Sensor signal processing

Data Augmentation for Inertial Sensor Data in CNNs for Cattle Behavior Classification

Chao Li[1], Korkut Kaan Tokgoz[1], Masamoto Fukawa[1,2], Jim Bartels[1], Takumi Ohashi[1], Ken-ichi Takeda[3], and Hiroyuki Ito[1]

[1] *Tokyo Institute of Technology, Yokohama 226-8503, Japan.*
[2] *TechnoPro Design Company, Tokyo 106-6135, Japan.*
[3] *Shinshu University, Matsumoto 390-8621, Japan.*

Manuscript received September 25, 2021; accepted October 5, 2021. Date of publication October 11, 2021; date of current version October 26, 2021.

Abstract—Cattle behavior monitoring is critical for understanding cattle welfare and health status. One of the most powerful and cost-effective monitoring methods is a neural network-based monitoring system that analyzes time series data from inertial sensors attached to cows. However, while deep learning has achieved many successes in pattern recognition, large-scale datasets are often required. When given a limited number of data, data augmentation is an extremely useful and low-cost preprocessing step for neural network-based systems. Data augmentation for inertial sensor data, however, has yet to be thoroughly investigated. This letter proposes several inertial sensor data augmentation methods in a manner that fits the characteristics of cattle behavioral data. The proposed approaches are applied to the task of cattle behavior classification with convolutional neural networks, which is challenging given limited data. The classification performance increases from 83.07 to 94.43% with appropriate augmentation steps. In conclusion, the data augmentation approaches presented here can help to improve deep learning performance regarding cattle behavior classification and decrease the overall system cost stemming from data acquisition and labeling.

Index Terms—Cattle behavior classification, data augmentation, neural networks, sensor signal processing.

I. INTRODUCTION

The quality and quantity of end products in the dairy and livestock industries are directly related to animal welfare and health status, which can be understood through animal behaviors [1]. As such, cattle behavior monitoring has a significant role in improving the amount and quality of dairy and beef products. Thus, the accurate assessment of a cattle's behavior is essential to ascertain its health status. Conventional cattle state assessment methods rely on visual observations by the farmer. Studies have proposed automating the assessment process based on changes in activity levels detected with inertial sensors. One of the most powerful and cost-effective methods is a neural network-based monitoring system that analyzes time series data obtained from inertial sensors attached to cows.

Convolutional neural networks (CNNs) have proven to be quite effective in many challenging pattern recognition applications. Part of the latest success achieved by deep learning is due to the recent availability of a massive amount of data. On the other hand, applying CNNs to problems presents a challenge when only a small labeled dataset is available. For instance, acquiring and annotating a vast number of behavioral data can be problematic due to the high cost of the data collection process. As a result, applying CNNs to activity classification with small-scale behavioral data is challenging.

One potential solution to address this issue is to perform data augmentation [2]–[9]. However, the generation of new data that are consistent with the correct labels is challenging and usually necessitates domain expertise. In certain domains, such as those involving inertial sensor data, how to augment data with preserved labels is not obvious. For instance, acceleration data generated by scaling an original dataset might cause some labels to be changed since some labels are discriminated by their movement intensity.

This letter specifically focuses on cattle behavior classification. Cattle behavior monitoring is a challenging task due to the presence of extraneous activity disturbances, substantial individual variability, noisy labels, and restricted access to labeled data [8], [10], [11]. In this work, several augmentation approaches for inertial sensor data are proposed, thereby successfully addressing the challenging cattle behavior classification task with low-cost data collection based on CNNs.

Our contributions are as follows. We work with small-scale behavioral data using CNNs applied to the task of cattle behavior classification. In addition, we propose several label-preserving data augmentation techniques for inertial sensor data. Furthermore, we perform an experimental comparison of the proposed approaches using a CNN-based behavioral model.

II. RELATED WORK

Data augmentation is a critical preprocessing step used to attain maximum performance in deep learning. Most time series data augmentations are based on random transformations of the given training data, e.g., jittering, window slicing, permutation, scaling, and random warping in the time dimension, magnitude dimension, or frequency dimension. Such augmentation approaches have been applied in a wide variety of time series data. For example, Le Guennec *et al.* [12] used a window slicing and warping technique to augment time series data, extracting multiple small-size slices by speeding up/slowing down a randomly selected slice of the input time series. The problem with random transformation-based approaches is that they do not

Corresponding author: Hiroyuki Ito (e-mail: ito@pi.titech.ac.jp).
Associate Editor: Solmaz Kia.
Digital Object Identifier 10.1109/LSENS.2021.3119056

consider the different property characteristics of every dataset, and the transformed data may not maintain the correct labels. For example, acceleration data are generated by magnitude warping because some labels are discriminated by their magnitude levels.

More recently, pattern mixing and generative model-based methods have been proposed. In particular, generative adversarial networks (GANs) [13] have received significant attention in many fields and have been proven to be effective for generating synthetic samples [14]. However, GANs require domain-specific properties for building mathematical models or for external training. For pattern mixing, i.e., mixing several time series belonging to the same class, one approach was provided by Takahashi *et al.* [15], who added together random sequences of interclass sounds at various ratios. Nevertheless, adding two segments together might change the labels for behavioral data. In this work, we creatively recombine two segments at a random ratio to augment data.

Furthermore, we propose a new data augmentation approach to compensate for data loss, thus avoiding the loss of valuable data that may occur with the conventional approaches. Additionally, by considering the repeatability of behavioral data, a reversal (REV) approach is also proposed. The difference between the proposed approaches and the existing approaches is that instead of focusing on the generation of new data, we leverage limited data as much as possible so that the data labels are preserved, and we attempt to address problems with few complexities.

III. CATTLE BEHAVIOR CLASSIFICATION

A. Challenges in Behavioral Data

We consider five frequent cattle behaviors: feeding, walking, salting, ruminating, and resting. Notably, the datasets collected contain 530 485 data rows at 25 Hz, i.e., the data are approximately 5.89 h long; such datasets tend to be tiny. By comparison, in [16], approximately 63 h in total were captured, and a CNN model with long short-term memory (CNN-LSTM) was used to detect five basic behaviors (drinking, ruminating, walking, standing, and lying). In [17], the active video was approximately 68 h long and applied to an LSTM model. In [8], the video was approximately 32 h long, and random rotation (ROT) based augmentation was proposed to augment the data and address the imbalanced learning problem.

In addition to the difficulties of obtaining a vast amount and many varieties of behavioral data, creating a high-quality dataset necessitates the manual classification and annotation of the collected data, which is time consuming and laborious. Thus, to fully exploit the potential of deep learning and enable it to work effectively on small-scale datasets, accurate and effective augmentation approaches for behavioral data are urgently needed. On the other hand, the behavioral complexity of animals introduces challenges in real-world applications because different activities may generate similar sensor data, e.g., acceleration data. The reason for this phenomenon is that various activities involve similar gestures, for example, standing rumination and standing resting. Additionally, individual differences increase interclass variability. As a result, cattle activity classification is highly challenging, especially when the amount of input data is relatively small.

B. Data Augmentation Methods for Inertial Sensor Data

Data augmentation can be conceived as an injection of prior information about data attributes that are invariant against certain transformations. Augmented data attempt to cover the unknown and unfamiliar spaces of input patterns, thereby alleviating the effect of the overfitting problem and improving the generalization ability of neural network (NN). Minor changes in image data, such as scaling and rotating, are known to have negligible impacts on data labels because such changes are likely to occur in real-world observations. However, for inertial sensor data, label-preserving transformations are not immediately apparent and intuitively recognizable.

One aspect that may introduce label-invariant variability into inertial sensor data involves sensor position differences during monitoring. It is desirable to attach the sensor device at exactly the same position on the cow's neck for each measurement. However, sensor displacement is inevitable in real-world testing scenarios. Not only is there a risk of the sensor device slipping along the direction of the collar but also after the farm staff detaches the sensor device and attaches it again, although efforts are made to maximize positioning reproducibility. When the device rotates, the raw accelerometer data readings change, and the accuracy of the algorithm probably decreases considerably [18]. Therefore, different sensor device rotating states are simulated, and this type of ROT is considered an augmentation approach to cover additional data possibilities while maintaining the correct labels. Given an original time series x, the transformation formula for ROT is shown in the following equations:

$$x = x_1, x_2, \ldots, x_t, \ldots, x_T \tag{1}$$

$$x' = Rx_1, Rx_2, \ldots, Rx_t, \ldots, Rx_T \tag{2}$$

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \tag{3}$$

where $x_t = [a_{x_t}, a_{y_t}, a_{z_t}]^T$, indicating a three-axis data point at each time step t, T is the total number of time steps, and R is a three-dimensional ROT matrix [8]. Unlike the 30° ROT interval used in [8], a random angle θ in the range of 0–360° is applied here.

Perturbing the time position of the activity level in the input window is also considered another approach that can introduce label-invariant variability. Considering that cow movements, e.g., rumination, resting, and lying, are mostly repetitive, as well as the fact that window segmentation is arbitrary, the time positions of the activities in the input window may not convey useful information. Therefore, we may augment data by perturbing the time positions in the window. Recombining (REC) two sequences at a random ratio is a simple method to randomly perturb the time positions of two or more sequences. To perturb the time positions of the data between windows, we first randomly select two samples, slice the data with a ratio of N: $(1-N)$, with N ranging from 0 to 1, and recombine the segments to create a new sample. REV is another method used to perturb time positions. Assuming that the network learns data features that are time invariant, a potential strategy is to reverse the samples about the time axes and generate new time series. REV can be defined as follows:

$$x' = x_T, \ldots, x_t, \ldots, x_2, x_1. \tag{4}$$

In addition, if the length of a time series does not meet the imposed window size requirement, we generally directly abandon that part; as a result, some useful data information is lost. Under the assumption of repetitive cattle movements, a method to compensate for information loss (CIL) is proposed to fill the insufficient part and augment data by looping the existing period until the length requirement is reached. When looped twice, the augmented time series x' becomes

$$x' = x_1, \ldots, x_t, \ldots, x_E, x_1, \ldots, x_t, \ldots, x_E \tag{5}$$

where E is the number of data points in an insufficient sequence.

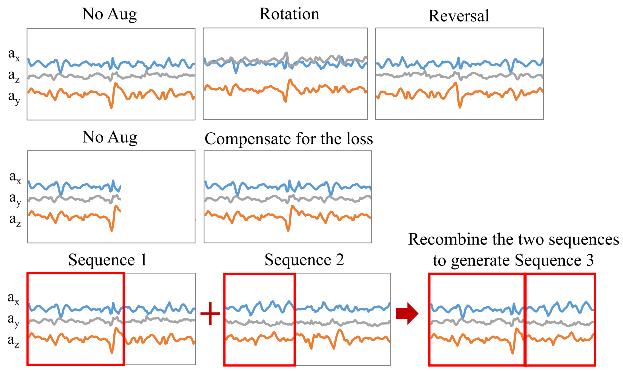

Fig. 1. Examples of time series generated in a 5-s window with the proposed data augmentation methods: ROT, REV, compensation for loss, and the recombination of two sequences. Note that the proposed compensatory approach compensates for data loss by looping the existing period. A combination of multiple augmentation methods can also be applied.

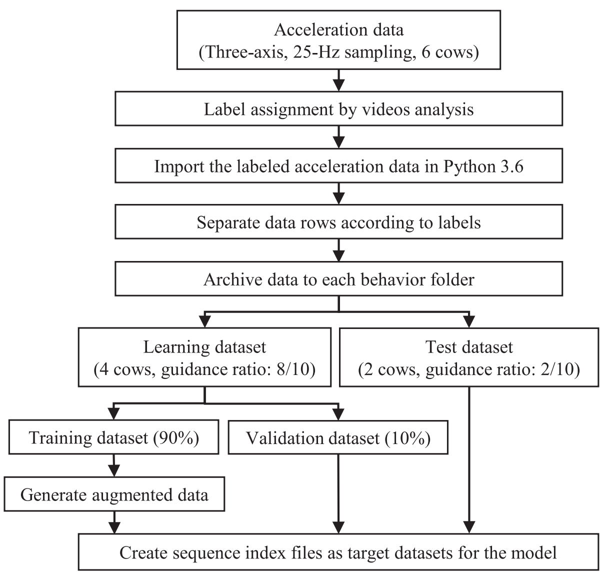

Fig. 2. Flow of data preparation process in this study.

In summary, ROT, REV, recombining two sequences at a random ratio, and the compensatory approach (see Fig. 1) are applied to augment inertial sensor data. The performance of cattle behavior classification achieved when using CNNs in conjunction with the proposed data augmentation approaches is evaluated in Section IV.

IV. EXPERIMENT

A. Data Preparation

A dataset containing six different cows under normal living conditions was collected using an acceleration sensor without human disturbance. The data were collected at a sampling frequency of 25 Hz by an accelerometer. The dataset is publicly available and can be accessed in [19]. In this letter, the dataset is randomly split into three parts. First, the dataset is divided into learning and independent test sets at an 80/20 guidance ratio, and then, 10% of the learning set (8% of the whole dataset) is used as a validation dataset (see Fig. 2). The training dataset is used to establish an initial behavioral model, the validation dataset is used to tune the model parameters during the training process, and the independent test dataset is used to evaluate the performance of the trained model.

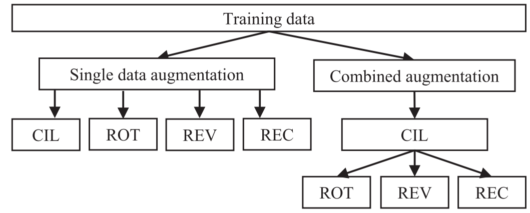

Fig. 3. Various augmentation scenarios used in this study.

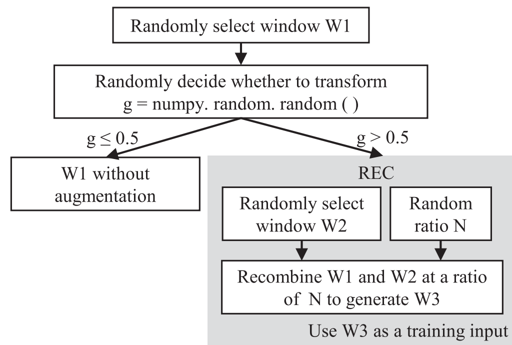

Fig. 4. Flow of the proposed REC augmentation used in this study.

B. CNN Architecture

In this letter, CNNs are utilized for the task of cattle behavior classification with limited available labeled data. CNNs are widely used to detect local features with filters sliding over time series data [20] and have the advantage of automatic feature extraction with only limited computational complexity [21]. A depth of eight layers is adopted for the CNNs to grasp the high variability of the small-scale behavioral data.

C. Results

Cattle behavior classification is performed using a CNN-based behavioral model with various data augmentation approaches. All experiments are conducted using 10-s random sliding windows for 1500 epochs. The number of training instances in every scenario, as shown in Fig. 3, is the same. For the baseline result, a CNN is applied to raw acceleration data without augmentation.

Different random parameter values are applied in the experiment. For ROT, a random ROT matrix is created for every input sequence. For REV, whether to apply the method to a given instance is decided randomly. For the recombination of mixed patterns (see Fig. 4), the recombination ratio is determined by random selection within the range from 0 to 1. For the compensatory method, we complement the input sequences with a window possessing a size less than 10 s. Notably, the compensatory method augments data by complementing the original data that would otherwise be discarded, while ROT, REV, and recombination generate new data by transforming the original training data. In this letter, we not only review the accuracy changes induced by every data augmentation approach but also evaluate the combined augmentation results of the other three data augmentation methods after compensating for the loss.

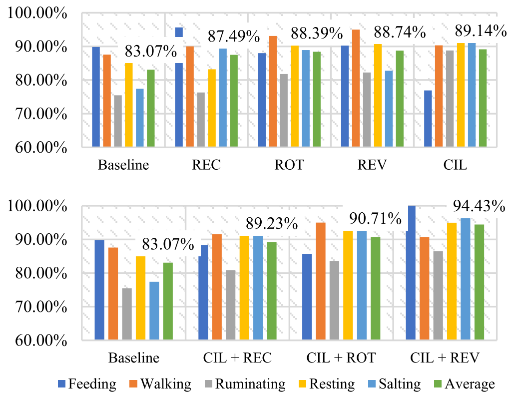

Fig. 5. Results of cattle activity classification with various data augmentation methods; the top and bottom panels correspond to single and double data augmentation, respectively. The numbers in the figure show the average F1 score for each scenario.

TABLE 1. Classification Performance on the Test Dataset With CIL + REV

	Precision	Recall	F1 score
Feeding	100.00%	100.00%	100.00%
Walking	87.50%	94.23%	90.74%
Ruminating	98.77%	76.92%	86.49%
Resting	90.86%	99.38%	94.93%
Salting	100.00%	100.00%	100.00%

The main results are presented in Fig. 5, where the numbers in the figure show the average F1 score for each scenario. The combinations of the proposed compensatory approach achieve better performance than a single data augmentation approach. Specifically, the combined augmentation methods outperform the baseline behavioral model by 6.16–11.36%. The best classification performance is achieved by CIL + REV, with an average F1 score of 94.43% (see Table 1). Although problems may remain in real-life applications, the proposed data augmentation approaches provide opportunities to improve the results obtained in other studies.

V. CONCLUSION

The task of cattle behavior classification is challenging due to the limited amount of available data. Using a CNN-based behavioral model and the proposed data augmentation approaches, this challenging task is successfully addressed in this work. The combination of the CIL + REV augmentation achieves the best performance, improving the baseline accuracy of 83.07 to 94.43%.

ACKNOWLEDGMENT

This work was supported in part by the Center of Innovation Program from JST-COI under Grant JPMJCE1309 and in part by the Research Program on the Development of Innovative Technology (funding agency: Bio-oriented Technology Research Advancement Institution, NARO). The work of Chao Li was supported by the program of the China Scholarships Council (No. 201706060192). The authors would like to thank the EISESiV members for their support.

REFERENCES

[1] R. Doyle and J. Moran, *Cow Talk: Understanding Dairy Cow Behaviour to Improve Their Welfare on Asian Farms*. Collingwood, VIC, Australia: CSIRO, 2015.

[2] D. She, X. Lou, and W. Ye, “RadarSpecAugment: A simple data augmentation method for radar-based human activity recognition,” *IEEE Sensors Lett.*, vol. 5, no. 4, pp. 1–4, Apr. 2021.

[3] X. Zhang, Z. Wang, D. Liu, Q. Lin, and Q. Ling, “Deep adversarial data augmentation for extremely low data regimes,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 1, pp. 15–28, Jan. 2021.

[4] T. Hasegawa, “Smartphone sensor-based human activity recognition robust to different sampling rates,” *IEEE Sensors J.*, vol. 21, no. 5, pp. 6930–6941, Mar. 2021.

[5] J. Ding, B. Chen, H. Liu, and M. Huang, “Convolutional neural network with data augmentation for SAR target recognition,” *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 364–368, Mar. 2016.

[6] J. Salamon and J. P. Bello, “Deep convolutional neural networks and data augmentation for environmental sound classification,” *IEEE Signal Process. Lett.*, vol. 24, no. 3, pp. 279–283, Mar. 2017.

[7] X. Cui, V. Goel, and B. Kingsbury, “Data augmentation for deep neural network acoustic modeling,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 9, pp. 1469–1477, Sep. 2015.

[8] C. Li *et al.*, “A data augmentation method for cow behavior estimation systems using 3-axis acceleration data and neural network technology,” *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, to be published.

[9] Q. Wen *et al.*, “Time series data augmentation for deep learning: A survey,” in *Proc. Int. Jt. Conf. Artif. Intell.*, pp. 4653–4660, 201, *arXiv:2002.12478*.

[10] H. W. Neave, D. M. Weary, and M. A. G. von Keyserlingk, “Review: Individual variability in feeding behaviour of domesticated ruminants,” *Animal*, vol. 12, pp. 419–430, 2018.

[11] R. Arablouei, L. Currie, B. Kusy, A. Ingham, P. L. Greenwood, and G. Bishop-Hurley, “In-situ classification of cattle behavior using accelerometry data,” *Comput. Electron. Agriculture*, vol. 183, 2021, Art. no. 106045.

[12] A. Le Guennec, S. Malinowski, and R. Tavenard, “Data augmentation for time series classification using convolutional neural networks,” in *Proc. ECML/PKDD Workshop Adv. Analytics Learn. Temporal Data*, 2016.

[13] I. J. Goodfellow *et al.*, “Generative adversarial networks,” *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020, *arXiv:1406.2661*.

[14] K. Nikolaidis, S. Kristiansen, V. Goebel, T. Plagemann, K. Liestøl, and M. Kankanhalli, “Augmenting physiological time series data: A case study for sleep apnea detection,” in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases*, 2020, pp. 376–399.

[15] N. Takahashi, M. Gygli, B. Pfister, and L. van Gool, “Deep convolutional neural networks and data augmentation for acoustic event detection,” in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, pp. 1982–2986, 2016, *arXiv:1604.07160*.

[16] D. Wu *et al.*, “Using a CNN-LSTM for basic behaviors detection of a single dairy cow in a complex environment,” *Comput. Electron. Agriculture*, vol. 182, 2021, Art. no. 106016.

[17] Y. Peng *et al.*, “Classification of multiple cattle behavior patterns using a recurrent neural network with long short-term memory and inertial measurement units,” *Comput. Electron. Agriculture*, vol. 157, pp. 247–253, 2019.

[18] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, “Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities,” *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–40, 2021.

[19] H. Ito *et al.*, Japanese black beef cow behavior classification dataset, Zenodo, 2021. [Online]. Available: https://doi.org/10.5281/zenodo.5399259

[20] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, “Convolutional neural networks for time series classification,” *J. Syst. Eng. Electron.*, vol. 28, no. 1, pp. 162–169, 2017.

[21] C.-L. Liu, W.-H. Hsaio, and Y.-C. Tu, “Time series classification with multivariate convolutional neural network,” *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4788–4797, Jun. 2019.