Translational Animal Science, 2024, **8**, txae167
https://doi.org/10.1093/tas/txae167
Advance access publication 3 December 2024
Technology in Animal Science

Determining the posture and location of pigs using an object detection model under different lighting conditions

Alice J. Scaillierez,[†] Tomás Izquierdo García-Faria,[‡] Harry Broers,[‖] Sofie E. van Nieuwamerongen - de Koning,[†] Rik P.P.J. van der Tol,[$] Eddie A.M. Bokkers,[†] and Iris J.M.M. Boumans[†]

[†]Animal Production Systems group, Wageningen University & Research, P.O. Box 338, 6700 AH Wageningen, The Netherlands
[‡]Wageningen Livestock Research, Wageningen University & Research, P.O. Box 338, 6700 AH Wageningen, The Netherlands
[‖]Signify Research, Signify, High Tech Campus 7, 5656 AE Eindhoven, The Netherlands
[$]Agricultural Biosystems Engineering group, Wageningen University & Research, P.O. Box 16, 6700 AA Wageningen, The Netherlands

Abstract

Computer vision techniques are becoming increasingly popular for monitoring pig behavior. For instance, object detection models allow us to detect the presence of pigs, their location, and their posture. The performance of object detection models can be affected by variations in lighting conditions (e.g., intensity, spectrum, and uniformity). Furthermore, lighting conditions can influence pigs' active and resting behavior. In the context of experiments testing different lighting conditions, a detection model was developed to detect the location and postures of group-housed growing-finishing pigs. The objective of this paper is to validate the model developed using YOLOv8 detecting standing, sitting, sternal lying, and lateral lying pigs. Training, validation, and test datasets included annotation of pigs from 10 to 24 wk of age in 10 different light settings; varying in intensity, spectrum, and uniformity. Pig detection was comparable for the different lighting conditions, despite a slightly lower posture agreement for warm light and uneven light distribution, likely due to a less clear contrast between pigs and their background and the presence of shadows. The detection reached a mean average precision (mAP) of 89.4%. Standing was the best-detected posture with the highest precision, sensitivity, and F1 score, while the sensitivity and F1 score of sitting was the lowest. This lower performance resulted from confusion of sitting with sternal lying and standing, as a consequence of the top camera view and a low occurrence of sitting pigs in the annotated dataset. This issue is inherent to pig behavior and could be tackled using data augmentation. Some confusion was reported between types of lying due to occlusion by pen mates or pigs' own bodies, and grouping both types of lying postures resulted in an improvement in the detection (mAP = 97.0%). Therefore, comparing resting postures (both lying types) to active postures could lead to a more reliable interpretation of pigs' behavior. Some detection errors were observed, e.g., two detections for the same pig were generated due to posture uncertainty, dirt on cameras detected as a pig, and undetected pigs due to occlusion. The localization accuracy measured by the intersection over union was higher than 95.5% for 75% of the dataset, meaning that the location of predicted pigs was very close to annotated pigs. Tracking individual pigs revealed challenges with ID changes and switches between pen mates, requiring further work.

Lay Summary

In this study, a computer vision model was developed to detect growing-finishing pigs and determine their location and posture in a conventional farm setting with different lighting conditions. This model managed to locate and accurately detect standing and lying pigs, however, detecting sitting pigs and differentiating pigs lying either on the side or on the belly was more complicated. The presence of unevenly distributed and warm light slightly reduced the percentage of postures correctly predicted. The position of pigs in the pen was accurately predicted by the model, although some errors occurred, such as pigs not being detected when occluded or in close proximity to other pigs, detection of non-existing pigs due to dirt on the camera and double detection of pigs in ambiguous postures for the model. Lastly, the model showed high accuracy in differentiating lying pigs (i.e., on the side or on the belly) from non-lying pigs (i.e., standing or sitting).

Key words: behavior, computer vision, deep learning, precision livestock farming, single-stage detector, swine

Introduction

The technologies used in the domain of precision livestock farming aim at automatically monitoring animals, and the use of such technologies has intensified over the last three decades (Tzanidakis et al., 2021). These technologies are based on a wide variety of sensors; for example radio-frequency identification tags, load cells, photoelectric sensors, microphones, thermal cameras, and RGB cameras (Gómez et al., 2021). Though embedded sensors can provide accurate information, for example, on active and nest-building behavior in sows (Cornou et al., 2011; Oczak et al., 2015), mounting embedded sensors on animals could possibly impair the animals' integrity or cause stress due to handling (Gómez et al., 2021). Moreover, the use of embedded sensors is barely possible in group-housed pigs due high risk of sensor destruction or loss. These issues are avoided when using camera technology in pig farming. The more recent introduction of computer vision

Received November 19, 2024 Accepted December 2, 2024.

Table 1. Description of the different lighting conditions tested during the photophase

Light treatment (corresponding experiment)	Average light intensity (in lux)	Average light uniformity coefficient (no unit)	Average correlated color temperature (in K)
Low intensity (exp. 1)	45	0.77	8072
Medium intensity (exp. 1)	198	0.83	8014
High intensity (exp. 1)	968	0.83	7597
Spatial gradient (exp. 1)	198 (from 71 to 330)	0.33	7452
Reference TL (exp. 2 to 3)	72	0.56	3787
Warm white LED (exp. 2)	67	0.79	2594
Forest white LED (exp. 2)	68	0.81	4336
Cool white LED (exp. 2)	68	0.83	6235
Dynamic daylight LED (exp. 3)	129	0.79	4950

techniques using cameras allows us to automatically retrieve information on animals, e.g., count the number of pigs, estimate their weight, localize them in a pen, and determine their behavior (Li et al., 2022).

In pig behavioral studies, RGB cameras are widely used alongside manual observation and scoring of video recordings, which is time-consuming (Tzanidakis et al., 2021). Recent developments of computer vision in research allow the analysis of a larger amount of video data and require less labor once the technology is ready to use. Those techniques continue to improve and appear as an alternative to manual observations (Li et al., 2022). Research in computer vision focuses on better detection of pig behaviors such as feeding, aggression, tail-biting, lying, and postural changes (Gómez et al., 2021). Detection of pig postures has been used in studies to assess the behavioral responses to immune challenges or environmental changes, for example in thermal conditions (Nasirahmadi et al., 2015, 2017; Fernández-Carrión et al., 2017). So far, the detection of pigs' posture has been performed mainly with two methods: (1) image segmentation, which involves identifying and outlining the object (Nasirahmadi et al., 2015, 2017; Riekert et al., 2020, 2021), and (2) object detection, which involves locating the object and placing a bounding box around it (Alameer et al., 2020; Ji et al., 2022; Jeon et al., 2024). Although both methods show similar performances, object detection is the preferred method as it is less time-consuming to train and is known to be faster and more efficient in computer resources (Chen et al., 2023).

A limitation in the currently developed detection models is their sensitivity to lighting conditions that introduce noise, impairing their performance to detect pigs (Tian et al., 2019; Li et al., 2022). As part of a project investigating the effect of light on pig welfare, video data of group-housed pigs were collected under different lighting conditions. Domestic pigs usually show diurnal activity patterns (Ingram and Dauncey, 1985; Simonsen, 1990), which can be influenced by lighting conditions e.g., light intensity (Christison, 1996; Kim et al., 2021; Liu et al., 2022). In this project, various lighting conditions were tested, necessitating the development of a detection model capable of accurately detecting pig postures that reflect their activity regardless of the different light intensities, light spectrum, and light uniformity. Therefore, the objective of this study was to develop and validate an object detection model that identifies individual group-housed growing-finishing pigs and provides their location in the pen and posture under different lighting conditions.

Material and Methods

The detection model was developed using data collected in a conventional commercial pig farm during experiments investigating the effects of different light treatments on pig welfare (see below for details). As these experiments did not include invasive treatments or measurements, no authorization from the Animal Welfare Authority of Wageningen University & Research (Wageningen, the Netherlands) was required, in compliance with the Dutch (Article 1 Wet op de Dierproeven, Rijksoverheid (2021)) and European legislations (European Directive 2010/63/EU (European Union, 2010)).

Animals And Data Collection

In three experiments, five batches (two batches in experiment 1, two batches in experiment 2, and one batch in experiment 3) of 224 pigs (TN70 × TN Tempo crossbreds, Topigs-Norsvin, the Netherlands) were studied from 9 to a maximum of 24 wk of age. Pigs were housed in groups of either seven females or males per pen, balanced for body weight per light treatment. All pigs had pink skin without black spots and were marked on the back with line patterns and numbers using blue and green livestock spray for individual recognition during manual behavioral observations from video recordings. Pigs were housed in four rooms containing eight pens each. Each pen included three different floor types, i.e., metal slats, solid concrete, and concrete slats, and each floor type accounted for about one-third of the pen. A one-space feeder with a drinking nipple giving semi-ad libitum access to feed, a metal chain, and a cotton rope were present in each pen. More details about the layout of the growing-finishing unit are available in Scaillierez et al. (2024).

Nine different lighting conditions were used during the three experiments conducted from September 2021 to July 2023. Light intensities, light uniformity coefficients, and correlated color temperatures of every lighting condition are presented in Table 1. For all light treatments, the photoperiod was between 07:00 to 18:00 (11L:13D) and no daylight entered the rooms as windows were covered with cardboard. During the scotophase, the light intensity was 0 lux.

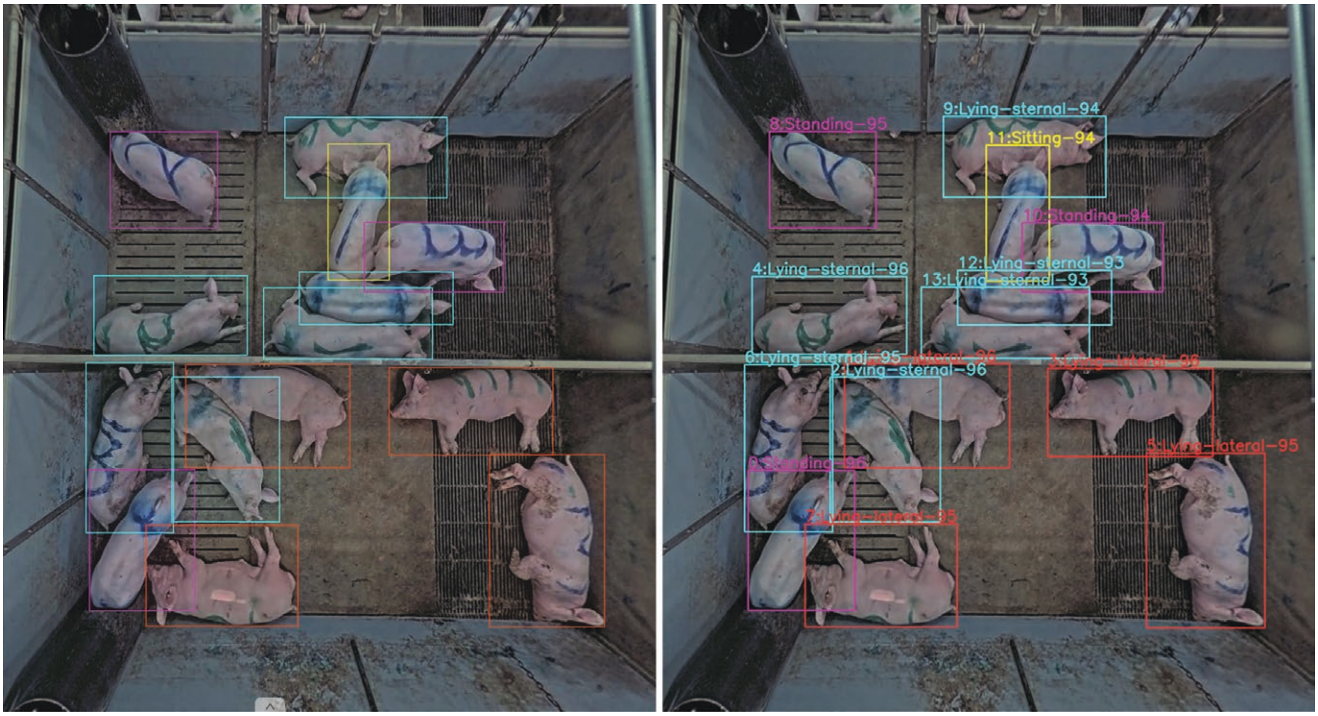

Figure 1. Example of an annotated frame with two pens on the left (red box: lateral lying, blue box: sternal lying, purple box: standing, yellow box: sitting), and example of output provided by the detection model on the same frame on the right. The caption above each box represents the number of the bounding box, the assigned posture label, and the bounding box confidence value.

In each room, two RGB cameras (Axis M3057-PLVE Mk II Network Camera, Axis Communications, Sweden) were mounted on the ceiling (height: 2.4 m) resulting in a top-down view with a resolution of 1920 × 1440 pixels and a frame rate of 25 FPS with images corrected for lens distortion and white balance. During the scotophase, camera-embedded infrared illumination was activated to allow night vision in black and white. The view of each of the eight cameras covered two adjacent pens (Figure 1), thus video data were collected in 16 pens in total. Each week, pens were recorded for 30 h continuously, from 18:00 until 00:00 the day after. During this period of recording no other activities took place than the daily routines of the farmer.

Description Of The Detection Model

The framework used to detect pigs' posture was the YOLOv8x (You Only Look Once, version eight extra-large) model since it is the most precise model of the YOLOv8 series according to Ultralytics (2023). YOLOv8x is a single-stage detector, meaning that object detection is performed in one step. Its architecture consists of two parts, both using convolutional neural network, to first extract image features and second predict the presence of an object, i.e., a pig (Sohan et al., 2024). Further information on the specifications and architecture of the YOLO framework can be found in Sohan et al. (2024). YOLOv8x was developed by Ultralytics and was by default pre-trained on the object detection COCO (Common Objects in Context) dataset containing 330,000 images comprising objects from 80 categories like cars, persons, boats, etc. (Ultralytics, 2023). The steps to train the detection model and assess it, along with the data used for each step are schematized in Figure 2.

Data Annotation And Model Training

To train the model, the postures of pigs were manually annotated on a total number of 1,243 frames selected from videos with various lighting conditions, ranging from light (as described in Table 1) to dark during nighttime (low intensity: 157 frames, medium intensity: 122 frames, high intensity: 100 frames, spatial gradient: 256 frames, reference TL (fluorescent tube light): 75 frames, warm white: 75 frames, forest white: 75 frames, cool white: 75 frames, darkness: 308 frames). A minimum of a hundred annotated frames in each of the four light conditions differing in intensity were selected, and 75 for the four light conditions differing in light spectrum. In addition, more frames were annotated for light conditions in which the detection performance of the model was expected to be lower (e.g., in darkness, low intensity, and spatial gradient). Frames were collected at different times of the day from different rooms (room A: 236 frames, room B: 457 frames, room C: 250 frames, room D: 300 frames) with growing-finishing pigs at different ages (from 10 to 24 wk old, with each age week ranging from 10 to 319 frames), to ensure that the model could detect all pig sizes under all experimental conditions.

Manual annotations were performed using the annotation platform CVAT.ai and consisted of drawing a bounding box around each visible pig and labeling the box with the pig posture (see Figure 1, left picture). Bounding boxes included all the visible parts of a pig and were in contact with the extreme left, right, highest, and lowest edges of the annotated pig. Only pigs in the region of interest (the two recorded pens) were annotated. The posture labels of bounding boxes were assigned according to the ethogram in Table 2.

In the dataset of 1,243 frames, the first 540 manually annotated frames by three observers (first author and 2

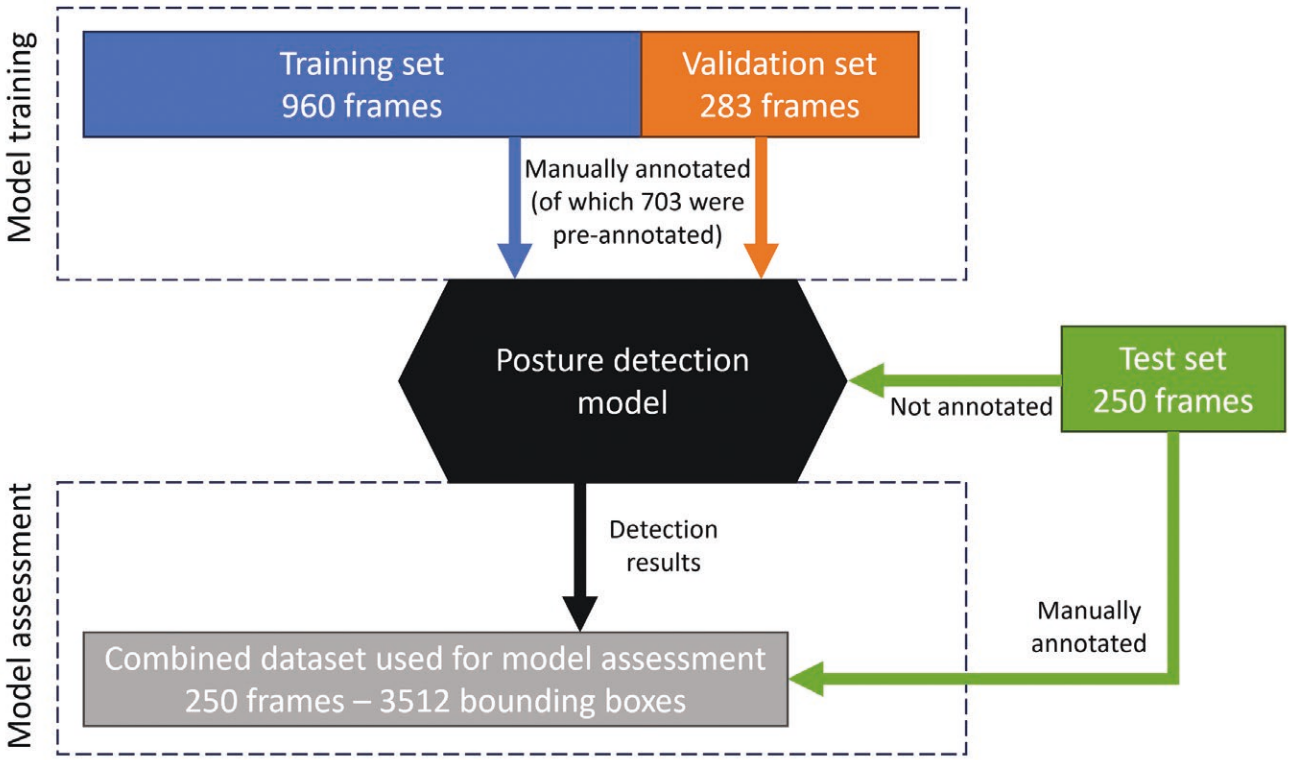

Figure 2. Schematic overview of the dataset used for model training and model assessment. Frames in the training, validation, and test sets were obtained from multiple videos including different lighting conditions, rooms, pens, and pigs at different ages. Rectangles represent datasets, dotted outlines represent phases to train and assess the model, and the black hexagon represents the model.

Table 2. Ethogram for the annotated postures

Posture (label of the bounding box)	Description
Lateral lying	Pig is lying on its side, with all limbs extended. The back of the pig is not visible or hardly visible.
Sternal lying	Pig is lying on its belly with at least one limb tucked below the body. The back of the pig is visible.
Sitting	Pig has its front legs extended and the hind quarter is in contact with the floor.
Standing	Pig is in an upright position with extended legs, and contact with the floor only via hooves.

trained students) were used for a preliminary model. This preliminary model provided some pre-annotations with suggested bounding boxes and posture labels for 703 frames to accelerate the annotation process. Thereafter, all the annotated and pre-annotated bounding boxes and their posture labels of the 1,243 frames were reviewed and corrected when deemed necessary by the first author to limit method and observer bias. Two subsets of the data were created during model training (Figure 2), the training set contained 960 frames to learn to draw the bounding box and associate the features of an image with the assigned labels (Eelbode et al., 2021). The remaining 283 frames were included in the validation set. This validation set allows to assess the performance of the learning process of the model and adjust its hyperparameters (architecture settings; e.g., number of convolutional layers or training epochs) for model improvement (Eelbode et al., 2021).

The YOLO model to detect posture was trained using the PyTorch-based library of YOLOv8x from Ultralytics. The model development was performed on a computer with 12th Gen Intel® Core™ i9-12900 processor 2.40 GHz and 32 GB, with a NVIDIA Geforce RTX3080 graphic processing unit and a memory of 10 GB. The final hyperparameters of the model were an image size of 1280, a learning rate of 0.01, a batch size of 2 frames, a number of epochs of 200 with early-stopping when the performance of the model did not improve for 50 epochs in a row (patience parameter), and the optimizer in auto mode. The loss curves for training and validation are provided in Figure 3, and the model loss function included three components; the focal loss for classification (addressing class imbalance), the IoU loss for localization (addressing the precision to localize objects), the objectness loss (addressing the likelihood to contain object in parts of the image). After running the model on a video, the output of YOLO contains for each frame, the number of pigs detected and the number of pigs for each of the four postures. For each pig, information is given on the detected posture, the x and y coordinates of the bounding box center calculated from the average between the minimum and maximum x or y coordinates of the box, the box confidence reflecting the likelihood of a bounding box to contain an object of a certain class, i.e., a pig in one of the four postures, and the number of the bounding box (new number generated when a new bounding box is created).

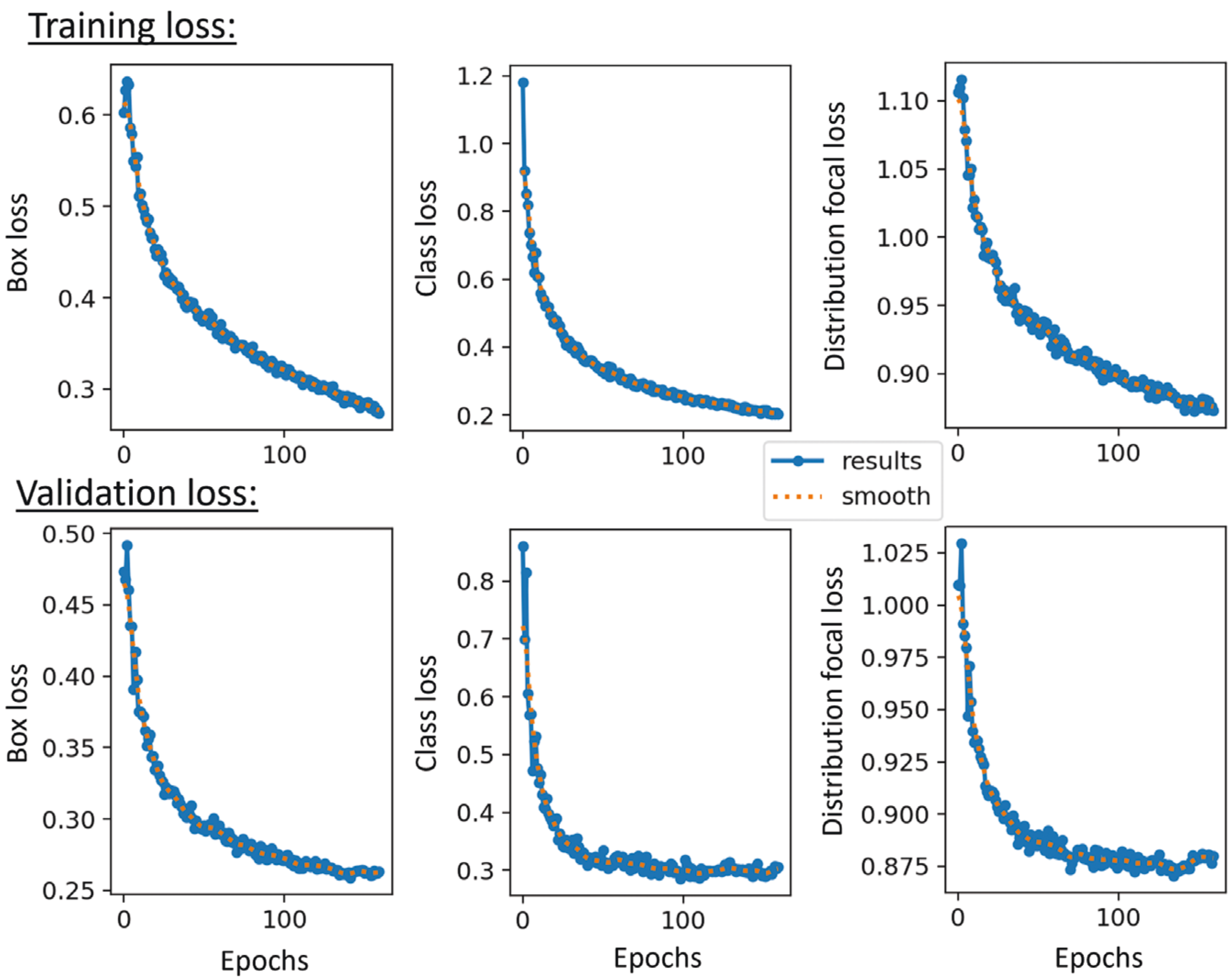

Figure 3. Evolution of each component of the loss function during training (upper plots) and validation (lower plots). Blue thicker points represent results for each epoch and the orange dotted line represents the smoothed trend.

Model Assessment

The model performance was assessed using a third independent dataset created and called the test set (Figure 2), allowing to reporting of the model performance in an unbiased way (Eelbode et al., 2021). The test set consisted of 250 frames with 25 frames from each of the nine lighting conditions (Table 1) and from darkness, selected on random videos with pigs at different ages (as explained above for the training and validation sets). These frames were independent of the training and validation to assess the model performance on completely unknown frames.

Frames of this test set were first manually annotated following the annotation method, as described above for the training and validation set, and the developed model was then run on the test frames (Figure 2). Data from the manual annotations are referred to as "ground truth" and were obtained from CVAT, while data from the model are further mentioned as "prediction" and consist of the raw output of the detection model. Data processing was performed in R Studio (R Core Team (2020), v.4.1.0; R Studio Team (2021), v.1.4.1717).

The metrics used to evaluate the ability of the model to assign the correct posture label were: agreement between the ground truth and the predicted boxes (also known as accuracy), precision, sensitivity (also known as recall), and F1 score.

Agreement was obtained by attributing a score of 1 when both the ground truth and the prediction had the same posture label, and 0 when labels did not match. The precision of the detection reflects how many of the detected postures were correct and is calculated with this formula: $True\ Positive \div (True\ Positive + False\ Positive)$ (Padilla et al., 2021). The precision of the detection model can be averaged across all postures giving the mean average precision (mAP) using an Intersection over Union (IoU, explained below) threshold of 0.5. Sensitivity informs on the ability to detect all ground truth bounding boxes of a certain posture and is calculated as: $True\ Positive \div (True\ Positive + False\ Negative)$ (Padilla et al., 2021). Lastly, the F1 score reflects on how good the model can detect a posture in an imbalanced dataset, e.g., when the prevalence of each posture is not equal, as was the case in our dataset. The F1 score is calculated as follows: $2 \times [(Precision \times Sensitivity) \div (Precision + Sensitivity)]$ (Padilla et al., 2021).

The ability of the model to localize a detected pig was assessed with the Intersection over Union (IoU). The IoU

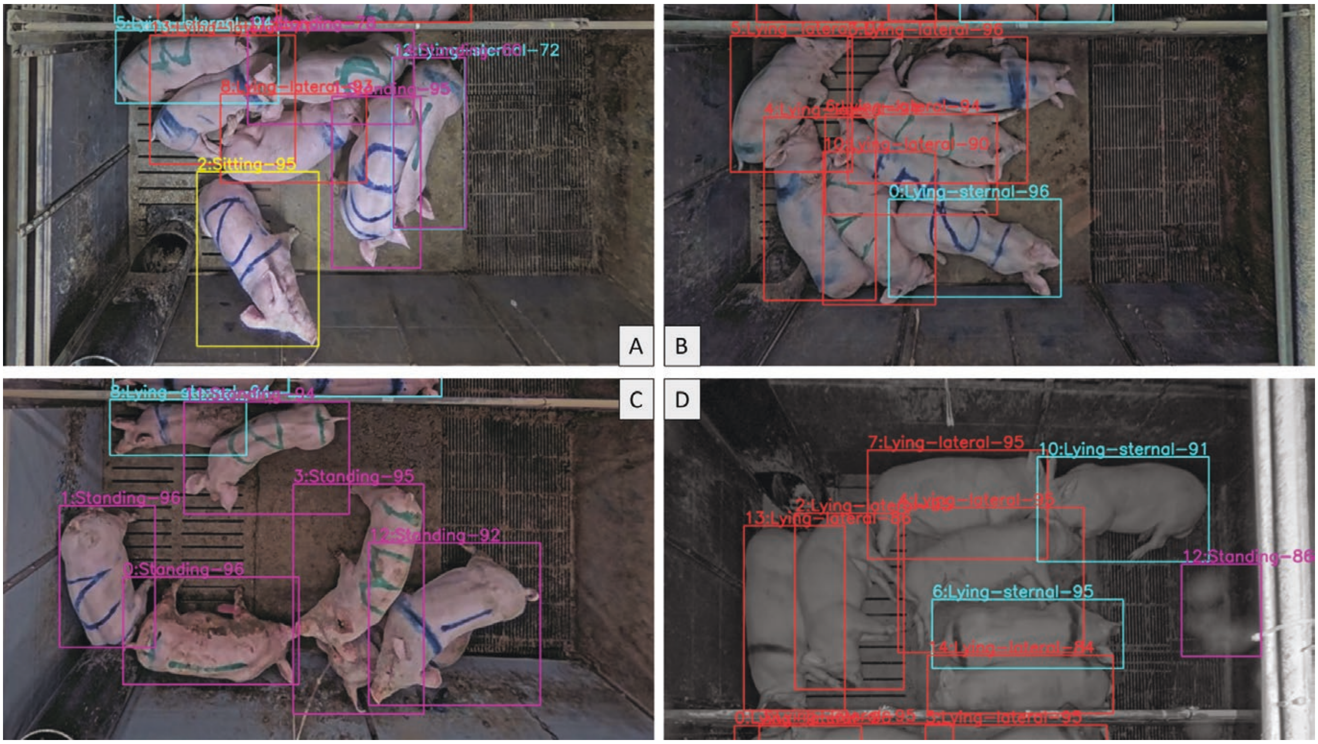

Figure 4. Examples of frames with detection errors. Bounding box colors represent postures as explained in Figure 1. A (top left): double bounding box on the pig on the right, B (top right): pig not detected lying in close proximity with pen mates at the top, C (bottom left): lying pig not detected due to occlusion by pen mates at the bottom right, D (bottom right): fly on the camera detected as a pig in night vision on the right.

represents the area of overlap between the ground truth bounding box and the predicted bounding box and is obtained by dividing the area of intersection between these two boxes by the area of the union of both boxes. The IoU ranges between 0, indicating no overlap, and 1, indicating a full overlap. In addition, more selective versions of the average precision (AP) can be calculated by adding a constraint of IoU above which the bounding box is considered correct (i.e., a true positive; Padilla et al., 2021). The IoU thresholds used are 0.5 and 0.75 (AP@0.5 and AP@0.75, respectively), meaning that, e.g., for AP@0.5, only bounding boxes with the right posture that are overlapping by more than 50% of their union are considered correct and this is the threshold used in this study for the AP calculation (see above).

Results and Discussion

Detection And Localization Of Pigs

For 449 of the 500 pens in the test set, all pigs in the pen were detected i.e., seven pigs or six pigs in case of mortality. When the number of detected pigs did not match the actual number of pigs in the pen, the model either did not detect pigs or bounding boxes were created without being associated with an existing pig. In total, 35 pigs labeled as ground truth were missed by the detection model, among which 19 pigs were sternal lying, 11 were lateral lying and 5 were sitting. A higher proportion of missed-lying pigs was found when pigs were lying in close proximity partially occluding their pen mates (Figure 4B), for example when performing huddling behavior (later discussed in relation to pig's age) or in crowded areas (Huang et al., 2023). Occlusion of lying pigs by standing pigs also occurred especially around the cotton rope attached to the wall (Figure 4C) or around the feeder, as also found by Li et al. (2022). Also, 17 non-existing pigs were detected by the model, in which boxes were labeled as lateral lying (ten boxes), standing (five boxes), sternal lying (one box), and sitting (one box). In most cases, additional bounding boxes were created on an already detected pig when the posture was not clearly visible, resulting in two overlapping boxes with different posture labels around the same pig (Figure 4A). Another reason for boxes around non-existing pigs was the presence of dirt or flies on the camera lens that were sometimes detected as pigs on grayscale images due to night vision (Figure 4D). A similar challenge with soiled camera lenses and flies was experienced by Zhang et al. (2019), Riekert et al. (2021), and Kühnemund et al. (2023). Strategies to tackle these issues are discussed later.

The model was trained on diverse light conditions to ensure a robust model to detect pigs. The detection was considered high for all light conditions as the percentage of correctly detected pigs in each light treatment ranged from 97.4% to 99.7% (Table 3). The lowest percentages of correctly detected pigs were found in darkness (during nighttime, 97.4%), in the spatial gradient (97.4%), in the reference TL light (97.7%), and in the warm white LED light (97.7%). Detection errors reported in the warm white LED treatment were explained by occlusion by pen mates of two undetected pigs, by the attribution of two overlapping bounding boxes with different posture labels on three pigs and three pigs undetected for unknown reasons. The two treatments that had the lowest light uniformity and showed more detection errors were TL light and the spatial gradient. While most errors in the TL treatment occurred due to two overlapping boxes on the same animal for five pigs, this reason was only reported for two pigs for the spatial gradient. Most of the detection errors in the

Table 3. Percentage of detected pigs and agreement between the ground truth and the predicted postures per light condition

	Percentage of pigs detected	Posture agreement
Low intensity	99.4	89.2
Medium intensity	98.9	90.9
High intensity	98.6	90.1
Spatial gradient	97.4	83.5
Reference TL	97.7	85.6
Warm white LED	97.7	84.1
Forest white LED	99.7	88.1
Cool white LED	99.7	89.1
Dynamic daylight LED	98.6	90.0
Darkness	97.4	90.3

spatial gradient were linked to five pigs huddling or lying in contact with pen mates. Therefore, if heterogeneous light intensity seems to induce more errors, there is no clear pattern in the type of detection error created. Lastly, in darkness, most of the detection errors were caused by six pigs huddling or lying next to pen mates. Such errors are logical as pigs mostly show a diurnal pattern in activity (Ingram and Dauncey, 1985) and pigs are more likely to lie at nighttime, when darkness was provided in the context of this study.

The ability of the model to detect and accurately locate a pig (regardless of the posture attributed) is indicated by the IoU between the ground truth and the predicted bounding boxes. On average, the predicted bounding boxes were overlapping by 94.6% with their respective ground truth box (mean IoU = 0.946), and 75% of the 3,512 bounding boxes had an IoU higher than 0.955. The maximum IoU was 0.998, while the lowest IoU value was 0 for 52 bounding boxes corresponding to 35 pigs annotated in the ground truth missed by the model and 17 non-existing pigs that were detected by the model. The distribution of IoU is shown in Figure 5.

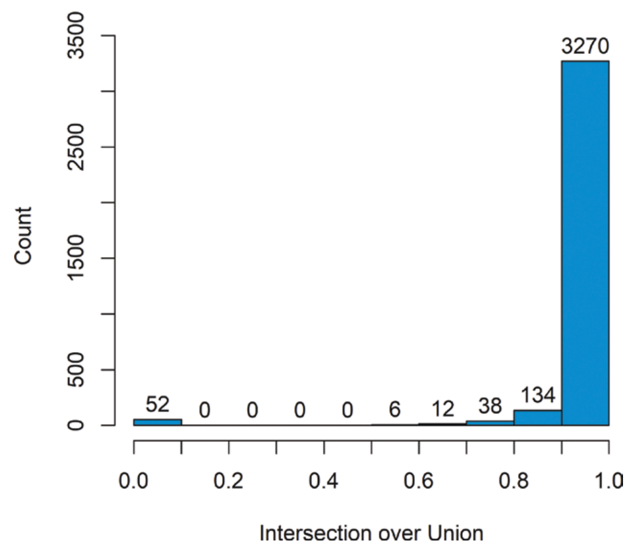

Figure 5. Distribution of Intersection over Union between the ground truth bounding box from manual annotations and the corresponding predicted bounding box by the model.

Postures Of Pigs

In the test set, the average confidence of the detection model to predict postures was 93.2%. However, the agreement between predicted and ground truth postures was 88.1%, suggesting that despite high confidence, the model misclassified some predicted postures. The model confidence ranged from the lowest at 91.2% for sitting, followed by 92.3% for sternal lying, 93.9% for lateral lying, and the highest at 95.1% for standing.

The model confidence to predict a posture was comparable in all light conditions since it ranged from 92.3% to 94.0%. The agreement in posture for each light condition was similar under low, medium, high intensities, forest white and cool white LED, and ranged from 88.1% to 90.9% (Table 3). However, the agreement in posture was slightly lower for warm white LED (84.1%), reference TL (85.6%), and the spatial gradient (83.5%). A possible explanation for the reduced posture agreement under warm white light could be due to a lower correlated color temperature resulting in less contrast between the pigs and their background and less pronounced pigs' features on the image. Furthermore, the cameras used were running on an automatic mode that could slightly adjust the image saturation, contrast, brightness, and white balance. Thus the automatic correction of these parameters may have been moderately less efficient to correct images collected under warm white light. For the reference TL and the spatial gradient, the lower light uniformity could have affected the agreement in posture due to the presence of shadows (Zhang et al., 2019). However, the observers indicated that shadows or color temperature did not influence their manual annotation of postures.

The detection model had a mAP of 89.4% for all postures combined. The AP@0.5 was also 89.4% because no predicted bounding box had an IoU lower than 53.1%. The more restrictive AP@0.75 was 88.7%, as 25 bounding boxes had an IoU of less than 0.75, and were therefore not overlapping enough to be considered as correct. When looking more specifically at how the model performed for each posture (Table 4), standing was the posture with the highest precision, sensitivity, and F1 score, meaning that standing was the best-detected posture and can be accurately predicted by the model. Concerning lying types, the precision was lower for lateral lying compared to sternal lying, but the sensitivity of lateral lying was higher than for sternal lying. This opposite result between the precision and sensitivity can be explained by 205 sternal lying pigs (ground truth) that were detected as lateral lying pigs, decreasing the sensitivity of sternal lying and the precision of lateral lying. In comparison, 57

Table 4. Summary of the performance metrics of the detection model for each posture (in percentage)

Posture	Prevalence	Precision	Sensitivity	F1 score
Lateral lying	32.1	82.9	93.9	88.0
Sternal lying	43.4	91.8	83.8	87.6
Lying (lateral and sternal)	75.5	97.5	98.0	97.8
Sitting	4.9	88.6	54.4	67.4
Standing	19.6	94.1	98.7	96.4

lateral-lying pigs (ground truth) were labeled as sternal lying, thus the sensitivity of lateral lying and the precision of lateral lying were less reduced. As also highlighted by Jeon et al. (2024), this confusion between sternal and lateral lying may have arisen from pigs resting in ambiguous postures, such as lying on their side with at least one leg tucked under the body or transitions between sternal and lateral lying. Visibility difficulties due to the camera position also occurred with occlusion of leg positions when pigs were in contact, or faced away from the camera (Figure 4A). Despite this confusion between the two forms of lying, the F1 score (taking into consideration the precision, the sensitivity, and the prevalence with an imbalanced dataset here) was rather similar for both types of lying. Furthermore, when grouping both sternal and lateral lying into one common lying category, the detection model performed substantially better as precision, sensitivity, and F1 score improved (97.5%, 98.0%, and 97.8%, respectively, see Table 4). Thus, data collected on sternal and lateral lying separately have to be interpreted with care, while results on overall lying are more robust.

Finally, sitting was the least accurately detected posture with a low sensitivity and F1 score (Table 4). The confusion matrix in Figure 6 is normalized meaning that, for each ground truth posture (columns), the percentage of predicted bounding boxes for each posture (rows) is displayed. Values in the diagonal from top left to bottom right reflect the sensitivity of the model for each posture. This figure also shows that sitting pigs (third column) were more often confused as sternal lying or as standing pigs. The main reason explaining the limited model performance to detect sitting pigs is the low prevalence of sitting pigs compared to other postures (3.7% of the training and validation sets and 4.9% of the test set, see Table 4). This imbalance problem is also called foreground-foreground class imbalance and leads to the underfitting of the model on the under-represented posture (Oksuz et al., 2020; Crasto, 2024). Imbalanced posture datasets were reported in various studies (Alameer et al., 2020; Riekert et al., 2020, 2021; Jeon et al., 2024). In some studies (Riekert et al., 2020, 2021; Ji et al., 2022; Jeon et al., 2024), this issue was tackled by using data augmentation generating new bounding boxes, e.g., by mirroring frames with low occurrence posture, to train the model on more data (Oksuz et al., 2020; Crasto, 2024). In addition, the limited performance to detect sitting pigs was caused by confusion of sitting with sternal lying and standing, occurring for 28.1% and 13.5% of sitting pigs annotated as ground truth (Figure 6). It could be hypothesized that, as sitting occurs when pigs are transitioning between standing and sternal lying, especially when heavier, the model might have difficulties to differentiate during these lying-standing transitions. Similar assumptions were reported by Jeon et al. (2024). Furthermore, observers remarked that distinguishing sitting pigs from sternal lying pigs when they were located right under the camera was sometimes difficult during manual annotations. Due to the camera's top view, the observation of whether front legs were stretched under the body or perfectly tucked under the body when sternal lying was not always easy to judge. A similar issue was experienced during manual annotations for stretched or bent hind legs when observers had to differentiate between sitting and standing. In those cases, observers paid attention to the depth of shoulders or hips of the pig of interest compared to surrounding pigs to judge how high the body part was from the floor and determine the posture. Such difficulties in posture distinction have been reported for both manual annotation and model detection in Nasirahmadi et al. (2019) and Jeon et al. (2024).

Effect Of Pig Age On Posture Detection

Regarding the age and size of pigs, a higher agreement was obtained at 13, 15, and 17 wk of age (92.1%, 90.7%, and 88.0%) and agreement was lower for the young (86.5% at 11 wk of age) and old pigs (86.4% at 19 wk of age). Younger pigs showed sometimes huddling behavior at the start of the growing-finishing phase, making it more difficult to distinguish individual pigs and their respective postures. The limited space allowance in conventional housing systems ($0.93 m^2$/pig in this study) likely contributed to the lowest agreement for old pigs, as more occlusions could occur when other pigs were standing (Figure 4C). Moreover, the reduced visible background and closer proximity between pigs made it more difficult to distinguish individual pigs and their postures (Figure 4B).

Limitations Of The Detection Method And Potential For Tracking

As explained previously, the detection model can generate errors by confusing posture labels, by missing the detection of pigs (1.0% of all bounding boxes in the test set), creating bounding boxes for non-existing pigs (0.03%), or by already detected pigs (0.45%). While pigs missed by the model cannot be recovered, the detection of non-existing pigs can be filtered afterward with data cleaning. For example, data obtained from identified dirty cameras could be manually excluded from the dataset (Riekert et al., 2021) or data obtained from blurry video frames could be automatically filtered out as explained in Kühnemund et al. (2023). If two highly overlapping bounding boxes with similar centers and different postures are obtained for one pig, the box with the lowest confidence could be filtered out. This post-processing data cleaning stage, known as "non-maximum suppression", was performed manually on the test set and was expected to

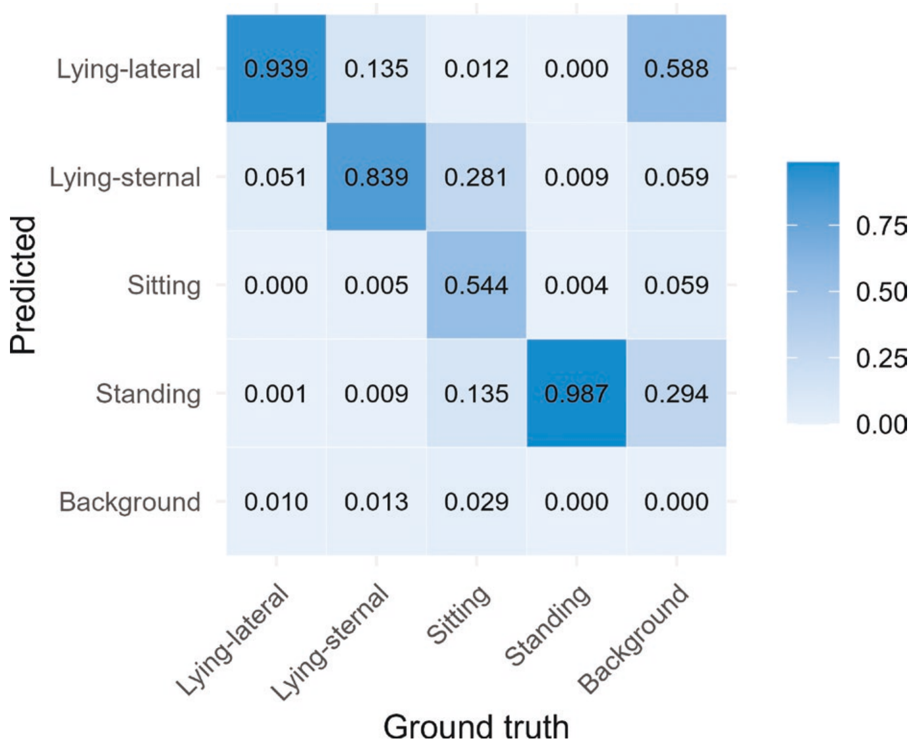

Figure 6. Normalized confusion matrix between the ground truth and predicted classes, values represent the proportion of predicted boxes over the total of ground truth boxes in each ground truth class. The label "Background" on the lowest row refers to undetected pigs, and in the right column, it refers to bounding boxes created for non-existing pigs.

improve the performance of the object detection model (Liu et al., 2023). In the test set, 16 pigs received two overlapping bounding boxes each, and removing one of the two overlapping boxes with the lowest confidence improved the mean agreement between the ground truth and the prediction and the mAP slightly (from 88.1% to 88.3% and from 89.4% to 89.7%). Perhaps these metrics did not improve more substantially as in some cases the bounding box with the lowest confidence, that was removed, was actually in agreement with the ground truth. Nevertheless, this illustrates the importance of data inspection and cleaning after collecting data from computer vision techniques when they are being applied to large datasets (Riekert et al., 2020).

An additional feature of YOLOv8x compared to prior YOLO versions is the possibility to include tracking, and this feature was enabled in our model. The tracking involves identifying a pig that retains the same bounding box number (or ID) over several consecutive video frames, allowing to study of posture changes and movement of individual pigs over time (Zhang et al., 2019; Chen et al., 2023). In literature, multi-object tracking when applied to pigs is known to be prone to errors, namely in ID switches (Zhang et al., 2019; Van der Zande et al., 2021; Chen et al., 2023). ID switches occur when the same pig is not recognized in consecutive frames, resulting in a new pig ID being created. Reasons for ID switches are numerous, including occlusion of pigs, close proximity with pen mates, and similar appearance between pigs (Zhang et al., 2019; Van der Zande et al., 2021; Chen et al., 2023). The performance of the multi-object tracking feature of our model could not be assessed because our test set did not contain consecutive frames. Independent non-consecutive frames were chosen in the test set to represent a large diversity of lighting conditions, pig postures, location, and age and accurately assess our model performance in locating and detecting postures. However, to illustrate this point, we ran the model on an additional 6-h video (from midnight until 06:00 in the morning) at 1 FPS. In total, 1,222 IDs were generated for 14 pigs over the 6 h and a pig retained an ID for an average of 4 min. This suggests that, as reported in the literature, there are still challenges to collect temporal information on individual pigs when using computer vision techniques (Zhang et al., 2019; Chen et al., 2023). Post-processing data inspection and cleaning could enhance these results. For example, investigating the creation of new IDs and correcting them to previous IDs based on, for example, the timing and/or the location where the ID change occurs might improve tracking of individual pigs.

Comparison Of The Detection Model Performance

When comparing the overall performances of the model with other studies, the detection model developed by Jeon et al. (2024) was the most comparable, as it was also based on the YOLO framework (v7) and detected the same set of postures. The F1 score of the model of Jeon et al. (2024) was higher than in our study (91% vs 85%), which was likely due to the use of data augmentation (i.e., image flip, rotation, and translation) that had improved the sitting sensitivity, the sternal lying precision, and the standing precision, ultimately resulting in a higher F1 score. A similar conclusion can be drawn when comparing our results with Alameer et al. (2020) and Ji et al. (2022), who respectively found an mAP of 97.2% and 95.7% (vs 89.4% in our study) using data augmentation and a substantially larger training set (image flip, scaling and HSV augmentation with 735,094 annotated pigs in Alameer et al. (2020), HSV augmentation, affine transformation, image flip and mosaic augmentation with 20,105 annotated pigs in Ji et al. (2022), while we used no augmentation and 17,403 annotated pigs in our study). Furthermore, different computer vision techniques were used to detect similar postures, for example, Nasirahmadi et al. (2019) used RFCN and found a mAP of 93%. Additionally, less detailed ethograms to detect whether pigs were lying or not were used with R-CNN in Riekert et al. (2020) and Riekert et al. (2021). In these two studies, the mAP reached 80.2% during the daytime (Riekert et al., 2020), and 84% during the day and 58% at night (Riekert et al., 2021). When combining both lateral and sternal lying classes as "lying," and sitting and standing classes as "not lying," our model reached a mAP of 97.0%, a sensitivity of 95.5%, and a F1 score of 96.2%. This shows that grouping these postures together results in a robust model detecting "lying" and "not lying" postures.

To conclude, locations and postures of group-housed growing-finishing pigs of different ages can be identified by an object detection model under different lighting conditions. The performance of the model to detect pigs was similar between lighting conditions, but the agreement in posture between the model and the human observer was slightly lower under warm and less uniform lighting conditions, and agreement was lower in young and old growing-finishing pigs. The developed model reached a mAP of 89.4% to classify lateral lying, sternal lying, sitting, and standing in pigs. This precision can be improved by grouping (lateral and sternal) lying postures, as well as non-lying (standing and sitting) postures, leading to better precision (mAP = 97.0%). The location of the predicted bounding boxes was also precise as 75% of the predicted boxes had an IoU of at least 95.5% with the ground truth. Despite the accurate location, errors occurred with undetected pigs, detection of non-existing pigs, or attribution of double overlapping boxes to the same pig. This stresses the importance of including diverse conditions during the model training. Future studies using posture detection models could include data augmentation of postures with a low occurrence to balance training and test sets and develop more robust detection models. Furthermore, even though multi-object tracking is a feature available in YOLO, some challenges remain regarding the change of ID for an individual pig or ID switches between pen mates.

Study Funding

This research received funding from the Netherlands Organization for Scientific Research (NWO) in the framework of the ENW PPP Fund and private parties Signify and De Hoeve Innovatie (ENWSS.2018.005).

Conflict of interest statement

The authors report no conflicts of interest.

Author Contributions

Alice Scaillierez (Conceptualization, Data curation, Investigation, Methodology, Resources, Visualization, Writing—original draft), Tomás Izquierdo García-Faria (Investigation, Methodology, Resources, Software, Writing—review & editing), Harry Broers (Investigation, Methodology, Resources, Software, Writing—review & editing), Sofie E. van Nieuwamerongen—de Koning (Conceptualization, Investigation, Resources, Writing—review & editing), Rik van der Tol (Conceptualization, Funding acquisition, Methodology, Supervision, Writing—review & editing), Eddie Bokkers (Conceptualization, Funding acquisition, Project administration, Supervision, Writing—review & editing), and Iris Boumans (Conceptualization, Data curation, Funding acquisition, Methodology, Supervision, Writing—review & editing)

Literature Cited

Alameer, A., I. Kyriazakis, and J. Bacardit. 2020. Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs. Sci. Rep. 10:13665. doi: 10.1038/s41598-020-70688-6

Chen, Z., J. Lu, and H. Wang. 2023. A review of posture detection methods for pigs using deep learning. Appl. Sci. 13:6997. doi: 10.3390/app13126997

Christison, G. I. 1996. Dim light does not reduce fighting or wounding of newly mixed pigs at weaning. Can. J. Anim. Sci. 76:141–143. doi: 10.4141/cjas96-019

Cornou, C., S. Lundbye-Christensen, and A. R. Kristensen. 2011. Modelling and monitoring sows' activity types in farrowing house using acceleration data. Comput. Electron. Agric. 76:316–324. doi: 10.1016/j.compag.2011.02.010

Crasto, N. 2024. Class imbalance in object detection: an experimental diagnosis and study of mitigation strategies. arXiv preprint arXiv 2403.07113. doi: 10.48550/arXiv.2403.07113

Eelbode, T., P. Sinonquel, F. Maes, and R. Bisschops. 2021. Pitfalls in training and validation of deep learning systems. Best Pract Res Clin Gastroenterol 52-53:101712. doi: 10.1016/j.bpg.2020.101712

European Union. 2010. *Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes 170*. p 61. Official Journal of the European Union.

Fernández-Carrión, E., M. Martínez-Avilés, B. Ivorra, B. Martínez-López, M. Ramos, and J. M. Sánchez-Vizcaíno. 2017. Motion-based video monitoring for early detection of livestock diseases: the case of African swine fever. PLoS One 12:e0183793. doi: 10.1371/journal.pone.0183793

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

Huang, Y., D. Xiao, J. Liu, Z. Tan, K. Liu, and M. Chen. 2023. An improved pig counting algorithm based on YOLOv5 and DeepSORT model. Sensors. 23:6309. doi: 10.3390/s23146309

Ingram, D. L., and M. J. Dauncey. 1985. Circadian rhythms in the pig. Comp. Biochem. Physiol. A Comp. Physiol. 82:1–5. doi: 10.1016/0300-9629(85)90695-4

Jeon, C., H. Kim, and D. Kim. 2024. Classifying Pig Poses for Smart Pigsties Using Deep Learning. ICIC-ELB 15:187–193. doi:10.24507/icicelb.15.02.187

Ji, H., J. Yu, F. Lao, Y. Zhuang, Y. Wen, and G. Teng. 2022. Automatic position detection and posture recognition of grouped pigs based on deep learning. Agriculture. 12:1314. doi: 10.3390/agriculture12091314

Kim, Y. J., M. H. Song, S. I. Lee, J. H. Lee, H. J. Oh, J. W. An, S. Y. Chang, Y. B. Go, B. J. Park, M. S. Jo, et al. 2021. Evaluation of pig behavior changes related to temperature, relative humidity, volatile organic compounds, and illuminance. J. Anim. Sci. Technol. 63:790–798. doi: 10.5187/jast.2021.e89

Kühnemund, A., S. Götz, and G. Recke. 2023. Automatic detection of group recumbency in pigs via ai-supported camera systems. Animals. 13:2205. doi: 10.3390/ani13132205

Li, J., A. R. Green-Miller, X. Hu, A. Lucic, M. R. M. Mohan, R. N. Dilger, I. C. F. S. Condotta, B. Aldridge, J. M. Hart, and N. Ahuja. 2022. Barriers to computer vision applications in pig production facilities. Comput. Electron. Agric. 200:107227. doi: 10.1016/j.compag.2022.107227

Liu, D., A. Parmiggiani, E. Psota, R. Fitzgerald, and T. Norton. 2023. Where's your head at? Detecting the orientation and position of pigs with rotated bounding boxes. Comput. Electron. Agric. 212:108099. doi: 10.1016/j.compag.2023.108099

Liu, K., L. Liu, M. Tai, Q. Ding, W. Yao, and M. Shen. 2022. Light from heat lamps affects sow behaviour and piglet salivary melatonin levels. Animal. 16:100534. doi: 10.1016/j.animal.2022.100534

Nasirahmadi, A., S. A. Edwards, S. M. Matheson, and B. Sturm. 2017. Using automated image analysis in pig behavioural research: assessment of the influence of enrichment substrate provision on lying behaviour. Appl. Anim. Behav. Sci. 196:30–35. doi: 10.1016/j.applanim.2017.06.015

Nasirahmadi, A., U. Richter, O. Hensel, S. Edwards, and B. Sturm. 2015. Using machine vision for investigation of changes in pig group lying patterns. Comput. Electron. Agric. 119:184–190. doi: 10.1016/j.compag.2015.10.023

Nasirahmadi, A., B. Sturm, A. -C. Olsson, K. -H. Jeppsson, S. Müller, S. Edwards, and O. Hensel. 2019. Automatic scoring of lateral and sternal lying posture in grouped pigs using image processing and Support Vector Machine. Comput. Electron. Agric. 156:475–481. doi: 10.1016/j.compag.2018.12.009

Oczak, M., K. Maschat, D. Berckmans, E. Vranken, and J. Baumgartner. 2015. Classification of nest-building behaviour in non-crated farrowing sows on the basis of accelerometer data. Biosyst. Eng. 140:48–58. doi: 10.1016/j.biosystemseng.2015.09.007

Oksuz, K., B. C. Cam, S. Kalkan, and E. Akbas. 2020. Imbalance problems in object detection: a review. IEEE Trans. Pattern Anal. Mach. Intell. 43:3388–3415. doi: 10.1109/tpami.2020.2981890

Padilla, R., W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva. 2021. A comparative analysis of object detection metrics with a companion open-source toolkit. Electronics. 10:279. doi: 10.3390/electronics10030279

R Core Team. 2020. R A language and environment for statistical computing. R Foundation for Statistical. Computing.

Riekert, M., A. Klein, F. Adrion, C. Hoffmann, and E. Gallmann. 2020. Automatically detecting pig position and posture by 2D camera imaging and deep learning. Comput. Electron. Agric. 174:105391. doi: 10.1016/j.compag.2020.105391

Riekert, M., S. Opderbeck, A. Wild, and E. Gallmann. 2021. Model selection for 24/7 pig position and posture detection by 2D camera imaging and deep learning. Comput. Electron. Agric. 187:106213. doi: 10.1016/j.compag.2021.106213

Rijksoverheid. 2021. Artikel 1 Wet op de dierproeven– [accessed September 1, 2024]. https://www.rijksoverheid.nl/onderwerpen/dierproeven/regels-dierproeven

R Studio Team. 2021. RStudio: Integrated Development Environment for R. accessed 16 September 2023. http://www.rstudio.com/.

Scaillierez, A. J., S. E. van Nieuwamerongen - de Koning, I. J. M. M. Boumans, P. P. J. van der Tol, S. K. Schnabel, and E. A. M. Bokkers. 2024. Effect of light intensity on behaviour, health and growth of growing-finishing pigs. Animal. 18:101092. doi: 10.1016/j.animal.2024.101092

Simonsen, H. B. 1990. Behaviour and distribution of fattening pigs in the multi-activity pen. Appl. Anim. Behav. Sci. 27:311–324. doi:10.1016/0168-1591(90)90127-y

Sohan, M., T. Sai Ram, R. Reddy, and C. Venkata. 2024. A review on yolov8 and its advancements. In: I. J. Jacob, S. Piramuthu, P. Falkowski-Gilski, editors. Data intelligence and cognitive informatics. ICDICI 2023. Algorithms for Intelligent Systems. Springer, Singapore; p 529–545. doi:10.1007/978-981-99-7962-2_39.

Tian, M., H. Guo, H. Chen, Q. Wang, C. Long, and Y. Ma. 2019. Automated pig counting using deep learning. Comput. Electron. Agric. 163:104840. doi: 10.1016/j.compag.2019.05.049

Tzanidakis, C., P. Simitzis, K. Arvanitis, and P. Panagakis. 2021. An overview of the current trends in precision pig farming technologies. Livest. Sci. 249:104530. doi: 10.1016/j.livsci.2021.104530

Ultralytics. 2023. Ultralytics YOLO Docs. YOLO Vision– [accessed September 1, 2024]. https://docs.ultralytics.com/

Van der Zande, L. E., O. Guzhva, and T. B. Rodenburg. 2021. Individual detection and tracking of group housed pigs in their home pen using computer vision. Front. Anim. Sci. 2:669312. doi: 10.3389/fanim.2021.669312

Zhang, L., H. Gray, X. Ye, L. Collins, and N. Allinson. 2019. Automatic individual pig detection and tracking in pig farms. Sensors (Basel, Switzerland). 19:1188. doi: 10.3390/s19051188