IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 21 April 2023, accepted 2 June 2023, date of publication 12 June 2023, date of current version 20 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3285144

RESEARCH ARTICLE

Video Analysis Using Deep Learning for Automated Quantification of Ear Biting in Pigs

ANICETUS ODO[1], (Member, IEEE), RAMON MUNS[2], LAURA BOYLE[3], AND ILIAS KYRIAZAKIS[1]

[1]Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, BT9 5DL Belfast, U.K.
[2]Sustainable Agri-Food Science Division, Livestock Production Science Branch, Agri-Food and Biosciences Institute, BT26 6DR Hillsborough, U.K.
[3]Animal & Grassland Research and Innovation Centre, Teagasc Pig Development, Moorepark, Fermoy, P61 C996 Ireland

Corresponding author: Anicetus Odo (A.odo@qub.ac.uk)

This research was part of the European Union (EU)-China HealthyLivestock project. The authors wish to acknowledge that HealthyLivestock is funded by the EU H2020 research and innovation program under grant agreement number 773436.

This work involved animals in its research. Ethics approval for the experiments conducted at Teagasc was granted by the Teagasc Animal Ethics Committee under Approval Number TAEC 40/2013. Similarly, experiments at the Agri-Food Bioscience Institute (AFBI) were carried out under the Project Licence Number PPL2851 as approved by the Animal Welfare Ethical Review Body and in accordance with the Animals Scientific Act (1986).

ABSTRACT Ear biting is a welfare challenge in commercial pig farming. Pigs sustain injuries at the bite site paving the way for bacterial infections. Early detection and management of this behavior are important to enhance animal health and welfare, increase productivity, and minimize inputs from medication. Pig management using physical observation is impractical because of the scale of modern pig production systems. The same applies to the manual analysis of videos captured from pigsty. Therefore, a method of automated detection is desirable. In this study, we introduce an automatic detection pipeline based on deep learning for the quantification of ear biting outbreaks. Two state-of-the-art detection networks, YOLOv4 and YOLOv7, were trained to localize the regions of ear biting. The detected regions were tracked over multiple video frames using DeepSORT and Centroid tracking algorithms. Tracking provided the association between detected instances in video frames, enabling the computation of the frequency and duration of occurrence. The frequency and duration of ear biting were expressed as the cumulative performance of each group of pigs. The pipeline was evaluated using two datasets from experimental and commercial farms with diverse management and monitoring settings. The detection networks achieved comparable average precision values of 98% & 97.5% and 85.6% & 80.9% on the respective datasets. The tracking algorithms produced 14% and 34% False-Alarm rates, respectively. The results show that automated detection and tracking of ear biting is possible. Subsequently, we applied our method to videos in which pigs were managed in a manner that was expected to affect the frequency of ear biting to different degrees. This method can be used as the basis of an early warning system for the detection of ear-biting in commercial farms.

INDEX TERMS Animal behavior, animal welfare, deep learning, image analysis, object detection, object tracking.

I. INTRODUCTION

Pigs housed in commercial conditions exhibit damaging behaviors that reflect welfare challenges and cause health and welfare problems in recipient pig [1]. Damaging behaviors include tail, ear, and flank biting [2]. Risk factors include poor health status, limited possibilities to explore the environment, high stocking density, and mixing of previously established groups [3]. Attacks on body parts result in injuries, and when farmers detect them, they may employ antimicrobials (AM), particularly if there is evidence of secondary bacterial infections. Pigs with severe lesions often require systemic AM to prevent the spread of bacterial infections locally or systemically [4]. The increased use of AM increases cost,

The associate editor coordinating the review of this manuscript and approving it for publication was Abdel-Hamid Soliman.

reduces the quality of farm products, and contributes to drug resistance [5], [6].

Tail biting is the most widely studied damaging behavior and a major welfare problem in pig production [3], [7]. Ear biting is generally a less studied damaging behavior than tail biting. However, ear biting is considered a main risk factor for ear necrosis [8], a growing and highly prevalent problem in weaned pigs in some countries [9], and is associated with disease and poor performance [10]. While ear biting generally starts with gentle manipulation, it can escalate into more damaging interactions [11]. The identification of ear manipulation could serve as a precursor to an outbreak. Ear biting may or may not be detected by cursory observation during routine pig management. Most frequently, ear injuries are the most common indication that ear biting occurs within a pen, which may be too late for a remedial action. Clearly, there is a substantial advantage in detecting ear biting early and intervening appropriately. Owing to the enormous scale of modern pig farms, where thousands of pigs can be housed in a single building at any one time, detection through direct visual inspection is impractical or even impossible. This highlights the need to develop a system for the automated (early) detection of ear-biting.

This study focused on the automatic detection of ear biting as a component of the development of an early warning system for managing pigs. Our method uses deep learning to detect any behavior targeting the ear of a pen-mate including sniffing, gentle manipulation, head knocking (the reaction of the recipient pig to ear-biting), and ear-chewing. We do not distinguish between different behaviors involving ear contact but treat all interactions as ear-biting events. Specifically, we describe a pipeline of deep learning-based methods and demonstrate the effectiveness of the models by testing them on two experimental datasets (Section III). These datasets were previously acquired for the inspection of behaviors such as feeding and posture; however, ear biting was observed during the process. They also included a dataset used for the development of an ethogram for ear-biting [11]. The results of detection and tracking are presented in Section IV. Finally, we discuss the results and some of their implications in Section V.

The contributions of this paper are summarized as follows:

1) This study applies computer vision methods for the automatic identification of challenging behaviors in pig production and management. The proposed pipeline adopts a state-of-the-art object detection network to localize the contact regions in images and then aggregates the detection to quantify the behavior. Tracking was used to determine the frequency and duration of occurrence at the group level to enable the identification of affected pens. Pens with a higher occurrence of ear-biting would receive remedial actions.
2) Importantly, we created a new dataset to facilitate the advancement of studies in this area.
3) To the best of our knowledge, this is the first published study on the automatic quantification of ear-biting behavior in pigs.
4) Investigate the impact of management manipulations on frequency and duration of ear biting events.

II. LITERATURE REVIEW

The high prevalence of ear lesions associated with biting is in agreement with [12] in which ear biting and manipulation of body parts were scored more frequently than tail biting. In both [12], [13], pigs were tail docked, which may explain the higher incidence of ear biting, likely because docked tails are not the most attractive or accessible part of the body to bite.

There is little information on the aetiology of ear biting, although it has been suggested that ear biting or chewing and tail biting may be linked. Brunberg et al. [14] observed that pigs housed in control pens exhibited a wider variety of pig-directed abnormal behavior and that not all pigs in a pen were performers of the behavior. Tail-biting pigs performed a higher frequency of ear-biting than non-performers. A survey of Dutch producers suggested that biting both tails and ears was identified by farmers as a welfare problem in pig farming. Similarly, a survey in Ireland identified that of the farms that had experienced tail biting over a period of a year (51 of 58 participating farms), 86% had also experienced ear biting. Of these incidences, farmers experience the greatest amount of ear biting in the second stage post-weaning and the highest incidence of tail biting in the finishing stage [15].

Computer vision tools are effective in detecting different behaviors in pigs [16], [17]. Viazi et al. [16] extracted the mean intensity of motion and occupation index from video frames and processed these features using a Linear Discriminant Analysis method. This method enabled the classification of every behavior episode as aggressive or otherwise. Behaviors with diagnostic relevance, such as standing, were detected using depth information to track pig positions in the videos [17]. Gaussian Mixture Models of 3D point clouds were developed to classify standing and non-standing postures. The average standing time can be used to detect health and welfare challenges. Traditional computer vision techniques are limited by the specific features used for development, and the quality of hand-crafted features depends on the quality of the image sources. Farm settings vary (different camera settings, lighting conditions, animals constantly moving about, and their orientation will change), and these important variables should be considered for a more generalized method.

Artificial intelligence is increasingly gaining popularity as a key component of precision livestock farming especially for learning and managing animal behavior in images [18]. Convolutional neural networks (CNN) provide suitable alternatives to feature engineering. CNNs can learn the diverse features required for solving computer vision tasks directly from data sources (particularly from images). Odo et al., [19] presented the localization of small components in aerial

(a) AFBI dataset showing: (i) perpendicular camera view (ii) low pig density, and (iii) good lighting condition.

(b) Teagasc dataset showing: (i) lateral camera view (ii) high pig density (iii) high occlusion, and (iv) poor illumination.

FIGURE 1. **Images showing characteristics of the datasets.**

images. While the targets in [19] were stationary, the sensor (camera) was placed on a moving aircraft. In image-based pig management, the targets (pigs) move while the camera is fixed. Both scenarios present additional challenges because of the motion component. Several studies have demonstrated the monitoring of behavior in pigs using CNNs [20], [21], [22], [23], [24], [25], [26]. Pipelines often involve identifying the location and orientation of each animal and its body parts. The detected parts are used to monitor interactions, for example, tail-mouth and ear-mouth pairs. Psota et al., [20] applied a fully CNN to detect various parts of a pig, for example, ears, shoulders, and tails, and [21] combined a CNN and a recurrent neural network (RNN) to recognize tail-biting. Considering that tail-biting and ear-biting outbreaks are correlated in terms of their triggers and damaging effects, automating ear-biting detection will be a useful addition to the management of pig systems.

Several well-established deep learning-based object detection models have achieved state-of-the-art performance in popular computer vision applications, such as the common objects in context (COCO) challenge [27]. The choice of method is often determined by the application-specific requirements. Such requirements often include a trade-off between speed and accuracy. For example, while a two-stage detector such as the Faster R-CNN [28] may be more accurate, one-stage networks such as the Single Shot multi-box detector (SSD) [29] and RetinaNet [30] are faster. RetinaNet uses focal loss to address the challenges of foreground and background imbalances in images. Advancements in one-stage detection networks have led to incremental improvements in the detection capabilities. The "You Only Look Once" (YOLO) models are a good example of the development in this area. Specifically, YOLOv4 [31] used feature aggregation including the "bag-of-freebies and bag-of-specials" modules and improved the speed and accuracy of object detection compared to previous versions. Recently, YOLOv7 [32] was introduced and built on the successes of the previous versions. This state-of-the-art object detection method incorporates a "trainable bag of freebies". The focus was on reducing the cost of training without loss of speed at the inference time. The effectiveness of YOLO for the automatic detection and management of pigs has been demonstrated [22], [23], [24], [25].

CNNs have been applied to quantify different pig postures, such as standing and lying postures, and can identify feeding and drinking behaviors [26]. These behaviors are useful for measuring the health and well-being of farm animals. In addition to posture and feeding behavior, the identification of potentially damaging interactions, such as pigs engaging in ear and tail biting may also be important. An approach for the automatic recognition of tail biting was presented using a combination of an SSD and a Long-Short-Term-Memory algorithm [21]. Oczak et al., [33] and Viazzi et al., [16] identified several aggressive interactions in videos. For the specific task of ear biting, [11] developed an ethogram of biter and bitten pig during an ear biting-event. Pigs vocalize in response to bites. Combining different sounds and visual information helped identify the behaviors. Some non-vocal behaviors that describe ear-biting are listed in Table 1. To date, ear-biting has been observed manually in videos. Manual observation is challenging considering the scale of the farms and the number of parameters to be monitored simultaneously. Automating the process of ear biting detection is required for continuous monitoring and management of pens.

Tracking is an important component for monitoring objects in a video. A detection network localizes objects in frames and the tracking system associates instances between consecutive frames. Simple Online and Real-time Tracking (SORT) [34] associates objects using a metric that relies on the overlap of bounding boxes. The limitation of this tracking technique is the high rate of identity switches (IDS). IDS refers to the number of times the reported identity of a ground-truth track changes. DeepSORT [35]

TABLE 1. Description of ear biting behaviors [11].

Behavior	Description
Gentle manipulation	Soft chewing of the pen mate's ear without any response from the pig being bitten.
Quick bite	Short duration bites directed towards pen mate's ears, usually accompanied by a response of the pig being bitten.
Chewing	Prolonged mastication of the pen mate's ear accompanied by a response of the pig being bitten
Ear pulling	Taking the ear of the pen mate into its mouth and exerting force to move it toward itself, accompanied by a response of the pig being bitten.
Biting	Forceful and rapid bite toward the face of the biter pig with or without a vocalization response.
Head knocking	Forceful and rapid or quick vertical action or pushing of the head against the body of the recipient pig with its head going up and down.

was developed with a focus on reducing the issue of identity switching. The contribution of [35] includes the replacement of the association metrics with one that combines motion and appearance information based on a deep appearance descriptor.

III. MATERIALS AND METHODS

A. DATASETS

Figure 1 shows the video frames from the two datasets used in this study. These datasets were previously collected for the study of other pig behaviors, and ear biting was observed during the investigation. Dataset I was derived from a controlled experiment in which the conditions could be adapted to suit the needs of data collection (e.g., lighting and position of the camera). Dataset II was derived from a commercial farm, where little deviation from the commercial routine was allowed (e.g., the farmer dictated the position of the video camera). Variations in management and monitoring settings are considered relevant to the future application of this method.

1) DATASET I

The first dataset comprised videos from an experiment conducted at the Agri-Food and Biosciences Institute (AFBI) in Hillsborough, Northern Ireland, between November 2019 and April 2020. The work was carried out under Project License Number PPL2851 in accordance with the Animals (Scientific Procedures) Act 1986 (The Parliament of the United Kingdom, 1986). The dataset involved weaned piglets at 28 ± 1 days of age with body weight of 9.47 kg ± 1.20 (mean ± SD). At weaning, the pigs were housed across five rooms, where they spent six weeks. There were six pens in each room, comprising ten mixed-sex pigs. Each pen had a plastic slatted floor and a dimension of 2.7m by 1.4m. No intentional health challenge was imposed on the animals. Video cameras (4 M.P. Fixed Bullet Network Cameras, HiLook IPC-B140H(-M), Hikvision, Hangzhou, China) were connected to a network video recorder and installed on the ceiling above the pen. This enabled the monitoring of the entire pen, as shown in Figure 1a. Different regions of the pen appeared to be evenly illuminated, ensuring that the pigs were clearly visible in the image. Each group was provided with environmental enrichment in the form of a suspended wooden block and flavored plastic biting toy (Porcichew, Nutrapet Ltd., U.K.). The pen temperature was initially set at 28°C but decreased 0.5°C/day, stabilizing at 21°C, and artificial lighting was provided daily when routine stock checks were carried out. Pigs also had access to natural light through windows. Artificial light was switched off daily from 0.00 a.m.-2:00 a.m. We selected eight videos at random, that is, three videos from Room 4 and five videos from Room 5 for training. Additionally, we selected two videos, one from each of Rooms 2 and 3, for validation. Each 2h video at 25FPS produced 180,000 frames (i.e., $2 \times 60 \times 60 \times 25$). This amounted to 1,440,000 images of the selected videos. To facilitate data annotation, we sampled three frames per second for each video (i.e., $2 \times 60 \times 60 \times 3 = 21{,}600$ frames). Table 2 shows only key frames, that is, images in which an expert identified ear-biting behaviors.

The dataset was derived from an experiment that tested the effects of different nutritional treatments on pig behavior. Pigs received a starter diet 1 followed by Starter diet 2: Starter 1 diet (S1) contained 16.25 MJ/kg digestible energy (DE), 20% crude protein (CP); 1.65% Lysine (Lys), 2.11% crude fiber (CF) and was offered for 13 days post weaning. Starter 2 (S2) contained 16.25 MJ/kg DE, 20% CP, 1.54% Lys, 2.29% CF and was offered for 16 days. There were three treatments of interest in this study.

- T1: Control - composed of the starter diets (S1, S2) with 20% CP and 2.11% CF, without any antimicrobial (ZnO)
- T2: Both S1 and S2 contained 18% CP and 3-4% non-fermentable fiber.
- T3: Both S1 and S2 contained 16% CP and 3-4% non-fermentable fiber.

The hypothesis was that different treatments would result in different frequencies and durations of ear-biting events. A decrease in the CP of the weaned pig diet may decrease the incidence of digestive disorders, and is currently recommended by the pig industry, but it has been shown to increase the occurrence of damaging behaviors, such as tail biting [36], which may or may not be alleviated by the inclusion of dietary fiber [37]. It has been suggested that the increased risk of damaging behaviors arises from the fact that foraging behavior and redirected biting behavior increase as pen mate blood becomes a more attractive source of nutrients [12]. Hence, the second data sample comprised 30 videos for the quantification of behavior. We selected one pen per treatment from each of the five batches of the AFBI experiment for a total of 15 pens. We analyzed data from two days (days 2 and 4) of the experiment to determine the possible effect of the treatments on ear-biting (Table 6).

Data points were extracted between 11:00 h and 12:00 h of daily observations. This time of day is when the pigs are the most active. In addition, this period was selected to minimize the effect of spontaneous behaviors associated with pig management, for example, when a staff member entered the pen early in the morning and before the end of the day.

The videos had a resolution of 1920×1080 pixels, which was reduced to 1024×580 pixels for easy computation without losing information. Additional pre-processing was required for the dataset to remove adjacent pens within the field of view to ensure that pigs within a pen were analyzed separately. The adjoining pen was masked as shown in Figure 1a.

2) DATASET II

The second dataset was created from a 300 sow farrow-to-finish commercial farm with a history of ear biting behavior, located in Co. Cork, Ireland. The farmers were willing to cooperate with the video data collection required for the study, and the procedure was approved by the Teagasc Animal Ethics Committee (TAEC 40/2013). As this was a purely observational study, pigs were managed according to the usual farming practices. Pigs were weaned at 28 $\pm$ 2 days of age and spent 4 weeks in pens measuring 3 m by 2.4 m, holding about 35 pigs, and subsequently to a larger pen measuring 6 m by 2.9 m, as per commercial practice to accommodate the increase in pig size. Two cameras (Panasonic®, model HC-V250EB-K) were used, one for each pen, placed above the pen in a lateral position at a height of approximately 2 m, so that a full-view of the experimental pen was attained.

As shown in Figure 1b, the lateral view of the pen is such that pigs closer to the camera are more visible in the image than those farther away, and the occlusion is high. Additionally, the videos were collected from a commercial setting and did not include the day or time of observation. We selected twenty-two videos from first-stage weaner pigs for development (19 videos for training and 3 videos for validation). Unlike Dataset I, we extracted and used all the video frames from Dataset II because they were relatively shorter (average length of 22 min). In addition, we analyzed three videos to determine the frequency (i.e., the average number of events in the observation) and duration of the interaction. The resolution of the Teagasc dataset was 1280×720 pixels and was used without reduction because it is more computationally manageable.

3) COMMON METHODOLOGY

Feeders were present in each pen and were provided ad libitum. We collected two samples from each dataset for the development (i.e., training and evaluation of the detection networks) and quantification of ear-biting in the video. Videos from both datasets were recorded at 25 frames per second (FPS). The key frames, that is, video frames where ear biting was visible, were manually annotated using the Visual Geometry Group (VGG) Image Annotator [38]. Descriptions of the contacts that represent ear-biting are listed in Table 1. Each region of ear-biting was defined by a bounding box overlaying the head to the forearm of the interacting pigs. The size of the bounding boxes varies based on the orientation of the biter and the bitten pig. It is important to emphasize that the ground truths were extracted from videos of different pens and different times of the day to ensure a fair representation of the pen conditions and scenarios. To avoid bias, the videos used for testing were different from those used for training our models, and the ground truth was extracted on a double-blind basis. Table 2 summarizes the data points from both datasets for the training and validation of the detection models, which include the number of video frames and ground truth bounding boxes. As there were two datasets with different pen settings and management conditions, we tested the detection networks on the individual datasets and presented a more representative model by combining the datasets.

B. DETECTION

Figure 2 shows a pipeline of the processes involved in the automatic quantification of ear-biting in pigs. The two main components in the pipeline are object detection and tracking systems. The first stage involves the detection of regions in images that exhibit ear-biting behavior. The detected regions of interest are associated over multiple frames. In this study, we trained and validated object detection networks. An ear-biting event is often observed as a sequence of interactions that may include the different behaviors listed in Table 1. We detected interactions, but without discrimination between the biter and bitten pigs, quick or gentle manipulation. All contact with the ear in the mouth was detected as ear-biting.

We utilized the baseline versions of YOLO, namely, YOLOv4-CSP [31] and YOLOv7 [32]. These are one-stage detection models that can perform localization and classification tasks using a single dense layer. They are faster than most state-of-the-art detection networks, and achieves real-time processing without compromising the accuracy. These features are particularly useful for the proposed application, which requires repeated detection over a long period of time. The models were implemented using PyTorch framework. Network parameters vary in depth and width, which determine the resources needed for its usage. YOLOv4-CSP has over sixty-four million parameters and is based on the CSPDarknet53 architecture [39] for feature extraction. On the other hand, YOLOv7 [32] has 36.9 million parameters, representing 43% reduction in computational requirements. These networks were previously trained on the MS-COCO dataset to detect 80 classes of objects. We changed the number of nodes in the output layers of the networks to match a single target, that is, ear-biting. The MS-COCO weights were used for network initialization during the training time. Table 2 lists the number of images and annotations (i.e., bounding boxes of ear-biting regions) used for training and validation. Similar hyper-parameters were used in training both networks (base learning rate of 1 × 10-3 and input size of 640 × 640). Batch sizes

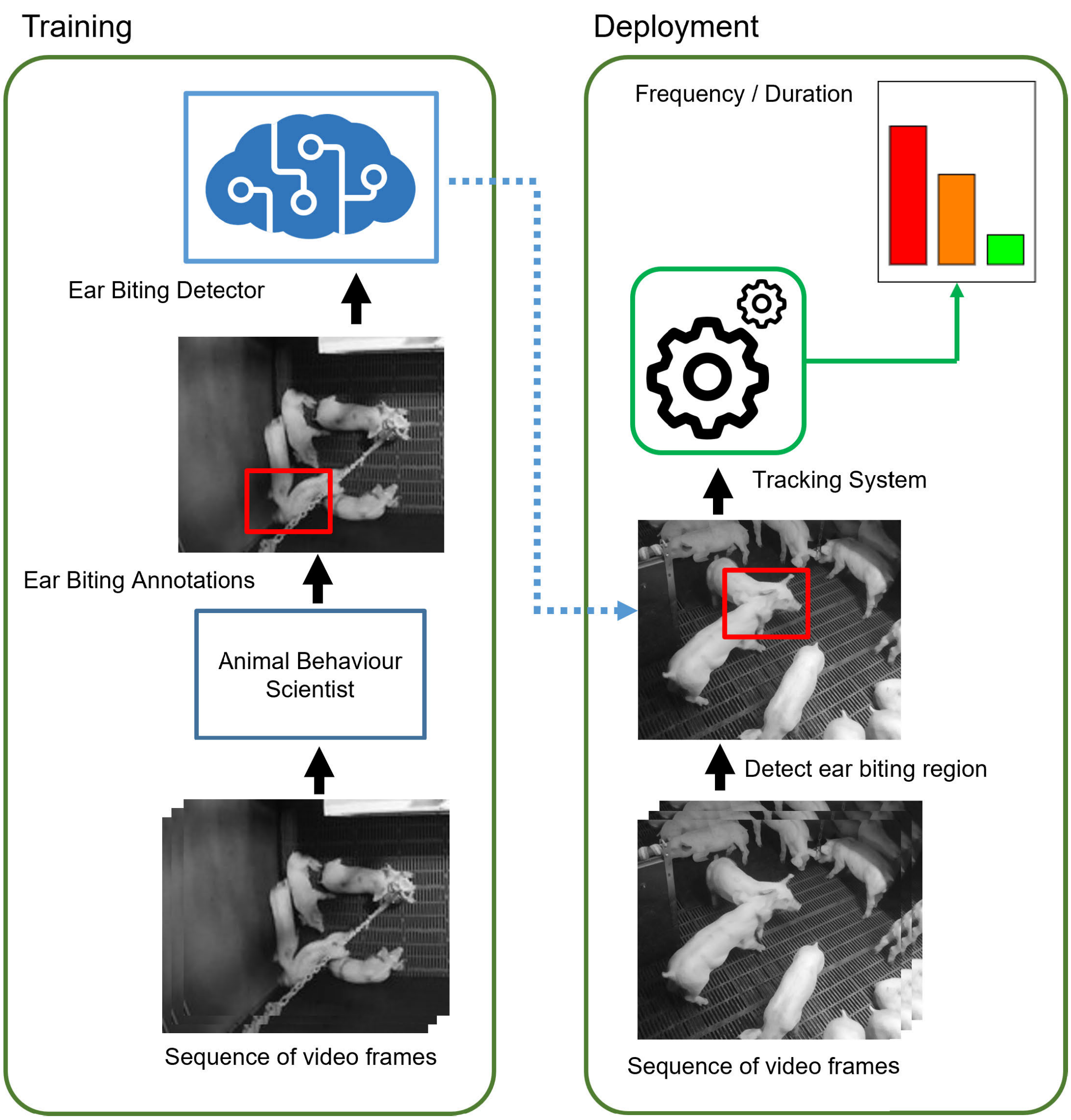


FIGURE 2. **Pipeline for the detection and quantification of ear biting in pigs.**

of 8 and 16 were used to train YOLOv4 and YOLOv7, respectively. The batch size varied because of the size of the networks and the available resources. Additionally, YOLOv7 used mosaic data augmentation to increase the variability of the feature space. This technique combines four random images into a single input image. All networks were trained for a maximum of 100 epochs. The trained weights were saved after each epoch, and the best model was selected for testing. Training and evaluation were performed using a GeForce RTX 2080 Ti GPU with a memory capacity of 10GB.

We evaluated the detection models using a set of output parameters: bounding boxes, that is, the coordinates of the detected region, top-left (x_1, y_1) and bottom-right (x_2, y_2), and the probability score. A higher probability implies better detection confidence. Bounding boxes with a probability of 0.25 and above were used for evaluation. Performance depends on the degree of overlap between the ground-truth bounding boxes and detected bounding boxes. We measured the degree of overlap by determining the intersection over union (IoU) of the corresponding bounding boxes. An IoU ≥ 0.5 was treated as a true positive prediction. Precision was computed as $\frac{TP}{TP+FP}$ and recall as $\frac{TP}{TP+FN}$ where TP, FP, and FN denote the counts of true positives, false positives, and false negatives, respectively. Average precision (AP) is a standard metric for estimating the performance of detection models. AP is the area under the precision-recall curve and is computed using the interpolation method [40].

C. TRACKING

The frequency and duration of episodes are important parameters for measuring the prevalence of ear-biting in

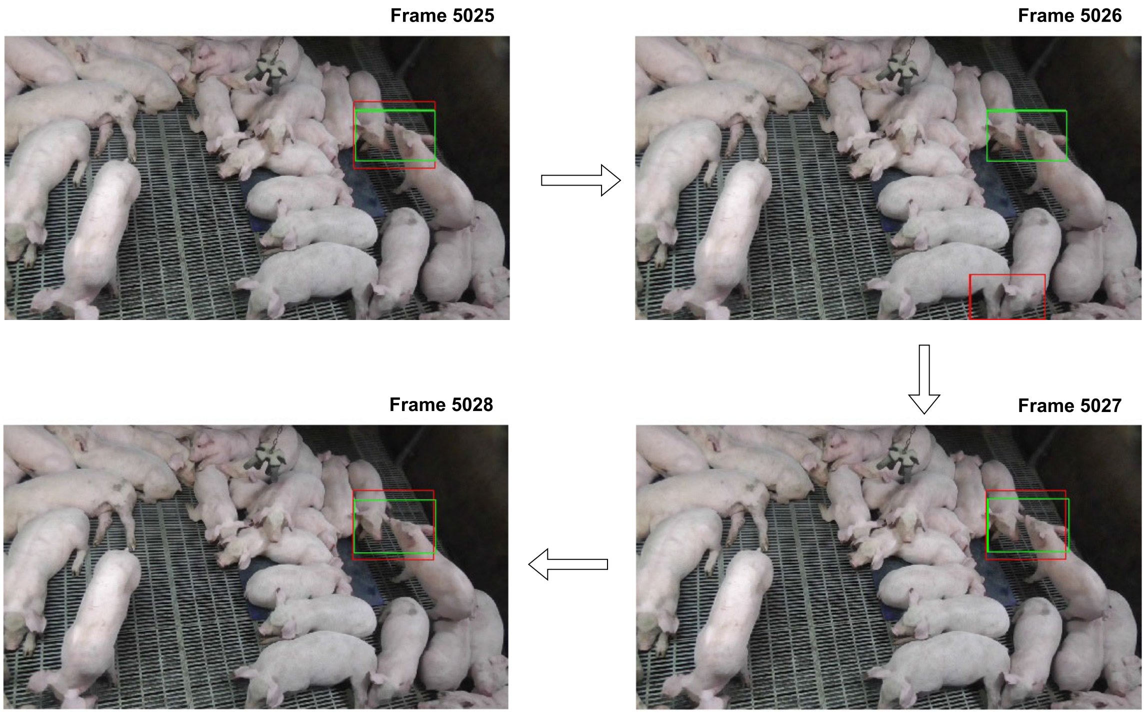


FIGURE 3. **Qualitative results: detection of ear biting event in consecutive video frames with bounding boxes overlaid. The ground truth (GREEN bounding boxes) and detected contact (RED bounding boxes).**

a pen. The interaction(s) between the biter and bitten pig was detected, and the duration of the interaction(s) was measured. Tracking enables the association of detections between multiple frames as members of the same event. An object detection network identifies the region(s) of interest in the current frame, and the tracker attempts to associate these with the objects in the previous frame. Tracking algorithms complement detection by following tracks until the detection fails for at least 200 consecutive frames.

We compared DeepSORT [35], an established multiple object tracker that uses bounding boxes, motion and appearance features, and a simple centroid tracking algorithm. The centroid tracker relies only on bounding box information to associate object instances, which results in a lower computational overhead.

The centroid-tracking method assumes that the displacement between centroids of the same event in consecutive video frames is shorter than the displacement between different events centroids. This technique involves measuring the Euclidean distances between interactions. Regions closer in consecutive video frames belonged to the same event. Figure 4 illustrates the possible trajectories of pigs involved in ear biting in two consecutive frames. The interactions e_1 and e_2 in the current frame, f_{k-1} can move in any direction in the next frame f_k. The possible trajectories of event e_1 are p and p' leading to point A or B. The distance p is shorter than p'. Hence, it is more likely that the interaction at A progressed from e_1. Similarly, the interaction at point B was

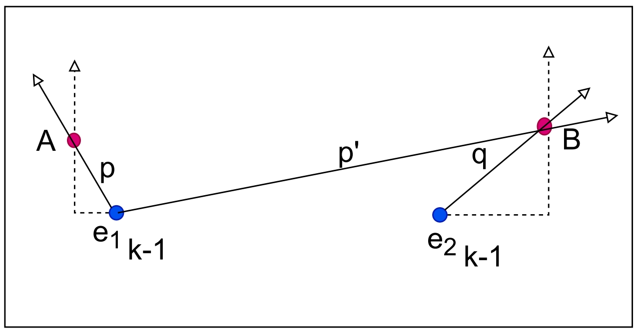


FIGURE 4. **e_1 and e_2 are the centroids of the detected regions (BLUE). Two points were detected in the next frame: A and B (RED). The event trajectories from the current video frames, f_{k-1} to the next are e_1A, e_1B, e_2A, e_2B.**

most likely from event e_2. A new event is recorded when there are unpaired detection(s).

Bashir et al., [41] presented several metrics for evaluating object detection and tracking including Tracker-Detection-Rate, Object-Tracking-Error and False-Alarm-Rate. Equation (1) represents Object-Tracking-Error (OTE) which is a measure of the average discrepancy between the ground truth bounding box centroid and the centroid of tracked region, where N_{rg} represents the total number of overlapping frames between ground truth and tracker, x_i^g represents the x-coordinate of the centroid of object in i^{th} frame of ground truth, x_i^r represents the x-coordinate of the centroid of object in i^{th} frame of the tracker, and y_i^g represents the y-coordinate of the centroid of object in i^{th} frame of ground truth,

TABLE 2. Datasets used for training and evaluation of the detection networks.

Dataset	Data Split	Video frame	Ground truth
I	Training	4672	5,087
	Validation	525	524
II	Training	6,984	7,066
	Validation	844	844

y_i^r represents the y-coordinate of the centroid of object in i^{th} frame of the tracker.

$$\text{OTE} = \frac{1}{N_{rg}} \sum \sqrt{(x_j^g - x_j^r)^2 + (y_j^g - y_j^r)^2} \tag{1}$$

$$\text{Tracker-Detection-Rate} = \frac{TP}{TG} \tag{2}$$

$$\text{False-Alarm-Rate} = \frac{FP}{TP + FP} \tag{3}$$

Tracker-Detection-Rate and False-Alarm-Rate are shown in (2) and (3), respectively, where true-positive (TP) = number of frames where both the ground truth and tracker agree on the presence of ear biting, ground truth (TG) = number of frames for the ground truth objects, false positive (FP) = number of frames where the tracker contains at least one object, but the ground truth does not contain any object or none of the ground truth overlaps, that is, IoU is less than 0.5.

D. FREQUENCY AND DURATION OF EAR BITING BEHAVIOR

Equation (4) shows the frequency of ear biting during the inspection window. The duration of an event is shown as a function of the number of frames in which the event occurs as in (5). where f_k is the k^{th} frame of an event and, e_k represents the number of events detected in k^{th} frame. N is the total number of frames in the observation and r is the camera frame rate. We introduced a wait period of 8 s between events to allow sufficient time for the interacting pigs to disengage. It also applies when the biter and bitten pigs go outside the field of view, for example, when cut-off by the video frame. Given a camera frame rate of, $r = 25$ FPS, 8 s is equivalent to 200 frames. An event was tracked until it disappeared for up to 200 consecutive frames.

$$\text{Frequency} = \frac{\sum_{k=1}^{N} e_k}{N} \tag{4}$$

$$\text{Duration} = \frac{\sum_{k=1}^{n} f_k}{r} \tag{5}$$

E. STATISTICAL ANALYSIS

The frequency and duration of ear biting from Dataset I (AFBI experiment) shown in Table 4 were further analyzed to determine if the effect of the dietary treatments differed between the two observation days (interaction between treatment x day). The data were normally distributed on each day, as assessed using the Shapiro-Wilk test ($p > 0.05$). The pigs in each pen were treated in the same way on the two days, and dietary treatments may have had some effect on the interactions between pigs. The analysis was carried out using RStudio ©2022.07.2 Build 576, 2009-2022, The R Foundation for Statistical Computing for Windows. In the model, dietary treatment and day were used as fixed effects, and batch was added as a random effect. The model was solved using the Linear Mixed Effect and the test statistics were extracted using analysis of variance with Satterthwaite's method.

IV. RESULTS

Table 3 presnets the performance of the detection networks for the two datasets (Table 2). The average precision of YOLOv4 and YOLOv7 on Dataset I were 0.98 and 0.975 AP@0.5, respectively. The performances of the models on Dataset II are listed in Table 3. On a combined dataset (I & II), YOLOv4 and YOLOv7 recorded 0.918 and 0.919 AP@0.5, respectively. Figure 3 shows the detection in a video sequence with overlaid bounding boxes of both the ground truth and object detection results. In this example, a detector missed a region of ear biting in frame 5026 and detected it again in frame 5027 and 5028. The tracker treats these as belonging to the same event until the detection network fails in 200 consecutive frames.

The complete system comprising detection and tracking was evaluated on a test video with a total of 3416 frames where ear-biting was manually identified in 2113 frames (ground truth). The precision and recall of DeepSORT-tracker were 0.859 and 0.63, respectively. The Centroid-tracker achieved precision and recall of 0.66 and 0.965, respectively. Table 5 summarizes tracker performance. Figure 5 and 6 show plots of regions tracked by the DeepSORT and Centroid tracker, respectively. The object tracking error, which represents the average discrepancy between the ground truth bounding box centroid and the centroid of the trackers, is 16.8 and 20.4 for the DeepSORT and Centroid trackers, respectively.

Tracking provides the event identification number and video frames in which they are detected. The interaction frequency and duration were calculated using these parameters. Figure 7 (a) and (b) show the event duration (i.e., the spikes) for a typical one-hour observation of a pen during days 2 and 4. Table 6 shows a summary of the data extracted from Dataset I (frequency and duration) for the two days observed, and Figure 9 captures the variation in frequencies and durations grouped by treatment and day. These results were obtained using the YOLOv4-based tracking methods. The average frequency and duration of ear biting episodes for Datasets I and II were (0.59 & 32 min) and (2.54 & 153 min), respectively, during a one-hour inspection window. It should be highlighted here that the duration is cumulative of all events and for this reason it can be more than one hour.

The statistical analysis of Dataset I (Table 6) showed that neither dietary treatment nor the day of the experiment,

TABLE 3. Results by the two detection networks used on two different datasets and showing Precision, Recall, AP@0.5 and AP@0.5:0.95.

Model	Network parameters	Dataset	Speed (ms)	Precision (P)	Recall (R)	AP@0.5	AP@0.5:0.95
Yolov4-CSP	52M	I	13.4	0.702	0.990	0.980	0.618
Yolov4-CSP	52M	II	13.6	0.746	0.815	0.856	0.506
Yolov7	36M	I	8.0	0.919	0.946	0.975	0.623
Yolov7	36M	II	9.0	0.771	0.753	0.809	0.426

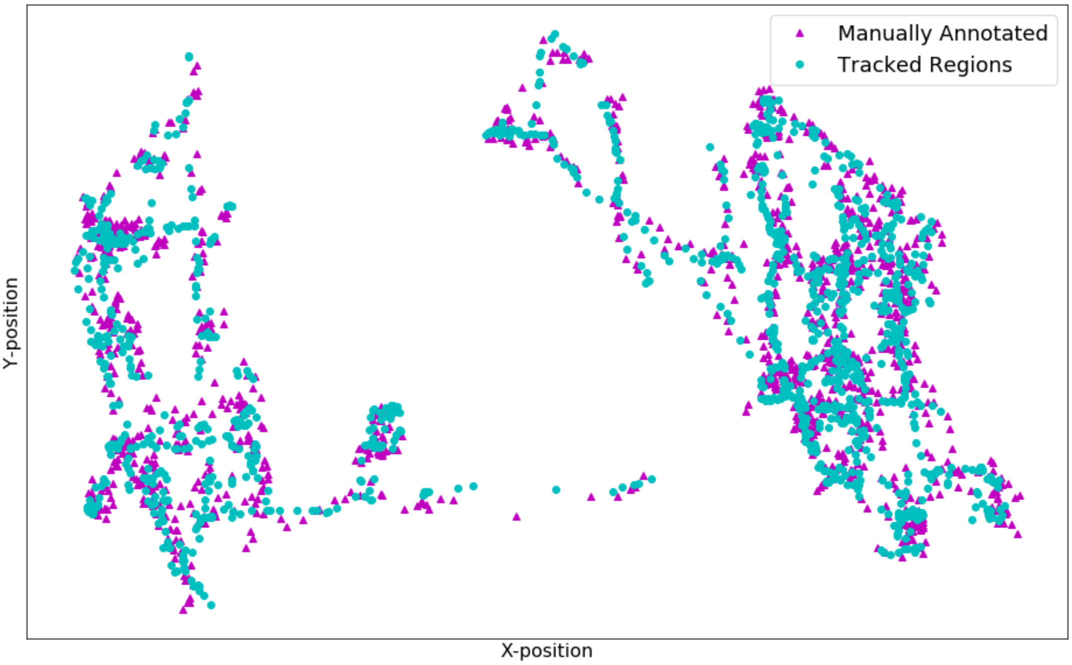


FIGURE 5. Trajectory of regions detected by DeepSORT. The plot shows the degree of overlap between the ground truth bounding box centroid and the centroid of the tracker with Object Tracking Error of 16.8.

TABLE 4. Detection results on combined dataset (I & II) by the two detection networks used, showing Precision, Recall, AP@0.5 and AP@0.5:0.95.

Model	Precision (P)	Recall (R)	AP@0.5	AP@0.5:0.95
Yolov4-CSP	0.742	0.916	0.918	0.573
Yolov7	0.883	0.853	0.919	0.568

TABLE 5. Performance of YOLOv4 detection with the Centroid and DeepSORT tracker.

Metrics	Centroid Tracker	DeepSORT Tracker
True-Positive	2040	1385
False-Positive	1033	228
False-Negative	73	728
Tracker Detection-Rate	0.97	0.66
False-Alarm-Rate	0.34	0.14
Object-Tracking-Error	20.4	16.8

nor the interaction between treatment and day, had a significant effect on the frequency of ear-biting (treatment: $F_{2,20} = 0.4028$ $p = 0.6738$ eta2 $= 0.04$; day: $F_{1,20} = 0.2405$ $p = 0.6292$ eta2 $= 0.01$ interaction between the two: $F_{2,20} = 0.3479$ $p = 0.7104$ eta2 $= 0.03$) measured on day 2 and 4. The same applies to their effects on duration as dietary treatments, day and the interactions between treatments and day did not have a significant effect on the duration of ear-biting episodes: (treatment: $F_{2,20} = 0.2922$ $p = 0.7497$ eta2 $= 0.03$; day: $F_{1,20} = 0.2517$ $p = 0.6213$ eta2 $= 0.01$ interaction between the two: $F_{2,20} = 0.2402$ $p = 0.7887$ eta2 $= 0.02$) measured on day 2 and 4.

V. DISCUSSION

Automatic detection of damaging behaviors, such as ear biting, is important for effective management of outbreaks through intervention. This could facilitate the introduction of remedial measures to improve pig health and welfare. In addition, the automated detection of the features of the behavior would encourage further research and understanding of its occurrence. Although direct observation is widely practiced, such observations are impractical in commercial farm settings because of their large scale. Multiple incidents

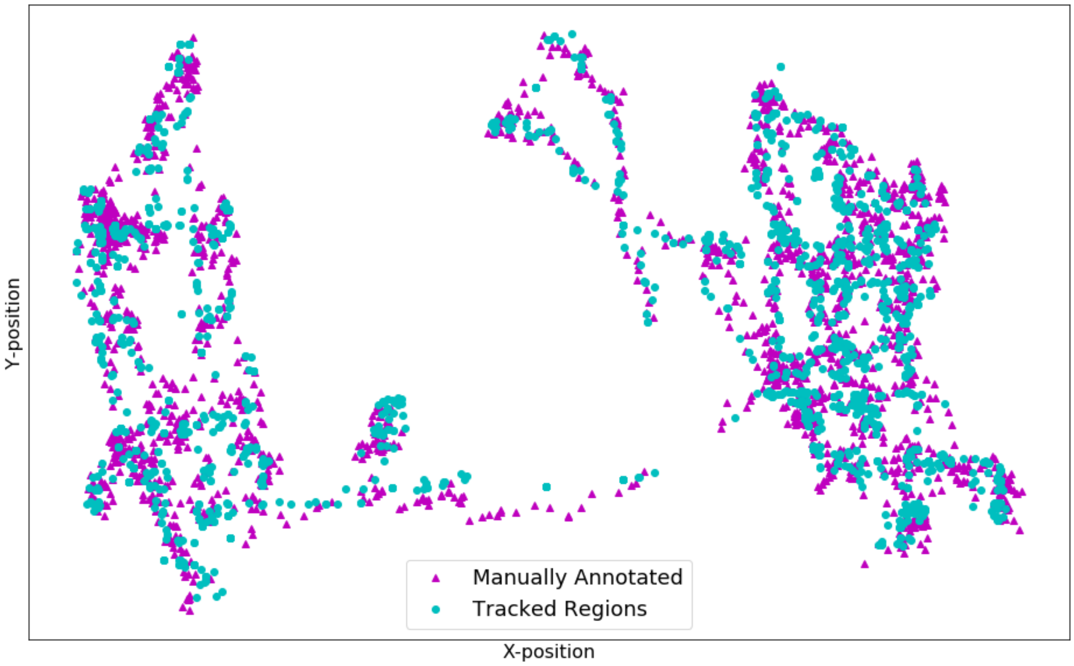
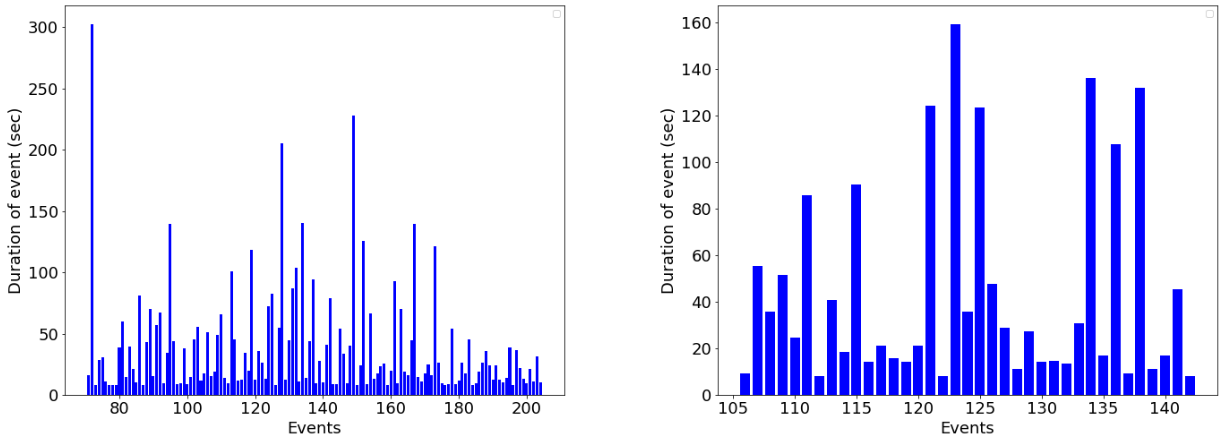


FIGURE 6. Trajectory of regions tracked by Centroid- tracking. The plot shows the degree of overlap between the ground truth bounding box centroid and the centroid of the tracker with Object Tracking Error of 20.4.



(a) Day 2 (b) Day 4

FIGURE 7. The duration of events in days 2 and 4 of the same pen (AFBI video) during one-hour (11:00-12:00) observation.

could occur simultaneously and could be missed because of human limitations, such as, limited field of view and fatigue. Ear biting is usually detected by its consequences, that is, injuries in the affected areas. Such a detection may be too late, and suggests that an outbreak has already started. Therefore, the development of alternative methods for automatic identification and management of abnormal behaviors in pigs is encouraged. Current research efforts are targeting the use of images and videos as they provide non-invasive and cost-effective techniques to detect and manage pigs [42].

As shown in Figure 1, pen configurations varied between the two datasets and affected the way the pigs were monitored. Dataset I (AFBI dataset) was derived from a controlled experiment, and Dataset II (Teagasc dataset) was obtained from a commercial farm where the opportunity for ideal camera position and other modifications were limited. The models showed superior AP on Dataset I, where the pens were viewed from above, reducing occlusion (Figure 1a). In addition, the low density of ten pigs per pen, in accordance with regulations for experimentation, minimized body contact. In contrast, the pigs in Dataset II were monitored from a lateral view, which introduces several challenges. The camera position presented a situation where pigs closer to it obstructed the visibility of those farther away. In addition to the occlusion caused by viewpoint, there were thirty-five pigs in the pen that constantly came into contact with each other, reflecting commercial conditions. There were also some lighting challenges that resulted in uneven illumination as shown in Figure 1b.

Centroid tracking detected more ear-biting events, but also recorded higher false alarms (34%) than DeepSORT (14%). Tracker performance depends on the outputs of the detection network. An instance associated with an episode of ear biting may be missed in a frame, but detected in the subsequent frame. Consecutive detection misses lead to the fragmentation of tracks. To reduce fragmentation, we set the maximum number of missed frames to 200, that is, an event was tracked until the detector failed in 200 consecutive frames. A typical case in which this is useful is when interacting pigs pause momentarily during an ear biting episode. If the waiting time is within the set threshold, the events belong to the same episode. High false alarms would result in the use of scarce resources to monitor pens where ear-biting did not occur. Future work will deal with cases of false positive detection using metrics that would ensure a detected region has at least two pig heads present such that detecting ear biting in a region with a head-to-rear of pigs is invalid.

Our focus was to identify all frames in which ear-biting occurred and use this information to determine the frequency and duration of the episodes. We presented frequency as the number of events within the inspection window expressed as a function of the frames in which the events occurred. While most of the events were short-lived (5-25 s), some continued for over 120 s (Figure 8). Although not directly comparable, the frequency of ear biting episodes was higher in Teagasc than in the AFBI dataset. There are several reasons that may have contributed to this, including differences in management, but discussing these is beyond the scope of this study. Ear biting is often spontaneous, and the biter often receives an immediate reaction from the recipient, such as a head-knock. We also detected sniffing or gentle manipulation of the ear of pen mates, which was identified as a precursor for biting [33]. Early detection of all forms of ear-biting is vital for preventive interventions.

We did not detect any effect of dietary treatment on the frequency and duration of ear-biting episodes. This was contrary to our original hypothesis, as we expected to see an increase in ear-biting behavior, due to the reduction in the protein content of the feeds offered to the pigs in the AFBI experiment, as suggested by [12], [36], and [43]. It is possible that this was due to the time window of our investigation, which occurred during the first four days post-weaning. The first week after weaning is associated with changes in pig housing, social and nutritional management, and pigs appear more unsettled [44]. On the one hand, this may lead to the manifestation of damaging behavior such as ear biting [10]. However, as pigs settle down differences in ear-biting episodes due to dietary treatment may become more apparent, and this is where dietary manipulation benefits may arise.

Frequency is often expressed as the percentage of occurrence of categories of ear biting, for example, the frequency of gentle manipulation, quick bite or pulling of the ear of pen mates [11], [33]. In this study, we detected all contacts targeting the ear of a pen mate as an ear-biting event. Description of the individual types of ear-biting behaviors are presented in Table 1. Different types of ear biting behavior may pose different degrees of injury risk. An event often involves multiple contacts of the different behavior types. For example, the interaction may begin with gentle manipulation and progress to chewing, followed by head knock. Although distinguishing between different ear biting types was beyond the scope of this study, it is possible that the methodology developed here could be extended to account for the greater granularity of the behavior. Ear biting detection will be a useful management tool as it could introduce

TABLE 6. Frequency and duration for two days observation of the AFBI dataset (one-hour inspection window 11:00-12:00). Frequency is the average number of events in the observation. Duration represents the total time of interaction during the inspection window.

Batches	Pens	Treatments	Frequency Day 2	Frequency Day 4	Duration (min) Day 2	Duration (min) Day 4
1	2	2	0.37	0.30	21.92	13.34
	4	4	0.46	1.10	27.57	55.89
	5	5	0.45	0.76	27.05	38.13
2	2	5	0.32	1.14	15.88	57.51
	3	4	0.06	0.01	3.82	0.69
	5	2	0.18	0.14	10.76	8.33
3	2	2	0.84	0.90	45.52	54.00
	5	4	0.68	0.30	40.85	17.84
	6	5	0.41	0.46	19.73	27.62
4	1	4	0.63	0.66	34.49	36.04
	3	5	0.28	0.51	16.88	30.55
	5	2	0.20	0.71	11.92	42.45
5	1	2	1.42	0.45	72.90	27.13
	2	4	1.19	1.72	55.73	86.60
	6	5	0.78	0.20	47.09	12.22

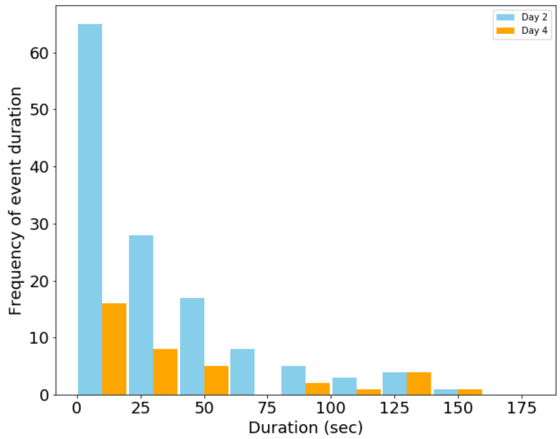
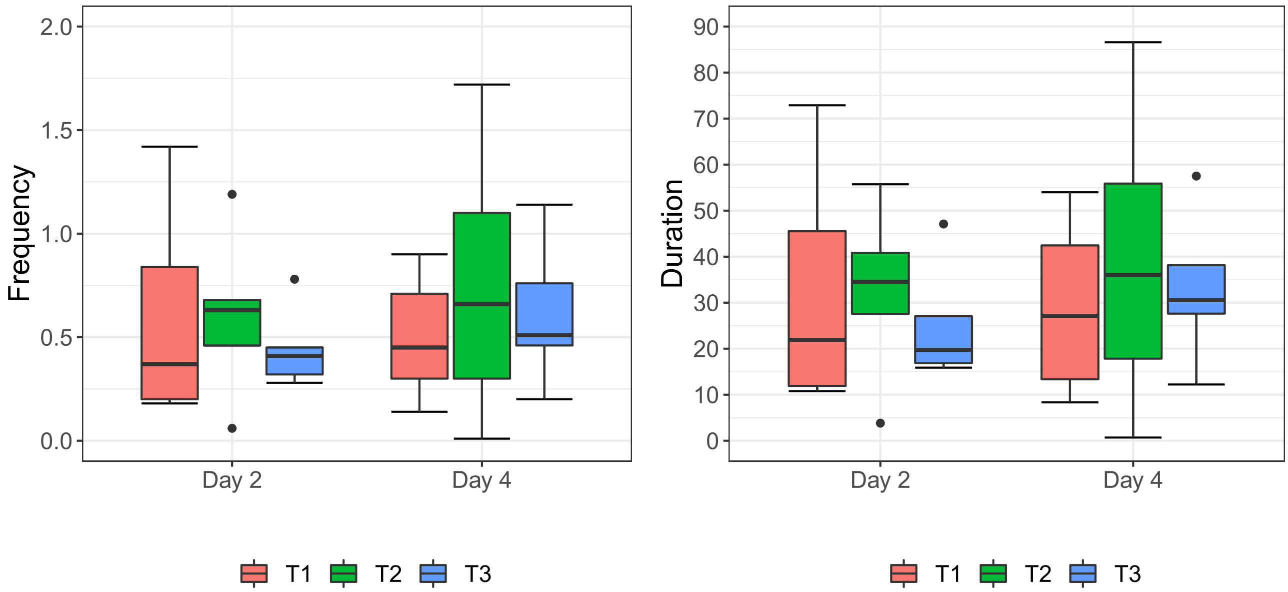


FIGURE 8. Frequency histograms of event duration from AFBI (Batch 5 - Pen 1) for one-hour (11:00-12:00) observation.



(a) Frequency grouped by treatment and day

(b) Duration grouped by treatment and day

FIGURE 9. The effect of different dietary treatments (T1-3) on the frequency and duration of ear biting events on days 2 and 4 post weaning.

a management practice to stop the behavior before injury is caused. Extensions of our method to identify consistent offenders (biter pigs) would require continuous observations for longer periods of time. The identification of such pigs may lead to management interventions, such as isolation.

VI. CONCLUSION

Ear-biting outbreaks present serious health and welfare challenges to pigs in commercial settings, and early detection is critical for efficient management that enhances their health and welfare while minimizing input from antimicrobials. This paper describes an approach to automatically quantify interactions involving ear-biting between pigs using video imaging. The main contributions of this study include the detection and tracking of ear-biting events. We determined the frequency and duration of events by considering all frames in which the event was detected. The test results on two independent datasets demonstrate the effectiveness of the models for identifying ear-biting interactions directly from the images. YOLOv4 and YOLOv7 achieved average detection speeds of 13.5 ms and 8.5 ms, respectively, demonstrating the potential of the proposed pipeline for real-time application.

Our method detected and quantified ear biting at group-level. We consider this the first step in the development of an on-farm early warning system for ear-biting detection. Clearly there are several barriers to overcome when up-scaling this deployment. These include training and testing of the method with more diverse training data arising from, for example, different management settings and pig sizes. If the desire is to operate the deployment in near-real time, then the system needs to detect the individual performer to enable immediate remedial actions such as the removal of the offender from the pen before real damage occurs. This tool has the potential for efficient management of behavior by ensuring the deployment of resources to a few pens where there is a high frequency and duration of occurrence instead of monitoring all pens within a farm. Quantifying performance at the group level meant that individual performers was not considered. The identification of individual pig will facilitate fine-grained intervention, such as the isolation of the offender. In this study, we are interested in group-level performance, that is, the aggregate performance of activities in the pen. Future work will investigate individual-level measurements, which will require pig identification. This can be achieved by assigning permanent IDs to all detected pigs such that ear-biting regions are associated with individual performers.

Ear-biting outbreaks present serious health and welfare challenges for pigs in commercial settings. Early detection enables early intervention, which maximizes the likelihood of success [17]. This is not only critical for efficient management but will also enhance pig health and welfare while minimizing the input from antimicrobials.

ACKNOWLEDGMENT

The authors would like to thank the AFBI Pig Unit staff for their input in the experiments reported here and also would like to thank the Pig Producer and his staff for facilitating the on-farm recording, which resulted in the Teagasc dataset. Dr. Niall McLaughlin of the School of Electronics, Electrical Engineering and Computer Science commented on an earlier version of this manuscript. The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflects only the views of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

[1] L. A. Boyle, S. A. Edwards, J. E. Bolhuis, F. Pol, M. Z. Šemrov, S. Schütze, J. Nordgreen, N. Bozakova, E. N. Sossidou, and A. Valros, "The evidence for a causal link between disease and damaging behavior in pigs," *Front. Vet. Sci.*, vol. 8, Jan. 2022, Art. no. 771682.

[2] N. van Staaveren, A. Hanlon, and L. A. Boyle, "Damaging behaviour and associated lesions in relation to types of enrichment for finisher pigs on commercial farms," *Animals*, vol. 9, no. 9, p. 677, Sep. 2019.

[3] J.-Y. Chou, R. B. D'Eath, D. A. Sandercock, N. Waran, A. Haigh, and K. O'Driscoll, "Use of different wood types as environmental enrichment to manage tail biting in docked pigs in a commercial fully-slatted system," *Livestock Sci.*, vol. 213, pp. 19–27, Jul. 2018.

[4] College of Veterinary Medicine Administration, Iowa State University of Science and Technology. (2022). *Vices (Tail Biting, Ear Biting, Flank Biting, Navel Sucking)*. Accessed: Oct. 28, 2022. [Online]. Available: https://vetmed.iastate.edu/vdpam/FSVD/swine/index-diseases/ear-biting

[5] D. Murphy et al., "EMA and EFSA joint scientific opinion on measures to reduce the need to use antimicrobial agents in animal husbandry in the European union, and the resulting impacts on food safety (RONAFA)," *EFSA J.*, vol. 15, no. 1, Jan. 2017, Art. no. e04666.

[6] S. J. More, "European perspectives on efforts to reduce antimicrobial usage in food animal production," *Irish Vet. J.*, vol. 73, no. 1, pp. 1–12, Dec. 2020.

[7] A. Valros and M. Heinonen, "Save the pig tail," *Porcine Health Manage.*, vol. 1, no. 1, pp. 1–7, 2015.

[8] M. Malik, K. Chiers, F. Boyen, S. Croubels, and D. Maes, "Porcine ear necrosis," *Vet. J.*, vol. 271, May 2021, Art. no. 105655.

[9] N. van Staaveren, J. A. C. Díaz, E. G. Manzanilla, A. Hanlon, and L. A. Boyle, "Prevalence of welfare outcomes in the weaner and finisher stages of the production cycle on 31 Irish pig farms," *Irish Vet. J.*, vol. 71, no. 1, pp. 1–9, Dec. 2018.

[10] J. Pessoa, C. McAloon, M. R. da Costa, E. G. Manzanilla, T. Norton, and L. Boyle, "Adding value to food chain information: Using data on pig welfare and antimicrobial use on-farm to predict meat inspection outcomes," *Porcine Health Manage.*, vol. 7, no. 1, p. 55, Dec. 2021.

[11] A. Diana, L. Carpentier, D. Piette, L. A. Boyle, D. Berckmans, and T. Norton, "An ethogram of biter and bitten pigs during an ear biting event: First step in the development of a Precision Livestock Farming tool," *Appl. Animal Behav. Sci.*, vol. 215, pp. 26–36, Jun. 2019.

[12] Y. V. D. Meer, W. J. J. Gerrits, A. J. M. Jansman, B. Kemp, and J. E. Bolhuis, "A link between damaging behaviour in pigs, sanitary conditions, and dietary protein and amino acid supply," *PLoS ONE*, vol. 12, May 2017, Art. no. e0174688.

[13] D. L. Teixeira, L. C. Salazar, D. Enriquez-Hidalgo, and L. A. Boyle, "Assessment of animal-based pig welfare outcomes on farm and at the abattoir: A case study," *Front. Vet. Sci.*, vol. 7, Oct. 2020, Art. no. 576942.

[14] E. Brunberg, P. Jensen, A. Isaksson, and L. J. Keeling, "Behavioural and brain gene expression profiling in pigs during tail biting outbreaks–evidence of a tail biting resistant phenotype," *PLoS ONE*, vol. 8, Jun. 2013, Art. no. e66513.

[15] A. Haigh and K. O'Driscoll, "Irish pig farmer's perceptions and experiences of tail and ear biting," *Porcine Health Manage.*, vol. 5, no. 1, p. 30, Dec. 2019.

[16] S. Viazzi, G. Ismayilova, M. Oczak, L. T. Sonoda, M. Fels, M. Guarino, E. Vranken, J. Hartung, C. Bahr, and D. Berckmans, "Image feature extraction for classification of aggressive interactions among pigs," *Comput. Electron. Agricult.*, vol. 104, pp. 57–62, Jun. 2014.

[17] S. G. Matthews, A. L. Miller, T. PlÖtz, and I. Kyriazakis, "Automated tracking to measure behavioural changes in pigs for health and welfare monitoring," *Sci. Rep.*, vol. 7, no. 1, p. 17582, Dec. 2017.

[18] S. Wang, H. Jiang, Y. Qiao, S. Jiang, H. Lin, and Q. Sun, "The research progress of vision-based artificial intelligence in smart pig farming," *Sensors*, vol. 22, no. 17, p. 6541, Aug. 2022.

[19] A. Odo, S. McKenna, D. Flynn, and J. B. Vorstius, "Aerial image analysis using deep learning for electrical overhead line network asset management," *IEEE Access*, vol. 9, pp. 146281–146295, 2021.

[20] E. Psota, M. Mittek, L. Pérez, T. Schmidt, and B. Mote, "Multi-pig part detection and association with a fully-convolutional network," *Sensors*, vol. 19, no. 4, p. 852, Feb. 2019.

[21] D. Liu, M. Oczak, K. Maschat, J. Baumgartner, B. Pletzer, D. He, and T. Norton, "A computer vision-based method for spatial–temporal action recognition of tail-biting behaviour in group-housed pigs," *Biosyst. Eng.*, vol. 195, pp. 27–41, Jul. 2020.

[22] H. Ahn, S. Son, H. Kim, S. Lee, Y. Chung, and D. Park, "EnsemblePigDet: Ensemble deep learning for accurate pig detection," *Appl. Sci.*, vol. 11, no. 12, p. 5577, Jun. 2021.

[23] Z. Sha, H. Feng, X. Rui, and Z. Zeng, "PIG tracking utilizing fiber optic distributed vibration sensor and YOLO," *J. Lightw. Technol.*, vol. 39, no. 13, pp. 4535–4541, Jul. 1, 2021.

[24] J. Kim, Y. Suh, J. Lee, H. Chae, H. Ahn, Y. Chung, and D. Park, "EmbeddedPigCount: Pig counting with video object detection and tracking on an embedded board," *Sensors*, vol. 22, no. 7, p. 2689, Mar. 2022.

[25] A. Alameer, I. Kyriazakis, and J. Bacardit, "Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs," *Sci. Rep.*, vol. 10, no. 1, p. 13665, Aug. 2020.

[26] A. Alameer, I. Kyriazakis, H. A. Dalton, A. L. Miller, and J. Bacardit, "Automatic recognition of feeding and foraging behaviour in pigs using deep learning," *Biosyst. Eng.*, vol. 197, pp. 91–104, Sep. 2020.

[27] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2014, pp. 740–755.

[28] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[29] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, Sep. 2016, pp. 21–37.

[30] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 318–327, Feb. 2020.

[31] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Scaled-YOLOv4: Scaling cross stage partial network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13024–13033.

[32] C.-Y. Wang, A. Bochkovskiy, and H.-Y. Mark Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," 2022, *arXiv:2207.02696*.

[33] M. Oczak, G. Ismayilova, A. Costa, S. Viazzi, L. T. Sonoda, M. Fels, C. Bahr, J. Hartung, M. Guarino, D. Berckmans, and E. Vranken, "Analysis of aggressive behaviours of pigs by automatic video recordings," *Comput. Electron. Agricult.*, vol. 99, pp. 209–217, Nov. 2013.

[34] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3464–3468.

[35] N. Wojke and A. Bewley, "Deep cosine metric learning for person re-identification," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2018, pp. 748–756.

[36] N. R. Taylor, D. C. J. Main, M. Mendl, and S. A. Edwards, "Tail-biting: A new perspective," *Vet. J.*, vol. 186, no. 2, pp. 137–147, Nov. 2010.

[37] J.-Y. Chou, K. O'Driscoll, D. A. Sandercock, and R. B. D'Eath, "Can increased dietary fibre level and a single enrichment device reduce the risk of tail biting in undocked growing-finishing pigs in fully slatted systems?" *PLoS ONE*, vol. 15, no. 10, Oct. 2020, Art. no. e0241619.

[38] A. Dutta and A. Zisserman, "The VIA annotation software for images, audio and video," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 2276–2279.

[39] C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 1571–1580.

[40] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," in *Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Jul. 2020, pp. 237–242.

[41] F. Bashir and F. Porikli, "Performance evaluation of object detection and tracking systems," in *Proc. 9th IEEE Int. Workshop PETS*, Jun. 2006, pp. 7–14.
[42] Q. Yang and D. Xiao, "A review of video-based pig behavior recognition," *Appl. Animal Behav. Sci.*, vol. 233, Dec. 2020, Art. no. 105146.
[43] M. McAuley, S. Buijs, R. Muns, A. Gordon, M. Palmer, K. Meek, and N. O'Connell, "Effect of reduced dietary protein level on finishing pigs' harmful social behaviour before and after an abrupt dietary change," *Appl. Animal Behav. Sci.*, vol. 256, Nov. 2022, Art. no. 105762.
[44] I. Kyriazakis, A. Alameer, K. Bučková, and R. Muns, "Toward the automated detection of behavioral changes associated with the post-weaning transition in pigs," *Front. Vet. Sci.*, vol. 9, Jan. 2023, Art. no. 1087570.

ANICETUS ODO (Member, IEEE) received the B.Eng. degree in computer engineering, in 2001, the M.Eng. degree, in 2009, and the Ph.D. degree in engineering from the University of Dundee, in 2022. He is currently a Postdoctoral Research Fellow with the School of Biological Sciences, Institute for Global Food Security, Queen's University Belfast. His research interests include image analysis and deep learning. He has been involved in interdisciplinary research collaborations with people in the academia and industry. Recently, he worked on the application of deep learning for image-based monitoring of overhead electrical towers. His works were presented at major conferences and published in journals.

RAMON MUNS is currently a Principal Research Officer and the Head of the Monogastric Research Unit, Agri-Food and Biosciences Institute (AFBI), and a Honorary Lecturer with Queen's University Belfast. He has nine years of expertise in pig production research, involving management and nutritional strategies to improve health and welfare, and minimize the environmental impact of pig systems. He has published 35 peer-reviewed papers and over 45 scientific abstracts from national and international conference proceedings. He co-supervised one Ph.D. student. He and his team conduct their work in collaboration with a wide range of research institutions at national and international levels as well as with industry partners. His research interests include management and nutritional strategies to support animal health and resilience, especially in pigs born with smaller weights in large litters.

LAURA BOYLE received the M.Agr.Sc. and Ph.D. degrees. She is currently a Senior Research Officer with Teagasc with over 25 years of research expertise in farm animal behavior and welfare science. She is currently in the educational and advisory roles with the Animal and Grassland Research and Innovation Center, Moorepark. She is also an Adjunct Professor with the School of Veterinary Medicine, University College Dublin. Additionally, she informing policy at the national level, she was also an Expert with the European Food Safety Authority, from 2020 to 2022, working on the new Scientific Opinion for Pig Welfare. She has published almost 130 peer-reviewed papers, has over 300 scientific abstracts in national and international conference proceedings, and co-supervised almost 30 Ph.D. and master's students. Her research interests include link between animal health and welfare and the contribution animal welfare can make to the sustainability of agriculture.

ILIAS KYRIAZAKIS is currently a Professor in animal and veterinary science with the School of Biological Sciences, Institute for Global Food Security, Queen's University Belfast. He is also a veterinarian who is interested in the introduction of disruptive technologies to livestock systems, to improve animal health, welfare, and productivity while reducing the environmental impact of their systems. His approach and his team are characterized by multi-disciplinarity and he has collaborated extensively with computer scientists and engineers to achieve this, with a focus thus far being on the application of the technologies to pig and poultry systems. A substantial part of his research is conducted in collaboration with a variety of stakeholders, including the industry, so that the solutions developed by his team enjoy applications.