Non-parametric	Scikit-Learn	0.85	0.87
Multi-Layer Perceptron (MLP) (Murtagh, 1991)	Neural network	PyTorch	0.90	0.91
Gradient boosting (Friedman, 2001)	Ensemble trees	Scikit-Learn	0.90	0.92
K-Nearest neighbors (Mucherino et al., 2009)	Non-parametric	Scikit-Learn	0.84	0.85
FT-Transformer (Gorishniy et al., 2021)	Transformer	PyTorch	0.92	0.93

Table 2
Comparison of performance metrics for vision-based image classification models.

Backbone relations	Model	Framework	Accuracy	Model characteristics
Prior work (Gan et al., 2021)	ResNet34	PyTorch	0.87	Neural network
Classical model	AlexNet	PyTorch	0.81	Neural network
Advanced model	EfficientNet-B0	PyTorch	0.89	Neural network
Prior work (Pan et al., 2023; Yin et al., 2023)	VGG16	PyTorch	0.86	Neural network
Advanced model (Jocher, 2020)	YOLOv5-classification	PyTorch	0.85	Neural network
Proposed model	Our multi-modal model	Scikit-Learn	0.97	Ensemble trees

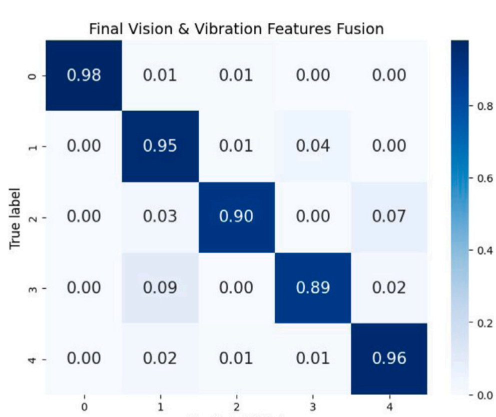

Fig. 9. Confusion matrices of the proposed multi-modal fusion framework for nursing stage classification. This confusion matrix is classified into five distinct nursing stages for analysis: 0 (no nursing), 1 (light nursing with sleeping), 2 (light nursing without sleeping), 3 (high nursing with sleeping), and 4 (high nursing without sleeping). The results are significantly better than the baselines.

section (Section 5) further elaborates on this comparison, highlighting the strengths and limitations of vision-based methods relative to the proposed multimodal approach.

4.3. Long-term continuous monitoring of piglet nursing behaviors

The proposed method also enables long-term continuous monitoring of piglet nursing behaviors with the development of our unsupervised active feature sampling approach that identifies the potential nursing periods. Fig. 10 shows the daily nursing period percentage trend over one week, generated using the unsupervised active feature sampling strategy with the K-Means clustering algorithm introduced in Section 3.3. From Fig. 10, the percentages of nursing period range from 16% to 37% percent of time over each day through observation, with slight variations from day to day. It is worth noting that the percentages of nursing time appear to exhibit a cyclic trend with an increase and decrease every two consecutive days, demonstrating the intrinsically adaptive behavior in piglet nursing. The unsupervised approach provides an efficient method to output the daily nursing and non-nursing duration to reflect the nursing duration variations.

In addition, the nursing stage prediction results of the proposed framework using multi-modal fusion are shown in Fig. 11. The figure demonstrates the predicted trend of nursing intensity over one week, quantified by the nursing percentile defined in Section 2. In Fig. 11, higher nursing intensity represents higher nursing percentile and lower sleeping percentile. This figure offers an alternative perspective on

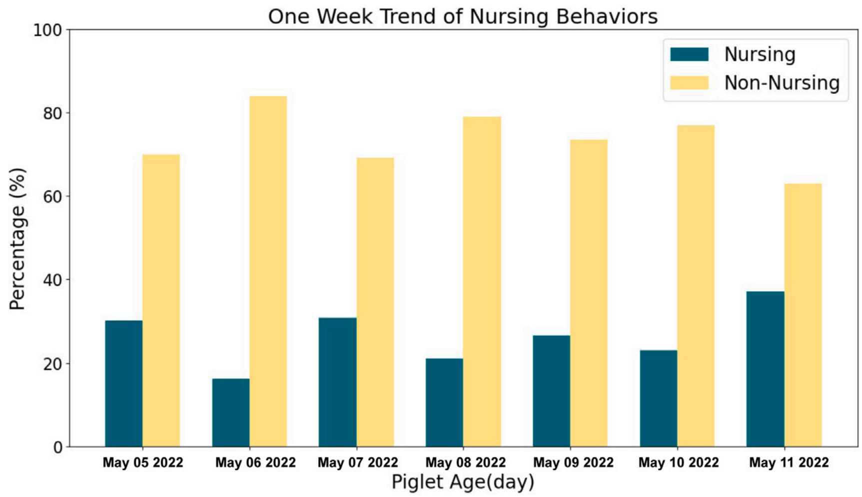

Fig. 10. One-week trend of nursing period identification results from the unsupervised active feature sampling, demonstrating the ability of the framework to provide essential long-term, continuous piglet nursing information without any training data.

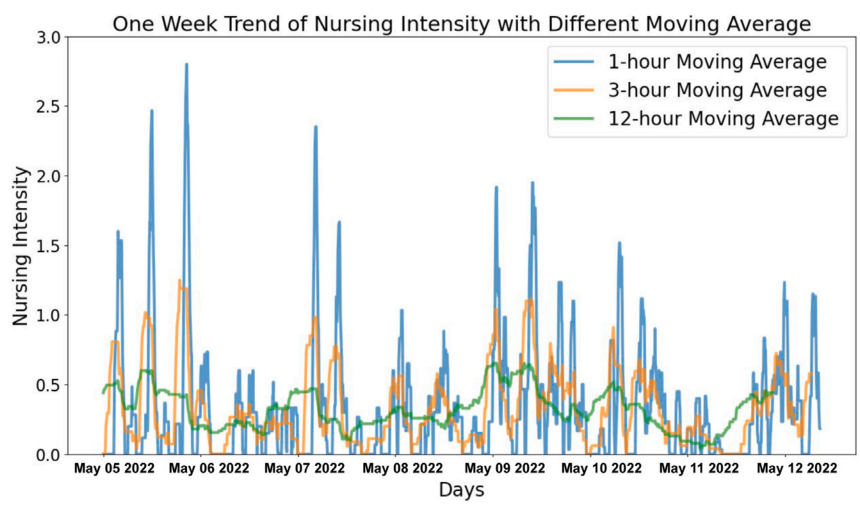

Fig. 11. One-week trend of nursing intensity from supervised multi-modal nursing stage prediction with 97% accuracy. The nursing intensity represents piglet nursing levels: higher nursing intensity indicates more piglets are nursing. The proposed system provides different temporal resolutions according to the monitoring purpose, demonstrating the ability of the framework to provide multi-resolution and continuous piglet nursing intensity information over time.

the trend in nursing intensity transition. While the proposed model produces results for each minute, the results are smoothed by 1-h, 3-h, and 12-h, respectively, to observe the long-term variations despite the noise in piglet behaviors. Among them, the 1-h trend has more detailed changes in nursing behaviors, showing alternating high and low nursing intensity peaks over time. The 3-h trend demonstrates the overall changes within a day, showing 2–3 high-intensity peaks within each day. The 12-h trend shows the day-to-day variations, which align with the cyclic trend observed in Fig. 10. These observations allow us to visualize the overall nursing pattern and the variation of nursing time over a longer term. In practical scenarios, the visualization of nursing duration trends can help producers monitor piglets' growth and health status by simply checking if there are unusual decreases or increases in nursing time.

5. Discussion

5.1. Discussion on results

This subsection aimed to analyze the strengths and limitations of each sensing modality — vision and vibration — as well as to justify the effectiveness of their fusion and the choice of Random Forest as the optimal model for tabular classification. The vision-only baseline (with a 0.82 F1 score) has better performance than the vibration-only baseline (with a 0.77 F1 score), especially when classifying the non-nursing case (label 0). This is because the vision features capture the piglets' location within the nursing zone precisely, which can help to determine if piglets are nursing or not. In contrast, the vibration data only captures the overall behavior of the piglets in a pen, which are sensitive to the noises during the non-nursing periods when there is high uncertainty and heterogeneity in the type and intensity of piglet activities. On the other hand, the vibration-only model has less misclassification among various stages than that of the vision-only model. This is because the vibration features are sensitive to the slight changes in the nursing activity intensity, while such detailed changes are often obstructed by the piglet bodies from the top-down cameras.

The main reason for the improvement over the baseline models is that the features extracted from vision and vibration complement each other's limitations in observation. Specifically, the vision features are effective in distinguishing between nursing and non-nursing, but have limited information to distinguish between different nursing stages. On the other hand, the vibration features are effective in classifying various nursing stages while having difficulty in distinguishing between nursing and non-nursing. When fusing two modalities, the vision features provide the location information while vibration features compensate for the movement information to help with the classification among various nursing stages.

Random forest demonstrates superior performance and offers several advantages: it is less sensitive to hyperparameters, making it easier to tune; it is less prone to overfitting compared to alternative models; and it is computationally efficient due to its ability to train multiple trees in parallel. Additionally, recent research (Grinsztajn et al., 2022) indicated that tree-based models, such as Random Forest, typically outperform deep learning models on tabular datasets due to their robustness to uninformative features, superior ability to capture irregular patterns, and preservation of data orientation. This aligns with our approach, in which both video and vibration data were first condensed into tabular features for reduced storage and computation in practical settings. Moreover, Random Forest exhibits better scalability and generalization on high-dimensional data, enabling producers to easily scale up the system according to their requirements. Therefore, the proposed framework employed Random Forest as the optimal tabular data classification model.

5.2. Comparison with existing work

This subsection aims to evaluate the individual and combined contributions of vision and vibration modalities to classification performance, benchmark the proposed multimodal framework against existing vision-based models, and justify the design choices based on both technical efficiency and practical considerations for real-world deployment. Previous research on sow nursing behavior detection has primarily relied on vision-only approaches. However, the lack of open-source models and code from these studies makes direct comparisons challenging. Therefore, the backbone model architectures from prior work were re-implemented in this study to be adopted as vision-only baselines. Additionally, several recent competitive vision-based image classification models were included in the comparison to assess the relative performance of the proposed multimodal method.

As illustrated in Fig. 1, instances where nursing and sleeping behaviors occur concurrently near the sow's belly present significant visual ambiguity, leading to reduced accuracy in vision-only classification methods. The proposed multimodal approach addresses this limitation by integrating vision features with vibration sensor features, thereby incorporating additional physical context that enhances the system's ability to distinguish between visually similar but behaviorally distinct events. By leveraging the complementary strengths of each modality, the multimodal framework achieves improved accuracy and robustness compared to methods relying solely on video data.

To mitigate the complexity introduced by fusing two data modalities, the framework simplifies processing by transforming the large video dataset into a compact, tabular format using SAM. This design significantly reduces hardware and storage requirements, enabling the system to operate without the need for GPUs during training or inference. From a deployment perspective, this tabular representation supports near real-time processing, as raw video data is analyzed on the fly and not stored. Unlike the existing vision-based methods that require full video storage, the proposed method only saves the extracted vision features locally (e.g., segmented piglet locations and contour), substantially lowering data storage demands with minimal sacrifice in interpretability.

Furthermore, this approach is designed with real-world challenges in mind. Camera-based systems deployed in farms are prone to lighting fluctuations, occlusions, and other visual artifacts that can compromise model performance. In contrast, vibration sensors are unaffected by such issues and provide a stable source of behavioral data. By fusing these two modalities, the system enhances its resilience to environmental variability and ensures more consistent performance in practical farm settings. In particular, multimodal fusion distinctly outperforms single-modality approaches under challenging scenarios involving sub-optimal lighting conditions, partial occlusions, or inconsistent video quality, conditions frequently encountered in practical farm operations. Additionally, by reducing reliance on manual video labeling, the proposed multi-modal approach mitigates systematic errors and significantly enhances classification robustness and reliability.

6. Considerations of practical variations

In this section, we discuss the performance of our framework under varying scenarios, including varying lighting conditions, different sensor locations, and various pig pen layouts. Finally, we summarize the potential limitations of this work and identify areas for future improvements.

6.1. Lighting condition variations

Lighting condition variations can be an issue that affects the visibility of video data and further influences the accuracy of monitoring piglet nursing behaviors. The proposed framework demonstrates robustness against variations in lighting conditions, where effective vision features can still be extracted. For example, the upper part of Fig. 12 shows the variability in lighting conditions, which is influenced by the status of the heat lamp — whether it is on or off, as well as by the natural light changes throughout the day. To evaluate the effectiveness of the system on varying lighting conditions, the segmentation performance for images across two extreme lighting scenarios encountered is tested, including low light and overexposed lighting conditions. As shown in the lower part of Fig. 12, although there are limitations of SAM segmentation results such as cutting one piglet into two segments or combining two piglets as one segment due to low light or overexposed lighting conditions respectively, such limitation is compensated by the multi-modal fusion and does not significantly affect the final results. Overall, the proposed system provides satisfactory piglet segments under these two extreme conditions. The low-light example represents the situation when the heat lamp is off during a cloudy day with less natural light. On the opposite side, the overexposed example represents the scenario with strong natural light or intense light from the heat lamp. The challenging lighting conditions are overcome by SAM through the identification of object boundaries to provide satisfactory image segments. Therefore, the proposed framework is able to process the segmentation results to extract necessary features, which remain unaffected by variations in lighting conditions.

6.2. Sensor location variations

Sensors are deployed at different locations in the crate, and different sensors exhibit varying percentages of data loss. To evaluate the influence of such variations on the model performance, each sensor in the crate is tested separately.

During the experiments, three sensors were installed along the diagonal line of the crate 2a to provide redundancy in data collection and allow performance comparison across sensor locations. In practical settings, a single sensor is typically sufficient to monitor nursing behavior. Fig. 13 shows the effect of sensor locations on nursing stage monitoring accuracy, along with the percentage of data loss during transmission. There is less than 2% variance in prediction accuracy among the three different sensors, meaning that the location does not contribute significantly to the model performance. In addition, a correlation between accuracy and data loss — Sensor 1 has the best prediction accuracy (97.5%) and the least data loss (5%), while Sensor 2 has the worst accuracy (96%) with the most data loss (10%) are observed. This implies that the minor variations in accuracy differences may be caused by the amount of data loss during transmission. This comparison suggests future improvements in data transmission are warranted, while all three sensor locations tested were similar in prediction accuracy.

Fig. 12. Segmentation results with lighting condition variations. The segmentation results show that the proposed framework is robust since satisfactory image segments are provided under the two typical extreme lighting conditions at the pig farm, including low light and overexposed conditions.

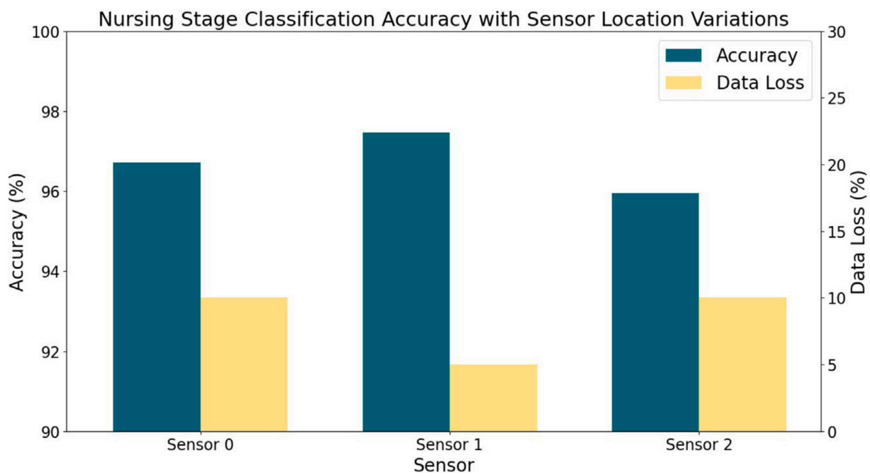

Fig. 13. Nursing stage classification accuracy with respect to sensor locations and data loss. The figure shows that sensors that have higher accuracy tend to have less data loss.

6.3. Pig pen layout variations

The proposed framework is designed to accommodate the variations in crate structures or layouts of diverse pig farms. There are three types of layout used in this particular nursing pen deployment: (1) offset, (2) diagonal, and (3) centered, as shown in Fig. 14. Each has its advantages and disadvantages, as discussed in existing studies (Ramirez et al., 2022; Condous et al., 2016). When adapting the proposed framework across various pig pens, the only change required is to redefine the location of the nursing zone. Based on the location of the sow, the area near the sow's belly is typically defined as a piglet nursing zone regardless of the crate layout. As shown in Fig. 14, the yellow boxes next to and parallel to the sow's belly are defined as nursing zones. Although the crate layouts are different, the nursing zones are always defined the same way as the rectangular area next to and parallel to the sow's belly. Once the nursing zones are defined, the nursing activities can be isolated from the other activities by focusing on the nursing zone, enabling the proposed system to process and monitor piglet nursing behavior effectively for different pig farm crate layouts.

6.4. Discussion on limitations and future improvements

While the framework has satisfactory accuracy and demonstrates robustness among various situations, there are several areas remain to be improved for future work, including (1) reducing onsite data storage and processing requirements, (2) adaptation to changes in floor properties, darkness, and data loss, and (3) evaluating long-term nursing behavior monitoring for the entire lactation period.

First, the data storage and processing requirements need to be reduced due to the large amount of video data (1.2 GB per hour) collected by the current system. Therefore, data compression and down-sampling methods need to be explored to address the trade-off between accuracy and computational expenses. While the current clustering and SAM-based method have significantly reduced the labeling effort, the five-stage nursing behavior classification still requires manual labeling on nursing stages. Future work may explore semi-supervised or self-supervised learning to mitigate this effort, enabling fine-grained continuous monitoring without the need for training data.

Second, potential data distribution shifts need to be considered thoroughly to ensure the robustness of the method. For example, changes in floor property and piglet behaviors are known to affect the data and thus may lead to a decrease in monitoring accuracy. In this study, the pig pens are made of the same materials and are relatively isolated, so the vibration data are observed to be consistent across pens. Future work needs to further explore the effect of the variations in floor properties and noises from the sow or adjacent pens. In addition, a strategy to handle data loss in the farm environment is needed, as occasional network and hardware failures did happen during our experiment in pig farms. Therefore, future efforts need to adapt the model through parameter fine-tuning or domain adaptation to make it resilient in coping with time-varying scenarios when one or more modalities are not available or have low-quality data.

Finally, while the current system has assessed one week of piglet nursing data, longer term evaluation of our method is needed to observe piglet nursing behavior over the entire lactation cycle (around 30 days). We will explore the changes in piglet nursing behaviors and identify unique patterns in behavior variations over time to provide more insights to the farm owners.

7. Conclusions

In this paper, a multi-modal framework is developed for robust piglet nursing behavior monitoring through the integration of computer vision and floor vibration. To overcome the challenge of heterogeneous nursing behaviors among multiple piglets, domain knowledge of piglet nursing patterns is incorporated by formulating various nursing stages

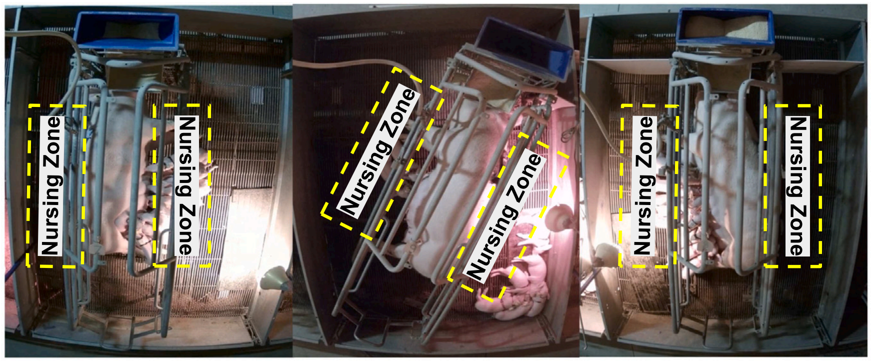

Fig. 14. Nursing zone characterization for various crate layouts. Three types of layout are shown in this figure: (1) offset, (2) diagonal, and (3) centered. The yellow box represents the nursing zone to accommodate the variations in crate layouts.

and defining nursing zones. The main insight behind the multi-modal fusion is the complementary nature of these two sensing modalities — while the cameras capture individual piglet locations, it has limited accuracy in identifying whether the piglet is nursing or sleeping near the sow's belly when they are at the same location; the vibration data compensate for this limitation by sensing activity-induced floor vibration, allowing accurate tracking among various nursing stage transitions. The proposed framework is evaluated through a real-world deployment at a pig farm and obtained 97% accuracy in identifying five different nursing stages (3× and 3.8× error reduction when compared with vision- and vibration-only baselines). The results demonstrate the effectiveness of the proposed framework and exhibit robustness across various lighting conditions, sensing locations, and pen layouts. Future work will extend the system for more automatic, adaptive, and long-term nursing behavior monitoring.

CRediT authorship contribution statement

Yiwen Dong: Writing – review & editing, Visualization, Supervision, Resources, Methodology, Formal analysis, Conceptualization, Writing – original draft, Validation, Software, Project administration, Investigation, Data curation. **Zihao Song:** Writing – original draft, Validation, Methodology, Formal analysis, Writing – review & editing, Visualization, Software, Investigation, Data curation. **Jesse R. Codling:** Methodology, Conceptualization, Resources, Data curation. **Gary Rohrer:** Supervision, Project administration, Conceptualization, Writing – review & editing, Resources, Investigation. **Jeremy Miles:** Writing – review & editing, Project administration, Conceptualization, Resources, Data curation. **Sudhendu Sharma:** Resources, Writing – review & editing, Project administration. **Tami Brown-Brandl:** Supervision, Conceptualization, Writing – review & editing, Resources. **Pei Zhang:** Supervision, Project administration, Conceptualization, Resources, Funding acquisition. **Hae Young Noh:** Writing – review & editing, Resources, Investigation, Conceptualization, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded in part by the U.S. National Science Foundation (under grant numbers NSF-CMMI-2026699), Stanford Blume Fellowship, and Cisco Systems. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of any University, the National Science Foundation, or the United States Government or any of its agencies. The USDA prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program (Not all prohibited bases apply to all programs.). Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720–2600 (voice and TDD). USDA is an equal opportunity provider and employer.

Data availability

The authors do not have permission to share data.

References

Baxter, E.M., Edwards, S., 2021. Optimising sow and piglet welfare during farrowing and lactation. In: Understanding the Behaviour and Improving the Welfare of Pigs. Burleigh Dodds Science Publishing, pp. 121–176. http://dx.doi.org/10.1201/9781003048220.

Blavi, L., Solà-Oriol, D., Llonch, P., López-Vergé, S., Martín-Orúe, S.M., Pérez, J.F., 2021. Management and feeding strategies in early life to increase piglet performance and welfare around weaning: A review. Animals 11 (2), 302. http://dx.doi.org/10.3390/ani11020302.

Bonde, A., Codling, J.R., Naruethep, K., Dong, Y., Siripaktanakon, W., Ariyadech, S., Sangpetch, A., Sangpetch, O., Pan, S., Noh, H.Y., Zhang, P., 2021. PigNet: Failure-tolerant pig activity monitoring system using structural vibration. In: Proceedings of the 20th International Conference on Information Processing in Sensor Networks. Co- Located CPS- IoT Week 2021, Association for Computing Machinery, http://dx.doi.org/10.1145/3412382.3458902.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. http://dx.doi.org/10.1023/A:1010933404324.

Condous, P.C., Plush, K., Tilbrook, A., Van Wettere, W., 2016. Reducing sow confinement during farrowing and in early lactation increases piglet mortality. J. Anim. Sci. 94 (7), 3022–3029. http://dx.doi.org/10.2527/jas.2015-0145.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297. http://dx.doi.org/10.1007/BF00994018.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255. http://dx.doi.org/10.1109/CVPR.2009.5206848.

Dong, Y., Bonde, A., Codling, J.R., Bannis, A., Cao, J., Macon, A., Rohrer, G., Miles, J., Sharma, S., Brown-Brandl, T., et al., 2023. PigSense: Structural vibration-based activity and health monitoring system for pigs. ACM Trans. Sens. Netw. 20 (1), 1–43. http://dx.doi.org/10.1145/3604806.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232. http://dx.doi.org/10.1214/aos/1013203451.

Gan, H., Guo, J., Liu, K., Deng, X., Zhou, H., Luo, D., Chen, S., Norton, T., Xue, Y., 2023. Counting piglet suckling events using deep learning-based action density estimation. Comput. Electron. Agric. 210, 107877. http://dx.doi.org/10.1016/j.compag.2023.107877.

Gan, H., Li, S., Ou, M., Yang, X., Huang, B., Liu, K., Xue, Y., 2021. Fast and accurate detection of lactating sow nursing behavior with CNN-based optical flow and features. Comput. Electron. Agric. 189, 106384. http://dx.doi.org/10.1016/j.compag.2021.106384.

Gorishniy, Y., Rubachev, I., Khrulkov, V., Babenko, A., 2021. Revisiting deep learning models for tabular data. In: NeurIPS.

Grinsztajn, L., Oyallon, E., Varoquaux, G., 2022. Why do tree-based models still outperform deep learning on tabular data? arXiv:2207.08815. URL https://arxiv.org/abs/2207.08815.

Hojgaard, C.K., Bruun, T.S., Theil, P.K., 2020. Impact of milk and nutrient intake of piglets and sow milk composition on piglet growth and body composition at weaning. J. Anim. Sci. 98 (3), skaa060. http://dx.doi.org/10.1093/jas/skaa060.

Jensen, P., Stangel, G., Algers, B., 1991. Nursing and suckling behaviour of semi-naturally kept pigs during the first 10 days postpartum. Appl. Anim. Behav. Sci. 31 (3–4), 195–209. http://dx.doi.org/10.1016/0168-1591(91)90005-I.

Jocher, G., 2020. Ultralytics YOLOv5. http://dx.doi.org/10.5281/zenodo.3908559, URL https://github.com/ultralytics/yolov5.

Kim, J.H., Ni, J.-Q., Ogundare, W., Schinckel, A.P., Minor, R.C., Johnson, J.S., Casey, T.M., 2025. Sow and piglet behavior characterization using visual observation, sensor detection, and video recording. Appl. Sci. 15 (6), 3018. http://dx.doi.org/10.3390/app15063018.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment anything. http://dx.doi.org/10.48550/arXiv.2304.02643, arXiv:2304.02643.

Kumar, N., Thakur, R., Harini, K., Ajay, A., Das, A., 2025. Chapter 11 - care and management of piglets. In: Chauhan, A., Tarafdar, A., Gaur, G.K., Jadhav, S.E., Tiwari, R., Dutt, T. (Eds.), Commercial Pig Farming. Academic Press, pp. 173–184. http://dx.doi.org/10.1016/B978-0-443-23769-0.00011-7, URL https://www.sciencedirect.com/science/article/pii/B9780443237690000117.

Lay, Jr., D., Matteri, R., Carroll, J., Fangman, T., Safranski, T., 2002. Preweaning survival in swine. J. Anim. Sci. 80 (E-suppl_1), E74–E86. http://dx.doi.org/10.2527/animalsci2002.0021881200800ES10011x.

Li, B., Xu, W., Chen, T., Cheng, J., Shen, M., 2023. Recognition of fine-grained sow nursing behavior based on the SlowFast and hidden Markov models. Comput. Electron. Agric. 210, 107938. http://dx.doi.org/10.1016/j.compag.2023.107938.

Manteuffel, C., Hartung, E., Schmidt, M., Hoffmann, G., Schön, P.C., 2017. Online detection and localisation of piglet crushing using vocalisation analysis and context data. Comput. Electron. Agric. 135, 108–114. http://dx.doi.org/10.1016/j.compag.2016.12.017.

Martin, C.E., Wagner, W.C., Elmore, R.G., Ross, R.F., 1978. Mastitis, metritis, agalactia (MMA).

Mucherino, A., Papajorgji, P.J., Pardalos, P.M., 2009. K-nearest neighbor classification. In: Data Mining in Agriculture. Springer New York, New York, NY, pp. 83–106. http://dx.doi.org/10.1007/978-0-387-88615-2_4.

Murtagh, F., 1991. Multilayer perceptrons for classification and regression. Neurocomputing 2 (5), 183–197. http://dx.doi.org/10.1016/0925-2312(91)90023-5.

Pan, Z., Chen, H., Zhong, W., Wang, A., Zheng, C., 2023. A CNN-based animal behavior recognition algorithm for wearable devices. IEEE Sens. J. 23 (5), 5156–5164. http://dx.doi.org/10.1109/JSEN.2023.3239015.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Puppe, B., Meunier-Salaün, M.-C., Otten, W., Orgeur, P., 2008. The welfare of piglets. Welf. Pigs 97–131.

Ramirez, B.C., Hayes, M.D., Condotta, I.C., Leonard, S.M., 2022. Impact of housing environment and management on pre-/post-weaning piglet productivity. J. Anim. Sci. 100 (6), skac142. http://dx.doi.org/10.1093/jas/skac142.

Rempel, L.A., Keel, B.N., Oliver, W.T., Wells, J.E., Lents, C.A., Nonneman, D.J., Rohrer, G.A., 2022. Dam parity structure and body condition during lactation influence piglet growth and gilt sexual maturation through pre-finishing. J. Anim. Sci. 100 (4), skac031. http://dx.doi.org/10.1093/jas/skac031.

Sadeghi, E., Kappers, C., Chiumento, A., Derks, M., Havinga, P., 2023. Improving piglets health and well-being: A review of piglets health indicators and related sensing technologies. Smart Agric. Technol. 5, 100246. http://dx.doi.org/10.1016/j.atech.2023.100246.

Valros, A., Rundgren, M., Špinka, M., Saloniemi, H., Rydhmer, L., Algers, B., 2002. Nursing behaviour of sows during 5 weeks lactation and effects on piglet growth. Appl. Anim. Behav. Sci. 76 (2), 93–104. http://dx.doi.org/10.1016/S0168-1591(02)00006-0, URL https://www.sciencedirect.com/science/article/pii/S0168159102000060.

Yang, A., Huang, H., Yang, X., Li, S., Chen, C., Gan, H., Xue, Y., 2019. Automated video analysis of sow nursing behavior based on fully convolutional network and oriented optical flow. Comput. Electron. Agric. 167, 105048. http://dx.doi.org/10.1016/j.compag.2019.105048.

Yin, M., Ma, R., Luo, H., Li, J., Zhao, Q., Zhang, M., 2023. Non-contact sensing technology enables precision livestock farming in smart farms. Comput. Electron. Agric. 212, 108171. http://dx.doi.org/10.1016/j.compag.2023.108171.