Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 9, No. 5, October, 2025

Deep Learning-Based Behavior Recognition for Group-Housed Pigs: Advancing Livestock Management with Segmentation Techniques

Arsanchai Sukkuea [1, 2], Pensiri Akkajit [3*]

[1] *School of Engineering and Technology, Walailak University, Nakhon Si Thammarat 80160, Thailand.*

[2] *Research Center for Intelligent Technology and Integration, School of Engineering and Technology, Walailak University, Nakhon Si Thammarat 80160, Thailand.*

[3] *Faculty of Technology and Environment, Prince of Songkla University, Phuket Campus, Phuket 83120, Thailand.*

Abstract

The increasing demand for sustainable, welfare-oriented livestock management necessitates innovative solutions for behavior monitoring, particularly in group-housed settings, where challenges such as animal density and overlapping bodies hinder traditional observation methods. This study introduces a Convolutional Neural Network (CNN)-based model enhanced with segmentation techniques to accurately classify behaviors among group-housed pigs, a context in which individual monitoring is crucial for welfare assessment, disease prevention, and production efficiency. By leveraging segmentation, the model isolates individual pigs in video footage, overcoming occlusion issues and significantly improving classification accuracy. This approach not only advances the analysis of animal behavior in dense environments but also aligns with the principles of innovation, promoting the adoption of AI-driven monitoring solutions in livestock management. In comparison with various models, YOLOv11m-augmentation achieved the highest mAP@0.5 score of 0.969 and a notable precision of 0.925. This CNN and segmentation-based method effectively identifies key behaviors, including eating, drinking, sleeping, and standing, with particularly high precision for behaviors most indicative of animal welfare. This research contributes to sustainable livestock practices by offering a scalable, cost-effective technology for real-time welfare assessment, potentially reducing labor requirements, enhancing farm management decisions, and promoting animal health. The study's findings underscore the potential of integrating innovation principles with AI in agriculture, presenting a viable pathway toward sustainable livestock management practices that balance productivity with animal welfare.

Keywords:

Behavior Classification;

Deep Learning;

Group-housed Pigs;

Livestock Management;

Segmentation.

Article History:

Received:	19	March	2025
Revised:	02	August	2025
Accepted:	11	August	2025
Published:	01	October	2025

1- Introduction

Livestock management is increasingly under pressure to adopt sustainable and welfare-oriented practices to meet the growing demand for ethical and efficient food production. Within commercial pig farming, the monitoring of animal behavior is crucial, as it provides insights into welfare conditions, health status, and production efficiency. The accurate detection and analysis of pig behaviors can provide valuable insights into the overall well-being of the animals, helping farmers and producers to make informed management decisions [1, 2]. Traditional methods for monitoring behaviors in group-housed pig systems, however, remain limited by several inherent challenges. In these systems, pigs are often housed in large, densely populated pens, where overlapping and occlusion among animals complicate individual

* **CONTACT**: pensiri.a@phuket.psu.ac.th

DOI: http://dx.doi.org/10.28991/ESJ-2025-09-05-013

identification and behavior tracking, making manual observation both labor-intensive and prone to inaccuracies [3, 4]. Advancements in artificial intelligence (AI), particularly deep learning, offer new opportunities to address these limitations. These technologies offer significant advantages over traditional methods, providing more accurate and real-time data that can be used to improve animal welfare and farm management practices.

Convolutional Neural Networks (CNNs) have shown significant potential in image recognition tasks, including animal behavior analysis, due to their capacity to process complex visual data and identify patterns [5-7]. Previous studies have focused on advanced deep learning models such as GoogLeNet, Faster R-CNN, You Only Look Once (YOLO) variants, Long Short-Term Memory (LSTM), Joint Detection and Embedding (JDE), Fair Multi-Object Tracking (FairMOT), and Deep Learning-based SORT (DeepSORT) to analyze or monitor the behavior of pigs housed in groups [8-17]. The primary objective of these studies was to identify various aspects, including feeding behavior, posture, body parts, ear-biting outbreaks, behavioral abnormalities, individual tracking, and specific interactions. However, the existing automated methods often lack real-time capabilities (e.g., Faster R-CNN, LSTM), encounter challenges in complex group-housed environments, and fail to provide actionable insights for farm management.

The YOLO model family has emerged as the preferred choice for these tasks [18], including the detection of animal behavior [19]. YOLO models are highly regarded for their speed and accuracy, rendering them suitable for applications necessitating real-time monitoring [20, 21]. Recent studies have demonstrated the effectiveness of various YOLO models, including YOLOv8, in agricultural applications. For instance, the YOLOv8 framework is shown to be highly effective in real-time detection tasks, including those related to agricultural products and livestock behavior [22, 23]. CNN models alone are often subject to limitations in group-housed settings due to their inability to separate individual animals in cases of high-density proximity, leading to behavior misclassification [21, 24]. The integration of segmentation techniques with CNNs provides a promising solution to these challenges by enabling the model to isolate individual animals within the group, thereby overcoming issues associated with occlusion and overlap [17, 25]. This advancement allows for more precise classification of behaviors such as feeding, resting, and social interaction, which are essential for welfare monitoring and early detection of health issues [26, 27]. By applying this framework to livestock management, particularly through AI-driven behavior classification models, there is an opportunity to develop efficient, welfare-focused monitoring systems, contributing to sustainable farming practices [28, 29].

Addressing the aforementioned challenges, this study explores the application of CNNs in conjunction with segmentation techniques and various YOLO models (YOLOv8n, YOLOv8m, YOLOv8x, YOLOv9c, YOLOv9e, YOLOv11n, YOLOv11m, YOLOv11x) to accurately classify the behaviors of group-housed pigs, thereby contributing to improved animal welfare and informed farm management decisions. The findings highlight the transformative potential of AI in agriculture. By isolating and accurately identifying individual behaviors, this approach offers a scalable and adaptable solution, aligning with innovative principles. This research provides several key contributions, including the capability for real-time monitoring and the facilitation of management decisions, enhancing both animal welfare and productivity.

2- Related Work

In recent years, innovative approaches in livestock management have increasingly integrated advanced technologies to address challenges in animal welfare, productivity, and environmental sustainability. One significant area of development is the use of deep learning for behavior classification in group-housed pigs, a task requiring precise monitoring to ensure animal health, welfare, and efficient resource management. The integration of Internet of Things (IoT) technologies further enhances this process by enabling real-time monitoring and data collection, which is crucial for the effective management of livestock health and productivity [30]. Traditional methods for observing pig behavior rely on manual monitoring, which is labor-intensive and prone to human error. Recent advancements in computer vision and deep learning have shown promise in automating behavior analysis, providing more consistent and objective insights [31]. A review of studies on the automated detection of pig behaviors is presented in Table 1. Existing research on pig behavior monitoring has utilized a range of deep learning-based object detection and tracking models, each exhibiting distinct strengths and limitations.

Many studies have focused on single or limited behaviors, such as feeding, standing, or aggression detection, but lacked a comprehensive classification of multiple behaviors (e.g., drinking, eating, sleeping, and standing) in a unified framework. Additionally, most studies have not explored the impact of data augmentation on model generalization, potentially limiting performance in diverse farm conditions. Despite many studies focusing on model accuracy, few have addressed real-world deployment challenges, such as adaptability to farm environments and web-based monitoring for continuous tracking. Alameer et al. demonstrated high accuracy (99.4% and 96.8%, respectively) in detecting specific behaviors using GoogLeNet and CNN+LSTM. However, these models primarily focused on binary or predefined behaviors rather than comprehensive multi-behavior classification. Faster R-CNN [15] and YOLOv4 [9, 13] achieved notable detection performance, with mAP scores reaching 92.65%. While these models excelled in detecting pig postures and interactions, their real-time capabilities were limited, making them less suitable for continuous on-farm monitoring. Segmentation models, including U-Net and Mask R-CNN, effectively isolate individual pigs from both the background and each other, thereby enhancing tracking precision and behavior classification, even in densely populated environments [32]. These segmentation techniques serve as a foundation for advanced analyses utilizing YOLO-based detection models or CNNs, which classify specific behaviors such as eating, drinking, lying, or standing [33]. Research demonstrates that the integration of segmentation with behavior classification significantly improves the detection accuracy and reliability of monitoring systems, particularly in group housing scenarios, where overlapping movements among pigs can complicate observations [34].

Table 1. Advances in Deep Learning for Livestock Monitoring

Study	Topic	Key Findings	Technologies/Methods Used	Benefits/ Challenges
Alameer et al. [8]	Feeding and foraging behavior recognition	Successfully recognized pig behaviors with high accuracy.	Deep learning (CNNs)	Accurate behavior recognition; requires quality video data.
Alameer et al. [9]	Contact behavior detection	Quantified social contact with precision.	Deep learning, YOLOv4	Enhances social behavior understanding; limited by occlusion.
Guo et al. [10]	Pig detection and tracking	Improved tracking of individual pigs on farms.	Camera-based tracking, AI models	Enables individual monitoring; computational demands.
Kühnemund et al. [12]	Group recumbency detection in pigs	Detected group lying behavior accurately.	AI-supported camera systems	Useful for welfare monitoring; may misclassify under crowding.
Ocepek et al. [13]	Body, head, and tail detection	Developed DigiPig system for automated monitoring.	Deep learning, vision-based system, YOLOv4	Supports early issue detection; initial phase, needs further testing.
Riekert et al. [15]	Pig position and posture detection	Achieved a mean Average Precision (mAP) of 80.2%.	Faster R-CNN	The applicability of DL models for posture detection in night near-infrared recordings
Rajendran et al. [30]	IoT in livestock management	Real-time data collection enhances productivity.	Internet of Things (IoT)	Efficient management; setup costs are high.
Kuzior et al. [31]	Deep learning in livestock management	Integrates advanced tech for better animal welfare and productivity.	Deep learning	Reduces labor, enhances monitoring; data sharing issues.
Dambaulova et al. [32]	Pig behavior monitoring	U-Net and Mask R-CNN used for precise tracking in groups.	U-Net, Mask R-CNN	High tracking accuracy; computationally intensive.
Chae et al. [33]	Behavior classification	Improved behavior analysis using segmentation and YOLO models.	YOLO models, CNNs	Higher accuracy; challenged by overlapping behaviors.
Vladimirov et al. [34]	Combined behavior analysis	Enhanced detection with segmentation and classification.	YOLO-based models	Boosts detection; complex for real-time use.

3- Research Methodology

3-1- Experiment Setup

A comprehensive camera system was installed at a pig farm in Trang Province, Southern Thailand, to systematically collect video footage of pig behavior in a controlled setting. Each pig pen measured five meters by five meters and accommodated approximately eight pigs, in line with group housing recommendations designed to reduce stress and promote natural behaviors essential for pig welfare [35]. Cameras were strategically placed to monitor various behaviors, such as drinking, eating, sleeping, and standing, ensuring that the pigs remained comfortable and in a natural state for accurate observation. The work was conducted following animal use protocol approved under Project License Number Ref. AG057/2023, submitted and reviewed by the Institutional Animal Care and Use Committee, Prince of Songkla University, and protocol number WU-ACUC-66055, reviewed and approved by the Walailak University Institutional Animal Care and Use Committee (WU-IACUC). Both protocols followed the guidelines of the Ethical Review Board of the Office of the National Research Council of Thailand (NRCT) for the care and use of animals in scientific research.

Figure 1 shows the experimental setup for the enhanced behavior classification of group-housed pigs employing advanced deep learning techniques, particularly utilizing the YOLO family of models renowned for their real-time detection capabilities. The process begins with the collection of video data from strategically positioned cameras to capture a variety of pig behaviors, including eating, drinking, sleeping, and standing, under different lighting and activity conditions. These data are meticulously annotated to ensure accurate behavior classification, with image segmentation techniques used to isolate individual pigs within frames, minimizing overlap in group-housed environments [36].

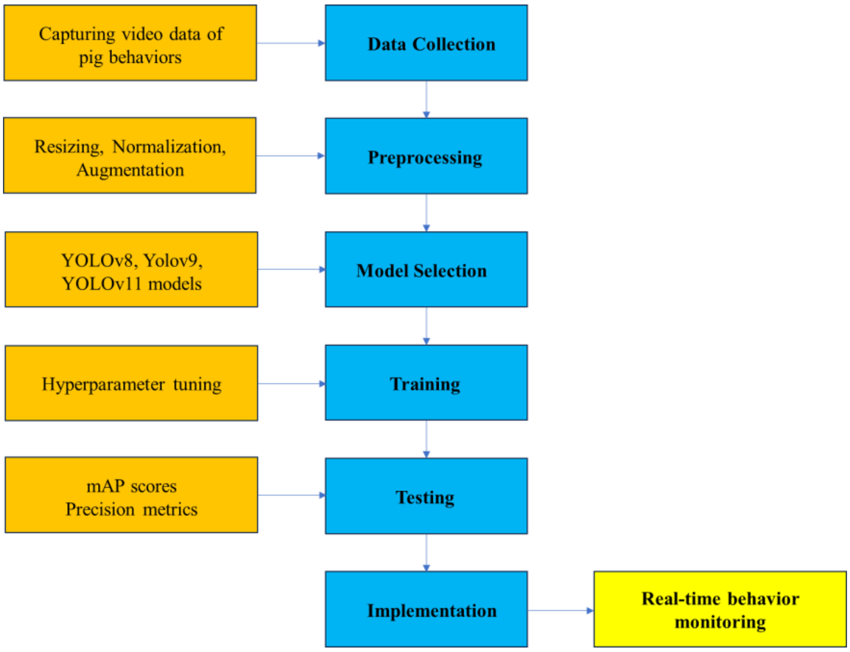


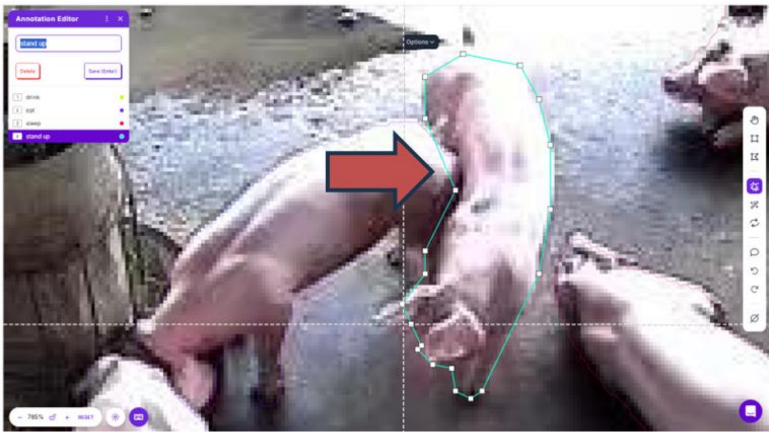
Figure 1. Experiment Setup for Enhanced Behavior Classification in Group-Housed Pigs

3-2- Data Pre-Processing

To enhance model performance, the raw data undergoes pre-processing steps, including resizing (640 × 640 pixels) with a batch size of 16, normalization, and data augmentation (random rotation up to 45°, scaling (80% to 120%), and horizontal flipping). Pre-processing techniques are crucial for improving the robustness of the model against variations in the input data, addressing issues such as overfitting and data imbalance [37]. This study employs a semi-automatic annotation process for dataset labeling (annotation). Figure 2 illustrates the data annotation process using Smart Polygon in Roboflow, generating initial automated polygon masks for the objects of interest. These initial masks are manually reviewed and adjusted by three trained annotators to ensure accuracy and consistency. The use of Smart Polygon facilitates consistent object boundaries, while manual refinement ensures high-quality labels suitable for training the segmentation model. An annotation protocol was developed to clearly define each behavior to ensure high-quality labeling of pig behaviors (e.g., drinking, eating, sleeping, standing).

The labeled pig behaviors for drinking were defined as follows: snout in water source for ≥ 3 seconds, with 300 clips ranging from 5 to 15 seconds; eating (mouth on feed for ≥ 5 seconds, 350 clips ranging from 10 to 20 seconds), sleeping (lying still for ≥ 10 seconds, 400 clips ranging from 15 to 30 seconds), and standing (standing stationary on four legs, the pig's head is positioned either neutrally or elevated for ≥ 3 seconds, 320 clips ranging from 5 to 10 seconds). These behaviors were carefully labeled from video images by trained annotators to ensure precision and consistency across the dataset. In this study, behaviors were categorized as mutually exclusive, each pig being assigned a single behavior label at any given moment. Multi-label classification was not utilized due to the rapid changes in the pigs' postures, which can occur within milliseconds. These transitions were recorded frame-by-frame, facilitating a clear distinction between each pig's posture and the assignment of one dominant behavior. The overlap in behavior during transition was minimal, rendering multi-label classification unnecessary. Challenges such as partial occlusion and group interactions were addressed by employing a tracking-by-detection approach. Each pig was assigned a unique detection ID across frames, facilitating the linkage of behavior predictions over time.

Following the annotation process, the dataset was divided into training (80%) and test (20%) sets. The training set was utilized for model training, and the test set for assessing the final performance of the trained model [21]. The annotated videos were systematically organized into folders for each dataset, with each folder containing video files and an accompanying annotation file detailing the start and end times for each behavior. Two distinct setups were analyzed: one without data augmentation, using a training and validation dataset of 451 images, and another with data augmentation, which expanded the dataset to 1,353 images. All models maintained a consistent test dataset of 100 images, ensuring performance evaluations were conducted under uniform conditions, thereby facilitating fair comparisons across the different YOLO architectures.



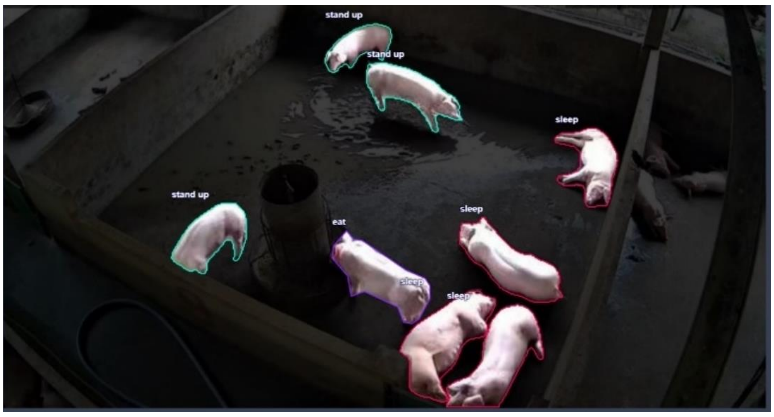
a) Image using Smart Polygon

b) Image after manual review

Figure 2. **Data Annotation**

3-3- Segmentation

The segmentation variant of YOLOv11 was employed in this study to enhance the original object detection capabilities of YOLO by integrating instance segmentation. It integrates object detection and segmentation into a unified pipeline, producing both bounding boxes and segmentation masks for each detected object. This is accomplished by incorporating a mask head within the detection architecture, allowing the network to learn the spatial details essential for accurate segmentation tasks. Compared to traditional segmentation models such as Mask R-CNN or U-Net, YOLOv11-seg delivers real-time performance with competitive accuracy, making it particularly suitable for the application in this study, where both speed and segmentation precision are critical. Figure 3 presents a structured methodology for segmenting and analyzing pig behaviors through distinct classes and layers. In section (a), behaviors such as drinking, eating, sleeping, and standing are categorized, with each representing a unique behavioral class, while (b) adds further granularity by introducing segmentation levels for isolating features like posture, movement, and spatial orientation. This multi-layered approach improves labeling accuracy and supports a deeper understanding of pig behavior.



a) classes

b) layers

Figure 3. Sample Segmentation of Pig Behaviors: a) classes, b) layers

3-4- Model Selection and Configuration

This study focuses on the selection and configuration of suitable YOLO (You Only Look Once) model architectures for pig behavior detection, specifically YOLOv8, YOLOv9, and YOLOv11. The choice of model architecture is critical as it directly impacts both detection accuracy and computational efficiency. Recent advancements in YOLO models, particularly YOLOv8, YOLOv9, and YOLOv11, have introduced various configurations that cater to different processing needs, allowing for a nuanced approach to model selection based on specific application requirements [38-40]. A comprehensive overview of the hyperparameters across the YOLOv8, YOLOv9, and YOLOv11 models are presented to facilitate analysis, including configurations with different augmentation levels, as shown in Table 2, offering insights into the optimization of model training by balancing detection accuracy with training time and computational resources. This trade-off underscores the significance of implementing thoughtful augmentation strategies tailored to specific application requirements [39, 40].

Variants including YOLOv8n (nano), YOLOv8m (medium), and YOLOv8x (extra-large) are examined for YOLOv8. Each model version is designed with specific sizes: the nano models are optimized for lightweight processing, making them suitable for applications requiring rapid inference times, while the extra-large models offer enhanced detection capabilities, albeit at the cost of increased computational resources [40, 41]. Conversely, YOLOv8m strikes a balance between speed and accuracy, making it a versatile option for applications requiring moderate precision without excessive computational demands [42]. On the other hand, YOLOv8x provides the highest accuracy and performance but necessitates significant computational resources, making it ideal for scenarios where precision is critical and resources are plentiful [43]. Similar configurations are available for YOLOv9 and YOLOv11, allowing for a comprehensive evaluation of each model's performance under varying computational and accuracy constraints. For instance, YOLOv9 has been recognized for its enhancements in real-time object detection capabilities, making it suitable for diverse applications, including agricultural monitoring and behavior detection in complex environments [44, 45]. The introduction of YOLOv11 has further refined these capabilities, offering improved detection performance and efficiency [40].

The careful selection of hyperparameters and the strategic use of data augmentation are crucial for improving the performance of the YOLO model. This approach provides a thorough understanding of each version's capabilities and sets the stage for future advancements in object detection technologies [40, 46, 47]. Experimental optimization of hyperparameters, such as learning rate, batch size, and training epochs, was carried out to maximize performance in monitoring and analyzing controlled pig behavior. This standardization of hyperparameters across various YOLO versions facilitates a systematic evaluation of how changes in model architecture and data augmentation impact performance metrics, including accuracy and speed. Each model utilizes the AdamW optimizer, an advanced variant of the Adam optimizer, integrating weight decay, thereby improving the generalization capabilities and stability of the models during training [48].

Training was conducted using Google Colab Pro, providing access to high-performance GPUs, specifically the NVIDIA Tesla T4, along with 25–52 GB of system RAM, depending on the session allocation. The training process for the YOLO-seg model required approximately 20 hours to complete 100 epochs on a dataset consisting of 1,353 labeled images at a resolution of 640 × 640 pixels. Google Colab Pro was chosen for this task, and the trained model operated efficiently on the Jetson Nano, achieving an inference speed of 25–30 frames per second (FPS) at the same resolution for real-time monitoring. The system's computational efficiency and real-time feasibility were evaluated on two CPU-based setups. On Google Colab, the average inference time was 0.3399 seconds per frame, while on a local Intel Core i5 desktop (no GPU), it achieved 0.1580 seconds per frame. These results demonstrate the system's capability to operate within 0.5 seconds per frame, supporting near real-time performance on modest hardware—particularly in applications where detection every second or half-second is sufficient, such as monitoring animal behavior changes not requiring millisecond precision. Therefore, the system appears feasible for real-time monitoring in commercial farm environments, including deployments on moderately powered hardware such as standard desktop CPUs.

Table 2. Model Selection and Configuration

Model	Train Images	Test Images	Image size (pixels)	Batch size	Optimizer
YOLOv8n	451	100	640 × 640	16	AdamW
YOLOv8n-augmentation	1,353	100	640 × 640	16	AdamW
YOLOv8m	451	100	640 × 640	16	AdamW
YOLOv8m-augmentation	1,353	100	640 × 640	16	AdamW
YOLOv8x	451	100	640 × 640	16	AdamW
YOLOv8x-augmentation	1,353	100	640 × 640	16	AdamW
YOLOv9c	451	100	640 × 640	16	AdamW
YOLOv9c-augmentation	1,353	100	640 × 640	16	AdamW
YOLOv9e	451	100	640 × 640	16	AdamW
YOLOv9e-augmentation	1,353	100	640 × 640	16	AdamW
YOLOv11n	451	100	640 × 640	16	AdamW
YOLOv11n-augmentation	1,353	100	640 × 640	16	AdamW
YOLOv11m	451	100	640 × 640	16	AdamW
YOLOv11m-augmentation	1,353	100	640 × 640	16	AdamW
YOLOv11x	451	100	640 × 640	16	AdamW
YOLOv11x-augmentation	1,353	100	640 × 640	16	AdamW

3-5- Evaluation and Performance Metrics

The performance evaluation for detecting pig behaviors was conducted using key metrics, including mean Average Precision at an IoU threshold of 0.50 (mAP@0.5) and precision, calculated using the equations provided below. These metrics assess the model's overall accuracy and its effectiveness in correctly identifying each behavior. The mAP@0.5 was computed by averaging the precision across all behavior classes, as defined by the following equations [49]:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{1}$$

$$mAP@0.5 = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{2}$$

4- Results and Discussion

4-1- Training Phase Performance Evaluation

The performance of various YOLO models across key metrics, including mean Average Precision (mAP@0.5), precision, total training time, epochs, and time per epoch during the training phase were evaluated (Table 3). This assessment provided insights into each model's accuracy and efficiency, comparing standard models and those with data augmentation. Across all models, augmentation consistently improved both mAP@0.5 and precision, although it required more training time. According to the results, the YOLOv8n achieved a high mAP@0.5 of 0.971 and a precision of 0.964, with relatively fast training times of 25.02 minutes for 100 epochs, equating to 15.01 seconds per epoch. However, when applying data augmentation (YOLOv8n-augmentation), the model's mAP@0.5 improved to 0.978, and precision increased to 0.977. In contrast, the training time rose significantly to 70.45 minutes, or 42.27 seconds per epoch, thus highlighting the trade-off between performance and training time associated with augmentation. Similar trends were observed in the YOLOv8m and YOLOv8x models. The YOLOv8m achieved a mAP@0.5 of 0.981 and a precision of 0.981, with a moderate training time of 37.74 minutes. Additionally, the YOLOv8m-augmentation demonstrated a substantial improvement in both metrics, reaching a mAP@0.5 of 0.989 and a precision of 0.992; however, the training time increased to 111.54 minutes, with an epoch time of 66.92 seconds. Similarly, the YOLOv8x model exhibited a comparable performance to YOLOv8m but had longer training times—88.02 minutes for the base model and 245.53 minutes for the augmented version. The YOLOv9c and YOLOv9e models demonstrated competitive performances, balancing accuracy with training efficiency. YOLOv9c achieved a mAP@0.5 of 0.979 and a precision of 0.975 within 52.86 minutes. In contrast, its augmented version required 152.14 minutes, resulting in improved performance with a mAP@0.5 of 0.987 and a precision of 0.988. The augmented version of YOLOv9e demonstrated a training duration of 289.87 minutes, achieving a mAP@0.5 of 0.984 and a precision of 0.974. The YOLOv11 models demonstrated similar performance trends. Specifically, YOLOv11m achieved a mAP@0.5 of 0.99 and a precision of 0.971, while its augmented version required 120.09 minutes. YOLOv11x had a training duration of 73.38 minutes, extending to 210.45 minutes when augmented.

Table 3. Training Phase Performance Evaluation

Model	mAP@0.5	Precision	Training Time (min)	Epochs	Train Time (s/Epoch)
YOLOv8n	0.971	0.964	25.02	100	15.01
YOLOv8n-augmentation	0.978	0.977	70.45	100	42.27
YOLOv8m	0.981	0.981	37.74	100	22.64
YOLOv8m-augmentation	0.989	0.992	111.54	100	66.92
YOLOv8x	0.981	0.984	88.02	100	52.81
YOLOv8x-augmentation	0.989	0.993	245.53	100	147.32
YOLOv9c	0.979	0.975	52.86	100	31.716
YOLOv9c-augmentation	0.987	0.988	152.14	100	91.28
YOLOv9e	0.977	0.963	103.02	100	61.81
YOLOv9e-augmentation	0.984	0.974	289.87	100	173.92
YOLOv11n	0.97	0.981	26.40	100	15.84
YOLOv11n-augmentation	0.977	0.990	73.56	100	44.14
YOLOv11m	0.990	0.971	42.72	100	25.63
YOLOv11m-augmentation	0.997	0.981	120.09	100	72.05
YOLOv11x	0.979	0.962	73.38	100	44.03
YOLOv11x-augmentation	0.986	0.972	210.45	100	126.27

Recent studies have shown that augmented models typically improve mean Average Precision (mAP) and precision; however, this often results in longer training times [7, 49]. These findings highlight the trade-offs in model selection, necessitating a balance between performance accuracy and computational demands to optimize detection capability in practical applications. More recent research has shifted to YOLO variants (YOLOv4, YOLOv7, and YOLOv5s) for behavior recognition and tracking. Odo et al. (2023) achieved high mAP scores, with 98% for ear-biting detection [14, 50], while Tran & Thanh (2023) [16] and Yang et al. (2024) [51] successfully tracked pigs using YOLOv7 with a precision exceeding 90%. However, these models did not assess the impact of data augmentation on detection performance, limiting their adaptability to varying farm conditions. In the present study, the models YOLOv11m-augmentation provided an optimal balance between high performance and reasonable training times, making them suitable for applications requiring high accuracy without imposing excessive computational burdens.

4-2- Testing Performance and Precision in Pig Behavior Classification

This section evaluates the testing performance and precision of each YOLO variant in identifying specific pig behaviors, such as drinking, eating, sleeping, and standing, using metrics mAP@0.5 and precision (Tables 4 and 5). The comparison of mAP@0.5 scores for pig behavior detection across YOLO models (both with and without augmentation), as shown in Table 4, offers valuable insights into their efficacy in detecting different behavioral categories: drinking, eating, sleeping, and standing. The YOLOv11m-augmentation model achieved the highest overall performance with a mAP@0.5 of 0.969, outperforming all other models. This model also excelled in detecting sleeping behaviors, with an impressive mAP@0.5 of 0.998, the highest score among all models. Additionally, YOLOv11m-augmentation demonstrated a strong performance in drinking (0.952), eating (0.978), and standing (0.947) behaviors, making it a highly versatile model for pig behavior detection. Among the YOLOv11 models, YOLOv11n-augmentation and YOLOv11x-augmentation both achieved solid results with mAP@0.5 of 0.962 and 0.958, respectively, but did not outperform YOLOv11m-augmentation in terms of overall accuracy. However, the YOLOv11x-augmentation model excelled in detecting eating behaviors (0.983) and performed admirably in all other behaviors, demonstrating its strong ability in specific behavior classifications.

Recent studies have corroborated these findings, indicating that advanced YOLO models, particularly those employing data augmentation techniques, significantly enhance detection capabilities in various animal behaviors, including those of pigs [52, 53]. The results highlight the value of data augmentation in improving detection accuracy across all behavior types. The YOLOv8m-augmentation model demonstrated exceptional results, achieving a mAP@0.5 of 0.965, with a similarly strong performance in detecting behaviors such as drinking, eating, and standing. Its performance in detecting sleeping behaviors (0.996) ranked second only to the YOLOv11m-augmentation model, highlighting YOLOv8m-augmentation as a formidable option, particularly for applications requiring high accuracy across diverse behavior types. Additionally, the YOLOv8n-augmentation and YOLOv8x-augmentation models exhibited notable enhancements due to augmentation, achieving mAP scores of 0.962 and 0.958, respectively. Notably, YOLOv8x-augmentation also demonstrated strong performance in detecting sleeping behaviors (0.996), comparable to that of YOLOv8m-augmentation. These models are shown to be highly effective, especially in scenarios where augmented data can significantly improve performance across all behavioral categories [54, 55]. On the other hand, YOLOv9c and YOLOv9e, although delivering good precision, lagged behind in terms of overall mAP@0.5. YOLOv9c-augmentation provided the best result among the YOLOv9 variants with a mAP@0.5 of 0.966, indicating that augmentation also contributed to improved performance, particularly in drinking (0.972) and standing (0.941) behaviors. However, the YOLOv9e model, even with augmentation, had a slightly lower mAP@0.5 of 0.958 and was outperformed by other models, particularly in detecting eating and standing behaviors. Augmented models outperformed non-augmented ones, demonstrating the importance of simulating diverse real-world scenarios through data manipulation. These findings align with existing literature showing that augmentation enhances model robustness and generalization for complex visual tasks [56].

Table 4. **The mAP@0.5 Performance of YOLO Models on Pig Behavior Classification**

Model	All	Drinking	Eating	Sleeping	Standing
YOLOv8n	0.954	0.95	0.958	0.985	0.924
YOLOv8n-augmentation	0.962	0.958	0.966	0.993	0.932
YOLOv8m	0.953	0.947	0.953	0.981	0.93
YOLOv8m-augmentation	0.965	0.955	0.971	0.996	0.938
YOLOv8x	0.949	0.94	0.939	0.981	0.934
YOLOv8x-augmentation	0.958	0.948	0.947	0.996	0.942
YOLOv9c	0.955	0.964	0.948	0.976	0.933
YOLOv9c-augmentation	0.966	0.972	0.966	0.984	0.941
YOLOv9e	0.947	0.945	0.943	0.981	0.919
YOLOv9e-augmentation	0.958	0.953	0.961	0.991	0.927
YOLOv11n	0.951	0.938	0.945	0.985	0.936
YOLOv11n-augmentation	0.962	0.946	0.963	0.993	0.944
YOLOv11m	0.956	0.944	0.96	0.983	0.939
YOLOv11m-augmentation	0.969	0.952	0.978	0.998	0.947
YOLOv11x	0.947	0.916	0.966	0.979	0.929
YOLOv11x-augmentation	0.958	0.924	0.983	0.987	0.937

For precision, the YOLOv11 models, especially YOLOv11x-augmentation, achieved the highest precision, particularly in detecting eating and sleeping behaviors, with an overall precision of 0.937. These findings suggest that YOLOv11 models, particularly when augmented, provide the most reliable and accurate detection of pig behaviors, making them well-suited for real-world applications in behavior monitoring systems. Alternatively, the YOLOv8n-augmentation and YOLOv8m-augmentation models demonstrated exceptional performance, achieving scores of 0.932 and 0.934, respectively (Table 5). The YOLOv8m-augmentation model excelled in detecting sleeping behavior, achieving a precision of 0.943, surpassing most other models. It also exhibited high precision across all behavioral categories, highlighting its robustness and versatility. Meanwhile, the YOLOv8x-augmentation model demonstrated comparable performance, with an overall precision of 0.933 and a slight improvement in detecting sleeping behaviors at 0.956. These findings suggest that augmentation techniques significantly enhance detection accuracy, especially for complex behaviors such as sleeping, which are challenging due to subtle variations in motion and posture. In contrast, non-augmented models such as YOLOv8n and YOLOv8m demonstrated a slight reduction in precision compared to their augmented counterparts.

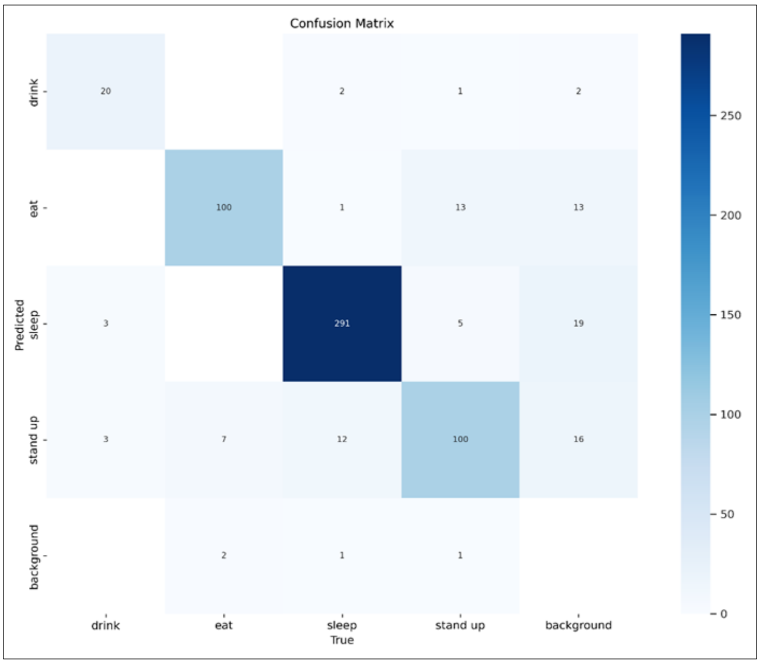
The YOLOv8n model achieved an overall precision of 0.926 across all behaviors, excelling in the detection of drinking and eating behaviors. However, it encountered difficulties with standing behaviors, resulting in a lower precision of 0.865. Similarly, YOLOv8m also attained an overall precision of 0.926, yet its performance for standing behaviors was even lower at 0.898. This highlights the challenges associated with detecting certain behaviors, particularly when the pigs are relatively still [16, 57]. The YOLOv9e and YOLOv9c models demonstrated commendable precision; however, their performance was not as robust when compared to the YOLOv8 and YOLOv11 models. Specifically, YOLOv9e recorded the lowest overall precision at 0.917, with significant declines in its ability to detect eating and standing behaviors. Minimal pig movement, resulting in fewer distinct key points, may lead to insufficient cues for differentiating standing from other behaviors. Additionally, standing pigs exhibit numerous visual similarities with other postures, which can result in confusion and misclassification. The results of this study are inconsistent with those of Scaillierez et al. [58], who found that standing was the best-detected posture, demonstrating the highest precision, sensitivity, and F1 score when using YOLOv8 to detect standing, sitting, sternal lying, and lateral lying pigs. Additionally, Jahn et al. [59] found that detecting drinking involving dynamic interactions is often complicated by significant occlusion compared to detecting standing. The enhanced version of YOLOv9e showed some improvement, achieving a precision of 0.924; nonetheless, it still lagged behind other models, particularly YOLOv8m-augmentation.

Table 5. The Precision of YOLO Models on Pig Behavior Classification

Model	All	Drinking	Eating	Sleeping	Standing
YOLOv8n	0.926	0.997	0.921	0.923	0.865
YOLOv8n-augmentation	0.932	0.997	0.929	0.931	0.87
YOLOv8m	0.926	0.946	0.926	0.935	0.898
YOLOv8m-augmentation	0.934	0.954	0.934	0.943	0.906
YOLOv8x	0.925	0.938	0.913	0.948	0.9
YOLOv8x-augmentation	0.933	0.946	0.921	0.956	0.908
YOLOv9c	0.925	0.967	0.921	0.944	0.869
YOLOv9c-augmentation	0.933	0.972	0.929	0.952	0.877
YOLOv9e	0.917	0.909	0.904	0.95	0.903
YOLOv9e-augmentation	0.924	0.917	0.909	0.958	0.911
YOLOv11n	0.916	0.968	0.926	0.939	0.829
YOLOv11n-augmentation	0.923	0.976	0.934	0.944	0.837
YOLOv11m	0.917	0.937	0.919	0.936	0.877
YOLOv11m-augmentation	0.925	0.945	0.924	0.944	0.885
YOLOv11x	0.929	0.896	0.947	0.964	0.908
YOLOv11x-augmentation	0.937	0.904	0.955	0.972	0.916

4-3- Confusion Matrix

In addition, the YOLOv11m-augmentation was evaluated to determine how well it classified pig behaviors. The multiclass classification performance was evaluated quantitatively on a normalized confusion matrix (Figure 4), with the horizontal axis representing the true labels and the vertical axis representing the predicted labels of pig behaviors. The evaluation illustrates the misclassification and confusion existing among the categories of pig behaviors—eating, drinking, sleeping, and standing. The YOLOv11m-augmentation significantly improved the efficiency and accuracy of pig behavior classification. The YOLOv11m-augmentation demonstrated the ability to distinguish between various pig behaviors, leading to the precise identification of different forms of pig behavior. According to the results, the YOLOv11m-augmentation exhibited strong discriminative capabilities for categorization.



a) Confusion Matrix

b) Misclassification Examples

Figure 4. Confusion Matrix of YOLOv11m-augmentation model

Instances of false positives were found where the ground truth label was "background," but the model incorrectly predicted behaviors like "standing," "sleeping," or "eating." This was likely due to duplicate predictions in overlapping areas with multiple pigs, causing the model to link background regions with pig behaviors. Additionally, there were cases where the true label was "eating," but the model predicted other behaviors. This could occur when pigs move away from or toward the feeding area, making visual cues ambiguous. To address this, the researchers plan to implement a tracking ID system to assign a consistent identity to each pig. By correlating predictions with a unique pig ID across frames, the system can reduce redundant behavior predictions in overlapping areas, thereby decreasing false positives in background regions.

4-4- Application of Innovative Technologies in Agriculture

The performance evaluation of the behavioral analysis concerning group-housed pigs, utilizing training and testing datasets, demonstrates that YOLOv11m-augmentation effectively achieves a balance between high precision and detection accuracy across various behaviors. Therefore, this section evaluates the predictions of the YOLOv11m-augmentation model on labeled and unlabeled data to demonstrate their ability to accurately detect and classify various behaviors exhibited by pigs, such as drinking, eating, sleeping, and standing (Figure 5). As illustrated in Figure 5, the image presents multiple frames of pigs within a pen, annotated with behaviors including drinking, eating, sleeping, and standing. Each behavior is accompanied by a confidence score, reflecting the model's certainty in identifying these actions. The YOLOv11m-augmentation automates the detection of behaviors within a labeled dataset, demonstrating advanced capabilities in identifying pig behaviors. The majority of detected behaviors exhibit high confidence scores (≥ 0.9), particularly for actions such as drinking and eating, highlighting the model's strong predictive accuracy. High confidence predictions suggest it is well-trained and reliable, though further fine-tuning could improve borderline cases.

The findings of this study demonstrate that YOLOv11m-augmentation models are highly effective for large-scale implementation in smart farming systems, where real-time monitoring of animal behavior is critical to ensuring welfare. These models enable the accurate and timely detection of pig behaviors, supporting early intervention strategies for health management and welfare optimization. The application of deep learning in this context assists farmers in maintaining optimal living conditions, potentially leading to increased productivity and reduced stress levels among livestock [53]. Moreover, the study highlights the effectiveness of advanced deep learning models, such as YOLO, in enhancing animal welfare through precise behavior recognition. The incorporation of data augmentation techniques further enhances model robustness, enabling a reliable performance across diverse operational environments. Collectively, these advancements offer practical and scalable solutions for improving livestock monitoring and management in modern agriculture. Recent studies have corroborated these findings, indicating that YOLO models, especially with augmentation, significantly improve the detection accuracy of various animal behaviors, including those of pigs, by leveraging advanced deep learning techniques [36, 53, 56, 60, 61].

a) Behavior Prediction for Group-housed Pigs

b) Comparison between labeled and predicted images for pig behavior classification

Figure 5. **Deep Learning for Behavior Classification in Livestock with YOLOv11m-augmentation**

5- Conclusion

The application of deep learning in livestock management, particularly through segmentation, marks a significant advancement in monitoring the welfare of group-housed pigs. The YOLOv11m-augmentation model demonstrates that integrating machine learning with data augmentation techniques substantially enhances behavior classification accuracy. The model achieved an outstanding mAP@0.5 of 0.969 and a strong precision score of 0.925. This CNN- and segmentation-based approach can accurately identify key pig behaviors—eating, drinking, sleeping, and standing—with especially high precision for those behaviors most indicative of animal welfare. These capabilities are crucial for making timely decisions related to feeding, health monitoring, and behavioral interventions, ultimately promoting better animal welfare and more efficient farm operations. The integration of segmentation techniques in deep learning models enhances behavioral analysis by enabling the precise identification of specific actions within group housing environments. Furthermore, the open-source nature of these technologies fosters rapid innovation in agriculture by encouraging collaboration among farmers, researchers, and technologists to share data and improve model performance. This collaborative ecosystem supports continuous advancement, allowing AI technologies to adapt to the diverse needs of farming environments, particularly in terms of computational efficiency and scalability.

6- Declarations

6-1- Author Contributions

Conceptualization, A.S. and P.A.; methodology, A.S. and P.A.; software, A.S. and P.A.; validation, A.S. and P.A.; formal analysis, A.S. and P.A.; investigation, P.A.; resources, A.S. and P.A.; data curation, A.S.; writing—original draft preparation, A.S. and P.A.; writing—review and editing, A.S. and P.A.; visualization, A.S.; supervision, P.A.; project administration, P.A; funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

6-2- Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6-3- Funding

This project is funded by National Research Council of Thailand (NRCT): Contract number N42A660827 and Walailak University [grant number WU66249]. The authors would like to thank Price of Songkla University, Phuket Campus for providing partial funding for this project.

6-4- Acknowledgements

The authors would like to thank Price of Songkla University, Phuket Campus for providing partial funding for this project.

6-5- Institutional Review Board Statement

The work was conducted following animal use protocol approved under Project License Number Ref. AG057/2023, submitted and reviewed by Institutional Animal Care and Use Committee, Prince of Songkla University, and protocol number WU-ACUC-66055, reviewed and approved by Walailak University Institutional Animal Care and Use Committee (WU-IACUC) in accordance with the guidelines of animal care and use under the Ethical Review Board of the Office of National Research Council of Thailand (NRCT) for the conduction of the scientific research.

6-6- Informed Consent Statement

Not applicable.

6-7- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

[1] Hassan, M., Park, J. H., & Han, M. H. (2023). Enhancing livestock management with IoT-based wireless sensor networks: a comprehensive approach for health monitoring, location tracking, behavior analysis, and environmental optimization. Journal of Sustainable Urban Futures, 13(6), 34-46.

[2] Kleen, J. L., & Guatteo, R. (2023). Precision Livestock Farming: What Does It Contain and What Are the Perspectives? Animals, 13(5), 779. doi:10.3390/ani13050779.

[3] Min, P. K., Mito, K., & Kim, T. H. (2024). The Evolving Landscape of Artificial Intelligence Applications in Animal Health. Indian Journal of Animal Research, 58(10), 1793–1798. doi:10.18805/IJAR.BF-1742.

[4] Ali AlZubi, A., & Al-Zu'bi, M. (2023). Application of Artificial Intelligence in Monitoring of Animal Health and Welfare. Indian Journal of Animal Research, 57(11), 1550–1555. doi:10.18805/IJAR.BF-1698.

[5] Melak, A., Aseged, T., & Shitaw, T. (2024). The Influence of Artificial Intelligence Technology on the Management of Livestock Farms. International Journal of Distributed Sensor Networks, 2024, 1–12. doi:10.1155/2024/8929748.

[6] Zhang, X., Xiong, Y., Pan, Y., Xu, D., Kawsar, I., Liu, B., & Hou, L. (2023). Deep-learning-based inverse structural design of a battery-pack system. Reliability Engineering & System Safety, 238, 109464. doi:10.1016/j.ress.2023.109464.

[7] Akkajit, P., Sukkuea, A., & Thongnonghin, B. (2023). Comparative analysis of five convolutional neural networks and transfer learning classification approach for microplastics in wastewater treatment plants. Ecological Informatics, 78, 102328. doi:10.1016/j.ecoinf.2023.102328.

[8] Alameer, A., Kyriazakis, I., Dalton, H. A., Miller, A. L., & Bacardit, J. (2020). Automatic recognition of feeding and foraging behaviour in pigs using deep learning. Biosystems Engineering, 197, 91–104. doi:10.1016/j.biosystemseng.2020.06.013.

[9] Alameer, A., Buijs, S., O'Connell, N., Dalton, L., Larsen, M., Pedersen, L., & Kyriazakis, I. (2022). Automated detection and quantification of contact behaviour in pigs using deep learning. Biosystems Engineering, 224, 118–130. doi:10.1016/j.biosystemseng.2022.10.002.

[10] Guo, Q., Sun, Y., Orsini, C., Bolhuis, J. E., de Vlieg, J., Bijma, P., & de With, P. H. N. (2023). Enhanced camera-based individual pig detection and tracking for smart pig farms. Computers and Electronics in Agriculture, 211, 108009. doi:10.1016/j.compag.2023.108009.

[11] Han, X. Le, Jiang, N. J., Hata, T., Choi, J., Du, Y. J., & Wang, Y. J. (2023). Deep learning based approach for automated characterization of large marine microplastic particles. Marine Environmental Research, 183, 105829. doi:10.1016/j.marenvres.2022.105829.

[12] Kühnemund, A., Götz, S., & Recke, G. (2023). Automatic Detection of Group Recumbency in Pigs via AI-Supported Camera Systems. Animals, 13(13), 2205. doi:10.3390/ani13132205.

[13] Ocepek, M., Žnidar, A., Lavrič, M., Škorjanc, D., & Andersen, I. L. (2022). Digipig: First developments of an automated monitoring system for body, head and tail detection in intensive pig farming. Agriculture (Switzerland), 12(1), 2. doi:10.3390/agriculture12010002.

[14] Odo, A., Muns, R., Boyle, L., & Kyriazakis, I. (2023). Video Analysis Using Deep Learning for Automated Quantification of Ear Biting in Pigs. IEEE Access, 11, 59744–59757. doi:10.1109/access.2023.3285144.

[15] Riekert, M., Klein, A., Adrion, F., Hoffmann, C., & Gallmann, E. (2020). Automatically detecting pig position and posture by 2D camera imaging and deep learning. Computers and Electronics in Agriculture, 174, 105391. doi:10.1016/j.compag.2020.105391.

[16] Tran, D. D., & Thanh, N. D. (2023). Pig Health Abnormality Detection Based on Behavior Patterns in Activity Periods using Deep Learning. International Journal of Advanced Computer Science and Applications, 14(5), 603–610. doi:10.14569/IJACSA.2023.0140564.

[17] Yang, Q., Hui, X., Huang, Y., Chen, M., Huang, S., & Xiao, D. (2024). A Long-Term Video Tracking Method for Group-Housed Pigs. Animals, 14(10), 1505. doi:10.3390/ani14101505.

[18] Lu, S., Wang, B., Wang, H., Chen, L., Linjian, M., & Zhang, X. (2019). A real-time object detection algorithm for video. Computers and Electrical Engineering, 77, 398–408. doi:10.1016/j.compeleceng.2019.05.009.

[19] Wang, J., Wang, N., Li, L., & Ren, Z. (2020). Real-time behavior detection and judgment of egg breeders based on YOLO v3. Neural Computing and Applications, 32(10), 5471–5481. doi:10.1007/s00521-019-04645-4.

[20] Jha, S., Seo, C., Yang, E., & Joshi, G. P. (2021). Real time object detection and tracking system for video surveillance system. Multimedia Tools and Applications, 80(3), 3981–3996. doi:10.1007/s11042-020-09749-x.

[21] Sukkuea, A., & Akkajit, P. (2024). Using Convolutional Neural Network for Behavior Classification of Group-Housed Pigs. Proceedings of the 9th World Congress on Civil, Structural, and Environmental Engineering, 120. doi:10.11159/iceptp24.120.

[22] Lei, C., Zeng, J., Xia, Y., & Pang, F. (2024). Aircraft type recognition based on YOLOv8. Journal of Physics: Conference Series, 2787(1), 12047. doi:10.1088/1742-6596/2787/1/012047.

[23] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Machine Learning and Knowledge Extraction, 5(4), 1680–1716. doi:10.3390/make5040083.

[24] Mahato, S., & Neethirajan, S. (2024). Integrating Artificial Intelligence in dairy farm management − biometric facial recognition for cows. Information Processing in Agriculture, 12(3), 312-325. doi:10.1016/j.inpa.2024.10.001.

[25] Luo, Y., Xia, J., Lu, H., Luo, H., Lv, E., Zeng, Z., Li, B., Meng, F., & Yang, A. (2024). Automatic Recognition and Quantification Feeding Behaviors of Nursery Pigs Using Improved YOLOV5 and Feeding Functional Area Proposals. Animals, 14(4), 569. doi:10.3390/ani14040569.

[26] Lagua, E. B., Mun, H.-S., Mark, K., Ampode, B., Chem, V., Park, H.-R., Kim, Y.-H., & Yang, C.-J. (2023). A Comprehensive Monitoring System for Health Status and Behavior of Pigs Under Different Environmental Conditions Using Vision and Audio Technologies. doi:10.20944/preprints202309.0695.v1.

[27] Lagua, E. B., Mun, H. S., Ampode, K. M. B., Chem, V., Kim, Y. H., & Yang, C. J. (2023). Artificial Intelligence for Automatic Monitoring of Respiratory Health Conditions in Smart Swine Farming. Animals, 13(11), 1860. doi:10.3390/ani13111860.

[28] Zhang, L., Guo, W., Lv, C., Guo, M., Yang, M., Fu, Q., & Liu, X. (2024). Advancements in artificial intelligence technology for improving animal welfare: Current applications and research progress. Animal Research and One Health, 2(1), 93–109. doi:10.1002/aro2.44.

[29] Tzanidakis, C., Tzamaloukas, O., Simitzis, P., & Panagakis, P. (2023). Precision Livestock Farming Applications (PLF) for Grazing Animals. Agriculture (Switzerland), 13(2), 288. doi:10.3390/agriculture13020288.

[30] Rajendran, J. G., Alagarsamy, M., Seva, V., Dinesh, P. M., Rajangam, B., & Suriyan, K. (2023). IoT based tracking cattle healthmonitoring system using wireless sensors. Bulletin of Electrical Engineering and Informatics, 12(5), 3086–3094. doi:10.11591/eei.v12i5.4610.

[31] Kuzior, A., Sira, M., & Brożek, P. (2023). Use of Artificial Intelligence in Terms of Open Innovation Process and Management. Sustainability (Switzerland), 15(9), 7205. doi:10.3390/su15097205.

[32] Dambaulova, G. K., Madin, V. A., Utebayeva, Z. A., Baimyrzaeva, M. K., & Shora, L. Z. (2023). Benefits of automated pig feeding system: A simplified cost–benefit analysis in the context of Kazakhstan. Veterinary World, 16(11), 2205–2209. doi:10.14202/vetworld.2023.2205-2209.

[33] Chae, J. W., Sim, H. S., Lee, C. W., Choi, C. S., & Cho, H. C. (2024). Video-Based Analysis of Cattle Behaviors: Improved Classification Using FlowEQ Transform. IEEE Access, 12, 42860–42867. doi:10.1109/ACCESS.2024.3379277.

[34] Vladimirov, N., Brui, E., Levchuk, A., Al-Haidri, W., Fokin, V., Efimtcev, A., & Bendahan, D. (2023). CNN-based fully automatic wrist cartilage volume quantification in MR images: A comparative analysis between different CNN architectures. Magnetic Resonance in Medicine, 90(2), 737-751. doi:10.1002/mrm.29671.

[35] Campler, M., Pairis-Garcia, M., Kieffer, J., & Moeller, S. (2019). Sow behavior and productivity in a small stable group-housing system. Journal of Swine Health and Production, 27(2), 76–86. doi:10.54846/jshap/1080.

[36] Huang, L., Xu, L., Wang, Y., Peng, Y., Zou, Z., & Huang, P. (2022). Efficient Detection Method of Pig-Posture Behavior Based on Multiple Attention Mechanism. Computational Intelligence and Neuroscience, 2022, 1–12. doi:10.1155/2022/1759542.

[37] Tan, S. X., Ong, J. Y., Goh, K. O. M., & Tee, C. (2024). Boosting Vehicle Classification with Augmentation Techniques across Multiple YOLO Versions. International Journal on Informatics Visualization, 8(1), 45–54. doi:10.62527/joiv.8.1.2313.

[38] Pham, D. A., & Han, S. H. (2024). Deploying a Computer Vision Model Based on YOLOv8 Suitable for Drones in the Tuna Fishing and Aquaculture Industry. Journal of Marine Science and Engineering, 12(5), 828. doi:10.3390/jmse12050828.

[39] Ali, M. L., & Zhang, Z. (2024). The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection. Computers, 13(12), 336. doi:10.3390/computers13120336.

[40] Sapkota, R., Meng, Z., Churuvija, M., Du, X., Ma, Z., & Karkee, M. (2024). Comprehensive performance evaluation of yolov12, yolo11, yolov10, yolov9 and yolov8 on detecting and counting fruitlet in complex orchard environments. arXiv preprint, arXiv:2407.12040. doi:10.48550/arXiv.2407.12040.

[41] Sapkota, R., Flores-Calero, M., Qureshi, R. et al. YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series. Artificial Intelligence Review, 58, 274. doi:10.1007/s10462-025-11253-3.

[42] Casas, E., Ramos, L., Bendek, E., & Rivas-Echeverria, F. (2024). YOLOv5 vs. YOLOv8: Performance Benchmarking in Wildfire and Smoke Detection Scenarios. Journal of Image and Graphics, 12(2), 127–136. doi:10.18178/joig.12.2.127-136.

[43] Zhang, M., Wang, Z., Song, W., Zhao, D., & Zhao, H. (2024). Efficient Small-Object Detection in Underwater Images Using the Enhanced YOLOv8 Network. Applied Sciences (Switzerland), 14(3), 1095. doi:10.3390/app14031095.

[44] Li, J., Feng, Y., Shao, Y., & Liu, F. (2024). IDP-YOLOV9: Improvement of Object Detection Model in Severe Weather Scenarios from Drone Perspective. Applied Sciences (Switzerland), 14(12), 5277. doi:10.3390/app14125277.

[45] Lu, D., & Wang, Y. (2024). MAR-YOLOv9: A multi-dataset object detection method for agricultural fields based on YOLOv9. PLoS ONE, 19(10 October), 307643. doi:10.1371/journal.pone.0307643.

[46] Yilmaz, B., & Kutbay, U. (2024). YOLOv8-Based Drone Detection: Performance Analysis and Optimization. Computers, 13(9), 234. doi:10.3390/computers13090234.

[47] Duong-Trung, H., & Duong-Trung, N. (2024). Integrating YOLOv8-agri and DeepSORT for Advanced Motion Detection in Agriculture and Fisheries. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, 11(1), 4618. doi:10.4108/eetinis.v11i1.4618.

[48] Ding, P., Zhan, H., Yu, J., & Wang, R. (2024). A bearing surface defect detection method based on multi-attention mechanism Yolov8. Measurement Science and Technology, 35(8), 86003. doi:10.1088/1361-6501/ad4386.

[49] Akkajit, P., Alahi, M. E. E., & Sukkuea, A. (2024). Enhanced detection and classification of microplastics in marine environments using deep learning. Regional Studies in Marine Science, 80. doi:10.1016/j.rsma.2024.103880.

[50] Orlando, J. I., Gerendas, B. S., Riedl, S., Grechenig, C., Breger, A., Ehler, M., ... & Schmidt-Erfurth, U. (2020). Automated quantification of photoreceptor alteration in macular disease using optical coherence tomography and deep learning. Scientific reports, 10(1), 5619. doi:10.1038/s41598-020-62329-9.

[51] Yang, Q., Hui, X., Huang, Y., Chen, M., Huang, S., & Xiao, D. (2024). A Long-Term Video Tracking Method for Group-Housed Pigs. Animals, 14(10), 1505. doi:10.3390/ani14101505.

[52] Ho, K. Y., Tsai, Y. J., & Kuo, Y. F. (2021). Automatic monitoring of lactation frequency of sows and movement quantification of newborn piglets in farrowing houses using convolutional neural networks. Computers and Electronics in Agriculture, 189, 106376. doi:10.1016/j.compag.2021.106376.

[53] Mao, R., Shen, D., Wang, R., Cui, Y., Hu, Y., Li, M., & Wang, M. (2024). An Integrated Gather-and-Distribute Mechanism and Attention-Enhanced Deformable Convolution Model for Pig Behavior Recognition. Animals, 14(9), 1316. doi:10.3390/ani14091316.

[54] Jin, M., Yang, B., & Wang, C. (2023). Research on Identification and Classification Method of Imbalanced Data Set of Pig Behavior. Engenharia Agricola, 43(2), e20220014. doi:10.1590/1809-4430-ENG.AGRIC.V43N2E20220014/2023.

[55] Wei, J., Tang, X., Liu, J., & Zhang, Z. (2023). Detection of Pig Movement and Aggression Using Deep Learning Approaches. Animals, 13(19), 3074. doi:10.3390/ani13193074.

[56] Tu, S., Zeng, Q., Liang, Y., Liu, X., Huang, L., Weng, S., & Huang, Q. (2022). Automated Behavior Recognition and Tracking of Group-Housed Pigs with an Improved DeepSORT Method. Agriculture (Switzerland), 12(11), 1907. doi:10.3390/agriculture12111907.

[57] Oliveira, R. F. D., Soares, R. D. T. R. N., Moreira, R. H. R., Andrade, R. P. D., Rosenfield, D. A., & Pizzutto, C. S. (2023). Effects of the environmental enrichment on pigs' behavior and performance. Revista Brasileira de Zootecnia, 52, e20210123. doi:10.37496/rbz5220210123.

[58] Scaillierez, A. J., García-Faria, T. I., Broers, H., van Nieuwamerongen-De Koning, S. E., Van Der Tol, R. P. P. J., Bokkers, E. A. M., & Boumans, I. J. M. M. (2024). Determining the posture and location of pigs using an object detection model under different lighting conditions. Translational Animal Science, 8, 167. doi:10.1093/tas/txae167.

[59] Jahn, S., Schmidt, G., Bachmann, L., Louton, H., Homeier-Bachmann, T., & Schütz, A. K. (2024). Individual behavior tracking of heifers by using object detection algorithm YOLOv4. Frontiers in Animal Science, 5, 1499253. doi:10.3389/fanim.2024.1499253.

[60] Huang, Y., Xiao, D., Liu, J., Tan, Z., Liu, K., & Chen, M. (2023). An Improved Pig Counting Algorithm Based on YOLOv5 and DeepSORT Model. Sensors, 23(14), 6309. doi:10.3390/s23146309.

[61] Wang, L., Li, L., Wang, H., Zhu, S., Zhai, Z., & Zhu, Z. (2023). Real-time vehicle identification and tracking during agricultural master-slave follow-up operation using improved YOLO v4 and binocular positioning. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 237(6), 1393–1404. doi:10.1177/09544062221130928.