

## Article

# A System for Monitoring Animals Based on Behavioral Information and Internal State Information

Taro Shibasaki <sup>1,\*</sup> , Yuugo Yamazaki <sup>2</sup> and Hideyuki Tonooka <sup>2</sup> 

<sup>1</sup> Department of Intelligent Mechanical Systems, Faculty of Environmental, Life, Natural Science and Technology, Okayama University, Okayama 700-8530, Japan

<sup>2</sup> Major in Computer and Information Sciences, Graduate School of Science and Engineering, Ibaraki University, Hitachi 316-8511, Japan

\* Correspondence: shibanoki@okayama-u.ac.jp; Tel.: +81-86-251-8232

**Simple Summary:** In the current study, we proposed a system that discriminates the state of an animal by combining two types of features: behavioral information, which determines the kind of movement performed by the animal, and internal information, which reflects the animal's internal state. The proposed system measures biological information from camera images and extracts features to discriminate states using machine learning. In addition, a neural network was used to increase the accuracy of detecting the target, and an anomaly detection method was used to perform discrimination. In the experiments performed, we prepared video images containing routine behaviors and generated a model that can detect non-routine behaviors using the proposed system. We also prepared video images including non-routine behavior in which the hamster was stimulated by clapping to generate an abnormal sound. In the video, the behavior changed significantly before and after the stimulus presentation, and the extracted feature values also changed according to the behavior. As a result of detection, the system discriminated the behavior as non-routine behavior after the stimulation. In conclusion, the results supported the possibility of using a pet monitoring system to immediately inform the owner when the animal is under load.

**Abstract:** Managing the risk of injury or illness is an important consideration when keeping pets. This risk can be minimized if pets are monitored on a regular basis, but this can be difficult and time-consuming. However, because only the external behavior of the animal can be observed and the internal condition cannot be assessed, the animal's state can easily be misjudged. Additionally, although some systems use heartbeat measurement to determine a state of tension, or use rest to assess the internal state, because an increase in heart rate can also occur as a result of exercise, it is desirable to use this measurement in combination with behavioral information. In the current study, we proposed a monitoring system for animals using video image analysis. The proposed system first extracts features related to behavioral information and the animal's internal state via mask R-CNN using video images taken from the top of the cage. These features are used to detect typical daily activities and anomalous activities. This method produces an alert when the hamster behaves in an unusual way. In our experiment, the daily behavior of a hamster was measured and analyzed using the proposed system. The results showed that the features of the hamster's behavior were successfully detected. When loud sounds were presented from outside the cage, the system was able to discriminate between the behavioral and internal changes of the hamster. In future research, we plan to improve the accuracy of the measurement of small movements and develop a more accurate system.

**Keywords:** monitoring system; image processing; mask R-CNN; anomaly detection; one-class SVM; rodents



**Citation:** Shibasaki, T.; Yamazaki, Y.; Tonooka, H. A System for Monitoring Animals Based on Behavioral Information and Internal State Information. *Animals* **2024**, *14*, 281. <https://doi.org/10.3390/ani14020281>

Academic Editor: Vera Baumans

Received: 20 November 2023

Revised: 31 December 2023

Accepted: 6 January 2024

Published: 16 January 2024



**Copyright:** © 2024 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

In the current period of reduced social interaction because of the coronavirus disease 2019 (COVID-19), pet ownership may be helpful for improving quality of life. Pet ownership has been reported to have positive benefits, such as reducing psychological distress caused by loneliness [1] and improving health [2]. The pet population in the US is around 157 million, and the number of pet owners has increased in recent years because of increased income [3] and the greater proportion of time spent at home because of the COVID-19 pandemic. The number of pet owners is predicted to increase in the future [4,5] even after COVID-19. Pets are often thought of as members of the family, and it is desirable to be able to regularly monitor animals' conditions considering the risk of injury or illness as well as their contribution to zoology.

Monitoring animals can be performed in various ways, such as by analyzing changes in behavior associated with changes in the living environment [6,7]. However, manual data collection can miss important information and has time constraints. Therefore, systematic and continuous monitoring is important [8]. In contrast, many monitoring methods using video cameras [9–12] and sensors such as thermography [13], as well as camera-based behavior analysis have been developed with recent improvements in deep neural network technology [14–23], making it possible to track animals with high accuracy in a non-contact manner. For example, for the care of orangutans in a zoo, a system used for human behavior analysis was adapted to perform individual identification, location tracking, and posture estimation using neural networks. These methods enable the detection of the locations in which each animal spends time in detail, and they can be applied to the optimal arrangement of the environment. However, it is not possible to determine animals' internal state, which limits the understanding of the causes of their behavior.

A previous study measured the heartbeat of animals to determine their internal state [24–30]. However, the effects of stress caused by the continuous wearing of sensors cannot be eliminated. In addition, various risks, such as contact infection, are associated with the use of sensors. Here, some systems perform heart rate detection by capturing color changes caused by blood flow from facial images taken by a video camera [25] or by measuring weak bodily movements caused by heartbeats with thermography [13]. Al-Naji et al. proposed the animal's cardiopulmonary signal detection method from video images and achieved non-contact monitoring of animals in a zoo [26]. However, there are few systems that monitor the condition of animals by considering not only external behavioral information but also the internal state of each animal.

We previously developed a monitoring system that simultaneously performs behavior analysis via video image analysis and internal state extraction in a non-contact manner [31]. This system also detects behaviors that are not performed in everyday life, and which can detect conditions that differ from those in everyday life. However, the system requires specialized knowledge in the setting of threshold values and detection accuracy degradation because of objects in the environment.

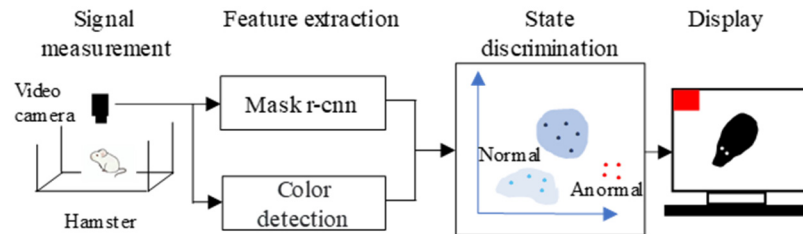
Therefore, the current paper proposes a new monitoring system for pets that combines a new feature extraction method using deep neural network technology and an anomaly detection method. The proposed system simultaneously extracts behavioral and internal information from the detected region using a mask R-CNN. The system then automatically detects unusual behaviors using a one-class support vector machine (SVM), which is widely used for anomaly detection.

This paper is organized as follows: Section 2 presents the proposed monitoring system, data acquisition, and the experimental setup. The experimental results of the proposed system with a hamster are discussed in Section 3 and Section 4, respectively. Finally, Section 5 outlines a conclusion.

## 2. Materials and Methods

Figure 1 shows the proposed monitoring system. A digital hi-vision video camera (HC-W870M, Panasonic Corporation, Osaka, Japan) is used to analyze a hamster in a

cage. Real-time monitoring can be performed using a PC with an HDMI capture device. When the system detects a state that is different from typical daily behavior, the system alerts the owner. The system consists of four parts of (1) signal measurement, (2) feature extraction, (3) state discrimination, and (4) display. The following subsections describe detailed information about each part.



**Figure 1.** The structure of the proposed rodent monitoring system.

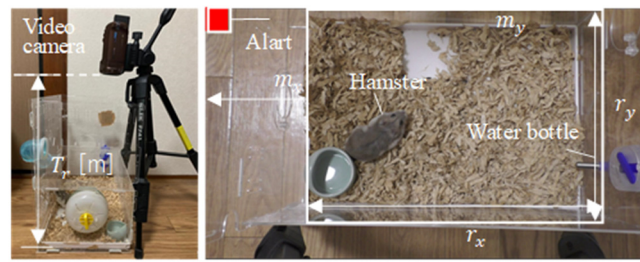
### 2.1. Signal Measurement

A video camera is set up parallel to the ground without any tilt above a typical cage, and the height of the camera  $T_r$  is adjusted to maximize the hamster's view angle as much as possible to measure its daily behavior ( $T_r = 0.4$  m). In this paper, the behavior of hamsters, including states such as the state of eating, was measured for approximately 3 min per day ( $2.8 \pm 0.7$  min, sampling frequency:  $f_s = 60$  Hz). A total of six days of video images were stored, and a clapping sound was generated at 90.7 s on one day of data collection.

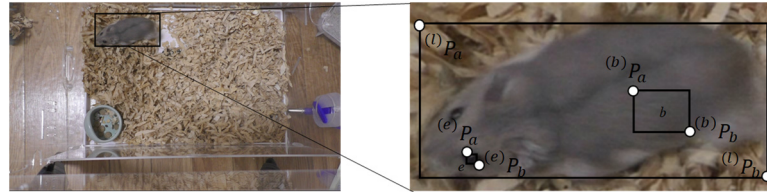
From the video image measured by the video camera, a  $r_x \times r_y = 1370 \times 880$  rectangular area  $A$  with margins  $m_x = 350$  and  $m_y = 0$  pixels in each axis direction is set from each frame to omit the living environment in the cage (see Figure 2). Here, the system first extracts the maximum length of the hamster from the image and color information to identify the hamster and the hamster's eyes from the measured video image. The user clicks on the image of frame  $t$  that corresponds to each region, and the system automatically detects the minimum/maximum thresholds for hamster and hamster eye region detection:  $(^b)h_{th}^{max} = 95$ ,  $(^b)h_{th}^{min} = 60$ ,  $(^b)s_{th}^{min} = 30$ ,  $(^b)s_{th}^{max} = 0$ ,  $(^b)v_{th}^{max} = 75$ ,  $(^b)v_{th}^{min} = 125$ ,  $(^e)h_{th}^{max} = 65$ ,  $(^e)h_{th}^{min} = 30$ ,  $(^e)s_{th}^{max} = 100$ ,  $(^e)s_{th}^{min} = 50$ ,  $(^e)v_{th}^{max} = 40$ , and  $(^e)v_{th}^{min} = 10$  (corresponding to each element in the HSV color system), and the body length  $a_x = 680$  and  $a_y = 150$  pixels of the hamster along the  $x/y$  axes are also determined in this paper. This process is performed for images in which the hamster is lying in the  $y$ -axis direction. Figure 3 shows an example. The user performs three processing steps on the image. The first is the extraction of color information. The user clicks diagonal points such as  $(^b)P_a$  and  $(^b)P_b$  in an area like  $b$  that contains the color of the hamster on the screen, and the color information in the rectangle is used as the hamster's color information. In this case, we select a region that is as close to a single color as possible so that there is no variation in the color information of the region containing the hamster. Next, the color information of the hamster's eyes is extracted. For the eye color information,  $(^e)P_a$  and  $(^e)P_b$  are specified to create a region that includes only the eyes, such as  $e$ . Finally, the body length of the hamster is extracted. The diagonal points  $(^l)P_a$  and  $(^l)P_b$  are specified so that the entire hamster is enclosed.

### 2.2. Feature Extraction

The proposed system uses both mask R-CNN and previously reported color information detection [31] to extract features related to posture and internal states. The behavioral information detects the amount of movement and the change in body and movement direction per frame, the internal state is obtained from the color information change from the surrounding pixels of the detected hamster area, and the heartbeat information is extracted.



**Figure 2.** Overview of the proposed monitoring system.



**Figure 3.** Regions for color information extraction.

### 2.2.1. Behavioral Information

First, the coordinates of the hamster's center of mass at frame  $t$  are calculated from extracted information using mask R-CNN [32], which is capable of object detection and classification as well as object shape analogy. Mask R-CNN consists of three main components: the backbone, region proposal network (RPN), and head. The first part, the backbone, is a network consisting of multiple convolutional layers and is responsible for extracting features from the input image. A convolutional neural network (CNN) is a neural network with a convolutional layer and a pooling layer, which can be combined with a pooling layer to extract more specific information (e.g., the complex shape of an object) by deeply constructing the convolutional layer. Here, Inception ResNet v2 is used in this study. After this step, the RPN takes the features output from the CNN and selects candidate object-like regions. The features extracted by the CNN are sent to the box classification layer, which detects the likelihood that an object exists in the candidate region, and the box regression layer, which represents the deviation from the correct object region. The candidate regions are subject to deviation from the actual position in the original image due to compression by the CNN. The RoIAlign layer divides the candidate region into  $3 \times 3$  feature maps. The values of each pixel point are interpolated by bilinear interpolation based on the four pixel points on the feature map, and these are interpolated by max or average pooling to obtain the reference points of the RoI region. Finally, classification, which determines what kind of object is in the candidate region obtained from the RPN, and region extraction, which detects the location of the object, can be used to extract what kind of object exists in what region of the image.

The mask image constructed using instance segmentation is obtained by inputting the image of region A to the model trained on the COCO dataset. Here, there are no hamsters (rodents) in the dataset, and it is usually difficult to classify the target from the mask image. Other than the hamster, the measurement environment (cage) contained only flooring materials and playground equipment. Thus, the mask images with animals are regarded as regions that include the hamster, and the largest of the masked regions is classified as the hamster region B. From the region B (instance mask) obtained here, the point set of the mask  $\vec{x}_n = (M_{xi}(t), M_{yi}(t))(i = 1, 2, \dots, I)(n = 1, 2, \dots, N)$  is extracted. To reduce processing, only the set of points  $\vec{x}_i = (M_{xi}(t), M_{yi}(t))(i = 1, 2, \dots, I)(I < N)$  representing the hamster's contour from the mask point  $(M_{xi}(t), M_{yi}(t))$  is extracted and the center of gravity  $[G_x(t), G_y(t)]^T$  of the coordinates of the point set of the mask is calculated from  $\vec{x}_i$ . Here, the following procedures were used to extract the amount of movement  $m(t)$ , the body shape change  $r'(t)$ , and the change in direction  $\theta'(t)$  as behavioral information.

The amount of movement is calculated as the time difference of the extracted center of gravity according to the following equation:

$$m(t) = \|[G_x(t), G_y(t)]^T - [G_x(t-1), G_y(t-1)]^T\|. \quad (1)$$

This extracts the degree to which the hamster is moving in the cage, and the larger the value, the greater the movement in the cage.

The proposed system also extracts changes in the body shape and direction of movement as feature values for behavioral information. The body shape change  $r'(t)$  is calculated using the difference between the hamster's center of gravity and eye position.

$$r'(t) = r(t) - r(t-1). \quad (2)$$

$$r(t) = \sqrt{(G_x(t) - F_x(t))^2 + (G_y(t) - F_y(t))^2}.$$

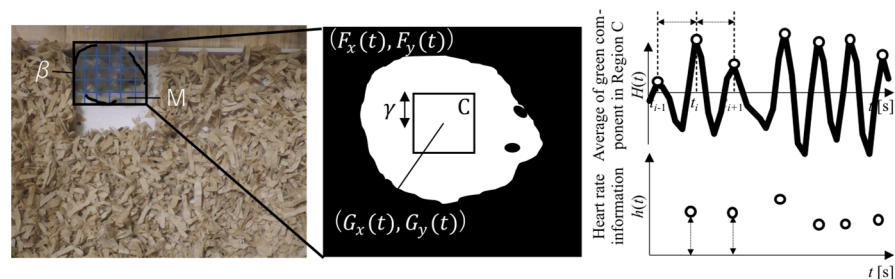
Here,  $[F_x(t), F_y(t)]^T$  is the position of the hamster's eyes extracted from the color information in the bounding box circumscribing area B [31]. If no target hamster is detected in the masked region by mask R-CNN, or if the masked region itself cannot be extracted at time  $t'$ , no information is extracted. Since  $r(t)$  is the distance between the center of gravity and the head, a large value indicates that the hamster's body is being stretched and a small value indicates that the hamster's body is shrunken. Therefore, by calculating the time difference,  $r'(t)$ , we can extract the change in the hamster's body shape per unit time.

The change in movement direction is calculated from the direction of the hamster extracted by the following processes. Because the shape of a hamster is typically characterized by an elongated head as shown in Figure 4, the first principal component, from principal component analysis, and the vector indicating the direction of the hamster can be obtained. The set of point  $\vec{x}_i$  representing the hamster's contour is used to calculate the first principal component using the following equation:

$$L(\lambda, w) = w^T \Sigma w - \lambda (w^T w - 1), \quad (3)$$

where  $\Sigma$  is the variance–covariance matrix of normalized  $\vec{x}_i$  and  $w = [w_x, w_y]^T$  is the eigenvector corresponding to the eigenvalue. The slope of the first principal component  $w_1$  corresponding to the largest eigenvalue is obtained from Equation (3). Then, the following equation is used to extract the direction in the image where the  $x$ -axis positive direction is 0 rad:

$${}^{(m)}\theta(t) = \tan^{-1}(w_{1y}/w_{1x}). \quad (4)$$



**Figure 4.** Region extraction based on mask R-CNN.

The direction  ${}^{(c)}\theta(t)$  calculated using a previously reported method [31] is also extracted as follows:

$${}^{(c)}\theta(t) = \tan^{-1}((G_y(t) - F_y(t)) / (G_x(t) - F_x(t))). \quad (5)$$

The mean value  $\theta(t)$  is obtained to extract information about the change in direction  $\theta'(t)$  as follows:

$$\begin{aligned}\theta'(t) &= \theta(t) - \theta(t-1) \\ \theta(t) &= \frac{{}^{(m)}\theta(t) + {}^{(c)}\theta(t)}{2}\end{aligned}\quad (6)$$

This allows for the evaluation of sudden movements, such as a hamster suddenly turning around and making a large change in its body direction.

### 2.2.2. Internal Body Information

Using a previously described method [31], a square region  $C$  of  $\gamma$  pixels around the hamster's center of gravity was defined, and the average value  $H(t)$  of the green component in region  $C$  was calculated (see Figure 4). The peak was detected from  $H(t)$  after applying a second-order digital Butterworth filter with a cutoff frequency of ( $f_l = 4.76$ ,  $f_h = 7.14$  Hz). The mountain climbing method was applied to extract the peak-to-peak time interval  $T(t)$  of  $H(t)$ .

$$T(t) = t_i - t_{i-1}. \quad (7)$$

where  $t_i$  is the time when the  $i$ -th peak of the detected  $H(t)$  appears. The inverse of the time difference was then multiplied by  $t_H = 0.7$  to detect the heart rate information  $h(t)$  around  $t_H$  seconds.

Additionally, if there is no target hamster in the masked region, or if the masked region itself cannot be extracted, the center of gravity and eye position are calculated, using the respective coordinate values. Here, if the area is less than the area of  $\alpha_{th}$  calculated from the maximum body length of the hamster set in advance, the heartbeat information is not extracted.

### 2.3. Anomaly Detection

On the basis of the extracted features, a one-class SVM is used to discriminate normal and anomalous activities of the hamster in daily life (Equation (8)).

$$f(t) = \mathbf{w}^T \varnothing(\mathbf{Z}(t)) - p. \quad (8)$$

Given training data  $\mathbf{Z}^q = [m^q, r^q, r'^q, \theta'^q, h^q]^T$  ( $q = 1, \dots, Q$ ),  $\mathbf{Z}^q$  is mapped to a high-dimensional feature space using a base function  $\varnothing(\mathbf{Z}^q)$  and linear separation is performed in the mapped space. Here, anomaly values are mapped near the origin and normal values are mapped far from the origin in the feature space using a Gaussian kernel that represents the distance between the data. A hyperplane is then set up such that the margin for discriminating data groups near the origin from other data groups is maximized. Therefore, the parameters of discrimination function  $\mathbf{w}$  and  $p$  (Equation (8)) are set such that Equation (9) is minimized so that  $\nu$  data of the training data are near the origin.

$$\frac{1}{2} \|\mathbf{w}\|^2 - p + \frac{1}{Q\nu} \sum_{j=1}^Q \xi_j. \quad (9)$$

where  $\xi_j$  is a slack variable for soft margin maximization. After training, the newly extracted feature vectors can be input to the discriminant function (Equation (8)) to determine the state (e.g., normality is zero and anomaly is one).

On the basis of the discrimination results, the user can judge the status of the hamster by displaying the results on the monitor. This system enables detailed monitoring of hamsters by using object detection as well as video image processing.

## 3. Experimental Results

To verify the effectiveness of the proposed system, we monitored the daily behavior of a hamster using the proposed system. We tried to reduce stress and prevent abnormal behavior caused by stress by making the environment the same in which the pets were

actually kept and the measurement environment during the experiment. The environment is the same as the actual pet-keeping environment and a video camera is placed on top so that the movements of the hamsters can be observed (see Figure 2). In addition, to discriminate in the living environment, we conducted measurements every evening when nocturnal hamsters started their activities.

As described in Section 2.1, the number of video images was six and a clapping sound was generated on one day of recording. The behavioral changes before and after the clapping sound were evaluated using the data (10,924 frames; 5062 samples were used to extract features). The data without sound generation were used as the training dataset (remaining 5 days; 48,255 frames; 23,179 samples were used to extract features). Five-fold cross-validation was used to determine the model parameters for which data without sound generation can be accurately discriminated as normal ( $\nu$  from 0.0001 to 0.01,  $\gamma$  [which is the parameter for the Gaussian kernel function to determine the decision boundary] from 0.01 to 1 in increments of 0.01).

Figure 5 shows scenes from the experiment and Figure 6 shows an example of the discrimination. Here, after the cross-validation, the highest discrimination rate was obtained when the parameters  $\nu$  and  $\gamma$  were 0.0001 and 0.25, respectively, and the results were obtained using the parameters.



Figure 5. Scenes from the experiment.

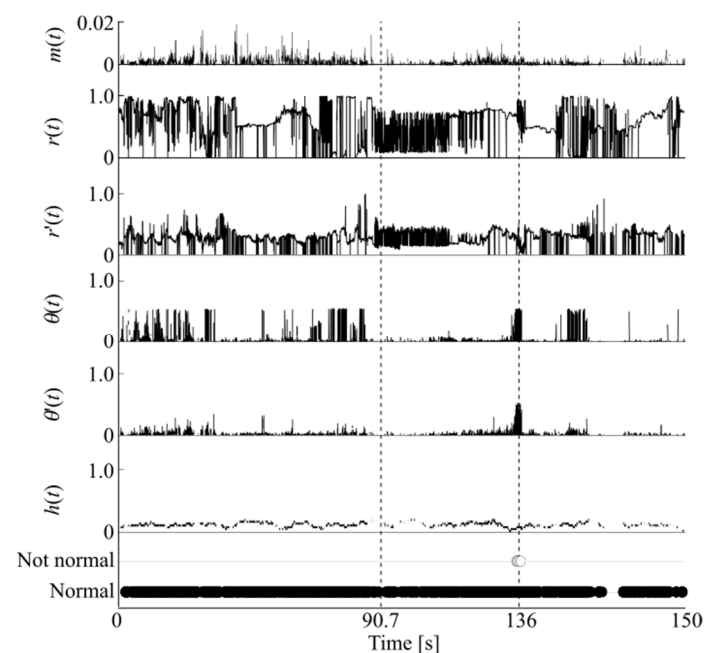


Figure 6. An example of experimental result.

Table 1 shows the confusion matrix, with the results of two-class classifications using SVM, linear discriminant analysis (LDA), and artificial neural network (MLP: multi-layer perceptron with two hidden layers of twenty hidden units) for the data of the daily activity class (before sound generation,  $f(t) = 0$ ) and the data for the entire time after sound occurrence as the non-daily activity class ( $f(t) = 1$ ) for comparison.

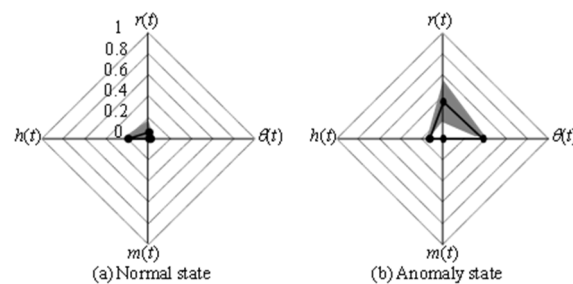
**Table 1.** Comparison results of state estimation.

	OCSVM		SVM		LDA		MLP	
	$f(t) = 0$	$f(t) = 1$	$f(t) = 0$	$f(t) = 1$	$f(t) = 0$	$f(t) = 1$	$f(t) = 0$	$f(t) = 1$
Normal	2674	0	2674	0	1145	1529	2667	7
Not normal	2377	11	2388	0	793	1595	2333	55

#### 4. Discussion

From Figure 6, the discrimination was performed as normal before the sound was generated; however, at approximately 136 s after the sound was generated (90.7 s), the hamster's state was judged as not normal. At this time, the hamster suddenly changes direction and its heart rate tends to decrease, indicating that the system is able to detect conditions different from normal daily activities simultaneously using behavioral information and internal states. Table 1 shows that SVM failed to discriminate the data well because all data were classified as the normal state while OCSVM can detect anomaly states. In both anomaly detection using OCSVM and when considering the two-class classification problem using SVM, the discrimination rate of non-daily activities did not have a high level of accuracy. This was mainly because the hamster shrank during the daily activities, and the direction vector was often inverted because of the noise caused by the hamster's shadow. Additionally, because the training data also include infrequent situations in daily activities, it is possible that the discriminative boundary was not accurately determined. In fact, in the results of the other discrimination methods shown in Table 1, linear discriminant analysis hardly discriminated between normal and anomaly states. In addition, when an artificial neural network was used, it can be seen that misdiscrimination occurred in each of the normal and non-normal states. Two-class discrimination except for OCSVM does not discriminate correctly even though data from the non-normal state was also used for training.

Figure 7 shows feature vectors extracted in (a) a normal state and (b) an anomaly state during the experiment. The solid line in the figure represents the mean value of each feature, and the shading represents the standard deviation. The figure shows that the variation in the behavioral information is large in all states. On the other hand, in the anomaly state, the value of the behavioral information is larger than in the normal state, and the value of  $h(t)$ , which represents the internal state, appears to be smaller. In the continuous measurement of hamster behavior, hamsters were often observed to exhibit slight changes in response to the occurrence of abnormal sounds, but they continued their daily behavior for a short period of time afterward. Although such characteristics were often observed in general, when data that differed from the daily state that can be measured continuously were obtained using the anomaly detection method, it was possible to discriminate the anomalous state as an unusual state and to produce a warning. Therefore, the proposed method appeared to be extremely effective for extracting behaviors and internal states and identifying them via anomaly detection.

**Figure 7.** An examples of extracted feature vectors.



These results indicate that the proposed system can detect conditions that are different from the everyday life of the animals by using just a video camera and this approach has been shown to be extremely useful in monitoring animals, including pets.

## 5. Conclusions

In this study, we proposed a monitoring system for hamsters using a single video camera. The system is capable of extracting different states from daily life using features of behavioral information, which determines the kind of movement performed by the animal, and internal information, which reflects the animal's internal state. Thus, the system has enabled the user to analyze pets' conditions without specialized knowledge. In an experiment, we investigated whether it was possible to detect behavioral changes in a hamster by stimulating it. In this experiment, a hamster's daily life activity was measured using the proposed system, and video images containing routine behavior were stored. We also prepared videos of the hamsters being stimulated by clapping during the recording, and the changes in behavior before and after the generation of abnormal sounds caused by the clapping were evaluated. The results confirmed that the proposed system was able to discriminate daily activities before the clapping sound was generated with an accuracy of 100% by adjusting the parameters of the one-class SVM. This system also detected anomaly states after the clapping sound was generated.

In the future, we aim to improve the discrimination accuracy by extracting more detailed behaviors, such as hand movements. The current results were obtained using only a few validation datasets with optimized parameters, and further validation with additional data is necessary. Furthermore, we plan to measure daily activities at night using sensors, such as a depth camera and thermography, so that the system can consider a wider variety of conditions.

**Author Contributions:** Conceptualization, T.S.; methodology, T.S.; software, T.S. and Y.Y.; validation, T.S. and Y.Y.; formal analysis, T.S., Y.Y. and H.T.; investigation, T.S. and Y.Y.; resources, T.S., Y.Y. and H.T.; data curation, T.S., Y.Y. and H.T.; writing—original draft preparation, Y.Y.; writing—review and editing, T.S. and H.T.; visualization, T.S. and Y.Y.; supervision, T.S. and H.T.; project administration, T.S.; funding acquisition, T.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by JSPS KAKENHI Grant Numbers: 17K12723 and 20K20212.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study due to the hamsters to be measured as being pets kept by the author, and the measurements, including the conditions for sound generation, were taken only under the conditions that occur in daily life using a consumer-use video camera.

**Informed Consent Statement:** Not applicable, as this research did not involve humans.

**Data Availability Statement:** Data available on request due to privacy and ethical restrictions.

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

## References

1. Damberg, S. *"Furry Tales": Pet Ownership's Influence on Subjective Well-Being During COVID-19 Times*; Springer: Berlin/Heidelberg, Germany, 2021.
2. Headey, B. Pet Dogs Benefit Owners' Health: A 'Natural Experiment' in China. *Soc. Indic. Res.* **2008**, *87*, 481–493. [CrossRef]
3. Alexander, P.; Berri, A.; Moran, D.; Reay, D.; Rounsevell, M.D. The global environmental paw print of pet food. *Glob. Environ. Change* **2020**, *65*, 102153. [CrossRef]
4. "Japan Pet Food Association". Available online: <https://petfood.or.jp/data/chart2020/index.html> (accessed on 1 December 2021).
5. Cabinet Office Government of Japan. Available online: <https://www5.cao.go.jp/keizai2/wellbeing/covid/index.html> (accessed on 1 December 2021).
6. Freeman, G.B.; Johnson, J.D.; Killinger, J.M.; Liao, S.C.; Davis, A.O.; Ruby, M.V.; Chaney, R.L.; Lovre, S.C.; Bergstrom, P.D. Bioavailability of arsenic in soil impacted by smelter activities following oral administration in rabbits. *Fundam. Appl. Toxicol.* **1993**, *21*, 83–88. [CrossRef]

7. Hettiarachchi, G.M.; Pierzynski, G.M.; Oehme, F.W.; Sonmez, O.; Ryan, J.A. Treatment of contaminated soil with phosphorus and manganese oxide reduces lead absorption by Sprague-Dawley rats. *J. Environ. Qual.* **2003**, *32*, 1335–1345. [[CrossRef](#)]
8. Dupuis-Desormeaux, M.; Davidson, Z.; Mwololo, M.; Kisio, E.; MacDonald, S.E. Comparing motion capture cameras versus human observer monitoring of mammal movement through fence gaps. *Afr. J. Ecol.* **2016**, *54*, 154–161. [[CrossRef](#)]
9. Wark, J.D.; Cronin, K.A.; Niemann, T.; Shender, M.A.; Horrigan, A.; Kao, A.; Ross, M.R. Monitoring the Behavior and Habitat Use of Animals to Enhance Welfare using the ZooMonitor App. *Anim. Behav. Cogn. ABC* **2019**, *6*, 158–167. [[CrossRef](#)]
10. Miller, N.E. Some psychophysiological studies of motivation and of the behavioural effects of illness. *Bull. Br. Psychol. Soc.* **1964**, *17*, 1–20.
11. Lasselin, J.; Schedlowski, M.; Karshikoff, B.; Engler, H.; Lekander, M.; Konsman, J.P. Comparison of bacterial lipopolysaccharide-induced sickness behavior in rodents and humans: Relevance for symptoms of anxiety and depression. *Neurosci. Biobehav. Rev.* **2020**, *115*, 15–24. [[CrossRef](#)] [[PubMed](#)]
12. Kulbeth, H.J.; Fukuda, S.; Brents, L.K. Automated quantification of opioid withdrawal in neonatal rat pups using Ethovision? XT software. *Nat. Neurosci.* **2018**, *21*, 1281–1289. [[CrossRef](#)]
13. Edgar, J.L.; Paul, E.S.; Nicol, C.J. Protective mother hens: Cognitive influences on the avian maternal response. *Anim. Behav.* **2013**, *86*, 223–229. [[CrossRef](#)]
14. Ni, A.; Azarang, A.; Kehtarnavaz, N. A Review of Deep Learning-Based Contactless Heart Rate Measurement Methods. *Sensors* **2021**, *21*, 3719. [[CrossRef](#)]
15. Mathis, A.; Mamidanna, P.; Cury, K.M.; Abe, T.; Murthy, V.N.; Mathis, M.W.; Bethge, M. DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **2018**, *21*, 1281–1289. [[CrossRef](#)]
16. Pereira, T.D.; Aldarondo, D.E.; Willmore, L.; Kislin, M.; Wang, S.S.-H.; Murthy, M.; Shaevitz, J.W. Fast animal pose estimation using deep neural networks. *Nat. Methods* **2019**, *16*, 117–125. [[CrossRef](#)]
17. Graving, J.M.; Chae, D.; Naik, H.; Li, L.; Koger, B.; Costelloe, B.R.; Couzin, I.D. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *eLife* **2019**, *8*, e47994. [[CrossRef](#)] [[PubMed](#)]
18. Bohoslav, J.P.; Wimalasena, N.K.; Clausing, K.J.; Dai, Y.Y.; Yarmolinsky, D.A.; Cruz, T.; Kashlan, A.D.; Chiappe, M.E.; Orefice, L.L.; Woolf, C.J.; et al. DeepEthogram, a machine learning pipeline for supervised behavior classification from raw pixels. *eLife* **2021**, *10*, e63377. [[CrossRef](#)]
19. Jia, Y.; Li, S.; Guo, X.; Lei, B.; Hu, J.; Xu, X.H.; Zhang, W. Selfee, self-supervised features extraction of animal behaviors. *eLife* **2022**, *11*, e76218. [[CrossRef](#)] [[PubMed](#)]
20. Lauer, J.; Zhou, M.; Ye, S.; Menegas, W.; Schneider, S.; Nath, T.; Rahman, M.M.; Di Santo, V.; Soberanes, D.; Feng, G.; et al. Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat. Methods* **2022**, *19*, 496–504. [[CrossRef](#)] [[PubMed](#)]
21. Congdon, J.V.; Hosseini, M.; Gading, E.F.; Masousi, M.; Franke, M.; MacDonald, S.E. The Future of Artificial Intelligence in Monitoring Animal Identification. *Health Behav.* **2022**, *12*, 1711.
22. Yang, X.; Chai, L.; Bist, R.B.; Subedi, S.; Wu, Z. A Deep Learning Model for Detecting Cage-Free Hens on the Litter Floor. *Animals* **2022**, *12*, 1983. [[CrossRef](#)] [[PubMed](#)]
23. Helwatkar, A.; Riordan, D.; Walsh, J. Sensor Technology for Animal Health Monitoring. *Int. J. Smart Sens. Intell. Syst.* **2014**, *7*, 1–6. [[CrossRef](#)]
24. Unakafov, A.M.; Möller, S.; Kagan, I.; Gail, A.; Treue, S.; Wolf, F. Using imaging photoplethysmography for heart rate estimation in non-human primates. *PLoS ONE* **2018**, *13*, e0202581. [[CrossRef](#)]
25. Chen, X.; Cheng, J.; Song, R.; Liu, Y.; Ward, R.; Wang, Z.J. Video-Based Heart Rate Measurement: Recent Advances and Future Prospects. *IEEE Trans. Instrum. Meas.* **2019**, *68*, 3600–3615. [[CrossRef](#)]
26. Al-Naji, A.; Gibson, K.; Lee, S.-H.; Chahl, J. Monitoring of cardiorespiratory signal: Principles of remote measurements and review of methods. *IEEE Access* **2017**, *5*, 15776–15790. [[CrossRef](#)]
27. Al-Naji, A.; Gibson, K.; Lee, S.-H.; Chahl, J. Real-time apnoea monitoring of children using the Microsoft Kinect sensor: A pilot study. *Sensors* **2017**, *17*, 286. [[CrossRef](#)] [[PubMed](#)]
28. Al-Naji, A.; Tao, Y.; Smith, I.; Chahl, J. A Pilot Study for Estimating the Cardiopulmonary Signals of Diverse Exotic Animals Using a Digital Camera. *Sensors* **2019**, *19*, 5445. [[CrossRef](#)] [[PubMed](#)]
29. Froesel, M.G. Automated video-based heart rate tracking for the anesthetized and behaving monkey. *Sci. Rep.* **2020**, *10*, 17940. [[CrossRef](#)]
30. Wang, D.; Eckert, J.; Teague, S.; Al-Naji, A.; Haun, D.; Chahl, J. Estimating the cardiac signals of chimpanzees using a digital camera: Validation and application of a novel non-invasive method for primate research. *Behav. Res.* **2023**, 1–19. [[CrossRef](#)] [[PubMed](#)]
31. Yamazaki, Y.; Shibasaki, T.; Tonooka, H. A Monitoring System of a Hamster Based on Video Image Analysis. In Proceedings of the 2022 International Conference on Artificial Life and Robotics (ICAROB2022), Oita, Japan, 20–23 January 2022.
32. 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; pp. 2961–2969.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.