Deep Learning of Nanopore Sensing Signals Using a Bi-Path Network

Dario Dematties,¹ Chenyu Wen,¹ Mauricio David Pérez, Dian Zhou, and Shi-Li Zhang*

ABSTRACT: Temporal changes in electrical resistance of a nanopore sensor caused by translocating target analytes are recorded as a sequence of pulses on current traces. Prevalent algorithms for feature extraction in pulse-like signals lack objectivity because empirical amplitude thresholds are user-defined to single out the pulses from the noisy background. Here, we use deep learning for feature extraction based on a bi-path network (B-Net). After training, the B-Net acquires the prototypical pulses and the ability of both pulse recognition and feature extraction without a priori assigned parameters. The B-Net is evaluated on simulated data sets and further applied to experimental data of DNA and protein translocation. The B-Net results are characterized by small relative errors and stable trends. The B-Net is further shown capable of processing data with a signal-to-noise ratio equal to 1, an impossibility for threshold-based algorithms. The B-Net presents a generic architecture applicable to pulse-like signals beyond nanopore currents.

KEYWORDS: neural network, deep learning, nanopore sensors, pulse-like signals, feature extraction

Nanopore sensing technology finds a wide scope of applications exemplified with DNA sequencing,¹ protein profiling,² small chemical molecule detection,³ and nanoparticle characterization.⁴ When analytes pass through a nanopore, characteristic pulses or spikes are generated on monitoring current traces.⁵ Properties of the analytes including size, shape, concentration, charge, and dipole moment can be inferred from the amplitude, width (duration), frequency, and waveform of such spikes.²,⁴,⁶,⁷ Traditional procedures in several different variants to recognize and extract translocation events, i.e., spikes, from noisy current traces are typically based on a user-defined amplitude threshold as a criterion to separate the spikes from background noise fluctuations.⁸,⁹ The flow of data processing for nanopore signals, as well as related algorithms, is a widely accepted establishment (see Note 1 of the Supporting Information (SI)). The determination of spikes is, thus, highly dependent on how the threshold is defined. There is apparently risk that this approach becomes subjective. Although progress has been made in diminishing the subjectivity with empirical selection of the threshold, e.g., defining the threshold by referring to the background noise level, user intervention cannot be totally avoided.¹⁰ Thus, conventional techniques for extracting features from raw data have been historically limited by their capacity. Machine learning based algorithms have, therefore, been considered for processing the nanopore sensing data. To isolate the translocation spikes from the background noise, a hidden Markov model has been adopted to discriminate the open-pore state with a relatively high current level from the blockade state in the spike form with a lower current level. The spikes associated with the current blockade events, as well as their basic features, can, then, be extracted automatically.¹¹,¹² In addition, neural network based deep learning (DL) algorithms have been adapted to the nanopore sensing data of one-dimensional nature. A convolutional neural network (CNN) has been developed to identify the fine patterns in the current blockade state that are caused by a sequence of tailor-designed hairpin loops placed at predefined positions as the barcode of the target DNA.¹³ However, these developments target specific applications and lack the universality for general translocation spike signals. Although several other DL-based algorithms, such as CNN, feed-forward fully connected neural network (FFNN), and long short-term memory, have also been used, most of them have a destined function as a classifier to distinguish different kinds of analytes by the features of their
translocation spikes.\textsuperscript{14,15} They are, hence, incapable of spike reorganization and feature extraction that this work is set to achieve. Specifically, an advanced algorithm based on a DL architecture in the form of a bi-path network (B-Net) is proposed in this work. The B-Net is capable of directly transforming raw data into an appropriate representation, from which certain ending subsystems, such as a classifier, can detect patterns in the input. In other words, the three important features of the spikes, \emph{i.e.}, amplitude, duration, and frequency, can be extracted as a package solution covering the demands for an appropriate nanopore sensing technology.

The B-Net is based on a highly consolidated DL architecture, the residual neural network (ResNet), as depicted in Figure 1. As an integral part of the ResNet architecture, the CNN can be utilized in any dimensional space, although dimensions up to three are the most commonly used depending on the application. For example, three-dimensional CNNs are normally used for volumetric data in medical imaging.\textsuperscript{16} On the other hand, two-dimensional CNNs are the most popular for images and matrices.\textsuperscript{16} As is the case of the B-Net introduced here, one-dimensional CNNs have also been used for signals and time series such as automated atrial fibrillation detection\textsuperscript{17} and sleep arousal detection.\textsuperscript{18}

The B-Net is first evaluated on artificially generated data sets. It is, then, employed for experimental data of \textit{λ}-DNA and streptavidin translocation in solid-state nanopores. Compared to its traditional algorithm counterparts, the B-Net shows improved performance from the perspective of robustness, objectivity, and stability. It is anticipated that the developed B-Net, with necessary adjustment and optimization, is applicable to feature extraction of other pulse-like signals, \emph{e.g.}, transverse tunneling current for single-molecule detection, spikes from neural cells and systems, and electrocardio pulses. It should be clarified that the B-Net concept is designed for handling pulse-like signals. Although DL algorithms have been widely used for treating DNA/RNA sequencing data generated from a nanopore sequencer, the B-Net in its present form is not meant for that. The sequencing data possess essentially different informative features usually as random stepwise jumps among different levels. Processing such sequencing data belongs to a totally different field.

**RESULTS/DISCUSSION**

**Neural Network Architecture.** The holy grail in all DL architectures is depth. The rationale behind these successful methods is built on experimental evidence, which suggests that adding more layers to a deep neural network (DNN) provides us with more abstract output vectors that would better represent the hidden features from raw input data. Such rich features would finally allow us to better perform the task for which the network is trained. Yet, this desirable phenomenon has important limitations since adding more layers to a network comes with penalties, \emph{e.g.}, vanishing and exploding gradients\textsuperscript{19,20} that are often encountered when training artificial neural networks with gradient-based methods. In each training iteration, such methods update the weights of the neural network proportionally to the partial derivatives of error functions with respect to the given weights. In an \textit{n}-layer network, a small/large gradient could end up with becoming vanishingly small/explosively large in early layers of the pipeline, provided that the gradient decreases/increases...
This undesired positive feedback can prevent weights from properly updating their values and potentially impeding further improvement of the neural network. ResNet is adopted in the B-Net to mitigate such risks. ResNet is an artificial DNN that uses CNNs and skips connections or bypasses some layers. The main motivation for bypassing/skipping layers is to avoid the problem of vanishing/exploding gradients. The architecture of the B-Net is visualized in Figure 1. It is a bi-path architecture composed of two parallel ResNets. Both ResNets receive the same input that is simply a temporal window, segmented from a complete nanopore translocation current trace. They also return the predicted number of pulses and average amplitude and duration of such pulses in the window.
The window width should be carefully selected according to the translocation frequency in the target data sets. The window should not be too small such that most of the windows are blank without any spike, which may cause an imbalance during the training process. The window should not be too large to include an excessive number of spikes either because some detailed information may become compromised in the averaged values over one window as the output. Although averaging in each window to obtain the feature values may compromise revealing some details of individual events, our algorithm is capable of capturing the statistical properties, e.g., mean, standard deviation, and distribution function. Most of the commonly used methods are based on the statistical properties of features, instead of the features of single events, to infer the physical parameters of the analytes or to identify/classify the analytes. In contrast, our algorithm advantageously uses the averaged feature values as the output, and the fineness of the results can be tuned by varying the window width. Furthermore, an autocorrelation procedure for window selection based on a few prior tests can be easily implemented to retain the objectivity of the B-Net approach.

As shown in Figure 1, ResNet 1 is assigned to predict the number of translocation spikes (pulses) in the window, while ResNet 2 forecasts the average amplitude and duration of all the pulses found in the same window. ResNet 1 also provides its output as an internal input for the FFNN in ResNet 2. An additional description of the B-Net architecture is provided in Methods and Note 2 of the SI.

**Training and Validation.** ResNet 1 and 2 in Figure 1 were implemented by utilizing another DNN, ResNet 18 in the Pytorch DL framework (https://www.pytorch.org). Since ResNet is originally designed for image processing in two dimensions, necessary modifications were implemented on the architecture to adapt it for the one-dimensional data. Additionally, the last linear layer was replaced by a multilayer perceptron with two layers, and the necessary adaptations were adopted to feed an extra internal input from the ResNet 1 output into the ResNet 2 FFNN, in its final layer. Finally, some batch normalization mechanisms in certain layers of the network were replaced by the group normalization strategy. Our implementation is publicly available.

The B-Net was trained and validated using artificially generated data. It is worth noting that the generated data sets are physics-based involving a set of well-established physical models for nanopore-based sensors. They include the nanopore resistance model, spike generation model, and noise model, by entailing stochastic variations of corresponding parameters in accordance with the related physical mechanisms (Methods and Note 3 of the SI). The B-Net was subsequently evaluated on and applied to both artificial and experimental data (Methods and Note 4 of the SI). Further, five different instances of the network were trained for five different signal-to-noise ratios (SNRs) in the artificial data sets. In all cases, smooth $l_1$-loss and stochastic gradient descent (SGD) optimization were adopted. Additional details about training, such as batch sizes, learning rate schedules, number of epochs, time consumed, curves of loss, and errors for all B-Net instances in this work can be found in Note 5 of the SI.

**Features Extracted from Generated Data Sets.** The general output of the B-Net and the process of performance evaluation are depicted in Figure 2. The B-Net receives a temporal window from a signal trace and returns the predicted
average amplitude and duration of all the pulses as well as the number of pulses in the window (Figure 1). Typical windows with the ground truth are shown in Figure 2b along with the resultant predicted average values of amplitude and duration as well as the count of pulses for artificially generated traces with SNR = 4.

Performance Evaluation. The procedure to performance-evaluate the B-Net is shown in Figure 2a (more details in Methods and Note 4 of the SI). The test data set contains traces for different configurations (Figure 2b) at specified values of concentration of nanospheres ($C_{np}$, correlated with translocation frequency), diameter of nanospheres ($D_{np}$, correlated with spike amplitude), and translocation dwell time (correlated with duration). The relative error of each data window is calculated for each parameter (i.e., number of spikes, amplitude, and duration) referred to the ground truth recorded during the data generation, following the formula shown in Figure 2a (details in Methods). Then, the relative errors are averaged throughout the duration in each situation of $C_{np}$ and $D_{np}$ as displayed in Figure 2c. The blue surface corresponds to the mean value, while the translucent red spike-like surfaces above and beneath the error depict the standard deviation (STD) of the error. The average errors and STDs for each duration in the data set are plotted in Figure 2d. Finally, the total average errors and STDs are computed and shown above each plot (complete evaluation results in Note 4 of the SI).

The outputs and relative errors of the B-Net are compared to those of the traditional algorithm in Figure 3. Different values of $n$ in $th_{n}$ denote the user-defined thresholds of the amplitude as a criterion to distinguish true translocation-generated spikes from fluctuations (noise) in a current trace. Here, $th$ stands for threshold and $n$ is defined as the number of multiples of the peak-to-peak value of the background noise (more details of the implementation of the traditional algorithm in Methods and SI). The comparison has a focus on processing the artificially generated test data set with SNR = 4.

The B-Net and the traditional algorithm agree well in yielding similar variations of amplitude, frequency, and duration (Figure 3a−c). Here, frequency is the ratio of the number of pulses predicted by the B-Net and the time extension of a temporal window (0.5 s). This computation was repeated for each window in the data set where the statistical features were collected. For the traditional algorithm, frequency is defined as the reciprocal of the time interval between the current and the immediate next event.\textsuperscript{10}

The average values of frequency, amplitude, and duration features resulting from the traditional algorithm obviously depend on the selection of the amplitude threshold, as shown by the deviations among the dot-on lines in the figures. Only if the threshold is properly selected can the traditional algorithm return a reasonable result. Small thresholds can lead to assignment of fluctuations as translocation spikes, thereby incorrectly increasing frequency count and lowering both amplitude and duration numerical values. On the other hand, large thresholds can give rise to missing translocation spikes, rendering misleadingly high average amplitude and duration. Furthermore, the feature extraction resulting from the traditional algorithm is clearly less stable than that produced from the B-Net. For example, two small peaks appear in the $th_{10}$ curve at $D_{np} = 8$ and 10 nm in Figure 3a. Such instability arising from the oversensitivity of the traditional algorithm to data generation is obviously absent with the B-Net approach.

The prediction errors for the different features for both algorithms, the B-Net and the traditional, are compared in Figure 3d−f. Except for the prediction of duration in the duration = 5 ms case (Figure 3f), relative errors of the traditional method are notoriously larger than those of the B-Net. Even in the exception, the error of the B-Net is still quite competitive. Strikingly, the relative errors of the B-Net for frequency count are 4 orders of magnitude smaller than those of the traditional algorithm for most $C_{np}$ values (Figure 3d). For amplitude (Figure 3e) and duration (Figure 3f), the relative errors of the B-Net are up to 3 and 2 orders of magnitude, respectively, lower than those of the traditional algorithm. In addition, the relative errors for the traditional algorithm are again highly dependent on the selection of the threshold, confirming the subjectivity of the algorithm. Comprehensive comparisons of frequency, amplitude, and duration for each data set, as well as their relative errors (Note 6 of the SI) affirm the observations above.

Signal-to-Noise Limitation. To explore the capability of the B-Net in identifying signals from a noisy background of various noise levels, five different networks with the same architecture based on ResNet 18 were trained by data sets with SNR = 4, 2, 1, 0.5, and 0.25. The performance of the B-Net in response to the different noise levels is displayed in Figure 4.

![Figure 4. Total average relative errors of frequency, amplitude, and duration vs SNR for the B-Net.](https://doi.org/10.1021/acsnano.1c03842)
The performance of the B-Net is compelling, given that it is almost impossible for the traditional algorithm to correctly recognize translocation spikes in a noisy background with a noise level having a similar amplitude to the spikes, i.e., SNR = 1. Therefore, this capability grants the B-Net the potential to process data from high-bandwidth systems that are usually characterized by high noise levels. By employing larger architectures in combination with labeled generated data sets that are closer to real measured data, it is possible to further reduce errors and facilitate correct processing of signals with even smaller SNRs.

In the B-Net, the features of spikes are acquired by the algorithm during the training process. They contain as much of the original information as possible, indicating not only picking up obvious features considered in the traditional algorithm, e.g., amplitude and duration, but also noticing details in the pulses, e.g., the waveform. These extra features assist the B-Net to better appreciate the difference between a real translocation spike and a background noise peak, even when the two have the same amplitude with SNR = 1. Furthermore, the decision process is flexible and probabilistic, indicating a powerful method with robust performance. As the entire process minimizes the participation and intervention of users, it warrants maximum objectivity; see below.

Objectivity Analysis of Experimental Data. The B-Net is also applied to two experimental data sets for nanopore translocation of λ-DNA and streptavidin from our previous work.10 The outputs produced by the B-Net and those produced by the traditional algorithm are compared in Figure 5a–c for the λ-DNA data set. There is a clear correlation regarding the variation in different features returned by both algorithms. For the B-Net outputs, both frequency and amplitude increase, while the duration decreases with bias voltage. Data analysis shows a linear relationship for all (Figure S20 of the SI), which, to the first-order, is physically plausible. In detail, the linear dependence of amplitude on bias voltage (Figure 5b and Figure S20a) agrees with earlier reports.27,28 This is an expected relationship since for a blockade resistance generated by analyze translocation, a higher voltage induces proportionally a larger current. The translocation frequency also displays a linear dependence on voltage (Figure 5a and Figure S20b), which indicates that the analyze capture probability is limited by diffusion, instead of by the energy barrier of the nanopore since the latter would give rise to an exponential dependence on voltage.29 It has been shown experimentally that long double-stranded DNA (<10^4 bps), such as λ-DNA, usually displays a diffusion-limited capture behavior,30 in agreement with our finding here. Furthermore, a linear decrease of duration with increasing voltage (Figure 5c and Figure S20c) is simply a result of increased electric field giving rise to a proportionally increased rate of translocation with an invariant ionic mobility, because the DNA translocation is electrophoretically governed.31

However, significantly larger dispersions and fluctuations in the output values are evident with the traditional algorithm whose outputs heavily depend on the subjective threshold determined by the user. The deviation among the results from the traditional algorithm with different thresholds is more significant than that for the generated data (Figure 3a–c). Hence, the features extracted from the traditional algorithm lose their objectivity to a certain extent. A similar comparison for streptavidin translocation (Figure 5d,e and Note 7 of the SI) displays also a clear correlation of the variation of the outputs with bias voltage from both algorithms. Compared with the λ-DNA translocation data, the dependence of amplitude and duration on bias voltage is weaker though also linear, in agreement with other reports.7,32 The duration is insensitive to bias voltage, which could be
related to the limited bandwidth at 10 kHz of the amplifier for data acquisition. When the translocation time is close to or shorter than the time constant defined by the cutoff frequency of the amplifier, the difference in the width of pulses is often smeared out.\textsuperscript{33,34} In addition, the long-duration events primarily caused by the interaction between streptavidin and the nanopore are easy to pick up by the readout electronics. They may contribute certain populations to the total events. However, such interaction-caused spikes are independent of bias voltage. According to the simple physical model of diffusive transport,\textsuperscript{34} the translocation frequency is estimated to be 193 Hz. The extracted frequency from the B-Net corresponds to 5–20\% of the theoretical value, which is a reasonable ratio in agreement with other reports\textsuperscript{34} (see Note 8 of the SI for details). Furthermore, the dispersion presented by the traditional algorithm is worse than the one with the \( \lambda \)-DNA translocation data. This difference can be related to a lower SNR of the streptavidin translocation data. It is shown in Figure 5d,e that the frequency and amplitude predicted by the traditional algorithm are even more dependent on the choice of subjective voltage threshold.

CONCLUSIONS

The analyses and comparisons in preceding sections confirm that the B-Net meets the essential and critical requirements of being objective and avoiding subjective parameters determined by the user. The adverse effects of a subjective and often blind input parameter adjustment are clearly appreciated in Figures 3 and 5, where the predictions of the traditional algorithm sensitively depend on a threshold adjusted beforehand. In contrast to the traditional algorithm relying on user-defined input parameters, the advantages of the B-Net lie also in its clear, stable, and consistent predictions as well as negligibly small relative errors. All this is indicative of the robustness of the B-Net in being able to analyze noisy data thereby easing the otherwise strict demands on the control of experimental conditions. The impressive performance of the B-Net in singling out pulses from a noisy background with relative errors below 1\% for count and 5\% for amplitude and duration for input data of SNR = 1 provides a validating example. Such a performance is not anticipated for traditional algorithms since when thresholds are used for recognizing translocation spikes, a noise peak, with similar amplitude, can easily be misclassified as a spike originated from a translocation event. Furthermore, the bi-path architecture in the B-Net assigns different categories of tasks (\textit{i.e.}, pulse count and average feature predictions) to distinct network branches, while the information processed in one branch (\textit{i.e.}, the pulse counter) is used by the other to predict extra average features. This strategy is naturally in agreement with the architecture of the human brain.\textsuperscript{35}

Regardless of all favorable features reviewed above, the B-Net is a DL-based method. As such, it is inherently a data-hungry strategy that works better when there are thousands, millions, or even billions of training examples.\textsuperscript{36} In problems with limited data sources, DL is not an ideal solution. In the specific area of concern, real traces collected from nanopore translocation experiments could be abundant, but they are not labeled. Recruiting staff for labeling such data is not viable given the extensiveness of the data sets needed to train the B-Net. Instead, we have generated our own artificial data sets in this work in order to train, validate, and, then, put in use the B-Net. The comparison experiments conducted against the traditional approaches by processing experimental traces clearly demonstrate that our generated data sets retain a high statistical correlation with the experimental traces collected in the laboratory. Yet, it is imperative to clarify that beyond such favorable results it is impossible to perfectly mimic the statistical distribution immersed in real traces. General palliative methods to solve this problem are available to DL. There are pretraining stages of the networks with which looser requirements are demanded, and smaller labeled data sets can later be used to fine-tune them. Pretraining could be tackled using alternative generated labeled data sets. In this work, we have shown the relevance of the data sets generated to train our B-Net. Our artificially generated data sets can also be used as a pretraining resource in the B-Net. Afterward, adding labels to a much smaller data set of traces collected in the laboratory can be a much more viable endeavor. This could result in a highly qualified network, fine-tuned with real traces. Such a network would show prominent performance differences from a network trained only using artificially generated data sets as the one introduced in this work.

In conclusion, our B-Net algorithm is highly flexible to the input signal and it is not limited to signals from nanopore sensors. A myriad of pulse-like signals, found in biotechnology, medical technology, physical sciences and engineering, information and communication technology, environmental technology, etc., can be processed by implementing the robust and objective B-Net. Therefore, the B-Net is a generalizable and flexible platform owing to the flexibility of DL strategies.

METHODS/EXPERIMENTAL

Data Preparation. Artificial Data Generation. The artificially generated data are composed of three parts: (1) randomly appearing translocation spikes, (2) background noise, and (3) baseline variations. The baseline current level, \( \text{i.e.} \), the open-pore current, is determined using the resistance model,\textsuperscript{24} with given geometry properties, electrolyte concentration, and bias voltage (more details in Note 3 of the SI). In this system, differently sized nanospheres are used to represent analytes. In the signal generation program, the sampling rate is set at 10 kHz to determine the time step of the signal. According to our previous work,\textsuperscript{37} the probability of appearance of translocation spikes at each time step is correlated to the concentration of nanospheres. The amplitude of spikes is assigned by our translocation model based on the resistance change by steric blockade during the translocation.\textsuperscript{38} The waveform of translocation spikes is approximated using an asymmetrical triangle with adjustable ramping slopes (details in Note 3 of the SI).

According to the related studies, colored Gaussian noise is adopted as the background noise\textsuperscript{37} whose power spectrum density is determined by our integrated noise model.\textsuperscript{39} At frequencies below 5 kHz (confined by the 10 kHz sampling rate), the noise has four components: flicker noise, electrode noise, white thermal noise, and dielectric noise. Their respective importance increases successively from low to high frequencies. The related parameters are selected as the typical values of SiN\( _x \). The amplitude of the background noise can be tuned for data sets with different SNR (more details in Note 3 of the SI).

In addition, two kinds of variations of the baseline, \( \text{i.e.} \), sudden jumps and slow fluctuations, are introduced to represent the perturbation. The former generates randomly appearing steps in the baseline to mimic the temporal adsorption–desorption of some objects near and in the pore.\textsuperscript{39,40} The latter simulates the instability of the nanopore, which can be caused by the fluctuation of the pore membrane, dynamics of contamination, etc.\textsuperscript{41,42} It is described by the superposition of eight terms of sine and cosine functions. The probability of appearance of the perturbation steps, the amplitude of
the steps, and the amplitude of the slow fluctuations are carefully tuned to approach the real common situations in experiments (more details in Note 3 of the SI). Involving the baseline variations can largely enhance the robustness of our neural network. Accustomed to complicated scenarios in the training process, the B-Net can handle real measured data more competently.

In signal generation, both diameter and thickness of a nanopore are fixed to 20 nm. Typical experimental conditions are adopted, i.e., a bias voltage of 300 mV, a 100 mM KCl electrolyte, and surface charge density of −0.02 C/m². Three parameters in the signal generation program, i.e., concentration of the nanospheres (aforementioned $C_\text{np}$), diameter of translocating nanospheres (aforementioned $D_\text{np}$), and duration of translocation, are systematically varied for each data set since these three parameters have a direct correlation with the three most important features of the translocation signal, i.e., frequency, amplitude, and duration. In each data set, $C_\text{np}$ is varied from 0.01 to 1 nM, changing in logarithmic scale (i.e., 20 different values in total), while $D_\text{np}$ is varied from 3 to 17 nm, with a 1 nm step (i.e., 15 different values in total). The duration of translocation is directly assigned to 0.5, 1, 1.5, 3, and 5 ms (i.e., five different values in total). To evaluate the performance of the B-Net for different SNR conditions, SNR is varied from 0.25 to 4 (i.e., five different values in total). It is worth noting that SNR is defined as the ratio of spike amplitude to the peak-to-peak value of the background noise. The peak-to-peak value of the Gaussian noise is estimated to be six times of its root-mean-square (RMS) value, while the noise RMS can be calculated by the root square of the integration of the noise power spectrum density in the range of the bandwidth. All the data are generated using a homemade program implemented in MATLAB.

**Experimental Data.** In order to evaluate the performance of our algorithm on experimental data from the laboratory, two groups of translocation experiments are implemented in truncated pyramid nanopores (TPPs). The TPPs are formed in the single-crystal silicon layer of a silicon-on-insulator wafer. Their shape naturally adopts a truncated pyramid by exploiting the difference of wet etching rate of the (100) and (111) crystal orientations of single-crystal silicon. Details about the nanopore fabrication can be found in our previous work.10 Two different TPPs were used in our experiments, one with a side length of 7.5 nm and another 16 nm, both in a 55-nm-thick silicon layer, for DNA and protein streptavidin translocation, respectively. λ-DNA and streptavidin were selected as two typical examples of translocating analytes, representing two typical examples of long and spherical objects, respectively. A λ-DNA strand has 48 502 nucleobase pairs in a double-stranded helix structure with a total length of about 16 μm.11 It carries a high density of negative charge, making its translocation usually governed by the electrophoretic force. Streptavidin is sphere-like with a diameter of 6 nm.1 It is weakly negatively charged in a pH neutral environment, and the translocation is usually dominated by the electrosomotic flow. They were purchased from Merck KGaA, Darmstadt, Germany, and used without further purification. The λ-DNA and streptavidin were dispersed in 500 mM KCl electrolyte with a concentration of 78 pM and 84 nM, respectively.

The nanopore chip was mounted on a custom-made poly(methyl methacrylate) (PMMA) flow cell and sealed using two polydimethylsiloxane (PDMS) O-rings (8 mm in inner diameter) on the two sides.10 Two compartments both filled with KCl solution were separated by the chip, and the only path of ionic current was through the nanopore. A pair of Ag/AgCl pseudoreference electrodes (2 mm in diameter, Warner Instruments, LLC) was used to apply a bias voltage and to collect the ionic current. Electrical measurements were controlled using a patch clamp amplifier (Axopatch 200B, Molecular Devices Inc.). During the translocation experiments, analytic dispersions were added to both compartments. The ionic current was converted to the digital signal by Axon Digidata 1550A (Molecular Devices LLC) and recorded by the software Axon pCLAMP 10 (Molecular Devices LLC). The λ-DNA translocation was measured at a 10 kHz sample rate with 2 kHz analog bandwidth, while the streptavidin translocation was detected at a 20 kHz sampling frequency with a 10 kHz bandwidth.

**Traditional Data Processing Method.** The recognition of translocation events by means of traditional data processing methods is based on the amplitude of spikes. Here, we use a homemade MATLAB program to locate the translocation spikes in current traces and extract the three parameters: translocation frequency, amplitude, and duration. In the program, the function findpeaks is adopted with the MinPeakProminence method. An amplitude threshold is defined by the user regarding the RMS of the background noise level. In the following discussion, this threshold is tuned from 4 to 25 multiples of the background noise RMS to demonstrate the dependence of the results on the threshold selection.

**Standard Database.** A standard database for testing the performance of the translocation signal has been established.14 There are two categories of data, i.e., generated and experimental. For each data set of the generated data, a current trace without background noise is also added, apart from its counterpart trace with noise. Furthermore, the spike information, i.e., amplitude, start time, end time, and duration, is given as the standard answer for corresponding traces. For the experimental data, the current traces are trimmed with the same length. In addition, all the other important information, such as simulation and experiment conditions, is offered. This database is completely open to the public and can be used to train, validate, and test different algorithms on the same page. The database will be enriched, updated, and maintained continually.

**B-Net Training Process Using Artificially Generated Train and Validation Data Sets.** In order to train and validate the B-Net, the artificially generated data introduced in the section Data Preparation were used. Afterward, the B-Net was evaluated on both artificial and experimental data also introduced in the section Data Preparation. The traces in each data set were segmented in temporal windows of 0.5 s. In regard to the artificially generated data sets, we ended up with 60 000 temporal windows for training, 30 000 for validation, and 30 000 for testing. Regarding experimental real data sets, we ended up with 852 temporal windows from our λ-DNA experimental data set and 1512 windows from our streptavidin experimental data set.

The B-Net is composed of two networks, ResNet 1 and ResNet 2, which were trained separately. Five different B-Net instances were trained using five different SNRs in the artificial data sets (SNR = 4, 2, 1, 0.5, and 0.25). In all cases, smooth $l_1$-loss and SGD optimization were adopted. Smooth $l_1$-loss can be seen as a blend of $l_1$-loss and $l_2$-loss. It returns the $l_1$-loss when the absolute value of the argument is high and the $l_2$-loss when such an argument is close to zero. In eq 1, the quadratic segment smoothly transitions the $l_1$-loss near $lx - l y = 0$. This loss is less sensitive to outliers than other measures such as the mean-square error loss. In some cases, it could prevent exploding gradients, which is desirable in networks with architectures like the one used in this work.15

$$\text{loss} = \begin{cases} 0.5(x - y)^2, & \text{if } lx - ly < 1 \\ lx - ly - 0.5, & \text{otherwise} \end{cases}$$

(1)

Smooth $l_1$-loss combines the advantages of $l_1$-loss (producing steady gradients for large values of $lx - ly$) and $l_2$-loss (producing fewer oscillations during updates when $lx - ly$ is small). For SNR equal to 4, 2, and 1, a batch size of 32 temporal windows and an initial learning rate of 0.001 were used. The learning rate was decreased by 90% every 10 epochs for ResNet 1 and every 20 epochs for ResNet 2. For SNR equal to 0.5 and 0.25, a batch size of 8 windows and an initial learning rate of 0.001 were used. The learning rate was decreased by 20% every 10 epochs for ResNet 1 and every 20 epochs for ResNet 2. During the ResNet 2 training procedure, temporal windows without translocation events were ignored. To predict average translocation amplitude and duration, ResNet 2 always received windows containing at least one translocation. On the other hand, ResNet 1 processed all temporal windows during training with and without translocation events. During training, ResNet 2 did not receive the prediction of the number of pulses in the window from...
ResNet 1; it instead received such information from the ground truth in the data set. Such information was injected in ResNet 2 in its FFNN.

Validation after each epoch in the training process was implemented on each network. For ResNet 2, the network with the minimum relative errors of amplitude and duration won and was saved in each epoch iteration. The relative error is defined as

\[ \text{Err} = \frac{|x - x_0|}{x_0} \times 100\% \]  
(2)

where \( x \) is the predicted value and \( x_0 \) the true value. ResNet 1, on the other hand, may process temporal windows without translocation spikes. Consequently, it has a division-by-zero risk if the relative error in eq 2 is applied. Therefore, for the ResNet 1 validation, relative percent difference (RPD), defined by eq 3, was adopted to represent the relative error of the predicted pulse number.

\[ \text{RPD} = \frac{2|x - x_0|}{|x_0| + |x|} \]  
(3)

**Evaluation (Testing) Process Using Artificially Generated Test Data Set.** In order to evaluate the B-Net, a held-out artificially generated data set, used for neither training nor validation, was used. This data set has the same size as the validation data set. Since it is artificially generated, it is completely labeled. Therefore, for each temporal window from the traces, the ground-truth features were known during data generation, including the real number of pulses and the real average amplitude and duration of the pulses in the temporal windows. Then, the relative error between the labeled signals and the predictions produced by the B-Net was calculated, following eq 2.

Unlike training and validation, during evaluation, the B-Net works as follows. First, ResNet 1 processes the temporal window and outputs an estimation of the number of pulses in the window. If the number of pulses predicted by ResNet 1 is 0, ResNet 2 will not process the input and the B-Net will predict 0 for pulses, 0 average amplitude, and 0 average duration. On the other hand, if ResNet 1 predicts one or more pulses in the window, then ResNet 2 will process the input and predict two aforementioned features, i.e., the average amplitude and duration of the pulses in the window. It is important to highlight that ResNet 2 was trained using only temporal windows with one or more translocation events (pulses). It never received windows without pulses, consequently, and only in the case of correct prediction of number of pulses did ResNet 1 prevent ResNet 2 from processing windows without pulses, for which it was not trained.

In addition, considering the division-by-zero risk, those two rules below were followed for the evaluation stage:

- If ResNet 1 correctly predicts 0 pulses in a window, then we consider 0% error for all features, i.e., number of pulses, average amplitude, and average duration.
- If ResNet 1 erroneously predicts more than 0 pulses in a window, when the ground truth is actually 0, then we consider 100% error for all features, i.e., number of pulses, average amplitude, and average duration.

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.1c03842.

Signal processing flow for nanopore sensing; network architecture; physical models and data generation; evaluation (testing) results of an artificially generated test data set for different SNRs; training history using artificially generated train and validation data sets for different SNRs; comparison between the results from our neural network and the traditional algorithm; translocation features of \( \lambda \)-DNA and streptavidin extracted by the B-Net; translocation frequency of streptavidin (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**

Shi-Li Zhang – Division of Solid-State Electronics, Department of Electrical Engineering, Uppsala University, SE-751 03 Uppsala, Sweden; orcid.org/0000-0003-2417-274X; Email: shili.zhang@angstrom.uu.se

**Authors**

Dario Dematties – Instituto de Ciencias Humanas, Sociales y Ambientales CONICET Mendoza Technological Scientific Center, Mendoza M5500, Argentina; orcid.org/0000-0002-8726-7837

Chenyu Wen – Division of Solid-State Electronics, Department of Electrical Engineering, Uppsala University, SE-751 03 Uppsala, Sweden; orcid.org/0000-0003-4395-7905

Mauricio David Pérez – Division of Solid-State Electronics, Department of Electrical Engineering, Uppsala University, SE-751 03 Uppsala, Sweden

Dian Zhou – Department of Electrical and Computer Engineering, University of Texas at Dallas, Richardson, Texas 75080, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acsnano.1c03842

**Author Contributions**

⊥D.D. and C.W. contributed equally.

**Author Contributions**

D.D.: conceptualization, data curation, neural network development and implementation, formal analysis, investigation, validation, visualization, writing (original draft), and writing (review and editing). C.W.: conceptualization, data curation, data generation, translocation experiment implementation, formal analysis, investigation, validation, visualization, writing (original draft, review, and editing). M.P.: investigation, discussion, writing (review and editing). D.Z.: result analysis and discussion, writing (review and editing). S.-L.Z.: project administration, supervision, resources, investigation, writing (original draft, review, and editing).

**Notes**


The generated data sets for all the combinations of the parameters and the experimental detests of \( \lambda \)-DNA and streptavidin translocation are available at https://zenodo.org/record/5013856#.YNYA7xvzZY. The code for B-Net is available at https://github.com/dariodematties/NANOPORE_TRANSLOCATIONS and https://zenodo.org/record/5031817#.YNYAS_kzZY.

**ACKNOWLEDGMENTS**

The authors thank Dr. Shuangshaung Zeng for providing the nanopore chips. They acknowledge the Graphical Processing Units (GPUs) publicly provided by the Google Collaboratory service for training, validating, and evaluating the neural
network architecture proposed in this work. This work was partially financially supported by Stiftelsen Olle Engkvist Byggmästare (No. 194-0646).

REFERENCES


