Review Article

A Review in the Automatic Detection of Pigs Behavior with Sensors

Ziting Zhang, Hang Zhang, Yuxiang He, and Tonghai Liu

College of Computer and Information Engineering, Tianjin Agricultural University, Tianjin, China

Correspondence should be addressed to Yuxiang He; heyuxiang@tjau.edu.cn and Tonghai Liu; tonghai_1227@126.com

Received 11 September 2021; Revised 17 March 2022; Accepted 31 March 2022; Published 30 April 2022

Academic Editor: Domenico Caputo

Copyright © 2022 Ziting Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the trend of large-scale and intensive livestock farming, it has become difficult to monitor the physical health and breeding performance of animals accurately, quickly, and comprehensively in the traditional feeding management mode, and with the development of computer technology, machine vision has been widely used as a non-contact sensor technology. In this work, it reviews the current automatic behavior detection methods. The chapters are summarized with machine vision technology as the core and classical behavior detection techniques are added for comparison. It is classified into three aspects: breeding performance, disease and health, and social behavior, and then discusses in detail daily behaviors (lying, drinking, feeding, lameness, etc.) and complex behaviors (mounting, aggression, tail biting, etc.).

1. Introduction

The global demand for livestock products is expected to increase further with the growth of population size and the acceleration of urbanization[1]. Traditional livestock in a single mode presents disadvantages such as low breeding efficiency, high labor cost, and large workload that hinder the development of agricultural modernization. Therefore, breeding oriented to scale, intensive, and digitalization has become an inevitable requirement for the development of precise livestock farming. Meanwhile, large-scale and intensive farming also brings higher requirements for management specifications. The traditional breeding model relies on empirical knowledge to analyze behavioral changes, to monitor the health status of animals, and to make appropriate policy. However, precision animal husbandry can integrate individual animal information at multiple stages such as feeding, breeding, and slaughtering. Combined with animal’s body condition status, feeding environment, and other individual information, it can perform personalized feeding, health monitoring, and timely breeding for animals of different growth cycles.

With the emergence of animal health, man-made stress, and environmental problems in scale farming, higher demands have been placed on the energy-efficient feeding management of enterprises. Animal welfare has received more scholarly attention in recent years, and welfare management aims to reduce the intensity of negative impacts critical to survival to tolerable levels [2]. Timely detection of behaviors that are contrary to animal welfare and health is beneficial to the profitability as well as sustainability of farming systems. Animals express emotions such as joy, anger, and anxiety through behavior, and the emotional expression can often provide the basis for daily feeding, such as disease monitoring, estrus detection, and prenatal and postnatal monitoring. Due to the subtle behavioral changes, it is expensive and impossible to rely on manual observation for long periods of time in large-scale farming. Some abnormal behaviors that occur regularly are often accompanied by diseases and other problems, and the machine can replace the manual real-time monitoring, such as abnormal excretion [3], abnormal water intake [4], abnormal activity [5], and abnormal breathing [6] to predict the occurrence of diseases and reduce the risk to reduce economic losses.

This paper summarizes the technologies related to automatic animal behavior detection from earlier studies, focuses on the current research status of domestic and international
2. Automatic Behavior Detection Systems

2.1. Accelerometer-Based in Behavior Detection. Accelerometer and gait scoring as a means of studying early motor behavior are effective in detecting gait abnormalities, changes in activity levels, and gait status during eating and drinking [7]. The accelerometer expresses the velocity change in each directional axis as vector information and determines the speed and direction of the motion behavior by the voltage value of the sensor [8]. In addition to measuring linear velocity, it measures the earth’s gravitational pull by determining the angle at which the device is tilted to describe the angular velocity of the motion. Measurements are usually made with single or double sensors on the neck, leg, back, and ear of the pig. For example, Main et al. [9] attached accelerometers to the hind legs of pigs to explore gait patterns to detect lameness. Cornou et al. [10] obtained lying variation data from an accelerometer placed on the neck of the pig to classify the pig’s posture specifically. Escalante et al. [11] also installed the device on the neck to quantify the feeding behavior of the pigs. The behavior is classified by different vector variations on the three axes. The different parts of the device placement also affect the trend of the data. Exploring multiple behaviors often requires multiple device acquisitions for analysis, which are expensive in commercial farming.

2.2. Vision Systems in Behavior Detection. Computer vision constructs an explicit and meaningful description of physical objects from images [12]. The vision system in behavior detection consists of vision sensors and computer hardware and software. Behavior detection by machine vision is mainly embodied in the following process: animal behavior images are obtained by vision sensors and transmitted to dedicated computer hardware for behavior image processing and analysis by software. The purpose is to perform the visual task of behavior detection. The application of vision systems in pig behavior detection is shown in Table 1.

<table>
<thead>
<tr>
<th>Visual sensor type</th>
<th>Composition of visual system</th>
<th>Image type</th>
<th>Visual angle</th>
<th>Behavioral feature</th>
<th>Feature description</th>
<th>Conclusion</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visible light sensor</td>
<td>Single CCD</td>
<td>Grayscale image</td>
<td>Top view</td>
<td>Drinking behavior</td>
<td>Use dynamic data-based model estimate half-hourly water use</td>
<td>92%</td>
<td>[13]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lying behavior</td>
<td>An ellipse fitting technique applied to find pig lying positions</td>
<td>93-95%</td>
<td>[14]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Locomotion</td>
<td>Calculation of pixel-level motion after ellipse fitting of pigs</td>
<td>89.8%</td>
<td>[15]</td>
</tr>
<tr>
<td>Infrared sensor</td>
<td>Infrared thermography</td>
<td>Thermography</td>
<td>Top view</td>
<td>Lying behavior</td>
<td>Scoring the average leg temperature to identify lameness</td>
<td>—</td>
<td>[16]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Locomotion</td>
<td>Comparison of changes in vulvar skin temperature and buttock skin temperature</td>
<td>—</td>
<td>[17]</td>
</tr>
<tr>
<td>Depth sensor</td>
<td>Kinect</td>
<td>Depth image</td>
<td>Top view</td>
<td>Locomotion</td>
<td>Extraction of skin temperature changes in the area of interest</td>
<td>—</td>
<td>[18]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Five postures (standing, sitting, sternal recumbency, ventral recumbency, and lateral recumbency)</td>
<td>Deep learning recognition five types of gestures</td>
<td>—</td>
<td>[19]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aggressive behavior</td>
<td>Clustering training kinetic energy difference between adjacent frames</td>
<td>95.8%</td>
<td>[20]</td>
</tr>
</tbody>
</table>
Sensors with three-dimensional (3D) imaging systems are or fail to satisfy the posture recognition. Therefore, depth view; the standing behavior and kneeling behavior that are coupled Device (CCD) camera is above the pig pen in top of detecting groups of pigs, such as when the Charge-coupled Device (CCD) camera is above the pig pen in top view; the standing behavior and kneeling behavior that are right below the camera are likely to cause recognition errors right below the camera are likely to cause recognition errors or fail to satisfy the posture recognition. Therefore, depth sensors with three-dimensional (3D) imaging systems are easy to extract height features and establish world coordinates more accurately compared to two-dimensional (2D) imaging systems. Among them, Time-of-Flight (ToF) and Kinect cameras are often used in animal behavior detection. ToF cameras output images with depth information by measuring the light pulses reflected from the surface of an object. The longer it takes for the reflected light pulse to reach the sensor, the greater the distance to the object. The principle of data acquisition by the Kinect depth sensor is similar to that of the ToF [24]. The researchers utilized depth sensors to obtain animal orientation and 3D spatial information to better identify behavioral postures. Zheng et al. [19] classified sows into five categories of lactation postures (standing, sitting, sternal recumbent, ventral recumbent, and lateral recumbent) based on the combination of position, orientation, and connection relationships of the body parts of the sow acquired by Kinect, which provided basic information to study the behavioral characteristics and patterns of sows. Some researchers have used depth images to obtain 2D image information. Chen et al. [20] acquired depth image information of pigs by Kinect and built a kinetic model to classify attack and non-attack behaviors with an accuracy of 95.8%. The accurate acquisition of depth images by depth sensors such as Kinect poses computational challenges for hardware devices, yet the precision of low-consumption depth sensors decreases as the distance used increases. Lee et al. [25] proposed a method with complementary depth and infrared images, where the depth image compensates the disadvantage of lower accuracy in the dark of the infrared image due to less influence by light, but the advantage of accurate pixel values of the infrared image can be obtained. The method has an execution time of 8.71 ms real-time detection accuracy of 95% and an execution time of 14.65 ms with 80% accuracy using YOLO in the same experimental environment.

Infrared imagers capture the heat information emitted by an object by receiving and measuring infrared radiation from the surface of the object, and the sensor then converts it into radiometric temperature data [23]. Can be non-contact, real-time access to thermal information at different behavior to avoid the stress response of the traditional contact monitoring body temperature, thereby improving animal welfare based on how it works. Infrared imagers can make up for the insufficiency of visible light cameras that are limited to collect data at night, instead of manual monitoring of animal behavior day and night without interruption and reduces manual consumption. Amezcue et al. [16] proposed a method to measure the average temperature of the sow’s legs through an infrared imager, which provided conditions for detecting the lame behavior of the sow. Scolari et al. [17] measured vulvar temperature changes by thermography for estrus detection. Boileau et al. [18] using a thermal imager acquired a plan view of the thermal imaging pig, and the pig was observed failure or retreat one feel pain when the temperature dropped, the thermal imaging evidence based on the back as aggressive behavior when fighting.

The two-dimensional images acquired by the visible light sensor can produce errors due to the complex environment of detecting groups of pigs, such as when the Charge-coupled Device (CCD) camera is above the pig pen in top view; the standing behavior and kneeling behavior that are right below the camera are likely to cause recognition errors or fail to satisfy the posture recognition. Therefore, depth sensors with three-dimensional (3D) imaging systems are easy to extract height features and establish world coordinates more accurately compared to two-dimensional (2D) imaging systems. Among them, Time-of-Flight (ToF) and Kinect cameras are often used in animal behavior detection. ToF cameras output images with depth information by measuring the light pulses reflected from the surface of an object. The longer it takes for the reflected light pulse to reach the sensor, the greater the distance to the object. The principle of data acquisition by the Kinect depth sensor is similar to that of the ToF [24]. The researchers utilized depth sensors to obtain animal orientation and 3D spatial information to better identify behavioral postures. Zheng et al. [19] classified sows into five categories of lactation postures (standing, sitting, sternal recumbent, ventral recumbent, and lateral recumbent) based on the combination of position, orientation, and connection relationships of the body parts of the sow acquired by Kinect, which provided basic information to study the behavioral characteristics and patterns of sows. Some researchers have used depth images to obtain 2D image information. Chen et al. [20] acquired depth image information of pigs by Kinect and built a kinetic model to classify attack and non-attack behaviors with an accuracy of 95.8%. The accurate acquisition of depth images by depth sensors such as Kinect poses computational challenges for hardware devices, yet the precision of low-consumption depth sensors decreases as the distance used increases. Lee et al. [25] proposed a method with complementary depth and infrared images, where the depth image compensates the disadvantage of lower accuracy in the dark of the infrared image due to less influence by light, but the advantage of accurate pixel values of the infrared image can be obtained. The method has an execution time of 8.71 ms real-time detection accuracy of 95% and an execution time of 14.65 ms with 80% accuracy using YOLO in the same experimental environment.

Besides imaging objects in the visible (VIS) color region, some machine vision systems are also able to inspect these objects in light invisible to humans, such as infrared (IR). It is very useful in morphological signs, body temperature, and behavioral information. It also helps to explore further applications such as disease, reproduction, and psychosocial.

**Figure 1:** Different levels in the image processing process [26].
2.2.2. Image Processing and Analysis in Behavior Detection. Image processing and analysis involves a series of steps, which can be broadly divided into three levels: low-level processing, mid-level processing, and high-level processing indicated in Figure 1 [26].

Low-level processing is the acquisition and pre-processing of behavioral images. The vision sensors will acquire raw images of variable quality. The region of interest (ROI) in behavioral detection requires enhancement to improve the image quality. Commonly used pre-processing includes correction of geometric distortion, noise removal, grayscale correction, and blur correction [27]. Xue et al. [28] adopted median filtering to remove noise and restricted contrast adaptive histogram equalization to enhance depth images.

Hierarchical image segmentation process is described, embodied as a target extraction and detection feature extraction behavior. Target extraction is a technology of image segmentation, which extracts research objects in a static single frame image or a dynamic continuous frame image. The result greatly affects the precision of behavior detection at a later stage. Segmentation can be achieved by three different techniques: thresholding, edge-based segmentation, and region-based segmentation as shown in Figure 2. Feature extraction is a processing method of image description. It needs to extract quantitative information from previously segmented images. Feature extraction as a processing method for image description requires quantitative information to be extracted from previously segmented images. Behavioral changes in animals are characterized as continuous and difficult to detect. Different behaviors have similar features in physical appearance, and features such as geometry, color, texture, grayscale, and contour are quantified by specific algorithms to facilitate recognition by classifiers. Nasirahmadi et al. ([14, 22]) used Otsu method [29] for global thresholding segmentation, setting thresholds to convert grayscale images to binary images and then using morphological closure operations to remove noise. The pig target was successfully extracted from the pig pen background and then ellipse fitting was used to describe the pig location and orientation based on the geometric features of the ellipse. Zhou et al. [30] used the Otsu method to segment the pig and the background after converting the image to grayscale and combine texture features and color features based on the Camshift algorithm to describe the pig movement status. The above methods are based on grayscale images for segmentation. Other researchers have performed image segmentation process based on color information. Xiao et al. [31] designed a dynamic color channel selection method to segment pigs using a threshold selection strategy that combines the maximum variance between classes and the minimum variance within classes. In order to ensure that the pigs have smaller targets when shooting the pigs in a panoramic view, it is not conducive to the observation of detailed sexual behaviors such as abdominal breathing behaviors. Ma et al. [32] proposed to automatically extract the contour of the pig from the video, segment the contour of the pig through the Sobel operator, and then use the Fourier descriptor to describe the contour feature. Sobel operator is based on edge detection segmentation method, and Robert operator, Prewitt operator, Canny operator, etc. are also commonly used. The edges are important information of an image [33] and segmenting the target contour from the background can obtain more detailed information which were used to describe the behavior. Some developers have used specific algorithms, such as deep learning-based algorithms for image segmentation and feature extraction. Song [33] used a modified RestNet-based deep convolutional network structure for individual pig detection with an accuracy of 96.4%. Han et al. [34] proposed a decision tree image segmentation model based on color features using CART algorithm to achieve target and background segmentation. Gao et al. [35] used an improved Mask R-CNN network for the segmentation of adherent pigs. The accuracy of segmenting the target and background classes was 86.15%, and the computing time was reduced by 30 ms compared with that the algorithm improved.

The high-level processing is recognition. In behavior detection, the classifier or deep neural network designed for the region of interest performs behavior detection based on the extracted features. With the development of computer technology, deep learning algorithms can automatically extract image features driven by data, which has accelerated the development of machine vision in automatic behavior detection. It is possible to learn deeper features from a large number of image samples of animal behaviors, but it is precisely because of the self-learning nature of features that it is difficult to explain them and requires in-depth exploration. Zeng [36] proposed a Faster R-CNN algorithm with improved Anchor and improved the detection speed by deleting some target frames in forward propagation. The accuracy of identifying pig delivery behavior and determining the moment of birth of piglets was 97%. Zhuang et al. [37] used an improved convolutional neural network AlexNet to identify the estrus behavior of large Landrace sow with an accuracy of 93.33% and an average single detection time of 26.28 ms.

The interaction with each stage at all stages of the entire process is essential for more precise decision-making and is
seen as an integral part of the image processing process, as shown in Table 2. The effectiveness of intelligent decision-making is more dependent on the integrity of the computer system, and the addition of machine vision technology has accelerated the achievement of system integrity. Genetic algorithms, fuzzy logic, neural networks, and other algorithms provide system control for recognition detection through constructed image understanding and decision-making capabilities.

3. Pig Behavior Patterns

3.1. Daily Behavior Patterns. Daily behavior reflects the most basic physiological needs of pigs including feeding, drinking, lying, and locomotion behaviors. The appearance of disease, estrus, and other abnormal states causes changes in behavioral patterns. The maintenance of pig body condition and timely breeding play a key role in production efficiency. Genetic algorithms, fuzzy logic, neural networks, and other algorithms provide system control for recognition detection through constructed image understanding and decision-making capabilities.

3.1.1. Feeding and Drinking Behavior. Pigs are supplemented with daily nutrition through feeding behavior and drinking behavior. Changes in feeding and drinking behavior are a key symptom of health and animal welfare [1, 2]. Precise quantification of eating behaviors at an early stage could prevent the occurrence of diseases and other problems to reduce the risk. Quantification can be expressed in a variety of ways, such as recording the duration of time spent chewing/biting food (drinking water) or recording the amount of time and/or frequency that the head of the animal is in the food trough or drinking system. The determination of feeding behavior relies on the extraction of features of the surrounding environment such as drinker/feeding troughs. Therefore, the determination of non-nutritive visits (NNVs) such as pigs exploring the environment becomes the key to improve the accuracy [43, 44]. The eating behavior of healthy pigs occurs at a fixed time. The appearance of abnormal behaviors such as decrease/increase in the number of feeding behaviors and decrease in feeding time could be used as a basis for determining the abnormal status of pigs, already widely used in large-scale farming for prediction of disease and estrus behavior, as shown in Table 3.

In early studies, feeding and drinking behavior detection was performed by electronic flow meters and accelerometers. Madsen et al. [4] installed electronic water flowmeters for measurements of 24 h drinking monitoring. The flow rate was measured at discrete intervals with a period of 2 mins. The V-mask arms were set to avoid the accumulation and error caused by the variation of growth rate. Practical application showed that data changes could be observed 17 h before the disease outbreak. The error mentioned in the method and based on a priori knowledge, in order to reduce the error caused by subjective factors, needed to be automated methods to extract features. Escalante et al. [11] used LIS3L02DS to measure acceleration data. The 4D vector was extracted from the measurements which consist of the three-dimensional axes and the length of the acceleration vector. After multiple classifier comparisons, the logitboost classifier reached the best prediction with 90.79% for feeding behavior. When dividing the time activity sequence in 2 min units, feeding behavior was correctly classified by the logitboost classifier with 100%. These experimental results indicated that feeding behavior could be effectively identified in behavioral sequences.

In order to monitor water drinking rate, Kashiha et al. [13] installed CCDs for recording drinking nipple visits with top-view images. The pig body was extracted from the binarized image using ellipse fitting. The body contours were then separated by reference to the centroid of body image. Water usage and water dispenser access times were evaluated through a dynamic data-based model with discrete unit times of half an hour. The analysis described the half-hourly water usage of pigs with an accuracy of 92%. This method could be applied to monitor the occurrence of abnormal events such as disease outbreaks and feed quality, while indicating that video surveillance could be used as a drinking water behavior detection. Zhu et al. [51] analyzed whether it was a drinking behavior based on individual pig identification. Histogram equalization was used to improve the image quality brought about by light. After the maximum entropy global segmentation and morphological noise reduction of the enhanced image, drinking pigs were obtained. Three low-order color moments of drinking pigs were extracted as color features and five geometric features (connection area, contour perimeter, distance of the center of mass from the drinker, external rectangle aspect ratio, hip circularity) as individual features to identify the corresponding pigs. The drinking behavior of each pig was identified by calculating the distance to drink nipple with an accuracy of 90.7%. Qiamei Yang et al. [52] monitored the drinking behavior of pigs by combining machine vision with RFID technology. Pigs’ detections were performed based on threshold segmentation. Then, the learning network of GoogleNet was used to identify the drinking behavior of pigs. The drinking time of pigs, the duration of water consumption was recorded, and the accuracy of drinking behavior recognition reached 92.11%. Some studies also used deep images to quantify drinking and feeding behaviors. Lao et al. [53] processed voids in images by moving average filter and binarized depth images for features extraction. After dividing the pig body into 7 parts, the average depth of each part was calculated for behavioral classification. The feeding behavior and drinking behavior were correctly classified by 97.4% and 92.7%.

In summary, an early system for detecting feeding behavior by electronic sensors. Results of behavioral
Table 2: Image processing and analysis in behavioral detection.

<table>
<thead>
<tr>
<th>Level</th>
<th>Method</th>
<th>Image type</th>
<th>Key technology</th>
<th>Conclusion</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low level: image pre-processing</td>
<td>Noise removal</td>
<td>Depth image</td>
<td>Median filtering denoising and adaptive histogram equalization are used to enhance the image</td>
<td>—</td>
<td>[28]</td>
</tr>
<tr>
<td></td>
<td>Threshold-based segmentation</td>
<td>Grayscale image</td>
<td>Global threshold segmentation based on Otsu method, described by elliptical geometry features</td>
<td>—</td>
<td>[14]</td>
</tr>
<tr>
<td>Mid-level: image segmentation and feature extraction</td>
<td>Threshold-based segmentation</td>
<td>Grayscale image</td>
<td>Threshold segmentation based on Otsu method, joint texture features, and color features description in Camshift algorithm</td>
<td>Better than Camshift algorithm</td>
<td>[30]</td>
</tr>
<tr>
<td></td>
<td>Segmentation based on edge detection</td>
<td>Color image</td>
<td>Segmentation based on dynamic selection of color channels</td>
<td>92.6%</td>
<td>[31]</td>
</tr>
<tr>
<td></td>
<td>Deep learning</td>
<td>RGB image</td>
<td>Segmentation of contours based on Sobel operator with Fourier descriptor</td>
<td>91.7%</td>
<td>[32]</td>
</tr>
<tr>
<td></td>
<td>Deep learning</td>
<td>RGB image</td>
<td>Detection of individual pigs based on improved RestNet</td>
<td>96.4%</td>
<td>[33]</td>
</tr>
<tr>
<td></td>
<td>Deep learning</td>
<td>RGB image</td>
<td>Segmentation of adherent pigs using modified Mask R-CNN network</td>
<td>86.15%</td>
<td>[35]</td>
</tr>
<tr>
<td>Mid-level: image segmentation and feature extraction</td>
<td>Specific algorithm</td>
<td>Color image</td>
<td>Decision tree image segmentation model based on color features using CART algorithm</td>
<td>Better than Otsu et al.'s method</td>
<td>[34]</td>
</tr>
<tr>
<td></td>
<td>Deep learning</td>
<td>Color image</td>
<td>Improving faster R-CNN algorithm to identify childbirth behavior</td>
<td>97% (Zeng, 2018)</td>
<td></td>
</tr>
<tr>
<td>High-level: feature classification and behavior recognition</td>
<td>Deep learning</td>
<td>Color image</td>
<td>Improving AlexNet algorithm to identify estrus behavior</td>
<td>93.33%</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Color image</td>
<td>Behavior classification using MobileNet algorithm</td>
<td>93.4% (mAP)</td>
<td>[38]</td>
</tr>
</tbody>
</table>

Quantification relying on statistics and a priori knowledge were oriented. This led to a more limited application orientation. With the application of 2D and 3D sensors, the enrichment of data features has improved. In particular, the determination of film and television behavior relies on region-specific feature segmentation. Image data better supports this work. The depth image, due to its spatial feature of "depth," facilitates the delineation of more detailed actions such as "sitting and drinking" and "standing and feeding" [53]. However, data-driven detection methods are less studied.

3.1.2. Lying Behavior. Pigs spend more than half of daily time in a lying position [54]. Exploring the physiological and psychological needs of pigs in combination with the lying state and position can help improve animal welfare. The different combination of position, orientation, and connecting relation of various parts of the body is referred to different lying posture of a sow. Among lying postures, the "recumbency" is subdivided into "sternal," "ventral," and "lateral" recumbent positions as shown in Table 4. The combination of behavioral characteristics with the physiological characteristics has been the subject of many studies. However, these researchers have generally been explored under experimental conditions. For example, lying behavior was used as a stationary reference state in the behavioral sequence to obtain a template of pig behavior, which was combined with time series to analyze daily activity. The lying preference of pigs is related to the material structure of the lying area, ambient temperature and humidity, stocking density, outdoor climate stimuli, and other factors [55]. Moreover, during farrowing and lactation of sows, different lying positions affect the life and health of piglets. Real-time monitoring of farrowing room is also important to improve piglet survival rate. Cornou et al. [10] placed an accelerometer below the neck for behavioral data collection. A multivariate dynamic linear model (DLM) is constructed by combining the triaxial observation vector with potential time variables. The variance parameter of the vector is used as the learning feature. Then, the Kalman filter is used to automatically classify the sequences. The results of DLM analysis showed that the probability of lying behavior being correctly classified was 97%, due to the need to observe the state of the mean axis when analyzing behavioral sequences; otherwise, it could lead to misclassification. It can also lead to deviations in axis values when the accelerometer is worn loosely. This is a problem that is difficult to avoid with contact sensors.

In addition to contact sensors, vision sensors could avoid device errors caused by motion variations. Nasirahmadi et al. [57] used machine vision system to score and classify pigs in lying posture. After background subtraction, the
resulting grayscale image is binarized. This is followed by a series of noise reduction operations of watershed transformation, size filtering, and hole filling. Finally, the adjacent frame pixel movement method is used to distinguish between standing and lying behavior. Extracting lying pig contours to build a boundary convex hull feature set. Classification of lying posture scores using SVM with linear kernel with 94.4% accuracy. Feature extraction in pattern recognition relies on pre-processing of images. The model stability is affected when facing the complex and variable light environment. The gathering of pigs or their proximity to the enclosures makes it difficult to segment lying pigs. This is one of the key studies in group pig behavior identification. Depth images avoid color-related problems such as shadows; Kim et al. [58] used Kinect camera to detect lying pigs. In this study, spatial interpolation method was applied to reduce the noise. The background image without the target pig is subtracted from the depth information of the input image. Subsequently, the frame difference image was binarized by the Otsu algorithm. Adjust the overlapped pigs by applying Connected Component Analysis (CCA) before output in an accuracy 80%.

Table 3: Other applications in behavior detection.

<table>
<thead>
<tr>
<th>Field of application</th>
<th>Variants</th>
<th>Detection device</th>
<th>Method</th>
<th>Conclusion</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease and health surveillance</td>
<td>Lameness behavior</td>
<td>Infrared sensor</td>
<td>Comparison of the average temperature of healthy and affected limbs</td>
<td>—</td>
<td>[16]</td>
</tr>
<tr>
<td></td>
<td>Excretion behavior</td>
<td>Depth sensor</td>
<td>Tracking of neck markers enables markerless tracking of other parts</td>
<td>—</td>
<td>[45]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wireless sensor</td>
<td>Monitor excretion time and frequency to detect abnormal behavior of sick pigs</td>
<td>78.38%</td>
<td>[3]</td>
</tr>
<tr>
<td></td>
<td>Drinking behavior</td>
<td>Electronic flow meter</td>
<td>Monitor daily water intake</td>
<td>Predict diarrhea symptoms a day in advance</td>
<td>[4]</td>
</tr>
<tr>
<td></td>
<td>Breathing behavior</td>
<td>Machine vision and RFID</td>
<td>GoogleNet-based monitoring of head drinking behavior and recording of drinking time and duration</td>
<td>92.11%</td>
<td>[46]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine vision</td>
<td>Extraction of target contours from video sequences and calculation of ridge and ventral line fluctuation frequencies</td>
<td>91.1%</td>
<td>[6]</td>
</tr>
<tr>
<td>Social behavior</td>
<td>Aggressive behavior</td>
<td>Infrared sensor</td>
<td>Recording back temperature changes with the intensity and phase of the competition</td>
<td>Classification based on mean motion intensity and booth index using linear discriminant</td>
<td>89.0%</td>
</tr>
<tr>
<td></td>
<td>Aggressive behavior</td>
<td>Machine vision</td>
<td>Automatic detection by activity indices and multi-layer feed-forward neural networks</td>
<td>99.8%</td>
<td>[48]</td>
</tr>
<tr>
<td>Social behavior</td>
<td>Tail-biting behavior</td>
<td>Depth sensor</td>
<td>Extracting active features and detecting attacks using SVM classification</td>
<td>95.7%</td>
<td>[25]</td>
</tr>
<tr>
<td>Lying behavior</td>
<td>Machine vision</td>
<td>Automatic pigtail posture detection</td>
<td>73.9%</td>
<td>[49]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Judging the relationship with ambient temperature by the degree of aggregation</td>
<td>—</td>
<td>[50]</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Definition and description of postures.[56].

<table>
<thead>
<tr>
<th>Behavior level</th>
<th>Posture</th>
<th>Shorter from</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeding</td>
<td>—</td>
<td>—</td>
<td>Lowering head into the feeder. The sow may ingest food, or merely demonstrate its interest in food.</td>
</tr>
<tr>
<td>Drinking</td>
<td>—</td>
<td>—</td>
<td>Touching the nipple water drinker with snout, regardless of whether the sow is drinking or merely manipulating the nipple.</td>
</tr>
<tr>
<td>Sternal recumbency</td>
<td>StR</td>
<td>—</td>
<td>Lying on abdomen/sternum with front and hind legs folded under the body; udder is totally obscured.</td>
</tr>
<tr>
<td>Ventral recumbency</td>
<td>VtR</td>
<td>—</td>
<td>Lying on abdomen/sternum with front legs folded under the body and visible hind legs (right side, left side); udder is partially obscured.</td>
</tr>
<tr>
<td>Lateral recumbency</td>
<td>LlR</td>
<td>—</td>
<td>Lying on either side with all four legs visible (right side, left side); udder is totally visible.</td>
</tr>
</tbody>
</table>
Accelerometers were applied to the lying variation of individual pigs. Specific scoring of lying posture was performed by different axis values. However, group monitoring is difficult to realize with wearable sensors due to the aggregated activity habits of group pigs. Machine vision technology faces new challenges in improving group monitoring problems while overlapping pig segmentation. 3D sensors have been less studied for lying behavior detection.

3.1.3. Locomotion Behavior. Pigs’ locomotion behavior is closely related to health status and behavioral disturbances [22]. The locomotor process consists of continuous sequences of movements such as standing and walking. Researchers analyzed and explored locomotor behavior in different expressions in an experimental environment. It was verified that a relationship exists between locomotor behavior and disease health and breeding applications. The gait variation, gait overlap, and stride length of pigs in abnormal conditions would differ from normal pigs. Lame-ness in pigs varies in visual presentation due to differences in pain levels. Early lameness behavior might be difficult to detect promptly. Therefore, after visual scoring by the gait scoring system, the lameness level is assessed with the sensors such as accelerometers, force plates, and combined with kinematics. The gait scoring system set the score from 0 to 5 (0: normal gait, ability to change direction and accelerate easily; 1: movements appear stiff, abnormal stride length; 2: walking with a large body twist, shortened stride; 3: minimum weight-bearing on affected limb, still able to gallop and trot; 4: reliance on non-affected limb for movements; 5: non-movements.) [9]. Gait scoring system normally uses machine vision as the initial determination of lameness, which is subjective compared to automatic detection by means of intrusive sensors. Significant claudication behavior achieves the expected accuracy. However, there are challenges in the determination of early claudication and regarding the degree of claudication/pain. The sensors are able to quantify behavioral information visually through values such as number of steps. Conte et al. [59] placed single accelerometer (Hobo Pendant G Data Logger, Onset Computer Corporation, Pocasset, MA, USA) on the hind leg. The accelerometer recorded the x-axis acceleration was <0.6 g or >1.4 g while a pig is standing. The result presented that lame sow made a greater number of steps per min than sound sows (P = 0.013). The change in the weight value of each leg was obtained from a force plate (Pacific Industrial Scale Co. Ltd., Richmond, BC, Canada). Two cameras were then used to identify whether each leg was on the correct platform. The frequency of weight transfer (WS), the average percentage of body weight per leg (%BW), the percentage of time WS (%time WS), and other variables were subsequently obtained by calculation. The negative correlations of front and hind legs were calculated to measure asymmetry. The result initiated that the lame sows also had lower contralateral hind leg weight ratios than the sound sows (P = 0.062). The level of lameness was obtained by combining the results of accelerometer, force plate, and kinematic analysis. Lameness is influenced by multiple factors, and unconsidered variables may become important factors influencing the classification of lameness levels, for example, the weight distribution of pigs and clinical signs manifestation. Therefore, automatic detection methods are required to prevent competent factors to the maximum extent. Thompson et al. [60] used two Axivity AX3 logging triaxial accelerometers, which were attached separately to the rear and the neck. After data cleaning, the extracted features were trained on a Support Vector Machine (SVM) classifier with a Radial Basis Function (RBF) kernel. After the optimization operation such as standard sequential minimal and grid-search, the standing and walking behavior obtained F1 scores of 0.77 and 0.84. One of the problems that this paper focuses on is how to correctly classify behaviors in a continuous sequence of behaviors. Sliding discrimination of different behaviors in an active sequence in three consecutive frames is used to avoid smoothing over short transitions. As mentioned before, pattern recognition based on SVM, KNN, Bayes, and other methods could extract features in a targeted manner. Adequate description of the semantics of the behavior is accompanied by feature selectivity.

In addition to intrusion sensors, Kashiha et al. [15] used a top view CCD to automatically quantify locomotion behavior. After eliminating the light effect, the image filtered the background by Otsu method and used morphological closing operator to eliminate noise. After the segmented pigs were fitted with ellipses, the movement status of the pigs was monitored by ellipse parameters. The overall accuracy reached 89.9% in a farming density of 1.23 pig/m². This both achieves the detection of the number of pigs at the same time. The method that simultaneously achieves quantification of movement behavior provides the possibility for automation in commercial farming. Ellipse fitting achieves better tracking and localization of pigs than the way of marking color or id on pigs in previous studies. Segmentation of more accurate pigs in binary images naturally facilitates ellipse fitting. However, the segmentation results are easily affected by the contrast with the background and the ambient light intensity, which could be solved by setting a separate threshold. Furthermore, the change of farming density also affects the locomotion/resting routine of pigs. The emergence of complex tasks such as overlapping pigs near each other due to increased farming density creates difficulties in splitting pigs. Lind et al. quantified pig motility by processing continuous frames. After the fisheye distorted image is rectified by the geometric transformation, a moving object is extracted by the image-subtraction. It is worth mentioning that the threshold is automated selected, while balancing the contrast among the object, the background, and the noise. The noise was referred the primary reason causing the bias, which was suppressed by the median value computed in the last n frames. It has been experimentally shown that the track distance was overestimated by 9%. The image-subtraction method resolved the problem caused by low contrast even the similar colors. However, without automatically updating a reference frame was one source of errors in video frame tracking. Automatically updating the reference may alleviate this problem at the beginning of the process, but the objects would gradually become part of the reference frame over time. In addition to image subtraction, the
optical flow (OF) method is often applied to quantify motion behavior by video [61]. Gronskyte et al. [62] used OF tracking individual pigs with no marks. Through OF preliminary estimated and filtered the continuous frames, the OF vectors would be further estimated by modified angular histograms (MAH). Due to a high correlation present in the MAH vector sets, SVM is used to exclude unnecessary vectors. The MAHs indicated that partial identified OF vectors can describe actual direction and velocity. The proposed method is helpful in detecting slow abnormal behavior. However, monitoring results were susceptible to density effects, and abnormal behavior vectors were averaged out in the process of low-level statistics. And the variation in pig size creates a challenge in tracking individuals. OF was also applied in another study about the process of transporting pigs to the slaughterhouse to avoid health hazards in pigs. The model classified by SVM, which reached 93.5% with the sensitivity and 90% with the specificity.

Due to the difference in spatial characteristics between locomotion behavior and flying behavior, such as height variation, 3D vision sensors are also often used in this application. Stavrakakis et al. [45] used kinematic gait analysis (Vicon T20, Oxford, UK) as a reference system to demonstrate that the Kinect system performed lameness detection. The experiments were separately compared by marker tracking and marker-free tracking. The mark height difference was $8 \pm 1.1$ mm between the Kinect system and the Vicon. The difference of neck marker trajectories amounted $5 \pm 1.5$ mm. Similarly, the mean of vertical displacement amplitudes was $5 \pm 2.8$ mm. It concluded that the Kinect system could detect the lameness by tracking the neck mark. However, the application of marker-free tracking needs further exploration. The Kinect system is mainly applied in locomotion analysis, such as skeleton tracking [63] and action recognition [64]. Skeletal tracking of quadrupeds is more difficult in practical applications, owing to the reason that some skeletal points of the body structure can be obscured. Therefore, innovations based on marker tracking, back curve detection, and other methods. While ensuring the accuracy of locomotion behavior recognition, the spatial features captured by the 3D sensor are also preserved.

In conclusion, locomotion behavior is one of the main forms of daily activity. Lameness behavior is more intuitive and affects the health of pigs. Lameness detection is an early tool to improve animal welfare and has reached maturity in gait analysis techniques, such as Axivity AX3, ADXL330, and LIS3L02DS, which are widely applied to investigate the advanced behavior of pigs in relation between physiology and psychology. The unique identification of the devices also provides a method for individual pig identification and health status tracking. The 2D sensor tracks individuals through continuous frames and allows for abnormal behavior detection for individual frames. This reduces the loss of wearable sensors during pig movement but increases the computational cost and the challenge of tracking and detection due to the similarity of the pig body shape and color. The body shifts in vertical height during pig movement, which is difficult to capture by 2D sensors. 3D sensors apply spatial features such as height changes. The detection of lameness behavior using the pattern of back curve change, the magnitude of vertical height change, etc., but the computational cost in the algorithm needs to be improved.

3.2. Advanced Behavior Patterns. Advanced behaviors are used to infer the current state of the pig by combining the underlying behaviors, such as estrus behavior that causes a sudden decrease in food intake, an increase in activity. However, there are also specific behaviors, such as estrus behavior that can produce climbing and straddling and aggression that can result in violent collisions. Advanced behavior is mainly a relationship between social behavior and psychological profiling of pigs. It helps to understand and explore the potential link between pig behavior and psychology to improve breeding patterns. A farming environment that better fits the physiological and psychological requirements of pigs could be provided. The economic value is increased while promoting welfare farming.

3.2.1. Estrus Behavior. The reproductive performance of sows can affect the production efficiency of pig farms. Estrus detection is critical for improving reproductive performance; it is very important to correctly determine the time of estrus and conduct effective scientific breeding in a timely manner. Sows in estrus will produce a number of symptoms, such as reduced feed intake, increased activity, and swollen vulva with mucus discharge, and reactions such as static reaction and erect ears. Contact IoT devices including accelerometers, posture sensors, and RFID are used in estrus detection in a traditional manner. To reduce the occurrence of stress, infrared imaging technology, visible camera technology, and depth sensor technology are also used in estrus detection is shown in Table 5.

Some researchers based on non-visual sensors for estrus detection. Ostersen et al. [65] proposed an automatic estrus detection method based on visiting boars to sows. The duration was modeled by a multi-process dynamic linear model with first-order Markov probability by reading the individual information of the visiting sow, the start time of the visit, and the duration through an RFID sensor at the visit window. The specificity of the test results was 99.4% better than the detection of 87.4% of sows entering estrus. The number of visits per 6 hours was defined as the boar visit frequency to build a dynamic generalized linear model with a specificity of 98.5%. A method of combining such information sources through Bayesian networks is also proposed in the study as an option for future research. Gies et al. [66] developed a device based on RFID technology that was applied in estrus detection. When boars and sows are kept in separate pens, the sows can recognize the boar scent through the sniffing hole on the detector. A sensor on the top or side wall of the detector records the electronic ear tag number of the sow near the sniff hole and the residence time, while an infrared sensor further verifies whether the sow is in estrus. Hu [67] recorded daily feeding, body weight, and estrus index (the number of times the sow contacted the boar on that day) of sows marked as in estrus through RFID reader on the feeding pen to monitor the condition of sows in estrus to avoid abnormalities through feeding behavior and
body condition changes. Other studies have used wearable devices such as accelerometers and posture sensors for activity frequency as well as activity counts to detect estrus behavior. Cornou and Lundbye-Christensen [68] used time series (values for the three-dimensional axes $x$, $y$, and $z$) of acceleration measurements to automatically classify activity types performed by group-housed sows; the results indicated that multivariate models are well suited to categorize activity types. An example of direct application of the modeling is to detect the onset of estrus, by monitoring the activity level of

<table>
<thead>
<tr>
<th>Field of application</th>
<th>Variants</th>
<th>Detection method</th>
<th>Method</th>
<th>Conclusion</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrus detection</td>
<td>Boar visit</td>
<td>RFID</td>
<td>Dynamic linear modeling of sow visit times with first-order Markov probabilities</td>
<td>87.4%</td>
<td>[65]</td>
</tr>
<tr>
<td></td>
<td>Feeding volume</td>
<td>RFID</td>
<td>Record the length of sow visits and verify estrus with an infrared detector</td>
<td>—</td>
<td>[6]</td>
</tr>
<tr>
<td></td>
<td>Activity level</td>
<td>Triaxial accelerometer (LIS3L02DS)</td>
<td>Detection of estrus by feeding behavior and body condition</td>
<td>—</td>
<td>[67]</td>
</tr>
<tr>
<td></td>
<td>Activity level and mounting behavior</td>
<td>Infrared sensor</td>
<td>Time series on three axes are selected for behavioral classification based on a multi-process Kalman filter to detect outbreaks</td>
<td>Active category: 96%; passive category: 94%</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td>Vaginal hormone</td>
<td>Vaginal resistance</td>
<td>Translating temperature changes generated by behavioral changes into voltage changes to detect estrus</td>
<td>80%</td>
<td>[5]</td>
</tr>
<tr>
<td></td>
<td>Temperature variation</td>
<td>Infrared sensor</td>
<td>Optimizing LSTM to identify estrus behavior using MFO</td>
<td>86.57%</td>
<td>[69]</td>
</tr>
<tr>
<td></td>
<td>Vulvar skin</td>
<td>Infrared sensor</td>
<td>Comparison of impedance changes before and after estrus to detect estrus</td>
<td>$P &lt; 0.01$</td>
<td>[70]</td>
</tr>
<tr>
<td></td>
<td>Ears erect</td>
<td>Visible light sensor</td>
<td>Monitoring the resistivity of the vaginal mucosa to detect estrus</td>
<td>—</td>
<td>[71]</td>
</tr>
<tr>
<td></td>
<td>Mounting behavior</td>
<td>Attitude sensor</td>
<td>Geometric features and optical flow features of the target region are extracted, SVM is used as a classifier to detect the hair situation</td>
<td>90.9%</td>
<td>[74]</td>
</tr>
<tr>
<td></td>
<td>Posture detection</td>
<td>Posture detection</td>
<td>Analysis of vulva surface temperature to detect estrus</td>
<td>—</td>
<td>[73]</td>
</tr>
<tr>
<td></td>
<td>Posture detection</td>
<td>Triaxial accelerometer (MPU6050)</td>
<td>Improved AlexNet recognition of binaural erection time to detect estrus</td>
<td>93.33%</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>Posture detection</td>
<td>Depth sensor</td>
<td>Establishment of SBDA-DL model to detect estrus</td>
<td>92.3%</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td>Analysis of prenatal and postnatal behavior</td>
<td>Visible light sensor</td>
<td>Geometric features and optical flow features of the target region are extracted, SVM is used as a classifier to detect the hair situation</td>
<td>94.5%</td>
<td>[75]</td>
</tr>
<tr>
<td></td>
<td>Posture detection</td>
<td>Depth sensor</td>
<td>Analysis of vulva surface temperature to detect estrus</td>
<td>—</td>
<td>[76]</td>
</tr>
<tr>
<td></td>
<td>Analysis of prenatal and postnatal behavior</td>
<td>Triaxial accelerometer (MPU6050)</td>
<td>Improved AlexNet recognition of binaural erection time to detect estrus</td>
<td>93.33%</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>Posture detection</td>
<td>Depth sensor</td>
<td>Establishment of SBDA-DL model to detect estrus</td>
<td>92.3%</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td>Piglet</td>
<td>Visible light sensor</td>
<td>Improved Mask R-CNN model to detect seizures after region segmentation</td>
<td>94.5%</td>
<td>[75]</td>
</tr>
<tr>
<td></td>
<td>Piglet</td>
<td>Ultrasonic sensor</td>
<td>Setting motion energy thresholds and building action classifiers to detect high-risk behaviors</td>
<td>81.7%</td>
<td>[75]</td>
</tr>
</tbody>
</table>
individual sows. [69]) placed posture sensors on the neck of sows and identified estrus behavior by mounting behavior and volume of activity. MFO was used to optimize the number of first and second hidden layer nerves, maximum training period, block size, and learning rate of Long Short-Term Memory (LSTM). The error rate was 13.43% and the recall rate was 90.63% when 30 minutes was used as the recognition time of estrus behavior.

In addition to the behavioral parameters of activity, feeding behavior, and frequency of boar visits, methods have been proposed to detect estrus behavior based on physical and biological characteristics, such as ear temperature, vaginal temperature, rectal temperature, and hormonal changes in the vagina [79]. Dusza et al. [70] reported that the measure of vaginal impedance helps in predicting the LH surges, which occurred 16.9 h (±17.8) after the first signs of estrus. Řezáč et al. [71, 80] reported that the impedance changes in the vaginal vestibule during peri-estrus are considerably different from those described earlier in the vagina.

In the early estrus detection, the uniqueness of the device identification enables better tracking of individual information, i.e., it can effectively combine individual identification with behavioral identification to achieve precise raising of animals. But at the same time, both the estrus detection device on the feeding pen and the sensor near the boar visit window could only record the individual information of one sow at a time. Contact sensors such as accelerometers, posture sensors, and vaginal resistance devices are prone to costly problems such as loss of equipment and wear and tear when animals move or crash with friction.

Machine vision-based estrus monitoring: Digital infrared thermal imaging (DITI) technology could detect temperature gradients over a surface area due to its ability to detect temperature gradients. The body changes in temperature are easily recognized and detected by DITI and may provide a means for the discovery and monitoring of normal and abnormal physiological events [72]. Sykes et al. [73] used DITI to discriminate between estrus and diestrus phases of the porcine estrous cycle. Vulva thermal images from defined regions of interest were analyzed for maximum (MAX), minimum (MIN), and average (AVG) temperatures. After performed by the gate level model (GLM) procedure, Pearson’s correlations were used to determine relationships among MAX, MIN, and AVG vulva, rectal, and ambient temperatures. Based on the current data, DITI can discriminate between vulva surface temperatures during estrus and diestrus in gilts. Frenson et al. [5] proposed a method of using infrared sensors to obtain daily average activity to determine whether a sow was in heat. The change in behavior is transformed into a change in body temperature and finally reflected in the form of voltage. According to the change in voltage, 4 parameters of exercise volume (average daily activity, standard deviation of average daily activity, minimum, and peak) are set. Research has shown that the accuracy rate of judging the issuance behavior based on the average daily activity amount is 80%. Based on this method, Sun et al. [81] designed a microprocessor-based infrared monitoring system to identify the estrus behavior of sows. The experimental results showed an accuracy of more than 80%.

To improve the accuracy of machine vision in estrus detection, deep learning models have shorter detection time and higher model integration compared to traditional vision models [13, 82], which provide new research directions for this purpose. Zhuang et al. [37] identified the binaural erection behavior of large white sows by setting a binaural erection time threshold in estrus to identify whether they are in estrus or not. A CNN model AlexNet with a 7-layer structure was designed by reducing the number of convolutional layers and the size and number of convolutional kernels in each layer of the AlexNet network structure. From the data, the range of binaural erection time at estrus was found and a threshold of 76 s was set to determine whether the sow was in estrus or not, and the accuracy of the test reached 93.33%. Zhang et al. [38] proposed a SBDA-DL model based on SBDA-DL for monitoring the climbing behavior of sows. Improving on VGG16 + SSD, MobileNet was used as the classification module instead of VGG16. The detection module uses SSD with the last two convolutional layers deleted to reduce the detection frame to 1/4 of the original VGG16 + SSD network structure. Experiments showed an average accuracy of 92.3% in identifying mounting behavior. Zhang et al. (2018) [83] used video analysis for estrus detection with median and mean filtering for noise removal and homomorphic filtering for image enhancement. The geometric features and optical flow features of the target region were used as behavioral detection features, and the average accuracy of the SVM classifier was 90.9% for recognizing mounting behavior. Li et al. [75] proposed a Mask R-CNN based algorithm for pig mounting behavior recognition, using Mask R-CNN network to segment the region of pigs in the detected images, and then based on the mask pixel area of pigs with mounting behavior and non-mounting behavior in the dataset. Then, the defined threshold of the pixel area of the pigs with mounting behavior was obtained. Finally, the method was used to classify the mounting behavior with 94.5% accuracy.

Estrus detection is different from daily behavioral monitoring and is more dependent on complex behavioral sequences of multiple individuals. It relies on more complex sequence analysis involving more than one animal and is therefore more challenging than simple shape or location detection tasks which can be used for other behavioral categories. The information contained in the data collected based on IoT technology is more homogeneous, compared to image data that can convey more behavioral meaning.

3.2.2. Aggressive Behavior. Aggressive behavior revealed that it was a complex and gradual behavior. Initial behaviors are characterized by slower movement (walking) while those occurring in the final phase are vigorous, rapid, and dynamic [83]. Boileau et al. [18] used 1284 thermal images taken from 46 pigs in a controlled test environment. From this thermal window, the average, minimum, and maximum temperature, standard deviation, and coefficient of variation (CV) were analyzed in relation to contest phase. The study showed that peripheral temperature, as recorded by IRT, responded to the intensity and phases of a contest and may allow new insight into the physiological and welfare
outcomes of aggressive behavior. Viazzi et al. [47] developed a method for continuous automated detection of aggressive behavior among pigs by means of image processing. Based on two features, the mean intensity of motion and the occupation index, the Linear Discriminant Analysis was used to classify aggressive interactions in every episode. The accuracy of the system was 89.0%, and the result showed that it was possible to use image analysis to automatically detect aggressive behaviors among pigs. Oczak et al. [48] proposed a method to automatically detect aggressive behavior in pigs by using an activity index and a multi-layer feed-forward neural network. Five features of activity index were calculated on the recorded videos (average, maximum, minimum, sum, and variance) during 14 time periods. The results revealed that ANNs, calculated on 241 s time intervals, classified high aggression events with an accuracy of 99.8% whereas medium aggression events were classified with an accuracy of 99.2%. Lee et al. [25] extracted activity features (minimum, maximum, average, standard deviation of velocity, and distance between the pigs) from the Kinect depth information in order to detect aggressive activity. The method employed two binary-classifier support vector machines in a hierarchical manner with a better accuracy, and the accuracies of detection and classification were over 95.7% and 90.2%, respectively.

Since the attack occurs in a standing position, the depth information of the individual is more significant. Depth images can identify pigs as standing or lying based on depth criteria in the data processing stage to facilitate target tracking. Compared with visible light cameras, depth sensors can also circumvent the problems of target occlusion, light, and texture transformation, which are less used in practice.

3.2.3. Delivery and Lactation Behavior. Sows need to receive special management at different stages after estrus. Negative energy balance in sows during gestation increases the probability of abortion, so the monitoring of body condition and abnormal behavior of sows during gestation needs to be enhanced. The traditional farming method relies on manual prediction of farrowing time for sows approaching farrowing, which is prone to inefficiency and piglet mortality. Some researchers have analyzed the clinical signs of sows at the time of farrowing, such as rising body temperature, rising respiration and heart rate, udder initiation of milk production, and swelling of the pubic area. Also, some investigators have automatically determined the time of parturition by monitoring the sow’s resident litter behavior, etc. Under non-human interference conditions, the maternal ability of the sow determines the survival rate and daily weight gain of the piglets during lactation [76]. Thus, behavioral performance during the prenatal gestation and prodromal periods and during postpartum lactation can also affect pig production.

Existing methods for analysis of prenatal and postnatal behaviors mainly include three-axis acceleration-based detection methods, ultrasonic detection methods, and machine vision processing methods. Yan et al. [76] used MPU 6050 sensor for recognition of lactating sow posture, and Haar wavelet extraction features were used to reconstruct the basic profile of acceleration curve. The behavior was recognized using support vector machine after combining the posture feature information. A subsequent study [84] found that the high-risk movement of lactating sows switching from standing and sitting to lying down was the main cause of death in preweaned piglets. The frequency of movement occurrence and transformation was determined as the maternal index of sows, and the data from MPU 6050 sensors were used to set the movement energy threshold and construct a high-risk movement classifier. The accuracy of the results reached 81.7%, providing data basis for achieving scientific evaluation of maternal ability and reproduction. Zhang et al. [77] used ultrasonic sensors to collect distance information on the amounts of head, back, and tail activity of sows before farrowing. The correct rate of behavior detection and classification using K-means clustering algorithm reached 90.47%. Liu et al. [78] analyzed sow farrowing video image features for motion target detection based on improved single Gaussian model with background subtraction. Target recognition of piglets was performed based on the color and area features of motion regions. Xue et al. [28] proposed an improved Faster R-CNN based on lactating sow pose recognition algorithm using deep video images as the data source and introduced the residual structure into the ZF network. The ZF-D2R network was designed to improve the accuracy and maintain the real-time performance, and then the Center Loss supervised signal was introduced into the Faster R-CNN training to enhance the cohesiveness of intra-class features. The recognition speed is 0.058 s/frame, which is 0.034 s faster than the VGG16 network.

The method based on the IoT technology can avoid the effects of diurnal light changes and hot light illumination, but at the same time it can lead to misjudgment of behavioral data because of the variability among different individual sows. Although the visual processing method is affected by the light problem, the introduction of depth data can relatively overcome the impact of this problem. The performance of the model is guaranteed while bringing the problem of a larger model. Further consideration needs to be given to compress the model for better migration deployment to embedded systems.

3.2.4. Tail Biting Behavior. Tail biting occurs mostly in commercial indoor farming, usually due to high farming densities, poor farming environments, inadequate ventilation, and poorer feed quality consumption rates [85–89], as shown in Table 3. In addition to direct damage caused to the tail and caudal region of the pig, it may also lead to blood loss or traumatic death, and wound infection can develop and spread throughout the body, causing abscesses and sepsis leading to paralysis and death, thus affecting animal welfare and production efficiency. D’Eath et al. [49] used Time-of-Flight 3D cameras and machine vision algorithms to automate the measurement of pig tail posture. Validation of the 3D algorithm found an accuracy of 73.9% at detecting low vs. not low tails. Yuzhi et al. [61] used a camera to automatically monitor changes in activity levels prior to outbreaks of tail-biting behavior. The rate of change of pixel blocks was estimated by the optical flow method, which calculates the image brightness of blocks between two consecutive frames as well as the spatial average of the blocks. The
result indicated that pigs increased activity level 3 d before the first outbreak of tail biting ($P < 0.01$).

In summary, tail biting behavior can be monitored by tail posture and behavioral changes. The tail is smaller relative to the overall target, which is difficult to measure by sensors such as accelerometers, while visual sensors can enhance attention to localized parts. However, for monitoring the activity level, the time of tail biting outbreak can be predicted by analyzing the activity trend of each behavior over a long period of time, and the choice of sensors can be chosen from accelerator, 2D, and 3D camera.

4. Discussion

(i) In addition to the behaviors mentioned in the previous section, there are also detections on breathing behavior and excretion behavior. Zhu et al. [3] proposed an automatic monitoring system to identify suspected diseased pigs by excretion behavior. The system was based on ARM platform to monitor the excretion time and frequency of fattening York pigs, and the excretion line was judged as abnormal behavior when it was too frequent. The test results showed that the correct detection rate of sick pigs was 78.38%. Ji et al. [6] captured the spine and abdominal line contours of pigs to warn sick pigs suffering from shortness of breath. The target contour was obtained by binarizing the acquired video sequence frames, opening the operation, removing orphan noise, and smoothing the target edges. The range of the fluctuation of the shape center is tracked to identify pigs standing still and moving, and the location of the ridge line is determined based on the shape center. The accuracy of automatic video detection of sharply breathing pigs was 91.1%. In machine vision-based disease surveillance, chronic diseases do not initially show obvious signs, such as hoof-limb infections. Target diseases are hard to be monitored by remote vision systems because of their small size relative to the animal body but can be identified by detection of specific parts [16]. Behavioral information such as drinking, feeding, and excretion of multiple target animals can be detected simultaneously, especially after processing by deep learning algorithms, which can extract deep feature information from the input data to improve recognition accuracy. Some specific sensors such as electronic flow meter, vaginal resistance, pressure pad, and other sensors are also used in behavior for detection but there are some problems. For example, electronic flow meters need to be combined with RFID technology if they track the drinking behavior of individual pigs. Vaginal resistance can easily cause stress during measurement. Incorrect treatment could also cause disease in pigs or detachment of the sensor to the detriment of animal welfare. In addition, the main reason affecting the invasive sensors including accelerometers is the limited battery technology. The main purpose of behavior monitoring is to be real-time and accurate for timely prevention. Real-time data transmission requires a power supply which requires a long-lasting battery life. The high frequency load of the device also reduces the service life.

(ii) Advanced behavior is also a way of expressing social behavior. When an unfamiliar counterpart joins the group, pigs will initiate interactions to gain hierarchical status and access to resources, such as space and feed [90–95], which can lead to excessive behaviors such as aggression and tail biting among individuals. Interactions between behaviors are sometimes negative, such as causing skin damage leading to infection, which can be fatal in extreme cases [96]. These negative effects have either direct impacts or costs of recovery from injuries and associated infections [97]. Over-aggressive behavior can also lead to economic losses and the disadvantaged party can suffer from problems such as slow growth and weight loss due to lack of access to adequate food; therefore, monitoring of over-aggressive behavior between animals is key to improve animal welfare in intensive farming to solve disease health problems [98]. There are also changes in social behavior that can be used as indicators for monitoring the psychological or physiological health of animals; for example, the aggregation behavior of pigs can indicate whether the surrounding environment is suitable for growth [99–101], while high ambient temperature influences lying behavior with pigs spreading out [50]. Social behavior can reflect changes in animal psychology and deep emotions, which can be affected to some extent by human intervention. The application of machine vision technology can not only record the social behavior of animals in the form of video, but also can replace part of the farm operators’ behavior observation work.

(iii) With the advent of 5G technology, edge computing devices can be centrally managed and automated through cloud infrastructures. Large amounts of data are stored while the required information can be dynamically accessed and recalled [102]. The amount of behavioral data collected is large and strongly correlated with each other. It tends to have temporal continuity and regularity. Edge computing technology is applied in which historical data can be retained and visualized for analysis, facilitating accurate management. However, the long transmission distance of traditional wireless communication systems leads to high latency. Edge computing technology shortens the distance between devices, thus reducing latency and improving the reliability of data transmission.

5. Conclusion and Prospect

This paper reviews the methods of behavior analysis from target extraction, behavior classification, and behavior monitoring applications. The monitoring devices used gradually
shift from contact to non-contact devices to reduce the psychological stress on the animals, while machine vision monitoring as non-contact monitoring can be of great help to improve animal welfare. Especially in behavior monitoring, the traditional wearable monitoring equipment with the movement of the animal will have data bias, equipment damage, and other problems. Video monitoring avoids these problems well by target tracking from a distance, but machine vision monitoring research is still in its infancy and faces more challenges.

Accuracy of vision system detection: A single vision system is mostly affected by diurnal shift, bright and dark light, and system accuracy, which affects the later behavior detection, and more methodological investigations are needed to compensate for the problems in the vision system. Reliability of individual recognition algorithms. The information contained in animal behavior needs to be explored more. In addition to behavior detection itself, individual recognition of animals is also very important, especially multi-target individual recognition is the focus of future research. In the refinement of farming, the management of individuals in the group breeding method has put forward more stringent requirements, requiring more innovation and integration in target detection, accuracy of target tracking algorithms, and arithmetic power.

Applicability of behavioral detection algorithms. The application direction of behavior detection mostly stays in the research of general behaviors such as standing and lying down. Some behavioral detection such as estrus behavior and aggressive behavior with more complex judgment conditions also play a key role in production but the research results are still few and should be the direction of further research.

This paper reviews pig target extraction, behavioral image analysis, and behavioral detection methods based on machine vision technology, presents the problems of single vision system, difficult fusion of individual information, and applicability of behavioral detection algorithms, and proposes optimization strategies for these problems.

The use of multi-sensor vision system: the combination of two or more vision sensors to make up for the light transformation, 2D images lack of depth spatial information, and other problems to achieve all-day detection; The use of multiple sensors to collect data: video monitoring combined with sound monitoring, individual information collection equipment integration, and other diversified methods, but at the same time the data fusion of diversified sensors, process synchronization, and application to production after the device integration and other issues also come up. Machine vision for animal behavior motion detection in many types of research objects, complex application scenarios, variable environmental variables (temperature, humidity, etc.) will also become a challenge for machine vision algorithms, but machine vision as one of the methods of non-contact detection for improving animal welfare still has the potential to be explored.

Conflicts of 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.

Authors’ Contributions

Hang Zhang is the co-first author.

Acknowledgments

This work was supported by the Tianjin Municipal Key R&D Program (grant number 20VFZCSN00220), Central Government Guides Local Science and Technology Development Projects (grant number 21ZYCGSN00590), Inner Mongolia Autonomous Region Key R&D Program (grant number 2020GG0068), and Tianjin Municipal Postgraduate Research and Innovation Project (grant number 2021XY026).

References

[33] Song, Pig Target Based on Deep Learning Inspection and Status Analysis, 2019.


[82] Z. Zhang, Research on detection of cow estrus information based on video analysis, Northwest A&F University, Xianyang, 2018.


