Denoising and decoding spontaneous vagus nerve recordings with machine learning

Mafalda Ribeiro[1,2], Ryan G. L. Koh[3], Tom Donnelly[1,2], Christof Lutteroth[1,2], Michael J. Proulx[1,2,4], Paulo R. F. Rocha[5], Benjamin Metcalfe[1,2]

Abstract— **Neural interfaces that electrically stimulate the peripheral nervous system have been shown to successfully improve symptom management for several conditions, such as epilepsy and depression. A crucial part for closing the loop and improving the efficacy of implantable neuromodulation devices is the efficient extraction of meaningful information from nerve recordings, which can have a low Signal-to-Noise ratio (SNR) and non-stationary noise. In recent years, machine learning (ML) models have shown outstanding performance in regression and classification problems, but it is often unclear how to translate and assess these for novel tasks in biomedical engineering. This paper aims to adapt existing ML algorithms to carry out unsupervised denoising of neural recordings instead. This is achieved by applying bandpass filtering and two novel ML algorithms to in-vivo spontaneous, low-SNR vagus nerve recordings. The performance of each approach is compared using the task of extracting respiratory afferent activity and validated using cross-correlation, MSE, and accuracy in terms of extracting the true respiratory rate. A variational autoencoder (VAE) model in particular produces results that show better correlation with respiratory activity compared to bandpass filtering, highlighting that these models have the potential to preserve relevant features in complex neural recordings.**

I. Introduction

Neural implants that electrically stimulate the central or peripheral nervous systems have been an active topic in neural engineering research and in industry, given the possibility to treat neurological conditions in a more localised way [1]. There are commercially-available systems capable of stimulating the nervous system for the treatment of epilepsy and depression (i.e. brain, vagus nerve), for bladder control (i.e. sacral roots), and other applications [2], [3]. However, given the variability between different patients, and even within the same patient under different conditions, initial stimulation parameters may not always deliver optimal treatment, requiring tuning by a clinician. To address this, various approaches and technologies for conducting electrical recordings of the brain and peripheral nerves are being proposed in order to close the loop.

Cuff electrodes have been a popular choice for electrical stimulation given that they surround peripheral nerves without penetrating the epineurium, whilst still being able to selectively recruit specific fibre types. However, given their positioning on the nerve, additional sources of non-stationary noise and interference, such as biological and instrumentation noise will also be recorded. Therefore, spontaneous electroneurogram (ENG) signals recorded using cuff electrodes typically have a very low signal-to-noise ratio (SNR), rendering them difficult to filter and process.

Currently, the most common approach for denoising neural recordings is through the use of analogue or digital bandpass filters centered around frequencies of action potentials (APs), which typically range from $100\,\mathrm{Hz}$-$10\,\mathrm{kHz}$ [4]. Wavelet-based denoising has also been proposed, yet this requires the fine-tuning once again of parameters such as the mother wavelet, the decomposition level, the threshold definition, among others [5]. In other domains, such as biomedical image denoising and segmentation, machine learning (ML) techniques have started to be proposed and deployed successfully. Of particular interest for denoising neural recordings are models which do not require exact ground-truth labels, given the difficulty in ascertaining these for less-invasive recordings, such as those using cuff electrodes.

Bearing these developments in mind, this work details the implementation and comparison of two ML algorithms for denoising problems. The first was a variational autoencoder (VAE) model inspired by the Coordinate-VAE proposed by [6], which is a recent unsupervised model specifically aimed at cuff recordings. The second model examined was the Noise2Noise model proposed by [7], which is a supervised regression model that works on noisy samples only, previously applied to image denoising problems. This model can be modified for the task of denoising neural recordings by changing 2D convolutions for processing noisy images into 1D convolutions for processing noisy signal sequences. Both of the proposed ML models were compared to conventional bandpass filtering, and applied in a spontaneous cuff recording task.

In this work, respiratory activity is also derived from blood pressure measurements and used to validate denoising performance for the first time. In comparison, previous ML-based denoising approaches have relied on human labels for specific types of neural activity.

II. Methods

A. Experimental data collection

The procedures for collecting single animal recordings for this work were performed in accordance with the Danish An-

[1] Centre for Accountable, Responsible, and Transparent AI (ART-AI), Department of Computer Science, University of Bath, UK
[2] University of Bath, UK
[3] KITE Research Institute, Toronto Rehabilitation Institute – University Health Network, Toronto, Canada
[4] Reality Labs Research, Meta, Redmond, Washington, USA
[5] Centre for Functional Ecology-Science for People & the Planet, Associated Laboratory TERRA, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

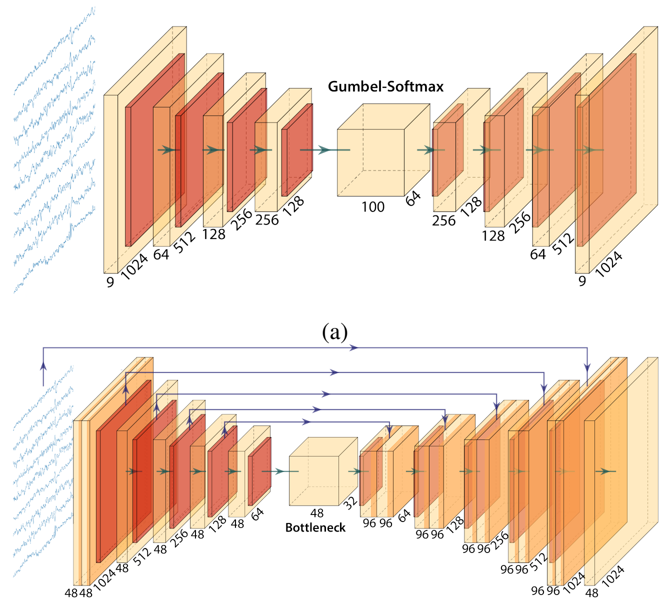
imal Experiments Inspectorate (approval no. 2013-15-2934-00753), as well as the care and use of laboratory animals as described by the U.S National Institutes of Health, and are described in full in [8]. Spontaneous neural activity was obtained from cuff recordings from pig right cervical vagus nerve, using a cuff with ten platinum-iridium electrodes [9]. The electrodes were connected as nine bipolar channels, and a bandpass filter from $100\,\mathrm{Hz}$-$10\,\mathrm{kHz}$ was applied to the data. A sampling rate of $100\,\mathrm{kHz}$ was chosen given the small propagation delays in faster APs, which were as low as $175\,\mu\mathrm{s}$. In total, the full dataset of spontaneous neural activity consisted of 2 minutes of continuous recording. Arterial blood pressure was also monitored.

B. Denoising models

Convolutional neural networks (CNNs) are useful for processing 1D or 2D sequences of data, and improve an ML system through sparse interactions, parameter sharing, and equivariant representations [10]. This implies fewer parameters are required, only one set of parameters needs to be learnt for each location on the input, and changes at the input modify the representation in the same way. In terms of denoising, encoder/decoder CNNs leverage smaller latent spaces to encode inputs and sometimes allow "skip" connections to allow some leakage of information from encoding to decoding layers (typically referred to as U-Nets), both of which can benefit the model in learning the required task [11].

The Coordinate VAE model proposed by [6] is aimed specifically at denoising PNS data collected from cuff electrodes. As shown in Figure 1(a), VAE models comprise of encoder and decoder blocks, with the main objective of minimising the error between the original input and reconstructed output. However, unlike conventional autoencoders, VAEs encode the input as a distribution over the latent space. In the Coordinate-VAE model applied specifically to cuff recordings, this is attained by using small windows of data ($\sim 8.5\,\mathrm{ms}$) as the input to the encoder, which then produces a corresponding one-hot vector in the latent space using a Gumbel-Softmax activation [12]. The loss was defined as a weighted sum between the mean squared error (MSE) between the original input and reconstruction, and the negative of the Kullback-Leibler (KL) divergence. An additional block added to this model consisted of a coordinate encoder, which leveraged information about time samples near the baseline to minimise the size of the encoded one-hot vector.

The Noise2Noise model proposed by [7] has been predominantly applied to various imaging tasks such as denoising or undersampling. This model follows a U-Net structure as shown in Figure 1(b), which contains contracting and expanding paths for downsampling and subsequently reconstructing the input. The upsampling layers get concatenated with outputs from prior contracting layers to relay context information to higher resolution layers [11]. A key contribution from the Noise2Noise model was the ability to denoise input data using corrupted targets only. This is achieved by using pairs of the same image, corrupted by the same type of



(a)

(b)

Fig. 1: ML model architectures, (a) Variational Autoencoder (VAE) and (b) Noise2Noise model

noise at different locations. The authors further highlight that the optimal network parameters remain unchanged if input-conditioned target distributions $p(y|x)$ are replaced with arbitrary distributions with the same conditional expected values [7]. In turn, this implies networks can be trained with targets containing zero-mean noise without changing the learnt reconstruction.

C. Data processing

A fourth-order bandpass filter with cutoff frequencies $250\,\mathrm{Hz}$ and $10\,\mathrm{kHz}$ was applied to the raw recording data, with the frequency range corresponding to a broad interval where neural activity is expected to lie. The data were then split into training, validation, and test sets (70%, 10%, 20%). Furthermore, the raw vagus ENG recordings contained artifacts occurring periodically, likely owing to interference from neighbouring equipment. As such, any points above two standard deviations of the data were replaced with the channel median. The data were then standardised and rescaled to the range of [-1, 1] as described in [6]. Non-overlapping windows of $10.24\,\mathrm{ms}$ (1024 samples) from all nine bipolar channels were extracted both from the raw and bandpass-filtered data to be used as inputs to the ML models. Data augmentation was also applied by extracting windows from the training data at two different offsets.

D. Algorithm training and evaluation

The implemented VAE model contained CNN-based encoder and decoder blocks, and a Gumbel-Softmax activation in the latent space [12]. For this study, the coordinate encoder described in the previously mentioned Coordinate VAE model [6] was omitted since plausible reconstructions could not be obtained with any parameter combinations. To

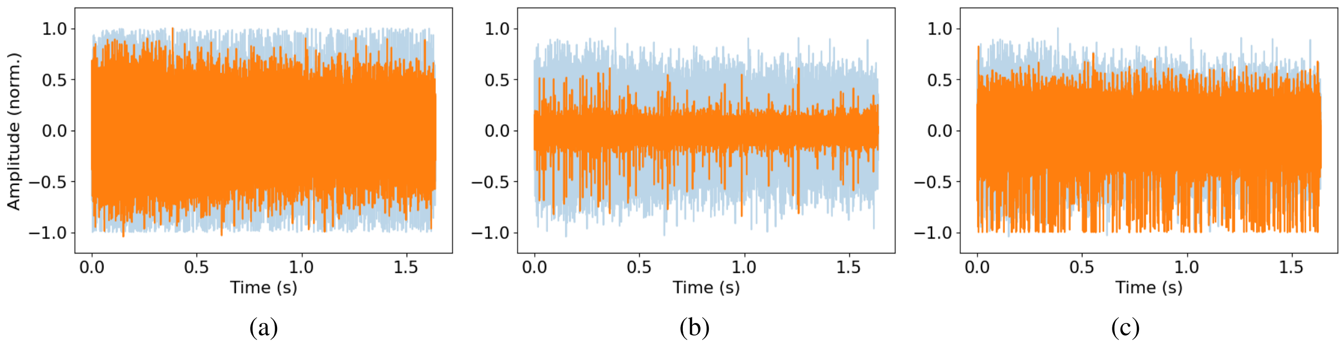


Fig. 2: Comparison of bandpass filtering and two ML denoising approaches. (a) Window of original vagus recording before (faint blue) and after bandpass filtering with cutoff frequencies 250-10 kHz (orange), (b) Unsupervised VAE reconstruction (orange) compared with bandpass filter (faint blue), (c) Noise2Noise model reconstruction (orange) compared with bandpass filter (faint blue).

compensate for this and the larger input size, the latent space was made larger with 50 filters. For this network, the loss was defined as the weighted sum of the MSE between the input and output waveforms (reconstruction error), the KL divergence, and the MSE between the blood pressure respiratory envelope and the moving root mean square (RMS) plot of the ENG data. For these experiments the weights for each term were set to 65%, 20%, and 15%, respectively.

The Noise2Noise model followed the same structure described in its original publication [7], but it was adapted for performing 1D convolutions instead. The loss function was also updated to include an MSE term between the respiratory envelope and moving RMS ENG data, weighed at 50% of the total loss. Finally, given the relatively small dataset being used, the sizes of the encoding and decoding layers were decreased and a lower learning rate was used. In both the VAE and Noise2Noise model, kernel sizes of 5 and 13 were used respectively, and the AdamW optimiser was used with a learning rate of 0.005 [13].

The VAE model used bandpass filtered data as input, whereas the Noise2Noise model used both the raw and bandpass filtered datasets. Finally, to quantify the complexity of these models, the parameter counts and training times are reported in the results section. Training was carried out on a machine with an NVIDIA RTX 3090.

E. Respiration activity task

Respiration activity can be extracted from the envelope of a blood pressure waveform. This relationship between blood pressure and respiration has been previously quantified as being proportional to changes in intrathoracic pressure during inspiration and expiration [14]. For this work, blood pressure windows recorded alongside ENG data in the training and test sets were obtained, and the positive envelope was extracted. During training, consecutive blood pressure windows were combined for each batch and compared against a moving RMS window of 1 s from each denoised ENG recording. This approach was also applied to windows from an unshuffled test set, which were similarly combined into a longer recording and compared against a moving RMS window of corresponding ENG recordings. In both cases, the envelope and moving RMS data were normalised using z-score normalisation. The MSE was calculated for training purposes, and both cross-correlation at zero lag (Pearson correlation coefficient) and MSE were used for validating test results. Given the animal was mechanically ventilated at a rate of 0.25 Hz, the respiratory rate was also extracted from the moving RMS recordings and compared.

III. Results

Figures 2(a)-(c) highlight the differences in the different denoising methods in the time domain, with examples being drawn from a hold-out, unshuffled test set. Figure 3 also shows a comparison the respiratory envelope extracted from blood pressure and moving RMS plots of the denoised ENG signals. In the time domain, it can be seen that the VAE and Noise2Noise models produce significantly different results in terms of the activity that is preserved. Particularly, the VAE removes any baseline wander and preserves less of the original signal content compared to the Noise2Noise model. From Figure 3, it can also be seen that both ML models track the respiratory envelope differently to the bandpass filter. Three performance metrics were also quantified for each approach, namely the cross-correlation and MSE between the respiratory envelope and denoised ENG signals, and the accuracy in extracting the mechanical ventilation rate of 0.25 Hz. Bearing these metrics in mind, the VAE shows performance improvements in comparison to conventional bandpass filtering in terms of cross-correlation and MSE. In particular, it is able to reconstruct ENG signals whilst preserving detail also seen in respiratory data from blood pressure. The Noise2Noise model only outperformed bandpass filtering and the VAE model in extracting the underlying respiratory rate of the recording.

IV. Discussion

In this work, two recent ML models were implemented, trained, and tested on the task of denoising very low-SNR recordings of pig vagus ENG data. Alongside this data, blood pressure recordings were also collected, which provided a separate physiological signal to compare with the denoising results. From the plots in Figures 2 and 3, and Table I, it can

Model	Cross-correlation with BP (Mean (SD))	Respiratory rate error (%) (Mean (SD))	MSE (Mean (SD))	Parameter count	Training time (min)
Bandpass filtering	0.813 (0.00911)	11.4 (0.231)	0.371 (0.0183)	-	-
VAE	**0.825** (0.00600)	12.2 (0.211)	**0.351** (0.0120)	470,957	82
Noise2Noise	0.788 (0.0160)	**11.0** (0.293)	0.422 (0.0315)	71,625	20

TABLE I: Quantification of performance and complexity for each model, with the results presented as the mean and standard deviation across all nine channels of data. The latter two categories were omitted for bandpass filtering given it is a non-ML method that does not require prior training.

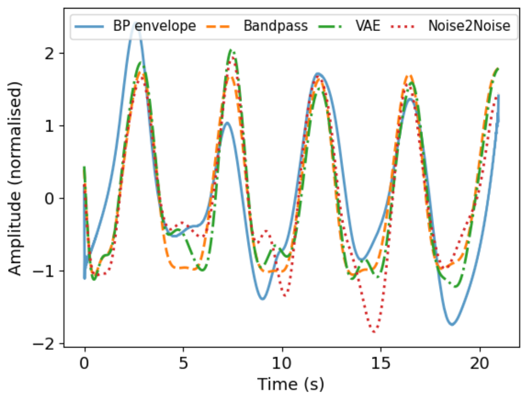


Fig. 3: Blood pressure recording and average of all channels of moving RMS plots for the test set signals using a moving window size of 1s. Blood pressure recording is in blue, bandpass filtering with cutoff frequencies 250-10 kHz in orange, VAE model results in green, and Noise2Noise model results in red

be seen that the VAE in particular poses improvements over conventional bandpass filtering in terms of cross-correlation and MSE between the moving RMS of ENG data and blood pressure data. More specifically, the ability to incorporate respiratory activity as part of the loss function provides a potential advantage to ML models, which are then able to denoise whilst preserving relevant features for the task at hand; in this case, extracting respiratory activity from vagus nerve recordings. It is hypothesised that ML models such as those described in this work could be further improved with larger datasets with more diverse and realistic data, since the current dataset was limited to a few minutes of activity from a single animal at a constant respiratory rate of 0.25 Hz.

The primary metric used to quantify the performance for the task was the cross-correlation between the respiration envelope and moving RMS plots for each denoising method. It could also be worthwhile investigating the resulting velocity spectra from the denoised recordings for this task [8]. This approach would offer a more fine-grained view of whether particular conduction velocities, or afferent and efferent activity, can be more easily extracted with differing denoising methods.

V. Conclusions

Although ML models have been used successfully in a number of domains, it can be unclear how to translate and assess these for novel problems in biomedical engineering. This work describes the process of selecting, adapting, and analysing two candidate ML models to denoise low-SNR neural signals. In particular, the models take into account additional physiological signals from an organ innervated by the nerve in question, in this case the lungs and the vagus nerve. The results from the denoising task were examined in the time-domain, as well as validated against multiple metrics to assess the performance of each denoising approach. In short, the VAE model showed performance improvements for the task of respiration activity extraction in comparison to a bandpass filter. In future, it would be worthwhile to investigate whether a more complex unsupervised model could be trained on larger volumes of more diverse data. From the perspective of extracting information from ENG recordings, it would also be worthwhile comparing the performance of different denoising approaches in the velocity domain.

References

[1] C. E. Larson *et al.*, "A review for the peripheral nerve interface designer," *Journal of Neuroscience Methods*, vol. 332, p. 108523, 2020. [Online]. Available: https://doi.org/10.1016/j.jneumeth.2019.108523
[2] J. Yang *et al.*, "The present and future of vagus nerve stimulation," *Journal of Korean Neurosurgical Society*, vol. 62, pp. 344–352, 2019.
[3] T. Guiho *et al.*, "Sacral anterior root stimulation and visceral function outcomes in spinal cord injury–a systematic review of the literature over four decades," pp. 218–232.e14, 1 2022.
[4] K. Yoshida, "Peripheral nerve signal processing, denoising," pp. 1–5, 2013. [Online]. Available: https://doi.org/10.1007/978-1-4614-7320-6_215-1
[5] G. Baldazzi *et al.*, "Systematic analysis of wavelet denoising methods for neural signal processing," *Journal of Neural Engineering*, vol. 17, 12 2020.
[6] T. J. Hardcastle *et al.*, "Coordinate-vae: Unsupervised clustering and de-noising of peripheral nervous system data," 2019.
[7] J. Lehtinen *et al.*, "Noise2noise: Learning image restoration without clean data," 3 2018. [Online]. Available: http://arxiv.org/abs/1803.04189
[8] B. W. Metcalfe *et al.*, "First demonstration of velocity selective recording from the pig vagus using a nerve cuff shows respiration afferents," *Biomedical Engineering Letters*, vol. 8, pp. 127–136, 2018.
[9] M. Haugland, "Flexible method for fabrication of nerve cuff electrodes," *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp. 359–360, 1996.
[10] I. J. Goodfellow *et al.*, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, http://www.deeplearningbook.org.
[11] O. Ronneberger *et al.*, "U-net: Convolutional networks for biomedical image segmentation," 5 2015. [Online]. Available: http://arxiv.org/abs/1505.04597
[12] E. Jang *et al.*, "Categorical reparameterization with gumbel-softmax," 11 2016. [Online]. Available: http://arxiv.org/abs/1611.01144
[13] D. P. Kingma *et al.*, "Adam: A method for stochastic optimization," 12 2015. [Online]. Available: http://arxiv.org/abs/1412.6980
[14] G. Lorenzi-Filho *et al.*, "Entrainment of blood pressure and heart rate oscillations by periodic breathing," *Am J Respir Crit Care Med*, vol. 159, pp. 1147–1154, 1999. [Online]. Available: www.atsjournals.org