# Measuring equine respiration in the field: an exploration of microphone data and deep learning detectors

J.I.M. Parmentier[1,2,3*], R.M. Aarts[2], F.M. Serra Bragança[2,4], B.J. van der Zwaag[1,3]

**[1]Pervasive Systems Group, Edge Research Centre, EEMCS, University of Twente, Enschede, The Netherlands,**
**[2]Clinical Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands,**
**[3]Inertia Technology B.V., Enschede, The Netherlands**
**[4]Sleip AI, Stockholm, Sweden**
***Corresponding author: j.i.m.parmentier@utwente.nl**

## Introduction

An important physiological aspect in exercising horses is respiration and its associated parameters (e.g., breathing pattern, respiratory rate, and locomotor-respiratory coupling). Quantifying variations during training in these parameters has a huge potential on understanding mechanisms linked to equine performance [1], but also to potentially detect respiratory anomalies, and ultimately improve equine welfare. Related to this, several studies have explored using microphones to study respiratory sounds showing its potential to detect upper-airway disorders [2]–[4]. This supports the use and exploration of microphone data to monitor equine respiration during training.

However, segmentation of respiratory cycles was limited to manually counting respiratory sounds and most studies were conducted on treadmills, where environmental noises are not representative of training conditions and using a sulky is impossible. Due to increasing available computational power and new data processing techniques (e.g., deep learning), there are opportunities to develop near-real-time quantifications of the respiratory patterns and respiratory rates of (harness) horses during training.

In this study, we explore deep learning models to automatically detect exhalation events in microphone data obtained on exercising harness trotters. We also evaluate the effect of downsampling the microphone data on the performances of our exhalation events detectors.

## Materials and Methods

Fifteen warmblood harness trotters in full training were equipped with an omni-directional microphone (ECM-LV1, Sony, 44100 Hz) taped between the nostrils and a voice recorder (ICD-Px470, Sony) attached to the bridle. All horses underwent a standardized exercise test for harness trotters on an oval track, with increasing trotting speeds. The last segment of high-speed trot (37.9±1.5km/h on average) was used for this study as respiration sounds were more distinguishable than at lower speeds. High-speed trot is also the condition under which the respiratory system of the trotters is under pressure and during which respiratory sounds are the loudest.

As exhalation sounds were generally louder and more distinct than inhalation sounds, especially in presence of environmental noise (tack, hoof beats, wind), exhalation events were chosen for this study. Exhalation events were labelled by one user in Audacity 3.4.2. [5] through sound and visually inspecting both microphone channels and their Mel spectrogram and labelled microphone data were then imported in MATLAB R2022b (MathWorks, Natick, Massachusetts, USA) for training the detectors.

Temporal Convolutional Network (TCN) detectors [6] of different depths (4, 8 and 12 dilation blocks) were trained to classify sound signals into exhalation (1) and no exhalation (0) in a sequence-to-sequence classification fashion. The detectors were trained with either one of the two channels or both channels together, with different downsampling factors (1, 2, 5, 10 and 40, thus sampling frequencies of 44100 Hz, 22050 Hz, 8820 Hz, 4410 Hz and 1102.5 Hz respectively). Each model was trained with the data of thirteen horses, with one horse withheld for validation (early stopping with validation patience of five epochs) and another one for test (leave-one-out performance evaluation). The F1-scores were then computed to compare the performance of the different TCNs. Detectors with higher median F1-score with lower interquartile range (IQR) were considered performing better.

## Results

When training with shallower TCN detectors, the best performance was obtained with lower sampling frequencies (see Table 1). However, at higher sampling frequencies, most false positives were loud sound events detected as exhalations. Increasing the depth of the TCNs decreased the sensitivity to the sampling frequencies, as shown in Table 2. This is explained by the learning mechanisms lying behind the TCN architecture: deeper TCN learn from a wider receptive field in the training data. For a receptive field of ten samples, the duration of the observed events is of 0.0002 seconds at 44100Hz against 0.001 seconds at 8820Hz. Overall, better results were obtained when training with Channel 2 or Both Channels compared to Channel 1. This was regardless of the downsampling factor (see Table 2). Exhalation events were more recognisable in Channel 2 compared to Channel 1 during the labelling process, explaining the different results and further showing the importance of correct labelling in deep learning methodologies.

Table 3, Median and interquartile ranges (IQR) F1-scores for 4 blocks TCN detectors trained with either Channel 1, Channel 2 or Both Channels with high sampling frequencies (22050Hz and higher) or low sampling frequencies (8820Hz and lower).

| Sampling frequency (TCN 4 Blocks) | Channel 1 | Channel 2 | Both Channels |
|---|---|---|---|
| High | 0.72 IQR [0.66-0.77] | 0.77 IQR [0.73-0.80] | 0.79 IQR [0.66-0.82] |
| Low | 0.81 IQR [0.73-0.85] | 0.86 IQR [0.84-0.90] | 0.86 IQR [0.85-0.90] |

Table 4, Median and interquartile ranges (IQR) F1-scores for TCN detectors trained with either Channel 1, Channel 2 or Both Channels with different depths, for all sampling frequencies.

| TCN depth (All sampling frequencies) | Channel 1 | Channel 2 | Both Channels |
|---|---|---|---|
| 4 blocks | 0.77 IQR [0.71-0.82] | 0.84 IQR [0.77-0.87] | 0.84 IQR [0.80-0.88] |
| 8 blocks | 0.84 IQR [0.79-0.89] | 0.89 IQR [0.85-0.93] | 0.90 IQR [0.86-0.92] |
| 12 blocks | 0.86 IQR [0.78 -0.90] | 0.90 IQR [0.87-0.93] | 0.91 IQR [0.88-0.92] |

## Conclusion

This work shows that by using a microphone simply taped to the nose of horses it is possible to automatically detect exhalation events at high-speed trot with deep learning models, despite the environmental noise. We also showed that decreasing the sampling frequency improves the performance of the detectors, especially for less complex models which are less prone to overfitting. Using only one channel of data seems to be sufficient for the automatic detection of equine exhalation events in microphone data, but we would like to be cautious with the application of this result as some measured sound events might be buried in environmental noise. This would be depending on where and how the microphone is placed on the horse's nose. Further comparisons of microphone placements need to be conducted to conclude on the most robust and reliable channel(s) to use for this purpose. In the future, we will explore the performance of our detectors at lower trotting speeds, as well as evaluating the detectors' outputs to compute physiological information like respiratory rate and compare our results to a simpler signal processing method published in the human literature [7].

## Ethical Statement

All animal experiments in this work were approved by the Swedish Ethics Committee (diary number 5.8.18-04197/2022) in agreement with Swedish legislation on animal experiments.


## Acknowledgements

The authors would like to thank Zala Zgank, Ebba Zetterberg and Guilherme de Camargo Ferraz for their help during the data collection, as well as the Menhammar Stuteri drivers and grooms. They would also like to thank Elin Hernlund, Marie Rhodin, Johan Hellander, Mihai Marin-Perianu and Raluca Marin-Perianu for their input in the project. This study was partially funded by the research project "Varenne" E! 114697.


## References


1. C. Cotrel, C. Leleu, and A. Courouce-Malblanc, 'Factors influencing variation in locomotor-respiratory coupling in Standardbred Trotters in the field', *Equine Vet. J.*, vol. 38, no. S36, pp. 562–566, 2006, doi: 10.1111/j.2042-3306.2006.tb05605.x.
2. G. R. G. Barnes, M. Brennan, B. E. Goulden, and J. Kirkland, 'Sound spectography in the diagnosis of equine respiratory disorders: a preliminary report', *N. Z. Vet. J.*, vol. 27, no. 7, pp. 145–146, Jul. 1979, doi: 10.1080/00480169.1979.34629.
3. F. J. Derksen, S. J. Holcombe, W. Hartmann, N. E. Robinson, and J. A. Stick, 'Spectrum analysis of respiratory sounds in exercising horses with experimentally induced laryngeal hemiplegia or dorsal displacement of the soft palate', *Am. J. Vet. Res.*, vol. 62, no. 5, pp. 659–664, May 2001, doi: 10.2460/ajvr.2001.62.659.
4. S. H. Franklin, S. G. Usmar, J. G. Lane, J. Shuttleworth, and J. F. Burn, 'Spectral analysis of respiratory noise in horses with upper airway disorders', *Equine Vet. J.*, vol. 35, no. 3, pp. 264–268, May 2003, doi: 10.2746/042516403776148228.
5. 'Audacity ® | Free Audio editor, recorder, music making and more!' Accessed: Dec. 31, 2023. [Online]. Available: https://www.audacityteam.org/
6. S. Bai, J. Z. Kolter, and V. Koltun, 'An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling'. arXiv, Apr. 19, 2018. doi: 10.48550/arXiv.1803.01271.
7. C. Romano *et al.*, 'Respiratory Rate Estimation during Walking and Running Using Breathing Sounds Recorded with a Microphone', *Biosensors*, vol. 13, no. 6, Art. no. 6, Jun. 2023, doi: 10.3390/bios13060637.