Deep-learning-based quantification of sow activity during farrowing

Johann Wahmhoff[1], Martin Wutke[1] and Imke Traulsen[1]

Abstract: The activity of sows after farrowing significantly influences the occurrence of crushing losses. For this reason, sow activity is often quantified in scientific research through video analysis. To improve practicality, this study developed an automated approach based on the YOLOv10 object detection framework, addressing the big data challenge. The core idea of this approach was to calculate an activity score based on the number of postural changes. Therefore, the model detects the sow's posture for each frame in the videos. The raw data from the model is then filtered, and the number of postural changes is calculated, from which an activity score is derived for each video. For comparison with a human-generated activity assessment, the videos were finally categorized into three classes based on the activity score. The final performance evaluation across activity classes shows an accuracy of 0.945. Based on these results, it can be concluded that this approach provides an automated alternative to manual video analysis for determining sow activity.

Keywords: deep learning, YOLOv10, video analysis, sow activity, farrowing, crushing losses

1 Introduction

The seventh amendment of the Animal Husbandry Regulation in Germany, particularly the reduction in the duration of sow fixation around farrowing, has made the issue of piglet crushing even more important in practice and science. In this context, the activity of the sow after birth significantly impacts the number of crushing events, making it a key factor for further investigation.

One common approach for assessing sow activity is the use of video recordings. However, the large amount of data generated often poses challenges, as manual evaluation is impractical and cost-intensive. As a result, automated methods are increasingly being employed, such as activity analysis using optical flow techniques [Gr16]. While optical flow has proven useful for detecting changes in the scenery, it is not without limitations. Issues such as illumination changes and moving shadows can significantly affect accuracy [Ga21; LHG21].

To overcome these challenges, this study employs an alternative approach using a deep-learning-based object detection model to quantify farrowing activity. Specifically, a

[1] Christian-Albrechts-University Kiel, Institute of Animal Breeding and Husbandry, Hermann-Rodewald-Straße 6, 24116 Kiel, jwahmhoff@tierzucht.uni-kiel.de, mwutke@tierzucht.uni-kiel.de, itraulsen@tierzucht.uni-kiel.de

YOLOv10 architecture [Wa24] was used to estimate the sow's body posture within the farrowing pen. Based on the model results, postural changes can be identified, which form the basis of the activity assessment. We further calculate an activity score, which can provide valuable insights into sow behavior. The presented approach provides an alternative way to analyze large video datasets and to improve our understanding of sow activity patterns and their relationship to piglet crushing, thus contributing to more effective strategies to reduce losses in modern pig farming.

2 Material and methods

2.1 Data processing

The video footage was collected as part of the Zissau project (funding code: 28N305603) between October 2022 and October 2023. Sows were recorded from an overhead perspective using RGB cameras (Reolink RLC-520A) installed 3.2 meters above the farrowing pens. The video resolution was 2560 x 1920 pixels, with frame rates varying from 10 to 30 frames per second. For model training and evaluation, 4387 individual frames, captured both pre- and post-parturition, were randomly extracted (standing: 1165, sitting: 1035, lying: 1016, lying lateral: 1171) and manually annotated using LabelMe software (Version 5.2.1). Four distinct classes were defined for body posture annotation: standing, sitting, lying, and lying lateral, with each posture identified using bounding boxes.

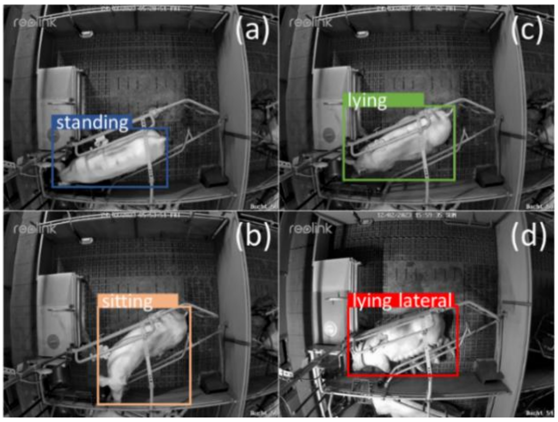

Fig. 1: Overhead view of the farrowing pen with annotated body posture classes: standing (a), sitting (b), lying (c) and lying lateral (d)

To ensure consistency within the training dataset, all images were converted to grayscale, as the dataset included night-time images, which were inherently grayscale. To enhance the information richness of the dataset, image augmentation was applied, including horizontal, vertical, and combined horizontal-vertical rotations. This procedure expanded the dataset to a total of 17548 images, improving model robustness. Figure 1 shows an example of the farrowing pen and the annotated body position classes. To support the community, an extensive dataset, including additional husbandry environments, is being built and will be published soon to allow reproducibility.

2.2 Model training and activity quantification

The object detection model used in this study is based on a YOLOv10 architecture and was trained over 250 epochs using the SGD-optimizer and a batch rate of eight images. For the training process, we used 17048 images and evaluated the model with 500 images. After training, the model achieved a Precision, Recall and mAP50-95 value of 0.991, 0.999 and 0.995, respectively.

The sow activity estimation is conducted in two steps. First, the model results below a confidence score of 0.96 are removed to reduce potential false positive predictions and to ensure a high model quality. Subsequently, a sliding window approach corresponding to a five-second period is applied, where the mode of the postures within the window size is considered the correct posture. In the final filtering stage, frames with postures held for only one second after the sliding window are excluded. Figure 2 illustrates a comparison between the unfiltered and filtered results for a single video.

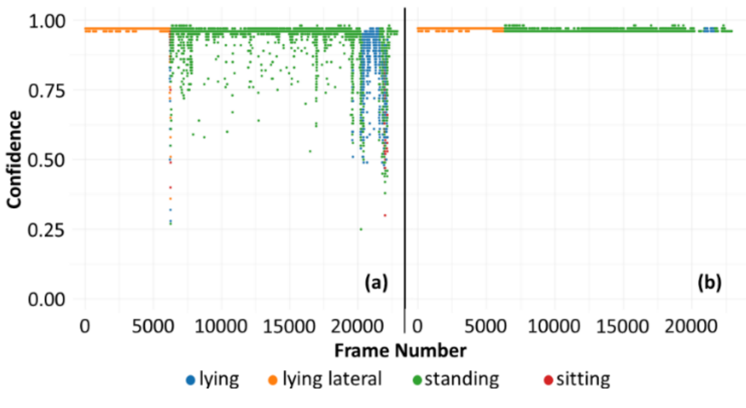

Fig. 2: Unfiltered (a) and filtered (b) model output for a single video

In the second step, the final activity score for each video is calculated in the form of posture changes detected by the model. As the videos to be analyzed have different runtimes, ranging from one minute to 60 minutes, the video duration had to be considered to adjust the number of posture changes. Here, a min-max transformation of the runtimes in minutes is used to normalize the temporal information. As a result, the activity score is defined as:

$$Activity\ score = \frac{number\ of\ changes}{1 + \left(\frac{runtime\ - runtime_{min}}{runtime_{max} - runtime_{min}}\right)}$$

2.3 Evaluation of the activity quantification

Since even small phases of false detections can have an effect on the estimated activity score, we further evaluated our approach to assess the suitability of the proposed approach for determining sow activity through tracking posture changes. Therefore, an additional three-stage evaluation process was conducted. The first evaluation stage focused on assessing the model's performance in estimating the sow's body posture for each frame. For this purpose, 200 randomly selected videos were analyzed by the model and manually evaluated to calculate the accuracy for the four posture classes.

In the second evaluation stage, the focus was on determining whether the proposed methodology could accurately track the number of postural changes, as this is essential for calculating a valid activity score. For this purpose, the number of real posture changes obtained from the manual evaluation and the number of estimated posture changes based on the processed model output were compared using the well-known evaluation metrics Recall, Precision and F1-score [Ho21].

The final evaluation stage aims to assess the activity value described above. For this purpose, the activity value generated by the model was compressed and compared with a human activity classification. To enable this comparison, the activity values of the 200 randomly selected videos were normalized using a min-max transformation and divided into three classes. If the normalized activity score was zero, the video was classified as low activity. If the score was between 0 and 0.5, the video was classified as medium activity. If the score was greater than 0.5, the video was classified as high activity.

3 Results and discussion

In this chapter, the results of the evaluation will be presented. Table 1 shows the results of the first evaluation stage. These findings demonstrate that the model is capable of reliably detecting the individual postures. Frames where the posture was not accurately identified often occurred during transitions when the sow moves from one posture to another. However, these frames are filtered out during the activity evaluation process, preventing any distortion of the activity score. Notably, the high accuracy for the "lying lateral" posture is positive, as sows spend most of their time in this position following parturition [Yu19]. The relatively low accuracy for the "sitting" class could potentially be improved by increasing the number of training images, especially including challenging images.

Posture class	Frequency	Accuracy
Standing	139 216	0.936
Sitting	100 985	0.831
Lying	626 640	0.874
Lying lateral	1 934 396	0.988
All	2 824 149	0.955

Tab. 1: Evaluation results of the first evaluation stage for posture detection

To additionally assess the approaches' ability to quantify posture changes, we compared the detected posture changes with the manual assessment. The findings from the second evaluation stage are summarized in Table 2. The results indicate a good, though not perfect, detection of the exact number of posture changes. This highlights the challenge of accurately identifying the number of posture changes, as even a few false detections can lead to inaccuracies in the results, despite the model's strong performance in recognizing individual postures.

Dataset	# Videos	Real changes	Detected changes	Delta changes	F1	Recall	Precision
Without change	116	0	7	7	0	-	0
With change	84	279	284	83	0.852	0.860	0.845
All	200	279	284	90	0.842	0.860	0.825

Tab. 2: Evaluation results of the second evaluation stage for tracking the number of postural changes

		Predicted Group				
	Activity	low	medium	high	All	Accuracy
True Group	low	114	2	0	116	0.983
	medium	6	68	1	74	0.906
	high	0	2	7	10	0.777
	All	120	72	8	200	0.945

Tab. 3: Evaluation results of the activity groups

To finally assess our performance in quantifying the sows' activity level, we compared the automatically generated results with the manual ratings in a compressed way, where the respective activity score was classified into one of three activity categories. The results are shown in Table 3. As can be seen in the results, the activity quantification shows a strong performance. Out of 200 videos, only 11 were misclassified, with a maximum deviation of one activity group. Notably, the classification was particularly accurate for low activity levels, which is beneficial since sows tend to have low activity and spend much of their time lying down, especially after farrowing [Yu19]. The comparatively lower performance in high-activity cases may be attributed to the smaller sample size for this category.

4 Conclusion

Video analysis is a crucial tool in the study of animal behavior. Automating the evaluation process can save significant labor resources in livestock research. In this study, a deep-learning-based approach for determining sow activity in the farrowing area was successfully applied. Additionally, this automated analysis enables the generation of movement profiles and behavioral patterns related to sow postural changes. In the ongoing Zissau project, the model will be further utilized to investigate the effect of organic nesting material on the occurrence of piglet crushing losses.

Acknowledgments

This work is funded by the German Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany, granted by the Federal Office for Agriculture and Food (BLE; grant number 28N305603).

Bibliography

[Ga21] Gan, H. et al.: Fast and accurate detection of lactating sow nursing behavior with CNN-based optical flow and features. Comp. and Electr. in Agriculture 189, S. 106384, 2021.

[Gr16] Gronskyte, R. et al.: Monitoring pig movement at the slaughterhouse using optical flow and modified angular histograms. Biosystems Engineering 141, S. 19-30, 2016.

[Ho21] Ho, K. et al.: Automatic monitoring of lactation frequency of sows and movement quantification of newborn piglets in farrowing houses using convolutional neural networks. Computers and Electronics in Agriculture 189, S. 106376, 2021.

[LHG21] Liao, B.; Hu, J.; Gilmore, R. O.: Optical flow estimation combining with illumination adjustment and edge refinement in livestock UAV videos. Computers and Electronics in Agriculture 180, S. 105910, 2021.

[Wa24] Wang, A. et al.: YOLOv10: Real-Time End-to-End Object Detection, In arXiv:2405.14458 [cs.CV], 2024.

[Yu19] Yun, J. et al.: Factors affecting piglet mortality during the first 24 h after the onset of parturition in large litters: effects of farrowing housing on behaviour of postpartum sows. Animal 13, S.1045-1053, 2019.