

Article

Thermal Imaging-Based Detection of Sow and Piglet Behaviors Under Imbalanced Conditions in Commercial Farrowing Pens

Chia-Ying Chang

Bachelor of Program in Scientific Agriculture, National Pingtung University of Science and Technology,
Pingtung 91201, Taiwan; csiefly@mail.npust.edu.tw

Featured Application

This study provides a practical thermal imaging and deep learning framework for automated sow and piglet monitoring in commercial farrowing pens. The proposed approach can support nighttime livestock surveillance, piglet safety monitoring, and precision livestock farming under real farm conditions with severe class imbalance and animal occlusion.

Abstract

Monitoring sow and piglet behaviors during farrowing is important for reducing piglet mortality and improving animal welfare in commercial swine production systems. However, accurate behavioral monitoring remains challenging because neonatal piglets are small, frequently occluded by the sow, and difficult to detect under low-light conditions, particularly for black-coated piglets during nighttime observation. In this study, thermal imaging combined with RGB imaging from a conventional visible-light video camera and deep learning-based object detection was applied for automated livestock monitoring under practical farrowing conditions. The dataset contained diverse sow postures and piglet-related activities acquired in a commercial farrowing-house environment. In addition, the dataset exhibited substantial class imbalance, with piglet-related behaviors representing only a small proportion of the total annotations. Three lightweight YOLO-based detection models were evaluated for livestock monitoring performance under commercial farm conditions. Results showed that lightweight detectors provided stable overall detection performance for general sow behavior monitoring, while anchor-free detection models demonstrated improved sensitivity for small piglet detection under highly imbalanced conditions. The findings demonstrate the feasibility of thermal imaging for continuous farrowing monitoring under variable lighting and occlusion conditions and highlight the importance of evaluating detection systems based on their ability to identify small neonatal targets relevant to practical farm management. This study provides useful insights for the development of automated precision livestock monitoring systems for commercial farrowing operations.

Academic Editor: Douglas
O'Shaughnessy

Received: 21 May 2026
Revised: 26 June 2026
Accepted: 27 June 2026
Published: 30 June 2026

Copyright: © 2026 by the author.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: YOLO object detection; livestock behavior detection; small object recognition; precision livestock farming

1. Introduction

According to the Food and Agriculture Organization (FAO), livestock contributes significantly to global food security but also faces challenges related to productivity, animal welfare, and labor shortages [1]. Precision Livestock Farming (PLF) has emerged as a data-driven approach to the continuous monitoring of animals using sensing technologies in conjunction with information systems and artificial intelligence [2,3]. Rapid population growth has intensified the demand for sustainable and efficient livestock production systems. Recent bibliometric analyses have also demonstrated a rapidly increasing research interest in PLF technologies and AI-based livestock monitoring systems worldwide [4].

An accurate indication of animal welfare requires the close monitoring of pig behavior, particularly in farrowing environments [5], where maternal posture and piglet activity are a direct indication of neonatal survival based on the risk of piglets being crushed. Conventional monitoring relies heavily on manual observation by farm personnel, which is labor-intensive, time-consuming, and prone to subjective interpretation [6].

Recent advances in computer vision and deep learning have enabled the use of image and video data to automate the analysis of animal behavior [7,8]. Convolutional neural networks (CNNs) and object detection models have been applied to livestock monitoring tasks, including cattle identification, analysis of poultry activity, and the recognition of pig posture [9,10]. Continuous, non-invasive monitoring can alleviate the need for human interventions and enhance the depth of data used to manage decision-making.

Most early livestock detection systems adopted anchor-based object detection frameworks such as Faster R-CNN, SSD [11,12], and particularly YOLOv2, which provides a practical balance between inference speed and detection accuracy [13]. Nevertheless, anchor-based detectors are sensitive to anchor box design and small-scale object overlap, which is a common occurrence when neonatal piglets cluster around the sow.

Anchor-free object detection models have been shown to improve detection performance under close-quarter conditions. The decoupled head design and advanced label assignment strategy of YOLOX promote general object detection by reducing reliance on predefined anchor boxes [14,15]. Nonetheless, this model has seldom been applied to thermal-image monitoring of livestock.

Thermal imaging is highly robust under low-light conditions and nighttime operations [16]. Unlike RGB cameras, thermal devices provide stable contrast based on variations in temperature, when poor illumination precludes reliance on conventional imaging. Previous studies have demonstrated the feasibility of thermal imaging for assessing the health and posture of pigs [17,18]. This is among the first studies to integrate thermal imaging with lightweight, anchor-free object detectors for multi-class pig behavior recognition.

Severe class imbalance is another practical challenge in livestock detection. In farrowing environments, the primary focus is on sow posture, with relatively little attention paid to critical neonatal interactions. This imbalance can have a profound effect on detector optimization and reduce minority-class recall, particularly when dealing with small objects presenting indistinct object boundaries.

The objective of this study was to determine the suitability of different YOLO-based object detectors for thermal-image-based livestock monitoring under practical farrowing-house conditions. The objective of this study was to determine the suitability of different YOLO-based object detectors for thermal-image-based livestock monitoring under practical farrowing-house conditions. Specifically, the performance of YOLOv2, YOLOv4-Tiny, and YOLOX was compared for six predefined detection categories comprising four sow postures (lying down, standing, sitting, and prone) and two piglet-related activity categories (unsuckling and suckling). Model performance was evaluated using average precision (AP), mean average precision (mAP), precision, recall, and class-wise sensitivity.

The remainder of this paper is organized as follows: Section 2 describes the dataset construction, preprocessing procedures, presents the model architectures and evaluation methods. Section 3 reports the experimental results. Section 4 discusses the findings and Section 5 concludes the study.

2. Materials and Methods

2.1. Dataset Acquisition and Annotation

The dataset used in this study was collected from a commercial farrowing house using a single fixed dual-spectrum (RGB–thermal) camera mounted above an individual pig pen. Each farrowing pen was monitored individually, and the same camera was repositioned above each pen during the corresponding recording period. Thermal imaging was adopted to enable continuous livestock monitoring under practical farming environments, including low-light and nighttime conditions where conventional RGB imaging may exhibit reduced visibility, particularly for black-coated piglets. In addition to thermal images, paired RGB images were simultaneously collected for comparison of imaging characteristics under different illumination conditions.

This study used data from the Southern Branch of the Livestock Research Institute. Each farrowing pen measured 221 cm × 196 cm, as shown in Figure 1. A Dahua DHI-TPC-BF1241-T-S2 dual-spectrum camera (Dahua Technology Co., Ltd., Hangzhou, China) was installed 389 cm above the floor to simultaneously acquire RGB and thermal infrared images. The RGB camera provided a field of view (FOV) of 71.2° × 52.0°, while the thermal imaging camera provided an FOV of 50.7° × 37.8°. The resulting imaging area covered approximately 388.57 cm × 190 cm. The flooring consisted of silver-colored stainless-steel slatted panels commonly used in commercial farrowing systems.



Figure 1. Schematic diagram of livestock housing environment and pen dimensions.

The experimental animals consisted of two Landrace (L) sows and four Taiwan Indigenous Black Pig sows. The two Landrace sows produced 10 and 7 piglets, respectively, while the four Taiwan Indigenous Black Pig sows produced 6, 10, 6, and 5 piglets, resulting in a total of 44 piglets included in this study.

For each farrowing event, continuous image acquisition started 24 h before the birth of the first piglet and continued until 24 h after complete placental expulsion. Therefore, each sow was monitored continuously throughout the periparturient period using the same camera configuration.

The farrowing house was a conventional open-sided facility without evaporative cooling pads or enclosed environmental control systems, relying on natural ventilation

and daylight. Average barn temperatures measured in May were 28.54 °C at noon and 25.56 °C at night.

Thermal observations indicated that the body surface temperatures of both sows and piglets remained above 35 °C during both daytime and nighttime periods, generating a clear thermal contrast against the surrounding environment and supporting reliable animal detection under varying illumination conditions.

The acquired dataset included multiple practical farm scenarios, such as different sow postures, piglet activities, overlapping animals, partial occlusions, and varying animal densities during both daytime and nighttime periods. To establish the detection dataset, all thermal images were manually annotated using bounding boxes according to six predefined behavioral categories. The annotation process was performed using image-level labeling to identify the spatial locations of sows and piglets under realistic farm conditions. Particular attention was given to small piglet targets and partially occluded animals to improve annotation consistency and detection robustness. After annotation, the dataset was divided into training, validation, and testing subsets using a fixed image-level partitioning strategy to ensure consistency across all experiments.

2.1.1. Thermal Imaging Data Acquisition

Thermal imaging was chosen to ensure robust data acquisition under variable illumination conditions, including nighttime and low-light situations. Unlike RGB cameras, thermal sensors provide consistent image quality, regardless of ambient lighting.

Figure 2 illustrates representative imaging conditions used during dataset acquisition. The dataset includes pairs of RGB and thermal images collected from white pigs and black pigs under daytime and nighttime conditions. In the figure, the image on the left is an RGB image, and the image on the right is the corresponding thermal image.

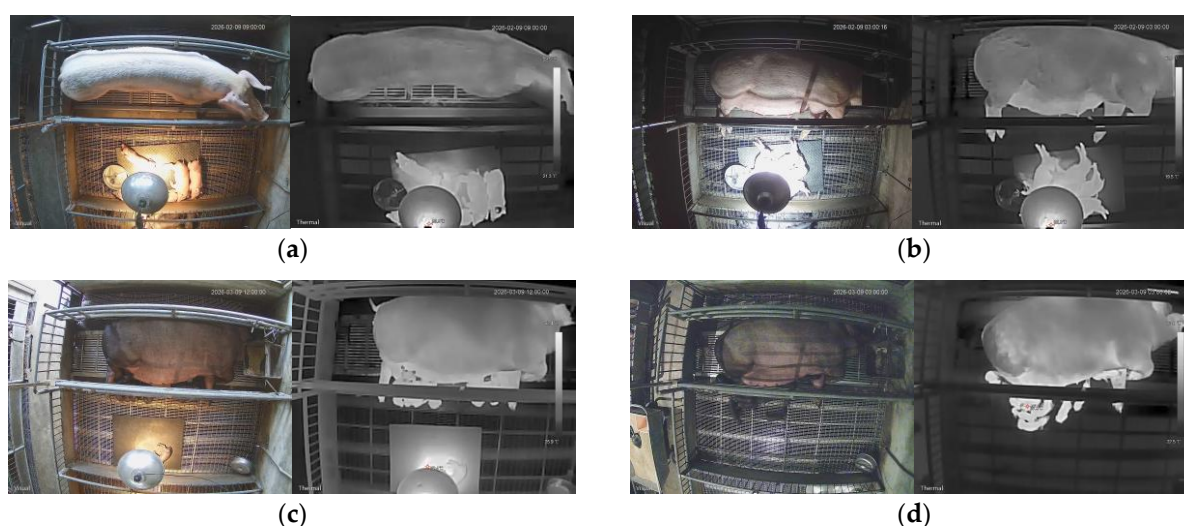


Figure 2. Representative pairs of RGB (left) with corresponding thermal images (right): (a) white pig during daytime, (b) white pig during nighttime, (c) black pig during daytime, and (d) black pig during nighttime.

During the daytime, RGB images generally provided clear visual information for both white and black pigs; however, the quality of RGB images captured at night depended strongly on coat color and contrast against the surrounding environment, such that white pigs remained visible and black pigs were difficult to distinguish against the dark background. Thus, we also employed thermal images, the contrast of which is based on body temperature rather than visible light. The thermal images provided clear indica-

tions of sow posture, piglet location, and sow–piglet interactions under low-light conditions, regardless of coat color.

The resulting dataset differentiated coat color, lighting conditions, posture, and animal interactions, reflecting the conditions encountered in commercial livestock environments, which was invaluable in characterizing the performance of the detection models outside controlled laboratory settings. The recorded images included overlapping animals, partial occlusions, and groups of varying densities, all of which are common in real-world farrowing pens. Note that all images were captured at a resolution of 640×512 pixels and stored in standard formats compatible with deep learning workflows.

2.1.2. Behavior Categories

Based on established veterinary standards and previous livestock behavior studies, six representative pig behaviors were selected as target classes [3,4]:

Sow-posture categories:

1. Lying down—a sow resting in a lateral or dorsal recumbent position.
2. Standing—a sow maintaining an upright posture supported by all four limbs.
3. Sitting—a sow resting on the hindquarters with the forelimbs extended.
4. Prone—a sow lying in a ventral position with the abdomen contacting the floor.

Piglet-related categories:

5. Unsuckling—piglets present within the farrowing pen but not actively nursing.
6. Suckling—piglets actively nursing from the sow.

The sow-posture categories were selected because they represent the most frequently observed postural states associated with animal welfare, health status, and maternal behavior in commercial farrowing systems. The piglet-related categories were included to facilitate monitoring of neonatal activity and nursing behavior, which are important indicators of piglet survival and growth performance.

Although the acquired thermal images contained diverse posture variations and piglet activities under practical farm conditions, the behavioral categories were intentionally simplified into six representative classes. This simplification was adopted to ensure annotation consistency, maintain sufficient sample numbers for each category, and facilitate objective comparison among different object detection architectures. More detailed ethological classifications and social interaction analyses were beyond the scope of the present study and will be investigated in future work.

2.1.3. Annotation Procedure

Ground truth annotations were created using MATLAB's Image Labeler tool. All images were manually annotated using the Image Labeler application provided in MATLAB R2024a (MathWorks Inc., Natick, MA, USA) to establish the ground-truth dataset required for object detection model training. Each image was manually annotated by trained personnel with domain knowledge in animal behavior to ensure labeling accuracy and consistency. Bounding boxes were drawn tightly around each pig instance, and corresponding behavior labels were assigned according to the predefined categories.

To minimize annotation bias, ambiguous cases were cross-checked by multiple annotators with consensus-based corrections. Images with severe motion blurring or unrecognizable behaviors were excluded from the dataset to preserve annotation quality.

All annotations were exported in MATLAB ground truth format (.mat), containing image paths and bounding box information for each behavior class. This structured format facilitated seamless integration with the YOLO-based training and evaluation pipelines.

2.1.4. Dataset Split and Statistics

The annotated dataset was randomly divided into training (80%) and testing (20%) subsets at the image level to avoid data leakage. Class distribution across both subsets was preserved to ensure a fair and reliable evaluation of model performance.

Table 1 summarizes the number of annotated instances for each behavior category in the dataset.

Table 1. Distribution of the six target categories used for thermal-image-based livestock behavior detection.

Animal Group	Behavioral Category	Images	Behavior Class
Sow	Lying down	2631	2631
	Standing	244	244
	Prone	41	41
	Sitting	72	72
Piglets	Unsuckling	1416	4539
	Suckling	2528	14,583

Note that this dataset exhibited a class imbalance representative of livestock behavior in the real world. This study opted not to artificially balance the dataset. We sought to preserve the original distribution in order to evaluate model robustness under realistic deployment conditions.

The dataset construction process followed recommendations reported in recent surveys of computer-vision datasets for precision livestock farming, emphasizing annotation consistency and reproducibility [19].

2.1.5. Dataset Availability and Reproducibility

The dataset was obtained under commercial farrowing-house conditions and is not publicly available due to farm privacy and management restrictions. However, the annotation protocol, behavioral category definitions, dataset partitioning strategy, and evaluation procedures are fully described in this study to facilitate experimental reproducibility. The standardized annotation workflow also supports future dataset expansion and cross-study comparisons in livestock monitoring applications.

2.2. Deep Learning-Based Detection Framework

Three representative one-stage object detection models, namely YOLOv2, YOLOv4-Tiny, and YOLOX, were employed in this study to evaluate livestock behavior detection performance under practical farrowing conditions. These models were selected because they represent different object detection strategies and computational complexities commonly used in real-time monitoring applications.

YOLOv2 was included as a conventional anchor-based detector with stable detection capability and a relatively simple architecture. YOLOv4-Tiny was selected as a lightweight version of YOLOv4 designed for fast inference and real-time deployment in resource-limited environments. YOLOX was adopted as an anchor-free detector to evaluate its effectiveness in detecting small piglet targets under severe class imbalance and partial occlusion conditions.

All models were trained using the same annotated dataset and identical training-validation-testing partitions to ensure fair comparison among architectures. During training, the models learned to predict object locations and behavior classes directly from thermal images. Model performance was subsequently evaluated using standard object de-

tection metrics, including precision, recall, average precision (AP), mean average precision (mAP), and confusion matrix analysis.

2.2.1. Overall Framework

Thermal images acquired from pig pens were manually annotated according to the six predefined target categories. The annotated dataset was subsequently divided into training, validation, and test subsets. YOLOv2, YOLOv4-Tiny, and YOLOX were independently trained using the same data partitions.

Following model training, object detection performance was assessed using average precision (AP), mean average precision (mAP), precision, recall, confusion matrices, and small-object sensitivity analysis.

2.2.2. YOLO-Based Object Detection Architectures

Three representative one-stage object detection models, namely YOLOv2, YOLOv4-Tiny, and YOLOX, were employed in this study to evaluate livestock behavior detection performance under practical farrowing conditions. These models were selected because they represent different design strategies for real-time object detection and provide varying balances between computational efficiency and small-object detection capability.

YOLOv2 is a conventional anchor-based detector that formulates object detection as a unified regression problem by simultaneously predicting object locations and class probabilities from grid-based image regions. In this study, the network was fine-tuned using thermal livestock images to detect sow postures and piglet-related behaviors under complex farm conditions. Multiple anchor scales were configured to accommodate the substantial size differences between adult sows and neonatal piglets.

YOLOv4-Tiny is a lightweight version of the YOLOv4 framework designed for efficient real-time inference. The model incorporates Cross-Stage Partial (CSP) connections and multi-scale feature extraction mechanisms to improve feature reuse and computational efficiency. Due to its lightweight architecture, YOLOv4-Tiny is suitable for edge-based or embedded livestock monitoring systems where computational resources may be limited.

YOLOX is an anchor-free object detector that predicts object centers directly without predefined anchor boxes. The architecture employs decoupled classification and localization branches together with multi-scale feature fusion to improve the representation of small objects. This design is advantageous for detecting neonatal piglets because piglets frequently appear as small, sparse, and partially occluded targets in farrowing environments.

2.2.3. Model Training Strategy

All detection models were trained using GPU acceleration under identical dataset partitioning conditions to ensure fair comparison among architectures.

For YOLOv2 training, stochastic gradient descent with momentum (SGDM) was employed together with transfer learning from pre-trained backbone weights. The training configuration included an input image size of 448×448 pixels, a mini-batch size of 32, an initial learning rate of 0.003, and 600 training epochs. Training samples were randomly shuffled at each epoch to reduce overfitting and improve model generalization.

YOLOv4-Tiny was trained using transfer learning and GPU-based optimization with automatic image resizing and augmentation procedures. A mini-batch size of 8 and 80 training epochs were used. Due to the use of pre-trained feature representations and a lightweight architecture, YOLOv4-Tiny demonstrated faster convergence compared with YOLOv2.

YOLOX training was performed within an anchor-free detection framework using separate classification and localization optimization branches. Multi-scale feature fusion and automatic label assignment strategies were adopted to improve small-object learning performance. During training, intermediate checkpoint models were periodically saved for monitoring model convergence and minimizing overfitting risk.

All model development, training, and evaluation procedures were implemented in MATLAB R2024a using the Deep Learning Toolbox and Computer Vision Toolbox (MathWorks Inc., Natick, MA, USA). The YOLOv2, YOLOv4-Tiny, and YOLOX object detection architectures were trained under identical dataset partitions and training conditions to ensure a fair comparison.

Training and inference were performed on a workstation equipped with an NVIDIA GeForce RTX 5080 Laptop GPU (NVIDIA Corporation, Santa Clara, CA, USA) using GPU acceleration to improve computational efficiency.

3. Results

3.1. Overall Detection Performance

The primary objective of this study was to compare the detection performance of YOLOv2, YOLOv4-Tiny, and YOLOX for thermal-image-based monitoring of sow and piglet behaviors under practical livestock production conditions. Therefore, the results first focus on the overall detection accuracy of each model, followed by analyses of class imbalance effects and small-object sensitivity.

Table 2 summarizes the overall detection performance of the three models. YOLOv4-Tiny achieved the highest mAP of 0.8767 with a precision of 0.88 and a recall of 0.86, indicating a good balance of localization accuracy and detection completeness across the six behavioral categories. YOLOX achieved a moderate mAP of 0.7772 with the highest precision of 0.9335, indicating fewer false-positive detections. Limited recall of 0.4536 suggested the incomplete retrieval of minority-class targets under highly imbalanced data conditions. YOLOv2 returned the lowest mAP of 0.4760 and the lowest recall of 0.2524, despite high precision. These results indicate the efficacy of lightweight anchor-based detection for livestock behavior monitoring under strong dataset imbalance.

Table 2. Overall performance of three detection models.

Model	mAP	Precision	Recall
Yolo v2	0.4760	0.9401	0.2524
Yolo v4-Tiny	0.8767	0.8800	0.8600
Yolo x	0.7772	0.9335	0.4536

3.2. Dataset Distribution and Class Imbalance

As shown in Table 3, the dataset used in this study exhibited substantial class imbalance. Based on the dataset split used for model training and evaluation, the prone class was the dominant category, with 14,583 annotated instances, accounting for 65.69% of the total dataset. This is followed by *lying down* with 4539 instances (20.53%) and *sitting* with 2630 instances (11.90%).

Minority classes were severely underrepresented, with only 244 annotated instances of *standing* (1.10%), 41 instances of *piglets* (0.19%), and 72 instances of *suckling* (0.33%). Imbalances of this severity pose serious challenges for detector optimization, as dominant postural classes contribute disproportionately to model learning, with minority classes contributing very little.

The ground-truth distribution revealed that some minority classes appeared in relatively few images. However, these classes are important because they often occur in criti-

cal situations when the piglets are in close proximity to the sow, with body overlap, close contact, and partial occlusion making detection more difficult, particularly in thermal images.

Table 3. Dataset composition.

Class	Train	Validation	Test	Total	Percentage (%)
Sitting	1834	395	401	2630	11.90
Lying_down	3271	644	624	4539	20.53
Piglets	20	10	11	41	0.19
Suckling	53	10	9	72	0.33
Standing	182	33	29	244	1.10
Prone	10,118	2188	2277	14,583	65.96

This imbalance is particularly important for small-object detection. Piglets occupy only a small part of the image and are often surrounded or partly hidden by larger body structures. Detection performance for the *piglet* and *suckling* classes is an important indicator of the reliability of the model under practical livestock monitoring conditions.

3.3. Small Object Sensitivity Analysis

A small-object sensitivity analysis was conducted for piglets and suckling classes.

The AP values for sparse neonatal target detection (piglets) were as follows: YOLOX (0.90), YOLOv4-Tiny (0.84), and YOLOv2 (0.24), as shown in Figure 3. This suggests that the anchor-free detection mechanism improves localization flexibility for sparse targets.

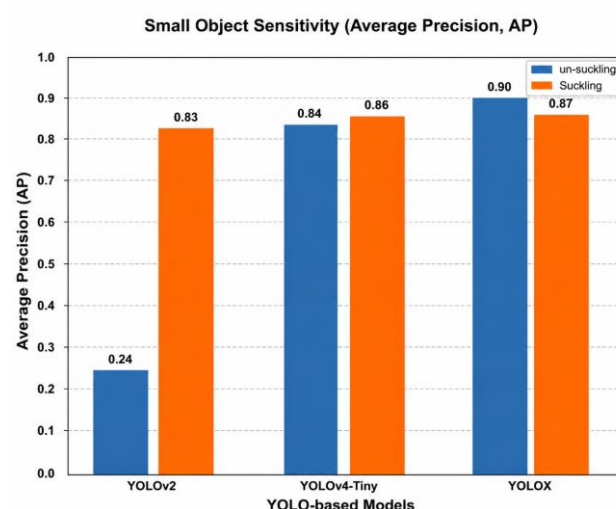


Figure 3. Comparison of YOLOv2, YOLOv4-Tiny, and YOLOX detectors in terms of average precision (AP) for small-object classes (unsuckling (blue bar) and suckling (orange bar)).

The AP values for suckling were as follows: YOLOv4-Tiny (0.86), YOLOX (0.875), and YOLOv2 (0.83). These results indicate that lightweight anchor-based architectures are more effective for close-contact interactions.

3.4. Result Summary

In experiments, YOLOv4-Tiny achieved the best overall detection performance, whereas YOLOX performed better in detecting small objects (e.g., piglets) under highly imbalanced conditions. These results suggest that lightweight anchor-based detectors are suitable for general livestock monitoring, while anchor-free detectors may be useful when the target objects are small, sparse, or unevenly represented in the dataset.

Overall, the results indicate that YOLOv4-Tiny achieved the best balance between precision, recall, and mAP, making it the most suitable detector for general thermal-image-based livestock behavior monitoring. In contrast, YOLOX demonstrated superior sensitivity for small-object detection, particularly for piglets, thereby partially fulfilling the study objective of identifying suitable detectors for different livestock monitoring scenarios.

4. Discussion

Dataset analysis revealed a severe imbalance in our annotation structure, with 14,583 instances for the prone posture (65.96% of all bounding boxes) and only 41 instances for piglets and 72 instances for suckling. An imbalance of this magnitude makes it difficult for the model to learn the minority classes effectively.

The high overall mAP of YOLOv4-Tiny (0.8767) demonstrates the effectiveness of lightweight anchor-based feature aggregation. The high precision of YOLOv2 (0.9401) with low recall (0.2524) indicates conservative detection with limited generalizability across minority classes.

The low overall mAP of YOLOX (0.7772) with high precision (0.9335) indicates high sensitivity for small-object detection. In piglet detection, the AP of YOLOX was 0.90, while the AP of YOLOv2 was only 0.24. This indicates that anchor-free prediction improves localization for sparse targets with significant variance in scale and position.

The low recall of YOLOX (0.4536) indicates that rare classes suffer from incomplete detection under severe imbalance, due perhaps to the limited availability of positive samples for minority categories, such as piglets and suckling, which together accounted for less than 1% of the entire dataset. These findings point to the need for data augmentation, focal loss adjustment, or resampling strategies to improve recall stability.

The selection of detectors for livestock monitoring depends on both global mAP and class-specific robustness, particularly in situations requiring the close monitoring of small animals.

The novelty of this study lies in the systematic comparison of YOLOv2, YOLOv4-Tiny, and YOLOX using thermal images collected under practical farrowing-house conditions. Unlike many previous livestock-monitoring studies that relied primarily on RGB imagery or balanced experimental datasets, the present study evaluated detector performance under severe class imbalance and realistic production environments. The results provide practical evidence regarding the strengths and limitations of both anchor-based and anchor-free detection frameworks for precision livestock monitoring.

The findings are consistent with previous studies reporting that YOLOv4-based architectures provide strong overall detection performance in agricultural and livestock applications due to efficient multi-scale feature aggregation [20]. In contrast, the superior piglet-detection performance observed in YOLOX suggests that anchor-free detectors may offer advantages when detecting sparse or small targets exhibiting large variations in position and scale. This trend has also been reported in recent object-detection studies involving small-object recognition tasks.

From a practical perspective, the results suggest that YOLOv4-Tiny is suitable for routine behavioral monitoring in commercial farrowing facilities, where robust overall performance and computational efficiency are important. Conversely, YOLOX may be more appropriate for applications focusing on neonatal piglet monitoring, particularly when early detection of vulnerable piglets is required. Such information can support the development of intelligent livestock-monitoring systems capable of assisting farm managers in improving animal welfare and reducing production losses.

A limitation of this study is that the behavioral complexity observed in commercial farrowing environments was represented using a limited number of broad behavioral cat-

egories. Although the collected thermal images contained diverse posture variations and piglet-related activities, the behaviors were simplified into six representative classes to ensure annotation consistency and maintain sufficient sample numbers for model training and evaluation. While this approach facilitated an objective comparison among different object detection architectures, it may reduce sensitivity to subtle behavioral differences and social interactions among animals. Consequently, the interpretability of animal behavioral dynamics may be limited compared with more detailed ethological analyses. Future studies should investigate finer behavioral classifications to support more comprehensive welfare assessment and precision livestock monitoring.

5. Conclusions

This study compared YOLOv2, YOLOv4-Tiny, and YOLOX in the detection of livestock farrowing behavior using thermal images under a highly imbalanced dataset. In experiments, YOLOv4-Tiny achieved the best overall detection accuracy, while YOLOX demonstrated superior performance for piglet detection.

The strong class imbalance in the dataset reduced recall for minority classes and showed the difficulty of detecting less frequent behaviors under real farm conditions. Although the anchor-free detector did not achieve the highest overall mAP, it performed better for sparse neonatal targets. This suggests that anchor-free detection may be useful when the main task is to detect small piglets rather than larger or more frequent posture classes.

In addition, practical observations indicated that nighttime monitoring using RGB images was more challenging for dark-coated sows and piglets, particularly Taiwan Indigenous Black Pigs, due to reduced image contrast and lower visibility under low-light conditions. Although white-coated Landrace sows and piglets were less affected by nighttime illumination, image quality and target distinguishability still decreased compared with daytime conditions. In contrast, thermal imaging maintained stable target visibility regardless of coat color or ambient lighting, demonstrating its potential as a reliable sensing modality for continuous livestock monitoring in commercial production environments.

Future work will focus on improving the detection of minority classes through training methods that account for class imbalance. Further analysis of thermal images may also help estimate piglet distribution and movement patterns in real production environments.

Funding: This research was funded by the National Science and Technology Council under grant number NSTC 114-2313-B-020-002 and the Ministry of Agriculture under grant number 115 農科-17.1.5-牧-01.

Institutional Review Board Statement: The animal study was conducted in accordance with the ethical guidelines of the Kaohsiung Breeding Farm, Livestock Research Institute, Ministry of Agriculture. All procedures involving animals were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of the Livestock Research Institute, Council of Agriculture, Taiwan (protocol No. IACUC #KAPS-LSI #110-15, approved on 10 November 2020). The study was carried out in accordance with approved animal welfare and ethical standards.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request. The data are not publicly available due to commercial farm privacy restrictions and animal facility confidentiality agreements.

Acknowledgments: The author would like to thank the staff of the Kaohsiung Breeding Farm, Livestock Research Institute, Ministry of Agriculture, for their assistance in animal management, experimental support, and thermal image data collection.

Conflicts of Interest: The author declares no conflicts of interest.

References

- Food and Agriculture Organization of the United Nations. *The Future of Food and Agriculture—Trends and Challenges*; FAO: Rome, Italy, 2017. Available online: <https://www.fao.org/3/i6583e/i6583e.pdf> (accessed on 2 October 2025).
- Berckmans, D. Precision Livestock Farming Technologies for Welfare Management in Intensive Livestock Systems. *Rev. Sci. Tech.* **2014**, *33*, 189–196. <https://doi.org/10.20506/rst.33.1.2273>.
- Kleen, J.L.; Guatteo, R. Precision Livestock Farming: What Does It Contain and What Are the Perspectives? *Animals* **2023**, *13*, 779. <https://doi.org/10.3390/ani13050779>.
- Marino, R.; Petrer, F.; Abeni, F. Scientific Productions on Precision Livestock Farming: An Overview of the Evolution and Current State of Research Based on a Bibliometric Analysis. *Animals* **2023**, *13*, 2280. <https://doi.org/10.3390/ani13142280>.
- Rajagukguk, R.A.; Lee, S.-Y.; Park, J.-Y.; Daniel, K.F.; Lee, C.-R.; Chen, Z.; Liu, D.; Norton, T.; Park, J.; Hong, S.-W. Deep learning for visual animal monitoring (detection, tracking, pose estimation, and behavior classification): A comprehensive review. *Smart Agric. Technol.* **2025**, *12*, 101539. <https://doi.org/10.1016/j.atech.2025.101539>.
- Matthews, S.G.; Miller, A.L.; Plötz, T.; Kyriazakis, I. Automated tracking to measure behavioural changes in pigs for health and welfare monitoring. *Sci. Rep.* **2017**, *7*, 17582.
- Kamilaris, A.; Prenafeta-Boldú, F.X. Deep Learning in Agriculture: A Survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. <https://doi.org/10.1016/j.compag.2018.02.016>.
- Wang, Y.; Muecher, S.; Wang, W.; Guo, L.; Kooistra, L. A Review of Three-Dimensional Computer Vision Used in Precision Livestock Farming for Cattle Growth Management. *Comput. Electron. Agric.* **2023**, *206*, 107687. <https://doi.org/10.1016/j.compag.2023.107687>.
- Alameer, A.; Kyriazakis, I.; Dalton, H.A.; Miller, A.L.; Bacardit, J. Automatic recognition of feeding and foraging behaviour in pigs using deep learning. *Biosyst. Eng.* **2020**, *197*, 91–104. <https://doi.org/10.1016/j.biosystemseng.2020.06.013>.
- Zhang, K.; Li, D.; Huang, J.; Chen, Y. Automated Video Behavior Recognition of Pigs Using Two-Stream Convolutional Networks. *Sensors* **2020**, *20*, 1085. <https://doi.org/10.3390/s20041085>.
- Girshick, R. Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015*; IEEE: New York, NY, USA, 2015; pp. 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>.
- Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In *Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016*; Springer: Cham, Switzerland, 2016; pp. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2.
- Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017*; IEEE: New York, NY, USA, 2017; pp. 7263–7271. <https://doi.org/10.1109/CVPR.2017.690>.
- Ge, Z.; Liu, S.; Wang, F.; Li, Z.; Sun, J. YOLOX: Exceeding YOLO Series in 2021. *arXiv* **2021**, arXiv:2107.08430. <https://doi.org/10.48550/arXiv.2107.08430>.
- Lin, T.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017*; IEEE: New York, NY, USA, 2017; pp. 2980–2988. <https://doi.org/10.1109/ICCV.2017.324>.
- McManus, C.; Tanure, C.B.; Peripolli, V.; Seixas, L.; Fischer, V.; Gabbi, A.M.; Menegassi, S.R.; Stumpf, M.T.; Kolling, G.J.; Dias, E.; et al. Infrared thermography in animal production: An overview. *Comput. Electron. Agric.* **2016**, *123*, 10–16. <https://doi.org/10.1016/j.compag.2016.01.027>.
- Ricci, G.D.; da Silva-Miranda, K.O.; Titto, C.G. Infrared thermography as a non-invasive method for the evaluation of heat stress in pigs kept in pens free of cages in the maternity. *Comput. Electron. Agric.* **2019**, *157*, 403–409. <https://doi.org/10.1016/j.compag.2019.01.017>.
- Liu, X.; Zeng, X.; Chu, M.; Li, Q.; Liu, G. Deep learning-based temperature detection for group-raised pigs via thermal images. *Smart Agric. Technol.* **2025**, *12*, 101206. <https://doi.org/10.1016/j.atech.2025.101206>.

19. Bhujel, A.; Wang, Y.; Lu, Y.; Morris, D.; Dangol, M. A Systematic Survey of Public Computer Vision Datasets for Precision Livestock Farming. *Comput. Electron. Agric.* **2024**, *226*, 109718. <https://doi.org/10.1016/j.compag.2024.109718>.
20. Lai, J.; Liang, Y.; Kuang, Y.; Xie, Z.; He, H.; Zhuo, Y.; Huang, Z.; Zhu, S.; Huang, Z. IO-YOLOv5: Improved Pig Detection under Various Illuminations and Heavy Occlusion. *Agriculture* **2023**, *13*, 1349. <https://doi.org/10.3390/agriculture13071349>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.