Real-time implementation of computer vision based farrowing prediction in pens with a possibility of temporary sow confinement

M. Oczak[1,2,*] and J. Baumgartner[2]
[1]*Precision Livestock Farming Hub, The University of Veterinary Medicine Vienna (Vetmeduni Vienna), Veterinärplatz 1, 1210 Vienna, Austria*
[2]*Institute of Animal Husbandry and Animal Welfare, University of Veterinary Medicine Vienna (Vetmeduni Vienna), Veterinärplatz 1, 1210 Vienna, Austria*
[*]Corresponding author: Maciej Oczak, maciej.Oczak@vetmeduni.ac.at

Abstract

Due to growing concern over the welfare of livestock and changing legislation in the European Union pig farmers will have to adopt new management methods in farrowing pens. An important challenge related to implementation of temporary sow confinement is the optimal timing of confinement in crates, considering the sow welfare and piglet survival. In total, 71 sows and four types of farrowing pens were included in the observational study. To automatically detect the optimal timing of sow confinement we applied computer vision model You Only Look Once X (YOLOX) to detect sows' locations, calculated activity level of sows' based on detected locations and detected changes in sows activity trends with Kalman filtering and fixed interval smoothing algorithm. Results indicated the beginning of nest-building behaviour with a median of 12 h 51 min and ending with a median of 2 h 38 min before the beginning of farrowing based on centroids of sows extracted with YOLOX-large object detection model. It was possible to predict farrowing with good performance on validation dataset i.e. for 29 out of 44 sows, considering that the object detection model had to be trained on the farrowing pens present in the validation dataset. The developed method could be applied to warn the farmer when nest-building behaviour starts and then to confine the sow in a crate when the end of nest-building behaviour is detected. This could reduce labour costs otherwise required for regular control of sows in farrowing compartments.

Keywords: sow, nest-building, computer vision, hay rack use, automated monitoring, deep learning

Introduction

In societies of European Union (EU) member states there is a growing concern over the welfare of livestock. Specifically, phasing out and finally prohibition of the use of cage systems in the EU was proposed in the Citizens' Initiative "End the Cage Age", which was signed by over 1.4 million EU citizens and supported by the European Commission (EU 2021). Moreover, in recent years national legislation on animal welfare in two EU member states imposed limitations on the use of crates in farrowing pens i.e. to only the critical period of piglet lives in Austria (BMGÖ 2012) and to a maximum of 5 days in Germany (BMEL 2021). These limitations will become mandatory for pig farmers in both countries in 2033 and 2036, respectively. As a consequence pig farmers in the EU will have to adopt new management methods in farrowing pens i.e. temporary sow confinement.

An important challenge related to implementation of such a management system is the optimal timing of sow confinement in crates, considering the sow welfare and piglet survival. One of the possibilities is to confine sows in crates after the nest-building behaviour is finished but before farrowing starts (Goumon et al., 2022). However, in practical farm conditions, due to the biological variability in gestation length (Sasaki and Koketsu 2007), such precision in confining sows is only possible if farm staff performs time-consuming observations of sow's behaviour, including in the night time.

Precision Livestock Farming (PLF) has potential to automatise monitoring sows' behaviour in farrowing pens and to indicate the right timing of sow confinement in crates i.e. at the end of nest-building behaviour (Oczak, Maschat, and Baumgartner 2020). As a first step in development of computer vision methods for farrowing prediction we aim to test if similar performance of farrowing prediction as with other sensor technology i.e. ear-tag accelerometer (Oczak, Maschat, and Baumgartner 2020), can be achieved with computer vision applied for monitoring the activity level of sows.

In this study we aim to select the optimal method of You Only Look Once X (YOLOX) object detection algorithm i.e. nano, tiny, small, medium, large or extra large, for activity monitoring of sows in real-time (Ge et al., 2021). YOLOX also surpasses RetinaNet, used by us in the previous research, in terms of speed and accuracy (Lin et al., 2017). We hypothesise that YOLOX will provide an optimal trade-off between the speed and accuracy for activity monitoring in farrowing pens. We aim to test the performance of YOLOX methods on unseen farrowing pens and animals. The second objective was to validate the previously tested Kalman filtering and fixed interval smoothing (KALMSMO) algorithm, for farrowing prediction (Oczak, Maschat, and Baumgartner 2020). This aims to test if the KALMSMO algorithm can achieve similar performance for farrowing prediction independently of whether activity of sows was estimated based on an accelerometer data, as in our previous research (Oczak, Maschat, and Baumgartner 2020), or image data.

Materials and methods

Animals and housing

The observation was conducted in 2 stages at Medau, the pig research and teaching farm (VetFarm) of the University of Veterinary Medicine Vienna, Vienna, Austria. The first dataset was collected between June 2014 and May 2016, while the second one between December 2021 and July 2022. In total, 71 Austrian Large White sows and Landrace × Large White crossbreds sows were included in the trials. The sows were kept in four types of farrowing pens with the possibility of keeping a sow either unconfined or in a farrowing crate. Out of 71 sows, 9 were kept in SWAP (Sow Welfare and Piglet Protection) pens (Jyden Bur A/S, Vemb, Denmark), 9 in trapezoid pens (Schauer Agrotronic GmbH, Prambachkirchen, Austria), 9 in wing pens (Stewa Steinhuber GmbH, Sattledt, Austria) and 44 in BeFree pens (Schauer, Prambachkirchen Austria). None of the animals included in the observations was confined in a farrowing crate from the introduction to the farrowing pen until the end of farrowing.

The observational period was from the introduction of the sow to the farrowing room until the end of farrowing. The farrowing pens that were recorded as part of the first dataset i.e. SWAP, trapezoid and wing, were located in the testing unit of the farm. The second dataset contained only BeFree pens which were located in the production unit of the farm.

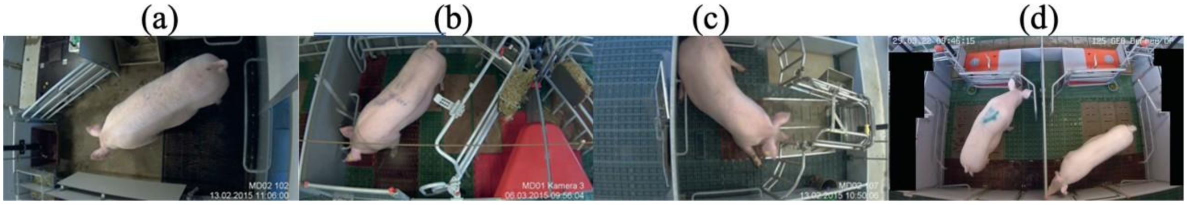

Figure 1; Farrowing pens with possibility of temporary crating. (a) SWAP pen, (b) trapezoid pen, (c) wing pen and (d) 2 BeFree pens

Video recording

The behaviour of sows was video recorded from introduction to the farrowing pens until the end of farrowing with two-dimensional (2D) cameras in order to create a dataset that could be labelled. Each pen in the first dataset (SWAP, trapezoid and wing pens) was equipped with one IP camera (GV-BX 1300-KV, Geovision, Taipei, China) locked in protective housing (HEB32K1, Videotec, Schio, Italy) hanging 3 m above the pen, giving an overhead view. In dataset 2 each IP camera (GV-BX2700, Geovision) was installed with a top view on 2 farrowing pens (BeFree). Additionally, above each farrowing pen in both datasets infrared spotlights (IR-LED294S-90, Microlight, Moscow, Russia) were installed in order to allow night recording. The images were recorded with 1280 × 720 pixels resolution, in MPEG-4 format, at 30 fps for dataset 1, while for dataset 2 at 25 fps.

The cameras used for recording of the first dataset (SWAP, trapezoid and wing pens) were connected to a PC on which Multicam Surveillance System (8.5.6.0, Geovision, Taipei, China) was installed. The system allowed simultaneous recording of images from 9 cameras. The PC had a processor Intel5, CPU 3330, 3 GHz (Intel, Santa Clara, CA, USA) with 4 GB of physical memory. The operating system was Microsoft Windows 7 Professional (Redmond, WA, USA). The first dataset was stored on exchangeable, external 2 and 3 TB hard drives. The cameras used for recording of the second dataset (BeFree pens) were connected to a server for storage of video data (Synology, Taipei, Taiwan) with 4 cores, 8GB memory and 260TB storage.

Dataset

The dataset was composed of video recordings collected in a period from introduction to farrowing pen until 24 h after the end of farrowing. The dataset 1, which contained recordings of 27 sows in SWAP, trapezoid and wing pens, was used for training of YOLOX object detection models and contained 4,667 h of video recordings. The dataset 2, which contained recordings of 44 sows in BeFree pens, was used for testing of YOLOX object detection models, calculation of activity of sows and implementation of KALMSMO farrowing prediction models. It contained 17,713 h of video recordings. The division of recorded videos on dataset 1 and 2 was motivated by the first objective of this study i.e. to test the performance of YOLOX methods on unseen farrowing pens and animals. This dataset division allowed simulation of the expected performance of the models when implemented in a new environment.

Data labelling

To create a reference dataset on the basis of which further data analysis could be performed we labelled the time of the beginning of farrowing of each individual sow (n = 71). It was defined as the point in time when the body of the first-born piglet dropped on the floor. The time of birth of the last piglet indicated the end of farrowing. Labelling software Interact (version 9 and 14, Mangold International GmbH, Arnstorf, Germany) was used to label the beginning and ending of farrowing in dataset 1. For labelling of dataset 2 we used labelling software Boris (version 7.9.15).

For dataset 1 frame selection was performed according to the procedure described in Oczak et al. (2022). For the purpose of selection of specific frames to be used for labelling we applied the k-means algorithm described in Pereira et al. (2019). K-means algorithm was used to select images with the least correlation. In the dataset the k-means algorithm identified 14 242 frames that were the most different between each other (Table 1). One object class was labelled by a trained human labeller on each frame out of selected 14,242 frames i.e. body of the sows. Computer Vision Annotation Tool (version 3.17.0) was used to label the frames (Sekachev et al., 2020). Sow's body was labelled with a rectangle so that the centre of an object was placed in the centre of the rectangle.

Table 1: Selected frames with k-means algorithm for dataset 1 and 2.

Dataset	Duration of video recordings [h]	Frames selected from periods	N. frames selected
1	4,667	introduction to farrowing pen, one day before farrowing, day of farrowing	14,242
2	17,713	from introduction to farrowing pen to one day after farrowing	1,000

For dataset 2 frame selection was performed similarly as for dataset 1 with the same k-means algorithm. However, we selected 500 frames from all the videos recorded for all sows in dataset 2, recorded in a period from introduction to the farrowing pen until one day after farrowing. Because in dataset 2 there were 2 sows under one camera view (Figure 1d) the number of frames used for labelling of sows was increased to 1000 by masking the view on either right or left BeFree pen. In dataset 2 one object class was labelled on each frame out of selected 1,000 frames i.e. body of the sows. Coco annotator was used to label the frames (version 0.11.1) (Brooks 2019).

YOLOX object detection model

The OpenMMLab toolbox was used to train, validate and test the methods of YOLOX i.e. nano, tiny, small, medium, large and extra large. Parametrization of the methods of YOLOX was used as implemented in MMDetection i.e. optimizer stochastic gradient descent (SDG) with learning rate 0.01 and momentum 0.9. Similarly images were augmented as implemented in MMDetection with mosaic, random affine, mixup, random horizontal flip and colour jitter. No changes were made to the architecture of YOLOX methods, optimizer or augmentations provided in MMDetection. Python version 3.8 was used with MMDetection.

We designed 2 experiments to test the performance of methods of the YOLOX algorithm in terms of generalisation ability and inference speed. In both experiments out of 15,242 labelled images, 9,969 (65.4%) were randomly selected for the training set, 4,273 (28%) for the validation set and 1,000 (6.6%) for the test set. In experiment 1 training and validation sets included images from dataset 1, while test set from dataset 2. Thus, in experiment 1 it was possible to test the generalisation ability of the YOLOX on new unseen sows and farrowing pens (BeFree). In experiment 2 all 4 pen types and sows were represented in training, validation and test sets (Figure 2).

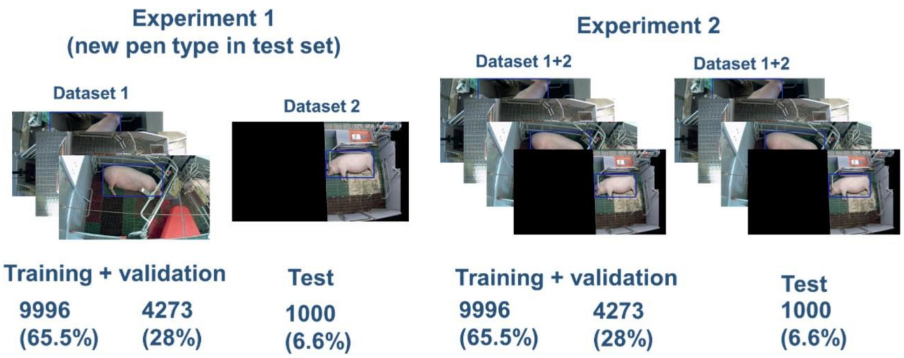

Figure 2: Experiments

Training was set to 100 epochs and was done on RTX Titan (NVIDIA, Santa Clara, US) with evaluation of validation and test set performances with every 5 epochs. Performance of the models was evaluated with standard 12 COCO evaluation metrics e.g. Average Precision (AP) and Average Recall (AR) (Lin et al., 2014). The optimal model was selected by the highest value of the primary COCO challenge metric Average Precision AP on the test set in both experiments. Speed of inference for each YOLOX method was estimated by inferring sows' locations with MMDetection function inference_detector on 1000 frames in the test set.

Activity level of sows

Two best performing models, one from experiment 1 and a second one from experiment 2, were used to extract sows' locations in the videos recorded in the sow observational period, in BeFree pens. Bounding boxes indicating sows' location were extracted in 1 fps out of videos recorded in 25 fps. In the next step euclidean distance was calculated between centroids of extracted bounding boxes. Euclidean distance was further smoothed with a mean calculated on a sliding window of 24 h with 15 min steps, similarly as in Oczak et al. (2020) where standard deviation was used on the same window size and the same step to process ear-tag accelerometer data. This allowed elimination of variation in activity related to diurnal rhythms. Activity level was not labelled by a human labeller but only automatically detected by computer vision technique.

Farrowing prediction

To estimate the dynamics of activity of sows the Kalman filtering and fixed interval smoothing (KALMSMO) algorithm was used as described in Oczak et al. (2020) with the same values of hyper-parameters of the model. The KALMSMO algorithm was fitted to an input variable i.e. euclidean distance at a fixed interval of 48 h and expanded recursively by 15 min steps until the trend in animal activity changed to significantly increasing. The increase in activity trend was indicated by euclidean distance reaching a higher value than the upper confidence interval of the estimated trend. Then the "first-stage" alarm was raised. The preferred time frame for the "first-stage" alarms was within 48 h before the onset of farrowing, and the alarm was not supposed to be generated after the onset of farrowing (Oczak, Maschat, and Baumgartner 2020). The "second-stage" was raised when the trend in animal activity changed to significantly decreasing. This was indicated by the input variable reaching a lower value than the lower confidence interval of the estimated trend. This alarm could be interpreted as an indication that nest-building behaviour had ended. The preferred time frame for the "second-stage" alarm was after the "first-stage" alarm (within 48 h before the onset of farrowing) and not later than the end of farrowing (Oczak, Maschat, and Baumgartner 2020).

Analysis was performed with a commercial software package (MATLAB 2019b, The MathWorks, Inc., Natick, US) and function irwsm of CAPTAIN toolbox (Young 2006) was used to fit the KALMSMO algorithm.

Results and discussion

Performance of models in experiment 1 was generally worse than in experiment 2. The best model for detection of sows in farrowing pens, on unseen environments i.e. BeFree pens was YOLOX-medium, which had the highest AP of 84.2 on the test set in comparison to the other models. In experiment 2, the model with the highest AP of 95.4 on the seen environment was YOLOX-large. Thus, for extraction of centroids of sows these two models were used, YOLOX-medium trained for 70 epochs and YOLOX-large trained for 100 epochs (Table 2).

In the study of Küster et al. (2021) YOLOv3 was used to detect different parts of sows' bodies in farrowing pens i.e. heads, tails, legs and udder but not the whole sows' bodies as in our study. They detected heads with 97 AP_{50}, tails with 78 AP_{50}, legs with 75 AP_{50} and udder with 66 AP_{50}. Performance of YOLOX-large used in our study was better in detecting the whole bodies of sows with a perfect result of 100 AP_{50}. YOLOv3 was also used in the study of van der Zande et al. (2021). However in this study the model was used to detect whole bodies of piglets in a group pen. Nearly perfect AP_{50} of 99.9 was achieved. In both studies, the first of Küster et al. (2021) and the second of van der Zande et al. (2021), validation of object detection models was performed on the same pens as the training of the models. This revealed the performance of the trained models in these pens. However, in PLF we aim to implement the trained models on multiple farms, in different pens and environments (Berckmans 2017). To test YOLOX's generalisation ability we validated the models not only on seen but also unseen farrowing pens i.e. BeFree. The difference between performance of both models on the test set with BeFree pens, the first trained on BeFree and the second trained on the other farrowing pens was 11.2 AP, indicating the better performance of the model which was trained on BeFree pens. Results of farrowing prediction with the KALMSMO models indicated the "first-stage" alarms with median of 10 h 46 min and 12 h 51 min based on centroids of sows extracted with YOLOX-medium and YOLOX-large, respectively (Figure 3). "First-stage" alarms were raised slightly later (2 h 5 min) based on application of the YOLOX-medium model trained in experiment 1. This was confirmed by further analysis of distribution of alarms with the 1st quartile of "first-stage" alarms at 4 h 8 min in comparison to 6 h 2 min and the 3rd quartile at 16 h 17 min in comparison to 19 h 43 min.

Table 2: Best performance of YOLOX methods evaluated by AP

Experiment	Dataset	Method	Epoch	AP	AP_{50}	AP_{75}
1	validation	YOLOX-large	100	96.9	99.0	98.9
1	test	YOLOX-medium	70	84.2	99.0	98.9
2	validation	YOLOX-medium	100	96.5	100	99.0
2	test	YOLOX-large	100	95.4	99.0	98.9

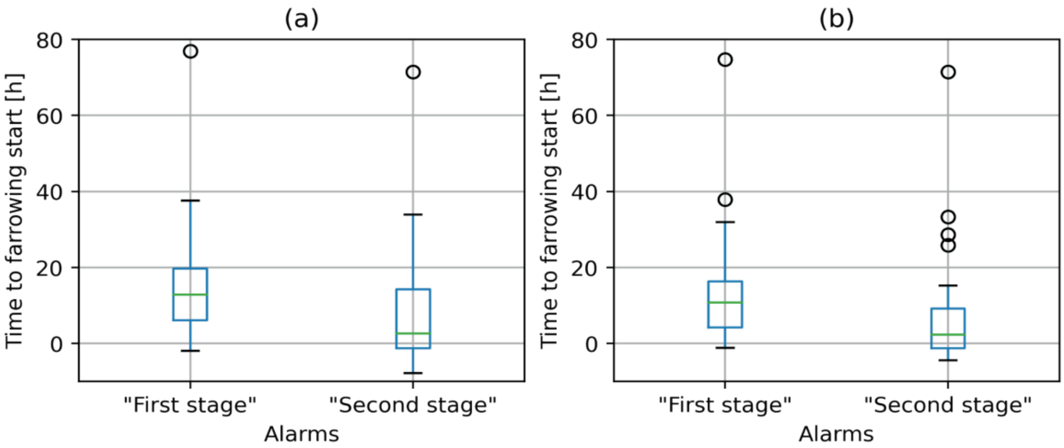

Figure 3: Distribution of duration between time of alarms and beginning of farrowing. (a) Centroids of sows were extracted with YOLOX-medium trained in experiment 1. (b) Centroids of sows were extracted with YOLOX-large trained in experiment 2.

Similarly "second-stage" alarms were raised slightly later based on application of the YOLOX-medium model trained in experiment 1 with median very near to the beginning of farrowing at 2 h 17 min in comparison to 2 h 38 min based on the YOLOX-large trained in experiment 2. The other metrics of distribution of "second-stage" alarms represented a similar difference in timing between both models i.e. first and third quartiles (Figure 3). These results showed that the difference in performance of 11.2 AP between both YOLOX models, first trained in experiment 1 and second in experiment 2 had little impact on the difference in timing of "first and second-stage" alarms.

However, the analysis also revealed that the "first-stage" alarms were not raised for 20 sows out of 44 (45%) when predictions of farrowing were based on centroids extracted with YOLOX-medium trained in experiment 1. This was much worse results in comparison to only 13 sows out of 44 (30%) when YOLOX-large trained in experiment 2 was used.

Performance of the farrowing prediction model KALMSO applied for the first time on ear-tag accelerometer data in Oczak et al. (2020) was confirmed in our current study when model YOLOX-large trained in experiment 2 was used. In both studies the parameters of KALMSMO e.g. nose-to-variance ratio, CI limits, were the same. In Oczak et al. (2020) the "first-stage" alarms were raised in the 48 h period before the beginning of farrowing for 18 out of 26 sows (69%), while in this study for 29 out of 44 sows (66%). The "second-stage" alarms were raised for 17 out of 26 sows (65%) within 48 h before the beginning of farrowing until the end of farrowing in results of Oczak et al. (2020) and 28 out of 44 (63%) in our current study. The results of both studies were also very similar in terms of timing of the "first and second-stage" alarms.

Conclusions

For implementation of the farrowing prediction methodology on the other farms than VetFarm Medau we recommend application of YOLOX-medium trained in experiment 2 of our study for 70 epochs. This model seemed to generalise better than the other models on new unseen farrowing pens and it was trained on all 4 types of farrowing pens available in our dataset. Presence of all farrowing pens in the training set might improve the performance of the model in new, unseen environments. The developed method could be applied to warn the farmer when nest-building behaviour starts and then to confine the sow in a crate when the end of nest-building behaviour is detected. This could reduce labour costs otherwise required for the regular control of sows in farrowing compartments. The future work will be focused on estimation of benefits for animal welfare from implementation of this monitoring system.

References

Bellamy, R., Castiglione, D., and Shaw, J. (2006) *Making European Citizens*. Springer.

Berckmans, D. (2017) General introduction to precision livestock farming. *Animal Frontiers* 7(1), 6-11.

BMGÖ. (2012) Verordnung des bundesministers für gesundheit, mit der die 1. tierhaltungsverordnung geändert wird. *Federal Law Journal for the Republic of Austria,* 61.

des Bundesrates, B. (2020) Siebte verordnung zur änderung der tierschutz-nutztierhaltungsverordnung. 992. *Sitzung des Bundesrates am* 3, 0301-0400.

Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J. (2021) Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv 2107, 08430*.

Goumon, S., Illmann, G., Moustsen, V.A., Baxter, E.M., and Edwards, S.A. (2022) Review of temporary crating of farrowing and lactating sows. *Frontiers in Veterinary Science* 9, 811810.

Brooks, J. (2019) *COCO annotator*. *URL: https://github.com/jsbroks/coco-annotator/*.

Küster, S., Nolte, P., Meckbach, C., Stock, B., and Traulsen, I. (2021) Automatic behavior and posture detection of sows in loose farrowing pens based on 2d-video images. *Frontiers in Animal Science*, 64.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017) Focal loss for dense object detection. In: *Proceedings of the IEEE international conference on computer vision*.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C.L. (2014) Microsoft coco: Common objects in context. In: *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings,* V 13.

Oczak, M., Bayer, F., Vetter, S., Maschat, K., and Baumgartner, J. (2022) Comparison of the automated monitoring of the sow activity in farrowing pens using video and accelerometer data. *Computers and Electronics in Agriculture* 192, 106517.

Oczak, M., Maschat, K., and Baumgartner, J. (2019) Dynamics of sows' activity housed in farrowing pens with possibility of temporary crating might indicate the time when sows should be confined in a crate before the onset of farrowing. *Animals* 10(1), 6.

Pereira, T.D., Aldarondo, D.E., Willmore, L., Kislin, M., Wang, S.S.-H., Murthy, M., and Shaevitz, J.W. (2019) Fast animal pose estimation using deep neural networks. *Nature methods* 16(1), 117-125.

Sasaki, Y., and Koketsu, Y. (2007) Variability and repeatability in gestation length related to litter performance in female pigs on commercial farms. *Theriogenology* 68(2), 123-127.

Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., Osmanov, D.K., Zankevich, A., Dmitriy, M.M., and Johannes, M. (2020) Chenuet a andre telenachos. *Opencv/cvat* 1.

van der Zande, L.E., Guzhva, O., and Rodenburg, T.B. (2021) Individual detection and tracking of group housed pigs in their home pen using computer vision. *Frontiers in Animal Science* 2, 669312.

Young, P.C., Tych, W., and Taylor, C.J. (2009) The captain toolbox for matlab. In: *IFAC Proceeding* 42(10), 758-763.