

## Article

# Breathing Pattern Analysis in Cattle Using Infrared Thermography and Computer Vision

Sueun Kim \* and Yuichi Hidaka

Laboratory of Veterinary Surgery, Department of Veterinary Science, University of Miyazaki, 1-1 Gakuen Kibana-dai Nishi, Miyazaki 889-2192, Japan; yhidaka@cc.miyazaki-u.ac.jp

\* Correspondence: kim\_su\_eun@med.miyazaki-u.ac.jp; Tel.: +81-985-58-7791

**Simple Summary:** Breathing patterns are commonly used to assess cattle health and welfare parameters such as stress, pain, and disease. Infrared thermography has recently been accepted as a non-invasive tool for breathing pattern measurement. In this study, we applied a computer vision method (Mask R-CNN) to infrared thermography and made it possible to automatically estimate the breathing pattern in cattle. Breathing patterns identified by computer vision were highly correlated with those measured through thermal image observation. As this method is not labor-intensive and can handle numerable big data, it might be possible to analyze breathing patterns from various angles in the future.

**Abstract:** Breathing patterns can be considered a vital sign providing health information. Infrared thermography is used to evaluate breathing patterns because it is non-invasive. Our study used not only sequence temperature data but also RGB images to gain breathing patterns in cattle. Mask R-CNN was used to detect the ROI (region of interest, nose) in the cattle RGB images. Mask segmentation from the ROI detection was applied to the corresponding temperature data. Finally, to visualize the breathing pattern, we calculated the temperature values in the ROI by averaging all temperature values in the ROI. The results in this study show 76% accuracy with Mask R-CNN in detecting cattle noses. With respect to the temperature calculation methods, the averaging method showed the most appropriate breathing pattern compared to other methods (maximum temperature in the ROI and integrating all temperature values in the ROI). Finally, we compared the breathing pattern from the averaging method and that from the thermal image observation and found them to be highly correlated ( $R^2 = 0.91$ ). This method is not labor-intensive, can handle big data, and is accurate. In addition, we expect that the characteristics of the method might enable the analysis of temperature data from various angles.

**Keywords:** breathing pattern; infrared thermography; cattle health and welfare; computer vision; machine learning



**Citation:** Kim, S.; Hidaka, Y. Breathing Pattern Analysis in Cattle Using Infrared Thermography and Computer Vision. *Animals* **2021**, *11*, 207. <https://doi.org/10.3390/ani11010207>

Received: 6 December 2020

Accepted: 14 January 2021

Published: 16 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Breathing patterns are a vital sign that can provide valuable information relating to disease [1], stress [2], and pain [3] in cattle. It is common to evaluate breathing with the aid of a stethoscope, but the cattle can be excited through contact and this can affect the breathing patterns. The breathing pattern can also be measured by observing flank movements [4]; however, it can be difficult to observe in cold environments as the movements are typically slower and less pronounced [5]. To overcome these limitations, infrared thermography (IRT) has recently been used to measure the breathing characteristics of cattle [5,6]. In fact, IRT has been used not only to measure breathing but also in various fields such as stress assessment by eye temperature [7–9] or nose temperature [10]. However, the previous methods have utilized it in a limited way. For example, they used non-sequence temperature data [7], such as measurements taken three times a day for three days, but this

does not provide a temperature pattern. In other cases, the observer monitored the color change around the region of interest (ROI) [5,6] to evaluate the temperature change, which is labor-intensive. Some methods used fixed ROI settings and gained sequence temperature data [8,9,11], but, as we cannot control the animal's movement, these methods require resetting the ROI or moving the camera in order to maintain the ROI in the desired region.

The recent development of computer vision has broadened the horizons of research, and it has been used in various biomedical fields, including animal pose estimation [12], breathing pattern analysis in humans [13], and head pose estimation [14]. In particular, computer vision was also used for IRT research, where human breathing patterns were analyzed not only with IRT but also with computer vision [13]. They used thermal images and a tracking algorithm, the Median Flow algorithm [15], to automatically detect the ROI (nose) in the thermal images. Then, they made the breathing pattern visible by integrating all temperature values in the ROI, not by averaging them. This was because they detected the ROI (nose) in the thermal image and the detected nose size varied according to the breathing strength. In this case, the average value could not distinguish between deep breathing and shallow breathing.

The application of the same method to cattle has some limitations. First, a better ROI detection algorithm is necessary because the cattle's movements are too uncontrolled to detect in thermal images. Secondly, infrared studies so far have set the ROI to a square shape, including unnecessary parts other than the nostrils. This effect is significant, especially in cattle, because the distance between the two nostrils is greater than in many other species. Finally, in terms of the temperature calculation, the integrated method could be problematic. The reason for this is that, in the case of cattle, the distance between the two nostrils is great and anatomically three-dimensional, so they are often detected alternately, one nostril or both nostrils. If integration is performed, both detections indicate twice the value of one detection.

To overcome these limitations, we used Mask R-CNN, a state-of-the-art algorithm for object detection. Mask R-CNN is a two-stage detection algorithm which prioritizes detection accuracy over inference speed and has been used in animal welfare evaluation [16–19] because of its high-quality segmentation mask for each detection. Using Mask R-CNN means we used RGB images, not thermal images, for the ROI detection. RGB images have more clearly defined characteristics than thermal images. In addition, as the algorithm provides a high-quality segmentation mask for each detection, we could set the ROI according to the shape of the nose. Finally, as we used RGB images for the ROI detection, the nose size is not related to the breathing strength. This allowed us to use the thermal value averaging method in the ROI, not the integrating method.

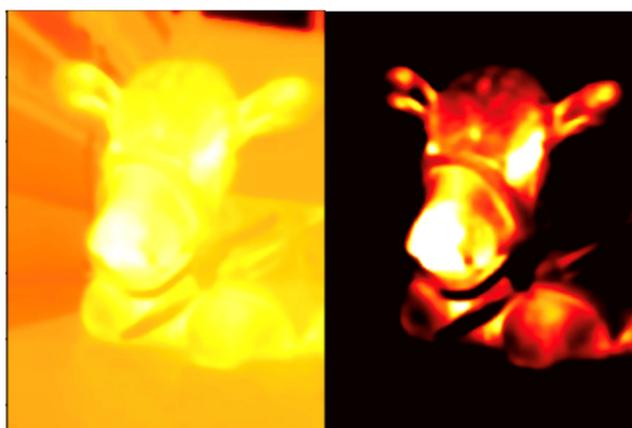
To sum up, in this paper, our purpose is to incorporate computer vision into IRT in breathing pattern analysis in cattle. To assess the computer vision method's applicability, we compared it with the previously assessed method, thermal image observation [6]. In addition, for the temperature value calculation, we evaluated three different methods: (1) maximum temperature in the ROI, which is the most commonly used method in previous IRT research; (2) integrating all temperature values in the ROI, which was used for computer-vision-based IRT breathing pattern analysis in humans; and (3) averaging all temperature values in the ROI.

## 2. Materials and Methods

### 2.1. Breathing Pattern Analysis through Observation

The temperature distribution on the skin changes with the ambient temperature. These changes could lead to difficulties in identifying the ROI on thermal images because the information about the facial morphology becomes lost. To overcome this problem, the process of translating from a continuous temperature value to its digital color-mapped equivalent, which is called quantization, should be applied. We used the optimal quantization technique [13,20] before observation. This technique dynamically finds the threshold value that best separates objects from background and maximizes the contrast in the thermal image,

making it possible to detect subtle signals (Figure 1). After the optimal quantization, the observer monitored the color change that occurred as a result of the thermal fluctuations around the nose during inhalation and exhalation. The time taken to complete five breaths was converted into the number of breaths per minute. In other words, the number of breaths was counted through the color change in the nostril; if it took 10 s to breathe five times, the respiration rate was 30 per minute. This manually calculated respiratory rate was used to evaluate the following algorithm-determined respiratory rate.



**Figure 1.** An example of optimal quantization: (left) visualization of the temperature values before the optimal quantization; and (right) the image after the optimal quantization shows the separation of the object from the background, maximizing thermal image contrast.

## 2.2. Observation through a Computer Vision Method

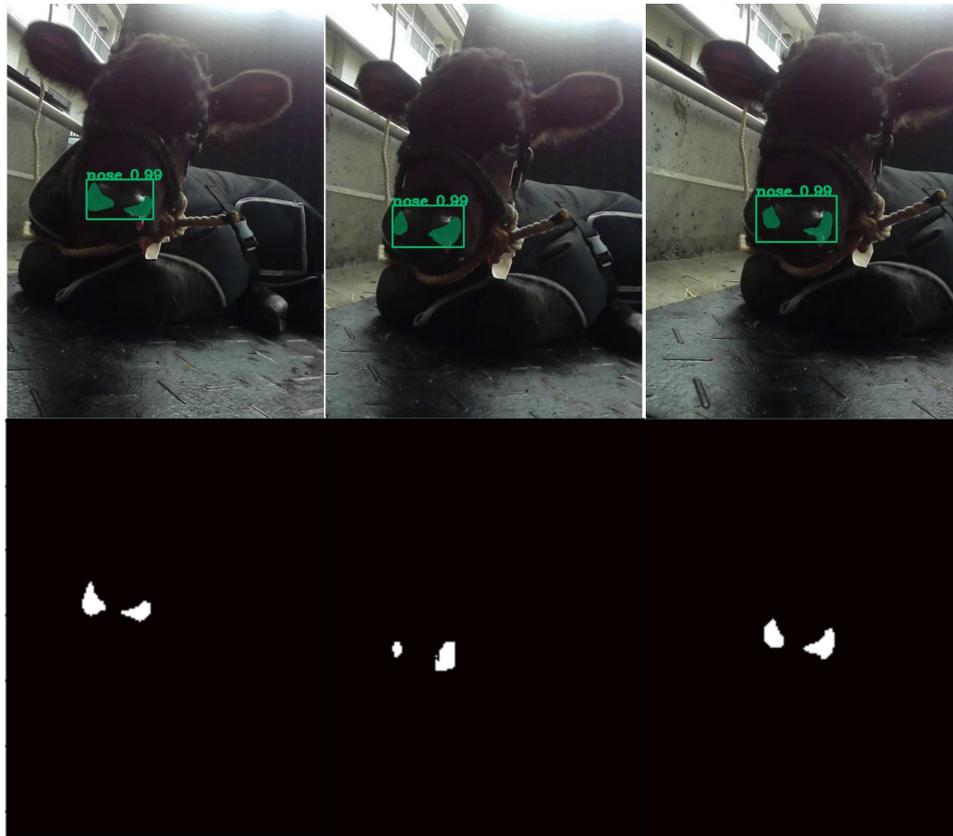
### 2.2.1. Object Detection Training

To automatically detect the ROI (nose) in RGB images, it is necessary to train a nose-part annotated dataset. The dataset then includes images and image annotations. We used VIA (the VGG Image Annotator) [21] to annotate the nose part for 1400 images: 1000 images for training and 400 images for validation. For the detection training task, we used Mask R-CNN [22] and transfer learning with a pre-trained model. Transfer learning is the improvement of learning through the transfer of knowledge from a related task that has already been learned [23]. When a dataset is not large enough to train, a model learned from a similar existing dataset is used. Our dataset of 1000 images for training and 400 images for validation was not big enough to train. COCO is a large-scale object detection dataset made by Microsoft [24]. There were parts similar to the COCO dataset [24], which includes animals, so we used a pre-trained backbone weight based on the COCO dataset to accelerate the training process. ResNet 101 [25] was used as the backbone. With ResNet 101 and the COCO dataset, we fine-tuned a new classifier for our nose detection and gained a weight for the detection.

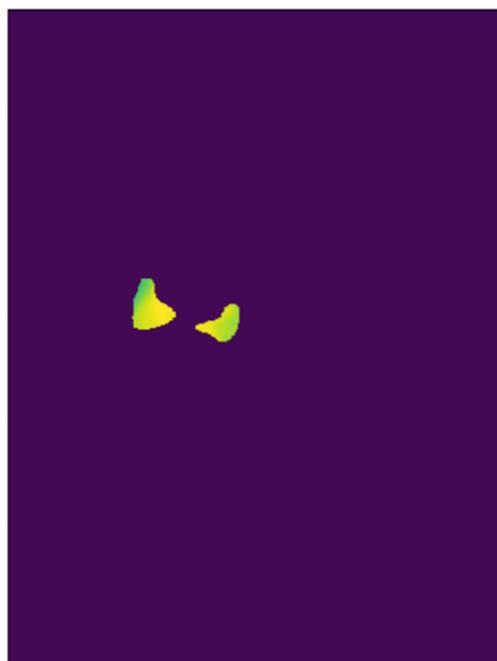
### 2.2.2. ROI Detection and Temperature Extraction

For the purposes of the present study, 50 samples comprising RGB images (video) and temperature sequence data were collected from five Japanese Black calves hospitalized in a veterinary teaching hospital (University of Miyazaki, Miyazaki, Japan). We used a mobile thermal imaging camera (FLIR One Pro for iOS (accuracy:  $\pm 3.0$  °C, sensitivity: 0.07 °C; field of view:  $55^\circ \times 43^\circ$ ; resolution:  $80 \times 60$ ; temperature range:  $-20$  to  $400$  °C), FLIR Systems Inc., Santa Barbara, CA, USA, [www.flir.com](http://www.flir.com)). FLIR One Pro is a basic entry level camera with some limitations compared to other thermal imaging camera, such as accuracy, sensitivity, battery performance and emissivity setting. However, we used this device considering its portability and commerciality. The temperature sequence data and RGB video were recorded simultaneously using the camera. The temperature sequence data contained all pixel temperature data. For example, if one minute of temperature data were

obtained, 184,320,000 temperature values would be included ( $184,320,000 = 640$  (frame height)  $\times 480$  (frame width)  $\times 10$  frames per second  $\times 60$  s). The emissivity was set to 0.95 (the proper emissivity for animal tissue should be 0.98, not 0.95; however, it was set to 0.95 due to the device's emissivity setting limitation) and the recordings began at 1 m in front of the calf; however, considering its applicability, afterwards, the distance between the camera and the calf was lessened or increased during recording. The RGB images were used to detect the ROI (nose) and gain the mask. After detecting the nose region in the RGB images with the weight obtained from the detection training detailed in Section 2.2.1 (Figure 2), the masks were applied to the thermal temperature data of the corresponding frames and all temperature data were obtained for the ROI (Figure 3). More specifically, in the RGB images, we set the nose part pixels (using the mask) to 1, set the not-nose part pixels to zero, and constructed the matrix (npz file, 640 (frame height)  $\times 480$  (frame width)  $\times$  frame numbers). By multiplying it with the corresponding temperature sequence data, only the temperature values for the exact part of the nostril were obtained. In addition, the average temperature value of the nostril could be obtained by dividing it by the number of pixels corresponding to the mask. We used Python for all these calculations.



**Figure 2.** Sample results of nose detection using the proposed method: (**top**) green bounding boxes (detection) and green masks (segmentation) show the ROI (nose) detection; and (**bottom**) the detected masks were extracted for application to the corresponding thermal data.



**Figure 3.** Sample image representing the multiplication of the mask and the corresponding thermal data. On this frame, the average temperature inside the mask was 33.85 °C.

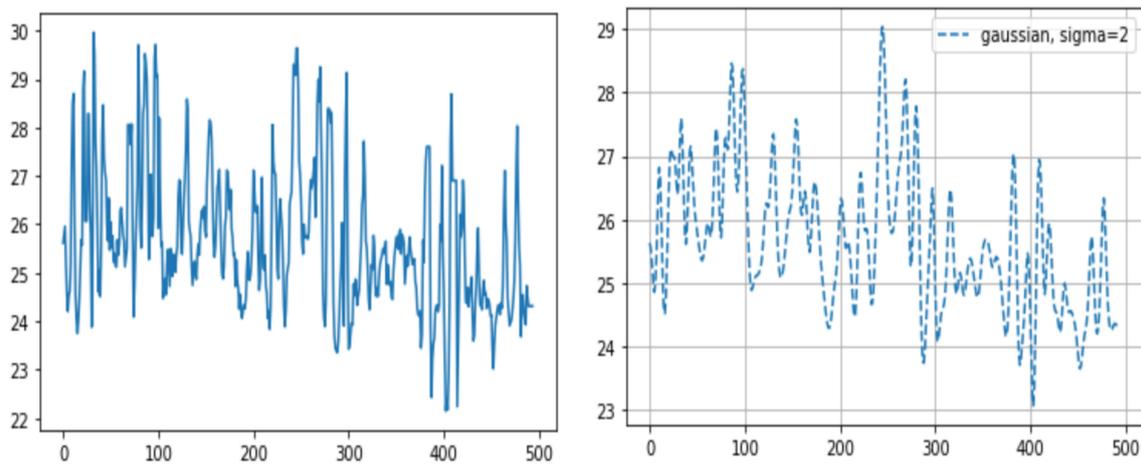
### 2.3. Statistical Analysis

#### 2.3.1. mAP

To evaluate the accuracy of the algorithm, the mAP (mean average precision) for object detection was used. AP (average precision) is a popular metric for measuring the accuracy of object detectors. AP is a calculation of the precision and recall based on IoU (intersection over union). IoU measures the overlap between two boundaries: the predicted boundary and the ground truth boundary. We predefined the IoU threshold as 0.5; when the ratio of two bounding boundaries exceeds 50%, precision and recall are calculated. Precision measures the number of predicted trues out of the total number of detections. Recall measures the number of predicted trues out of the number of ground truths. For example, if there is an image with five objects, the algorithm detects three, and the number of accurate object detections is two, then the precision is 66.6% (2/3) and the recall is 40% (2/5). AP is obtained by integrating the precision–recall curve. Finally, the average of the AP becomes the mAP. The average precision value is computed over 0–1. The higher is the value between 0 and 1, the better is the accuracy.

#### 2.3.2. Statistical Analysis of Temperature Calculations and Overall Accuracy

We calculated the temperature data for all pixels of the ROI via three different methods: (1) determining the maximum temperature value in each frame ROI; (2) averaging the temperature values in each frame ROI (after adding the temperature values for all pixels corresponding to the nostril and dividing the sum by the number of pixels corresponding to the mask); and (3) integrating temperature values in each frame ROI (same as the averaging temperature values method but without dividing by the number of pixels corresponding to the mask). Through this process for all frames, we visualized the breathing patterns as temperature changes over time. To decrease ripples in the breathing pattern, we used the Gaussian window (Figure 4). As mentioned above, the time taken to complete five breaths was converted into the number of breaths per minute.



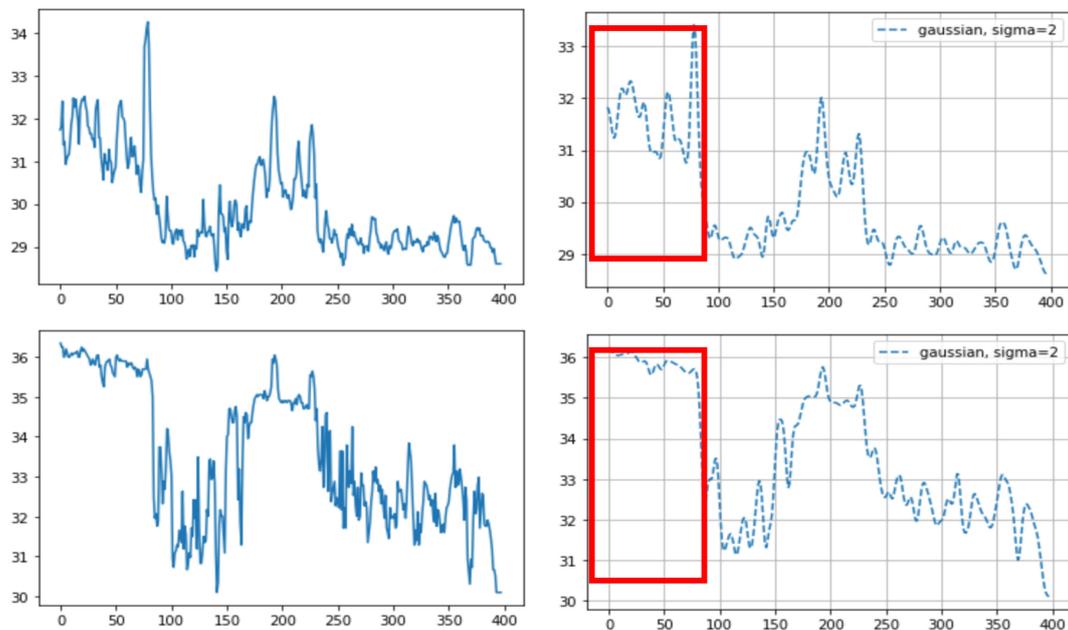
**Figure 4.** Sample graphs representing the breathing pattern. After multiplication of the masks and the corresponding thermal data through all frames, we were able to gain the breathing pattern. The X-axis represents the frame number (10 fps, frames per second) and the Y-axis represents temperature. (right) The graph after applying the Gaussian window.

Once the respective breaths per minute was determined via thermal image observation and the computer vision method, a regression analysis was carried out in order to assess the level of agreement between the two types of recording. The temperature value averaging method was selected out of the three different methods for this analysis.

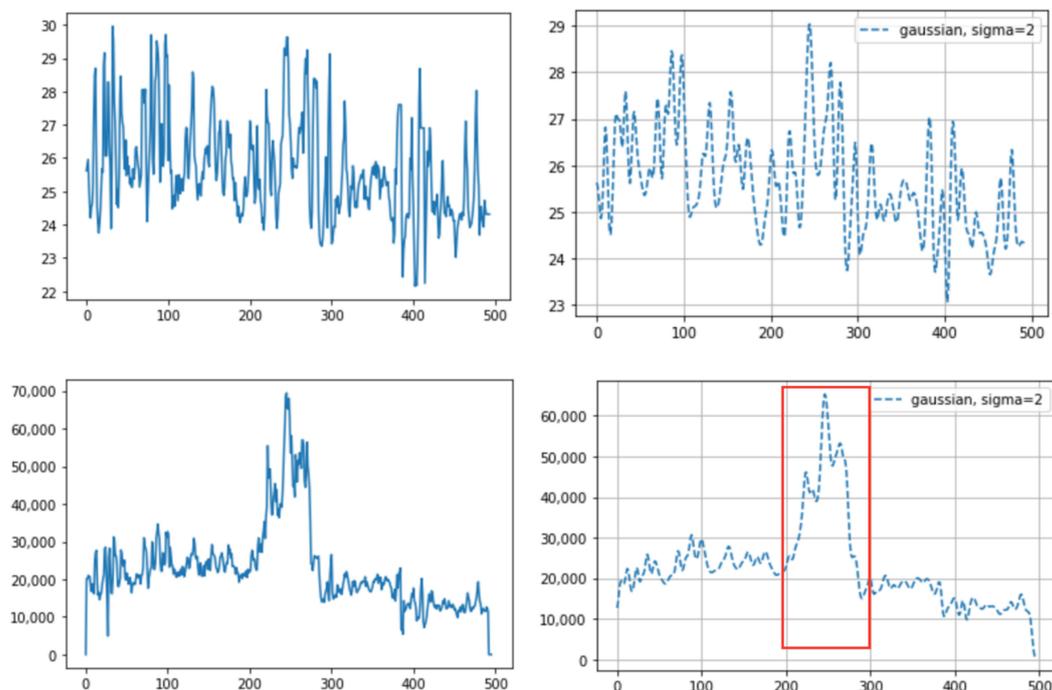
### 3. Results

In the case of human-related object detection, datasets containing large amounts of image data have already been made [24], providing enough data for algorithm training. In this study, we tried to make our dataset to apply computer science to the animal field. It was relatively small compared to the human datasets, and transfer learning was thus applied to maximize efficiency. The training was conducted on these datasets, and we evaluated the weights' accuracy through mAP. We set the IoU threshold to 0.5, and the mAP for the nose dataset achieved 76%. This value is comparable to other general object detection accuracy values [22]. The nose was actually detectable when we saw the nose-detected video. In addition, it was well detected regardless of whether one or both nostrils were in the image.

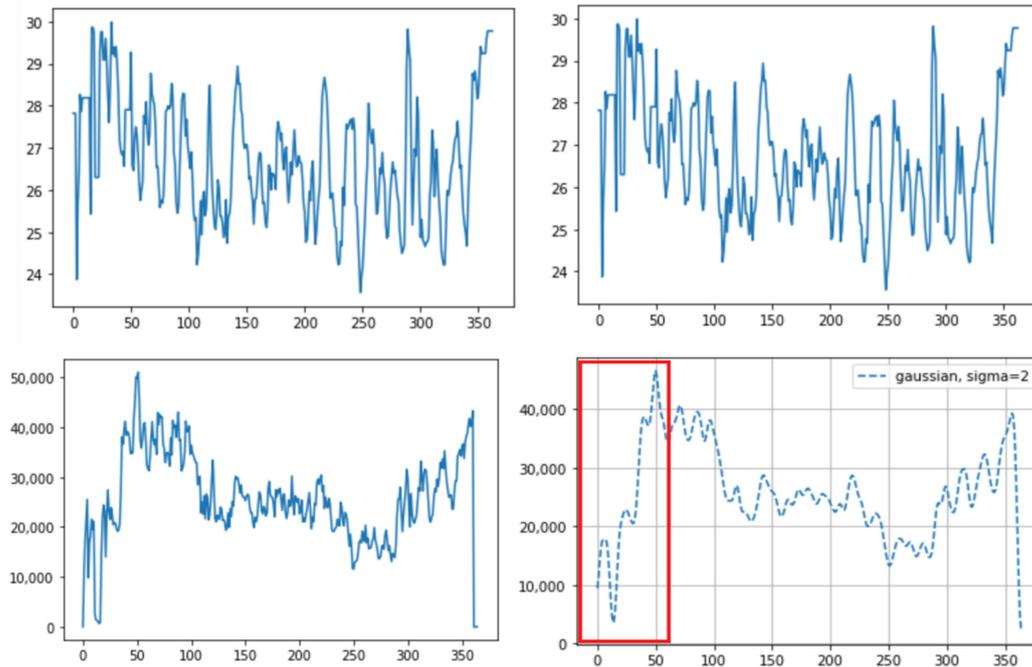
Of the three different analysis methods—maximum temperature, integrating, and averaging—averaging all temperature values in the ROI was found to provide a strong correlation with the actual breathing pattern. For the first method, although the temperature changes depend on the breathing patterns, the maximum ROI temperature showed a range of change that was sometimes small (Figure 5). With the second method, the integration of all temperature values in the ROI, the integrated values sometimes suddenly or gradually increased regardless of the breathing strength (Figures 6 and 7). On the other hand, the results of averaging all the pixel temperature values were the most similar to a normal breathing pattern and were found to be highly correlated with those measured via thermal image observation ( $R^2 = 0.91$ ,  $p < 0.001$ , Figure 8).



**Figure 5.** An example of different thermal data calculation methods on the same sample: **(top)** averaging of all the pixel temperature data, before **(left)** and after **(right)** the Gaussian window; and **(bottom)** the maximum temperature in each frame ROI, before **(left)** and after **(right)** the Gaussian window. The temperature fluctuation does not represent the breathing pattern in the case of the maximum temperature method (red box (bottom)). The X-axis represents the frame number (10 fps, frames per second) and the Y-axis represents temperature.



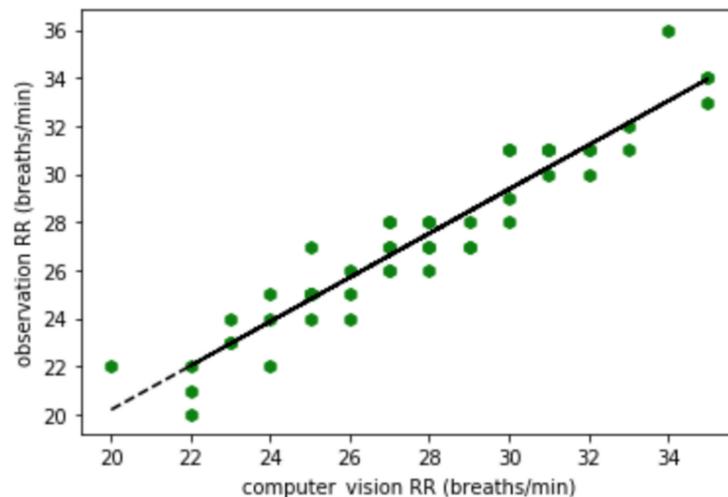
**Figure 6.** An example of different thermal data calculation methods on the same sample: **(top)** averaging of all the pixel temperature data: before **(left)** and after **(right)** the Gaussian window; and **(bottom)** integrated values of all the pixel temperature data: before **(left)** and after **(right)** the Gaussian window. The integrated values suddenly increased as they approached around the 200th frame (red box (bottom)). In this case, the camera approached the cattle during the time of the 200th frame. The X-axis indicates the frame number (10 fps, frames per second) and the Y-axis indicates temperature (top) or the integrated value (bottom).



**Figure 7.** An example of different thermal data calculation methods on the same sample: **(top)** averaging of all the pixel temperature data: before **(left)** and after **(right)** the Gaussian window; and **(bottom)** integrated values of all the pixel temperature data: before **(left)** and after **(right)** the Gaussian window. The integrated values suddenly increased as they approached around the 30th frame (red box (bottom)). This occurred due to the difference between only one nostril detected before the 30th frame and both nostrils detected after the 30th frame due to the cattle head direction. The X-axis represents the frame number (10 fps, frames per second) and the Y-axis represents the temperature (top) or the integrated value (bottom).

$$y = 0.9171x + 1.8422$$

$$R\_squared : 0.911$$



**Figure 8.** Correlation between respiration rates (RRs) measured via thermal image observation and a computer vision method (averaging all temperature values in the ROI) for 50 recordings.

#### 4. Discussion

This paper overcomes the limitations of previous thermal image research: (1) Some previous infrared research used naked-eye observation to analyze the breathing pattern [5,6]. However, this is labor-intensive. We used an automatic object detection algorithm, making

it possible to get a long-time breathing pattern. (2) Other infrared research, not using automatic detection algorithms, used the maximum value to detect the region [8]. This is because the eyes usually show the maximum temperature in the whole body or in the head, so the temperature change in the area can be observed through the maximum temperature. However, chances are high that it will not be representative of temperature values within the ROI. For example, two different breaths cannot be differentiated using only the maximum temperature; shallow breathing accompanies a small region temperature change and deep breathing accompanies a larger region temperature change. In this study, as we used an object detection algorithm, not the maximum temperature, to detect the ROI, we could avoid the problem of using the maximum value and could use all temperature data within the ROI. (3) As we used the Mask R-CNN for ROI detection, we obtained better detection accuracy, and optimal quantization, which was used to maximize the contrast within the thermal image, was no longer necessary. In human breathing pattern research [13], thermal images were used to track the ROI and optimal quantization was used to maximize the detection accuracy. Compared to humans, animals move their body and head much more randomly and actively, so a more accurate detection tool is necessary to apply computer vision methods to animal practice. Mask R-CNN basically uses RGB images, which contain more image characteristics than thermal images, and thus shows greater accuracy. In addition, as the segmentation mask only covered the exact ROI, a more accurate temperature pattern was obtained. Previous research only set the ROI as a square format and unnecessarily included parts that did not correspond to the ROI. Especially in cattle, since the distance between both nostrils is relatively greater than that in humans, chances are high that many unnecessary parts are included together when set in a square shape. This can lead to inaccurate temperature values. (4) In a prior study [13], a more accurate result was obtained when the authors integrated all the temperature values corresponding to the nose region than when they averaged the values. This was because the ROI was detected on thermal images, not RGB images, and the ROI size varied depending on the breathing strength. Thus, the average values between deep breathing and shallow breathing were not different. However, if the integrated method is used for cattle, some problems could appear. First, values can vary depending on the distance between the camera and the object, regardless of the breathing strength. If the distance between the object and the camera is smaller, even if the breathing strength is equal, the number of pixels that the ROI takes up in the image increases and the integrated value increases as well. In practice, we saw an increased value regardless of breathing strength when we brought the camera closer to the cattle or the cattle came closer to the camera (Figure 6). Secondly, the integrated values were different between when there were two nostrils on the image and when there was just one nostril (Figure 7). Even if the breathing strength is equal, when the number of nostrils in the image increases from one to two, it doubles the integrated value. This problem could be critical in cattle because the frequencies of both or one side nostril are more alternating due to their activity. In addition, the fact that the nostrils of cattle are more anatomically three-dimensional than those of humans worsens the problem. On the other hand, using average temperature values, which was problematic in the prior study, did not present problems in our study. This is because we detected the ROI on RGB images and the ROI size was not related to the breathing strength. Thus, we could gain a stable breathing pattern by avoiding the disadvantages of using integral values and solved the problem of using average values in the previous study.

This study was intended to evaluate cattle breathing patterns non-invasively for a long period of time so that fast or shallow breathing patterns could be utilized to detect pain or stress. However, there are some limitations to evaluating other pulmonary diseases, such as pneumonia, and subsequent abnormal pulmonary sounds with this method. We think that auscultation sound is still necessary for these diseases' diagnosis. In addition, we decided to use the FLIR One Pro because it was most likely to be commercialized considering its price and was able to obtain the sequence temperature data. However, it is lacking in resolution, battery performance, emissivity settings for biological tissue, and

accuracy when compared to other infrared cameras. To apply this method in dairy farms and make it a real-time method, it must be developed through further research.

## 5. Conclusions

Breathing patterns are a vital sign that provide health information. Infrared thermography has recently been used to evaluate breathing patterns. We combined an object detection algorithm, Mask R-CNN, with infrared thermography. We recorded RGB images and sequence temperature data simultaneously. The extracted ROI masks from Mask R-CNN were applied to the corresponding temperature data, and we were able to gain stable breathing patterns by averaging the ROI temperature values. The characteristics of this method will make it possible to analyze not only the breathing rate but also diverse breathing information from various aspects.

**Author Contributions:** Conceptualization, S.K.; methodology, S.K.; data collection and validation, S.K.; Software, S.K.; formal statistical analysis, S.K.; writing—original draft preparation, S.K.; writing—review and editing, Y.H.; and supervision, Y.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Acknowledgments:** The authors gratefully acknowledge the assistance from Veterinary Surgery staff at the University of Miyazaki.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Verhoeff, J.; Ban, M.V.; Nieuwstadt, A.V. Bovine respiratory syncytial virus infections in young dairy cattle: Clinical and haematological findings. *Vet. Rec.* **1984**, *114*, 9–12. [[CrossRef](#)] [[PubMed](#)]
2. Gaughan, J.B.; Holt, S.M.; Hahn, G.L.; Mader, T.L.; Eigenberg, R.A. Respiration rate—Is it a good measure of heat stress in cattle. *Asian-Australas. J. Anim. Sci.* **2013**, *13*, 329–332.
3. Stewart, M.; Shepherd, H.M.; Webster, J.R.; Waas, J.R.; Mcleay, L.M.; Schütz, K.E. Effect of previous handling experiences on responses of dairy calves to routine husbandry procedures. *Animal* **2012**, *7*, 828–833. [[CrossRef](#)] [[PubMed](#)]
4. Milan, H.F.; Maia, A.S.; Gebremedhin, K.G. Technical note: Device for measuring respiration rate of cattle under field conditions. *J. Anim. Sci.* **2016**, *94*, 5434–5438. [[CrossRef](#)] [[PubMed](#)]
5. Stewart, M.; Wilson, M.; Schaefer, A.; Huddart, F.; Sutherland, M. The use of infrared thermography and accelerometers for remote monitoring of dairy cow health and welfare. *J. Dairy Sci.* **2017**, *100*, 3893–3901. [[CrossRef](#)] [[PubMed](#)]
6. Lowe, G.; Sutherland, M.; Waas, J.; Schaefer, A.; Cox, N.; Stewart, M. Infrared Thermography—A Non-Invasive Method of Measuring Respiration Rate in Calves. *Animals* **2019**, *9*, 535. [[CrossRef](#)] [[PubMed](#)]
7. Valera, M.; Bartolomé, E.; Sánchez, M.J.; Molina, A.; Cook, N.; Schaefer, A. Changes in Eye Temperature and Stress Assessment in Horses During Show Jumping Competitions. *J. Equine Vet. Sci.* **2012**, *32*, 827–830. [[CrossRef](#)]
8. Jerem, P.; Jenni-Eiermann, S.; McKeegan, D.; McCafferty, D.J.; Nager, R.G. Eye region surface temperature dynamics during acute stress relate to baseline glucocorticoids independently of environmental conditions. *Physiol. Behav.* **2019**, *210*, 112627. [[CrossRef](#)] [[PubMed](#)]
9. Bartolomé, E.; Sánchez, M.J.; Molina, A.; Schaefer, A.L.; Cervantes, I.; Valera, M. Using eye temperature and heart rate for stress assessment in young horses competing in jumping competitions and its possible influence on sport performance. *Animal* **2013**, *7*, 2044–2053. [[CrossRef](#)]
10. Cho, Y.; Bianchi-Berthouze, N.; Oliveira, M.; Holloway, C.; Julier, S. Nose Heat: Exploring Stress-induced Nasal Thermal Variability through Mobile Thermal Imaging. In Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction (ACII), Cambridge, UK, 3–6 September 2019. [[CrossRef](#)]
11. Stewart, M.; Stafford, K.J.; Dowling, S.K.; Schaefer, A.L.; Webster, J.R. Eye temperature and heart rate variability of calves disbudded with or without local anaesthetic. *Physiol. Behav.* **2008**, *93*, 789–797. [[CrossRef](#)] [[PubMed](#)]
12. Nath, T.; Mathis, A.; Chen, C.A.; Patel, A.; Bethge, M.; Mathis, W.M. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat. Protoc.* **2019**, *14*, 2152–2176. [[CrossRef](#)] [[PubMed](#)]

13. Cho, Y.; Bianchi-Berthouze, N.; Julier, S.J. DeepBreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In Proceedings of the Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), San Antonio, TX, USA, 23–26 October 2017. [[CrossRef](#)]
14. Yang, T.-Y.; Chen, Y.-T.; Lin, Y.-Y.; Chuang, Y.-Y. FSA-Net: Learning Fine-Grained Structure Aggregation for Head Pose Estimation From a Single Image. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–21 June 2019. [[CrossRef](#)]
15. Kalal, Z.; Mikolajczyk, K.; Matas, J. Forward-Backward Error: Automatic Detection of Tracking Failures. In Proceedings of the 20th International Conference on Pattern, Istanbul, Turkey, 23–26 August 2010. [[CrossRef](#)]
16. Gomes, R.A.; Monteiro, G.R.; Assis, G.J.F.; Busato, K.C.; Ladeira, M.M.; Chizzotti, M.L. Technical note: Estimating body weight and body composition of beef cattle through digital image analysis. *J. Anim. Sci.* **2016**, *94*, 5414–5422. [[CrossRef](#)] [[PubMed](#)]
17. Hansen, M.F.; Smith, M.L.; Smith, L.N.; Abdul Jabbar, K.; Forbes, D. Automated monitoring of dairy cow body condition, mobility and weight using a single 3D video capture device. *Comput. Ind.* **2018**, *98*, 14–22. [[CrossRef](#)] [[PubMed](#)]
18. Stajko, D.; Brus, M.; Hočevar, M. Estimation of bull live weight through thermographically measured body dimensions. *Comput. Electron. Agric.* **2008**, *61*, 233–240. [[CrossRef](#)]
19. Nilsson, M.; Herlin, A.H.; Ardö, H.; Guzhva, O.; Åström, K.; Bergsten, C. Development of automatic surveillance of animal behaviour and welfare using image analysis and machine learned segmentation technique. *Animal* **2015**, *9*, 1859–1865. [[CrossRef](#)] [[PubMed](#)]
20. Picture Thresholding Using an Iterative Selection Method. *IEEE Trans. Syst. Man Cybern.* **1978**, *8*, 630–632. [[CrossRef](#)]
21. Dutta, A.; Zisserman, A. The VIA Annotation Software for Images, Audio and Video. In Proceedings of the 27th ACM International Conference on Multimedia—MM '19, Nice, France, 21–25 October 2019. [[CrossRef](#)]
22. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017. [[CrossRef](#)]
23. Yang, Q.; Zhang, Y.; Dai, W.; Pan, S.J. *Transfer learning*. Cambridge; Cambridge University Press: Cambridge, UK, 2020.
24. Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. *Lect. Notes Comput. Sci.* **2014**, 740–755. [[CrossRef](#)]
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016. [[CrossRef](#)]