



**HAL**  
open science

# Integrating computer vision algorithms and RFID system for identification and tracking of group-housed animals: an example with pigs

Mónica Mora, Miriam Piles, Ingrid David, Guilherme Rosa

► **To cite this version:**

Mónica Mora, Miriam Piles, Ingrid David, Guilherme Rosa. Integrating computer vision algorithms and RFID system for identification and tracking of group-housed animals: an example with pigs. Journal of Animal Science, 2024, 10.1093/jas/skae174 . hal-04621669

**HAL Id: hal-04621669**

**<https://hal.science/hal-04621669>**

Submitted on 25 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Integrating computer vision algorithms and RFID system for identification and tracking of group-housed animals: an example with pigs

Mónica Mora<sup>\*,†</sup>, Miriam Piles<sup>\*</sup>, Ingrid David<sup>#</sup> Guilherme J. M. Rosa<sup>†</sup>

\* Institute of Agrifood Research and Technology (IRTA) - Animal Breeding and Genetics, Caldes de Montbui, Barcelona, Spain 08140

† Department of Animal and Dairy Sciences, University of Wisconsin-Madison, Wisconsin, USA 53706

# GenPhySE, Université de Toulouse, INRAE, ENVT, Castanet Tolosan, France 31326

Corresponding author: [grosa@wisc.edu](mailto:grosa@wisc.edu)

## Funding

This study was part of the project PID2021-128173OR-C21 (GENEF3) and MMF is a recipient of a “Formacion de Personal Investigador (FPI)” associated with the research project RTI2018-097610R-I00. GJMR acknowledges funding from USDA NIFA FACT CIN award# 2021-67021-34150.

## Lay Summary

In precision livestock farming, monitoring animal activity is crucial to ensure their health, well-being, and productivity. While digital cameras and computer vision algorithms offer a promising solution for this task, tracking individual animals of similar appearance when housed in groups can be challenging. Close interaction among animals can lead to a loss of individual identity, which affects tracking accuracy. To overcome this problem, we developed a framework that combines camera images with radio frequency identification (RFID) ear tags. This methodology was applied to a pen housing 12 pigs, with an RFID reader located inside the feeder. Among the pigs, three had unique coat markings, enabling them to be tracked most of the time without losing their identity (87% of the time). The remaining pigs could not be visually distinguished from each other, so information from the RFID system was used to recover lost IDs every time pigs entered the feeder. The framework achieves 97% accuracy in tracking, offering a reliable solution for monitoring group-housed pigs.

## Teaser Text

The combination of computer vision and radio frequency identification can enhance tracking performance of group-housed animals.

Accepted Manuscript

## Abstract

Precision livestock farming aims to individually and automatically monitor animal activity to ensure their health, well-being, and productivity. Computer vision has emerged as a promising tool for this purpose. However, accurately tracking individuals using imaging remains challenging, especially in group housing where animals may have similar appearances. Close interaction or crowding among animals can lead to the loss or swapping of animal IDs, compromising tracking accuracy. To address this challenge, we implemented a framework combining a tracking-by-detection method with a radio frequency identification (RFID) system. We tested this approach using twelve pigs in a single pen as an illustrative example. Three of the pigs had distinctive natural coat markings, enabling their visual identification within the group. The remaining pigs either shared similar coat color patterns or were entirely white, making them visually indistinguishable from each other. We employed the latest version of the You Only Look Once (YOLOv8) and BoT-SORT algorithms for detection and tracking, respectively. YOLOv8 was fine-tuned with a dataset of 3,600 images to detect and classify different pig classes, achieving a mean average precision of all the classes of 99%. The fine-tuned YOLOv8 model and the tracker BoT-SORT were then applied to a 166.7-min video comprising 100,018 frames. Results showed that pigs with distinguishable coat color markings could be tracked 91% of the time on average. For pigs with similar coat color, the RFID system was used to identify individual animals when they entered the feeding station, and this RFID identification was linked to the image trajectory of each pig, both backward and forward. The two pigs with similar markings could be tracked for an average of 48.6 min, while the seven white pigs could be tracked for an average of 59.1 min. In all cases, the tracking time assigned to each pig matched the ground truth 90% of the time or more. Thus, our proposed framework enabled reliable tracking of group-housed pigs for extended periods, offering a promising alternative to the independent use of image or RFID approaches alone. This approach represents a significant step forward in combining multiple devices for animal identification, tracking, and traceability, particularly when homogeneous animals are kept in groups.

## Keywords

BoT-SORT, electronic ear tags, 2D camera, PLF, tracking-by-detection, YOLO

## List of abbreviations

AP@0.5IOU, average precision with 0.5 IoU threshold

CAP, *Centre de Control y Avuació de Porcí*

CNN, convolutional neural network

COCO, Common Objects in Context

FP, false positive

GPS, global positioning system

IoU, Intersection over Union

IRTA, Institute of Agrifood Research and Technology

mAP@0.5IOU, mean average precision with 0.5 IoU threshold

MOT, Multiple object tracking

PLF, precision livestock farming

RFID, radio frequency identification

RGB, Red Green Blue

SQL, structured query language

YOLO, You Only Look Once

Accepted Manuscript

## Introduction

One of the biggest challenges of livestock farming nowadays is the development of systems to monitor animals automatically and individually. The continuous monitoring of individual animals is crucial because it allows for more efficient and effective management, leading to improved animal welfare, economic benefits, and sustainability in livestock production (Anzai and Hirata, 2021). In this context, the concept of precision livestock farming (PLF) has emerged as an approach to creating smart farms that utilize technology and data-driven solutions to continuously and automatically collect livestock information (Berckmans, 2017; Rosa, 2021).

Over the years, different sensors have been developed to monitor individual activity levels in livestock, such as accelerometers that measure the acceleration in the three axes, which can be transformed into individual activity levels using machine learning algorithms (Ringgenberg et al., 2010; Riaboff et al., 2019; Turner et al., 2022). However, the type of behavior they can measure depends on where they are attached to the animal, potentially even affecting the readings and adding noise to the raw data (Chapa et al., 2020). Other types of sensors are GPS trackers or the use of radio frequency identification (RFID) tags, which, on their own, do not provide information directly about the activity levels of an individual. The GPS trackers can provide information on the spatial distribution and movement of individual animals (Bailey et al., 2018; Riaboff et al., 2020), while the RFID tags only can provide information about the occurrence of the animals at the locations where the receivers are installed, which are typically the feeders and drinkers (Williams et al., 2019; Kapun et al., 2020). Wearable sensors for livestock monitoring face some challenges related to durability, battery life, and cost. In addition, their implementation requires animal handling and potential changes in their behaviour and/or damages since wearable sensors have to be attached to the animals.

Computer vision strategies are an alternative or a complement to these devices, allowing for the measurement of animal activity through image and video analysis without the need for animal handling (Fernandes et al., 2020). Moreover, one single camera can monitor multiple animals simultaneously. However, before computing the individual activity levels of each animal, they have to be identified, unlike the sensors previously mentioned, which have a unique ID incorporated in their hardware that allows linking them to the respective animal. The process of identifying individual animals using computer vision algorithms is a complex process, involving two tasks: object detection and tracking. Object detection is the process through which objects in an image are localized and classified. Object tracking algorithms, on the other hand, are responsible for monitoring individual objects. They link objects detected across consecutive frames in a scene while keeping a unique identifier for each one. Tracking-by-detection has emerged as the predominant method for addressing object tracking due to the rapid development of reliable object detectors. In the context of livestock, each individual present in the scene will be an object and if they cannot be differentiated from each other, they will all belong to the same object class (for example, the pig class). Generally, the animals of interest are housed in groups which means that many objects can be detected in a single frame or image. Multiple object tracking (MOT) is the most complex type of tracking since it involves detecting and tracking more than one object at the same time. During this process, animals interact among themselves and can be occluded by others. When this happens, the individual identification may be lost, and the tracking algorithm needs to re-identify it. The ability to re-identify an object breaks down when the objects under study are homogeneous, as is often the case for livestock populations with breeds of individuals

of similar color and size. Pigs, for example, are difficult to distinguish individually thus, in some research applications, artificial marks have been used to differentiate and identify the animals (Kashiha et al., 2013; Yang et al., 2018). However, to ensure good individual identification during a long tracking process, the marks must be re-painted frequently. This limits such applications to selection nuclei or small farms. Another approach to distinguish the pigs in a pen is through face recognition which has been applied by Hansen et al. (2018), but this approach requires entering the pen and obtaining images of pig faces which may involve human intervention. Hence, the goal is to develop a framework for tracking-by-detection pigs housed in groups, in a non-invasively way and without using any artificial mark to differentiate between them.

The YOLO (You Only Look Once) is a powerful detection algorithm that has gained a lot of popularity in recent years, outperforming the results provided by other previous architectures (Lee et al.; Jiang et al., 2020). Different versions of this algorithm have been released over the years, with YOLOv8 being the latest version that scores substantially better than its previous versions. Information on YOLOv8 can be found in the Ultralytics GitHub repository (<https://github.com/ultralytics/ultralytics>). This last version supports a full range of different tasks, including detection, segmentation, pose estimation, tracking and classification. YOLOv8 supports two multi-object trackers, with BoT-SORT (Aharon et al., 2022) being one of them. It ranks among the top in the datasets of MOTChallenge (Milan et al., 2016; Dendorfer et al., 2020), a benchmark for evaluating the different MOT algorithms.

Most MOT methods have been developed for pedestrians, vehicles, or stationary objects. However, there have been few studies conducted on tracking animals. In swine production, some approaches have been investigated in the last years to track individual pigs housed in groups without using artificial marks (Shirke et al., 2021; van der Zande et al., 2021; Guo et al., 2023). However, a system is still needed to recover the identity of the animals when the tracking algorithm loses it after an occlusion or interaction to track them for a longer time. In this context, Guo et al. (2023) suggested that a combination of video-based tracking results with data from sensors like RFID could enhance long-duration tracking performance.

With this in mind, in this study we developed a framework to combine a tracking-by-detection method on video recordings with the information provided by an RFID system. The proposed approach was implemented on a pen with twelve pigs with the aim to identify individual animals and measure how long they could be tracked up to ~3 hours.

## Material and Methods

### Animals and video recording

The experimental procedures used in the research complied with the European directive 2010/63/EU and the Spanish guidelines for the care and use of animals in research (B.O.E. number 34, Real Decreto 53/2013). The animals involved in this experiment come from a Pietran x (Landrace x Duroc) cross. After weaning (70 days of age), they were transported to the swine control center (*Centre de Control y Avuació de Porcí, CAP*) of the Institute of Agrifood Research and Technology (IRTA) located in Girona, Spain. They remained in these facilities until the end of the fattening period (160 days of age). During this time, the pigs were housed in groups of a maximum of 12 animals per pen (0.89 pig/m<sup>2</sup>). Each pen was equipped with an electronic feeder and a single drinking trough, with the total pen dimension being

13.5 m<sup>2</sup> (5.0 m long and 2.7 m wide). Animals had *ad libitum* access to food and water for the whole experiment. The daylight hours corresponded with the natural light hours.

Each pig was fitted with a low-frequency electronic ear tag (Allflex; Merck & Co., Inc., Rahway, NJ, USA) with a unique identification number (animal ID). The RFID reader was located inside the feeding station, the animal ID was recorded every time an animal entered and left it. The reader transmitted the data to a central computer system for processing and storage. The information was saved in an SQL database along with the information collected by the electronic feeder. All together the systems are able to track how many visits each pig performed to the feeding station, how long was each visit, and how much feed was consumed in each visit. Our goal in this study was to leverage on this RFID system and, together with images, develop a framework to identify and track the animals when they are not at the feeding station. Such a system can be useful, for example, to monitor animal activity and associate it with animal growth and feed efficiency.

The analysis was conducted using the information from a single pen where the animals arrived at CAP in May 2023 and completed the fattening period in August 2023. A total of 12 pigs participated in the study, where three of them had natural marks on their coat that allowed them to be distinguished at a glance (pigs A, B and C), two had similar natural marks and were referred to as 'Pair', and the remaining seven of them were white pigs impossible to distinguish from each other, which were named 'White' (Figure 1).

Throughout the study period, the animals were recorded for one day once a week (9 days of recordings). An RGB camera was installed at the center of the pen and at the height of 3.10 m perpendicularly to the floor offering a top view of the entire pen. The camera was a Hikvision model DS-2CD2345G0P-I, with an image resolution of 1280 x 720 and a frame rate of 10 fps.

### **Dataset preparation**

The detection algorithms need to be trained with a wide variety of images to provide satisfactory performance. To prepare our training dataset, we randomly selected 100 frames (images) from each day of the total recording period (i.e., 9 days). Before annotating the images, all of them were cropped to exclude pigs from neighboring pens and resized to the same size of 900x720 pixels. Images were manually annotated using the Windows version of LabelImg (Tzutalin, 2015). The contours of the pig were labeled by a bounding box, where each side of the bounding box touched the pig. Five different pig classes were annotated in each image depending on their coat: A, B, C, Pair and White (Figure 1). After that, a dataset with 900 image files (.jpg) and 900 text files (.txt) was created. Each text file contained the coordinates of the bounding box and the class of each pig present in its corresponding image. All this information was saved in YOLO format.

Data augmentation was used to expand and generalize the dataset, enhancing its diversity and representativeness. Each frame was transformed dynamically including photometric distortion such as brightness adjustments (increasing the brightness 1.5 points), as well as geometric distortion like horizontal and vertical flipping and 180° rotation. This way, the dataset increased by 4 times, being the final dataset of 3,600 images.



## Training the detection model

YOLO is an object detection algorithm that utilizes a single neural network to simultaneously predict bounding boxes along with class probabilities for each object in an image. It utilizes a convolutional neural network (CNN) with three main parts: the backbone, the neck and the head. The backbone is composed of the CNN CSPDarknet53. This CNN consists of 53 layers that are responsible for extracting the initial features of the input image. Then, this information is passed to the neck where all the extracted features are collected and sent to the prediction layer which is the head of YOLO. It consists of multiple convolutional layers followed by a series of fully connected layers to output the predicted bounding boxes, the confidence scores, and the class probabilities for all the objects detected in the image.

An essential characteristic of YOLOv8 involves the incorporation of a self-attention mechanism. This mechanism allows the model to focus on different parts of the image and adjust the importance of different features based on their relevance to the task. Another important characteristic of YOLOv8 is its ability to perform multi-scaled object detection. This means that the model utilizes a feature pyramid network to detect objects of different sizes and scales within an image. Finally, YOLOv8 applies mosaic augmentation during training, which combines four images to give the model more context. This technique is stopped in the last 10 epochs to improve performance.

In this study, the YOLOv8-tiny model pre-trained on the Microsoft Common Objects in Context (COCO) dataset (Lin et al., 2014) was fine-tuned to detect the five different classes of pigs. The 3,600 images were randomly divided into 70% training, 20% validation and 10% testing. The model was trained using the default hyperparameters, except for the number of epochs, which was set to 600, and the optimizer, for which *Adam* was selected.

Intersection over Union (IoU) was used to evaluate the performance of object detection, by comparing the ground truth bounding box (i.e., the annotated bounding box) with the predicted bounding box:

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

If the IoU is larger than a threshold (by default, 0.5), the predicted bounding box is considered a correct prediction. A prediction is classified as true positive, false positive (FP), or false negative (FN) depending on the classification and IoU with the ground truth. A true positive (TP) is a prediction correctly classified and having an IoU higher than the threshold. A FP occurs when a prediction is either incorrectly classified or has an IoU lower than the threshold. A FN refers to ground-truth annotations that do not correspond to any prediction. In this study, the average precision with 0.5 IoU threshold (AP@0.5) for each class and the mean average precision with 0.5 IoU threshold (mAP@0.5) for all the classes were computed, which is the area under the precision-recall curve. The confusion matrix was also evaluated.

## Tracking-by-detection

The finetuned YOLOv8 model and the tracker BoT-SORT (Aharon et al., 2022) were applied to a video lasting 166.7 min (100,018 frames). This video was taken one morning (from 8:37 to 11:24) during the middle of the fattening period. It is important to take into account that the frames corresponding to this video were not used for fine-tuning the YOLO.

In the tracking-by-detection method, the detections produced by the finetuned YOLOv8 were used as input for the tracking algorithm. It is not perfect, and in certain images, it might detect more bounding boxes than the expected number of animals (i.e., FP). An instance of a FP could be when two bounding boxes are generated for the same pig with practically the same coordinates or when a bounding box is generated in the background. Images with FP were identified when more than the expected number of pigs were detected. In these images, the extra bounding boxes with lower confidence scores were considered FP and were excluded from the analysis to prevent interference during tracking. Similarly, in some frames, the finetuned YOLOv8 model can detect fewer bounding boxes than the expected number of animals (i.e., FN). Thus, the tracking algorithm is fully dependent on the quality of the detections. The objective of a tracking algorithm is to assign a unique ID to each detected object in the image and keep this ID over time, i.e. linking detections across frames to form trajectories. In the first context, the algorithm assigns a unique ID to each object, ranging from 1 to the total number of objects in the image. The tracker will estimate the positions of the unique objects in the remaining images or frames to obtain the trajectories of each individual. When the objects are homogeneous, the tracker can fail and be unable to re-identify the same object. In this case, the initial ID of that object is deleted and a new one is created, creating a new trajectory. Thus, the same object can present many trajectories with a different ID for each one. In other words, a trajectory is a track where an object has a consistently associated ID in sequential frames. (Figure 2). To associate the detections across frames, two approaches are implemented in the tracking algorithm: the Kalman filters and Hungarian algorithm. The Kalman filters are responsible for estimating the position of the object in the next frame while the Hungarian algorithm determines whether an object detected in the current frame is the same as the one detected in the previous frame.

BoT-SORT introduces improvements in Kalman filters compared to their predecessors (SORT, DeepSORT) by modifying the parameters of the filter matrix to better match the target. In addition, BoT-SORT adds camera motion compensation which is important in dynamic camera situations, where static objects can appear to move, and objects that are moving can appear to be static. Finally, this tracker proposes a new way of solving the association step by combining the motion and appearance information. A deep explanation of this tracker can be found in the literature (Aharon et al., 2022).

From now on, we will refer to the unique ID assigned by the tracker as the “track ID” to distinguish it from the animal ID (the unique identifier of each pig presented in its RFID tag). At the end of the tracking-by-detection process, all the detected pigs associated with the same track ID will correspond to the same animal. However, multiple track IDs (i.e., multiple trajectories) can be attributed to the same pig. A pig with many track IDs will be one that the algorithm has lost many times. When this happens in the homogenous pigs (White class), the individual identification of the animal is lost.

### **Assignment of an animal ID to each pig**

The process of assigning an animal ID to each pig depended on its classification as a pig with natural marks (pig A, B or C), a White pig, or a Pair pig. For the assignment of the animal ID for the pigs of the White and Pair classes, the information from the RFID system was used. In addition, this information was used to recover the identity of these animals when the tracker lost their identity. This will allow us to extend the duration for which we can track these individual pigs.

*Pigs classified as A, B or C class.* The corresponding animal IDs for each marked pig are as follows: pig A → 7926, pig B → 7940, and pig C → 7993. This assignment is fully dependent on the classification during the detection process. The total tracking time assigned to each of these pigs will be the total time when pigs were classified as A, B or C.

*Pigs classified as White class.* For this class, the animal ID is assigned and recovered every time the pig entered the feeding station, where the RFID antenna was located. The feeding station was situated in a specific location within the pen, allowing us to verify the presence of a particular pig by cross-referencing the coordinates of its predicted bounding box with those of the feeder. Only one pig could enter the feeding station at a time. The database clock of the electronic feeder was synchronized with the clock of the recording system to facilitate the assignment of an animal ID to the pig that was in the feeder. The database was accessed each time a pig was detected in the feeder to perform this assignment. Therefore, for a specific frame X in which a pig was detected at the feeding station, we accessed the database and assigned the corresponding animal ID to this pig detection. As observed in the previous section, as a result of the tracking algorithm, the detection bounding box corresponding to this pig had been assigned a unique track ID. Consequently, we linked the animal ID obtained from the RFID reader to this track ID. Subsequently, for the bounding boxes identified in both the previous and following frames where this track ID appeared, we linked them to the same animal ID. The same animal ID was assigned to the whole trajectory of the animal. That means that the animal ID was recovered both backward and forward, relative to the time it entered the feed station and had its RFID read. The total number of trajectories or track IDs assigned to each white pig depended on how many times it visited the feeder. Pigs with some track IDs never entered the electronic feeder during the tracking process and remained unlinked to an animal ID. In Figure 3, a representation of this process is provided. For simplicity, this process is illustrated using only one of the pigs (ID 7909), for which two trajectories could be assigned (track IDs 2 and 4).

The total tracking time assigned to white pigs was the sum of the time of all the trajectories assigned to the same pig. It is important to note that when the detected objects are homogeneous, occasionally, one object can take over the track ID of another one. This swap can be bidirectional or unidirectional. It can happen after an interaction between objects or in crowded situations and it affects the total time assigned to the pig.

*Pigs classified as Pair class.* For the assignment of the animal ID for the two pigs classified as Pair, information from the classification during detection was used. A pig classified as Pair either has animal ID 7992 or animal ID 8003. To distinguish and assign the correct animal ID to each Pair pig, information from RFID system was used (similar to the pigs classified as White). Therefore if an animal ID was assigned for one of the Pair pigs, the individual identity of the other Pair pig was known. The total tracking time assigned to each Pair pig was the total time the pigs were classified as Pair and they could be distinguished from each other.

## **Evaluation of the animal ID assignment**

To evaluate the animal ID assignment, the ground truth trajectories and classifications of the detected pigs obtained from the tracking-by-detection process were manually annotated for all 12 pigs across all the frames (100,018). Hence, the real total tracking time and the number of real trajectories for each

individual were computed. These metrics were compared with the number of trajectories and the total tracking assigned to each pig calculated in the previous step.

## Computer hardware setup

The computer hardware configuration used to perform the analysis was a Intel® Xeon® Gold 5218R CPU @ 2.10GHz. It was equipped with two graphic cards NVIDIA RTX 6000/8000 and with a memory size of 192 GB. The CUDA version was 12.0. All the codes were run with the programming language Python v3.10.11.

## Results

### YOLOv8 finetuned

The detection accuracy of the finetuned YOLOv8 model on the test set was 99% mAP@0.5, with an AP@0.5 of 99.5% for class A, 99.1 % for class B, 98.8% for class C, 98.9% for the class Pair and 98.7% for the class White. The confusion matrix shows that almost 100% of the classes were classified correctly with this model. Some of the detections were not made on the objects but occurred in the background (FP). Most of these detections were classified as White pigs, with 66% of the detections in the background corresponding to this class (Figure 4). This prediction performance indicates that the model successfully detected and classified the five different classes of pigs described in this study. Thus, this model was applied in the tracking-by-detection process.

### Tracking-by-detection

Figure 5.A shows the distribution of the number of pigs detected per frame after the tracking-by-detection process. Twelve is the expected number of pigs in all frames, and the plot indicates that in the majority of the frames (72%), the expected number of pigs was detected. All the frames with less than 12 pigs contained FN detections while all the frames with more than 12 pigs contained FP detections. In this case, 1.6% of the total detections were FN and 0.8% were FP. This resulted in deletion of 9,909 detections out of 1,190,904 total. Most of the FP detections were classified as pig C or White class. The confidence score distribution of the bounding boxes after removing the FP is displayed in Figure 5.B, where the larger part of the predicted bounding boxes had a confidence score greater than 0.8.

Before assigning an animal ID to each trajectory, the trajectories that included the pig entering the feeding station must be identified. Figure 6 shows all the track IDs of these kind of trajectories for all the pigs. For each trajectory, the number of frames that the pig spent inside and outside the feeder is depicted. More frames outside the feeder indicate that that pig could be tracked for a longer time before/after entering the feeder. For track ID 2143, for example, only 7% of the time of that trajectory the pig was outside the feeder, being the lowest time. This indicates that this trajectory disappeared almost immediately after the pig left the feeding station. However, 75% of the trajectories contained more than 50% of the time outside the feeder. The largest trajectories outside the feeder were for track IDs 582, 719, 836, 1830, 2452 and 2894, which persisted outside the feeder for more than 95% of the time, allowing for longer tracking of these animals. The total length of the trajectories ranged from 620 frames (1.03 min) for the trajectory with track ID 1202 to 45,095 frames (70.15 min) for the trajectory with track ID 719. The mean length (SE) for all these trajectories was 21.55 min (3.35 min)

For each of these trajectories, a unique animal ID was assigned to it, except for one trajectory for which two animal IDs were assigned (track ID: 1202). That means that one pig took over the track ID of another pig, and that the pig also visited the feeding station. In these cases, the animal ID associated with this trajectory depends on the moment in the video, assigning the animal ID recorded the last time the pig with that trajectory was at the feeder.

### **Assignment of the animal ID**

The ground truth trajectories, tracking time and the mean length of the trajectories for all pigs are presented in Table 1. The duration of the video where the tracking-by-detection method was applied was 166.7 min. In the ground truth, the tracking time was not exactly that duration because detection failed in some frames. The minimum tracking time in the ground truth was 157.1 min for the pig classified as C, meaning that in 5% of the video, this pig was not detected. It can be observed that pigs with natural marks presented a higher number of distinct trajectories compared to White pigs, with averages of 104.7 and 26.9, respectively. Similarly, the trajectories were longer for White pigs with an average of 6.95 min compared to marked animals (1.67 min). The trajectories for the pigs corresponding to the Pair class had intermediate time lengths between the distinguished natural marked (pigs A, B and C) and the White pigs, with a mean number of trajectories of 39.5 and a mean length of 4.17 min.

After the process of assigning the corresponding animal ID to each detection, 55.96% of the pig detections had an animal ID associated with it. The animal ID assignment for the pigs belonging to the classes A, B, C or Pair depended on the quality of the detections and classifications. In Table 2, the animal ID, the detection time assigned to each individual, and the percentage of that time that matched the ground truth are presented for all the naturally marked pigs (A, B, C and Pair class). For all pigs, more than 94% of the assigned tracking time corresponded to those individuals in the ground truth. The detection time assigned for pigs in classes A, B and C was slightly but similarly lower than the corresponding tracking time in the ground truth (161.0 min compared to 165.3 min for pig 7909, 148.3 min compared to 158.6 min for pig 7940, and 146.91 compared to 157.1 min for pig 7993). That means that the classifications only failed in a few instances. For pigs in the Pair class (pigs 7992 and 8003), the assigned detection time was much lower than the corresponding time in the ground truth because there was not always enough information available to distinguish them in the process of animal ID assignment. However, the 98% and 94% of tracking time assigned to these pigs correctly corresponded to them.

The assignment of the animal ID to each White pig depended on the tracking and the information provided by the RFID system. In Table 3, the animal ID of each White pig, the number of trajectories assigned, the percentage of trajectories that matched with the ground truth, the total tracking time assigned, and the percentage of that time that matched the ground truth are presented. The number of trajectories assigned to each one depended on the number of times the pig visited the feeder. For all pigs, 100% of the trajectories matched with the ground truth. This indicates that every time the pig visited the feeder, its RFID tag was correctly read. The total tracking time assigned to each pig was the sum of the time for its trajectories. After an interaction or occlusion, the track ID of the trajectory of a pig can be taken over by another. In that case, the total tracking time assigned to that pig would be overestimated. For two out of the seven White pigs, 100% of the assigned tracking time matched with the ground truth. For animal IDs 7909, 7922 and 7957, the 99%, 99% and 92% of the tracking time assigned respectively matched with the ground truth. Only for one pig (animal ID: 8047), 18% of the

total tracking time assigned matched with the ground truth. In this case, one of its two-track IDs was taken by two natural marked pigs (animal ID: 7940 and 7993) where the classification failed, and they were classified as White pigs. For the first one, the trajectory with this track ID lasted only 2.3 min, but for the second one, this trajectory was 22.3 min long, which caused an overestimation of the assigned tracking time for the white pig with the animal ID 8047.

The most trackable pigs were those that exhibited more trajectories (i.e., visited the feeder more often), these being the pigs with IDs 7957 and 8020. The total tracking time assigned to those pigs was 96.4 and 119.6 min respectively out of the 166-min video duration. The least trackable pig was ID 7030, with only one trajectory of 9.1 min. The mean of the total tracking time assigned to the white pigs was 59.1 min.

## Discussion

The objective of the present study was to present a strategy to track homogeneous animals housed in groups by combining computer vision algorithms with an RFID system. The proposed approach consists of tracking objects (i.e. animals) using a tracking-by-detection methodology, and assigning each trajectory to an individual animal with the RFID system. Although all data analysis in the current study was performed after data collection (or video recording), the system can also be implemented in real-time. The methodology was applied to a pen with 12 pigs, in which three animals had natural coat marks that allowed for easy visual distinction to identify them as different classes. In addition, two other pigs were distinguishable from others but similar within the pair. Lastly, the remaining seven pigs were all white and indistinguishable within their group. This pen configuration enabled us to compare the tracked time between naturally marked and homogeneous animals comprising the seven pigs which were completely white. The test images were recorded in the period from 8:37 h to 11:24 h (166 min) when animals are usually active and spend time moving, and interacting with each other, as pointed out by Xu et al. (2023) and Huang et al. (2024). In addition, the animals were in groups of twelve (0.89 pig/m<sup>2</sup>) which can be considered a common density under commercial conditions and were 115 days old (the second half of the fattening period) which means that the space available for each of them at the start of this stage has reduced considerably. All of the above conditions make individual identification and tracking difficult but could be considered representative of commercial conditions.

In tracking-by-detection, pig detection accuracy is important, given that the tracking algorithm's effectiveness depends on the quality of these detections. In the present study, YOLOv8 was fine-tuned to detect pigs with specific marks and to detect white pigs. This model was trained to detect specific marks on the pigs in the pen under study; however, it is trained to detect the "pig" class, so it could be used in other pens to detect pigs irrespective of whether they have different colors in their coat or not. In that case, all individuals in the pen would be detected and classified into the same class, the "pig" class. In the present study, the finetuned YOLOv8 model presented a mAP of 99%. Nonetheless, these metrics cannot be directly compared to those of other studies because of the significant variability in recording conditions. These conditions involve various factors including group size, density, lighting and background. Various studies have used different versions of YOLO to detect animals housed in groups. For example, van der Zande et al. (2021) used YOLOv3 to detect pigs (1.07 pig/m<sup>2</sup>) and reported a mAP of 99% in all the tested datasets and environments. Qiao et al. (2023) used YOLOv5 to detect cow and beef and reported a mAP of 94.7% in the test set (300 images). (Chen et al., 2023) fine-tuned YOLOv7 to

detect broilers with 1,000 images and reported a mAP of 98.2%. These studies demonstrate the adaptability of YOLO models for animal detection in different environments.

In the present study, we took advantage of the presence of pigs with different coat colors to compare the tracking results of distinguishable, colored pigs and homogeneous, white pigs. In a pen with another type of natural or artificial marks, the model would need to be re-trained. Similarly, Yang et al. (2018) marked the pigs artificially and trained a custom detection model to recognize the different letters painted on the four pigs in the pen. Thus, the model was trained to detect the letters A, B, C, and D. A more efficient approach for distinguishing between different pigs without training a detection model depending on the painted marks is presented by Kashiha et al. (2013). In that study, individual pigs were identified by their respective paint patterns using pattern recognition techniques. All these studies would mean having to repaint the pigs frequently to ensure good detection and identification. Another approach was proposed by Psota et al. (2020), who equipped the pigs with different distinguishable tags and used a classification network to assign a unique identification to each pig.

When the pigs have marks that distinguish them from their pen mates, they can be detected as separate objects, and their tracking time will depend on the accuracy of the detection model. In the present study, the AP@0.5 of the pigs classified as A, B and C were 99.5%, 99.1 % and 98.8%, respectively. When this finetuned model was applied to the long video lasting 166 min, these pigs could be tracked, on average, 88% of the time, and more than 95% of the assigned tracking time matched the ground truth. Thus, the results show the effectiveness of the model to detect these different pigs. Although the tracking time assigned to these pigs depended on detection, their trajectories were also analyzed. In the ground truth, the pigs A, B and C presented more trajectories than the White pigs. This occurred because Bot-SORT failed to establish the connection of these pigs across consecutive frames, resulting in the creation of new trajectories whenever it failed. Such failure could be attributed to the creation of shadows depending on the lighting or background, leading to the algorithm failure. Consequently, the duration of these trajectories is shorter than in the case of the White pigs.

With homogenous animals or when all the individuals of the same pen are detected as the same object, the number of trajectories is very important. Ideally, a single trajectory should be assigned to each individual. This would mean that the tracking algorithm does not fail in establishing the connection of the detected objects across consecutive frames. However, this remains a challenge for tracking algorithms. In the present study, we applied the tracker BoT-SORT which has been evaluated in two MOT datasets achieving good results using the main MOT metrics. To the best of our knowledge, this tracking algorithm has not been used before to track pigs. The progress made in the detection and tracking of pigs has been reviewed by (Wang et al., 2022). Another baseline tracker like SORT has been employed to track pigs housed in groups. For example, van der Zande et al. (2021) tracked white pigs housed in groups of 6 (1.07 pig/m<sup>2</sup>) for an average of 24.8 min applying a tracking-by-detection method (SORT + YOLOv3). The less available space a pig has to move, the easier it is for an occlusion to occur and for the pig's identity to be lost, making it more likely to create trajectories for the same individual. In our study, the pen density was slightly lower (0.89 pig/m<sup>2</sup>) than in van der Zande et al. (2021). However, our results are not entirely comparable with the results of that study due to differences in the execution of the experiment, including variations in lighting conditions, image quality, and tracking hours. The choice of the time of the day when the tracking-by-detection method is applied is very important because the accuracy of the tracking depends on the state of the animals. During moments of the day when the animals are mostly resting, it will be easier to track them without losing their identity. In the present

study, we used a video in the morning where the pigs were very active (from 8:37 to 11:24). In this scenario, our framework of combining the tracking-by-detection method with the information provided by RFID system was applied. Results showed that up to three different trajectories could be assigned to the same individual pig. If the identity of the pigs were not recovered every time they enter the feeder, the tracking time assigned to each individual would be that of its first trajectory. Once the algorithm loses the individual and cannot assign the same track ID of the trajectory to the same pig between two consecutive frames, the identity of the individual is lost. With our framework, the tracking time assigned to each pig corresponds to the sum of all their trajectories that included entering the feeder. Hence, the average tracking time assigned to the white pigs was 59.1 min. Once the tracking time is assigned to each individual, it is important to check whether that time truly corresponds to that individual or if it is overestimated, as the track ID of one individual's trajectory can be taken by another pig after an occlusion or animal interaction. Thus, the assigned tracking time was compared with the real tracking time in the ground truth. Only for one pig (animal ID: 8047), the assigned tracking time was overestimated. For the remaining white pigs, more than 92% of the assigned tracking time matched with the ground truth. This indicates that the proposed framework presents a good performance. Identifying when there is a loss of identity or a swap of track ID is complicated since there is no visual pattern. Doornweerd et al. (2024) applied a tracking-by-detection method to analyze broiler tracking. In that study, the kinematics were examined to determine whether the velocity and acceleration at which the animals moved caused more identity losses or swaps, but it was concluded that the kinematics did not influence this phenomenon.

Another approach to tracking animals, different from using tracking-by-detection algorithms, was proposed by Parmiggiani et al. (2023), with the application of graph convolutional networks. This method outperformed DeepSORT when tracking 16 pigs housed in a group, achieving an average tracking time of 12 min compared to 4 min with DeepSORT. However, there is no recovery system to regain the animal ID when it is lost. Chen et al. (2022) presented a semi-supervised pipeline to positions of multiple pigs with a top view RGB camera where the average tracking error was 17.99 cm.

The RFID receiver was located inside the feeding station, where only one pig could enter at a time, allowing the detection of electronic ear tags of one single animal without confusion with another. The limitation of the system was that we had only one RFID reader available, recovering animal IDs only when pigs were feeding. Nevertheless, if more receivers were installed in different locations inside the pen, the identity of the pigs could be recovered more frequently, which would imply more time for individual tracking. Consequently, the longer an animal keeps its identification, the longer we will be able to take measurements of its activity. In some studies, to simplify the identification process, global activity in the pen was measured (Ott et al., 2013; Jiang et al., 2020). However, individual activity is preferred because this could be considered a new phenotype that could be taken into account in breeding programs for improving feed efficiency or resilience to diseases among other economically important traits if a substantial genetic correlation exists. In other studies, animals were housed individually, so they could be monitored without the need for tracking. This would be an example for lactating sows, where only the sow and piglets are present in any image, and the sow is distinguishable, allowing monitoring of her activity (Yang et al., 2020).

With the framework developed in this study, the identity of the animal and consequently its activity is not recovered until the pig enters the feeder and the electronic ear tag is read by the receiver. At that time, the animal ID is assigned to the detection of that pig for both past and future track instances (i.e.,



to the entire trajectory). In addition, the report of the individual activity could be performed daily since the electronic ear tags of all pigs can be read by the receiver many times. In that way, activities such as feeding, drinking, resting, and even interactions between pigs could be monitored daily leading to improvements in animal welfare, feed management, and breeding strategies, with overall economic benefits.

## Conclusion

The finetuned YOLOv8 model trained in this study enabled the detection and classification of five classes of pigs defined in this work. Thus, the distinguishable pigs could be tracked for almost all of the follow-up period. For the homogeneous individuals (white pigs), the combination of the tracking-by-detection method (BoTSORT + finetuned YOLOv8) with RFID system information extended their tracking time. This was possible because animal ID was recovered every time they entered the feeder, enabling association with their entire trajectory, both backward and forward. There was an almost perfect correspondence between tracking time with the ground truth for each pig. However, additional efforts are necessary to further extend tracking time in homogeneous livestock populations possibly combining information from several devices.

Accepted Manuscript

## **Acknowledgments**

We are grateful for the assistance in data collection for this study provided by Neus Torrentó, and to Juan Pablo Sánchez for facilitating the use of the cameras.

## **Conflict of interest statement**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## **Author Contributions**

Individual author contributions to this research article are as follows: Conceptualization and design: MM, MP, GJMR, methodology: MM, formal analysis: MM, writing the original draft: MM, writing-review editing: MP, ID, GJMR, supervision: MP, GJMR and funding acquisition: MP, GJMR.

Accepted Manuscript

## References

- Aharon, N., Orfaig, R., and Bobrovsky, B. 2022. BoT-SORT: Robust associations multi-pedestrian tracking. arXiv:2206.14651. <https://doi.org/10.48550/arXiv.2206.14651>
- Anzai, H., and M. Hirata. 2021. Individual monitoring of behavior to enhance productivity and welfare of animals in small-scale intensive cattle grazing systems. *Frontiers in Sustainable Food Systems* 5:694413. <https://doi.org/10.3389/fsufs.2021.694413>
- Bailey, D. W., M. G. Trotter, C. W. Knight, and M. G. Thomas. 2018. Use of GPS tracking collars and accelerometers for rangeland livestock production research. *Translational Animal Science* 2:81-88. <https://doi.org/10.1093/tas/txx006>
- Berckmans, D., 2017. General introduction to precision livestock farming. *Animal Frontiers* 7:6-11. <https://doi.org/doi:10.2527/af.2017.0102>
- Chapa, J. M., K. Maschat, M. Iwersen, J. Baumgartner, and M. Drillich. 2020. Accelerometer systems as tools for health and welfare assessment in cattle and pigs - A review. *Behavioral Processes* 181:104262. <https://doi.org/10.1016/j.beproc.2020.104262>.
- Chen, B.-L., T.-H. Cheng, Y.-C. Huang, Y.-L. Hsieh, H.-C. Hsu, C.-Y. Lu, M.-H. Huang, S.-Y. Nien, and Y.-F. Kuo. 2023. Developing an automatic warning system for anomalous chicken dispersion and movement using deep learning and machine learning. *Poultry Science* 102:103040. doi: <https://doi.org/10.1016/j.psj.2023.103040>
- Chen, C.-P. J., G. Morota, K. Lee, Z. Zhang, and H. Cheng. 2022. VTag: a semi-supervised pipeline for tracking pig activity with a single top-view camera. *Journal of Animal Science* 100(6): skac147. <https://doi.org/10.1093/jas/skac147>
- Dendorfer, P., H. Rezatofighi, A. Milan, J. Shi, D. Cremers, I. Reid, S. Roth, and L. Leal-Taixé. 2020. MOT20: A benchmark for multi object tracking in crowded scenes. arXiv:2003.09003. <https://doi.org/10.48550/arXiv.2003.09003>

- Doornweerd, J. E., R. F. Veerkamp, B. de Klerk, M. van der Sluis, A. C. Bouwman, E. D. Ellen, and G. Kootstra. 2024. Tracking individual broilers on video in terms of time and distance. *Poultry Science* 103(1):103185. <https://doi.org/10.1016/j.psj.2023.103185>
- Fernandes, A. F. A., Dórea, J. R. R., and Rosa, G. J. M. 2020. Image analysis and computer vision applications in animal sciences: an overview. *Front. Vet. Sci.* 7:551269. doi: 10.3389/fvets.2020.551269.
- Guo, Q., Y. Sun, C. Orsini, J. E. Bolhuis, J. de Vlieg, P. Bijma, and P. H. N. de With. 2023. Enhanced camera-based individual pig detection and tracking for smart pig farms. *Computers and Electronics in Agriculture* 211:108009. doi: <https://doi.org/10.1016/j.compag.2023.108009>
- Hansen, M. F., M. L. Smith, L. N. Smith, M. G. Salter, E. M. Baxter, M. Farish, and B. Grieve., 2018. Towards on-farm pig face recognition using convolutional neural networks. *Computers in Industry* 98:145-152. <https://doi.org/10.1016/j.compind.2018.02.016>
- Huang, Y., D. Xiao, J. Liu, Y. Liu, Z. Tan, X. Hui, and S. Huang. 2024. Analysis of pig activity level and body temperature variation based on ear tag data. *Computers and Electronics in Agriculture* 219:108768. doi: <https://doi.org/10.1016/j.compag.2024.108768>
- Jiang, M., Y. Rao, J. Zhang, and Y. Shen. 2020. Automatic behavior recognition of group-housed goats using deep learning. *Computers and Electronics in Agriculture* 177:105706. <https://doi.org/10.1016/j.compag.2020.105706>
- Kapun, A., F. Adrion, and E. Gallmann. 2020. Case study on recording pigs' daily activity patterns with a UHF-RFID system. *Agriculture* 10:542. <https://doi.org/10.3390/agriculture10110542>

- Kashiha, M., C. Bahr, S. Ott, C. P. H. Moons, T. A. Niewold, F. O. Ödberg, and D. Berckmans. 2013. Automatic identification of marked pigs in a pen using image pattern recognition. *Computers and Electronics in Agriculture* 93:111-120. <https://doi.org/10.1016/j.compag.2013.01.013>
- Lee, S., W. Lee, and J. Park. 2023. A study on tracking moving objects: pig counting with YOLOv5 and StrongSORT. In: 29th International Workshop on Frontiers of Computer Vision, Yeosu - South Korea. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://iwfcv2023.github.io/assets/Poster/P1-10%20A%20Study%20on%20Tracking%20Moving%20Objects%20Pig%20counting%20with%20YOLOv5%20and%20StrongSORT\_Seungwan%20Lee.pdf
- Lin, T.-Y., M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. 2014. Microsoft coco: Common objects in context. In: *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V* 13. p 740-755.
- Milan, A., L. Leal-Taixé, I. Reid, and S. Roth. 2016. MOT16: A benchmark for multi-object tracking. arXiv:1603.00831. <https://doi.org/10.48550/arXiv.1603.00831>
- Ott, S., C. Moons, M. Kashiha, C. Bahr, F. Tuytens, D. Berckmans, and T. Niewold. 2013. Automated video analysis of pig activity at pen level highly correlates to human observations of behavioural activities. *Livestock Science* 160:132-137. <https://doi.org/10.1016/j.livsci.2013.12.011>
- Parmiggiani, A., D. Liu, E. Psota, R. Fitzgerald, and T. Norton. 2023. Don't get lost in the crowd: Graph convolutional network for online animal tracking in dense groups. *Computers and Electronics in Agriculture* 212:108038. <https://doi.org/10.1016/j.compag.2023.108038>

- Psota, E., T. Schmidt, B. Mote, and L. C. Pérez. 2020. Long-term tracking of group-housed livestock using keypoint detection and MAP estimation for individual animal identification. *Sensors* 20:13. <https://doi.org/10.3390/s20133670>
- Qiao, Y., Y. Guo, and D. He. 2023. Cattle body detection based on YOLOv5-ASFF for precision livestock farming. *Computers and Electronics in Agriculture* 204:107579. <https://doi.org/10.1016/j.compag.2022.107579>
- Riaboff, L., S. Aubin, N. Bédère, S. Couvreur, A. Madouasse, E. Goumand, A. Chauvin, and G. Plantier. 2019. Evaluation of pre-processing methods for the prediction of cattle behaviour from accelerometer data. *Computers and Electronics in Agriculture* 165:104961. <https://doi.org/10.1016/j.compag.2019.104961>
- Riaboff, L., S. Couvreur, A. Madouasse, M. Roig-Pons, S. Aubin, P. Massabie, A. Chauvin, N. Bedere, and G. Plantier. 2020. Use of predicted behavior from accelerometer data combined with GPS data to explore the relationship between dairy cow behavior and pasture characteristics. *Sensors* 20:4741. <https://doi.org/10.3390/s20174741>
- Ringgenberg, N., R. Bergeron, and N. Devillers. 2010. Validation of accelerometers to automatically record sow postures and stepping behaviour. *Applied Animal Behaviour Science* 128:37-44. <https://doi.org/10.1016/j.applanim.2010.09.018>
- Rosa, G. J. M. 2021. Grand challenge in precision livestock farming. *Frontiers Animal Science* 2:650324. <https://doi.org/10.3389/fanim.2021.650324>
- Shirke, A., A. Saifuddin, A. Luthra, J. Li, T. Williams, X. Hu, A. Kotnana, O. Kocabalkanli, N. Ahuja, A. Green-Miller, I. Condotta, R. N. Dilger, and M. Caesar. 2021. Tracking grow-finish pigs across large pens using multiple cameras. *arXiv:2111.10971* <https://arxiv.org/abs/2111.10971>

- Turner, K. E., A. Thompson, I. Harris, M. Ferguson, and F. Sohel. 2022. Deep learning based classification of sheep behaviour from accelerometer data with imbalance. *Information Processing in Agriculture* 10:377-390. <https://doi.org/10.1016/j.inpa.2022.04.001>
- Tzutalin, D. 2015. LabelImg Git code. <https://github.com/HumanSignal/labelImg>
- van der Zande, L., O. Guzhva, and B. Rodenburg. 2021. Individual detection and tracking of group housed pigs in their home pen using computer vision. *Frontiers in Animal Science* 2:669312. doi: 10.3389/fanim.2021.669312
- Wang, S., H. Jiang, Y. Qiao, S. Jiang, H. Lin, and Q. Sun. 2022. The research progress of vision-based artificial intelligence in smart pig farming. *Sensors* 22(17):6541. <https://doi.org/10.3390/s22176541>
- Williams, L. R., D. R. Fox, G. J. Bishop-Hurley, and D. L. Swain. 2019. Use of radio frequency identification (RFID) technology to record grazing beef cattle water point use. *Computers and Electronics in Agriculture* 156:193-202. <https://doi.org/10.1016/j.compag.2018.11.025>
- Xu, J., J. Ye, S. Zhou, and A. Xu. 2023. Automatic quantification and assessment of grouped pig movement using the XGBoost and YOLOv5s models. *Biosystems Engineering* 230:145-158. doi: <https://doi.org/10.1016/j.biosystemseng.2023.04.010>
- Yang, A., H. Huang, B. Zheng, S. Li, H. Gan, C. Chen, X. Yang, and Y. Xue. 2020. An automatic recognition framework for sow daily behaviours based on motion and image analyses. *Biosystems Engineering* 192:56-71. <https://doi.org/10.1016/j.biosystemseng.2020.01.016>
- Yang, Q., D. Xiao, and S. Lin. 2018. Feeding behavior recognition for group-housed pigs with the faster R-CNN. *Computers and Electronics in Agriculture* 155:453-460. <https://doi.org/10.1016/j.compag.2018.11.002>

## Figures

Figure 1. Pen under study with the five different classes (A, B, C, Pair and White) of pigs included in the study.

Figure 2. Representation of the tracking-by-detection method with two different objects (blue and orange) in each frame.

Figure 3. Illustration of the animal ID assignment process for one pig classified as white.  $x$ : center-x,  $y$ : center-y,  $w$ : width, and  $h$ : height.

Figure 4. Confusion matrix for YOLOv8 fine-tuned model evaluation.

Figure 5. Distribution of the number of pigs detected per frame and the confidence score distribution after removing false positives (FP).

Figure 6. Trajectories for all the pigs that at some point enter the feeding station.



Table 1. Trajectories of the ground truth for each pig.

Pig class	Tracking time (min)	Number of trajectories	Mean length (SE) (min)
A	165.28	67	2.46 (0.83)
B	158.61	124	1.28 (0.33)
C	157.08	123	1.28 (0.36)
Pair	164.69	44	3.74 (1.16)
Pair	161.22	35	4.60 (1.31)
White	161.64	26	6.21 (1.66)
White	164.70	17	9.69 (4.06)
White	164.51	25	6.58 (2.41)
White	160.83	46	3.50 (0.94)
White	164.63	19	8.66 (2.99)
White	164.75	17	9.69 (4.33)
White	166.33	38	4.38 (0.88)

Accepted Manuscript

Table 2. Evaluation of the animal ID assignment for all the naturally marked pigs (A,B, C and Pair).

<b>Pig class</b>	<b>Animal ID</b>	<b>Detection time assigned</b>	<b>Percentage of real detection time</b>
A	7926	161.0	98%
B	7940	148.3	95%
C	7993	146.9	97%
Pair	7992	46.8	98%
Pair	8003	50.3	94%

Accepted Manuscript

Table 3. Evaluation of the animal ID assignment for each White pig.

<b>Animal ID</b>	<b>Number of trajectories assigned</b>	<b>Percentage of trajectories correctly assigned</b>	<b>Total tracking time assigned</b>	<b>Percentage of real tracking time</b>
7909	2	100%	34.59	99%
7911	2	100%	87.95	100%
7922	1	100%	36.29	99%
7930	1	100%	9.11	100%
7957	3	100%	96.38	92%
8020	3	100%	119.60	99%
8047	2	100%	29.98	18%

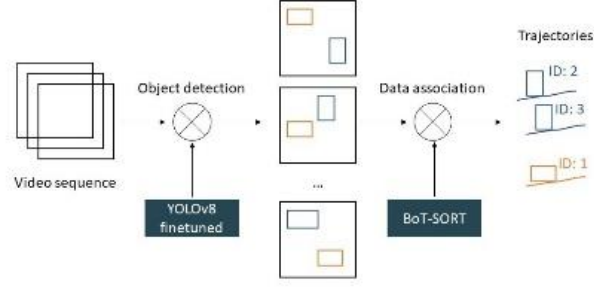
Accepted Manuscript

Figure 1



Accepted Manuscript

Figure 2



Accepted Manuscript

Figure 3

Frame	Class	x	y	w	h	Track ID	Feeding	Animal ID assigned
1	White	Boundig box coordinates				1	No	Unknown
2	White	Boundig box coordinates				1	No	Unknown
3	White	Boundig box coordinates				2	No	7909
4	White	Boundig box coordinates				2	Yes	7909
5	White	Boundig box coordinates				2	Yes	7909
6	White	Boundig box coordinates				2	No	7909
7	White	Boundig box coordinates				3	No	Unknown
8	White	Boundig box coordinates				4	Yes	7909
...	...	...	...	...	...	...	...	...

Number of different track IDs assigned to the pig 7909: 2

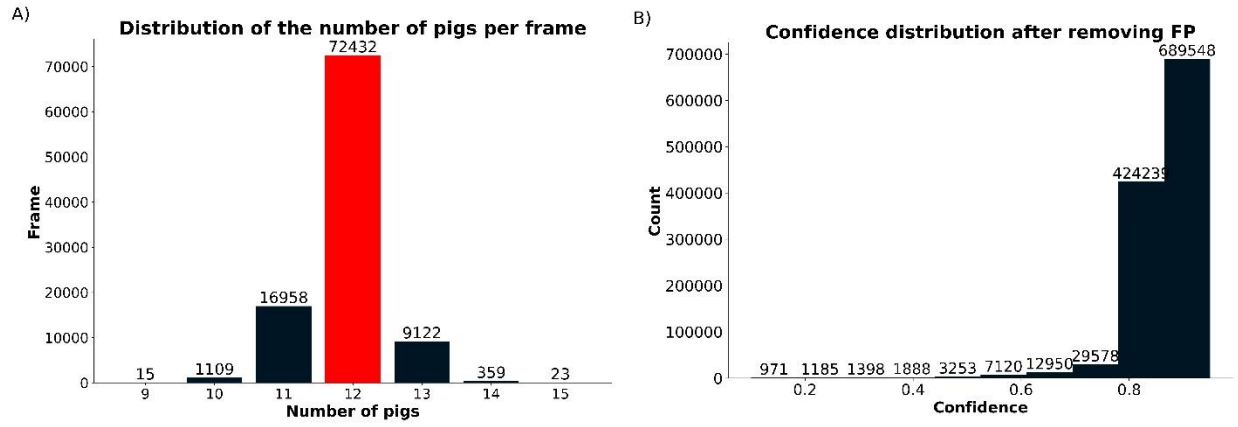
Accepted Manuscript

Figure 4

Predicted	True					
	A	B	C	Pair	White	background
A	0.99					0.07
B		0.98	0.01			0.06
C		0.02	0.98			0.06
Pair				0.99		0.14
White				0.01	0.99	0.66
background			0.01			

Accepted Manuscript

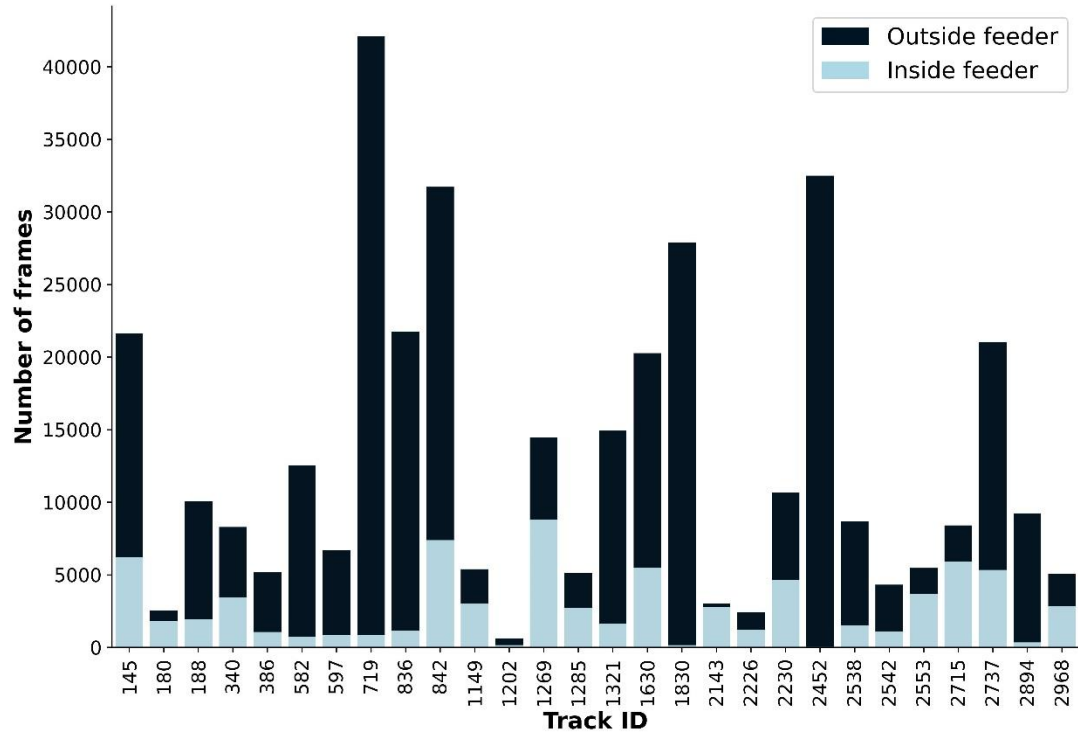
Figure 5



Accepted Manuscript



Figure 6



Accepted