Engineered Science
DOI: https://dx.doi.org/10.30919/es1808

Improved Detection and Classification of Precise Behaviors in Group-Housed Pigs Using Deep Learning Models

Arsanchai Sukkuea[1,2] and Pensiri Akkajit[3,4,*]

Abstract

Accurate monitoring of pig behavior is crucial for enhancing animal welfare and production efficiency in agriculture. Traditional observation limits the early detection of health issues and results in ineffective farm management. This study evaluates the performance of the You Only Look Once (YOLO) family, specifically YOLOv8n (nano), YOLOv8m (medium), YOLOv8x (extra-large), and a novel approach, YOLOv8ma (modified architecture), in classifying key behaviors—drinking, eating, sleeping, and standing—that correlate with health status for accurate behavioral analysis in group-housed pigs. This is based on overall performance, behavior detection accuracy, the impact of 2x and 3x augmentation, real-time processing, generalization capability, adaptability to environmental conditions, model size, and processing speed. The results indicate that the YOLOv8ma model's performance increased from 0.947 without augmentation to 0.957 with 2x augmentation, while precision improved from 0.896 to 0.910. Compared to the YOLOv8 models, the YOLOv8ma also increased mAP@0.5 and precision by 6.32% and 4.59%, respectively. The model features shorter training times of 33.24 to 86.70 seconds per epoch and processing speeds of 40 frames per second, making it well-suited for fast-paced scenarios. PigSenseAI is a scalable, cost-effective web application that enables real-time behavior detection, classification, and automated alerts through an intuitive interface.

Keywords: Agricultural technology; Data augmentation; Deep learning; Group-housed pigs; Pig behavior.

Received: 25 March 2025; Revised: 27 May 2025; Accepted: 08 October 2025.
Article type: Research article.

1. Introduction

Livestock behavior monitoring is crucial for ensuring animal welfare, health, and productivity in modern agricultural practices.[1-3] Among livestock, pigs are particularly important due to their economic significance; however, managing their behavior in group housing systems poses unique challenges.[4] Traditional methods of monitoring pig behavior, such as direct observation, are labor-intensive, time-consuming, and often subjective.[5] Recently, there has been growing interest in leveraging advanced technologies, such as deep learning (DL), to automate the monitoring and analysis of pig behavior.[6-9] 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. The proposed technology effectively extracts key behavioral features by accurately detecting and analyzing relevant regions of the pigs. This enables insights into well-being, helping farmers make informed decisions.[10,11]

Deep learning, a subset of artificial intelligence (AI), has shown remarkable success in various computer vision tasks, including image classification,[12,13] object detection,[14-16] forecasting models for environmental quality,[17] and behavior analysis.[18] Convolutional neural networks (CNNs), a type of deep learning architecture, have demonstrated superior performance in detecting and classifying objects in images and videos.[19,20] One of the key advantages of CNNs

[1] *School of Engineering and Technology, Walailak University 222 Thaiburi, Thasala, 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*
[4] *Hub of Waste Management for Sustainable Development, Center of Excellence on Hazardous Substance Management, Chulalongkorn University, Bangkok 10330, Thailand*
* *E-mail:* pensiri.a@phuket.psu.ac.th (P. Akkajit)

is their ability to automatically learn and extract relevant features from raw data, making them well-suited for complex tasks such as behavior detection in animals.[21,22] Object detection is a critical technology in video surveillance, facilitating the identification and localization of multiple objects in real-time. The You Only Look Once (YOLO) model family has emerged as a preferred choice for these tasks,[23] including animal behavior detection and animal phenotype application.[24-27] YOLO models are highly regarded for their speed and accuracy, rendering them suitable for applications that necessitate real-time monitoring.[28,29]

In group-housed environments, pigs often cluster together, which presents challenges for monitoring and detecting individual behaviors.[30] This study investigates the detection of pig behavior to monitor the activities of group-housed pigs. The behaviors of interest include drinking, eating, sleeping, and standing, as they provide essential insights into both the physical and mental health of the pigs.[4,31] These behaviors are vital for effective livestock management and serve as indicators of pig health and welfare, enabling a comprehensive assessment of the pigs' daily activities and overall well-being. The drinking behavior of pigs is crucial to ensure they have regular access to water, which is essential for their health and well-being.[32,33] Eating behavior helps assess their nutritional intake and indicates their health status or any feeding-related issues.[34,35] Sleeping behavior is important for understanding rest and recovery times, which are vital for their overall health and growth.[36] Standing behavior provides insights into the pigs' activity levels, mobility, and general physical condition. Detecting and classifying the behavior patterns of group-housed pigs enables farmers to monitor and manage their daily activities more effectively. This approach provides valuable insights for assessing pig health based on observed behaviors.[37] Consequently, advanced models capable of managing dense and complex scenes are essential for effective behavior analysis and health monitoring.

Previous studies have employed deep learning techniques to analyze and monitor pig behavior, focusing on aspects such as movement, aggression,[38,39] ear-biting outbreaks,[40] and other activities that may indicate health and welfare issues.[26] Existing automated methods often encounter challenges in complex group-housed environments. Many studies have focused on a single behavior or a narrow range of behaviors, resulting in a lack of a comprehensive classification that includes multiple behaviors within a unified framework. Additionally, many systems rely on extensive tracking infrastructure and require substantial computational resources, which limits their practical application on farms. Several approaches have utilized the Vision Transformer (ViT) to monitor pig behavior, improving the extraction of spatiotemporal features and enhancing performance in pig behavior recognition tasks. The ViT was employed for real-time detection of aggression in pigs within a pen, achieving F1-scores of 82.8% and 82.7%, along with an AUC of 85%.[39] Additionally, ViTAM-SlowFast was used to analyze the behaviors of group-housed pigs by integrating a Video Transformer Encoder with a Temporal Adaptive Module. This methodology achieved a mean Average Precision (mAP) of 92.68%, representing a 2.68% improvement over the SlowFast model.[37] The MobileNetV2 and Autoformer were proposed for pig aggression detection, enhancing feature extraction. The model achieved a recall of 98.08%, precision of 94.44%, accuracy of 96.23%, and an F1-score, with a total of 10.41 million parameters.[26]

Despite recent advances, most existing models require substantial computational resources for training and inference. This poses a challenge for practical automated pig detection, making their use in real-time pig behavior recognition systems often impractical.[41] To address the complexities of monitoring pigs on farms, Zhang *et al*. (2025) introduced YOLOv8-PigLite, a lightweight model designed for identifying pig behavior, and showed that the model maintains consistent precision, recall, and a mean average precision of 50%, while reducing the number of parameters and floating-point operations by 59.80% and 39.50%, respectively. YOLOv8n was utilized to classify pigs as lying, standing, or walking for individual monitoring purposes, aiming to enhance animal welfare. The model achieved an accuracy of 90.66%, with a precision of 90.91% and a recall of 90.66%.[42]

Xiang *et al*. (2025) utilized the oriented bounding box (OBB) target detection model, YOLO v11m-OBB model for real-time body temperature monitoring in group-housed pigs by detecting body temperature through the identification and pairing of the left and right ear roots in thermal images.[30] Liao *et al*. (2025) developed the YOLOv8 integration of the ADown module for segmentation and detection (YOLOv8A-SD), which leverages a dual-task strategy combining segmentation and detection. It incorporates an ADown attention mechanism to improve its ability to extract features from crowded or cluttered farm scenes, achieving 96.1% precision, 96.3% mAP50, and an 83.1% IoU for segmentation.[43] Huang *et al*. (2025) introduced PR-SegFormer, a pig body semantic segmentation model using a Polarized Self-Attention (PSA) module for fine-grained feature extraction. YOLOv8-cls-MCM integrates a parameter-efficient MobileNetV3 backbone with a coordinate attention (CA) mechanism, resulting in a 40% reduction in parameters, 93.16% pose recognition accuracy in complex backgrounds, 93.7% average precision, and 92.95% recall.[44] Therefore, YOLO remains the optimal choice for fast, end-to-end object detection, making it suitable for a variety of applications. It also facilitates continuous tracking, timely anomaly detection, and automated alerts through a web application, thereby enhancing its practicality for smart livestock farming.

Addressing the aforementioned challenges, this study evaluated various models within the YOLO family—YOLOv8n (nano), YOLOv8m (medium), YOLOv8x (extra-large), and a novel approach, YOLOv8ma (modified architecture)—optimized for detecting small and medium objects that has not been extensively explored in the context of group-housed pigs. YOLOv8's architectural enhancements, particularly in feature extraction and generalization, result in significantly improved robustness and reliability. By mitigating bounding box overlap, it provides a more dependable solution for real-world detection scenarios.[45] With specific architectural changes, YOLOv8ma enhances performance while reducing computational demands, making it ideal for applications where both accuracy and resource efficiency are critical. Each model is optimized for different trade-offs between accuracy and computational efficiency. The assessment was conducted using varying levels of augmentation (2x and 3x), thereby expanding the training dataset by two to three times through transformations. The object detection relies on spatial information and does not account for temporal dependencies. To address this limitation, the system classifies behaviors using a series of consecutive frames, facilitating stable predictions over time. Furthermore, post-processing techniques, such as moving average smoothing and temporal consistency filtering, reduce false detections and enhance the reliability of classifications. This approach aimed to identify the most effective model for detecting behavior in group-housed pigs. By leveraging the capabilities of deep learning models, we aim to advance technology-driven solutions for livestock management, ultimately improving outcomes for both animals and producers. This research makes several key contributions:

Methodological Advancement: An efficient algorithm, YOLOv8ma, with a modified architecture is proposed to accurately detect and recognize pig behaviors in intensive farming environments. We conducted a performance comparison of behavior classification using different variations of the YOLOv8 model: YOLOv8n, YOLOv8m, YOLOv8x and YOLOv8ma. Each variation offers a distinct trade-off between computational requirements and performance capabilities.

Performance Evaluation: The different levels of augmentation (2x and 3x) provide valuable insights into their effectiveness in detecting key behaviors in group-housed pigs. This is especially true for 2x augmentation, which can be used to develop highly accurate and efficient animal welfare monitoring systems in agricultural settings.

Web-Based Application: The PigSenseAI allows farmers to self-assess pig behaviors to enhance the efficiency and effectiveness of pig farming while promoting sustainable practices. This application benefits operators in the agricultural industry and offers potential improvements for animal welfare, production efficiency, and farm management.

This innovation addresses the limitations of prior research that has not been applied to real-time monitoring of group-housed pigs. This approach balances accuracy and computational efficiency, making it practical for pig farming. The model features reduced training times and enhanced processing speeds, making it suitable for dynamic environments. This study distinguishes itself by incorporating Gradient-weighted Class Activation Mapping (Grad-CAM) analysis to identify features in both correctly and misclassified images. Furthermore, it enhances classification by removing the concatenate, upsample, and C2fBlock layers from the YOLOv8 model. This refinement enables the model to concentrate on critical features for detecting small and medium objects—an approach that has not been previously utilized in studies of pig behavior detection. These advancements highlight the potential of deep learning in monitoring animals and improving welfare outcomes.

2. Methodology

2.1 Dataset acquisition

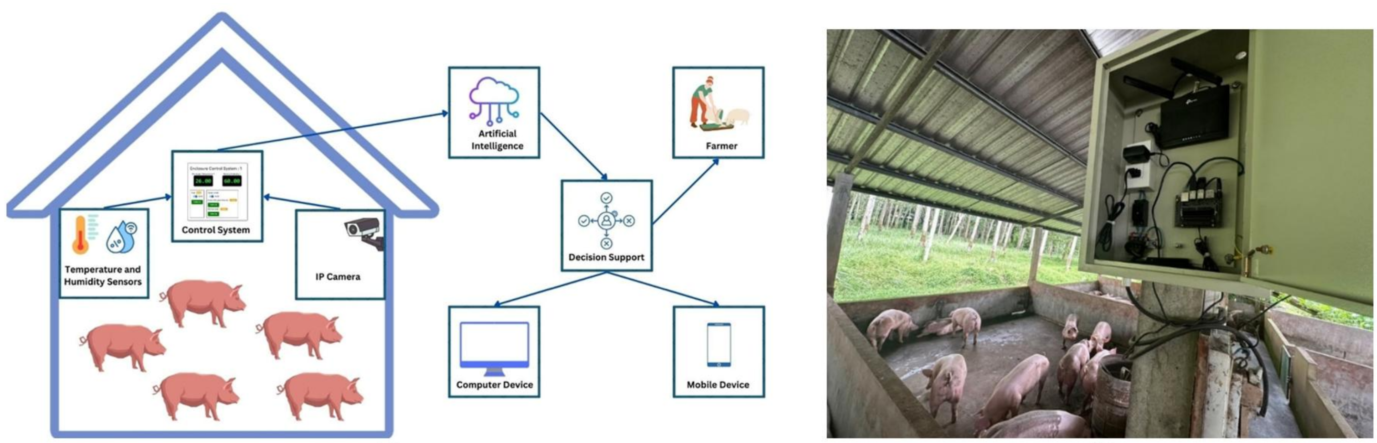
A camera setup was established at a pig farm in Trang Province, Southern Thailand, to collect a dataset of videos capturing pig behavior in a controlled environment. Trang is renowned for its grilled pork (Moo Yang Trang), a local delicacy famous across the country. This tradition is deeply rooted in the region's cultural heritage, passed down through generations. Cameras were strategically positioned to capture a range of behaviors, including drinking, eating, sleeping, and standing. The videos were recorded 24 hours a day throughout October 2023. The farms were divided into three areas, each measuring 5 m × 5 m, accommodating a total of 25 Large White pigs for data collection (Pen 1: 8 pigs, Pen 2: 8 pigs, Pen 3: 9 pigs). Fig. 1 illustrates the monitoring and decision support system developed for group-housed pigs. The dataset consists of a total of 1,300 video recordings, each with a duration of 60 seconds. Each video was captured at a resolution of 2560 × 1440 pixels and recorded at a frame rate of 30 fps. A total of 5,300 images were selected from the recordings, then extracted and annotated to construct the dataset.

Cameras were positioned on the roof to monitor and record pig behaviors, including drinking, eating, sleeping, and standing. This arrangement provided a comprehensive view of the pigpen and captured the animals in a natural environment. The work was conducted following an 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. Fig. 1 illustrates the monitoring and decision support system developed for group-housed pigs. The Micro Processor Unit (MCU) in the farm is powered by the NVIDIA Jetson Nano board, designed for AI Edge systems. This compact, powerful board handles real-time AI tasks such as object detection and classification, enabling local processing without relying on the cloud. The system includes an RGB camera with a Sony CMOS sensor (Resolution: 1920 × 1080), which captures high-resolution color images of farm activities. This setup allows for efficient, real-time monitoring and data analysis, improving farm management and decision-making.

2.2 Data preparation

The recorded videos were carefully reviewed, with each frame annotated to capture the specific behaviors of interest. Fig. 2 provides sample frames illustrating the different behaviors observed in pigs, such as drinking, eating, sleeping, and standing. The criteria for labeling the data were based on visual cues observed in the footage and were applied consistently across all videos (Fig. 2). Drinking behavior was characterized by the pig's snout making contact with the water source. Eating was defined as the pig's mouth making contact with the feed. Standing referred to an upright posture in which the pig remained stationary on all four legs, exhibiting no leg movement or change in location. Sleeping was defined as the pig's body resting on the ground, which included lying down, lying laterally (on its side), or lying sternally (on its belly). The proportion of each annotated behavior is as follows: Drinking: 2,080 frame images (5.47%), Eating: 8,060 frame images (21.29%), Sleeping: 18,610 frame images (48.94%), and Standing: 9,280 frame images (24.40%).

For dataset annotation, tools were used to mark the start and end times of each behavior in the videos. Anchor boxes were configured to align with the typical aspect ratios and scales of pig behaviors observed in the footage, optimizing the model's ability to predict bounding boxes accurately. After annotation, the annotated videos were organized into folders, and the dataset was split into a training and validation set (80%) and a testing set (20%). The training set was used to train the model, the validation set was utilized to tune hyperparameters and evaluate performance, and the testing set was employed to assess the final performance of the trained model for pig behavior detection in a controlled environment.[46] To ensure annotation accuracy, an inter-annotator agreement assessment was conducted utilizing two independent annotators who



(a) (b)

Fig. 1: a) Schematic of the smart monitoring system for pig farming and b) On-site implementation of the smart monitoring system in a pig pen.

labeled a random subset of our video clips. The Kappa coefficient yielded a score of 0.82, thereby affirming the quality and consistency of our behavioral labels.

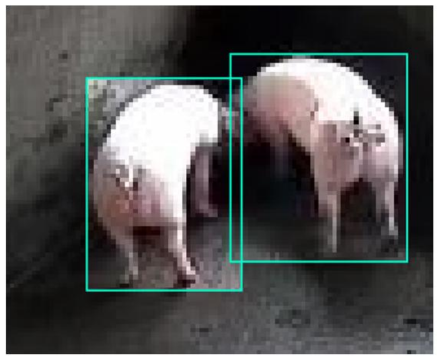
2.3 Data augmentation

To enhance the model's performance and robustness, we applied data augmentation techniques, including random rotation (up to 45°), scaling (80% to 120%), and horizontal flipping. By employing 2x and 3x augmentation, we significantly improved the model's generalization ability on unseen data, expanding our training and validation dataset from 3,710 images to 7,420 images with 2x augmentation and to 11,130 images with 3x augmentation. A fixed testing set of 927 images ensured consistent performance assessment across all models. These augmentations increased the dataset's diversity, exposing the model to a broader range of input variations, thereby enhancing its ability to accurately detect pig behaviors under diverse conditions. This is crucial for evaluating the model's adaptability and performance in real-world scenarios. To maintain experimental consistency, standardized training parameters, including a 640 × 640 input image size, a batch size of 32, and the AdamW optimizer, are applied to all model instantiations. K-means clustering, using Intersection over Union (IoU) as a distance metric, was employed to determine optimal anchor boxes. Hyperparameter optimization, including adjustments to learning rate, batch size, and training epochs, was conducted experimentally to achieve maximum performance in the monitoring and analysis of controlled pig behavior.

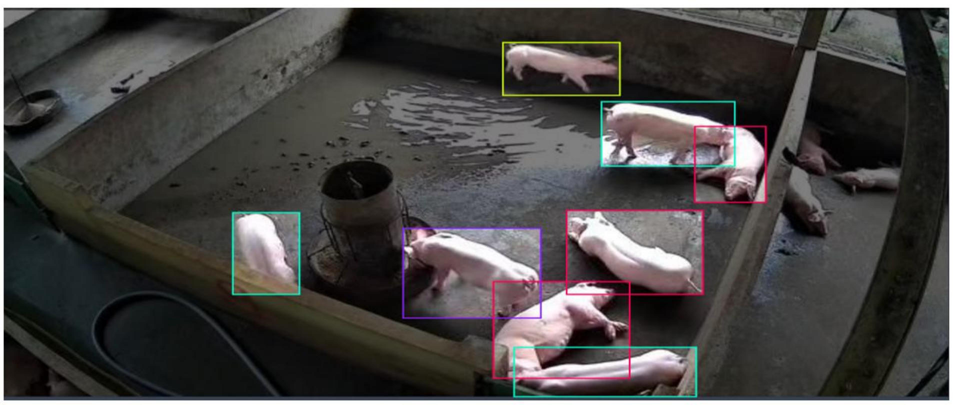
a) Drinking: snout in water source

b) Eating: mouth on feed

c) Sleeping: lying still on its side or stomach



d) Standing: upright position, stationary on 4 legs



d) All

Fig. 2: Sample frames of the pig behaviors a) drinking, b) eating, c) sleeping, d) standing, and e) all the aforementioned behaviors.

2.4 Model selection and configuration

YOLOv8 contains five different configurations, YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x, which gradually increase according to the depth and width of the network.[45,47] While newer versions such as YOLOv9 and YOLOv11 were taken into account, they were either unavailable or insufficiently validated at the time of this study, thereby reinforcing the selection of YOLOv8 as the most suitable framework for our evaluation. In this study, we consider three variations of the YOLOv8 model—YOLOv8n, YOLOv8m, and YOLOv8x—along with a modified version known as YOLOv8ma. Each offers distinct trade-offs between computational demands and performance.[48] Table 1 provides a comprehensive overview of the hyperparameter configurations across these augmentation levels.

YOLOv8n is the smallest, lightest, and fastest variant, specifically designed for low-resource environments. While it delivers high speed, this compact model may sacrifice some accuracy due to its reduced size and fewer parameters.[49,50] YOLOv8m balances speed and accuracy. It is larger than YOLOv8n but still maintains efficient processing, making it suitable for a wider range of tasks that require more precision without significant computational overhead.[51] YOLOv8x, the largest and most powerful variant, provides enhanced accuracy and detection performance; however, it requires significantly more computational power and memory. This study presented YOLOv8ma, a modified version of YOLOv8, which balances accuracy and efficiency. With specific architectural changes to optimize small and medium object detection, it offers improved performance with reduced computational demands, making it suitable for applications where both accuracy and resource efficiency are important. This highlights the trade-off between augmentation-induced performance gains and the associated increase in training time and computational resources. This trade-off underscores the need for careful consideration of augmentation strategies based on the specific requirements and constraints of the application.

2.4.1 YOLOv8

The architecture of YOLOv8 consists of a backbone and a head, which work together to perform tasks such as object detection, segmentation, pose estimation, tracking, and classification (Fig. 3). The backbone extracts features from an input image with dimensions of 640 × 640 × 3. As the image progresses through five Conv blocks, its height and width decrease while the channel depth increases. Between these Conv blocks, the C2f block is strategically placed, representing an advanced version of the CSPLayer from YOLOv5. This cross-stage partial bottleneck with two convolutions effectively merges high-level features with contextual information, enhancing detection accuracy.[52] In the head architecture, feature maps from Conv 3, 4, and 5 are processed through several layers, including concatenate,

Table 1: Hyperparameters comparison of YOLOv8 models with augmentation levels.

Model	Train+Val Images	Test Images	Image size (pixels)	Batch size	Optimizer
YOLOv8n	3710	790	640 × 640	32	AdamW
YOLOv8n-augmentation (2x)	7420	790	640 × 640	32	AdamW
YOLOv8n-augmentation (3x)	11130	790	640 × 640	32	AdamW
YOLOv8m	3710	790	640 × 640	32	AdamW
YOLOv8m-augmentation (2x)	7420	790	640 × 640	32	AdamW
YOLOv8m-augmentation (3x)	11130	790	640 × 640	32	AdamW
YOLOv8x	3710	790	640 × 640	32	AdamW
YOLOv8x-augmentation (2x)	7420	790	640 × 640	32	AdamW
YOLOv8x-augmentation (3x)	11130	790	640 × 640	32	AdamW
YOLOv8ma	3710	790	640 × 640	32	AdamW
YOLOv8ma-augmentation (2x)	7420	790	640 × 640	32	AdamW
YOLOv8ma-augmentation (3x)	11130	790	640 × 640	32	AdamW

upsample, C2f, and Conv, before reaching the detection block, as shown in Fig. 4. At this stage, losses related to bounding boxes and classifications are calculated to facilitate object detection. YOLOv8 employs an anchor-free model with a decoupled head, separately addressing objectness and classification. It uses the sigmoid function for objectness and softmax for class probabilities, leveraging Complete Intersection over Union (CIoU) and Distribution Focal Loss (DFL) to improve its capability, particularly in detecting smaller objects.[53]

2.4.2 YOLOv8ma

While YOLOv8 achieves robust behavior classification and mitigates information loss, its increased model complexity, characterized by a high parameter count, negatively impacts training and inference latency.[54] To improve computational efficiency, we propose YOLOv8ma, a modified architecture for pig behavior recognition. Key components of the YOLOv8ma, including the concatenate, upsample, and C2fBlock layers, were removed to streamline the model (Fig. 4). The concatenate operation, originally used to merge multi-scale feature maps, was eliminated to focus the network on features directly relevant to smaller objects. Removing the upsample layer, which aligns smaller feature maps with larger ones, reduced computational complexity and enhanced efficiency by avoiding unnecessary feature map enlargement. Additionally, the C2fBlock, responsible for feature fusion across two stages, was removed to simplify feature extraction, ensuring the model emphasizes critical features for small and medium object detection.[55] These modifications result in a more focused architecture, leading to faster inference times and improved accuracy for detecting small and medium objects. However, the reduced multi-scale detection capability makes this approach most effective for applications targeting these object sizes, while also decreasing computational demands.[56]

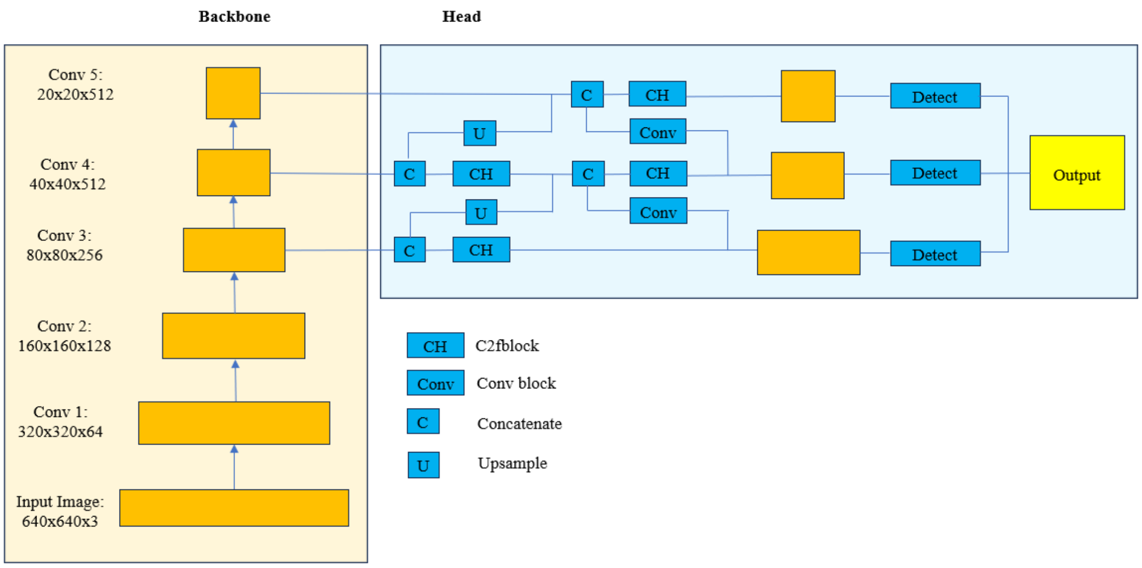


Fig. 3: The architecture of YOLOv8.

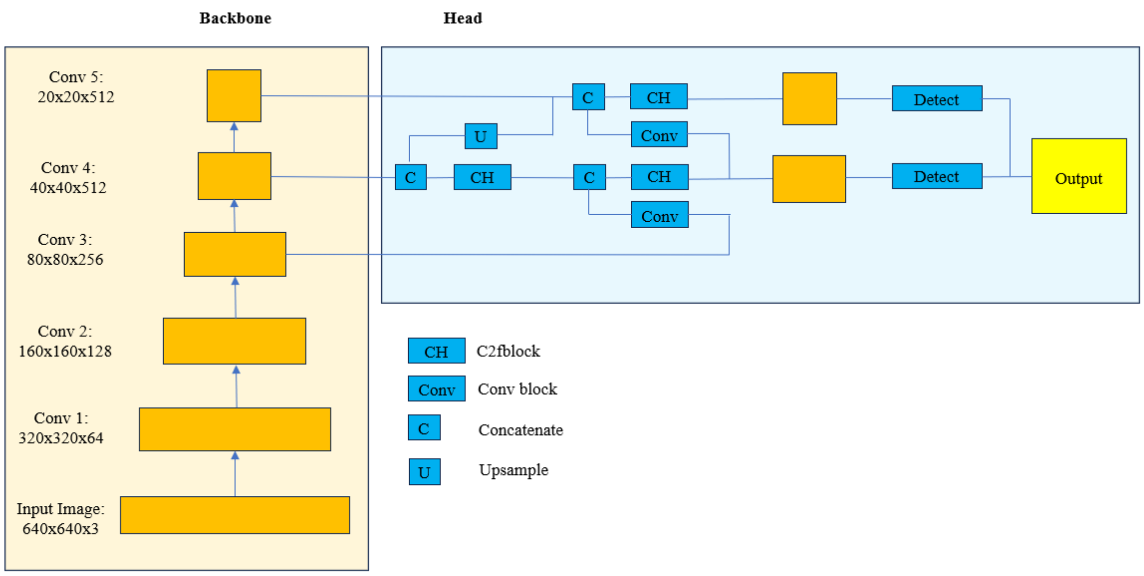


Fig. 4: The architecture of YOLOv8ma (modified architecture).

2.5 Evaluation and performance metrics

The performance of pig behavior detection was evaluated using mean average precision at an Intersection over Union (IoU) threshold of 0.50 (mAP@0.5), precision, and recall using Eqs. (1-3).[57]

$$Precision = \frac{True\ Prositive}{True\ Positive+False\ Positive} \quad (1)$$

$$Recall = \frac{True\ Prositive}{True\ Positive+False\ Negative} \quad (2)$$

$$mAP@0.5 = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (3)$$

Additionally, to gain deeper insights into the decision-making processes of the most effective model, we employed Gradient-weighted Class Activation Mapping (Grad-CAM) to visualize the regions in the images that were most influential in predicting specific pig behaviors. Grad-CAM generates heatmaps that highlight critical areas contributing to the model's decisions by creating a weighted sum of the feature maps, with gradients emphasizing regions relevant to the target class.[58] These heatmaps, overlaid on the original images, provide insights into the model's decision-making process. ReLU activation was applied to focus on positive contributions, resulting in a heatmap that visually indicates where the model is concentrating its attention.

3. Results

3.1 Overall performance on the training dataset

We evaluated the performance of YOLOv8 variants (YOLOv8n, YOLOv8m, YOLOv8x, and YOLOv8ma) in detecting pig behaviors, both with and without augmentation, using mAP@0.5, precision, training time, and training time per epoch (Table 2). All models were trained for 100 epochs on the same training and test datasets. From the table, it can be observed that YOLOv8ma-augmentation (2x) achieved the highest mAP@0.5, while YOLOv8ma-augmentation (3x) attained the highest precision. In comparison to all YOLOv8 variants (YOLOv8n, YOLOv8m, YOLOv8x, and YOLOv8ma), both with and without data augmentation, the YOLOv8ma model enhanced the mAP@0.5 by 0.1% (YOLOv8m - augmentation 2x) to 6.32% (YOLOv8n - augmentation 3x). Furthermore, the YOLOv8ma model demonstrated a precision improvement ranging from 0.11% (YOLOv8m - augmentation 3x) to 4.59% (YOLOv8x).

The YOLOv8ma model has a baseline mAP@0.5 of 0.964 and precision of 0.908, with a training time of 55.4 minutes for 100 epochs, averaging 33.24 seconds per epoch. The 2x augmentation improves mAP@0.5 to 0.976 and precision to 0.922, but training time increases to 114.3 minutes. The 3x augmentation results in a mAP@0.963 and precision of 0.935, requiring 144.5 minutes of training time. Overall, the YOLOv8ma-augmentation (2x) stands out as the most effective configuration, offering the best balance of high accuracy (mAP@0.5 of 0.976) and manageable training time (114.3 minutes), demonstrating the importance of model and augmentation selection to optimize performance for specific applications. Alternatively, the YOLOv8m model also showed superior performance, achieving the high mAP@0.5 and precision (Table 2). Alternatively, the YOLOv8m model has a baseline mAP@0.5 of 0.964 and precision of 0.907, with a training time of 73.2 minutes for 100 epochs, averaging 43.92

Table 2: Performance evaluation of YOLOv8 models with augmentation.

Model	mAP@0.5	Precision	Training Time (min)	Train Time (s/Epoch)
YOLOv8n	0.957	0.925	57.0	34.20
YOLOv8n-augmentation (2x)	0.930	0.902	119.4	71.64
YOLOv8n-augmentation (3x)	0.918	0.890	157.2	94.32
YOLOv8m	0.964	0.907	73.2	43.92
YOLOv8m-augmentation (2x)	0.975	0.921	139.2	83.52
YOLOv8m-augmentation (3x)	0.962	0.934	181.2	108.72
YOLOv8x	0.958	0.894	149.4	89.64
YOLOv8x-augmentation (2x)	0.961	0.921	237.6	142.56
YOLOv8x-augmentation (3x)	0.958	0.920	296.8	178.08
YOLOv8ma	0.964	0.908	55.4	33.24
YOLOv8ma-augmentation (2x)	0.976	0.922	114.3	68.58
YOLOv8ma-augmentation (3x)	0.963	0.935	144.5	86.70

seconds per epoch. Augmentation improves performance, with 2x augmentation achieving a mAP@0.5 of 0.975, the highest among all configurations, though it requires 139.2 minutes of training time. The 3x augmentation results in a slight drop to 0.962 mAP@0.5 but boosts precision to 0.934, requiring 181.2 minutes of training time.

3.2 Behavior-specific performance on the training dataset

To select an appropriate model for pig detection, we performed a behavioral analysis of group-housed pigs across various categories—Drinking, Eating, Sleeping, and Standing—utilizing the training dataset, and the results are summarized in Table 3. According to the results, YOLOv8ma-

Table 3: The mAP@0.5 and precision during the training of object detection models.

	Models	All	Drinking	Eating	Sleeping	Standing
mAP@0.5	YOLOv8n	0.957	0.951	0.973	0.967	0.941
	YOLOv8n-augmentation (2x)	0.930	0.915	0.970	0.970	0.865
	YOLOv8n-augmentation (3x)	0.918	0.889	0.961	0.946	0.873
	YOLOv8m	0.964	0.964	0.986	0.981	0.927
	YOLOv8m-augmentation (2x)	0.975	0.963	0.989	0.987	0.954
	YOLOv8m-augmentation (3x)	0.962	0.957	0.984	0.976	0.931
	YOLOv8x	0.958	0.951	0.959	0.987	0.939
	YOLOv8x-augmentation (2x)	0.961	0.937	0.980	0.970	0.954
	YOLOv8x-augmentation (3x)	0.958	0.955	0.975	0.967	0.936
	YOLOv8ma	0.964	0.965	0.986	0.981	0.928
	YOLOv8ma-augmentation (2x)	0.976	0.963	0.990	0.988	0.955
	YOLOv8ma-augmentation (3x)	0.963	0.957	0.985	0.977	0.931
Precisions	YOLOv8n	0.925	0.939	0.970	0.955	0.844
	YOLOv8n-augmentation (2x)	0.902	0.916	0.946	0.930	0.818
	YOLOv8n-augmentation (3x)	0.890	0.856	0.956	0.937	0.830
	YOLOv8m	0.907	0.873	0.970	0.972	0.829
	YOLOv8m-augmentation (2x)	0.921	0.866	0.970	0.975	0.889
	YOLOv8m-augmentation (3x)	0.934	1.000	0.960	0.937	0.834
	YOLOv8x	0.894	0.798	0.937	0.967	0.880
	YOLOv8x-augmentation (2x)	0.921	0.879	0.970	0.967	0.880
	YOLOv8x-augmentation (3x)	0.920	0.939	0.957	0.938	0.868
	YOLOv8ma	0.908	0.874	0.971	0.973	0.830
	YOLOv8ma-augmentation (2x)	0.922	0.867	0.970	0.976	0.890
	YOLOv8ma-augmentation (3x)	0.935	1.000	0.961	0.938	0.836

Table 4: The mAP@0.5 and precision during the testing of object detection models.

	Models	All	Drinking	Eating	Sleeping	Standing
mAP@0.5	YOLOv8n	0.940	0.933	0.954	0.948	0.925
	YOLOv8n-augmentation (2x)	0.914	0.901	0.951	0.953	0.850
	YOLOv8n-augmentation (3x)	0.902	0.876	0.944	0.929	0.858
	YOLOv8m	0.946	0.945	0.967	0.961	0.910
	YOLOv8m-augmentation (2x)	0.956	0.949	0.970	0.967	0.936
	YOLOv8m-augmentation (3x)	0.944	0.939	0.964	0.957	0.915
	YOLOv8x	0.941	0.935	0.940	0.968	0.922
	YOLOv8x-augmentation (2x)	0.944	0.922	0.962	0.955	0.936
	YOLOv8x-augmentation (3x)	0.941	0.938	0.957	0.948	0.920
	YOLOv8ma	0.947	0.946	0.968	0.962	0.911
	YOLOv8ma-augmentation(2x)	0.957	0.950	0.971	0.968	0.937
	YOLOv8ma-augmentation (3x)	0.945	0.940	0.965	0.958	0.916
Precisions	YOLOv8n	0.912	0.923	0.950	0.941	0.832
	YOLOv8n-augmentation (2x)	0.889	0.900	0.932	0.916	0.806
	YOLOv8n-augmentation (3x)	0.879	0.842	0.939	0.922	0.815
	YOLOv8m	0.895	0.860	0.952	0.953	0.816
	YOLOv8m-augmentation (2x)	0.909	0.854	0.950	0.961	0.872
	YOLOv8m-augmentation (3x)	0.921	1.00	0.944	0.921	0.819
	YOLOv8x	0.881	0.786	0.923	0.948	0.866
	YOLOv8x-augmentation (2x)	0.909	0.867	0.950	0.952	0.866
	YOLOv8x-augmentation (3x)	0.908	0.923	0.940	0.924	0.854
	YOLOv8ma	0.896	0.861	0.953	0.954	0.817
	YOLOv8ma-augmentation (2x)	0.910	0.855	0.951	0.962	0.872
	YOLOv8ma-augmentation (3x)	0.922	1.00	0.945	0.921	0.820

augmentation (2x) and YOLOv8ma-augmentation (3x) achieved the highest mAP@0.5 and precision across all behavior categories. This highlights the effectiveness of YOLOv8ma in balancing high precision and detection accuracy across diverse behaviors, making it a preferred choice for complex object detection tasks (Table 3).

The YOLOv8ma model, with an mAP@0.5 of 0.964, performed comparably to YOLOv8m, excelling in Eating (0.986) and Sleeping (0.981). The YOLOv8ma-augmentation (2x) achieved the highest mAP@0.5 (0.976) and precision (0.922), balancing detection accuracy and reduced false positives. The 3x augmented model showed a slight performance drop (mAP@0.5 of 0.963) but maintained high precision (0.935), making it a strong contender for complex object detection tasks. In contrast, the YOLOv8m model demonstrated strong performance by achieving the high mAP@0.5 of 0.975 with 2x augmentation, particularly excelling in Eating (0.989) and Sleeping (0.987). The YOLOv8m-augmentation (3x) model followed closely with a mAP@0.5 of 0.962, maintaining robustness across all categories, though slightly less effective than the 2x version. Additionally, the YOLOv8m-augmentation (2x) achieved the highest precision at 0.921, demonstrating its ability to minimize false positives while maintaining high detection accuracy

3.3 Behavior-specific performance on the testing dataset

In order to validate the behavioral analysis of group-housed pigs across various categories (Drinking, Eating, Sleeping, and Standing), we evaluated each model using images from the testing dataset (Table 4). These results, consistent with the training outcomes shown in Table 3, demonstrate the effectiveness of YOLOv8ma in achieving a balance between high precision and detection accuracy. Specifically, both YOLOv8ma-augmentation (2x) and YOLOv8ma-augmentation (3x) attained the highest mAP@0.5 and precision across all behavior categories. The results highlight that the YOLOv8ma model with augmentation surpasses the original YOLOv8 in both overall precision and mAP@0.5. As shown in Table 4, YOLOv8m consistently achieves a high mAP@0.5, especially with 2x augmentation, which enhances performance across all activities. YOLOv8m's best result is 0.956 for the All category with 2x augmentation, demonstrating its strong accuracy and robust detection. Data augmentation, whether at 2x or 3x, generally improves model performance, particularly in Eating and Sleeping, with YOLOv8m and YOLOv8ma benefiting most from these enhancements. However, the standing activity remains the most challenging across all models, with lower mAP@0.5 scores, especially in the augmented YOLOv8n variants.

In addition, YOLOv8ma achieves a precision of 0.953 for Eating and 0.954 for Sleeping, but Standing detection remains challenging at 0.817. Augmentation improves performance, particularly in Drinking (1.00 with 3x augmentation), but does not address the issues with Standing detection. Overall, YOLOv8ma offers the most consistent performance across behaviors, despite ongoing challenges with Standing detection. Alternatively, YOLOv8m performs better overall, with high precision for Eating and Sleeping (0.952 and 0.953), and shows significant improvement with augmentation, particularly in Drinking (1.00 with 3x augmentation). However, standing detection remains a challenge, with precision at 0.816 in the base version and 0.819 with 3x augmentation. When pigs are grouped together, their bodies may overlap or be partially obscured, making it harder to identify individual postures. This frequently leads to model confusion and misclassification, consistent with findings reported in prior studies.[54,59]

Moreover, a review of prior studies on the automated recognition of pig behaviors was conducted (Table 5). Many of these studies focused on model accuracy, while few addressed practical deployment challenges, such as the impact of data augmentation on model generalization and integration with web-based systems for continuous monitoring. Guo *et al*. (2023) and Yang *et al*. (2024), used *Joint Detection and Embedding* (JDE) *tracker*, FairMOT, and DeepSORT for individual pig tracking, achieving multi-object tracking accuracy (MOTA) of 88.99%. While effective, these methods often require additional tracking infrastructure and may not generalize well to dynamic, large-scale farm settings.[60,61] Han *et al*. (2023) evaluated a CNN + LSTM (Long short-term memory) model to detect agonistic pig behaviors in a single-space feeding stall. While the model achieved high accuracy in random validation (96.8%), its performance declined in blocking-by-time and blocking-by-feeder validations.[62] These results highlight the need for improved generalization to ensure robustness in real-world scenarios. Our models have demonstrated superior performance, showcasing their potential for practical applications in livestock management by surpassing traditional methods in accuracy, efficiency, and scalability. Therefore, given the detection accuracy and model parameters, we selected the YOLOv8ma-augmentation (2x) model for detecting and identifying pig behaviors in practical applications.

3.4 Implementation results

The performance evaluation of the behavioral analysis of group-housed pigs, utilizing training and testing datasets, shows that YOLOv8ma-augmentation (2x) effectively balances high precision and detection accuracy across various behaviors. Therefore, this section assesses the model's predictions for YOLOv8ma-augmentation (2x) on both labeled and unlabeled data to demonstrate the model's capability to accurately detect and classify various behaviors exhibited by pigs, including drinking, eating, sleeping, and standing (Figs. 5 and 6). As shown in Fig. 5, the image displays several frames of pigs in a pen, labeled with behaviors such as drinking, eating, sleeping, and standing. Each behavior includes a confidence score, indicating the model's certainty in detecting these actions. The YOLOv8ma-augmentation (2x)

Table 5: Summary of deep learning studies on group-housed pigs.

Study	Model used	Detection task	Dataset	Performance metric	Results
[63]	GoogLeNet	Feeding and non-nutritive visits in group-housed pigs	34375 images	Accuracy	99.4% ± 0.6%
[64]	Faster R-CNN	Pig position and posture detection	7277 images	mAP	80.2%
[7]	YOLOv4	Pigs' bodies, heads and tails	583 images	Precision	Bodies: 96%, Tails: 77%, Heads: 66%
[65]	YOLOv4	Pig's head to another pig's rear	2781 images	mAP	92.65% ± 3.74%
[40]	YOLOv4 YOLOv7	Ear-biting outbreak quantification	Dataset 1: 5,197 images Dataset 2: 7,828 images	mAP	Dataset 1: YOLOv4 =98% YOLOv7 =97.5% Dataset 2: YOLOv4 =85.6% YOLOv7 =80.9%
[66]	Velocity threshold	Recumbent and standing	9,634 images	mAP	80.2%
[67]	YOLOv7	Behavioral Abnormalities	3,069 images	mAP	90%
[61]	Joint Detection and Embedding (JDE), FairMOT, and YOLOv5s with DeepSORT	Individual pig tracking	22,384 frames	Multi-object tracking accuracy	JDE: 83.56%, FairMOT: 88.55%, YOLOv5s + DeepSORT: 88.99%
[62]	CNN + LSTM	Contact, Ear-to-body, Head-to-body, Levering, Mounting	15,679 frame videos	Accuracy	96.8% ± 0.1
[68]	YOLOv7-tiny_Pig	Pig tracking	5,400 images	Higher-Order Tracking Accuracy, Multiple Object Tracking Precision, and Identification F1 Score	83.16%, 97.6%, and 91.42%

automates the detection of behaviors within a labeled dataset, demonstrating advanced capabilities in identifying pig behaviors (Fig. 5). However, the standing behavior is observed in multiple frames, yet the confidence scores are not consistently high. This indicates challenges in differentiating standing from other movements, possibly due to minimal movement and feature similarity, which complicate precise detection.

Leveraging the success of our pre-trained YOLOv8ma-augmentation (2x) model, we further investigated its potential by employing it to autonomously process an unannotated dataset (Fig. 6). The findings suggest that the images clearly illustrate the model's ability to detect and classify pig behavior without requiring manual annotations. The results highlight the model's ability to generalize well to unseen data and its potential for application in automated pig behavior monitoring systems. This decrease in manual effort conserves time and enhances the consistency and reliability of the generated data, making the process more efficient and effective. The augmentation process generates variations of the data, enhancing the model's generalization and accuracy in real-world conditions. This method not only saves time and resources but also improves the system's capacity to recognize pig behaviors with greater efficiency and precision.

Fig. 5: Prediction of YOLOv8ma-augmentation (2x) on testing dataset.

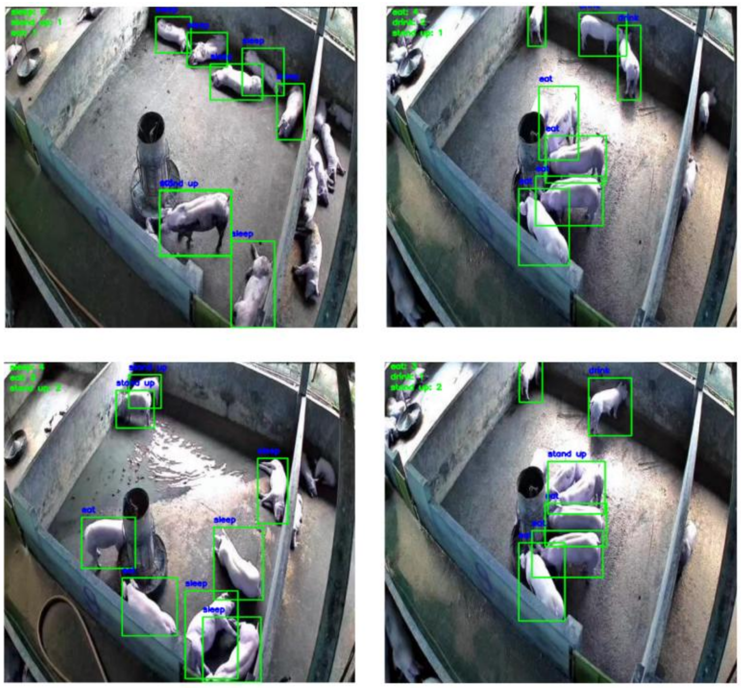


Fig. 6: Prediction of YOLOv8ma-augmentation (2x) on unlabeled data.

3.5 Grad-cam visualization

Fig. 7 visualizes the important regions for detecting pig behavior using the YOLOv8ma-augmentation (2x) model, as enhanced by Grad-CAM. It presents the prediction results for various pig behaviors, thereby enhancing the model's credibility and reliability through visual evidence of its decision-making process. The model successfully extracted key behavioral features by focusing on the right areas of the pigs' bodies for each behavior, showcasing the practical application of YOLOv8ma with augmentation in pig farming. Heatmaps illustrating sleeping behavior reveal that the model concentrates on the pigs' bodies as they lie down, effectively capturing their posture and position. This demonstrates the model's ability to recognize subtle cues, such as orientation and body alignment, essential for detecting sleeping behavior. Similarly, the Grad-CAM visualizations for the standing behavior indicate that the model concentrates on the torso and legs of the pigs. These features are crucial for identifying

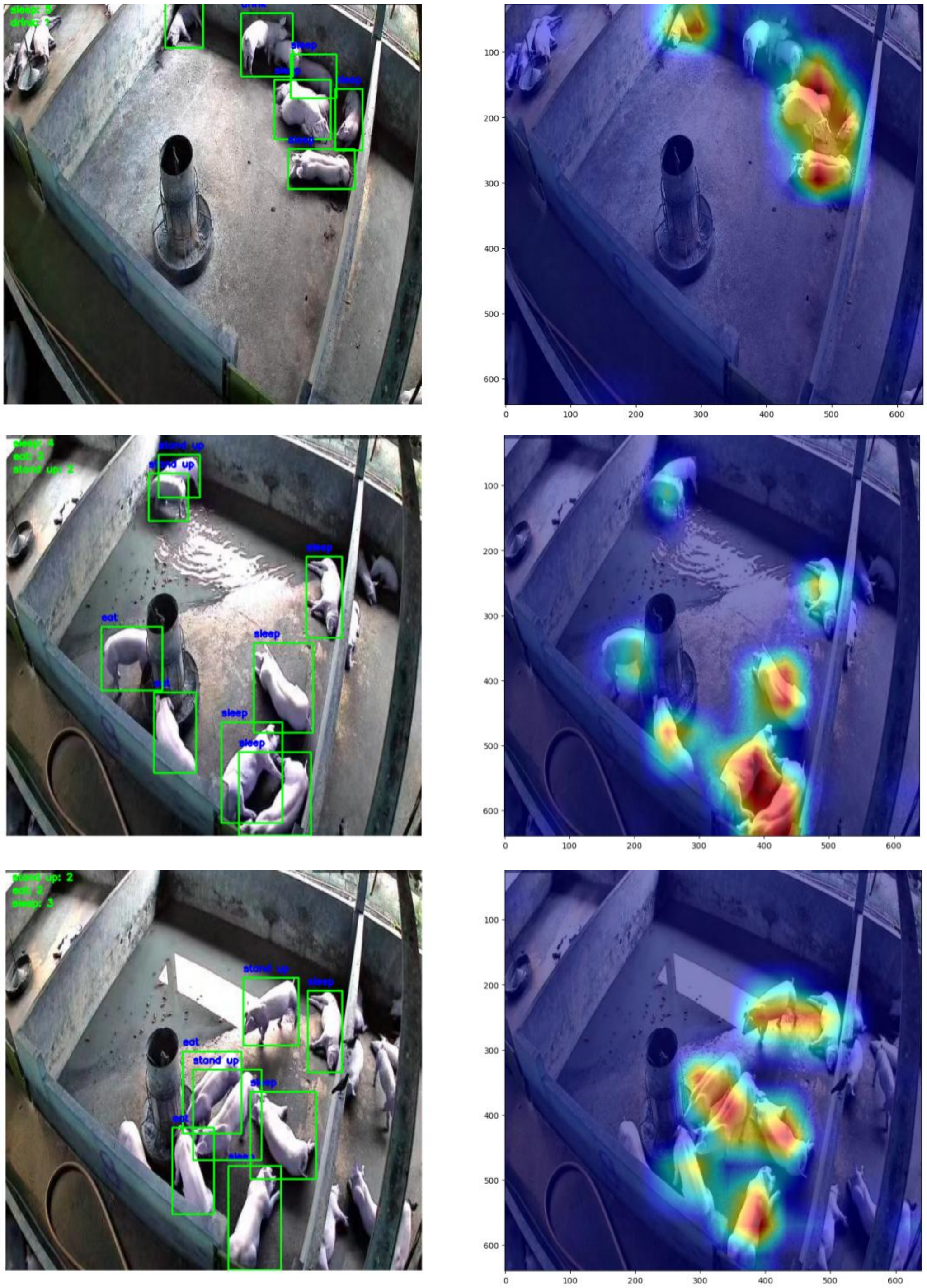


Fig. 7: Visualization of important regions for pig behavior detection using Grad-CAM.

standing activity, highlighting the model's attention to key physical characteristics that differentiate standing from other behaviors. For eating behavior, the heatmaps reveal that the model concentrates on areas around the feeding troughs and the pigs' mouths. These regions are vital for recognizing the act of eating, further confirming the model's capability to accurately identify and classify eating activities by focusing on the appropriate visual cues. This shows that the optimized model fully utilizes contextual information to capture pig behavioral states, extract key features, and improve recognition accuracy.

4. Discussion

4.1 Effectiveness of YOLOv8 models

Table 6 presents a summary of the models' effectiveness in detecting and analyzing pig behaviors across various aspects. The proposed YOLOv8ma algorithm shows notable enhancements in mAP@0.5, average precision, real-time processing, generalization capability, adaptability to environmental conditions, model size and speed, and best use case. YOLOv8ma demonstrated excellent effectiveness in almost all aspect and outperformed the other models with an overall mAP@0.5 score of 0.957 ± 0.02, followed by YOLOv8m with a score of 0.956 ± 0.02, YOLOv8x with a score of 0.942 ± 0.02 and YOLOv8n with 0.919 ± 0.02. In addition, real-time processing was optimal with YOLOv8ma, which achieved 45 frames per second (FPS), making it ideal for fast-paced applications. YOLOv8ma also demonstrated the highest generalization capability (0.92 ± 0.02) and adaptability to environmental conditions (0.93 ± 0.02), underscoring its robustness in varied agricultural environments. In terms of model size and speed, YOLOv8n and YOLOv8ma were the smallest and fastest, making it suitable for resource-limited settings, while YOLOv8x, despite being slightly slower (0.05 sec per frame), provided the most detailed analysis, which could be crucial for in-depth behavior studies.

Table 6: Summary of YOLOv8 models' effectiveness in detecting and analyzing pig behaviors.

Aspect	YOLOv8n (nano)	YOLOv8m (medium)	YOLOv8x (extra-large)	YOLOv8ma (modified architecture)
Overall mAP@0.5 Score	0.919 ± 0.02 (Moderate performance)	0.956 ± 0.02 (Highest performance)	0.942 ± 0.02 (High performance)	0.957 ± 0.02 (Highest performance)
Behavior Detection Accuracy				
Drinking	0.905 ± 0.03 (Good)	0.949 ± 0.03 (Excellent)	0.932 ± 0.02 (Very good)	0.950 ± 0.03 (Excellent)
Eating	0.950 ± 0.02 (Good)	0.967 ± 0.01 (Excellent)	0.962 ± 0.02 (Very good)	0.968 ± 0.01 (Excellent)
Sleeping	0.933 ± 0.02 (Moderate)	0.967 ± 0.02 (Excellent)	0.968 ± 0.01 (Good)	0.968 ± 0.02 (Excellent)
Standing	0.878 ± 0.04 (Good)	0.936 ± 0.03 (Excellent)	0.936 ± 0.02 (Very good)	0.937 ± 0.03 (Excellent)
Impact of 2x Augmentation	0.914 ± 0.02 (Improved performance)	0.956 ± 0.02 (Significantly improved performance)	0.944 ± 0.01 (Improved performance)	0.957 ± 0.02 (Significantly improved performance)
Impact of 3x Augmentation	0.902 ± 0.02 (Improved performance)	0.944 ± 0.01 (Further improved performance)	0.941 ± 0.01 (Improved performance)	0.945 ± 0.01 (Further improved performance)
Real-time Processing	35 FPS (Efficient)	30 FPS (Very efficient)	25 FPS (Efficient)	40 FPS (Very efficient)
Generalization Capability	Moderate (0.88 ± 0.03)	High (0.91 ± 0.02)	High (0.89 ± 0.03)	High (0.92 ± 0.02)
Adaptability to Environmental Conditions	Moderate (0.85 ± 0.03)	High (0.92 ± 0.02)	High (0.89 ± 0.03)	High (0.93 ± 0.02)
Model Size and Speed	Smaller, faster (35 MB, 0.03 sec per frame)	Balanced size and speed (75 MB, 0.04 sec per frame)	Larger, slightly slower (120 MB, 0.05 sec per frame)	Balanced size and speed (35 MB, 0.03 sec per frame)
Best Use Case	General behavior detection	Comprehensive behavior analysis with high accuracy	Detailed behavior analysis requiring high accuracy	Comprehensive behavior analysis with high accuracy

It is important to note that both YOLOv8m and YOLOv8ma are effective in detecting and analyzing pig behaviors, showing similar performance levels. However, YOLOv8ma demonstrates superior performance in real-time applications, with training times for YOLOv8m (2x, 3x) ranging from 43.92 to 108.72 seconds per epoch and a processing speed of 30 FPS. In contrast, YOLOv8ma (2x, 3x) achieves training times of 33.24 to 86.70 seconds per epoch and a processing speed of 40 FPS, making it particularly suited for fast-paced applications. This underscores the trade-offs between model speed, size, and accuracy for specific use cases. This study suggests that the improved model based on the YOLOv8ma algorithm can provide technical support for pig behavior detection in group-raised environments. Its exceptional performance, combined with the advantages of augmentation, positions it as a valuable tool for enhancing animal welfare and improving production efficiency.

4.2 Web application

This study presents a significant advancement in the monitoring of group-housed pigs through the development of a web application for the detection and classification of their behaviors, utilizing the YOLOv8ma-augmentation (2x) model. Fig. 8 shows the web application interface of PigSenseAI, designed to identify and classify essential behaviors such as drinking, eating, sleeping, and standing, enabling real-time monitoring. The deployment of a web application on Vercel hosts detects pig behavior and offers a user-friendly interface for seamless interaction. This method ensures model accessibility from anywhere, providing rapid performance due to Vercel's global network. The invention of a device for classifying group-housed pig behavior introduces technology for quickly and accurately detecting and classifying the behavior of pigs raised in groups. Once the behavior data is classified and processed, the results are displayed as graphs in the web application. In addition to detection and classification capabilities, the web application features a comprehensive

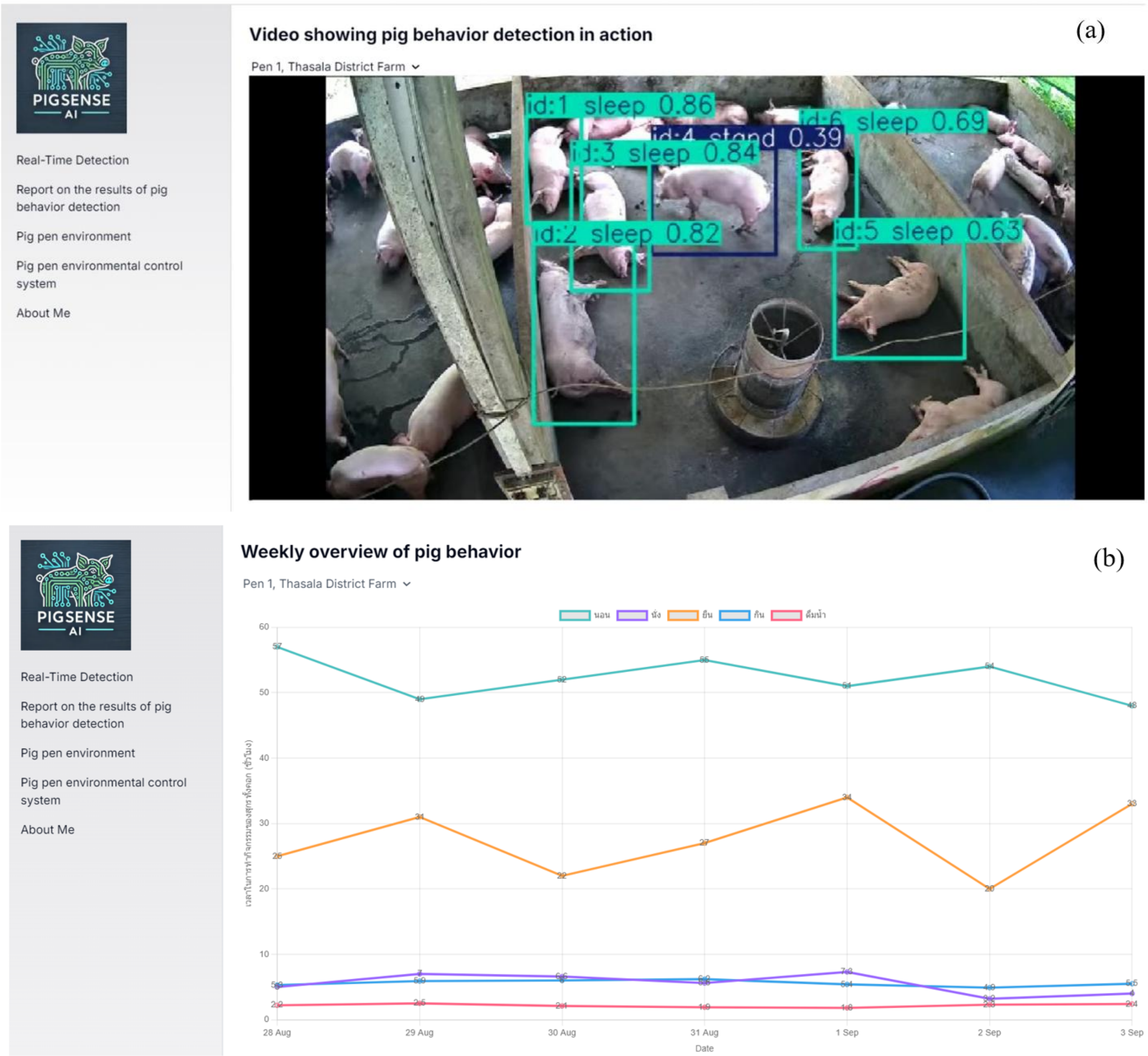


Fig. 8: Web interface of PigSenseAI.

activation of the system both indoors and outdoors. The control interface continuously monitors and manages control system designed for livestock environments, which includes temperature and humidity management. Users can switch the fan between automatic and manual modes. Similar control options are available for water spraying, allowing environmental conditions within the pen to ensure optimal settings for the livestock, providing real-time updates on the status of these systems. This allows for effective monitoring and management of group-housed pig behavior, enhancing the ability to manage pigs efficiently on farms and serving as a critical tool in improving pig farming to be more efficient, environmentally friendly, and sustainable for agricultural operators.

4.3 Limitations of the study

Despite the successful detection of pig behavior using YOLOv8ma-augmentation (2x) model, this study identifies limitations that may affect its real-world application. The absence of temporal information indicates that the integration of tracking or hybrid models could enhance detection reliability. Moreover, the generalizability of the model across diverse farm environments and the optimal camera deployment for high-density settings remain untested. Future work should address these aspects to advance AI-driven livestock monitoring.

5. Conclusion

This study demonstrates the effectiveness of YOLO-based object detection models for accurate and real-time behavior classification in group-housed pigs. By evaluating multiple YOLOv8 variants and introducing the modified YOLOv8ma architecture, the research highlights significant improvements in detection accuracy, precision, and processing speed. The results show that YOLOv8ma with 2x augmentation outperforms other models, achieving a 0.957 mAP@0.5 and 0.910 precision, making it a highly efficient solution for pig behavior monitoring while significantly reducing the number of parameters, the amount of computation, and the model size. The model's ability to process 40 frames per second and its reduced training time further enhance its suitability for real-world agricultural applications. To bridge the gap between research and practical implementation, YOLOv8ma has been integrated into the PigSenseAI web application, creating an automated, real-time system for monitoring drinking, eating, sleeping, and standing behaviors. This study presents a scalable, high-performance alternative to manual observation, enhancing animal welfare, facilitating early detection of health issues, and improving farm management efficiency. Future research should prioritize the integration of temporal modeling techniques, particularly recurrent neural networks (RNNs), to enhance behavior analysis and more effectively capture time-dependent behavior patterns. Additionally, performance optimization can be achieved by improving frames per second (FPS) processing. Incorporating additional sensors, integrating multi-scale features, and utilizing multimodal data will significantly enhance behavior monitoring systems.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

Acknowledgements

This project is funded by National Research Council of Thailand (NRCT): Contract number N24A680482 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.

CRediT Statement

Arsanchai Sukkuea: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project administration, Funding acquisition. **Pensiri Akkajit**: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

References

[1] E. Arulmozhi, A. Bhujel, B. E. Moon, H. T. Kim, The application of cameras in precision pig farming: an overview for swine-keeping professionals, *Animals*, 2021, **11**, 2343, doi: 10.3390/ani11082343.

[2] Y. Gómez, A. H. Stygar, I. J. M. M. Boumans, E. A. M. Bokkers, L. J. Pedersen, J. K. Niemi, M. Pastell, X. Manteca, P. Llonch, A systematic review on validated precision livestock farming technologies for pig production and its potential to assess animal welfare, *Frontiers in Veterinary Science*, 2021, **8**, 660565, doi: 10.3389/fvets.2021.660565.

[3] P. Racewicz, A. Ludwiczak, E. Skrzypczak, J. Składanowska-Baryza, H. Biesiada, T. Nowak, S. Nowaczewski, M. Zaborowicz, M. Stanisz, P. Ślósarz, Welfare health and productivity in commercial pig herds, *Animals*, 2021, **11**, 1176, doi: 10.3390/ani11041176.

[4] H. S. Mluba, O. Atif, J. Lee, D. Park, Y. Chung, Pattern

mining-based pig behavior analysis for health and welfare monitoring, *Sensors*, 2024, **24**, 2185, doi: 10.3390/s24072185.
[5] E. M. Bortoluzzi, M. J. Goering, S. J. Ochoa, A. J. Holliday, J. M. Mumm, C. E. Nelson, H. Wu, B. E. Mote, E. T. Psota, T. B. Schmidt, M. Jaberi-Douraki, L. E. Hulbert, Evaluation of precision livestock technology and human scoring of nursery pigs in a controlled immune challenge experiment, *Animals*, 2023, **13**, 246, doi: 10.3390/ani13020246.
[6] K. Zhang, D. Li, J. Huang, Y. Chen, Automated video behavior recognition of pigs using two-stream convolutional networks, *Sensors*, 2020, **20**, 1085, doi: 10.3390/s20041085.
[7] M. Ocepek, A. Žnidar, M. Lavrič, D. Škorjanc, I. L. Andersen, DigiPig: first developments of an automated monitoring system for body, head and tail detection in intensive pig farming, *Agriculture*, 2022, **12**, 2, doi: 10.3390/agriculture12010002.
[8] J. M. Siegford, J. P. Steibel, J. Han, M. Benjamin, T. Brown-Brandl, J. R. R. Dórea, D. Morris, T. Norton, E. Psota, G. J. M. Rosa, The quest to develop automated systems for monitoring animal behavior, *Applied Animal Behaviour Science*, 2023, **265**, 106000, doi: 10.1016/j.applanim.2023.106000.
[9] J. Wei, X. Tang, J. Liu, Z. Zhang, Detection of pig movement and aggression using deep learning approaches, *Animals*, 2023, **13**, 3074, doi: 10.3390/ani13193074.
[10] M. Hassan, J. H. Park, M. H. Han, 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*, 2023, **13**, 34-46.
[11] J. L. Kleen, R. Guatteo, Precision livestock farming: what does it contain and what are the perspectives? *Animals*, 2023, **13**, 779, doi: 10.3390/ani13050779.
[12] P. Akkajit, A. Sukkuea, B. Thongnonghin, Comparative analysis of five convolutional neural networks and transfer learning classification approach for microplastics in wastewater treatment plants, *Ecological Informatics*, 2023, **78**, 102328, doi: 10.1016/j.ecoinf.2023.102328.
[13] A. Sukkuea, J. Inpun, P. Cherdsukjai, P. Akkajit, Automatic microplastic classification using dual-modality spectral and image data for enhanced accuracy, *Marine Pollution Bulletin*, 2025, **213**, 117665, doi: 10.1016/j.marpolbul.2025.117665.
[14] P. Akkajit, M. E. E. Alahi, A. Sukkuea, Enhanced detection and classification of microplastics in marine environments using deep learning, *Regional Studies in Marine Science*, 2024, **80**, 103880, doi: 10.1016/j.rsma.2024.103880.
[15] L. Shen, H. Tao, Y. Ni, Y. Wang, V. Stojanovic, Improved YOLOv3 model with feature map cropping for multi-scale road object detection, *Measurement Science and Technology*, 2023, **34**, 045406, doi: 10.1088/1361-6501/acb075.
[16] X. Xu, Y. Liu, L. Lyu, P. Yan, J. Zhang, MAD-YOLO: a quantitative detection algorithm for dense small-scale marine benthos, *Ecological Informatics*, 2023, **75**, 102022, doi: 10.1016/j.ecoinf.2023.102022.
[17] A. Sukkuea, P. Akkajit, K. Suwannarat, P. Foithong, N. Afsarimanesh, M. E. E. Alahi, AI-driven time series forecasting of coastal water quality using sentinel-2 imagery: a case study in the gulf of Thailand, *Water*, 2025, **17**, 1798, doi: 10.3390/w17121798.
[18] M. W. Mathis, A. Mathis, Deep learning tools for the measurement of animal behavior in neuroscience, *Current Opinion in Neurobiology*, 2020, **60**, 1-11, doi: 10.1016/j.conb.2019.10.008.
[19] P. Akkajit, A. Sukkuea, Rapid classification of microplastics by using the application of a convolutional neural network, *World Congress on Civil, Structural, and Environmental Engineering*, March 29-31, Avestia Publishing, 2023, doi: 10.11159/iceptp23.126.
[20] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, Y. Miao, Review of image classification algorithms based on convolutional neural networks, *Remote Sensing*, 2021, **13**, 4712, doi: 10.3390/rs13224712.
[21] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, L. Farhan, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, *Journal of Big Data*, 2021, **8**, 53, doi: 10.1186/s40537-021-00444-8.
[22] Y. Liu, H. Pu, D. Sun, Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices, *Trends in Food Science & Technology*, 2021, **113**, 193-204, doi: 10.1016/j.tifs.2021.04.042.
[23] S. Lu, B. Wang, H. Wang, L. Chen, L. Ma, X. Zhang, A real-time object detection algorithm for video, *Computers & Electrical Engineering*, 2019, **77**, 398-408, doi: 10.1016/j.compeleceng.2019.05.009.
[24] J. Wang, N. Wang, L. Li, Z. Ren, Real-time behavior detection and judgment of egg breeders based on YOLO v3, *Neural Computing and Applications*, 2020, **32**, 5471-5481, doi: 10.1007/s00521-019-04645-4.
[25] K. Zhang, Y. Zhang, H. Xu, Lightweight domestic pig behavior detection based on YOLOv8, *Applied Sciences*, 2025, **15**, 6340, doi: 10.3390/app15116340.
[26] Y. Pu, Y. Zhao, H. Ma, J. Wang, A lightweight pig aggressive behavior recognition model by effective integration of spatio-temporal features, *Animals*, 2025, **15**, 1159, doi: 10.3390/ani15081159.
[27] G. Li, R. Jian, X. Jun, G. Shi, A review of you only look once algorithms in animal phenotyping applications, *Animals*, 2025, **15**, 1126, doi: 10.3390/ani15081126.

[28] S. Jha, C. Seo, E. Yang, G. P. Joshi, Real time object detection and trackingsystem for video surveillance system, *Multimedia Tools and Applications*, 2021, **80**, 3981-3996, doi: 10.1007/s11042-020-09749-x.

[29] A. Sukkuea, P. Akkajit, Using convolutional neural network for behavior classification of group-housed pigs, *World Congress on Civil, Structural, and Environmental Engineering*, April 14-16, Avestia Publishing, 2024, doi: 10.11159/iceptp24.120.

[30] R. Xiang, Y. Zhang, H. Lin, Y. Fu, X. Rao, J. Pan, C. Pan, Body temperature detection of group-housed pigs based on the pairing of left and right ear roots in thermal images, *Animals*, 2025, **15**, 642, doi: 10.3390/ani15050642.

[31] A. Mao, E. Huang, X. Wang, K. Liu, Deep learning-based animal activity recognition with wearable sensors: Overview, challenges, and future directions, *Computers and Electronics in Agriculture*, 2023, **211**, 108043, doi: 10.1016/j.compag.2023.108043.

[32] M. M. Dione, P. Oba, Z. Nsadha, K. Asmare, T. J. D. Knight-Jones, R. E. Doyle, The status of pig welfare in selected districts of Uganda: implications for health and productivity interventions, *Frontiers in Animal Science*, 2022, **3**, 878359, doi: 10.3389/fanim.2022.878359.

[33] C. Liu, H. Ye, L. Wang, S. Lu, L. Li, Novel tracking method for the drinking behavior trajectory of pigs, *International Journal of Agricultural and Biological Engineering*, 2023, **16**, 67-76, doi: 10.25165/j.ijabe.20231606.7450.

[34] L. Blavi, D. Solà-Oriol, P. Llonch, S. López-Vergé, S. M. Martín-Orúe, J. F. Pérez, Management and feeding strategies in early life to increase piglet performance and welfare around weaning: a review, *Animals*, 2021, **11**, 302, doi: 10.3390/ani11020302.

[35] M. Kim, Y. Choi, J.-N. Lee, S. Sa, H.-C. Cho, A deep learning-based approach for feeding behavior recognition of weanling pigs, *Journal of Animal Science and Technology*, 2021, **63**, 1453-1463, doi: 10.5187/jast.2021.e127.

[36] B. C. Ramirez, M. D. Hayes, I. C. F. S. Condotta, S. M. Leonard, Impact of housing environment and management on pre-/ post-weaning piglet productivity, *Journal of Animal Science*, 2022, **100**, skac142, doi: 10.1093/jas/skac142.

[37] J. Wu, Q. Yang, D. Xiao, M. Wu, Z. Chen, Q. Hong, Quantifying behavioural patterns for group-housed pigs based on deep learning and statistical analysis, *Computers and Electronics in Agriculture*, 2025, **237**, 110521, doi: 10.1016/j.compag.2025.110521.

[38] K. Yan, B. Dai, H. Liu, Y. Yin, X. Li, R. Wu, W. Shen, Deep neural network with adaptive dual-modality fusion for temporal aggressive behavior detection of group-housed pigs, *Computers and Electronics in Agriculture*, 2024, **224**, 109243, doi: 10.1016/j.compag.2024.109243.

[39] G. Taiwo, S. Vadera, A. Alameer, Vision transformers for automated detection of pig interactions in groups, *Smart Agricultural Technology*, 2025, **10**, 100774, doi: 10.1016/j.atech.2025.100774.

[40] A. Odo, R. Muns, L. Boyle, I. Kyriazakis, Video analysis using deep learning for automated quantification of ear biting in pigs, *IEEE Access*, 2023, **11**, 59744-59757, doi: 10.21227/qy7j-cq11.

[41] Y. Luo, K. Lin, Z. Xiao, E. Lv, X. Wei, B. Li, H. Lu, Z. Zeng, PBR-YOLO: a lightweight piglet multi-behavior recognition algorithm based on improved yolov8, *Smart Agricultural Technology*, 2025, **10**, 100785, doi: 10.1016/j.atech.2025.100785.

[42] F. Deutch, M. G. Weiss, S. R. Wagner, L. S. Hansen, F. Larsen, C. Figueiredo, C. Moers, A. K. Keller, Machine-learning-based activity tracking for individual pig monitoring in experimental facilities for improved animal welfare in research, *Sensors*, 2025, **25**, 785, doi: 10.3390/s25030785.

[43] Y. Liao, Y. Qiu, B. Liu, Y. Qin, Y. Wang, Z. Wu, L. Xu, A. Feng, YOLOv8A-SD: a segmentation-detection algorithm for overlooking scenes in pig farms, *Animals*, 2025, **15**, 1000, doi: 10.3390/ani15071000.

[44] N. Huang, D. Zhu, L. He, Y. Li, The pig pose recognition method based on pig body semantic segmentation, *Smart Agricultural Technology*, 2025, **12**, 101309, doi: 10.1016/j.atech.2025.101309.

[45] E. Casas, L. Ramos, E. Bendek, F. Rivas-Echeverria, YOLOv5 *vs*. YOLOv8: performance benchmarking in wildfire and smoke detection scenarios, *Journal of Image and Graphics*, 2024, **12**, 127-136, doi: 10.18178/joig.12.2.127-136.

[46] G. Varoquaux, O. Colliot, Evaluating machine learning models and their diagnostic value, *Machine Learning for Brain Disorders*, New York, NY: Springer US, 2023, 601-630, doi: 10.1007/978-1-0716-3195-9_20.

[47] J. Wei, X. Tang, J. Liu, Z. Zhang, Detection of pig movement and aggression using deep learning approaches, *Animals*, 2023, **13**, 3074, doi: 10.3390/ani13193074.

[48] B. Yilmaz, U. Kutbay, YOLOv8-based drone detection: performance analysis and optimization, *Computers*, 2024, **13**, 234, doi: 10.3390/computers13090234.

[49] A. Khan, A. Sohail, U. Zahoora, A. S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, *Artificial Intelligence Review*, 2020, **53**, 5455-5516, doi: 10.1007/s10462-020-09825-6.

[50] F. Saeed, S. Aldera, A. A. Al-Shamma'a, H. M. Hussein Farh, Rapid adaptation in photovoltaic defect detection: Integrating CLIP with YOLOv8n for efficient learning, *Energy Reports*, 2024, **12**, 5383-5395, doi: 10.1016/j.egyr.2024.11.033.

[51] A. Elhanashi, P. Dini, S. Saponara, Q. Zheng, TeleStroke: real-time stroke detection with federated learning and YOLOv8

on edge devices, *Journal of Real-Time Image Processing*, 2024, **21**, 121, doi: 10.1007/s11554-024-01500-1.
[52] J. Terven, D. M. Córdova-Esparza, J. A. Romero-González, A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS, *Machine Learning and Knowledge Extraction*, 2023, **5**, 1680-1716, doi: 10.3390/make5040083.
[53] K. Wu, Y. Chen, Y. Lu, Z. Yang, J. Yuan, E. Zheng, SOD-YOLO: a high-precision detection of small targets on high-voltage transmission lines, *Electronics*, 2024, **13**, 1371, doi: 10.3390/electronics13071371.
[54] S. Jahn, G. Schmidt, L. Bachmann, H. Louton, T. Homeier-Bachmann, and A. K. Schütz, Individual behavior tracking of heifers by using object detection algorithm YOLOv4, *Frontiers in Animal Science*, 2025, **5**, 1499253.
[55] A. Betti, M. Tucci, YOLO-S: a lightweight and accurate YOLO-like network for small target detection in aerial imagery, *Sensors*, 2023, **23**, 1865, doi: 10.3390/s23041865.
[56] W. Fang, L. Wang, P. Ren, Tinier-YOLO: a real-time object detection method for constrained environments, *IEEE Access*, 2019, **8**, 1935-1944, doi: 10.1109/ACCESS.2019.2961959.
[57] R. Thangaraj, S. Anandamurugan, V. K. Kaliappan, Automated tomato leaf disease classification using transfer learning-based deep convolution neural network, *Journal of Plant Diseases and Protection*, 2021, **128**, 73-86, doi: 10.1007/s41348-020-00403-0.
[58] V. Kumar, H. M. Azamathulla, K. V. Sharma, D. J. Mehta, K. T. Maharaj, The state of the art in deep learning applications, challenges, and future prospects: a comprehensive review of flood forecasting and management, *Sustainability*, 2023, **15**, 10543, doi: 10.3390/su151310543.
[59] A. J. Scaillierez, T. Izquierdo García-Faria, H. Broers, S. E. van Nieuwamerongen-de Koning, R. P. P. J. van der Tol, E. A. M. Bokkers, I. J. M. M. Boumans, Determining the posture and location of pigs using an object detection model under different lighting conditions, *Translational Animal Science*, 2024, **8**, txae167, doi: 10.1093/tas/txae167.
[60] Q. Yang, X. Hui, Y. Huang, M. Chen, S. Huang, D. Xiao, A long-term video tracking method for group-housed pigs, *Animals*, 2024, **14**, 1505, doi: 10.3390/ani14101505.
[61] Q. Guo, Y. Sun, C. Orsini, J. E. Bolhuis, J. de Vlieg, P. Bijma, P. H. N. de With, Enhanced camera-based individual pig detection and tracking for smart pig farms, *Computers and Electronics in Agriculture*, 2023, **211**, 108009, doi: 10.1016/j.compag.2023.108009.
[62] J. Han, J. Siegford, D. Colbry, R. Lesiyon, A. Bosgraaf, C. Chen, T. Norton, J. P. Steibel, Evaluation of computer vision for detecting agonistic behavior of pigs in a single-space feeding stall through blocked cross-validation strategies, *Computers and Electronics in Agriculture*, 2023, **204**, 107520, doi: 10.1016/j.compag.2022.107520.
[63] A. Alameer, I. Kyriazakis, H. A. Dalton, A. L. Miller, J. Bacardit, Automatic recognition of feeding and foraging behaviour in pigs using deep learning, *Biosystems Engineering*, 2020, **197**, 91-104, doi: 10.1016/j.biosystemseng.2020.06.013.
[64] M. Riekert, A. Klein, F. Adrion, C. Hoffmann, E. Gallmann, Automatically detecting pig position and posture by 2D camera imaging and deep learning, *Computers and Electronics in Agriculture*, 2020, **174**, 105391, doi: 10.1016/j.compag.2020.105391.
[65] A. Alameer, S. Buijs, N. O'Connell, L. Dalton, M. Larsen, L. Pedersen, I. Kyriazakis, Automated detection and quantification of contact behaviour in pigs using deep learning, *Biosystems Engineering*, 2022, **224**, 118-130, doi: 10.1016/j.biosystemseng.2022.10.002.
[66] A. Kühnemund, S. Götz, G. Recke, Automatic detection of group recumbency in pigs *via* AI-supported camera systems, *Animals*, 2023, **13**, 2205, doi: 10.3390/ani13132205.
[67] D. D. Tran, N. D. Thanh, Pig health abnormality detection based on behavior patterns in activity periods using deep learning, *International Journal of Advanced Computer Science and Applications*, 2023, **14**, 603-610, doi: 10.14569/ijacsa.2023.0140564
[68] Q. Yang, X. Hui, Y. Huang, M. Chen, S. Huang, D. Xiao, A long-term video tracking method for group-housed pigs, *Animals*, 2024, **14**, 1505, doi: 10.3390/ani14101505.