Applied Animal Behaviour Science 265 (2023) 106000

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

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

The quest to develop automated systems for monitoring animal behavior

Janice M. Siegford[a,*], Juan P. Steibel[b], Junjie Han[a,c,1], Madonna Benjamin[d], Tami Brown-Brandl[e], Joao R.R. Dórea[f], Daniel Morris[g], Tomas Norton[h], Eric Psota[i], Guilherme J.M. Rosa[f]

[a] Department of Animal Science, Michigan State University, East Lansing, MI 48824, USA
[b] Department of Animal Science, Iowa State University, Ames, IA 50011, USA
[c] Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA
[d] Department of Large Animal Clinical Sciences, Michigan State University, East Lansing, MI 48824, USA
[e] Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, NE 68583, USA
[f] Department of Animal and Dairy Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA
[g] Department of Biosystems and Agricultural Engineering and Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA
[h] Animal and Human Health Engineering, KU Leuven, Belgium
[i] PIC North America, USA

ARTICLE INFO

Keywords:
Animal behavior
Automated monitoring technology
Sensors
Observation

ABSTRACT

Automated behavior analysis (ABA) strategies are being researched at a rapid rate to detect an array of behaviors across a range of species. There is growing optimism that soon ethologists will not have to manually decode hours (and hours) of animal behavior videos, but that instead computers will process them for us. However, before we assume ABA is ready for practical use, it is important to take a realistic look at exactly what ABA is being developed, the expertise being used to develop it, and the context in which these studies occur. Once we understand common pitfalls occurring during ABA development and identify limitations, we can construct robust ABA tools to achieve automated (ultimately even continuous and real time) analysis of behavioral data, allowing for more detailed or longer-term studies of behavior on larger numbers of animals than ever before. ABA is only as good as it is trained to be. A key starting point is having manually annotated data for model training and assessment. However, most ABA developers are not trained in ethology. Often no formal ethogram is developed and descriptions of target behaviors in ABA publications are limited or inaccurate. In addition, ABA is also frequently developed using small datasets, which lack sufficient variability in animal morphometrics, activities, camera viewpoints, and environmental features to be generalizable. Thus, ABA often needs to be further validated before being used satisfactorily on different populations or under other conditions, even for research purposes. Multidisciplinary teams of researchers including ethologists and ethicists as well as computer scientists, data scientists, and engineers are needed to help address problems when applying computer vision ABA to measure behavior. Reference datasets that can be used for behavior detection should be generated and shared that include image data, annotations, and baseline analyses for benchmarking. Also critical is the development of standards for creating such reference datasets and descriptions of best practices for methods for validating results from detection tools to ensure they are robust and generalizable. At present, only a handful of publicly available datasets exist that can be used for development of ABA tools. As we work to realize the promise of ABA (and subsequent precision livestock farming technologies) to detect animal behavior, a clear understanding of best practices, access to accurately annotated datasets, and networking among ethologists and ABA developers will increase our chances for rapid and robust successes.

* Corresponding author.
E-mail address: siegford@msu.edu (J.M. Siegford).
1 Present address: Bayer Crop Science, Chesterfield, MO, 63017

https://doi.org/10.1016/j.applanim.2023.106000
Received 16 January 2023; Received in revised form 3 July 2023; Accepted 12 July 2023
Available online 17 July 2023

1. Introduction

Automated behavior analysis (ABA) strategies are being developed at a rapid rate to detect an array of behaviors across a range of species. There is growing optimism that soon, ethologists will not have to manually decode hours (and hours) of animal behavior videos, but that instead a computer can process visual, audiological, and other animal-based information for scientists working to gain insight into animal lives. However, despite this optimist outlook, it is important to take a realistic look at the ABA that is being developed, who is developing it, and the context in which these studies occur.

To set the stage, ABA[2] is here defined as use of technology to detect and observe the behavior of animals—collecting data, perhaps even continuously and in real time and at the level of the individual—in ways that require minimal human labor (Zamansky et al., 2021). Also called computational ethology (Anderson and Perona, 2014) or computational analysis of behavior (Egnor and Branson, 2016), ABA has potential to reduce both the human labor and subjectivity that currently impede scientific analysis of animal behavior. A recent search for peer-reviewed articles on ABA of animals revealed nearly 25,000 publications on the topic. The diversity of contexts and species in which work is underway to automate the observation of behavior is impressive, ranging from studies of mussels responding to environmental contaminants (Shen and Nugegoda, 2022) to social interactions of fruit flies (Dankert et al., 2009) to grazing of dairy cattle on pasture (Deniz et al., 2017) to fish caught in hook-and-line fisheries (Knotek et al., 2022).

Clearly, there is interest in and enthusiasm for solving the problems posed by detecting animals and the array of behaviors they can perform. The motivations of researchers driving their work in this area are nearly as diverse as the species being examined or the purposes for which these tools can be used. For example, some researchers wish to use technology as a tool to scientifically understand animals more deeply (Anderson and Perona, 2014), including their ecology (Dell et al., 2014; Weinstein, 2018) in ways that could inform conservation efforts (Bain et al., 2021; Tuia et al., 2022). Others may be using ABA to search for insights into human psychology or health through studies of animals (Voikar and Gaburro, 2020). ABA could also be used to recognize pain and emotion (Jourdan et al., 2001; Broome et al., 2023) and to understand the workings of the brain (Mathis and Mathis, 2021; Tecott and Nestler, 2004).

Of particular importance to applied ethologists, data from ABA could be used for behavioral research (Noldus et al., 2001; Valletta et al., 2017), to inform animal management practices (i.e., husbandry (Buller et al., 2020) or veterinary treatment (Kaplun et al., 2019)), or to assess and improve animal welfare (Broome et al., 2023; Dawkins et al., 2009; Matthews et al., 2016; McLoughlin et al., 2019; Rushen et al., 2012). As technology becomes a lynchpin in managing food and fiber animals more efficiently and precisely, there is a need to develop ABA that can closely monitor large numbers of animals on farms and detect changes in their behavior (Norton et al., 2019). Similarly, ABA tools can enhance transparency about farm animal welfare (Buller et al., 2020; Larsen et al., 2021) or create connections between zoo animals and the public (Clay et al., 2011)—playing a role in our social license to keep animals for human purposes.

Our objectives with this review are to help applied ethologists develop a critical approach to evaluating ABA tools for their own use and to facilitate ethologists' engagement as experts to help develop high quality ABA. Many of the specific examples presented and literature cited focus on livestock (pigs in particular) and poultry. ABA used by applied ethologists in the context of animal agriculture must be robust as livestock and poultry species are often housed in large numbers (hundreds to thousands), at high densities, and groups are typically comprised of animals of the same breed, sex, and life stages—rendering them very homogeneous in appearance (Nasirahmadi et al., 2017). The difficulty in distinguishing between nearly identical animals coupled with frequent occlusion and overlap makes using ABA in these conditions particularly challenging (Chen et al., 2021). This is not to say that these conditions never arise in other populations, including wild ones (i.e., large herds of ungulates or seabirds for example), however, these conditions are almost universal in animal agriculture.

2. Methods

As the body of literature surrounding ABA is extensive and growing rapidly and the intent of this paper was not to provide an exhaustive systematic review of any one type or use of ABA, a targeted search of the peer-reviewed literature was conducted. Initially, the authors drew upon their experience conducting research and previous literature reviews on ABA to provide examples illustrating particular situations. To objectively broaden the scope of the review, a search of the scientific literature was conducted. The first search focused on detecting review articles published within the last 10 years using 'automated' AND 'behavior-' AND 'analysis' alone and coupled with 'animal' and 'review'. Similar searches were conducted using synonyms of ABA such as 'computational' AND 'analysis' AND 'behavior' as well as 'computational' AND 'etholog-' alone and in combination with 'animal' and 'review'. Next, these searches were repeated by combining ABA and its synonyms with 'livestock', 'poultry', 'cattle', 'pig'. Articles were first reviewed to ensure they were 1) published peer-reviewed articles and 2) written in English, and then were selected for inclusion based on their relevance to the topic of ABA in general and applied ethology more particularly. Finally, a snowballing strategy was used whereby additional peer-reviewed articles were discovered through examination of the references cited in the articles discovered by the initial searches. In this way, several early articles that discussed the potential to use ABA in applied ethology were discovered (e.g., Jourdan et al., 2001; Dawkins et al., 2009; Leroy et al., 2006; Noldus and Jansen, 2004; Rushen et al., 2012).

Table 1 presents reviews of ABA from the peer-reviewed scientific literature that describe potential applications for ABA, evaluate performance of programs and approaches, and summarize the research conducted to date. These reviews present thoughtful perspectives on opportunities and constraints to using ABA that complement the points raised throughout this article. The reviews compiling and comparing specific ABA approaches provide a useful index of the literature describing technological developments.

3. ABA is only as good as it is trained to be

For the purposes of this paper, the focus will be less on why monitoring tools are developed or specific uses to which they can be put but rather on some elements that are critical to consider when evaluating or helping create tools to automatically monitoring behavior. For a technological ABA solution to be accurate, robust, and usable in the real world, it must be developed with consideration from the formation of the development team through final testing of the product (Fig. 1). As described below, failure at any of these stages is likely to result in ABA that is less than fully automated or that does not deliver useful outputs.

[2] The ABA discussed in this paper is not synonymous with Precision Livestock Farming (PLF). ABA data is not necessarily converted into actionable information for use either by a human caretaker or to cause an automated change in management. When ABA is integrated into a commercialized technology aimed at assisting a farmer working in a production setting, then it becomes part of a precision livestock farming (PLF) approach. PLF detects animals and monitors their responses—ideally continuously and in real time—as well as provides actionable information to the farmer that can be used to make management decisions (Berckmans, 2017). At PLF's core is likely technology that does ABA—though in other cases, it could be physical characteristics of animals that are being monitored such as weight, body condition, temperature, injury, or even measurements taken beyond the level of the animal (Banhazi et al., 2012)

Table 1
Peer-reviewed literature describing and evaluating automated behavioral analysis (ABA) methods and technologies and potential uses for these tools.

Citation	Focus	Reasons to Read
Anderson and Perona (2014)	ABA overview	Exploration of how ABA could allow better collection of behavioral data needed to understand how the brain works. Reviews importance and history of ethology and necessity of understanding behavior and its value as an animal-based indicator
Broome et al. (2023)	ABA for emotion, pain	Review of automated behavior analysis approaches to study animal emotions and pain. Discusses practical aspects such as data collection, annotation, recording equipment, evaluation of ABA performance, and provides best practice recommendations
Egnor and Branson (2016)	ABA overview	Review of how computational approaches (i.e., machine learning) can be used to quantitatively analyze behavior. Discusses key considerations of experimental design to achieve biologically relevant data while facilitating use of technology. Overview of types of machine vision and learning mistakes and their strengths and limitations
Mathis and Mathis (2020)	Deep learning, pose estimation	Review of how deep learning methods have potential to improve ABA, with a focus on pose estimation (which can be transformed into actions or kinematic information) using an open source program
McVey et al. (2023)	Unsupervised learning, case studies	Examination of potential for unsupervised learning to improve ABA in precision livestock applications. Discusses difference between data and information, data compression and information loss, model-dependent and model-free approaches and potential to discover complex and unexpected behavior signals from a range of data streams
Rushen et al. (2012)	ABA for welfare	Exploration of using automated on farm technology to provide behavior data as an aspect of assessing welfare. Explanation of basic terms and rationale for using automation as well as drawbacks to consider
Valletta et al. (2017)	Machine learning for analysis	Examination of using machine learning rather than traditional statistics to turn automatically collected data into information with applications ranging from behavioral pattern detection to detecting emotional state to understanding social network structure to welfare assessment
Wurtz et al. (2019)	Machine vision for ABA	Systematic review assessing 108 ABA studies conducted on livestock and poultry housed on farm using machine vision approaches. Provides information about equipment used, species studied, and types of behaviors analyzed. Emphasizes need for detailed metadata to allow for evaluation of ABA outcomes and its potential to be used under other conditions
Zamansky et al. (2021)	ABA overview	Succinct discussion of use of ABA to remove error and human subjectivity from behavioral studies as well as to increase volume of data that can be analyzed, which increases potential for discovering knowledge, including complex patterns as well as behavioral features humans are not able to perceive. Covers importance of how we use language or symbols to describe behavior to improving ABA
Dell et al. (2014)	Program evaluation	Discuss potential uses of programs developed to track animal movement beyond the laboratory. Reviewed 16 open-source programs and noted their pros and cons (but seePanadeiro et al., 2021 for more recent evaluation)
Panadeiro et al. (2021)	Program evaluation	Reviewed 28 free animal tracking software programs and characterized their abilities and limitations, including user-friendliness, whether the programs have been updated, group size, length of tracking and how well they preserve animal identification. Discussion of importance of reliable tracking of individuals as a precursor to more complex ABA
Wurtz et al. (2022)	Program evaluation	Assessed performance of 4 open source animal tracking programs on pigs housed on farm. Describes strengths and weaknesses and constraints to consider
Arulmozhi et al. (2021)	Pigs	Summary of studies using 2D, 3D, and infrared cameras to acquire information, including behavior from pigs. Includes section on ABA. Includes section on limitations of cameras with images illustrating problems
Chen et al. (2021)	Pigs, Cattle	Technical discussion of how deep learning methods can improve ABA compared to traditional computer vision approaches. Focuses on behaviors of importance to health and welfare such as aggression, drinking, feeding, etc
Gómez et al. (2021)	Pigs	Systematic review of 111 validated sensor technologies used to detect animal welfare (many of which do ABA). Description of different types of internal and external validation. Description of number of animals used in each study (most studies use < 50 animals to develop technologies)
Jourdan et al. (2001)	Pain	Describes how automation of behavior analysis could bring objectivity and depth to studies of animal pain and reduce time and labor
Larsen et al. (2021)	Pigs	Systematic review of 101 publications (1989–2020) relating to information technologies developed to detect welfare, including ABAs. Tables present technologies grouped by welfare quality parameters, many of which rely on behavior for assessment. Descriptions of type of technology and stage of technology development and what stage of pig production was studied
Li et al. (2022)	Pigs	Examination of barriers related to using computer vision on pig farms, including summary of challenges related to ABA of pig behavior using vision. Discusses problems related to moving from human to machine recognition of animal behavior related to describing, defining, and labeling. Emphasizes importance of dataset construction and need for benchmarking datasets and methods to speed construction of ABAs

(continued on next page)

Table 1 (*continued*)

Citation	Focus	Reasons to Read
Matthews et al. (2016)	Pigs	Evaluation of technologies aimed at automatically detecting changes in pig behavior, with an emphasis on those that facilitate detection of health and welfare outcomes. Tables depicting useful behavior targets as well as state of ABA technologies. Covers publications from 1993 to 2016. Explanation of importance of using behavior measures validated as health and welfare indicators
Nasirahmadi et al. (2017)	Pigs, Cattle	Review of use of 2D and 3D cameras for ABA of cattle and pigs. Describes criteria used to evaluate ABA system performance (e.g., sensitivity, specificity). Describes features of state of the art imaging technology and processing techniques. Specific sections on feeding and drinking, lying, locomotion and lameness, aggression, and mounting
Yang and Xiao (2020)	Pigs	Review of approaches to video based ABA of pig behavior that breaks down various aspects necessary to the process with accompanying review of the literature
Zhang et al. (2022)	Pigs	Summary of studies using sensors to detect pig behaviors. Includes simplified description of types of processing and analysis. Presents some performance metrics from various studies. Covers behaviors ranging from feeding, drinking, locomotion and aggression an other social behavior through estrus detection, lactation, and parturition as well as tail biting
Abd Aziz et al. (2020)	Poultry	Review of open access literature published from 2010 to 2020 on computer vision in poultry (chicken/broilers). Includes detailed descriptions of types of hardware, software, and data processing and analysis methods
Li et al. (2020)	Poultry	Review of publications from 2015 to 2020 describing technologies (including body worn and remote) used to detect behavior of chickens (broilers and layers) with potential for on farm use. Image processing section heading includes a useful table and description of recent studies focused on detecting individual bird behavior
Ojo et al. (2022)	Poultry	Systematic literature review of studies published between 2010 and 2022 presenting information technologies developed for use with poultry. Includes a table with a comprehensive list of studies on monitoring poultry welfare using a range of device types
Okinda et al. (2020)	Poultry	Technical review of computer vision approaches used in processing images, segmenting, feature extraction, shape analysis, kinematics, optical flow, and statistical approaches. Includes a table of computer vision monitoring systems used for behavior and welfare applications
Olejnik et al. (2022)	Poultry	Examination of technologies used to improve broiler management; includes sections on sound analysis and locomotion and activity tracking
Rowe et al. (2019)	Poultry	Review of automated measures that could be used in precision livestock farming of poultry to provide welfare information. Includes an overview of technologies that monitor behavior
Vieira Rios et al. (2020)	Poultry	Systematic review of 57 papers describing technologies used to detect broiler welfare (many of which rely on ABA). Describes which phase of life or type of production system the technology could be used in and whether studies were experimental or had practical application
Wu et al. (2022)	Poultry	Review of importance of information technologies in modern poultry farming. Includes a specific section on behavior recognition describing various studies done toward recognizing behaviors such as feeding and drinking, movement, and egg laying using a variety of types of sensors and ABA approaches.

Fig. 1. Creating useful ABA requires several important inputs. The development team conceiving ABA must have domain expertise (e.g., ethologists) as well as necessary technical skills (e.g., computer science). Sufficient high-quality data are needed for training and testing ABA, and independent data are needed for validation. If supervised learning is used, manual decoding must be done by an expert following consistent high standards that connect to previous research on the behavior. Resulting datasets also need to include metadata that describes features of data collection, animal details, and environmental context.

3.1. *Developers with the right expertise and standards*

A key point to start with when evaluating how good an ABA technology is likely to be is who helped develop it (Li et al., 2022). Most ABA systems are developed by a team, at the center of which are the computer or data scientists who create algorithms capable of detecting the behavior from the data that is collected by the technology. Engineers are also involved, particularly when sensors used to capture data are specifically developed for the ABA instead of using off the shelf cameras or accelerometers or in deploying the technology in the real world. While many teams may include an agricultural engineer, not all teams include animal scientists or biologists. Therefore, domain experts in ethology might be missing from development teams. Thus, when evaluating whether the ABA is capable of recognizing behaviors of interest, it is necessary to read the paper the work was published in or documentation accompanying the technology to learn about the expertise underlying its development.

Critical questions to ask when evaluating the expertise behind ABA development include:

- How well do the humans training the ABA understand the behavior they are attempting to monitor? Was the ABA produced by a team that included domain experts (i.e., ethologists with species-specific expertise)?
- Do they understand the specific postures or actions that make up the functional behavior? Or what the full expression of the functional

behavior looks like? For example, dust bathing is composed of many small elements (e.g., movemes such as side lying) and actions (e.g., head rubbing or wing shaking) that come together in a particular way into a functional dust bathing activity that results in cleaner feathers (see Olsson and Keeling, 2005 for a review).

- Do they give behaviors the same names to postures or activities that have been commonly used in ethology or animal science literature? (See Zamansky et al., 2021 for a discussion on the importance of how behavioral phenomena are represented in language or symbols when doing ABA.) Is a detailed description of the behavior provided (or an image or video) to allow for verification that the behavior they detected is what a reader assumes it to be (Li et al., 2022). For example, when developing ABA to recognize 'aggression' in pigs (e.g., Chen et al., 2019), there are many forms of aggressive behavior that must be identified such as head to head knocking, head to body knocking, parallel pressing, inverse parallel pressing, and biting of the head, neck, or body which must be defined and identified in ways consistent with previous important work (e.g., Desire et al., 2015, O'Connell et al., 2005; O'Malley et al., 2022).
- If they try to interpret motivations or meanings underlying the behavior, are these in line with the scientific research on these behaviors? For example, tail biting in pigs is not motivated by aggression but is caused by a variety of housing and management factors (Prunier et al., 2020; Sonoda et al., 2013; Taylor et al., 2010). Thus, AMB systems that are designed to predict tail biting outbreaks (e.g., D'Eath et al., 2018) cannot also be used to predict aggression nor can presence of tail injuries be used as a proxy for aggressive behavior the way wounds to other areas of the body might be (Turner et al., 2006).
- Do they assume that proximity to a resource (such as a feeder) means that the animal is engaging in a particular behavior (eating)? Or if one animal is near another that it is socially interacting? It is important to consider accompanying motions and orientations as well to determine if the functional behavior of interest is occurring (Yang and Xiao, 2020). Particularly in situations where animals are housed at high densities, they may not have the opportunity to avoid being near resources or other animals, even if they are not performing behaviors often associated with such proximity (Alameer et al., 2020).

3.2. Good data are essential for training and validating ABA

As important as considering the expertise and engagement of the team developing the ABA are the data that are used to train and test the approach to detecting behavior.

Most published research relating to use of ABA with livestock and poultry was conducted using small datasets and/or small numbers of animals (e.g., as reviewed by Gómez et al. (2021), Li et al. (2020), Li et al. (2022), Wurtz et al. (2019). While this means the developers can quickly generate a solution that works for that particular data set, this also means that it is unlikely the solution will be useful in other contexts or on different animals without additional training and testing (Fig. 2, Gómez et a., 2021). This is true even if we are thinking about using the technology in a similar small scale research study—let alone scaling up to a commercial facility where both animal number and density increase. Why? Small datasets do not capture the variability needed to enable ABA to adapt to variations in individual animals (appearance and behavior), environments, or data capture devices (Arulmozhi et al., 2021; Wurtz et al., 2019).

Key factors to evaluate with respect to datasets used for ABA development include:

- Over what length of time data were collected, the times of day or season; ages, breeds, and sexes of animals and details about group sizes (and how long groups were together)
- Did the developers describe the number of images or length of video clips used? Were enough independent videos/images used? It is important to note, that there can be many aspects that need to be considered in terms of independence including temporally, spatially, genetically, etc.
- Was the number of replicates (animals and pens and farms) sufficient to capture a wide range of individual variation in behavior?
- What was the context? Is it a very specific type of pen or arena? Or did they include a large number of physical environments suggesting the technology will generalize well to other situations?

3.3. Manual decoding needs to be rigorous

Of course, the best sensor data in the world can still lead to development of a poor tool if the initial decoding of the data that were used to train the model was done poorly (Li et al., 2022). Here, decoding refers to accurately labeling sensor data, whether it be images, videos, audio recordings, or accelerometer data, etc., with respect to behaviors of interest. When such decoding is done using human observers, it is often described as the 'gold standard', which implies that this manual decoding is the accurate standard against which we will measure other

Fig. 2. Data used to develop ABA must include enough variability to be robust enough to be used in other situations. The top panel illustrates variability in ages and genders of domestic chickens (*Gallus gallus domesticus*) with images (from left to right) of a chick, pullets, a hen, and a rooster. The lower panel illustrates only some of the variability that can be present in ages, breeds, group sizes, and housing conditions of *Bos taurus* and *Bos indicus* cattle.

ways of capturing the data, such as using ABA tools. Unfortunately, manual decoding by a human is not always as good as gold. We need to critically evaluate the individual that manually decoded the data and the process they used to train the ABA to detect the behaviors of interest. Ethologists can play a critical role on development teams in constructing ethograms and manual decoding approaches and assisting computer science and engineering colleagues in better understanding behavior.

Key aspects to evaluate when examining the standard that was used to train the algorithm are:

- How much data was decoded?
- Who did the decoding? Were decoders trained? Against what standard? What was the inter- or intra-observer reliability?
- Was the decoding protocol built using an ethogram or detailed descriptions or images/videos of the behavior? Often no formal ethogram is developed and descriptions of target behaviors in ABA publications are limited or inaccurate.
- Was there a decoding protocol that was followed that guided the observer in how to record the data? How did the observer know when to consider that a behavior had stopped or started or if a pause constituted a stop in the behavior?
- At what time scale? Is it the right time scale for the approach you are using for automation?

3.4. Datasets must be well documented

Datasets underlying ABA development that are published must be accompanied by the right type of contextual information for others to be able to understand how the data were used previously and to enable reuse of the dataset (Heil et al., 2021). For example, there need to be clear explanations of what the labels mean, including spelling out all acronyms. All behaviors should be clearly described using an ethogram or pictorial examples so that everyone knows exactly what is meant by each behavior. If data are labeled following a systematic, phenomenological approach that takes into account the spatial and temporal scales over which behavior occurs, this would facilitate integration of this information into ABA approaches (Zamensky et al., 2021).

Beyond carefully describing the type of data collected, additional information (metadata) describing the context in which the data were collected must also be provided (Li et al., 2022). Unfortunately, such information is often missing from published studies (as described in Wurtz et al., 2019). At a minimum, metadata should include the total number of animals observed along with their age, sex, and breed. Many other meta-variables can also be useful such as coat color patterns, body condition score, weight, parity number, lactation status, and so on depending on the species or life stage under observation. Although some of these may be hard to record, when available, they should be captured when possible and considered when training and/or evaluating models for ABA. Additional useful data (as per Wurtz et al., 2019) include the objective of the study, a detailed description of housing conditions (e.g., lighting type, light-dark schedule, flooring type, pen size, and number of animals per pen), type of recording equipment used (e.g., camera, lens, and recording device specifications), management and production interventions (e.g., medical treatments or mixing animals into new social groups), problems or unusual conditions that arose during the experiment (e.g., feed or water line failures or fluctuations in temperature), recording settings used during data capture (e.g., resolution, frame rate, lens type, field of view, and shutter speed), and whether individual animals were marked (e.g., ear tags, hair dye, and back marks). Additional information that should accompany datasets previously used for developing are the data processing method, the method used to validate automatic detection, and performance evaluation results such as accuracy and precision (Broome et al., 2023; Gómez et al., 2021).

3.5. ABA performance must be appropriately validated

In ABA model development, a training-validating-testing data split is a common strategy used to divide available data into subsets for model training and testing purposes (May et al., 2010; Reitermanova, 2010). A training set is typically employed for model development and initial fitting, while a validation set is reserved for periodically testing the model performance during training and/or for fine tuning the model (May et al., 2010). A testing set is often a second hold-out dataset for validating the optimal model, which should be independent of the initial training-validation dataset. We acknowledge that the training-validating-testing split is widely accepted in the computer vision domain. However, we need to point out that, in this review, the validation strategy focuses on the training-validation split instead, where the validation strategy refers to the way to split the available data into a training set for the model development purpose and a validation set for the model evaluation purpose.

To demonstrate that ABA does detect behaviors of interest accurately and reliably, its performance must be evaluated (Broome et al., 2023). This is commonly referred to as 'validation'. A range of validation strategies exist (Gómez et al., 2021; Han et al., 2023b), and, importantly, depending on the validation strategy employed, even the best results may not mean the ABA will work under different conditions. A typical practice in applications using computer vision to detect animal behavior (Li et al., 2019; Nasirahmadi et al., 2019; Chen et al., 2020; Liu et al., 2020; Zhang et al., 2020) is to use a random validation approach to evaluate model performance, where the training sets and their corresponding validation sets are split at random.

However, in practical animal farming, there are unavoidable underlying structures present in the data that make a random validation strategy problematic. For example, temporal (e.g., animals growing over time) or spatial (e.g., differences between pens or rooms) dependence structures typically exist in data collected from animals. A random data split may ignore these dependence structures, leading to biased results (Han et al., 2023b; Roberts et al., 2017; Ferreira et al., 2022). Furthermore, if the data subsets are split inappropriately, the training and validation sets may not be representative of the problem domain (May et al., 2010).

For instance, Han et al. (2023) developed a computer vision-based pig agonistic behavior classifier and validated the model using three different validation strategies including random validation, block-by-time validation (training and validation sets were independent in terms of time), and block-by-social-group validation (training and validation sets contained different social groups of pigs). They reported that compared to random validation, evaluation metrics were substantially worse in block-by-time and block-by-social-group validations. In another application of machine learning-based cattle grazing activity prediction, Coelho Ribeiro et al. (2021) employed leave-one-animal-out (LOAO; all data points of given individuals were reserved for validation while the remaining data were used for training), leave-one-day-out (LODO; all data points of given days were reserved for validation), and random validation. The authors reported that LOAO and LODO validations yielded lower accuracy compared to random validation, implying the overfitting effect holdout. Therefore, the appropriate validation strategy needs to be considered given the research question and future use of the ABA algorithm rather than just defaulting to the typical, but often inappropriate, random validation strategy. The use of this strategy is one of the main explanations for why even under seemingly similar conditions or with similar species, ABA may perform more poorly than expected.

4. Current ABA solutions have limited utility

While there seems to be an endless array of exciting ABA being developed, much of it is not yet ready for mainstream use by ethologists to answer research questions for several reasons. Some of these reasons are related to the technological abilities of the ABA while others are related to commercial availability of products that can be used off the shelf. However, some reasons have to do with the skill sets of the ethologists, who may not be trained to fine tune software, adjust device settings, or manipulate the resulting data.

4.1. ABA solutions are very focused

Most ABA focuses on one problem or type of behavior at a time, such as detecting feeding (Fig. 3; Arulmozhi et al., 2021). This is an excellent strategy for developing tools that can be used in research aimed at understanding a specific behavior or when the purpose is to detect one key problem. However, it is not good for research or applications where a more holistic view is necessary, such as when looking at changes in behavioral profiles or time budgets or when trying to evaluate a social network that includes both proximity and what the animals are doing when they are close to each other. As ethologists, what we usually need is ABA with the ability to detect the performance of several (or even many) different types of behavioral activities (e.g., Leonard et al., 2019; Schmidt et al., 2022) and transitions between behaviors. Further, if we want to link behavior to health or welfare outcomes, we also need finer resolution. For example, not just recognition that a reduction in activity or feeding that indicates illness has occurred, but also the detection of other indicators that allow for the diagnosis of a specific problem.

4.2. Limited ready to use ABA available

Current ABA solutions have limited capacity to be used without modification by your average ethologist who does not possess the ability to code. Many ABA programs have been developed for use under laboratory conditions (i.e., good lighting, high contrast backgrounds, and one or few animals being detected at a time). Even under these conditions, these programs still have limitations in their abilities or are not intuitive for use without computer science expertise (as described in Panadeiro et al. (2021) and Dell et al. (2014) who both reviewed programs with respect to ease of use and robustness). When Wurtz et al. (2022) tested four open source laboratory animal tracking programs with pairs or small groups of pigs under commercial conditions, they encountered difficulties in maintain pig identities related to occlusion of the target pig(s) by other pigs or pen features, subtracting the background from the pig(s) of interest, lighting, and field of view.

As mentioned above, most ABA, including that intended specifically for use with livestock and poultry, has been developed with very little biological replication (i.e., small numbers of animals, all within the same set of pens) and cannot be generalized to other contexts (i.e., different breeds/ages animals or other environmental configurations and situations such as light levels (e.g., Sa et al., 2019) and weather conditions). This means that even if the source code for an ABA approach is freely available, it would require revision of the code to adapt it to a purpose or context beyond the one it was created in.

Finally, as this is a relatively young field, many ABA approaches are not at that stage of technological readiness that would allow them to be turned into commercial products. Unfortunately, due to the relatively small market for commercial ABA, it will not be economically feasible to create commercial products for sale or complete with technical support, at least not in the near future.

4.3. ABA has not yet reached its full potential

Most ABA is not yet as good as humans at adapting to the variability of animals they are detecting and observing. Moving from a human expert's qualitative recognition of a behavior into a quantitative description used by a machine (distance, angle, speed, pixels) is time-consuming and labor intensive (Norton et al., 2019). Behavioral detection is complicated by differences between how the data are collected, how ethologists record behavior, and the information machines need for training algorithms (Li et al., 2022).

- Animals, as living beings, have more complexity than is often acknowledged or that is present in data sets used to develop ABA (e.g., they are complex, individually different, time-varying and dynamic as described in Berckmans, 2017). In animal experiments data collection is constrained by the budget and time available to carry out a study. This limits the opportunity to capture the level of variability in animal behavior needed to train accurate and precise ABA technology (Gomez et al., 2021; Wurtz et al., 2019).
- Artificial Intelligence (AI) is not yet as good as humans at adapting to the variability of animals they are detecting and observing (e.g., wide range of size, shape, color, and coat phenotypes associated with different species, life stages or sexes of animals). This may be due in part to limitations in supervised learning approaches that require use of previous knowledge to train AI. However, unsupervised methods (Yang and Xiao, 2020) are likely to speed the process, provided they perform well. Unsupervised methods may also be useful in revealing

Fig. 3. ABA is often initially developed to detect one or few specific behaviors. This is a useful starting point but to be useful for ethological research or later commercial applications, ABA will need to detect and monitor multiple behaviors in a single, combined solution.

behavior patterns hidden in messy, complicated real-world data—either allowing for creation of better models or enabling visualization of systematic features of data sets that were not anticipated (McVey et al., 2023).
- Coding (data annotation) procedures for ABA development and ethological studies do not always correspond (e.g., different times scales or features used by humans versus machines to detect behavior; Leoni et al., 2020). As Zamansky et al. (2021) point out, "human understanding and interpretation of behavior is in itself subjective and sometimes inconsistent. In fact, there is still no agreement among biologists on the definition of 'behavior'." To better integrate our knowledge into ABA tools, we should to begin by using a systematic approach to representing behavioral phenomena using language or symbols that allow for multiple scales of resolution in time and space (Anderson and Perona, 2014). This approach needs to encompass the hierarchy of behavioral elements that builds from a moveme (a basic unit of behavior that cannot be further broken down, such as step) to an action (which is composed of movemes occurring in the same sequence, such as walk (step + step + step)) to an activity (a more complex and variable sequence of movemes and actions, such as foraging).
- Processes often used when training AI to detect, identify, and track static objects (e.g., car or soda can) do not always translate well when applied to living beings, which are much more variable in shape, orientation, or capacity for deformation (Egnor and Branson, 2016; Grondin et al., 2022; Marks et al., 2022). Deep learning strategies such as self-attenuation mechanisms or adaptive spatial feature fusion may help overcome issues related to animal bodies changing shape as they move or occlusions when they interact with each other (Chen et al., 2021). However, a consequence of using deep learning is that it is harder to interpret why classifications are made because the underlying computations become increasingly complex (Broome et al., 2023).
- All or part of animals' bodies may be blocked from a sensor (i.e., occluded) by parts of housing or other animals, and the animal may leave and re-enter the sensor's field of detection (e.g., outside the field-of-view of a camera or range of a tag reader) (Egnor and Branson, 2016).
- Humans can quickly put together information from various senses to recognize behaviors and the individuals doing it (e.g., observers can listen to and observe pigs simultaneously to detect a coughing pig. ABA would require a pen level cough monitor combined with computer vision that recognizes a specific pig showing movements associated with coughing). Humans combine several types of comprehension and memory to recognize behavior (most animal behaviors cannot be distinguished from a single image but require video (Liu et al., 2020; Yang et al., 2021). In other words, the visual recognition of behaviors requires a human's comprehensive abilities. The complex judgments require long-term memory, several reviews, and more than one type of human intelligence (induction, deduction, and prediction).

5. Moving forward to achieve useful ABA

As described above, the options are currently limited for using technology to automatically detect animal behaviors—particularly if we are looking to use a single solution to detect multiple behaviors across a range of environmental context and breeds (or even species). How do we get to ready to use ABA that can do what we need it to do, whether for purposes of understanding animals better or for creating management tools that monitor animals usefully? Can we have ABA that accurately identifies and tracks unmarked individual animals—even when housed in large, dense groups? Will it be possible to use ABA that identifies and records frequencies, durations, intensities, patterns, and variations in all the behaviors the animals do across the day? Can ABA be developed that localizes where the animal was doing the behavior or who it was interacting with?

5.1. Ethologists need the ability to use ABA

One approach to getting closer to ABA that is easier for ethologists to use is for us to expand our skill sets (Fig. 4). We have long appreciated the value of being cross trained in subjects like statistics, genetics, or ecology in order to answer more complex research questions. Now we need to consider developing strategies to gain proficiency in practical computer and data science skills to enhance our ability to work with technology and the large, complex data that are generated.

For example, if we (or our students) had better computer science training, we could do some rudimentary coding to adapt ABA for our purposes or to trouble shoot when something does not work. If we were trained in data science, we could develop some of the modeling, data processing approaches, or machine learning algorithms to better capture the complexities of behavior that are present the data but often obscured by the volume of data we collect from ABA. We would also be more comfortable interpreting how well the ABA performed (e.g., sensitivity, specificity, and confusion matrices, see Fig. 5) to understand the quality of information produced by the ABA.

Additionally, ABA can also be developed in ways that make it easier to use (Fig. 4). In some cases, apps have been, or can be, developed for ABA with user-friendly interfaces that can make interaction with the technology intuitive. In other cases, it may be possible to develop ABA that can simply do what we ask of it without the need for coding skills or knowledge of specific jargon. Developments in natural language could help here—making it easier for novices to tell the detection software what to do using simple verbal directions (Chang et al., 2020; Mishra and Kumar, 2020). For example, if a user can simply tell the ABA to "detect tail biting and record which pigs are tail biting," more ethologists may be comfortable incorporating ABA into their research.

Beyond just using AI to create ABA that makes the task of collecting behavioral data easier, we should also consider how we might be able to use AI to us ask better questions of animal behavior. Questions that are not biased by our anthropomorphic perspective or the cultural legacy we carry with us from our scientific training and the scientific paradigms we have been trained to operate within (Packard et al., 1990). Using computational animal behavior analysis approaches that apply AI techniques to analysis of ethological data can help us more objectively characterize what is occurring (Zamansky et al., 2021) or reveal subtle, complex, or longer-term patterns in the data that allow us to better understand phenomena such as welfare (Rufener et al., 2018; Gómez et al., 2022).

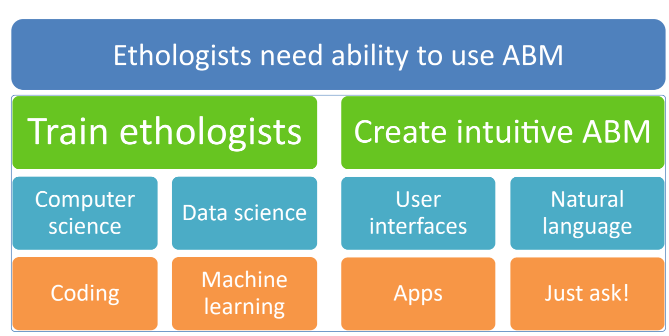

Fig. 4. Ethologists need ABA solutions they can use out of the box for their own purposes. At present many ABA approaches require continued work with the technical developer team to be adapted for new projects or contexts. One solution to this (shown on the left) would be for ethologists to be trained in computer and data science to enable them to perform their own programming to adjust ABA to suit their needs. ABA could also be created in ways that are more intuitive for end users to manipulate (shown on the right) through graphical user interfaces or the ability to give natural language commands.

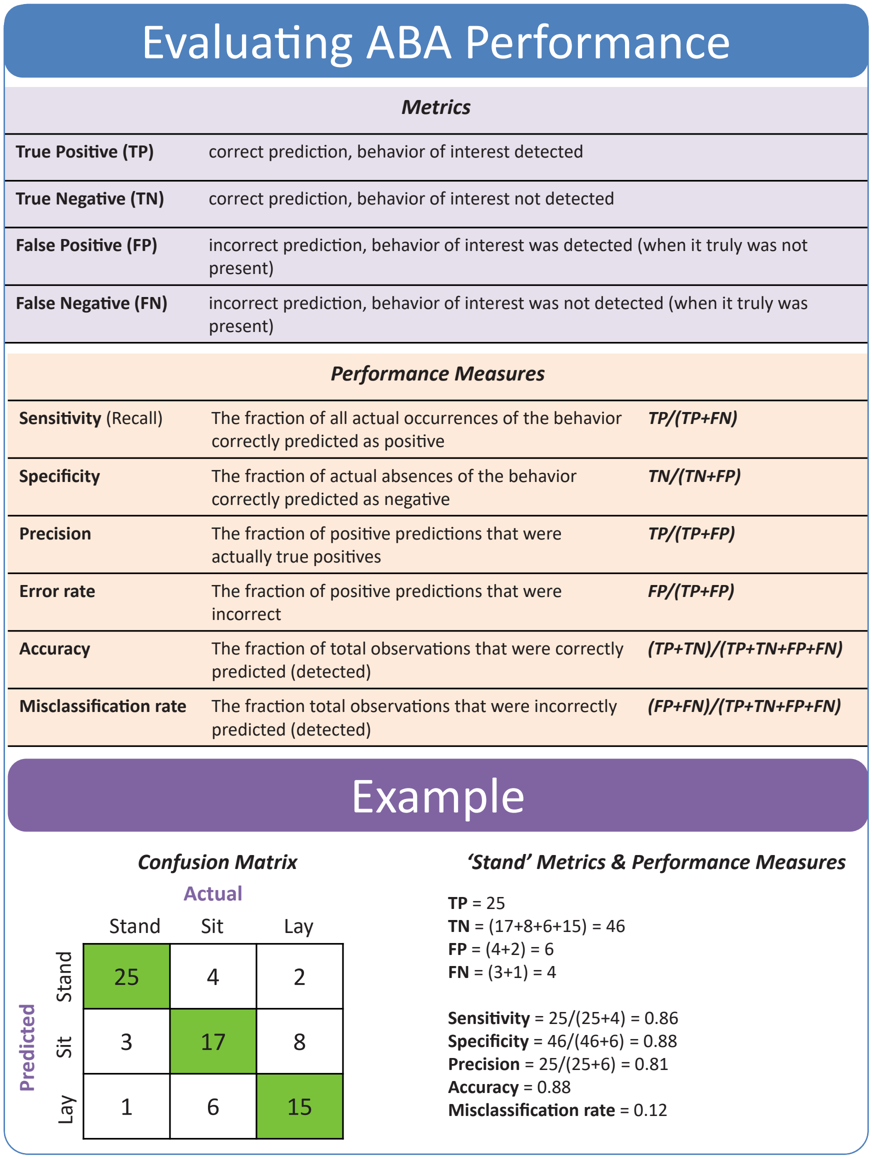

Fig. 5. Basic metrics and performance measures can be used to evaluate the performance of ABA. Confusion matrixes can be used to visualize the number of times behaviors are correctly detected (predicted) by the ABA compared to the actual number of times the behaviors occurred.

5.2. Make more datasets available to train ABA

To achieve robust and generalizable ready to use ABA tools, more and bigger datasets of animals of all types are needed, including of livestock and poultry, so that we have enough variation in training datasets to drive certainty for modeling (Fig. 6; Abd Aziz et al., 2020; Broome et al., 2023; Egnor and Branson, 2016; Han et al., 2023a; Li et al., 2022).

Video, images, sound, etc. from many species of livestock, poultry, fish, companion animals, etc.

- Various ages (stages of growth), sexes, groups sizes, breeds, colors
- Various types of environmental conditions to capture variation in dust, occlusion, flooring, lighting, etc.

Ethologists regularly collect such data as part of experimental studies and could contribute these valuable resources toward furthering ABM development, and funding agencies and scientific journals increasingly require providing others with access to this data (i.e., open data). While the size of video data sets in particular has made it prohibitive to directly

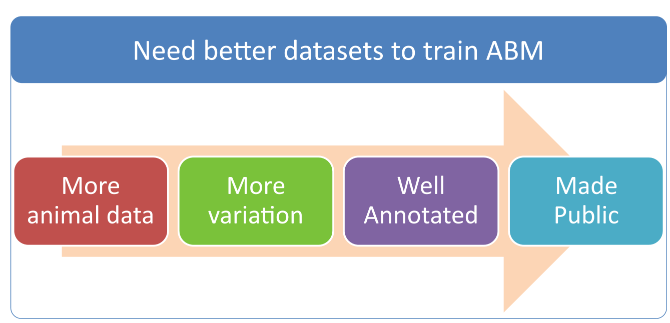

Fig. 6. More and better publicly available animal-based datasets are needed for ABA creation. Access to more animal data (images, sounds, etc.) will allow developers to train and test ABA to handle more variability, making end products more robust across different scientific studies or real-world contexts. Datasets that are rigorously decoded and contain sufficient metadata to explain how and where data were collected, annotated, and processed will also allow others to build new ABA without starting all over.

share video files, better compression and increasingly large and inexpensive online storage options may help overcome this problem. Any shared datasets should be annotated and labeled with metadata in ways that makes them usable by others beyond their creators (as described above). They also need to be shared and managed using rigorous standards so that they are accessible and usable by all. Publicly available image datasets, for example, exist for other AI applications and have been very useful for things such as facial recognition or autonomous driving (Cordts et al., 2016). Unfortunately, most of these datasets have relatively few animal images and those that are present are from non-livestock species or animals in non-agricultural types of settings.

In a recent search of the literature, we identified datasets with images of animals in agriculture that were publicly available and had been used for a computer vision (CV) application (Han et al., 2023b). What we found illustrates the limitations of the data that are available to help others develop and test ABA.

1. Most publicly available datasets for livestock concentrated on pig (9) and cattle (11), while there was only one publicly available dataset each for poultry, horses, sheep, and goats
2. It is challenging to remain up-to-date on new CV algorithms given how rapidly they are being developed and could become the state-of-the-art. However, in the ABA domain, it is even more challenging to adapt novel algorithms to animal specific application because of the lack of publicly available datasets (Han et al., 2023a; Li et al., 2022). We found few datasets available for specific CV tasks including open-set animal identification, interactive behavior recognition, and (object) segmentation—making it hard to quickly test or adapt new algorithms for other uses.
3. Animal-specific parameters are as important as image-specific parameters for ABA datasets. A big dataset must not only contain many images, but also be a variable dataset with a large number of animals across multiple environmental setups and/or across several production phases (Yang and Xiao, 2020).
4. Providing animal-specific metadata along with the images elevates the value/re-usability of the dataset. This allows other users to quickly understand the detailed conditions under which the data were collected so they can evaluate whether it will be appropriate for their intended use.
5. Most reviewed public image datasets have utilized top-down or angled-down camera views, and there is a lack of side-view and front-view datasets. Further, region of interest (ROI) is an important attribute for CV datasets and providing the ROI annotation along with the raw image is encouraged, rather than providing only the cropped ROI region.
6. There is a lack of standards/guidelines/protocols specifically designed for sharing data for use in CV applications for use with animals. FAIR principles (Go Fair, 2023) provide good general guidelines for data sharing. However, we need explicit guidance related to inclusion of more granular items/details that could be useful for ABA or precision livestock farming applications (Yang and Xiao, 2020).

5.3. Make collaboration a goal

One way we can speed the development of ABA that works for monitoring behavior of animals of various species, across a range of real-world contexts, and for purposes ranging from basic research to automated application is to work collaboratively. As noted above, development teams should include experts from a range of backgrounds (Li et al., 2022), which may not be present within a single institution. Collaborative projects and platforms that allow others to meet and interact, learn new techniques from each other, and open doors for deeper, longer-term work on funded research or commercial projects will be beneficial to ABA development. In addition to growing numbers of conferences that focus on automated monitoring of behavior, collaborative projects are underway around the globe that engage people from multiple institutions, host symposia, train personnel, offer courses, and sponsor webinars, discussion forums or problem-solving challenges. The use of real-world challenges posed by animal scientists, veterinarians, and farmers will provide computer scientists and engineers the opportunity to develop solutions to user-defined problems. ABA that is grounded in useful outcomes has potential to be used for more than detecting behavior and monitoring changes over time, and can move to management and intervention including tasks such as diagnosis and treatment of disease (Buller et al., 2020; Norton et al., 2019). As ABA moves from being used in research to animal management, it is imperative that ethicists, industry stakeholders, government representatives, certification organization, consumers, and members of the public are also engaged in the development conversation to consider the social, economic and environmental context in which such technology will operate (Akinyemi et al., 2023; Dawkins, 2021; Guzhva et al., 2021).

Another element helpful in facilitating development of ABA is the generation and sharing of reference datasets for behavior detection, including image data, annotations, metadata, and baseline analyses for benchmarking. Standards are necessary for the development and sharing of such datasets that stipulate inclusion of key pieces of information describing the context of the data, such as the technology used to collect it, parameters related to the data itself, and details of the animals and environment (Li et al., 2022). Once such datasets are available, analytical challenges built around the shared data can also spur discovery of solutions. Finally, standards for evaluating ABA results are needed to ensure high quality end products—as typical validation and assessment approaches are currently over-optimistic when it comes to how robust or generalizable an ABA solution will be.

6. Conclusion

As we work to realize the promise of ABA (and subsequent precision livestock farming technologies) to detect animal behavior, a clear understanding of best practices, access to accurately annotated datasets, and networking among ethologists and ABA developers will increase our chances for rapid and robust successes. Once we understand common pitfalls occurring during ABA development and identify limitations, we can construct robust ABA to achieve automated (ultimately even continuous and real time) analysis of behavioral data, allowing for more detailed or longer-term studies of behaviour on larger numbers of animals than ever before.

CRediT authorship contribution statement

Janice M Siegford, Juan P Steibel, Junjie Han: the conception and design of the study, or acquisition of data, or analysis and interpretation of data, drafting the article or revising it critically for important intellectual content; **Madonna Benjamin, Tami Brown-Brandl, Joao R.R. Dorea, Daniel Morris, Tomas Norton, Eric Psota:** drafting the article or revising it critically for important intellectual content; **Guilherme J. M. Rosa**.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Authors of this paper have been involved in the development of automated approaches to detect animal behavior or recognize animal appearance. None of the resulting products are endorsed in this manuscript. The authors declare that they have no other known conflicts of interests, whether personal or financial, that influenced the work reported in this paper.

Acknowledgements

USDA NIFA FACT CIN award# 2021-67021-34150 and USDA NIFA award# 2022-67021-37858 supported this work.

References

Abd Aziz, N.S.N., Daud, S.M., Dziyauddin, R.A., Adam, M.Z., Azizan, A., 2020. A review on computer vision technology for monitoring poultry farm—application, hardware, and software. IEEE Access 9, 12431–12445. https://doi.org/10.1109/ACCESS.2020.3047818.

Akinyemi, B.E., Vigors, B., Turner, S.P., Akaichi, F., Benjamin, M., Johnson, A.K., Pairis-Garcia, M.D., Rozeboom, D.W., Steibel, J.P., Thompson, D.P., Zangaro, C., Precision Livestock Farming: A Qualitative Exploration of Key Swine Industry Stakeholders. Frontiers in Animal Science, 4, p.47. https://doi.org/10.3389/fanim.2023.1150528.

Alameer, A., Kyriazakis, I., Dalton, H.A., Miller, A.L., Bacardit, J., 2020. Automatic recognition of feeding and foraging behaviour in pigs using deep learning. Biosyst. Eng. 197, 91–104. https://doi.org/10.1016/j.biosystemseng.2020.06.013.

Anderson, D.J., Perona, P., 2014. Toward a science of computational ethology. Neuron 84 (1), 18–31. https://doi.org/10.1016/j.neuron.2014.09.005.

Arulmozhi, E., Bhujel, A., Moon, B.E., Kim, H.T., 2021. The application of cameras in precision pig farming: an overview for swine-keeping professionals. Animals 11 (8), 2343. https://doi.org/10.3390/ani11082343.

Bain, M., Nagrani, A., Schofield, D., Berdugo, S., Bessa, J., Owen, J., Hockings, K.J., Matsuzawa, T., Hayashi, M., Biro, D., Carvalho, S., 2021. Automated audiovisual behavior recognition in wild primates. Sci. Adv. 7, eabi4883. https://doi.org/10.1126/sciadv.abi4883.

Banhazi, T.M., Lehr, H., Black, J.L., Crabtree, H., Schofield, P., Tscharke, M., Berckmans, D., 2012. Precision livestock farming: an international review of scientific and commercial aspects. Int. J. Agric. Biol. Eng. 5 (3), 1–9.

Berckmans, D., 2017. General introduction to precision livestock farming. Anim. Front. 7, 6–11. https://doi.org/10.2527/af.2017.0102.

Broome, S., Feighelstein, M., Zamansky, A., Carreira Lencioni, G., Haubro Andersen, P., Pessanha, F., Mahmoud, M., Kjellström, H., Salah, A.A., 2023. Going deeper than tracking: a survey of computer-vision based recognition of animal pain and emotions. Int. J. Comput. Vis. 131, 572–590. https://doi.org/10.1007/s11263-022-01716-3.

Buller, H., Blokhuis, H., Lokhorst, K., Silberberg, M., Veissier, I., 2020. Animal welfare management in a digital world. Animals 10 (10), 1779. https://doi.org/10.3390/ani10101779.

Chang, P., Liu, S., Chen, H., Driggs-Campbell, K., 2020. Robot sound interpretation: Combining sight and sound in learning-based control, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 5580–5587. IEEE. https://doi.org/10.1109/IROS45743.2020.9341196.

Chen, C., Zhu, W., Liu, D., Steibel, J., Siegford, J., Wurtz, K., Han, J., Norton, T., 2019. Detection of aggressive behaviours in pigs using a RealSence depth sensor. Comput. Electron. Agric. 166, 105003 https://doi.org/10.1016/j.compag.2019.105003.

Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., Norton, T., 2020. Recognition of feeding behaviour of pigs and determination of feeding time of each pig by a video-based deep learning method. Comput. Electron. Agric. 176, 105642 https://doi.org/10.1016/j.compag.2020.105642.

Chen, C., Zhu, W., Norton, T., 2021. Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning. Comput. Electron. Agric. 187, 106255 https://doi.org/10.1016/j.compag.2021.106255.

Clay, A.W., Perdue, B.M., Gaalema, D.E., Dolins, F.L., Bloomsmith, M.A., 2011. The use of technology to enhance zoological parks. Zoo. Biol. 30, 487–497. https://doi.org/10.1002/zoo.20353.

Coelho Ribeiro, L.A., Bresolin, T., Rosa, G.J., de, M., Rume Casagrande, D., Danes, M., de, A.C., Dórea, J.R.R., 2021. Disentangling data dependency using cross-validation strategies to evaluate prediction quality of cattle grazing activities using machine learning algorithms and wearable sensor data. J. Anim. Sci. 99 (skab206) https://doi.org/10.1093/jas/skab206.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, W., Roth, S., Schiele, B., 2016. The cityscapes dataset for semantic urban scene understanding. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 3213–3223.

D'Eath, R.B., Jack, M., Futro, A., Talbot, D., Zhu, Q., Barclay, D., Baxter, E.M., 2018. Automatic early warning of tail biting in pigs: 3D cameras can detect lowered tail posture before an outbreak. PLoS ONE 13, e0194524. https://doi.org/10.1371/journal.pone.0194524.

Dankert, H., Wang, L., Hoopfer, E.D., Anderson, D.J., Perona, P., 2009. Automated monitoring and analysis of social behavior in Drosophila. Nat. Methods 6, 297–303.

Dawkins, M.S., 2021. Does smart farming improve or damage animal welfare? Technology and what animals want. Front. Anim. Sci. 2, 736536 https://doi.org/10.1038/nmeth.1310.

Dawkins, M.S., Lee, H.J., Waitt, C.D., Roberts, S.J., 2009. Optical flow patterns in broiler chicken flocks as automated measures of behaviour and gait. Appl. Anim. Behav. Sci. 119, 203–209. https://doi.org/10.1016/j.applanim.2009.04.009.

Dell, A.I., Bender, J.A., Branson, K., Couzin, I.D., de Polavieja, G.G., Noldus, L.P., Pérez-Escudero, A., Perona, P., Straw, A.D., Wikelski, M., Brose, U., 2014. Automated image-based tracking and its application in ecology. Trends Ecol. Evol. 29, 417–428. https://doi.org/10.1016/j.tree.2014.05.004.

Deniz, N.N., Chelotti, J.O., Galli, J.R., Planisich, A.M., Larripa, M.J., Rufiner, H.L., Giovanini, L.L., 2017. Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals. Comput. Electron. Agric. 138, 167–174. https://doi.org/10.1016/j.compag.2017.04.024.

Desire, S., Turner, S.P., D'Eath, R.B., Doeschl-Wilson, A.B., Lewis, C.R.G., Roehe, R., 2015. Analysis of the phenotypic link between behavioural traits at mixing and increased long- term social stability in group-housed pigs. Appl. Anim. Behav. Sci. 166, 52–62. https://doi.org/10.1016/j.applanim.2015.02.015.

Egnor, S.R., Branson, K., 2016. Computational analysis of behavior. Annu. Rev. Neurosci. 39, 217–236. https://doi.org/10.1146/annurev-neuro-070815-013845.

Go Fair, 2023. FAIR Principles. ⟨https://www.go-fair.org/fair-principles/⟩ (Accessed 16 April 2023).

Ferreira, R.E.P., Bresolin, T., Rosa, G.J.M., Dórea, J.R.R., 2022. Using dorsal surface for individual identification of dairy calves through 3D deep learning algorithms. Comput. Electron. Agric. 201, 107272 https://doi.org/10.1016/j.compag.2022.107272.

Gómez, Y., Stygar, A.H., Boumans, I.J., Bokkers, E.A., Pedersen, L.J., Niemi, J.K., Pastell, M., Manteca, X., Llonch, P., 2021. A systematic review on validated precision livestock farming technologies for pig production and its potential to assess animal welfare. Front. Vet. Sci. 8, 660565 https://doi.org/10.3389/fvets.2021.660565.

Gómez, Y., Berezowski, J., Jorge, Y.A., Gebhardt-Henrich, S.G., Vögeli, S., Stratmann, A., Toscano, M.J., Voelkl, B., 2022. Similarity in temporal movement patterns in laying hens increases with time and social association. Animals 12, 555. https://doi.org/10.3390/ani12050555.

Grondin, V., Fortin, J.-M., Pomerleau, F., Giguère, P., 2022. Tree detection and diameter estimation based on deep learning. For.: Int. J. For. Res., cpac043. https://doi.org/10.1093/forestry/cpac043.

Guzhva, O., Siegford, J.M., Lunner Kolstrup, C., 2021. The Hitchhiker's guide to integration of social and ethical awareness in precision livestock farming research. Front. Anim. Sci. 2, 725710 https://doi.org/10.3389/fanim.2021.725710.

Han, J., Dorea, J., Norton, T., Morris, D., Siegford, J., Steibel, J., 2023a. Publicly available datasets for computer vision in precision livestock farming: A review, in: Conference Proceedings of the 2nd US Precision Livestock Farming Conference, pp. 618–625.

Han, J., Siegford, J., Colbry, D., Lesiyon, R., Bosgraaf, A., Chen, C., Norton, T., Steibel, J. P., 2023b. Evaluation of computer vision for detecting agonistic behavior of pigs in a single-space feeding stall through blocked cross-validation strategies. Comput. Electron. Agric. 204, 107520 https://doi.org/10.1016/j.compag.2022.107520.

Heil, B.J., Hoffman, M.M., Markowetz, F., Lee, S.I., Greene, C.S., Hicks, S.C., 2021. Reproducibility standards for machine learning in the life sciences. Nat. Methods 18, 1132–1135. https://doi.org/10.1038/s41592-021-01256-7.

Jourdan, D., Ardid, D., Eschalier, A., 2001. Automated behavioural analysis in animal pain studies. Pharmacol. Res. 43, 103–110. https://doi.org/10.1006/phrs.2000.0760.

Kaplun, D., Sinitca, A., Zamansky, A., Bleuer-Elsner, S., Plazner, M., Fux, A., van der Linden, D., 2019. Animal health informatics: towards a generic framework for automatic behavior analysis. In Proceedings of the 12th International Conference on Health Informatics (HEALTHINF 2019).

Knotek, R.J., Brooks, E.J., Howey, L.A., Gelsleichter, J.G., Talwar, B.S., Winchester, M. M., Jordan, L.K., Williams, S., Bond, M.E., Brooks, A.M., Shipley, O.N., 2022. Merging technologies and supervised classification methods to quantify capture behavior on hook-and-line. J. Exp. Mar. Biol. Ecol. 555, 151782.

Larsen, M.L., 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.

Leonard, S.M., Xin, H., Brown-Brandl, T.M., Ramirez, B.C., 2019. Development and application of an image acquisition system for characterizing sow behaviors in farrowing stalls. Comput. Electron. Agric. 163, 104866.

Leoni, J., Tanelli, M., Strada, S.C., Berger-Wolf, T., 2020. Ethogram-based automatic wild animal monitoring through inertial sensors and GPS data. Ecol. Inform. 59, 101112.

Leroy, T., Vranken, E., Van Brecht, A., Struelens, E., Sonck, B., Berckmans, D., 2006. A computer vision method for on-line behavioral quantification of individually caged poultry. Trans. ASABE 49, 795–802.

Li, D., Chen, Y., Zhang, K., Li, Z., 2019. Mounting behaviour recognition for pigs based on deep learning. Sensors 19, 4924.

Li, J., Green-Miller, A.R., Hu, X., Lucic, A., Mohan, M.M., Dilger, R.N., Condotta, I.C., Aldridge, B., Hart, J.M., Ahuja, N., 2022. Barriers to computer vision applications in pig production facilities. Comput. Electron. Agric. 200, 107227 https://doi.org/10.1016/j.compag.2022.107227.

Li, N., Ren, Z., Li, D., Zeng, L., 2020. Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming. Animal 14 (3), 617–625. https://doi.org/10.1017/S1751731119002155.

Liu, D., Oczak, M., Maschat, K., Baumgartner, J., Pletzer, B., He, D., Norton, T., 2020. A computer vision-based method for spatial-temporal action recognition of tail-biting behaviour in group-housed pigs. Biosyst. Eng. 195, 27–41. https://doi.org/10.1016/j.biosystemseng.2020.04.007.

Marks, M., Jin, Q., Sturman, O., von Ziegler, L., Kollmorgen, S., von der Behrens, W., Mante, V., Bohacek, J., Yanik, M.F., 2022. Deep-learning-based identification, tracking, pose estimation and behaviour classification of interacting primates and mice in complex environments. Nat. Mach. Intell. 4, 331–340.

Mathis, M.W., Mathis, A., 2020. Deep learning tools for the measurement of animal behavior in neuroscience. Curr. Opin. Neurobiol. 60, 1–11.

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.

May, R.J., Maier, H.R., Dandy, G.C., 2010. Data splitting for artificial neural networks using SOM-based stratified sampling. Neural Netw. 23, 283–294.

McLoughlin, M.P., Stewart, R., McElligott, A.G., 2019. Automated bioacoustics: methods in ecology and conservation and their potential for animal welfare monitoring. J. R. Soc. Interface 16, 20190225.
McVey, C., Hsieh, F., Manriquez, D., Pinedo, P., Horback, K., 2023. Invited Review: Applications of unsupervised machine learning in livestock behavior: Case studies in recovering unanticipated behavioral patterns from precision livestock farming data streams. Appl. Anim. Sci. 39, 99–116. https://doi.org/10.15232/aas.2022-02335.
Mishra, B.K., Kumar, R. (Eds.), 2020. Natural Language Processing in Artificial Intelligence. CRC Press.
Nasirahmadi, A., Edwards, S.A., Sturm, B., 2017. Implementation of machine vision for detecting behaviour of cattle and pigs. Livest. Sci. 202, 25–38. https://doi.org/10.1016/j.livsci.2017.05.014.
Nasirahmadi, A., Sturm, B., Edwards, S., Jeppsson, K.H., Olsson, A.C., Müller, S., Hensel, O., 2019. Deep learning and machine vision approaches for posture detection of individual pigs. Sensors 19, 3738. https://doi.org/10.3390/s19173738.
Noldus, L.P., Spink, A.J., Tegelenbosch, R.A., 2001. EthoVision: a versatile video tracking system for automation of behavioral experiments. Behav. Res. Methods, Instrum., Comput. 33, 398–414.
Noldus, L.P.J.J., Jansen, R.G., 2004. Measuring broiler chicken behaviour and welfare: prospects for automation. Measuring and Auditing Broiler Welfare. CABI Publishing,, Wallingford UK, pp. 267–290.
Norton, T., Chen, C., Larsen, M.L.V., Berckmans, D., 2019. Precision livestock farming: building 'digital representations' to bring the animals closer to the farmer. Animal 13, 3009–3017.
O'Connell, N.E., Beattie, V.E., Watt, D., 2005. Influence of regrouping strategy on performance, behaviour and carcass parameters in pigs. Livest. Prod. Sci. 97, 107–115.
O'Malley, C.I., Steibel, J.P., Bates, R.O., Ernst, C.W., Siegford, J.M., 2022. The social life of pigs: changes in affiliative and agonistic behaviors following mixing. Animals 12, 206. https://doi.org/10.3390/ani12020206.
Ojo, R.O., Ajayi, A.O., Owolabi, H.A., Oyedele, L.O., Akanbi, L.A., 2022. Internet of things and machine learning techniques in poultry health and welfare management: a systematic literature review. Comput. Electron. Agric. 200, 107266 https://doi.org/10.1016/j.compag.2022.107266.
Okinda, C., Nyalala, I., Korohou, T., Okinda, C., Wang, J., Achieng, T., Wamalwa, P., Mang, T., Shen, M., 2020. A review on computer vision systems in monitoring of poultry: a welfare perspective. Artif. Intell. Agric. 4, 184–208. https://doi.org/10.1016/j.aiia.2020.09.002.
Olejnik, K., Popiela, E., Opaliński, S., 2022. Emerging precision management methods in poultry sector. Agriculture 12 (5), 718. https://doi.org/10.3390/agriculture12050718.
Olsson, I.A.S., Keeling, L.J., 2005. Why in earth? Dustbathing behaviour in jungle and domestic fowl reviewed from a Tinbergian and animal welfare perspective. Appl. Anim. Behav. Sci. 93, 259–282.
Packard, J.M., Folse, L.J., Stone, N.D., Makela, M.E., Coulson, R.N., 1990. Applications of artificial intelligence to animal behavior. In: Bekoff, M., Jamieson, D. (Eds.), Interpretation and Explanation in the Study of Animal Behavior. Routledge, pp. 147–191.
Panadeiro, V., Rodriguez, A., Henry, J., Wlodkowic, D., Andersson, M., 2021. A review of 28 free animal-tracking software applications: current features and limitations. Lab Anim. 50, 246–254.
Prunier, A., Averos, X., Dimitrov, I., Edwards, S.A., Hillmann, E., Holinger, M., Ilieski, V., Leming, R., Tallet, C., Turner, S.P., Zupan, M., 2020. Early life predisposing factors for biting in pigs. Animal 14, 570–587.
Reitermanova, Z., 2010. Data splitting, in: WDS' 10 Proceedings of Contributed Papers, Part I, Matfyzpress, Prague, 10, 31–36.
Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40, 913–929.
Rowe, E., Dawkins, M.S., Gebhardt-Henrich, S.G., 2019. A systematic review of precision livestock farming in the poultry sector: iIs technology focussed on improving bird welfare? Animals 9 (9), 614. https://doi.org/10.3390/ani9090614.
Rufener, C., Berezowski, J., Maximiano Sousa, F., Abreu, Y., Asher, L., Toscano, M.J., 2018. Finding hens in a haystack: consistency of movement patterns within and across individual laying hens maintained in large groups. Sci. Rep. 8, 12303.
Rushen, J., Chapinal, N., de Passile, A.M., 2012. Automated monitoring of behavioural-based animal welfare indicators. Anim. Welf. 21, 339–350.
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. https://doi.org/10.3390/sym11020266.
Schmidt, T.B., Lancaster, J.M., Psota, E., Mote, B.E., Hulbert, L.E., Holliday, A., Woiwode, R., Pérez, L.C., 2022. Evaluation of a novel computer vision-based livestock monitoring system to identify and track specific behaviors of individual nursery pigs within a group-housed environment. Transl. Anim. Sci. 6, txac082.
Shen, H., Nugegoda, D., 2022. Real-time automated behavioural monitoring of mussels during contaminant exposures using an improved microcontroller-based device. Sci. Total Environ. 806, 150567.
Sonoda, L.T., Fels, M., Oczak, M., Vranken, E., Ismayilova, G., Guarino, M., Viazzi, S., Bahr, C., Berckmans, D., Hartung, J., 2013. Tail Biting in pigs—causes and management intervention strategies to reduce the behavioural disorder. A review. Berl. und Munch. Tierarztl. Wochenschr. 126, 104–112.
Taylor, N.R., Main, D.C., Mendl, M., Edwards, S.A., 2010. Tail-biting: a new perspective. Vet. J. 186, 137–147.
Tecott, L.H., Nestler, E.J., 2004. Neurobehavioral assessment in the information age. Nat. Neurosci. 7, 462–466.
Tuia, D., Kellenberger, B., Beery, S., Costelloe, B.R., Zuffi, S., Risse, B., Mathis, A., Mathis, M.W., van Langevelde, F., Burghardt, T., Kays, R., 2022. Perspectives in machine learning for wildlife conservation. Nat. Commun. 13, 792.
Turner, S.P., Farnworth, M.J., White, I.M.S., Brotherstone, S., Mendl, M., Knap, P., Penny, P., Lawrence, A.B., 2006. The accumulation of skin lesions and their use as a pre- dictor of individual aggressiveness in pigs. Appl. Anim. Behav. Sci. 96, 245–259.
Valletta, J.J., Torney, C., Kings, M., Thornton, A., Madden, J., 2017. Applications of machine learning in animal behaviour studies. Anim. Behav. 124, 203–220. https://doi.org/10.1016/j.anbehav.2016.12.005.
Vieira Rios, H., Waquil, P.D., Soster de Carvalho, P., Norton, T., 2020. How are information technologies addressing broiler welfare? A Syst. Rev. Based Welf. quality® Assess. Sustain. 12, 1413. https://doi.org/10.3390/su12041413.
Voikar, V., Gaburro, S., 2020. Three pillars of automated home-cage phenotyping of mice: novel findings, refinement, and reproducibility based on literature and experience. Front. Behav. Neurosci. 14, 575434.
Weinstein, B.G., 2018. A computer vision for animal ecology. J. Anim. Ecol. 87, 533–545.
Wu, D., Cui, D., Zhou, M., Ying, Y., 2022. Information perception in modern poultry farming: a review. Comput. Electron. Agric. 199, 107131 https://doi.org/10.1016/j.compag.2022.107131.
Wurtz, K., Camerlink, I., D'Eath, R.B., Peña Fernández, A., Norton, T., Steibel, J., Siegford, J., 2019. Recording behaviour of indoor-housed farm animals automatically using machine vision technology: A systematic review. PLoS ONE 14, e0226669. https://doi.org/10.1371/journal.pone.0226669.
Wurtz, K., Norton, T., Siegford, J., Steibel, J., 2022. Assessment of open-source programs for automated tracking of individual pigs within a group. In: Banhazi, T., Halas, V., Maroto-Molina, F. (Eds.), Practical Precision Livestock Farming: Hands-on Experiences with PLF Technologies in Commercial and R&D Settings. Wageningen Academic Publishers, The Netherlands, pp. 213–230.
Yang, Q., Xiao, D., 2020. A review of video-based pig behavior recognition. Appl. Anim. Behav. Sci. 233, 105146 https://doi.org/10.1016/j.applanim.2020.105146.
Yang, Q., Xiao, D., Cai, J., 2021. Pig mounting behaviour recognition based on video spatial–temporal features. Biosyst. Eng. 206, 55–66. https://doi.org/10.1016/j.biosystemseng.2021.03.011.
Zamansky, A., Sinitca, A., van der Linden, D., Kaplun, D., 2021. Automatic animal behavior analysis: opportunities for combining knowledge representation with machine learning. Procedia Comput. Sci. 186, 661–668. https://doi.org/10.1016/j.procs.2021.04.187.
Zhang, K., Li, D., Huang, J., Chen, Y., 2020. Automated video behavior recognition of pigs using two-stream convolutional networks. Sensors 20, 1085. https://doi.org/10.3390/s20041085.
Zhang, Z., Zhang, H., He, Y., Liu, T., 2022. A Review in the automatic detection of pigs behavior with sensors. J. Sens. 2022, 4519539 https://doi.org/10.1155/2022/4519539.