



OPEN ACCESS

EDITED BY

Nursen Ozturk,
Istanbul University-Cerrahpasa, Türkiye

REVIEWED BY

Ahmet Çağdaş Seçkin,
Adnan Menderes University, Türkiye
Olalekan Chris Akinsulie,
Washington State University, United States

*CORRESPONDENCE

Takehige Otoi
✉ otoi@tokushima-u.ac.jp

RECEIVED 11 March 2024

ACCEPTED 15 April 2024

PUBLISHED 07 May 2024

CITATION

Nagahara M, Tatemoto S, Ito T, Fujimoto O,
Ono T, Taniguchi M, Takagi M and Otoi T
(2024) Designing a diagnostic method to
predict the optimal artificial insemination
timing in cows using artificial intelligence.
Front. Anim. Sci. 5:1399434.
doi: 10.3389/fanim.2024.1399434

COPYRIGHT

© 2024 Nagahara, Tatemoto, Ito, Fujimoto,
Ono, Taniguchi, Takagi and Otoi. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Designing a diagnostic method to predict the optimal artificial insemination timing in cows using artificial intelligence

Megumi Nagahara^{1,2}, Satoshi Tatemoto³, Takumi Ito^{1,2},
Otoha Fujimoto^{1,2}, Tetsushi Ono⁴, Masayasu Taniguchi⁴,
Mitsuhiro Takagi⁴ and Takehige Otoi^{1,2*}

¹Bio-Innovation Research Center, Tokushima University, Tokushima, Japan, ²Faculty of Bioscience and Bioindustry, Tokushima University, Tokushima, Japan, ³Agricultural and Horticultural Research Division, Tokushima Prefectural Technical Support Center for Agriculture, Forestry and Fisheries, Tokushima, Japan, ⁴Joint Faculty of Veterinary Science, Yamaguchi University, Yamaguchi, Japan

Dairy farmers and beef cattle breeders aim for one calf per year to optimize breeding efficiency, relying on artificial insemination of both dairy and beef cows. Accurate estrus detection and timely insemination are vital for improving conception rates. However, recent challenges such as operational expansion, increased livestock numbers, and heightened milk production have complicated these processes. We developed an artificial intelligence (AI)-based pregnancy probability diagnostic tool to predict the optimal timing for artificial insemination. This tool analyzes external uterine opening image data through AI analysis, enabling high conception rates when inexperienced individuals conduct the procedure. In the initial experimental phase, images depicting the external uterine opening during artificial insemination were acquired for AI training. Static images were extracted from videos to create a pregnancy probability diagnostic model (PPDM). In the subsequent phase, an augmented set of images was introduced to enhance the precision of the PPDM. Additionally, a web application was developed for real-time assessment of optimal insemination timing, and its effectiveness in practical field settings was evaluated. The results indicated that when PPDM predicted a pregnancy probability of 70% or higher, it demonstrated a high level of reliability with accuracy, precision, and recall rates of 76.2%, 76.2%, and 100%, respectively, and an F-score of 0.86. This underscored the applicability and reliability of AI-based tools in predicting optimal insemination timing, potentially offering substantial benefits to breeding operations.

KEYWORDS

artificial insemination, artificial intelligence, bovine, image capturing, external uterine opening

1 Introduction

Artificial insemination is a contemporary method for inducing pregnancy in fertile female animals, particularly cows, where its prevalence approaches 100% in Japan (Yoshioka, 2014). However, a discernible decline in conception rates via artificial insemination, concomitant with an extension of the calving period, has been observed in both dairy and beef cows in recent years (Dobson et al., 2008; Yaginuma et al., 2019). Accurate estrus detection and timely artificial insemination are imperative for enhancing conception rates. The standing behavior exhibited by cows at the peak of estrus serves as the primary indicator of estrus. However, challenges arise owing to the shortened duration of estrus and an increase in weak estrus cases, where cows ovulate without exhibiting discernible estrus signs. This is attributed to rising herd sizes and augmented milk production (Arkan et al., 2023). While daily assessments of ovarian dynamics through rectal examination or ultrasound imaging offer a reliable means for determining the optimal insemination time, practicality diminishes as operations scale up and labor resources decrease. Consequently, these methods are less feasible for dairy farmers and breeders to oversee large herds. Presently, determining the optimal insemination time relies heavily on subjective judgments by farmers and inseminators, drawing upon their seasoned expertise and intuition. However, there is a growing need to develop and incorporate objective judgment techniques and indicators to enable accurate determination, even in the absence of highly skilled personnel.

The adoption of sophisticated digital solutions to address these issues has increased recently (Achour et al., 2019; Eckelkamp, 2019; García et al., 2020). One technological innovation acknowledges that the electrical characteristics of bovine intravaginal mucus, specifically vaginal electrical conductivity and electrical resistance (VER), undergo swift alterations during estrus (Glencorse et al., 2023). Research indicates that the VER on estrus days exhibits a significant decrease compared to that on non-estrus days (Hockey et al., 2010). However, this approach remains at the investigational stage and has yet to be implemented in practical applications.

In the present study, a program using artificial intelligence (AI) was developed as a novel tool to assess the appropriate stage of insemination. This program discerns a suitable stage of insemination based solely on video and images captured from the external opening of the uterus. The initial step involved creating a pregnancy probability diagnostic model (PPDM). The AI was trained using images of the external uterine opening during both the estrus and non-estrus phases, and subsequent enhancements to PPDM were achieved by including and selecting additional images. This iterative process aimed to refine the model and ensure its practical applicability in the field.

Furthermore, the accuracy of PPDM was scrutinized through a field test involving the utilization of a web-based pregnancy probability prediction judgment application (judgment APP) developed based on PPDM. Predictive judgment of pregnancy probability was assessed to validate the efficacy and reliability of the developed tool in practical scenarios.

2 Materials and methods

2.1 Development environments

The PPDM system was created based on the methods used in Tatemoto et al. (2019). Image processing and model training was conducted using PyTorch (Meta Platforms, Inc., California, USA) on a GPU-equipped Colab (Google Inc., California, USA).

2.2 Image capture and data collection

A camera (Figure 1) was used to capture images of the external uterine opening. Videos were recorded from each head by three cooperative farmers during the artificial insemination procedures involving Holstein (n=151) and Japanese Black (n=221) cows. The camera operator uploaded videos directly to a dedicated website. A total of 316 videos recorded from August 2021 to October 2022 used for training were supplemented with an additional 56 videos in May 2023, resulting in a total of 372 videos used for training. The sample size used for machine learning was consequently beyond the statistical range (n = 172) (Eng, 2003).

2.3 Image preprocessing

The program predetermined the number of still images, which were subsequently extracted from the video at approximately one-second intervals. Images displaying unclear focal points were excluded through visual inspection and subsequently categorized into two classes, pregnant and non-pregnant, based on pregnancy diagnosis results, facilitating the training process. Images with moderate variability were selected to construct the dataset to minimize unnecessary noise and prevent the potential overfitting of the convolutional neural network (CNN) (Krizhevsky et al., 2017).

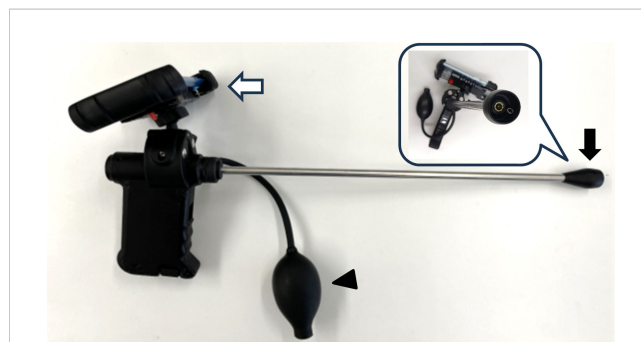


FIGURE 1

A camera device for capturing images of the external uterine opening. The black arrow indicates the endoscope, and the white arrow indicates the mobile phone for photo capturing. If mucus interfered with the lens, the balloon (arrowhead) could be pushed to release a jet of air to remove the mucus.

2.4 Model creation and evaluation

The EfficientNet network was used for model creation (Tan and Le, 2019), utilizing the smallest networks (B0 to B1 and B2) in the experiments. A classifier model was constructed with the following training parameters: 30 epochs, no pre-training model, and a stochastic gradient descent (SGD) optimization algorithm. Throughout the training, the images were resized to 224×224 pixels, the learning rate was set at 0.01, and a batch size of 32 was used. The models were evaluated using a dataset split of 8:2 between the training and validation sets (Val). Specifically, the assessment focused on accuracy (Acc) and loss (Loss) metrics derived from the validation data for EfficientNet-B0, B1, and B2. Points of interest within the bovine external uterine opening were visualized through a heat map generated using Grad-CAM (Krizhevsky et al., 2017) to enhance the understanding of the finer features essential for decision-making in this study.

2.5 Image selection and selection criteria

Visually unclear images were excluded from the dataset to enhance the learning accuracy. Furthermore, images of the external uterine opening, for which pregnancy outcomes are known, were scrutinized by a skilled inseminator to accurately identify cases of obvious estrus and non-estrus. The selection criteria for images indicative of estrus were based on the method of Sumiyoshi et al. (2014) and selected according to features such as abundant mucus, a conspicuous uterine opening, and a reddish coloration. Images designated as non-estrous and unsuitable for artificial insemination were chosen based on characteristics such as the absence of mucus, closed external uterine opening, post-ovulation bleeding, and absence of reddish coloration.

2.6 Data collection using the judgment APP

In total, 65 cows including Holstein ($n=32$) and Japanese Black ($n=33$) cows were tested by four cooperative farmers. Data collection involved placing a smartphone on the camera during

artificial insemination to capture images of an external uterine opening (Figure 2). These images were then uploaded to a web-based Judgment APP, where the pregnancy probability was immediately displayed as a percentage. All onsite judgment data were systematically saved within the application, including individual identification numbers and the date and time of image capture.

2.7 Verification of PPDM

Individuals subjected to artificial insemination were later assessed for success or failure of conception, and the accuracy of PPDM was verified. This evaluation was conducted in both pregnant and non-pregnant individuals. The means and standard errors of the percentages indicated by the judgment APP were calculated for each group. The threshold value was established as the upper limit that exhibited the highest value among all the calculated parameters. The accuracy of the predictive judgments based on predetermined thresholds was computed, and the PPDMs were evaluated using metrics encompassing accuracy, precision, recall rates, and F-score (Chou et al., 2023).

3 Results and discussion

Three models were developed and comparatively assessed for accuracy, revealing that EfficientNet-B1 achieved the highest Acc at 78%, whereas EfficientNet-B0 exhibited the lowest Acc at 71%, and Loss ranged from 0.4 for EfficientNet-B1 to 0.77 for EfficientNet-B2 (Figure 3). Notably, EfficientNet-B1 demonstrated superior capacity for extracting detailed features using Grad-CAM (Figure 4). Consequently, the AI-based diagnostic model developed in this study proved effective in predicting pregnancy based on images of the external uterine opening, with EfficientNet-B1 emerging as the optimal model with 78% Acc.

Images were added to strengthen the PPDM, expanding the dataset from the 513 images utilized in initial model to 7,642 images. Subsequently, a skilled person performed image sorting, selecting 1,392 final images from the initial 7,642. The outcomes of this retraining revealed that the highest Acc values for the PPDMs

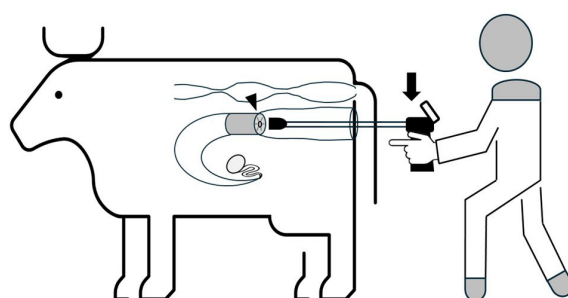


FIGURE 2

The operator inserts the camera (arrow) into the vagina through the vulva after disinfection, sets it in front of the external uterine opening (arrowhead), and presses the button on the grip to capture images.

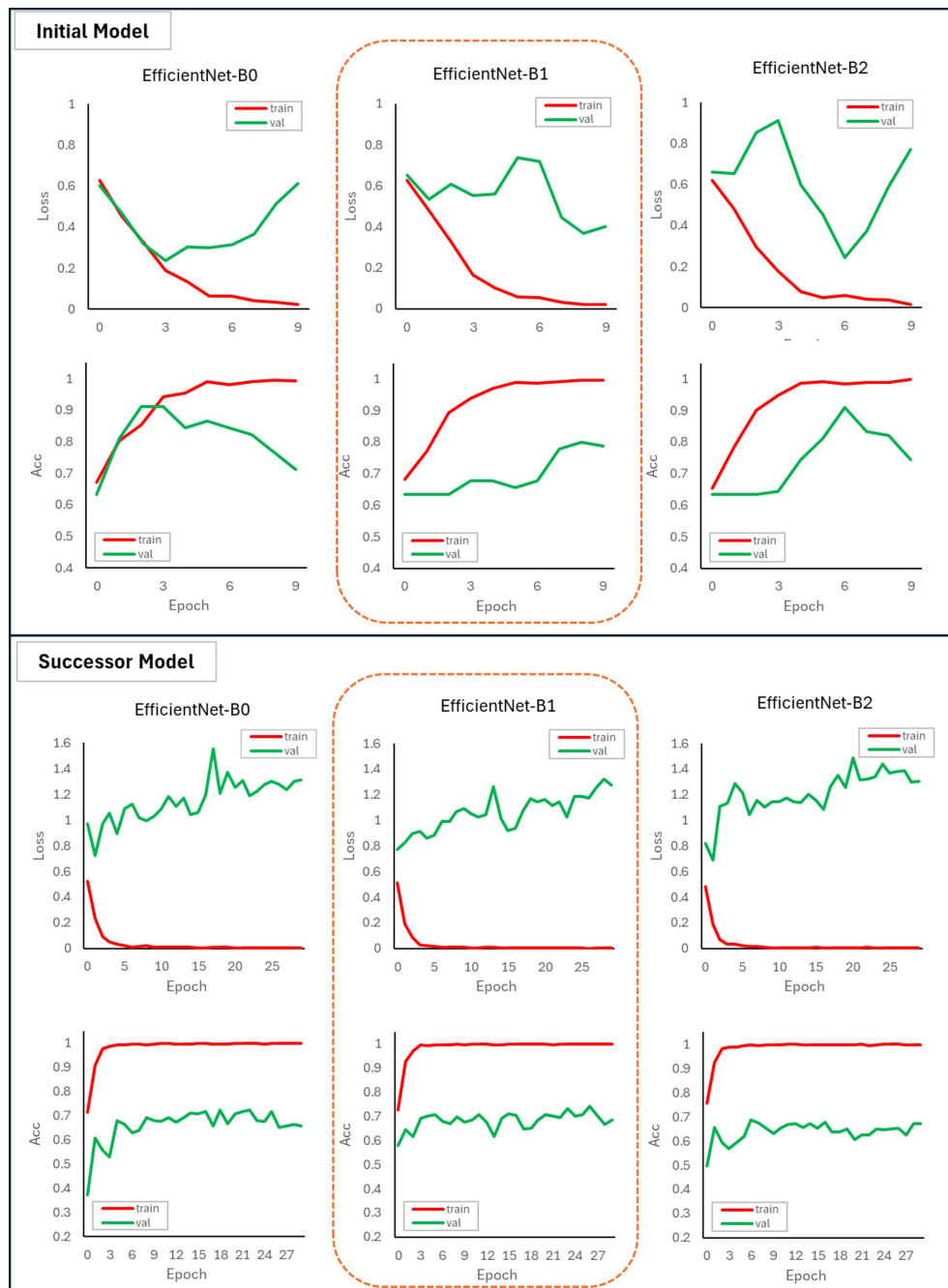


FIGURE 3

Parameter of loss (Loss) and accuracy (Acc) in EfficientNet-B0, EfficientNet-B1, and EfficientNet-B2 after machine training of initial (upper graphs) and successor (lower graphs) model. The models were assessed using a dataset split 8:2 between the training (train: red line) and validation sets (val: green line). Before and after machine training, EfficientNet-B1 emerged as the most effective model, with lower Loss and higher Acc than the other models.

was 74% for EfficientNet-B1, and the lowest was 68% for EfficientNet-B2. Correspondingly, the Loss for EfficientNet-B0 was minimized to 1.27, whereas that for EfficientNet-B1 reached a maximum of 1.30 (Figure 3). Consequently, EfficientNet-B1 emerged as the most effective model.

An assessment based on AI-predicted judgment values indicated that optimal accuracy, precision, and recall rates were observed when the judgment APP suggested a probability of 70% or

higher, yielding percentages of 76.2%, 76.2%, and 100%, respectively, along with an F-value of 0.86 (Figure 5).

We developed an AI-based pregnancy probability prediction system using images of the external uterine opening in this study and successfully created a PPDM based on the methods used in Tatemoto et al. (2019) who used deep learning methods to determine fruit ripeness based on training with sequential images depicting fruit growth. In addition, we successfully established an

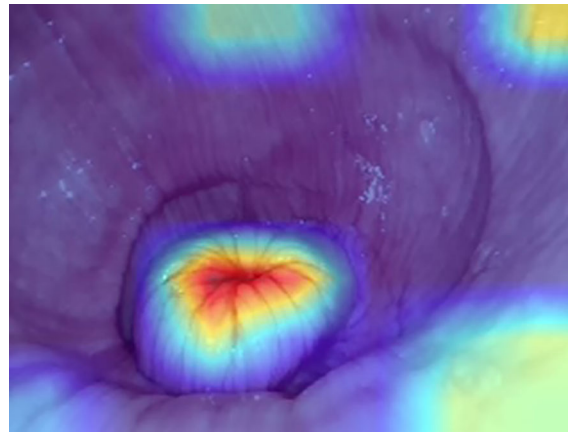


FIGURE 4
Points of interest of the external uterine opening were visualized through a heat map generated using Grad-CAM (CNN). The center of the external uterine opening, shown in the red area, indicates the most focused area for segregation of judgment.

APP that could diagnose pregnancy probability in real-time by photographing the external uterine opening during artificial insemination. Images were incorporated, organized, and subjected to retraining by the AI to augment the performance of the system. Among the three models considered, EfficientNetB1 exhibits the highest proficiency. Notably, the model size of EfficientNetB1 was reduced to one-third of the parameters of inception-v3, maintaining similar accuracy levels (Tan and Le, 2019). This reduction in model size enhances its applicability for deployment on servers and edge devices owing to its lower computational load. Notably, Genno and Kobayashi (2020) reported improved accuracy through data

sorting, a practice used in this study to augment the diagnostic model. Whereas initial model utilized 513 images, successor model incorporated 1,392 images, indicating a deliberate effort to enhance the training dataset. The observed variability in accuracy may stem from the ongoing need for a more extensive pool of training data to stabilize the accuracy levels despite incorporating additional training images into the AI model. Insufficient training images could contribute to the observed variability, emphasizing the crucial role of image allocation in AI training data generation. A reduction in the number of images is inevitable during data sorting, as those that are slightly unclear or significantly deviate from the

Judgment (%)	Accuracy (%)	Precision (%)	Recall (%)	F-score
0~	52	70	49	0.58
10~	52	51	70	0.59
20~	52	54	70	0.61
30~	57	64	70	0.67
40~	60	75	70	0.72
50~	64	88	70	0.78
60~	74	74	100	0.85
70~	76	76	100	0.86
80~	75	75	100	0.86
90~	75	75	100	0.86

Evaluation of machine learning

	Result (+)	Result (-)
AI Judgement (+)	True Positive (TP)	False Negative (FN)
AI Judgement (-)	False Positive (FP)	True Negative (TN)

Accuracy = $(TP+TN)/(all)$
 Precision = $TP/(TP+FP)$
 Recall = $TP/(TP+FN)$
 F-score = $2 \times Precision \times Rcall / (Precision+Recall)$

FIGURE 5
Evaluation of PPDM. The result indicates when the Judgment APP suggested a probability of 70% or higher, the accuracy, precision, and recall rates were 76.2%, 76.2%, and 100%, respectively, along with an F value of 0.86.

characteristics of estrous and normal periods, as discerned by a skilled observer, are excluded. Thus, augmenting the sample data is imperative to enhance the accuracy of the diagnostic model. Moreover, the Grad-CAM visualization method used in creating this model (Krizhevsky et al., 2017) aligns with findings in deep learning models using Faster R-CNN for fruit detection on mango, apple, and almond trees (Bargoti and Underwood, 2017). Increased network layering enhances the extraction of finer features. A heatmap representation of the influential points of the CNN forming the basis for image decisions was visualized in pursuit of heightened image recognition accuracy. Further refinement of the accuracy through adjustments to Grad-CAM warrants further exploration in future studies.

In contrast, machine learning evaluation metrics, including accuracy, precision and recall rates, and F-score, were calculated in the field test. The AI indicated the highest probability of pregnancy prediction at 70% or higher. In simple terms, the study verified that when the judgment result reached 70% or higher, the likelihood of accurately identifying fertility was 76.2%. While AI proved effective as a reference for assessing suitability for artificial insemination, discrepancies were observed. Despite the assumption that higher AI-assigned probabilities correspond to increased correctness percentages, instances of incorrect predictions, even for probabilities exceeding 80%, were observed in this study. Three primary factors contributing to these inaccuracies were identified: bovine-side issues, such as abnormal follicle development, ovulation timing deviations, and fallopian tube blockages; inseminator technique-related problems, including insertion site quality, freezing and thawing method proficiency, and injector insertion hygiene; and environmental factors, such as elevated body temperature due to heat. These variables may have contributed to the misinterpretation of the AI (Salisbury and VanDemark, 1961; Dobson et al., 2008).

On the other hand, the advantage of external measurement method is that even beginners in artificial insemination can learn the optimum timing for insemination by capturing images of the external uterine opening using this device. However, the disadvantage is that it requires skills in capturing sufficient images. If the image is out of focus or does not focus the center position, it may be judged as an incorrect prediction (Figure 6). Future improvements are necessary, such as automatic capturing of the best shots from the video (Takasaki et al., 2021). In the terms of animal welfare, the device is thinner than a vaginal speculum and has a rounded tip to avoid injured, however it is not certified as a medical device and requires adequate care for use on animals.

Moreover, this device is portable and can be used anywhere in the world with an internet connection. Since no statistical data on the external uterine opening during estrus has been collected so far, the development of this device can contribute to the collection of big data around the world. One of the objectives is to collect and utilize data from any breed worldwide. The conception rate of Holstein cows has been declining over the years, and blunt estrus is increasing (Dobson et al., 2008; Yaginuma et al., 2019). Therefore, we selected Holstein and Japanese Black cows, the main breeds in Japan, to collect data. The difference in accuracy between the breeds predicted as more than 70% of pregnancy was 61.5% and 100% for Holstein and Japanese Black cows, respectively (data not shown). The differences in accuracy by breed should be further investigated in the future.

4 Conclusion

The results underscore that the pregnancy probability prediction application based on AI-driven PPDM effectively anticipates high pregnancy probabilities of 70% or higher. This application is a valuable tool for identifying the optimal timing of

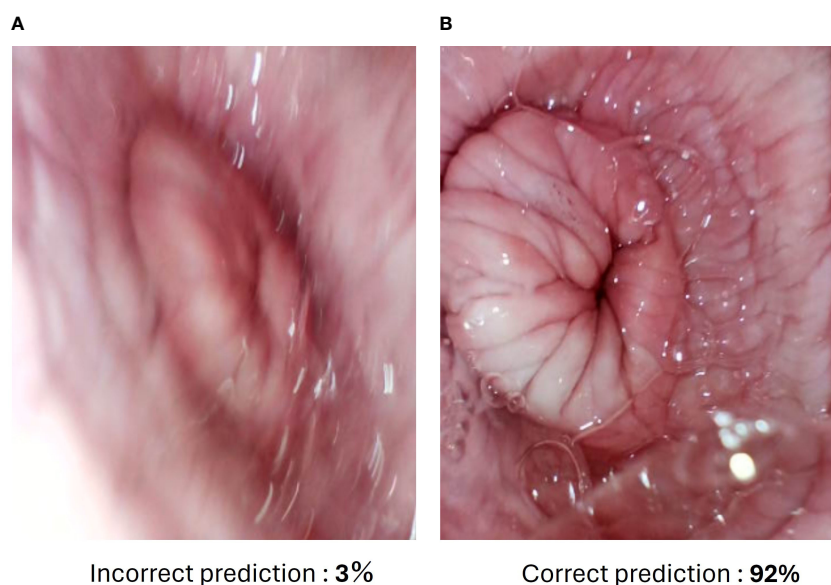


FIGURE 6

The different results of captured images by judging app. (A) Incorrect prediction image due to out of focus. (B) Correct prediction image of the same external uterine opening. The percentage indicates a probability of pregnancy prediction.

artificial insemination. Furthermore, the key feature of the device is its minimal technological demands, requiring only a dedicated camera and smartphone for installation, thus simplifying its deployment across production sites. Future endeavors will focus on refining the AI accuracy through expanded sample data and precise data assignment. Additionally, efforts will be directed towards promoting the deployment of the developed judging application in the field to enhance bovine conception rates.

Data availability statement

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

Ethics statement

The animal studies were approved by the Institutional Animal Care and Use Committee of Tokushima University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent was obtained from the owners for the participation of their animals in this study.

Author contributions

MN: Conceptualization, Formal analysis, Writing - original draft. ST: Data curation, Formal Analysis, Methodology, Software, Writing - original draft. TI: Investigation, Writing - original draft. OF: Investigation, Writing - original draft. TeO: Writing - review & editing. MaT: Conceptualization, Writing - review & editing. MiT: Funding acquisition, Supervision, Writing - review & editing. TaO:

Project administration, Supervision, Data curation, Writing - review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was partially supported by the Japan Racing Association (R3-no. 18).

Acknowledgments

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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

- Achour, B., Belkadi, M., Aoudjit, R., and Laghrouche, M. (2019). Unsupervised automated monitoring of dairy cows' behavior based on Inertial Measurement Unit attached to their back. *Comput. Electron. Agric.* 167. doi: 10.1016/j.compag.2019.105068
- Arikan, İ., Ayav, T., Seçkin, A.Ç., and Soygazi, F. (2023). Estrus detection and dairy cow identification with cascade deep learning for augmented reality-ready livestock farming. *Sensors (Basel)*. 23, 9795. doi: 10.3390/s23249795
- Bargoti, S., and Underwood, J. (2017). "Deep fruit detection in orchards," in 2017 IEEE International Conference on Robotics and Automation (ICRA). Singapore: IEEE Press. 3626–3633.
- Chou, C. Y., Hsu, D. Y., and Chou, C. H. (2023). Predicting the onset of diabetes with machine learning methods. *J. Pers. Med.* 13, 406. doi: 10.3390/jpm13030406
- Dobson, H., Walker, S. L., Morris, M. J., Routly, J. E., and Smith, R. F. (2008). Why is it getting more difficult to successfully artificially inseminate dairy cows? *Animal* 2, 1104–1111. doi: 10.1017/S175173110800236X
- Eckelkamp, E. A. (2019). INVITED REVIEW: Current state of wearable precision dairy technologies in disease detection. *Appl. Anim. Sci.* 35, 209–220. doi: 10.15232/aas.2018-01801
- Eng, J. (2003). Sample size estimation: how many individuals should be studied? *Radiology* 227 (2), 309–313. doi: 10.1148/radiol.2272012051
- García, R., Aguilar, J., Toro, M., Pinto, A., and Rodríguez, P. (2020). A systematic literature review on the use of machine learning in precision livestock farming. *Comput. Electron. Agric.* 179. doi: 10.1016/j.compag.2020.105826
- Genno, H., and Kobayashi, K. (2020). Data cleansing using deep learning and its application to apple fruit images. *Agric. Inf. Res.* 29, 47–61. doi: 10.3173/air.29.47
- Glencorse, D., Grupen, C. G., and Bathgate, R. (2023). Vaginal and vestibular electrical resistance as an alternative marker for optimum timing of artificial insemination with liquid-stored and frozen-thawed spermatozoa in sows. *Sci. Rep.* 13, 12103. doi: 10.1038/s41598-023-38803-5
- Hockey, C. D., Norman, S. T., Morton, J. M., Boothby, D., Phillips, N. J., and McGowan, M. R. (2010). Use of vaginal electrical resistance to diagnose oestrus, dioestrus and early pregnancy in synchronized tropically adapted beef heifers. *Reprod. Domest. Anim.* 45, 629–636. doi: 10.1111/j.1439-0531.2008.01320.x
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Commun. ACM*. 60, 84–90. doi: 10.1145/3065386
- Salisbury, G. W., and VanDemark, N. L. (1961). *Physiology of reproduction and artificial insemination of cattle* (San Francisco (& London: W. H. Freeman and Company).
- Sumiyoshi, T., Tanaka, T., and Kamomae, H. (2014). Relationships between the appearances and changes of estrous signs and the estradiol-17 β peak, luteinizing hormone surge and ovulation during the periovulatory period in lactating dairy cows kept in tie-stalls. *J. Reprod. Dev.* 60 (2), 106–114. doi: 10.1262/jrd.2013-119
- Takasaki, C., Takefusa, A., Nakada, H., and Oguchi, M. (2021). Action recognition using pose data in a distributed environment over the edge and cloud. *Transact. Info. Syst.* E104.D, 539–550. doi: 10.1587/transinf.2020DAP0009
- Tan, M., and Le, Q. (2019). "EfficientNet: Rethinking model scaling for convolutional neural networks," in MA: *International conference on machine learning* (PMLR), 6105–6114.

Tatemoto, S., Harada, Y., and Imai, K. (2019). Image-based determination of plum “tsuyuakane” ripeness via deep learning. *Agric. Inf. Res.* 28 (3), 108-114. doi: 10.3173/air.28.108

Yaginuma, H., Funesima, N., Tanikawa, N., Miyamura, M., Tsuchiya, H., Noguchi, T., et al. (2019). Improvement of fertility in repeat breeder dairy cattle by embryo

transfer following artificial insemination: Possibility of interferon tau replenishment effect. *J. Reprod. Dev.* 65, 223-229. doi: 10.1262/jrd.2018-121

Yoshioka, K. (2014). “Artificial Insemination” in *Handbook of stock raising*. K. Ogimoto, K. Nirasawa, M. Kuwabara, F. Terada, Y. Nakai and K. Sugiura (Tokyo, JP: Kodansha) 105-111.