

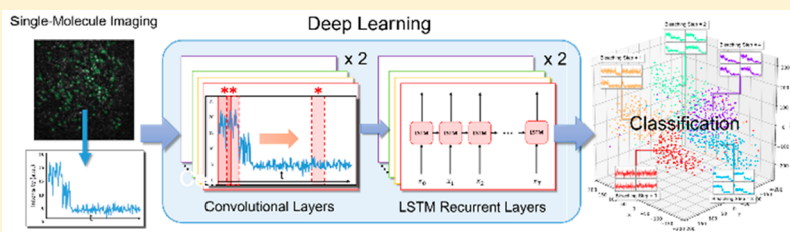
Automated Stoichiometry Analysis of Single-Molecule Fluorescence Imaging Traces via Deep Learning

Jiachao Xu,^{†,‡} Gege Qin,^{†,‡} Fang Luo,^{†,‡} Lina Wang,^{†,‡} Rong Zhao,^{†,‡} Nan Li,^{†,‡} Jinghe Yuan,^{*,†,‡} and Xiaohong Fang^{*,†,‡}

[†]Key Laboratory of Molecular Nanostructure and Nanotechnology, CAS Research/Education Center for Excellence in Molecular Sciences, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100190, China

[‡]University of Chinese Academy of Sciences, Beijing 100049, China

Supporting Information



ABSTRACT: The stoichiometry of protein complexes is precisely regulated in cells and is fundamental to protein function. Single-molecule fluorescence imaging based photobleaching event counting is a new approach for protein stoichiometry determination under physiological conditions. Due to the interference of the high noise level and photoblinking events, accurately extracting real bleaching steps from single-molecule fluorescence traces is still a challenging task. Here, we develop a novel method of using convolutional and long-short-term memory deep learning neural network (CLDNN) for photobleaching event counting. We design the convolutional layers to accurately extract features of steplike photobleaching drops and long-short-term memory (LSTM) recurrent layers to distinguish between photobleaching and photoblinking events. Compared with traditional algorithms, CLDNN shows higher accuracy with at least 2 orders of magnitude improvement of efficiency, and it does not require user-specified parameters. We have verified our CLDNN method using experimental data from imaging of single dye-labeled molecules in vitro and epidermal growth factor receptors (EGFR) on cells. Our CLDNN method is expected to provide a new strategy to stoichiometry study and time series analysis in chemistry.

1. INTRODUCTION

Many important proteins function as multimers or protein complex with different subunits to play crucial roles in a variety of cellular processes including signal transduction, ion transportation, and immune reaction. For example, most receptors on cell membrane such as growth factor receptors, G protein-coupled receptors (GPCRs), and cytokines undergo ligand-induced dimerization or oligomerization to trigger specific cell signaling pathways.¹ The protein stoichiometry is fundamental to protein function and is precisely regulated in cells.² The change in their oligomerization states is also associated with different human diseases including cancers.^{3–7} Therefore, quantitative characterization of protein stoichiometry is of great importance for the investigation of structure–function relationship of proteins as well as protein complex based molecular machines.

Single-molecule fluorescence imaging based photobleaching event counting is an emerging technique for protein stoichiometry study with distinct advantages over the traditional bioanalytical methods.³ The principle of photobleaching event counting, as its name indicated, is based on the assumption that the fluorophore-tagged proteins are photo-

bleached in a discrete steplike way. Each drop of the fluorescence intensity is regarded as a single photobleaching event. The number of the drops reveals the number of subunits in a protein complex; e.g., a monomer shows one-step photobleaching and has two states, and dimer shows two-step photobleaching and three states.⁴ This method enables the real-time monitoring of compositions and transitions of protein complex in living cells under physiological conditions, which is very important for membrane protein study. Moreover, counting at the single-molecule level uncovers the heterogeneous distribution of protein stoichiometry that is often obscured in other ensemble-averaged measurements. Thus, the method has been successfully used to analyze various important proteins such as $\alpha 1E$ Ca^{2+} channel, glutamate-gated NMDA receptors,² transforming-growth-factor- β type II receptor (T β RII),⁵ $\beta 2$ -adrenergic receptors,⁶ epidermal growth factor receptor,⁷ and other molecule complexes including pRNAs of phi29 DNA-packaging motor⁸ and even the single-walled carbon nanotube (SWNT)-based sensors.⁹

Received: January 20, 2019

Published: April 5, 2019

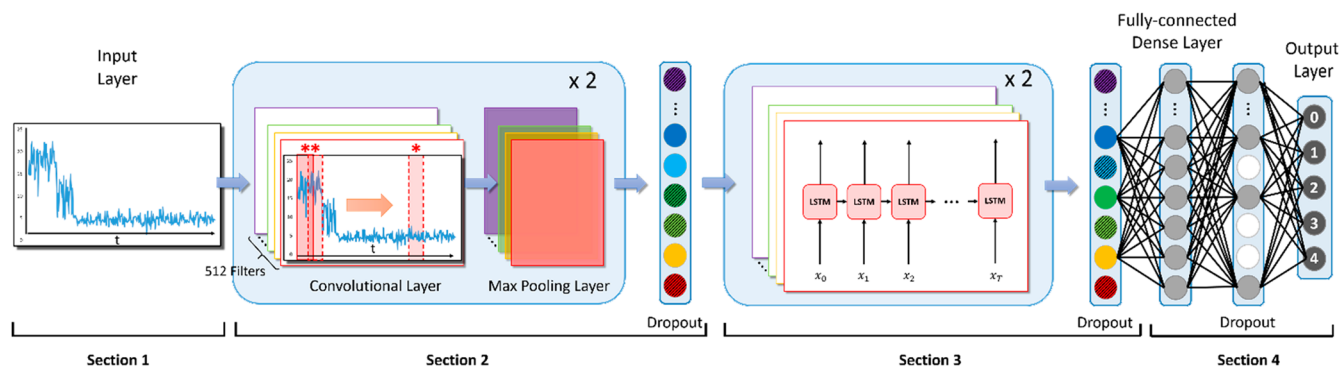


Figure 1. Architecture of the CLDNN for photobleaching event counting from single-molecule imaging data. Different colors represent different filters. Convolution operators are displayed as red “*”.

However, the accurate characterization of protein stoichiometry by single-molecule photobleaching is still challenging. The major reason is that single-molecule fluorescence signals obtained from the complexed biological systems such as living cells are very weak and always obscured by different kinds of noises, including from both the detection devices and biological systems.¹⁰ The noise level often changes stochastically, making it difficult to distinguish the real bleaching steps.⁴ Another problem is the interference of photoblinking, a common property of the fluorophores, which is the phenomenon of random switching between ON (bright) and OFF (dark) states of the emitter.¹¹ In addition, single-molecule photobleaching traces obtained from images under different conditions and from different laboratories are usually quite different in signal-to-noise level and signal change pattern. Even in the same batch of experiment, where thousands of photobleaching traces are obtained in one time, the difference is obvious due to the stochastic changes in biological systems. Thus, the counting method should be objective, efficient, and user-friendly and have extensive applicability.

Up to now, much effort has been made to develop suitable algorithms to solve the above problems. A simple way is to use filters such as Chung and Kennedy’s nonlinear filter¹² and Haar wavelet-based filter¹³ to reduce noise and judgment difficulty of fluorescence drop.^{2,5,14–17} While the photobleaching steps in those filtered fluorescence intensity traces still need to be determined manually one by one, which is subjective and quite time-consuming, more objective and automatic approaches including thresholding against the displacement distribution, running *t*-test algorithm, step-fitting based on χ^2 calculation, and other algorithms have been reported.^{13,18–21} However, the photoblinking events or stochastically varying noises are not under consideration for these methods. To overcome this problem, the hidden Markov model (HMM) has been proposed and widely used,^{22,23} in which the transitions of emitter states can be assumed to be a Markov process. It describes a sequence of possible events in which the probability of each event depends only on the state attained in the previous event.²⁴ Nevertheless, the HMM methods show a weak ability to correlate long-term events and require users to preset several parameters such as initial states and a transition matrix,²⁵ even though iHMM is not completely parameter-free.²⁶ Moreover, all these methods described above need user-defined parameters such as sampling window length, step height, and thresholds. These parameters have significant impacts on the accuracies of these

methods, yet they are usually difficult for users to preset reasonably. This is because the users generally have limited knowledge on either the subunit number of a protein complex or the algorithms designed by others. In practice, we often find problems in the applicability of these methods, as they perform well with some data sets but not others.

Considering the core of single-molecule photobleaching event counting method is to design a feature extraction algorithm which can transform the fluorescence intensity change into correct internal representations or features of steplike drops, the booming deep learning technologies would provide a new approach. Deep learning is a representation-learning method referring to a cascade of multiple layers of nonlinear processing units for feature extraction and transformation.^{27,28} Compared with traditional machine learning methods and other algorithms, deep learning technology has more powerful capacity in feature extraction because it is composed of amounts of simple and nonlinear units that each transforms the representation at one level into a representation at a more abstract level. In principle, with the composition of enough such transformations, very complex functions can be learned to extract features with high precision.²⁹ Moreover, the deep learning process needs little data preprocessing with hand-engineered filters and does not rely on the prior knowledge and human effort in feature design. Meanwhile, benefiting from the recent advances in parallel graphics processing unit (GPU)-accelerated computing, huge volumes of data can be put into the deep learning architecture for training to make it suitable for analyzing data obtained from different situations. Therefore, users do not need to set any parameter to analyze the data with personal subjectivity. The deep learning technologies have shown state-of-the-art performances in a wide variety of fields such as information sciences²⁹ and medical diagnosis.³⁰ In the field of chemistry, application of deep learning technologies is a relatively new trend with only a few reports.³¹ For example, it has been successfully applied to predict the crystal structure,³² toxicity of compounds,³³ and retention time of peptide in chromatography.³⁴ However, these applications did not process information contained in the time dimension. Moreover, the features in single-molecule study could be more difficult to extract since it has much weaker signals and higher noise levels than the other measurements.²⁸ In this work, we have developed a deep learning architecture, which is called convolutional and long-short-term memory deep learning neural network (CLDNN), for the single-molecule data

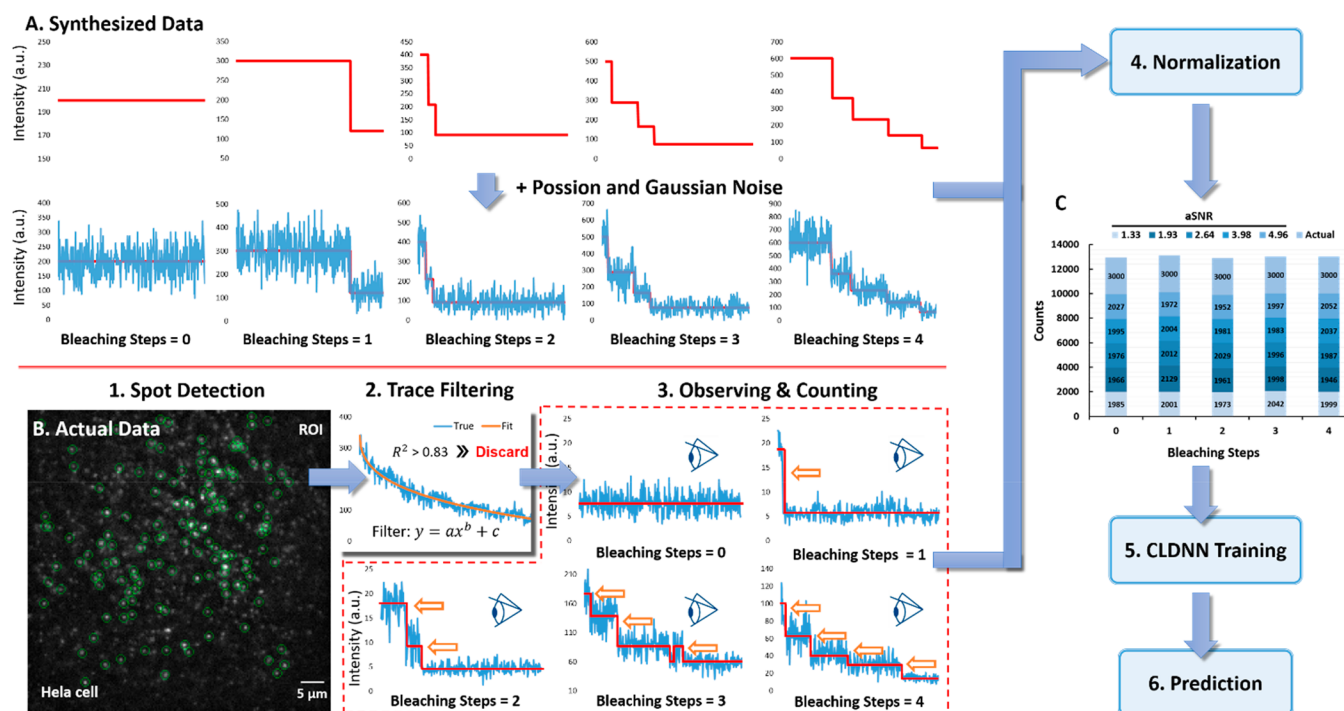


Figure 2. Construction and processing of training data for CLDNN. (A) The construction procedure of synthesized data set. (B) The procedure of extracting and labeling photobleaching traces from single-molecule imaging movies for training. (C) Percentages of synthesized data with different mean aSNRs and actual experimental data. The process of CLDNN training is the following: Step one, cell samples were imaged using a total internal reflection fluorescence microscope (TIRFM) and recorded as movies. Then, the single-molecule spots were detected to extract traces (see the [Supporting Information](#)). Step two, a polynomial function ($y = ax^b + c$) was used to exclude slopelike traces and keep steplike traces. Step three, photobleaching traces were observed and labeled manually. Step four, the data of both synthesized and actual data sets were normalized. Step five, the training data sets were used to train CLDNN. Step six, parameters of the trained model were recorded, and the model was ready to analyze other experimental data.

processing and protein stoichiometry study. This new strategy can be also explored to the analysis of other time series data, such as molecular motor movement, etc.

The CLDNN is composed of both convolutional layers with a set of convolutional operators to extract abstract features of steplike drops, and longshort-term memory (LSTM) recurrent layers to “remember” values over arbitrary time intervals like “memory” to distinguish between photobleaching and photoblinking events. Since the single-molecule data has limited signal-to-noise levels, we designed the convolutional layers with many convolution kernels to ensure accurate feature extraction. Furthermore, we trained the CLDNN with not only synthesized data but also data set including actual samples from experiments to make it more practical. We have verified our CLDNN method with both simulated data and experimental data from single-molecule imaging of Cy5-labeled DNA sample in vitro and membrane proteins on cells. According to our result, the CLDNN method outperformed other algorithms with enhanced accuracy and at least 2 orders of magnitude improvement of efficiency.

2. RESULTS AND DISCUSSION

2.1. Architecture of CLDNN. The framework of CLDNN which combines convolutional and LSTM recurrent layers is shown in [Figure 1](#). All the layers of CLDNN can be grouped into five sections as following.

In the first section, the input to the CLDNN is fluorescence intensity sequence against time with a sampling rate at f Hz for T s (e.g., 10 Hz for 40 s used in our experiments).

The second section acts as feature detectors to provide abstract representations of the raw data in feature maps. It consists of five layers, that is C (convolutional layer)–M (max pooling layer)–C–M–D (dropout layer).

The convolutional layer convolves the input or the previous layer’s output to extract a feature map with one-dimension (1D) kernels (to be learned in the training process, see the [Supporting Information](#)). The shallow convolutional layer extracts the local and simple features such as lines, corners, and edges, while the deep convolutional layer can abstract more global and complicated features.

The max-pooling is a process of subsampling designed to progressively reduce the spatial size of the representation. It finds the maximum feature map over a set of nearby temporal sequence data. Therefore, it can reduce amount of computation and control overfitting (see the [Supporting Information](#)).

Dropout, as a type of regularization technique, randomly abandons some neurons (units) from the layer to reduce the reliance on specific units and prevent overfitting.³⁵

The third section of CLDNN further improves the classification accuracy with temporal information and especially works as a photobleaching or photoblinking event discriminator. It includes two LSTM recurrent layers and one dropout layer. Although the convolutional layer can recognize the feature of steplike drop, it cannot distinguish the drop from photobleaching or photoblinking. In contrast, the LSTM network takes into account the information contained in the time dimension. It can remember the photobleaching events but ignore or forget photoblinking events to recognize the true

steplike drops. Briefly, LSTM unit, as an extension of recurrent neural network (RNN) unit, can be considered as multiple copies of the same neurons, passing message from one to successor with long-term memory. The structures of LSTM units are shown in [Supporting Information Figure S1](#). It makes full use of the concept of gating and uses the input gate, forget gate, and output gate to control what information should be forgotten, remembered, and updated (see the [Supporting Information](#)).

Assuming in one scenario, a blinking event including a sudden drop and rise of fluorescence intensity happens. The state of LSTM unit C_{t-m} records the drop event at moment $t - m$ and the value is passed to the C_{t-1} . After judging the value of C_{t-1} , forget gate closes and reserves the drop event (eq ES3). At the moment t , a rise event occurs to open input gate and \tilde{C}_t (the candidate value for the state of LSTM unit) records this rising event (eqs ES4–S5). The whole blinking events is recorded as an updated new state value C_t (eq ES6). Based on the judgment of a well-trained CLDNN, the output gate is closed (eq ES7). According to eqs ES8–S9 for output, this blinking event is not recorded at moment t .^{36–40}

The fourth section of CLDNN is designed to map the values from the previous layer into output classes, which determines the number of bleaching steps. The fully connected dense layer is a standard multilayer perceptron neural network to achieve the parametric-concatenation (see the [Supporting Information](#)). For the photobleaching data, researchers only care about the actual number of steps. The output layer is governed by the softmax function ($Z_i = \frac{e^{z_i}}{\sum_{k=1}^C e^{z_k}}$), where C is the number of output classes (bleaching steps). $f(Z_i)$ is the output of the i th neuron. This softmax function provides the posterior probability of the classification results and outputs the class with the highest probability.

2.2. Model Training. The CLDNN was implemented using Keras,⁴¹ an open source neural network library written in Python, to build and train neural networks. The model training and classification were run with a general hardware requirement on a GeForce 1080Ti GPU with 3584 CUDA cores, 1708 MHz clock speed, and 11 GB RAM.

The training data consisted of two subsets, synthesized data and experimental data.

The major subset (50 000 traces) was synthesized according to the method described in our previous report.⁴² Briefly, each synthesized fluorescence intensity trace included 400 sampling points. The step number (0–4), locations, and heights of the drops were initiated randomly. Each state included at least five sampling points. Next, Poisson distributed shot noise and random Gaussian distribution noise were added to the noiseless trace ([Figure 2A](#)). Data sets with different signal-to-noise values (SNRs, equal proportions of samples with aSNR from 1.33, 1.93, 2.64, 3.98 to 4.96) were synthesized for training ([Supporting Information Figure S2B](#)). We defined the value of SNR⁴ as

$$\text{aSNR} = \begin{cases} 0 & k = 1 \\ \frac{\sum_{i=1}^{k-1} \frac{2\mu_i}{\sigma_i + \sigma_{i+1}}}{k-1} & k = 2, 3, 4, 5 \end{cases} \quad (1)$$

where k is the number of states (s). μ_i and σ_i are the mean and standard deviation (SD) of the i th state ([Supporting Information Figure S2A](#)). An exemplary three-step trace with

different aSNR is shown in [Supporting Information Figure S3](#). As we could observe a few emitters which did not experience photobleaching during the single-molecule imaging experiment, we also set zero-step traces in our synthesized data.

The data of the minor subset (15 000 samples) came from the experimental data of actual single-molecule images ([Figure 2B](#)). The reason for training CLDNN with experimental data is even though we had elaborately designed the synthesized data sets to simulate the actual data, it was still hard to include all the features of the real photobleaching traces. With this in mind, we collected mounts of photobleaching traces including bleaching and blinking events of proteins labeled with different fluorophores and used them to train the CLDNN to make the network more practical. The photobleaching traces were extracted from movies using the program as previously reported^{5,43} (see the [Supporting Information](#)), and the photobleaching steps were counted manually according to the uniform judgment of four independent researchers in this work. For the experiment-derived traces used in the training process of CLDNN, we only selected those showing easy-recognizable photobleaching steps by eye. After training, CLDNN is expected to judge traces without clear step drops more accurately and objectively than human.

It is noted that according to our previous experimental data, the proportion of protein complexes with more than four subunits was no more than 1% and had little impact on the statistic distribution of bleaching steps. Moreover, reports from other groups confirmed that for multimers of more than four subunits, the results from photobleaching step counting by eyes became unreliable due to the decreased SNR.^{44,45} Therefore, we only set the step number from 0 to 4 in our CLDNN.

Finally, the data sets from both synthesized data and experimental data were normalized to values ranging from 0 to 1 and then put into the CLDNN for training. The detailed percentages of synthesized data with different aSNRs and actual data are shown in [Figure 2C](#).

The model was trained in a fully supervised way. The network parameters were optimized by minimizing the cross-entropy loss function using mini-batch gradient descent with the Adam update rule.⁴⁶ We gradually adjusted the number of parameters in each layer to achieve the high accuracy and computational efficiency of CLDNN. The model was trained with an optimized learning rate of 0.001. Weights are randomly initialized orthogonally. The parameters presented in the network are detailed in [Supporting Information Table S1](#).

2.3. Model Evaluation. We trained the CLDNN for 20 epochs in total with train-val-test splitting method (holdout method). In every mini-batch or epoch in the process of training, 90% data in the training data set was used for training while the other 10% data for validation. Then a different test data set ([Supporting Information Figure S4](#)) was used to evaluate the trained CLDNN. The loss and accuracy curves for training and validation are shown in [Figure 3A](#). The model reached 90.56% accuracy on the validation data set after 12 epochs and 92.07% after 20 epochs. The validation accuracy was greater than the training accuracy in every epoch, indicating that the model did not overfit the training data set.⁴⁷ The confusion matrix was used to find classification error sources of this model by using the test data set. As illustrated in [Figure 3B](#), CLDNN performed very well on the step-determination of traces with few errors. The majority of errors came from the classification between three-step and four-step

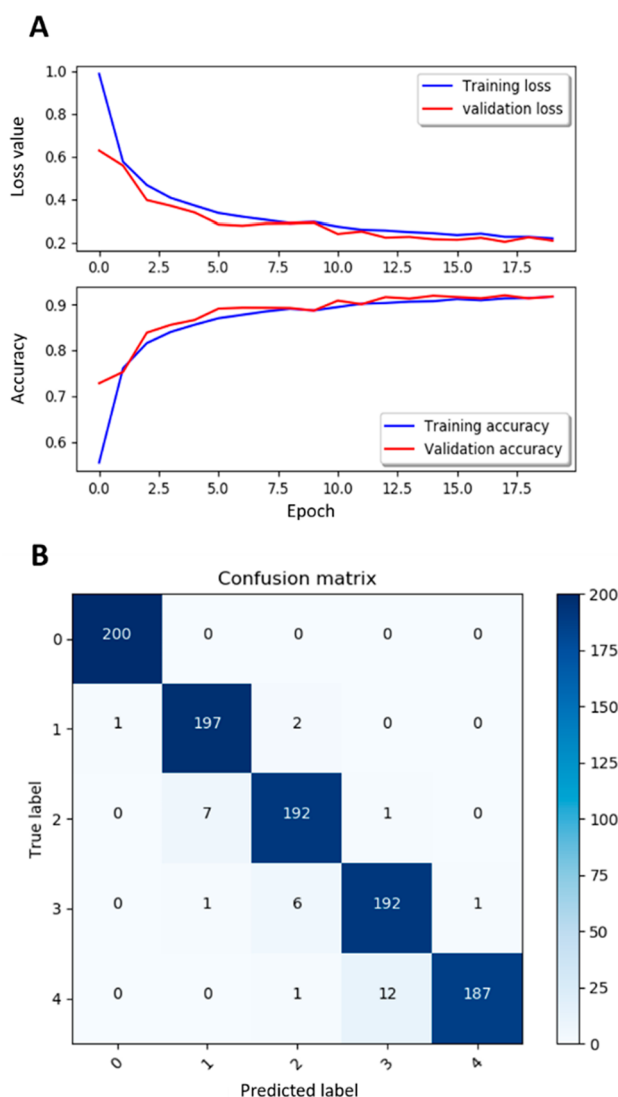


Figure 3. Evaluation results of CLDNN. (A) The loss and accuracy curves for training and validation. (B) The confusion matrix of the evaluation results. The numbers of zero, one, two, three, and four-step traces are all 200 in the test data set (Supporting Information Figure S4) for evaluation.

traces. With the increasing number of emitters in one fluorescent spot, the intensity fluctuation caused by noise became larger and even obscured the real “drop” of the signal, which might make it more difficult for the model to identify the actual bleaching steps. However, the proportion of these errors was only 6.5% $((1 + 12)/200)$ and thus had little impact on the statistic distribution of bleaching steps.

We also visualized output of some layers in the CLDNN to check whether these layers implemented their functions as we expected. As showed in Figure 4 and Supporting Information Figure S5, a two-step trace (the drops of signals occurred at moment t_{152} and t_{350} , respectively) with a blinking event (from t_{20} to t_{26}) was put into the CLDNN for prediction. The feature maps outputted by the second convolution layer (Conv1d_2) showed that the features, “drops”, have been extracted and transformed into peaks through the convolution layers. While there were some interferential minor peaks, the characteristic peaks were prominent and corresponded to drop moments ($t_9 - t_{20}$ (a drop from blinking), $t_{75} - t_{152}$ (a drop from

bleaching), and $t_{173} - t_{350}$ (a drop from bleaching). (Note: the total number of frames outputted by Conv1d_2 transformed into 200 because of the previous operation of max pooling.) However, the drop feature from blinking could hardly be excluded by convolution layers. Next, the feature maps extracted were put into LSTM layers. As expected, the LSTM layers worked well as photobleaching or photoblinking event discriminators. The drop feature from blinking was not recorded, and interferential minor peaks were also excluded for most outputs of LSTM layer 1. The final classification result (bleaching steps) from output layer was two (the corresponding largest probability was $9.99\text{E}-1$) to this trace as we expected (see the Supporting Information). From the outputs of these randomly chosen layers, we confirmed the CLDNN layers worked well as we designed. Therefore, the CLDNN could effectively extract and distinguish photobleaching/photoblinking event.

In addition, we examined the internal feature vectors learned by the CLDNN using *t*-SNE (*t*-distributed stochastic neighbor embedding).⁴⁹ As shown in Figure 5, each point represented a photobleaching trace projected from the 512-dimensional output of the CLDNN’s last dense layer into three dimensions. Obviously, traces with different steps could cluster into different groups, indicating that features vectors extracted by CLDNN were effective to the data classification.

It is noted that we trained CLDNN with five mean aSNRs in the range from 1.3 to 5, as the SNR values of most single-molecule imaging experiments and data processing algorithms are reported in this range.^{42,48} We also tested the generalization of this five aSNR-trained CLDNN with a wider SNR range (Supporting Information Figure S6A). When the mean aSNR is above 3, the accuracy reached a platform which was above 99%. When the mean aSNR was less than 2, the accuracy showed an obvious aSNR-dependent decrease. As the SNR in single-molecule imaging studies is generally above 2.5, the accuracy of our method is high enough. Even when the mean aSNR was 1.5, the accuracy was higher than 80%. This indicated that the CLDNN can be used for different SNR levels (Supporting Information Figure S6B).

2.4. Comparison of Different Algorithms. Synthesized test data sets with means of aSNRs at 1.37, 1.91, 2.67, 3.89, and 4.91 (1000 samples for each data set without zero-step traces) were used to compare the accuracy and computational efficiency of different algorithms. Compared with the HMM combined with maximum-likelihood clustering,⁴² a method our group developed previously, CLDNN showed higher accuracies to the data sets with different aSNRs and performed better especially for low-aSNR data sets. In the case of low aSNR (1.37), CLDNN achieved $78.4 \pm 1.0\%$ accuracy (mean \pm s.d.), which was 26.6% higher than that of the HMM method ($51.8 \pm 1.0\%$). Furthermore, CLDNN achieved $96.5 \pm 0.2\%$ accuracy when the mean aSNR of data sets was 2.67, which was about the same signal-to-noise level in our single-molecule imaging data (Figure 6A).

In addition, we tested two other methods including NoRSE⁹ and PIF⁴⁸ using recommend parameters in previous reports and also tuned the parameters to achieve the best performance. The results showed that the accuracies of these two methods were lower than CLDNN under all the aSNRs (Supporting Information Figure S7). The possible reason is that their feature detectors (NoRSE used intensity-peak discretion and PIF used intensity drop degree as feature detectors) were weaker than that of CLDNN. Moreover, while CLDNN can

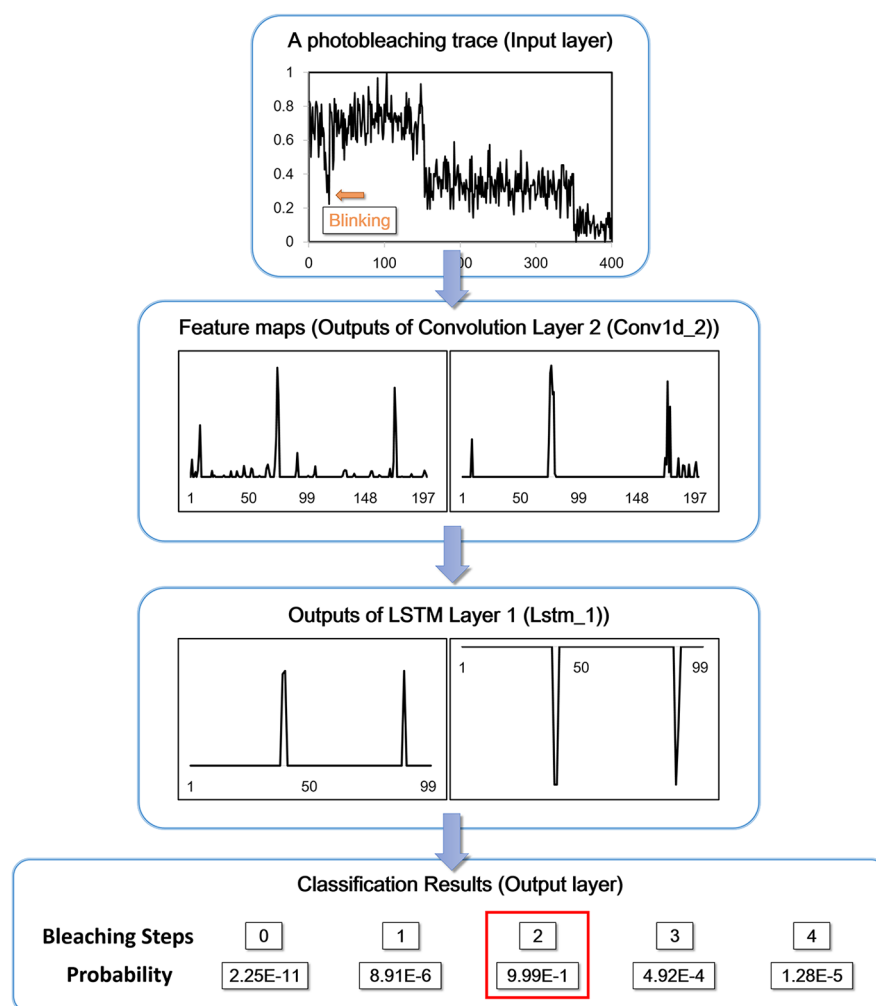


Figure 4. Visualization of the outputs of some layers in our CLDNN. The orange arrow indicates the location of blinking event. Two feature maps outputted by the convolution layer 2 are illustrated in the second box. The output layer returns the corresponding probabilities to different classes (bleaching steps = 0–4) and outputs the class with largest probability (red box).

update its feature detectors in an autolearning way, users must set reasonable thresholds to feature detectors of the other two methods, which may be unfriendly to the novices.

In the photobleaching data analysis, besides the accuracy, researchers also care about the statistic distribution of bleaching steps, which indicates the population of protein with different stoichiometry. Herein, 1000 synthesized samples without zero-step traces (the mean of aSNR was 2.64) were tested. The proportions of one, two, three, and four-step traces were 23.6%, 25.4%, 24.8%, and 26.2%, respectively. We defined a function to evaluate deviation degrees of proportion (S_M) as below.

$$S_M = \sum_{k=0}^{S+} (|P_k^M - P_k^T|),$$

$$M = \text{CLDNN, HMM, PIF, NORSE} \quad (2)$$

where M represents different methods including CLDNN, HMM, PIF, and NORSE. T is the actual distribution. k is the number of steps. P_k^M represents the proportion of k -step analyzed by method M .

The S_M of CLDNN was 3.0%, significantly lower than that of HMM (15.4%) and other methods (Supporting Information

Figure S8). All methods including CLDNN displayed the weakest classification ability in four-step samples. Nevertheless, the proportion of four-step samples analyzed by CLDNN was 24.8% (only 1.4% lower than the actual value), while 18.7% analyzed by HMM. This result indicated that when the data included more bleaching steps, it became more difficult for these methods to distinguish the real drops from noise. CLDNN adapts numerous convolution kernels to adequately extract features with no information loss caused by profiteering and learns them in a deep way, leading to a more powerful capacity of extracting real events from details (Figure 6B).

The computational efficiencies of different methods were also compared using the same synthesized test data sets (1000 samples without zero-step traces), and the results are listed in Supporting Information Table S2. CLDNN only needed 0.41s (even ran on the same hardware platform using CPU as other methods), which was 3 orders of magnitude faster than the HMM method (678.37 s). Another significant advantage for deep neural network is that unlike other methods, it does not need to update and optimize parameters for different data after training.

2.5. Analysis of Experimental Data. After evaluation of CLDNN with synthetic data, we moved to apply it to real-sample experiment. Two samples were used. The first was a

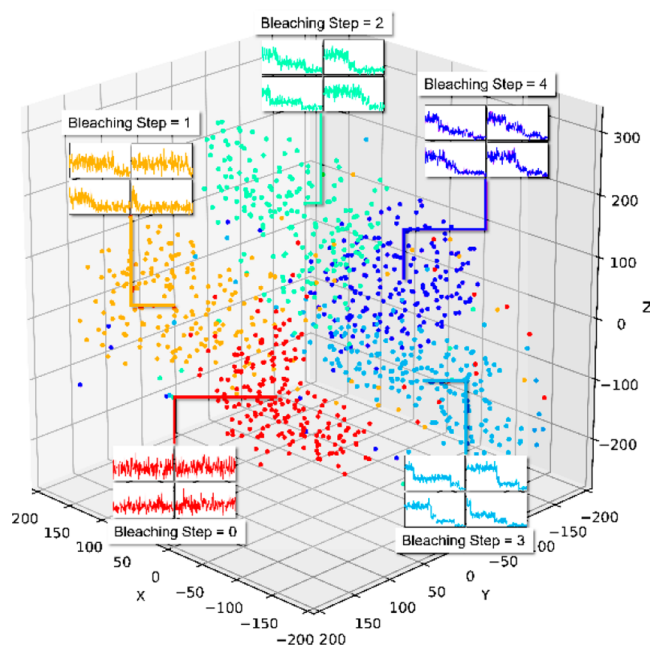


Figure 5. *t*-SNE visualization of the feature vectors outputted by the last dense layer in the CLDNN for different bleaching steps. The high-dimensional data were visualized in the CLDNN of the validation data set (1000 samples). Different colored point clouds represent the traces with different bleaching steps. Insets show typical traces corresponding to different classes.

standard one-step bleaching sample of dye-labeled ssDNA which was immobilized on surface. In brief, a flow chamber composed of a microscope slide and a cover slide was passivated with methoxy polyethylene glycol (mPEG)⁵⁰ to prevent nonspecific adsorption. Then, the cover slide was immobilized with biotinylated PEG and streptavidin, followed by biotinylated single-strand DNA (ssDNA) labeled with Cy5 (Figure 7A). The density of fluorescent spots was <0.3 molecules per μm^2 to ensure single-molecule imaging within optical resolution. The sample was imaged with a custom total internal reflection fluorescence microscope (TIRFM), and the data were analyzed by CLDNN.

As shown in Figure 7B, 92% of Cy5-labeled ssDNA fluorescent spots were photobleached with single drops, as expected. A few spots exhibited two-step bleaching, probably because two fluorescent molecules were located within the region of diffraction limit. The CLDNN-derived distribution of bleaching steps was similar to that of monomeric YFP immobilized on the glass in previous report (the percentage of one-step molecule was 93%).⁵⁰ The results confirmed the validation of our CLDNN.

Next, we applied CLDNN to analyze the dimerization of epidermal growth factor receptors (EGFR) upon EGF stimulation on cell membrane. It is well-known that the activation process of EGFR is initiated with ligand-induced dimerization of the monomeric receptors on the membrane⁵¹ (Figure 7C).

We imaged the Hela cells expressing EGFP-tagged EGFR with or without EGF stimulation. It was shown that the majority (61.45%) of EGFR is monomer on the plasma membrane of resting cells. After EGF stimulation, the proportion of monomers decreased from 61.45% to 37.40%, while that of dimers increased from 39.24% to 56.25% (Figure 7D). Dimers became a dominant population as expected.

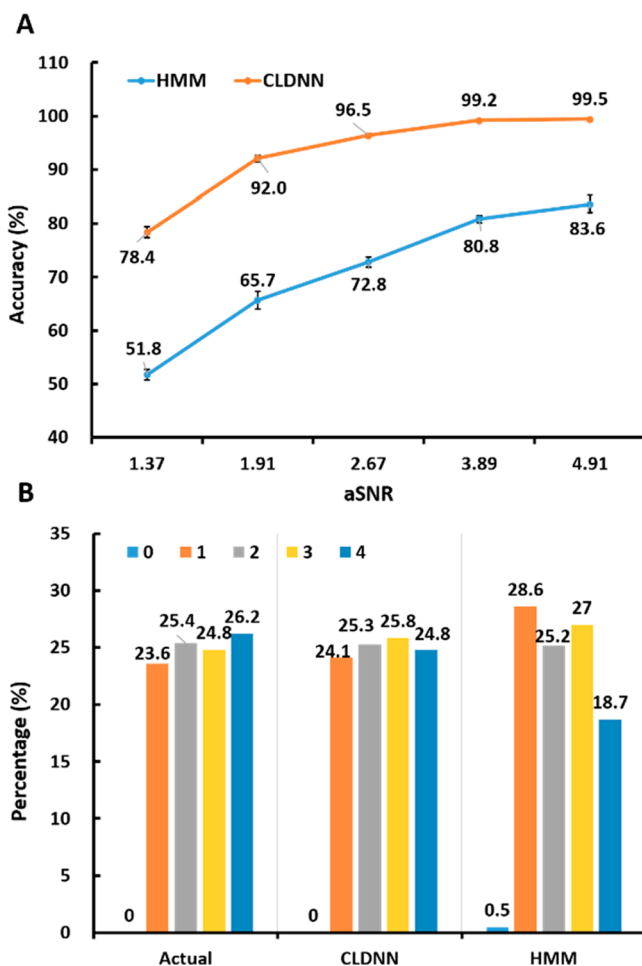


Figure 6. Comparison of different methods on the synthesized test data sets. (A) The analysis accuracies of CLDNN and HMM to the synthesized data sets with means of aSNRs at 1.37, 1.91, 2.67, 3.89, and 4.91 (750 traces were randomly sampling from 1000 samples three times). (B) The true distribution and outputted distributions of bleaching steps analyzed by different algorithms. The synthesized data sets only include one, two, three, and four-step with mean of aSNR being 2.64. S_{CLDNN} is 3.0%, and S_{HMM} is 15.4%.

Therefore, the results analyzed by CLDNN were reasonable. The applicability of this method to the experimental data has been demonstrated.

3. CONCLUSIONS

In summary, we have developed a novel photobleaching event counting method using deep learning technologies. We proposed, for the first time, a neural network architecture, CLDNN, composed of convolutional and LSTM recurrent layers to analyze single-molecule fluorescence photobleaching data. This method does not need preprocessing such as filtering, which is benefited from the powerful feature extraction capability of convolutional layers. Meanwhile, LSTM layers with excellent memory ability are successfully applied to distinguish between photobleaching and photoblinking events, which further improve the accuracy.

Compared with previous methods, CLDNN presents many advantages. First, it effectively eliminates the interference of blinking events to the analysis and shows higher accuracy even with low-SNR data. Second, it exhibits amazing computational efficiency to analyze data in a high-through way. Third, this

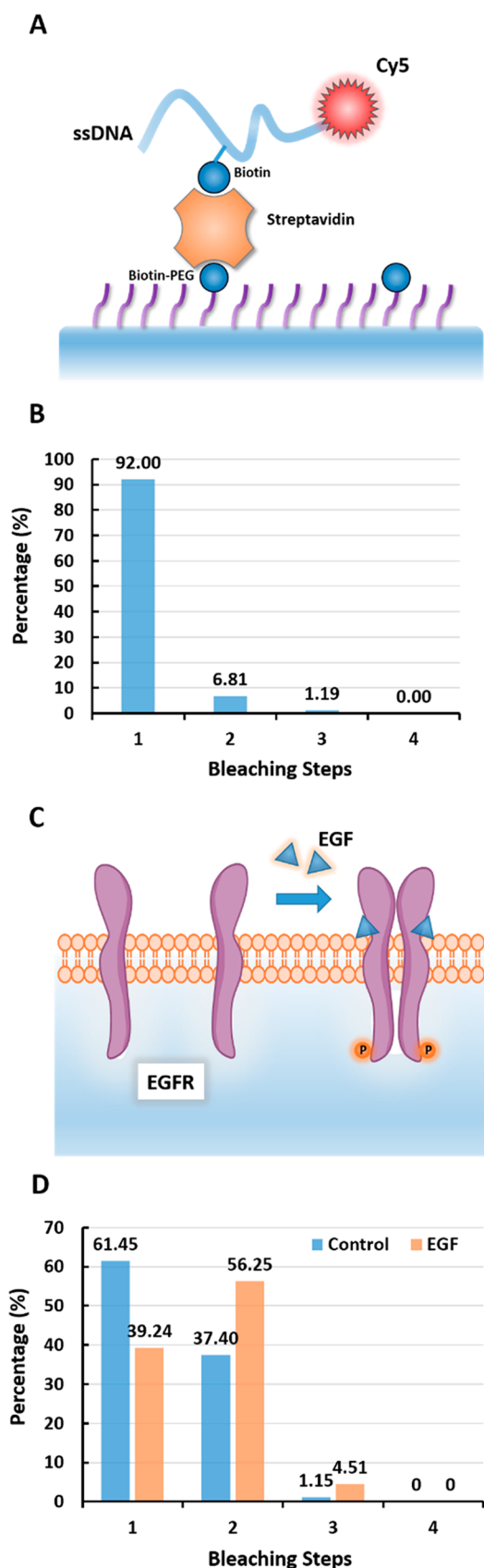


Figure 7. Results of CLDNN for experimental data analysis. (A) The schematic diagram of a standard one-step bleaching sample of Cy5-labeled ssDNA. (B) The distribution of bleaching steps of the ssDNA sample analyzed by CLDNN. (C) A model of EGFR activation upon

Figure 7. continued

EGF stimulation. (D) The distribution of bleaching steps of EGFR on cell membrane analyzed by CLDNN.

method is user-friendly and completely parameter-free for users once the training is done, which ensures the objectivity. Fourth, users can also adopt their specific data set to train the CLDNN to meet their individual analysis requirements with extensive applicability. Moreover, this new strategy can be explored to the analysis of other time series data in electroanalytical chemistry, chemical physics, and analytical chemistry such as HPLC traces for classification of pharmaceutical fingerprinting, etc.

Worthy of mention is that although most single-molecule stoichiometry studies based on photobleaching step counting are applied for proteins with low subunit number, it is advantageous to develop methods enabling the quantitation of protein oligomers or complex at high-order states.^{52,53,54} According to our evaluation of CLDNN in the analysis of more steps with synthesized data sets, the accuracy decreased to be lower than 75% when the step is larger than 20 (Supporting Information Figure S9). In this case, more layers might be needed for deeper learning of CLDNN to improve accuracy.

We will further improve this method in future, with the ability to analyze data with more bleaching steps and provide kinetic properties such as the dwell time of each state. It is believed that CLDNN will be widely applied to the study of protein stoichiometry and analysis of stepwise events.

4. MATERIALS AND METHODS

4.1. Plasmid Construction. The DNA fragments encoding full-length EGFR were subcloned into the *Hind*III and *Bam*HI sites of pEGFP-N1 (Clontech), yielding the EGFR-EGFP expression plasmids. The plasmids were confirmed by DNA sequencing.

4.2. Cell Culture and Transfection. Hela cells were maintained in DMEM, supplemented with 10% fetal bovine serum (Gibco) and 100 U/ml penicillin–streptomycin (Invitrogen) at 37 °C with 5% CO₂. On reaching ~80% confluency, cells were split and seeded onto 35 mm glass-bottom dishes (Shengyou Biotechnology) which were coated with enhanced cell adhesion reagent (ECAK-3) (Beijing xigong biological technology) to a density of 5×10^4 cells/dish. After 24 h, cells were transfected with 0.5 μ g plasmids per dish using Lipofectamine 3000 (Invitrogen). To achieve a low expressing level of proteins, cells were incubated with plasmids for only 4.5 h. For EGF stimulation group, EGF was added into the medium to a final concentration of 200 pM and incubated for 10 min at 37 °C. Then the medium was removed, and the transfected cells were washed with cold PBS (4 °C) twice and fixed with 3% (wt/vol) paraformaldehyde and 0.1% (wt/vol) glutaraldehyde/PBS solution for 20 min before fluorescence imaging.

4.3. Preparation of Dye-Labeled ssDNA. The ssDNA, 5'-AGCATTAAT-biotin-CACCGTGGCCATTA-Cy5-3' (the biotin was modified on the ninth T base), was synthesized by Sangon Biotech (Shanghai), according to the previous report.⁵⁵

4.4. Single-Molecule Fluorescence Imaging. Single-molecule fluorescence imaging was performed using a custom-made total internal reflection fluorescence (TIRF) microscopy according to previous reports.^{5,7,42} An Olympus IX71 microscope equipped with a total internal reflective fluorescence illuminator, a $\times 100/1.45$ NA Plan ApoChromatic TIR objective, and an electron-multiplying charge-coupled device (EMCCD) camera (Andor iXon DU-897D BV) was used. EGFP was excited at 488 nm and Cy5 was excited at 647 nm with the power of 1 mW (17 W/cm²) measured after the laser passing through the objective. The collected fluorescent signals were passed

through the suitable filters (HQ 525/50 from Chroma Technology for EGFP and BL 676/29 from Semrock for Cy5) before being directed to the EMCCD camera. The gain of the EMCCD camera was set at 300. Movies of 400 frames were acquired for each sample at a frame rate of 10 Hz.

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b00688.

Supporting information about CLDNN and supporting methods; Figures S1–S9 and Tables S1–S2 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*xfang@iccas.ac.cn

*jhyuan@iccas.ac.cn

ORCID

Rong Zhao: 0000-0003-2169-3953

Jinghe Yuan: 0000-0001-6456-7696

Xiaohong Fang: 0000-0002-2018-0542

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (Nos. 21735006, 21890742, 21127901) and Chinese Academy of Sciences.

■ REFERENCES

- (1) Atanasova, M.; Whitty, A. Understanding cytokine and growth factor receptor activation mechanisms. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47* (6), 502–30.
- (2) Ulbrich, M. H.; Isacoff, E. Y. Subunit counting in membrane-bound proteins. *Nat. Methods* **2007**, *4* (4), 319–321.
- (3) Aurousseau, M. R. P.; McGuire, H.; Blunck, R.; Bowie, D. A Step-by-Step Guide to Single-Subunit Counting of Membrane-Bound Proteins in Mammalian Cells. In *Ionotropic Glutamate Receptor Technologies*; Popescu, G. K., Ed.; Springer New York: New York, 2016; pp 15–30.
- (4) Tsekouras, K.; Custer, T. C.; Jashnsaz, H.; Walter, N. G.; Presse, S. A novel method to accurately locate and count large numbers of steps by photobleaching. *Mol. Biol. Cell* **2016**, *27* (22), 3601–3615.
- (5) Zhang, W.; Jiang, Y.; Wang, Q.; Ma, X.; Xiao, Z.; Zuo, W.; Fang, X.; Chen, G. Single-molecule imaging reveals transforming growth factor-beta-induced type II receptor dimerization. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (37), 15679–15683.
- (6) Sun, Y.; Li, N.; Zhang, M.; Zhou, W.; Yuan, J.; Zhao, R.; Wu, J.; Li, Z.; Zhang, Y.; Fang, X. Single-molecule imaging reveals the stoichiometry change of beta(2)-adrenergic receptors by a pharmacological biased ligand. *Chem. Commun.* **2016**, *52* (44), 7086–7089.
- (7) Zhang, M.; He, K.; Wu, J.; Li, N.; Yuan, J.; Zhou, W.; Ye, Z.; Li, Z.; Xiao, H.; Lv, Z.; Zhang, Y.; Fang, X. Single-molecule imaging reveals the stoichiometry change of epidermal growth factor receptor during transactivation by beta(2)-adrenergic receptor. *Sci. China: Chem.* **2017**, *60* (10), 1310–1317.
- (8) Shu, D.; Zhang, H.; Jin, J.; Guo, P. Counting of six pRNAs of phi29 DNA-packaging motor with customized single-molecule dual-view system. *EMBO J.* **2007**, *26* (2), 527–537.
- (9) Reuel, N. F.; Bojo, P.; Zhang, J.; Boghossian, A. A.; Ahn, J. H.; Kim, J. H.; Strano, M. S. NoRSE: noise reduction and state evaluator for high-frequency single event traces. *Bioinformatics* **2012**, *28* (2), 296–7.
- (10) Aggarwal, T.; Materassi, D.; Davison, R.; Hays, T.; Salapaka, M. Detection of Steps in Single Molecule Data. *Cell. Mol. Bioeng.* **2012**, *5* (1), 14–31.
- (11) Stefani, F. D.; Hoogenboom, J. P.; Barkai, E. Beyond quantum jumps: Blinking nanoscale light emitters. *Phys. Today* **2009**, *62* (2), 34–39.
- (12) Chung, S. H.; Kennedy, R. A. Forward-backward non-linear filtering technique for extracting small biological signals from noise. *J. Neurosci. Methods* **1991**, *40* (1), 71.
- (13) Knight, A. E.; Molloy, J. E. Coupling ATP hydrolysis to mechanical work. *Nat. Cell Biol.* **1999**, *1* (4), E87–E89.
- (14) Ji, W.; Xu, P.; Li, Z.; Lu, J.; Liu, L.; Zhan, Y.; Chen, Y.; Hille, B.; Xu, T.; Chen, L. Functional Stoichiometry Of The Unitary Calcium-release-activated Calcium Channel Revealed By Single-molecule Imaging. *Biophys. J.* **2009**, *96* (3), 560a–560a.
- (15) Zhang, W.; Yuan, J.; Yang, Y.; Xu, L.; Wang, Q.; Zuo, W.; Fang, X.; Chen, Y. Monomeric type I and type III transforming growth factor-beta receptors and their dimerization revealed by single-molecule imaging. *Cell Res.* **2010**, *20* (11), 1216–1223.
- (16) Demuro, A.; Penna, A.; Safrina, O.; Yeromin, A. V.; Amcheslavsky, A.; Cahalan, M. D.; Parker, I. Subunit stoichiometry of human Orai1 and Orai3 channels in closed and open states. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (43), 17832–17837.
- (17) Nakajo, K.; Ulbrich, M. H.; Kubo, Y.; Isacoff, E. Y. Stoichiometry of the KCNQ1-KCNE1 ion channel complex. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (44), 18862–18867.
- (18) Wang, Y. Jump and Sharp Cusp Detection by Wavelets. *Biometrika* **1995**, *82* (2), 385–397.
- (19) Sadler, B. M.; Swami, A. Analysis of multiscale products for step detection and estimation. *IEEE Trans. Inf. Theory* **1999**, *45* (3), 1043–1051.
- (20) Carter, N. J.; Cross, R. A. Mechanics of the kinesin step. *Nature* **2005**, *435* (7040), 308–12.
- (21) Kersemakers, J. W.; Munteanu, E. L.; Laan, L.; Noetzel, T. L.; Janson, M. E.; Dogterom, M. Assembly dynamics of microtubules at molecular resolution. *Nature* **2006**, *442* (7103), 709.
- (22) Andrec, M.; Levy, R. M.; Talaga, D. S. Direct determination of kinetic rates from single-molecule photon arrival trajectories using hidden Markov models. *J. Phys. Chem. A* **2003**, *107* (38), 7454–7464.
- (23) Messina, T. C.; Kim, H. Y.; Giurleo, J. T.; Talaga, D. S. Hidden Markov model analysis of multichromophore photobleaching. *J. Phys. Chem. B* **2006**, *110* (33), 16366–16376.
- (24) Gagnieu, P. A. *Markov Chains: From Theory to Implementation and Experimentation*; John Wiley & Sons: New York, 2017.
- (25) Tavakoli, M.; Taylor, J. N.; Li, C. B.; Komatsuzaki, T.; Presse, S. Single Molecule Data Analysis: An Introduction. *Adv. Chem. Phys.* **2017**, *162*, 205–305.
- (26) Beal, M. J.; Ghahramani, Z.; Rasmussen, C. E. The infinite hidden Markov model. *Adv. Neur. In.* **2002**, *14*, 577–584.
- (27) Deng, L.; Yu, D. Deep Learning: Methods and Applications. *Foundations & Trends in Signal Processing* **2013**, *7* (3), 197–387.
- (28) Ordonez, F. J.; Roggen, D. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors* **2016**, *16* (1), 115.
- (29) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436.
- (30) Esteva, A.; Kuprel, B.; Novoa, R. A.; Ko, J.; Swetter, S. M.; Blau, H. M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542* (7639), 115.
- (31) Goh, G. B.; Hodas, N. O.; Vishnu, A. Deep learning for computational chemistry. *J. Comput. Chem.* **2017**, *38* (16), 1291–1307.
- (32) Ryan, K.; Lengyel, J.; Shatruk, M. Crystal Structure Prediction via Deep Learning. *J. Am. Chem. Soc.* **2018**, *140* (32), 10158–10168.
- (33) Fernandez, M.; Ban, F. Q.; Woo, G.; Hsing, M.; Yamazaki, T.; LeBlanc, E.; Rennie, P. S.; Welch, W. J.; Cherkasov, A. Toxic Colors: The Use of Deep Learning for Predicting Toxicity of Compounds Merely from Their Graphic Images. *J. Chem. Inf. Model.* **2018**, *58* (8), 1533–1543.

- (34) Ma, C.; Ren, Y.; Yang, J.; Ren, Z.; Yang, H.; Liu, S. Improved Peptide Retention Time Prediction in Liquid Chromatography through Deep Learning. *Anal. Chem.* **2018**, *90* (18), 10881–10888.
- (35) Kalchbrenner, N.; Grefenstette, E.; Blunsom, P. A Convolutional Neural Network for Modelling Sentences. *Eprint Arxiv* **2014**, *1*, 655.
- (36) Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput* **1997**, *9* (8), 1735–1780.
- (37) Gers, F. A.; Schmidhuber, J.; Cummins, F. Learning to forget: Continual prediction with LSTM. *Neural Comput* **2000**, *12* (10), 2451–2471.
- (38) Bergstra, J.; Breuleux, O.; Bastien, F.; Lamblin, P.; Pascanu, R.; Desjardins, G.; Turian, J.; Warde-Farley, D.; Bengio, Y. In *Theano: a CPU and GPU math expression compiler* **2010**, *4*.
- (39) Bastien, F. d. r.; Lamblin, P.; Pascanu, R.; Bergstra, J.; Goodfellow, I.; Bergeron, A.; Bouchard, N.; Warde-Farley, D.; Bengio, Y., Theano: New features and speed improvements. *arXiv preprint arXiv:1211.5590*, **2012**.
- (40) Graves, A. *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer Berlin Heidelberg: Berlin, Heidelberg, Germany, **2012**.
- (41) Chollet, F. *Keras: The python deep learning library*. Astrophysics Source Code Library, **2018**.
- (42) Yuan, J.; He, K.; Cheng, M.; Yu, J.; Fang, X. Analysis of the Steps in Single-Molecule Photobleaching Traces by Using the Hidden Markov Model and Maximum-Likelihood Clustering. *Chem. - Asian J.* **2014**, *9* (8), 2303–2308.
- (43) Li, P.; Miao, Y.; Dani, A.; Vig, M. alpha-SNAP regulates dynamic, on-site assembly and calcium selectivity of Orail channels. *Mol. Biol. Cell* **2016**, *27* (16), 2542–2553.
- (44) Das, S. K.; Darshi, M.; Cheley, S.; Wallace, M. I.; Bayley, H. Membrane protein stoichiometry determined from the step-wise photobleaching of dye-labelled subunits. *ChemBioChem* **2007**, *8* (9), 994–999.
- (45) Coffman, V. C.; Wu, J. Q. Counting protein molecules using quantitative fluorescence microscopy. *Trends Biochem. Sci.* **2012**, *37* (11), 499–506.
- (46) Kingma, D. P.; Ba, J., Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*, **2014**.
- (47) Prechelt, L. Automatic early stopping using cross validation: quantifying the criteria. *Neural Networks* **1998**, *11* (4), 761–767.
- (48) McGuire, H.; Aurousseau, M. R. P.; Bowie, D.; Blunck, R. Automating Single Subunit Counting of Membrane Proteins in Mammalian Cells. *J. Biol. Chem.* **2012**, *287* (43), 35912–35921.
- (49) van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *J. Mach Learn Res.* **2008**, *9*, 2579–2605.
- (50) Jain, A.; Liu, R. J.; Ramani, B.; Arauz, E.; Ishitsuka, Y.; Ragunathan, K.; Park, J.; Chen, J.; Xiang, Y. K.; Ha, T. Probing cellular protein complexes using single-molecule pull-down. *Nature* **2011**, *473* (7348), 484–U322.
- (51) Cochet, C.; Kashles, O.; Chambaz, E. M.; Borrello, I.; King, C. R.; Schlessinger, J. Demonstration of Epidermal Growth Factor-Induced Receptor Dimerization in Living Cells Using a Chemical Covalent Cross-Linking Agent. *J. Biol. Chem.* **1988**, *263* (7), 3290–3295.
- (52) Lenn, T.; Gkekas, C. N.; Bernard, L.; Engl, C.; Jovanovic, G.; Buck, M.; Ying, L. Y. Measuring the stoichiometry of functional PspA complexes in living bacterial cells by single molecule photobleaching. *Chem. Commun.* **2011**, *47* (1), 400–402.
- (53) Engl, C.; Jovanovic, G.; Lloyd, L. J.; Murray, H.; Spitaler, M.; Ying, L. M.; Errington, J.; Buck, M. In vivo localizations of membrane stress controllers PspA and PspG in Escherichia coli. *Mol. Microbiol.* **2009**, *73* (3), 382–396.
- (54) Li, N.; Zhao, R.; Sun, Y.; Ye, Z.; He, K.; Fang, X. Single-molecule imaging and tracking of molecular dynamics in living cells. *National Science Review* **2017**, *4*, 739–760.
- (55) Hannestad, J. K.; Sandin, P.; Albinsson, B. Self-Assembled DNA Photonic Wire for Long-Range Energy Transfer. *J. Am. Chem. Soc.* **2008**, *130* (47), 15889–15895.