2023 Design, Automation & Test in Europe Conference (DATE 2023)

Towards Smart Cattle Farms: Automated Inspection of Cattle Health with Real-Life Data

Yigit Tuncel*, Toygun Basaklar*, Mackenzie Smithyman†, João Dórea‡,
Vinícius Nunes De Gouvêa§, Younghyun Kim* and Umit Ogras*
*Department of Electrical and Computer Engineering, University of Wisconsin - Madison, USA
†Department of Animal and Range Sciences, New Mexico State University, USA
‡Animal and Dairy Sciences, University of Wisconsin - Madison, USA
§Texas A&M AgriLife, Texas A&M University, USA

I. Introduction

Cattle diseases have a significant negative impact not only on the animals' welfare but also on the economic performance of the cattle industry [1], [2]. For example, Bovine Respiratory Disease is responsible for approximately 75% of the morbidity and 57% of the mortality in US feedlots, which is estimated to cost the agriculture industry about $1B annually [1], [2]. The current management practice to diagnose and select cattle for treatment is a widespread clinical scoring system called DART (Depression, Appetite, Respiration, and Temperature). DART requires manual labor and skilled personnel, which is a limiting factor due to labor-shortage in several industry sectors, including agriculture [3]. Therefore, a continuous and automated IoT solution to predict the health state of a cow is a critical tool for the cattle industry.

This paper presents a wearable smart cattle health monitoring system that can completely operate at the edge. Furthermore, this is the first study that learns how to use sensor data to predict DART scores that differentiate healthy and sick animals. We thoroughly analyze accelerometer data collected from 54 cows to identify patterns that relate to their daily behavior. Then, we systematically construct 33 features that enable us classify healthy and sick animals using eleven shallow decision tree (DT) classifiers and majority voting. Our approach outperforms *thirteen state-of-the-art (SOTA) time series classifiers* in the literature, with 78% accuracy in differentiating healthy and sick cows. Moreover, the proposed approach uses only 1 KB on-chip SRAM and consumes 29 µJ in a day on a prototype wearable device.

II. Methodology

A. Dataset Collection and Preprocessing

This experiment was conducted with 54 cows at a Livestock Research Center over 25 days following a protocol approved by Institutional Animal Care and Use Committee. A 4-point scale technique based on DART as defined by [4] was used by trained personnel to track animal health every day throughout the trial. In addition, HOBO Pendant G acceleration data loggers were fitted to all 54 cows on the same day of feedlot arrival. They were mounted to the right rear leg to record 3-axis acceleration data at 1-minute intervals.

After pre-processing the accelerometer data from all 54 cows, we obtain 350 segments; 180 and 170 segments for "Healthy" and "Sick" classes, respectively. Each segment is a (3, 1440) matrix, where 3 is the number of accelerometer channels and 1440 is the minutes in a day.

B. Feature Construction

The 3-axis acceleration data consists of three distinct clusters that generalize to all cows in the dataset. We can infer that these three clusters correspond to different activities, such as, lying down, standing, and eating or a similar activity. Therefore, we use a density-based clustering algorithm (`HDBSCAN` [5]) to map each sample to one of the three clusters, and generate a one-dimensional *tri-state* time series data. This data is valuable since it correlates with the daily cattle behavior, e.g., how long they lie down, feed, and how often they change their state. Thus, we transform the raw accelerometer data to a one-dimensional time series to reduce redundancy and improve correlation with daily behavior that is representative of cattle health.

Inspecting the one-dimensional data reveals consistent observations, which motivated us to construct a set of features that encode the daily behavior. We first divided each day into three 8-hour segments since the cow behavior show variations during night, morning, and later in the day. For each segment, we count the number of minutes spent in each state and the number of state transitions. As a result, we end up with 27 features that summarize the daily behavior of the animal. Moreover, we add six more features: The max, min, $|max\text{-}min|$, norm, variance and the mean of the feature vector.

III. Experiments

A. Experimental Setup

We aim for a system that generalizes to all cows rather than focusing on a single subject. Therefore, we first randomly sample 10 cows out of the 54 cows for the test set. Then, data from the remaining 44 cows are merged into a training set. Furthermore, we repeat the procedure 7 times for each classifier to minimize the effects of random sampling, i.e., we apply 7-fold cross validation.

According to the results presented in a recent review [6], we choose *thirteen* SOTA time series classifiers that perform

978-3-9819263-7-8/DATE23/© 2023 EDAA

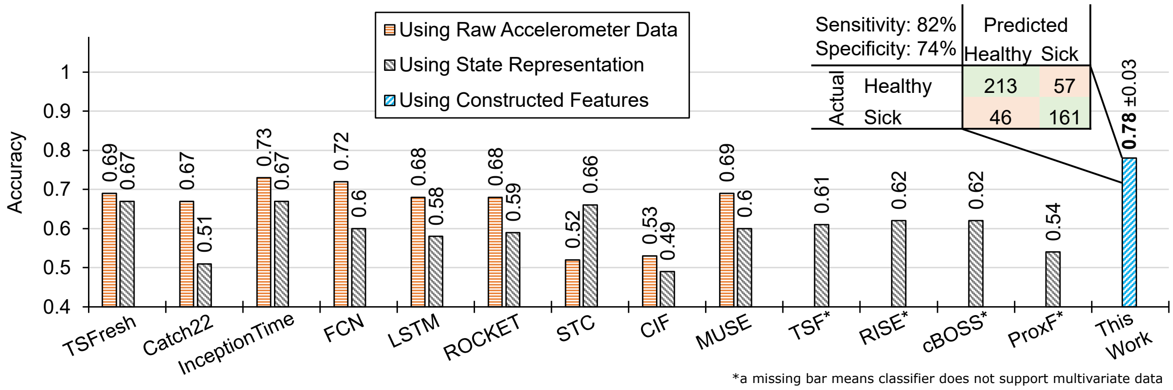

Fig. 1: 7-fold cross validation results of all classifiers. Compared to InceptionTime, our technique does not only increase the accuracy from 73% to 78%, it also uses over 36× less memory and has negligible execution time and energy consumption.

statistically better than the rest. LSTM, FCN, and InceptionTime are deep learning approaches, implemented in the `tsai`[1] python package. TSFresh and Catch22 are automatic feature extractors that extract predefined sets of features from the input data. We use the `tsfresh`[2] and `sktime`[3] python packages for the TSFresh and Catch22 implementations, respectively. TSF, RISE, ROCKET, STC, CIF, MUSE, cBOSS and ProxF are statistical methods that apply various transformations to the data and extract features from the transformed data. We use their implementations in the `sktime`[3] python package.

B. Experimental Results

Using the 3-axis raw accelerometer data: The raw accelerometer data has three channels, i.e., it is a multivariate data. Nine of the thirteen classifiers support multivariate signals. Their average accuracy (over 7-fold CV) are shown by orange striped bars in Figure 1. Among them, InceptionTime obtains the highest accuracy with 73% and FCN follows closely with 72%. CIF and STC perform very poorly with 53% and 52% classification accuracy, respectively.

Although InceptionTime and FCN can obtain relatively good accuracy, using these classifiers directly on the raw data has several drawbacks. First, they are hard to interpret, as we do not understand the implicit representation of the data within the classifiers. Second, their computational complexities are high, rendering them unsuitable for resource-constrained edge devices. Finally, the data processing overhead is significant since they use all of the data.

Using the extracted state representation: All thirteen classifiers support univariate signals. Figure 1 shows the accuracy of all of them with diagonally striped black bars. Using the extracted state representation needs less processing effort than using the raw data, but the classification accuracy deteriorates significantly. The accuracies for this case are lower than 70% across the stack. In addition, most of the multivariate classifiers perform worse, STC being the only exception. For example, the accuracy of InceptionTime decreases from 73% to 67%.

[1] `tsai`. [https://github.com/timeseriesAI/tsai]
[2] `tsfresh`. [https://github.com/blue-yonder/tsfresh]
[3] `sktime`. [https://github.com/alan-turing-institute/sktime]

Proposed Technique: Our proposed approach uses a Random Forest classifier with 11 decision trees, each with a maximum depth of 10. We achieve 78% classification accuracy with 82% sensitivity and 74% specificity using our constructed feature set, as illustrated in Figure 1.

Our approach not only achieves the highest accuracy, but also has significantly smaller memory footprint and energy consumption. We implemented the proposed real-time technique on TI CC2652R Microcontroller, which has 48 MHz operating frequency, 352 KB flash memory, and 80 KB SRAM. Our approach requires only 10 KB flash (for program) and 1 KB SRAM (for data) memory, over 36× smaller memory requirement than the approaches that operate on raw data. Furthermore, the proposed technique is extremely light-weight with negligible runtime overhead and only 29 µJ (2.44 nAh at 3.3V) daily energy consumption.

IV. Conclusion

This paper presented a novel IoT system that predicts animal health using accelerometer data. Thorough evaluation on a real dataset collected from 54 cows shows that the proposed approach achieves 78% accuracy in classifying healthy and sick cows, which is higher than thirteen state of the art time series classifiers. Furthermore, the proposed system is extremely lightweight, using only 1 KB on-chip SRAM and less than 29 µJ in a day.

References

[1] G. A. Hanzlicek *et al.*, "Serial evaluation of physiologic, pathological, and behavioral changes related to disease progression of experimentally induced mannheimia haemolytica pneumonia in postweaned calves," *American jrnl. of veterinary research*, vol. 71, no. 3, pp. 359–369, 2010.

[2] R. E. Centeno-Martinez *et al.*, "Identification of bovine respiratory disease through the nasal microbiome," *Animal microbiome*, vol. 4, no. 1, pp. 1–18, 2022.

[3] AGAmerica, "The U.S. Farm Labor Shortage," https://agamerica.com/blog/the-impact-of-the-farm-labor-shortage, accessed Sep 2022.

[4] D. Step *et al.*, "Effects of commingling beef calves from different sources and weaning protocols during a forty-two-day receiving period on performance and bovine respiratory disease," *Journal of Animal Science*, vol. 86, no. 11, pp. 3146–3158, 2008.

[5] L. McInnes, J. Healy, and S. Astels, "hdbscan: Hierarchical density based clustering," *Jrnl. of Open Source Soft., 2(11)*, p. 205, 2017.

[6] A. P. Ruiz *et al.*, "The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances," *Data Min. Knowl. Discov.*, vol. 35, no. 2, pp. 401–449, 2021.