frontiers | Frontiers in Animal Science

TYPE Original Research
PUBLISHED 14 December 2022
DOI 10.3389/fanim.2022.846893

Check for updates

OPEN ACCESS

EDITED BY
Julio Giordano,
Cornell University, United States

REVIEWED BY
Tiago Bresolin,
University of Illinois at Urbana-Champaign, United States
Bradley Heins,
University of Minnesota Twin Cities, United States
Jacquelyn Boerman,
Purdue University, United States

*CORRESPONDENCE
John W.M. Bastiaansen
john.bastiaansen@wur.nl

SPECIALTY SECTION
This article was submitted to Precision Livestock Farming, a section of the journal Frontiers in Animal Science

RECEIVED 31 December 2021
ACCEPTED 17 November 2022
PUBLISHED 14 December 2022

CITATION
Bastiaansen JWM, Hulsegge I, Schokker D, Ellen ED, Klandermans B, Taghavi M and Kamphuis C (2022) Continuous real-time cow identification by reading ear tags from live-stream video.
Front. Anim. Sci. 3:846893.
doi: 10.3389/fanim.2022.846893

COPYRIGHT
© 2022 Bastiaansen, Hulsegge, Schokker, Ellen, Klandermans, Taghavi and Kamphuis. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Continuous real-time cow identification by reading ear tags from live-stream video

John W.M. Bastiaansen*, Ina Hulsegge, Dirkjan Schokker, Esther D. Ellen, Bert Klandermans, Marjaneh Taghavi and Claudia Kamphuis

Animal Breeding and Genomics, Wageningen University & Research, Wageningen, Netherlands

In precision dairy farming there is a need for continuous and real-time availability of data on cows and systems. Data collection using sensors is becoming more common and it can be difficult to connect sensor measurements to the identification of the individual cow that was measured. Cows can be identified by RFID tags, but ear tags with identification numbers are more widely used. Here we describe a system that makes the ear tag identification of the cow continuously available from a live-stream video so that this information can be added to other data streams that are collected in real-time. An ear tag reading model was implemented by retraining and existing model, and tested for accuracy of reading the digits on cows ear tag images obtained from two dairy farms. The ear tag reading model was then combined with a video set up in a milking robot on a dairy farm, where the identification by the milking robot was considered ground-truth. The system is reporting ear tag numbers obtained from live-stream video in real-time. Retraining a model using a small set of 750 images of ear tags increased the digit level accuracy to 87% in the test set. This compares to 80% accuracy obtained with the starting model trained on images of house numbers only. The ear tag numbers reported by real-time analysis of live-stream video identified the right cow 93% of the time. Precision and sensitivity were lower, with 65% and 41%, respectively, meaning that 41% of all cow visits to the milking robot were detected with the correct cow's ear tag number. Further improvement in sensitivity needs to be investigated but when ear tag numbers are reported they are correct 93% of the time which is a promising starting point for future system improvements.

KEYWORDS

deep learning, image analysis, precision farming, animal identification, number recognition

1 Introduction

With the developments in precision dairy farming, on-farm data streams are increasing rapidly (Kamphuis et al., 2012; Cerri et al., 2020). Data are collected by devices such as neck collars (Rutten et al., 2017), by farm equipment such as milking systems and feeding stalls (Foris et al., 2019; Deng et al., 2020), and from measuring devices that analyze milk or other cow output such as manure and breath air (Negussie et al., 2017). When sensors are attached to the cow they will collect data on that specific animal (e.g., activity) and the connection of these data to the cow identification is easily controlled by properly recording which sensor is fitted to which cow. However, data is desired in many locations that can be inside the barn or outside in pasture. For a sensor that is continuously measuring in a fixed location it is difficult to connect the collected data to a specific cow. For instance, when a sensor is measuring breath methane (Huhtanen et al., 2015) in a barn where cows are moving around, it is difficult to know which cow is producing the methane being measured by the sensor at a specific moment (van Breukelen et al., 2022).

Individual identification of dairy cows in the Netherlands is possible by the mandatory ear tags that all farm animals have. In addition, lactating dairy cows are typically fitted with Radio-Frequency IDentification (RFID) transponders. Standard farm equipment, like milking robots and feeding stations, identify which cow is present by reading this transponder. Often this transponder information is stored, but accessing and matching this transponder identification to data from other sources is not always straightforward, and making the cow identification available in real-time is very difficult or impossible. Identification by reading the ear tags with imaging equipment would be a universal solution in precision dairy farming, both for places where RFID readers are not present, or for animals that are not fitted with RFID transponders such as dry cows or young animals. A method for reading ear tag numbers from images was described by Ilestrand (2017) who tested several optical character recognition methods. Zin et al. (2020) used selected images from 4k videos and developed an algorithm for head detection, ear tag detection, and reading the ear tag number using a Convolutional Neural Network (CNN). These methods have been applied to static datasets of previously collected images. To monitor or manage a dairy farm with precision, a system is needed that continuously reads the ear tag number and makes this cow identification available in combination with other data that are collected in real-time.

Real-time data on the production of milk, methane, or manure, or observed deviations from expected production or expected behavior, may lead to alerts to inspect the cow, or to immediate interventions in the feeding regime or other management factors for a specific cow. Also, camera systems are already studied to extract data for precision dairy farming by monitoring locomotion (Van Hertem et al., 2018; Russello et al., 2022) or assessing individual feed intake (Lassen et al., 2022). In these cases identification by reading ear tag numbers may not require additional hardware, only additional software to read the ear tag numbers from the images. Precision dairy farming is all about retrieving information about individual cows. Cow identification by reading ear tag numbers allows for quick adaptable and real-time cow identification. Addition of identification to image based data collection allows targeted interventions at cow level. The aim of this study was to develop a system that makes the ear tag identification of the cow continuously available from a live-stream video and to test the performance of such a system on a commercially operated dairy farm.

2 Materials and methods

Section 2.1 describes the development of the ear tag detection and digit recognition model using images of ear tags collected at two commercial Dutch dairy farms. Section 2.2 describes the infrastructure that was set up to test the model performance for predicting ear tag numbers in real-time from live-stream video on a different dairy farm at the Wageningen University & Research Dairy Campus. No specific animal care protocol was needed for this study. The data collection was done by video equipment installed in commercially operated dairy farms without interaction with the cows present.

2.1 Ear tag detection and digit recognition

The ear tag reading model was developed on images extracted from recorded video footage. The recording was produced with an experimental setup at two commercial dairy farms in the Netherlands with 90 and 95 milking cows respectively. All cows on these farms were fitted with ear tags that display a black number on a yellow background, as part of the national identification and registration program. A video station with a Foscam FI9901 camera (Shenzhen FOSCAM Intelligent Technology Co., China) was installed inside a milking robot on these two farms. The milking robot unit was chosen as the location to test the video station because all milking cows are known to pass through here several times per day. The video stream content was cached in Audio Video

Interleave (AVI) format and stored on a Solid State Drive (SSD). The ear tag reading model was developed and implemented in Python 3.8 using the Open Computer Vision Library (OpenCV 4.1.2). The video stream was read using the OpenCV command `VideoCapture` which transmits every frame of the video as an image for processing. Images are processed in three consecutive steps: (1) detecting the presence of an ear tag and segmenting the ear tag image from the image; (2) processing the ear tag image; and (3) digit number recognition (Figure 1). Each of these steps are described in more detail below.

2.1.1 Detecting and segmenting ear tag images

Ten images were extracted from the video per second. Detecting the presence or absence of an ear tag in an image was based on the presence of the yellow background color of the ear tag. To extract the color, each image was converted from Red,

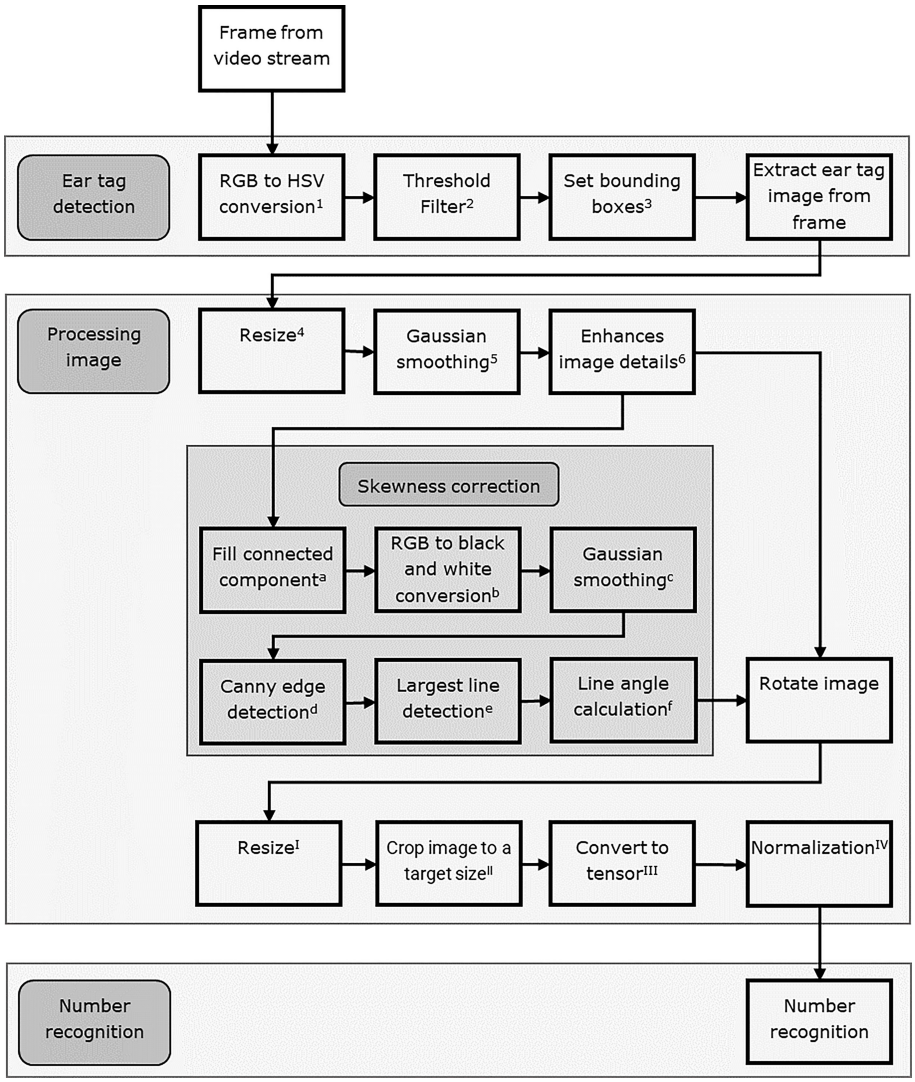

FIGURE 1
Steps in ear tag recognition algorithm. Applied functions (openCV python): [1]cvtColor(frame, cv2.COLOR_BGR2HSV), [2]inRange(frame, [22, 95, 90], [40, 255, 255]) and threshold(image, 127, 255, 0), [3]findContours(frame, cv2.RETR_TREE, cv2.CHAIN_APPROX_SIMPLE), [4]resize(crop_img, dim, interpolation = cv2.INTER_LINEAR), [5]GaussianBlur(image, (5, 5), 0), [6]detailEnhance(image, sigma_s = 20, sigma_r = 0.10), [a]floodFill(image, [height + 2, width + 2], (100, 100), (0, 255, 255), (60, 60, 110), (70, 150,200), cv2.FLOODFILL_FIXED_RANGE), [b]cvtColor(image, cv2.COLOR_BGR2GRAY) and threshold(grayImage, 127, 255, cv2.THRESH_BINARY), [c]GaussianBlur(blackAndWhiteimage, (21, 21), 2), [d]Canny (blackAndWhiteImage, threshold1 = 20, threshold2 = 30,apertureSize = 7), [e]HoughLinesP(edges, 1, np.pi/180, 30, minLineLength = 50, maxLineGap = 50), [f]rad2deg(np.arctan2(y2 - y1, x2 - x1), [I]transforms.Compose([transforms.Resize([64, 64]), [II]transforms.CenterCrop([54, 54]), [III]transforms.ToTensor(), [IV]transforms.Normalize([0.5, 0.5, 0.5], [0.5, 0.5, 0.5]).

Green, Blue (RGB) to Hue, Saturation, Value (HSV) color space. Yellow pixels in the converted image were identified by having Hue between 27 and 40, Saturation between 95 and 255, and Value between 90 and 255. A rectangular bounding box was drawn around the yellow pixels by the model, to identify the area of the ear tag. When the bounding box had a minimum width and height of 50 pixels it was considered to have detected an ear tag in the image. This threshold of 50 pixels was determined empirically and is specific for the current setup of video resolution and the distance between the camera and the head of the cow. A detected ear tag was segmented from the image by extracting the area within the bounding box. This segmented area is subsequently called the "ear tag image" and submitted to the ear tag image processing step.

2.1.2 Processing ear tag images

Pre-processing the ear tag images consisted of applying Gaussian smoothing and image details enhancement using *cv2.GaussianBlur* and *cv2.detailEnhance* functions, respectively. To correct skewness of the ear tag in the image, flood fill was applied (*cv2.floodFill*) to fill the holes in the pre-processed image with the yellow color of the ear tag. The image was then converted to a gray scale image using *cv2.cvtColor* and to a black and white image using *cv2.threshold*. Gaussian smoothing and Canny edge detection were applied to the black and white image using *cv2.GaussianBlur* and *cv2.Canny* respectively, to find the contours of the ear tag. The contours were necessary to rotate each ear tag image such that its baseline was horizontal, allowing for effective number recognition. The longest straight line in the shape of the ear tag was used to find the base of the ear tag. A probabilistic Hough transform was applied using *cv2.HoughLinesP* to identify the baseline. For number recognition, the images were rotated, resized, cropped and normalized using the *Torchvision.transforms* functions *ToTensor*, *resize*, *CenterCrop*, and *Normalize*, respectively. Further details related to the image pre-processing are shown in Figure 1, including the parameter settings for all the functions.

2.1.3 Ear tag number recognition

The ear tag image processing steps resulted in images that were formatted to 64 by 63 pixels. A set of 750 of these ear tag images were randomly selected for developing the ear tag number recognition model. This set included ear tag images from 140 different cows. Each digit on the selected ear tag images was annotated with the correct value using labelImg (https://github.com/tzutalin/labelImg), resulting in 3,000 annotated digits (i.e. 4 digits per image). These annotated ear tag images were randomly divided in a training set (n = 600 images; 2,400 digits), a validation set (n = 75 images; 300 digits) and a test set (n = 75 images; 300 digits).

The model for recognition of digits in the processed ear tag images was developed by retraining an existing house number model (model-54000.pth, Goodfellow et al., 2013). This deep CNN model was trained on the Street View House Number dataset (Netzer et al., 2011). This model was chosen because of the similarity of house numbers to ear tag numbers. The house number model was retrained using the training and validation set of ear tag images. The retraining was performed using a batch size of 32 with the stochastic gradient descent (SGD) optimizer (Yang and Yang, 2018) with learning rate set to 0.01, momentum term set to 0.9 and weight decay set to 0.0005. A step decaying learning rate scheduler (StepLR) was used with a step size of 10.000 epochs and gamma of 0.9. The loss function employed during retraining was cross-entropy loss. The early stopping patience was set to 100 epochs, meaning that the retraining stopped when no improvement was observed in the last 100 epochs. Full details of the model are described in the github repository (https://github.com/potterhsu/SVHNClassifier-PyTorch/blob/master/README.md).

To assess the improvement of the existing model when retrained on ear tag images, we first applied the pretrained model-54000.pth without retraining, to predict the 4 digits on the 75 ear tags in the test set. Then, the model-54000.pth was retrained using the training and validation set of ear tag images. The retrained model was subsequently applied to predict the 4 digits on the 75 ear tags in the test set. For both models the probability threshold for reporting a predicted digit was set to 0.95. The performance of the pretrained and retrained models were compared based on the number of correctly identified digits per ear tag.

2.2 Model implementation and evaluation in live-stream video

The ear tag recognition system including the detection, processing, and number recognition step with the retrained model was installed on a Raspberry Pi 4 Model B in the experimental setup at the Wageningen University & Research facility Dairy Campus (Leeuwarden, the Netherlands). The Raspberry Pi was connected to a DS-2CD2T43G0 camera (Hikvision, Hangzhou, China) through ethernet (LAN) connection. The camera was installed in the milking robot unit such that when a cow is present in the milking robot the camera records the head from above and behind (See Figure 2 for the view of the camera). The model analyzed the live-stream video and continuously reported ear tag numbers in real-time when an ear tag was detected. These reported ear tag numbers were transmitted through an Unshielded Twisted Pair (UTP) cable into the local network. The reported ear tag numbers were recorded with a timestamp for a 40-hour period between March 13, 2022, 8:00 am and March 14, 23:59 pm. During the same time period the milking robot was recording the RFID transponder number and the time when a cow entered the robot. Based on the transponder numbers, the ear tag numbers

FIGURE 2
Direction of view from the video camera in the milking robot.

of cows that entered the robot during the test period were retrieved from the farm management system and added to a list called the "cow list". The cow list contains the ear tag numbers from all the cows that were present in the experiment.

The ability to correctly identify a cow was evaluated based on the comparison between the predicted ear tag number from the image and the known ear tag number of the cow from the RFID transponder readings of the milking robot. To evaluate the reported ear tag numbers, the numbers transmitted by the model were first filtered using the cow list. Ear tag numbers where one or more of the 4 digits was not predicted and mistakes in digit prediction that resulted in an ear tag number that is not on the cow list were not reported. Each milking robot visit was classified as either a True Positive (TP) when only the correct ear tag number was reported, a False Positive (FP) when incorrect ear tag number(s) were reported, or a False Negative (FN) when no ear tag number was reported. False positives included the visits with a mix of correct and incorrect ear tag numbers. The *precision* of the system was defined as the proportion of milking robot visits with a TP result out of all the visits with a positive result, i.e. TP/(TP+FP). The *sensitivity* of the system was defined as the proportion of all the cow visits to the robot for which the correct ear tag was reported by the model, i.e. TP/(TP+FP+FN).

TABLE 1 Number of ear tag images with 0 to 4 correctly identified digits in the test set.

Correct digits[1]	Pretrained	Retrained
4	36	48
3	21	19
2	14	3
1	4	5
0	–	–
Accuracy[2]	79.7%	86.7%

Pretrained model pretrained with house numbers, or retrained with ear tag images. [1]Number of correctly identified digits, [2]percentage of correctly identified digits.

3 Results

3.1 Digit recognition model

Image processing segmented the ear tag area from each of the 750 images in the development dataset. In the test dataset the pretrained house number model had an accuracy of digit prediction of 80% (Table 1). The retrained model predicted 87% of the 300 digits on the 75 ear tag images (i.e. 4 digits predicted per ear tag) in the test dataset correctly (Table 1), and 98% was predicted correctly in the validation dataset. With the retrained model, 48 out of 75 ear tags (65%) in the test dataset had all four digits predicted correctly, compared to 36 out of 75 ear tags (48%) for the pretrained model. Neither of the two models resulted in a completely failed prediction where none of the 4 digits in an ear tag number were predicted correctly (Table 1).

3.2 Model evaluation in live-stream video

During the 40 hour test period the milking robot recorded 248 separate visits made by 60 unique cows. The cow list was compiled of the ear tags of these 60 unique cows. The median number of visits per cow was 4, with a range of 1 to 6 visits.

In the 40 hour test period, an ear tag number with 4 recognized digits was reported 25,448 times or 102.6 numbers reported per visit. Of these reported numbers 11,021 (43%) were found to match an ear tag in the cow list. Based on the reported ear tag numbers, 214 separate visits to the milking robot by 52 different cows from the 60 cows on the cow list were predicted.

Out of the 11,021 reported ear tag numbers 10,265 (93%) corresponded to the cow that was present in the milking robot according to the recorded transponder number. The 11,021 ear tag numbers were reported during 156 out of the 248 actual visits to the milking robot. This means that 37% of the visits to the milking robot were not detected by the system. Out of the 248 milking robot visits 102 resulted in a TP outcome, 47 in a FP outcome, and 92 in a FN outcome. The sensitivity of the system to detect the correct ear tag during a visit was 102 out of 248 visit or 41% and the precision was 102 out of 156 detected visits or 65%.

4 Discussion

A system was developed and implemented in the milking robot unit at Dairy Campus (Leeuwarden, The Netherlands) that is producing a real-time and continuous reading of ear tag identification numbers from live-stream video. The system was build using off-the-shelf video and networking equipment. The ear tag numbers reported by the system were correct 93% of the

time. While the reported ear tag numbers were very often correct, the sensitivity for detecting the correct ear tag of the cow present during a milking event was much lower at 41%.

The number recognition model applied here was obtained by retraining. The model-54000.pth (Goodfellow et al., 2013) was used as the pretrained model. A pretrained model based on street numbers was expected to be a good choice given our similar objective of reading numbers from an object within a varied background. Retraining was found successful because digit level accuracy was improved from 80% to 87% with a small dataset of the target images consisting of only 675 ear tags. Obtaining a model with 87% digit level accuracy with a dataset that is less than 0.5% the size of the dataset that was used for the pretrained model showed that retraining should be preferred in this situation. Training of the pretrained model took approximately 6 days (Goodfellow et al., 2013) while retraining was very quick (< 5 hours) which shows the value of retraining for limiting the computational requirements. Most importantly, the alternative of capturing and annotating a much larger dataset to train the model from scratch would require a prohibitive amount of work. Annotating a dataset of 200k ear tag images at a high speed of 1 digit per second would already take 27.8 working days.

Before the number recognition model is applied, the system captures ear tag images from a video live-stream. Clearly, identifying the ear tag in the images is an important requirement for the system to work. The current system uses the ear tag color to identify it, and the shape (straight bottom) to rotate the images to the desired orientation. While ear tags come in many colors, the dairy cattle approved ear tags in the Netherlands are all yellow. A change in color will require a modification of currently implemented ear tag detection algorithm. For ear tags of different colors modifying the color thresholds will be easy to implement with some tuning of these parameters. Other studies have used alternative methods to find the ear tag or recognize digits such as template matching (Ilestrand, 2017; Zin et al., 2020). For instance, Zin et al. (2020), used the shape of the head to find the location of ear tags, combined with color detection in a following step. They found 100% accuracy in detection cow heads and subsequently 92.5% accuracy for digit recognition which is similar but higher than the 87% found in this study for digit recognition. Ilestrand (2017) shows accuracy for different subsets of images, with the best images in the "A data" resulting in digit recognition accuracy of 98% and the poor images in the "C data" resulting in 81% accuracy using a Support Vector Machine (SVM) approach. Template matching for digit recognition performed worse than the SVM with 94% and 67% accuracy in the A data and C data, respectively.

Colors of ear tags can vary and therefore it may seem beneficial to have an algorithm that doesn't rely on specific colors. However, we do recommend the use of color because it is a very distinctive feature of the ear tags, and color ranges in the ear tag reading algorithm can easily be adjusted to those used on a specific farm. Technical differences between ear tags can be addressed by small changes to the algorithm when they are fixed differences between farms. However the system will also encounter variation within farms due to environmental conditions. Both Ilestrand (2017) and Zin et al. (2020) discuss variation in lighting conditions and the presence of dirt, hair or other objects that cause occlusion of ear tags. Ilestrand (2017) reports results for the best images and poor images separately and shows large differences, especially in the ability to correctly read a complete ear tag number. Lighting conditions in our current study are fairly standardized because artificial light is on for 24 hours per day in the milking robot. However, outside conditions such as bright sunlight or reflection of light can affect the performance of the ear tag reading models. An important factor for the system presented in this study is the visibility of the ear tag within the image. The camera is currently positioned behind the cow (Figure 2) and therefore the position of the cows head can easily cause occlusion of the ear tag.

The digit level accuracy of 87% achieved with the retrained number recognition model is quite a bit higher than accuracy with the pretrained model on the same test set (80%). However, higher accuracies are reported in other studies. For example, the pretrained model was reported to have a 97.8% digit level accuracy when applied to street numbers (Goodfellow at al., 2013). Achieving such a high accuracy may only be possible with very large training datasets on the target images. However, our aim was not to obtain high digit level accuracy but rather to identify the correct cow by its ear tag number, so we aim for ear tag number accuracy. Therefore we focused on improvements of the processes that are downstream from the number recognition model. For example, we applied a comparison of the predicted ear tag number with a list of currently present ear tags on the farm compiled in the cow list. Using this cow list, it is easy to filter out a part of the erroneous predictions. Zin et al. (2020) described this same process in their "Ear tag confirmation process". For a dairy farm with a finite number of ear tags present, this step can be very effective. The accuracy of getting the correct 4 digit ear tag number when an ear tag is detected was already at 93% in the current system. Further improvements of the downstream process still need to be investigated. During a cow visit to the milking robot many images are processed and for each image the predicted ear tag number can be compared to those predicted from previous images for consistency. Such a system could be developed by using a majority vote of the prediction in a sliding window of consecutive images. A voting system needs to be tuned such that a real shift from one cow to another is detected while it is robust against temporary differences from errors in reading the ear tags. The aforementioned potential improvements still have to be

investigated but we expect that these will resolve at least some of the 35 visits where both correct and incorrect ear tag numbers were reported to a unique and correctly assigned ear tag number.

Precision dairy farming benefits from and may even require up to date or real-time information on cows. Collecting this information from automated systems is rapidly increasing. Cow identification from ear tag reading will especially be useful in situations where other identification is not easily implemented. Other approaches have been reported using images to recognize individual cows from their appearance (Shen et al., 2020). The disadvantage of such a system is that it is more likely to work for cow breeds with a pattern (e.g., Holstein Frisian), but perform more poorly for breeds that are unicolored (e.g., Jersey). More importantly, a model based on coat pattern or other appearance traits needs to be trained for recognizing each new cow entering the herd, whereas the ear tag model stays the same and reads new numbers with equal accuracy.

The system described here is providing real-time ear tag numbers from a video live-stream in a commercially operating dairy farm. The use of retraining, in combination with a simple technical setup of a camera and a small single board computer, provides an easy to implement addition to other equipment for identifying cows in real-time and at a low cost. The sensitivity of the system is 41%, which may be too low for practical implementation at this point. However, data from the installed system will be used to further optimize the camera positions and other technical parameters. Because the accuracy of the reported ear tag numbers is already high at 93% most improvements are expected from changing the video capturing setup to obtain the right images for the retrained number recognition model. In addition, the accuracy of reported ear tag numbers will also benefit further from improvements in the downstream processing.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the animal study because No specific animal care protocol was needed for this study. The data collection was done by video equipment installed in commercially operated dairy farms without interaction with the cows present.

Author contributions

CK designed and coordinated the study. EE and BK provided data. CK, IH, DS, JB and EE contributed to the aims and approaches. IH, CK, BK, MT, and DS developed the methodology and IH and MT carried out the modelling and analysis. JB drafted the manuscript. All authors contributed to the writing and read and approved the final manuscript.

Funding

This research was conducted by Wageningen Livestock Research, commissioned and funded by the Ministry of Agriculture, Nature and Food Quality, within the framework of Policy Support Research theme "Data driven & High Tech" (project number KB-38-001-004).

Acknowledgments

We acknowledge Esther van der Heide, Manya Alfonso, and Gerrit Seigers for their contributions in the early stages of this project.

Conflict of interest

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Cerri, R. L. A., Burnett, T. A., Madureira, A. M. L., Silper, B. F., Denis-Robichaud, J., LeBlanc, S., et al. (2020). Symposium review: Linking activity-sensor data and physiology to improve dairy cow fertility. *J. Dairy Sci.* 104 (1), 1220–1231. doi: 10.3168/jds.2019-17893

Deng, Z., Hogeveen, H., Lam, T. J., van der Tol, R., and Koop, G. (2020). Performance of online somatic cell count estimation in automatic milking systems. *Front. Veterinary Sci.* 7. doi: 10.3389/fvets.2020.00221

Foris, B., Thompson, A. J., von Keyserlingk, M. A. G., Melzer, N., and Weary, D. M. (2019). Automatic detection of feeding-and drinking-related agonistic behavior and dominance in dairy cows. *J. dairy Sci.* 102 (10), 9176–9186. doi: 10.3168/jds.2019-16697

Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., and Shet, V. (2013). "Multi-digit number recognition from street view imagery using deep convolutional neural networks," in *arXiv preprint*. https://arxiv.org/abs/1312.6082

Huhtanen, P., Cabezas-Garcia, E. H., Utsumi, S., and Zimmerman, S. (2015). Comparison of methods to determine methane emissions from dairy cows in farm conditions. *J. Dairy Sci.* 98 (5), 3394–3409. doi: 10.3168/jds.2014-9118

Ilestrand, M. (2017). Automatic eartag recognition on dairy cows in real barn environment. master's thesis (Linköping, Sweden: Linköping University).

Kamphuis, C., DelaRue, B., Burke, C. R., and Jago, J. (2012). Field evaluation of 2 collar-mounted activity meters for detecting cows in estrus on a large pasture-grazed dairy farm. *J. Dairy Sci.* 95, 3045–3056. doi: 10.3168/jds.2011-4934

Lassen, J., Thomasen, J. R., and Borchersen, S. (2022)CFIT–cattle feed InTake–a 3D camera based system to measure individual feed intake and predict body weight in commercial farms. In: *Proc. of the 12th WCGALP* (Rotterdam, The Netherlands). Available at: https://www.wageningenacademic.com/pb-assets/wagen/WCGALP2022/10_011.pdf (Accessed October 2, 2020).

Negussie, E., de Haas, Y., Dehareng, F., Dewhurst, R. J., Dijkstra, J., Gengler, N., et al. (2017). Invited review: Large-scale indirect measurements for enteric methane emissions in dairy cattle: A review of proxies and their potential for use in management and breeding decisions. *J. Dairy Sci.* 100 (4), 2433–2453. doi: 10.3168/jds.2016-12030

Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., and Ng, A. Y. (2011). "Reading digits in natural images with unsupervised feature learning," in *In NIPS workshop on deep learning and unsupervised feature learning*. https://research.google/pubs/pub37648.pdf

Russello, H., van der Tol, R., and Kootstra, G. (2022). T-LEAP: Occlusion-robust pose estimation of walking cows using temporal information. *Comput. Electron. Agric.* 192, 106559. doi: 10.1016/j.compag.2021.106559

Rutten, C. J., Kamphuis, C., Hogeveen, H., Huijps, K., Nielen, M., and Steeneveld, W. (2017). Sensor data on cow activity, rumination, and ear temperature improve prediction of the start of calving in dairy cows. *Comput. Electron. Agric.* 132, 108–118. doi: 10.1016/j.compag.2016.11.009

Shen, W., Hu, H., Dai, B., Wei, X., Sun, J., Jiang, L., et al. (2020). Individual identification of dairy cows based on convolutional neural networks. *Multimedia Tools Appl.* 79 (21), 14711–14724. doi: 10.1007/s11042-019-7344-7

van Breukelen, A. E., Aldridge, M. A., Veerkamp, R. F., and de Haas, Y. (2022). Genetic parameters for repeatedly recorded enteric methane concentrations of dairy cows. *J. Dairy Sci.* 105 (5), 4256–4271. doi: 10.3168/jds.2021-21420

Van Hertem, T., Tello, A. S., Viazzi, S., Steensels, M., Bahr, C., Romanini, C. E. B., et al. (2018). Implementation of an automatic 3D vision monitor for dairy cow locomotion in a commercial farm. *Biosyst. Eng.* 173, 166–175. doi: 10.1016/j.biosystemseng.2017.08.011

Yang, J., and Yang, G. (2018). Modified convolutional neural network based on dropout and the stochastic gradient descent optimizer. *Algorithms* 11 (3), 28. doi: 10.3390/a11030028

Zin, T. T., Pwint, M. Z., Seint, P. T., Thant, S., Misawa, S., Sumi, K., et al. (2020). Automatic cow location tracking system using ear tag visual analysis. *Sensors* 20 (12), 3564. doi: 10.3390/s20123564