A deep learning framework for automated and generalized synaptic event analysis

Philipp S. O'Neill[1,2], Martín Baccino-Calace[1], Peter Rupprecht[2,3], Rainer W. Friedrich[4,5], Martin Müller[1,2,6], and Igor Delvendahl[1,2,*]

[1]Department of Molecular Life Sciences, University of Zurich (UZH), 8057 Zurich, Switzerland
[2]Neuroscience Center Zurich, 8057 Zurich, Switzerland
[3]Brain Research Institute, University of Zurich, 8057 Zurich, Switzerland
[4]Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland
[5]Faculty of Natural Sciences, University of Basel, 4003 Basel, Switzerland
[6]University Research Priority Program (URPP), Adaptive Brain Circuits in Development and Learning (AdaBD), University of Zurich, 8057 Zurich, Switzerland
[*]Correspondence: igor.delvendahl@uzh.ch

ABSTRACT

Quantitative information about synaptic transmission is key to our understanding of neural function. Spontaneously occurring synaptic events carry fundamental information about synaptic function and plasticity. However, their stochastic nature and low signal-to-noise ratio present major challenges for the reliable and consistent analysis. Here, we introduce miniML, a supervised deep learning-based method for accurate classification and automated detection of spontaneous synaptic events. Comparative analysis using simulated ground-truth data shows that miniML outperforms existing event analysis methods in terms of both precision and recall. miniML enables precise detection and quantification of synaptic events in electrophysiological recordings. We demonstrate that the deep learning approach generalizes easily to diverse synaptic preparations, different electrophysiological and optical recording techniques, and across animal species. miniML provides not only a comprehensive and robust framework for automated, reliable, and standardized analysis of synaptic events, but also opens new avenues for high-throughput investigations of neural function and dysfunction.

Keywords: Synaptic transmission, Machine learning, Data analysis, Neurons, Electrophysiology, Imaging

INTRODUCTION

Synaptic communication serves as the fundamental basis for a wide spectrum of brain functions, from computation and sensory integration to learning and memory. Synaptic transmission either arises from spontaneous or action potential-evoked fusion of neurotransmitter-filled synaptic vesicles (Kaeser and Regehr, 2014) resulting in an electrical response in the postsynaptic cell. Such synaptic events are a salient feature of all neural circuits and can be recorded using electrophysiological or imaging techniques.

Random fluctuations in the release machinery or intracellular Ca^{2+} concentration cause spontaneous fusions of single vesicles ('miniature events') (Kavalali, 2015), which play an important role in synaptic development and stability (Banerjee et al., 2021; Kaeser and Regehr, 2014; Kavalali, 2015; McKinney et al., 1999). Measurements of amplitude, kinetics, and timing of these events provide essential information about the function of individual synapses and neural circuits. Miniature events are therefore key to our understanding of fundamental processes, such as synaptic plasticity or synaptic computation that support neural function (Abbott and Regehr, 2004; Holler et al., 2021). For example, amplitude changes of miniature events are a proxy of neurotransmitter receptor modulation, which is thought to be the predominant mechanism driving activity-dependent long-term alterations in synaptic strength (Huganir and Nicoll, 2013; Malinow and Malenka, 2002) and homeostatic synaptic plasticity (O'Brien et al., 1998; Turrigiano et al., 1998). In addition to spontaneous vesicle fusions, synaptic events may also result from presynaptic action potentials in neural networks. Alterations in spontaneous neurotransmission have been observed in models of different neurodevelopmental and neurodegenerative disorders (Alten et al.,

2021; Ardiles et al., 2012; Miller et al., 2014). A comprehensive analysis of synaptic events is thus paramount for studying synaptic function and for understanding neural diseases. However, the detection and quantification of synaptic events in electrophysiological or fluorescence recordings remains a major challenge. Synaptic events are often small in size, resulting in a low signal-to-noise ratio (SNR), and their stochastic occurrence further complicates reliable detection and evaluation.

Several techniques have been developed for synaptic event detection: Finite-difference approaches use crossings of a predefined threshold, typically in either raw data (Kim et al., 2021), baseline-normalized data (Kudoh and Taguchi, 2002), or their derivatives (Ankri et al., 1994). Template-based methods use a predefined template and generate a matched filter via a scaling factor (Clements and Bekkers, 1997; Jonas et al., 1993), deconvolution (Pernía-Andrade et al., 2012) or an optimal filtering approach (Shi et al., 2010; X. Zhang et al., 2021). In addition, Bayesian inference can be used to detect synaptic events (Merel et al., 2016). These techniques have facilitated the analysis of synaptic events. However, they also have several relevant limitations, such as a strong dependence of detection performance on a threshold and other decisive hyperparameters, or the need for visual inspection of results by an experienced investigator to avoid false positives. The widespread use of synaptic event recordings and the difficulty in obtaining results that are reliable across investigators and laboratories highlight the need for an automated, accurate, efficient, and reproducible synaptic event analysis method.

Artificial intelligence (AI) technologies such as deep learning (LeCun et al., 2015) can significantly enhance biological data analysis (Richards et al., 2022) and thus contribute to a better understanding of neural function. Convolutional neural networks (CNNs) are especially effective for image classification, but can also be used for one-dimensional data (Fawaz et al., 2019; Wang et al., 2016). CNNs have been successfully applied in neuroscience to segment brain regions (Iqbal et al., 2019), detect synaptic vesicles in electron microscopy images (Imbrosci et al., 2022), identify spikes in Ca^{2+} imaging data (Rupprecht, Carta, et al., 2021), and localize neurons in brain slices (Yip et al., 2021), or neurons with fluorescence signals in time-series data (Denis et al., 2020; Sitá et al., 2022).

Here, we present *miniML*, a novel approach for detecting and analyzing spontaneous synaptic events using supervised deep learning. The miniML model is an end-to-end classifier trained on an extensive dataset of annotated synaptic events. When applied to time-series data, miniML provides high performance event detection with negligible false-positive rates, outperforming existing methods. The method is fast, virtually parameter-free, threshold-independent, and generalizable across diverse data types, allowing for efficient and reproducible analysis, even for large datasets. We anticipate that AI-based synaptic event analysis will greatly enhance the investigation of synaptic function, plasticity, and computation from individual synapse to network level.

RESULTS

miniML enables highly accurate classification of synaptic events

To investigate whether an AI model can detect stochastic synaptic events in inherently noisy time-series data, we designed a deep neural network consisting of CNN, long short-term memory (LSTM), and fully connected dense layers ('miniML', Figure 1A). The CNN-LSTM model takes a section of a univariate time-series recording as input and outputs a label for that section of data (Figure 1B). The miniML model is trained to classify short segments of electrophysiological data as either positive or negative for a synaptic event using supervised learning. The trained classifier can then be applied to unseen time-series data to localize events.

To train the miniML model, we first extracted a large number of synaptic events and corresponding event-free sections from previous voltage-clamp recordings of cerebellar mossy fiber to granule cell (MF–GC) miniature excitatory postsynaptic currents (mEPSCs) (Delvendahl et al., 2019). All samples were then visually inspected and labeled to generate the training dataset. We applied data augmentation techniques to include examples of typical false positives in the training data (Materials and Methods). In total, the training data comprised ~30,000 samples that were split into training and validation sets (0.75/0.25). Across training epochs, loss decreased and accuracy increased, stabilizing after ~30 epochs (Figure 1C). The model with the highest validation accuracy was selected for further use, achieving 98.4% (SD 0.1, fivefold cross-validation). Saliency map analysis (Simonyan et al., 2013) indicated that the AI model mainly relied on the data sections around the peak of synaptic events to discriminate with respect to the labels (Figure S1). The trained miniML model achieved an area under the receiver operating characteristic (ROC) curve close to 1 (Figure 1D), indicating almost perfect separability of the

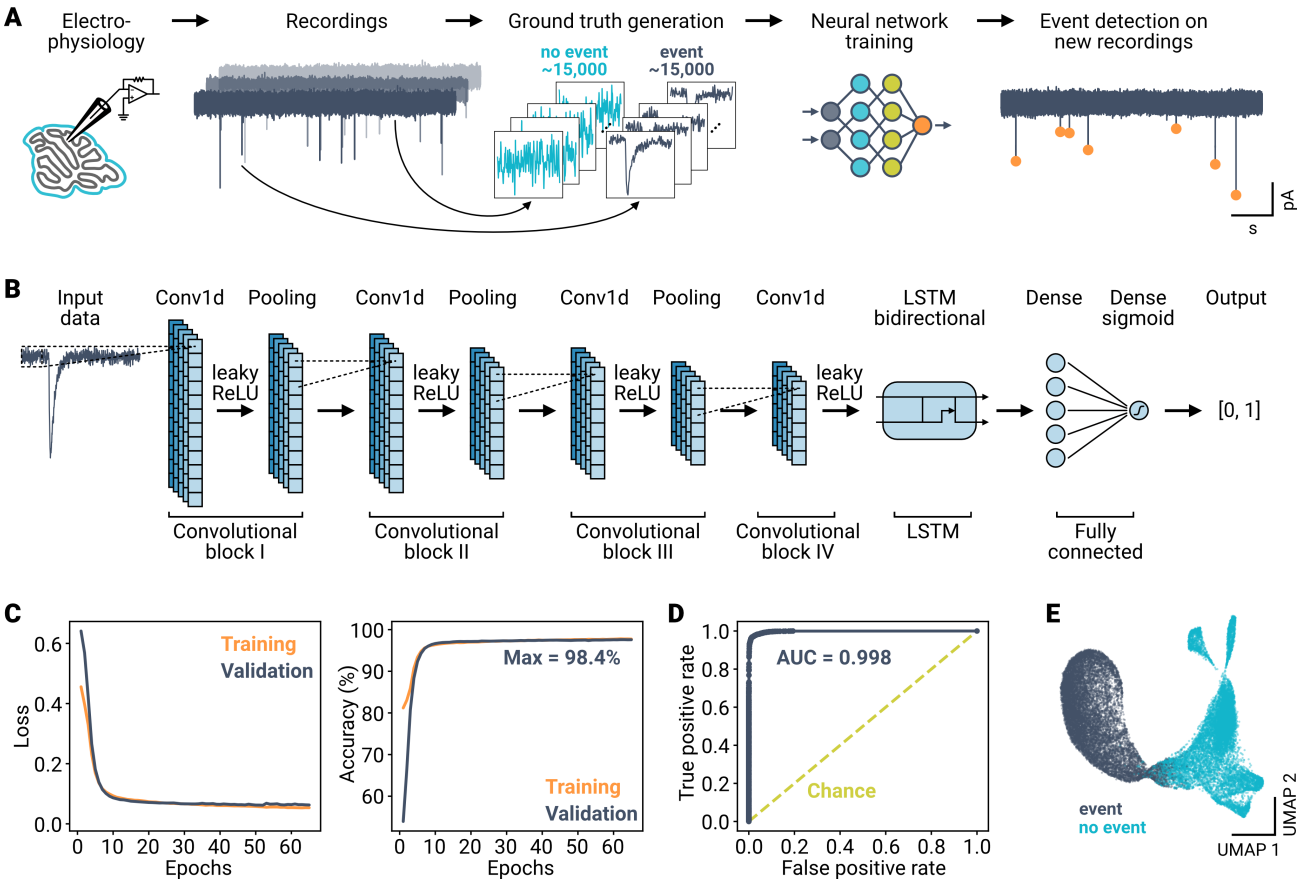

Figure 1. High performance classification of synaptic events using a deep neural network.
(**A**) Overview of the analysis workflow. Data segments from electrophysiological recordings are extracted and labeled to train an artificial neural network. The deep learning-based model is then applied to detect events in novel time-series data. (**B**) Schematic of the model design. Data is input to a convolutional network consisting of blocks of 1D convolutional, ReLU, and average pooling layers. The output of the convolutional layers is processed by a bidirectional LSTM block, followed by two fully connected layers. The final output is a label in the interval [0, 1]. (**C**) Loss (binary crossentropy) and accuracy of the model over training epochs for training and validation data. (**D**) Receiver operating characteristic of the best performing model. Area under the curve (AUC) is indicated; dashed line indicates performance of a random classifier. (**E**) UMAP representation of the training data as input to the final layer of the model, indicating linear separability of the two event classes after model training. Also see Figure S1.

classes (Figure 1E, Figure S1). Deep learning typically requires large datasets for training (van der Ploeg et al., 2014). To investigate how miniML's classification performance depended on the dataset size, we systematically increased the number of training samples. As expected, the accuracy increased with larger datasets. However, the performance gain was marginal when exceeding 5,000 samples (<0.2%, Figure S2), indicating that relatively small datasets suffice for effective model training (Bailly et al., 2022). These results demonstrate the efficacy of supervised deep learning in accurately classifying segments of neurophysiological data containing synaptic events.

The miniML model robustly detects synaptic events in electrophysiological recordings

Synaptic events are typically recorded as continuous time-series data of either membrane voltage or current. To apply the trained classifier to detect events in arbitrarily long data, we used a sliding window approach (Figure 2A). Time-series data are divided into sections corresponding to the input shape of the CNN-LSTM classifier, using a given stride, which reduces the number of inferences needed and speeds up computation time while maintaining high detection performance (Figure S3). By reshaping the data into overlapping sections, model inference can be run in batches and employ parallel processing techniques, including graphics processing unit (GPU) computing, resulting in analysis times of a few seconds for minute-long recordings (Figure S3). The miniML detection method is thus time-efficient and can be easily integrated into (high-throughput) data analysis

pipelines.

The miniML model output predicts the label—no event or event—for each time step (i.e., stride) with a numerical value ranging from zero to one (Figure 1B). These output values can be interpreted as the confidence that the corresponding data segment contains a synaptic event. Model inference thus outputs a prediction trace that has a slightly lower temporal resolution compared to the original input. Using peak finding on the prediction trace allows extracting data segments with individual events from a recording. While 0.5 represents a reasonable minimum peak value, the exact choice is not critical to detection performance (see below). To quantify events, the extracted sections are checked for overlapping events, and detected events are aligned by the steepest slope, allowing calculation of individual event statistics. Figure 2B illustrates the sliding window approach to detect spontaneous events in time-series data, such as a continuous voltage-clamp recording of spontaneous mEPSCs in a cerebellar GC. miniML provided a clear peak for all synaptic events present, without false positives (Figure 2C, Figure S4), allowing fast and reliable event quantification (Figure 2D). These data demonstrate that a deep learning model can be applied to detect synaptic events in electrophysiological time-series data.

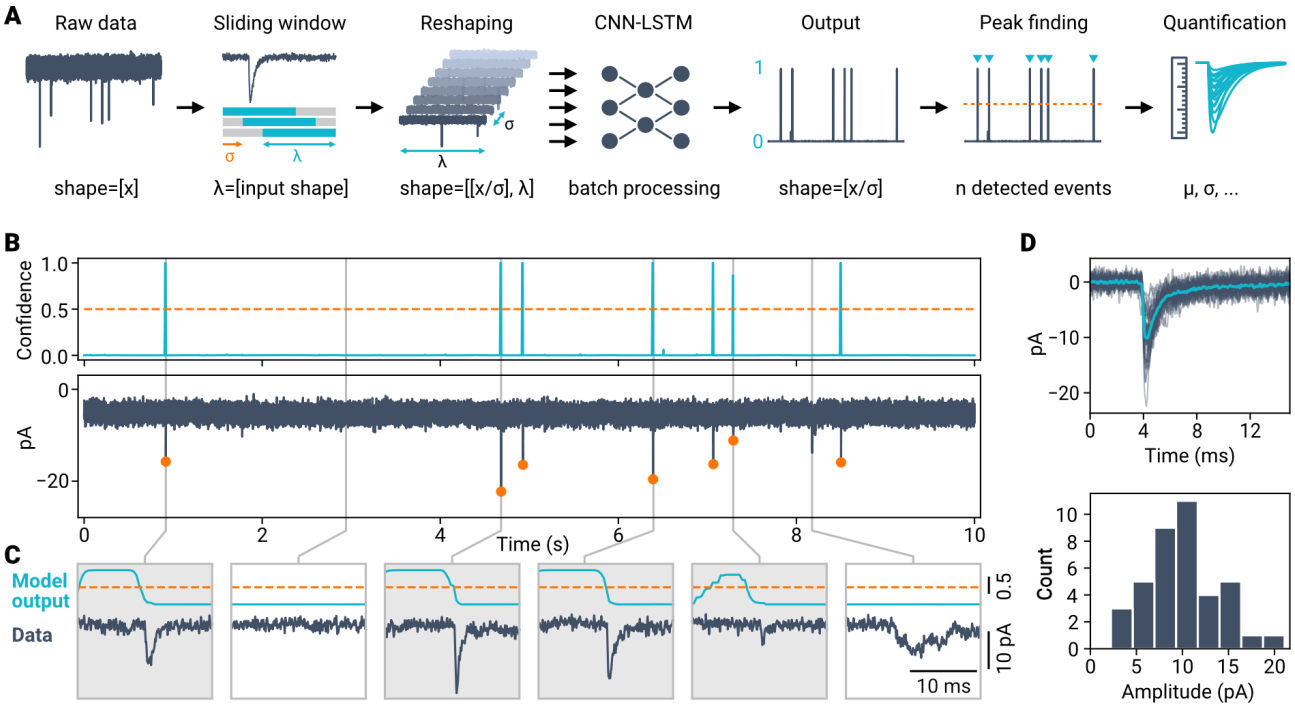

Figure 2. Applying AI-based classification to robustly detect synaptic events in electrophysiological time-series data. (**A**) Event detection workflow using a deep learning classifier. Time-series data are reshaped with a window size corresponding to length of the training data and a given stride. Peak detection of the model output allows event localization (orange dashed line indicates peak threshold) and subsequent quantification. (**B**) Example of event detection in a voltage-clamp recording from a cerebellar granule cell. Top: Prediction trace (top) and corresponding raw data with synaptic events (bottom). Dashed line indicates minimum peak height of 0.5, orange circles indicate event localizations. (**C**) Zoom in for different data segments. Detected events are highlighted by gray boxes. (**D**) Event analysis for the cell shown in (**B**) (total recording time, 120 s). Top: Detected events with average (light blue). Bottom: Event amplitude histogram.

AI-based event detection is superior to previous methods

To benchmark the AI model's event detection performance, we compared it with commonly used template-based approaches (template-matching (Clements and Bekkers, 1997) and deconvolution (Pernía-Andrade et al., 2012)), and a finite-threshold-based method (Kudoh and Taguchi, 2002). Some of these—or similar—algorithms are implemented in proprietary software solutions for recording and analyzing of electrophysiological data. We also included a recently developed Bayesian event detection approach (Merel et al., 2016) and the automated event detection routine of MiniAnalysis software (Synaptosoft Inc.). Although no longer commercially available, MiniAnalysis is still being used in a large number of recent publications. To compare synaptic event detection methods, we developed a standardized benchmarking pipeline (Figure 3A). We first performed event-free

voltage-clamp recordings from mouse cerebellar GCs (in the presence of blockers of inhibitory and excitatory transmission; see Materials and Methods and Figure S4). Synthetic events with a two-exponential time course were then generated and superimposed on the raw recording to produce ground-truth data (Figure 3B–C). Event amplitudes were drawn from a log-normal distribution (Figure S5) with varying means to cover the range of SNRs typically observed in recordings of miniature events (2–15 dB, data from n = 170 GC recordings, Figure 3D). We generated events with kinetics that closely matched the template used for the matched-filtering approaches, thus ensuring a conservative comparison of miniML with other methods applied under optimal conditions (i.e., using the exact average event shape as template). To measure detection performance, we calculated recall (i.e., sensitivity), precision (fraction of correct identifications) and the F1 score. Recall depended on SNR for all methods, with miniML and deconvolution showing the highest values (Figure 3E). The precision was highest for miniML, which detected no false positives at any SNR, in contrast to all other methods (Figure 3F). When assessing overall performance, miniML provided the highest F1 scores across SNRs (Figure 3G). In addition, miniML showed superior results when changing event kinetics, indicating higher robustness to variable event shapes (Figure 3H and Figure S5), which may be particularly important in neurons with diverse synaptic inputs due to mixed receptor populations (Lesperance et al., 2020), or during pharmacological experiments (Ishii et al., 2020).

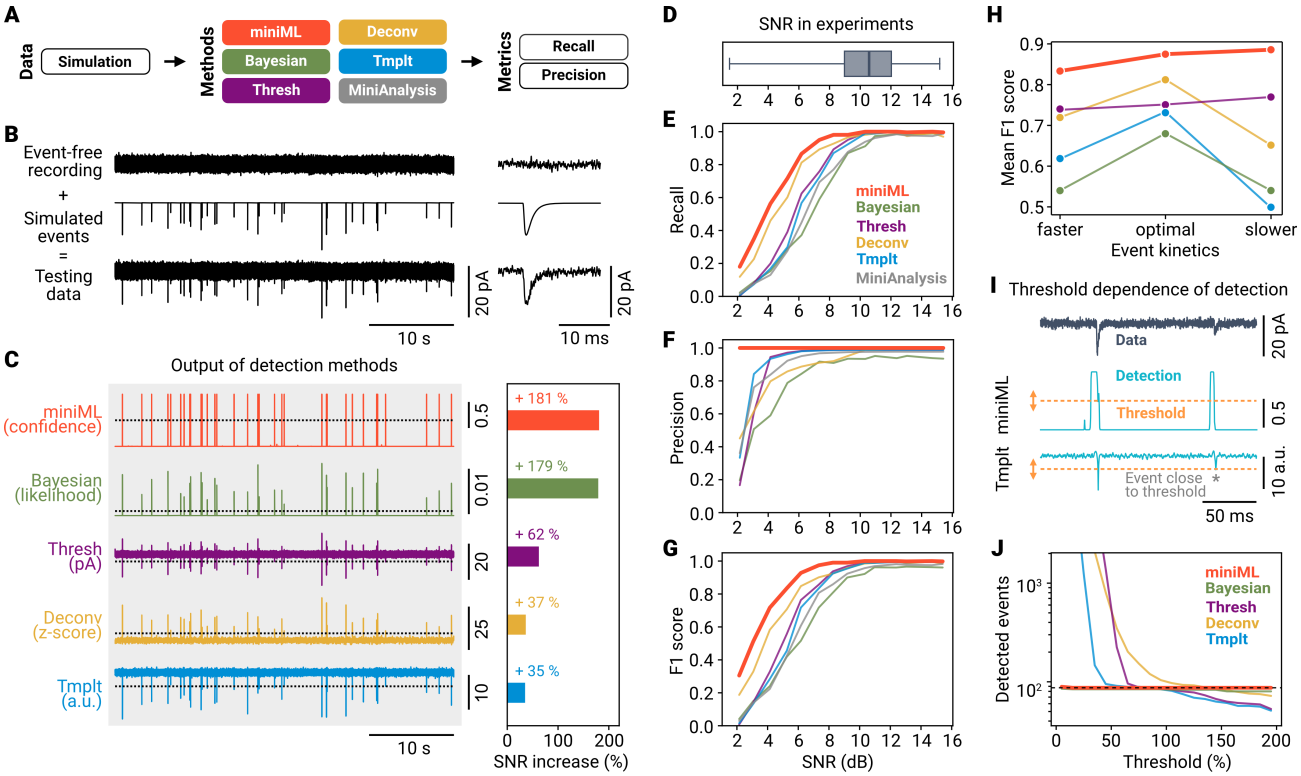

Figure 3. Systematic benchmarking demonstrates that AI-based event detection is superior to previous methods.
(**A**) Scheme of event detection benchmarking. Six methods are compared using precision and recall metrics. (**B**) Event-free recordings were superimposed with generated events to create ground-truth data. Depicted example data have a signal-to-noise ratio (SNR) of 9 dB. (**C**) Left: Output of the detection methods for data in (**B**). Right: Improvement in SNR relative to the data. Note that MiniAnalysis is omitted from this analysis because the software does not provide output trace data. (**D**) SNR from mEPSC recordings at MF–GC synapses (n = 170, whiskers cover full range of data). (**E–G**) Recall, precision, and F1 score versus SNR for the six different methods. Data are averages of three independent runs for each SNR. (**H**) Average F1 score versus event kinetics. Detection methods relying on an event template (template-matching, deconvolution and Bayesian) are not very robust to changes in event kinetics. (**I**) Evaluating the threshold dependence of detection methods. Asterisk marks event close to detection threshold. (**J**) Number of detected events vs. threshold (in % of default threshold value, range 5–195) for different methods. Dashed line indicates true event number.

Conventional event detection methods typically produce a detection trace with a shape identical to the input data and values in an arbitrary range (Figure 3C). In contrast, miniML generates output in the interval [0, 1], which can be interpreted as the confidence of event occurrence. The output of event detection methods (i.e., the

detection trace) significantly increases the SNR with respect to the original data. In our benchmark scenario, miniML and the Bayesian method provided the greatest discrimination from background noise (Figure 3C). To locate the actual event positions, a threshold must be applied to the detection trace, which can drastically affect the results, as peaks in the detection trace may depend on the event amplitude. Intriguingly, miniML's detection trace peaks did not depend on event amplitudes, setting it apart from other methods (Figure S5). For template-based methods, recommendations are provided on threshold selection (Clements and Bekkers, 1997; Pernía-Andrade et al., 2012), but users typically need to adjust this parameter according to their specific data. The choice of the threshold strongly influences detection performance, as even small changes can lead to marked increases in false positives or false negatives. To investigate the threshold dependence of different methods, we systematically varied the threshold and analyzed the number of detected events and the F1 score. Notably, miniML's detection performance remained consistent over a wide range (5–195%, Figure 3I–J and Figure S5), with false positives only occurring at the lower threshold limit (5%, corresponding to a cutoff value of 0.025 in the detection trace). Conversely, the other detection methods were very sensitive to threshold changes (Figure 3J and Figure S5) due to the comparatively low SNR ratio of the output traces (but note that the Bayesian method is only slightly threshold dependent). In addition, the miniML threshold (i.e., minimum peak) value is bounded with an intuitive meaning. These comparisons underscore that miniML requires no prior knowledge of the exact event shape and is virtually threshold independent, thus enabling reliable event detection.

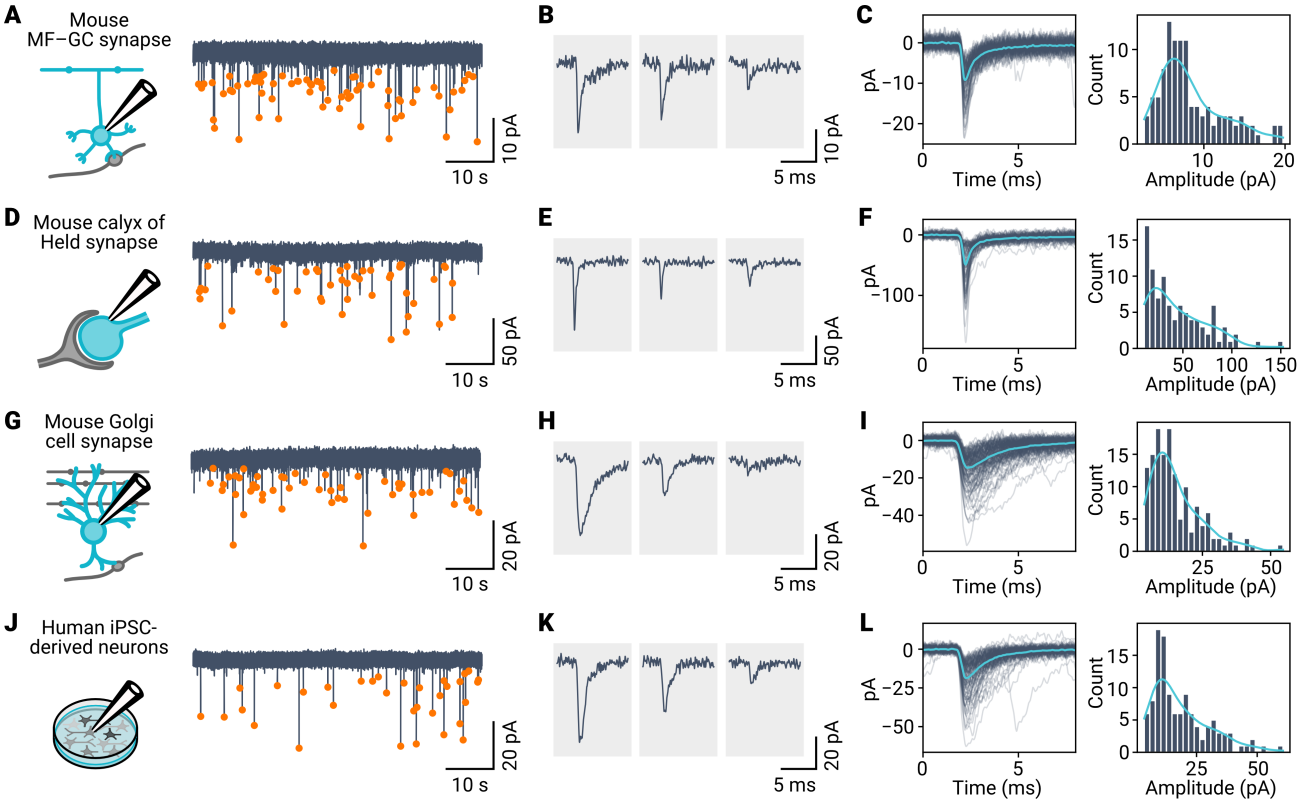

Figure 4. Application of miniML to electrophysiological recordings from diverse synaptic preparations.
(**A**) Schematic and example voltage-clamp recordings from mouse cerebellar GCs. Orange circles mark detected events. (**B**) Representative individual mEPSC events detected by miniML. (**C**) Left: All detected events from (**A**), aligned and overlaid with average (light blue). Right: Event amplitude histogram. (**D–F**) Same as in (**A–C**), but for recordings from mouse calyx of Held synapses. (**G–I**) Same as in (**A–C**), but for recordings from mouse cerebellar Golgi cells. (**J–L**) Same as in (**A–C**), but for recordings from cultured human induced pluripotent stem cell (iPSC)-derived neurons. miniML consistently retrieves spontaneous events in all four preparations.

miniML reliably detects spontaneous synaptic events in out-of-sample data

We trained the miniML model using data from cerebellar MF–GC synapses, which enabled reliable analysis of mEPSC recordings from this preparation (Figure 2, Figure 4A-C). However, synaptic event properties typically differ between preparations, for example in terms of event frequency, waveform and SNR. To test miniML's

generalizability, we evaluated the detection performance on out-of-sample data. We recorded and analyzed data from the mouse calyx of Held synapse, a large axosomatic synapse in the auditory brainstem that relays rate-coded information over a large bandwidth and with high temporal precision. Due to the large number of active zones, synaptic events are quite frequent. Despite being trained on cerebellar MF–GC data, miniML accurately detected mEPSCs in these recordings (Figure 4D–F). This result may be facilitated by the fast event kinetics at the calyx of Held, which approach those of MF–GC mEPSCs. To test whether miniML could also detect events with slower kinetics, we applied it to recordings from cerebellar Golgi cells (Figure 4G-H) (Kita et al., 2021). These interneurons in the cerebellar cortex provide surround inhibition to GCs and receive input from parallel fiber and MF synapses. miniML reliably detected synaptic events in these recordings (Figure 4G–H), although event decay kinetics were slower compared to the training data (Golgi cell, 1.83 ms (SD 0.44 ms, n = 10 neurons), GC training data, 0.9 ms). However, simulations indicated that larger differences in event kinetics may ultimately hinder detection when using the MF–GC mEPSC model (Figure S6).

Synaptic events are also commonly recorded from neuronal culture preparations. We determined whether miniML could be applied to this scenario using recordings from cultured human induced pluripotent stem cell (iPSC)-derived neurons. We patch-clamped eight-week-old neurons and recorded spontaneous synaptic events in voltage-clamp (Figure 4J-L). Prediction on these human iPSC-derived neuron data showed that miniML robustly detected synaptic events (Figure 4J–L), which had an average frequency of 0.15 Hz (SD 0.24 Hz, n = 56 neurons). Taken together, the consistent performance across different synaptic preparations indicates that the miniML model can be directly applied to out-of-sample data with similar kinetics.

Generalization of miniML to diverse event and data types via transfer learning

Synaptic events can vary considerably between preparations, for example due to differences in the number and type of postsynaptic neurotransmitter receptors, the number of synaptic inputs, postsynaptic membrane properties, transmitter content of synaptic vesicles or presynaptic release properties. In addition, different hardware and recording conditions may affect the characteristics of the recorded data. Because classical detection methods are comparatively sensitive to event shape and threshold settings, their application to novel preparations and/or different recording modes is challenging. We employed a transfer learning (TL) strategy to facilitate the application of miniML to different preparations, recording conditions, and data types. TL is a powerful technique in machine learning that allows for the transfer of knowledge learned from one task or domain to another (Caruana, 1994; Yosinski et al., 2014). TL is widely used with CNNs to take advantage of large pretrained models and repurpose them to solve new, unseen tasks (Theodoris et al., 2023). Importantly, only a part of the network needs to be trained for the novel task, which significantly reduces the number of training samples needed and speeds up training while avoiding overfitting (Yosinski et al., 2014). We therefore reasoned that TL based on freezing the convolutional layers during training of our pretrained network could be used to train a new model to detect events with different shapes and/or kinetics, using a lower number of training samples (Figure 5A).

We tested the use of TL for miniML with recordings of miniature excitatory postsynaptic potentials (mEPSPs) in mouse cerebellar GCs. These events have opposite polarity and slower kinetics compared to the original mEPSC training data. We compared TL-based model training to full training (with all layers trainable and reinitialized weights), varying the training sample size. Whereas accuracy increased and loss decreased with the number of samples (Figure 5B), TL models performed well with as few as 400 samples. Under these conditions, accuracy was only slightly lower than for full training with almost ten times the sample size (median accuracy, 95.4 versus 96.1%; Figure 5C). This suggests that TL can significantly reduce the sample size needed for training, minimizing the time-consuming process of event extraction and labeling. Across different datasets, TL-trained models performed comparably to those trained from scratch (Figure S7). Taken together, these data demonstrate that TL allows model training with an order of magnitude smaller amount of training data, making miniML easily transferable to new datasets.

To investigate how a TL-trained model performs in event detection, we analyzed data in different recording modes (voltage-clamp vs. current-clamp). The different noise conditions typically encountered in current-clamp and the difference between EPSP and EPSC waveforms necessitate distinct detection schemes for common detection methods. Recording miniature EPSPs and EPSCs in the same cerebellar GCs (Figure 5D–E) enabled us to compare detection performance via event frequency. mEPSPs could be approximated by a two-exponential time course, similar to postsynaptic currents. However, their kinetics were considerably slower due to the

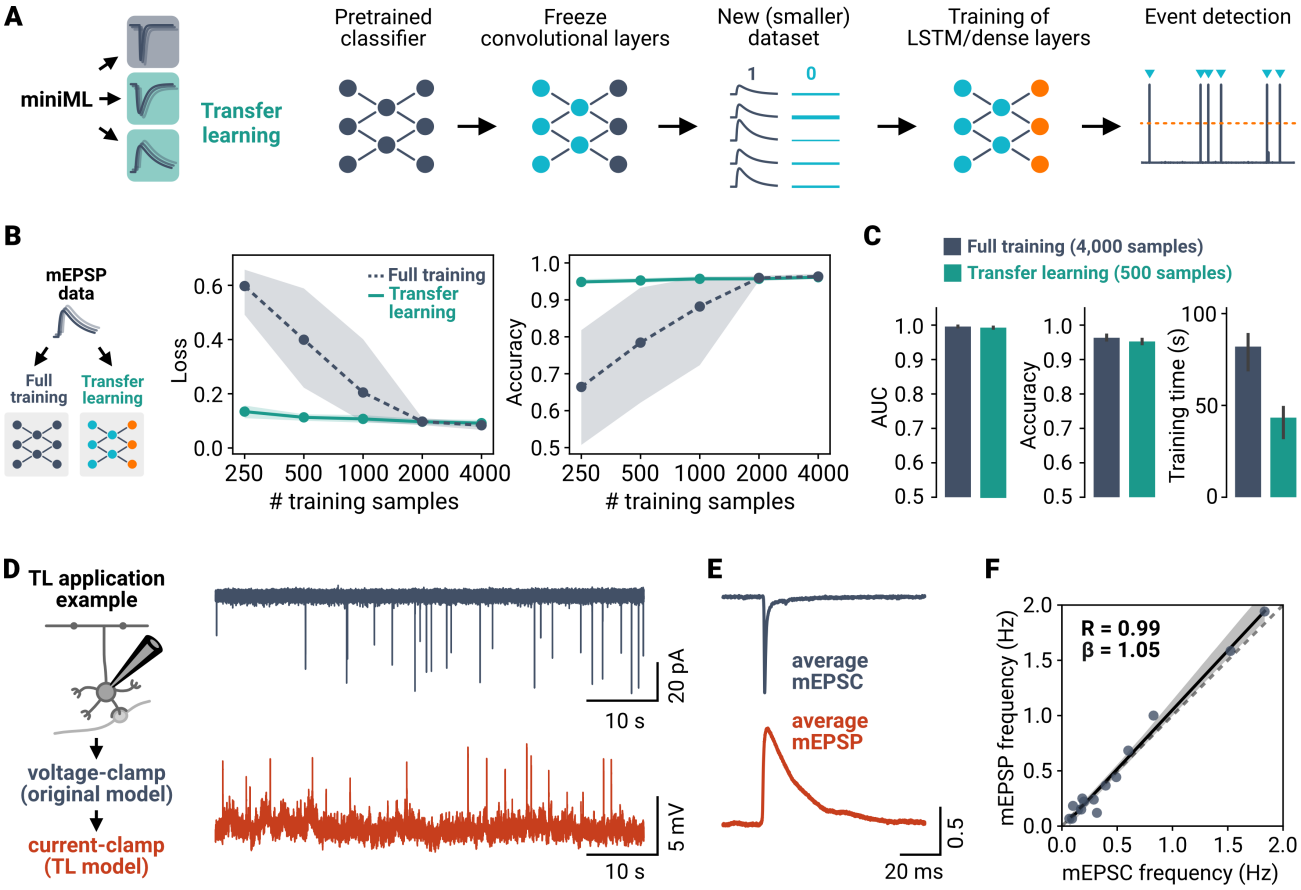

Figure 5. Transfer learning allows analyzing different types of events with small amounts of training data.
(**A**) Illustration of the transfer learning (TL) approach. The convolutional layers of a trained miniML model are frozen before retraining with new, smaller datasets. (**B**) Comparison of TL and full training for MF–GC mEPSP data. Shown are loss and accuracy versus sample size for full training and TL. Indicated subsets of the dataset were split into training and validation sets; lines are averages of fivefold crossvalidation, shaded areas represent 95% CI. Note the log-scale of the abscissa. (**C**) Average AUC, accuracy, and training time of models trained with the full dataset or with TL on a reduced dataset. TL models yield comparable classification performance with reduced training time using only 12.5% of the samples. (**D**) Example recordings of synaptic events in voltage-clamp and current-clamp mode consecutively from the same neuron. (**E**) Average peak normalized mEPSC and mEPSP waveform from the example in (**D**). Note the different event kinetics depending on recording mode. (**F**) Event frequency for mEPSPs plotted against mEPSCs. Dashed line represents identity, solid line with shaded area represents linear regression.

charging of the plasma membrane (Figure 5E). We used a TL-trained model for mEPSP detection and the standard miniML model for mEPSCs. Remarkably, the average event frequencies were very similar in the two different recording modes (voltage-clamp: 0.49 Hz, SD 0.53 Hz, current-clamp: 0.54 Hz, SD 0.6 Hz, n = 15 for both) and highly correlated across neurons (Figure 5F). This highlights the reliability of TL with small training sets for the consistent detection of synaptic events in datasets with varying characteristics.

miniML reveals diversity of synaptic event kinetics in an *ex vivo* whole brain preparation

The presence of overlapping and highly variable event shapes pose additional challenges for event detection methods. To evaluate miniML's performance in such complex scenarios, we analyzed a dataset recorded from principal neurons of the adult zebrafish telencephalon (Rupprecht and Friedrich, 2018). We focused on spontaneous excitatory inputs to these neurons, characterized by diverse event shapes and frequencies (Rupprecht and Friedrich, 2018). Training via TL (see Materials and Methods) yielded a model that enabled the reliable detection of spontaneous excitatory currents (Figure 6A). Analysis of event properties across cells revealed broad distributions of event statistics (Figure 6B), including a large diversity of event rise and decay kinetics (Figure 6B–C). Notably, miniML consistently identified synaptic events with diverse kinetics and shapes.

We next used the extracted event kinetics features of individual neurons (Figure 6D–H) to demonstrate

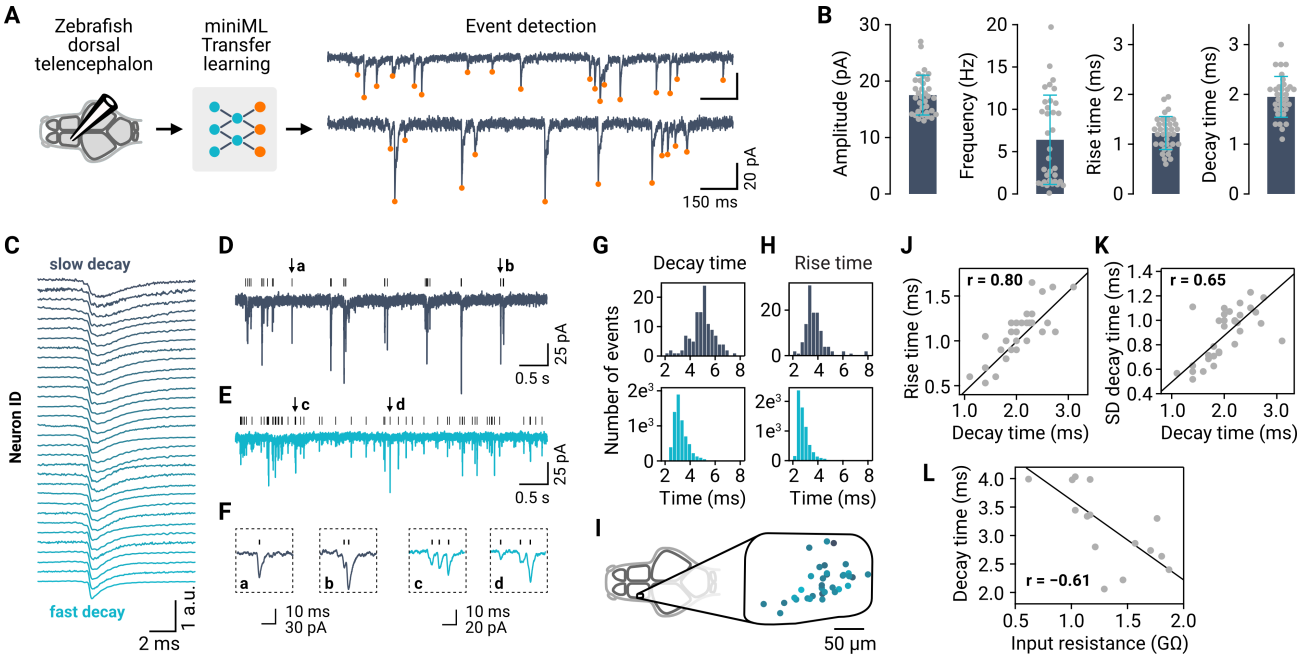

Figure 6. Synaptic event detection for neurons in a full-brain explant preparation of adult zebrafish.
(**A**) Application of TL to facilitate event detection for EPSC recordings. (**B**) Extraction of amplitudes, event frequencies, decay times, and rise time for all neurons in the dataset (n = 34). (**C**) Typical (mean) event kinetics for the analyzed neurons, ordered by decay times. (**D–E**) Example recordings with slow (dark blue) and fast (light blue) event kinetics. (**F**) Examples of events taken from (**D–E**), illustrating the diversity of kinetics within and across neurons. (**G**) Distribution of event decay kinetics across single events for two example neurons (traces shown in (**D–E**)). (**H**) Distribution of event rise kinetics across single events for same two example neurons. (**I**) Mapping of decay times (color-coded as in (**C**)) onto the anatomical map of the recording subregion of the telencephalon. (**J**) Mean decay and rise times are correlated across neurons. (**K**) Decay time distributions are broader (SD of decay time distributions) when mean decay times are larger. (**L**) Input resistance as a proxy for cell size is negatively correlated with the decay time.

miniML's utility in better understanding the diversity of an existing dataset. First, we explored whether the anatomical location of each neuron could predict event decay times. We mapped the recorded neurons to an anatomical reference and plotted decay times as a function of their position but did not find a strong relationship (Figure 6I; correlation with position $p>0.05$ in all 3 dimensions). Second, we tested the hypothesis that slower event kinetics are associated with larger cells. In large cells, EPSCs may undergo stronger filtering as they propagate from synaptic sites to the soma. Consistent with this idea, we observed correlations between decay and rise times across neurons (Figure 6J). Furthermore, the distribution of decay times (examples shown in Figure 6H) was broader for neurons with longer decay times (Figure 6K), suggesting a broader distribution of distances from synapses to the cell body. Finally, for a subset of neurons, we recorded input resistance, which approximates the cell membrane resistance and is therefore a proxy for cell size. Input resistance was negatively correlated with decay times across neurons (Figure 6L), consistent with the hypothesis that diverse event kinetics across neurons are determined by the conditions of synaptic event propagation to the soma and, more specifically, cell size. Taken together, this analysis underscores the versatility of miniML, as it can be successfully applied to new datasets with varying recording conditions. miniML consistently extracted synaptic events across a spectrum of event kinetics, enabling the identification and investigation of key factors determining event kinetics and other event-related properties across neurons.

miniML robustly detects mEPSC differences upon genetic receptor perturbation

Synaptic events can also be recorded at synapses of the peripheral nervous system. We next applied miniML to analyze data obtained from the *Drosophila melanogaster* larval neuromuscular junction (NMJ) (Baccino-Calace et al., 2022). This synapse is characterized by a higher frequency of spontaneous synaptic events with a slower time course compared with typical brain slice recordings. In addition, mEPSC recordings are performed using two-electrode voltage-clamp, which can be associated with large leak currents and rather low SNR. Because of the different event shapes in these data, we used a small dataset of manually extracted NMJ events to train a TL

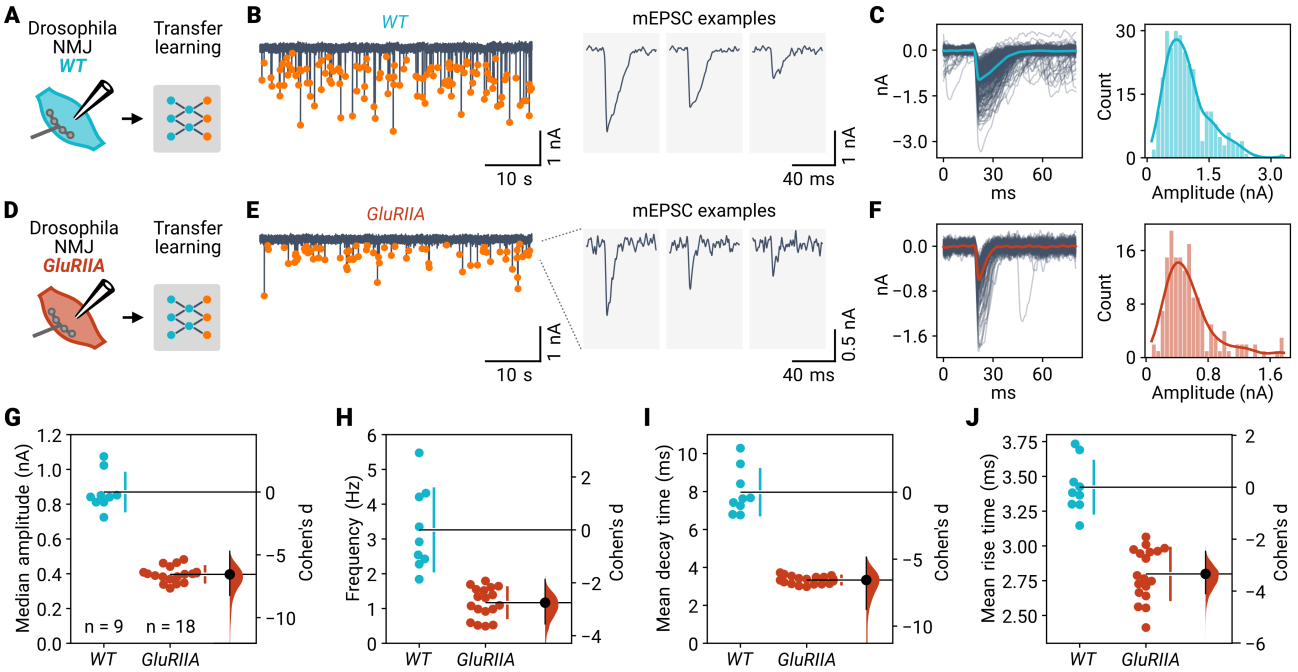

Figure 7. Event detection at *Drosophila* neuromuscular synapses upon altered glutamate receptor composition.
(**A**) Two-electrode voltage-clamp recordings from wild-type (WT) *Drosophila* NMJs were analyzed using miniML with transfer learning. (**B**) Left: Example voltage-clamp recording with detected events highlighted. Right: Three individual mEPSCs on expanded time scale. (**C**) Left: All detected events from the example in (**B**) overlaid with average (blue line). Right: Event amplitude histogram. (**D–F**) Same as in (**A–C**), but for *GluRIIA* mutant flies. (**G**) Comparison of median event amplitude for WT and GluRIIA NMJs. Both groups are plotted on the left axes; Cohen's d is plotted on a floating axis on the right as a bootstrap sampling distribution (red). The mean difference is depicted as a dot (black); the 95% confidence interval is indicated by the ends of the vertical error bar. (**H**) Event frequency is lower in *GluRIIA* mutant NMJs than in WT. (**I**) Knockout of *GluRIIA* speeds event decay time. (**J**) Faster event rise times in *GluRIIA* mutant NMJs.

model. The TL model was able to reliably detect synaptic events in wild-type (WT) NMJ recordings, with excellent recall and precision (Figure 7A–C). We next assessed event detection upon deletion of the non-essential glutamate receptor subunit *GluRIIA*, which causes a strong reduction in mEPSC amplitude and faster kinetics (DiAntonio et al., 1999). A separate TL model allowed reliable synaptic event detection in recordings from *GluRIIA* mutant larvae (Figure 7D–F). We observed a 54% reduction in mEPSC amplitude compared to WT (Figure 7G), consistent with previous reports (DiAntonio et al., 1999; Petersen et al., 1997). In addition, the event frequency was reduced by 64% (Figure 7H), probably due to some events being below the detection limit. Half decay and rise times were also shorter at *GluRIIA* than at WT NMJs (–58% and –18%, respectively) (Figure 7I–J), which can be explained by the faster desensitization of the remaining *GluRIIB* receptors (DiAntonio et al., 1999). Thus, miniML can be applied to two-electrode voltage clamp recordings at the *Drosophila* NMJ and robustly resolves group differences upon genetic receptor perturbation.

miniML enables reliable analysis of optical miniature events from glutamate imaging

The development of highly sensitive fluorescent probes enables the recording of synaptic events using fluorescence imaging techniques (Abdelfattah et al., 2023; A. Aggarwal et al., 2023; Ralowicz et al., 2024). To further assess the generalizability of miniML, we tested its ability to analyze events in time-series data derived from live fluorescence imaging experiments. These imaging datasets often feature a lower sampling rate and distinct noise profile compared with electrophysiological recordings. Nevertheless, the waveforms of synaptic release events or Ca^{2+} transients closely resemble those used to train miniML. Thus, we hypothesized that miniML could also be employed for event detection in imaging data. We analyzed a previous dataset (A. Aggarwal et al., 2023) from rat neuronal cultures expressing the glutamate sensor iGluSnFR3 (Figure 8A–B). These data, recorded in the presence of TTX, contain spontaneous transient increases in fluorescence intensity representing individual release events ('optical minis'). Initially, we selected a small subset of events from the data to train a TL model. Given the low sampling rate of the imaging (100 Hz), we upsampled the data by a factor of 10 to match the

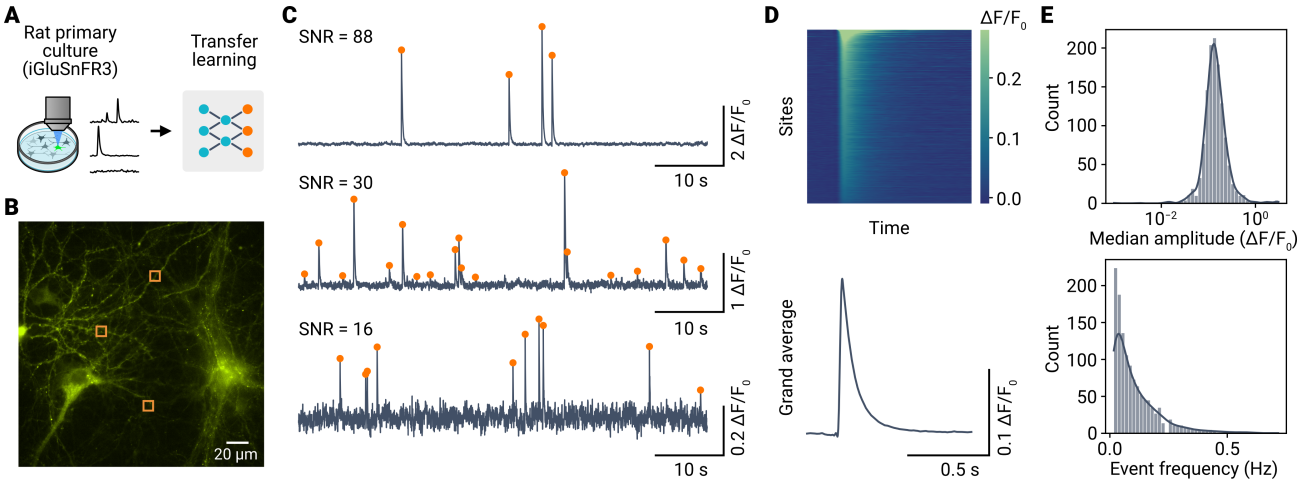

Figure 8. miniML reliably detects optical minis across a wide range of signal-to-noise ratios.
(**A**) The miniML model was applied to recordings from rat primary culture neurons expressing iGluSnFR3. Data from (A. Aggarwal et al., 2023). (**B**) Example epifluorescence image of iGluSnFR3-expressing cultured neurons. Image shows a single frame with three example regions of interest (ROIs) indicated. (**C**) $\Delta F/F_0$ traces for the regions shown in (**B**). Orange circles indicate detected optical minis. Note the different signal-to-noise ratios of the examples. (**D**) Top: Heatmap showing average optical minis for all sites with detected events, sorted by amplitude. Bottom: Grand average optical mini. (**E**) Top: Histogram with kernel density estimate (solid line) of event amplitudes for n = 1524 ROIs of the example in (**B**). Note the log-scale of the abscissa. Bottom: Inferred event frequency.

model's input shape. The TL model was subsequently applied to all detected sites within the widefield recording (A. Aggarwal et al., 2023). A qualitative assessment of the imaging traces showed excellent event detection, with miniML consistently localizing the iGluSnFR3 fluorescence transients at varying SNRs (Figure 8C). The detected optical minis had similar kinetics to those reported in (A. Aggarwal et al., 2023) (10–90% rise time, median 21.8 ms; half decay time, median 48.7 ms, Figure 8D). In addition, analysis of event frequencies across sites revealed a power-law distribution (Figure 8E), consistent with fractal behavior of glutamate release (Lowen et al., 1997). These results demonstrate that miniML generalizes beyond electrophysiological data and can support the evaluation of fluorescence imaging experiments, thus paving the way for high-throughput analyses of spontaneous events in diverse imaging datasets.

DISCUSSION

Here, we developed and evaluated miniML, a novel supervised deep learning approach for detecting spontaneous synaptic events. miniML provides a comprehensive and versatile framework for analyzing synaptic events in diverse time-series data with several important advantages over current methods.

miniML outperforms existing methods for synaptic event detection, particularly with respect to false positives, which is important for the accurate quantification of synaptic transmission. In addition to its high precision, miniML's detection performance is largely threshold independent. This effectively eliminates the trade-off between false positive and false negative rates typically present in event detection methods (Clements and Bekkers, 1997; Merel et al., 2016; Pernía-Andrade et al., 2012). Thus, miniML is highly reproducible and overcomes the need for laborious manual event inspection, enabling automated synaptic event analysis. Automated, reproducible data analysis is key to open science and the use of publicly available datasets (Ascoli et al., 2017; Ferguson et al., 2014). miniML may also be particularly well suited for analyzing large-scale neurophysiological datasets.

Our results from a wide variety of preparations, covering different species and neuronal preparations, demonstrate the robustness of miniML and its general applicability to the study of synaptic physiology. miniML generalizes well to different experimental preparations, conditions, and data types. Data with event waveforms that are similar to the original miniML training data (mouse cerebellar GC recordings) can be analyzed immediately. For more distinct event waveforms or data properties, data resampling and/or retraining of a classifier via TL enables event detection optimized for the respective recording conditions. Importantly, TL requires only a few hundred labeled events for training, facilitating its application to novel datasets. For example, an existing miniML model can be easily retrained to eliminate potential false positives, e.g., due to sudden changes in noise

characteristics.

miniML's approach to detect events is not restricted to electrophysiological data, but can also be applied to time-series data derived from live imaging experiments using, e.g., reporters of glutamate (A. Aggarwal et al., 2023), Ca^{2+} (Tran and Stricker, 2021), or membrane voltage (Li et al., 2020). The development of novel optical sensors with improved sensitivity and temporal resolution (Evans et al., 2023; Y. Zhang et al., 2023) necessitates efficient and robust event detection methods. Similarly, miniML can be used for the analysis of evoked postsynaptic responses, such as failure analysis or quantification of unitary events, and for functional connectivity studies (Campagnola et al., 2022). In addition, the method could be extended to other areas of biology and medicine where signal detection is critical, such as clinical neurophysiology or imaging.

Comprehensive performance comparison is essential for method evaluation and selection. Benchmarking requires standardized ground truth data, but unlike, e.g., spike inference from Ca^{2+} imaging (Theis et al., 2016), these are usually not available for spontaneous synaptic event recordings. Furthermore, the results can vary between simulated and real data (Theis et al., 2016). We therefore established a benchmarking pipeline using event-free recordings with simulated events, circumventing differences in the noise power spectrum of simulations vs. recordings (Merel et al., 2016; Pernía-Andrade et al., 2012). This approach may provide a general toolbox for evaluating the effectiveness of different synaptic event detection methods.

The supervised learning method of miniML requires labeled training data. Although TL requires only a few hundred training samples, these data need to be collected and annotated by the user. Future directions may explore unsupervised learning techniques such as autoencoders to reduce the dependence on annotated training data. In addition, at very high event frequencies, individual synaptic events may overlap, making their separation difficult. Although miniML can generally detect overlapping events, very close events may not always be detected and there may be a lower bound. Implementing additional techniques such as wavelet transformation, or spectral domain or shapelet analysis (Batal and Hauskrecht, 2009; Ye and Keogh, 2009) may improve the accuracy of event detection, especially for overlapping events.

miniML presents an innovative data analysis method that will advance the field of synaptic physiology. Its open-source Python software design ensures seamless integration into existing data analysis pipelines and enables widespread use of the method, fostering the development of new applications and further innovation. Remarkably, despite its deep learning approach, miniML runs at relatively rapid speed on commodity hardware. Thanks to its robust, generalizable, and unbiased detection performance, miniML allows researchers to perform more accurate and efficient synaptic event analysis. A standardized, more efficient, and reproducible analysis of synaptic events will promote important new insights into synaptic physiology and dysfunction (Lepeta et al., 2016; Zoghbi and Bear, 2012) and help improve our understanding of neural function.

MATERIALS AND METHODS

Electrophysiological recordings

Animals were treated according to national and institutional guidelines. All experiments were approved by the Cantonal Veterinary Office of Zurich (approval number no. ZH206/2016 and ZH009/2020). Experiments were performed in male and female C57BL/6J mice (Janvier Labs, France). Mice were 1–5-months-old, but for recordings from the Calyx of Held, which were performed in P9 animals. Animals were housed in groups of 3–5 in standard cages under a 12h-light/12h-dark cycle with food and water ad libitum. Mice were sacrificed by rapid decapitation after isoflurane anesthesia. The cerebellar vermis was quickly removed and mounted in a chamber filled with chilled extracellular solution. Parasagittal 300-µm thick slices were cut with a Leica VT1200S vibratome (Leica Microsystems, Germany), transferred to an incubation chamber at 35 °C for 30 minutes, and then stored at room temperature until experiments. The extracellular solution (artificial cerebrospinal fluid, ACSF) for slicing and storage contained (in mM): 125 NaCl, 25 $NaHCO_3$, 20 D-glucose, 2.5 KCl, 2 $CaCl_2$, 1.25 NaH_2PO_4, 1 $MgCl_2$, aerated with 95% O_2 and 5% CO_2.

Slices were visualized using an upright microscope with a 60×, 1 NA water immersion objective, infrared optics, and differential interference contrast (Scientifica, UK). The recording chamber was continuously perfused with ACSF. For event-free recordings, we blocked excitatory and inhibitory transmission using ACSF supplemented with 50 µM D-APV, 10 µM NBQX, 10 µM bicuculline, and 1 µM strychnine. Patch pipettes (open-tip resistances of 3–8 MΩ) were filled with solution containing (in mM): 150 K-D-gluconate, 10 NaCl, 10 HEPES, 3 MgATP, 0.3 NaGTP, 0.05 ethyleneglycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid (EGTA), pH adjusted to 7.3 with

KOH. Voltage-clamp and current-clamp recordings were made using a HEKA EPC10 amplifier controlled by Patchmaster software (HEKA Elektronik GmbH, Germany). Voltages were corrected for a liquid junction potential of +13 mV. Experiments were performed at room temperature (21–25 °C). Miniature EPSCs (mEPSCs) were recorded at a holding potential of −100 mV or −80 mV, and miniature EPSPs (mEPSPs) at the resting membrane potential. Data were filtered at 2.9 kHz and digitized at 50 kHz. Synaptic event recording periods typically lasted 120 s.

Spontaneous EPSCs were recorded from human iPSC-derived neurons in a similar fashion, but ACSF consisted of (in mM): 135 NaCl, 10 HEPES, 10 D-glucose, 5 KCl, 2 $CaCl_2$, 1 $MgCl_2$. We recorded from 8-week-old neuronal cultures as described in (Asadollahi et al., 2023). Synaptic events were observed in $\sim$ 51% of neurons.

Training data and annotation

We used synaptic event recordings from a previous publication (Delvendahl et al., 2019) to generate the training dataset. mEPSCs were extracted based on low-threshold template-matching. Corresponding sections of data without events were randomly selected from the recordings. The extracted windows had a length of 600 data points. Given our sampling rate of 50 kHz, this corresponds to 11.98 ms. We subsequently manually scored data sections as event-containing (label=1) or not event-containing (label=0). The ratio of events to non-events was kept close to one to ensure efficient training. Based on empirical observations of model performance, we included relatively small amplitude events that are often missed by other methods. Similarly, including negative examples that are commonly picked up as false positives, resulted in more accurate prediction traces. We used data augmentation techniques to further improve model discrimination. We simulated waveforms of non-synaptic origin, which are occasionally encountered during recordings, and superimposed them onto noise recordings. Examples included rapid current transients that can be caused by peristaltic perfusion pump systems often used in brain slices recordings, or slow currents with a symmetric rise and decay time course. A total of 4500 segments were created and labeled as 0. To maintain the ratio of negative to positive examples, we added an equivalent number of simulated synaptic events. The biexponential waveform described in the section *Benchmarking detection methods* was used for event simulation. The final training dataset contained 30,140 samples (21,140 from recorded traces and 9,000 simulated samples).

Deep learning model architecture

We built a discriminative end-to-end deep learning model for one-dimensional time-series classification (Fawaz et al., 2019). The neural network architecture comprised multiple convolutional layers, an LSTM layer, and a fully connected layer. The approach is related to networks designed for event detection in audio (Passricha and R. K. Aggarwal, 2019) and image data (Donahue et al., 2017; Islam et al., 2020), or classification of genomic data (Tasdelen and Sen, 2021). The classifier was built using TensorFlow, an open-source machine learning library for Python (Martín Abadi et al., 2015). The network takes batches of one-dimensional (1D) univariate time-series data as input, which are converted into a tensor of shape (batch size, data length, 1). The data is passed to three convolutional blocks. Each block consists of a 1D convolutional layer with a leaky Rectified Linear Unit (leaky ReLU) activation function followed by an average pooling layer. To avoid overfitting, each convolutional block uses batch normalization (Ioffe and Szegedy, 2015). Batch normalization reduced training time by about two times and improved the accuracy of the resulting model. We added a fourth convolutional block that includes a convolutional layer, batch normalization and a leaky ReLU activation function, but no average pooling layer. The output of the final convolutional layer passes through a bidirectional recurrent layer of LSTM units. The final layers consist of a fully connected dense unit layer and a dropout layer (Srivastava et al., 2014), followed by a single unit with sigmoid activation. The output of the neural network is a scalar between [0, 1]. The layers and parameters used, including output shape and number of trainable parameters, are summarized in Supplementary Table 1. The total number of trainable parameters was 190,913.

Training and evaluation

The network was trained with Tensorflow 2.12 and Python 3.10 with CUDA 11.4. Datasets were scaled between zero and one, and split into training and validation data (0.75/0.25). The model was compiled using the Adam optimizer (Kingma and Ba, 2014) with AMSGrad (Reddi et al., 2019). We trained the classifier using a learning rate η = 2E–5 and batch size of 128 on the training data. Training was run for maximum 100 epochs with early

stopping to avoid overfitting. Validation data was used to evaluate training performance. Early stopping was applied when the validation loss did not improve for eight consecutive epochs. Typically, training lasted for 20–40 epochs. We used binary cross-entropy loss and binary accuracy as measures of performance during training. The best performing model was selected from a training run, and a receiver-operating characteristic (ROC) was calculated. We used accuracy and area under the ROC curve to evaluate training performance. To accelerate training, we used a GPU; training time for the neural network was ~8 minutes on a workstation with NVIDIA Tesla P100 16 GB GPU (Intel Xeon 2.2 GHz CPU, 16 GB RAM).

Model and training visualization

To analyze the discriminative data segments, we calculated per-sample saliency maps with SmoothGrad using the Python package tf-keras-viz (Simonyan et al., 2013). Saliency maps were smoothed with a 10-point running average filter for display purposes. To illustrate the transformation by the model, we performed dimensionality reduction of the training data by Uniform Manifold Approximation and Projection (UMAP), using either the raw dataset or the input to the final layer of the deep learning model.

Applying the classifier for event detection

The trained miniML classifier takes sections of data with predefined length as input. To detect events in arbitrarily long recordings, a sliding window procedure is used. Time-series data from voltage-clamp or current-clamp recordings is segmented using a sliding window with stride. The resulting 2D-tensor is min-max scaled and then used as model input for inference. To overcome the potential limitation of long computation times, we used a sliding window with stride procedure. Using a stride > 1 significantly reduces the inference time of the model (Figure S3), especially for data with high sampling rates or long recording times. This approach results in a prediction trace with data being spaced at $samplinginterval * stride$. With $stride = 20$, a 120-s recording at 50 kHz sampling rate can be analyzed in ~15 s on a laptop computer (Apple M1Pro, 16 GB RAM, Figure S3).

Events in the input data trace result in distinct peaks in the prediction trace. We applied a maximum filter to enhance post-processing of the data. Thus, by applying a threshold to the prediction trace, synaptic event positions can be detected. Our analyses indicate that the absolute threshold value is not important in the [0.05, 0.95] range (Figure S5). Data following a threshold crossing in the prediction trace are cut from the raw data and aligned by steepest rise. To find the point of steepest rise, a peak-search is performed in the first derivative of the short data segment. If multiple peaks are detected, any peak that has a prominence $\geqslant 0.25$ relative to the largest detected prominence is treated as an additional event that is in close proximity or overlapping. The resulting 2D-array with aligned events can be further analyzed to obtain descriptive statistics on, e.g., amplitude, charge, or rise and decay kinetics of individual events.

Benchmarking detection methods

We compared the deep learning-based miniML with the following previously described detection methods: template-matching (Clements and Bekkers, 1997), deconvolution (Pernía-Andrade et al., 2012), a finite threshold-based approach (Kudoh and Taguchi, 2002), the commercial MiniAnalysis software (version 6.0.7, Synaptosoft), and a Bayesian inference procedure (Merel et al., 2016). Detection methods were implemented in Python 3.9, except for the Bayesian method, which was run using Matlab R2022a, and MiniAnalysis running as stand-alone software on Windows.

We used generated standardized data to benchmark the detection performance of different methods. Event-free recordings (see section *Electrophysiological Recordings* for details) were superimposed with simulated events having a biexponential waveform:

$$I(t) = (1 - e^{(\frac{-t}{\tau_{rise}})}) * e^{(\frac{-t}{\tau_{decay}})}$$

Where $I(t)$ is the current as a function of time, and τ_{rise} and τ_{decay} are the rise and decay time constants, respectively. Simulated event amplitudes were drawn from a log-normal distribution with variance 0.4. Mean amplitude was varied to generate diverse signal-to-noise ratios. Decay time constants were drawn from a normal distribution with mean 1.0 ms and variance 0.25 ms (Delvendahl et al., 2019). Generated events were randomly placed in the event-free recording with an average frequency of 0.7 Hz and a minimum spacing of 3 ms. Generated traces provided ground-truth data for the evaluation of the different methods.

To quantify detection performance of different methods over a range of signal-to-noise ratios, we calculated the number of true positives (TP), false positives (FP) and false negatives (FN). From these, the following metrics were calculated:

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

For all metrics, higher values indicate a better model performance.

We also evaluated detection performance under non-optimal conditions (i.e., events that did not precisely match the event template or the training data). To do this, we varied the kinetics of simulated events by either increasing (mean = 4.5 ms) or decreasing (mean = 0.5 ms) the average decay time constant (Figure S5).

Hyperparameter settings of detection methods

For all conditions, threshold settings were: –4 for template-matching, $5 * SD$ of the detection trace for deconvolution, and –4 pA for the finite-threshold method. Both template-based methods used a template waveform corresponding to the average of the simulated events. For the Bayesian detection approach, we used the code provided by the authors (Merel et al., 2016) with the following adjustments to the default hyperparameters: minimum amplitude = 3.5, noise $\phi = [0.90; -0.52]$, rate = 0.5. We chose a cutoff of $6 * SD$ of the event time posteriors as threshold. MiniAnalysis (Synaptosoft) was run in the automatic detection mode with default settings for EPSCs and amplitude and area thresholds of –4 pA and 4, respectively. We used a threshold of 0.5 for miniML.

Transfer learning

To make the miniML method applicable to data with different characteristics, such as different event kinetics, noise characteristics, or recording mode, we used a transfer learning (TL) approach (Pratt et al., 1991). We froze the convolutional layers of the fully trained MF–GC miniML model, resulting in a new model with only the LSTM and dense layers being trainable. Thus, the convolutional layers act as pretrained feature detectors and much fewer training samples are required. Hyperparameters and training conditions were the same as for full-training (see *Training and evaluation*) with the following exceptions: learning rate η = 2E–8, patience = 15, batch size = 32, dropout rate = 0.5. The training data were resampled to 600 data points to match the input shape of the original model.

To compare TL with full training, random subsets of the training data (mEPSPs in cerebellar GCs, spontaneous EPSCs in zebrafish, or mEPSCs at the *Drosophila* neuromuscular junction) with increasing size were generated and used to train models using a fivefold cross-validation. We always used the same size of the validation dataset. After comparing TL with full training (Figure 5, Figure S7), we trained separate models to analyze the different datasets using subsets of the available training data.

Quantifications

Computing times were quantified using the performance counter function of the built-in Python package *time* and are given as wall times. Statistical comparisons were made using permutation t-tests with 5000 reshuffles (Ho et al., 2019). Effect sizes are reported as Cohen's d with 95% confidence intervals, obtained by bootstrapping (5000 samples; the confidence interval is bias-corrected and accelerated) (Ho et al., 2019). Event amplitudes were quantified as difference between detected event peaks and a short baseline window before event onset. Decay times refer to half decay times, and rise times were quantified as 10–90% rise times.

Data and code availability

Datasets used for model training are available from Zenodo [Link will be made available upon publication]. miniML source code and pretrained models are available online [https://github.com/delvendahl/miniML], including analysis code. All code was implemented in Python with the following libraries: TensorFlow, Scipy, Numpy, Matplotlib, Pandas, h5py, scikit-learn, pyABF, dabest.

We used datasets from previous publications to generate training sets and to assess the application of miniML in zebrafish and *Drosophila* (Baccino-Calace et al., 2022; Delvendahl et al., 2019; Rupprecht and Friedrich, 2018). In addition, a published dataset (A. Aggarwal et al., 2023) was used to probe the application of miniML to imaging data [https://doi.org/10.25378/janelia.21985406].

ACKNOWLEDGMENTS

We thank Mark D. Robinson and Anu G. Nair for helpful discussions.

This work received funding by the Swiss National Science Foundation (grant PZ00P3_174018 to I.D., grant PZ00P3_209114 to P.R., grant 310030B_152833/1 to R.F.), the Novartis Research Foundation (to R.F.), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 742576 to R.F.), a fellowship from the Boehringer Ingelheim Fonds (to P.R.), and the UZH Alumni Research Talent Development fund (to I.D.). The funding bodies had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Abbott, L. F., & Regehr, W. G. (2004). Synaptic computation. *Nature*, *431*(7010), 796–803. https://doi.org/10.1038/nature03010

Abdelfattah, A. S., Zheng, J., Singh, A., Huang, Y.-C., Reep, D., Tsegaye, G., Tsang, A., Arthur, B. J., Rehorova, M., Olson, C. V., Shuai, Y., Zhang, L., Fu, T.-M., Milkie, D. E., Moya, M. V., Weber, T. D., Lemire, A. L., Baker, C. A., Falco, N., … Kolb, I. (2023). Sensitivity optimization of a rhodopsin-based fluorescent voltage indicator. *Neuron*, *111*(10), 1547–1563.e9. https://doi.org/10.1016/j.neuron.2023.03.009

Aggarwal, A., Liu, R., Chen, Y., Ralowicz, A. J., Bergerson, S. J., Tomaska, F., Mohar, B., Hanson, T. L., Hasseman, J. P., Reep, D., Tsegaye, G., Yao, P., Ji, X., Kloos, M., Walpita, D., Patel, R., Mohr, M. A., Tillberg, P. W., Looger, L. L., … Podgorski, K. (2023). Glutamate indicators with improved activation kinetics and localization for imaging synaptic transmission. *Nature Methods*, *20*(6), 925–934. https://doi.org/10.1038/s41592-023-01863-6

Alten, B., Zhou, Q., Shin, O.-H., Esquivies, L., Lin, P.-Y., White, K. I., Sun, R., Chung, W. K., Monteggia, L. M., Brunger, A. T., & Kavalali, E. T. (2021). Role of aberrant spontaneous neurotransmission in snap25-associated encephalopathies. *Neuron*, *109*(1), 59–72.e5. https://doi.org/10.1016/j.neuron.2020.10.012

Ankri, N., Legendre, P., Faber, D., & Korn, H. (1994). Automatic detection of spontaneous synaptic responses in central neurons. *Journal of Neuroscience Methods*, *52*(1), 87–100. https://doi.org/10.1016/0165-0270(94)90060-4

Ardiles, Á. O., Tapia-Rojas, C. C., Mandal, M., Alexandre, F., Kirkwood, A., Inestrosa, N. C., & Palacios, A. G. (2012). Postsynaptic dysfunction is associated with spatial and object recognition memory loss in a natural model of alzheimer's disease. *Proceedings of the National Academy of Sciences*, *109*(34), 13835–13840. https://doi.org/10.1073/pnas.1201209109

Asadollahi, R., Delvendahl, I., Muff, R., Tan, G., Rodríguez, D. G., Turan, S., Russo, M., Oneda, B., Joset, P., Boonsawat, P., Masood, R., Mocera, M., Ivanovski, I., Baumer, A., Bachmann-Gagescu, R., Schlapbach, R., Rehrauer, H., Steindl, K., Begemann, A., … Rauch, A. (2023). Pathogenic *SCN2A* variants cause early-stage dysfunction in patient-derived neurons. *Human Molecular Genetics*, *32*(13), 2192–2204. https://doi.org/10.1093/hmg/ddad048

Ascoli, G. A., Maraver, P., Nanda, S., Polavaram, S., & Armañanzas, R. (2017). Win–win data sharing in neuroscience. *Nature Methods*, *14*(2), 112–116. https://doi.org/10.1038/nmeth.4152

Baccino-Calace, M., Schmidt, K., & Müller, M. (2022). The e3 ligase thin controls homeostatic plasticity through neurotransmitter release repression. *eLife*, *11*. https://doi.org/10.7554/elife.71437

Bailly, A., Blanc, C., Francis, E., Guillotin, T., Jamal, F., Wakim, B., & Roy, P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine*, *213*, 106504. https://doi.org/10.1016/j.cmpb.2021.106504

Banerjee, S., Vernon, S., Jiao, W., Choi, B. J., Ruchti, E., Asadzadeh, J., Burri, O., Stowers, R. S., & McCabe, B. D. (2021). Miniature neurotransmission is required to maintain drosophila synaptic structures during ageing. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-24490-1

Batal, I., & Hauskrecht, M. (2009). A supervised time series feature extraction technique using DCT and DWT. *2009 International Conference on Machine Learning and Applications*. https://doi.org/10.1109/icmla.2009.13

Campagnola, L., Seeman, S. C., Chartrand, T., Kim, L., Hoggarth, A., Gamlin, C., Ito, S., Trinh, J., Davoudian, P., Radaelli, C., Kim, M.-H., Hage, T., Braun, T., Alfiler, L., Andrade, J., Bohn, P., Dalley, R., Henry, A., Kebede, S., ... Jarsky, T. (2022). Local connectivity and synaptic dynamics in mouse and human neocortex. *Science*, *375*(6585). https://doi.org/10.1126/science.abj5861

Caruana, R. (1994). Learning many related tasks at the same time with backpropagation. In G. Tesauro, D. Touretzky, & T. Leen (Eds.), *Advances in neural information processing systems*. MIT Press. https://proceedings.neurips.cc/paper_files/paper/1994/file/0f840be9b8db4d3fbd5ba2ce59211f55-Paper.pdf

Clements, J., & Bekkers, J. (1997). Detection of spontaneous synaptic events with an optimally scaled template. *Biophysical Journal*, *73*(1), 220–229. https://doi.org/10.1016/s0006-3495(97)78062-7

Delvendahl, I., Kita, K., & Müller, M. (2019). Rapid and sustained homeostatic control of presynaptic exocytosis at a central synapse. *Proceedings of the National Academy of Sciences*, *116*(47), 23783–23789. https://doi.org/10.1073/pnas.1909675116

Denis, J., Dard, R. F., Quiroli, E., Cossart, R., & Picardo, M. A. (2020). DeepCINAC: A deep-learning-based python toolbox for inferring calcium imaging neuronal activity based on movie visualization. *eneuro*, *7*(4), ENEURO.0038–20.2020. https://doi.org/10.1523/eneuro.0038-20.2020

DiAntonio, A., Petersen, S. A., Heckmann, M., & Goodman, C. S. (1999). Glutamate receptor expression regulates quantal size and quantal content at the *Drosophila* neuromuscular junction. *The Journal of Neuroscience*, *19*(8), 3023–3032. https://doi.org/10.1523/jneurosci.19-08-03023.1999

Donahue, J., Hendricks, L. A., Rohrbach, M., Venugopalan, S., Guadarrama, S., Saenko, K., & Darrell, T. (2017). Long-term recurrent convolutional networks for visual recognition and description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(4), 677–691. https://doi.org/10.1109/tpami.2016.2599174

Evans, S. W., Shi, D.-Q., Chavarha, M., Plitt, M. H., Taxidis, J., Madruga, B., Fan, J. L., Hwang, F.-J., van Keulen, S. C., Suomivuori, C.-M., Pang, M. M., Su, S., Lee, S., Hao, Y. A., Zhang, G., Jiang, D., Pradhan, L., Roth, R. H., Liu, Y., ... Lin, M. Z. (2023). A positively tuned voltage indicator for extended electrical recordings in the brain. *Nature Methods*, *20*(7), 1104–1113. https://doi.org/10.1038/s41592-023-01913-z

Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., & Muller, P.-A. (2019). Deep learning for time series classification: A review. *Data Mining and Knowledge Discovery*, *33*(4), 917–963. https://doi.org/10.1007/s10618-019-00619-1

Ferguson, A. R., Nielson, J. L., Cragin, M. H., Bandrowski, A. E., & Martone, M. E. (2014). Big data from small data: Data-sharing in the 'long tail' of neuroscience. *Nature Neuroscience*, *17*(11), 1442–1447. https://doi.org/10.1038/nn.3838

Ho, J., Tumkaya, T., Aryal, S., Choi, H., & Claridge-Chang, A. (2019). Moving beyond p values: Data analysis with estimation graphics. *Nature Methods*, *16*(7), 565–566. https://doi.org/10.1038/s41592-019-0470-3

Holler, S., Köstinger, G., Martin, K. A. C., Schuhknecht, G. F. P., & Stratford, K. J. (2021). Structure and function of a neocortical synapse. *Nature*, *591*(7848), 111–116. https://doi.org/10.1038/s41586-020-03134-2

Huganir, R. L., & Nicoll, R. A. (2013). AMPARs and synaptic plasticity: The last 25 years. *Neuron*, *80*(3), 704–717. https://doi.org/10.1016/j.neuron.2013.10.025

Imbrosci, B., Schmitz, D., & Orlando, M. (2022). Automated detection and localization of synaptic vesicles in electron microscopy images. *eneuro*, *9*(1), ENEURO.0400–20.2021. https://doi.org/10.1523/eneuro.0400-20.2021

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv*, 10.48550/ARXIV.1502.03167. https://doi.org/10.48550/ARXIV.1502.03167

Iqbal, A., Khan, R., & Karayannis, T. (2019). Developing a brain atlas through deep learning. *Nature Machine Intelligence*, *1*(6), 277–287. https://doi.org/10.1038/s42256-019-0058-8

Ishii, T., Stolz, J. R., & Swanson, G. T. (2020). Auxiliary proteins are the predominant determinants of differential efficacy of clinical candidates acting as AMPA receptor positive allosteric modulators. *Molecular Pharmacology*, *97*(5), 336–350. https://doi.org/10.1124/mol.119.118554

Islam, M. Z., Islam, M. M., & Asraf, A. (2020). A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using x-ray images. *Informatics in Medicine Unlocked*, *20*, 100412. https://doi.org/10.1016/j.imu.2020.100412

Jonas, P., Major, G., & Sakmann, B. (1993). Quantal components of unitary EPSCs at the mossy fibre synapse on CA3 pyramidal cells of rat hippocampus. *The Journal of Physiology*, *472*(1), 615–663. https://doi.org/10.1113/jphysiol.1993.sp019965

Kaeser, P. S., & Regehr, W. G. (2014). Molecular mechanisms for synchronous, asynchronous, and spontaneous neurotransmitter release. *Annual Review of Physiology*, *76*(1), 333–363. https://doi.org/10.1146/annurev-physiol-021113-170338

Kavalali, E. T. (2015). The mechanisms and functions of spontaneous neurotransmitter release. *Nature Reviews Neuroscience*, *16*(1), 5–16. https://doi.org/10.1038/nrn3875

Kim, Y. G., Shin, J. J., & Kim, S. J. (2021). Minhee analysis package: An integrated software package for detection and management of spontaneous synaptic events. *Molecular Brain*, *14*(1). https://doi.org/10.1186/s13041-021-00847-x

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv*, 10.48550/ARXIV.1412.6980. https://doi.org/10.48550/ARXIV.1412.6980

Kita, K., Albergaria, C., Machado, A. S., Carey, M. R., Müller, M., & Delvendahl, I. (2021). GluA4 facilitates cerebellar expansion coding and enables associative memory formation. *eLife*, *10*. https://doi.org/10.7554/elife.65152

Kudoh, S. N., & Taguchi, T. (2002). A simple exploratory algorithm for the accurate and fast detection of spontaneous synaptic events. *Biosensors and Bioelectronics*, *17*(9), 773–782. https://doi.org/10.1016/s0956-5663(02)00053-2

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lepeta, K., Lourenco, M. V., Schweitzer, B. C., Adami, P. V. M., Banerjee, P., Catuara-Solarz, S., de La Fuente Revenga, M., Guillem, A. M., Haidar, M., Ijomone, O. M., Nadorp, B., Qi, L., Perera, N. D., Refsgaard, L. K., Reid, K. M., Sabbar, M., Sahoo, A., Schaefer, N., Sheean, R. K., . . . Seidenbecher, C. (2016). Synaptopathies: Synaptic dysfunction in neurological disorders - a review from students to students. *Journal of Neurochemistry*, *138*(6), 785–805. https://doi.org/10.1111/jnc.13713

Lesperance, L. S., Yang, Y.-M., & Wang, L.-Y. (2020). Delayed expression of activity-dependent gating switch in synaptic AMPARs at a central synapse. *Molecular Brain*, *13*(1). https://doi.org/10.1186/s13041-019-0536-2

Li, B., Chavarha, M., Kobayashi, Y., Yoshinaga, S., Nakajima, K., Lin, M. Z., & Inoue, T. (2020). Two-photon voltage imaging of spontaneous activity from multiple neurons reveals network activity in brain tissue. *iScience*, *23*(8), 101363. https://doi.org/10.1016/j.isci.2020.101363

Lowen, S. B., Cash, S. S., Poo, M.-m., & Teich, M. C. (1997). Quantal neurotransmitter secretion rate exhibits fractal behavior. *The Journal of Neuroscience*, *17*(15), 5666–5677. https://doi.org/10.1523/jneurosci.17-15-05666.1997

Malinow, R., & Malenka, R. C. (2002). AMPA receptor trafficking and synaptic plasticity. *Annual Review of Neuroscience*, *25*(1), 103–126. https://doi.org/10.1146/annurev.neuro.25.112701.142758

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Jia, Y., Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, . . . Xiaoqiang Zheng. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems [Software available from tensorflow.org]. https://www.tensorflow.org/

McKinney, R. A., Capogna, M., Dürr, R., & Gähwiler, B. (1999). Miniature synaptic events maintain dendritic spines via AMPA receptor activation. *Nature Neuroscience*, *2*(1), 44–49. https://doi.org/10.1038/4548

Merel, J., Shababo, B., Naka, A., Adesnik, H., & Paninski, L. (2016). Bayesian methods for event analysis of intracellular currents. *Journal of Neuroscience Methods*, *269*, 21–32. https://doi.org/10.1016/j.jneumeth.2016.05.015

Miller, E. C., Teravskis, P. J., Dummer, B. W., Zhao, X., Huganir, R. L., & Liao, D. (2014). Tau phosphorylation and tau mislocalization mediate soluble aβ oligomer-induced ¡scp¿ampa¡/scp¿ glutamate receptor signaling deficits. *European Journal of Neuroscience*, *39*(7), 1214–1224. https://doi.org/10.1111/ejn.12507

O'Brien, R. J., Kamboj, S., Ehlers, M. D., Rosen, K. R., Fischbach, G. D., & Huganir, R. L. (1998). Activity-dependent modulation of synaptic AMPA receptor accumulation. *Neuron*, *21*(5), 1067–1078. https://doi.org/10.1016/s0896-6273(00)80624-8

Passricha, V., & Aggarwal, R. K. (2019). A hybrid of deep CNN and bidirectional LSTM for automatic speech recognition. *Journal of Intelligent Systems*, *29*(1), 1261–1274. https://doi.org/10.1515/jisys-2018-0372

Pernía-Andrade, A. J., Goswami, S. P., Stickler, Y., Fröbe, U., Schlögl, A., & Jonas, P. (2012). A deconvolution-based method with high sensitivity and temporal resolution for detection of spontaneous synaptic currents in vitro and in vivo. *Biophysical Journal*, *103*(7), 1429–1439. https://doi.org/10.1016/j.bpj.2012.08.039

Petersen, S. A., Fetter, R. D., Noordermeer, J. N., Goodman, C. S., & DiAntonio, A. (1997). Genetic analysis of glutamate receptors in drosophila reveals a retrograde signal regulating presynaptic transmitter release. *Neuron*, *19*(6), 1237–1248. https://doi.org/10.1016/s0896-6273(00)80415-8

Pratt, L. Y., Mostow, J., & Kamm, C. A. (1991). Direct transfer of learned information among neural networks. *AAAI Conference on Artificial Intelligence*.

Ralowicz, A. J., Hokeness, S., & Hoppa, M. B. (2024). Frequency of spontaneous neurotransmission at individual boutons corresponds to the size of the readily releasable pool of vesicles. *The Journal of Neuroscience*, e1253232024. https://doi.org/10.1523/jneurosci.1253-23.2024

Reddi, S. J., Kale, S., & Kumar, S. (2019). On the convergence of adam and beyond. *arXiv*, 10.48550/ARXIV.1904.09237. https://doi.org/10.48550/ARXIV.1904.09237

Richards, B., Tsao, D., & Zador, A. (2022). The application of artificial intelligence to biology and neuroscience. *Cell*, *185*(15), 2640–2643. https://doi.org/10.1016/j.cell.2022.06.047

Rupprecht, P., Carta, S., Hoffmann, A., Echizen, M., Blot, A., Kwan, A. C., Dan, Y., Hofer, S. B., Kitamura, K., Helmchen, F., & Friedrich, R. W. (2021). A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging. *Nature Neuroscience*, *24*(9), 1324–1337. https://doi.org/10.1038/s41593-021-00895-5

Rupprecht, P., & Friedrich, R. W. (2018). Precise synaptic balance in the zebrafish homolog of olfactory cortex. *Neuron*, *100*(3), 669–683.e5. https://doi.org/10.1016/j.neuron.2018.09.013

Shi, Y., Nenadic, Z., & Xu, X. (2010). Novel use of matched filtering for synaptic event detection and extraction. *PLoS ONE*, *5*(11), e15517. https://doi.org/10.1371/journal.pone.0015517

Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv*. https://doi.org/10.48550/ARXIV.1312.6034

Sitá, L., Brondi, M., de Leon Roig, P. L., Curreli, S., Panniello, M., Vecchia, D., & Fellin, T. (2022). A deep-learning approach for online cell identification and trace extraction in functional two-photon calcium imaging. *Nature Communications*, *13*(1), 1529. https://doi.org/10.1038/s41467-022-29180-0

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The journal of machine learning research*, *15*(1), 1929–1958.

Tasdelen, A., & Sen, B. (2021). A hybrid CNN-LSTM model for pre-miRNA classification. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-93656-0

Theis, L., Berens, P., Froudarakis, E., Reimer, J., Román Rosón, M., Baden, T., Euler, T., Tolias, A. S., & Bethge, M. (2016). Benchmarking spike rate inference in population calcium imaging. *Neuron*, *90*(3), 471–482. https://doi.org/10.1016/j.neuron.2016.04.014

Theodoris, C. V., Xiao, L., Chopra, A., Chaffin, M. D., Sayed, Z. R. A., Hill, M. C., Mantineo, H., Brydon, E. M., Zeng, Z., Liu, X. S., & Ellinor, P. T. (2023). Transfer learning enables predictions in network biology. *Nature*, *618*(7965), 616–624. https://doi.org/10.1038/s41586-023-06139-9

Tran, V., & Stricker, C. (2021). Spontaneous and action potential-evoked ca2+ release from endoplasmic reticulum in neocortical synaptic boutons. *Cell Calcium*, *97*, 102433. https://doi.org/10.1016/j.ceca.2021.102433

Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., & Nelson, S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, *391*(6670), 892–896. https://doi.org/10.1038/36103

van der Ploeg, T., Austin, P. C., & Steyerberg, E. W. (2014). Modern modelling techniques are data hungry: A simulation study for predicting dichotomous endpoints. *Bller Medical Research Methodology*, *14*(1). https://doi.org/10.1186/1471-2288-14-137

Wang, Z., Yan, W., & Oates, T. (2016). Time series classification from scratch with deep neural networks: A strong baseline. *arXiv*. https://doi.org/10.48550/ARXIV.1611.06455

Ye, L., & Keogh, E. (2009). Time series shapelets. *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. https://doi.org/10.1145/1557019.1557122

Yip, M. C., Gonzalez, M. M., Valenta, C. R., Rowan, M. J. M., & Forest, C. R. (2021). Deep learning-based real-time detection of neurons in brain slices for in vitro physiology. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-85695-4

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *NIPS*.

Zhang, X., Schlögl, A., Vandael, D., & Jonas, P. (2021). MOD: A novel machine-learning optimal-filtering method for accurate and efficient detection of subthreshold synaptic events in vivo. *Journal of Neuroscience Methods*, *357*, 109125. https://doi.org/10.1016/j.jneumeth.2021.109125

Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G. J., Tsang, A., Tsegaye, G., Narayan, S., Obara, C. J., Lim, J.-X., Patel, R., Zhang, R., Ahrens, M. B., Turner, G. C., Wang, S. S.-H., Korff, W. L., . . . Looger, L. L. (2023). Fast and sensitive gcamp calcium indicators for imaging neural populations. *Nature*, *615*(7954), 884–891. https://doi.org/10.1038/s41586-023-05828-9

Zoghbi, H. Y., & Bear, M. F. (2012). Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities. *Cold Spring Harbor Perspectives in Biology*, *4*(3), a009886–a009886. https://doi.org/10.1101/cshperspect.a009886

SUPPLEMENTARY FILES

Table 1. Overview of model architecture

Block	Layer	Output shape	Parameters	Settings	Values
Convolutional Block I	1D Convolutional Layer	(None, 600, 32)	320	Filters Kernel Size Padding	32 9 'same'
	Batch Normalization	(None, 600, 32)	128	All	default
	Leaky ReLU	(None, 600, 32)	0	α	0.3
	Avg. Pooling	(None, 200, 32)	0	Pool Size Strides	3 3
Convolutional Block II	1D Convolutional Layer	(None, 200, 48)	10'800	Filters Kernel Size Padding	48 7 'same'
	Batch Normalization	(None, 200, 48)	192	All	default
	Leaky ReLU	(None, 200, 48)	0	α	0.3
	Avg. Pooling	(None, 100, 48)	0	Pool Size Strides	2 2
Convolutional Block III	1D Convolutional Layer	(None, 100, 64)	15'424	Filters Kernel Size Padding	64 5 'same'
	Batch Normalization	(None, 100, 64)	256	All	default
	Leaky ReLU	(None, 100, 64)	0	α	0.3
	Avg. Pooling	(None, 50, 64)	0	Pool size Strides	2 2
Convolutional Block IV	1D Convolutional Layer	(None, 50, 80)	15'440	Filters Kernel Size Padding	80 5 'same'
	Batch Normalization	(None, 50, 80)	320	All	default
	Leaky ReLU	(None, 50, 80)	0	α	0.3
LSTM Layer	Bidirectional LSTM	(None, 96)	135'936	Units Dropout Rate Merge Mode Activation	96 0.2 'sum' tanh
Fully Connected Layers	Dense	(None, 128)	12'416	Units Activation	128 Leaky ReLU
	Dropout	(None, 128)	0	Dropout Rate	0.2
	Dense	(None, 1)	129	Units Activation	1 sigmoid

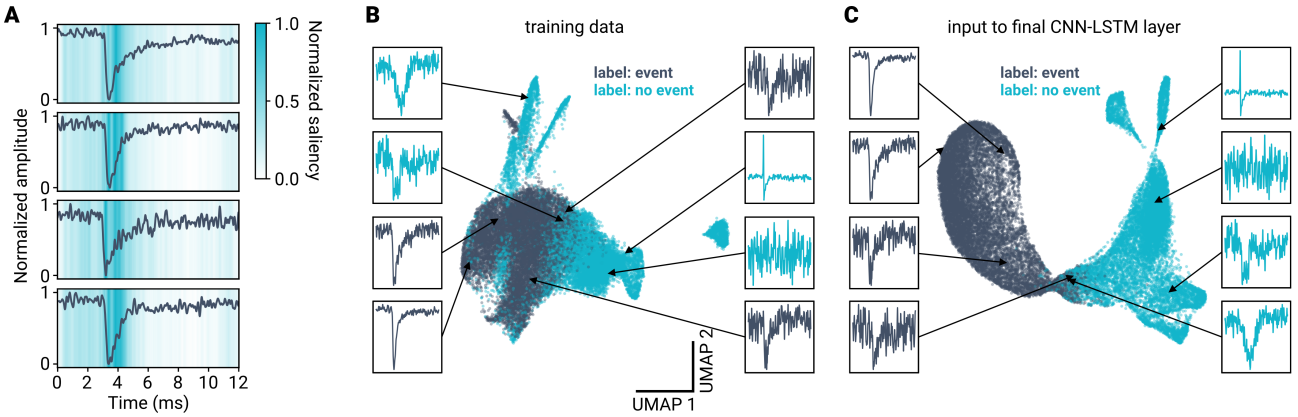

Figure S1. Visualization of model training.
(**A**) Saliency maps (Simonyan et al., 2013) for four example events of the training data. Darker regions indicate discriminative data segments. Data and saliency values are min-max scaled. (**B**) The miniML model transforms input to enhance separability. Shown is a Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction of the original training dataset. (**C**) UMAP of the input to the final ML model layer. Examples of labeled training samples are illustrated. Model training greatly improves linear separability of the two labeled event classes.

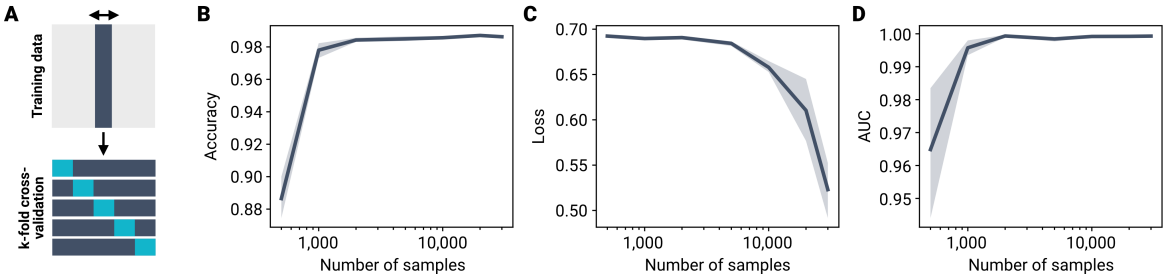

Figure S2. Dependence of training performance on dataset size.
(**A**) To test how size of the training dataset impacts model training, we took random subsamples from the MF–GC dataset and trained miniML models using fivefold cross validation. (**B–D**) Comparison of loss (**B**), accuracy (**C**) and area under the ROC curve (AUC; **D**) across increasing dataset sizes. Data are means of model training sessions with k-fold cross-validation. Shaded areas represent SD.

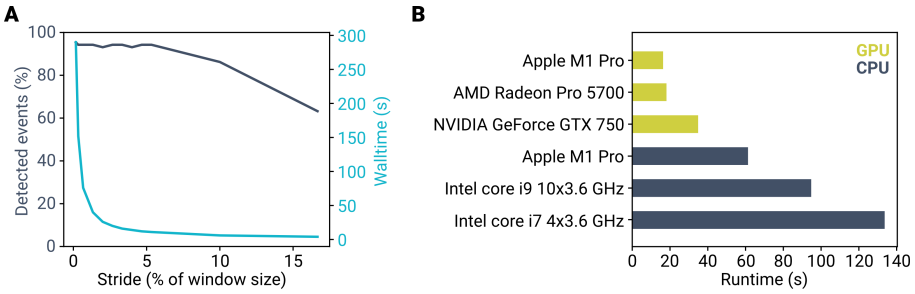

Figure S3. Fast computation time for event detection using miniML.
(**A**) Detected events and analysis runtime plotted versus stride. Note that runtime can be minimized by using stride sizes up to 5% of the event window size without impacting detection performance. (**B**) Analysis runtime with different computer hardware for a 120-s long recording at 50 kHz sampling (total of 6,000,000 samples). Runtime is given as wall time including event analysis. GPU computing enables analysis runtimes shorter than 20 s.

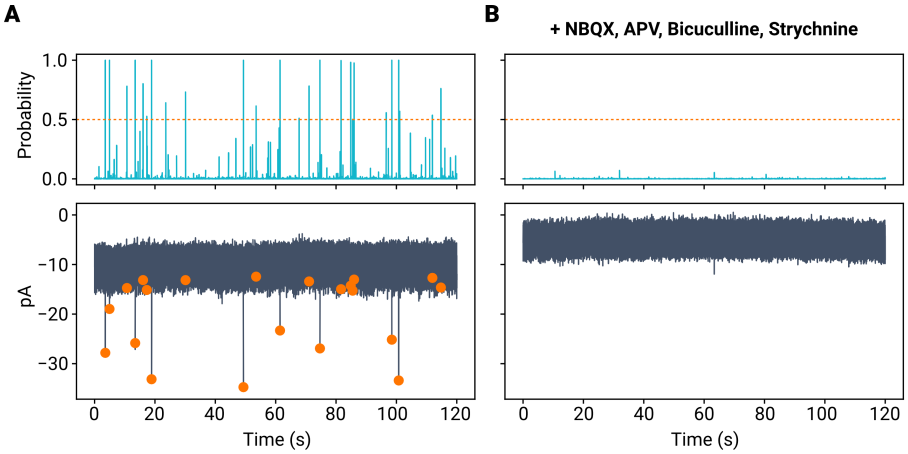

Figure S4. miniML performance on event-free data.
(**A**) Probability (top) and raw data with detected events from a MF–GC recording. Dashed line indicates miniML minimum peak height. (**B**) Recording from the same cell after addition of blockers of synaptic transmission (NBQX, APV, Bicuculline, Strychnine). miniML does not detect any events under these conditions. Note that addition of Bicuculline blocks tonic inhibition in cerebellar GCs, causing a reduction in holding current and reduced noise (Kita et al., 2021).

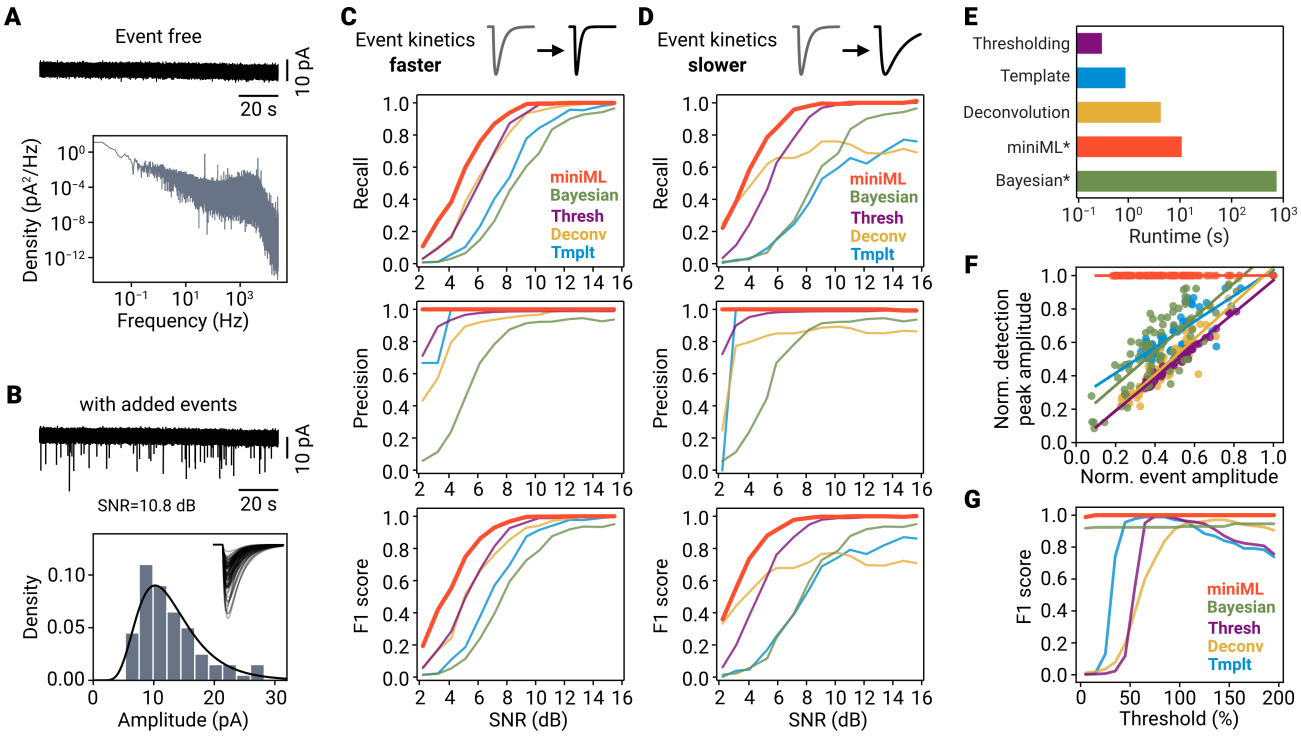

Figure S5. Extended benchmarking and threshold dependence of event detection.
(**A**) Example event-free recording (Top) and power density spectrum of the data (Bottom). (**B**) Top: Data trace from (**A**) superimposed with simulated synaptic events. Bottom: Amplitude histogram of simulated events for a signal-to-noise ratio (SNR) of 10.8 dB. Solid line represents the log-normal distribution from which event amplitudes were drawn; Inset shows individual simulated events. (**C**) Recall, precision and F1 score for five detection methods with simulated events that have a 2× faster decay than in Figure 3. (**D**) Same as in (**C**), but for events with a 4.5× slower decay time constant. Note that we did not adjust detection hyperparameters for the comparisons in (**C**) and (**D**). (**E**) Runtime of five different synaptic event detection methods for a 120-s section of data recorded with 50 kHz sampling rate. Average wall-time of five runs. * miniML was run using a GPU, and data were downsampled to 20 kHz for the Bayesian analysis. (**F**) Normalized amplitude of the peaks in the different methods' detection traces ("detection peak amplitude") versus normalized event amplitude. A linear dependence indicates that detection traces contain implicit amplitude information, which will cause a strong threshold dependence of the detection result. Note that miniML's output does not contain amplitude information. (**G**) F1 score versus threshold (in % of default threshold value, range 5–195) for different methods.

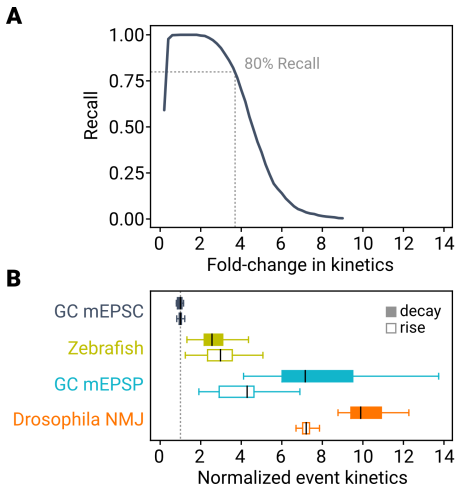

Figure S6. Recall depends on event kinetics.
(**A**) Recall versus event kinetics for the MF–GC model. Kinetics (i.e., rise and decay time constants) of simulated events were changed as indicated. miniML robustly detects events with up to ∼ 4-fold slower kinetics (dashed line indicates 80% recall). (**B**) Event kinetics for different preparations and/or recording modes. Data are normalized to MF–GC mEPSCs (dashed line).

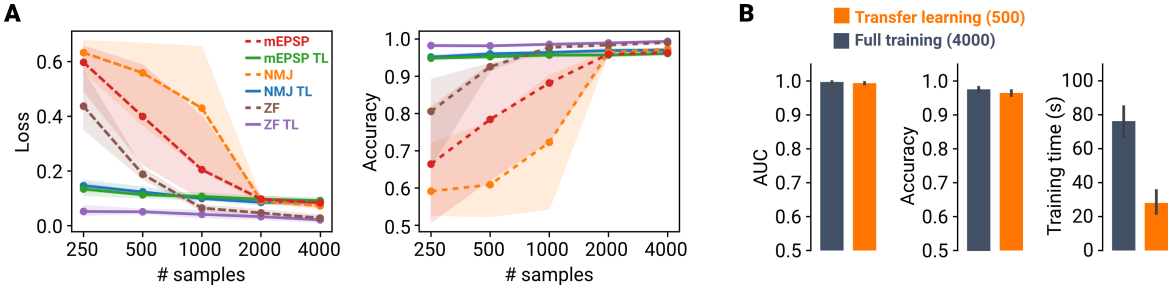

Figure S7. Transfer learning facilitates model training across different datasets.
(**A**) Loss and accuracy versus number of training dataset samples for three different datasets (mouse MF–GC mEPSPs, *Drosophila* NMJ mEPSCs, zebrafish (ZF) spontaneous EPSCs). Dashed lines indicate transfer learning (TL), whereas solid lines depict results from full training. Points are averages of fivefold cross-validation and shaded areas represent 95% CI. (**B**) Average AUC, accuracy, and training time for TL using 500 samples, and full training using 4,000 samples. Error bars denote 95% CI.