

A Convolutional Neural Network for Classification of Nerve Activity based on Action Potential Induced Neurochemical Signatures

Paul Roever*, Khalid B. Mirza†, Konstantin Nikolic†, Christofer Toumazou†

*Department of Bioengineering, Imperial College London

† Centre for Bio-Inspired Technology, Imperial College London, SW7 2BT, UK.

Email: {p.roever15,k.mirza,k.nikolic,c.toumazou}@imperial.ac.uk

Abstract—Neural activity results in chemical changes in the extracellular environment such as variation in pH or potassium/sodium ion concentration. Higher signal to noise ratio make neurochemical signals an interesting biomarker for closed-loop neuromodulation systems. For such applications, it is important to reliably classify pH signatures to control stimulation timing and possibly dosage. For example, the activity of the subdiaphragmatic vagus nerve (sVN) branch can be monitored by measuring extracellular neural pH. More importantly, gut hormone *cholecystokinin* (CCK)-specific activity on the sVN can be used for controllably activating sVN, in order to mimic the gut-brain neural response to food intake. In this paper, we present a convolutional neural network (CNN) based classification system to identify CCK-specific neurochemical changes on the sVN, from non-linear background activity. Here we present a novel feature engineering approach which enables, after training, a high accuracy classification of neurochemical signals using CNN.

I. INTRODUCTION

Closed-loop neuromodulation systems are the next generation frontier in neuromodulatory devices. The benefits include improved therapeutic, clinical and engineering outcomes such as improved amelioration of symptoms, possibly better response to neuroplasticity and adaption, improved compliance, lower power consumption and longer battery lifetimes in implant [1]. Furthermore, closed-loop neuromodulation offers the opportunity to develop a more personalised approach towards neuromodulation therapies. It has also been shown to offer better therapeutic outcome [1], [2]. However, a significant step towards achieving an effective closed-loop system is selection of an appropriate symptomatic biomarker [3].

Neurochemical signals can be further classified into those originating from ions such as potassium (K^+), sodium (Na^+) and those originating from neurotransmitters such as dopamine serotonin etc. The former is directly linked to electrical neural activity and can be observed in Central Nervous System (CNS) and Peripheral Nervous System (PNS). The latter is generally observed in CNS, more specifically in the presence of cell bodies where inter-neuron response transmission occurs. More research has been performed in observing neurotransmitters compared to ionic responses with the primary reason being difficulty in fabricating implantable, penetrating, ion-selective microelectrodes until recently [3].

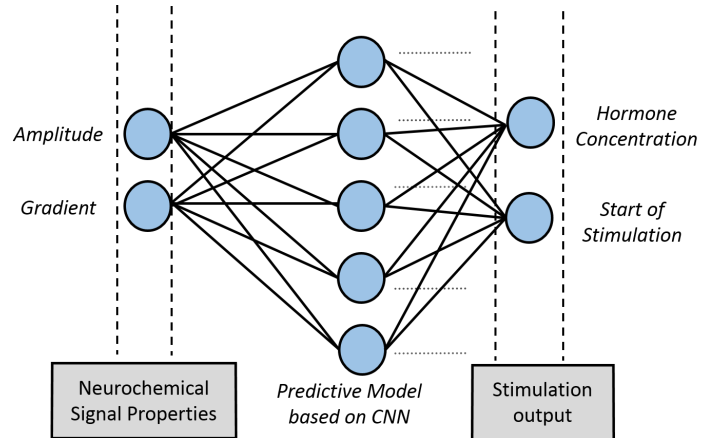


Fig. 1. A neural network can be used to accurately detect neurochemical response and predict target biomarker concentration followed by indication on whether to start stimulation or not.

While performing neurochemical recording *in vivo*, both linear and non-linear background signals have been observed with various background subtraction techniques being used previously [4], [5]. The next step after background subtraction is response classification. This is a binary classification stage where the response is determined to be originating from the target physiological biomarker, in this case CCK, or not. Based on the outcome of this classification, the next level consists of multi-level classification where the concentration of hormone is determined based on response characteristics.

Previous papers have explored modelling the relationship between neurochemical levels and stimulation parameters [6]. In this paper, a neural network based classification system is described while considering various trade-offs and data preparation techniques. After drift removal, the pre-processed signal is labelled and used to train a neural network. The network is then optimised and used to perform classification. The precision and recall results for this network are presented. This paper focusses on primarily methodology for identifying and enhancing features for noisy neurochemical signals.

In this paper, Section II provides a brief introduction into existing closed-loop systems, and the functional units of an

implant processing unit.

II. BACKGROUND

A. Neurochemical Changes in the Vagus Nerve

The vagus nerve innervates the subdiaphragmatic region on the ventral and dorsal gut regions. It links the gut to the brain and plays a crucial role in appetite control through interaction of several gut hormones [7]. It is well-established that cholecystokinin (CCK), secreted in the intestines, plays a crucial role in anorectic effects through the vagus nerve [8], [9]. In the experiments described in this paper, CCK is used to activate signalling in the vagus nerve by intravenous injection. This results in activation of the vagus nerve resulting in changes in neurochemical pH.

B. Closed Loop Systems

A closed-loop system consists of sensors/electrodes followed by data acquisition, pre-processing, possible dimensionality reduction followed by decision and stimulation dose tuning unit. The ultimate goal of a closed-loop system is to accurately and autonomously answer *when?* and *how much?* to stimulate. Closed-loop systems can be divided into different types, depending on functionality and type of control feedback. These are *responsive*: a partial closed-loop system stimulation a partial closed-loop system where one of the two questions (*when?*, *how much?*) is answered on the basis of a single physiological biomarker and a predetermined threshold of this biomarker, *adaptive*: where a single biomarker is monitored but both *when?* and *how much?* are answered, *complete closed-loop*: where multiple biomarkers are monitored and considered to deliver closed-loop functionality [3], [10].

In order to implement autonomous detection, algorithms need to be trained using an experimental dataset [11], [5], [3]. The training step is performed on *in vivo* or *in vitro* data. A variety of different algorithms have been used previously [11], [5], [3]. These include machine learning (ML) algorithms such as principal component regression (PCR), which is a combination of principal component analysis and linear regression. Artificial Neural networks (ANN) have been investigated previously for neurotransmitter release due to electrical stimulation [11], [6]. ANN were used to model both the relationship between stimulation-evoked dopamine release and vice-versa [6].

C. CNN vs RNN

For timeseries classification such as the one described in this paper, has been shown to be more suitable for Convolutional Neural Network (CNN) as compared to Recursive Neural Network (RNN). It is dependant on a number of factors such as size of pre-processing window.

Furthermore, CNNs train about 20 times faster than RNN. On a batch size 64, RNNs take on an average of 1 ms while CNNs take an average of 200 μ s. A hybrid of CNN and RNN was also tried. Even though a small improvement in classification was observed, it was slower to test and train. Hence, RNNs were not considered an optimal solution.

D. Processing Unit

A closed-loop implant comprises of a processing unit which determines *when* and *how much* to perform neural stimulation. The processing unit consists of a pre-processing unit which helps in removing data, followed by data labelling, dimensionality reduction and classification. This paper primarily focusses on classification method based on low latency and relatively low resource deep neural network (DNN).

III. METHODOLOGY

This section describes the methodology for training and validating a CNN for classification of neurochemical data. It consists of data preparation, labelling and optimisation for training and validation of the CNN. The experimental data consisted of neurochemical pH recorded *in vivo* consisting of one saline injection and four CCK injections of different dosages (50, 100, 300, 1000 pmol/kg) [12].

A. Data Preparation

The raw neurochemical data consisted of biomarker induced neurochemical changes, background neurochemical activity and electrode drift. The raw data is pre-processed to remove drift using the technique described in [4]. After drift removal, the outliers were identified and removed. For outlier removal a modified z-score process was used, first described in [13] (Figure III-A). This was used to remove biased scaling due to very large data values.

After outlier removal, the data was partitioned into temporal windows of 15 seconds with an overlap of about 5 seconds. Following this the data was normalised to be such that its value was between -1 to 0. The normalisation was done using the following formula.

$$y_{norm} = \frac{y_i - y_{min}}{y_{max} - y_{min}} \quad (1)$$

The data was then labelled either 1 or 0 depending on whether the datapoint was part of CCK-induced neurochemical response or not, respectively (See Figure III-A).

B. Class size Imbalance

Class imbalance is a common issue in ML-based classification problems, where the objects belonging to one class exceed the number of objects in other classes. This problem can be solved through either creation or oversampling of minority class objects or undersampling of the majority class. In this paper, the number of datapoints which are not part of the CCK-induced response far outnumber those which originate from the CCK-induced response. Hence, we use Synthetic Minority Oversampling Technique (SMOTE) to deal with this issue [14]. In this case, SMOTE was applied to the data, post-windowing, such that there are equal number of positives(1) and negatives(0) in the training dataset.

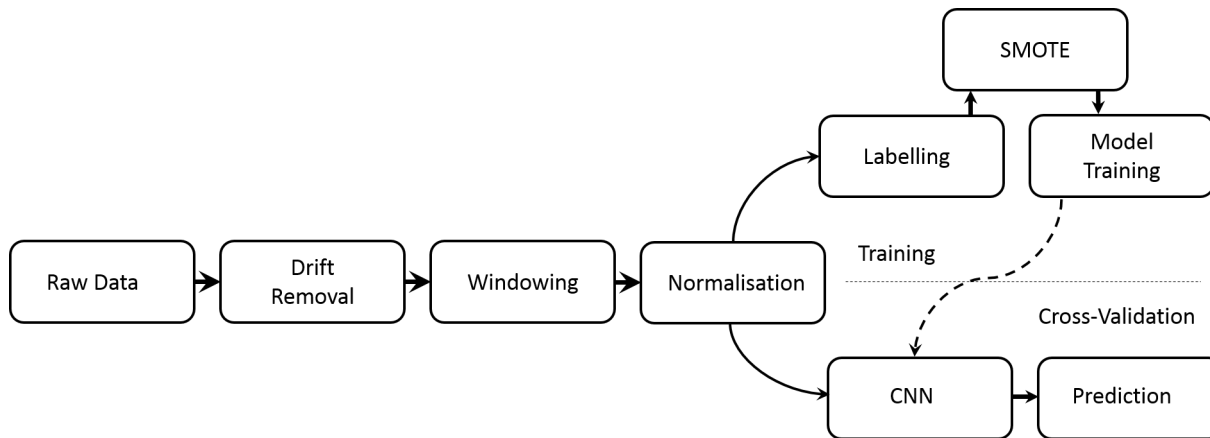


Fig. 2. Different steps used to train and cross-validate a CNN for detection and classification of neurochemical response.

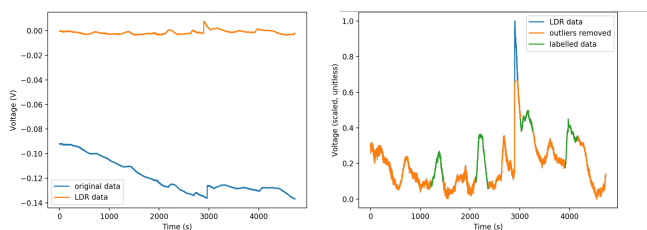


Fig. 3. (A) The drift in raw sensor data was removed using a technique described earlier ([4]). The outliers were removed using the modified z-score process [].

C. Model Training and Validation

A training data set of N experiments were prepared and provided to models generated using Keras 2.2.4 with TensorFlow 1.13.1 as a backend [15]. The model was provided with a training data tensor of shape $(N_{windows}, L_{windows}, 1)$, where $N_{windows}$ is the number of training windows, $L_{windows}$ is the length of each window. After a model architecture was selected, batches of training data were used to train the model. The batch size is used to determine the number of times per epoch, the weights are updated. After one complete run of training a model i.e epoch, the model was cross-validated on a cross-validation dataset. If the model performed better than the previous one, then the model was saved to be used later.

In order to prevent plateauing of model performance, a learning rate scheduler was used. Models in which the performance did not improve after 10 epochs had their learning rate halved. Models which did not improve after 40 epochs were discarded to reduce training time.

D. Model Architecture and Hyperparameter Optimisation

After completion of training and validation, objective evaluation of different model hyperparameters and architectures was validated with the same validation experiments using k-fold cross validation. Various model architectures, activation functions were tried ranging from dense neural networks

to convolutional and recurrent neural networks were implemented using open-source Neural Network Intelligence (NNI), Microsoft. Smaller changes were made further to manually improve model performance.

TABLE I
MODEL PARAMETERS OF CNN USED IN THIS PAPER

Parameters	Values
Convolutional Layers	4
Channels per layer	16
Kernels per Channel	3
Stride Length	1
Dropout Fraction	0.2
Activation Function	ReLU
Optimisation	ADAM
Batch Size	64

IV. RESULTS

A number of different parameters were tried to achieve optimal results with regards to training time and correct classification results.

A. Network Parameters: Batch Size, Window Length, Dropout Rate

1) *Batch Size*: There is a trade-off involved in the selection of batch size. If a large batch size is chosen i.e updating weights happens less frequently per epoch has less performance overhead compared to a small batch size. The impact of different batch sizes is seen in Figure 4. Figure 4 shows the model performance (validation loss) vs training time for each batch size. A batch size of 64 indicates optimal performance with fast training time and good model performance.

2) *Window Length*: The length of training windows were varied to determine an optimal window size. This is significant because it will help determine how early a CCK response is detected. A window length of 10 seconds with an overlap of 8 seconds is found to be optimal. Shorter time windows does not allow enough temporal features to be captured, hence

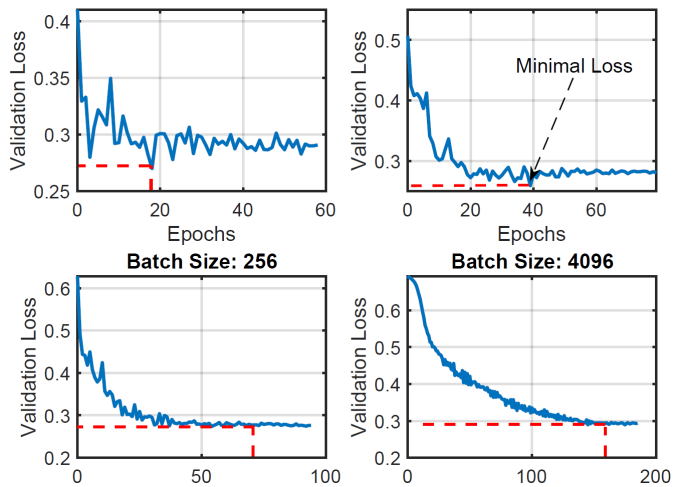


Fig. 4. An optimal batch size is chosen based on optimal validation loss values.

model performance drops. Longer windows results in excess data bleeding in from the non-CCK areas.

3) *Dropout rate*: Dropout is a tool used for regularisation by randomly deactivating a proportion of neurons in a model during each training step [16]. This is performed to prevent overfitting and also to deny neuron from specialising heavily on specific features. This also means readjustment of neuron weights as neurons with connections to dropped-out neurons will have to rely heavily on existing neurons. The effect on validation loss due to dropout is shown in Figure 5. A dropout of 20% was chosen as an optimal value.

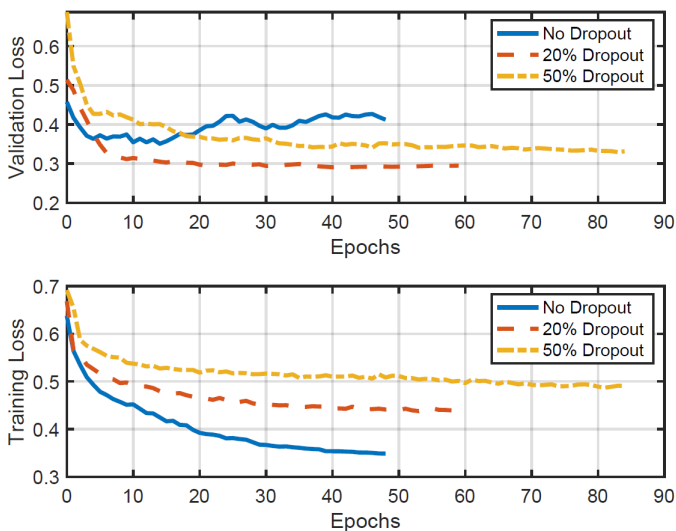


Fig. 5. The effect of different dropout rates show the effect on validation loss (A) and training loss (B). This is used to determine an optimal value for dropout rate while training the model.

B. Overall Performance and Improvements : Precision and Recall

The overall performance of the final model (parameters shown in Table I) indicates an accuracy of 93.27% (fraction

of correct predictions compared to total predictions). This was achieved with a minimum validation loss of 0.186. The optimisation algorithm used for the CNN was ADAM. A sample of prediction accuracy of the model is shown in Figure 6 and overall results shown in Table II. Table II indicates performance of the model on incoming real-time data where class-size imbalance negatively affects the model performance.

TABLE II
PRECISION AND RECALL RESULTS ON ENTIRE CROSS-VALIDATION DATASET

Predicted Label Target Label	Non CCK Event	CCK Event
Non CCK Event	23010 (TN)	204 (FP)
CCK Event	1151 (FN)	502 (TP)

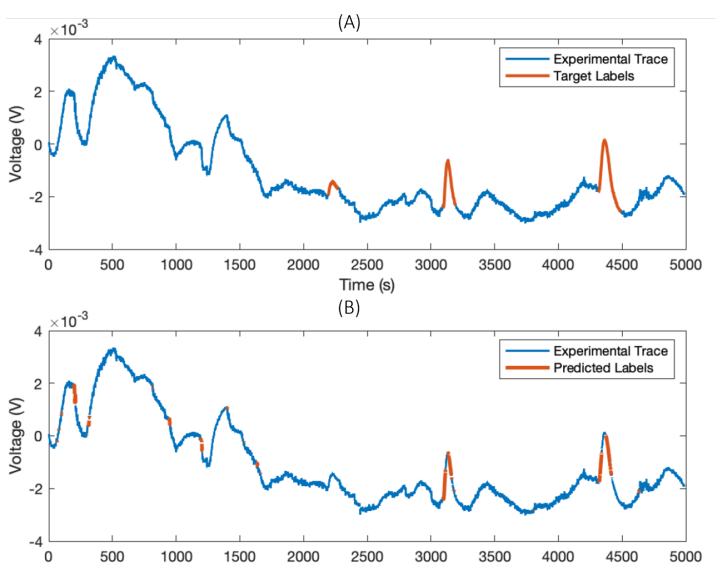


Fig. 6. (A) A cross-validation, experimental data set is labelled as shown. (B) Shows the datapoints predicted by the CNN model.

V. CONCLUSION

This paper describes data pre-processing steps, feature enhancement, development and optimisation of a convolutional neural network model for detecting neurochemical changes in a real-time application. The advantage of using a CNN is that it ensures early, more accurate detection of neurochemical changes. Although the model successfully detects neurochemical changes, however the accuracy is affected by class-size imbalance for real-time detection. Hence, there is a need either for better quality training data or additional feature information which can further improve the accuracy of the model.

ACKNOWLEDGMENT

This work was supported by the European Research Council (Synergy Grant No. 319818, i2MOVE), UK EPSRC grant EP/N002474/1, and EPSRC-Imperial Impact Acceleration Award (2019).

REFERENCES

- [1] B. Rosin, M. Slovik, R. Mitelman, M. Rivlin-Etzion, S. N. Haber, Z. Israel, E. Vaadia, and H. Bergman, "Closed-loop deep brain stimulation is superior in ameliorating parkinsonism," *Neuron*, vol. 72, no. 2, pp. 370–384, 2011.
- [2] F. T. Sun and M. J. Morrell, "Closed-loop Neurostimulation: The Clinical Experience," *Neurotherapeutics*, vol. 11, no. 3, pp. 553–563, 2014.
- [3] K. B. Mirza, N. Kulasekeram, Y. Liu, K. Nikolic, and C. Toumazou, "System on chip for closed loop neuromodulation based on dual mode biosignals," *Proceedings - IEEE International Symposium on Circuits and Systems*, vol. 2019-May, 2019.
- [4] T. Ahmed, K. B. Mirza, and K. Nikolic, "Resource Efficient Pre-processor for Drift Removal in Neurochemical Signals," in *IEEE International Symposium on Circuits and Systems*, 2018, pp. 1–5.
- [5] B. Bozorgzadeh, D. P. Covey, B. A. Heidenreich, P. A. Garris, and P. Mohseni, "Real-time processing of fast-scan cyclic voltammetry (FSCV) data using a field-programmable gate array (FPGA)," *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2014, pp. 2036–2039, 2014.
- [6] P. J. Grahn, G. W. Mallory, O. U. Khurram, B. M. Berry, J. T. Hachmann, A. J. Bieber, K. E. Bennet, H.-k. Min, S.-y. Chang, K. H. Lee, and J. L. Lujan, "A neurochemical closed-loop controller for deep brain stimulation : toward individualized smart neuromodulation therapies," vol. 8, no. June, pp. 1–11, 2014.
- [7] S. C. Cork, "The role of the vagus nerve in appetite control: Implications for the pathogenesis of obesity," *Journal of Neuroendocrinology*, vol. 30, no. 11, pp. 1–10, 2018.
- [8] S. Okano-Matsumoto, J. A. McRoberts, Y. Taché, and D. W. Adelson, "Electrophysiological evidence for distinct vagal pathways mediating CCK-evoked motor effects in the proximal versus distal stomach." *The Journal of physiology*, vol. 589, no. Pt 2, pp. 371–93, 2011. [Online]. Available: <http://doi.wiley.com/10.1113/jphysiol.2010.196832%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3043539&tool=pmcentrez&rendertype=abstract>
- [9] K. B. Mirza, A. Alenda, A. Eftekhar, N. Grossman, K. Nikolic, and R. Stephen, "INFLUENCE OF CHOLECYSTOKININ-8 ON COMPOUND NERVE ACTION POTENTIALS FROM VENTRAL GASTRIC VAGUS IN RATS," *International Journal of Neural Systems*, vol. Accepted, no. Accepted, pp. 1–15, 2018.
- [10] K. B. Mirza, K. Wildner, N. Kulasekeram, S. C. Cork, S. R. Bloom, K. Nikolic, and C. Toumazou, "Live Demo: Platform for Closed Loop Neuromodulation Based on Dual Mode Biosignals," in *IEEE 2017 Biomedical Circuits and Systems Conference, BioCAS 2017 - Proceedings*, Milan, 2017.
- [11] J. K. Trevathan, A. Yousefi, H. O. Park, J. J. Bartoletta, K. A. Ludwig, K. H. Lee, and J. L. Lujan, "Computational Modeling of Neurotransmitter Release Evoked by Electrical Stimulation: Nonlinear Approaches to Predicting Stimulation-Evoked Dopamine Release," *ACS Chemical Neuroscience*, vol. 18, no. 3, pp. 386–392, 2017.
- [12] S. C. Cork, A. Eftekhar, K. B. Mirza, C. Zuliani, K. Nikolic, J. V. Gardiner, S. R. Bloom, and C. Toumazou, "Extracellular pH monitoring for use in closed-loop vagus nerve stimulation," *Journal of Neural Engineering*, vol. 15, no. 1, p. 016001, feb 2018.
- [13] A. Schmidt and Z. Bandar, "A modular neural network architecture with additional generalization abilities for large input vectors," in *Artificial Neural Nets and Genetic Algorithms*. Springer, 1998, pp. 35–39.
- [14] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.
- [15] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.
- [16] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.