Deep-Channel: A Deep Convolution and Recurrent Neural Network for Detection of Single Molecule Events

Numan Celik[1], Fiona O'Brien[1], Yalin Zheng[1], Frans Coenen[2], Richard Barrett-Jolley[1*]

[1]Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, United Kingdom,

[2]Department of Computer Science, University of Liverpool, Liverpool, United Kingdom

***Correspondence:**

Corresponding Author

rbj@liverpool.ac.uk

Keywords: Single molecule, Ion Channel, smFRET, Markov Model, Deep Learning, CNN, LSTM

Abstract

Single molecule research delivers a unique biological insight by capturing the movement of individual proteins in real time, unobscured by whole-cell ensemble averaging. The critical first step in single molecule analyses is event detection, so called “idealisation”, where noisy raw data are turned into discrete records of protein movement. The most common type of single molecule research is electrophysiological patch-clamp recording of ion channel gating. To date, there have been practical limitations in the analyses pipelines for ion channel or other single molecule data; they are typically manually performed; laborious and require human supervision. In addition, this task can become infeasible with complex biological data containing many distinct native single ion channel proteins gating simultaneously. In this report we describe an “artificial intelligence” deep learning model for analyses of single molecule data, based on convolutional neural networks (CNN) and long short-term memory (LSTM) architecture. This network automatically idealises complex single molecule activity more accurately and faster than manual “threshold crossing” analyses. We believe this is the first use of deep learning to analyse single molecule datasets and such methods may revolutionise the unsupervised automatic detection of ion channel and other single-molecule transition events in the future.

Introduction

Ion channels produce "functional" data in the form of electrical currents typically recorded with the Nobel Prize winning patch-clamp electrophysiological technique[1,2]. The role of ion channels in generation of the nerve action potentials was first described in detail in the Nobel Prize winning work of Hodgkin and Huxley[3], but it is now known they sub serve a wide range of processes via control of the membrane potential[4]. Loss or dysregulation of ion channels directly underlies many human and non-human animal diseases (so called "channelopathies"); including cardiovascular diseases such as "LQT" associated Sudden Death[5]. The first step in analyses of ion channel or other single molecule data (which may, in fact, include several individual "single" proteins) is to idealise the noisy raw data. This is typically accomplished by human supervised threshold-crossing although other human supervised methods are available[6,7]. This produces time-series data with each time point binary classified as open or closed, with more complex data this is a categorical classification problem, with classifiers from zero to *n* channels open. Similar data are also acquired from other single molecule techniques such as lipid bilayer[8] or single molecule FRET[9-11]. These data can then be used to re-construct the hidden Markov stochastic models underlying the protein activity, using applications such as HJCFIT[12], QuB (SUNY, Buffalo[13]) or SPARTAN[11]. The initial idealisation step, however, is well recognised by electrophysiologists as a time consuming and labour-intensive bottleneck. This was perhaps best summarised by Professors Sivilotti and Colquhoun FRS[14] "[patch-clamp recording is] *the oldest of the single molecule techniques, but it* ***remains unsurpassed [in] the time resolution that can be achieved.*** *It is the richness of information in these data that allows us to study the behaviour of ion channels at a level of detail that is unique among proteins. [BUT] This quality of information comes at a price* […]. ***Kinetic analysis is slow and laborious****, and its success cannot be guaranteed, even for channels with good signals*". In the current report we show that the solution to these problems could

be to apply the latest artificial intelligence “deep learning” methodology to “single channel” patch-clamp data analyses.

Deep learning[15] is a machine learning development that has been used to extract features and/or to detect objects from different types of datasets for classification problems. Convolutional neural network layers (CNNs) are a powerful component of deep learning useful for learning patterns within complex data. 2-dimensional (2D) CNNs are most commonly applied to computer vision[16-18] and we have previously used them for automatic diagnosis of retinal disease in images[19]. An adaptation of the 2D CNN is the one-dimensional (1D) CNN. These have been specifically developed to bring the power of the 2/3D CNN to classification time series, although never previously patch clamp data. More commonly, the deep learning architecture known as recurrent neural networks (RNNs) have been applied to time series analyses[20,21]. General RNNs are a useful model for text and speech classification and object detection in time series data, however the model begins degrading once output information depends on long time scales due to a vanishing gradient problem[22]. Long short-term memory (LSTM) networks, are a type of RNN that resolve this problem[23-25]. While 1D-CNNs layers can effectively classify raw sequence data, in the current work we combine these with LSTM units to improve the detection of learn long term temporal relationships in time series data[26].

In the current work, we introduce a hybrid recurrent convolutional neural network (RCNN) model to idealise ion channel records, with up to 5 ion channel events occurring simultaneously. To train and validate models, we developed an analogue synthetic ion channel record generator system and find that our “Deep-Channel” model, involving LSTM and CNN layers, rapidly and accurately idealises/detects experimentally observed single molecule events without need for human supervision. To our knowledge, this work is the first deep learning model designed for the idealisation of single molecule events.

Results

Benchmarking Deep-Channel event detection against human supervised threshold crossing.

Our data generation workflow is illustrated in Fig. 1a,c. In training and model development we found that whilst LSTM models gave good performance, the combination with a time distributed CNN gave increased performance, a so called "RCNN" we call here Deep-Channel. After training and model development we used 17 newly generated datasets, previously unseen by Deep-Channel, and thus uninvolved with the training process. Authentic ion channel data (Fig.1b) were generated as described in the methods from two kinetic schemes, the first "M1" (see methods and Fig. 2a) with low channel open probability, and the second "M2" with a high open channel probability and thus an average of approximately 3 channels open at a time (Fig. 2b). Across the datasets we included data from both noisy, difficult to analyse signals and low noise (high signal to ratio samples) as would be the case in a any patch-clamp project. Examples of these data, together with ground truth and Deep-Channel idealisation are shown in Fig. 3.

In datasets where channels had a low opening probability (i.e., from model M1), the data idealisation process becomes close to a binary detection problem (Fig. 3a), with ion channel events type closed or open (labels '0' and '1' respectively). In this classification, the ROC area under the curve (AUC) for both open and closed event detection exceeds 96% (Figure 4, Table I). Full data for one example experiment is shown, with confusion matrix and ROC in Fig. 4a,b. Overall, in experiments with low open channel probability Deep-Channel returned a macro-F1 accuracy of 0.89±0.02 n=10, whereas the 50% threshold method in QuB gave a macro-F1 of 84.7±0.05, n=10.

In cases where datasets included highly active channels (i.e., from model M2, Fig. 2c,d, Fig. 3b) this becomes a multi-class comparison problem and here, Deep-Channel outperformed 50% threshold-crossing event detection considerably. The Deep-Channel macro-F1 accuracy for such events was 0.87±0.07, n=7, but 50% threshold-crossing macro-F1 fell to 0.47±0.37 (Student's paired t-test

between methods, p=0.0052). An example ROC for high activity channel detection, and associated confusion matrix is shown in Fig. 4c,d. Frequently, in drug receptor or toxicity studies biologists look for changes in open probability and so we also compared the open probability from manual threshold crossing and Deep-Channel against ground truth (fiducial). In the presence of several simultaneously opening channels in some quite noisy datasets, careful manual 50% threshold crossing sometimes essentially fails entirely, but Deep-Channel continues to be successful. For example, in some experiments with very noisy data threshold crossing open probability estimations were over 100% out. Nevertheless, overall there were highly significant correlations for both Deep-Channel vs ground truth (0.9998, 95% confidence intervals 0.9996 to 0.9999, n=17) and threshold crossing vs ground truth (0.95, 95% confidence intervals 0.87 to 0.98, $n=17$). In terms of speed, Deep-Channel consistently outperforms threshold crossing. Deep-channel analysed at the rate of approximately 10 seconds of data recording in under 4 seconds of computational time, whereas analysis time with threshold crossing in QuB was entirely dependent on the complexity of the signal. For multi-channel patches it took over 1 minute of analysis time *per second* of recorded data, and this was by very experienced QuB users.

Table I Deep-Channel ROC area under the curve (AUC) values achieved during 17 separate experiments, representative examples shown in Figure 4.

	Mean ROC AUC	
Number of Channels open simultaneously	**low Popen data model M1, *n*=10** **mean ± SD**	**High Popen data model M2, *n*=7** **mean ± SD**
0	0.98±0.02	0.9997±0.0002
1	0.96±0.05	0.9973±0.0022
2	-	0.9917±0.0065
3	-	0.9824±0.0139
4	-	0.9787±0.0169
5	-	0.9937±0.0050

Real (biological) protein data testing

In the final section of this work, we examined the ability of the Deep-Channel model to classify real "single-molecule" events. Firstly, ion channel events from real biological data previously recorded, analysed and published[27]. In this case, by definition *there cannot be any fiducial (ground truth) idealisation* and so experimenters typically examine the quality of their idealization by eye. Note these data had not previously been exposed to the network, there was no further training etc and so it forms an authentic real-world test. Fig. 5a illustrates the results of these experiments with real biological data, comparing raw to QuB idealisation and Deep-Channel idealisation. Lastly, we test Deep-Channel with another clear substrate for automated AI event classification, single molecule FRET (smFRET) data. These data were obtained with kind permission of Prof Scott Blanchard[11]. Again, just as with patch clamp data, *there can be by definition no ground truth* and so visual inspection (of the raw and

idealised records) is the gold standard method to assess the quality of idealisation. Figure 5B illustrates the idealisation of smFRET data with QuB and Deep-Channel.

Discussion

Single molecule research, both FRET and patch-clamp electrophysiology provide high resolution data on the molecular state of proteins in real time, but their analyses are usually time consuming and require expert supervision. In this report we demonstrate for the first time, that a deep neural network, Deep-Channel, combining recurrent and convolutional layers can detect single molecule events in data automatically. Deep-channel analysis is completely unsupervised and thus adds objectivity to single molecule data analyses. With complex data Deep-Channel also significantly outperforms traditional manual threshold crossing both in terms of *speed and accuracy*. We find this method works with very high accuracy across a variety of input datasets.

The most established single molecule method in biology is patch-clamp recording[2]. Its development led to the award of the Nobel Prize to Sakmann and Neher in 1991[28] and the ability to observe single channels gate in real time validated the largely theoretical model of the action potential developed in the earlier Nobel Prize winning work of Hodgkin and Huxley[3]. Whilst the power and resolution of single channel recording has never been questioned, it is well accepted to be a technically difficult technique to use practically since the data stream created requires a laborious supervised analyses procedure. In some cases, where several “single” channels “gate” simultaneously, it becomes impractical to analyse and data can be wasted. For practical purposes, drug screening etc, where subtle changes in channel activity could be crucial[5], this means that the typical method is to measure bulk activity from a whole-cell simultaneously. Average current can be measured which is useful, but does not contain the detailed resolution that individual molecular recording has[14]. Furthermore, new technologies are emerging which can record ion channel data automatically[29,30], but whilst whole-cell

currents are large enough to be analysed automatically, there are currently no solutions to do the same with single-channel events. In this report we show that the latest artificial intelligence methods, that of deep learning, including recurrent and convolutional neural network layers can address these limitations.

The fundamental limitation to applying deep learning to classification of biological data of all kinds is the prerequisite for training data. Deep learning is a form of supervised learning where during the training phase, the network must "be taught" at every single instant what the ground truth state is (it looks open, but is it really open or closed?). We considered two possible approaches to deal with this conundrum; (a) to collect data from easily analysable single molecule/patch clamp experiments and get a human expert to idealise this (classify or "annotate" it). This has two fundamental flaws that could be referred to as "Catch-22". Firstly, if you train a network only to detect "easy" to analyse data, the output will be a network that can only detect "easy" to detect events. Secondly, even then, if the events had to be human detected in the first place, it would mean that the final (trained) network would learn the same events as the human taught it; *and learn the same errors*. Analyses of ambiguous events would not tend toward detection with perfect accuracy, but inherit human biased errors. We therefore developed an alternative approach (b); single channels gate in a stochastic, Markovian manner and therefore an unlimited number of idealised records can be simulated. This approach has been successfully applied before with other analyses development studies[12,31]. The limitation is that there are inherent distortions and filtering that occur during collection of genuine data from a real analogue world. These can be imitated mathematically, but instead we used an entirely novel method of generating semi-synthetic training data; we played our idealised records out to a genuine patch-clamp amplifier using the dynamic-clamp approach[32] and used an established analogue "test cell" (resistors and capacitors equivalent to a patch pipette and membrane). Our first data figure (Fig. 1b) shows the authenticity of this data and the approach. In summary, our novel methodology allows the creation of 100,000s training sets with noisy data in parallel to a ground truth idealisation.

Since our aim was to classify a time series, we developed a network with the combined power of both 1D-CNN layers and RNN (LSTM) units. Deep-Channel has a 1D-CNN at its core, but whilst ion channel activity is Markovian, the presence of both short and long duration underlying states means that it is important for a detection network to also be able to learn long-term dependencies across and so accuracy is improved with the LSTM (*see preliminary data*). Similar approaches combining RNNs and convolution layers have previously been applied to various analysis of biological gene sequences[33] and cell detection in image classification[34], but this is its first use for single molecule activity detection.

We used a number of metrics that are commonplace in machine learning and patch-clamp recording. Initially, to test the ability of Deep-Channel to detect events, we compared detected ("predicted") idealised events against the fiducial idealised records. To compare against the human supervised methods (we analysed matching datasets with QuB and Deep-Channel and compared the summary parameter, open probability (*Po*) between them, together with F1 accuracies. Firstly, Deep-Channel was far quicker, to the degree that sometimes with complex multichannel data, manual analyses seems a stressful and near impossible task. With Deep-channel the complexity (within the datasets we used) made no difference to the speed or accuracy of performance. In terms of relative performance, note that whilst there was a strong correlation between open probabilities measured between Deep-Channel and threshold crossing, the F1-accuracy scores of the noisy data measured by threshold crossing fell of sharply. The significance of this is that whilst average open probability estimates from threshold crossing seem reasonable (some over estimates, some under-estimates cancelling out) the time-point by time-point accuracy *essential for kinetic analyses* is poor.

A perceived problem with a machine learning model is its generalisability. The concern is that the network would be good at detecting events in the exact dataset it was trained on, but fall short, when challenged with a different, but equally valid dataset; a problem known as “over fitting”. We chose to conclude this study by testing the ability of the network to detect events in existing published data, either from our own lab or freely available. We used archive patch-clamp recording data from skeletal muscle BK channels and a stretch of FRET data from[11]. Since there is, by definition, no way to have ground truth with native biological data the only way to assess the quality of detection with such data is by eye. Fig. 5 shows the success of Deep-Channel to detect events in these archive sections of biological data. It is critical to note these data were fed straight into the trained network and that there was no element of retraining necessary.

We have demonstrated here the effectiveness of Deep-Channel, an artificial deep neural network to detect events in single molecule datasets, especially, but not exclusively patch-clamp data, but the potential for deep learning convolution/LSTM networks to tackle other problems cannot be overestimated.

Figures and legends

Figure 1

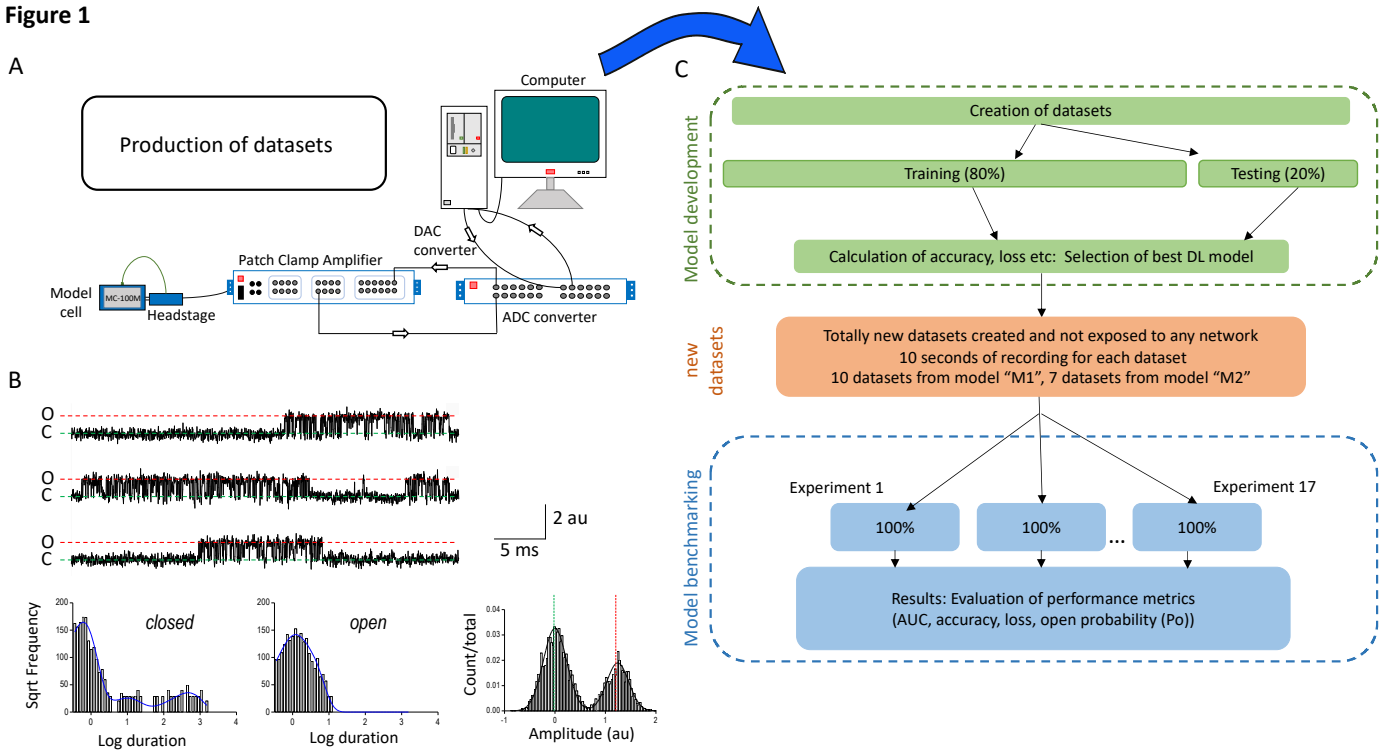


Fig. 1 | Workflow diagram: Generation of artificial analogue datasets. a, For both training, validation and benchmarking data were generated first as fiducial records with authentic kinetic models in MATLAB; these data were then "played" out through a CED digital to analogue converter to a patch clamp amplifier that sent this signal to a model cell and recorded the signal back (simultaneously) to a hard disk with CED Signal software via a CED analogue to digital converter. The degree of noise could be altered simply by moving the patch-clamp amplifier closer or further to the PC. In some cases, drift was added as an additional challenge via a separate Matlab script. **b,** Raw "single channel patch clamp data" produced by these methods are visually indistinguishable from genuine patch clamp data. To illustrate this point we show here a standard analyses work-up for one such experiment with **bi**, raw data, **bii,** all points amplitude histogram and **biii,** kinetic analyses of channel open and closed dwell times. The difference between this and standard ion channel data is that here we have a perfect fiducial record with each experimental dataset, which is impossible to acquire without simulation. **c,** Illustrates our over-all model design and testing workflow. The supplementary data includes training metrics from the initial validation and the main text here shows performance metrics acquired from 17 experiments with entirely new datasets. The training datasets typically contained millions of sample points and the 17 benchmarking experiments were sequences of 100,000 samples each.

Figure 3

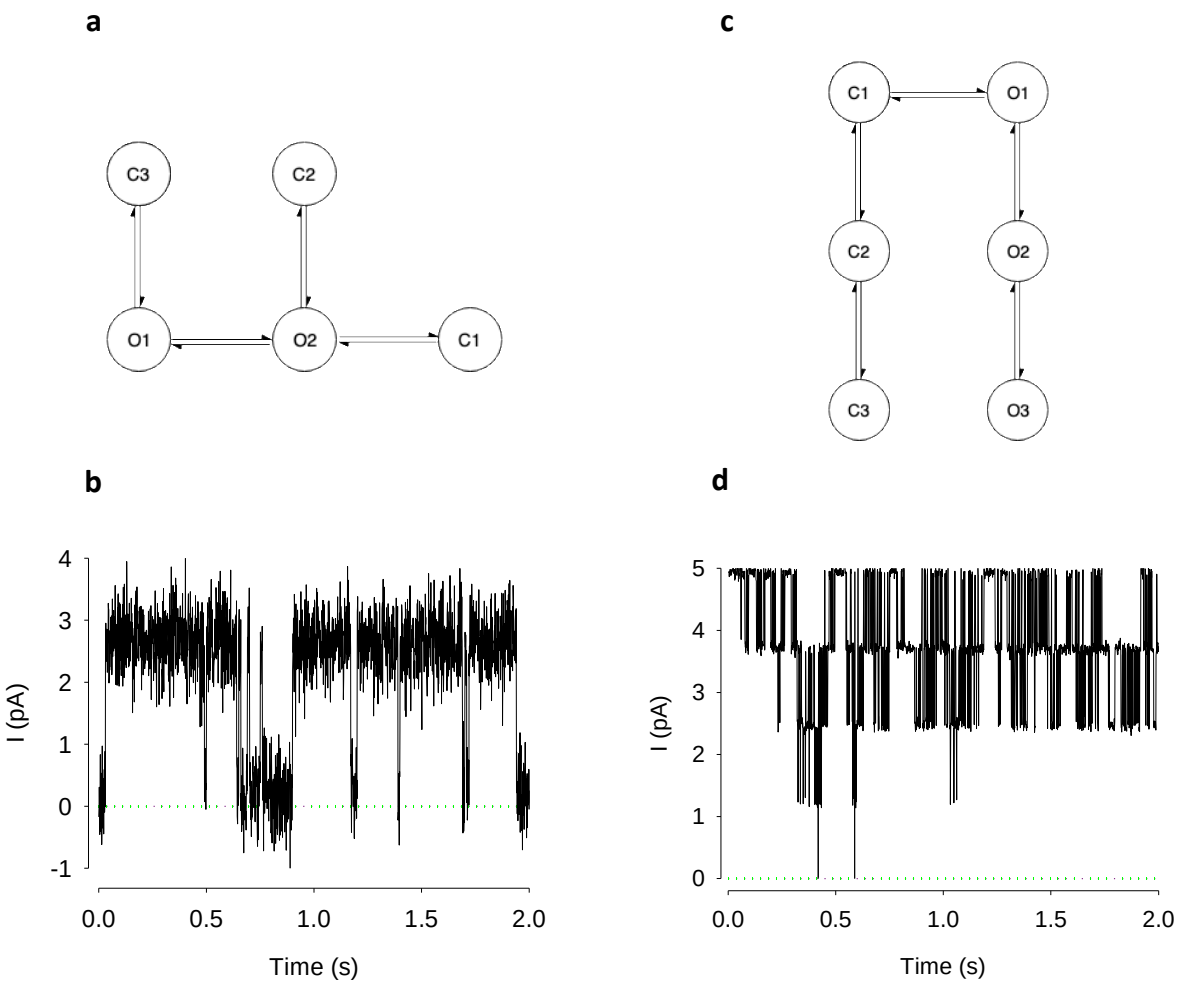


Fig.2 | "Patch clamp" data was produced from two different stochastic models. a, and **c,** are the Markovian models used for simulation of ion channel data. Ion channels typically move between Markovian states that are either "closed" (zero conductance) or "open" (unitary conductance, *g*). The current passing when the channel is closed is zero (aside from recording offsets and artefacts), whereas when open the current (*i*) passing is given by $i = g \times V$, where *V* is the "driving potential" (equilibrium potential for the conducting ion minus the membrane potential). In most cases there are several open and closed states ("O1", "O2", "O3", or "C1", "C2", "C3" respectively). The central dogma of ion channel research is that the *g* will be the same for O1, O2 or O3. Although substates have been identified in some situations, these are beyond the scope of the current models. **a,** Model "M1"; the stochastic model from Davies et al[35] and its output **b.** This model has a low open probability, and so the data is mostly a representation of zero or one channel open. **c,** Model "M2"; the stochastic model from [36] and its output data **d**, since open probability is high the signal is largely composed of 3 or more channels simultaneously open.

Figure 4

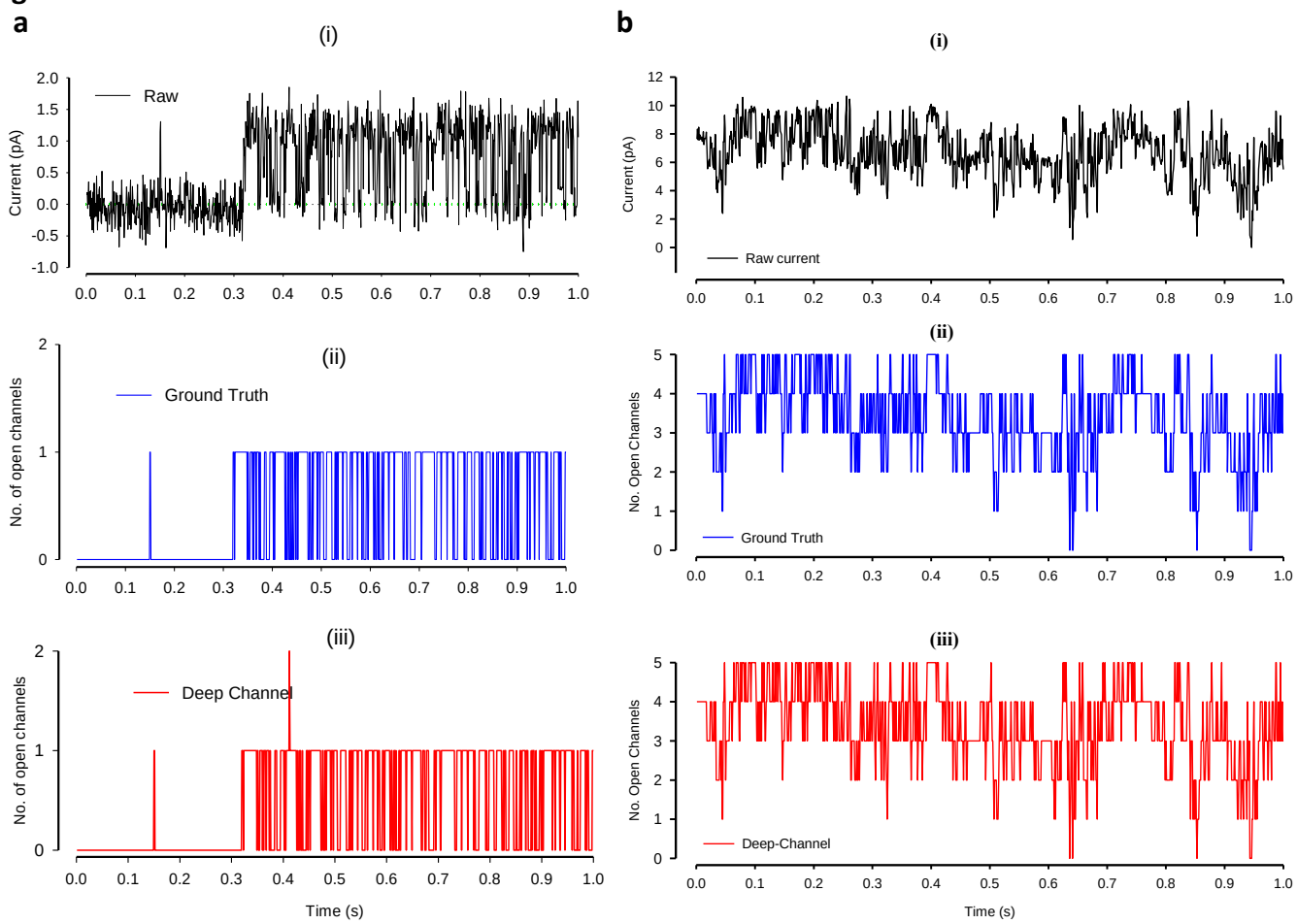
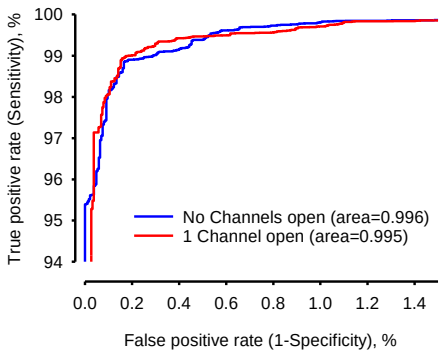


Fig. 3 | Qualitative performance Deep-Channel with previously unseen data. a, Representative example of Deep-Channel classification performance with low activity ion channels (data from model M1, Fig. 2a,b): **ai,** The raw semi-simulated ion channel event data (black). **aii,** The ground truth idealization/annotation labels (blue) from the raw data above in **ai**. **aiii,** The Deep-Channel predictions (red) for the raw data above **ai**. **b,** Representative example of Deep-Channel classification performance with 5 channels opening simultaneously (data from model M2, Fig. 2c,d). **bi,** The semi-simulated raw ion channel event data (black). **bii,** The ground truth idealization/annotation labels (blue) from the raw data above in **bi**. **biii,** The Deep-Channel label predictions (red) for the raw data above **bi.**

Figure 5

a **Low open probability model**

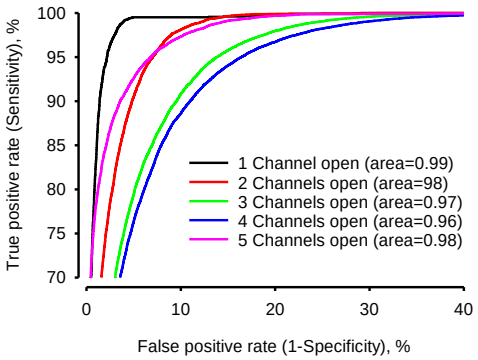


b

Deep-channel predicted label

True label	0	1	2	3	4	5
0	**96890**	726	0	0	0	0
1	6	**2365**	12	0	0	0
2	0	0	**0**	0	0	0
3	0	0	0	**0**	0	0
4	0	0	0	0	**0**	0
5	0	0	0	0	0	**0**

c **High open probability model**



d

Deep-channel predicted label

True label	0	1	2	3	4	5
0	**129**	15	0	0	0	0
1	130	**1993**	262	0	0	0
2	0	932	**9580**	1621	0	0
3	0	0	1884	**25241**	3054	0
4	0	0	0	2782	**31439**	2579
5	0	0	0	0	2050	**16308**

Fig. 4 | Quantitative performance of Deep-Channel with previously unseen data. a, Representative receiver operating characteristic (ROC) curve for ion channel event classification using the M1 stochastic gating model (Fig 2a.) and with only one channel present. **b,** The confusion matrix table for the example in **a**. Label 0 = no channels open, label 1= one channel open. Note that Deep-Channel is trained to recognise up to 5 channels opening at a time, however, with only one channel active at a time, the maximum ground truth class ("True Label") is label 1 (one channel open). When analysing this simulated patch, Deep-Channel only made 12 (incorrect calls) of labels 2 to 5 ("Deep-Channel predicted labels"). **c,** Representative receiver operating characteristic (ROC) curve for ion channel event classification using the M2 stochastic gating model (Fig. 2) and with five channels present. Mean AUC are given in Table I. **d,** The confusion matrix table for the example in **c**. Label 1= one channel open, label 2= 2 channels open etc. The mean ROC curve area under the curves (AUC) for all labels and all 17 experiments are given in Table I.

Figure 6

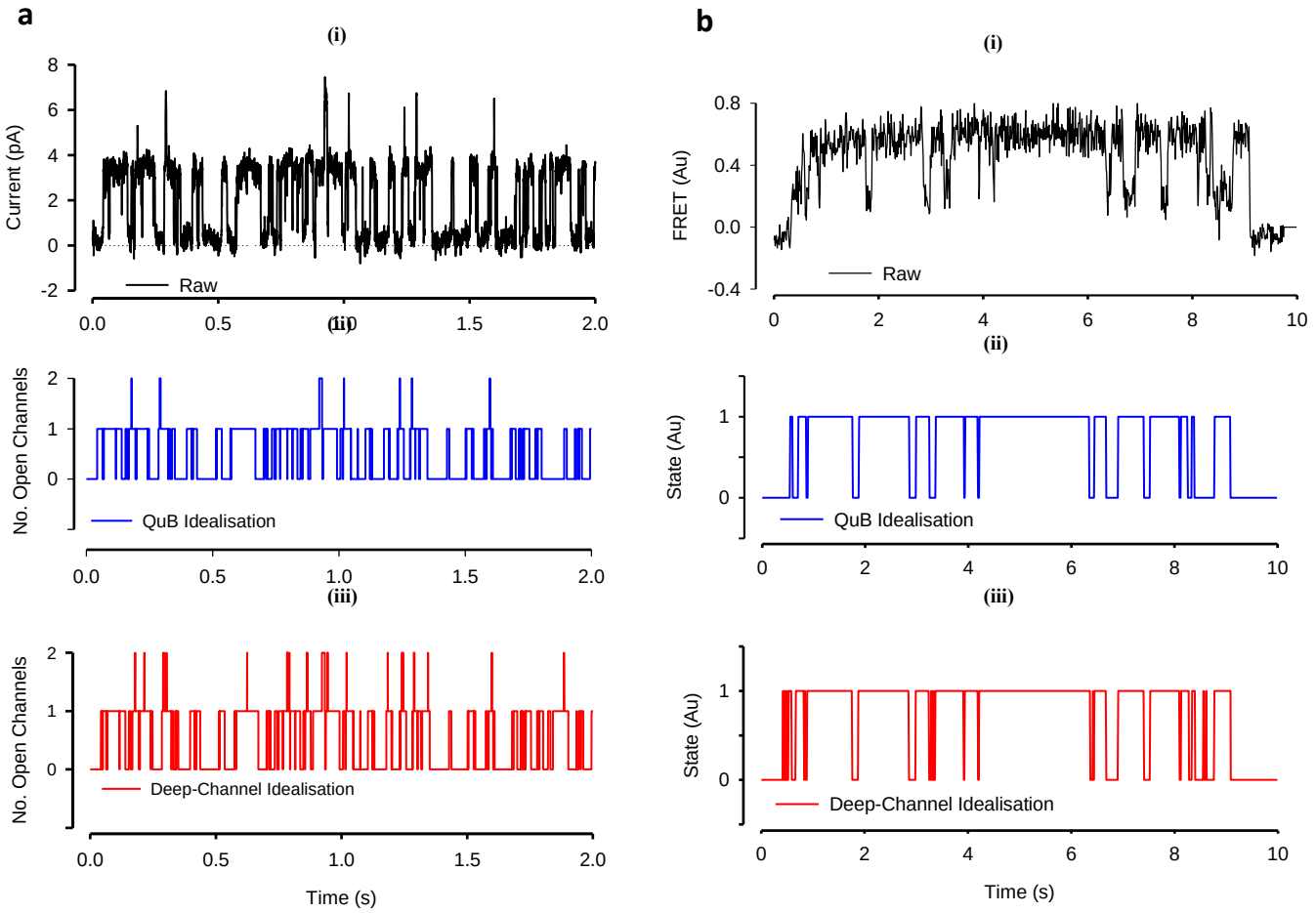


Fig. 5 | Performance of Deep-Channel with biological ion channel single and single molecule FRET data. **a,** Single channel data from[27] and idealised with Deep-Channel without adjustment. **ai,** Raw ion channel data (black). **aii,** QUB threshold idealized ion channel data (blue) **aiii,** Deep-Channel classified/idealised events (red). **b,** Single molecule FRET data from[11] fed to Deep-Channel without adjustment. **bi,** Raw FRET data (black)[11]. **bii,** Human supervised threshold idealisation using QuB (blue). **biii,** Deep-Channel classified states (red).

Methods

We develop a novel deep learning approach to automatically process large collections of single/multiple ion channel data series with detection of ion channel transition events, and re-construction of annotated idealised records. Datasets with pre-processing and analysis pipeline code will be made publicly available on GitHub once published in a peer review journal, including the model code to facilitate reproducibility. Fig. 1 shows an overall workflow and experimental design; creation of the digitised synthetic analogue datasets for developing a deep learning model, together with steps for training and testing (validating).

Data description and dataset construction

Ion channel dwell-times were simulated using the method of Gillespie[37] from published single channel models. Channels are assumed to follow a stochastic Markovian process and transition from one state to the next simulated by randomly sampling from a lifetime probability distribution calculated for each state. Authentic 'electrophysiological' noise was added to these events by passing the signal through a patch-clamp amplifier and recording it back to file with CED's Signal software via an Axon electronic "model cell". In some datasets additional "drift" was applied to the final data with Matlab. Two different stochastic gating models, (termed M1 and M2) were used to generate semi-synthetic ion channel data. M1 is a low open probability model from[35] (Fig.2a, b), typically no more than one ion channel opens simultaneously. Model M2 is from[36,38] and has a much higher open probability (Fig. 2c, d), consequently up to 5 channels opened simultaneously and there are few instances of zero channels open. The source code for generating a combination of different single/multiple ion channel recordings is also given along with the publicly available datasets. Using this system, we can generate any number of training datasets with different parameters such as number of channels in the "patch", number of open/close states, sampling frequency and temporal duration, based on published stochastic models. Fiducial, "ground truth" annotations for these datasets were produced

simultaneously using MATLAB. Recordings were sampled at 10 kHz and each record has 10 seconds duration. To validate the "Deep-Channel" model, 6 different validation datasets were used: 3 datasets for single; and 3 datasets for multi-channel recordings. Datasets for training typically contained 10,000 subsets of 10 seconds each. Each dataset includes raw current data and ground truth state labels from the stochastic model, which we refer to as the "idealisation". Within these training datasets, the third column is the fiducial record/ground truth and includes the class labels; '0', '1', '2', '3', '4' and '5'. Each label indicates the instantaneous number(s) of open channels at a given time.

Model training

In the model training stage, once the probability values are calculated, errors between the predicted values and true values were calculated with a sparse categorical cross entropy as a loss function. To optimize the loss value, stochastic gradient descent was applied as an optimizer with an initial learning rate of 0.001, momentum of 0.9, and the size of a mini-batch was set to between 256 and 2048 depending on the model. A learning rate decaying strategy was employed to the model to yield better performance. Based on this strategy, the learning rate (initially is 0.001) was decayed at each 10th epoch with decaying factor 0.01 of learning rate. The proposed Deep-Channel model was trained for 50 epochs. In the case of the training data an 80% -train and 20%-test split was performed.

Performance Metrics

One of the clearest quality indicators of a classification deep learning method is the confusion matrix, sometimes called as contingency table[39]. This consists of four distinctive parameters, which are true positives *(TP),* false positives *(FP),* false negatives *(FN),* and true negatives *(TN).* If the model's output accurately predicts the ion channel event, it is considered as TP. On the other hand, it is indicated as FP if the model incorrectly detected an ion channel event when there is no a channel opening. If the model output misses an ion channel event activity, then it is computed as FN. These metrics are used

in calculation of evaluation metrics such as precision (positive predicted value), recall (sensitivity), and F-score as described below in (3), (4), and (5) respectively:

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

$$F = \frac{2PR}{P + R} \tag{5}$$

where *P*, *R*, and *F* denote precision, recall and F-score, respectively. In addition, area under curve (AUC) and receiver operating characteristic (ROC) parameters are efficiently used to visualize the model performance in classification problems. The ROC shows the probability relations between true positive rate (sensitivity-recall), and false positive rate (1-specificity), while AUC represents a measure of the separability between classes.

As an additional metric, more familiar to electrophysiologists we also calculated the open probability (Po), and compared this metric between Deep-Channel, a traditional software package (QuB) and the ground truth. The equation for open probability is given in (equation 6) as follows:

$$P_o = \frac{\sum_{j=1}^{N} t_j j}{TN} \tag{6}$$

where *T* denotes total time, *N* is defined as numbers of channels in the patch, and t_j is referred to the time spent with *j* channels open[40]. Since true numbers of channels in a patch is always an unknown parameter this was estimated as the maximum number of simultaneous openings.

Computer hardware

In this work, we trained the Deep-Channel model on a workstation with an Nvidia Geforce GTX 1080Ti and 32GB RAM. The entire process of the proposed model including training, validating, evaluating, and visualizing process was employed within Python Spyder 3.6. Speed estimations were made with a similar PC, but without a GPU.

Statistics

Statistical analyses were performed with R-studio, means are quoted with standard deviations, *n* is the number of experiments.

Reporting Summary

Further information on research design is available in the Nature Research Supporting Summary linked to this article.

Data availability

Datasets with pre-processing and analysis pipeline code will be made publicly available on GitHub (https://github.com/RichardBJ/Deep-Channel.git)

Code availability

All code will be available on GitHub (https://github.com/RichardBJ/Deep-Channel.git)

Acknowledgements

This work was funded by a BBSRC Transformative Resources Development Fund award.

FRET data was provided by Prof Scott Blanchard.

References

1 Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. Improved Patch-Clamp Techniques for High-Resolution Current Recording from Cells and Cell-Free Membrane Patches. *Pflugers Archiv-European Journal of Physiology* **391**, 85-100 (1981).

2 Neher, E. & Sakmann, B. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature* **260**, 799-802, doi:10.1038/260799a0 (1976).

3 Hodgkin, A. L. & Huxley, A. F. A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve. *Journal of Physiology-London* **117**, 500-544 (1952).

4 Abdul Kadir, L., Stacey, M. & Barrett-Jolley, R. Emerging Roles of the Membrane Potential: Action Beyond the Action Potential. *Front Physiol* **9**, 1661, doi:10.3389/fphys.2018.01661 (2018).

5 Lehmann-Horn, F. & Jurkat-Rott, K. Voltage-gated ion channels and hereditary disease. *Physiological Reviews* **79**, 1317-1372 (1999).

6 Colquhoun, D. & Sigworth, F. in *Single-channel recording* 483-587 (Springer, 1995).

7 Qin, F., Auerbach, A. & Sachs, F. A direct optimization approach to hidden Markov modeling for single channel kinetics. *Biophysical Journal* **79**, 1915-1927 (2000).

8 O'Brien, F. *et al.* Enhanced activity of multiple TRIC-B channels: an endoplasmic reticulum/sarcoplasmic reticulum mechanism to boost counterion currents. *J Physiol* **597**, 2691-2705, doi:10.1113/JP277241 (2019).

9 Ha, T. Single-molecule fluorescence resonance energy transfer. *Methods* **25**, 78-86, doi:DOI 10.1006/meth.2001.1217 (2001).

10 Blanco, M. & Walter, N. G. Analysis of complex single-molecule FRET time trajectories. *Methods Enzymol* **472**, 153-178, doi:10.1016/S0076-6879(10)72011-5 (2010).

11 Juette, M. F. *et al.* Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat Methods* **13**, 341-344, doi:10.1038/nmeth.3769 (2016).

12 Colquhoun, D., Hatton, C. J. & Hawkes, A. G. The quality of maximum likelihood estimates of ion channel rate constants. *Journal of Physiology-London* **547**, 699-728, doi:10.1113/jphysiol.2002.034165 (2003).

13 Nicolai, C. & Sachs, F. Solving ion channel kinetics with the QuB software. *Biophysical Reviews and Letters* **8**, 191-211 (2013).

14 Sivilotti, L. & Colquhoun, D. In praise of single channel kinetics. *J Gen Physiol* **148**, 79-88, doi:10.1085/jgp.201611649 (2016).

15 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436-444, doi:10.1038/nature14539 (2015).

16 Angermueller, C., Parnamaa, T., Parts, L. & Stegle, O. Deep learning for computational biology. *Mol Syst Biol* **12**, 878, doi:10.15252/msb.20156651 (2016).

17 Krizhevsky, A., Sutskever, I. & Hinton, G. E. in *Advances in neural information processing systems.* 1097-1105.

18 Pratt, H. *et al.* Automatic Detection and Distinction of Retinal Vessel Bifurcations and Crossings in Colour Fundus Photography. *J Imaging* **4**, doi:UNSP 4

10.3390/jimaging4010004 (2018).

19 Al-Bander, B., Al-Nuaimy, W., Williams, B. M. & Zheng, Y. L. Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc. *Biomed Signal Proces* **40**, 91-101, doi:10.1016/j.bspc.2017.09.008 (2018).

20 Azizi, S. *et al.* Deep Recurrent Neural Networks for Prostate Cancer Detection: Analysis of Temporal Enhanced Ultrasound. *IEEE Trans Med Imaging* **37**, 2695-2703, doi:10.1109/TMI.2018.2849959 (2018).

21 Tang, D., Qin, B. & Liu, T. in *Proceedings of the 2015 conference on empirical methods in natural language processing.* 1422-1432.

22 Hochreiter, S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int J Uncertain Fuzz* **6**, 107-116, doi:Doi 10.1142/S0218488598000094 (1998).

23 Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Computation* **9**, 1735-1780, doi:DOI 10.1162/neco.1997.9.8.1735 (1997).

24 Graves, A., Mohamed, A.-r. & Hinton, G. in *2013 IEEE international conference on acoustics, speech and signal processing.* 6645-6649 (IEEE).

25 Sutskever, I., Vinyals, O. & Le, Q. V. Sequence to Sequence Learning with Neural Networks. *Advances in Neural Information Processing Systems 27 (Nips 2014)* **27** (2014).

26 Lipton, Z. C., Berkowitz, J. & Elkan, C. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019* (2015).

27 Mobasheri, A. *et al.* Characterization of a stretch-activated potassium channel in chondrocytes. *Journal of Cellular Physiology* **223**, 511-518, doi:Doi 10.1002/Jcp.22075 (2010).
28 Aldhous, P. Nobel prize. Patch clamp brings honour. *Nature* **353**, 487, doi:10.1038/353487a0 (1991).
29 Dunlop, J., Bowlby, M., Peri, R., Vasilyev, D. & Arias, R. High-throughput electrophysiology: an emerging paradigm for ion-channel screening and physiology. *Nat Rev Drug Discov* **7**, 358-368, doi:10.1038/nrd2552 (2008).
30 Yajuan, X., Xin, L. & Zhiyuan, L. A comparison of the performance and application differences between manual and automated patch-clamp techniques. *Curr Chem Genomics* **6**, 87-92, doi:10.2174/1875397301206010087 (2012).
31 Mukhtasimova, N., daCosta, C. J. & Sine, S. M. Improved resolution of single channel dwell times reveals mechanisms of binding, priming, and gating in muscle AChR. *J Gen Physiol* **148**, 43-63, doi:10.1085/jgp.201611584 (2016).
32 Sharp, A. A., O'Neil, M. B., Abbott, L. F. & Marder, E. Dynamic clamp: computer-generated conductances in real neurons. *J Neurophysiol* **69**, 992-995 (1993).
33 Lanchantin, J., Singh, R., Wang, B. L. & Qi, Y. J. Deep Motif Dashboard: Visualizing and Understanding Genomic Sequences Using Deep Neural Networks. *Biocomput-Pac Sym*, 254-265 (2017).
34 Kitrungrotsakul, T. *et al.* in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 1239-1243 (IEEE).
35 Davies, L. M., Purves, G. I., Barrett-Jolley, R. & Dart, C. Interaction with caveolin-1 modulates vascular ATP-sensitive potassium (K(ATP)) channel activity. *Journal of Physiology-London* **588**, 3254-3265, doi:10.1113/jphysiol.2010.194779 (2010).
36 O'Brien, F. & Barrett-Jolley, R. O'Brien, F. & Barrett-Jolley, R. (2018) CVS role of TRPV: from single channels to HRV assessment with Artificial Intelligence. *FASEB J* **32**, in press (2018).
37 Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical-Reactions. *J Phys Chem-Us* **81**, 2340-2361, doi:DOI 10.1021/j100540a008 (1977).
38 Feetham, C. H., Nunn, N., Lewis, R., Dart, C. & Barrett-Jolley, R. TRPV4 and KCa ion channels functionally couple as osmosensors in the paraventricular nucleus. *Br J Pharmacol* **172**, 1753-1768, doi:10.1111/bph.13023 (2015).
39 Story, M. & Congalton, R. G. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing* **52**, 397-399 (1986).
40 Barrett-Jolley, R., Comtois, A., Davies, N. W., Stanfield, P. R. & Standen, N. B. Effect of adenosine and intracellular GTP on K-ATP channels of mammalian skeletal muscle. *Journal of Membrane Biology* **152**, 111-116, doi:10.1007/s002329900090 (1996).