Using Deep Learning for automated birth detection during farrowing

Jan-Hendrik Witte[1], Johann Gerberding[2], Clara Lensches[3], Imke Traulsen[4]

Abstract: Pig livestock farming has been undergoing major structural change for years. The number of animals per farm is constantly increasing, while competition is becoming more intense due to volatile slaughter prices. Sustainable, welfare-oriented livestock farming becomes increasingly difficult under these conditions. Studies have shown that animal-specific birth monitoring of sows can significantly reduce piglet losses. However, continuous monitoring by human staff is inconceivable, which is why systems need to be created that assist farmers in these tasks. For this reason, this paper aims to introduce the first step towards an automated birth monitoring system. The goal is to use deep learning methods from the field of computer vision to enable the detection of individual piglet births based on image data. This information can be used to develop systems that detect the beginning of a birth process, measure the duration of piglet births, and determine the time intervals between piglet births.

Keywords: precision livestock farming, birth monitoring, deep learning, computer vision

Addresses Sustainable Development Goal 9: Industry, innovation and infrastructure

1. Introduction

The structures of modern pig livestock farming, and piglet production have changed significantly in recent years. The situation report of the German Farmers' Association shows the opposite trend of a steadily decreasing number of farms with a simultaneous increasing number of sows held per farm [De20]. A total of 70% of all sows housed in Germany are kept on the largest 2,000 individual farms, each with 250 breeding sows per farm. Meanwhile, the slaughter price has been highly volatile in recent years, which further intensifies competition and poses major challenges for the farmer now and in the future. At the same time, politics and society alike are calling for more sustainable and more animal-friendly husbandry [Be14], which creates additional pressure and makes economically profitable livestock farming increasingly difficult. These challenges cannot be met with conventional methods, which is why new and innovative solutions are needed. As a result, research in the domain of precision livestock farming (PLF) has increased in recent years. PLF describes systems that utilize modern camera and sensor technologies to enable automatic real-time monitoring in livestock production to supervise animal health, welfare and behaviour [Be14] [D'18]. This involves the automated acquisition,

[1] University of Oldenburg, jan-hendrik.witte@uni-oldenburg.de
[2] University of Oldenburg, johann.gerberding@uni-oldenburg.de
[3] University of Göttingen, clara.lensches@uni-goettingen.de
[4] University of Göttingen, imke.traulsen@uni-goettingen.de

processing, analysis and evaluation of sensor-based data like temperature, humidity, NH3 or CO2 concentration [Co18] as well as image and video data [Ch20]. These distinct types of information and data sources hold the potential to enable data-driven assistance systems that support farmers in their daily work and would help them adapt to the constantly changing conditions in sow livestock farming.

To build such systems, methods are first needed that allow the automated processing of these types of data streams in the form of image, video, and sensor data. Video and image data alone can be used for a variety of PLF related use cases, many of which can already be found in the literature. For example, methods from the field of computer vision (CV) can be utilized to detect changes in activity of sow behaviour during final gestation [Kü20], which contains valuable information for interpreting the sow's behaviour and could be used for various subsequent processes. Similarly, Lao et al. investigated the automatic behaviour detection of lactating sows based on image data [La16]. Especially the image-based detection of objects in the context of PLF poses a particular challenge with various problems such as the grouping, overlapping and occlusion of animals, their different postures, orientations and positions, as well as constantly changing environmental factors such as different lighting conditions, soiling of animals or occlusion caused by objects in the pen. Due to their ability to generalize, the use of deep learning (DL) methods from the field of CV has been proven effective in addressing these challenges.

One topic area that has not yet been considered in the literature using these techniques is the birth monitoring of sows. Various studies in the field of birth monitoring have already proven that constant and targeted observation can reduce piglet loss during the birth process [Ho95]. White et al. [Wh96] were able to reduce piglet losses from 18.2% to 10.1%, through targeted birth monitoring based on a custom protocol. However, intensive, permanent observation of the farrowing process of individual births is not feasible in practice. There is a need to create systems that allow automated monitoring of birth processes that informs the farmer as soon as individual problems like stillbirths or prolonged farrowing is detected. This paper aims to lay the foundation for the development of such systems by developing a model for automated birth detection based on video streams.

The paper is structured as follows: First, the current state of the art in birth monitoring in sows will be presented. The primary focus lies on papers that apply DL models and architectures as well as their respective performance. This is followed by the introduction of the approaches for automated frame-based birth detection considered in this paper. Subsequently, a description of the workflow for data collection, preparation and analysis as well as model selection for each respective approach will be presented. In addition, the data sets created for each approach are described, as well as the test environment in which the different models were instantiated, trained, and evaluated. To conclude, the current status of this research is described. Here, the problems and challenges are addressed, potential solutions are presented, and the future proceedings are described in more detail.

2. Related Work

There are currently no papers available which have addressed the topic of automated birth detection based on DL methods from the field of CV. So far, the literature has considered use cases that address the automatic detection of different body conditions of the sow as well as use cases that are located before and after the actual farrowing event. In terms of body condition, Cang et al. [Ca19] use a custom Faster-R-CNN with an additional regressive branch for initial sow detection and subsequent weight estimation with an average absolute error of 0.644 kg and a relative error of 0.374%, while Huang et al. [Hu19] apply convolutional neural networks (CNN) to determine body condition scores of individual sows. Behaviours such as nest building, which can be observed prior to the actual farrowing, have been addressed in the literature by using accelerometer data from sensors to classify nest building behaviour with a generalized linear model, achieving an accuracy of 85% on the applied test set [Oc15]. Kasani et al. use different DL architectures to detect and classify sow posture into laying left, laying right, sitting and standing. The authors evaluated variations of DenseNet, VGG and Inception architectures as well as MobileNet based on a custom data set, in which the DenseNet121 achieved an accuracy of 99,83% in the classification of sow posture [Ka21]. However, most papers in the literature address the automated behaviour and posture detection of lactating or nursing sows [La16], [Wa21], [Ya18], [Zh18], [Zh20a]. Zheng et al. [Zh20b]for example use Faster R-CNN for sow posture classification into standing, sitting, ventral lying and lateral lying and achieve a mean average precision (mAP) of 0.927, while Zhang et al. [Zh19] apply a combination of MobileNet and SSD network for sow behaviour detection in drinking, urinating and mounting behaviour with an accuracy of 0.965, 0.914 and 0.923 respectively and an overall mAP of 0.934.

We found one work in progress paper which introduces an embedded system to monitor farrowing, in which the actual birth detection of piglets is considered. Silapachate et al. [Si18] applied histogram equalization, background subtraction and edge detection for image pre-processing and plan to apply histogram of oriented gradients, different machine and DL models like support vector machines or CNNs to train a binary classification model that "distinguish video frames with a newly farrowed piglet and those without". The method presented in this paper differs in the following aspects:

- The piglet birth itself should be classified. Unlike Silapachote et al. [Si18], this should not be based on a newly detected piglet in the pen, but on the distinctive visual features in the area of the vulva during the birth event.
- In addition to the binary classification approach, the use case will also be addressed based on an object detection approach by using bounding boxes to recognize, localize and classify the farrowing event.

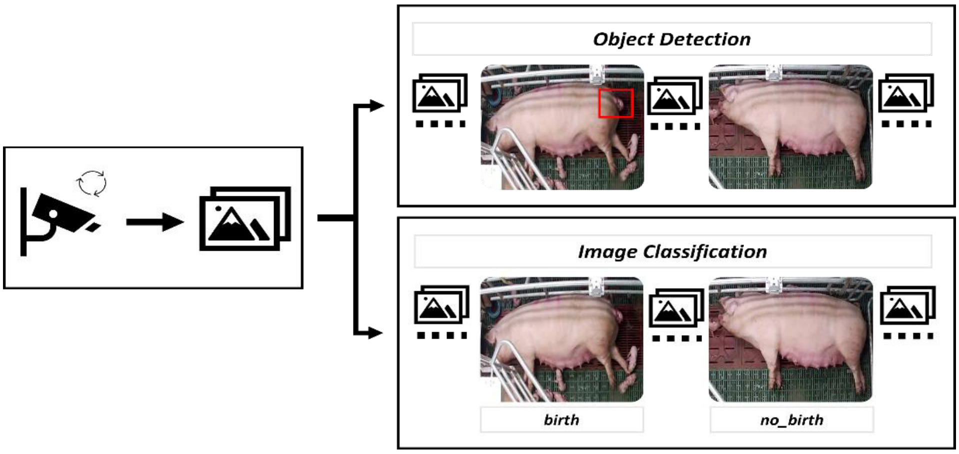

Fig. 1: Overview of approaches

3. Approaches for farrowing detection

Automated detection of a single birth event in the overall farrowing process is defined as both an object detection as well as an image classification use case. Object detection describes the detection and localization of objects of a defined class by enclosing bounding boxes around the respective object in the image, while image classification describes the task of assigning a given image to a defined label or class [Wa19]. From a system point of view, the basic idea is to split a given video stream into single frames and classify on each frame whether a birth event is taking place or not. Fig. 1 describes this process. Depending on the investigated method, the image data is processed differently. In the case of the object detection approach, the birth of a piglet is detected and localized using bounding boxes. Frames in which the object detection model predicts a bounding box with high confidence therefore in theory contains a farrowing event, frames in which no bounding box was placed correspondingly do not contain a birth event. In the case of the image classification approach, the images are processed as a whole and classified into *birth* and *no_birth* as a binary classification task. Both approaches have immediate advantages and disadvantages in terms of implementation and data preparation. While data preparation for an object detection task requires manual placement of bounding box annotations on each image, preparation of the image classification dataset only requires the categorization in one of the two defined classes, which can be done much more efficiently. However, the bounding box annotation provides a direct bounding of the context to be considered within the image, which is beneficial for the actual detection and localization of the birth event. Since the image classification model processes the image as a whole, the corresponding approach does not have this property, which could make the classification of the frames more difficult. Data preparation, dataset creation, model selection and evaluation are performed individually for each approach.

4. Materials and methods

4.1 Data Collection

To address the considered use case of frame-based automatic birth detection, an interdisciplinary data collection workflow has been defined to obtain the necessary data basis for dataset creation, model training and evaluation. Within the DigiSchwein project [Ga21], camera recordings of individual birth processes of sows are recorded and stored on a data platform. So far, experiments were conducted between May and October 2021 at the agricultural research farm for pig husbandry of the Chamber of Agriculture Lower Saxony in Wehnen. Of the planned 96 farrowing processes, 26 have already been conducted and recorded. So far, six of these farrowings were analyzed. To expand the database, four more farrowing processes have been added that were recorded within the InnoPig project at the agriculture research farm Futterkamp (Chamber of Agriculture in Schleswig-Holstein, Germany). Each farrowing pen was equipped with a commercial camera system (AXIS M3024-LVE Network Camera) which was installed in top view above the rear part of the sow. Each sow was recorded during the entire farrowing and lactation period. The videos from the DigiSchwein project were recorded at 10 fps, while the from the Futterkamp research farm have 5 fps. Both video recordings have a resolution of 1280x800. These recordings were analyzed by animal scientists to provide time stamps indicating the points in the video at which a birth event occurred. The evaluation of the video files was performed by using the open source Behavioral Observation Research Interactive Software (BORIS) [Fr16]. The starting point of the continuous observation was an hour before the beginning of each farrowing process which was defined as the birth of the first piglet of a litter. Video observation stopped at the end of the post-partum phase. The annotated time stamps in the BORIS software contain the start and end points of individual birth events, which are then used to extract the corresponding frames from the video recordings in which the respective birth event was identified. Each frame was then manually reviewed to determine if a birth event could be detected so that farrowing events could be accurately described on a frame-by-frame basis. A frame was annotated as soon as the content matched the following criteria:

- Visibility of parts of the newborn piglet in the area of the sow's vulva.
- Visibility of the expansion or extension of the sow's vulva.
- A combination of both criteria.

Fig. 2 shows an example of each of these criteria. In total, a number of 176 single birth events were extracted from the collected recordings, which were subsequently used for data analyses, exploration, and preparation.

Fig. 2: Example of birth definitions

4.2 Data Exploration

Exploratory data analysis was conducted to extract specific indicators and key metrics about the respective farrowings. The duration of individual piglet births as well as the time intervals between the birth of two successive piglets were extracted and examined as relevant indicators and are presented in Fig. 3. The results show that a substantial proportion of piglet birth durations are within the range of one to five seconds. Frame-by-frame analysis of birth events has also revealed that birth events are less than one second long. At 5 or 10 fps, this would mean that the shortest observed birth event of 0.8 seconds is 4 or 8 consecutive frames. At the same time, there are also significant outliers when considering the duration of individual births. Cases were identified in which a single birth event birth event was up to 24 seconds long. Conversely, this means that these partial birth sequences are up to 120 and 240 frames long at 5 and 10 fps, respectively. The same applies to the time intervals between individual birth events in the overall farrowing process. It can be observed that most of the intervals are between 10 and 20 minutes long, while there are also exceptions in which the intervals are up to two hundred minutes long. This information can be used to define specific thresholds at which the farmer could be informed about, for example, delayed subsequent births or similar complications. Within the DigiSchwein project, further farrowing processes will be analyzed and examined so that this database can be steadily expanded over time.

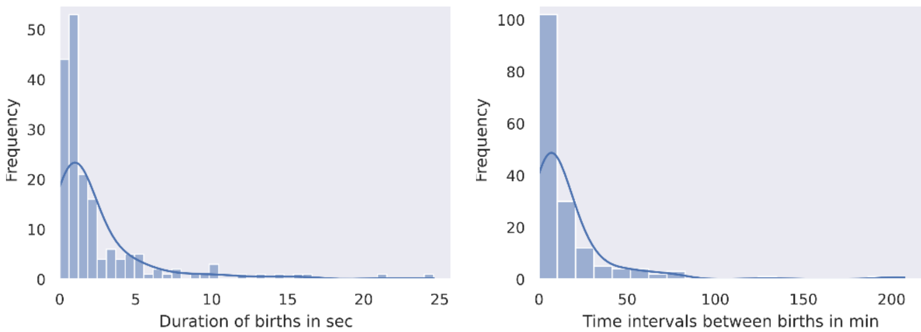

Fig. 3: Data Exploration

4.3 Model Selection

The model selection for the respective approach was conducted by defined selection criteria. These criteria are based both on models and architectures that were already used in literature as well as on the requirements for PLF systems that have been mentioned in the PLF literature. The following criteria were defined:

- **Prediction accuracy**: The prediction of the respective models should be as accurate as possible [No19].
- **Prediction speed**: Model inference should be in real-time [Le19].
- **Cost efficiency**: The respective models should be as resource efficient as possible to allow a potential deployment to low cost hardware [Ba12].

The website paperswithcode[5] provides an overview of all published real-time object detection architectures and their benchmark results on the COCO test-dev, a popular dataset on which model performance is evaluated and benchmarked. This overview served as a basis for selecting the object detection model as well as the image classification model. The following models and architectures were selected for the image classification and object detection approach:

EfficientNet: EfficientNets are among the top performers in image classification on benchmark datasets such as ImageNet, while being smaller and faster than other architectures such as ResNet or Inception [Ta19]. At the core, EfficientNets are based on a traditional CNN architecture. By applying the introduced compound scaling method to uniformly scale network depth, width and resolution as well as a neural architecture search, different EfficientNet variants were created depending on the selected compound coefficient [Ta19]. In the context of this paper, EfficientNet-B0 was used since is the smallest of the EfficientNet variants and therefore fits the specified criteria.

YOLOv5: Since the YOLO architecture has already been used in the PLF literature for various use cases [Al20a], [Sh21], [Le19], has high performance, an active developer community and also meets the defined criteria as it has a fitting balance between speed, performance and hardware requirements, it was selected as the object detection model for the initial prototyping process. YOLOv5 is the latest instalment of the YOLO architecture, but there is currently no official paper for this version. The latest paper release is YOLOv4 by Bochkovskiy et al. [Bo20], which applies specific methods and concepts summarized under the terms *bag of freebies* and *bag of specials* to improve accuracy and execution speed compared to YOLOv3 and other architectures such as EfficientDet. The comparison of the two official implementations of YOLOv4 [Al20b] and YOLOv5 [Jo21] resulted in the selection of the YOLOv5 implementation, as it was more suitable for the context of this paper.

In addition to EfficientNet for the image classification approach, following architectures

[5] https://paperswithcode.com/sota, last visited: 01.04.2022

have been evaluated as well: ResNet [He15] and SwinTransformer [Li21]. ResNet represents the baseline approach in various PLF related publications, while SwinTransformer, based on the Transformer architecture [Va17] is currently the baseline for various state of the art models in the natural language processing (NLP) as well as the CV domain.

4.4 Dataset and test environment

Overall, a total of 3.216 images were extracted from the 176 individual birth events of the examined farrowing processes, in which a birth event could be detected based on the defined visual criteria in Sec. 4.1. These were used as the foundation to create the training and test dataset for the object detection as well as the image classification approach.

Object detection dataset: In total, all 3.216 images were annotated with bounding boxes. The open source tool Labelme was used to annotate the images for model training and evaluation [Wa16].

Image Classification dataset: For the classification approach, different sampling strategies were conducted for frame selection. The best model performance was achieved by using data sets that were generated using of a hard sampling strategy based on inspired by Shrivastava et al. [Sh16]. Based on the assumption that frames immediately after and before a birth event are more difficult for the model to classify, a total of 3.216 negative examples were included to the image classification dataset in addition to the 3.216 positive examples, with one-third of the 3.216 negative examples representing frames found immediately before and after a birth event. This results in a dataset containing 6.432 images.

Both data sets were split into train and test set using an 80/20 ratio. To ensure that the test set contains only or limited data that the respective model has not yet seen, images of two birth processes were used exclusively for testing purposes and are therefore not included in the training data set. Both data sets were split into train and test set using an 80/20 ratio. To ensure that the test set only contains only data that the respective model has not yet seen, images of two complete birth sequences were used exclusively for testing purposes and are therefore not included in the training set.

Model training was performed on a desktop workstation with two Nvidia RTX 3090 with 24 GB VRAM each, a Threadripper 3960X and 64 GB RAM. For the object detection task of, the YOLOv5 implementation of Jocher et al. [Jo21] was applied. Standard parameters were used for training. The model was trained for 20 epochs with a batch size of 16 and the images were scaled to 640×640. Based on the selection criteria, the smallest checkpoint, YOLOv5s, was used for initial training and to enable transfer learning.

For the image classification task, the PyTorch Image Models [Wi19] framework was applied for model training and testing. Images were resized to 224×224 pixels and model training was set to 20 epochs with a batch size of 64 and a learning rate of 0.0001. Cross entropy loss was used as the loss function, Adam [Ki14] for the optimizer. Image augmentation was also applied by randomly rotate the image within a given degree, horizontal flipping, RGB-shifts as well as changes in brightness and contrast.

4.5 Challenges and limitations

At the current state of this research, there are several challenges and limitations that may limit the generalizability and transferability of the results of this paper, which will be addressed in this section. The data recording was conducted in several pens, but since the pens are all located at the Wehnen site in the Lower Saxony Chamber of Agriculture and are therefore very similar in structure and visual layout, both data diversity and transferability or generalizability could be limited. This cannot be resolved by adding new training data to the already annotated dataset presented in this paper, unless video recordings from other pens would be added to the dataset. Furthermore, the annotation effort to create the training dataset is very high. First, birth recordings, which are usually several hours long, must be analysed by skilled personnel and birth starts or other important events must be tagged. Then, the individual images extracted from these tagged timestamps must also be annotated manually, which can, depending on how the data should be annotated, take several seconds per image. In this case, either with bounding boxes to create an object detection dataset, or with the respective class to create an image classification dataset. Although the manual labelling effort required to annotate the images with bounding boxes could be reduced by having the previously trained algorithm pre-label the unlabelled data and then manually inspect it, however, the manual inspection of the video recordings will be difficult to substitute.

5. Current results

The results for both the object detection as well as image classification approach are summarized in Tab. 1. So, far, the results show that none of the examined approaches can produce convincing results. In the image classification approach, the best model achieves an F1-score of 67,06% on the test set, which is clearly insufficient for operational usage. The same can be observed with the object detection approach, where an AP of 0.577 and an overall mAP 0.246 of can be achieved. Compared to the precision, the low recall also shows that the model has difficulties in detecting actual positive samples in the test set. Although the EfficientNet has by far the smallest number of parameters compared to the other models, it achieved the best accuracy on the test set with a value of 66.20% in the classification task. Considering the much higher number of parameters, the SwinTransformers perform on average worse compared to the other models. The deficient performance of both the image classification approach and object detection approach can be explained as follows:

- Insufficient data basis.
- Both the image classification approach and the object detection approach are inadequate.
- A combination of both.

Image Classification							
Model	**Inference Time (s)**		**Parameters (Mio.)**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
	GPU	CPU					
ResNet50	0.004	0.035	23.51	64.93 %	71.41 %	63.22 %	67.06 %
EfficientNet-B0	0.007	0.018	4.01	66.20 %	59.72 %	68.61 %	63.86 %
SwinTransformer	0.012	0.121	86.74	64.51 %	64.22 %	64.59 %	64.41 %
Object Detection							
Model	**Inference Time (s)**		**Parameters (Mio.)**	**Precision**	**Recall**	**$AP^{IoU=0.5}$**	**mAP**
	GPU	CPU					
YOLOv5	0.002	-	7.01	0.856	0.463	0.577	0.246

Tab. 1: Overview of results

The former could be confirmed by the fact that all models evaluated in this paper showed signs of overfitting. This could be a signal for an insufficient data basis. The second could be confirmed by the fact that the considered use case is too complex to be solved with these simple approaches. The video analyses have shown that certain behavioral patterns can be recognized in the sow shortly before the expulsion of a piglet, e.g., the flapping of the tail or the stretching of the rear legs. The results so far give reason to believe that these patterns, as well as the associated temporal context, need to be considered in the detection of birth events. In the further research development, these aspects will be further investigated and elaborated. The analysis and evaluation of additional farrowing processes will show whether the problem of insufficient performance is due to the data basis or whether novel approaches must be considered in order to effectively detect birth events based on image data. Consideration of other model architectures to capture temporal context and identified behavioral patterns based on, for example, action recognition models could also be explored in this context.

Bibliography

[Al20a] Alameer, A.; Kyriazakis, I.; Bacardit, J.: Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs. Scientific reports 1/10, S. 13665, 2020.

[Al20b] Alexey et al.: AlexeyAB/darknet: YOLOv4 pre-release. Zenodo, 2020.

[Ba12] Banhazi et al.: Precision Livestock Farming: An international review of scientific and commercial aspects. International Journal of Agricultural and Biological Engineering 3/5, S. 1–9, 2012.

[Be14] Berckmans, D.: Precision livestock farming technologies for welfare management in intensive livestock systems. Revue scientifique et technique (International Office of Epizootics) 1/33, S. 189–196, 2014.

[Bo20] Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y. M.: YOLOv4: Optimal Speed and Accuracy of Object Detection, 2020.

[Ca19] Cang, Y.; He, H.; Qiao, Y.: An Intelligent Pig Weights Estimate Method Based on Deep Learning in Sow Stall Environments. IEEE Access 99/7, S. 164867–164875, 2019.

[Ch20] Chen, C. et al.: Recognition of aggressive episodes of pigs based on convolutional neural network and long short-term memory. Computers and Electronics in Agriculture 169, S. 105166, 2020.

[Co18] Cowton, J. et al.: A Combined Deep Learning GRU-Autoencoder for the Early Detection of Respiratory Disease in Pigs Using Multiple Environmental Sensors. Sensors (Basel, Switzerland) 8/18, S. 2521, 2018.

[D'18] D'Eath, R. B. et al.: Automatic early warning of tail biting in pigs: 3D cameras can detect lowered tail posture before an outbreak. PloS one 4/13, e0194524, 2018.

[De20] Deutscher Bauernverband e.V.: Situationsbericht 2020/21. Trends und Fakten zur Landwirtschaft. Kapitel 3 Agrarstruktur.

[Fr16] Friard, O.; Gamba, M.: BORIS a free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution 11/7, S. 1325–1330, 2016.

[Ga21] Gandorfer, M. et al. Hrsg.: Informatik in der Land-, Forst- und Ernährungswirtschaft. Fokus: Informations- und Kommunikationstechnologie in kritischen Zeiten Referate der 41. GIL-Jahrestagung, 08.-09. März 2021, Leibniz-Institut für Agrartechnik und Bioökonomie e.V., Potsdam. Gesellschaft für Informatik, Bonn, 2021.

[He15] He, K. et al.: Deep Residual Learning for Image Recognition, 2015.

[Ho95] Holyoake, P. K. et al.: Reducing pig mortality through supervision during the perinatal period. Journal of Animal Science 12/73, S. 3543–3551, 1995.

[Hu19] Huang, M.-H.; Lin, E.-C.; Kuo, Y.-F.: Determining the body condition scores of sows using convolutional neural networks. American Society of Agricultural and Biological Engineers, S. 1, 2019.

[Jo21] Jocher, G. e. a.: ultralytics/yolov5: v5.0 - YOLOv5-P6 1280 models, AWS, Supervise.ly and YouTube integrations. Zenodo, 2021.

[Ka21] Kasani, P. H. et al.: A computer vision-based approach for behavior recognition of gestating sows fed different fiber levels during high ambient temperature. Journal of Animal Science and Technology 2/63, S. 367–379, 2021.

[Ki14] Kingma, D. P.; Ba, J.: Adam: A Method for Stochastic Optimization, 2014.

[Kü20] Küster, S. et al.: Usage of computer vision analysis for automatic detection of activity changes in sows during final gestation. Computers and Electronics in Agriculture 169, S. 105177, 2020.

[La16] Lao, F. et al.: Automatic recognition of lactating sow behaviors through depth image processing. Computers and Electronics in Agriculture 125, S. 56–62, 2016.

[Le19] Lee, S. et al.: Practical Monitoring of Undergrown Pigs for IoT-Based Large-Scale Smart Farm. IEEE Access 7, S. 173796–173810, 2019.

[Li21] Liu, Z. et al.: Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows: ICCV 2021, 2021.

[No19] Norton, T. et al.: Review: Precision livestock farming: building 'digital representations' to bring the animals closer to the farmer. animal 12/13, S. 3009–3017, 2019.

[Oc15] Oczak, M. et al.: Classification of nest-building behaviour in non-crated farrowing sows on the basis of accelerometer data. Biosystems Engineering 140, S. 48–58, 2015.

[Sh16] Shrivastava, A.; Gupta, A.; Girshick, R.: Training Region-Based Object Detectors with Online Hard Example Mining: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), S. 761–769, 2016.

[Sh21] Shao, H.; Pu, J.; Mu, J.: Pig-Posture Recognition Based on Computer Vision: Dataset and Exploration. Animals an open access journal from MDPI 5/11, 2021.

[Si18] Silapachote, P., Srisuphab, A., Banchongthanakit, W.; A. Srisuphab; W. Banchongthanakit: An Embedded System Device to Monitor Farrowing: 2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), S. 208–213, 2018.

[Ta19] Tan, M.; Le, Q.: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In (Chaudhuri, K.; Salakhutdinov, R. Hrsg.): Proceedings of the 36th International Conference on Machine Learning. PMLR, S. 6105–6114, 2019.

[Va17] Vaswani, A. et al.: Attention Is All You Need, 2017.

[Wa16] Wada, K.: labelme: Image Polygonal Annotation with Python, 2016.

[Wa19] Wang, S.; Su, Z.: Metamorphic Testing for Object Detection Systems, 2019.

[Wa21] Wang, M. et al.: A PCA-based frame selection method for applying CNN and LSTM to classify postural behaviour in sows. Computers and Electronics in Agriculture 189, 2021.

[Wh96] White, K. R.; Anderson, D. M.; Bate, L. A.: Increasing piglet survival through an improved farrowing management protocol. Canadian Journal of Animal Science 4/76, S. 491–495, 1996.

[Wi19] Wightman, R.: PyTorch Image Models. GitHub, 2019.

[Ya18] Yang, A. et al.: Automatic recognition of sow nursing behaviour using deep learning-based segmentation and spatial and temporal features. Biosystems Engineering 175, S. 133–145, 2018.

[Zh18] Zheng, C. et al.: Automatic recognition of lactating sow postures from depth images by deep learning detector. Computers and Electronics in Agriculture 147, S. 51–63, 2018.

[Zh19] Zhang, Y. et al.: Real-time sow behavior detection based on deep learning. Computers and Electronics in Agriculture 163, S. 104884, 2019.

[Zh20a] Zhu, X. et al.: Automatic recognition of lactating sow postures by refined two-stream RGB-D faster R-CNN. Biosystems Engineering 189, S. 116–132, 2020.

[Zh20b] Zheng, C. et al.: Automatic posture change analysis of lactating sows by action localisation and tube optimisation from untrimmed depth videos. Biosystems Engineering 194, S. 227–250, 2020.