

AI BASED DEVELOPMENT OF A LOW COMPUTATIONAL INTENSITY ALGORITHM FOR CATTLE HEART RATE (HR) ESTIMATION

Éva HAJNAL

Óbuda University, Alba Regia Technical Faculty, 45 Székesfehérvár Budai Str., H8000, Hungary

Corresponding author email: hajnal.eva@amk.uni-obuda.hu

Abstract

The main goal is to estimate the HR value from the activity sensor 3D acceleration measurements of the cattle rumen bolus. During the development of the algorithm it was intended to execute the primary calculations on the device's microcontroller and the additional calculations could be performed on the server. The proposed HR estimation algorithm is based on simple data cleaning and peak detection, but the validation and postprocessing of the detection uses AI methods, namely MLP artificial neural network with different cell numbers. The accuracy of the period estimation (IBI) was ± 50 ms, which means an 8% error. This allows basic alerts to be implemented.

Key words: 3D accelerometer data, Artificial Intelligence methods, cattle rumen bolus, HR estimation.

INTRODUCTION

PLF is one of the leading technologies in agriculture (Knight, 2020b). One of the main directions of this is to collect relevant information from the animals (Szabo & Alexy, 2022) and their environment with the appropriate sensors, process it with an information system (Cabrera & Fadul-Pacheco, 2021; Caja et al., 2016; Dado & Allen, 1994; Daum et al., 2022; El Bilali et al., 2020; Khanal et al., 2010) and use the obtained results for the purpose of automation and/or decision support. These processes are of great economic importance, as they make animal husbandry more profitable, and at the same time solve the partial replacement of the missing human workforce and help to ensure the well-being of the animals (Alsaad et al., 2012; Caja et al., 2016; Knight, 2020a; Michie et al., n.d.). A rumen bolus sensor can be used as a sensor in such a system, and it is already used in dairy cattle. Rumen bolus sensors typically measure temperature, pH and/or activity, as this is usable technical solution for operating such sensors at the current technological level (Borchers et al., 2017; Cabrera & Fadul-Pacheco, 2021; Caja et al., 2021; Campos et al., 2018; Dado & Allen, 1994; Hajnal et al., 2022; Hamilton et al., 2019; Hanušovský et al., 2017; Ipema et al., 2008; Knauer et al., 2016; Knight, 2020b; Mottram, 2010; Vakulya et al., 2022;

Zhang et al., 2018). The current systems can also be used for alarms, but I thought it was possible to expand the range of measured characteristics with new ones. In this article, it was examined whether there is a realistic possibility to determine the heart function parameters (Heart Rate –HR, Interbeat Interval –IBI) with the help of the rumen bolus. This intention is meaningful, because the heart function in cows (Caja et al., 2021; Kovács et al., 2014), just like in humans (Piros et al., 2023), is an important characteristic that gives information about the state of health, indicates the level of stress and certain events, such as the start of calving. As a result, this information can be extremely valuable in PLF systems. In the case of cows, no such tests have yet been carried out, but in the case of humans, there is a lot of literature with results available, as a common target function of modern wearable technology is the examination of heart function (Bruser et al., 2013; Curone et al., 2010; Galli et al., 2018; Hernandez et al., 2015; Kwon et al., 2011; Lahdenoja et al., 2016; Nakano et al., 2012; Zhao et al., 2021). The publications show that it is very difficult to accomplish the task precisely and efficiently. For human use, the heart rate is typically determined from optical data from pulse oximeters, and in case of activity measurement, the devices work with a high sampling frequency and often with an additional sensor. Many publications deal with

the methodology of data processing, which often uses signal processing methods with high computational demands due to the significant noise in the measurement (Alzahal et al., 2009; Galli et al., 2018; Nakano et al., 2012).

The aim of this research is to develop an algorithm with low computational requirements that can be run on the rumen bolus microcontroller and is able to estimate cardiac IBI values. Since a real device to be developed does not necessarily have a way to continuously measure and process data, it was necessary to determine how large a series of measured data could be processed with the highest accuracy. The aim is therefore to establish the optimal period durations on the basis of shorter measured data series of a few seconds. According to previous experience, the algorithms can be made sufficiently accurate with the help of some post-processing step. In this paper, an artificial neural network was used for post-processing. The research question is, with which parameters a simple algorithm gives the best estimation, as well as whether it is possible to post-process the data with the help of the neural network and whether it is possible to achieve a usable result with the relatively low sampling frequency and processing steps with little computational demand.

MATERIALS AND METHODS

In this paper, 3D acceleration data measured by the rumen bolus sensor were used. The experimental set-up and the method of data processing are described in details in the previous publications (Vakulya et al., 2022). The accelerometer was used to measure 3D acceleration data with a sampling frequency of 25Hz. The data was sent via radio communication to the receiver, which recorded the measurements supplemented them with a time stamp. Parallel ECG monitoring measurements were carried out (Hajnal et al., 2022; Kovács et al., 2014). We had no way of synchronizing the timer of the two devices, but at the same time we confirmed with another experiment that the difference was within 2s, so the ECG values can be considered as actual control values.

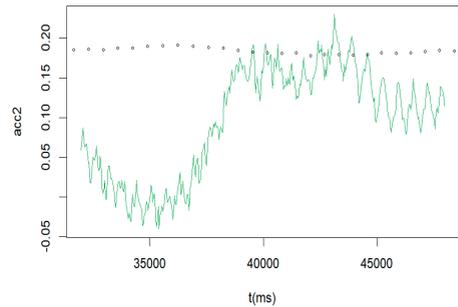


Figure 1. Raw accelerometer data of 20s long data acquisition period (green) and the control Interbeat Interval (IBI) values (dots)

The algorithm was developed using R scripts within the RStudio 2022.07.1+554 environment and Orange DataMining 3.35 software (*Orange DataMining*, n.d.).

The Figure 1. shows the raw data. It can be seen that the periods belonging to the heart can be clearly discerned in some parts of the curve, for example in the last 3s, but they are almost barely perceptible in certain parts of the curve (at 35s). It can be seen that the movement activity of the animal is strongly superimposed on the curve, and elsewhere the curve becomes detached, often resulting in completely false period length detection. The basic idea behind detection is a method based on standard pre-processing, peak detection and post-processing steps (Zhao et al., 2021). The details of the algorithm and the method of data processing can be seen in Figure 2. During the data processing, in the first step, it was a trial to reduce motion artefacts (MAs) originates from the animal's body movement using a thresholding procedure of the acceleration data. Values that differed from the average by more than the threshold were cut to the threshold level. The thresholds were established on the basis of the distribution statistics of the acceleration derivative. This was followed by a low-pass filter to remove the high-frequency signal components, it was resolved by a moving mean computation with a window of a given width. For the peak detection the two highest acceleration axes were used, because technically the cow's heart is located almost vertically upwards to the rumen, so the

accelerations resulting from the heart's action are mostly indicated on the vertical axis. The principle of peak detection was to locate the zero points of the derivative. The peak detection was performed on the basis of the minimums and maximums of the function for each data series of a few seconds in length, so final two data series were resulting in a series of IBI values detected based on the minima and maxima.

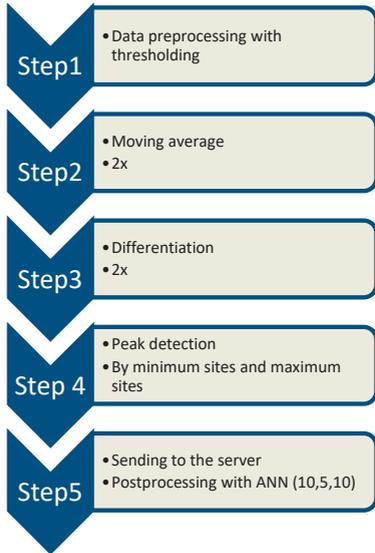


Figure 2. Data processing workload

Unfortunately, higher frequency harmonics and period lengths multiplied due to detection errors can also be included in this series. It has been described in the literature that MAs can also generate false period lengths. In this way, the minimum, maximum, average and median values from each data series were determined, so total of 16 data for the two channels were calculated which in an ideal data series are all the equal and according to the 25Hz sampling frequency the 1/40 part of the IBI values. This data set was used for post-processing. This system was investigated for the ideal data series length in which the detection based on 3-4-5 s long data series. The accuracy after detection and the accuracy after postprocessing of trained neural network were determined. An important pre-requisite for neural networking is that the dataset must be balanced. Figure 3 shows the histogram of the control ECG data set, and it is concluded that the prerequisite is not met.

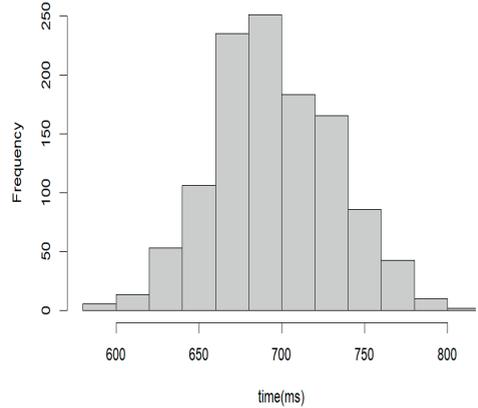


Figure 3. Histogram of control ECG Interbeat Intervals (IBI)

The training dataset of the neural network was balanced using the random walk oversampling (RWO) method (H. Zhang & Li, 2014). The neural network used is an MLP with three hidden layers with 10,5,10 neurons, RELU activation function and ADAM solver. The data was share to teaching (70%) and testing (30%) datasets. The neural network training was performed as a classification and evaluated in the usual way (ACC, MAE, ROC, F1, Confusion matrix). The presented metrics were get in all cases on testing data.

RESULTS AND DISCUSSIONS

Figure 4 shows a sample dataset after preprocessing and peak detection. We can see that in certain cases it was possible to find the real peak very accurately. Since the curves have small subpeaks (shoulders), we set the peak detection so that it does not detect two peaks close to each other, but because of this, the algorithm often does not find the most prominent peak, but instead produces the shoulder as a result. It can be concluded that the peaks are detected, but the accuracy with the specified parameters should be further improved in the future, or perhaps supplemented with an algorithm that helps to separate the main and secondary peaks. If the algorithm detects several secondary peaks, then the correct value can be obtained from the average of two or three sections, so the average of the peak distances can provide an

approximately good solution in this case. If several peaks are included in the analysis, the false peaks cause the shortening of the period length on one side and the increase of the period length on the other side. If there is a correctly identified period in the data set, it will be located near to the median of the series. As more ideal the data series, as more homogeneous the resulting IBI series, ideally all members are equal, and contrary, if there are large differences in the detection, it can definitely indicate a detection error. In such cases, it is possible to discard the data series as a bad detection, or to try to decide what the real value is in the post-processing (with the MLP artificial neural network in this paper). The distribution of errors after the first detection is shown in Figure 7 B. It can be seen that the error function has a roughly normal distribution, because many parameters cause the error, but the function has shifted a little to the left, so typically the IBI values are underestimated by the algorithm. The average error is unacceptably large, and in this form it cannot be used even for a rough estimation.

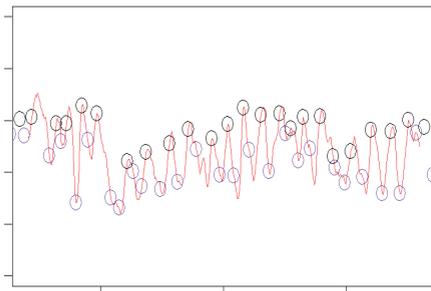


Figure 4. Result of the data processing algorithm. The red curve is the preprocessed acceleration data supplemented with the detected maximum and minimum peaks

The primarily obtained values were processed with a neural network. The primary detection can be implemented on the microcontroller of the bolus sensor and the result of the measurement is a data set of a size that can already be sent using a radio data transmission standard protocol. Post-processing using an intelligent method can already be implemented on the server side. During the implementation, we consciously tried to use a relatively small neural network in order to avoid overtraining as

much as possible, taking the risk that the obtained results would be weaker. I solved the neural network data processing as a classification task, because in this way the processing of individual IBI period categories can be evaluated in more detail. Table 1 contains the evaluation of the post-processing of data sets of size 3s, 4s, 5s. It can be seen that the Classification Accuracy (CA) value is relatively small. The best case is the processing of a 4s long data series, but even here the detection rate accuracy is only 61.5%. The same is true for other metrics like F1, Precision, Recall, MCC. At the same time, the AUC (Area Under Curve) value, which assesses the validity of the detection, the size of the area under the ROC curve, produced significantly better results. The best case here also occurred in case of processing the 4s long data series, with a value of 0.874, which is quite close to the ideal value of 1. The explanation for the two types of results is that, although in many cases it is not possible to accurately categorize the data, at the same time, in most cases, the result is not fundamentally bad at the end of the processing, only one category mistake, which is approximately corresponds to an error of 40ms. The actual IBI values fall between 600ms and 1200ms, that is, the difference of the detection means an error of 5-10%.

Table 1. Quality metrics of the ANN processing of 3s, 4s, 5s long data series primary results

Mode I	AUC	CA	F1	Prec	Reca II	MCC
NN3s	0.845	0.552	0.545	0.567	0.552	0.456
NN4s	0.874	0.615	0.616	0.635	0.615	0.532
NN5s	0.859	0.603	0.598	0.612	0.603	0.518

A detailed evaluation of the ANN post-processing can be done with the help of Figures 5 and 6. Figure 5 shows the confusion matrix for the processing of 3s, 4s, 5s long data series. Naturally, here too, the processing of the 4s long data series is the most effective in terms of almost all values. In the main diagonal, the proportion of correctly categorized data is better in all values compared to the other two

data sets. The ratio of the worst categorized data can be seen in the bottom left and upper right corners, in this respect the processed data series of 5s is the best, it contains the fewest gross errors. It can be seen that the largest numerical values are found in the main diagonal and in the band immediately next to it, which means that the majority of the data can be classified with no more than 1 category error. Detection of extreme categories is better than the average. The detection of category 15 (600ms IBI) is 93.7-97.4% accurate.

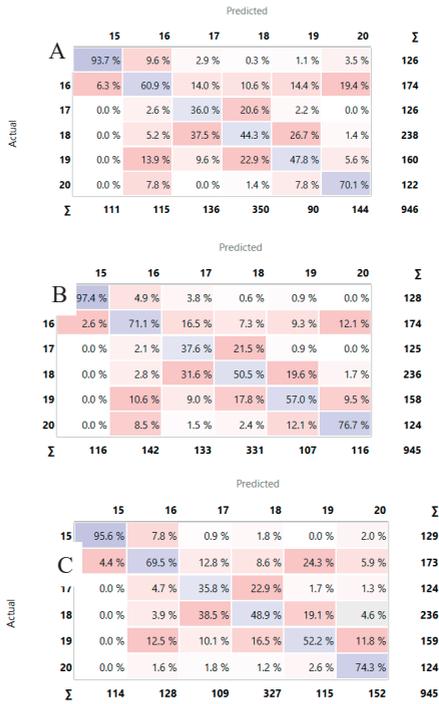


Figure 5. Confusion matrix of (A) 3s, (B) 4s, (C), 5s long data series. Each category is 40ms wide

This category is particularly important for implementing alarms related to elevated heart rate. It should be mentioned that this category can only be mistaken in one direction in the system. Figure 6 shows the ROC curve belonging to class 15 (600ms IBI). It can be seen that the shape of the curve is almost ideal and enables a good categorization of the class. Unfortunately, we cannot ignore the fact that this category was rather underrepresented in the learning dataset, and the required amount of learning data is the result of oversampling. In

such cases, relying on statistical considerations, we can hope that the processing gave a real result, but this can only be confirmed with further measurements.

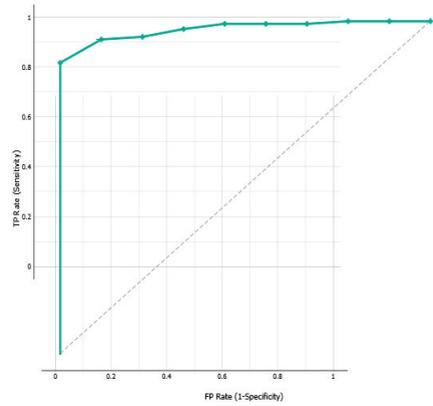


Figure 6 ROC curve for class 15 which corresponds to the 600ms IBI

Figure 7 shows the histogram of error values after primary data processing (A) and after neural network post-processing. It can be seen that the detection accuracy has improved by almost an order of magnitude.

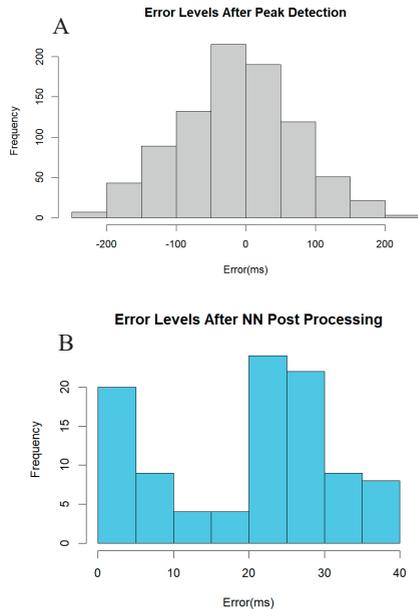


Figure 7. Histogram of IBI detection errors: (A) errors of the 4s long detection period by the median of detected maxima, (B) errors of the ANN result

Table 2. Detected IBI categories and the corresponding HR values

Cat	IBI (ms)	HR (BPM)
1	800	75
2	760	79
3	720	83
4	680	88
5	640	93
6	600	100
7	560	107
8	520	115
9	480	125

Table 2 shows the detectable IBI and HR values. Among them, it was possible to detect categories 1-6 in the current experiment.

The average error value is 20-40 ms, which is not suitable for very accurate animal health measurements, but it is already suitable for the implementation of important health warnings or alarms related to significant HR changes (disease, increased stress, initiation of calving) in PLF-related animal husbandry support applications.

CONCLUSIONS

In this paper, an algorithm development process was presented, of which the IBI and the HR values can be calculated from the measured 3D accelerometer data in a cow rumen bolus sensor. The algorithm to be implemented must be of low computational intensity that it can also be implemented on the sensor's microcontroller. In such a case, the first question is the quantity of data to be collected, for the optimum calculation can be carried out. Experiments have revealed that the collection of 4-5 seconds of data is optimal for calculations. It could be seen from the results that 4s are ideal according to most metrics, and 5s is the best in terms of gross error rate, but processing a longer sequence with such a simple algorithm no longer improves the result. The 4s-5s data series can mean 100 or 125 measured data, considering the 25 Hz sampling frequency. The measurement error is 40ms in this case, which theoretically can be improved to a 10 ms detection accuracy assuming 3-4 cycles of detection, but unfortunately, due to

the significant noise of the measurement, such good results cannot be obtained on real data. Measurements were evaluated with a simple O (N) data cleaning and peak detection algorithm. The error of the results obtained is of a magnitude greater than the theoretical minimum. Based on the results, it can be concluded that it would be worthwhile to use a slightly more skillful data cleaning and peak detection algorithm in the future, but it also appears that due to the high noise ratio of the measurement, this alone is not sufficient. It is absolutely necessary to subject the measurement to some kind of post-processing. In this work, a neural network was examined as a post-processing method and it was found that the neural network is a suitable solution for post-processing. A relatively small sized neural network was consciously chosen to avoid the risk of over teaching. Postprocessing with a neural network significantly improved detection accuracy and enabled detection with a ± 40 ms error. The resulting method is thus appropriate for classifying cardiac function into six categories. The accuracy of the obtained results can be improved in the future and the current accuracy is not yet suitable for veterinary purposes. On the other hand, the system is suitable for use in information systems related to PLF, it is probably suitable auxiliary data for detection of stress on animals, for alarms related to calving and diseases.

The HR values above 100 could not be measured by the control ECG measurement, so there is currently no information on the detectability of these values.

The next step of the research is to conduct additional measurements and gather the corresponding control data. After that, the refinement of the data cleaning and peak detection algorithm and the examination of the neural network solution in the detection of values that are underrepresented or not measured at all in the current experiment.

ACKNOWLEDGEMENTS

Thanks for Albacomp Zrt for creation of the bolus and for Levente Kovács (MATE) for help in control measurements. Lot of thanks for Etyek Farm to support us as an experimental place.

REFERENCES

- Alsaood, M., Römer, C., Kleinmanns, J., Hendriksen, K., Rose-Meierhöfer, S., Plümer, L., & Büscher, W. (2012). Electronic detection of lameness in dairy cows through measuring pedometric activity and lying behavior. *Applied Animal Behaviour Science*, *142*(3–4), 134–141.
- Alzahal, O., Steele, M. A., Valdes, E. V., & McBride, B. W. (2009). *Technical note: The use of a telemetric system to continuously monitor ruminal temperature and to predict ruminal pH in cattle*. 5697–5701. <https://doi.org/10.3168/jds.2009-2220>
- Borchers, M. R., Chang, Y. M., Proudfoot, K. L., Wadsworth, B. A., Stone, A. E., & Bewley, J. M. (2017). Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle. *Journal of Dairy Science*, *100*(7), 5664–5674.
- Bruser, C., Winter, S., & Leonhardt, S. (2013). How speech processing can help with beat-to-beat heart rate estimation in ballistocardiograms. *MeMeA 2013 - IEEE International Symposium on Medical Measurements and Applications, Proceedings*, 12–16. <https://doi.org/10.1109/MeMeA.2013.6549696>
- Cabrera, V. E., & Fadul-Pacheco, L. (2021). Future of dairy farming from the Dairy Brain perspective: Data integration, analytics, and applications. *International Dairy Journal*, *121*, 105069. <https://doi.org/10.1016/j.idairyj.2021.105069>
- Caja, G., Castro-Costa, A., & Knight, C. H. (2016). Engineering to support wellbeing of dairy animals. *Journal of Dairy Research*, *83*(2), 136–147.
- Caja, G., Castro-Costa, A., & Knight, C. H. (2021). *Engineering to support wellbeing of dairy animals Background and current scenario*. <https://doi.org/10.1017/S0022029916000261>
- Campos, D. P., Abatti, P. J., Bertotti, F. L., Hill, J. A. G., & da Silveira, A. L. F. (2018). Surface electromyography segmentation and feature extraction for ingestive behavior recognition in ruminants. *Computers and Electronics in Agriculture*, *153*, 325–333.
- Curone, D., Tognetti, A., Secco, E. L., Anania, G., Carbonaro, N., De Rossi, D., & Magenes, G. (2010). Heart rate and accelerometer data fusion for activity assessment of rescuers during emergency interventions. *IEEE Transactions on Information Technology in Biomedicine*, *14*(3), 702–710.
- Dado, R. G., & Allen, M. S. (1994). Variation in and Relationships Among Feeding, Chewing, and Drinking Variables for Lactating Dairy Cows. *Journal of Dairy Science*, *77*(1), 132–144.
- Daum, T., Ravichandran, T., Kariuki, J., Chagunda, M., & Birner, R. (2022). Connected cows and cyber chickens? Stocktaking and case studies of digital livestock tools in Kenya and India. *Agricultural Systems*, *196*, 103353. <https://doi.org/https://doi.org/10.1016/j.agry.2021.10.3353>
- El Bilali, H., Bottalico, F., Ottomano Palmisano, G., & Capone, R. (2020). Information and communication technologies for smart and sustainable agriculture. *IFMBE Proceedings*, *78*. https://doi.org/10.1007/978-3-030-40049-1_41
- Galli, A., Narduzzi, C., & Giorgi, G. (2018). Measuring Heart Rate during Physical Exercise by Subspace Decomposition and Kalman Smoothing. *IEEE Transactions on Instrumentation and Measurement*, *67*(5), 1102–1110.
- Hajnal, É., Kovács, L., & Vakulya, G. (2022). Dairy Cattle Rumen Bolus Developments with Special Regard to the Applicable Artificial Intelligence (AI) Methods. *Sensors*, *22*(18), 6812.
- Hamilton, A. W., Davison, C., Tachtatzis, C., Andonovic, I., Michie, C., Ferguson, H. J., Somerville, L., & Jonsson, N. N. (2019). Identification of the Rumination in Cattle Using Support Vector Machines with Motion-Sensitive Bolus Sensors. *Sensors*, *19*(1165), 1–14.
- Hanušovský, O., Biro, D., Šimko, M., Gálik, B., Juráček, M., Rolíneck, M., & Herkeľ, R. (2017). Drinking regime evaluation with continuous ruminal monitoring boluses. *Acta Fytotechnica et Zootechnica*, *20*(1), 1–5.
- Hernandez, J., McDuff, D., & Picard, R. W. (2015). Biowatch: Estimation of heart and breathing rates from wrist motions. *Proceedings of the 2015 9th International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth 2015*, 169–176.
- Ipema, A. H., Goense, D., Hogewerf, P. H., Houwers, H. W. J., & van Roest, H. (2008). Pilot study to monitor body temperature of dairy cows with a rumen bolus. *Computers and Electronics in Agriculture*, *64*(1), 49–52.
- Khanal, A. R., Gillespie, J., & MacDonald, J. (2010). Adoption of technology, management practices, and production systems in US milk production. *Journal of Dairy Science*, *93*(12), 6012–6022.
- Knauer, W. A., Godden, S. M., & McDonald, N. (2016). Preliminary evaluation of an automated indwelling rumen temperature bolus measurement system to detect pyrexia in preweaned dairy calves. *Journal of Dairy Science*, *99*(12), 9925–9930.
- Knight, C. H. (2020a). Review: Sensor techniques in ruminants: more than fitness trackers. *Animal*, *14*, s187–s195.
- Knight, C. H. (2020b). Review: Sensor techniques in ruminants: More than fitness trackers. *Animal*, *14*(S1), S187–S195.
- Kovács, L., Jurkovich, V., Bakony, M., Szenci, O., Póti, P., & Tözsér, J. (2014). Welfare implication of measuring heart rate and heart rate variability in dairy cattle: Literature review and conclusions for future research. *Animal*, *8*(2), 316–330.
- Kwon, S., Lee, J., Chung, G. S., & Park, K. S. (2011). Validation of heart rate extraction through an iPhone accelerometer. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 5260–5263.
- Lahdenoja, O., Humanen, T., Tadi, M. J., Pankaala, M., & Koivisto, T. (2016). Heart rate variability estimation with joint accelerometer and gyroscope sensing. *Computing in Cardiology*, *43*, 717–720.

- Michie, C., Andonovic, I., Davison, C., Hamilton, A., Tachtatzis, C., Jonsson, N., Duthie, C.-A., Bowen, J., & Gilroy, M. (n.d.). The Internet of Things enhancing animal welfare and farm operational efficiency. *Journal of Dairy Research*, 87(S1), 20–27.
- Mottram, T. T. (2010). *Is A Lifetime Rumen Monitoring Bolus Possible?* <http://precisiondairy.com/proceedings/s11mottram.pdf>
- Nakano, M., Konishi, T., Izumi, S., Kawaguchi, H., & Yoshimoto, M. (2012). Instantaneous heart rate detection using short-time autocorrelation for wearable healthcare systems. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 6703–6706. <https://doi.org/10.1109/EMBC.2012.6347532>
- Orange DataMining. (n.d.). Retrieved May 10, 2023, from <https://orangedatamining.com/download/#windows>
- Piros, P., Fleiner, R., Jánosi, A., & Kovács, L. (2023). Further Evolution of Mortality Prediction with Ensemble-based Models on Hungarian Myocardial Infarction Registry. *Acta Polytechnica Hungarica*, 20(4).
- Szabo, S., & Alexy, M. (2022). Practical Aspects of Weight Measurement Using Image Processing Methods in Waterfowl Production. *Agriculture*, 12, 1869. <https://doi.org/10.3390/agriculture12111869>
- Vakulya, G., Hajnal, É., & Udvardy, P. (2022). Experimental Bolus Sensor for Dairy Cattle. *2022 IEEE 20th Jubilee International Symposium on Intelligent Systems and Informatics (SISY)*, 157–162.
- Zhang, H., & Li, M. (2014). RWO-Sampling: A random walk over-sampling approach to imbalanced data classification. *Information Fusion*, 20, 99–116.
- Zhang, L., Lu, J., Nogami, H., Okada, H., Itoh, T., & Arai, S. (2018). Solid-state pH sensor prototype for real-time monitoring of the rumen pH value of Japanese cows. *Microsystem Technologies*, 24(1), 457–463.
- Zhao, C., Zeng, W., Hu, D., & Liu, H. (2021). Robust Heart Rate Monitoring by a Single Wrist-Worn Accelerometer Based on Signal Decomposition. *IEEE Sensors Journal*, 21(14), 15962–15971.