



Article

DigiPig: First Developments of an Automated Monitoring System for Body, Head and Tail Detection in Intensive Pig Farming

Marko Ocepek ^{1,*} , Anja Žnidar ^{1,2}, Miha Lavrič ³, Dejan Škorjanc ²  and Inger Lise Andersen ¹

¹ Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, 1432 Ås, Norway; anja.znidar@student.um.si (A.Ž.); inger-lise.andersen@nmbu.no (I.L.A.)

² Faculty of Agriculture and Life Sciences, University of Maribor, Pivola 10, 2311 Hoce, Slovenia; dejan.skorjanc@um.si

³ Research Group Ambient Intelligence, Saxion University of Applied Sciences, P.O. Box 70.000, 7500 KB Enschede, The Netherlands; m.lavric@saxion.nl

* Correspondence: marko.ocepek@nmbu.no; Tel.: +47-414-30-972

Abstract: The goal of this study was to develop an automated monitoring system for the detection of pigs' bodies, heads and tails. The aim in the first part of the study was to recognize individual pigs (in lying and standing positions) in groups and their body parts (head/ears, and tail) by using machine learning algorithms (feature pyramid network). In the second part of the study, the goal was to improve the detection of tail posture (tail straight and curled) during activity (standing/moving around) by the use of neural network analysis (YOLOv4). Our dataset ($n = 583$ images, 7579 pig posture) was annotated in Labelbox from 2D video recordings of groups ($n = 12-15$) of weaned pigs. The model recognized each individual pig's body with a precision of 96% related to threshold intersection over union (IoU), whilst the precision for tails was 77% and for heads this was 66%, thereby already achieving human-level precision. The precision of pig detection in groups was the highest, while head and tail detection precision were lower. As the first study was relatively time-consuming, in the second part of the study, we performed a YOLOv4 neural network analysis using 30 annotated images of our dataset for detecting straight and curled tails. With this model, we were able to recognize tail postures with a high level of precision (90%).

Keywords: pig; welfare; image processing; object detection; deep learning; smart farming



Citation: Ocepek, M.; Žnidar, A.; Lavrič, M.; Škorjanc, D.; Andersen, I.L. DigiPig: First Developments of an Automated Monitoring System for Body, Head and Tail Detection in Intensive Pig Farming. *Agriculture* **2022**, *12*, 2. <https://doi.org/10.3390/agriculture12010002>

Academic Editor: Michael Schutz

Received: 31 October 2021

Accepted: 12 December 2021

Published: 21 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

To monitor animal behaviour, classical approaches are used which involve real-time manual observation or the manual analysis of recorded animal behaviours. These methods are labour-intensive and a video from one experiment may take several months to analyse. Although standardized methods and analysis programs are used, most of the data still need to be collected using manual methods. Thus, automatic monitoring is desirable and urgently needed. Several methods using 2D cameras for detection and tracking have been investigated [1–3]. As is the case with image analysis techniques, these methods suffer the same problem: visual cues are unreliable and similar objects might be difficult to differentiate. With the blooming of machine learning/deep learning (ML/DL) research in images in recent years, significant improvements have been made in animal shape detection [4,5] and behavioural sequence detection [6]. Individual pigs can be identified on the basis of their inherent dimensions and colour [7].

Several sensor modalities are now available for the automatic monitoring of behaviour. For instance, deviations in drinking and feeding and the frequency of coughs and vocalisations have been registered by using such systems [8]. Deviations from behavioural synchrony in groups of pigs is important as they tend to show very synchronous activity

patterns, and the individuals deviating from this pattern could potentially be suffering health or welfare issues. However, if one wants to monitor several pens with many pigs in each and gain an insight into their welfare status, we need cameras from above with a slight side angle, which would thus make it easier to recognize pigs based on shapes rather than faces as their faces are usually oriented towards the ground. The face will thus be less visible than their body posture, shape, tail or ears. Detecting individual pigs and their body parts by using deep learning-based computer vision has great potential as a welfare assessment tool to define a positive/negative affective state in individual pigs and the interaction between them (head to tail/ear proximity to define ear/tail biting). A two-dimensional imaging system supported by deep learning can be successfully utilized to detect the standing and lying (belly and side) postures of pigs under commercial farming conditions [4]. Data from different commercial farms were used for the training and validation of the proposed models. Experimental results show that, for instance, the R-FCN ResNet101 DL-network was able to detect lying and standing postures with a mean precision of more than 93%. This is extremely interesting as both positive behaviours, such as play and exploration, and negative behaviours such as aggressive conflicts are associated with certain postures that can most likely be recognized from images. Some have used deep learning for the automatic recognition of sows' nursing behaviours in 2D images, with a precision of 97.6% [5]. Faster R-CNN and ZFnet were applied to recognize the individual feeding behaviours of pigs [9], where each pig in the barn was labelled with a letter. Their proposed method was able to recognise pigs' feeding behaviours with a precision of 99.6%. Image analysis techniques using fully convolutional networks (FCNs) appear to be among the most promising methods for the automatic recognition of sow behaviours from video sequences. In a study of lactating sows [6], features that evaluated the temporal motions of the animals were extracted, and these spatial and temporal features were then put into a hierarchical classifier for behavioural recognition. Based on the 468,000 frames of three sows, the accuracies of behavioural classification compared to manual scoring was: 98% for drinking, 95% for feeding and 88% for nursing, respectively.

The most reliable and preventive way of ensuring the positive welfare state of the pig is to understand how species-specific signals can serve as immediate non-invasive indicators of the individuals' affective state (i.e., their mental and physical condition). Pigs react (behave) differently to signals in their environment, for example, in harmful (suffering) or rewarding situations (pleasure) [10]. Behaviour expressions can be used to describe the affective states of domestic pigs [10,11]. Thus, pig behaviour might be the most powerful and efficient early warning tool to monitor welfare at an individual level as they can predict more serious welfare and health problems that can occur at a later stage. They are honest signals (pig postures) and responses to the physical and social environment as well as the caretaker. Thus, implementing a camera-based monitoring system can serve as an important tool for on-farm preventive animal welfare work as behaviours represent early warning signs of a positive (good) vs. negative (poor) welfare state in pigs.

In such automated monitoring of pigs' behaviour, a focus should be on individual recognition while the pig is in lying or standing position. As pigs are social animals, they spend most of their time lying in proximity to or over pen-mates, which makes them less detectable. Therefore, it is still a problem to detect and recognise individual pigs at every point in their life span. Furthermore, it is of great importance to monitor pig body parts such as head with ears and tail. In a barren environment, pigs are likely to manipulate the ears and tail of pen-mates, a precursor to injurious ear and tail biting [12]. Tail and ear injuries can be sources of infection resulting in further suffering, weight loss and can potentially lead to carcass condemnation at slaughter [13]. Therefore, the monitoring and identification of individual pig tail and head/ears are of great importance for the future detection of individual pigs and biting outbreaks on farms. In addition, tail posture (straight down vs. curled) is associated with affective state in pigs. While a straight tail in an individual pig is linked to a negative affective state, a curled tail is linked with a neutral-to-positive state [12]. Thus, it is important to develop a robust automated monitoring system of individual pig

body parts (head with ears and tail) which could potentially lead to a better understanding of pig needs in their environment, to prevent tail/ear biting and determine the welfare (negative vs. positive) status of pigs on the farm.

The goal of the present paper was to develop an automated monitoring system for pig body, head and tail detection for future behavioural study applications. In the first part of this study, the aim was to recognize individual pigs in groups (in lying or standing position) and their body parts (head/ears, and tail) by the use of machine learning algorithms for object detection based on the feature pyramid network (FPN) architecture. In the second part of this study, the goal was to improve the detection of tail posture (tail straight and curled) using a YOLOv4 neural network analysis.

2. Materials and Methods

2.1. Animal and Housing

Housing and management routines were described in detail in [12]. The experiment took place at the Pig Research Unit of the Norwegian University of Life Sciences (Animal Research Centre, Ås, Norway). We included 10 litters of 5-week-old pigs (crossbred Norsvin Landrace \times Yorkshire sows inseminated with Duroc boar semen) for a 5-week period. Litter sizes at weaning varied from 12 to 15 siblings for a total of 140 pigs (males, $n = 71$; females, $n = 69$) and each litter was housed in a 7.7 m² pen (Figure 1). Pens were divided into a nest area with a solid floor covered with a 3-cm-thick rubber mattress and hayrack, and an activity/dunging area with a plastic slatted floor, two nipple drinkers, and a feeder (SowComfort pen [14]). Males were surgically castrated by veterinarian between 10 and 14 days of age. Teeth and tails were kept intact.

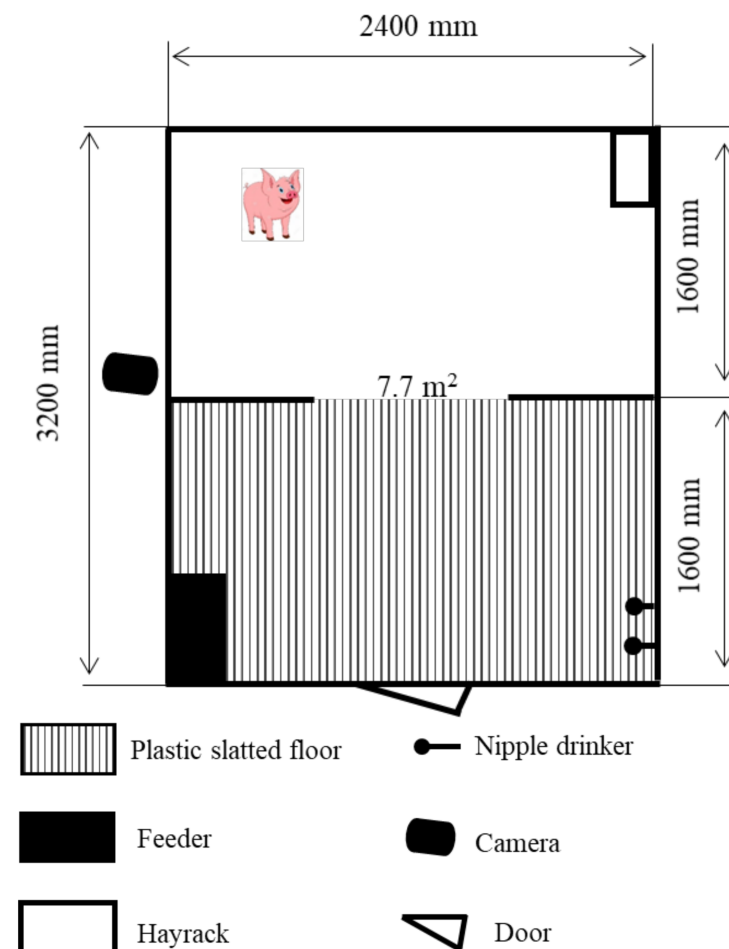


Figure 1. Pen design.

Pen cleaning and the provision of sawdust was performed twice daily at 08:00 and 13:00 h. Pigs were fed *ad libitum*. They also had free access to water from two nipple drinkers. Ambient temperature was initially 22 °C which decreased by 0.5 °C weekly to 19 °C in the last week. Artificial light was provided between 06:00 and 16:00 h.

2.2. Video and Data Collection

A high-definition 2D camera (Foscam FI9,821W, 1280 × 720 P, 25 frames per second, ShenZhen Foscam Intelligent Technology Co., Ltd., Shenzhen, China) was mounted on the roof with a slight side angle above each pen (Figure 1). Video recording started for an hour after the morning (08:30–09:30) and afternoon (13:30–14:30 h) cleans of the pen, from Tuesday to Friday each week for five weeks (5–10 weeks of pig age). The collected 400 h of video were inspected to see whether the quality from any of the cameras in the various bins stood out in a negative or positive direction by a multi-person inspection committee focusing on image sharpness, lighting, colour and contrast, as well as evaluating the visual presence of objects, including pigs, pens and floors. As none of the video sources were singled out as particularly negative or positive, it was decided to select images according to the criteria of: (a) covering as great a variation of postures as possible; (b) selected images spread maximally throughout the filming period; and (c) at least one image from each recording. As a result, a collection of 583 images was made, with an average of 13 pigs per image, for a total of 7579 individual pig postures visible on the images.

2.3. Image Pre-Processing and Manual Labelling

In the 583 images selected, individual pigs in groups and their body parts (heads, tails) were manually labelled (Table 1) using the Labelbox annotation tool (<https://labelbox.com/product/platform/annotate>, accessed on 1 August 2019). Before labelling the images, we tested the best program to proceed with. Out of the three tested programs Labelbox, Imglab, and Supervisely, only Labelbox allowed us to create the right training data, manage the process data in one place and made it easy to use due to a better interface (displayed panel on the left side with the relevant classes and their hot keys). The disadvantage of Labelbox is that it does not support export to Common Objects in Context (COCO) format, and therefore requires extra development. Nonetheless, we used Labelbox in order to label 583 images from a total of 7579 detected pig postures and 23,202 detected objects.

Table 1. Description and definition of individual pig labelling in the first part of the study.

Pig	Visibility	Definition
Lying—good visibility	More than 80% of a pig body, tail, head and feet are visible.	Lying prone or on the side.
Lying—bad visibility	Less than 80% of pig body, tail, head and feet are visible.	
Standing—good visibility	More than 80% of the pig is visible.	Standing while displaying any other behaviours.
Standing—bad visibility	Less than 80% of the pig is visible.	
Head—good visibility	Eyes, snout and mouth are easily visible.	The head with eyes, snout and mouth.
Head—bad visibility	Head is not clearly visible or is facing the other direction. Snout or mouth cannot be labelled.	
Tail curl—good visibility	Curled tail is clearly visible.	Tail coiled up. Only recorded when active (not lying down). Tail hangs straight down or held tucked between the hindlegs. Only recorded when active (not lying down). Tail of lying pig or standing pig with no visible tail.
Tail straight	Straight tail is clearly visible.	
Tail uncertain	Tail is not clearly visible.	

In addition, good and bad visibility was determined and indicated as part of the label (Table 1; Figures 2 and 3). Image annotation was performed by three trained observers (A.Z., M.O., I.L.A.).

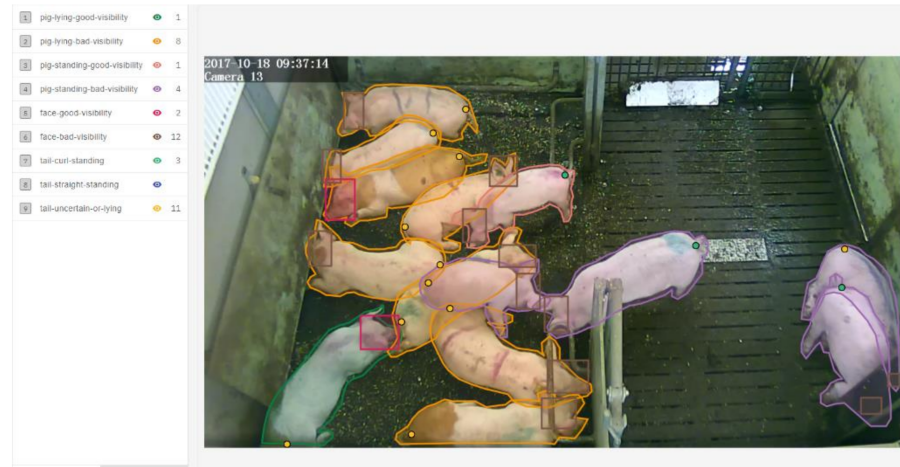


Figure 2. Image labelling overview in the Labelbox annotation tool.

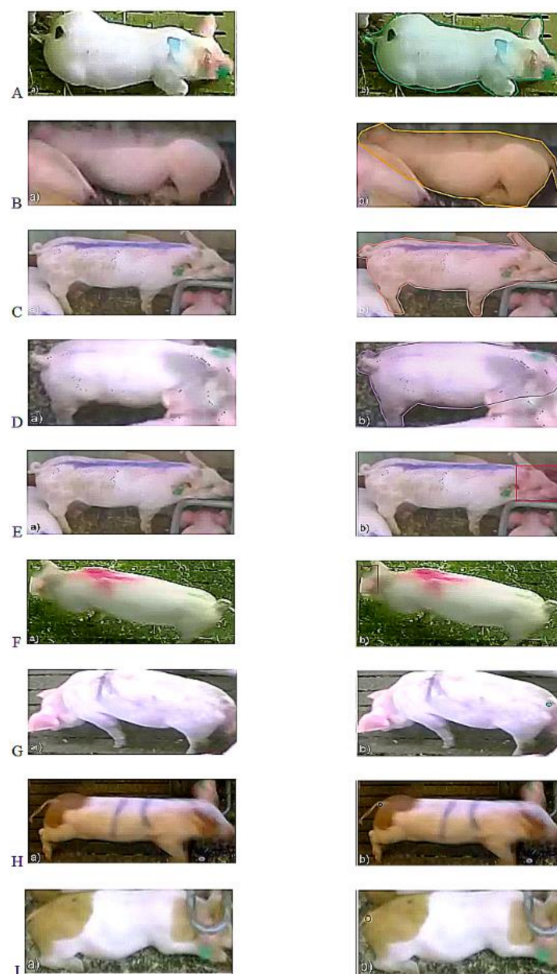


Figure 3. Enlarged examples of good and bad visibility of pigs (lying or standing), faces and tails (curled or straight) for annotation (left, (a) = original; and right, (b) = annotated), from top to bottom: (A) pig lying—good visibility; (B) pig lying—bad visibility; (C) pig standing—good visibility; (D) pig standing—bad visibility; (E) face—good visibility; (F) face—bad visibility; (G) curled tail—standing; (H) straight tail—standing; and (I) tail uncertain or lying.

For the subsequent deep learning application, the COCO format (<http://cocodataset.org/#format-data>, accessed on 20 December 2021) was deemed most appropriate; however, Labelbox does not support this export format. Due to the combination of polygons, rectangles and points in the labelled data, none of the readily available conversion algorithms proved useful, so a custom conversion script was written. The script used a custom-made Python library, based on the json, PIL, os, numpy and re Python libraries, to convert all Labelbox annotations into COCO formatted files. The script created COCO-compatible polygons from the Labelbox point data and COCO-compatible bounding boxes from Labelbox polygons. Furthermore, the option to include/exclude quality segments was added (e.g., pig lying–bad visibility, pig lying–good visibility, etc.), while also removing uncomplete elements that are often left out in annotation.

2.4. Pig, Head and Tail Detection with Mask R-CNN

For the pig, head and initial tail detection, a Mask R-CNN-based deep network model for instance segmentation [15] was used (Figure 4). More specifically, this was a Mask R-CNN 180 Matterport implementation on Python 3.4, Keras 2.0.8 and TensorFlow 1.3 (https://github.com/matterport/Mask_RCNN, accessed on 20 December 2021) [16], based on the feature pyramid network (FPN) and a ResNet101 backbone. A transfer learning approach was applied using pre-trained weights from the MS-COCO dataset which were used as an outpoint for deep learning initialization. Although COCO did not contain images of “Pig”, it contained images of several animals and ‘pig-like’ objects. The Mask R-CNN was trained on 533 images out of 583 and tested on the remaining 50 images for performance (Figure 5).

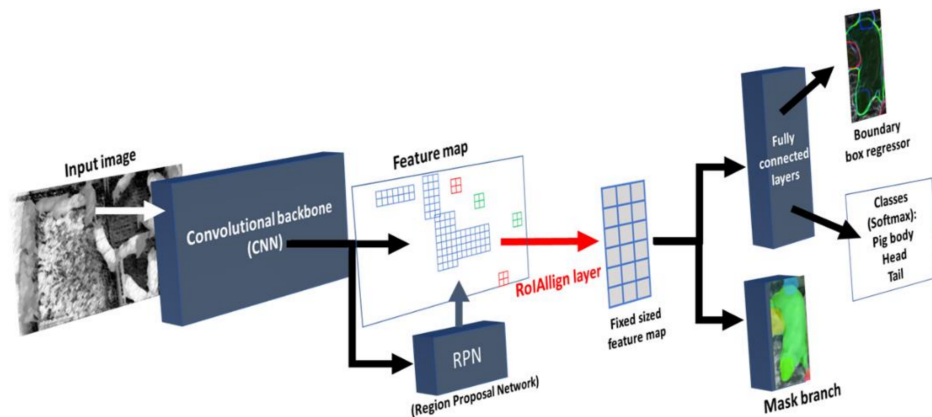


Figure 4. The Mask R-CNN framework for instance segmentation.

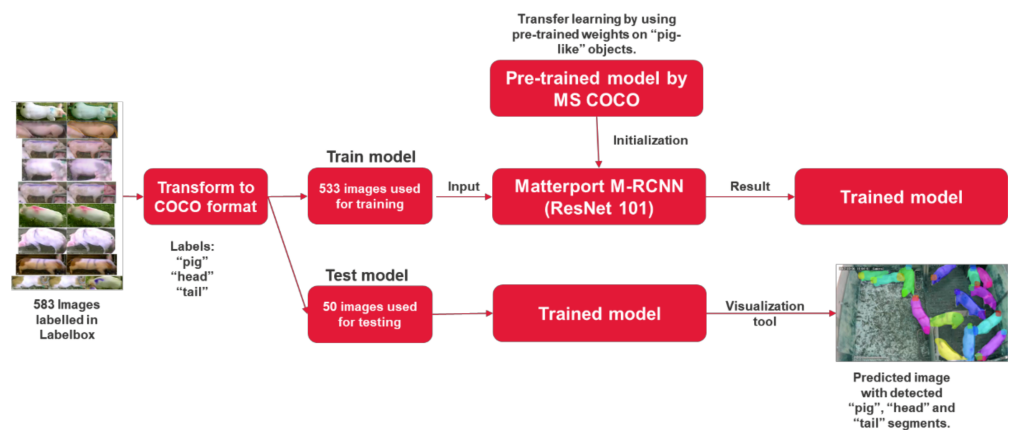


Figure 5. The Digital Pig Matterport M-RCNN-based segmentation framework architecture.

2.5. Tail Detection with YOLOv4

As the Mask R-CNN object detection framework did not perform well for tail detection, a YOLOv4 deep learning alternative [17] was trained. A one-hour-long video (25fps) of 12 pigs in a pen was used for training and validating pig tail detection deep learning network. From the video, a batch of thirty images with on average six pig tails visible per image (range 3–9) were selected and labelled as straight and curled (Table 2: tail labelling) and used to train the YOLOv4 object detector (Figure 6). The object detector was validated for performance on the remaining video frames.

Table 2. Description and definition of straight and curled tail labelling for YOLOv4 tail detection.

Pig Behaviour	Description
Tail:	
Straight	Tail hangs straight down or held tucked between the hind legs. Only recorded when active (not lying down).
Curled	Tail coiled up. Only recorded when active (not lying down).

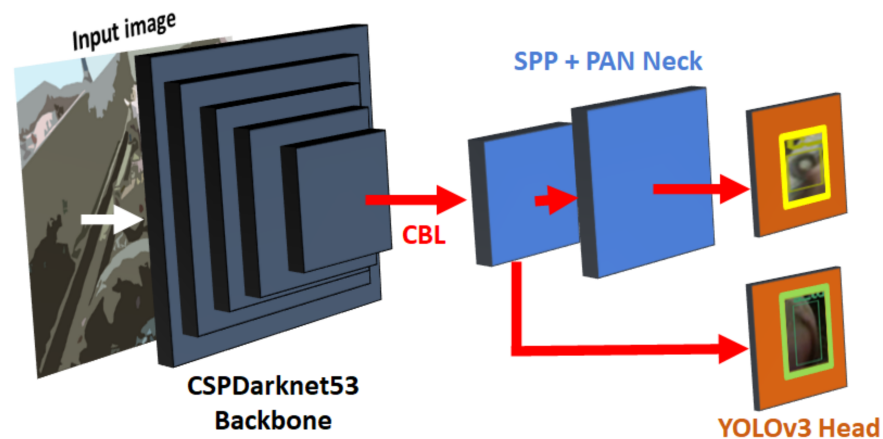


Figure 6. YOLOv4 object detector.

3. Results and Discussion

3.1. Experiment 1—Pig, Head and Tail Labelling and Detection

Out of 23,202 detected objects, there were 3307 showing pigs in lying posture (pig lying–good visibility, $n = 1030$; pig lying–bad visibility, $n = 2277$), 4436 in standing posture (pig standing–good visibility, $n = 1518$; standing–bad visibility, $n = 2918$), 7717 heads (face–good visibility, $n = 1649$; face–bad visibility, $n = 6068$) and 7742 tails (tail–curled–good visibility, $n = 3651$, tail–straight, $n = 273$; tail–uncertain, $n = 3818$).

The Mask R-CNN Matterport implementation was able to recognize pigs with a precision of 96%, tails with 77% and heads with 66% precision—thereby already achieving human-level precision (Figure 7, Table 3) when compared to three independent human observers (Figure 7, Table 3).



Figure 7. Visualization of the object detection of pigs, their heads and their tails by Mask R-CNN.

Table 3. Detection precision and visualization of object detection of pigs, heads and tails by Mask R-CNN.

Test = 50 Images	Pigs	Heads	Tails
Precision	0.96	0.66	0.77
Recall	0.86	0.47	0.60
True positives	579	317	404
False positives	21	159	118
False negatives	91	354	261

The stated precision and recall values were in fact the average precision and average recall calculated over all positive detections in the test set, using intersection over union (IoU) thresholds. IoU is a term mainly used in applications related to object detection, where we train a model to output a bounding box that fits around an object of interest [18]. The IoU describes the extent of overlap between two bounding boxes, one marking the ground truth (the actual labelled object), the other describes the bounding box predicted by our model. The greater the region of overlap, the greater the IoU. Positive detection IoU thresholds are between 0.5 and 0.95. Positive detections or true positives (TPs) indicate the instances in which the object detection algorithm correctly identified pigs, heads or tails on the image. False detections or false positives (FPs) indicate objects wrongly identified by the object detection algorithm as pigs, heads or tails. Finally, missed detections or false negatives (FNs) are instances in which pigs, their heads or tails were labelled on the image, but they were not detected by the object detection algorithm. Average precision is the ratio of how many positives were correctly classified as positive among all positives, so this was calculated as the number of all TPs divided by the sum of all TPs and FPs. The average recall (also known as sensitivity), on the other hand, refers to the proportion of positive detections out of all actual positive detections, and is calculated as the number of all TPs divided by the sum of TPs and FNs.

While the precision of individual pig detection in groups was the highest, the head and tail detection was the lowest. The most important result from our first experiment was that we were able to distinguish individuals in groups of 12 until 15 pigs, which is the most common group size in Norwegian and other European farms. While studies have previously documented that detection precision with the R-FCN ResNet101 DL-network for both standing/lying posture was of 93%, our study revealed that with the feature pyramid network (FPN) architecture, we can achieve almost human-level precision. We showed that our model is even better than previous ones for pigs in standing/lying posture but not the best one for head and tail detection. While the detection of tail posture (straight vs. curled) was correct in 77% of occasions, the head was correct only 66% of the time. As both ears

and tails are relatively small compared to the rest of the body, it is sometimes problematic even for human detection.

Pigs prefer to lie/sleep in close proximity to pen-mates—or even more frequently—lying over them [19]. This makes it harder to identify tails or heads. Out of 7.717 pig heads labelled, their eyes, snout, mouth and ears were only visible in 21% of them. Even though tail posture detection is crucial not only in terms of determining negative welfare (tail biting) but also grading positive status (affective state), we were not able to gain enough variation in our dataset. While we labelled 47% of curled tails, there was only 3.6% of the time with tail straight down. One possible strategy to improve the model performance would be to use a separate script to increase the size of the bounding boxes, as some annotations of the head do not cover the entire head with the face and ears, as seen in Figure 7. This can adversely impact network performance. Another strategy would be to exclude “bad_visibility” annotations from the training batch, especially if those head/tail annotations are hidden behind an annotation belonging to another individual. Bad visibility mostly occurs while pigs are lying down. From our data, we can pinpoint that we should not focus on the head/tail while the pig is lying, but only while it is standing/moving. Even though pigs spend 80% of their time resting/lying in traditional barren environments [20], it is not necessary to have a constantly running automated monitoring system. It is more efficient to have an on-demand monitoring that can scan activities (lying/standing) with high precision, with further focus on only a certain time interval, focusing on the periods where the pigs are most active, to detect the pig body parts (head and tail). This would improve the precision of the model as pigs’ parts would be easier to detect. Ear and tails biting only appears during active periods [12]. As the feature pyramid network (FPN) architecture is time consuming and demands a high number of labelled images, we decided to start with a new method in the second experiment.

3.2. Experiment 2—Tail Detection during Active Phase with YOLOv4

In the second part of this study, the goal was to improve the detection of tail posture (straight or curled) during the active phase using a YOLOv4 neural network analysis. We tested the tail detection, curved or straight, in YOLOv4 as an alternative to Mask R-CNN, as YOLO often beats Mask R-CNN in object detection performance. YOLOv4 was initialized via a standard framework protocol (see Figure 6) and trained on a custom dataset of 30 images with on average of 6 pigs with visible tails. The algorithm detected straight or curled tails with an average precision of 90% (Figure 8).

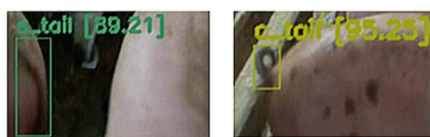


Figure 8. Images from the video showing the detection precision of curved and straight tails from YOLOv4 object detector.

With this new method, we used only 30 annotating images and by focusing on tail detection during the active phase, we were able to improve the precision from 77% to 90%. In this case, we were able to more precisely recognize the pig affective state at the farm level, showing the importance of the proper definition of a golden standard (i.e., tail posture during active phase). In the future, we are looking into possibilities to achieve even higher precision by retraining the model on additional images in an iterative fashion, potentially assisting with the labelling of new images (Figure 9) or perhaps by using semi-supervised learning techniques [21,22] to even further reduce the workload.

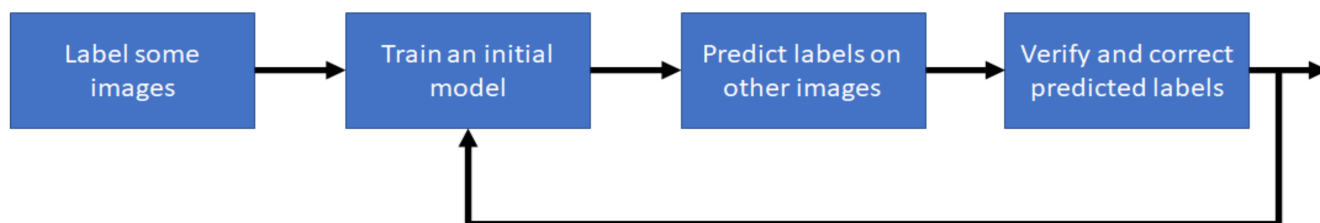


Figure 9. The workflow used to accelerate the labelling process.

Developing an automated monitoring system for body, head and tail detection in pigs is a new time-consuming work method that needs a lot of preparation in testing annotation tools, the preparation/annotation of images, choosing the best/most appropriate model (based on, e.g., Mask R-CNN or YOLOv4 neural networks), training the chosen models and running the models to gain a precision as high as possible, preferably at the human level. However, if the model is developed with a detailed focus on solving all problems such as overcoming bad visibility or focusing on the active rather than passive stage, a high precision model can be the result, such as in our case. This means that, in the methodological part of this study, it was crucial to develop the best performing models based on “golden standards” traits to detect under different circumstances to obtain the best-quality data on those traits one would like to monitor. Therefore, in the next stage, we can begin to focus on welfare assessment problems with tail/ear biting by using YOLOv4 neural network with greater certainty and confidence. As this is again a novel approach in the systematic assessment of pig welfare, there will be a lot of work invested before it can be used at the farm level. In addition, digital solutions are still expensive to be used on all farms; thus, farmers may avoid implementing them. However, after using our developed model in the assessment of welfare, it would be easier, quicker and cheaper than current classical approaches of gathering data based on manual observation in real time or the manual analysis of recorded animal behaviours. Therefore, developing novel digital models leading to welfare improvement is of great importance. Furthermore, with decreasing prices of equipment while increasing operational efficiency together with our future goal of defining a complete digital concept that would work under farm conditions, such a system concept would most likely be implemented as soon as possible. Farmers are namely in need of having better control over their pigs, so bringing this to their attention is most crucial. Currently, farmers do not have the possibility of 24/7 pig monitoring. Our novel digital solutions with their thorough methodological contribution and future implementation in welfare assessment have the potential to be valuable tools for farmers to reduce workload and costs. The main problem of farmers nowadays is that they still have to produce more and more pigs yearly only to survive, meaning that control and contact with each individual pig is being reduced and only a proper digital monitoring concept can ameliorate this trend and consequently improve farmer and pig welfare.

4. Conclusions

In conclusion, out of three tested annotation programs (Labelbox, Imglab and Supervisely), we decided to use Labelbox (since it allowed the creation of the right training data, managing the process data in one place and was easy to use due to a better interface). With the use of machine learning algorithms for object detection based on feature pyramid network (FPN) architecture, the precision of individual pig detection in groups was almost the same as human-level precision. However, the method was time consuming and not optimal for head and tail detection during the passive (lying) and the active (standing/moving phase) phases. By the use of the YOLOv4 neural network analysis, we were able to reduce the human workload and improve tail posture (straight vs. curled) detection precision during the active phase—which is most crucial for reducing the incidence of tail biting and in the evaluation of the welfare state of the pigs on the farm. Our new method can be further explored in detecting behavioural sequences, group synchrony as well as quantifying positive welfare (play, exploration, tail curled and wagging). Most of these behaviours are

associated with certain body postures and by defining such golden standards with high, human-level precision, we could improve the welfare status of the pig. This means that the current classical approaches of gathering data based on manual observation in real time or the manual analysis of recorded animal behaviours in research will be time- and labour-expensive and will be replaced by a cheaper, real-time digital monitoring system. Furthermore, with the implementation of a digital system, we will be able to gather more information about pig behaviour (positive and negative), and thus have better control over them and be able to provide optimal conditions in real time at the farm level.

Author Contributions: M.O.—conceptualization, methodology, investigation, writing for draft, supervising, pig labelling; A.Ž.—pig labelling, investigation, methodology; M.L.—conceptualization, formal analysis, validation, writing—first draft; D.Š.—conceptualization, writing—review and editing, supervising; I.L.A.—supervising, conceptualization, investigation, writing—review and editing, pig labelling. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed by the Norwegian Research Council and Fossli AS (grant number 268158).

Institutional Review Board Statement: The present research was conducted in accordance with the Norwegian laws and regulations controlling experiments and procedures on live animals in Norway. Approval from an ethical review board was not required for this study.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors wish to acknowledge staff at the Pig Research Unit for their technical assistance and for taking good care of the animals. We are especially grateful to the DIGI pig group (Ingrid Melkild and Kristine Hov Martinsen) for their support of this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, Y.; Sun, L.; Zou, Y.; Li, Y.; Sun, L. Individual pig object detection algorithm based on Gaussian mixture model. *Int. J. Agric. Biol. Eng.* **2017**, *10*, 186–193. [[CrossRef](#)]
2. Lee, J.; Jin, L.; Park, D.; Chung, Y. Automatic recognition of aggressive behavior in pigs using a kinect depth sensor. *Sensors* **2016**, *16*, 631. [[CrossRef](#)] [[PubMed](#)]
3. Kashihara, M.; Bahr, C.; Ott, S.; Moons, C.P.H.; Niewold, T.A.; Odberg, F.O.; Berckmans, D. Automatic identification of marked pigs in a pen using image pattern recognition. *Comput. Electron. Agric.* **2013**, *93*, 111–120. [[CrossRef](#)]
4. Nasirahmadi, A.; Sturm, B.; Edwards, S.; Jeppsson, K.H.; Olsson, A.C.; Müller, S.; Hensel, O. Deep learning and machine vision approaches for posture detection of individual pigs. *Sensors* **2019**, *19*, 3738. [[CrossRef](#)] [[PubMed](#)]
5. Yang, A.; Haung, H.; Zhu, X.; Yang, X.; Chen, P.; Li, S.; Xue, Y. Automatic recognition of sow nursing behaviour using deep learning-based segmentation and spatial and temporal features. *Biosyst. Eng.* **2018**, *175*, 33–145. [[CrossRef](#)]
6. Yang, A.; Haung, H.; Zheng, B.; Li, S.; Gan, H.; Chen, C.; Yang, X.; Xue, Y. An automatic recognition framework for sow daily behaviours based on motion and image analyses. *Biosyst. Eng.* **2020**, *192*, 56–71. [[CrossRef](#)]
7. Zhu, W.X.; Guo, Y.Z.; Jiao, P.P.; Ma, C.H.; Chen, C. Recognition and drinking behaviour analysis of individual pigs based on machine vision. *Livest. Sci.* **2017**, *205*, 129–136. [[CrossRef](#)]
8. Matthews, S.G.; Miller, A.L.; Clapp, J.; Plotz, T.; Kyriazakis, L. Early detection of health and welfare compromises through automated detection of behavioural changes in pigs. *Vet. J.* **2016**, *217*, 43–51. [[CrossRef](#)] [[PubMed](#)]
9. Yang, Q.M.; Xiao, D.Q.; Lin, S.C. Feeding behavior recognition for group-housed pigs with the Faster R-CNN. *Comput. Electron. Agr.* **2018**, *155*, 453–460. [[CrossRef](#)]
10. Rudolf, N.; Cardinal, J.A.P.; Hall, J.; Barry, J.E. Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* **2002**, *26*, 321–352. [[CrossRef](#)]
11. Wemelsfelder, F.H.E.; Mendl, M.T.; Lawrence, A.B. The spontaneous qualitative assessment of behavioural expressions in pigs: First explorations of a novel methodology for integrative animal welfare measurement. *Appl. Anim. Behav. Sci.* **2000**, *67*, 193–215. [[CrossRef](#)]
12. Ocepek, M.; Neberry, R.C.; Andersen, I.L. Which types of rooting material give weaner pigs most pleasure? *Appl. Anim. Behav. Sci.* **2020**, *231*. [[CrossRef](#)]
13. D' Eath, R.B.; Niemi, J.K.; Ahmadi, B.V.; Rutherford, K.M.D.; Ison, S.H.; Turner, S.P.; Lawrence, A.B. Why are most EU pigs tail docked? Economic and ethical analysis of four pig housing and management scenarios in the light of EU legislation and animal welfare outcomes. *Animal* **2016**, *10*, 687–699. [[CrossRef](#)] [[PubMed](#)]

14. Andersen, I.L.; Morland, E. Piglet survival and causes of mortality in the Sowcomfort pen: Data from two commercial loose-housed sow herds. In Proceedings of the 50th Congress of the International Society for Applied Ethology; Dwyer, C., Haskell, M., Sandilands, V., Eds.; International Society for Applied Ethology: Edinburgh, UK, 2016; pp. 1–456.
15. He, K.G.G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; O’Conner, L., Ed.; The Institute of Electrical and Electronics Engineers, Inc.: Venecia, Italy, 2017; pp. 2961–2969.
16. Mask r-cnn for Object Detection and Instance Segmentation on Keras and Tensorflow. Available online: https://github.com/matterport/Mask_RCNN (accessed on 12 March 2019).
17. Bochkovskiy, A.C.; Wang, C.Y.; Liao, M.H.Y. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
18. Rezatofighi, H.T.N.; Gwak, J.; Sadeghian, A.; Reid, I.; Savarese, S. Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; Gupta, A., Hoiem, D., Hua, G., Tu, Z., Eds.; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2019; pp. 658–666.
19. Ocepek, M.; Goold, C.; Busančić, M.; Aarnink, A.J.A. Drinker position influences the cleanness of the lying area of pigs in a welfare-friendly housing facility. *Appl. Anim. Behav. Sci.* **2017**, *198*, 44–51. [[CrossRef](#)]
20. Ocepek, M.; Goold, C.; Busančić, M.; Aarnink, A.J.A. Maize silage as enrichment material improves the welfare of growing-finishing pigs in environmentally-friendly pens. *Appl. Anim. Behav. Sci.* **2020**, *230*, 105043. [[CrossRef](#)]
21. Jeong, J.L.S.; Kim, J.; Kwak, N. Consistency-based semi-supervised learning for object detection. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 10759–10768.
22. Sohn, K.Z.Z.; Li, C.L.; Zhang, H.; Lee, C.Y.; Pfister, T. A simple semi-supervised learning framework for object detection. *arXiv* **2020**, arXiv:2005.04757.