Applied Animal Behaviour Science 263 (2023) 105923

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim

Validation of non-invasive sensor technologies to measure interaction with enrichment material in weaned fattening pigs

Fleur Veldkamp [a,b,*], Tomas Izquierdo Garcia-Faria [a], Vivian L. Witjes [c], Johanna M.J. Rebel [b], Ingrid C. de Jong [a]

[a] Wageningen Livestock Research, Wageningen University and Research, De Elst 1, NL-6708 WD Wageningen, the Netherlands
[b] Adaptation Physiology Group, Wageningen University and Research, De Elst 1, NL-6708 WD Wageningen, the Netherlands
[c] Farm Animal Health, Department Population Health Sciences, Veterinary Medicine, Utrecht University, Yalelaan 7, NL-3584 CL Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Pigs
Enrichment
Passive infra-red
Accelerometer
Algorithm
Behaviour

ABSTRACT

Measuring animal behaviour is important in the assessment of animal welfare. When interaction with the enrichment material (EM) can be measured, it can be used for detecting an increasing/decreasing interest in a certain EM. In this study, non-invasive sensor technologies were validated for measuring interaction with EM in pens with weaned fattening pigs. The experiment was carried out in four pens with six weaned pigs per pen (until a body weight of ± 25 kg) at a semi-commercial farm. Pens were provided with EM (ball and piece of wood (and a rope in two of the four pens) connected to a chain). Different sensor technologies were tested: passive infra-red detectors (PIDs), tri-axial accelerometers (TAA) and neural network model algorithms (NNMA 1 and NNMA 2) based on video images. Per pen, a PID was placed above the EM which measured the movement of body heat around the chain (⌀20 cm) in volts per second. A TAA was attached to the EM (at the top of the chain) and measured acceleration based on X-, Y- and Z-axis co-ordinates every second. A video camera was placed above each pen to record video images that were used to feed the NNMAs and for behavioural observations. Interaction with EM (shake, carry, nose, bite, chew or root) was manually scored per second per pig (pooled per pen afterwards) for 30 min of video footage per pen per week and was compared with data from PIDs, TAAs and NNMAs. F1 score (F1) and Matthews Correlation Coefficient (MCC) were calculated to measure the performance of the sensor technologies. PIDs (F1 = 0.380, MCC = 0.192), as well as TAAs (X-axis: F1 = 0.482, MCC = 0.345; Y-axis: F1 = 0.524, MCC = 0.401; Z-axis: F1 = 0.465, MCC = 0.320; XYZ-axis: F1 = 0.474, MCC = 0.333), overestimated interaction with the EM which might be due to the relatively small pen size, resulting in piglets touching the EM without intentional interaction with the EM. NNMAs achieved the highest performance parameters (NNMA 1: F1 = 0.554, MCC = 0.466; NNMA 2: F1 = 0.540, MCC = 0.445). Overall, only moderate F1s and MCCs were reached. The results indicated that the individual sensor technologies are not yet appropriate to measure interaction with the EM. However, there is potential to measure interaction with EM by applying a multi-sensor approach (combination of PID, TAA and NNMA), but this merits further study.

1. Introduction

High welfare standards and reliable methods to assess these standards are nowadays of great importance in the livestock sector (Alonso et al., 2020; Chapa et al., 2020). When high standards are not met and health and welfare are likely to be compromised, it should be detected at an early stage, so that timely interventions can be taken, negative welfare impacts can be reduced, and sustainable pig production can be promoted (Matthews et al., 2016). One way to detect the presence of these welfare compromises is by observing behavioural changes (Blackshaw, 1986; Matthews et al., 2016). For example, significant changes in activity were found in pigs after infection (Escobar et al., 2007; Reiner et al., 2009) and stress induction (Salak-Johnson et al., 2004). Besides the detection of health and welfare compromises, behavioural changes can be used as a prediction tool for impaired welfare (Matthews et al., 2016). Ursinus et al. (2014) and Zonderland et al. (2011) suggested that observing manipulative behaviour, such as chewing activity, directed at enrichment devices could be used as a tool

* Corresponding author at: Wageningen Livestock Research, Wageningen University and Research, De Elst 1, NL-6708 WD Wageningen, the Netherlands.
E-mail address: fleur1.veldkamp@wur.nl (F. Veldkamp).

https://doi.org/10.1016/j.applanim.2023.105923
Received 29 November 2022; Received in revised form 5 April 2023; Accepted 11 April 2023
Available online 13 April 2023

in the prediction of tail biting outbreaks. Observing interaction with enrichment materials can also be used for measuring an increasing or decreasing interest in certain enrichment materials. This is important from the perspective of ensuring that effective enrichment material is being used to combat animal welfare problems (Mkwanazi et al., 2019).

However, scoring of behaviour by using live observations requires direct observation by humans and only provides information at selected time points on a predefined sample size of animals (Czycholl et al., 2016; Pfeifer et al., 2019; Tuyttens et al., 2014). In addition, early signs of reduced welfare (e.g., visible as certain behavioural changes) may occur infrequently or may be very subtle (Matthews et al., 2017, 2016), complicating the detection of these signs with live observations alone (Wemelsfelder and Mullan, 2014). Although some of these drawbacks of live observations can be overcome by using video for behavioural observations, sensor technologies can be used to monitor behavioural indicators associated with changes in the animals' well-being in a more continuous, reproducible, and objective manner (Berckmans, 2014; Chapa et al., 2020; Larsen et al., 2021; Matthews et al., 2016).

Various non-invasive sensor technologies such as passive infra-red detectors (PID), tri-axial accelerometers (TAA) and neural network model algorithms (NNMA) have been applied to automatically monitor the behaviour of farm animals. Von Jasmund et al. (2020) used PIDs, that measure movement of body heat, focused on specific areas of the pen to record the activity of fattening pigs and stated that PIDs, combined with data from climate-related sensors, could serve as a monitoring tool for the early detection of behavioural changes. Chapa et al. (2020) reviewed TAAs, which record acceleration on three axes (X-, Y- and Z-axis), as a tool for health and welfare assessment in cattle and pigs. They stated that high sensitivities have been achieved in sows for lying (in ventral and lateral positions), walking, standing, posture changes, nest-building, exploratory behaviour and stepping behaviours (Chapa et al., 2020). However, only body-worn TAAs (i.e., ear-tag TAAs and neck collar TAAs) were used in these studies (Cornou et al., 2011; Oczak et al., 2015). Tzanidakis et al. (2021) stated that computer vision approaches (i.e., using artificial intelligence to train computers to capture and interpret information from image and video data) have been used in multiple livestock behaviour analyses and classification-related applications such as weight estimation of growing-finishing pigs (Kashiha et al., 2014a; Stygar et al., 2018), monitoring drinking behaviour (Kashiha et al., 2013a), automatic detection and counting the numbers of pigs (Tian et al., 2019), identifying resting behaviour of pigs (Kashiha et al., 2013b), detection of pigs' locomotion (Kashiha et al., 2014b), tracking of pigs' movement (Nasirahmadi et al., 2016) and automatic detection of aggressive behaviours (Chen et al., 2020b).

As described above, sensor technologies are able to measure different behavioural indicators. If manipulative behaviour directed towards the enrichment material (EM) can be monitored by sensor technologies, it can be used to measure an increasing or decreasing interest in certain EMs and in this way, possibly predict outbreaks of tail biting and determine preferences for and suitability of certain EMs. However, NNMAs, PIDs and TAAs have not been used previously for measuring interaction with EMs. Therefore, this study aimed to validate NNMAs, PIDs and TAAs for measuring interaction with EMs in pens with weaned fattening pigs by comparing sensor technologies with manual behavioural observations (gold standard). Different types of enrichment materials (with or without a rope attached) were used to validate the sensor technologies for different situations. Next to this, the performance parameters of the sensor technologies were given by time (i.e., age) and pens to validate the sensor technologies and compare the performance parameters for different situations.

2. Materials and methods

All data collection procedures were screened by the Animal Welfare Body of Wageningen Research, the Netherlands, and were indicated as non-invasive. Therefore, no ethical approval of the experiment was required.

2.1. Animals and housing

The study was carried out on a semi-commercial farm in the Netherlands from June 2021 until September 2021. Pigs (Topigs 20/TN70 × Tempo) were weaned around 28 days of age and relocated to the weaned pigs' compartment. Pens at the weaned pigs' compartment were 1.80 m × 1.45 m (2.61 m^2) with 40 % solid flooring (concrete with underfloor heating) and 60 % grid flooring (profiled steel triangular grid). Per production round (5.3 weeks per production round), four pens (pens 2, 3, 23 and 32) with six weaned pigs per pen were observed from weaning until they reached a body weight of ± 25 kg. The location of the pens used in this study was based on the aim of another study that was conducted at the same time. Pigs were subjected to standard management procedures on the farm. Within each pen, enrichment material (EM) was provided (a hanging metal chain with a ball and piece of wood connected to it). In two of the four pens, a rope was attached to the chain to validate the sensor technologies for different types of EMs.

Pigs had continuous access to one drinker and were fed ad libitum (commercial diet; weaner feed (week 1–2), rearing feed (week 3–5.3)). The ambient temperature was set at 28 °C at the time of weaning and was then gradually decreased to a fixed level of 23 °C at week 5.3. Underfloor heating was set at 36 °C at the time of weaning, decreased to 33 °C on day 3 after weaning and switched off on day 4 after weaning. The lighting regime was 10 h of light (from 7:00 h until 17:00 h) and 14 h of darkness, but daylight could enter through the side of the compartment (from the visitors' corridor).

2.2. Data collection

Data used in the analysis were collected from all four pens on Sundays from 14:00 h until 14:30 h (no persons present in the pens) during week one (of production round two), three and five (of production round one) after weaning (to compare different ages). Two production rounds were needed due to technical issues in week one of production round one. Therefore, we included a second production round to collect data for week one.

2.3. Behavioural observations

To be able to recognise the individual pigs on the video images, pigs were marked with a livestock colour spray two or three times a week (depending on the visibility of the colour spray) with an individual colour-symbol combination on their backs. The interaction with EM was scored by using the Observer software (version XT 14, Noldus Information Technology, Wageningen, The Netherlands). The ethogram for the behavioural observations is presented in Table 1. Behavioural observations were performed by one animal scientist trained in doing behavioural observations. All six pigs per pen were observed for 30 min by using instantaneous scan sampling with an interval of one second (1801 samples/pig/30 min), resulting in a total of 10806 samples per pen per 30 min. Samples of individual pigs were then pooled per pen afterwards because the sensor technologies measured at pen level (1801 samples/pen/30 min). The same observational data (i.e., gold standard) was used for all sensor technologies. Several combinations of behaviours were made after pooling to be more specific about the type of interaction with the EM (Table 2). To determine the performance of the sensor technologies in more detail, three different categories were made (Table 2), namely "Intentional interaction with enrichment material" (i.e., "shake", "carry", "nose", "bite", "chew", "root" and/or more than one type of these behaviours), "Interaction with enrichment material" (i.e., intentional interaction with enrichment material by one or more pigs in combination with behaviours "lie" and/or "body" by one or more pigs at the same time) and "All contact with enrichment material". By using these categories, it was possible to assess in more detail what the specific

Table 1
Ethogram used for behavioural observation for scoring the interaction with the enrichment material.

Behaviour	Description
Shake	While holding the object in its mouth, the animal energetically moves the enrichment material from side to side using its neck and head.
Carry	Animal securely holds the enrichment material in its mouth, while moving in a forward/backward/sideward direction.
Nose	Animal moves snout along or close to enrichment material, without holding the enrichment material in its mouth
Bite	Animal bites the enrichment material once, without keeping the enrichment material in its mouth.
Chew	Animal chews on the enrichment material, without moving the enrichment material forward/backward/sideward.
Root	Animal nudges or lifts enrichment material with movement of the snout, without keeping the enrichment material in its mouth.
Lie	Physical contact with the enrichment material other than mouth or snout (limbs, body, etc.) while lying down.
Body	Physical contact with the enrichment material other than mouth or snout (limbs, body, etc.) while standing/walking/running.
No action	No physical contact with the enrichment material.

Table 2
Criteria for additional behavioural categories after pooling of the manual observations of the interaction with the enrichment material, and interaction categories (Intentional interaction with the enrichment material, interaction with enrichment material, and all contact with enrichment material) defined to be more specific about the performance of the sensor technologies.

Behavioural categories after pooling	Criteria	Intentional interaction with enrichment material	Interaction with enrichment material	All contact with enrichment material
Shake, carry, nose, bite, chew, root and/or more than one type of the above behaviours	Only “Shake”, only “Carry”, only “Nose”, only “Bite”, only “Chew”, only “Root”, and/or more than one type of the above behaviours	X	X	X
Use of enrichment material plus lie	“Shake”, “Carry”, “Nose”, “Bite”, “Chew” and/or “Root” and “Lie”		X	X
Use of enrichment material plus body	“Shake”, “Carry”, “Nose”, “Bite”, “Chew” and/or “Root” and “Body”		X	X
Only lie	Only “Lie”			X
Only body	Only “Body”			X
Lie and body	“Lie” and “Body”			X
Only no action	Only “No action”			

sensor technology measured.

2.4. Sensor technologies

All four pens were equipped with a video camera, two passive infra-red detectors (PID), a tri-axial accelerometer (TAA) and a radio frequency identification (RFID) antenna (RFID data not discussed in this paper).

2.4.1. Video camera and neural network model algorithms

The four pens were equipped with one camera each (HIKVision, Hangzhou, China; Type DS-2CE16H5T-ITE 2.8 mm), attached to the ceiling (height = 3.00 m), capturing a birds'-eye view of the pen. The camera works with a resolution of 2556 × 1359 at a rate of 30 frames per second. Video images were recorded continuously and stored on a recorder (HIKVision, Hangzhou, China; Type DS-7204HUHI-K1/P). Video images were used for behavioural observations (gold standard) and training and feeding of the NNMAs. For the validation of the NNMAs, three datasets were used namely one for training (to make the models learn), one for validation (to validate how the models were learning) and one to measure the performance of the NNMAs (to compare with the gold standard). The dataset for the training of the NNMAs was obtained by stratified random sampling with age and pen as strata out of 24/7 video recordings, excluding periods used for comparing with the gold standard. For the development of the models, NNMA 1 used 1000 frames for training and 95 frames for validation whereas NNMA 2 used 1000 frames for training and 158 frames for validation. Both NNMAs were trained using a batch size (i.e., number of training examples used in one iteration) of 2 for 100 epochs (i.e., number of times the entire training dataset was fed to the NNMAs). For the annotation process, Computer Vision Annotation Tool (Sekachev et al., 2020) was used to create bounding boxes around the regions of interest.

The approach of NNMA 1 was to detect the pigs and the EM and to create a logic, based on intersection over union, on top of the predictions to determine whether the interaction with the EM could be considered active or passive. For this approach, the heads of the pigs (not the whole body) and the EM were detected by using an intersection over union-based logic (which indicates the degree of overlap between bounding boxes in one frame and another frame) to predict the activity of the pig (Fig. 1). The deterministic logic consists of packing the positions of the bounding boxes for two sequential frames and determining if the instances are moving and if they are touching. In the case of the EM, the distance moved from its average position, which was close to the position of the EM when resting, was calculated. A threshold for each of these four variables (intersection of instances, movement of the pig, movement of the EM and its position respective to the resting position) was set and used to determine the final prediction (Fig. 2).

The approach of NNMA 2 was to create bounding boxes for only the pigs (and not the EM). NNMA 2 was trained to detect the following behaviours: interaction with EM, standing, lying, drinking and eating (Fig. 1). For this study, the last four activities were interpreted as “No interaction” for simplicity.

The models used when comparing with the gold standard generated a binary value “Interaction” or “No interaction” per second. For this process, an Quadro RTX 5000 high-end graphics card with Max-Q Design was used along with a Xeon Inside Intel CPU. For the object detection algorithm, Mmdetection2 (Chen et al., 2019), a PyTorch (Paszke et al., 2019) based framework was used to compare the models Deformable-DETR (Zhu et al., 2020) and Tood (Feng et al., 2021). All the training was done using a 16 cores Ryzen 7 CPU and a single GPU Nvidia GeForce GTX 1080Ti with 11 Gb of RAM. Models took from four (Deformable-DETR) to twelve (Tood) hours of training.

2.4.2. Passive infra-red detector

Each pen was equipped with two PIDs (Technical Development Studio, Wageningen UR). One PID was focused on the EM (Fig. S1) whereas the other PID was focused on the whole pen (the latter not discussed in this paper). Plastic tubes were attached to the ceiling directly above the EM so that the PIDs could be attached to them to position the lens at the correct height. Tubes were also attached around the lens of the PID to ensure that the view on the ground had a diameter of 20 cm (radius of 10 cm around EM). The PIDs were based on a

Fig. 1. Screenshots of NNMA 1 with bounding boxes around the enrichment material and the heads of the pigs (right side of figure) and of NNMA 2 with bounding boxes around the pigs (left side of figure).

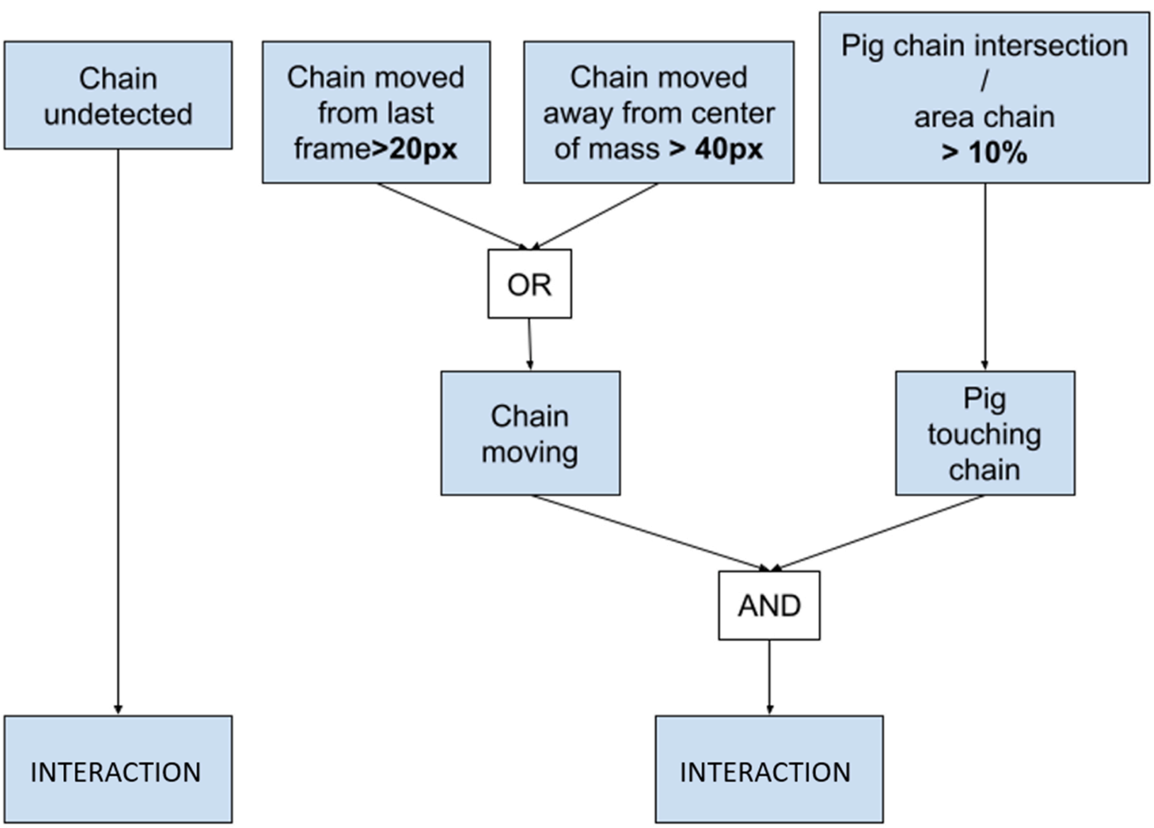

Fig. 2. Overview of the threshold values in the determination of interaction with the enrichment material. The EM was considered as moving when the EM moved 20px from the last frame OR when the EM was detected 40px away from the centre of mass (average position of the centre of the EM bounding box). A pig was considered as interacting with the EM when the area of the intersection between the pig bounding box and the EM bounding divided by the area of the EM was larger than 10 %.

Panasonic EKMB1301112K motion sensor. These sensors had a digital output of 3.3 V and detected the movement of body heat with an IR-sensitive semiconductor and a segmented lens. The electric pulsed output was then put into a low pass filter, resulting in a ratio output voltage range of 0–3.3 V. This signal was measured and recorded with a standard Lascar EL-USB-3 single-channel datalogger (Lascar Electronics Inc., USA; EasyLog EL-USB-3). Data was stored every second from 8:00 h until 17:00 h one day per week and was downloaded with EasyLog software (Lascar Electronics Inc., USA; EasyLog USB software version 7.7). The voltage (V) per second was merged with the pooled observational data per second. PID data was made binary, using ≤ 0.05 V as "no interaction" and > 0.05 V as "interaction". This threshold was based on the movement of body heat whereas ≤ 0.05 V did not indicate movement of body heat and > 0.05 V did indicate movement of body heat.

2.4.3. Tri-axial accelerometer

Each pen was equipped with a TAA (SOWNet Technologies, Pijnacker, the Netherlands) which was attached to the EM (at the top of the chain) (Fig. S2). The TAA was protected by a hard plastic box and then attached to the chain using tie wraps and duct tape to prevent damage by the pigs. The TAA measured acceleration based on X-, Y- and Z-components, with a sampling frequency of 1 Hz (one record per second), during a continuous period of 5.3 weeks. Data was stored on an SD card and exported to the PC after each production round. For each axis (X, Y and Z) and the average of the three axes (XYZ), stationary positions were determined. Minimum and maximum threshold values were determined for each X, Y and Z and for the XYZ to create ranges which were used to classify a value as "interaction" (outside range) or "no interaction" (inside range). Several minimum and maximum thresholds were tested during the validation. Following minimum (e.g., $X - 10$) and maximum (e.g., $X + 10$) thresholds resulted in the highest performance; $X - 10$, $X + 10$, $Y - 2$, $Y + 2$, $Z - 14$, $Z + 14$, $XYZ - 5$, $XYZ + 5$. Data were merged with the observational data by second.

2.5. Analysis

The dataset in this study was unbalanced (the number of samples in the "no interaction" class was larger than the number of samples in the "interaction" class). For this reason, F1 scores (F1), which perform well on issues involving imbalanced classification, and Matthews Correlation Coefficients (MCC), to overcome class imbalance issues, were used as performance metrics. The F1 is the harmonic mean of the precision and recall ($F1 = 2 * ((Precision * Recall) / (Precision + Recall))$) and can range from 0 to 1, with 1 representing a model that perfectly classifies each observation into the correct class and 0 representing a model that is unable to classify any observation into the correct class. High precision and low recall can result in the same F1 as low precision and high recall. For this reason, a comparison of the different F1s should be made with caution. MCCs give a correlation between predicted classes and ground truth ($MCC = (True\ Positives * True\ Negatives - False\ Positives * False\ Negatives) / \sqrt{((True\ Positives + False\ Positives)(True\ Positives + False\ Negatives)(True\ Negatives + False\ Positives)(True\ Negatives + False\ Negatives))}$) and can range from -1–1, with -1 representing total disagreement between predicted classes and actual classes, 0 representing prediction no better than random and 1 representing total agreement between predicted classes and actual classes.

F1s and MCCs were calculated for three categories ("Intentional interaction with enrichment material", "Interaction with enrichment material" and "All contact with enrichment material") (Table 2) by using R 4.1.1 (R Core Team, 2021), the caret (v6.0–92; Kuhn, 2022) and the mccr (v0.4.4; Iuchi, 2017) packages. Metrics were calculated for the complete dataset (all records) and for subcategories (per pen-week combination, per pen, per week, per presence/absence of an RFID antenna and per presence/absence of a rope attached to the EM) to validate the sensor technologies for different situations. In terms of sensor performance, the main interest was in the category "Intentional interaction with the enrichment material", since this category requires the best performance of the sensor technologies regarding the aim of this study.

3. Results

This study aimed to validate NNMAs, PIDs and TAAs for measuring interaction with the EM in pens with weaned fattening pigs by comparing sensor technologies with manual behavioural observations (gold standard). F1 scores and MCCs are presented for the NNMAs, PID and TAA (all axis) in Table 3–5. In the results section, we will focus on the performance of the sensors regarding the category "Intentional interaction with enrichment material" as this was the main focus of our study. The percentage of time piglets were performing "Intentional interaction with enrichment material" per pen and per week is presented in Fig. S3.

3.1. Neural network model algorithm 1

The overall performance of NNMA 1 was moderate for the categories "Intentional interaction with enrichment material" ($F1 = 0.554$) and "Interaction with enrichment material" ($F1 = 0.578$), but low for "All contact with enrichment material" ($F1 = 0.489$) (Table 3). The highest performance was achieved in pen 32 in week 5 ($F1 = 0.703$) for the category "Intentional interaction with enrichment material". In this category, NNMA 1 performed better in week 5 compared to weeks 3 and 1, respectively ($F1 = 0.604$ versus $F1 = 0.492$ and $F1 = 0.528$). When an RFID antenna was present above/around the EM, NNMA 1 had lower performance than in pens without an RFID antenna ($F1 = 0.517$ versus $F1 = 0.599$) in the category "Intentional interaction with the enrichment material". In the same category, NNMA 1 performed better in pens without a rope attached to the EM compared to pens with a rope ($F1 = 0.607$ versus $F1 = 0.500$).

3.2. Neural network model algorithm 2

The overall performance of NNMA 2 was lower than the overall performance of NNMA 1 for the categories "Intentional interaction with enrichment material" ($F1 = 0.540$ versus $F1 = 0.554$), "Interaction with enrichment material" ($F1 = 0.541$ versus $F1 = 0.578$) and "All contact with enrichment material" ($F1 = 0.451$ versus $F1 = 0.489$) (Table 3). In the category "Intentional interaction with the enrichment material", the highest performance was achieved for pen 3 in week 1 ($F1 = 0.718$). Contrary to NNMA 1, NNMA 2 showed the highest performance in weeks 1 and 3 versus week 5 ($F1 = 0.561$ and $F1 = 0.592$ versus $F1 = 0.481$) in the category "Intentional interaction with the enrichment material". In the pens where a rope was attached to the EM, NNMA 2 achieved better performance in the category "Intentional interaction with the enrichment material" than in pens without a rope ($F1 = 0.580$ versus $F1 = 0.496$) which is the opposite as for NNMA 1.

3.3. Passive infra-red detector

The overall performance of the PID was low for the categories "Intentional interaction with enrichment material" ($F1 = 0.380$) and "Interaction with enrichment material" ($F1 = 0.410$) but moderate for the category "All contact with enrichment material" ($F1 = 0.566$). For the category "Intentional interaction with the enrichment material", the highest performance was seen in pen 3 in week 1 ($F1 = 0.672$) (Table 4). Within the same category, it was found that PIDs performed better in pens with RFID antenna ($F1 = 0.450$) than in pens without RFID antenna ($F1 = 0.329$).

3.4. Tri-axial accelerometer

For the complete dataset, the highest performance for TAA in the category "Intentional interaction with enrichment material" was

Table 3
F1-score and Matthews Correlation Coefficient (MCC) for Neural Network Model Algorithm 1 and Neural Network Model Algorithm 2 for categories "Intentional interaction with enrichment material", "Interaction with enrichment material" and "All contact with enrichment material" for different subsets of data (P = pen (pen 2, 3, 23 and 32 respectively); W=week (week 1, 3, 5 respectively)). F1 scores and MCC > 0.5 are highlighted in bold.

	Intentional interaction with enrichment material				Interaction with enrichment material				All contact with enrichment material			
	F1-score		**MCC**		**F1-score**		**MCC**		**F1-score**		**MCC**	
	NNMA 1	**NNMA 2**	**NNMA 1**	**NNMA 2**	**NNMA 1**	**NNMA 2**	**NNMA 1**	**NNMA 2**	**NNMA 1**	**NNMA 2**	**NNMA 1**	**NNMA 2**
All records	**0.554**	**0.540**	0.466	0.445	**0.578**	**0.541**	0.496	0.445	0.489	0.451	0.421	0.348
P2W3	0.338	0.433	0.362	0.422	0.338	0.433	0.362	0.422	0.252	0.403	0.309	0.419
P3W3	**0.525**	**0.664**	0.394	**0.504**	**0.549**	**0.714**	0.437	**0.571**	0.457	**0.620**	0.339	0.425
P23W3	0.378	0.462	0.305	0.408	0.397	**0.524**	0.304	0.450	0.393	0.369	0.296	0.204
P32W3	**0.557**	**0.579**	**0.504**	**0.504**	**0.563**	**0.569**	**0.515**	0.486	0.393	0.472	0.358	0.360
P2W5	**0.617**	0.446	0.317	0.184	**0.675**	0.423	0.370	0.118	**0.654**	0.421	0.254	0.041
P3W5	0.338	**0.527**	0.243	0.445	0.361	0.467	0.284	0.384	0.235	0.303	0.124	0.188
P23W5	**0.687**	**0.556**	**0.540**	0.415	**0.720**	**0.553**	**0.578**	0.411	**0.643**	0.431	0.421	0.230
P32W5	**0.703**	-[a]	**0.689**	-0.014	**0.709**	-[a]	**0.703**	-0.015	0.326	0.009	0.386	0.014
P2W1	**0.541**	0.436	0.460	0.377	**0.536**	0.438	0.454	0.381	**0.505**	0.385	0.443	0.359
P3W1	0.306	**0.718**	0.277	**0.672**	0.309	**0.722**	0.282	**0.676**	0.305	**0.745**	0.282	**0.702**
P23W1	**0.666**	**0.599**	**0.564**	0.449	**0.697**	**0.614**	**0.604**	0.451	**0.599**	**0.594**	0.498	0.362
P32W1	0.337	0.395	0.332	0.336	0.363	0.407	0.364	0.349	0.345	0.409	0.391	0.357
P2	**0.588**	0.442	0.480	0.350	**0.630**	0.427	**0.530**	0.335	**0.602**	0.411	**0.521**	0.353
P3	0.430	**0.642**	0.346	**0.550**	0.445	**0.640**	0.377	**0.545**	0.342	**0.510**	0.278	0.395
P23	**0.625**	**0.556**	**0.517**	0.427	**0.652**	**0.573**	**0.541**	0.438	**0.566**	0.473	0.427	0.267
P32	**0.527**	0.456	**0.501**	0.404	**0.542**	0.450	**0.523**	0.396	0.362	0.346	0.383	0.304
W1	**0.528**	**0.561**	0.465	0.475	**0.547**	**0.571**	0.489	0.485	**0.501**	**0.543**	0.467	0.454
W3	0.492	**0.592**	0.421	**0.515**	**0.509**	**0.628**	0.445	**0.551**	0.416	**0.505**	0.378	0.408
W5	**0.604**	0.481	0.481	0.372	**0.632**	0.456	**0.505**	0.343	**0.524**	0.360	0.373	0.239
With antenna	**0.517**	**0.551**	0.415	0.455	**0.544**	**0.544**	0.452	0.449	0.473	0.466	0.398	0.378
Without antenna	**0.599**	**0.527**	**0.527**	0.438	**0.623**	**0.538**	**0.552**	0.445	**0.507**	0.434	0.445	0.318
With rope	**0.500**	**0.580**	0.404	0.486	**0.516**	**0.589**	0.424	0.493	0.433	0.463	0.353	0.345
Without rope	**0.607**	0.496	**0.527**	0.401	**0.641**	0.487	**0.568**	0.390	**0.550**	0.437	0.491	0.349

[a] Could not be computed due to the absence of true positives

Table 4
F1-score and Matthews Correlation Coefficient (MCC) for the Passive Infra-red Detector for categories "Intentional interaction with enrichment material", "Interaction with enrichment material" and "All contact with enrichment material" for different subsets of data (P = pen (pen 2, 3, 23 and 32 respectively); W=week (week 1, 3, 5 respectively)). F1 scores and MCC > 0.5 are highlighted in bold.

	Intentional interaction with enrichment material		Interaction with enrichment material		All contact with enrichment material	
	F1	**MCC**	**F1**	**MCC**	**F1**	**MCC**
All records	0.380	0.192	0.410	0.189	**0.566**	0.251
P2W3	0.232	0.273	0.232	0.273	0.382	0.391
P3W3	0.485	0.182	**0.501**	0.166	**0.586**	0.188
P23W3	0.207	0.119	0.265	0.144	**0.582**	0.118
P32W3	0.264	0.066	0.287	0.074	**0.520**	0.106
P2W5	0.498	-0.005	**0.516**	-0.116	**0.686**	-0.106
P3W5	0.321	0.140	0.410	0.147	0.476	0.086
P23W5	0.483	0.167	**0.513**	0.185	**0.665**	0.167
P32W5	0.247	0.267	0.284	0.292	0.485	0.303
P2W1	0.372	0.229	0.369	0.224	0.380	0.203
P3W1	**0.672**	**0.626**	**0.677**	**0.631**	**0.664**	**0.617**
P23W1	0.425	0.156	0.446	0.143	**0.584**	0.164
P32W1	0.304	0.229	0.316	0.241	0.447	0.329
P2	0.436	0.250	0.452	0.241	**0.585**	0.362
P3	0.464	0.295	0.495	0.298	**0.547**	0.318
P23	0.373	0.112	0.410	0.122	**0.612**	0.131
P32	0.268	0.206	0.293	0.224	0.497	0.293
W1	0.425	0.287	0.437	0.292	**0.520**	0.334
W3	0.305	0.131	0.337	0.134	**0.550**	0.220
W5	0.428	0.186	0.464	0.162	**0.604**	0.187
With antenna	0.450	0.272	0.268	0.268	**0.566**	0.337
Without antenna	0.329	0.165	0.361	0.179	**0.567**	0.227
With rope	0.393	0.208	0.431	0.211	**0.574**	0.261
Without rope	0.368	0.177	0.391	0.168	**0.558**	0.247

achieved on values of the Y-axis (F1 = 0.524) followed by the X-axis (F1 = 0.482), an average of X-, Y-, Z-axis (F1 = 0.474) and Z-axis (F1 = 0.465) (Table 5). For the category "Intentional interaction with the enrichment material", values of the X-axis performed best in pen 3 in week 1 (F1 = 0.710) whereas values of the Y-axis performed best in pen 23 in week 5 (F1 = 0.674) and values of the Z-axis and the average of X-, Y- and Z-axis in pen 2 in week 5, respectively (F1 = 0.603, F1 = 0.629). When no rope was attached to the EM, values of the X-, Y-, and Z-axis and the average of X-, Y-, and Z-axis, respectively, performed better than when there was a rope attached to the EM in the category "Intentional interaction with enrichment material" (F1 = 0.537 versus F1 = 0.436, F1 = 0.560 versus F1 = 0.487, F1 = 0.516 versus F1 = 0.416, F1 = 0.534 versus F1 = 0.415). When comparing the performance of the accelerometer in the different pens for the category "Intentional interaction with enrichment material", pen 32 had the lowest performance compared to pens 2, 3 and 23 on the X-axis (F1 = 0.394 versus F1 = 0.569, F1 = 0.462, F1 = 0.461), Y-axis (F1 = 0.437 versus F1 = 0.552, F1 = 0.479, F1 = 0.567), Z-axis (F1 = 0.361 versus F1 = 0.569, F1 = 0.407, F1 = 0.473) and the average of X-, Y-, and Z-axis (F1 = 0.357 versus F1 = 0.590, F1 = 0.422, F1 = 0.474). The TAA was best in the category "All contact with enrichment material" compared to "Intentional interaction with enrichment material" and "Interaction with enrichment material" on the X-axis (F1 = 0.718 versus F1 = 0.482 and F1 = 0.546), Y-axis (F1 = 0.676 versus F1 = 0.524 and F1 = 0.581), Z-axis (F1 = 0.709 versus F1 = 0.465 and F1 = 0.531) and the average of X-, Y- and Z-axis (F1 = 0.693 versus F1 = 0.474 and F1 = 0.531).

4. Discussion

In the present study, three non-invasive sensor technologies were compared with the gold standard for measuring interaction with the EM in pens with weaned fattening pigs. Results showed that within the main category of interest, "Intentional interaction with enrichment material", NNMAs performed best when compared to the gold standard, followed by TAA and PID, respectively. Overall, only moderate F1s and MCCs were reached. The results of this study indicated that the individual sensor technologies are not yet appropriate to record interaction with the EM. However, there is potential to measure interaction with the EM

Table 5
F1-score and Matthews Correlation Coefficient (MCC) for the x-axis (X), y-axis (Y), z-axis (Z) and the average of the x-, y- and z-axis (XYZ) of the tri-axial accelerometer for categories "Intentional interaction with enrichment material", "Interaction with enrichment material" and "All contact with enrichment material" for different subsets of data (P = pen (pen 2, 3, 23 and 32 respectively); W=week (week 1, 3, 5 respectively)). F1 scores and MCC > 0.5 are highlighted in bold.

	Intentional interaction with enrichment material								Interaction with enrichment material								All contact with enrichment material							
	F1				**MCC**				**F1**				**MCC**				**F1**				**MCC**			
	X	Y	Z	XYZ	X	Y	Z	XYZ	X	Y	Z	XYZ	X	Y	Z	XYZ	X	Y	Z	XYZ	X	Y	Z	XYZ
All records	0.482	**0.524**	0.465	0.474	0.345	0.401	0.320	0.333	**0.546**	**0.581**	**0.531**	**0.531**	0.397	0.447	0.376	0.375	**0.718**	**0.676**	**0.709**	**0.693**	**0.547**	**0.510**	**0.551**	**0.511**
P2W3	0.238	0.111	0.192	0.243	0.225	0.091	0.167	0.233	0.238	0.111	0.192	0.243	0.225	0.091	0.167	0.233	0.396	0.240	0.424	0.433	0.361	0.206	0.391	0.401
P3W3	**0.542**	**0.576**	**0.529**	0.496	0.229	0.349	0.276	0.193	**0.603**	**0.621**	**0.592**	**0.549**	0.276	0.394	0.346	0.237	**0.746**	**0.671**	**0.666**	**0.645**	0.367	0.447	0.435	0.297
P23W3	0.246	0.286	0.254	0.256	0.164	0.222	0.167	0.180	0.315	0.368	0.322	0.320	0.205	0.280	0.207	0.211	**0.645**	**0.616**	**0.645**	**0.651**	0.372	0.359	0.419	0.392
P32W3	0.419	0.417	0.431	0.449	0.306	0.302	0.328	0.356	0.433	0.437	0.458	0.470	0.307	0.313	0.345	0.361	**0.711**	**0.725**	**0.735**	**0.742**	**0.601**	**0.632**	**0.602**	**0.611**
P2W5	**0.648**	**0.635**	**0.603**	**0.629**	0.296	0.253	0.173	0.229	**0.735**	**0.712**	**0.699**	**0.717**	0.364	0.279	0.243	0.277	**0.793**	**0.816**	**0.836**	**0.850**	0.239	0.218	0.319	0.317
P3W5	0.293	0.356	0.279	0.302	0.059	0.184	0.036	0.091	0.482	**0.533**	0.465	0.466	0.166	0.293	0.145	0.172	**0.802**	**0.655**	**0.797**	**0.697**	0.422	0.301	0.476	0.343
P23W5	**0.524**	**0.674**	**0.565**	**0.532**	0.250	**0.513**	0.329	0.266	**0.551**	**0.702**	**0.597**	**0.562**	0.262	**0.537**	0.354	0.287	**0.775**	**0.722**	**0.794**	**0.776**	0.467	0.469	**0.559**	0.485
P32W5	0.435	**0.654**	0.316	0.273	0.429	**0.634**	0.321	0.288	0.464	**0.681**	0.329	0.288	0.442	**0.659**	0.310	0.277	0.376	0.380	0.386	0.355	0.269	0.403	0.216	0.142
P2W1	0.478	**0.533**	**0.558**	**0.572**	0.339	0.419	0.447	0.466	0.485	**0.540**	**0.559**	**0.575**	0.345	0.425	0.445	0.467	**0.662**	**0.676**	**0.699**	**0.699**	0.481	**0.504**	**0.558**	**0.559**
P3W1	**0.710**	0.498	0.443	0.494	**0.680**	0.413	0.357	0.411	**0.706**	0.495	0.453	**0.503**	**0.675**	0.409	0.368	0.421	**0.704**	**0.517**	0.467	**0.529**	**0.677**	0.430	0.382	0.448
P23W1	**0.554**	**0.659**	**0.520**	**0.576**	0.377	**0.535**	0.333	0.411	**0.597**	**0.711**	**0.567**	**0.620**	0.410	**0.585**	0.378	0.448	**0.761**	**0.782**	**0.726**	**0.738**	**0.576**	**0.650**	**0.568**	**0.543**
P32W1	0.293	0.336	0.231	0.259	0.227	0.273	0.172	0.190	0.313	0.344	0.241	0.274	0.249	0.280	0.182	0.205	0.344	0.324	0.257	0.319	0.287	0.258	0.207	0.255
P2	**0.569**	**0.552**	**0.569**	**0.590**	0.436	0.414	0.435	0.466	**0.632**	**0.606**	**0.636**	**0.652**	0.490	0.453	0.497	**0.521**	**0.733**	**0.728**	**0.779**	**0.787**	**0.567**	**0.536**	**0.648**	**0.653**
P3	0.462	0.479	0.407	0.422	0.278	0.310	0.203	0.224	**0.562**	**0.564**	**0.514**	**0.508**	0.361	0.384	0.301	0.293	**0.769**	**0.644**	**0.709**	**0.656**	**0.573**	0.450	**0.526**	0.437
P23	0.461	**0.567**	0.473	0.474	0.275	0.437	0.294	0.296	**0.503**	**0.616**	**0.518**	**0.517**	0.300	0.476	0.327	0.324	**0.732**	**0.709**	**0.730**	**0.728**	0.479	0.493	**0.524**	0.480
P32	0.394	0.437	0.361	0.357	0.322	0.370	0.287	0.286	0.414	0.457	0.382	0.375	0.336	0.387	0.300	0.294	**0.538**	**0.548**	**0.541**	**0.534**	0.446	0.490	0.416	0.392
W1	**0.524**	**0.552**	0.488	**0.526**	0.416	0.453	0.374	0.419	**0.545**	**0.576**	**0.511**	**0.549**	0.436	0.475	0.396	0.441	**0.674**	**0.655**	**0.628**	**0.647**	**0.562**	**0.538**	**0.519**	**0.532**
W3	0.420	0.412	0.412	0.403	0.304	0.288	0.288	0.278	0.471	0.460	0.465	0.449	0.341	0.324	0.331	0.309	**0.691**	**0.631**	**0.667**	**0.661**	**0.526**	0.464	**0.524**	0.491
W5	**0.501**	**0.582**	0.485	0.491	0.306	0.431	0.279	0.290	**0.594**	**0.662**	**0.578**	**0.572**	0.373	0.491	0.345	0.334	**0.757**	**0.715**	**0.769**	**0.735**	0.484	0.485	**0.510**	0.425
With antenna	**0.512**	**0.518**	0.488	**0.507**	0.354	0.362	0.319	0.346	**0.595**	**0.586**	**0.574**	**0.580**	0.424	0.415	0.397	0.406	**0.752**	**0.686**	**0.743**	**0.719**	**0.572**	0.484	**0.585**	**0.542**
Without antenna	0.443	**0.532**	0.437	0.436	0.325	0.438	0.316	0.318	0.479	**0.573**	0.474	0.471	0.350	0.473	0.343	0.341	**0.676**	**0.662**	**0.668**	**0.664**	**0.514**	**0.531**	**0.510**	0.479
With rope	0.436	0.487	0.416	0.415	0.273	0.347	0.247	0.243	**0.511**	**0.553**	0.493	0.479	0.334	0.402	0.314	0.292	**0.729**	**0.640**	**0.697**	**0.667**	**0.535**	0.454	**0.527**	0.464
Without rope	**0.537**	**0.560**	**0.516**	**0.534**	0.423	0.455	0.395	0.424	**0.587**	**0.609**	**0.570**	**0.583**	0.464	0.493	0.439	0.460	**0.705**	**0.714**	**0.721**	**0.722**	**0.556**	**0.569**	**0.576**	**0.562**

by applying a multi-sensor approach (combination of PID, TAA and NNMA), but this deserves further study.

4.1. Sensor technologies

4.1.1. Neural network model algorithms

The two NNMA models differed concerning the detection of interaction with the EM. The detector of NNMA 1 was only capable of detecting when a pig was interacting with the chain, whereas when using NNMA 2, the detector could also detect when a pig was interacting with the rope. This was due to the bounding boxes that were created around the EM in NNMA 1. Pigs were able to pull the rope outside the bounding box around the EM which was then not detected as interaction with the EM. This likely affected the performance of the NNMA models. In NNMA 1, bounding boxes were created around the complete head of the pigs while behaviours in the category "Intentional interaction with the enrichment material" were all performed with the snout. If bounding boxes had only been created around the snout of the pig instead of the complete head, the performance of NNMA 1 might have been better.

Overall, the performance for the category "All contact with enrichment material" was the lowest, compared to the other categories for both NNMA 1 and NNMA 2. This is encouraging since it demonstrates that the NNMAs are more effective at detecting intentional interaction with the EM than at detecting movement of the EM. Chen et al. (2020a) used a Hue, Saturation, Value colour space-based method to track objects in the region of interest and an InceptionV3 network and long short-term memory to automatically recognise episodes of enrichment engagement with a blue ball, golden ball and wooden beam in pigs and reached an accuracy of 96,5 %, 96.8 % and 97.6 %, respectively. Although different performance metrics were used (due to a balanced dataset of Chen et al., 2020a), a difference is seen in performance between the method used by Chen et al. (2020a) and this study. If more data will be used for training of NNMA 2, including data of different conditions such as different pen sizes, the performance is expected to be higher. However, when generalising the NNMA to other scenarios (e.g., other hanging EMs), we hypothesise that NNMA 1 will perform better since the algorithm can detect different objects. Thus, NNMAs are promising regarding interaction with EM in pigs, but the model, the attachment of other sensor technologies and the interaction between external factors needs to be refined to achieve optimal performance.

4.1.2. Passive infra-red detector

In general, the PIDs overestimated intentional interaction with the EM which might be due to the relatively small pen sizes resulting in pigs lying close to or standing/walking/running right past the EM without interaction. Von Jasmund et al. (2020) used passive infra-red detectors to record group activity and activity in certain focus areas in fattening pigs and found that PIDs generally overestimated activity, as compared to visual assessment of activity, which could also be the case in this study. Nevertheless, they found strong correlations up to $r = 0.87$ ($p < 0.01$) for the measurement of activity by PID compared to visual assessment (von Jasmund et al., 2020). Although different performance metrics were used, it can be assumed that PIDs have a better performance for measuring general activity compared to measuring activity directed towards the EM.

The height at which the PID was placed, and the threshold used in this study could also have influenced the results. The radius around the EM in this study was set at 10 cm. If the radius had been smaller, the performance would possibly have been better, i.e., fewer false positives. However, this could also have resulted in more false negatives because pigs would have pulled the EM even quicker outside the view of the PID. The threshold used in this study was based on detection (>0.05 V) or no detection of body heat (≤0.05 V). When the threshold would be higher, fewer false positives but also fewer true positives would be expected. This is based on the fact that the PID measures movement of body heat and some behaviours in the category "Intentional interaction with enrichment material" consisted of little movement of body heat. Increasing the threshold was also tested in this study but did not result in improved performance of the PID.

4.1.3. Tri-axial accelerometer

The highest performance of the TAA in the category "Intentional interaction with enrichment material" was achieved by the Y-axis. We expect this to be due to pigs wildly moving the EM and/or lifting/pushing/pulling the EM upwards during intentional interaction with the EM. For the category "All contact with enrichment material", the X-axis showed the highest performance. When pigs were lying in contact with or standing/walking/running against the EM without intentional interaction, the EM moved. Probably, at these moments, most of the acceleration occurred on the X-axis. For all the axes, the performance was higher for the category "All contact with enrichment material" compared to other categories. This was to be expected as this category included all contacts with the EM that made the EM move, and since the TAA measures acceleration, this was not surprising. The performance of all axes is not as high as expected which might be due to the continued swinging of the chain immediately after interaction with the EM, which was still measured by the TAA as acceleration (but without interaction). If this swinging can be filtered out of the data, higher performance is likely to be achieved. However, based on TAA data alone, it is difficult to discriminate between swinging and shaking of the EM. This can possibly be done with help of the other sensor technologies (to determine if a pig is present around the EM by e.g., using the PID) but this deserves further study.

4.2. Sensor performance over time and pens

The performance of sensor technologies differed over time (i.e., weeks 1, 3 and 5). NNMA 1 achieved higher performance in week 5 compared to weeks 1 and 3 whereas NNMA 2 achieved higher performance in weeks 1 and 3 compared to week 5 in the category "Intentional interaction with enrichment material". This may be due to the size of the piglets because when piglets were older (and larger), NNMA 1 may have had a better view of the pigs' heads compared to when they were younger (and smaller). The higher performance in weeks 1 and 3 compared to week 5 of NNMA 2 in the category "Intentional interaction with enrichment material" might be caused by the relative space per pig. Because pigs were larger during week 5, less space per pig was available compared to weeks 1 and 3. Due to this, NNMA 2 might have had more difficulties in detecting intentional interaction with the EM because pigs were more often lying down in contact with the EM or walked more against the EM without intentional interaction with the EM, which made it difficult to differentiate between standing, lying, and interaction with the EM. Next to this, in the category "Intentional interaction with enrichment material", NNMA 1 performed better in pens without an RFID antenna present (located around the EM) which was probably caused by the reduced visibility of the EM for the NNMA, as the RFID antenna took away some of the visibility.

It was expected that PIDs would achieve higher performance in pens without a rope attached to the chain compared to pens with a rope attached to the chain because pigs were able to pull the rope outside the focus area of the PID. However, in contrast to what was expected, PIDs performed better in pens with a rope than in pens without a rope. During the behavioural observations, it was observed that pigs were chewing on the rope within the focus area of the PID. Besteiro et al. (2018) validated measurements by PIDs against human observations of animal activity on a commercial weaner farm and concluded that animal weight affects the measurement capacity of the PIDs. They found that the PID performed better during the first weeks after weaning, whereas in this study, the performance of the PID was approximately the same in week 1 and week 5, respectively ($F1 = 0.425$ versus $F1 = 0.428$), for the category "Intentional interaction with enrichment material". The signal of the PID is proportional to the difference in temperature between the bodies and the background (Pedersen and Pedersen, 1995). Due to this, the detection capacity of the PID can be affected by the kg/m^2 ratio (Ni et al., 2017) because larger weaners occupy a larger area. However, within this study, the focus area of the PID was relatively small in proportion to the pen size, which might have resulted in no differences between ages because even with small pigs, a large proportion of the focus area was covered.

The performance of sensor technologies also differed between pens. In the present study, the performance of the PID and TAA was lowest in pen 32 compared to pens 2, 3 and 23. Due to synchronisation problems of the recorder time, external clocks were placed in the pens to see the actual time in the video footage. However, pen 32 was located more towards the centre of the compartment (more metal surroundings), compared to pens 2, 3 and 23 (which were located at the side of the compartment). Due to this, the external clock in pen 32 had more difficulty connecting with the transmission pole, resulting in this clock being one second behind the actual time at some moments and consequently a mismatch between observational data and sensor data, despite attempts to add a correction. For future studies (when exact time (i.e., exact seconds) is very important), it is therefore recommended to connect the video recorder to the network. The low performance of sensor technologies in pen 32 compared to pens 2, 3 and 23 might also be partially caused by the fact that piglets in pen 32 spent on average less time on "Intentional interaction with enrichment material" (9.86 %) compared to pigs in pens 2 (21.67 %), 3 (21.49 %) and 23 (21.60 %) which resulted in more imbalanced data for pen 32.

5. Conclusions

This paper compared two neural network model algorithms (NNMA 1 and NNMA 2), a passive infra-red detector (PID) and a tri-axial accelerometer (TAA) with manual behavioural observations (gold standard) to assess the performance of measuring interaction with the enrichment material in pens with weaned fattening pigs. Results showed that NNMA 1 and NNMA 2 performed best, when compared to the gold standard, followed by the Y-axis, X-axis, XYZ-average and Z-axis of the TAA and the PID, respectively, for measuring intentional interaction with the enrichment material (shake, carry, nose, bite, chew, root and/or more than one type of these behaviours). Overall, only moderate performance was reached. The results of this study indicated that the individual sensor technologies are not yet appropriate to measure interaction with the EM. However, there is potential to measure interaction with the EM by applying a multi-sensor approach (combination of NNMAs, PID and TAA), but this merits further study.

Funding

The research presented in this paper was supported by the Regio Deal Foodvalley.

CRediT authorship contribution statement

Fleur Veldkamp = planned and carried out the experiment, analysed the data, wrote the manuscript. Tomas Izquierdo Garcia-Faria = designed the neural network model algorithms. Vivian L. Witjes = planned and carried out part of the experiment. Johanna M.J. Rebel = supervised the project. Ingrid C. de Jong = was involved in planning and setup of the experiment and supervised the analysis and writing progress of the manuscript.

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.

Acknowledgements

The authors want to thank the staff of the Denkafarm (Denkavit Netherlands BV, Voorthuizen, The Netherlands) for taking care of the pigs and for their support during the study. We thank H. Gunnink (Wageningen Livestock Research) for the technical support and for assisting in the set-up of the sensor technologies and J.M.C.D. Meijer (Technical Development Studio, Wageningen University) for developing the passive infra-red detectors.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.applanim.2023.105923.

References

Alonso, M.E., González-Montaña, J.R., Lomillos, J.M., 2020. Consumers' concerns and perceptions of farm animal welfare. Animals 10, 385. https://doi.org/10.3390/ani10030385.

Berckmans, D., 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. Rev. Sci. Tech. 33, 189–196.

Besteiro, R., Rodríguez, M.R., Fernández, M.D., Ortega, J.A., Velo, R., 2018. Agreement between passive infrared detector measurements and human observations of animal activity. Livest. Sci. 214, 219–224. https://doi.org/10.1016/j.livsci.2018.06.008.

Blackshaw, J.K., 1986. Objective measures of welfare in farming environments. Aust. Vet. J. 63, 361–364. https://doi.org/10.1111/j.1751-0813.1986.tb02897.x.

Chapa, J.M., Maschat, K., Iwersen, M., Baumgartner, J., Drillich, M., 2020. Accelerometer systems as tools for health and welfare assessment in cattle and pigs – a review. Behav. Process. 181, 104262 https://doi.org/10.1016/j.beproc.2020.104262.

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. Comput. Electron Agric. 175, 105580 https://doi.org/10.1016/J.COMPAG.2020.105580.

Chen, C., Zhu, W., Steibel, J., Siegford, J., Wurtz, K., Han, J., Norton, T., 2020b. Recognition of aggressive episodes of pigs based on convolutional neural network and long short-term memory. Comput. Electron Agric. 169, 105166 https://doi.org/10.1016/J.COMPAG.2019.105166.

Chen, K., Wang, Jiaqi, Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., Zhang, Z., Cheng, D., Zhu, C., Cheng, T., Zhao, Q., Li, B., Lu, X., Zhu, R., Wu, Y., Dai, J., Wang, Jingdong, Shi, J., Ouyang, W., Loy, C.C., Lin, D., 2019. MMDetection: Open MMLab Detection Toolbox and Benchmark.

Cornou, C., Lundbye-Christensen, S., Kristensen, A.R., 2011. Modelling and monitoring sows' activity types in farrowing house using acceleration data. Comput. Electron Agric. 76, 316–324. https://doi.org/10.1016/j.compag.2011.02.010.

Czycholl, I., Kniese, C., Büttner, K., Beilage, E. grosse, Schrader, L., Krieter, J., 2016. Interobserver reliability of the 'Welfare Quality® animal welfare assessment protocol for growing pigs. Springerplus 5. https://doi.org/10.1186/S40064-016-2785-1.

Escobar, J., Van Alstine, W.G., Baker, D.H., Johnson, R.W., 2007. Behaviour of pigs with viral and bacterial pneumonia. Appl. Anim. Behav. Sci. 105, 42–50. https://doi.org/10.1016/J.APPLANIM.2006.06.005.

Feng, C., Zhong, Y., Gao, Y., Scott, M.R., Huang, W., 2021. TOOD: Task-aligned One-stage Object Detection, In: Proceedings of the IEEE International Conference on Computer Vision. pp. 3490–3499. https://doi.org/10.1109/ICCV48922.2021.00349.

Hitoshi Iuchi, 2017. mccr: The Matthews Correlation Coefficient.

von Jasmund, N., Wellnitz, A., Krommweh, M.S., Büscher, W., 2020. Using passive infrared detectors to record group activity and activity in certain focus areas in fattening pigs. Animals 10, 792. https://doi.org/10.3390/ani10050792.

Kashiha, M., Bahr, C., Haredasht, S.A., Ott, S., Moons, C.P.H., Niewold, T.A., Ödberg, F.O., Berckmans, D., 2013a. The automatic monitoring of pigs water use by cameras. Comput. Electron Agric. 90, 164–169. https://doi.org/10.1016/j.compag.2012.09.015.

Kashiha, M., Bahr, C., Ott, S., Moons, C.P.H., Niewold, T.A., Ödberg, F.O., Berckmans, D., 2013b. Automatic identification of marked pigs in a pen using image pattern recognition. Comput. Electron Agric. 93, 111–120. https://doi.org/10.1016/J.COMPAG.2013.01.013.

Kashiha, M., Bahr, C., Ott, S., Moons, C.P.H., Niewold, T.A., Ödberg, F.O., Berckmans, D., 2014a. Automatic weight estimation of individual pigs using image analysis. Comput. Electron Agric. 107, 38–44. https://doi.org/10.1016/J.COMPAG.2014.06.003.

Kashiha, M., Bahr, C., Ott, S., Moons, C.P.H., Niewold, T.A., Tuyttens, F., Berckmans, D., 2014b. Automatic monitoring of pig locomotion using image analysis. Livest. Sci. 159, 141–148. https://doi.org/10.1016/J.LIVSCI.2013.11.007.

Max Kuhn, 2022. caret: Classification and Regression Training.

Larsen, M.L. v, Wang, M., Norton, T., 2021. Information technologies for welfare monitoring in pigs and their relation to welfare quality®. Sustainability 13, 692. https://doi.org/10.3390/su13020692.

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. Vet. J. 217, 43–51. https://doi.org/10.1016/j.tvjl.2016.09.005.
Matthews, S.G., Miller, A.L., Plötz, T., Kyriazakis, I., 2017. Automated tracking to measure behavioural changes in pigs for health and welfare monitoring. Sci. Rep. 7 https://doi.org/10.1038/s41598-017-17451-6.
Mkwanazi, M.V., Ncobela, C.N., Kanengoni, A.T., Chimonyo, M., 2019. Effects of environmental enrichment on behaviour, physiology and performance of pigs — a review. Asian-Austral J. Anim. Sci. 32, 1. https://doi.org/10.5713/AJAS.17.0138.
Nasirahmadi, A., Hensel, O., Edwards, S.A., Sturm, B., 2016. Automatic detection of mounting behaviours among pigs using image analysis. Comput. Electron Agric. 124, 295–302. https://doi.org/10.1016/J.COMPAG.2016.04.022.
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. Biosyst. Eng. 158, 86–94. https://doi.org/10.1016/j.biosystemseng.2017.03.014.
Oczak, M., Maschat, K., Berckmans, D., Vranken, E., Baumgartner, J., 2015. Classification of nest-building behaviour in non-crated farrowing sows on the basis of accelerometer data. Biosyst. Eng. 140, 48–58. https://doi.org/10.1016/j.biosystemseng.2015.09.007.
Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems.
Pedersen, S., Pedersen, C.B., 1995. Animal activity measured by infrared detectors. J. Agric. Eng. Res. 61, 239–246. https://doi.org/10.1006/jaer.1995.1051.
Pfeifer, M., Eggemann, L., Kransmann, J., Schmitt, A.O., Hessel, E.F., 2019. Inter- and intra-observer reliability of animal welfare indicators for the on-farm self-assessment of fattening pigs. Animal 13, 1712–1720. https://doi.org/10.1017/S1751731118003701.
R Core Team, 2021. R: A Language and Environment for Statistical Computing.
Reiner, G., Hübner, K., Hepp, S., 2009. Suffering in diseased pigs as expressed by behavioural, clinical and clinical-chemical traits, in a well defined parasite model. Appl. Anim. Behav. Sci. 118, 222–231. https://doi.org/10.1016/J.APPLANIM.2009.02.010.
Salak-Johnson, J.L., Anderson, D.L., McGlone, J.J., 2004. Differential dose effects of central CRF and effects of CRF astressin on pig behavior. Physiol. Behav. 83, 143–150. https://doi.org/10.1016/J.PHYSBEH.2004.08.005.
Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., TOsmanov, Kruchinin, D., Zankevich, A., DmitriySidnev, Markelov, M., Johannes222, Chenuet, M., a-andre, telenachos, Melnikov, A., Kim, J., Ilouz, L., Glazov, N., Priya4607, Tehrani, R., Jeong, S., Skubriev, V., Yonekura, S., truong, vugia, zliang7, lizhming, Truong, T., 2020. opencv/cvat: v1.1.0. https://doi.org/10.5281/ZENODO.4009388.
Stygar, A.H., Dolecheck, K.A., Kristensen, A.R., 2018. Analyses of body weight patterns in growing pigs: a new view on body weight in pigs for frequent monitoring. Animal 12, 295–302. https://doi.org/10.1017/S1751731117001690.
Tian, M., Guo, H., Chen, H., Wang, Q., Long, C., Ma, Y., 2019. Automated pig counting using deep learning. Comput. Electron. Agric. 163, 104840 https://doi.org/10.1016/J.COMPAG.2019.05.049.
Tuyttens, F.A.M., de Graaf, S., Heerkens, J.L.T., Jacobs, L., Nalon, E., Ott, S., Stadig, L., Van Laer, E., Ampe, B., 2014. Observer bias in animal behaviour research: can we believe what we score, if we score what we believe. Anim. Behav. 90, 273–280. https://doi.org/10.1016/J.ANBEHAV.2014.02.007.
Tzanidakis, C., Simitzis, P., Arvanitis, K., Panagakis, P., 2021. An overview of the current trends in precision pig farming technologies. Livest. Sci. 249, 104530 https://doi.org/10.1016/j.livsci.2021.104530.
Ursinus, W.W., Van Reenen, C.G., Kemp, B., Bolhuis, J.E., 2014. Tail biting behaviour and tail damage in pigs and the relationship with general behaviour: predicting the inevitable? Appl. Anim. Behav. Sci. 156, 22–36. https://doi.org/10.1016/j.applanim.2014.04.001.
Wemelsfelder, F., Mullan, S., 2014. Applying ethological and health indicators to practical animal welfare assessment. Rev. Sci. Tech. . Int. Epiz 33, 111–120.
Zhu, X., Su, W., Lu, L., Li, B., Wang, X., Dai, J., 2020. Deformable DETR: deformable transformers for end-to-end object detection.
Zonderland, J.J., Schepers, F., Bracke, M.B.M., den Hartog, L.A., Kemp, B., Spoolder, H.A.M., 2011. Characteristics of biter and victim piglets apparent before a tail-biting outbreak. Animal 5, 767–775. https://doi.org/10.1017/S1751731110002326.