INNOVATIVE METHODOLOGY | *Neural Circuits*

Combining biophysical modeling and deep learning for multielectrode array neuron localization and classification

Alessio P. Buccino,[1,2]* Michael Kordovan,[3,4]* Torbjørn V. Ness,[5] Benjamin Merkt,[3,4] Philipp D. Häfliger,[1] Marianne Fyhn,[1] Gert Cauwenberghs,[2] Stefan Rotter,[3,4]* and Gaute T. Einevoll[1,5]*

[1]*Center for Integrative Neuroplasticity (CINPLA), Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway;* [2]*Department of Bioengineering, University of California, San Diego, California;* [3]*Bernstein Center Freiburg, Freiburg, Germany;* [4]*Faculty of Biology, University of Freiburg, Freiburg, Germany; and* [5]*Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway*

Buccino AP, Kordovan M, Ness TV, Merkt B, Häfliger PD, Fyhn M, Cauwenberghs G, Rotter S, Einevoll GT. Combining biophysical modeling and deep learning for multielectrode array neuron localization and classification. *J Neurophysiol* 120: 1212–1232, 2018. First published May 30, 2018; doi:10.1152/jn.00210.2018.—Neural circuits typically consist of many different types of neurons, and one faces a challenge in disentangling their individual contributions in measured neural activity. Classification of cells into inhibitory and excitatory neurons and localization of neurons on the basis of extracellular recordings are frequently employed procedures. Current approaches, however, need a lot of human intervention, which makes them slow, biased, and unreliable. In light of recent advances in deep learning techniques and exploiting the availability of neuron models with quasi-realistic three-dimensional morphology and physiological properties, we present a framework for automatized and objective classification and localization of cells based on the spatiotemporal profiles of the extracellular action potentials recorded by multielectrode arrays. We train convolutional neural networks on simulated signals from a large set of cell models and show that our framework can predict the position of neurons with high accuracy, more precisely than current state-of-the-art methods. Our method is also able to classify whether a neuron is excitatory or inhibitory with very high accuracy, substantially improving on commonly used clustering techniques. Furthermore, our new method seems to have the potential to separate certain subtypes of excitatory and inhibitory neurons. The possibility of automatically localizing and classifying all neurons recorded with large high-density extracellular electrodes contributes to a more accurate and more reliable mapping of neural circuits.

NEW & NOTEWORTHY We propose a novel approach to localize and classify neurons from their extracellularly recorded action potentials with a combination of biophysically detailed neuron models and deep learning techniques. Applied to simulated data, this new combination of forward modeling and machine learning yields higher performance compared with state-of-the-art localization and classification methods.

convolutional neural networks; deep learning; extracellular action potentials; multielectrode arrays; neuron localization and classification

INTRODUCTION

The neural circuits of the brain involve the interplay of many different types of neurons. On the most superficial level, neurons are classified by their effect on the neurons they project to as either excitatory or inhibitory. Extracellular recordings enable us to measure the activity of neurons as electrical potential deflections mainly due to ionic transmembrane currents. In recent years, many groups have been developing high-density multielectrode arrays (MEAs) for in vitro and in vivo applications, which allow measurements of neuronal activity with high spatiotemporal resolution (Berdondini et al. 2014; Müller et al. 2015; Neto et al. 2016; Schröder et al. 2015; Welkenhuysen et al. 2016). These probes provide something close to a functional electrical imaging (Vassanelli 2014) of the neural activity, and this high information density can potentially be exploited to obtain a better understanding of the neural circuits under study. Specifically, it might be possible to extract information that could be used to localize the individual neurons and to classify them into their respective cell types. The latest developments in manufacturing of high-density neural probes call for novel analysis methods to exploit the richness and detail in the recordings.

Neural localization from extracellular action potentials (EAPs) recorded on a MEA is an ill-posed problem, since solutions might not be unique for complex neural morphologies. Current approaches for localization assume simple neuronal models to facilitate the inverse problem and make the solution unique. Examples of models used in previous studies are monopole current-source models (Blanche et al. 2005; Chelaru and Jog 2005; Kubo et al. 2008), dipole current-source models (Blanche et al. 2005; Mechler et al. 2011; Mechler and Victor 2012), as well as line-source models (Somogyvári et al. 2012). More recently, Delgado Ruz and Schultz (2014) used a modified ball-and-stick model (Pettersen and Einevoll 2008) to predict somatic locations. In these approaches, the soma posi-

tion is estimated by minimizing the error between the electrical potential on the MEA sites predicted by the chosen model and the recorded potential. However, it is experimentally challenging to measure the correct position of the soma (Neto et al. 2016); therefore, detailed computational neuronal models are usually used and treated as simulated ground truth to evaluate the accuracy of the localization methods (Delgado Ruz and Schultz 2014; Somogyvári et al. 2005, 2012). The main limitations regarding neuron localization are that the models chosen to solve the inverse problem are often too simple to grasp complex spike waveforms (e.g., monopolar or bipolar current-source models) or are tuned to specific cell types (ball-and-stick model for pyramidal morphology).

Regarding classification of neurons, unsupervised clustering techniques are commonly applied to differentiate EAP shapes (Barthó et al. 2004; McCormick et al. 1985; Peyrache et al. 2012): narrow waveforms are considered to be fast-spiking inhibitory neurons and broad waveforms excitatory neurons. Also in this case, it is experimentally challenging, especially in vivo, to validate the classification methods. One approach is to measure a multitude of neurons and find putative monosynaptic connections based on the shape of spike train cross-correlograms. However, the rate of recorded monosynaptic connections is usually very low (~0.2%; Barthó et al. 2004; Peyrache et al. 2012), resulting in a small number of observations useful for validation. In neural classification, the complexity of spike shapes across the multiple recording sites is usually compressed by extracting spike widths (such as peak-to-peak and full-width half-maximum widths; Barthó et al. 2004; Peyrache et al. 2012) only from the electrode with the highest recorded amplitude.

In this study, we apply a powerful machine learning technique, namely, convolutional neural networks (CNNs), to classify excitatory and inhibitory neurons and to localize their somata from simulated EAPs. This approach—being a supervised learning algorithm—demands for a large amount of labeled data, in this case EAPs in combination with soma position and cell type of the neuron evoking the EAPs. The proposed method is schematically depicted in Fig. 1. First, compartmental cell simulations are performed (Fig. 1A) that yield EAP data sets with known simulated ground truth (forward modeling) (Fig. 1B). Relying on the simulations, CNNs are trained (Fig. 1C) on these data sets to predict position and cell type (Fig. 1D) of the neuron generating the simulated EAP. If the method is applied to experimental data (Fig. 1E), a spike-triggered average EAP (average waveform) serves as input to a CNN previously trained on simulated data to predict soma position and cell type. In addition to binary classification, we attempt to distinguish 11 different morphological types (m-type classification). The performance of the CNNs depending on different characteristics extracted from the EAP, different MEA designs, and different CNN configurations is explored. Finally, we evaluate the effect of varying neuron alignments with respect to the recording MEA. To put our approach into context, we compare its performance with established methods of cell localization and classification.

CNNs perform supervised machine learning and require a large data set to be trained. It would be experimentally challenging, if not unfeasible, to gather the required number of recordings of exact known position (used for localization) and morphology (used for classification) information. Therefore, we rely on detailed compartmental cell models to provide detailed simulated recordings and simulated ground-truth information. Forward biophysical modeling of extracellular potentials has been developed and refined over the last 30 years (Gold et al. 2006; Holt and Koch 1999; Lindén et al. 2014;
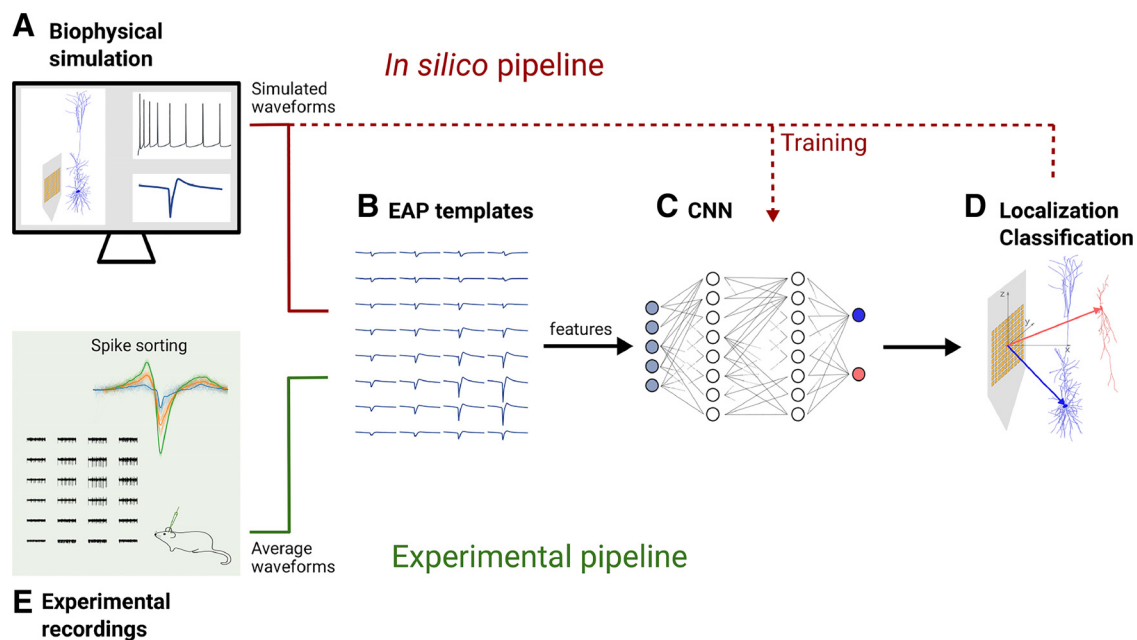

Fig. 1. Graphical representation of the presented method. Red arrows show our approach for training (dashed lines indicate error backpropagation) and validating the convolutional neural network (CNN) on simulated data. Green arrow shows how the method is applied within an experimental pipeline. Starting with the red path, biophysical simulations (A) are used to generate extracellular action potential (EAP) templates (B), from which features (e.g., amplitude and width; see *Classification and Localization Features*) are extracted and fed to a CNN (C) to localize and classify excitatory (blue) and inhibitory (red) neurons (D). When applied to experimental data (green arrow), recordings are first spike sorted (E), then features are extracted from spike-triggered average waveforms (B), and finally the CNN trained on simulated data (C) is used to localize and classify spike-sorted neurons (D).

Pettersen and Einevoll 2008; Rall and Shepherd 1968), and it is a still growing field. Therefore, we expect that computer simulations will become closer to real recordings as the detail and fidelity of neural models increase.

The aim of this study was to investigate the capability and feasibility of the proposed deep learning approach and to explore the large parameter space to guide future developments and experiments. At this stage, the method is a proof of concept, as we used some simplifications along the way. However, valuable information can be obtained and used in the future. Simulation software scripts are available at https://github.com/CINPLA/NeuroCNN.

MATERIALS AND METHODS

Cell Models

We first generated simulated spike recordings from various neuronal cell types (Fig. 1*A*). The neuronal models have been adopted from the Neocortical Microcircuit Collaboration (NMC) Portal (Markram et al. 2015; Ramaswamy et al. 2015, https://bbp.epfl.ch/nmc-portal), a database containing cell models from juvenile rat somatosensory cortex. We focused on neurons in layer (L)5, but the methods described can be applied to a larger variety of neuronal models. The cell models were used unmodified and are based on experimentally reconstructed morphologies, with up to 13 different types of active ion channels, depending on the cell type. The dynamics of these active ion channels have been fitted to reproduce specific features of somatic current-injection experiments, such as the amplitude and width of the evoked action potentials, the depth and timing of the following afterhyperpolarization, the mean interspike interval, spiking adaptations, and the amplitude of the backpropagating action potential at different positions along the apical dendrite of pyramidal cells (PC) (Markram et al. 2015). Importantly, this means that the source of the cell type-specific variability that we observe in, e.g., the spike widths is grounded in experimental data. For a more detailed description of the cell models the reader is referred to Markram et al. (2015). Note also that each of the cell models used can be explored interactively online through the NMC Portal (Ramaswamy et al. 2015, https://bbp.epfl.ch/nmc-portal/microcircuit#/layer/L5).

In principle, the data set contains 13 types of morphologies (m-types) in L5, in accordance with the NMC Portal. However, because of the limited data variety in the case of bipolar and neuroglial cells (BP and NGC), we excluded them from the analyses unless elsewhere specified. APPENDIX A describes the data set in detail and explains how we manipulated the original data set to extract unbiased sub-data sets for training and validating the results. A single-cell morphology of each m-type is displayed in Fig. 2, divided into inhibitory and excitatory neurons. Axons are not depicted because only their initial tract is included in the simulations.

Simulated Recordings

Extracellular action potential computation. Each of the multicompartment neuronal models was simulated separately, and extracellular potentials were calculated in two steps. First, transmembrane currents were computed by solving the cable equation (Holt and Koch 1999) with LFPy (Lindén et al. 2014) running on NEURON 7.4 (Carnevale and Hines 2006; Hines et al. 2009). A constant depolarizing current was applied to the soma to get the neuron firing at least 10 times (and not more than 30 times) in a simulation period of 1.2 s. Multiple spikes were simulated to account for spike-to-spike variations due to electrophysiological properties (e-types). All transmembrane currents for each compartment were stored within a time window $t = -2$ ms and $t = 5$ ms, where $t = 0$ is the time of the intracellular membrane voltage peak considered as spike time. Simulations were run with a time resolution of d$t = 2^{-5}$ ms, i.e., with a sampling frequency of 32 kHz, so that a single spike window of 7-ms duration (2 ms + 5 ms) consists of 224 samples.

Second, transmembrane currents were used within LFPy to calculate the extracellular potential at the recording site. Each transmembrane current, including the somatic one, was distributed over a line source with the length of its corresponding neural segment. With quasi-static approximation (Nunez and Srinivasan 2006) and assuming a homogeneous and isotropic neural tissue with conductivity $\sigma = 0.3$ S/m (Goto et al. 2010), the contribution of each compartment i at position r_i with transmembrane current $I_i(t)$ to the electric potential on an electrode at position r_j reads (Holt and Koch 1999; Lindén et al. 2014; Pettersen and Einevoll 2008)

$$\phi_i(r_j, t) = \frac{1}{4\pi\sigma} I_i(t) \int \frac{\mathrm{d}r_i}{\|r_j - r_i\|} \qquad (1)$$
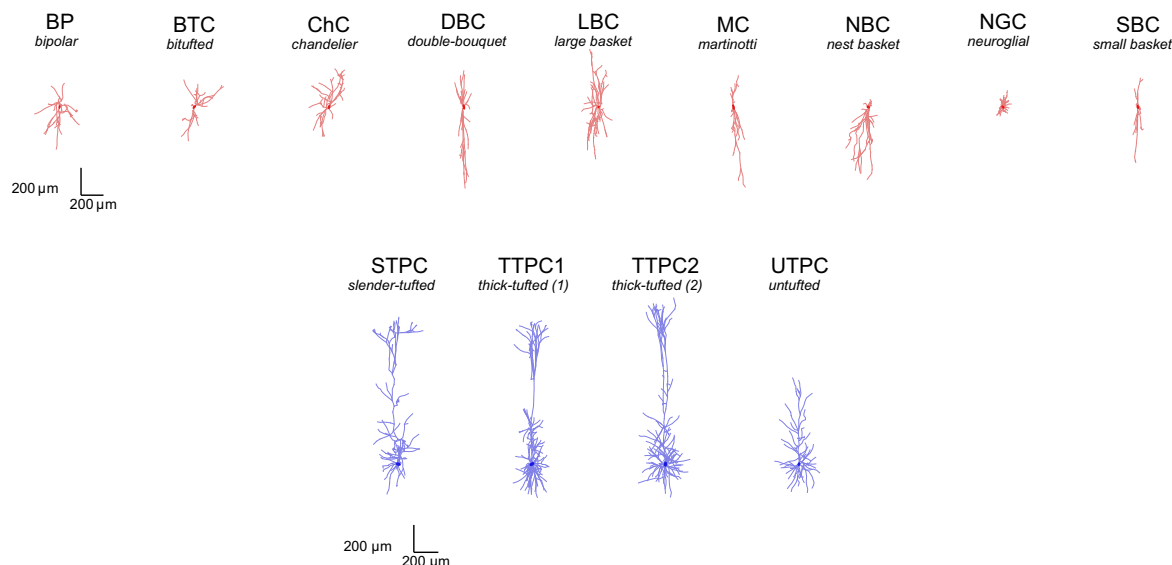
Fig. 2. One representative cell for each different morphology type in layer 5 [data from Neocortical Microcircuit Collaboration Portal (Ramaswamy et al. 2015)]. *Top*: 9 inhibitory cells (red). *Bottom*: 4 excitatory cells (blue). Dendrites are colored in lighter shades, and the soma is indicated with a darker circle. The same red/blue convention is used throughout article. For all cells C stands for cell, and for excitatory cells the P represents pyramidal.

The simulated EAP was obtained by summing up the contributions of all compartments. For each recording site, the electric potential was computed at a single point corresponding to the center of the recording electrode. These EAPs are associated with the templates in Fig. 1*B*.

Only spikes generating a notable waveform with a peak-to-peak amplitude exceeding 30 μV on at least one of the electrodes were included in the data set. The detection threshold of 30 μV was chosen to mimic experimental settings, where noise in the recordings due to equipment electronics and background neural signals does not allow detection of low-amplitude action potentials.

The coordinate system was fixed in reference to the MEA plane. Each recording site (different MEA designs are explained in *MEA designs*) lay within the *yz*-plane, and neuron somata were located within the semispace of the positive *x*-axis (the *x*-coordinate is thus the distance from the MEA). For each neuron, 1,000 EAP recordings above the detection threshold were simulated by randomly choosing one of the generated spikes and by placing the soma at random locations with distances *x* between 10 μm and 80 μm. The *y* and *z* boundaries were different for each MEA, and a neuron could exceed the *y* and/or *z* boundary of the MEA by a maximum of 30 μm. For the EAP, we considered the contributions of all dendritic compartments, including those crossing the MEA. This was done to force the sum of transmembrane currents to be zero (Pettersen and Einevoll 2008). In other words, we did not clip neurites entering the probe, but we made sure that their position did not coincide with a recording site within LFPy.

MEA designs. We investigated the performance of seven different MEA probes. Five of these were based on the prototype described in Schröder et al. (2015), in which the recording sites are arranged in a 16 $\times$ 16 matrix on a penetrating shank. Instead of considering a single interelectrode distance, we investigated five different pitches (i.e., distance between the centers of adjacent electrodes), namely, 10, 15, 20, 25, and 30 μm. The probe models were built in a way that they roughly covered the same area on the shank.

Hence, we ended up with the following list of squared probes: SqMEA-15-10: squared MEA with 15 $\times$ 15 electrodes and 10-μm pitch; SqMEA-10-15: squared MEA with 10 $\times$ 10 electrodes and 15-μm pitch; SqMEA-7-20: squared MEA with 7 $\times$ 7 electrodes and 20-μm pitch; SqMEA-6-25: squared MEA with 6 $\times$ 6 electrodes and 25-μm pitch; and SqMEA-5-30: squared MEA with 5 $\times$ 5 electrodes and 30-μm pitch.

Moreover, we simulated recordings on the Poly3-25s probe (Neuronexus Technologies), which has 32 electrodes in three columns with a hexagonal arrangement, a *y*-pitch of 18 μm, and a *z*-pitch of 22 μm. Another probe becoming popular is the NeuroPixels probe (Jun et al. 2017), with recording sites arranged in a checkerboard pattern with a *y*-pitch of 32 μm and a *z*-pitch of 20 μm. Although the probe has 384 recording channels, for convenience we decided to trim it to 128 channels. Finally, we constructed a model of the NeuroSeeker probe (http://www.neuroseeker.eu; used in Neto et al. 2016), a MEA with 128 recording sites arranged in a 4 $\times$ 32 configuration and a regular interelectrode distance of 22.5 μm. Figure 3 shows all the probes in the *yz*-plane.

The CNNs we used required a rectangular shape of the input data. The two dimensions of the data array correspond to the number of electrode sites N_y and N_z in *y*- and *z*-directions, respectively. Therefore, we cut the Neuronexus-32 MEA probe to a Neuronexus-30, which is shown in the fourth position from the left in Fig. 3.

Neuron-MEA alignment. We investigated different neuron-MEA alignments (or rotations) of neurons. Some neurons, such as pyramidal cells (PC) or bipolar cells (BP), have morphologies that follow a specific orientation (see Fig. 2) that might affect the classification and localization performance. For this reason, we generated three rotational data sets:

Norot: The orientation of the cells (e.g., the apical dendrite of PCs) was along the *z*-axis (same orientation as in Delgado Ruz and Schultz 2014 and Somogyvári et al. 2012).

Physrot: Neurons with a preferential orientation were randomly rotated such that after rotation their axis from white matter toward the pia pointed into a cone around the *z*-direction with an opening angle of 15° (the puncture point on the unit sphere is uniformly distributed in this spherical cap). Neurons without a preferential orientation were rotated randomly in the three-dimensional (3D) space. We considered all
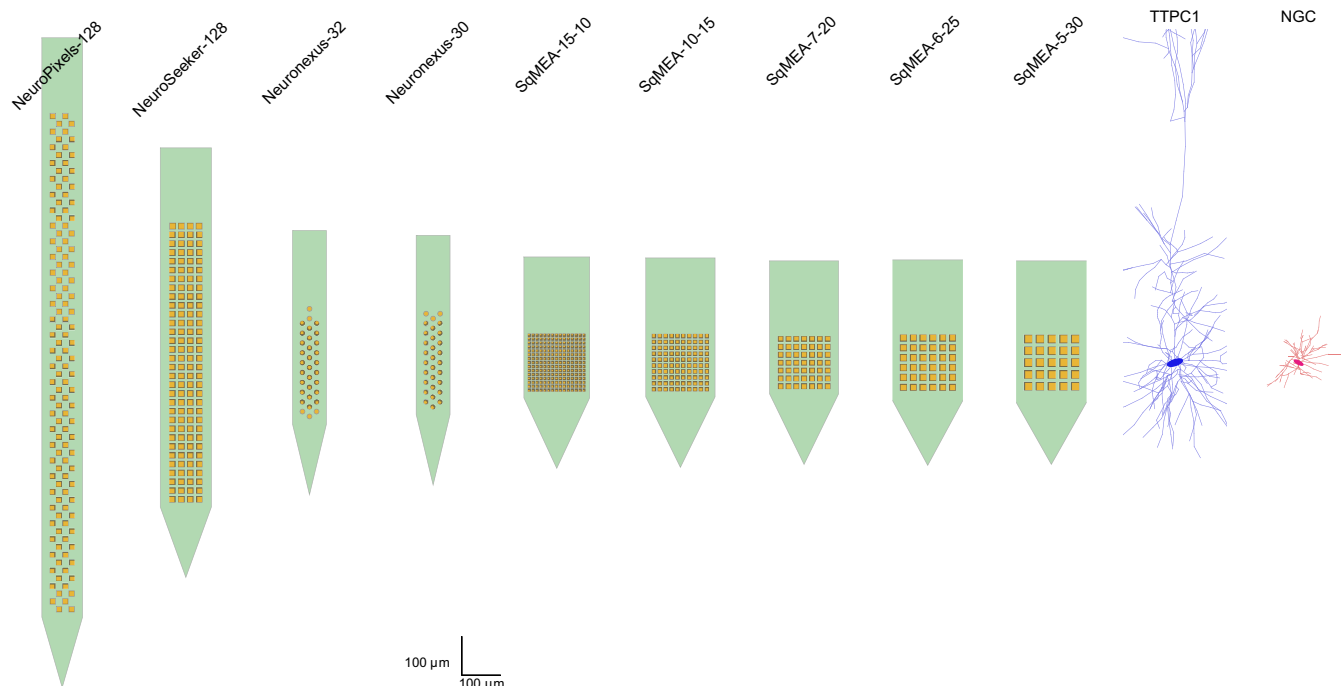
Fig. 3. Multielectrode array models used in the study. *Right*: plots show an excitatory cell [thick-tufted pyramidal cell (PC) with late bifurcating apical tuft (TTPC1)] and an inhibitory cell [neuroglial cell (NGC)] to compare probe and neuron sizes. PCs are on average much larger than inhibitory cells, and only a portion of the neuron is located directly in front of the probe (the apical dendrite is not fully shown, and it can be seen in Fig. 2).

Fig. 4. Neuron-multielectrode array (MEA) alignments: in each panel we show a sample orientation of a pyramidal cell [thick-tufted pyramidal cell with late bifurcating apical tuft (TTPC1)] placed at a fixed position of (50 μm, 0, 0). The MEA is the SqMEA-10-15, and the colors of each recording site (displayed as an image at *top left* of each plot) show the qualitative sodium peak image (Na image, explained in *Classification and Localization Features*). *A*: Norot: the pyramidal cell main axis is aligned to the *z*-direction. *B*: Physrot: the TTPC1 is rotated around the *z*-axis, and the apical dendrite is tilted at maximum 15° inside a cone around the *z*-axis. *C*: 3drot: the neuron can be rotated along all axes and might end up with its apical dendrite entering the MEA probe, as depicted in the plot. While the Na image is similar for *A* and *B*, it can become quite different in *C*, when the neuron is 3D-rotated.
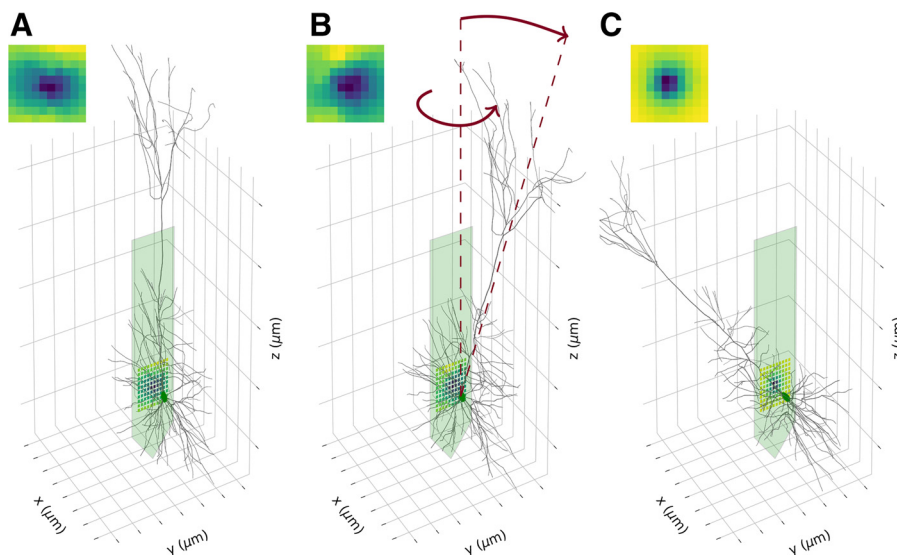
neurons apart from nest basket cells (NBC), small basket cells (SBC), and NGC to have a preferential orientation (DeFelipe et al. 2006; Markram et al. 2004; Overstreet-Wadiche and McBain 2015; Wang et al. 2002, 2004; Woodruff and Yuste 2008). NBC, SBC, and NGC were assumed to have no preferential axis.

3drot: Neurons were rotated randomly around all axes.

The soma positions corresponded to the intersection point of rotation axes and were shifted randomly inside the observation volume in all cases. Figure 4 displays a sample orientation with respect to the MEA of a PC [thick-tufted PC with late bifurcating apical tuft (TTPC1)] in each of the three data sets, Norot (Fig. 4*A*), Physrot (Fig. 4*B*), and 3drot (Fig. 4*C*).

Classification and Localization Features

We extracted features from the EAPs as input variables to a CNN for training. Since classification and localization of neurons from extracellular recordings are quite different tasks, we used different sets of features from the simulated spikes. The pipeline for extracting feature images is described in Fig. 5. First, neurons with known cell type and position were simulated and the spike traces on the MEA probe were obtained. Then, for each spike, a set of features was computed and these features were rearranged in a 2D shape according to the MEA arrangement, i.e., the feature image. In the following paragraphs N, N_y, and N_z are the total number of electrodes, the number of electrodes in *y*-direction, and the number of electrodes in *z*-direction, respectively.

Localization features. The goal of localization is to estimate the soma position with respect to the probe. Therefore, we considered only the EAP negative peak and the positive peak time points, during which negative and positive transmembrane currents are larger in proximity of the soma (Delgado Ruz and Schultz 2014; Gold et al. 2007; Somogyvári et al. 2005, 2012). For simplicity, we refer to the EAP negative peak as Na peak because, close to the soma, it is mainly attributed to the sodium currents flowing into the soma. The positive peak is referred to as Rep because it is associated with the cell repolarization phase. The peak values were computed with respect to a reference of 0 μV, i.e., the baseline, as follows:

NA. For each spike recording on N electrodes, the spike with the largest negative peak amplitude was identified. At the time instant when the minimum peak occurred (t_{min}) the recorded potential on all N electrodes was used to build the Na image (the amplitude values are sampled at the same time instant t_{min}).

REP. The time instant of the repolarization peak (t_{max}) was extracted from the spike trace with the largest negative trough (same electrode as Na) and a Rep image was built by probing all N electrodes at t_{max}.

Overall, the localization-specific (N_y, N_z)-dimensional sets of features are Na, Rep, and NaRep, where the last is a stacked version of both features having dimension (N_y,N_z, 2).

Classification features. From each spike, we extracted features that are commonly used for cell classification (Barthó et al. 2004; Peyrache et al. 2012): peak-to-peak width (*W*), full-width half-maximum (*F*), and peak-to-peak amplitude (*A*). The peak-to-peak amplitude *A*, despite not being one of the commonly used features for classifying neurons from extracellular recordings, was selected as a feature in combination with *F* and *W* because spike widths increase with increasing recording distance (Anastassiou et al. 2015; Hagen et al. 2015; Pettersen and Einevoll 2008) and therefore with decreasing amplitude.

The following is a list including a detailed description of the features:
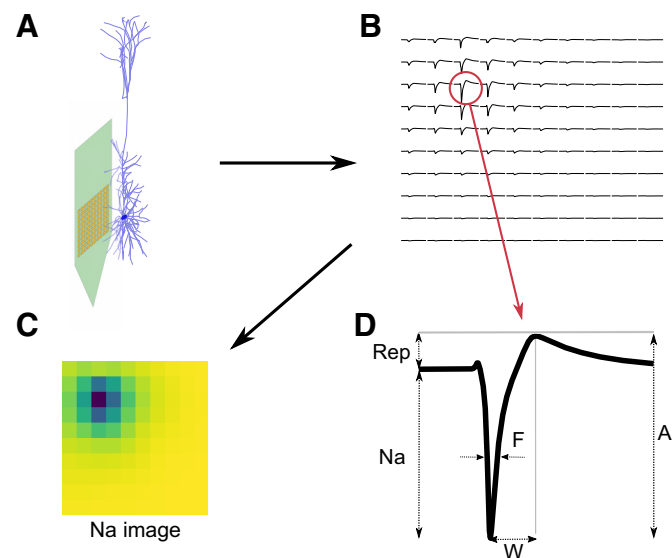
Fig. 5. Feature extraction pipeline: first, neuronal models (a pyramidal cell here) are simulated (*A*) and extracellular action potentials (EAPs; *B*) are computed on the multielectrode array (MEA) probe (SqMEA-10-15 here). Features (*D*) are then extracted from EAPs. Localization features are sodium negative peak (Na) and repolarization (positive) peak (Rep), and classification features are peak-to-peak amplitude (*A*), spike width (*W*), and full-width half-maximum width (*F*). The feature images (*C*) are finally used as input for convolutional neural networks.

A: For each recording site, the peak-to-peak amplitude was extracted as the absolute difference between the negative peak and the following positive repolarization peak. If the amplitude value of a recording site was less than a detection threshold of 5 μV, then the amplitude for that electrode was set to zero.

W: For each electrode site, the width was computed as the time difference between the negative peak and the following positive repolarization peak. If the amplitude of the corresponding electrode was below the detection threshold (i.e., when the amplitude feature was set to 0), the width was set to the duration of the spike window, which was 7 ms.

F: For each electrode site, the full-width half-maximum was computed as the width at 50% of the negative maximum amplitude (refer to Fig. 5 for a graphical visualization and further explanation). In this case, the reference voltage was the initial voltage on the selected electrode site at beginning of the spike window. If the amplitude of the corresponding electrode was below the detection threshold (when the amplitude feature was set to 0), *F* was set to the duration of the spike window, which was 7 ms.

For classification, we considered the feature combinations *AW* (N_y, N_z, 2), *FW* (N_y, N_z, 2), and *AFW* (N_y, N_z, 3), where the shapes of input arrays to the CNN are indicated in parentheses.

Waveform. We also investigated the performance using the entire spike waveform as input to the CNNs for localization and classification. While localization and classification features focused on amplitudes at specific time points (e.g., Na, Rep) or on extracting significant spike shape parameters (*A, F, W*), here we took into account the evolution of the action potential in time. As the additional third dimension (2D + time) increased the training time significantly, we downsampled the initial spike waveforms from 224 to 14 samples, i.e., with a downsampling factor of 16. We denoted this feature set, with a shape of (N_y, N_z, 14), as Waveform.

Convolutional Neural Network

CNNs are biologically inspired artificial neural networks, and they differ from standard artificial neural networks mainly by the use of convolutional layers. The biological inspiration originates from the information processing in the visual system (Krizhevsky et al. 2012; Zeiler and Fergus 2014). For our implementation, we used the opensource software TensorFlow to train and evaluate the network (see Abadi et al. 2016; software available from https://www.tensorflow.org). All computations were done on the HPC clusters NEMO in Freiburg and ABEL in Oslo.

Configuration. We investigated the performance of CNNs of different sizes, all having the same underlying configurations (except for Waveform input features, whose CNN morphology is explained at the end of this section). Five CNN sizes (XS, S, M, L, XL) were used, and they differ in the size k_i of convolutional kernels (the index $i \in \{1,2\}$ specifies the convolutional layer), convolutional layer depths d_i, and the number of nodes in the fully connected layer n_{FC}. The values used for different sizes are listed in Table B1 in APPENDIX B.

Feature images of dimension (N_y, N_z) were input to a d_1-deep convolutional layer with rectified-linear units that filter the input image with (k_1, k_1) kernels and a stride of 1. Then max-pooling was applied, and the image was shrunk to a (n_y, n_z) = (N_y/2, N_z/2) footprint. Another d_2-deep convolutional layer with rectified-linear units applied (k_2, k_2) kernels, and a second max-pooling operation reduced the output image features to a (m_y, m_z) = (n_y/2, n_z/2) size. The (m_y, m_z) features were input to a fully connected layer with n_{FC} nodes. The fully connected nodes were connected to the nodes in the output layer (see *Output layer and optimization*).

The Waveform feature set differed from the classification- and localization-specific sets as it included time as a dimension. Although some feature images for localization and classification were concatenated and thus had a 3D shape [for example, NaRep had a shape of (N_y, N_z, 2) and *AFW* is (N_y, N_z, 3)-dimensional], the optimized kernels

were the same for two or three dimensions. For the Waveform feature set, a 3D CNN was used, i.e., convolutional kernels were 3D with a shape of (k_1, k_1, k_t) and (k_2, k_2, k_t). Max-pooling was also applied in all three dimensions. For the Waveform feature set, we used a CNN with $k_t = 4$.

Output layer and optimization. The output layer of the network was different depending on whether localization or classification was performed. In case of localization, three output nodes linearly summed the fully connected node inputs and biases to output the *x*-, *y*-, and *z*-coordinates. Optimization minimized the mean squared error between the predicted *x*-, *y*-, and *z*-coordinates and the true soma positions of the training spikes. For classification, there were instead two output nodes in case of excitatory/inhibitory classification. For the m-type classification, we used 11 output nodes, 1 for each cell type under consideration (see *Cell Models* for a list of m-types). During training, softmax cross entropy was minimized (Goodfellow et al. 2016).

For both localization and classification we used an Adam optimizer (learning rate = 0.0005) (see Kingma and Ba 2014), and we ran 2,000 training epochs. At each iteration, 10% of the training observations were randomly sampled and used to update network weights with backpropagation. To limit overfitting, we used dropout on the fully connected layer (Srivastava et al. 2014) with a dropout rate of 0.3 (during training 30% randomly chosen fully connected nodes were dropped and not considered for updating the network weights).

Validation strategy. To estimate the accuracy of the CNNs, we divided the input data into a training set, used to estimate the CNN parameters, and a validation set, upon which the trained CNN's accuracy was tested. Before the training-validation set division, we preprocessed the data set so that morphologies in the training and validation sets were unique (APPENDIX A). Then, we balanced the occurrence of observations of the same cell type (m-type) by random undersampling. For excitatory/inhibitory classification, we further subsampled the inhibitory neuron observations to match the excitatory ones (in the data set, there are 7 inhibitory cell types—not counting BP and NGC—and 4 excitatory types). The class balancing was performed for training and validation sets separately. For localization (and m-type classification), the training and validation data sets consisted of 44,000 and 11,000 instances, respectively. For classification, we used 32,000 observations for training and 8,000 for validation.

Comparison with Other Models

Localization. In previous work on neural localization, the underlying idea has been to solve the inverse problem by choosing a generative model and minimizing the error between the true extracellular potential and the one predicted by the chosen model. The soma position has been among the model parameters that have been optimized. Several models were used in previous studies: monopole and dipole current-source models (Blanche et al. 2005), line-source models (Somogyvári et al. 2012), and ball-and-stick models (Delgado Ruz and Schultz 2014). We compared our localization approach to inverse problem solutions solved with the EAP negative peak (i.e., Na image) with the following models (σ denotes the extracellular conductivity):

MONOPOLAR CURRENT SOURCE. A negative monopolar current-source I_s placed at position r_s evokes a potential $\phi(r)$ at position r according to

$$\phi(r) = \frac{I_s}{4\pi\sigma\|r - r_s\|} \tag{2}$$

The somatic current and soma position are the only parameters to be optimized. The predicted soma position is r_s.

BIPOLAR CURRENT SOURCE. Placing a negative current-source I_s at position r_{neg} and its positive counterpart (absolute value of I_s) at r_{pos}, the potential at position r reads

$$\phi(\boldsymbol{r}) = \frac{I_s}{4\pi\sigma}\left(\frac{1}{\|\boldsymbol{r} - \boldsymbol{r}_{neg}\|} - \frac{1}{\|\boldsymbol{r} - \boldsymbol{r}_{pos}\|}\right). \tag{3}$$

In this case, the estimated soma position corresponds to the negative current-source location, which is $\boldsymbol{r}_{neg}$. This model is equivalent to the two-monopole model in Pettersen and Einevoll (2008).

BALL AND STICK. The ball-and-stick model combines a somatic point-like constant current-source I_s at position $\boldsymbol{r}_s$ with a dendritic stick of length d_{len} pointing in direction $\boldsymbol{d}$. We used a modified version of the ball-and-stick model described in Pettersen and Einevoll (2008), since we do not assume net currents to be zero (Delgado Ruz and Schultz 2014). The current along the stick $I(x)$ is assumed to decay exponentially, as confirmed by experimental data (Foust et al. 2010; Goldberg and Yuste 2005; Golding et al. 2001; Gulledge and Stuart 2003; Migliore et al. 1999; Sasaki et al. 2012; Waters et al. 2005). With initial negative value I_{dend} at $\boldsymbol{r}_s$, the current distribution along the stick reads

$$I(x_d) = I_{dend} \times \exp\left(-\frac{x_d}{d_\tau}\right) \tag{4}$$

where d_τ denotes the decay constant and $x_d \in [0, d_{len}]$ is the position along the dendritic segment (discretized in $N_{seg} = 50$ uniformly distributed points along the stick of length d_{len}). The predicted soma location corresponds to $\boldsymbol{r}_s$. The potential at position $\boldsymbol{r}$ is given by the summation of the somatic and dendritic contributions:

$$\phi(\boldsymbol{r}) = \phi_{soma}(\boldsymbol{r}) + \phi_{dend}(\boldsymbol{r}) = \frac{I_s}{4\pi\sigma\|\boldsymbol{r} - \boldsymbol{r}_s\|}$$
$$+ \sum_{i=1}^{N_{seg}} \frac{1}{4\pi\sigma} I(x_i) \int \frac{d\boldsymbol{r}_i}{\|\boldsymbol{r} - \boldsymbol{r}_i\|} \tag{5}$$

where each segment is modeled as a line current source (see *Eq. 1*).

Table 1 summarizes the parameters to be estimated for each described model.

GENETIC OPTIMIZATION. To estimate the model parameters, we minimized the sum of squared errors at each recording site between the extracellular potential predicted by the model and the extracted Na feature image of the true simulated extracellular potential. Optimization was performed with a genetic algorithm implemented with the Distributed Evolutionary Algorithms in Python (DEAP) package (Fortin et al. 2012). We used the $(\mu + \lambda)$ evolution strategy, which selects the next parents from the common set of the current parents (μ individuals) and the offspring (λ individuals). More precisely, the algorithm was implemented with the deap.algorithms.eaMuPlusLambda function. We used $\mu = 100$, $\lambda = 200$, crossover probability $p_{cx} = 0.8$, and mutation probability $p_{mut} = 0.2$. Furthermore, tournament selection (deap.tools.selTournament) and blend crossover (deap.tools.cxBlend) were used for selecting and mating individuals, respectively. Mutation was performed with a random Gaussian mutation (deap.tools.mutGaussian). When an individual was selected for mutation with probability p_{cx}, each of its elements was individually mutated with an individual probability of $p_{ind} = 0.3$. Gaussian means for all parameters were set to zero, and standard deviations (SDs) were

Table 1. *Inverse model parameters*

Model	Parameters	No. of Parameters
Monopolar	I_s, r_{s_x}, r_{s_y} r_{s_z}	4
Bipolar	I_s, r_{pos_x} r_{pos_y} r_{pos_z}, r_{neg_x}, r_{neg_y}, r_{neg_z}	7
Ball and stick	I_s, r_{s_x}, r_{s_y} r_{s_z}, d_x, d_y, d_z, I_{dend}, d_{len}, d_τ	10

Summary of parameters for the different inverse models involved in the study. I_s, current source; r_s, predicted soma position; r_{pos}, position of positive I_s; r_{neg}, position of negative I_s; I_{dend}, dendritic current; d_{len}, dendritic length; d_τ, decay constant.

Table 2. *Model parameter summary*

Parameter	Range	Gaussian σ
I_s, I_{dend}	$(-100, 0)$ nA	1 nA
x positions	$(10, 80)$ μm	10 μm
y, z positions	$(-180, 180)$ μm	10 μm
d_x, d_y, d_z	$(-1, 1)$	0.1
d_{len}	$(1, 500)$ μm	50 μm
d_τ	$(0.1, 500)$ μm	20 μm

Range for initialization and constraint and standard deviation σ for mutation Gaussian for the different parameters. I_s, current source; I_{dend}, dendritic current; d_{len}, dendritic length; d_τ, decay constant.

different depending on the parameter. The parameter values are summarized in Table 2, and we constrained the optimization to solutions within biophysically acceptable boundaries (shown in Table 2).

Classification. Besides applying a CNN, the problem of classifying neurons according to their EAP can be done by several other methods (Barthó et al. 2004; Delgado Ruz and Schultz 2014; Peyrache et al. 2012). It is a well-established observation that pyramidal excitatory cells present a broad spike waveform, while inhibitory cells have a narrow one (Barthó et al. 2004). Therefore, a standard way of classifying between the two classes is to plot spike width W and full-width half-maximum F (see *Classification and Localization Features* for feature extraction details) of the EAP with the maximum amplitude and then cluster the data points in this 2D space (Barthó et al. 2004; Peyrache et al. 2012). Once F and W were computed for the electrode with the maximum peak-to-peak amplitude, we applied two different clustering techniques to the point cloud: *k*-means and a mixture of Gaussians (MoG) clustering (Pedregosa et al. 2011). While *k*-means clustering iteratively assigns points to K clusters based on their distances to the cluster means and then recomputes the cluster means with new assignments until convergence, the MoG estimates K Gaussians to fit the data and then labels the data points based on the Gaussian shape. In this case, since the goal is excitatory/inhibitory classification, we set $K = 2$.

Statistical Analysis

For localization errors, statistical tests were run on the 11,000 validation observations: since all distributions did not satisfy the normality assumption, nonparametric tests were run (1-sided Mann-Whitney *U*-test; Mann and Whitney 1947). When sample sizes are large, statistical tests are prone to indicate that there is a significant difference (effect) between distributions, resulting in low *P* values. To characterize whether such difference is relevant, a measure of its magnitude, or effect size, should be included (Sullivan and Feinn 2012). To quantify the effect size, we used Cohen's *d* coefficient (Cohen 1992), i.e., the difference between population means normalized by the pooled SD. We considered significant differences (low *P* values) to be negligible (effect size < 0.2), small (effect size ≃ 0.2), medium (effect size ≃ 0.5), or large (effect size ≃ 0.8). Test results are shown in APPENDIX B, divided by group (data set rotation, Table B3; CNN size, Table B4; feature type, Table B5; probe type, Table B6; and localization method, Table B7). Each entry of the tables shows the Cohen's *d* coefficient (rounded to 2 decimals) and the significance of the Mann-Whitney *U*-test with the alternative hypothesis that column group < row group.

RESULTS

In the following sections, we show localization and excitatory/inhibitory classification results only of m-type cells included in the training data set. Therefore, unless otherwise specified, BP and NGC are excluded from the analysis. The performance measures were different for localization and clas-

sification. In case of localization we used the average total error and for classification the average classification accuracy (ratio between correctly classified observations and total number of observations). Moreover, we analyzed the cell-specific accuracy to get more insight on the classification performance. The average localization error in the x-, y-, and z-directions can be interpreted in the following way. Assuming normally distributed errors, the true soma location is with 32% probability inside a box centered at the predicted soma position with edge lengths of twice the average localization error in the corresponding dimension. The probability rises to 87% with a box edge length of four times the average localization error in the corresponding dimension.

Dependence on Neuron-MEA Alignment

The first question we investigate is how the neuron-MEA alignment affects the localization and classification performance. Three data sets were built, Norot, Physrot, and 3drot. To focus on the effects of alignments, we use fixed feature sets (NaRep for localization, *FW* for classification; see *Classification and Localization Features* for definitions), MEA probe (SqMEA-15-10), and CNN size L.

Localization. First, we show the mean and SD of the localization errors along the three axes as well as the total error in Table B2. Each row displays the performance of a rotational data set. Average errors and SDs are 7.3 ± 5.7 μm for Norot, 7.8 ± 6.3 μm for Physrot, and 8.9 ± 8.2 μm for 3drot.

The Norot data set results in significantly lower errors with respect to the 3drot data set (effect size = 0.21; Table B3). Negligible differences are found in the comparison between the Norot errors and the Physrot errors (effect size = 0.09) and between the Physrot and 3drot data sets (effect size = 0.13). Taking into account the finite size of the soma (~10–15 μm of diameter), we consider the resulting error values to be a sufficient performance for most applications. The errors along the three axes appear to be isotropically distributed, as they show similar values in all directions (but the observations in the x-direction are not uniformly distributed—as shown below in Fig. 9C—since we only considered spikes above 30 μV and spike amplitude decreases with distance).

In Fig. 6, *A*–*C*, we show the errors along x-, y-, and z-axes with respect to the x-, y-, and z-coordinates for the three rotational data sets. In these plots, we bin the true x, y, and z neuron positions in seven bins and treat them as categorical data (i.e., the positions can have discrete values depending on the bin they belong to). The data points are the mean of the error grouped by bin and rotation type and the error bar is the SD. Figure 6A shows that errors in the distance estimation (x-direction) tend to increase as the distance of the neuron increases for all three data sets, similarly to Delgado Ruz and Schultz (2014). Regarding the y- and z-dimensions (Fig. 6, *B* and *C*, respectively), it is interesting to note how the errors have a convex shape, meaning that errors tend to increase when the neuron is at the border of
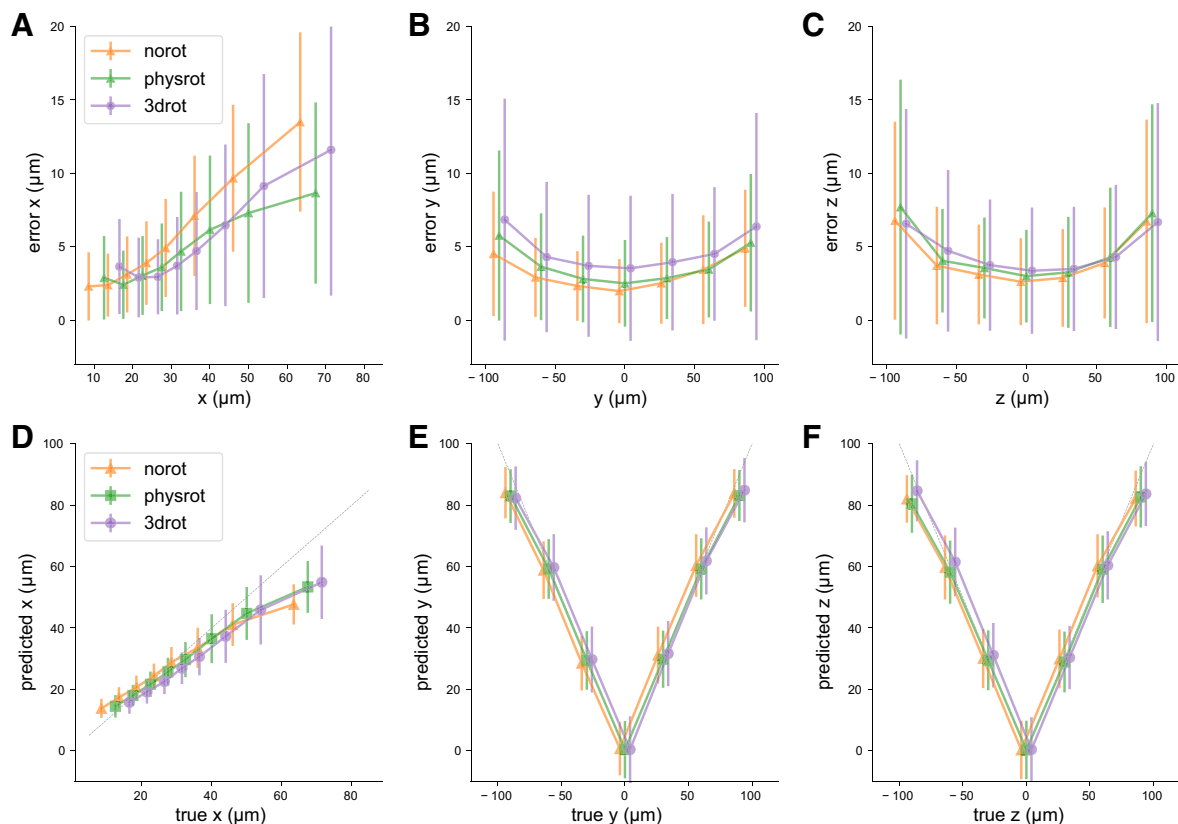

Fig. 6. *A*: x errors with respect to the x-coordinate (distance). *B*: y errors with respect to the y-coordinate. *C*: z errors with respect to the z-coordinate. *D*: x predictions with respect to the true x-coordinate (distance). *E*: absolute value of the y predictions with respect to the true y-coordinate. *F*: absolute value of the z predictions with respect to the true z-coordinate. All values are in μm. Orange lines indicate the *Norot* data set, green lines the *Physrot* data set, and purple lines the *3drot* data set. Gray lines correspond to a perfect prediction. Data are binned in 7 bins along x-, y-, and z-directions: points and error bars display the average errors and their SDs for each bin and each data set.

the probe, in which case only partial information about the spike is available.

When looking at the distribution of the predicted x-, y-, and z-coordinates with respect to the true coordinates in Fig. 6, *D–F*, one can note how the errors observed in Fig. 6, *A–C*, are caused by an underestimation of the soma distance in all dimensions: at large distances (x-direction) from the probe, neurons are predicted to be closer to the MEA; when they are close to the y and z borders of the probe, the predicted position is closer to the center of the MEA.

Next, we consider the variability of localization performance depending on cell types and alignment. In Fig. 7*A* the bar plots show the average total errors and their SDs grouped by neuron morphology type (11 training morphologies + BP and NGC; see Fig. 2 for representative cells) for the Norot, Physrot, and 3drot data sets. The range and distribution of distances taken into consideration are the same as in Fig. 6. Focusing on the Physrot data set, the minimum error of 4.7 ± 3.8 μm is obtained for the SBC morphology, while the worst performance is 15.9 ± 9.1 μm for slender-tufted PC (STPC) morphology. The difference in prediction performance with respect to cell type does not seem to be depending on excitatory/inhibitory morphologies (i.e., pyramidal and nonpyramidal cells), nor do they look to be clustered depending on morphological subclasses; for instance, among the different basket cells (names ending with BC) there is some variability among large basket cells (LBC), NBC, and SBC, and the same holds for pyramidal cells [STPC, TTPC1, thick-tufted PC with early bifurcating apical tuft (TTPC2), and untufted PC (UTPC)]. The performance of BP and NGC (Fig. 7*A*), which were not used for training, is in line with other cell types, with errors of 8.8 ± 2.2 μm and 9.4 ± 5.1 μm, respectively. This result confirms that the method is capable of dealing with diverse morphologies, as long as the training set contains a large representation of cell types.

Classification. The accuracy analysis of excitatory/inhibitory classification is based on the *FW* feature set. Table B8 in APPENDIX B shows the classification accuracies for each cell morphology plus the average accuracy and the SD for the different data sets. The cell-specific accuracies are also visualized by color coding in Fig. 7*B*.

The average classification accuracy is equally high for the Norot and Physrot data sets ($98.1 \pm 2.4\%$ and $98.0 \pm 3.9\%$, respectively) and lower for the 3drot data set ($97.6 \pm 3.9\%$). This is because neurons are rotated with more degrees of freedom; nevertheless, on average the accuracy remains very high in all cases. A closer examination of this result reveals that the main reason for the drop in classification accuracy was misclassification of the chandelier cells (ChCs). The lowest value is the ChC accuracy in the 3drot data set (84.1%). In Fig. 7*C*, we show the spike shapes in the *FW*-plane of the electrode site with largest amplitude. Inhibitory neurons mainly lie in the lower left part (narrow spikes). Excitatory neurons are almost perfectly classified as excitatory cells, as shown in Table B8 and Fig. 7*B*. The spike shapes of ChC in the *FW*-plane mainly lie at the interface with the excitatory neurons. This might explain why they are harder to classify correctly with respect to the other cell types.

Effect of CNN Size

We investigated how localization and classification performances vary with network size. The results shown in this section were obtained with the SqMEA-10-15 probe, NaRep features for localization, and *FW* for classifications. For the remaining analyses, boxplots and cumulative distribution functions (cdfs) are used to represent the performance of the localization models. In all boxplots, the box is the interquartile range (IQR), i.e., the 25th and 75th percentiles, the horizontal lines inside the box show the medians, and the red diamonds display the means. The whiskers (horizontal black lines) represent the highest and lowest data values within 1.5 times the IQR. Data points outside the whiskers are plotted as black dots and are regarded as outliers. We obtained the cdfs by sorting the sample and pairing each data point with its normalized rank (percentile). Hence, the point where the cdf crosses 0.5 represents the median of the localization error.

Fig. 7. *A*: localization error grouped by cell type for all rotational data sets. Bars are the average errors, and error bars show the SDs in μm. Orange bars display the *Norot* data set, green bars display the *Physrot* data set, and purple bars display the *3drot* data set. Ticks on the *x*-axis show the cell types: red ticks are inhibitory cells, blue ticks are excitatory, and yellow ticks are bipolar (BP) and neuroglial (NGC) cells (not used for training). *B*: excitatory/inhibitory classification accuracy in color code grouped by rotation and cell type. Each element of the matrix is the accuracy of a specific cell type (red, inhibitory neurons; blue, excitatory neurons) in the different rotational data sets (rows). For explicit values see Table B8. *C*: spike shapes for maximum peak electrode sites are plotted in the *FW*-plane. Red dots are inhibitory neurons, and they lie in the lower left part of the plot. Blue dots show excitatory cells, in the upper right part of the plane. The magenta dots are chandelier cells (ChC), and they lie at the intersection between excitatory and inhibitory cells. The opacity represents the number of occurrences; more opaque dots correspond to more frequent observations. The ellipses are built by following the principal axis of each distribution, and the lengths of their major and minor axes are proportional to the eigenvalues of the covariance matrix. FWHM, full-width half-maximum.
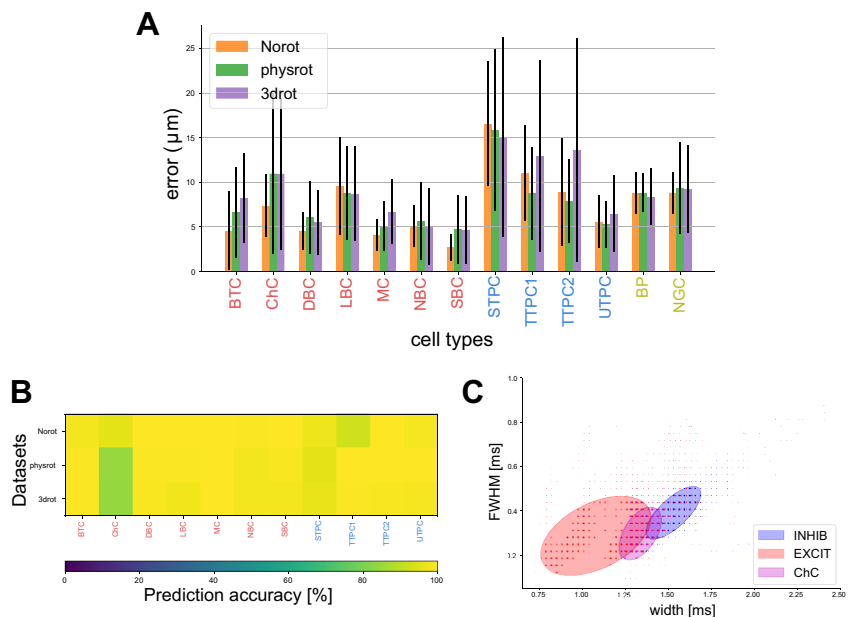
Localization. Figure 8*A* shows the localization errors grouped by CNN size (XS, S, M, L, and XL). Increasing the size of the network improves the performance significantly (Table B4), but for sizes L and XL the average localization error is almost the same [7.8 ± 6.3 μm for L and 7.3 ± 5.8 μm for XL (effect size = 0.09)]. If not stated otherwise, networks of size L have been chosen, as they provide a good compromise between performance and time required for training.

Classification. Table B9 in APPENDIX B shows the performance in classification into excitatory and inhibitory neuron types. The highest accuracy (98.6 ± 1.1%) is reached with a network of size M, while all others show a slightly lower performance. A possible explanation for the lower score of the XL network is overfitting to the training set because of the large number of parameters.

Feature Selection

In the previous sections, we have presented results with fixed feature sets (NaRep for localization and *FW* for classification), eliminating the effects caused by other factors, such as alignment, cell type and CNN size. The following results were obtained on SqMEA-10-15 probes using CNNs of size M (because of the long training time required by 3D CNNs for Waveform feature).

Localization. In Fig. 8, *C* and *D*, we display the boxplots and cdf of the errors with varying feature sets for localization. In other studies, either the sodium peak is the only feature used (Blanche et al. 2005; Delgado Ruz and Schultz 2014; Mechler et al. 2011; Mechler and Victor 2012; Somogyvári et al. 2005)

or the entire spike time course is modeled (Somogyvári et al. 2012). Here we show that all CNNs relying on peak input show roughly the same performance: average errors ± SDs are 8.8 ± 7.1 μm for Na, 8.8 ± 7.9 μm for Rep, and 8.8 ± 7.7 μm for NaRep. Negligible differences are found when comparing Na, Rep, and NaRep, with effect sizes close to zero (Table B5).

The Waveform CNN results in a lower average prediction error of 6.9 ± 6.5 μm, which is significantly better in comparison with Na (effect size = 0.28), Rep (effect size = 0.26), and NaRep (effect size = 0.27; Table B5). We speculate that the performance of the Waveform approach is only slightly increased (by ∼2 μm) for the following reason: when considering the peaks only, transmembrane currents are mainly concentrated around the soma (Delgado Ruz and Schultz 2014; Gold et al. 2007; Somogyvári et al. 2005, 2012); therefore, the peak features contain almost all information the CNN needs for soma location.

Classification. Classification performances are listed in Table B10 in APPENDIX B. The *AFW* feature set, with an accuracy of 98.6%, performs better than *AW* and *FW*, with accuracies of 98.1% and 97.0%, respectively. The Waveform feature set, which uses a downsampled version of the entire spike, performs almost perfectly on the classification task (accuracy 99.7%). Given these results, the Waveform feature set is better than the other approaches, at the expense of more computationally demanding training procedures.

Performance with Different MEA Probes

We built simulated spikes using eight different MEA models: five of them are square arrays with varying pitch, and the
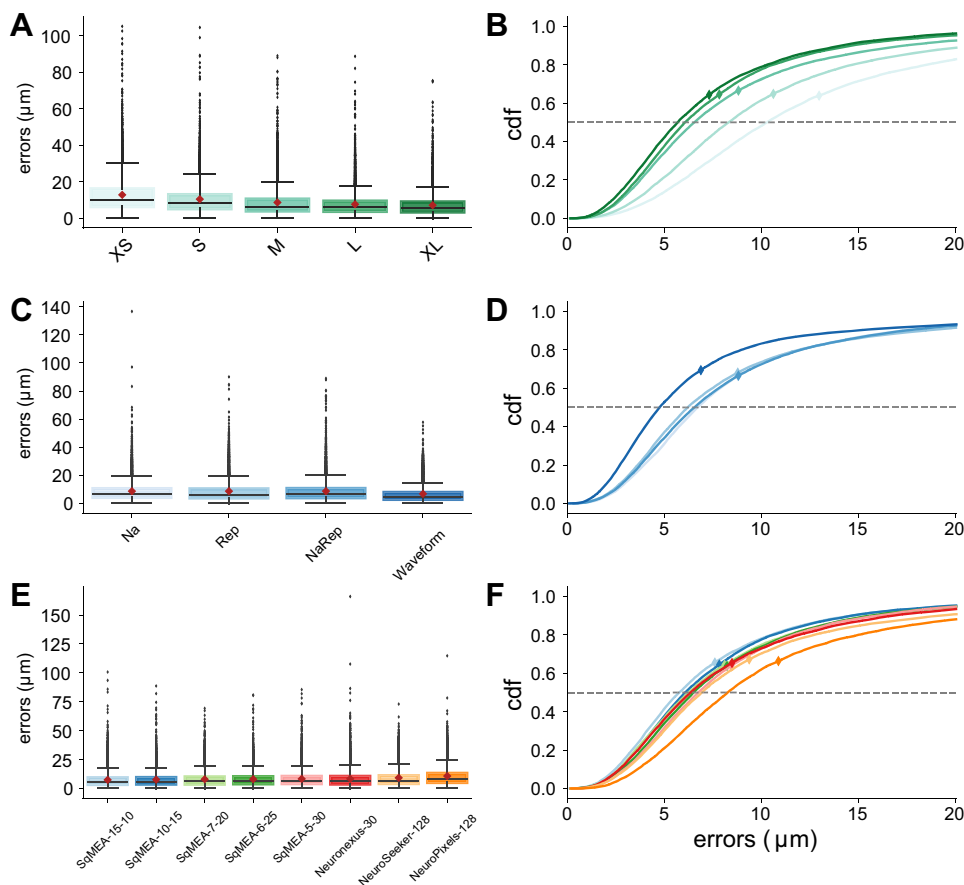
Fig. 8. *A*, *C*, and *E*: error boxplots grouped by convolutional neural network (CNN) size (*A*), feature type (*C*), and probe type (*E*). Red diamonds represent means, and black diamonds denote outliers. *B*, *D*, and *F*: cumulative distribution function (cdf) grouped by CNN size (*B*), feature type (*D*), and probe type (*F*). Gray dashed horizontal line at 0.5 defines the median. Diamonds on each curve show the means (their *x* values are the average error; the *y* values are the percentile of their occurrence). The error values are clipped to 20 μm to zoom in the distributions. Statistical analyses are shown in Table B4 (for *A* and *B*), Table B5 (for *C* and *D*), and Table B6 (for *E* and *F*).

other three are the NeuroSeeker (Neto et al. 2016), NeuroPixels (Jun et al. 2017) (trimmed to 128 channels), and Neuronexus-32 (clipped to 30 electrodes to make it rectangular) probes. In the following paragraphs, we present the capabilities in terms of neuron localization and classification for the different probes. All simulations shown in this section make use of CNNs of size L, NaRep features for localization, and *FW* for classification.

Localization. Figure 8, *E* and *F*, show localization errors for the eight different probes (boxplots and cdf). Although an error reduction can be observed from square MEA with 30-μm pitch to 10-μm pitch, as expected, even with a relatively low density (30-μm pitch) a CNN can learn localization models with an average error as low as 8.4 $\pm$ 6.4 μm for the SqMEA-5-30. As a comparison, the average error for the probe with 10-μm pitch (SqMEA-15-10) is 7.6 $\pm$ 6.4 μm. The errors are in the same order also for the Neuronexus probe (mean of 8.5 $\pm$ 7.2 μm) and for the NeuroSeeker probe (mean of 9.3 $\pm$ 7.7 μm). When evaluating the performance on 128 sites with the arrangement of the NeuroPixels probe, the average error is 10.8 $\pm$ 8.5 μm. One may note that even though the NeuroSeeker and Neuronexus probes have lower pitch (NeuroSeeker: 22.5 μm in *y*- and *z*-axes; Neuronexus: 18 μm in *y*-axis and 25 μm in *z*-axis) compared with SqMEA-5-30, their localization error is higher. The reason for this discrepancy might be in the arrangement of the electrodes: while the SqMEA-5-30 has an effective width (considering point electrode contacts) of 120 μm, for the NeuroSeeker the effective width (considering point electrode contacts) is 67.5 μm and for the Neuronexus it is 36 μm. Hence on the NeuroSeeker and Neuronexus probes there is less spatial information in the *y*-direction, which may explain the reduced localization accuracy.

In general, most comparisons show negligible differences (effect size < 0.2), except for the NeuroSeeker and NeuroPixels probes. The NeuroSeeker probe performs worse than the high-density square MEAs (SqMEA-15-10: effect size = 0.25, SqMEA-10-15: effect size = 0.22), while the NeuroPixels errors show effect sizes above 0.2 in all comparisons (ranging from 0.43 compared with SqMEA-15-10 to 0.3 compared with Neuronexus) except for the comparison with the NeuroSeeker probe (effect size = 0.19). In case of the NeuroPixels probe the checkerboard arrangement might pose additional difficulties, resulting in even lower performance.

Classification. Table B11 in APPENDIX B shows accuracies for classification with different probes. The average accuracies are very high and almost the same for all probes, from a minimum of 96.6% (SqMEA-6-25, SqMEA-15-10) to a maximum of 98.6% (NeuroPixels-128).

Comparison with Other Approaches

In this section, we compare the CNN approach to other state-of-the-art methods. For localization, we used the monopolar, bipolar, and ball-and-stick models described in *Comparison with Other Models* to solve the inverse problem on our simulated data sets. Hence the results obtained for other methods might be different with respect to ones in the literature because the number of cell models, the utilized probes, and the neuron-MEA alignments vary. For characterization of excitatory and inhibitory neurons, we compared with commonly used clustering techniques.

Localization. For localization, we use the validation data set on the SqMEA-10-15 probe. The CNN errors displayed in the plots are obtained with the NaRep feature set and a network of size L. In Fig. 9, we show the errors of the simplified models described in *Comparison with Other Models* and for the CNN method.

We found that the CNN performs significantly better than the inverse approach in all cases, with an average error and SD of 7.8 $\pm$ 6.3 μm. The large differences between the CNN and the other methods' error distributions are confirmed by the effect sizes: 0.9 for the monopolar approach, 0.68 for the bipolar approach, and 0.87 compared with the ball-and-stick approach (Table B7). Among the models used to solve the inverse problem, the monopolar has a mean error and SD of 21.7 $\pm$ 20.9 μm, the bipolar model of 15.6 $\pm$ 15.2 μm, and the ball-and-stick model of 22.6 $\pm$ 23.2 μm. The better performance of the bipolar model with respect to the monopolar model (and ball-and-stick model) can be due to the fact that it is the only model capable of predicting negative and positive potential values on the MEA. Dendritic branches act as current sources when the soma is depolarized, causing positive deflections in the extracellular potential (Pettersen and Einevoll 2008).

Studying the probability density function (pdf) of the predicted coordinates by different models in Fig. 9, *C–E*, the monopolar model tends to underestimate the distance (*x*-coordinate) from the MEA (note sharp peak in the distribution in Fig. 9*C*). In the *y*- and *z*-axes, instead the predictions are closer to the center of the MEA when observations lie outside the boundary of the probe (note the different steepness and shape of the monopolar pdf with respect to the true pdf in Fig. 9, *D* and *E*, close to -100 μm and 100 μm). Similarly, the bipolar model also underestimates distances in the *x*-direction, but the underestimation is less severe. In the *y*- and *z*-directions it nicely follows the true distribution. The ball-and-stick model has distributions very similar to the monopolar model in all three directions. The CNN approach, on the other hand, is the closest match to the true distribution in all three dimensions. Note that the distribution in the *x*-direction is not as uniform as in the *y*- and *z*-directions (the density decreases with increased distances) because we discarded spikes with a peak-to-peak amplitude below 30 μV here.

Classification. For excitatory/inhibitory classification we compared the performance of our CNN approach to standard clustering techniques in the *FW* space (*F*: full-width half-maximum, *W*: width). In Fig. 10*A*, we show the validation data with the SqMEA-10-15 probe and the excitatory/inhibitory balanced data sets (8,000 observations). Each point is computed from the recording site with largest amplitude. Although it is true that inhibitory cells cover the bottom left part of the cloud (narrower width and full-width half-maximum) and excitatory cells the top right (wider spike shape), we can observe that there is some overlap between the two groups. When we apply *k*-means clustering (Fig. 10*B*), the algorithm correctly assigns the bottom left part to inhibitory neurons and the top right part to excitatory neurons, but the overlap is mainly assigned to the excitatory class. This yields an accuracy of 99.9% for the excitatory class but only 60.7% for the inhibitory one, with an average of 80.3%. When it comes to the MoG, the data are fit to two multivariate Gaussians and labels are assigned based on the probability of an observation to belong
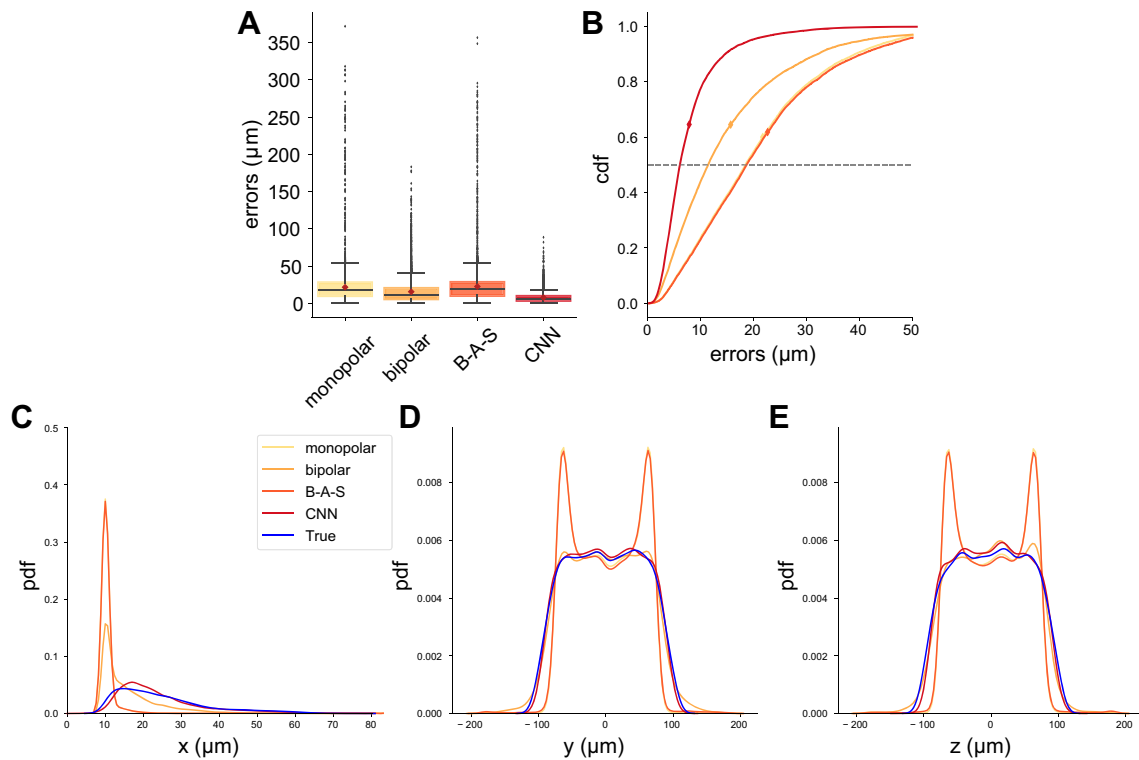
Fig. 9. *A*: error boxplots grouped by model used to solve the inverse problem. Red diamonds are means; black diamonds are outliers. B-A-S, ball-and-stick; CNN, convolutional neural network. *B*: cumulative distribution function (cdf) grouped by localization method. Gray dashed horizontal line at 0.5 defines the median. Diamonds on each curve show the means (their *x* values are the average error; the *y* values are the percentile of their occurrence). The error values are clipped to 50 μm to zoom in the distributions. *C–E*: probability density functions (pdfs) of predicted values by different models and true values in the *x* (*C*)-, *y* (*D*)-, and *z* (*E*)-directions. Statistical analysis is shown in Table B7.

to the two distributions. Figure 10*C* shows the estimated Gaussians (ellipses) and the labeling of the points. Although the MoG is capable of describing the diagonal shape of the excitatory cloud, the overlap between the observations cannot be untangled, resulting in an accuracy of 98.8% for the excitatory class and 54.2% for the inhibitory one, with an average of 76.5%. The CNN method, instead, is able to discern the overlap in the *FW* space. This is certainly due to its higher complexity, due both to the method itself and to the use of all electrodes' information, not only those with highest amplitude. The CNN result shown here (*FW* feature set, size L) allows us to correctly predict excitatory cells in 98.9% of the cases and inhibitory cells with an accuracy of 97.1%. This makes it the best-performing method among those compared here, with an overall accuracy of 98.0%. It could be argued that the comparison was somewhat unfair, as our CNN approach considers *F* and *W* images (computed on all recording sites), while the clustering is performed with values computed from the electrode with highest amplitude only. Nevertheless, it is not common practice to consider waveforms on all electrodes but only on the one with highest amplitude (Barthó et al. 2004; Peyrache et al. 2012).

m-Type Classification

In addition to separating excitatory cells from inhibitory ones by trained CNNs, we tried to make a finer subdivision and classify cells into morphology classes (m-type) based on the EAP. The approach is similar to that for excitatory/inhibitory classification, but instead of only 2 output classes we take the

11 m-type classes (cells of m-type BP and NGC are excluded, since only 1 morphology is available in the data set). We use a CNN of size L and consider Waveform features (in this case with a downsampling factor of 8, i.e., a sampling frequency of 4 kHz) on the SqMEA-10-15 with the Physrot data set. The resulting confusion matrix *E*, in which each entry E_{ij} represents the amount of observations of the true cell type *i* predicted as cell type *j*, is depicted in Fig. 11. We do not observe a striking diagonal, indicating that full identification of all cell types is not feasible from EAPs. But it is noteworthy that there is some block structure dividing excitatory neurons from inhibitory neurons. This division is learned intrinsically by the network, and inhibitory cells are classified within the inhibitory block in 100.0% of the cases and excitatory cells within the excitatory block in 95.7%. Concerning the mixing of TTPC1 and TTPC2, we do not expect to be able to differentiate between these two types because their only difference is the distance of the bifurcation point of the apical dendrite to the soma. Since the MEA is located close to the somatic region, recordings might not be sensitive to this delicate difference. Disregarding this mixing, the m-type classification performs well (chance would be 9.1%) on excitatory cells and inhibitory Martinotti cells (MC) (80.5%). Note that these well-classified cells make up a large proportion of cortical cells. The overall accuracy of 34.0% illustrates that the morphological details are partially resolved by the CNN. In cases in which the CNN is not able to extract the information about the morphological details, it is unclear whether the information is present at all in the EAP or an increased number of cell models could solve the problem. In
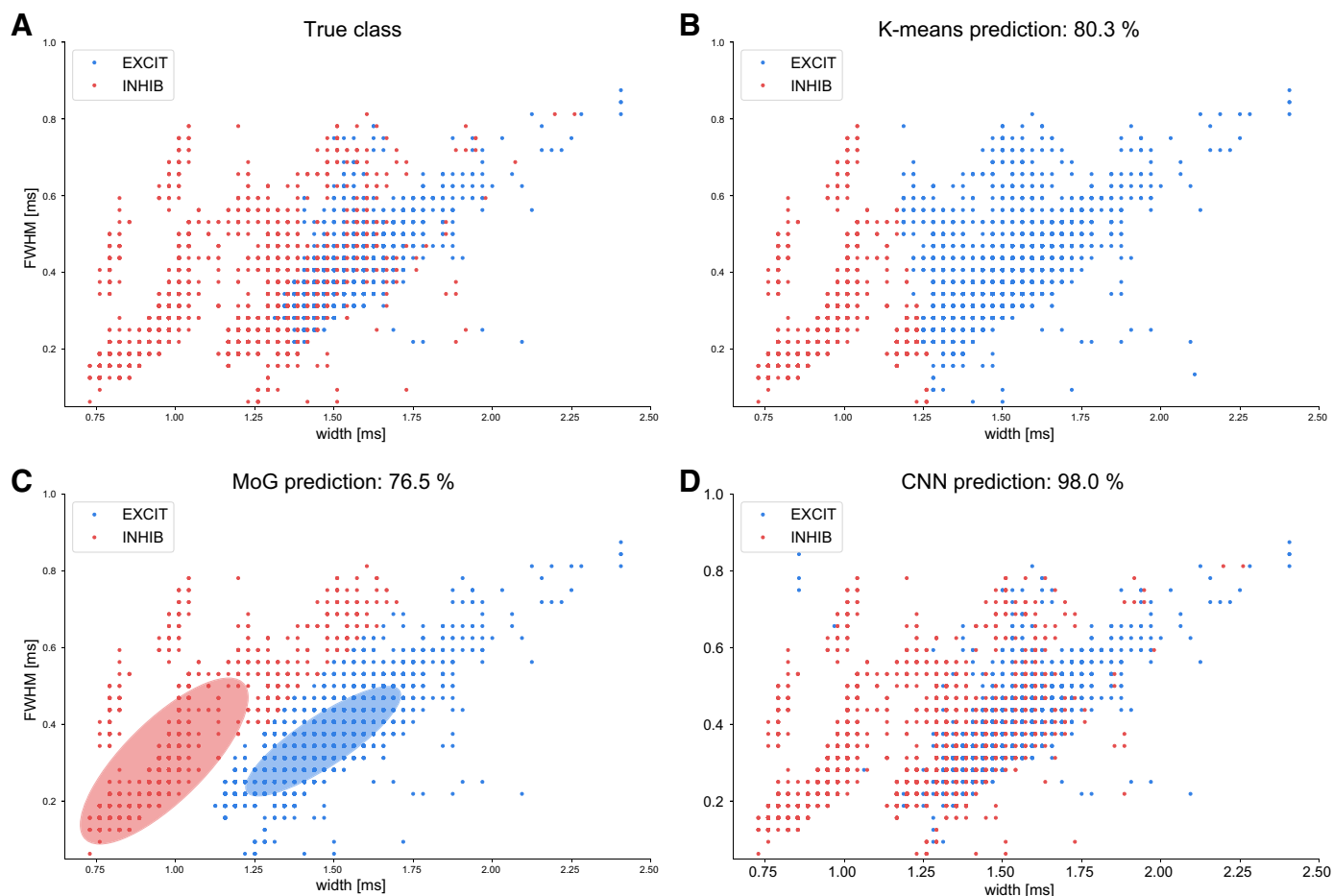
Fig. 10. Comparison with standard clustering approaches for excitatory/inhibitory classification. Each point is 1 spike in the *FW*-plane; red dots are inhibitory cells, and blue dots are excitatory cells. *A*: true excitatory/inhibitory classes: it is evident that there is some overlap between the groups in the center of the plot. *B*: *k*-means prediction: *k*-means clustering splits the data in 2 groups and cannot untangle the overlap (accuracy: 80.3%). *C*: mixture of Gaussians (MoG) prediction: MoG clustering correctly finds the diagonal shape of the excitatory population, but it cannot separate the overlap either (accuracy: 76.5%). *D*: convolutional neural network (CNN) prediction: the CNN output is able to correctly predict almost all the observations, with an accuracy of 98.0%. FWHM, full-width half-maximum.

conclusion, the results show promise for a more refined classification than only distinguishing excitatory cells from inhibitory cells.

Validation on Different Models

To investigate how general the trained CNN models are, we tested the performance of localization and classification on simulated EAPs from other neuronal models, namely, the cell model from Hay et al. (2011) and the models from the Allen Brain Institute (ABI) cell type database (Gouwens et al. 2018; http://celltypes.brain-map.org). For the following results, we used the SqMEA-10-15 probe, CNNs of size L, and NaRep and *FW* feature sets for localization and classification, respectively.

Hay model. The Hay cell models a neocortical pyramidal cell from L5b, and the techniques used to build the models were similar to the models from the NMC Portal. Therefore, we expect a relatively good performance in localization and classification with the CNNs trained on our standard NMC data sets. We built a Physrot data set of Hay cells consisting of 1,000 observations at random locations around the probe as described in *Simulated Recordings*, and we then evalu-

ated the performance of the CNNs in localization and classification.

For localization, the average error on the Hay data set is 8.7 ± 6.6 μm, perfectly in line with the average errors of TTPC models in the NMC validation data set (Fig. 7*A*). The average error over all cell types in the NMC validation data set is 7.8 ± 6.3 μm. For classification, we obtain an average accuracy of 76.4%, while the accuracy on the NMC validation set is 98.0%. The lower accuracy could be due to the fact that the Hay model includes other types of mechanisms, such as active calcium channels in the apical dendrites, that are not modeled in the NMC cell models.

Allen Brain Institute models. The cell models from the ABI that we selected are quite different from the NMC cell models at least for two reasons. First, the ABI neurons are from mice, whereas the NMC cells are from juvenile rats. Second, they are from visual cortex (19 cells) and postrhinal area (1 cell), whereas the NMC models are from somatosensory cortex. With CNNs trained on NMC data are expected to have lower accuracy when applied to the ABI data. We generated 1,000 EAPs for each of the 20 ABI cell models, according to the description in *Simulated Record-*
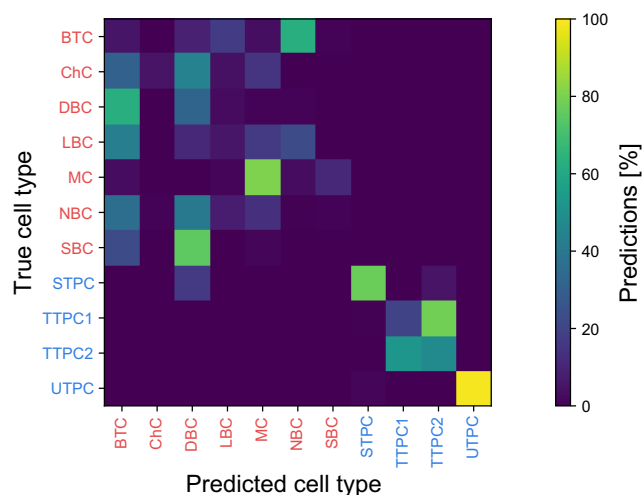
Fig. 11. Normalized confusion matrix of m-type classification based on Waveform features (here with downsampling factor 8, i.e., 28 sample points) for Physrot data set. The convolutional neural network size is L and extracellular action potentials are taken for the SqMEA-10-15 probe. BTC, bitufted cells; ChC, chandelier cells; DBC, double bouquet cells; LBC, large basket cells; MC, Martinotti cells; NBC, nest basket cells; SBC. small basket cells; STPC, slender-tufted pyramidal cells (PC); TTPC1, thick-tufted PC with late bifurcating apical tuft; TTPC2, thick-tufted PC with early bifurcating apical tuft; UTPC, untufted PC.

ings. We used the 3drot alignment because of the variability in the ABI cell models' orientation (for details about selection of cell models see APPENDIX A).

We ran the 3drot CNN for localization on the ABI data set, obtaining an average error of 19.3 ± 11.5 μm, larger than the 8.9 ± 8.2 μm obtained on the NMC validation set as expected. For classification, we distinguished excitatory and inhibitory cells in the ABI data set based on mouse transgenic lines (details in APPENDIX A). With the CNN for classification trained on NMC models, the average accuracy is 76.9%, while it is 97.6% on the NMC validation data set.

Since the cell models of the ABI come from a different species and are from a different cortical region, we trained a CNN on this data set only—16 models are used for training and 4 for validation (APPENDIX A). We used a CNN of size L and NaRep features, obtaining a localization average error of 5.9 ± 4.5 μm, which is in line with the performance we obtained on NMC models only. We did not run classification with so few models (only 20 cell models in total), because the CNNs need a larger diversity to find general features related to excitatory-inhibitory types (using the NMC data we trained on 192 cell models).

Test on Experimental Data

Although the method proposed in this report is at a proof-of-concept stage, we tested some CNNs trained with simulated data on experiments at least for plausibility.

We decided to use data (publicly available at http://www.kampff-lab.org/validating-electrodes) from paired juxtacellular and extracellular recordings (Neto et al. 2016) where, to a certain extent, the ground-truth location is known. The extracellular signals are measured with either the Neuronexus or the NeuroSeeker probe. Taking the amplitude threshold of our CNN training simulation (peak-to-peak amplitude of 30 μV on at least 1 electrode) into account and considering only

cells in front of the MEA, we were left with 10 data sets (see APPENDIX A for further details). After performing juxtacellular-triggered averaging, we fed the average EAP waveform into the CNN and predicted the soma position. The CNNs were trained with simulated data having the appropriate geometric alignment (MEA probes are rotated by $-48.2°$ along the *y*-axis). On average the prediction error is 42.2 ± 16.8 μm, assuming the true soma position is the tip of the juxtacellular probe. The experimentally determined positions for the *x*-, *y*-, and *z*-coordinates range from 27 μm to 129 μm, -48 μm to 6 μm, and -121 μm to 21 μm, respectively. Neto et al. (2016) report a distance uncertainty of 10.5 ± 5.2 μm. This uncertainty only applies to the tip position of the juxtacellular probe with respect to the MEA, but it is a drastic assumption to consider this position equal to the soma position (center of the soma). Without neglecting the soma diameter, one might favor a larger distance error by adding a soma radius uncertainty of 10–20 μm. Furthermore, the uncertainty of 10.5 ± 5.2 μm reflects misalignment errors caused by the manipulators used in the experiment and is investigated under free conditions, i.e., without entering any brain tissue. Additional misalignment originating from entering brain tissue with the probes or from the brain's pulsation due to breathing of the mouse are not taken into account.

Figure 12 shows predicted vs. true coordinates. Figure 12*A* demonstrates that the predicted soma distances (*x*-coordinate) are in a plausible range. Overall, the distance *x* is predicted with a mean error of 20.3 ± 16.1 μm. The *y*- and *z*-positions are in the same range of precision (22.3 ± 13.1 μm and 22.9 ± 15.7 μm, respectively, on average). Note that the horizontal error bars in the plots only represent the uncertainty due to the misalignment of the manipulators as reported by Neto et al. (2016) and all other previously mentioned uncertainties (which are not quantified) are not considered.

DISCUSSION

This work provides a deep learning approach for neuron localization and classification based on MEA recordings. We simulated in vivo-equivalent EAPs and built data sets for various probe designs, using a multitude of cell models from the NMC Portal (205 cell models from Ramaswamy et al. 2015). CNN models trained on these simulated spikes predict the soma position of the neuron and characterize whether it is excitatory or inhibitory. The accuracy depends on the neuron-MEA alignment, the specific cell types, the CNN size itself, and the input feature sets. For completeness, we compared the proposed method with existing strategies regarding both localization and classification of recorded spikes, we validated on cell models from other databases, and we tested the models on publicly available experimental data.

Localization

We showed that the CNN method is robust and accurate in predicting the 3D soma location from spikes generated by neurons with a physiological neuron-MEA alignment (Physrot, defined in *Neuron-MEA alignment*). The average errors are on the order of 7.6–11.7 μm for all probes involved in the study (Fig. 8*E*). We demonstrated the CNN approach to be robust with different cell models and to be able to generalize among cell types not used for training (BP and NGC). Finally, local-

Fig. 12. Soma position predictions of a convolutional neural network (CNN) based on experimental extracellular action potential (EAP) recordings. Experimental data are from paired juxtacellular-extracellular recordings (Neto et al. 2016) where the position of the soma is associated with the tip position of the juxtacellular probe. The CNN (size L, NaRep feature) is trained on simulated (3drot) EAP signals. Error bars are CNN prediction uncertainties for the predicted coordinates and 4.2 μm, 2.8 μm, and 8.5 μm (misalignment uncertainties reported by the experimenters of Neto et al. 2016) for true x-, y-, and z-coordinates, respectively.
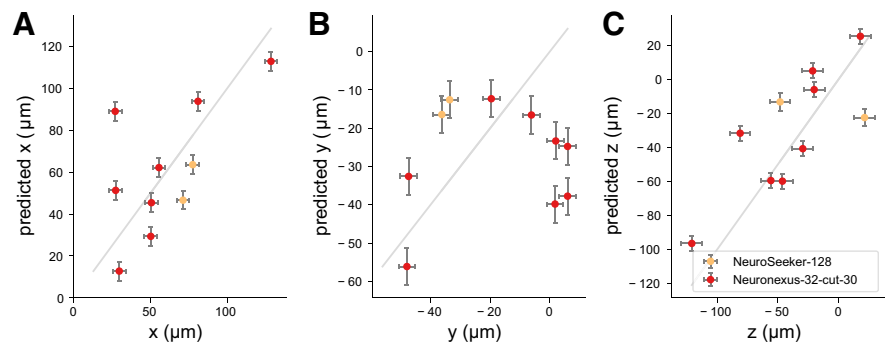
ization performances achieved with our approach are significantly better than solving the inverse problem with various generative models. With a SqMEA-10-15 data set the total error in three dimensions was 21.7 ± 20.9 μm for the monopolar current source, 15.6 ± 15.2 μm for the bipolar current source, and 22.6 ± 23.2 μm for the ball-and-stick model, whereas with our CNN approach we obtained an error of 7.8 ± 6.3 μm (as shown in Fig. 9).

In a recent study (Delgado Ruz and Schultz 2014), a Neuronexus-32 probe was used (shown in Fig. 3) with a modified ball-and-stick model to solve the inverse problem. For the five cell types considered in Delgado Ruz and Schultz (2014), they reached average errors of 6.26 ± 6.10 μm, 6.03 ± 7.68 μm, and 2.58 ± 4.75 μm along the x-, y-, and z-axes, respectively. Using the same probe on our Physrot data set, we obtained with CNNs average errors of 4.1 ± 4.5 μm, 4.3 ± 4.7 μm, and 4.3 ± 5.3 μm for the x-, y-, and z-axes, respectively (with NaRep feature set and size L).

Classification

The deep learning method was applied to excitatory/inhibitory classification with accuracies above 96.6% for all employed MEA models using the *FW* feature set and a CNN size L. An almost perfect outcome of 99.7% was obtained with the Waveform features on the SqMEA-10-15 probe. Compared with standard strategies using spike widths extracted from the spike shape, it showed a significant improvement (*k*-means clustering: 80.3%, MoG: 76.5%; *Comparison with Other Approaches*). We also attempted to distinguish among 11 cell morphologies (m-type classification). The overall accuracy of 34.0% is substantially better than the chance level of 9.1%. It is interesting to see that m-type classification performs a sort of unsupervised learning, as inhibitory cells were classified as inhibitory in 100.0% of the cases and excitatory cells as excitatory 95.7% of the time.

Overfitting and Stability

When evaluating the predictions of our CNNs on the validation data set, we observed a drop in accuracy compared with training accuracy. The drop is in an acceptable range for excitatory/inhibitory classification (0–3% with respect to the training accuracy) and localization (up to 3.7 μm prediction error increase). In case of m-type classification, the validation accuracy drops ~65% compared with training accuracy, clearly indicating overfitting. Since we do not have enough diversity in cell model data to build a third data set for implementing early-stopping regularization (i.e., stop training as soon as the

generalization error increases), we tracked the evaluation accuracies depending on the number of training epochs. In most cases, they reached a plateau after roughly 2,000 training epochs and did not decrease significantly afterwards, while training accuracies still increased. Therefore, we decided to stop training after 2,000 training epochs, assuming that the CNN has extracted most of the generalizable information provided by the EAPs at that point. Moreover, we tried to quantify the stability of the performance depending on different initial weights before the optimization process. To do so we ran the CNN training for localization and classification (on the SqMEA-10-15 probe, CNN size L, and with NaRep and *FW* features, respectively) six times with different random seeds. We obtained an average mean error of 7.6 ± 0.1 μm with an average SD of 6.3 ± 0.2 μm for localization (including the BP and NGC models) and an average mean accuracy of 97.9% with a SD of 0.2% for classification, indicating that performance is not dependent on the initial conditions of network weights and the convergence is robust.

Model-Based Approach

The findings presented in this study are based on simulations. Although this might be regarded as a limitation, we want to stress that the proposed method makes use of highly detailed cell models (Markram et al. 2015) and the complexity of such models is maintained and learned by CNNs. Previous approaches to localization and/or classification relied on simple forward models to solve the inverse models—monopolar, bipolar, ball-and-stick models, etc. (Blanche et al. 2005; Delgado Ruz and Schultz 2014; Somogyvári et al. 2012). We showed that CNNs outperform these models in estimating the soma positions. Another point that plays in favor of the use of neural simulations is the difficulty in gathering ground-truth data experimentally. Localizing and classifying neurons in real recordings requires advanced and highly accurate equipment, and the recorded labeled data would most likely still not be sufficient to train data-hungry machine learning algorithms such as CNNs. Nevertheless, validation on experimental data is definitely a required step and will be based on combined approaches with paired electrophysiological recordings and standard microscopy (Neto et al. 2016), or even involving more sophisticated and precise imaging techniques, such as two-photon imaging (Göbel and Helmchen 2007), which was paired with electrophysiological recordings in vivo in Shew et al. (2010). Paired electrophysiology and two-photon microscopy data, possibly in combination with intracellular voltage monitoring through patch clamping or voltage-sensitive dyes, could also represent a valuable tool to further validate and

improve the forward modeling schemes, providing morphological, intracellular, and extracellular recordings simultaneously.

Another advantage of using forward modeling is that the performance of the machine learning algorithm could be improved by building case-specific data sets that better match the real experimental scenarios. In this work, we assumed that the simulated probe was inserted in L5 of somatosensory cortex of a juvenile rat with a vertical insertion angle. However, somatosensory cortex can present large differences with respect to other brain regions (e.g., hippocampus, cerebellum, or other cortical regions) but also among animal species. Therefore, we do not envision a single universal model to localize and classify neurons but species- and brain area-specific CNNs to accurately deal with variability in neuronal types and functions. For example, when we fed mouse data from the ABI database, the localization CNN trained on rat data performed relatively poorly (19.3 ± 11.5 μm; *Allen Brain Institute models*), but trained on mouse models the performance is in line with what we obtained on rat data (5.9 ± 4.5 μm).

Effect of Probe Design

Regarding neural probes, a forward modeling-based approach can give important insights for the design and manufacturing of next-generation probes. For example, our results showed that even relatively low-density probes, such as the SqMEA-5-30, despite performing slightly worse than higher-density probes, still yield high accuracy in localizing and classifying neurons. Potentially, the pursuit of extremely high-density probes, which makes the design complicated and the data throughput very high, is not required for classification and localization tasks [although it might still be important for spike sorting (Franke et al. 2012; Rossant et al. 2016)]. However, for such simulation-driven MEA design, the simulations lack a more accurate electrode model considering finite size recording sites (in this work we used an ideal point electrode), electrode impedances, and transfer functions.

Future Extensions

The generative model for spike simulations could be improved in various ways. A straightforward improvement to obtain more accurate simulations could be including the MEA scar in the data generation, by clipping or bending neuronal branches in the proximity of the probe before simulating the recordings. Another refinement might be to take into account the finite size effects of the electrode contacts by means of the disk-electrode approximation (Lindén et al. 2014), which was shown to be appropriate for current sources positioned at distances larger than the contact radius (Ness et al. 2015). Moreover, here we assumed a tissue with homogeneous and isotropic electrical properties, but experimental findings suggest that in the cortex the conductive properties of the extracellular space are anisotropic (Goto et al. 2010). Anisotropy could be easily taken into account for the simulation of spikes (Ness et al. 2015; Pettersen et al. 2012). As the proposed approach strongly relies on high-fidelity simulations that reliably describe the neuron dynamics and volume conduction, another strategy could be using finite element method-based models, as in Agudelo-Toro and Neef (2013), Pods et al. (2013), and Tveito et al. (2017), which would result in more detailed simulations at the cost of a much higher computational

cost for data generation. Another layer of modeling is the electrode-tissue interface. The generated data should include electrical properties of the electrodes, such as the impedance, and account for their variability in experimental scenarios. In this work, we used polytrodes with a relevant size with respect to the neuron: although we assumed a homogeneous medium, the presence of the probe itself represents an obstacle for electrical signal propagation and can be modeled with either finite element method or analytical simplifications, such as the method of images (Ness et al. 2015).

In this work, we did not include any noise in the simulated recordings. The rationale behind this choice is that sorted spikes can be cleaned by applying spike-triggered averaging. With spike-triggered averaging, additive random noise is reduced by a factor of $\sqrt{N}$, where N is the number of occurrences of the sorted unit. Moreover, a common problem in spike sorting is electrode drift, in which the relative position between a neuron and the recording electrodes changes during the experiment. If drifting is detected from the spike sorting algorithm, one could feed different averaged EAPs computed in separate time windows and evaluate the drift over time, similarly to Delgado Ruz and Schultz (2014), in which windows of 5 min were used to compute the mean EAP.

Furthermore, the recording site area affects the amount of noise in the recordings, as the recording area is related to the impedance of the electrode. Here we assumed perfectly sorted spikes, from which a clean EAP can be computed. Clearly, with experimental data errors in spike sorting would affect the performance of localization and classification due to distorted waveforms from wrong assignments.

Outlook

Precise neural localization and classification from in vivo extracellular recordings has the potential of making electrophysiology an even more powerful technique to interact with neural tissue. Rather than only extracting spike trains, we could build a 3D representation of the recorded units and perform functional electrical imaging to study the spatial interactions among different cell types in neural microcircuits. On top of this, a precise localization of neuronal somata might enable the use of highly selective electrical stimulation patterns (Buccino et al. 2016) and represent an advancement in single-neuron stimulation from extracellular probes.

We strongly believe that computational approaches must go hand in hand with experimental ones, and an extension of this work might include the simulations of the entire pipeline from simulated MEA recordings, for example, with VISAPy (Hagen et al. 2015), to electrical stimulation including spike sorting, localization, classification, electrical stimulation, and evaluation of its effect on detailed neural morphologies.

APPENDIX A: DATA SELECTION

Neocortical Microcircuit Collaboration Portal Data Set

In this appendix we discuss the data set and the modifications that we applied to make sure that that training and validation set are completely disjointed.

In the original data set (https://bbp.epfl.ch/nmc-portal/welcome; L5 cells) there are nine inhibitory neuron types: BP, bitufted cells (BTC), ChC, double bouquet cells (DBC), LBC, MC, NBC, NGC, and SBC.

The four excitatory types, i.e., the PCs, are grouped into STPC, TTPC1, TTPC2, and UTPC. While belonging to the same m-type, neurons can have different electrophysiology properties (e-type) based on their firing patterns (Markram et al. 2015). In L5 the e-types are categorized into continuous accommodating (cAC), continuous stuttering (cSTUT), burst accommodating (bAC), burst stuttering (bSTUT), continuous nonaccommodating (cNAC), delayed stuttering (dSTUT), burst nonaccommodating (bNAC), continuous irregular (cIR), delayed nonaccommodating (dNAC), burst irregular (bIR), and continuous adapting (cAD). Since not all m-types express all e-types, the combination of morphological and electrical type gives rise to 52 morpho-electrical types (me-types) in L5. For each me-type, the NMC database contains five cell models; therefore, there are a total of 260 cell models in the data set.

In Markram et al. (2015), to extend the number of reconstructed models, an algorithm is used to clone morphologies: neural compartments are randomly scaled and rotated with respect to each other. Moreover, morphologies are also stretched and shrunk to make up new morphologies. We identified 54 different morphologies in the data set, listed in the Supplemental Material for this article. Although the cloned and/or scaled morphologies are indeed different than the original ones, their shape is quite similar.

The use of CNNs, which are among the most powerful machine learning algorithms, pushed us to pay particular attention in the training-validation splitting so that no information of the validation set is present in the training set (leakage). Hence, the presence of a cloned/scaled version of the same morphology in both training and validation has been avoided. We selected training and validation cell models so that all morphologies in the validation set are unique. In doing so, we had to remove all instances of BP and NGC from the training set, as all the models are derived from the same reconstructed morphology. For localization and excitatory/inhibitory classification, we kept a BP and an NGC model in the additional validation set.

After the manipulation, the training set consists of 192 cell models, while the validation set only contains 11 cell models, one for each m-type. Moreover, we use one BP and one NGC model, not used for training, as further validation. In total, we included 205 neuronal models out of the available 260. The cell models are listed in the Supplemental Material.

Allen Brain Institute Data Set

From the Allen Brain Institute cell type portal (http://celltypes.brain-map.org/data), we selected cell models according to three criteria: *1*) cells were from mice, *2*) cells were from L5 (to maintain consistency with the data from the NMC Portal), and *3*) cells had an all-active model. This search reduced the number of cell models available to 42. During the simulation process, we further discarded 22 models based on two extra rules: *1*) if adjusting the current-clamp amplitude to the soma could not induce a number of spikes between 10 and 30 in 10 iterations (in which the weight was multiplied by 0.75 if the number of intracellular spikes was >30 and by 1.25 when <10 spikes were detected) and *2*) if <5 EAP peaks had a peak-to-peak threshold of 30 μV in 500 random positioning of the neuron around the probe (meaning that the EAPs were mainly below the defined detection threshold). After this pruning, 20 cell models are left. To distinguish between excitatory and inhibitory cells, we used the transgenic line information: *Pvalb*, *Sst*, *Htr3*, and *Gad2* lines were considered inhibitory; *Rbp4*, *Scnn*, and *Rorb* were considered excitatory (Gouwens et al. 2018). After this division there were 11 inhibitory and 9 excitatory cell types.

To avoid overfitting, we randomly selected 4 models, 2 excitatory and 2 inhibitory, and we set them aside for validation, while we used the remaining 16 neuronal models for training.

The cell models are listed in the Supplemental Material.

Kampff Laboratory Data Set

Accompanying their article on paired juxtacellular-extracellular recordings (Neto et al. 2016), the laboratory of Adam Kampff publicly offers the data on http://www.kampff-lab.org/validating-electrodes. To extract an averaged extracellular waveform for each cell that can be fed into a trained CNN, some data processing was necessary. First, we detected spikes in the juxtacellular probe by thresholding the signal to get the cell's spike times. Second, we high-pass filtered the extracellular MEA recording with a third-order Butterworth filter in forward-backward mode with a band pass of 100–14,250 Hz. Afterwards, we averaged the EAP in windows of 7 ms around the spike times (2 ms preceding and 5 ms after the peak). This average waveform was then referred to as the juxtacellular-triggered average and was used as input for CNN predictions. After this preprocessing, 10 of the 29 available data sets fulfilled the criteria of having a peak-to-peak amplitude of 30 μV on at least one electrode and being in front of the extracellular probe (2014_03_26: Pair 2.0, 2014_03_20: Pair 3.0, 2014_03_26: Pair 2.1, 2014_10_17: Pair 1.1, 2014_10_17: Pair 1.0, 2014_11_25: Pair 3.0, 2014_11_25: Pair 1.0, 2014_11_25: Pair 2.0, 2015_09_04: Pair 5.0, 2015_09_03: Pair 6.0). These 10 data sets were used in *Test on Experimental Data* to test our deep learning approach.

APPENDIX B: ADDITIONAL INFORMATION

This appendix contains additional information on parameters and results.

Table B1 shows the specific CNN parameters for different network sizes.

The average localization errors and SDs for different rotational data sets are contained in Table B2, and the corresponding statistical analysis is depicted in Table B3.

Further results on significant differences and effect sizes of localization performances for different CNN sizes, features, MEA probes, and localization methods are listed in Tables B4, B5, B6, and B7, respectively.

The excitatory/inhibitory classification accuracies grouped by cell type for different rotational data sets, CNN sizes, feature sets, and MEA probes are shown in Tables B8, B9, B10, and B11, respectively.

Table B1. *CNN sizes*

Size	k_1	d_1	k_2	d_2	n_{FC}
XS	2	4	2	8	128
S	2	8	2	16	256
M	3	16	3	32	512
L	3	32	3	64	1,024
XL	3	64	3	128	2,048

Convolutional neural network (CNN) parameters of the different network sizes: layer 1 convolutional kernel size k_1, layer 1 convolutional kernel depth d_1, layer 2 convolutional kernel size k_2, layer 2 convolutional kernel depth d_2, and nodes in fully connected layer n_{FC}.

Table B2. *Localization errors grouped by rotational data set*

Data Set	x Error	y Error	z Error	Total Error
Norot	3.9 ± 3.9	3.0 ± 3.2	3.9 ± 4.6	7.3 ± 5.7
Physrot	3.8 ± 3.9	3.5 ± 3.8	4.3 ± 5.1	7.8 ± 6.3
3drot	4.4 ± 4.9	4.5 ± 5.7	4.4 ± 5.6	8.9 ± 8.2

Values (in μm) are average ± SD errors along x, y, and z dimensions and total errors grouped by rotational data sets. The average of total error is computed over the 3-dimensional distances and is not derived from the mean x, y, and z errors.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

A.P.B., M.K., T.V.B.N., S.R., and G.T.E. conceived and designed research; A.P.B., M.K., and T.V.B.N. performed experiments; A.P.B. and M.K. analyzed data; A.P.B. and M.K. interpreted results of experiments; A.P.B. and M.K. prepared figures; A.P.B. and M.K. drafted manuscript; A.P.B., M.K., T.V.B.N., B.M., P.D.H., M.F., G.C., S.R., and G.T.E. edited and revised manuscript; A.P.B., M.K., T.V.B.N., B.M., P.D.H., M.F., G.C., S.R., and G.T.E. approved final version of manuscript.

REFERENCES

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mane D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viegas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X. TensorFlow: large-scale machine learning on heterogeneous distributed systems (Preprint). ArXiv 1603.04467, 2016.

Agudelo-Toro A, Neef A. Computationally efficient simulation of electrical activity at cell membranes interacting with self-generated and externally imposed electric fields. *J Neural Eng* 10: 026019, 2013. doi:10.1088/1741-2560/10/2/026019.

Anastassiou CA, Perin R, Buzsáki G, Markram H, Koch C. Cell type- and activity-dependent extracellular correlates of intracellular spiking. *J Neurophysiol* 114: 608–623, 2015. doi:10.1152/jn.00628.2014.

Barthó P, Hirase H, Monconduit L, Zugaro M, Harris KD, Buzsáki G. Characterization of neocortical principal cells and interneurons by network interactions and extracellular features. *J Neurophysiol* 92: 600–608, 2004. doi:10.1152/jn.01170.2003.

Berdondini L, Bosca A, Nieus T, Maccione A. Active pixel sensor multielectrode array for high spatiotemporal resolution. In: *Nanotechnology and Neuroscience: Nano-electronic, Photonic and Mechanical Neuronal Inter-*

Table B3. *Localization by data set rotation: statistical analysis*

>	Norot	Physrot	3drot
Norot		ns	ns
Physrot	0.09***		ns
3drot	0.21***	0.13***	

Statistical analysis for localization errors grouped by rotational data set. Each entry shows Cohen's d and significance of the test column group < row group. ***$P < 0.001$. ns, Not significant.

Table B4. *Localization by CNN size: statistical analysis*

>	XS	S	M	L	XL
XS		0.26***	0.48***	0.63***	0.71***
S	ns		0.23***	0.38***	0.47***
M	ns	ns		0.14***	0.22***
L	ns	ns	ns		0.09***
XL	ns	ns	ns	ns	

Statistical analysis for localization errors grouped by convolutional neural network (CNN) size. Each entry shows Cohen's d and significance of the test column group < row group. ***$P < 0.001$. ns, Not significant.

facing, edited by Di Vittorio M, Martiradonna L, Assad J. New York: Springer, 2014, p. 207–238. doi:10.1007/978-1-4899-8038-0_7.

Blanche TJ, Spacek MA, Hetke JF, Swindale NV. Polytrodes: high-density silicon electrode arrays for large-scale multiunit recording. *J Neurophysiol* 93: 2987–3000, 2005. doi:10.1152/jn.01023.2004.

Buccino AP, Stöber T, Næss S, Cauwenberghs G, Häfliger P. Extracellular single neuron stimulation with high-density multi-electrode array. *2016 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. Shanghai, China, October 17–19, 2016, p. 520–523.

Carnevale NT, Hines ML. *The NEURON Book.* Cambridge, UK: Cambridge Univ. Press, 2006. doi:10.1017/CBO9780511541612.

Chelaru MI, Jog MS. Spike source localization with tetrodes. *J Neurosci Methods* 142: 305–315, 2005. doi:10.1016/j.jneumeth.2004.09.004.

Cohen J. A power primer. *Psychol Bull* 112: 155–159, 1992. doi:10.1037/0033-2909.112.1.155.

DeFelipe J, Ballesteros-Yáñez I, Inda MC, Muñoz A. Double-bouquet cells in the monkey and human cerebral cortex with special reference to areas 17 and 18. *Prog Brain Res* 154: 15–32, 2006. doi:10.1016/S0079-6123(06)54002-6.

Delgado Ruz I, Schultz SR. Localising and classifying neurons from high density MEA recordings. *J Neurosci Methods* 233: 115–128, 2014. doi:10.1016/j.jneumeth.2014.05.037.

Fortin FA, De Rainville FM, Gardner MA, Parizeau M, Gagné C. DEAP: evolutionary algorithms made easy. *J Mach Learn Res* 13: 2171–2175, 2012.

Foust A, Popovic M, Zecevic D, McCormick DA. Action potentials initiate in the axon initial segment and propagate through axon collaterals reliably in cerebellar Purkinje neurons. *J Neurosci* 30: 6891–6902, 2010. doi:10.1523/JNEUROSCI.0552-10.2010.

Franke F, Jäckel D, Dragas J, Müller J, Radivojevic M, Bakkum D, Hierlemann A. High-density microelectrode array recordings and real-time spike sorting for closed-loop experiments: an emerging technology to study neural plasticity. *Front Neural Circuits* 6: 105, 2012. doi:10.3389/fncir.2012.00105.

Göbel W, Helmchen F. In vivo calcium imaging of neural network function. *Physiology (Bethesda)* 22: 358–365, 2007.

Gold C, Henze DA, Koch C. Using extracellular action potential recordings to constrain compartmental models. *J Comput Neurosci* 23: 39–58, 2007. doi:10.1007/s10827-006-0018-2.

Gold C, Henze DA, Koch C, Buzsáki G. On the origin of the extracellular action potential waveform: a modeling study. *J Neurophysiol* 95: 3113–3128, 2006. doi:10.1152/jn.00979.2005.

Table B5. *Localization by feature type: statistical analysis*

>	Na	Rep	NaRep	Waveform
Na		0.0***	0.0**	0.28***
Rep	ns		ns	0.26***
NaRep	ns	0.0***		0.27***
Waveform	ns	ns	ns	

Statistical analysis for localization errors grouped by feature type. Each entry shows Cohen's d and significance of the test column group < row group. Na, extracellular action potential (EAP) negative peak (mainly attributed to sodium currents flowing into soma); Rep, EAP positive peak (associated with cell repolarization phase); NaRep, stacked version of Na and Rep; Waveform, downsampled EAP waveforms. ***$P < 0.001$, **$P < 0.01$. ns, Not significant.

Table B6. *Localization by MEA probe: statistical analysis*

>	Sq15-10	Sq10-15	Sq7-20	Sq6-25	Sq5-30	Neuronexus	NeuroSeeker	NeuroPixels
Sq15-10		ns	ns	ns	ns	ns	ns	ns
Sq10-15	0.04***		ns	ns	ns	ns	ns	ns
Sq7-20	0.08***	0.04**		ns	ns	ns	ns	ns
Sq6-25	0.1***	0.06***	0.02***		ns	ns	ns	ns
Sq5-30	0.13***	0.1***	0.06***	0.04***		0.0***	ns	ns
Neuronexus	0.13***	0.1***	0.06**	ns	ns		ns	ns
NeuroSeeker	0.25***	0.22***	0.18***	0.16***	0.13***	0.12***		ns
NeuroPixels	0.43***	0.41***	0.37***	0.36***	0.32***	0.3***	0.19***	

Statistical analysis for localization errors grouped by probe type. Each entry shows Cohen's d and significance of the test column group < row group. MEA, multielectrode array. ***$P < 0.001$, **$P < 0.01$. ns, Not significant.

Goldberg JH, Yuste R. Space matters: local and global dendritic Ca^{2+} compartmentalization in cortical interneurons. *Trends Neurosci* 28: 158–167, 2005. doi:10.1016/j.tins.2005.01.005.

Golding NL, Kath WL, Spruston N. Dichotomy of action-potential back-propagation in CA1 pyramidal neuron dendrites. *J Neurophysiol* 86: 2998–3010, 2001. doi:10.1152/jn.2001.86.6.2998.

Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press, 2016.

Goto T, Hatanaka R, Ogawa T, Sumiyoshi A, Riera J, Kawashima R. An evaluation of the conductivity profile in the somatosensory barrel cortex of Wistar rats. *J Neurophysiol* 104: 3388–3412, 2010. doi:10.1152/jn.00122.2010.

Gouwens NW, Berg J, Feng D, Sorensen SA, Zeng H, Hawrylycz MJ, Koch C, Arkhipov A. Systematic generation of biophysically detailed models for diverse cortical neuron types. *Nat Commun* 9: 710, 2018. doi:10.1038/s41467-017-02718-3.

Gulledge AT, Stuart GJ. Action potential initiation and propagation in layer 5 pyramidal neurons of the rat prefrontal cortex: absence of dopamine modulation. *J Neurosci* 23: 11363–11372, 2003. doi:10.1523/JNEUROSCI.23-36-11363.2003.

Hagen E, Ness TV, Khosrowshahi A, Sørensen C, Fyhn M, Hafting T, Franke F, Einevoll GT. ViSAPy: a Python tool for biophysics-based generation of virtual spiking activity for evaluation of spike-sorting algorithms. *J Neurosci Methods* 245: 182–204, 2015. doi:10.1016/j.jneumeth.2015.01.029.

Hay E, Hill S, Schürmann F, Markram H, Segev I. Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLOS Comput Biol* 7: e1002107, 2011. doi:10.1371/journal.pcbi.1002107.

Hines ML, Davison AP, Muller E. NEURON and Python. *Front Neuroinform* 3: 1, 2009. doi:10.3389/neuro.11.001.2009.

Holt GR, Koch C. Electrical interactions via the extracellular potential near cell bodies. *J Comput Neurosci* 6: 169–184, 1999. doi:10.1023/A:1008832702585.

Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, Lee AK, Anastassiou CA, Andrei A, Aydın Ç, Barbic M, Blanche TJ, Bonin V, Couto J, Dutta B, Gratiy SL, Gutnisky DA, Häusser M, Karsh B, Ledochowitsch P, Lopez CM, Mitelut C, Musa S, Okun M, Pachitariu M, Putzeys J, Rich PD, Rossant C, Sun WL, Svoboda K, Carandini M, Harris KD, Koch C, O'Keefe J, Harris TD. Fully integrated silicon probes for high-density recording of neural activity. *Nature* 551: 232–236, 2017. doi:10.1038/nature24636.

Kingma D, Ba J. Adam: a method for stochastic optimization (Preprint). ArXiv 1412.6980, 2014.

Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems 25*, edited by Pereira F, Burges CJ, Bottou L, Weinberger KQ. La Jolla, CA: NIPS, 2012, p. 1097–1105.

Kubo T, Katayama N, Karashima A, Nakao M. The 3D position estimation of neurons in the hippocampus based on the multi-site multi-unit recordings with silicon tetrodes. *Conf Proc IEEE Eng Med Biol Soc* 2008: 5021–5024, 2008. doi:10.1109/IEMBS.2008.4650341.

Lindén H, Hagen E, Łęski S, Norheim ES, Pettersen KH, Einevoll GT. LFPy: a tool for biophysical simulation of extracellular potentials generated by detailed model neurons. *Front Neuroinform* 7: 41, 2014. doi:10.3389/fninf.2013.00041.

Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat* 18: 50–60, 1947. doi:10.1214/aoms/1177730491.

Markram H, Muller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA, Ailamaki A, Alonso-Nanclares L, Antille N, Arsever S, Kahou GA, Berger TK, Bilgili A, Buncic N, Chalimourda A, Chindemi G, Courcol JD, Delalondre F, Delattre V, Druckmann S, Dumusc R, Dynes J, Eilemann S, Gal E, Gevaert ME, Ghobril JP, Gidon A, Graham JW, Gupta A, Haenel V, Hay E, Heinis T, Hernando JB, Hines M, Kanari L, Keller D, Kenyon J, Khazen G, Kim Y, King JG, Kisvarday Z, Kumbhar P, Lasserre S, Le Bé JV, Magalhães BR, Merchán-Pérez A, Meystre J, Morrice BR, Muller J, Muñoz-Céspedes A, Muralidhar S, Muthurasa K, Nachbaur D, Newton TH, Nolte M, Ovcharenko A, Palacios J, Pastor L, Perin R, Ranjan R, Riachi I, Rodríguez JR, Riquelme JL, Rössert C, Sfyrakis K, Shi Y, Shillcock JC, Silberberg G, Silva R, Tauheed F, Telefont M, Toledo-Rodriguez M, Tränkler T, Van Geit W, Díaz JV, Walker R, Wang Y, Zaninetta SM, DeFelipe J, Hill SL, Segev I, Schürmann F. Reconstruction and simulation

Table B7. *Localization with different methods: statistical analysis*

>	Monopolar	Bipolar	B-A-S	CNN
Monopolar		0.33***	ns	0.9***
Bipolar	ns		ns	0.68***
B-A-S	0.04*	0.36***		0.87***
CNN	ns	ns	ns	

Statistical analysis for localization errors grouped by localization method. Each entry shows Cohen's d and significance of the test column group < row group. B-A-S, ball and stick; CNN, convolutional neural network. ***$P < 0.001$, *$P < 0.05$. ns, Not significant.

Table B8. *Classification accuracies by rotation*

	Norot	Physrot	3drot
BTC	98.4	99.3	99.5
ChC	96.0	84.8	84.1
DBC	100.0	99.3	99.1
LBC	100.0	99.5	97.9
MC	100.0	100.0	100.0
NBC	99.5	98.1	98.2
SBC	100.0	98.9	98.4
STPC	97.4	96.0	96.5
TTPC1	92.6	100.0	99.5
TTPC2	100.0	100.0	99.1
UTPC	98.7	99.8	98.7
Average	98.1	98.0	97.6
SD	2.4	3.9	3.9

Values are excitatory/inhibitory classification accuracy (in %) grouped by rotation and cell type. BTC, bitufted cells; ChC, chandelier cells; DBC, double bouquet cells; LBC, large basket cells; MC, Martinotti cells; NBC, nest basket cells; SBC. small basket cells; STPC, slender-tufted pyramidal cells (PC); TTPC1, thick-tufted PC with late bifurcating apical tuft; TTPC2, thick-tufted PC with early bifurcating apical tuft; UTPC, untufted PC.

Table B9. *CNN size classification performance*

	XS	S	M	L	XL
BTC	96.8	98.4	99.3	99.3	98.8
ChC	73.6	79.5	97.0	84.8	87.4
DBC	97.4	99.1	98.9	99.3	98.6
LBC	96.1	98.1	99.6	99.5	98.6
MC	99.5	100.0	99.8	100.0	100.0
NBC	94.2	96.5	96.1	98.1	97.5
SBC	96.7	98.2	97.9	98.9	98.9
STPC	100.0	100.0	98.7	96.0	88.5
TTPC1	100.0	99.8	99.4	100.0	99.9
TTPC2	99.8	99.4	99.6	100.0	99.9
UTPC	97.6	97.9	97.8	99.8	99.8
Average	96.4	97.4	98.6	98.0	97.1
SD	6.6	5.1	1.1	3.9	4.5

Values are excitatory/inhibitory classification accuracy (in %) grouped by size and cell type. BTC, bitufted cells; ChC, chandelier cells; DBC, double bouquet cells; LBC, large basket cells; MC, Martinotti cells; NBC, nest basket cells; SBC. small basket cells; STPC, slender-tufted pyramidal cells (PC); TTPC1, thick-tufted PC with late bifurcating apical tuft; TTPC2, thick-tufted PC with early bifurcating apical tuft; UTPC, untufted PC.

of neocortical microcircuitry. *Cell* 163: 456–492, 2015. doi:10.1016/j.cell. 2015.09.029.

Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci* 5: 793–807, 2004. doi:10.1038/nrn1519.

McCormick DA, Connors BW, Lighthall JW, Prince DA. Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J Neurophysiol* 54: 782–806, 1985. doi:10.1152/jn.1985.54.4.782.

Mechler F, Victor JD. Dipole characterization of single neurons from their extracellular action potentials. *J Comput Neurosci* 32: 73–100, 2012. doi:10.1007/s10827-011-0341-0.

Mechler F, Victor JD, Ohiorhenuan I, Schmid AM, Hu Q. Three-dimensional localization of neurons in cortical tetrode recordings. *J Neurophysiol* 106: 828–848, 2011. doi:10.1152/jn.00515.2010.

Migliore M, Hoffman DA, Magee JC, Johnston D. Role of an A-type K^+ conductance in the back-propagation of action potentials in the dendrites of hippocampal pyramidal neurons. *J Comput Neurosci* 7: 5–15, 1999. doi:10.1023/A:1008906225285.

Müller J, Ballini M, Livi P, Chen Y, Radivojevic M, Shadmani A, Viswam V, Jones IL, Fiscella M, Diggelmann R, Stettler A, Frey U, Bakkum DJ, Hierlemann A. High-resolution CMOS MEA platform to study neurons at subcellular, cellular, and network levels. *Lab Chip* 15: 2767–2780, 2015. doi:10.1039/C5LC00133A.

Ness TV, Chintaluri C, Potworowski J, Łęski S, Głąbska H, Wójcik DK, Einevoll GT. Modelling and analysis of electrical potentials recorded in microelectrode arrays (MEAs). *Neuroinformatics* 13: 403–426, 2015. doi:10.1007/s12021-015-9265-6.

Neto JP, Lopes G, Frazão J, Nogueira J, Lacerda P, Baião P, Aarts A, Andrei A, Musa S, Fortunato E, Barquinha P, Kampff AR. Validating silicon polytrodes with paired juxtacellular recordings: method and dataset. *J Neurophysiol* 116: 892–903, 2016. doi:10.1152/jn.00013.2016.

Nunez PL, Srinivasan R. *Electric Fields of the Brain: the Neurophysics of EEG*. New York: Oxford Univ. Press, 2006. doi:10.1093/acprof:oso/9780195050387.001.0001.

Overstreet-Wadiche L, McBain CJ. Neurogliaform cells in cortical circuits. *Nat Rev Neurosci* 16: 458–468, 2015. doi:10.1038/nrn3969.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: machine learning in Python. *J Mach Learn Res* 12: 2825–2830, 2011.

Pettersen KH, Einevoll GT. Amplitude variability and extracellular low-pass filtering of neuronal spikes. *Biophys J* 94: 784–802, 2008. doi:10.1529/biophysj.107.111179.

Pettersen KH, Lindén H, Dale AM, Einevoll GT. Extracellular spikes and CSD. In: *Handbook of Neural Activity Measurement*, edited by Brette R, Destexhe A. Cambridge, UK: Cambridge Univ. Press, 2012, p. 92–135. doi:10.1017/CBO9780511979958.004.

Peyrache A, Dehghani N, Eskandar EN, Madsen JR, Anderson WS, Donoghue JA, Hochberg LR, Halgren E, Cash SS, Destexhe A. Spatio-temporal dynamics of neocortical excitation and inhibition during human sleep. *Proc Natl Acad Sci USA* 109: 1731–1736, 2012. doi:10.1073/pnas.1109895109.

Pods J, Schönke J, Bastian P. Electrodiffusion models of neurons and extracellular space using the Poisson-Nernst-Planck equations—numerical simulation of the intra- and extracellular potential for an axon model. *Biophys J* 105: 242–254, 2013. [Erratum in *Biophys J* 106: 769–770, 2014.] doi:10.1016/j.bpj.2013.05.041.

Rall W, Shepherd GM. Theoretical reconstruction of field potentials and dendrodendritic synaptic interactions in olfactory bulb. *J Neurophysiol* 31: 884–915, 1968. doi:10.1152/jn.1968.31.6.884.

Ramaswamy S, Courcol JD, Abdellah M, Adaszewski SR, Antille N, Arsever S, Atenekeng G, Bilgili A, Brukau Y, Chalimourda A, Chindemi G, Delalondre F, Dumusc R, Eilemann S, Gevaert ME, Gleeson P, Graham JW, Hernando JB, Kanari L, Katkov Y, Keller D, King JG, Ranjan R, Reimann MW, Rössert C, Shi Y, Shillcock JC, Telefont M, Van Geit W, Diaz JV, Walker R, Wang Y, Zaninetta SM, DeFelipe J, Hill SL, Muller J, Segev I, Schürmann F, Muller EB, Markram H. The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex. *Front Neural Circuits* 9: 44, 2015. doi:10.3389/fncir.2015.00044.

Rossant C, Kadir SN, Goodman DF, Schulman J, Hunter ML, Saleem AB, Grosmark A, Belluscio M, Denfield GH, Ecker AS, Tolias AS, Solomon S, Buzsáki G, Carandini M, Harris KD. Spike sorting for large, dense electrode arrays. *Nat Neurosci* 19: 634–641, 2016. doi:10.1038/nn.4268.

Sasaki T, Matsuki N, Ikegaya Y. Effects of axonal topology on the somatic modulation of synaptic outputs. *J Neurosci* 32: 2868–2876, 2012. doi:10.1523/JNEUROSCI.5365-11.2012.

Schröder S, Cecchetto C, Keil S, Mahmud M, Brose E, Dogan Ö, Bertotti G, Wolanski D, Tillack B, Schneidewind J, Gargouri H, Arens M, Bruns J, Szyszka B, Vassanelli S, Thewes R. CMOS-compatible purely capacitive interfaces for high-density in-vivo recording from neural tissue. *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. Atlanta, GA, October 22–24, 2015, p. 1–4.

Shew WL, Bellay T, Plenz D. Simultaneous multi-electrode array recording and two-photon calcium imaging of neural activity. *J Neurosci Methods* 192: 75–82, 2010. doi:10.1016/j.jneumeth.2010.07.023.

Somogyvári Z, Cserpán D, Ulbert I, Érdi P. Localization of single-cell current sources based on extracellular potential patterns: the spike CSD method. *Eur J Neurosci* 36: 3299–3313, 2012. doi:10.1111/j.1460-9568.2012.08249.x.

Somogyvári Z, Zalányi L, Ulbert I, Érdi P. Model-based source localization of extracellular action potentials. *J Neurosci Methods* 147: 126–137, 2005. doi:10.1016/j.jneumeth.2005.04.002.

Table B10. *Feature classification performance*

	AW	FW	AFW	Waveform
BTC	99.8	98.4	99.3	100.0
ChC	91.6	79.3	97.0	100.0
DBC	97.7	98.4	98.9	100.0
LBC	99.3	97.7	99.6	100.0
MC	100.0	100.0	99.8	100.0
NBC	91.9	93.7	96.1	100.0
SBC	99.6	98.4	97.9	99.8
STPC	99.9	95.6	98.7	98.1
TTPC1	99.5	100.0	99.4	100.0
TTPC2	99.7	99.9	99.6	100.0
UTPC	97.0	99.8	97.8	100.0
Average	98.1	97.0	98.6	99.7
SD	2.8	5.3	1.1	0.6

Values are excitatory/inhibitory classification accuracy (in %) grouped by feature set and cell type. *W*, peak-to-peak width; *F*, full-width half-maximum; *A*, peak-to-peak amplitude; BTC, bitufted cells; ChC, chandelier cells; DBC, double bouquet cells; LBC, large basket cells; MC, Martinotti cells; NBC, nest basket cells; SBC. small basket cells; STPC, slender-tufted pyramidal cells (PC); TTPC1, thick-tufted PC with late bifurcating apical tuft; TTPC2, thick-tufted PC with early bifurcating apical tuft; UTPC, untufted PC.

Table B11.   *MEA probe classification performance*

	Sq15-10	Sq10-15	Sq7-20	Sq6-25	Sq5-30	Neuronexus	NeuroSeeker	NeuroPixels
BTC	98.9	99.3	98.4	98.9	99.1	99.1	99.6	98.9
ChC	83.0	84.8	81.8	84.4	89.1	93.3	91.4	91.2
DBC	97.7	99.3	97.4	97.9	99.1	99.5	99.1	99.1
LBC	97.9	99.5	98.9	98.1	98.9	99.3	98.9	98.9
MC	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
NBC	95.3	98.1	93.5	93.7	98.1	99.1	96.3	96.7
SBC	97.0	98.9	97.7	97.5	99.8	99.1	98.6	99.1
STPC	90.6	96.0	93.4	90.6	84.1	83.6	97.0	98.4
TTPC1	100.0	100.0	100.0	99.7	99.8	100.0	99.8	100.0
TTPC2	100.0	100.0	100.0	99.6	99.6	99.6	99.8	100.0
UTPC	99.7	99.8	99.9	99.7	98.8	98.5	99.7	99.6
Average	96.6	98.0	96.9	96.6	96.7	97.0	98.4	98.6
SD	4.8	3.9	4.8	4.6	5.4	5.3	2.3	2.2

Values are excitatory/inhibitory classification accuracy (in %) grouped by probe and cell type. MEA, multielectrode array; BTC, bitufted cells; ChC, chandelier cells; DBC, double bouquet cells; LBC, large basket cells; MC, Martinotti cells; NBC, nest basket cells; SBC. small basket cells; STPC, slender-tufted pyramidal cells (PC); TTPC1, thick-tufted PC with late bifurcating apical tuft; TTPC2, thick-tufted PC with early bifurcating apical tuft; UTPC, untufted PC.

Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 15: 1929–1958, 2014.

Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. *J Grad Med Educ* 4: 279–282, 2012. doi:10.4300/JGME-D-12-00156.1.

Tveito A, Jæger KH, Lines GT, Paszkowski Ł, Sundnes J, Edwards AG, Māki-Marttunen T, Halnes G, Einevoll GT. An evaluation of the accuracy of classical models for computing the membrane potential and extracellular potential for neurons. *Front Comput Neurosci* 11: 27, 2017. doi:10.3389/fncom.2017.00027.

Vassanelli S. Multielectrode and multitransistor arrays for in vivo recording. In: *Nanotechnology and Neuroscience: Nano-electronic, Photonic and Mechanical Neuronal Interfacing*, edited by De Vittorio M, Martiradonna L, Assad J. New York: Springer, 2014, p. 239–267. doi:10.1007/978-1-4899-8038-0_8.

Wang Y, Gupta A, Toledo-Rodriguez M, Wu CZ, Markram H. Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex. *Cereb Cortex* 12: 395–410, 2002. doi:10.1093/cercor/12.4.395.

Wang Y, Toledo-Rodriguez M, Gupta A, Wu C, Silberberg G, Luo J, Markram H. Anatomical, physiological and molecular properties of Martinotti cells in the somatosensory cortex of the juvenile rat. *J Physiol* 561: 65–90, 2004. doi:10.1113/jphysiol.2004.073353.

Waters J, Schaefer A, Sakmann B. Backpropagating action potentials in neurones: measurement, mechanisms and potential functions. *Prog Biophys Mol Biol* 87: 145–170, 2005. doi:10.1016/j.pbiomolbio.2004.06.009.

Welkenhuysen M, Hoffman L, Luo Z, De Proft A, Van den Haute C, Baekelandt V, Debyser Z, Gielen G, Puers R, Braeken D. An integrated multi-electrode-optrode array for in vitro optogenetics. *Sci Rep* 6: 20353, 2016. doi:10.1038/srep20353.

Woodruff A, Yuste R. Of mice and men, and chandeliers. *PLoS Biol* 6: e243, 2008. doi:10.1371/journal.pbio.0060243.

Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. In: *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings Part I*, edited by Fleet D, Pajdla T, Schiele B, Tuytelaars T. Cham, Switzerland: Springer International, 2014, p. 818–833.