A Deep Learning Approach for the Classification of Neuronal Cell Types

Alessio P. Buccino*‡, Torbjørn V. Ness†, Gaute T. Einevoll*†, Gert Cauwenberghs‡, and Philipp D. Häfliger*.
*Center for Integrative Neuroplasticity (CINPLA), University of Oslo, Oslo, Norway
†Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway
‡Department of Bioengineering, University of California San Diego, La Jolla, CA, USA
Email: alessiob@ifi.uio.no, torbjorn.ness@nmbu.no, gaute.einevoll@nmbu.no, gert@ucsd.edu, hafliger@ifi.uio.no

***Abstract*—Classification of neurons from extracellular recordings is mainly limited to *putatively* excitatory or inhibitory units based on the spike shape and firing patterns. Narrow waveforms are considered to be fast spiking inhibitory neurons and broad waveforms excitatory neurons. The aim of this work is twofold. First, we intend to use the rich spatial information from high-density Multi-Electrode Arrays (MEAs) to make classification more robust; second, we hope to be able to classify sub-types of excitatory and inhibitory neurons. We first built, in simulation, a large dataset of action potentials from detailed neural models. Then, we extracted spike features from the simulated recordings on a high-density Multi-Electrode Array model. Finally, we used a Convolutional Neural Networks (CNN), to classify the different cell types. Compared with the ground truth data from the simulated dataset, the results show that this forward modelling/machine learning approach is very robust in recognizing excitatory and inhibitory spikes (accuracy $\geq$ 92.15%). Additionally, the approach can, to a certain extent, correctly classify different cell sub-types. As the detail and fidelity of neural models increase and high-density recordings become available, this approach could become a viable and robust alternative for classification of neural cell types from in-vivo extracellular recordings.**

I. Introduction

In neural tissue, excitatory and inhibitory neurons are organized in networks to allow neural circuits to perform processing of information [1]. Therefore, distinguishing between these two classes of cell types is essential to investigate their functions in neural circuits. With extracellular recordings, we are able to measure the activity of neurons as electrical potential deflections mainly due to ionic transmembrane currents. The largest signals are recorded when neurons fire action potentials (or spikes). Neurons are usually classified as excitatory or inhibitory units by comparing spike shapes [1], [2]: narrow waveforms are considered to be fast spiking inhibitory neurons and broad waveforms excitatory neurons.

Although differences in somatic current amplitudes and widths exist during spiking of excitatory and inhibitory neurons [1] (Fig. 1A), the extracellular action potential (EAP) is also affected by the dendritic transmembrane currents [3] (in Fig. 1B we show EAP at 15 μm away from the soma). When looking at simulated spike widths and amplitudes of a multitude of neurons located around a recording probe (Fig. 1C), the differences between excitatory and inhibitory neurons are not as clear as for the somatic currents. While the difference in amplitude is lost because the neuron location is unknown and the potential falls off depending on the recording distance [3], the spike width tends to increase with increasing distance from the recording site [4], yielding an extensive overlap between excitatory and inhibitory units. In this study, we propose a novel approach to classify neuron cell types from extracellular recordings; our method precisely categorize excitatory and inhibitory neurons and can further classify sub-types of inhibitory neurons.

The latest development in fabrication of advanced neural probes, such as high-density Multi-Electrode Arrays (MEAs) [5], provides rich spatial information, since the same spike is measured on multiple recording sites. We combine these devices with detailed neural models [6], which allow us to build large datasets of high-fidelity simulated recordings. Then, we apply powerful machine learning techniques, such as Convolutional Neural Networks (CNN), to classify excitatory and inhibitory neurons (*binary classification*), and to distinguish among 13 different morphological types (*m-type classification*). We also evaluate the effect of varying neurons alignment with respect to the recording MEA, rather than only shifting neurons in space, as in *Ruz et al.* [7].

This work is an extension of *Buccino et al.* [8], in which we used the same dataset and approach for neural localization. The article is organized as follows: in section II we describe

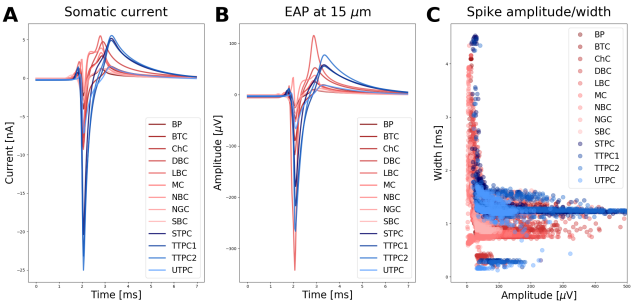

Fig. 1. (A) Somatic spike currents of 9 inhibitory neurons (red) and 4 excitatory neurons(blue). (B) Extracellular Action Potentials (EAP) 15 μmfrom the soma. There are differences with respect to (A), e.g. one of the inhibitory neuron's amplitude is larger than the excitatory ones, probably due to a dendrite close to the recording site. (C) Spike amplitude *versus* width of 13000 simulated neurons at different distances from the recording site. The excitarory (blue) and inhibitory (red) points greatly overlap in the center.

978-1-5386-3646-6/18/$31.00 ©2018 IEEE

the methods used for simulations and deep learning; in section III the results of the classification are presented; and in section IV we discuss this work and highlight potentials, limitations, and future directions.

II. Methods

A. Cell types and morphologies

The Neocortical Microcircuit Collaboration Portal [6], [9] has been chosen as source for the neuronal models that we used to generate simulated recordings. The models are from juvenile rat somatosensory cortex and we focused on neurons in layer 5 (L5). There are 13 types of morphologies (*m-type*) in L5: 9 types are inhibitory neurons (Bipolar Cells - BP, Bitufted Cells - BTC, Chandelier Cells - ChC, Double Boquet Cells - DBC, Large Basket Cells - LBC, Martinotti Cells - MC, Nest Basket Cells - NBC, Neuroglial Cells - NGC, Small Basket Cells - SBC) and 4 types are excitatory, i.e. the pyramidal cells (PC), (Small-Tufted PC - STPC, Thick-Tufted PC (type 1) - TTPC1, Thick-Tufted PC (type 2) - TTPC2, Untufted PC - UTPC). Inhibitory cells of the same *m-type* can show different electrophysiology firing properties (*e-type* - e.g. bAC stands for *burst accomodating*). However, since we use single spike waveforms to classify among neural types, we consider models of a unique *e-type* for each *m-type*. Table I summarizes the 13 neuron models involved in the study, available online at `https://bbp.epfl.ch/nmc-portal/welcome`.

B. Simulated Recordings

We used NEURON [10], within the LFPy Python package [11], to simulate the cell models. The 13 neuronal models are run separately: first, we apply a constant current to the soma so that the neuron fires 10 times in 1.2 s simulation period; second, spikes are clipped between t=-2 ms and t=5 ms, where t=0 is when the intracellular membrane voltage peak occurs; third, in LFPy, we used transmembrane currents computed in the previous steps to build the extracellular potential at the recording sites by summing the contribution of each segment using a line-source model (tissue conductivity σ=0.3 S/m - for further detail please refer to [7], [11]).

The probe recording sites are in a 2D configuration with an inter-electrode-distance of 15 μm, similarly to the MEA prototype described in *Schröder et al.* [5]. Instead of a 16x16 configuration, we used a 10x10 arrangement. The center of the MEA is located at the origin of the reference system, while the recording sites lie on the *yz*-plane. Neurons are located in the semispace having positive *x*-axis (the *x*-coordinate is the distance from the MEA). We generated 1000 extracellular action potential (EAP) recordings for each neuronal model, by randomly choosing 1 out of the 10 generated spikes and by placing the soma at random 3D locations within the following boundaries: $x \in$ [10 μm, 60 μm], $y \in$ [-75 μm, 75 μm], and $z \in$ [-75 μm, 75 μm] (the MEA covers the *yz* plane from around -75 μmto 75 μmin the *y* and *z* directions). In the calculation of the extracellular potential, LFPy automatically adjusts the position of neurites that are too close to a recording point and would, in turn, lead to an infinite electric potential [11]. Only sizeable spikes with a peak-to-peak amplitude greater than 30 μV on at least one recording site were included in the final dataset.

model name	*type*	*layer*	*m-type*	*e-type*	*cell_id*
L5_BP_bAC217_1	I	L5	BP	bAC217	1
L5_BTC_bAC217_1	I	L5	BTC	bAC217	1
L5_ChC_cACint209_1	I	L5	ChC	cACint209	1
L5_DBC_bAC217_1	I	L5	DBC	bAC217	1
L5_LBC_bAC217_1	I	L5	LBC	bAC217	1
L5_MC_bAC217_1	I	L5	MC	bAC217	1
L5_NBC_bAC217_1	I	L5	NBC	bAC217	1
L5_NGC_bNAC219_1	I	L5	NGC	bAC217	1
L5_SBC_bNAC219_1	I	L5	SBC	bAC217	1
L5_STPC_cADpyr232_1	E	L5	STPC	cADpyr232	1
L5_TTPC1_cADpyr232_1	E	L5	TTPC1	cADpyr232	1
L5_TTPC2_cADpyr232_1	E	L5	TTPC2	cADpyr232	1
L5_UTPC_cADpyr232_1	E	L5	UTPC	cADpyr232	1

TABLE I
CELL MODELS FROM NMC PORTAL: *model name* IS THE FULL MODEL NAME, *type* SHOWS EXCITATORY (E) AND INHIBITORY (I) CLASS, *layer* IS THE CORTICAL LAYER, *m-type* IS THE MORPHOLOGICAL TYPE, *e-type* IS THE ELECTRICAL TYPE, *cell_id* REFERS TO THE SPECIFIC CELL MODEL.

As anticipated in section I, we investigated different neuron-MEA alignments (or rotations). Some of the neurons that we used, such as pyramidal cells (PC) and bipolar cells (BP), show a selected orientation of their dendritic arbor. As all neurites, including dendrites, actively contribute to the extracellular potential [3], different alignments might affect the classification performance. For this reason, we generated four *rotational* datasets:

- **Norot:** neurons are only translated around the MEA and the pyramidal cells' apical dendrite is aligned to the *z*-axis
- **Zrot:** neurons are shifted around the MEA and randomly rotated around the *z*-axis (PC apical dendrite is still along the *z*-axis)
- **Xrot:** neurons are shifted around the MEA and rotated randomly around the *x*-axis
- **3drot:** neurons are shifted around the MEA and rotated randomly around all axes.

Although in previous studies simulated datesets were generated by shifting neuron models around MEAs [7], prediction models can be sensitive to different alignments and, therefore, it is interesting to observe the difference classification performance depending on rotation. Fig. 2 shows how a pyramidal cell's (TTPC1) orientation could change in each of the four datasets (A - *Norot*, B - *Zrot*, C - *Xrot*, D - *3drot*).

From the simulated spikes we can extract significant features from EAPs and classify the variety of neural morphologies depending on their spike *signatures*.

C. Deep Learning Approach

In order to exploit to the rich spatial information in the high-density MEA recordings we used a Convolutional Neural Network (CNN) as machine learning algorithm.

Spike widths and amplitudes served as input for the network and they were computed on each recording site as follows: the amplitude is the absolute value of the difference between the negative peak and the positive peak that follows it (initial positive peaks due to capacitive outward currents

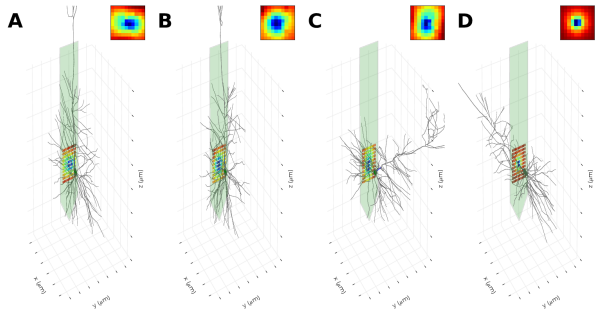

Fig. 2. 3D representation of the possible relative alignments of a pyramidal cell (TTPC1) with soma at (50 μm, 0, 0) with respect to the MEA in the Norot (A), Zrot (B), Xrot (C), and 3drot (D) datasets. The pseudocolors of the recording sites and the upper right axes are the negative peak *images*. The *images* of Norot, Zrot, and Xrot datasets have quite similar shapes and extensions, while the in the 3drot case the peak is more circumscribed mainly due to the different inclination of the soma.

cell type	*Norot*	*Zrot*	*Xrot*	*3drot*
BP	100%	100%	100%	100%
BTC	100%	98.25%	97.85%	97.85%
ChC	100%	98.99%	98.99%	81.82%
DBC	100%	99.04%	95.19%	85.58%
LBC	98.96%	96.875%	100%	94.79%
MC	100%	100%	100%	100%
NBC	100%	100%	100%	99.1%
NGC	100%	99.06%	100%	99.06%
SBC	100%	100%	100%	100%
STPC	100%	100%	100%	94.25%
TTPC1	100%	100%	100%	90.38%
TTPC2	100%	98.99%	98.99%	95.96%
UTPC	97.14%	89.52%	68.57%	60.95%
Average	**99.69%**	**98.54%**	**95.92%**	**92.15%**

TABLE II
BINARY CLASSIFICATION ACCURACY OF EACH CELL TYPE FOR THE FOUR DATASETS. THE LAST ROW SHOWS AVERAGE VALUES. A CORRECT CLASSIFICATION HERE MEANS THAT INHIBITORY/EXCITATORY CELLS ARE CLASSIFIED AS INHIBITORY/EXCITATORY, RESPECTIVELY.

are discarded); the width is computed as the difference in time between the identified peaks. When the amplitude of a certain recording site is below a threshold of 5 μV, it is set to 0 and the width is set to 7 ms, which is the entire spike window duration.

Tensorflow [12] was used to train the CNNs with the following configuration. The 10x10 amplitude and width *images* (see Fig. 2) are input to a 32-deep convolutional ReLU layer which filters the input image with 3x3 kernels with stride=1. Max pooling is then applied and the *image* is shrunken to a 5x5 footprint. Another 64-deep convolutional ReLU layer applies 3x3 kernels and max pooling reduces the output *image* features to a 3x3 size. The 3x3 features are input for a fully connected layer with 1024 artificial neurons and 2 output nodes for the *binary classification* and 13 for the *m-type classification*.

The dropout method is implemented to avoid overfitting [12], with a dropout rate of 0.7. Softmax cross entropy is minimized with the *Adam* optimizer during training, which is run for 5000 epochs, each time sampling 1000 observations from the dataset.

III. RESULTS

A. Binary Classification

The first question we investigated is whether we could untangle the overlap in excitatory and inhibitory amplitudes and widths by means of deep learning techniques (shown in Fig. 1C). We also evaluated how the *binary classification* performance is affected by different neuron-MEA alignments. Table II shows the average classification accuracy for the different datasets (last row) and for each cell morphology. The same results are also visualized in color-code in Fig. 3.

The accuracy, as expected, degrades from *Norot* to *3drot* datasets because neurons are rotated with more degrees of freedom; nevertheless on average the accuracy remains relatively high (above 92%) in all cases. A closer examination of this result reveal that the main reason for the drop in classification accuracy was misclassification of the pyramidal UTPC. Interestingly, from Fig. 1A the spike shape of UTPC (lighter blue) does not *cluster* with the other pyramidal cells (blue shades), but the shape is more similar to inhibitory ones (red colors). This peculiar shape might contribute to its high misclassification.

B. M-type Classification

The performance of *m-type classification*, i.e. recognition among the 13 different cell morphology, is shown in Fig. 4, in which each plot is related to a different rotation dataset. Table III summarizes the distribution of the accuracy obtained for the different cell types (i.e. the 13 values on the diagonals of the matrices of Fig. 4). While the *Norot* dataset is almost perfectly classified (mean accuracy of 98.96%), the accuracy decreases for the other datasets, as observed in the *binary classification*. For *Zrot* and *Xrot* the average accuracy is 93.31% and 87.56%, respectively, but for *3drot* it is 57.49% (for 13 classes the chance level is $^1/_{13}$ *x* 100 $\simeq$ 7.7%). Focusing on the *3drot* dataset (Fig. 4D) and observing the different cell types separately, it is clear that, depsite misclassification, inhibitory neurons are mainly labeled as inhibitory (red labels) and excitatory ones as excitatory (blue labels). Among inhibitory cells, SBC accuracy is 38.78%, but 41% of the times it is recognized as BP. Likewise, BP accuracy is 71.03%, and it is labeled as SBC 23% of the times. Similarly, MC (accuracy=50.55%) is classified as BTC in 30.77% of the cases and BTC (accuracy=56.99%) is labeled as MC

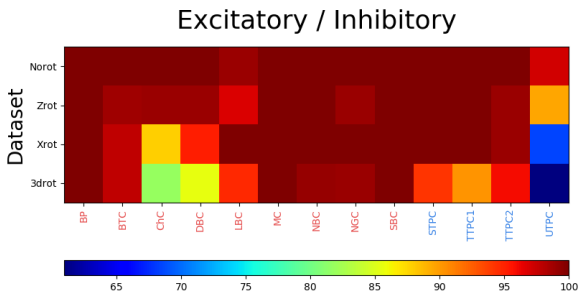

Fig. 3. Accuracy for binary classification divided by cell types. The drop in average accuracy is mainly due to misclassification of UTPC (Note that colorbar starts at around 60%)

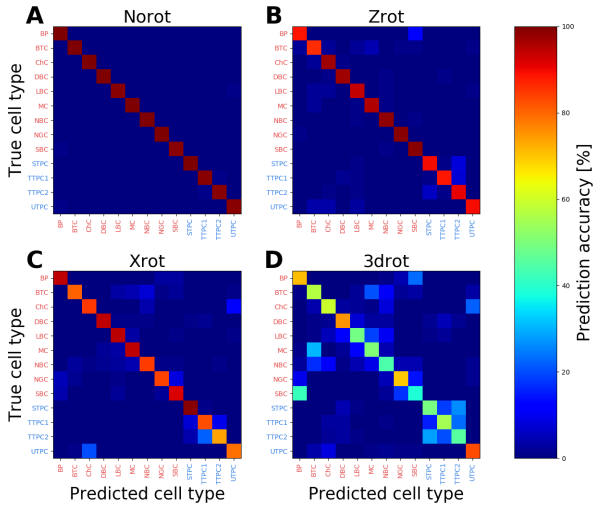

Fig. 4. Accuracy for *m-type classification* divided by cell types and dataset (A - *Norot*, B - *Zrot*, C - *Xrot*, D - *3drot*). Each row i shows the color-coded percent of cell i classified as cell j. A perfect classification would result in 100% values on the diagonal only. Inhibitory cells are labeled in red, excitatory in blue.

in 20.43% of observations. NBC and LBC misclassification is more uniform among other inhibitory units, while the sole excitatory/inhibitory confusion appears to be between ChC and UTPC cells (ChC as UTPC=22.22%, UTPC as ChC=8.57%). Excluding ChC and UTPC, inhibitory cells are classified as inhibitory in $\geq$91.34%, and excitatory as excitatory in $\geq$93.27% of the cases.

IV. Discussions and Conclusion

The present approach combines detailed biophysical modeling and powerful machine learning to classify neurons from MEA simulated recordings. The results show that *binary classification* between excitatory and inhibitory types is very robust, despite the overlap between excitatory and inhibitory cells regarding spike amplitudes and widths (Fig. 1C). With the same method, we were also able to classify 13 different cell types and we showed that the accuracy is dependent on the alignment between neurons and the MEA.

Our findings are only based on simulations. Validation with real data is a required and important step. In addition, electrophysiology will need to be matched with imaging techniques to reconstruct neuron morphology [6], needed to validate the model predictions. In this work we did not include any noise in the simulated recordings, with the rationale that sorted spikes can be *cleaned* by applying spike-triggered-averaging.

We are currently using a similar forward modeling/machine learning paradigm to localize neuronal somata in 3D from EAP traces [8]. This simulation-based approach can also give important insight to drive the design of new probes, for example in terms of number, density, size, and location of recording sites. Moreover, the performance of the machine learning algorithm could be improved by building case-specific datasets that better match the real recordings, depending on animal type, target layer(s), insertion angle, MEA characteristics and so on.

We believe that high-density next-generation MEAs will allow us to perform a *functional image* of the neural tissue surrounding the probe. Moreover, information about neural types and location could be used to optimize stimulation paradigms [13] in precisely controlled closed-loop experiments, resulting in a device/tissue interaction down to the single neuron resolution.

Dataset	*mean*	*median*	*max*	*min*	*sd*
Norot	99.62%	100%	98.96%	100%	0.48%
Zrot	93.31%	93.75%	86.02%	99.05%	4.41%
Xrot	87.56%	84.85%	73.74%	98.85%	7.34%
3drot	57.49%	54.81%	38.78%	82.86%	12.88%

TABLE III
Performance of m-type classification. Mean, median, max, min, and sd are computed on the 13 cell types.

Acknowledgments

Alessio P. Buccino is a doctoral fellow in the Simula-UCSD-University of Oslo Research and PhD training (SUURPh) program, an international collaboration in computational biology and medicine funded by the Norwegian Ministry of Education and Research.

References

[1] A. Peyrache, N. Dehghani, E. N. Eskandar, , *et al.*, "Spatiotemporal dynamics of neocortical excitation and inhibition during human sleep," *Proceedings of the National Academy of Sciences*, vol. 109, no. 5, pp. 1731–1736, 2012.

[2] P. Barthó, H. Hirase, L. Monconduit, , *et al.*, "Characterization of neocortical principal cells and interneurons by network interactions and extracellular features," *Journal of neurophysiology*, vol. 92, no. 1, pp. 600–608, 2004.

[3] K. H. Pettersen and G. T. Einevoll, "Amplitude variability and extracellular low-pass filtering of neuronal spikes," *Biophysical Journal*, vol. 94, no. 3, pp. 784–802, 2008.

[4] C. Gold, D. A. Henze, C. Koch, *et al.*, "On the origin of the extracellular action potential waveform: a modeling study," *Journal of neurophysiology*, vol. 95, no. 5, pp. 3113–3128, 2006.

[5] S. Schröder, C. Cecchetto, S. Keil, *et al.*, "CMOS-compatible purely capacitive interfaces for high-density in-vivo recording from neural tissue," in *Biomedical Circuits and Systems Conference (BioCAS), 2015 IEEE*, pp. 1–4, 2015.

[6] H. Markram, E. Muller, S. Ramaswamy, *et al.*, "Reconstruction and simulation of neocortical microcircuitry," *Cell*, vol. 163, no. 2, pp. 456–492, 2015.

[7] I. D. Ruz and S. R. Schultz, "Localising and classifying neurons from high density MEA recordings," *Journal of Neuroscience Methods*, vol. 233, pp. 115–128, 2014.

[8] A. P. Buccino, T. V. Ness, G. T. Einevoll, G. Cauwenberghs, and P. D. Häfliger, "Localizing neuronal somata from multi-electrode array in-vivo recordings using deep learning," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, pp. 974–977, IEEE, 2017.

[9] S. Ramaswamy, J.-D. Courcol, M. Abdellah, *et al.*, "The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex," *Frontiers in neural circuits*, vol. 9, 2015.

[10] N. T. Carnevale and M. L. Hines, *The NEURON book*. Cambridge University Press, 2006.

[11] H. Lindén, E. Hagen, S. Leski, *et al.*, "LFPy: a tool for biophysical simulation of extracellular potentials generated by detailed model neurons," *Frontiers in Neuroinformatics*, vol. 7, p. 41, 2014.

[12] M. Abadi, A. Agarwal, P. Barham, *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software available from tensorflow.org.

[13] A. P. Buccino, T. Stöber, S. Næss, *et al.*, "Extracellular single neuron stimulation with high-density multi-electrode array," in *2016 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pp. 520–523, Oct 2016.