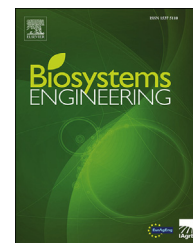




ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

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

Research Paper

Automatic recognition of feeding and foraging behaviour in pigs using deep learning



Ali Alameer^{a,b,*}, Ilias Kyriazakis^c, Hillary A. Dalton^a, Amy L. Miller^a,
Jaume Bacardit^b

^a School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, NE1 7RU, United Kingdom

^b School of Computing, Newcastle University, Newcastle Upon Tyne, NE4 5TG, United Kingdom

^c Institute for Global Food Security, School of Biological Sciences, Queen's University, Belfast, BT9 5DL, UK, Ireland

ARTICLE INFO

Article history:

Received 24 February 2020

Received in revised form

6 June 2020

Accepted 18 June 2020

Published online 12 July 2020

Keywords:

Automated detection

Consummatory behaviour

Deep learning

Pig behaviour

Feeding,

Foraging

Automated, vision-based early warning systems have been developed to detect behavioural changes in groups of pigs to monitor their health and welfare status. In commercial settings, automatic recording of feeding behaviour remains a challenge due to problems of variation in illumination, occlusions and similar appearance of different pigs. Additionally, such systems, which rely on pig tracking, often overestimate the actual time spent feeding, due to the inability to identify and/or exclude non-nutritive visits (NNV) to the feeding area. To tackle these problems, we have developed a robust, deep learning-based feeding detection method that (a) does not rely on pig tracking and (b) is capable of distinguishing between feeding and NNV for a group of pigs. We first validated our method using video footage from a commercial pig farm, under a variety of settings. We demonstrate the ability of this automated method to identify feeding and NNV behaviour with high accuracy ($99.4\% \pm 0.6\%$). We then tested the method's ability to detect changes in feeding and NNV behaviours during a planned period of food restriction. We found that the method was able to automatically quantify the expected changes in both feeding and NNV behaviours. Our method is capable of monitoring robustly and accurately the feeding behaviour of groups of commercially housed pigs, without the need for additional sensors or individual marking. This has great potential for application in the early detection of health and welfare challenges of commercial pigs.

© 2020 The Author(s). Published by Elsevier Ltd on behalf of IAGRE. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The accurate quantification of feeding and associated behaviours is an important challenge for the early detection of health and welfare challenges in livestock. Changes in feeding

behaviour are a key symptom of health and welfare problems (González et al., 2008). Subtler changes, linked to the way in which the animal consumes an amount of food, may be of value for the early detection of health and welfare compromises (González et al., 2008; Tolcamp et al., 2011).

* Corresponding author. School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, NE1 7RU, United Kingdom.

E-mail address: Ali.Alameer@Newcastle.ac.uk (A. Alameer).

<https://doi.org/10.1016/j.biosystemseng.2020.06.013>

1537-5110/© 2020 The Author(s). Published by Elsevier Ltd on behalf of IAGRE. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Nomenclature

F_i	Feeding index
FPF_k	Number of pigs feeding at a given time-frame
K	A given time-frame
N	Total number of frames in a video segment
NNV_i	Non-nutritive visits index
NPF_k	Number of non-nutritive visits at a given time-frame
AWERB	Animal welfare and ethical review body
CNN	Convolutional neural network
FPS	Frames per second
GPU	Graphical processing unit
NNV	Non-nutritive visits
PC	Personal computer
RAM	Random access memory
RFID	Radio frequency identification
RGB	Red, green and blue
T-SNE	T-distributed stochastic neighbour embedding
TL	Transfer learning

Feeding is a fundamental behaviour which can be quantified in a number of different ways when considering a group of pigs. These include recording the amount of food consumed, recording the duration of time spent chewing/biting food, or recording the amount of time and/or frequency that the head of the animal is in the food trough. Unlike actual consummatory behaviour, animals will also visit the feeding area without consuming any feed, to sample or explore the area where food is, or should be, distributed. This is classified as a non-nutritive visit (NNV) (Miller et al., 2019; Weary et al., 2009). The function of this behaviour may simply be to facilitate knowledge about when food is, or should usually be, available (Weary et al., 2009). For instance, when pigs experience food deprivation, their feeding motivation increases, leading to higher activity and heightened interest in the feeding area; an increased number of NNV may be observed (Day et al., 1995; Pastorelli et al., 2012). In other circumstances, where the health or welfare of an animal is compromised, a decrease in the frequency of NNV may be apparent prior to larger scale changes in behaviour, such as in daily food intake (González et al., 2008). To date, quantifying NNV behaviour in group housed animals has only been possible retrospectively via highly time-consuming manual analysis and therefore has limited use in a real world scenario.

Radio Frequency Identification (RFID) provides a suitable solution for detecting the feeding behaviour of pigs (Cornou et al., 2008; Marcon et al., 2015). Electronic ear tags are required to be attached to the pigs long-term, so that their individual food intake can be calculated when they enter the feeding area. On a commercial scale, RFID tags may not be a feasible option as the attachment and detachment of tags entails additional labour cost, and reduces the commercial value of pigs in certain international markets (i.e., value of pig ears). In addition, the implementation of the system imposes constraints on the feeding space/process. Despite

the low-cost and robustness of infrared sensors in quantifying pig activities (Ni et al., 2017), these systems may over-estimate the actual time spent feeding due to the inability to quantify and exclude non-nutritive visits (NNV) to the feeding area.

Video surveillance is a suitable alternative to RFID for detecting feeding behaviour with practical diagnostic value, due to its low cost and the simplicity of its implementation. The key challenge in this approach is how to extract formative features from the images from which actionable knowledge can be reliably extracted (Abolghasemi et al., 2018; Alameer et al., 2016, 2020). In recent years, there have been some relevant studies on how to accurately detect pigs housed in groups. In the context of utilising depth imaging, several researchers (Matthews et al., 2017; Mittek et al., 2017; Sa et al., 2019; Yang et al., 2018) have proposed systems that track individual pigs in a group-housed environment. Despite the accurate tracking of the latter methods, they have only been capable of providing short-term (< 20 min s) segments of behaviours per animal, which may be insufficient for the quantification of behaviours in a commercial context. RGB (red, green and blue) cameras have been used to distinguish the pigs from the background using handcrafted filters of feature extraction, e.g., Gabor filters (Huang et al., 2018; Nasirahmadi et al., 2019b). The main drawback for these image processing methods is the inability to cope with the variable farm environment (e.g., varying illumination) that may easily disrupt system performance. To tackle these challenges, researchers have used convolutional neural networks (CNNs) to accurately detect pigs (Nasirahmadi et al., 2019a; Psota et al., 2019; Zhuang and Zhang, 2019). The dynamic filter selection in CNNs allows invariance to different conditions, e.g., illumination (Ciaparrone et al., 2019; Yang et al., 2018; Zhu et al., 2020).

In this work, we have developed a 2D camera-based deep learning method to automatically detect the feeding behaviour of groups of pigs under commercial conditions, without the need for additional sensors or individual marking. The system operates on grayscale video images, and was trained to handle the constantly changing farm conditions, e.g., lighting conditions, problems of occlusion caused by other pigs, and insects occluding the image from the camera. Unlike previous attempts to detect the feeding behaviour of pigs using traditional pig tracking methods, GoogLeNet-like architectures were utilised to monitor a smaller predefined pen area covering two food troughs and a simple, clearly defined area in front of those troughs. In this way, the proposed system avoids short ID track-related issues, which can continuously distort the accumulative feeding-behaviour recognition process. Our proposed system also allows feeding to be accurately identified (i.e., the pig has its head inside the feeding trough inspected visually from the top of the pen) and separately, NNV behaviour (i.e., the pig has one front foot, plus a second foot, within the defined feeding zone but does not have its head inside the feeding trough) on frame-by-frame basis, see Fig. 1. As our system focuses only on a subset of available feeding troughs within a commercial context, we demonstrate that sufficient data can be collected from this subset to identify changes associated in feeding behaviours at group level.

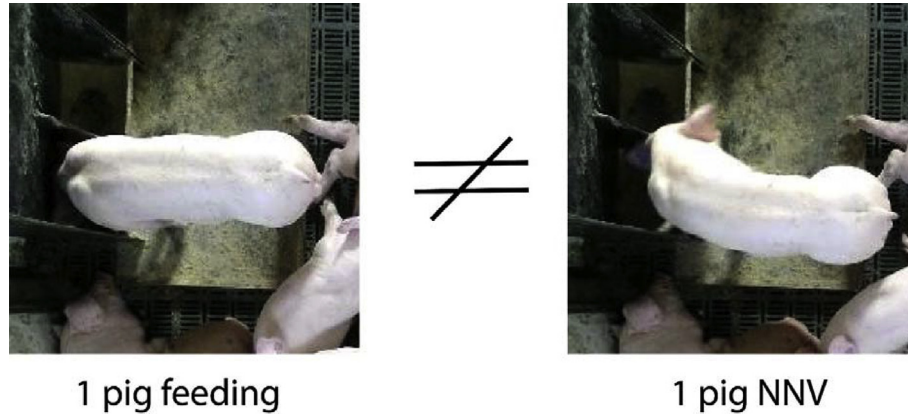


Fig. 1 – An example that illustrate the difference between feeding and non-nutritive visit (NNV) behaviours and how our proposed system was developed to tackle this problem.

2. Materials and methods

2.1. Animals and experimental design

All of the animal work was approved by the Animal Welfare and Ethical Review Body (AWERB) of Newcastle University. For this study, 15 pigs (Landrace/Large White synthetic sire line, Hermitage Seaborough Ltd., North Tawton, UK) were housed in a single fully-slatted pen (4 m × 2.4 m) from 9 to 14 weeks of age (mass range 33.6–51.0 kg at the start of the trial). This stocking density is representative of UK commercial conditions (0.67 m² per pig). Within the pen, water and food were available from four nipple drinkers and four feeding troughs (two of which were fully covered by our camera) with a black rubber mat (1 m × 0.4 m) covering the floor directly in front of the troughs. The design of the troughs allowed one pig of this age range to feed from a trough at any one time. The black mat area was designated as the feeding zone. A hanging chain with plastic pipes was also provided to meet commercial enrichment requirements. Every morning, any food remaining in the food troughs was removed, weighed and replaced with a known quantity of new food at approximately 09:30. All pigs were individually numbered using ear tags and had been previously vaccinated against pneumonia at 7 and 28 days of age (1 mL M + PAC, MSD Animal Health., Milton Keynes, UK), post-weaning multi systemic wasting syndrome at 28 days (1mL CircoFLEX, Boehringer Ingelheim GmbH, Ingelheim, Germany), and Glasser's disease when 9–10 weeks old (2mL Porcilis Glasser vaccine, MSD Animal Health., Milton Keynes, UK). During the study, the mean ambient temperature was 26.3 °C (range: 21.9–28.3 °C) and the relative humidity varied from 41 to 54% (mean: 47%).

Throughout the study, the pigs had free and continuous access to a commercial food suitable for this age and mass. However, during week 12 of age (approximately halfway through the experiment) the pigs were quantitatively food restricted receiving 80% of their daily *ad-libitum* food for 4 consecutive days. Water was available *ad-libitum* at all times.

2.2. Equipment set-up and behavioural observations

The full floor area of the pen was captured with two RGB cameras (Microsoft Kinect for Xbox One, Microsoft, Redmond, Washington, USA) attached to the ceiling within ingress protected enclosures and positioned perpendicularly to the pen floor (as described by Miller et al. (2019)). Videos of the pig behaviour were recorded at 25 frames s⁻¹ (FPS) with image frame width of 640 pixels and frame height of 360 pixels. Using the sampled frames of our video recordings, manual annotations of feeding and non-nutritive visits (NNV) to the feeding area were made by a single, highly-trained observer. We used scan sampling of daily activity for 10 min at the start of each 1/2 h from 06:00–11:40. As the observations focused on feeding-related behaviours, only the video footage around the feeding area, i.e., two of the food troughs covered by the camera and the feeding zone immediately in front of the troughs, and to the side of the outermost feeding trough, was used for behavioural analysis and the remaining pen area was excluded by an image size reduction factor of 4.6. Following Miller et al. (2019), a pig was considered to be feeding when it had its head inside a food trough. A NNV was scored when a pig entered the feeding area (i.e., on the black mat or side of outermost food trough) with two feet (one of which was a front foot) without ever consuming any food. Following the model validation, we calculated the feeding index and the NNV index as:

$$F_i = \frac{\sum_{k=1}^N FPF_k}{N} \quad (1)$$

$$NNV_i = \frac{\sum_{k=1}^N NPF_k}{N} \quad (2)$$

In Equations (1) and (2), F_i and NNV_i refer to the feeding and NNV indices, respectively. N is the total number of frames in a video segment, while FPF_k and NPF_k are the number of pigs feeding and NNVs at the k th frame, respectively. We obtained the indices for feeding and NNV to ensure that we have consistent measures across various data frames, e.g., dropping frames throughout recording. Indices of feeding and NNV were scored between 06:00–12:00 on the day immediately

prior to the period of restriction, the four days of food restriction, and the day immediately following the restriction period.

2.3. Algorithm for feeding behaviour recognition

Our method for detecting the feeding and NNV behaviours used a single deep learning network, based on the GoogLeNet architecture (Szegedy et al., 2015), which operated on a grayscale version of the images. We evaluated two variants of the architecture: one trained from scratch with a single channel, called Sc-GoogLeNet, and one architecture pre-trained on ImageNet and then adapted to work with grayscale images, called GoogLeNet in our experiments. We also compared the results of the above models using the manually annotated RGB images instead of grayscale ones.

2.3.1. Dataset

In order to build a robust system that generalises to a diverse farm setting (e.g., pigs with different body features or sizes),

we selected varied examples of pigs exhibiting feeding and NNV behaviour. We included images of pigs on top of each other and images with reduced quality due to direct exposure to sunlight or insects partially occluding the camera lenses.

Sample frames were selected from our database of video sequences to construct a data set for training, validation and testing. Our dataset comprised a total of 34375 images, divided into seven categories, where the number given (1 or 2) represents the number of pigs performing the listed behaviour. The behaviour classes were: 1 Pig Feeding (2270 images), 1 Feeding 1 NNV (378 images), 1 NNV (230 images), 2 Feeding (27736 images), 2 Feeding 1 NNV (933 images), 2 NNV (2688 images) and None (where none of the above scenarios occurred, 140 images). Figure 2 shows two examples from each class to reflect the richness of this dataset. It consists of a variety of pig postures, such as lying, standing, bowing, looking up, pigs standing on each other and pigs in direct contact with one another with different illumination conditions. As a result, the dataset used in this work can be considered diverse and representative of a commercial pig pen. This described

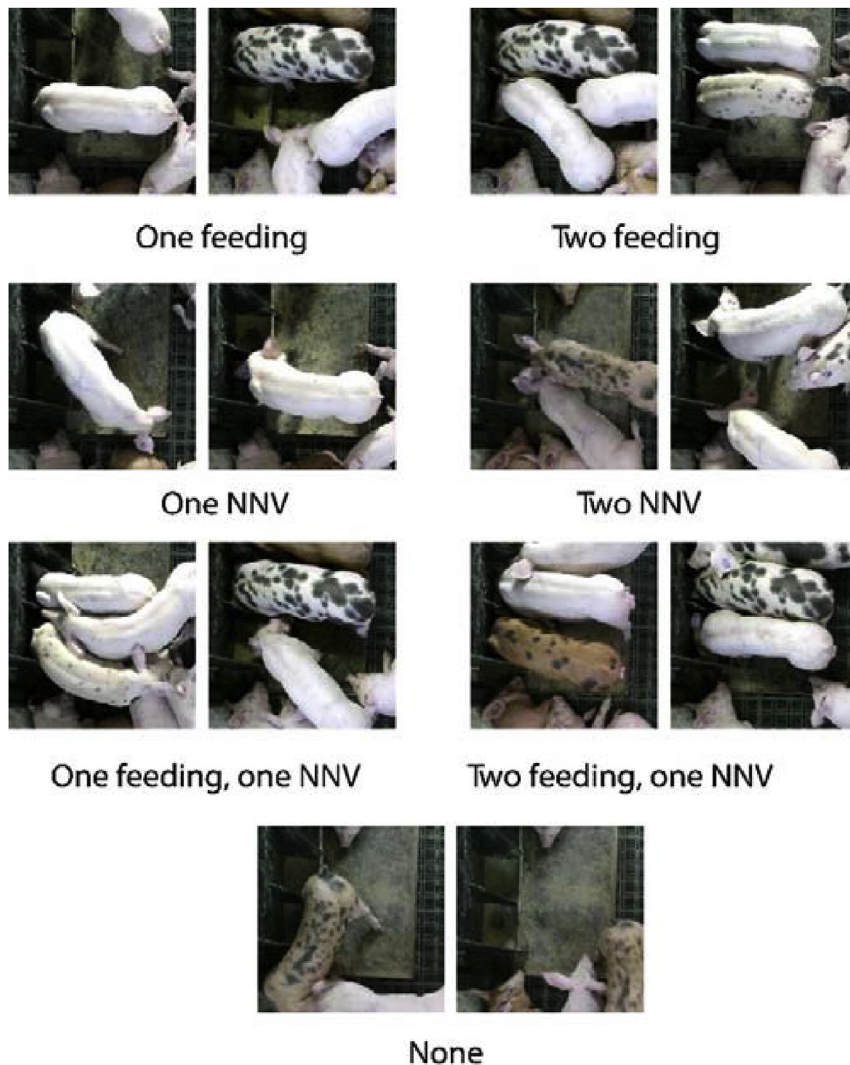


Fig. 2 – Examples of the image classes of pig behaviour in our dataset. The behaviours of interest were classified as feeding (i.e., a pig head in the trough), a pig performing non-nutritive visit (NNV) (i.e., a pig had two of its legs, including one front leg on the area defined by the mat under the trough) or none (i.e., a pig was not feeding or performing a NNV).

dataset was sampled from our food restriction trial in late spring/early summer and was only used for training and cross-validation purposes. In addition to the above dataset, we annotated a further dataset with a total of 7496 images for testing; this dataset was randomly sampled from the same study, however, at different dates to test for model generalisation to unseen dates of the same trial.

Furthermore, we collected and annotated two further datasets for validation purposes with the following characteristics:

- Data captured from two other commercial pig trials that were carried out at different time periods during the year, i.e., winter and early spring, resulting in changes in natural light between datasets. Variations in the data also include trough positioning within the pen and pig sizes i.e., mean mass of pigs (kg): 31.89 and 34.17. The total number of images in this collated dataset was 463; behaviour classes were: 1 pig Feeding (47 images), 1 Feeding 1 NNV (92 images), 1 NNV (10 images), 2 Feeding (267 images), 2 Feeding 1 NNV (15 images), 2 NNV (7 images) and None (where none of the above scenarios occurred, 25 images).
- A manually customised dataset with high exposure to sunlight was also used. Images were sampled from random days in the afternoon when the sunlight was illuminating the feeding area. The total number of images in this collated dataset was 444; behaviour classes were: 1 pig Feeding (128 images), 1 Feeding 1 NNV (3 images), 1 NNV (261 images), 2 Feeding (14 images), 2 NNV (16 images) and None (where none of the above scenarios occurred, 22 images).

2.3.2. System architecture

We used the pipeline in Fig. 3 to train and validate our system. To test our hypothesis that suggests the redundancy or negative impact of the colour (RGB) channels for detecting feeding postures of pigs (i.e., diverting the network attention to pig colours rather than feeding postures), we converted all input images into a single channel of grayscale representation. Consequently, we redesigned a network with a similar architecture to GoogLeNet (Szegedy et al., 2015), however with a single input channel, referred to as Sc-GoogLeNet, alongside the traditional GoogLeNet architecture with three input channels. In order to use the three-channel architecture we simply replicated the data of the grayscale channel three times. This apparent redundancy has an advantage: we can leverage a transfer learning (TL) strategy by using a network that has been pre-trained on the large ImageNet database (Deng et al., 2009). As the experiments will show, this strategy performs better than Sc-GoogLeNet or the traditional GoogLeNet architecture fed with RGB data. Similarly, we did not apply any augmentation to the input data due to the sensitivity of this task to common image transformations, such as reflection, rotation, scaling, translation and shearing. The rationale for selecting the GoogLeNet architecture is because its small size, compared to other standard CNN architectures, is translated into shorter prediction times, enabling our method to be close to real-time in prediction speed: the prediction time for processing a single image is 0.019–0.021 s.

Furthermore, this network architecture achieves high accuracies and low error rate on well-known datasets in machine vision, such as ImageNet dataset (Deng et al., 2009). The network depth, defined as the largest number of sequential convolutional or fully connected layers on a path from the

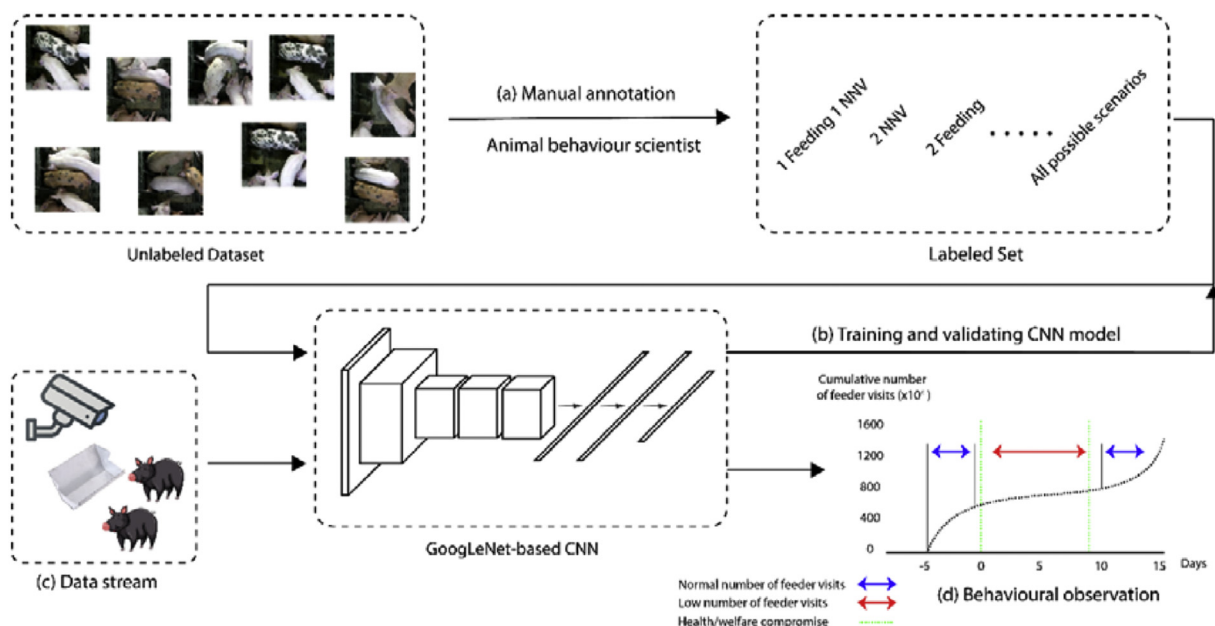


Fig. 3 – Architecture proposed for the automatic recognition of feeding and non-nutritive feeding behaviours in pigs. The CNN model was (a, b) trained/validated with the manual annotation of feeding and non-nutritive visit behaviours. It was then (c) utilised to detect changes in said behaviours from a data stream in days of control and food restriction. (d) Example graph on the behavioural observation that shows the cumulative number of feeder visits in periods of health/welfare compromise.

input layer to the output layer, is 22 layers with around 7 million parameters. We selected the hyper-parameters (*e.g.*, solver, learning rate schedule settings, batch size, and the maximum number of epochs) for training the network using nested cross-validation. Finally, a softmax layer was utilised to perform the final classification predictions.

The architecture of the GoogLeNet model consists of convolutional layers, max-pooling layers, relu layers, cross channel normalisation layers, dropout layers and a fully-connected layer. It also incorporates inception modules, see Fig. 4, which create a more in-depth network without serially stacking more layers. At the inception modules, varied sizes of convolutional filters were implemented to capture features with different levels of abstraction. Our network design utilises nine inception modules. Filters with larger size extract high-level features, while those with lower size extract features at a lower level. For example, the 11 convolution filters at the first stage of the module activate to correlated features in the same region. It is also used for dimensionality reduction where it can efficiently control the depth of the input features. Conversely, the 33 and 55 convolutions activate to more sophisticated features. Finally, the output is formed by concatenating the feature maps from all convolutions using the depth concatenation layer (Szegedy et al., 2015).

The model was implemented in Matlab R2019a on core i7 processor (2.5 GHz) PC using 16 G RAM and NVIDIA GeForce GTX 970 M Graphical processing unit (GPU).

2.3.3. Training and evaluation procedure

To validate our model performance, we used stratified 10-fold cross-validation. This means each fold has approximately the

same class distribution as in the whole set. This standard technique for validating the model performance produces accurate estimations for the generalisation to independent datasets (Kohavi, 1995; Alameer et al., 2015). The image dataset was randomly partitioned into ten equal-sized subsamples with similar class distribution as the whole set. Of the 10 subsamples, one subsample was used as the validation data for testing the model, and the remaining 9 subsamples were used as training data. The cross-validation process was then repeated 10 times, with each of the 10 subsamples used only once as a validation data. All models were trained and evaluated using the same data partitions.

In addition to the above cross-validation, we tested the model performance using sampled data frames from three different days of the trial that had not been used for training. We selected key dates where we predicted a change in feeding-associated behaviours to be present: *e.g.* baseline (standard feeding pattern for the animals is shown) vs. day 4 of food restriction (predicted change in feeding-associated behaviours due to known limited availability of food and therefore competition for resources) vs. day 1 of return to *ad-libitum* feeding (predicted change in feeding-associated behaviours as pigs can freely access food following a period of restriction and there is no longer competition for resources). This would demonstrate how the model generalises to different scenarios linked with *e.g.*, crowding at the feeder during food restriction when there is competition for resources.

Finally, we validated our primary model performance against other challenging conditions using the two customised datasets captured in different time periods of the year and with high exposure to sunlight.

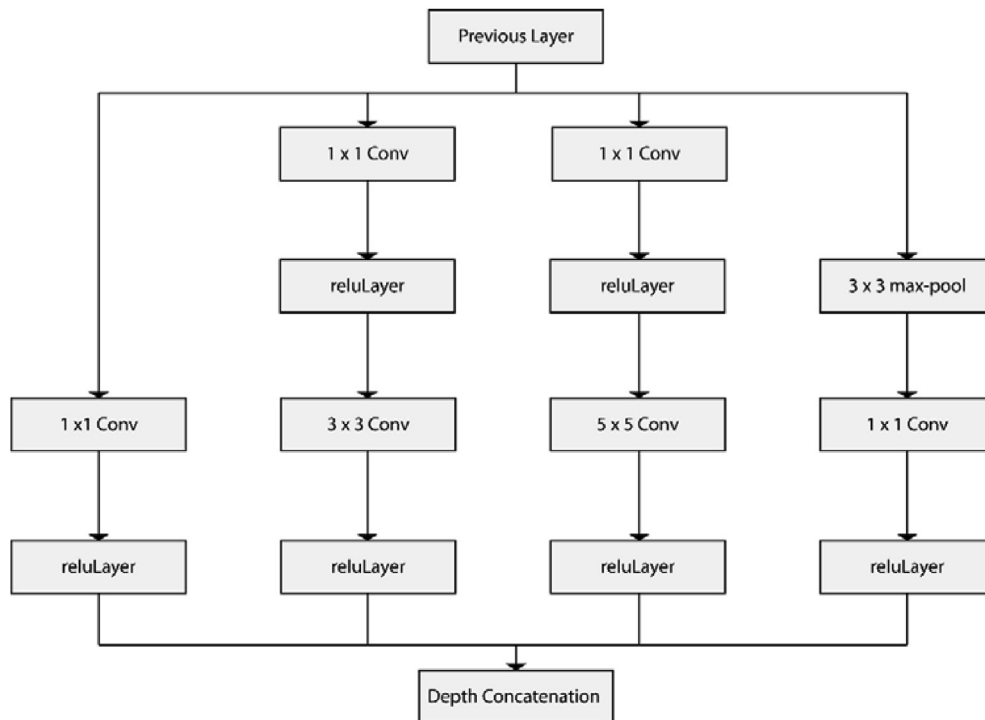


Fig. 4 – The structure of the inception modules used in this work. The above combination of convolutional filters, max-pool and relu layers encourages the network to capture features with different level of abstraction. A total of nine inception modules were used in all architectures provided in this work.

2.3.4. Visualisation methods

To understand how our trained model misclassified images, we examined the raw activations of higher layers of our model. We then inspected which features the network learned by comparing areas of activation with the original misclassified image. We normalised the activations such that the minimum activation is 0 and the maximum is 1. We investigated the output of the “interesting channels” by programmatically examining only channels with maximum activations.

In addition to directly visualising raw feature maps, we generated occlusion sensitivity maps (Zeiler and Fergus, 2014) to gain a high-level understanding of the biases of the network toward certain classes. We perturbed small segments of the image by applying a square-shaped occluding mask. We then moved the mask across the whole image, and measured the change in probability score for a given class as a function of mask position. When an indicative feature (to a certain class) of the image is occluded, the probability score of that class falls accordingly.

Finally, we visualised the high-dimensional activations of our model using t-distributed stochastic neighbour embedding (t-SNE) (Maaten and Hinton, 2008). We used this technique to visualise how our model changes the representation of input data as it passes through the network layers. It preserves distances such that points near each other in the high-dimensional space are also near each other in the 2-dimensional proximity.

3. Results

Ad-libitum food intake of the pigs at 12 weeks of age ranged from 0.059 to 0.070 kg [feed] kg⁻¹ [initial total pen body mass]. In contrast, the restricted daily intake was 0.047 kg [feed] kg⁻¹ [initial total pen body mass]. No adverse effects on health were recorded at any point before, during or after the food restriction protocol. Table 1 shows the parameters used to train our model. After training, our selected model accurately reported the number of pigs that exhibited feeding and NNV behaviour.

3.1. Behaviour observations

3.1.1. Feeding behaviour

We used our model to inspect the feeding and NNV behaviour during normal, baseline conditions and a planned period of

food restriction. Figure 5 shows the feeding index across the baseline day and across the 4th day of food restriction at an hourly level from 06:00–12:00. Between 06:00 and 07:00, pigs spend very low amounts of time eating on both baseline and food restriction days. However, across the rest of the morning, the pigs spend significantly more time feeding on the food restricted day than during the non-restricted day, with an increase shown after 9.30am. This coincides with when the pigs were provided with their allocated amount of food for the day. Feeding at this time would be a priority for the animals as they would anticipate the food would not be present in sufficient quantities to meet their needs later on the day, and thus the pigs are competing for resources at this time point. During baseline days, the food was topped up at exactly the same time of day, but as the food was freely available at all times, only limited feeding behaviour was observed during this specific time frame.

Figure 6a shows the calculated feeding index per day across the study period. Following the initial food restriction, the feeding index increased across the 4 day test period. A Wilcoxon Signed-Ranks test showed that the feeding index across day 4 of the food restriction period was significantly higher than at baseline ($p = 0.013$). This shows that despite less food being available, the pigs were spending an increasing amount of time with their heads in the food trough. This is probably due to the pigs checking the troughs thoroughly to see if food has been replenished when they are feeling hungry. The pigs may also be spending an increasing amount of time ensuring any small remaining amounts of food have been removed from the back and sides of the food trough. When feeding returned to *ad-libitum*, an immediate decrease in feeding index was seen ($p = 0.5186$) as the pigs were easily able to consume the full amount of food they required and thus had no need to make repeat trips to the food trough to check for further food availability. However, this decrease was relatively small, as the pigs consumed more food than on the first three restricted days. This could be in anticipation of resources again becoming limited.

3.1.2. Non-nutritive visits (NNV) behaviour

Figure 6b shows the calculated NNV index per day across the study period. Immediately following food restriction an increase can be seen in the duration of time spent performing NNV, as the pigs enter the feeding area and are unable to feed due to limited resource availability. Over the following three days when the food continued to be restricted, the duration spent performing NNVs decreased as the pigs learn that once the food has been consumed, no more will be made available until the following morning. This was supported by a Wilcoxon Signed-Ranks test that indicated that the NNV index during day 4 of the restricted food test period was significantly lower than that of the baseline ($p = 0.034$). When *ad-libitum* feeding was restored, the duration spent performing NNVs was equivalent to that of baseline period as this behaviour returns to control levels.

3.2. Behaviour-monitoring validation

Cross-validation was used to determine the prediction capacity of our automated feeding behaviour annotation

Table 1 – Parameter selection and their values used to train the proposed architectures. The learning rate was scheduled to decay by 0.05 every two epochs.

Parameter	Value
Solver	Adam optimizer
Initial Learn Rate	3.0 e–04
Learning Rate Schedule Settings	Drop by 0.05 every 2 epochs
L2 Regularization	1.0 e–04
Max number of epoch	10
Size of mini-batch	64
Input size	224 × 224 pixels

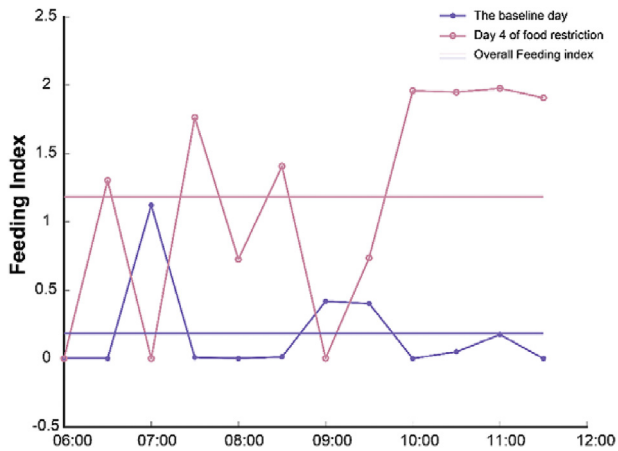


Fig. 5 – The feeding index per hour during the baseline day, when the pigs were fed *ad-libitum* and during the 4th day of food restriction. In the latter case the pigs were provided with their food allowance at 9.30. The higher the value of the feeding index the more time spent feeding. The overall feeding indices were 0.18 and 1.15 for the baseline and the 4th day of food restriction, respectively.

system. This validation showed that our models recognised the feeding and NNV behaviours of pigs with an accuracy of 99.4% (for pre-trained GoogLeNet) and 98.7% (for Sc-GoogLeNet) using stratified 10-folds cross-validation. With similar experimental settings, lower accuracies were scored using the RGB data: 99.2% with the pre-trained GoogLeNet and 96.46% with the non pre-trained GoogLeNet.

The performance of the pre-trained architecture with grey-scale imaging, with the highest accuracies, is described in further detail using the confusion matrix in Fig. 7, which shows the accumulative information of the actual and

predicted classifications. The average model accuracy in recognising the feeding behaviour was 99.5%, while the average accuracy in identifying NNV was 99.4%.

Interestingly, our proposed systems exhibited logical biases between certain classes. For example, the system misclassified “2 Feeding” in favour of “1 Feeding 1 NNV” or “2 Feeding 1 NNV”, each on five occasions. We also observed similar bias between “2 NNV” class and “None” class. Visually discriminating between these classes is challenging even for humans (*e.g.*, due to the head of the pig obscuring the front feet from some angles) and manual annotation of these examples required more attention. Additionally, the confusion matrix showed that the “2 Feeding” class comprised more images than other classes. This pattern would be expected to occur more often given there were more pigs in the pen than available feeding spaces and the feeder design allowing a maximum of one pig/feeder.

For each frame, we produced a label that described the current feeding-associated behaviours, and the predicted scores that reflected the system confidence in making decisions. Figure 8 shows the class scores produced in two scenarios taken from two different groups. In the first example, the model was fully confident of the feeding status. It predicted a maximum score of 1. In the second example, however, the model was less confident. It predicted scores of 0.67 and 0.33 for the classes “2 Feeding 1 NNV” and “2 Feeding”, respectively. In either case, the estimation for the number of feeder pigs was correct, *i.e.*, 2 feeders. Interestingly, the third pig was on the verge of the feeding trough area with only one leg visible.

3.3. Performance evaluation on the test set

To further validate our model performance, we used an independent dataset consisting of 7496 images from our surveillance video sequence for testing. The images were

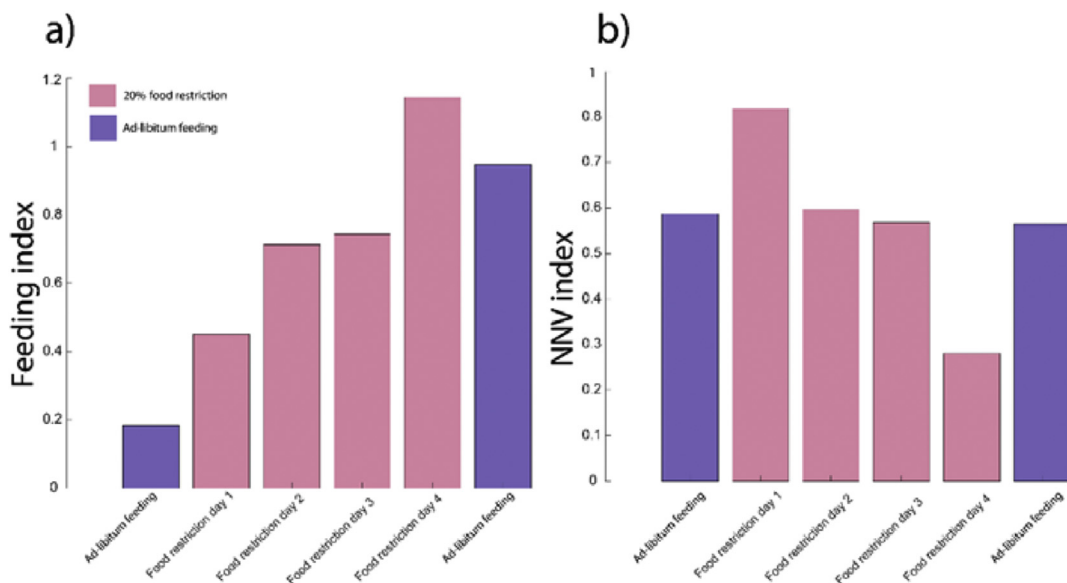


Fig. 6 – The calculated (a) feeding and (b) non-nutritive visit (NNV) behaviour indices per day across the study period. The *ad-libitum* feeding days (blue bars) correspond to the days immediately before and after the food restriction protocol period. The higher the Feeding and NNV index the more behaviour is shown during the day in question. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

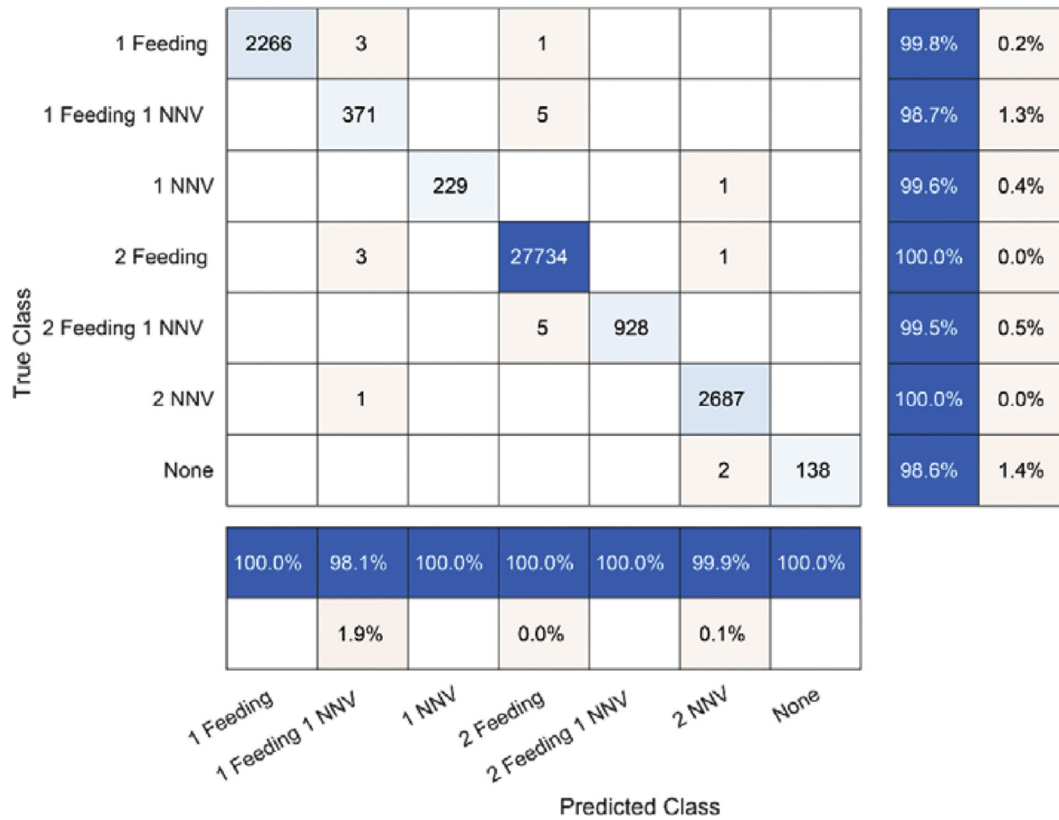


Fig. 7 – Confusion matrix chart. The true and predicted behaviours were classified as feeding (i.e., a pig head in the trough), a pig performing non-nutritive visit (NNV) (i.e., a pig had two of its legs, including one front leg on the area defined by the mat under the trough) or none (i.e., a pig was not feeding or performing a NNV). The dataset had a maximum of up to 15 pigs performing these behaviours.

sampled randomly from three alternative dates of the study period. Table 2 shows the average classification accuracy and standard deviation per class for these dates. The results were consistent for all dates tested. Once again, the model using the greyscale version of the data and a pre-trained GoogLeNet achieved the highest performance. Therefore, we utilised this model architecture as a primary model for all the experimental trials. Results from the primary model reflect its capacity in generalising to pigs of larger body sizes. For instance, the mean mass of the pigs was 41.56 ± 1.038 kg when we trained our model. It went up by 8% on day 2 of the food restriction period with apparent visual differences. Finally, the model showed consistent performance in classifying the two customised datasets for validation to (a) other batches with pigs of different sizes and (b) high exposure of sunlight, with an average classification accuracy of 97.1% ($\pm 1.98\%$) and 96.4% ($\pm 2.17\%$), respectively (Table 2).

3.4. Visualising activations

3.4.1. Raw feature maps

Each layer of our network architecture consists of many 2-D arrays called “channels”. Channels in the deeper layers had learned sophisticated features like the pig head, particularly when approaching the feeding trough. Here, we identified the location of the most prominent features to understand how

the network behaves under different circumstances, for instance, misclassification between particular classes, Fig. 9. The black pixels in this figure represent strong negative activations, while white pixels represent strong positive activations. We mapped pixel positions in the activation map such that it corresponds to the same position in the original image. The white pixels indicate that the channel is strongly activated at that position, for instance, at the pigs’ heads during feeding. To perform this visualisation method, we selected the activations in the second convolutional layer of the fourth inception module; with filter size of (5×5) . Empirical analysis suggested that the output of this layer was more informative than earlier and/or more advanced layers.

Figure 9 shows three examples of misclassified images. In Fig. 9a, the network misclassified the status of the pig exhibiting the NNV behaviour due to the excessive sunlight, therefore it was predicted as “2 Feeding”. Similar scenarios were observed in part b and c of Fig. 9, this time due to head and the body occlusion, respectively. These examples demonstrate how the network misclassified these rare scenarios of excessive light exposure and occlusion.

3.4.2. Occlusion sensitivity maps

In this experimental analysis, we applied artificial occlusions to further investigate our model behaviour. Figure 10 highlights the image regions with positive, or negative,



Fig. 8 – The labels and posterior probabilities for two image samples. The snapshot in (a) our system was 100% confident of the label “1 Feeding”, while in (b) it was 67% confident of the label “2 Feeding 1 NNV”.

contribution to the score for the “2 Feeding” class. Red areas indicate higher level of positive influence in the decision-making process. Occluding these areas negatively affected the model classification. Interestingly, both the occlusion sensitivity maps and raw-feature map equally indicate that pigs’ heads in the troughs provide the strongest evidence for identifying pigs’ feeding behaviour. On the other hand, occluding the blue areas of the image only increased the score for the “2 Feeding” class. This indicates that the blue areas of the map are evidence of different classes.

Despite the different functioning mechanism of the above visualisation techniques, the results indicate that our model is learning formative features to detect the feeding behaviour of pigs, thereby pulling its attention to pertinent spots of the image.

3.4.3. High-dimensional features

In Fig. 11, we visualised the high-dimensional activations of our model with t-SNE. Tight clusters in the t-SNE plot correspond to classes that the network classifies correctly. Activations from first layers, Fig. 11a, do not show apparent clustering by class as it does not contain semantic content. However, activations from deeper layers, Fig. 11b, clustered points more distinctly; in particular, the softmax layer shown in Fig. 11c. Interestingly, observations that are semantically similar, e.g., “1 NNV” and “2 NNV”, are near each other in the softmax activations space. This indicates that our model has formulated a high level of understanding of the feeding associated behaviours of pigs which is reflected in the t-SNE two-dimensional space.

4. Discussion

Overall, our paper makes five major contributions to the detection of feeding-associated behaviours in commercially housed pigs:

1. To enhance the speed and accuracy of existing methods, we reframed the task to directly infer behaviours from images, i.e., eliminating the detection and tracking stages.
2. We proposed a system that does not put particular emphasis on the pig head (e.g., tracking the pig head) to identify the feeding and NNV behaviour, as previous studies suggest that the detection of pig behaviours is not sustainable when the head is obscured (Psota et al., 2019; Yang et al., 2018).
3. We identified image components (RGB/grayscale) of most relevance and for the first time showed that images with one channel of grayscale are more effective in identifying animal feeding postures. Additionally, we identified an effective CNN-architecture to handle our data and we configured two variants of the GoogLeNet architecture to a) leverage transfer learning; b) train from scratch.
4. Using appropriate visualisation methods, we have shown how our deep learning architecture has learnt to capture patterns that are coherent from a domain perspective, being semantically relevant with the detection of feeding behaviour in pigs.

Table 2 – Average classification accuracy (%) \pm standard deviation (%) for the test data. TL (i.e., transfer learning) and no pre-training (i.e., training from scratch) correspond to the mechanisms used to train models. Our primary model with both TL and gray-scaling outperforms other proposed architectures by a large margin in all scenarios.

Time Point	Grey Scale		RGB Scale	
	TL	Scratch	TL	Scratch
Baseline	97.67 \pm 2.217	91.77 \pm 8.02	92.23 \pm 6.89	90.02 \pm 9.41
Day 4 food restriction	98.02 \pm 1.83	92.40 \pm 7.21	97.93 \pm 1.58	92.34 \pm 7.31
Day returned to <i>ad-libitum</i> feeding	97.15 \pm 3.35	92.32 \pm 6.00	94.20 \pm 8.65	86.25 \pm 18.12
Other batches	97.10 \pm 1.98	90.71 \pm 7.43	92.66 \pm 1.87	90.28 \pm 7.51
Sunlight exposure	96.40 \pm 2.17	93.47 \pm 3.85	94.82 \pm 2.93	90.77 \pm 4.16

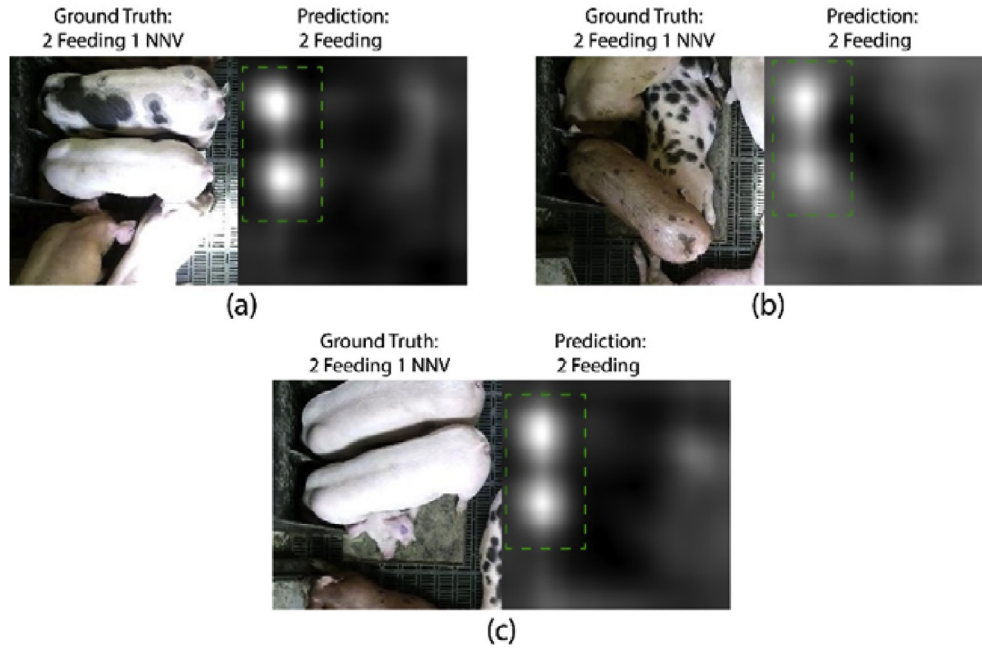


Fig. 9 – Visualising higher layer activation of misclassified images, due to (a) lighting conditions, (b) pig head occlusion and (c) pig full body occlusion. In all cases the ground truth annotation is 2 pigs feeding and 1 exhibiting non-nutritive visit (NNV) behaviour. In all cases the prediction is 2 pigs feeding.

- We demonstrated the capabilities of our method to detect changes in feeding-associated behaviours following a disruption to the pig feeding regime.

We showed that our method is robust enough to apply under a variety of conditions, as we applied it on a different batch that contained pigs of different sizes husbanded under different conditions, and under conditions of very different light intensities. When compared with currently available approaches used to detect and track pigs (Yang et al., 2018; Zhang et al., 2018), our system was faster (with a significant time reduction of ~95% per frame) and more robust to common challenges in commercial farm settings, such as the

short term tracking (i.e., losing track of a pig after a relatively short time) and alterations in farm conditions, e.g., lighting (Nasirahmadi et al., 2019b). To overcome the above challenges, we bypassed the tracking stage and directly inferred behaviours of pigs in the feeding area; we trained our system to generalise to a variety of conditions, e.g., lighting. Results show that our method provides sustainable and long-term segments of behaviour in “noisy” environments where pigs are more likely to be touching and frequently occluded by each other, overcoming problems associated with systems that rely on pig tracking to identify behaviours, e.g (Mittek et al., 2017). Furthermore, we tackled limitations associated with over-estimating the time spent feeding in pigs (Matthews

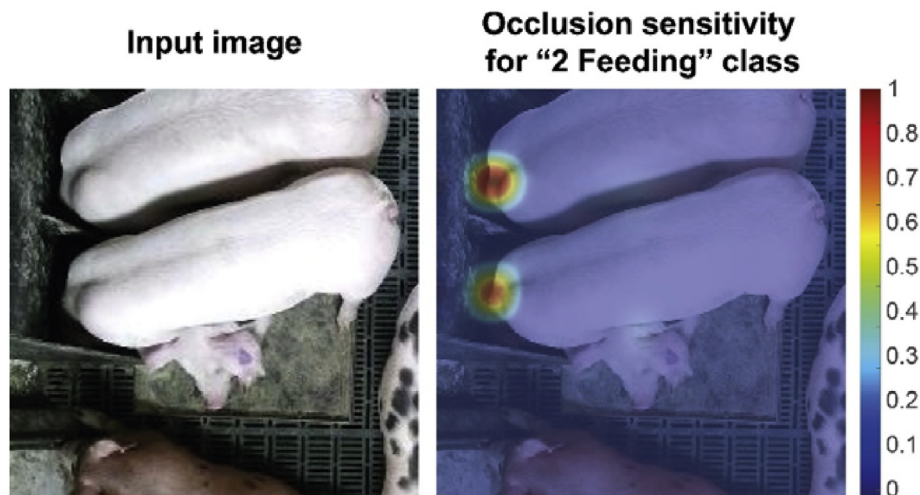


Fig. 10 – Occlusion sensitivity maps highlighting positive/negative areas in “2 Feeding” class. A minimum value of 0 corresponds to areas with a negative contribution, and a maximum value of 1 denotes areas with a positive contribution.

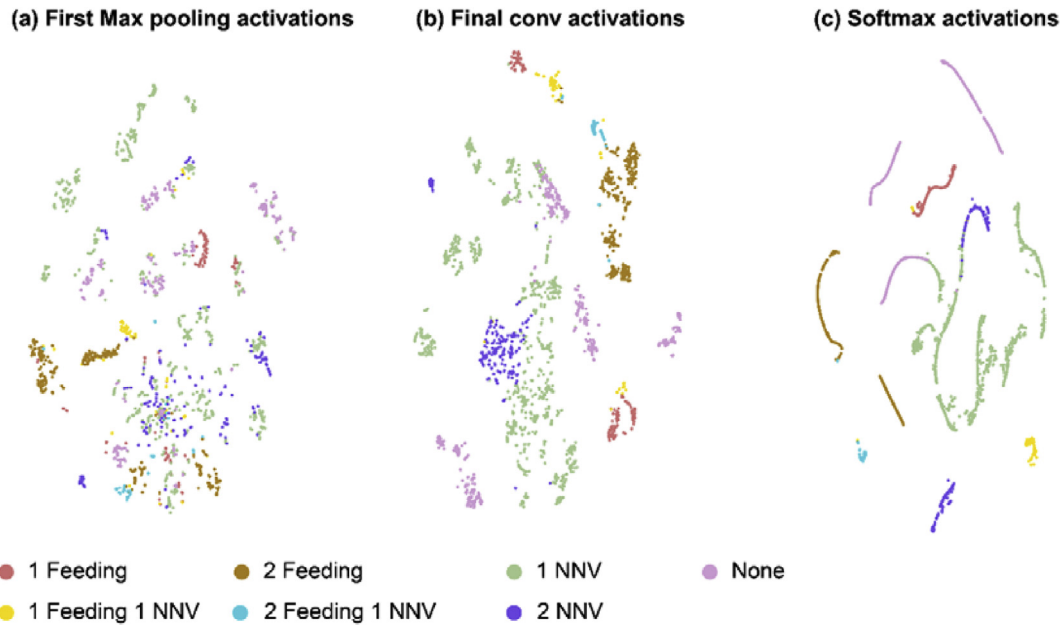


Fig. 11 – Representing network behaviour with t-SNE. Activations of the baseline dataset are visualised for (a) the first MAX pooling layer, (b) final convolutional layer and (c) the SoftMax layer. Each colour denotes a class of images; tight clusters of the same colour indicate correct classifications. The high-dimensional features of the network activations are mapped into two dimensions using t-SNE. This visualisation method allows tracking the network representation of input data while it sifts through the networks’ layers.

et al., 2017). The latter approach utilised the orientation and location of the pig to estimate feeding behaviour, therefore, it misclassifies scenarios where pigs are performing NNV behaviour (see Fig. 1). In contrast, our method performs classification on a frame-by-frame basis and therefore is capable of accurately distinguishing between feeding and NNV for a group of pigs.

The proposed system does not put particular emphasis on the pig head to identify behaviours, i.e., it does not require the location of the pig head to be detected in each frame sequence (Yang et al., 2018). Instead, our method detects the feeding-associated behaviours based on the whole structural features of the pig (e.g., while feeding) even when the head is entirely invisible (i.e., in the feeding trough). Moreover, an added benefit of the simpler system is efficiency: our method takes only 0.02 s on average to classify an image, which is about 2.5 x faster than the detection-based methods (Yang et al., 2018). Compared with segmentation-based approaches (Kashiha et al., 2013), our method handles situations where pigs are partially occluded or close to each other more efficiently. Our proposed method does not rely on segmenting pigs using its contour information that is sensitive to noise. It extracts high-level features of the entire pig posture to estimate their feeding status.

The developed system is capable of directly extracting the feeding associated behaviours of pigs without any post-processing stages, e.g., processing the trajectory of individual pigs. This mechanism allowed long-segments of behaviours to be obtained in real-time, i.e., 50 frames s^{-1} . In order to extrapolate this approach to encompass more pigs (+2 pigs feeding), we may define new classes of images, such as 3 feeding plus 1 NNV. Practically, this can be done by either utilising transfer learning (i.e., storing knowledge gained while

adding extra-classes), or by redefining the dataset and following the methods described in this paper. GoogLeNet architecture is capable of coping with an increased number of classes. The model has shown high performance in classifying a very large number of classes, e.g., 1000 classes in ImageNet dataset (Deng et al., 2009; Szegedy et al., 2015).

We trained/evaluated the model using a relatively large dataset of ~ 35000 annotated images. In comparison with other relevant methods in the field, where < 3000 images have been used (Nasirahmadi et al., 2019b; Psota et al., 2019; Yang et al., 2018), it is the largest annotated dataset by a big margin thus far. Using this dataset, we demonstrated that our proposed model can apply in a variety of conditions, such as fluctuations in natural lighting and pig body size. Furthermore, our proposed system does not require pigs to be individually marked (e.g., sprayed with numbers or tagged with an RFID) and therefore reduces the time and cost required for manual tasks. The introduced architectures, GoogLeNet and Sc-GoogLeNet, provide a trade-off between classification accuracy and network size. GoogLeNet provided superior performance by being able to leverage the knowledge captured by a network pretrained on ImageNet, despite being a more complex architecture with more weights to train. Finally, grey scaling the data has shown to be an effective pre-processing step leading to superior performance when compared to a network trained from the raw RGB data. Training our model with grey-scale data pulled the network attention exclusively to the pig feeding postures, rather than the colours and markings of individual pigs. Adopting a pre-trained GoogLeNet architecture to the grayscale data achieved the highest performances. This finding had led us to conclude that the colour channels may be redundant in identifying the feeding postures of pigs. The GoogLeNet architecture provided

optimum trade-off between classification accuracy (99.4%), speed (50 FPS) and size (21.8 MB). Faster network architectures, such as AlexNet (Krizhevsky et al., 2012) and SqueezeNet (Iandola et al., 2016), may compromise classification accuracy and/or network size.

The fast prediction time (0.02 s image⁻¹ or 50 frames s⁻¹; using a 2.5 GHz core i7 processor with NVIDIA GeForce GTX 970 M GPU) and a relatively simple architecture (a size of 21.8 MB with only 22 layers) of the developed GoogLeNet architecture, facilitates on-farm scale deployment. Practically, this can be done either: (a) embedding small PCs, e.g., Raspberry Pi, to each camera, with both devices housed in a protected enclosure, e.g., ingress (Matthews et al., 2017); or (b) by utilising a data capture infrastructure that sends the images from all deployed cameras to a centralised location (cloud computing or high-performance computing, e.g., core i9 processor (4.3 GHz) PC using (8 × 16) G RAM and NVIDIA GeForce RTX 2080 Ti GPU) where the method runs. The former approach may produce lower frame rate processing (< 50 frames s⁻¹) due to reduced capabilities of using GPU processing.

Quantifying feeding-associated behaviours is of great value for the early detection of compromises to the health and welfare of commercial pigs (González et al., 2008; Tolkamp et al., 2011). Distinguishing between feeding and NNV behaviours in such a quantification may have specific diagnostic value (Miller et al., 2019). The system developed here can distinguish between the feeding and NNV behaviours instantaneously in commercial stocking conditions without requiring knowledge on the previous locations of the pigs, and thus goes beyond previous work on the detection of feeding behaviour of livestock (Yang et al., 2018). Previous research has demonstrated the value of changes in NNV behaviour. Reduced NNV behaviour has been shown to be a sensitive indicator of declining health in transgenic mouse models of Alzheimer's and Huntington's disease (Codita et al., 2010; Oakeshott et al., 2011; Rudenko et al., 2009), and respiratory disease in calves (Svensson and Jensen, 2007). Therefore, changes in NNV behaviour may also have a value in the detection of health and welfare problems, over and above the changes in (consummatory) feeding behaviour. To our knowledge, no previous attempt has been made to detect the feeding and NNV behaviours of pigs directly from 2D images.

In this work, the black mat area was only used to identify the boundaries of the NNV area. This is relevant for both manual scoring by an animal behaviour scientist and also the automated method. If this system was to be implemented in another pen or on another farm, a simple indicator (e.g., spray paint) could be placed on the floor area to indicate the boundary of the NNV zone. Removing the mat would have no effect in quantifying the feeding behaviour of pigs.

We specifically designed our trial to provide a model akin to the early stages of a health/welfare compromise in a group of commercial pigs. When *ad-libitum* feeding stopped, the pen as a whole was provided with 80% of the food they would usually consume. We predicted that, at this level of food restriction, disruption to the behaviour would be present, but not at a level significant enough to result in overt, immediately identifiable changes in behaviour that would be seen pen side. Our study provided us with a data set that showed subtle changes in behaviour that would be of a similar

level to subtle behavioural changes in the early stage of health/welfare compromise (Kyriazakis and Tolkamp, 2010). Changes were detectable even when we monitored only a subset of the feeding troughs. Such changes are very difficult to detect by human visual inspection on large-scale commercial farms, thus warranting the development of this system that can monitor and detect such important changes in the patterns of feeding behaviour, without monitoring the entire pen.

5. Conclusions

Automation in animal husbandry is a tool that has the capability for capturing early changes in key behaviours that occur due to welfare and health compromise. Such changes are impractical to quantify manually and early detection, through automation, allows for timely intervention to prevent a further reduction in animal welfare and associated economic losses. This paper proposed a novel solution to resolve existing problems in automating the detection of feeding-associated behaviours in pigs. Using video surveillance, we have developed a method to automatically monitor and report the feeding and NNV behaviour of group-housed pigs under commercial settings. We demonstrated a novel automated system that can detect these subtle feeding behavioural changes with over 99.4% accuracy using only visual surveillance. The proposed method can operate in real-time processing up to 50 frames s⁻¹, and it does not require pigs to be fitted with sensors or individually marked. The paper provided a practical implementation for detecting the feeding behaviour of pigs using only video surveillance and suitable to be used in commercial settings, as it applied in a variety of husbandry and management conditions.

Acknowledgements

The animal trial was supported by the Biotechnology and Biological Sciences Research Council, UK, through an Agri-tech Research grant (BB/M011364/1) in conjunction with Zoetis Inc., Harbro Nutrition Ltd., Innovent UK Ltd., and RAFT solutions Ltd. This automated detection work was conducted under the Healthy Livestock project, which received funding from the European Commission under the European Union Framework Programme for Research and Innovation Horizon 2020 under grant agreement no: 773436; 633531. Jaume Bacardit is supported by the Engineering and Physical Science Research Council under grants EP/N031962/1 and EP/M020576/1.

REFERENCES

- Abolghasemi, V., Chen, M., Alameer, A., Ferdowsi, S., Chambers, J., & Nazarpour, K. (2018). Incoherent dictionary pair learning: Application to a novel open-source database of chinese numbers. *IEEE Signal Processing Letters*, 25, 472–476, 4.
- Alameer, A., Degenaar, P., & Nazarpour, K. (2020). Objects and scenes classification with selective use of central and

- peripheral image content. *Journal of Visual Communication and Image Representation*, 66, 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(8), 1062–1066.
- Alameer, A., Ghazaei, G., Degenaar, P., & Nazarpour, K. (2015). *2nd IET International Conference on Intelligent Signal Processing 2015 (ISP)*. IET.
- Ciaparrone, G., Sanchez, F. L., Tabik, S., Troiano, L., Tagliaferri, R., & Herrera, F. (2019). *Deep learning in video multi-object tracking: A survey*. Neurocomputing.
- Codita, A., Gumucio, A., Lannfelt, L., Gellerfors, P., Winblad, B., Mohammed, A. H., & Nilsson, L. N. (2010). Impaired behavior of female tg-ArcSwe app mice in the intellicage: A longitudinal study. *Behavioural Brain Research*, 215(1), 83–94.
- Cornou, C., Vinther, J., & Kristensen, A. R. (2008). Automatic detection of oestrus and health disorders using data from electronic sow feeders. *Livestock Science*, 118(3), 262–271.
- Day, J., Kyriazakis, I., & Lawrence, A. (1995). The effect of food deprivation on the expression of foraging and exploratory behaviour in the growing pig. *Applied Animal Behaviour Science*, 42(3), 193–206.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. *IEEE conference on computer vision and pattern recognition* (pp. 248–255). IEEE.
- González, L., Tolkamp, B., Coffey, M., Ferret, A., & Kyriazakis, I. (2008). Changes in feeding behavior as possible indicators for the automatic monitoring of health disorders in dairy cows. *Journal of Dairy Science*, 91(3), 1017–1028.
- Huang, W., Zhu, W., Ma, C., Guo, Y., & Chen, C. (2018). Identification of group-housed pigs based on gabor and local binary pattern features. *Biosystems Engineering*, 166, 90–100.
- Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). *Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size*. arXiv preprint arXiv:1602.07360.
- Kashiha, M., Bahr, C., Ott, S., Moons, C. P., Niewold, T. A., Ödberg, F. O., & Berckmans, D. (2013). Automatic identification of marked pigs in a pen using image pattern recognition. *Computers and Electronics in Agriculture*, 93, 111–120.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, pp. 1137–1145). Montreal: Canada.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* (pp. 1097–1105).
- Kyriazakis, I., & Tolkamp, B. (2010). Disease. In *The encyclopedia of applied animal behaviour and welfare* (pp. 176–177). Wallingford, Oxon: CAB International.
- Maaten, L.v. d., & Hinton, G. (2008). Visualizing data using tsne. *Journals of Machine Learning Research*, 9, 2579–2605.
- Marcon, M., Brossard, L., & Quiniou, N. (2015). Precision feeding based on individual daily body weight of group housed pigs with an automatic feeder developed to allow for restricting feed allowance. *Precision Livestock Farming*, 15, 592–601.
- 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(1), 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 groups behavior? *Journal of Animal Science*, 97(5), 1956–1966.
- Mittek, M., Psota, E. T., Carlson, J. D., Perez, L. C., Schmidt, T., & Mote, B. (2017). Tracking of group-housed pigs using multi-ellipsoid expectation maximisation. *IET Computer Vision*, 12(2), 121–128.
- Nasirahmadi, A., Sturm, B., Edwards, S., Jeppsson, K.-H., Olsson, A.-C., Müller, S., & Hensel, O. (2019a). Deep learning and machine vision approach for posture detection of individual pigs. *Sensors*, 19(17), 3738.
- Nasirahmadi, A., Sturm, B., Olsson, A.-C., Jeppsson, K.-H., Müller, S., Edwards, S., & Hensel, O. (2019b). 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.
- Ni, J.-Q., Liu, S., Radcliffe, J. S., & Vonderohe, C. (2017). Evaluation and characterisation of passive infrared detectors to monitor pig activities in an environmental research building. *Biosystems Engineering*, 158, 86–94.
- Oakeshott, S., Balci, F., Filippov, I., Murphy, C., Port, R., Connor, D., ... Brunner, D. (2011). Circadian abnormalities in motor activity in a bac transgenic mouse model of huntingtons disease. *PLoS currents*, 3.
- Pastorelli, H., Le Floc'h, N., Merlot, E., Meunier-Salaün, M., Van Milgen, J., & Montagne, L. (2012). Feed restriction applied after weaning has different effects on pig performance and health depending on the sanitary conditions. *Journal of Animal Science*, 90(13), 4866–4875.
- Psota, E. T., Mittek, M., Perez, L. C., Schmidt, T., & Mote, B. (2019). Multi-pig part detection and association with a fully-convolutional network. *Sensors*, 19(4), 852.
- Rudenko, O., Tkach, V., Berezin, V., & Bock, E. (2009). Detection of early behavioral markers of huntington's disease in r6/2 mice employing an automated social home cage. *Behavioural Brain Research*, 203(2), 188–199.
- Sa, J., Choi, Y., Lee, H., Chung, Y., Park, D., & Cho, J. (2019). Fast pig detection with a top-view camera under various illumination conditions. *Symmetry*, 11(2), 266.
- Svensson, C., & Jensen, M. B. (2007). Identification of diseased calves by use of data from automatic milk feeders. *Journal of Dairy Science*, 90(2), 994–997.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–9).
- Tolkamp, B., Allcroft, D., Barrio, J., Bley, T., Howie, J., Jacobsen, T., ... Kyriazakis, I. (2011). The temporal structure of feeding behavior. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 301(2), R378–R393.
- Weary, D., Huzzey, J., & Von Keyserlingk, M. (2009). Board-invited review: Using behavior to predict and identify ill health in animals. *Journal of Animal Science*, 87(2), 770–777.
- Yang, Q., Xiao, D., & Lin, S. (2018). Feeding behavior recognition for group-housed pigs with the faster r-cnn. *Computers and Electronics in Agriculture*, 155, 453–460.
- Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818–833). Springer.
- Zhang, L., Gray, H., Ye, X., Collins, L., & Allinson, N. (2018). *Automatic individual pig detection and tracking in surveillance videos*. arXiv preprint arXiv:1812.04901.
- Zhuang, X., & Zhang, T. (2019). Detection of sick broilers by digital image processing and deep learning. *Biosystems Engineering*, 179, 106–116.
- Zhu, X., Chen, C., Zheng, B., Yang, X., Gan, H., Zheng, C., Yang, A., Mao, L., & Xue, Y. (2020). Automatic recognition of lactating sow postures by refined two-stream rgb-d faster r-cnn. *Biosystems Engineering*, 189, 116–132.