Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Automated detection and quantification of contact behaviour in pigs using deep learning

Ali Alameer [a,b,*], *Stephanie Buijs* [c], *Niamh O'Connell* [b], *Luke Dalton* [d], *Mona Larsen* [e,f], *Lene Pedersen* [f], *Ilias Kyriazakis* [b]

[a] *School of Science, Engineering and Environment, University of Salford, Manchester M5 4WT, UK*
[b] *Institute for Global Food Security, Queen's University, Belfast BT9 5DL, UK*
[c] *Agri-Food and Biosciences Institute, Large Park, Hillsborough BT26 6DR, UK*
[d] *Zoetis International, Cherrywood, Loughlinstown, Dublin, Ireland*
[e] *Research Group of Animal Welfare, Department of Animal Science, Aarhus University, Tjele, Denmark*
[f] *M3-BIORES, Department of Biosystems, KU Leuven, Leuven, Belgium*

ARTICLE INFO

Article history:
Received 5 October 2021
Received in revised form
23 September 2022
Accepted 5 October 2022

Keywords:
Automated detection
Pig social interactions
Deep learning
Pig behaviour
Tail-biting

Change in the frequency of contact between pigs within a group may be indicative of a change in the physiological or health status of one or more pigs within a group, or indicative of the occurrence of abnormal behaviour, e.g. tail-biting. Here, we developed a novel framework that detects and quantifies the frequency of interaction, i.e., a pig head to another pig rear, between pigs in groups. The method does not require individual pig tracking/identification and uses only inexpensive camera-based data capturing infrastructure. We modified the architecture of well-established deep learning models and further developed a lightweight processing stage that scans over pigs to score said interactions. This included the addition of a detection subnetwork to a selected layer of the base residual network. We first validated the automated system to score the interactions between individual pigs within a group, and determined an average accuracy of 92.65% ± 3.74%, under a variety of settings, e.g., management set-ups and data capturing. We then applied the method to a significant welfare challenge in pigs, that of the detection of tail-biting outbreaks in pigs and quantified the changes that happen in contact behaviour during such an outbreak. Our study shows that the system is able to accurately monitor pig interactions under challenging farming conditions, without the need for additional sensors or a pig tracking stage. The method has a number of potential applications to the field of precision livestock farming of pigs that may transform the industry.

* *Corresponding author*. School of Science, Engineering and Environment, University of Salford, Manchester M5 4WT, UK.
E-mail addresses: A.Alameer1@salford.ac.uk, A.Alameer@qub.ac.uk (A. Alameer).
https://doi.org/10.1016/j.biosystemseng.2022.10.002

Nomenclature

II	Interaction index
IoU	Intersection over union
LR	Learning rate
mAP	Mean average precision
N	Total number of frames
NIF_k	Number of interactions of any two pigs at the *k*th frame
p	Precision
r	Recall
AFBI	Agri-Food and Biosciences Institute
AUF	Aarhus University Foulum
CIEL	Centre for Innovation Excellence in Livestock
CNN	Convolutional neural network
FPS	Frames per second
GPU	Graphical processing unit
HSV	Hue, saturation and value
ITC	Interaction threshold calibration
RAM	Random access memory
ResNet-50	Residual neural network
RFID	Radio frequency identification
RGB	Red, green and blue
YOLO	You only look once

1. Introduction

Pigs, as social animals engage in frequent contact with their conspecifics within a group, to express a number of behaviours, such as grooming, play or sexual interaction (Patbandha et al., 2013). The frequency of such contact may be affected by several factors, such as the age of the pig, its husbandry conditions (e.g., how feed is offered) and husbandry practices (e.g., age at weaning or mixing of groups) (Millman, 2007; Kyriazakis and Tolkamp, 2010). Change in the frequency and/or duration of contact between pigs may be indicative of a change in the physiological or health status of one or more pigs within a group, or indicative of the occurrence of abnormal behaviour, such as belly nosing or tail biting (Millman, 2007; Kyriazakis and Tolkamp, 2010). For example, both the frequency and duration of contact behaviour between pigs are expected to increase when pigs reach sexual maturity (Patbandha et al., 2013), and to decrease during health disturbances (Millman, 2007; Kyriazakis and Tolkamp, 2010). Almost all interactions between pigs are either preceded or followed by significant associations with nosing a certain body region; in particular, nosing the tail can proceed harmful behaviour, e.g., tail biting (Camerlink and Turner, 2013). As such, changes in these characteristics of contact behaviour may have predictive value to detect when, e.g., an infection spreads within a pen or when a behaviour, such as 'tail in mouth' (indicative of the more serious version of tail biting) takes place (Miller et al., 2019).

Currently, there is limited quantification of how contact behaviour changes during health and welfare disturbances, mainly because it is laborious to quantify such behaviour within a pen of pigs for long periods of time. Almost all our current knowledge on the changes in contact behaviour between 'sick' animals comes from laboratory rodent models (Millman, 2007) and therefore may not be applicable to the detection of health and welfare disturbances of livestock. Therefore, there would be benefits from automating the quantification of the behaviour, especially in commercial settings (Matthews et al., 2016).

Over the past few years, several vision-based methods have been developed to track and detect pig behaviours, e.g., mounting, lying, feeding and drinking, at an individual level (Nasirahmadi et al., 2019; Chen et al., 2020a,b,c; Yang et al., 2021); this has not been the case for contact or social behaviour in pigs (Arulmozhi et al., 2021). One of the main challenges associated with quantifying pig interaction-related behaviours, i.e., pig-to-pig, using 2D image data is the difficulty to maintain pig parts identification in cluttered farm conditions (Chen et al., 2019, 2021; Jorquera-Chavez et al., 2021). As a result, fewer studies have attempted to quantify interactions between pigs in groups to identify social behaviour, e.g., aggressive behaviour (Chen et al., 2018). For instance, researchers (Chen et al., 2020d; Gan et al., 2021; Liu et al., 2020) developed a hierarchy of processing stages that quantifies pig interaction, e.g., parallel pressing. Their approach involved cascading multiple processing stages to extract the temporal–spatial features and classify pig behaviour. A limitation associated with this approach is the dependency on predefined sets of consecutive frame sequences to score behaviours. This may be challenging to achieve in commercial farm settings whereby data capturing systems often drop incoming video frames (due to overheating related issues or when various system tasks compete for resources) (Matthews et al., 2017). The other limitation associated with extracting temporal-based features is the increased computational requirements that it imposes to score interactions, e.g., generating sub-videos of pairwise interaction following detection and tracking stages, to then feed a trained recurrent neural network (Liu et al., 2020).

Here, we developed an automated method that enabled us to quantify the frequency of the contact of one pig's head (including snout) with another pig's rear (including tail). There are several reasons why we concentrate on the automated detection and quantification of this behaviour, the main one being that such an interaction is expected to vary during tail biting outbreaks (D'eath et al., 2016) and therefore, it may have a predictive value. Another important reason is that the method can be extended to quantify the contact between a pig's head and another pig's body parts, such as the flank; the contact between snout and flank can also be indicative of abnormal behaviour (Camerlink and Turner, 2013). We propose a two-staged system that detects and identifies complex pig parts, i.e., head and rear, and quantifies these interactions using inexpensive 2D cameras. The proposed system does not rely on pig marking or additional invasive sensors. It only detects the relevant pig parts and scores interactions accordingly. It was trained to handle different camera recording outputs, RGB (red, green and blue) and infrared, constantly changing farm conditions (e.g., lighting conditions), problems of occlusion caused by other pigs and insects occluding the image from the camera. We use the automated method to quantify the head to rear contact during predefined outbreaks of tail-biting within a pen of pigs.

2. Materials and methods

2.1. *Datasets*

The dataset was collated from experimental trials at two sites: Agri-Food and Biosciences Institute (AFBI) in Northern Ireland and the pig research facilities at Aarhus University Foulum (AUF) in Denmark, as we aimed to construct a more holistic and representative dataset for training, validation and testing the proposed method.

2.1.1. *AFBI dataset*

Video recordings were collected during an experiment conducted at the Agri-Food and Biosciences Institute, Northern Ireland. The work was conducted following a protocol approved under Project Licence Number PPL2851 in accordance with the Animals (Scientific Procedures) Act 1986 (The Parliament of the United Kingdom, 1986). The videos were recorded between November 2019 and January 2020. The animals included in the experiment were finishing pigs (Duroc × (Landrace × Large White)), assigned to this study from 30 kg (±0.3 SEM) until slaughter. The pigs were born on-site, weaned at 4 weeks of age, and moved into the pens at 10 weeks of age. All pens were in the same room with a ceiling height of approximately 2 m. All pens were 1.71 × 4.00 m^2 and had plastic slatted floors. Each pen housed 10 pigs (at 1.08 m^2 per pig) and these groups were balanced for gender and body weight. The wall of the building constituted the back wall of each pen, with the other three walls being approximately 1.10 m high. Each pen was fitted with two nipple drinkers and dry feed was provided ad libitum through a single-space electronic feeding station (Schauer Compident MLP pro feeder). Pigs were provided with environmental enrichment in the form of a wooden block and flavoured plastic biting toy (Porcichew, Nutrapet Ltd., U.K.) suspended from the ceiling on a chain. Artificial lighting was provided between 08:00 and 16:00 and pigs also had access to natural light through windows. A trained stockman performed the general management of pigs, including a general daily health check. Temperature was set at 20 °C (19–21 °C), and housing was ventilated through fan-assisted natural ventilation. The experiment was part of a study on the effects of dietary protein level on performance, health and harmful social behaviour in pigs. Half of the pens were allocated to a diet with reduced dietary protein whereas the other half received a commercial diet (control). The pigs involved in this study were tail docked, with 50% of the tail removed within 24 h of birth.

2.1.1.1. AFBI Video Image Selection. GoPro cameras (HERO 5 Session) were installed onto the ceiling above the pen on the day that videos were recorded, resulting in a variable but mostly vertical recording angle, see pen layout in Fig. S1 of the Supplementary Materials. Videos were recorded from 10:30–15:30 on 7 different days throughout the duration of the finishing period. After recording the images were compressed in Adobe Premiere Pro., the stored dataset had an RGB standard format with a fixed frame rate of 15 frames per second (FPS) and an image resolution of 765 × 432 pixels. We used 4 pens across 4 days for the method development and validation.

2.1.2. *AUF dataset*

We utilised video recordings collected during a large scale experiment conducted at the Department of Animal Science, Aarhus University, Denmark. The experiment was conducted following a protocol approved by the Danish Animal Experiments Inspectorate (Journal no. 2015-15-0201-00593). The videos were recorded in June and July 2015. The animals included in the experiment were finisher pigs from 30 kg body weight until slaughter, bred from dams of Danbred Yorkshire × Danbred Landrace, all inseminated with Danbred Duroc semen. The pigs were born at a Danish commercial farm and arrived at the experimental unit as weaners. The pens were all in the same building, which was approximately 5 m high and divided into two sections with two concrete walls and a hallway between them. All pens were identical with dimensions of 5.45 × 2.48 m and with the flooring equally divided between solid, drained and slatted areas. The wall of the building constituted the back wall of each pen, with the other three walls being approximately 1 m high. The pigs were fed ad libitum with a commercial dry feed and the feeders were filled three times a day at 03:00, 10:00 and 18:30 h. Lights were switched on between 05:30 and 18:30. The general management was performed by trained stock people, which included a general daily health check. The experiment was part of a larger study on risk factors for tail biting, involving multiple batches of finisher pigs. The larger study was a 2 × 2 × 2 factorial experiment, meaning that each pen was assigned to one level of three treatments: (1) pigs with docked tails v. pigs with undocked tails, (2) provision of 150 g of straw per pig and day v. no provision of straw and (3) stocking density of 1.21 m^2/pig (11 pigs, 2 feeding spaces) v. 0.73 m^2/pig (18 pigs, 3 feeding spaces). More details about pen layout is provided in (Larsen et al., 2018). The tail-docked pigs had their tails shortened to half their original length, following Danish legislation (Larsen et al., 2018, 2020).

2.1.2.1. AUF video image selection. The AUF dataset was recorded using surveillance infrared cameras (Monacor, Type-TVCCD-1705, Bremen Germany). The cameras were positioned at a height of about 3 m on the wall of the building at the back of each pen, giving a horizontal (rather than vertical) view of the entire pen. The orientation/set-up of the cameras resulted in covering/recording parts of adjacent pens on the farm. To block the view from adjacent pens, we manually configured coordinations of polygon-shaped binary masks and composited these across pens, see (Supplementary Materials, Fig. S2).

We used recordings that took place from 8 am to 4 pm across all days of the trails. The data capturing system generated a non-standard format video switch (.vdm) format of variable frame rate. Before we applied our detection method, the dataset had to be reconstructed and converted, with Microsoft video codec, into a standard format with a resolution of (689 × 474) pixels.

We selected 4 pens without straw, each with 11 undocked pigs and the same stocking density (1.21 m^2/pig). In two of the pens a tail biting outbreak was recorded and on the two others there was no evidence of this (Tables S1 and S2 in the

Supplementary Materials). A pen was classified as having an outbreak when at least one pig in the pen was observed to have a bleeding tail, according to a predefined protocol (Larsen et al., 2018). Trained observers scored the tails in detail from within the pen every Monday, Wednesday and Friday. Furthermore, stock people scored tail injuries daily from outside the pen. Tail biting was not observed in all pens and at the same time point. For this reason, we chose to match pens where an outbreak was recorded, with pens where no outbreak was recorded over the same period of time. We focused on the video recordings that took place 3 days before the tail biting outbreak was identified in two pairs (each with an outbreak/control pens) with one month between them.

2.1.2.2. *Dataset annotation detection of pig parts*. The collected image dataset was annotated by two trained individuals with an animal behaviour background. It encompassed a variety of scenarios, for example, pigs in close contact with one another and under various lighting conditions (Alameer et al., 2015, 2016, 2020a; Ameri et al., 2022). We configured a set of pre-processing stages to augment the dataset, applying arbitrary scaling and horizontal flipping. We also manipulated the colour of the pixels and randomly altered the brightness and contrast using the hue, saturation, value (HSV) colour space (Kekre et al., 2010).

The detection dataset comprised a total of 51,193 instances (26,533 AFBI + 24,660 AUF) across 2781 images (1556 AFBI + 1225 AUF); each pig within an image was manually annotated into two parts: head and rear. A bounding box[1] was applied manually on the head and rear of all pigs in a pen. The bounding box denotes the location and size of each pig part.

2.1.2.3. *Contact between pigs*. An additional dataset was annotated to validate the interaction method, i.e., the processing stage that feeds from the detection method. This dataset consisted of images from both farms used in this framework. The total number of images of this dataset was 670 images; with sets of 376 and 294 images to represent AFBI and AUF datasets, respectively. A similar procedure was followed for selecting the image samples to diversify the dataset. This new dataset was annotated by an animal behaviour scientist who scanned all images to score interactions using a predefined ethogram, Table S3 (Supporting Materials). Any contact between one pig head and another pig rear was scored in a spreadsheet format. The entirety of the dataset comprised four classes (per image) as the following: no contact, 1 contact; 2 contacts; 3 or more contacts. Very few images (of the AUF dataset) included more than three contacts; therefore, these were combined in one class to achieve a more balanced class distribution.

2.2. Proposed methodology

The annotated dataset, provided by the trained data annotators, was fed to the developed deep learning-based pig parts detector for training and validation. We developed a new configuration from the YOLO (you-only-look-once) detector due to its high detection precision and speed (Redmon and Farhadi, 2018; Bochkovskiy et al., 2020). This model has achieved high mean average precision (*mAP*) and speed on standard detection tasks, such as PASCAL VOC (Everingham et al., 2010) and Microsoft Common Objects in Context (MS COCO) Lin et al. (2014). We also adopted relevant techniques for selecting anchor boxes i.e., sets of bounding boxes coordinates fine-tuned for the task (e.g., aspect ratio of pig parts). Finally, we developed a processing module that feeds from the detection network and accordingly scores interactions.

2.2.1. Detection method

The YOLO detector is a multi-scale network that uses a feature extraction network combined with a detection network to generate predictions. It applies a convolutional neural network (CNN) on an input image to produce network predictions from multiple feature maps. The object detector decodes generated predictions to produce bounding boxes with labels that identify objects of interest.

The hyper parameters (e.g., batch size, solver, learning rate schedule settings, and the maximum number of epochs) used to train the model were selected using nested cross-validation Alameer et al. (2019, 2020c). Furthermore, we utilised a transfer learning strategy by pre-training our feature extraction network on the ImageNet dataset (Deng et al., 2009).

In this work, we configured the architecture of both versions of YOLO (Redmon and Farhadi, 2018; Bochkovskiy et al., 2020) to tailor them to the task, i.e., detecting smaller pig parts. We developed and embedded additional detection heads (consisting of convolution, batch normalisation, and relu layers) within the structure of their baseline model, e.g., (the 10th addition layer) of the ResNet-50 model (He et al., 2016) in YOLO (Redmon and Farhadi, 2018), see (Fig. 1). In the latter scenario, this layer downsampled the input image (by a factor of 16) such that it achieved a suitable trade-off between feature depth and spatial resolution. Deeper layers encode higher-level image details at the cost of spatial resolution, an essential trade-off for detecting small objects such as pig heads. The additional detection sub-network is twice the size of the original detection network, therefore it is more suited to detecting smaller pig parts. Data processing and models were implemented in Matlab R2021a and tensorflow object detection API (v.1.13.1) within Ubuntu 20.04.2 LTS on core i9 processor (4.3 GHz) PC using (8×16) G RAM and NVIDIA GeForce RTX 2080 Ti GPU.

2.2.1.1. Anchor boxes

Selection of anchor boxes has a significant impact on the performance of detectors (Redmon and Farhadi, 2018). Here, we used the K-means clustering algorithm with the intersection over union (*IoU*) as a distance metric, to select anchor boxes. The *IoU* between any given two bounding boxes (BBox1 and Bbox2) is obtained using (Eq. (1)):

$$IoU = \frac{|Bbox_1 \cap Bbox_2|}{|Bbox_1 \cup Bbox_2|} \tag{1}$$

2.2.1,2. *Training and evaluation procedure*. In order to develop a system that can handle the diverse farm conditions (e.g.,

[1] A bounding box is represented by a vector in the format [x y width height], where x and y correspond to the upper left corner of the bounding box while the width and height defined the width and height of the rectangular-shaped box around each pig part.

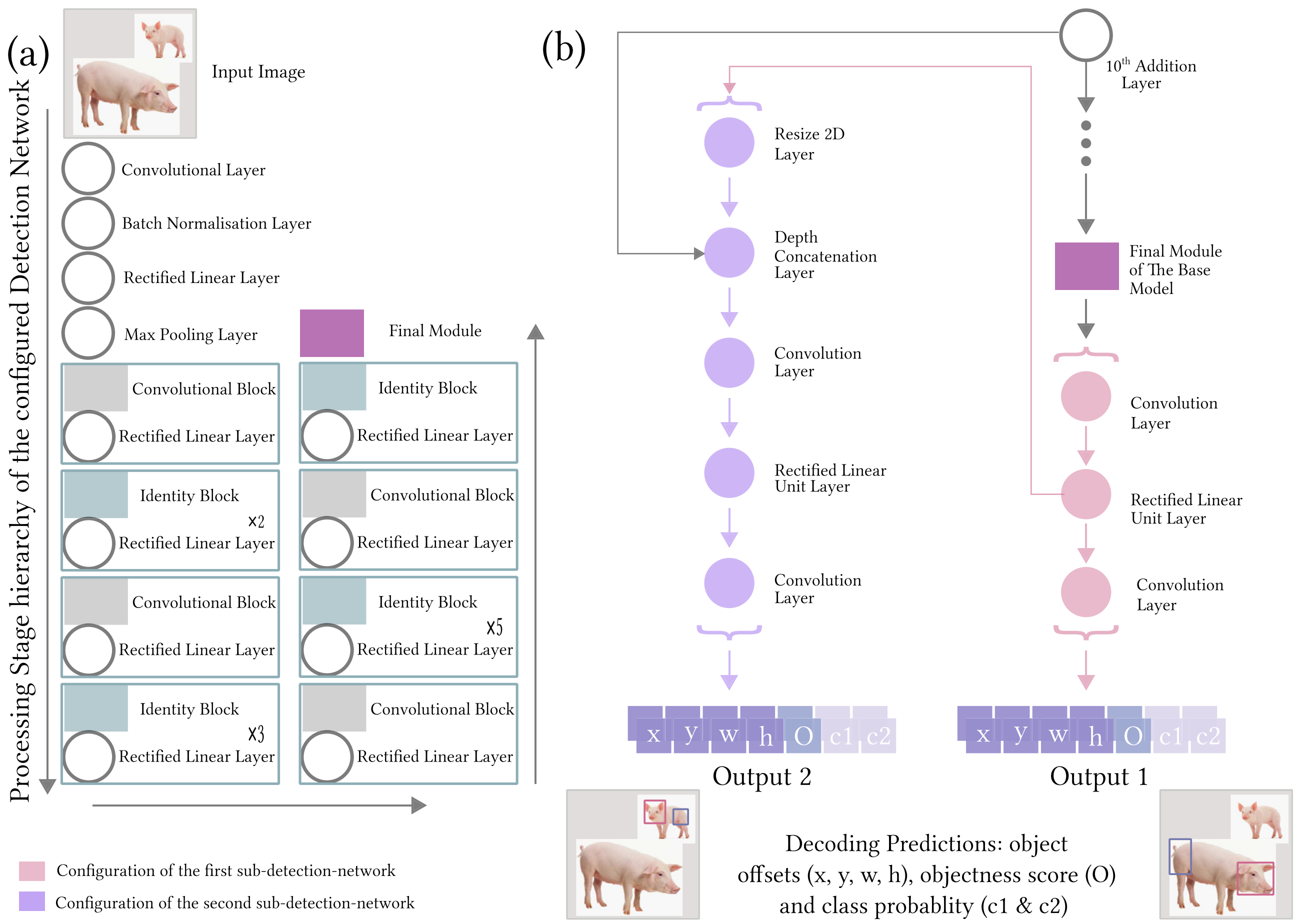

Fig. 1 – The configuration of the proposed detection method. (a) The architecture of the base network (ResNet-50). (b) The configuration of the first and second sub-detection networks. The first detection subnetwork (output 1) is more compatible with detecting larger objects, e.g., when pigs are closer to the camera. The second detection subnetwork (output 2) was developed to detect smaller objects, e.g., pig heads or when objects are further away from the camera. Further details regarding the configuration of the detection method architecture is provided in Fig. S5 and Table S4 of the Supplementary Materials.

pigs with different markings and size) and on pens with different settings, e.g., camera orientation and types of recording, we trained the network with many samples of the dataset using pigs exhibiting a diverse range of postures. The training dataset was constructed using thousands of sample frames at different days/time of the day during trial periods. To evaluate system performance, we used 2502 images (including 46,026 instances) for training and the rest for testing (279 images; 5167 instances). Table 1 shows the obtained parameters used to train and evaluate the detection method.

Further to the training dataset, we configured a set of augmentation routines and applied it to each image batch during the training phase. This process was repeated at each epoch to continually feed the network with various sets of images. This routine enabled the network to reduce the effect of over-fitting by learning similar features at each iteration/epoch. The variety in the validation dataset, which was collected from two different farms and many pens, enabled us to thoroughly examine the generalisation capacity of the proposed detector. For evaluation, we utilised the *mAP* (or AP^{50}) metric. The average precision for a particular class includes both the precision (*p*) and the recall (*r*); it denotes the area under the precision–recall curve (Eq. (2)) across all test image dataset.

$$Average\ Precision\ (AP) = \int_0^1 p(r)dr \tag{2}$$

2.2.2. *Interaction method*

To quantify interaction across all pigs in a pen, we developed a method that scans over all detected pig parts, searching for any possible contact. We computed the *IoU* between each detected pig head and pig rear as shown in (Eq. (1)). For instance, for pens with 12 pigs, i.e., 12 heads and 12 rears, we computed the overlap ratio 144 $\times$ per each frame. For each frame in a video sequence, we generated an interaction array ($\Re^{Heads \times Rears}$); each element of the array denoted the *IoU* between any two pigs (head-to-rear). An interaction is scored between any two pigs only when the corresponding *IoU*

Table 1 – Parameter selection used to train and evaluate the detection method; this includes anchor boxes. The selection of parameters (including the learning rate schedule, base network and the additional detection head) was performed using a nested-cross validation using an independent dataset. Mean average precision (*mAP*) is calculated with an *IoU* threshold of 0.5 (or AP^{50}) across all experiments.

Parameter	Value/Strategy
Solver	Stochastic gradient descent with momentum optimiser
Momentum	0.9
Learning rate (*LR*)	1×10^{-3}; with a warm-up period of 1×10^{3} iteration. This denotes the number of iterations that regulated the learning rate based on the following: $LR = LR \times (\frac{itiration}{warm-up\ period})^{4}$
Max number of epoch	100
Size of mini-batch	8
Network input size	[600, 600, 3] ResNet-50
Number of anchor boxes	2; boxes values calculated with *K*-means clustering algorithm
Feature extraction network	ResNet-50
Additional Detection Head	10th addition layer
ℓ_2 regularization factor	0.5×10^{-3}
Confidence threshold	0.5; keep only detections with confidence scores above this value
Overlap threshold (*IoU*)	0.5; remove overlapping detections

exceeds a pre-calibrated interaction threshold, see Figs. S3 and S4 of the Supplementary Materials. This mechanism enabled quantifying all interactions that may occur at any given frame, including interactions that involve more than two pigs, e.g., two pigs interacting with another. The method is also dynamic and it covers the entire pen area and it is not restricted to a certain part of the pen with a limited number of pigs. For calibrating an interaction threshold, we deployed an interaction threshold calibration (ITC) dataset (See Supplementary Materials; Contact Threshold Calibration) for each pen separately.

For method validation, we deployed the manually annotated head-to-rear contact dataset (see Dataset section). For each image of the dataset, we used the automated method to score interactions between pigs. We then obtained the confusion matrix to compare and report the method performance against that of the annotator.

2.2.3. *Testing for tail-biting outbreaks*

The interaction method was tested to detect tail biting outbreaks on the AUF dataset. The testing stage involved extracting the overall interactions across the selected 16-day tests. This included the two pairs of control/tail-biting outbreak pens (See Dataset section). In this part of the experiment, we fed the trained system with the video data (8 am–4 pm) across pens/dates to generate interaction indices. Each day of the test data consisted of around 26, 000 frames; with a total of 414,670 processed frames.

We compared the number of interactions in the control and ‘outbreak’ pens over 4 consecutive days. Days included in the statistical analysis were day −3, day −2, day −1 and day 0 relative to the outbreaks of tail biting, where day 0 was the day of the outbreak, with the same dates for the matched pairs of outbreak and control pens. The produced interaction metrics across pens/days with control/outbreak were then matched and analysed as a randomised block design.

To overcome issues related to variable frame rate, we constructed an index as a function of the corresponding cumulative interactions and the total number of frames, see (Eq. (3)). This approach enabled obtaining consistent measures across various data frames, e.g., recording drops frames due to pre-removed idle frames and hardware overheating issues.

$$II_i = \frac{\sum_{k=1}^{N} NIF_k}{N} \tag{3}$$

In (Eq. (3)), *II* refers to the interaction index in a given pen, *N* is the total number of frames in a video segment, NIF_k the number of interactions of any two pigs at the *k*th frame. The interaction index captures a whole-encompassing set of intentional/non-intentional interactions of head-to-rear between any two pigs, e.g., snout-to-tail and ear-to-tail. To mitigate the impact of large-scale variations of data frames between days of the trial, we adopted a customised sample rate of 0.25 FPS, i.e., the rate at which we scored interactions between animals. Furthermore, the II can be obtained over any given number of frames as shown in (Eq. (3)). In practice, we simply define the number of frames (*N*) to which we compute the II for any given period of time, e.g., hourly.

3. Results

3.1. *Detection of Pig Parts*

To assess the configured YOLO structures for detecting pig head and rear, we compared their performance with recent and well known deep learning-based object detectors, see Table 2. The comparison was based on model precision and speed to process an image. Both configurations of YOLO (Redmon and Farhadi, 2018; Bochkovskiy et al., 2020) showed high performance in terms of precision and speed. The configured YOLO (Bochkovskiy et al., 2020) achieved the highest average precision across all model configurations in Table 2. Generally, both YOLO model configurations have shown to be more compatible to pig parts detection and have outperformed other well-established models provided in Table 2. Therefore, the configuration of YOLO (Bochkovskiy et al., 2020) has been utilised for all subsequent experiments.

Table 2 – Detection results of the deployed detectors. All models were tested with identical parameters, including the training/validation image set. The table shows results in terms of the man average precision (AP^{50}) and inference speed (CPU-wise processing). The configured YOLO detectors achieved the highest performance and therefore have been deployed for the testing stage.

Detection Method	Base Network	Speed (second)	AP^{50}
EfficientDet (Tan et al., 2020)	EfficientNet (Tan and Le, 2019)	0.45	0.6134
SSD (Liu et al., 2016)	ResNet-50 (He et al., 2016)	0.09	0.5582
F-RCNN (Ren et al., 2017)	ResNet-50 (He et al., 2016)	0.60	0.6013
YOLO (Redmon and Farhadi, 2018)	DarkNet53 (Redmon, 2018)	0.32	0.8399
YOLO (Bochkovskiy et al., 2020)	Cross-Stage-Partial Darknet53 (Bochkovskiy et al., 2020)	0.91	0.8497
Configured YOLO (Redmon and Farhadi, 2018)	ResNet-50 (He et al., 2016)	0.35	0.8634
Configured YOLO (Bochkovskiy et al., 2020)	Cross-Stage-Partial Darknet53 (Bochkovskiy et al., 2020)	0.74	0.8735

Within the training phase of the configured model (Table 1), a learning schedule was set in order to improve the model learning in accordance with each phase of the process (Fig. 2a,b) (Zhang et al., 2019). The selected model achieved a high AP^{50} of (0.863 ± 0.013). The method has been shown to perform slightly better in detecting pig rears (0.872) than pig heads (0.853). The precision/recall (PR) curve showed that this model achieved high precision at varying levels of recall, see (Fig. 3). The PR curve also indicated that for both pig parts, i.e., head and rear, the model did not compromise the false positive rate, i.e., increasing the number of detected pig parts, to maintain a high recall. At high thresholds, we have few false/true positive samples therefore the difference of 1 positive classification could dramatically change precision. As a result, there were slight volatility (−0.05) at those regions (with high thresholds) as shown at the start of the curve in (Fig. 3b).

Figure 4 illustrates an example of detecting pig heads and rears for both the AFBI (Fig. 4a) and AUF (Fig. 4b) datasets. The method is capable of detecting and identifying pig parts under various challenging conditions, e.g., camera orientations and absence of colours with a high confidence score.

3.2. *Contact between pigs*

Using the annotated ITC dataset, we calibrated thresholds for each pen/camera; an interaction is scored only when the *IoU* value between pigs exceeds the given threshold. Expectedly, cameras positioned with a vertical angle facing down the pen, i.e., AFBI dataset, have shown to require lower values of the interaction threshold. Conversely, higher threshold values were computed to score contacts between pigs for pens with cameras positioned with a more horizontal angle.

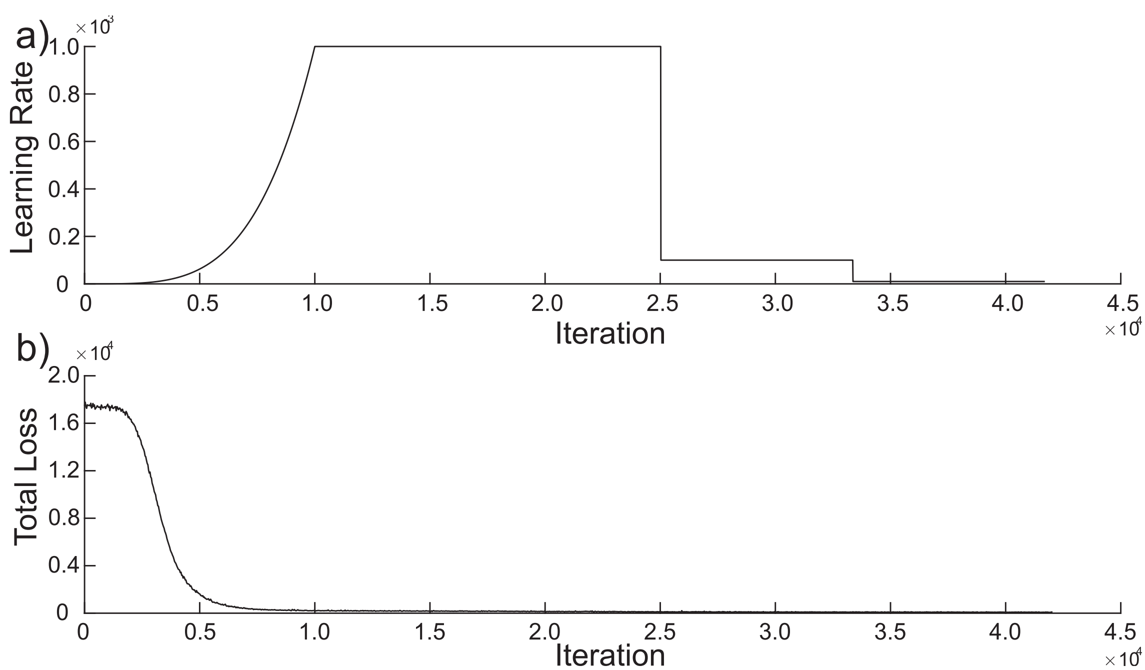

Fig. 2 – Learning stages of the detection model. (a) The dynamic selection and schedule of the learning rate. The value of the learning rate was repeatedly computed based on the iteration number. After a warm-up period, we kept the learning rate predefined value constant provided that the remaining number of epochs was less than 60%. Otherwise, we multiplied the learning rate by a factor of 0.1. Finally, if the remaining epochs were more than 90%, we multiplied the learning rate by 0.01. (b) Model learning, i.e., total loss, against iteration on the x-axis.

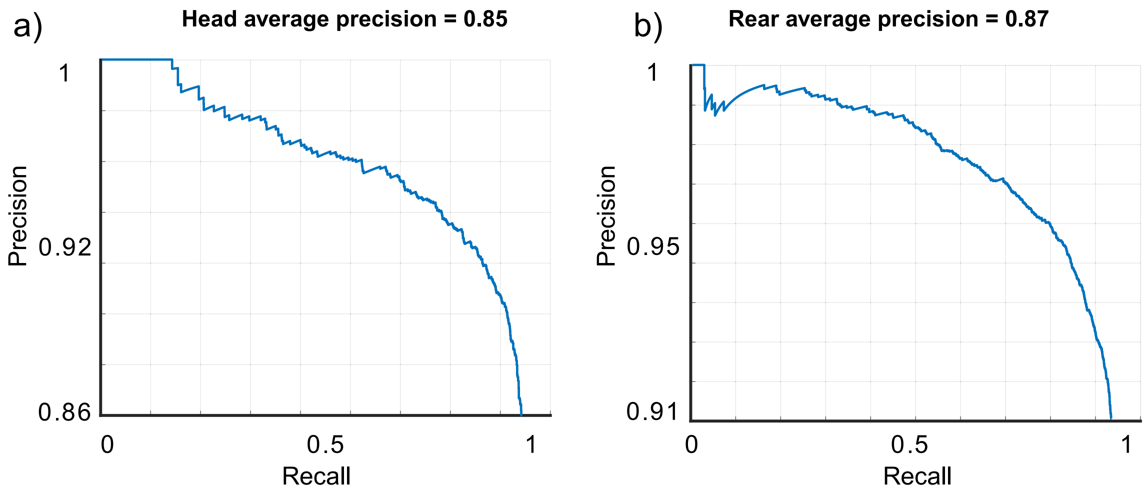

Fig. 3 – Precision-recall curve for detecting pig (a) head and (b) rear. The curve is calculated from the model's detection output for each class; this is by varying the model score threshold that determines the model-predicted positive detection of a given class.

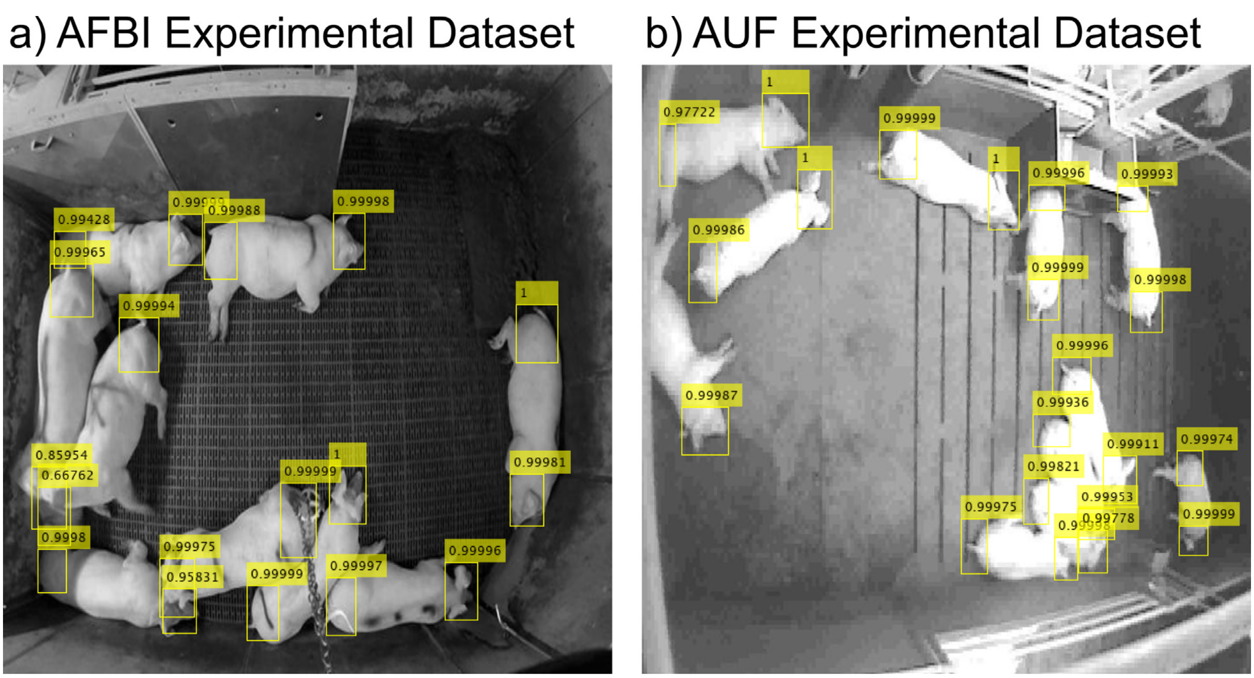

Fig. 4 – Example of detecting pig heads and rears across the experimental datasets using our validated model. Numbers on top of the bounding boxes represent the detection confidence score. A higher score denotes higher confidence in the detection. The method also generates an array of labels for each detected pig part. The generated labels were deployed in the next processing stage to score interactions.

Contact method produced high average classification accuracies of 92.65% ± 3.74%, as shown in Fig. 5, across both datasets, i.e., 95.4% ± 0.5% for AFBI and 90.0% ± 2.3% for AUF. A higher classification accuracy was achieved at AFBI due to, e.g., data capturing settings (see method). The method also achieved higher classification accuracies (95.8% for AFBI; 91.7% for AUF) in detecting the "No contact" class. Accuracy gradually reduces as the number of contacts per image increases, especially for the AUF, where it reduces to 86.7% for detecting 3 or more contacts.

3.3. *Interaction index*

Following the validation stages, we deployed the system to extract interaction indices across the AUF experimental dataset. Figure 6 shows the obtained interaction indices across pens designated 'control' and 'outbreak'. Day 0 represents the day in which the actual outbreak occurred, i.e., tail bleeding was detected. On the days the episodes were scored by pig staff, the method produced an interaction index averaged at 0.25 (±0.009), whereas 0.68 (±0.11) was

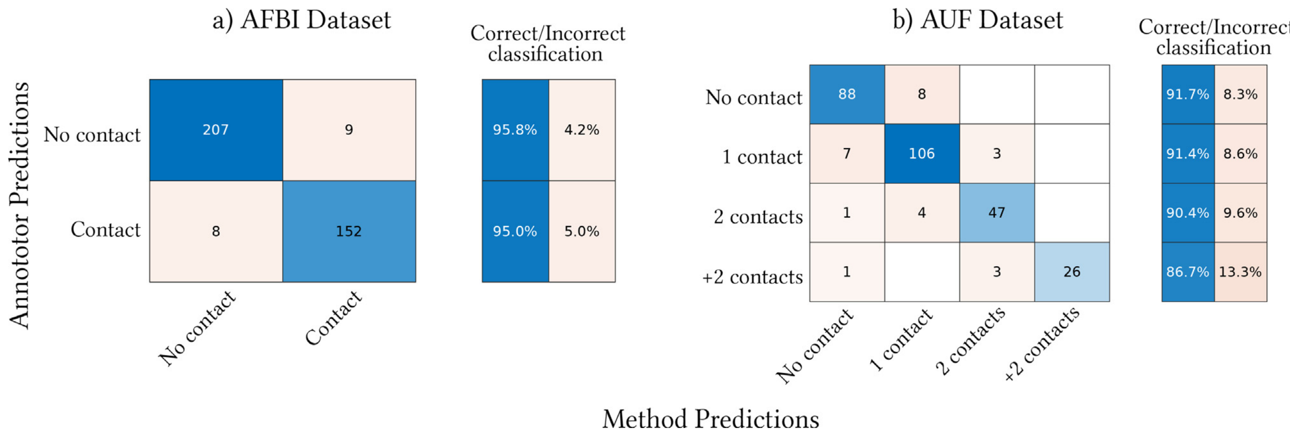

Fig. 5 – Confusion matrices for (a) AFBI and (b) AUF. The metric compares the method performance against that of the data annotator (animal behaviour scientist). For the AFBI dataset, few images (less than 5) have scored more than 2 contacts, therefore, we merged all images with contact into the "contact" class. Similarly, for the AUF dataset, a small number of images (less than 5) have scored more than 3 contacts, therefore, we merged all images with three or more contacts in the "+2 contacts" class. Darker blue colour denotes higher number of images in that class. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

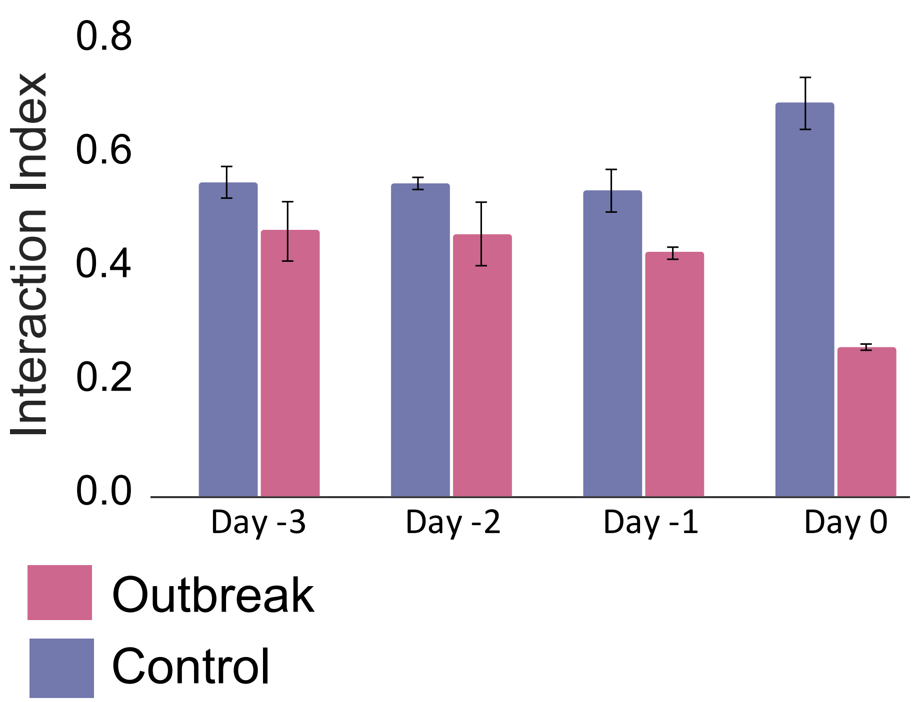

Fig. 6 – The scored interaction indices per day (08:00–16:00) across the study period. The indices represent the day-wise cumulative number of interactions (head and rear including intentional and unintentional) and were calculated to provide consistent measures across various data frames. Control and outbreak bars represent indices averaged across the two study pairs of the AUF dataset. These metrics were obtained using the configured and validated YOLO (Bochkovskiy et al., 2020) model.

scored in equivalent days of control pens. The results also show a significant change in the interaction between treatments (control vs outbreak; $p = 0.017$), which was due to the difference between the two treatments at day 0 (actual outbreak day). Despite the limited number of studied pens, this may suggest that this method is able to detect an outbreak at the same time/level as when a farmer detects a tail bleeding incident.

4. Discussion

We developed a method capable of automatically quantifying contact interactions between pigs in two different environments; each environment presented us with its own challenges. Changes in said interactions may be subtle indicators of reduced health and welfare in pigs that can be challenging to observe, in large commercial settings, at pen side by a stock person (Kyriazakis and Tolkamp, 2010; Matthews et al., 2016). The system operates with high accuracy and speed. If the interactions quantified automatically are capable of relating to change in the pig state, this would confer sustainability benefits to the pig farming industry through enhanced animal welfare, improved performance and timely intervention, e.g., medication and removal of affected pigs. Overall, our work offers a number of contributions to the automated detection of behaviours of indoor housed pigs:

1. We developed a system that detects high-level pig behaviours, i.e., interaction between any two pigs within a group, using only RGB and Infrared cameras.
2. We investigated the characteristics of existing detectors (Bochkovskiy et al., 2020; Liu et al., 2016; Redmon and Farhadi, 2018; Ren et al., 2017; Tan et al., 2020), in terms of speed and detection precision, showing that the configuration of the YOLO network is more suitable for our task.
3. We modified the YOLO network (Redmon and Farhadi, 2018; Bochkovskiy et al., 2020) architecture by adding an additional detection subnetwork to its baseline network, enabling the method to better detect smaller objects, in this case pig heads. We implemented a data-driven process to obtain the optimal layer of the baseline network for feature extraction. We then calculated relevant sets of anchor boxes using the K-means clustering algorithm.
4. We developed an additional processing module that feeds from the detection network and automatically scores interactions between pigs.
5. We applied the proposed system to a significant welfare challenge in the management of pigs, that of the detection of tail-biting outbreaks in pigs.
6. We produced and made publicly available a large annotated dataset for pig parts identification and pig interactions.

The configurations made on the model architecture (e.g., selection of: anchor boxes, detection heads and backbone network, as shown in Fig. 1) enabled better detection of smaller objects (e.g., pig heads as shown in Fig. 4b). The added sub-detection network receives features generated from earlier layers (10th addition layer; Fig. 1) of the feature extraction model, as a result providing a suitable trade-off between spatial resolution and feature depth. Furthermore, applying an unsupervised machine learning technique (k-means clustering algorithm) to generate anchor boxes (from the training dataset) has significantly enhanced method performance. The method gained prior knowledge of sizes/shapes of the detection targets (pig parts) providing enhanced representation to the dataset.

Following a manual inspection of the generated false negative detections, we found that a considerable proportion of these was caused by occlusions, e.g., the head being occluded by the feeding trough (Fig. 4a). These types of missed detection instances (to pig parts) have a reduced impact on the subsequent processing stage that quantifies interactions between pigs as it is highly unlikely for a pig to interact with other pigs (rear) whilst its head is inside the feeding trough. The latter processing stage has lightweight CPU-based computational requirements; it only involves calibrating a threshold for each pen/camera settings to score interactions between pigs. This contrasts with other methods that require extra processing stages with high computational requirements for tracking and extracting spatial and temporal features. The method scanned over all detected pig parts (frame-by-frame) and scored interactions using the *IoU* metric, without the need of computationally expensive processing units, e.g., recurrent neural networks (Chen et al., 2020d; Liu et al., 2020). Therefore, the method is not restricted by the data capturing infrastructure and does not require sets of consecutive frame sequences to quantify interactions (Gan et al., 2021).

In the developed framework, we directly detected pig parts, i.e., pig heads and rear, rather than the traditional detection of the entire pig body (large object), therefore, the target objects were in general of small/medium size. Specifically, the camera orientation for the AUF dataset (e.g., pig heads at the far end of the camera angle of view) and the occurrence of occlusions (to pig parts) have contributed to reducing target size, as shown in Fig. 4(b); to an average area of 30×32. The configured model architecture which adds a detection subnetwork to an earlier layer (the 10th addition layer) of the baseline (ResNet-50) model enabled better detection to these scenarios. The method has been shown to produce a slightly higher rate of incorrect classification for pens with a higher interaction rate (3 interactions or more per image; Fig. 5b), this is mainly due to the increased pig-to-pig occlusion rate caused by camera orientation. However, the method was shown to be robust in various farm settings, including lighting conditions, pig sizes, type of camera, e.g., infrared and pen design. This showed that the method has the potential of wider applicability beyond the presented framework.

The progress made in the automated detection of behaviours in pigs has been reviewed (Matthews et al., 2016). They concluded that little progress had been made in the automated quantification of social behaviour of pigs, and it has been suggested that this was due to the difficulties associated with the technical challenges for some sensors, such as cameras, to capture such behaviours. Some of the examples they provided to justify this statement, included the lack of consensus of what constitutes social behaviour, and in cases where this has been defined the challenges to detect interactions due, for example, to occlusions (Alameer et al., 2017; Alameer, 2018). Over the last 5 years, advances have been made in the quantification of social behaviour of pigs and other species (Munsterhjelm et al., 2019). Some of these approaches have value in the context of scientific research, but they are unlikely to be of relevance to pig farming. For example, Ohayon et al. (2013) have used artificial markings to automatically detect the social interactions between pairs of mice in laboratory settings. Hong et al. (2015) have used depth sensing and machine learning to detect mouse social behaviours. Such methods are unlikely to be of relevance to pig farming, where large numbers of animals are involved and the application of individual markings is impractical Alameer et al. (2020b, 2020c). The widespread use of radio-frequency identification (RFID) amongst livestock offers opportunities for automatic quantification of social interactions, such as social relationships between pigs (Kapun et al., 2020). However, their current use is associated with certain disadvantages, e.g., labour requirements to attach/detach the sensors (Arulmozhi et al., 2021) and the damage this may cause to economically valuable pig parts. The proposed method does not require pigs to be individually marked or fitted with sensors.

Recent research has been directed towards the automated quantification of behaviours of individual pigs (Riekert et al., 2020, 2021; Jensen and Pedersen, 2021), especially in the context of the injurious outcomes these behaviours may produce, e.g., tail biting (D'Eath et al., 2018). Quantification of social behaviour has been identified by several authors as a desirable metric for the assessment of health and welfare of animals (Ohayon et al., 2013; Matthews et al., 2016; Blut et al., 2017). This is because changes in social behaviour can be a useful indicator for the assessment of animal health and welfare, whether its expression increases or decreases. It has been suggested (Kyriazakis and Tolkamp, 2010) that during infection social interaction between farmed animals, such as pigs, is expected to decrease, mainly because there would be a decrease in animal activity, due to the animal feeling 'sick' (Hart and Hart, 2010). This hypothesis has yet to be quantified, in pigs at least. On the other hand, pigs may be expected to come to closer proximity with their pen-mates or even huddle during infectious disease, perhaps due to fever and the need for thermoregulation (Kyriazakis and Tolkamp, 2010). Several authors (examples given by (Larsen et al., 2020)) have suggested that there is an increase in activity and object manipulation prior to outbreaks of tail-biting in pigs, perhaps due to arousal or frustration.

Given the consensus that change in social behaviour may be a useful indicator of the changes in the animal health and welfare status, the question is which metric can be used to assess them. For instance, an attempt was made to investigate the group level movement of pigs (optical flow (Larsen et al., 2020), to predict tail-biting in pigs; hence a more targeted approach for assessing social behaviour in pigs may be desirable, especially to detect tail-biting in pigs and thus provide an additional precision farming tool for farmers. In this work we have used a similarly simple metric as an indicator of social interaction between pigs, the contact between the head of one pig with the rear (including tail) of another. It has been suggested (Camerlink and Turner, 2013) that almost all interactions between pigs are either preceded or followed by significant associations with nosing a certain body region. Here, we found that the frequency of head to rear contact had significantly changed on the day of the defined outbreak, i.e., the day in which the first tail bleeding incident was manually detected by stock people. The number of such interactions was decreased in pens where the outbreak was detected, compared with similar (control) pens where such an event did not take place. This is contrary to the expectation that an increase in activity would be expected to proceed tail biting outbreaks (Larsen et al., 2020). We suggest that such a decrease may be the outcome of contact-avoidance behaviour by a pig which already carries an injury as a result of a previous social interaction that has resulted in tail biting. At this stage, we do not claim that our method has predictive value for the occurrence of tail biting in pigs (due to the reduced size of the tail biting dataset), but it may be a promising method to detect behavioural changes in a pen of pigs. The next stage for this research will incorporate testing a larger dataset of tail biting to get more conclusive evidence regarding the value of the quantified interactions. Additionally, we will develop an additional processing stage in which we categorise and filter finer types of contacts (within head-and-rear) of pigs, e.g., snout-to-tail and non-intentional contacts.

The proposed system provides a novel tool that quantifies the frequency of contact between any pig head and another pig rear (including tails), using non-invasive 2D cameras and based on a bespoke deep learning framework. This novel metric has the potential to quantify the social interaction (contact) between pigs, whether this increases or decreases. The former case may arise in situations of redirected behaviour that may lead to compromises in animal health and welfare. Here we used the case of tail-biting, as a case in point, where we showed that the method detected changes in this metric on the days a tail biting outbreak was detected by manual observation. Decreases in the contact between pigs may occur during outbreaks of disease that lead to pig inactivity, which is one of the early signs of infection in livestock (Kyriazakis and Tolkamp, 2010). Therefore, the method has a number of potential applications to the field of precision livestock farming of pigs. The head-to-rear contact detection method can be modified to automatically quantify the contact between other pig parts, such as head-to-head and head-to-flank contact. Such behavioural metrics may be of relevance to the prediction of injurious behaviours, such as flank chewing (Kyriazakis and Tolkamp, 2010). Only the second processing stage would need to be slightly adjusted to compute *IoU* between pig heads (without any changes to the rest of the workflow). The proposed automated method has the following potential applications:

1. To be used as the basis of extending it to other interactions between different body parts of pigs within the pen.
2. To be applied to detect changes in contact behaviour when pig health and welfare are challenged, such as in the case of infection.

5. Conclusions

Automation in the pig farming industry has the potential to allow detection of early changes in behaviours that occur due to health and welfare compromises. Such changes are impossible to quantify manually and early detection, through automation, should enable timely intervention to reduce compromises in animal welfare and associated economic losses. This paper proposed a novel solution that enables quantifying interactions (head-to-rear contact) between group-housed pigs. The method was based on machine learning and image processing whereby highly established deep learning networks were developed to detect and associate pig parts. We developed a lightweight processing module that rapidly scores intersections between pigs. The paper introduced a practical implementation for detecting interactions between multiple pigs using only video surveillance (infrared and RGB) and suitable to be used in commercial settings, as it was applied in diverse conditions.

Declaration of competing interest

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.

Acknowledgment

We are grateful to Dr Katarina Buckova, Melanie McAuley, Zoe Tey and Joy McMillen for help with the collection and annotation of the datasets. This research was part of the EU-China HealthyLivestock project https://healthylivestock.net/. The authors wish to acknowledge that HealthyLivestock is funded by the European Union H2020 research and innovation program under grant agreement number 773436. The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

The automated detection work was supported in part by the UK Centre for Innovation Excellence in Livestock (CIEL) and Zoetis Inc. Zoetis did not influence the data selection, interpretation, conclusions drawn or the decision on how or what to publish.

The study was made possible by a grant from the Green Development and Demonstration Programme under the Ministry of Food, Agriculture and Fisheries, Denmark (project IntactTails j. nr. 34,009-13-0743).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biosystemseng.2022.10.002.

The code and dataset with all its annotations can be found online at https://doi.org/10.17866/rd.salford.21346767.

REFERENCES

Alameer, A. (2018). *Biologically-inspired hierarchical architectures for object recognition*. Ph.D. thesis. Newcastle University.

Alameer, A., Degenaar, P., & Nazarpour, K. (2017). Processing occlusions using elastic-net hierarchical max model of the visual cortex. In *2017 IEEE international conference on INnovations in intelligent SysTems and applications (INISTA)* (pp. 163–167). IEEE.

Alameer, A., Degenaar, P., & Nazarpour, K. (2019). Context-based object recognition: Indoor versus outdoor environments. In *Science and information conference* (pp. 473–490). Springer.

Alameer, A., Degenaar, P., & Nazarpour, K. (2020a). Objects and scenes classification with selective use of central and peripheral image content. *Journal of Visual Communication and Image Representation, 66*, Article 102698.

Alameer, A., Ghazaei, G., Degenaar, P., Chambers, J. A., & Nazarpour, K. (2016). Object recognition with an elastic net-regularized hierarchical max model of the visual cortex. *IEEE Signal Processing Letters, 23*, 1062–1066.

Alameer, A., Ghazaeil, G., Degenaar, P., & Nazarpour, K. (2015). An elastic net-regularized hmax model of visual processing. In *Intelligent signal processing (ISP) conference*. IET.

Alameer, A., Kyriazakis, I., & Bacardit, J. (2020b). Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs. *Scientific Reports, 10*, 1–15.

Alameer, A., Kyriazakis, I., Dalton, H. A., Miller, A. L., & Bacardit, J. (2020c). Automatic recognition of feeding and foraging behaviour in pigs using deep learning. *Biosystems Engineering, 197*, 91–104.

Ameri, R., Alameer, A., Ferdowsi, S., Nazarpour, K., & Abolghasemi, V. (2022). Labeled projective dictionary pair learning: Application to handwritten numbers recognition. *Information Sciences, 609*, 489–506.

Arulmozhi, E., Bhujel, A., Moon, B. E., & Kim, H. T. (2021). The application of cameras in precision pig farming: An overview for swine-keeping professionals. *Animals, 11*, 2343.

Blut, C., Crespi, A., Mersch, D., Keller, L., Zhao, L., Kollmann, M., Schellscheidt, B., Fülber, C., & Beye, M. (2017). Automated computer-based detection of encounter behaviours in groups of honeybees. *Scientific Reports, 7*, 1–9.

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*. https://doi.org/10.48550/ARXIV.2004.10934

Camerlink, I., & Turner, S. P. (2013). The pig's nose and its role in dominance relationships and harmful behaviour. *Applied Animal Behaviour Science, 145*, 84–91.

Chen, C., Zhu, W., Guo, Y., Ma, C., Huang, W., & Ruan, C. (2018). A kinetic energy model based on machine vision for recognition of aggressive behaviours among group-housed pigs. *Livestock Science, 218*, 70–78.

Chen, C., Zhu, W., Liu, D., Steibel, J., Siegford, J., Wurtz, K., Han, J., & Norton, T. (2019). Detection of aggressive behaviours in pigs using a realsence depth sensor. *Computers and Electronics in Agriculture, 166*, Article 105003.

Chen, C., Zhu, W., & Norton, T. (2021). Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning. *Computers and Electronics in Agriculture, 187*, Article 106255.

Chen, C., Zhu, W., Oczak, M., Maschat, K., Baumgartner, J., Larsen, M. L. V., & Norton, T. (2020a). A computer vision approach for recognition of the engagement of pigs with different enrichment objects. *Computers and Electronics in Agriculture, 175*, Article 105580.

Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., & Norton, T. (2020b). Classification of drinking and drinker-playing in pigs by a video-based deep learning method. *Biosystems Engineering, 196*, 1–14.

Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., & Norton, T. (2020c). Recognition of feeding behaviour of pigs and determination of feeding time of each pig by a video-based deep learning method. *Computers and Electronics in Agriculture, 176*, Article 105642.

Chen, C., Zhu, W., Steibel, J., Siegford, J., Wurtz, K., Han, J., & Norton, T. (2020d). Recognition of aggressive episodes of pigs based on convolutional neural network and long short-term memory. *Computers and Electronics in Agriculture, 169*, Article 105166.

D'Eath, R. B., Jack, M., Futro, A., Talbot, D., Zhu, Q., Barclay, D., & Baxter, E. M. (2018). Automatic early warning of tail biting in pigs: 3d cameras can detect lowered tail posture before an outbreak. *PLoS One, 13*, Article e0194524.

D'eath, R., Niemi, J., Ahmadi, B. V., Rutherford, K., Ison, S., Turner, S., Anker, H., Jensen, T., Busch, M., Jensen, K., et al. (2016). Why are most eu pigs tail docked? Economic and ethical analysis of four pig housing and management scenarios in the light of eu legislation and animal welfare outcomes. *Animal, 10*, 687–699.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *IEEE conference on computer vision and pattern recognition* (pp. 248–255). IEEE.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International Journal of Computer Vision, 88*, 303–338.

Gan, H., Ou, M., Huang, E., Xu, C., Li, S., Li, J., Liu, K., & Xue, Y. (2021). Automated detection and analysis of social behaviors among preweaning piglets using key point-based spatial and temporal features. *Computers and Electronics in Agriculture, 188*, Article 106357.

Hart, B., & Hart, L. (2010). *Encyclopedia of animal behavior*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).

Hong, W., Kennedy, A., Burgos-Artizzu, X. P., Zelikowsky, M., Navonne, S. G., Perona, P., & Anderson, D. J. (2015). Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. *Proceedings of the National Academy of Sciences, 112*, E5351–E5360.

Jensen, D. B., & Pedersen, L. J. (2021). Automatic counting and positioning of slaughter pigs within the pen using a convolutional neural network and video images. *Computers and Electronics in Agriculture, 188*, Article 106296.

Jorquera-Chavez, M., Fuentes, S., Dunshea, F. R., Warner, R. D., Poblete, T., Unnithan, R. R., Morrison, R. S., & Jongman, E. C. (2021). Using imagery and computer vision as remote monitoring methods for early detection of respiratory disease in pigs. *Computers and Electronics in Agriculture, 187*, Article 106283.

Kapun, A., Adrion, F., & Gallmann, E. (2020). Case study on recording pigs' daily activity patterns with a UHF-RFID system. *Agriculture, 10*, 542.

Kekre, H., Thepade, S. D., Lodha, V., Luthra, P., Joseph, A., & Nemani, C. (2010). Augmentation of block truncation coding based image retrieval by using even and odd images with sundry colour spaces. *Int. Journal on Computer Science and Engg, 2*, 2535–2544.

Kyriazakis, I., & Tolkamp, B. (2010). Disease. In *The encyclopedia of applied animal behaviour and welfare* (pp. 176–177). Wallingford, Oxon: CAB International.

Larsen, M., Andersen, H. L., & Pedersen, L. (2018). Which is the most preventive measure against tail damage in finisher pigs: Tail docking, straw provision or lowered stocking density? *Animal, 12*, 1260–1267.

Larsen, M. L. V., Pedersen, L. J., Edwards, S., Albanie, S., & Dawkins, M. S. (2020). Movement change detected by optical flow precedes, but does not predict, tail-biting in pigs. *Livestock Science, 240*, Article 104136.

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *European conference on computer vision* (pp. 740–755). Springer.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In *European conference on computer vision* (pp. 21–37). Springer.

Liu, D., Oczak, M., Maschat, K., Baumgartner, J., Pletzer, B., He, D., & Norton, T. (2020). A computer vision-based method for spatial-temporal action recognition of tail-biting behaviour in group-housed pigs. *Biosystems Engineering, 195*, 27–41.

Matthews, S. G., Miller, A. L., Clapp, J., Plötz, T., & Kyriazakis, I. (2016). Early detection of health and welfare compromises through automated detection of behavioural changes in pigs. *The Veterinary Journal, 217*, 43–51.

Matthews, S. G., Miller, A. L., PlÖtz, T., & Kyriazakis, I. (2017). Automated tracking to measure behavioural changes in pigs for health and welfare monitoring. *Scientific Reports, 7*, Article 17582.

Miller, A. L., Dalton, H. A., Kanellos, T., & Kyriazakis, I. (2019). How many pigs within a group need to be sick to lead to a diagnostic change in the group's behavior? *Journal of Animal Science, 97*, 1956–1966.

Millman, S. (2007). Sickness behaviour and its relevance to animal welfare assessment at the group level. *Animal Welfare*.

Munsterhjelm, C., Nordgreen, J., Aae, F., Heinonen, M., Valros, A., & Janczak, A. M. (2019). Sick and grumpy: Changes in social behaviour after a controlled immune stimulation in group-housed gilts. *Physiology & Behavior, 198*, 76–83.

Nasirahmadi, A., Sturm, B., Olsson, A. C., Jeppsson, K. H., Müller, S., Edwards, S., & Hensel, O. (2019). Automatic scoring of lateral and sternal lying posture in grouped pigs using image processing and support vector machine. *Computers and Electronics in Agriculture, 156*, 475–481.

Ohayon, S., Avni, O., Taylor, A. L., Perona, P., & Egnor, S. R. (2013). Automated multi-day tracking of marked mice for the analysis of social behaviour. *Journal of Neuroscience Methods, 219*, 10–19.

Patbandha, T., Kumaresan, A., Mohanty, T., Layek, S., & Pathak, R. (2013). Understanding pig behaviour for better management. *Pig Journal, 68*, 116–123.

Redmon, J. (2018). *Darknet: Open source neural networks in c*.

Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement*. arXiv preprint arXiv:1804.02767.

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*, 1137–1149.

Riekert, M., Klein, A., Adrion, F., Hoffmann, C., & Gallmann, E. (2020). Automatically detecting pig position and posture by 2d camera imaging and deep learning. *Computers and Electronics in Agriculture, 174*, Article 105391.

Riekert, M., Opderbeck, S., Wild, A., & Gallmann, E. (2021). Model selection for 24/7 pig position and posture detection by 2d camera imaging and deep learning. *Computers and Electronics in Agriculture, 187*, Article 106213.

Tan, M., & Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105–6114). PMLR.

Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10781–10790).

Yang, Q., Xiao, D., & Cai, J. (2021). Pig mounting behaviour recognition based on video spatial–temporal features. *Biosystems Engineering, 206*, 55–66.

Zhang, Z., He, T., Zhang, H., Zhang, Z., Xie, J., & Li, M. (2019). *Bag of freebies for training object detection neural networks*. arXiv preprint arXiv:1902.04103.