Horse foraging behavior detection using Recurrent Neural Networks

Leon Nunes[1], Yiannis Ampatzidis[1,*], Lucas Costa[1], and Marcelo Wallau[2]

1 Agricultural and Biological Engineering department, Southwest Florida Research and Education Center, University of Florida, IFAS, 2685 SR 29 North, Immokalee, FL 34142, USA; email: i.ampatzidis@ufl.edu; Tel: +01-239-658-3451

2 Agronomy department, University of Florida, IFAS, Gainesville, FL 32611, USA; email: mwallau@ufl.edu

* Correspondence: i.ampatzidis@ufl.edu; Tel.: +01-239-658-3451

Abstract

Wearable sensing technologies can be used for precision livestock applications and to study foraging behaviors. In this context, the monitoring of chews and bites events is a critical task for livestock management. This study presents a computational tool that utilizes wearable sensing and machine learning to distinguish chew and bite events. A micro camera with a microphone (0 – 18 kHz) attached was used to obtain video/audio data from horses during feeding. The audio data collected was used to train a recurrent neural network (RNN) to detect and distinguish chews and bites events. Initial evaluation of this system shows an accuracy of 87.8% for bite identification, and an accuracy of 93.2% for chew identification. The development of this tool can, if used with a GPS system, enable the identification of better types of pasture suitable for each animal. It can reduce data collection costs and improve livestock management.

Introduction

Foraging behavior evaluation, including bite and chew detections, is an important field of study that allows researchers and ranchers to analyze and improve livestock management systems; it is key to many important decisions in grazing environments (Ungar, 1996). This evaluation can be made using the animal jaw movements: bites and chews (Ungar et al., 2006). Additionally, monitoring of grazing behavior can be economically and ecologically compatible with conservation of resources (Del Curto et al., 2005). However, monitoring this behavior can be complex, including different series of steps in the foraging process (Clapham et al., 2006).

Grazing behavior analysis is a complicated task, mainly because of the animal's movement making the data acquisition challenging (e.g., difficult to monitor jaw movement). Direct observation of the animal can be a very time-consuming task (Ungar and Rutter, 2006). Electric resistance (Penning, 1983; Rutter et al., 1997) and pressure (Nydegger et al., 2011; Zehner et al., 2012; Ruuska et al., 2016) have been used with good accuracy to differentiate grazing from ruminating processes for cattle and sheep. Another data acquisition system was developed based on electromyography, using electrodes over the jaw muscles to measure the signals related to the masticatory movements (Büchel and Sundrum, 2014). Some of these methods use esophageal-fistulated animals, and, besides making the data acquisition expensive and time consuming, it is an invasive method for the animal (Laca and de Vries, 2000). A feasible alternative to mimic bites is via hand-plucking, where one observes the animals grazing and simulates the bites taken by the herbivore by picking grass by hand (de Vries, 1994). In general, all these methods are time-consuming and labor-intensive, and usually require precise sensor calibrations and observer training.

Acoustic monitoring methods have proven to be a good and easy way to identify, differentiate and count jaw movements (Laca et al., 1994); better than direct observation due to the presence of compound movements (i.e. chew and bite in the same movement). The sound recorded has usually valuable information that can be used in different ways to predict specific jaw movements (Penning et al., 1991) and acoustic biotelemetry has enough quality and usability to give good information about the animal foraging behavior (Alkon et al., 1989).

Deep and transfer learning algorithms have been increasingly used in precision agriculture and livestock applications (Ampatzidis et al., 2019; Ampatzidis et al., 2017; Kussul et al., 2017; Partel et al., 2019; Hansen et al., 2018; Gardner et al., 1999). The use of these methods allow a significant improvement on data analysis accuracy and are starting to attract the attention from research and industrial applications (LeCun et al., 2015). Dutta et al. (2014) used a multi-classifier pattern recognition system to classify cattle behavioral patterns, using collar systems. Kaixuan and Dongjian (2015) used video analysis method with a convolutional neural network to recognize individual cows in a feeding environment.

Rutter (2000) proposed the use of the software GRAZE® to develop an automate identification system for cattle, using the signal geometric characteristics to identify the ingestive behaviors. It was reported 91% correlation between visual observations of the foraging behavior and that ones identified by GRAZE® and collected by the IGER Behavior Record system (IBR, Institute of Grassland and Environmental Research, North Wyke, Devon, UK), a system to record animal foraging behavior. However, Ungar and Rutter (2006) verified that the GRAZE® software overestimates the bite rates detection in all foraging sessions confusing bites with chews, or counting bites in the rumination process, having a mean over-estimation of 24.6%.

Herein, an algorithm was developed to identify bites and chews events for horses using an audio obtained from a video recorded with a micro camera attached. The audio was collected, processed, filtered and used as input to train a Recurrent Neural Network model to identify bite and chew events.

Materials and Methods

Test Subject and Field

The test subject is a Paint horse, weighting approximately 450 Kg, and located at the Kings Ranch (Labelle, FL). The experiment and data collection took 20 min with the subject freely grazing on a pasture while data was being collected. The field used for the experiment was 872 m^2 in area with 15 cm canopy height bahiagrass (*Paspalum notatum* Flueggé). The animal was accustomed to the pasture and had free access to water and supplemental nutrients.

Hardware and Sensor Setup

A micro-camera was attached to a 1-inch nylon halter and placed on the bottom part of the animal's jaw (Fig. 1), to avoid breathing noises. The camera was used to record a video at the same time when the audio was recorded. The system recorded a total of four 5 min videos, totalizing 20 min of audio and video data. The camera recorded an AVI (Audio Video Interleave) video format from which the audio was taken using the open source library FFmpeg®.

Experimental Procedure

Prior to data collection, this same animal had been used for the initial sensor development which served as acclimation phase. Thus, there was no interfere of the sensors on natural grazing behavior. After fitting the equipment, the subject horse was released to graze freely on the block. The data collected from the microphone was then processed and used to train an Artificial Neural Network (ANN) for recognizing foraging behaviors. The video recording is used for supporting the operator to validate and label the sounds.

Data Separation

A Dell Optiplex XX personal computer (Dell Inc., One Dell-Way, Round Rock, TX, USA) was used in the data pre-processing, neural network training and prediction. From the 20 min data, a total of 5 min were used to train the neural network, as the subject did not spend all the time eating. The sound file was separated into small pieces of audio that cover a bite or chew event (Fig. 2).

Figure 1. Micro Camera used for recording the videos and audios.

To avoid having more than one event in each audio cut, a 0.5 s consecutive window was used to cut the audio files. Fig. 3 shows the process of identification of each event in the audio and how the file was cut.

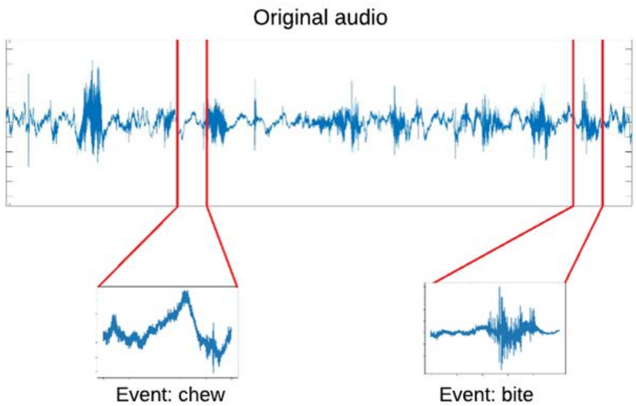


Figure 2. Workflow of the experiment.

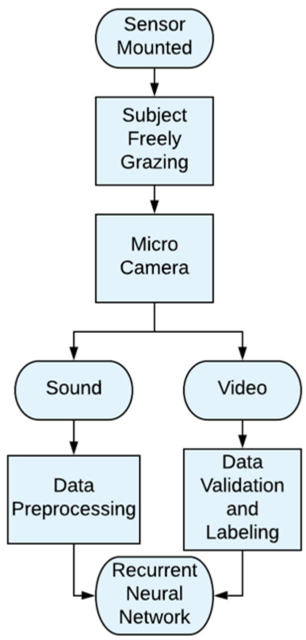


Figure 3. Original audio being cut in pieces of 0.5 s to identify each event.

Data Labeling Process

The system should identify and classify the sound as a bite or chew. Any other detected sounds (e.g., mosquitoes and birds sound, sounds coming from other animals, airplane engine noise, etc.) should be identified by the neural network as a noise to avoid misprediction. The strategy adopted creates two trained models: (a) chew and noise; (b) bite and noise. The model (a) was trained to detect "bite events" and everything else was considered noise. Similarly, the model (b) detects only "chew events" and everything else considered as noise.

The first step was to label the entire video and count the number of chews and bites. The main 5 min audio file was then divided in 0.5 s audio files and each of these were labelled accordingly with the video.

Data Filtering Process

The first stage of the filtering process is transforming the data from a time series in a frequency domain data. To do that a Short-Time Fourier Transform (STFT) was used, defined in Equation 1, which transforms the audio to a normalized magnitude on the frequency domain (Fig. 4).

$$F(k) = \int f(x) e^{-2\pi i k x} dx \qquad (1)$$

Where f(x) is the signal to be transformed (x represents time), i is the complex representation and 2πk is the angular velocity representation that can be obtained by the frequency.

Transforming the data in a frequency domain helps the neural network identify the patterns of each event. After the STFT, the data is filtered. The first step of the filter process is converting the data to the Mel scale taking the log of Fourier spectrum magnitude, shown in Equation 2. The Mel scale tries to capture small differences between two sounds, even if they are with close frequency. So, when applied, this coefficient will allow the sounds to be correctly represented.

$$M(f) = 1125 \ln(1 + f/700) \qquad (2)$$

Where *f* is the respective frequency of the signal. A discrete Cosine Transform (DCT) is used over the Filter Bank Coefficients and STFT is applied to generate the Mel Frequency Cepstrum Coefficients. It will be a more consistent input for the first layers of the Recurrent Neural Network. Fig. 5 shows the process of filtering the data. Fig. 6 presents the overall filtering workflow

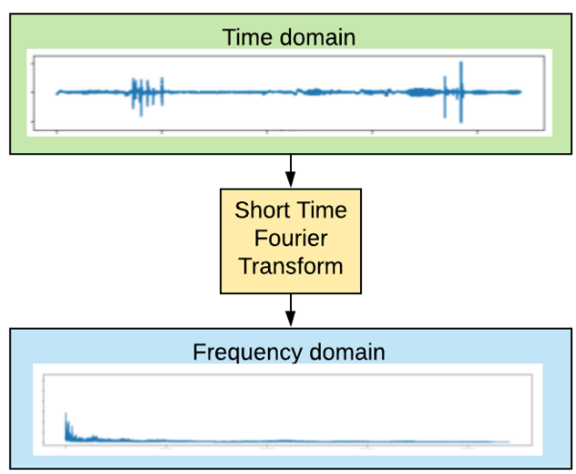


Figure 4. Short Time Fourier Transform applied to the audio file.

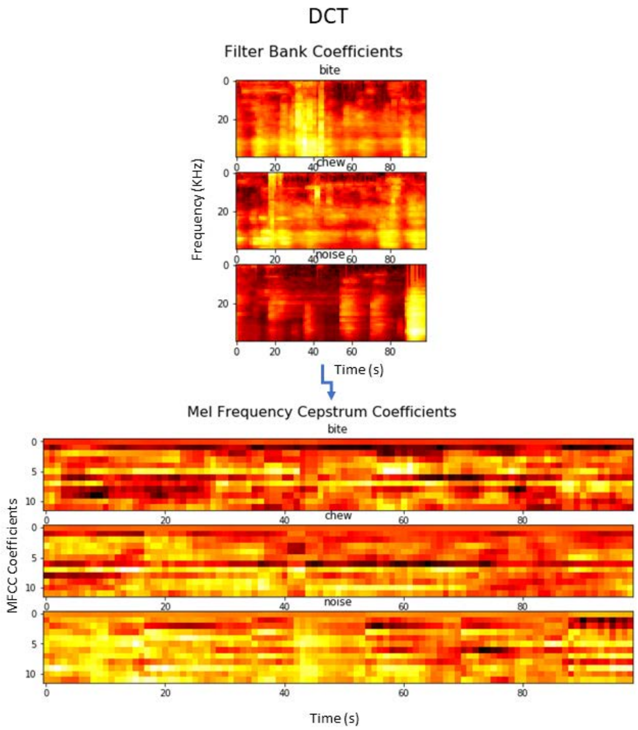


Figure 5. Filtering process of the audio data after the STFT.

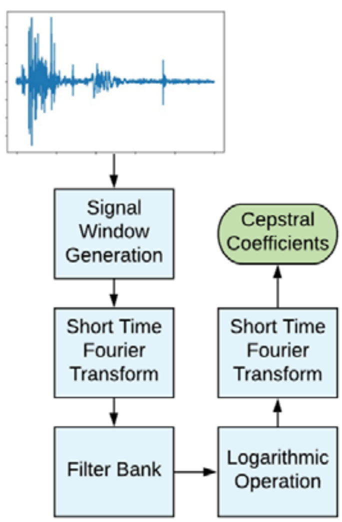


Figure 6. Filtering process workflow

Machine Learning Training Process

A Recurrent Neural Network (RNN) was utilized to generate a temporal sequence from the node's connection. It was designed in the Keras API with a Tensorflow backend and was used to train and generate an audio predictor model for chews and bites events. In this case, the neural network has a feedback inside the hidden layers (Fig. 7).

The mathematical form a recurrent neural network can be represented as:

$$ht = X2(UXt + Vh_{t-1}) \tag{3}$$

Where U and V are the weight matrices connecting the inputs and the recurrent outputs. This process creates an internal memory state which is added to the process input reducing small gradient multiplication. This concept, called Long short-term memory (LSTM), uses a "forget gate" to determine which state will remembered or forgotten.

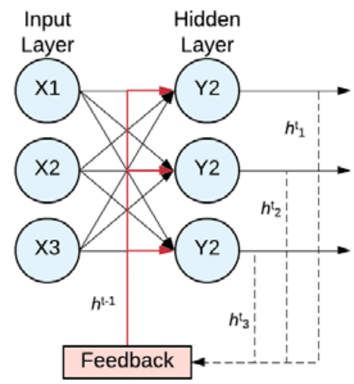


Figure 7. Representation of a Recurrent Neural Network with the feedback process shown.

The final network will be a combination of the LSTM feeding the hidden layers with a feedback (Figure 8). The RNN architecture used is shown in Figure 9.

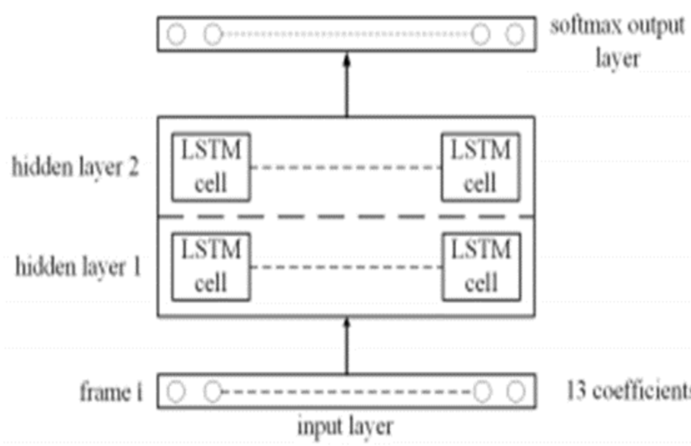


Figure 8. LSTM layers, Nahid et al (2017).

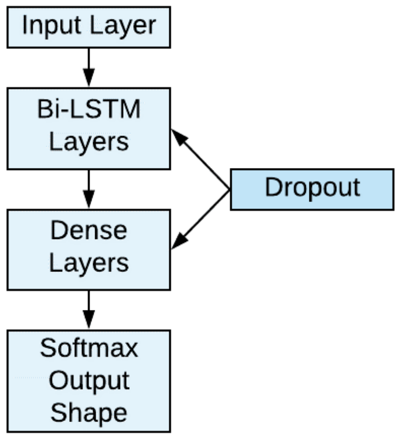


Figure 9. RNN architecture configuration.

Evaluation Metrics

To evaluate the performance of the proposed technique an experiment was proposed. The last 133 cut audio files were used to evaluate the models. Considering that there are both chews and bites in these files, the same files were used to test both models. Each of these audio files were labeled using the video recorded. The label is considered here the true positive event. To evaluate the bite model, everything besides bites were labeled as noise. To evaluate the chew model, everything besides chew were labeled as noise. Table 1 shows the event evaluation criteria.

Table 1. Events evaluation criteria.

Model	Error N °	Label	Model Predicts	Error Classification
Bite	1	bite	noise	False Negative (FN)
	2	noise	bite	False Positive (FP)
Chew	3	chew	noise	False Negative (FN)
	4	noise	chew	False Positive (FP)

True positive (TP) events are considered as the labeled events. Three evaluation metrics were used to evaluate the proposed technique: Precision (equation 4), Recall (equation 5), and F1 score (equation 6):

$$\boldsymbol{P} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP}+\boldsymbol{FP})} \tag{4}$$

$$\boldsymbol{R} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP}+\boldsymbol{FN}} \tag{5}$$

$$\boldsymbol{F1} = \boldsymbol{2}\,\frac{(\boldsymbol{P}*\boldsymbol{R})}{(\boldsymbol{P}+\boldsymbol{R})} \tag{6}$$

These three metrics were calculated for each model generated.

Results

Train and Test Curves

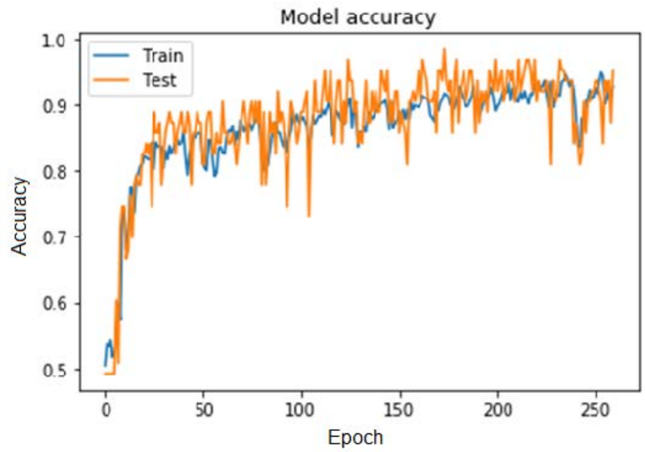
An accuracy plot was created during the training part to check the model convergence, shown in Fig. 10.

Horse bite/chew Detection Result Accuracy

Both models were tested, and the results metrics are presented in the Table 2:

Table 2. Results for the bite model and for the chew model

Model type	Precision	Recall	F1Score
Bite Model	80.0%	93.4%	87.8%
Chew Model	91.9%	94.4%	93.2%



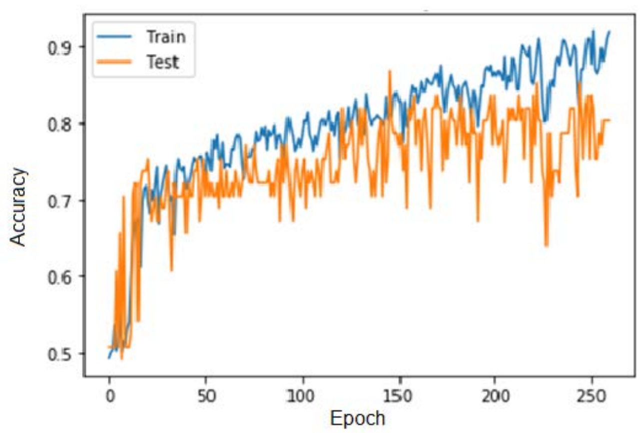


Figure 10. RNN Convergence curves for the: (a) bite identification model and (b) chew identification model.

Discussion

The convergence of test accuracy curves shows a flat region near 250 epochs. This shows that the number of epochs used was enough. The difference between the FNs and FPs for bite and chew model occurred mainly because the system confused some types of noise with the bite sound. Generally, this occurred when the horse hit some obstacle or when the halter started to move producing a hit sound that was sometimes confused with the bite sound. A solution for this problem is to find a better place for the camera, that allows avoid these noises and collect good audio data. Also, a better fixation of the halter can improve the audio data acquisition procedure.

The chew identification shown a high level of accuracy (93.2%), indicating that the tool proposed here can be effectively and quickly used to monitor horses foraging behavior. Compared to other identification systems, the neural network used here shown high accuracy for chew and bite identification. Werner et al (2019) presented a concordance correlation coefficient of 0.78 for bite identification and 0.94 for chew identification. The system proposed by Deniz et al (2017) presented an accuracy of 76.4%. Adriamandroso et al (2017) shown a 92% detection accuracy, with 95% accuracy for rumination detection and 91% accuracy for grass intake accuracy. Considering that further investigation will be performed, it is expected that this system will reach a reliable precision and consistency, being a powerful tool, even for other herbivores species, as cows, deer or sheep.

Conclusion

The presented method was able to identify chews and bites from a horse, with an accuracy of 87.8% for bite detection and 93.2% for chew detection. The chew model was better due to the problems with similarity between some noises and bites events. Even using filters, these noises kept similarities to the bite signals, and this led to some reduction in accuracy. Compared to other related works, the system described on this paper has shown great potential. The utilized Recurrent Neural Network accuracy can be improved by investigating the effect of several factors on detection accuracy, like the number of CNN layers, the number of neurons, and the dropout used between layers. Therefore, the use of Recurrent Neural Network to identify acoustic patterns in horses is a novelty and has a great potential to help horses stud owners and researchers to better study horses foraging behaviors.

Future Work

Several sensing systems can be utilized and evaluated to improve data acquisition quality. The best location to mount the wearable devise can be explored too. In the pre-processing step, several filters (e.g., high-pass and low-pass filters) should be investigated.

Acknowledgments

The authors thank Fabiane Rosa and Sonja Crawford for their support and assistance in conducting this study.

References

Andriamandroso, A. L. H., Lebeau, F., Beckers, Y., Froidmont, E., Dufrasne, I., Heinesch, B. & Bindelle, J. (2017). Development of an open-source algorithm based on inertial measurement units (IMU) of a smartphone to detect cattle grass intake and ruminating behaviors. *Computers and electronics in agriculture*, *139*, 126-137.

Alkon, P.U., Cohen, Y., and Jordan P.A. 1989. Towards an acoustic biotelemetry system for animal behavior studies. Journal of Wildlife Management

Ampatzidis, Y., Bellis, L.D., and Luvisi, A. 2017. iPathology: Robotic applications and management of plants and plant diseases. Sustainability 2017, 9, 1010.

Büchel, S., and Sundrum, A. 2014. Technical note: evaluation of a new system for measuring feeding behaviour of dairy cows. 2014. Comput. Electron. Agric. 108, 12–16.

Clapham, W. M., Abaye, A. O., Fedders, J. M., and Yarber, E. 2006. Sound spectral analysis of grazing steers. In Proceedings of the American Forage Grassland Conference (Vol. 15, pp. 139-143).

Deniz, N. N., Chelotti, J. O., Galli, J. R., Planisich, A. M., Larripa, M. J., Rufiner, H. L., & Giovanini, L. L. (2017). Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals. *Computers and electronics in agriculture*, *138*, 167-174.

DelCurto, T., Porath, M., Parsons, C. T., and Morrison, J. A. 2005. Management strategies for sustainable beef cattle grazing on forested rangelands in the Pacific Northwest. Rangeland Ecology & Management, 58(2), 119-127.

Dutta, R., Smith, D., Rawnsley, R., Bishop-Hurley, G., and Hills, J. 2014. Cattle behaviour classification using 3-axis collar sensor and multi-classifier pattern recognition. In SENSORS, 2014 IEEE (pp. 1272-1275). IEEE.

Gardner, J. W., Hines, E. L., Molinier, F., Bartlett, P. N., and Mottram, T. T. 1999. Prediction of health of dairy cattle from breath samples using neural network with parametric model of dynamic response of array of semiconducting gas sensors. IEE Proceedings-Science, Measurement and Technology, 146(2), 102-106.

Hansen, M. F., Smith, M. L., Smith, L. N., Salter, M. G., Baxter, E. M., Farish, M., and Grieve, B. 2018. Towards on-farm pig face recognition using convolutional neural networks. Computers in Industry, 98, 145-152.

Kaixuan, Z., and Dongjian, H. 2015. Recognition of individual dairy cattle based on convolutional neural networks. Transactions of the Chinese Society of Agricultural Engineering, 31(5).

Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A. 2017. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. IEEE Geosci. Remote Sens. Lett. 2017, 14, 778–782.

Laca, E.A., Ungar, E.D., and Demment M.W. 1994. Mechanisms of handling time and intake rate of a large mammalian grazer. Applied Animal Behaviour Science.

Laca, E. A., and de Vries M. F. W., 2000. Acoustic measurement of intake and grazing behaviour of cattle. Grass and Forage Science 55.2 (2000): 97-104.

LeCun, Y.; Bengio, Y.; Hinton, G. 2015. Deep learning. Nature 2015, 521, 436.

Logan, B. 2000. Mel Frequency Cepstral Coefficients for Music Modeling. In ISMIR (Vol. 270, pp. 1-11).

Md Mahadi Hasan Nahid, Bishwajit Purkaystha, Md Saiful Islam, Bengali speech recognition: A double layered LSTM-RNN approach, Conference: 20th International Conference of Computer and Information Technology (ICCIT), 22-24 December, 2017

At: Dhaka

Partel, V., Kakarla, S.C., Ampatzidis, Y. 2019. Development and Evaluation of a Low-Cost and Smart Technology for Precision Weed Management Utilizing Artificial Intelligence. Comput. Electron. Agric. 2019, 157, 339–350

Penning, P.D. 1983. A technique to record automatically some aspects of grazing and ruminating behaviour in sheep. Grass Forage Sci. 38, 89–96.

Penning, P.D., Rook, A.J. and Orr, R.J. 1991 Patterns of ingestive behaviour of sheep continuously stocked on monocultures of ryegrass or white clover. Applied Animal Behaviour Science.

Rutter, S.M., Champion, R.A., and Penning, P.D., 1997. An automatic system to record foraging behaviour in free-ranging ruminants. Appl. Anim. Behav. Sci. 54, 185–195.

Rutter, S. M. (2000). Graze: a program to analyze recordings of the jaw movements of ruminants. Behavior Research Methods, Instruments, & Computers, 32(1), 86-92.Ruuska, S., Kajava, S., Mughal, M., Zehner, N., and Mononena, J. 2016. Validation of a pressure sensor-based system for measuring eating, rumination and drinking behaviour of dairy cattle. Appl. Anim. Behav. Sci. 174, 19–23.

Ungar, E. D. 1996. Ingestive behavior. Pages 185–218 in The Ecology and Management of Grassland Systems. J. Hodgson and A. W. Illius, ed. CAB International, Wallingford, UK.

Ungar, E. D., Ravid, N., Zada, T., Ben-Moshe, E., Yonatan, R., Baram, H., and Genizi, A. 2006. The implications of compound chew–bite jaw movements for bite rate in grazing cattle. Applied animal behaviour science, 98(3-4), 183-195.

Ungar, E.D., and Rutter, S.M., 2006. Classifying cattle jaw movements: comparing IGER behaviour recorder and acoustic techniques. Appl. Anim. Behav. Sci. 98, 11–27.

de Vries, M. F. W., and Peter Schippers. Foraging in a landscape mosaic: selection for energy and minerals in free-ranging cattle. Oecologia 100.1-2 (1994): 107-117.

Werner, J., Leso, L., Umstatter, C., Niederhauser, J., Kennedy, E., Geoghegan, A., & O'Brien, B. (2018). Evaluation of the RumiWatchSystem for measuring grazing behaviour of cows. *Journal of neuroscience methods*, *300*, 138-146.

Zehner, N., Niederhauser, J.J., Nydegger, F., Grothmann, A., Keller, M., Hoch, M., Haeussermann, A., and Schick, M., 2012. Validation of a new health monitoring system (RumiWatch) for combined automatic measurement of rumination, feed intake, water intake and locomotion in dairy cows. In Proceedings of International Conference of Agricultural Engineering CIGR-AgEng2012, Valencia, Spain