

Implementation of Neural Networks for the Identification of Single Molecules

Benjamin P. Bowen,^{*,†} Allan Scruggs,[‡] Jörg Enderlein,[§] Markus Sauer,^{||} and Neal Woodbury[‡]

Department of Chemical and Materials Engineering, Arizona State University, Tempe, Arizona 85287, Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287, Institute for Biological Information Processing, Forschungszentrum Jülich, D-52425 Jülich, Germany, and Applied Laserphysics and Laserspectroscopy, University of Bielefeld, 33615 Bielefeld, Germany

Received: August 18, 2003; In Final Form: March 3, 2004

The effectiveness of neural networks and the optimization of parameters for implementing neural networks were evaluated for use in the identification of single molecules according to their fluorescence lifetime. The best network architecture and training parameters were determined for both ideal and nonideal single-molecule fluorescence data. The effectiveness of the neural network is compared to that of the maximum likelihood estimator on the basis of its ability to correctly identify single molecules. For ideal single-molecule data, it was found that the neural networks and the maximum likelihood estimator perform approximately equally well. For nonideal single-molecule fluorescence data, neural networks were able to correctly identify a larger percentage of single-molecule events than the MLE method.

Introduction

In recent years, neural network algorithms have received more and more use in pattern recognition applications.^{1,2} Some examples of such applications include predicting protein folds from sequence,³ weather pattern prediction,⁴ instrument and process monitoring⁵ and financial prediction.⁶ Advantages of neural networks in these applications include the ability of the algorithm to adapt as new data are presented, and their simplicity relative to the systems they are predicting. Neural networks utilize simple parallel elements to perform complex analysis and classification of data. Because of this feature, neural networks can be of particular use when an explicit mathematical model of behavior is not available.⁷

Pattern matching algorithms have been widely applied to the identification of single molecules on the basis of their detected fluorescence.^{8–10,11} These algorithms can include many different parameters measurable by single-molecule fluorescence spectroscopy including spectral information, polarization, fluorescence lifetime, and fluorescence intensity. Because of the inherently small signal detected from a single molecule, the ability to identify molecules is limited. In many current single-molecule fluorescence measurements, time-correlated single-photon counting (TCSPC) gives the ability to measure the excited state lifetime of individual molecules. To fully maximize the information content of a TCSPC single-molecule fluorescence data set, statistical estimator techniques have been shown to give the best results when dealing with the “stochastic” nature of single-molecule fluorescence data. In particular, for identifying molecules based on their fluorescence lifetime, pattern-matching using the maximum likelihood estimator (MLE) technique has been shown to give the theoretically best possible identification for typical fluorophores.^{12–15}

Though MLE performs at the theoretical optimum for distinguishing between simple (ideal) fluorophores, it may fail in cases where a single fluorophore can exist in either one of two possible states with widely different lifetimes, as it uses the average decay profile of a fluorophore (typically from bulk measurements) as a means of comparison. In this work, the use of neural networks in classifying fluorescence from individual molecules in a mixture is explored. Because a neural network should be able to extract the salient features (including population heterogeneity) of the data provided, and ignore nondiscriminating information, it could provide a significant advantage over the MLE technique in some cases. This type of approach was previously used by Sun et al. to identify individual bursts in a single-molecule data set that had been generated by a Monte Carlo algorithm. It was demonstrated that neural networks could empirically apply the distinguishing features of a data set including fluorescence lifetime, absorption cross-section, fluorescence quantum efficiency, and photodestruction efficiency for identification without having to develop a statistical theory of the features used for identification.¹⁶

It is often the case in measurements of biological systems that no physical model exists to describe the behavior. One advantage of neural networks is that they analyze complex systems without using a specific model and have shown robust accuracy in identification. Though the use of physical models in the analysis of data can sometimes lend insight into the underlying mechanism, if used inappropriately, such models can distort interpretation. To further explore the utility of neural networks in identifying nonideal fluorescent molecules, simulated data mimicking a system in which two fluorophores are mixed with one fluorophore existing in two states with distinct lifetimes is analyzed. The average fluorescence lifetime of the two fluorophores is the same. In this case, a neural network should be able to identify the two fluorophores using a form of pattern recognition less constrained than MLE.

Experimental Section

Single-Molecule Fluorescence Data. Single-molecule fluorescence data from the three rhodamine derivatives JF9,¹⁷

* Corresponding author. E-mail: bpb@asu.edu. Phone: 480-965-8162. Fax: 480-965-2747.

[†] Department of Chemical and Materials Engineering, Arizona State University.

[‡] Department of Chemistry and Biochemistry, Arizona State University.

[§] Forschungszentrum Jülich.

^{||} University of Bielefeld.

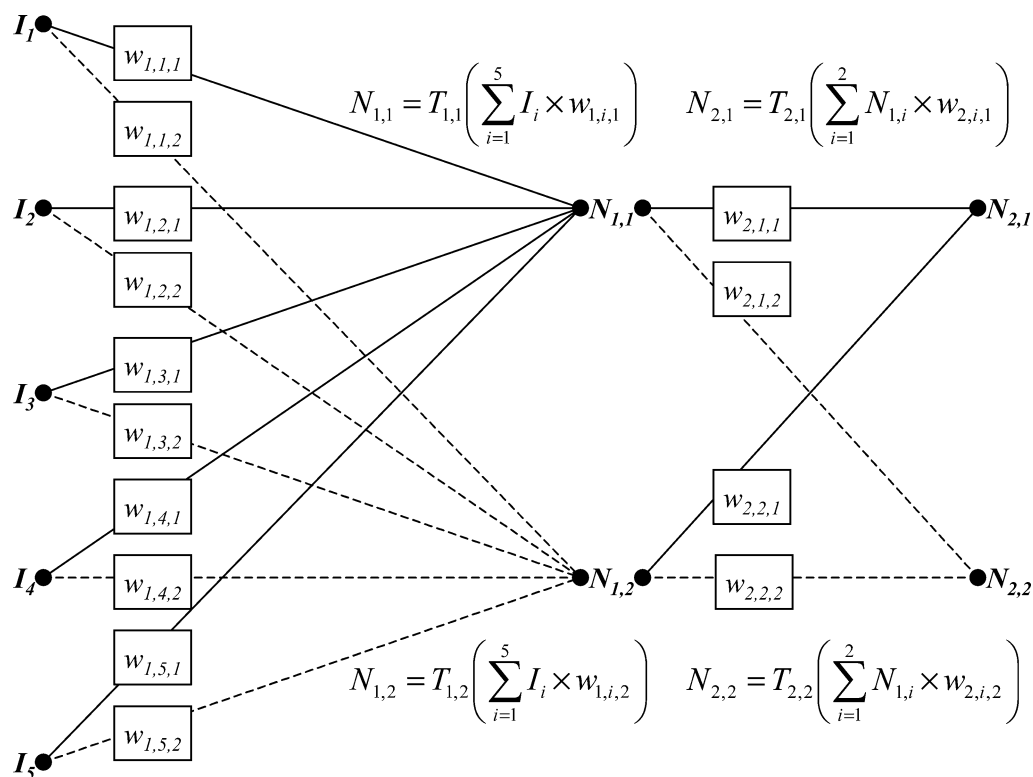


Figure 1. Two-layer network with two nodes (N) in each layer and five inputs (I). The transfer function (T) of each node operates on the sum of the weighted inputs that connect to each node. The output of the network is given by $N_{2,1}$ and $N_{2,2}$.

JA167,¹⁸ and JA53¹⁹ used in this work were taken from a previously published data set containing time-correlated single-photon (TCSPC) decay histograms of single-molecule events.¹² These dyes were used because they represent the ideal case of a group of fluorophores with simple, single-exponential decay times. The structure of each TCSPC histogram consists of sixty-four 200 ps wide time bins. The amplitude in each bin of the histogram is normalized by the total number of photons in the single-molecule event. The data for each dye were collected separately. The analysis below involves the identification of single molecules in mixtures of two dyes. The data sets for each dye pair of fluorophores were constructed by taking events at random from individual fluorophores and combining them in the computer. In this work, it is assumed that the single-molecule data are collected from a sample having a low enough concentration so that very few of the single-molecule events result from overlapping of two or more events. On the basis of this assumption, data sets created for dye-pair analysis did not include overlapping events.

The performance of single-molecule identification algorithms using nonideal fluorophores was also considered by generating hypothetical data sets containing two fluorophores as above, but in this case one fluorophore had two distinct conformations each with a different lifetime. The nonideal single-molecule data were generated by creating 6000 events for each hypothetical fluorophore. The number of photons in each theoretical single-molecule event was calculated by taking 20 photons plus a random number determined according to an exponential distribution function. The exponential distribution function used had a decay constant of 10 photons (Details of this methodology will be given later in Figure 5b). The arrival time, relative to the excitation pulse, for each photon was also determined using an exponential distribution function having a decay time corresponding to the hypothetical excited-state lifetime. For each molecule, arrival times were then histogrammed as discussed above into discrete time bins. The amplitude of each bin in the

histogram was normalized by the total number of photons in the histogram. Single-molecule events were generated having excited-state decay times of 1, 2, and 3 ns.

In this work, analysis using a maximum likelihood estimator (MLE) approach was compared to analysis using a neural network algorithm. The MLE methodology was the same as that described for the original analysis of these data.¹² The neural network approach is described in detail below. For both analysis techniques, the data were divided into a training set and a test set. The training set was used to generate the appropriate pattern base for single-molecule identification. The test set was the data used to evaluate the accuracy of single-molecule identification.

Construction and Training of Neural Networks for Single-Molecule Identification. A neural network is made up of interconnecting nodes or “neurons”. In this case, the input vector consists of the sixty-four bin values in the decay histogram of a single-molecule event described above. In Figure 1, an input vector with only five elements is used for clarity (I_1 – I_5 in Figure 1). In a standard “feed-forward” implementation, the input vector is multiplied by a weight matrix ($w_{1,j,k}$) to find the inputs to the first layer of the network ($N_{1,1}$ and $N_{1,2}$). Each node in this layer generates an output, based on what is known as its transfer function. The transfer function maps the input value of each node to an output value and is usually some nonlinear function bounded between 0 and +1 or –1 and +1. These output values are then weighted by a second weight matrix ($w_{2,j,k}$). This process continues for a set number of intermediate, or hidden layers. A final output is obtained from the output layer of nodes. In Figure 1, the output layer consists of $N_{2,1}$ and $N_{2,2}$. The values of the weight matrices are determined during the training period as described below.

In constructing the network, the number of layers and the number of nodes in each layer is defined, and weights are initially randomized. The network is presented with a set of input vectors whose desired output from the network is known, also called the training set. Deviations from this desired output

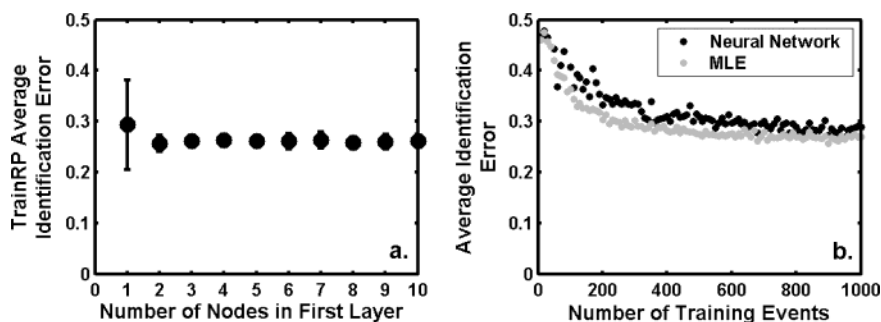


Figure 2. Average error in identifying events from JA53 and JA167 as a function of the number of nodes in the first layer (a) and number of training events (b). TrainRP was used as the training function for neural network analysis. In panel b, for comparison, the identification error when using MLE is also shown. The error bars show the standard error.

are accumulated as error, and after all inputs are presented, the overall error is propagated backward through the network and weights are changed according to a previously defined back-propagation algorithm. The algorithm changes the weights connecting the nodes to minimize the error seen in the training set using any of a number of minimization techniques. A second set of known inputs, called a validation set, can be used to then check the error of the network as well. Validation sets are used to avoid the network simply memorizing the training set and can improve training and test performance.

Training continues through multiple iterations until one of several possible conditions is met. In this work, training was stopped either when the error reached a certain lower boundary, when the error was no longer changing between iterations, or when the error on the validation set began to increase. The network architectures tested consisted of two layers (a first, hidden, layer receiving the weighted inputs, and a final layer producing the output). All networks tested were feed forward, and trained with a back-propagation method. Validation sets were used in all cases. The selection of events for the validation set, training set, and test set is performed at random. Events that are used for one set are not used for another set. The training set size was varied between 5 and 5000 single-molecule events. The number of nodes in the first layer was varied between 1 and 10, and three different training algorithms were tested with the largest training set size to determine which network architecture and training algorithm was most appropriate for this application. The transfer function used in every case was log-sigmoidal, though a sigmoidal transfer function and several other transfer functions were also explored with similar results. The output layer was structured to have two outputs that could each vary between 0 and 1. Output vectors for the training and validation set were defined to be [1,0] for the first dye and [0,1] for the second dye. The training algorithms used were a resilient back-propagation (TrainRP), a scaled conjugate gradient algorithm (TrainSCG), and the Levenberg–Marquardt algorithm (TrainLM). These training algorithms are part of the Matlab Neural Networks Toolbox. All neural network analysis was done using the Matlab programming platform.

Results

Application of Neural Networks to Single-Molecule Data.

Data from the JA167 and JA53 dye pair were used to determine the best configuration and training parameters of a neural network that identifies single molecules by their excited state decay characteristics. The excited state lifetimes of JA167 and JA53 are 2.0 and 2.4 ns, respectively, and in this data set, bursts having more than 20 photons were used. The bursts were selected from the data as described previously.¹² This data set

was previously analyzed using the maximum likelihood estimator technique in a study that showed that the MLE approach was able to achieve the best possible accuracy for identifying single molecules in this data set.¹² Here the same data set is analyzed using a neural network approach. The performance of either identification algorithm was measured in terms of the average identification error. The identification error is defined here as the fraction of the molecules of a particular fluorophore that were misidentified. The average identification error for a dye pair is the average of the identification errors for the individual fluorophores in the dye pair.

Shown in Figure 2a is the dependence of the number of nodes in the first layer of the network on the average identification error using TrainRP as a training function (see Methods for description of the network architecture used and definitions of training functions). The networks in this plot were trained with 1000 training cases and 1000 validation cases. The networks used in this study were two-layer networks. The use of three or more layered networks did not result in any improvement in performance. For each data point shown, 10 networks were trained and the performance of each was tested on 1000 test cases that were not part of the training data. The error bars show the standard deviation in the identification error for the network trained with TrainRP. Increasing the number of nodes used past 2 resulted in little or no improvement in either the average performance or the standard deviation. Other training functions were also tested including TrainLM and TrainSCG, but use of these functions resulted in a larger average identification error and a greater variability in the performance of networks trained under the same conditions (not shown). Because early stopping is used to prevent overtraining, only those training algorithms capable of early stopping were tested. The use of early-stopping resulted in shorter training times and a slightly improved identification accuracy. Without early stopping, the neural network can be trained to memorize the features of certain events instead of generalizing their salient features. The number of elements in the training set affected the performance of the network, as shown in Figure 2b. Larger training sets resulted in lower error in single-molecule identification. In Figure 2b, the dependence of the average identification error using the MLE technique is shown for comparison. The number of training events refers to the number of events used to either generate the patterns required for the MLE method or the number of training events used for the neural network based method.

Three dye pairs were chosen to validate the effectiveness of neural networks for single-molecule identification. On the basis of the initial tests described above, a two-layer network with two nodes in the first layer trained using the TrainRP training

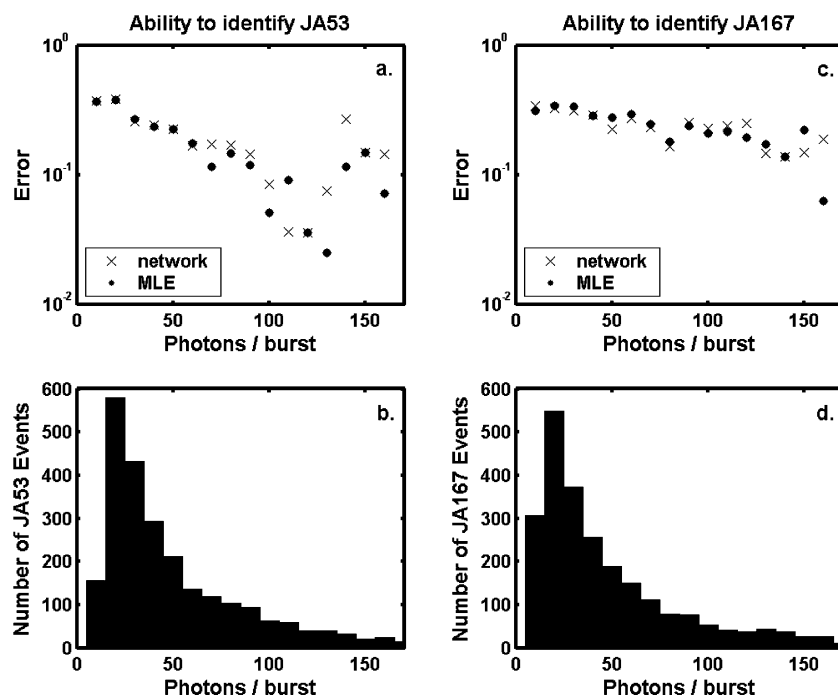


Figure 3. Identification error for individual fluorophores as a function of burst size for both neural network analysis and MLE analysis of a mixture of JA53 (a) and JA167 (c). Burst size distribution for these fluorophores in the same sample (b and d). Details are given in the text.

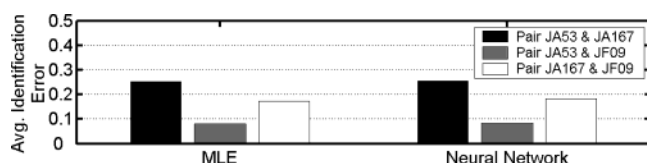


Figure 4. Average identification error for three dye-pairs using MLE and a trained neural network. For each dye tested, there were 1000 training events and 1000 validation events used to train the network. The trained network was evaluated using 2500 test events that were not among those used for training.

function was employed. In each case for each fluorophore, 1000 events were used as training cases and 1000 events were used as validation cases. A total of 2500 events were used to test the trained networks. In Figure 3, the identification error and number of single-molecule bursts are shown as a function of the number of photons per burst for fluorophores JA53 (Figure 3a,b) and JA167 (Figure 3c,d). In Figures 3a,c, the ability to identify single-molecule events as a function of the number of photons in a burst is plotted both using the MLE technique and using the trained neural network. As can be seen, the performance of the two methods is nearly identical. Because MLE has been shown to identify individual molecules using this data set with the theoretically best possible accuracy,¹² these results imply that properly trained neural networks can also perform optimally.

In Figure 4, the comparison between single-molecule fluorophore identification using the neural network vs the MLE approach is extended to two additional fluorophore pairs (JA53/JA167 is also shown for comparison). What is plotted is the average identification error for each dye pair. In all cases, the performance of the network is nearly identical to that of the MLE method.

Application of Neural Networks to Nonideal Single-Molecule Data Analysis. To explore the application of neural networks to the analysis of more complicated single-molecule data sets, single-molecule events were computer generated according to a scenario where one of the two fluorescing species (A) to be identified has multiple conformations with distinct fluorescence lifetimes. The other species to be identified (B)

has a fixed fluorescence lifetime and can be measured separately. An example of species A would be a DNA molecule with a fluorophore and a fluorescence quencher bound at different positions such that the efficiency of fluorescence quenching is a function of the DNA conformation. A subpopulation of the DNA molecules bound to a protein may change the conformation of the DNA molecule, and thus the degree of quenching. Species B could represent a fluorescent contaminant as one often finds in biological preparations.

The model data set includes an equal number of events from species A and species B. Species A has an excited-state lifetime of either 1 or 3 ns (depending on the conformation, see below) and species B has a lifetime of 2 ns. A distribution of fluorescence lifetimes histogrammed for these hypothetical molecules is shown in Figure 5a. For the case shown in this figure, species A consists of 35% molecules with an excited-state decay time of 1 ns and 65% molecules with an excited-state decay time of 3 ns. Though many other compositions of species A were used in the evaluation of the neural network approach, this ratio (35%:65%) is shown because it represented one of the worst cases for MLE analysis (see below). The burst size distributions for both species are shown in Figure 5b.

For the analysis of the nonideal data set described here, a two-layer neural network was used. The number of nodes in the output layer was held to 2 (because there were 2 possible outcomes of the identification). The number of nodes required in the first layer to achieve an optimal performance by the neural network was substantially greater than in the simpler (ideal) case described above. As shown in Figure 6, five nodes in the first layer were required to obtain the best performance. For each number of nodes tested in the first layer, a series of networks were trained and evaluated in which the percentage of species A having a 1 ns decay time varied from 0 to 100%. The standard deviation shown is based on the results from six networks trained under identical conditions. The ability to identify species A or B depends strongly on the fraction of species A in the 1 ns vs 3 ns conformation. A new network was trained using each of a series of different relative amounts

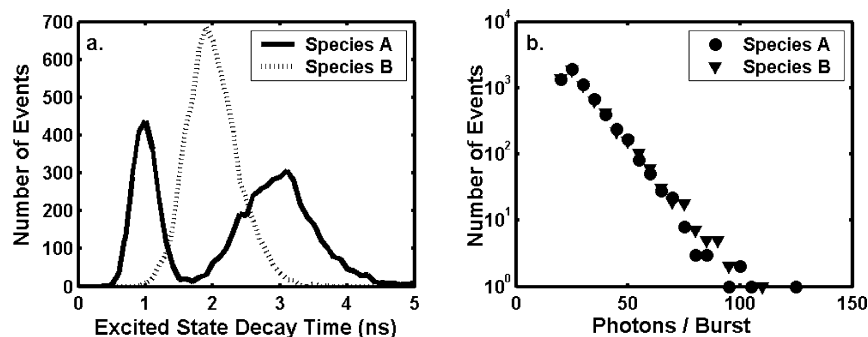


Figure 5. (a) Distribution of the fluorescence lifetimes measured for the simulated single-molecule events where species A consists of 35% molecules having a 1 ns decay time and 65% molecules having a 3 ns decay time. (b) Burst size distribution of the simulated single-molecule events for species A and species B.

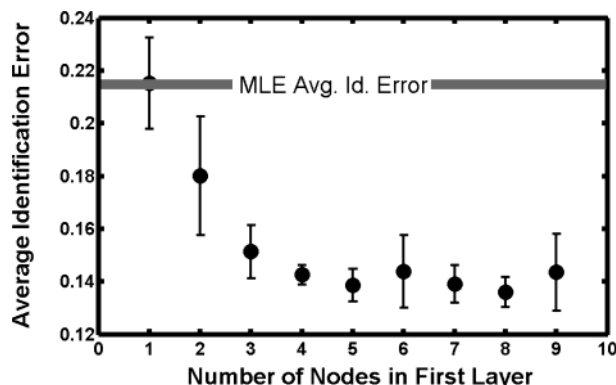


Figure 6. Average identification error over all compositions of species A tested (see text) as a function of number of nodes in the first layer. Neural network performance improves in this case as the number of nodes in the first layer is increased up to about 5 nodes. The error bars show the standard error.

of the 1 and 3 ns subspecies of species A and a range of different numbers of nodes in the first layer. In each case, the average identification error was measured by evaluating a randomly chosen mixture of 1000 test cases each from species A and species B. The average identification error shown in Figure 6 was determined by averaging the results across all the tested compositions (relative amounts of the 1 and 3 ns subspecies) of species A. This type of averaging was done to ensure that the network architecture gave the best performance over a variety of possible test compositions. For example, using a neural network with five nodes in the first layer, the average identification error ranges from about 0.05 to about 0.2 as a function of species A composition (this case is shown in Figure 7 as described below). The average of all the identification errors shown in Figure 7 corresponds to the point in Figure 6 where 5 nodes are in the first layer.

Network Performance. Networks having 5 nodes in the first layer were trained for a series of different relative amounts of 1 and 3 ns conformations of species A, and the average identification errors are shown in Figure 7 for each of these cases. These networks were trained with 1000 test events, 1000 training events, and 1000 validation events from each species for each case. MLE was also used to distinguish species A from B. Here the fluorescence pattern used in the MLE analysis was derived from the total fluorescence time course of each species, as described previously.¹² As expected, at the end points of the graph, where species A has only a single lifetime, the neural network and the MLE approach yielded the same results. These two cases are equivalent to distinguishing between two dyes each with a single lifetime as in Figures 3 and 4. In contrast, for cases in which there was a significant mixture of the two

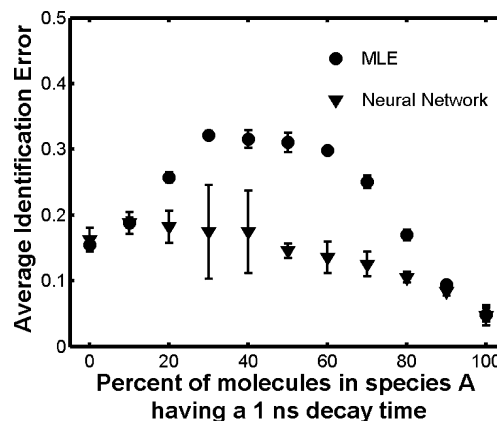


Figure 7. Average identification error using both MLE and neural network analysis as a function of the composition of species A (percent of the species A molecules having a 1 ns lifetime out of a mixture of molecules with 1 and 3 ns lifetimes). When there is a nearly equal mixture of the 1 and 3 ns subspecies in species A, the average identification error using MLE is much greater than it is using a neural network. The error bars show the standard error.

conformations of species A, the performance of the neural network was superior to that obtained with MLE. The accuracy of the MLE approach could be enhanced by applying a two population physical model to data analysis, dividing species A into two subspecies with different lifetimes, but what was tested here was the accuracy of identification algorithms that do not depend on a physical model (the neural network analysis never assumes a specific number of subspecies or a specific composition of subspecies). Primarily in the analysis of natural samples, there are cases where no reliable model can be imposed.

Discussion

Application of Neural Networks for Ideal Cases. For the identification of single-molecule events produced by ideal fluorophores, neural networks and MLE operate in essentially the same way. In both techniques, pattern recognition is used to determine if a given event has a lifetime more similar to one of the two molecules. The final result of both approaches is a number whose magnitude indicates how well the input decay curve resembles the lifetime of one fluorophore or the other. For the neural network, the pattern memory is contained in the weight matrices which transform the input histogram into a final output. For the MLE approach, fluorophore identification is performed by statistical comparison of each input histogram to the histogram of the bulk decay kinetics from the individual fluorophores. Neural networks having two nodes in the first layer function in approximately the same way as MLE in that in both methods only two transformation functions are applied to each

input data point. In the case of MLE, the two transformation functions are the probability evaluations based on the two independent fluorophore decay histograms. In the case of the neural network, these are the two input weight vectors applied to the input data. Consequently, the fact that neural network and MLE based approaches performed with equal accuracy is not surprising given their similarities.

This can be seen in Figures 3 and 4 where the ability of neural networks to identify single molecules is essentially equivalent to that of the MLE technique. For nearly ideal data such as this, MLE would be the method of choice for identifying molecules because of the ease with which it is implemented. However, for more complex samples common to biological systems, heterogeneous fluorescence characteristics are the norm. Consequently, because neural networks perform as well as MLE under ideal conditions, and better with heterogeneous fluorophores, they could be a more attractive option for single-molecule identification in certain cases.

Application of Neural Networks to Nonideal Cases. By increasing the number of nodes in the neural network, one enhances the ability of the network to extract salient features from nonideal data. As can be seen in Figure 6, for the nonideal cases described, the average identification error decreases as the number of nodes is increased. Even with only one node in the first layer, the network performs as well as MLE (Figure 6). As more nodes are added, the network can identify species A with greater accuracy because it can recognize the individual subspecies in species A. As a specific example, Rhodamine-6G covalently attached to DNA has been seen to have multiple distinctly different excited-state lifetimes.²⁰ This is because of the various interactions the fluorophore can have with the DNA bases. For identification of this molecule in a background of other fluorophores, the description of the rhodamine/DNA complex generated by a neural network would include all of the different excited-state lifetimes as possible, distinct states of the system, rather than using the ensemble averaged characteristics of the complex to identify it.

Figure 7 demonstrates the ability of neural networks to take into account multiple distinct fluorescent states of a molecule in performing identification. The identification error varies nearly monotonically between the two extremes of the composition of species A in the case of the neural network. In contrast, MLE identification is distinctly worse for mixed compositions of subspecies in species A than for extreme cases in which only the 1 ns or 3 ns subspecies were present. Regardless of the complexity of the data, a properly designed and trained neural network should be able to identify single molecules with an accuracy restricted only by the statistical limitation imposed by the signals available in single-molecule spectroscopy.

In complex cases of single-molecule fluorescence data, the direct application of MLE may fail. One could develop a physical model for the nonideality of the fluorescent probe and incorporate this model into a successful MLE approach. However, this requires a specific knowledge of the fluorescent probe's physical characteristics (including the number and frequency of distinct fluorescent states it can occupy). The advantage of using a neural network approach in such a case is that it requires no such model. As long as the two or more fluorescent species (or even groups of species) to be distinguished can be separated and used in the training of the network, in principle the neural network approach should be able to utilize all available information in performing species identification. Thus, the neural network approach is more general, allowing model-independent analysis of mixtures of fluorophores with either ideal or nonideal properties using a single methodology.

References and Notes

- (1) Boddy, L.; Wilkins, M. F.; Morris, C. W. *Cytometry* **2001**, *44*, 195.
- (2) Jain, A. K.; Duin, R. P. W.; Mao, J. C. *IEEE Trans. Pattern Anal. Machine Intelligence* **2000**, *22*, 4.
- (3) Zhang, L. X.; Li, J.; Jiang, Z. T.; Xia, A. *Polymer* **2003**, *44*, 1751.
- (4) Krasnopolsky, V. M.; Schiller, H. *Neural Networks* **2003**, *16*, 321.
- (5) Thissen, U.; Melssen, W. J.; Buydens, L. M. C. *Anal. Chim. Acta* **2001**, *446*, 371.
- (6) Buscema, M.; Sacco, P. L. *Expert Systems* **2000**, *17*, 149.
- (7) Meireles, M. R. G.; Almeida, P. E. M.; Simoes, M. G. *IEEE Trans. Industrial Electron.* **2003**, *50*, 585.
- (8) Eggeling, C.; Berger, S.; Brand, L.; Fries, J. R.; Schaffer, J.; Volkmer, A.; Seidel, C. A. M. *J. Biotechnol.* **2001**, *86*, 163.
- (9) Kuhnemuth, R.; Seidel, C. A. M. *Single Molecules* **2001**, *2*, 251.
- (10) Cotlet, M.; Hofkens, J.; Habuchi, S.; Dirix, G.; Van Guyse, M.; Michiels, J.; Vanderleyden, J.; De Schryver, F. C. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14398.
- (11) Cotlet, M.; Hofkens, J.; Kohn, F.; Michiels, J.; Dirix, G.; Van Guyse, M.; Vanderleyden, J.; De Schryver, F. C. *Chem. Phys. Lett.* **2001**, *336*, 415.
- (12) Enderlein, J.; Sauer, M. *J. Phys. Chem. A* **2001**, *105*, 48.
- (13) Maus, M.; Cotlet, M.; Hofkens, J.; Gensch, T.; De Schryver, F. C.; Schaffer, J.; Seidel, C. A. M. *Anal. Chem.* **2001**, *73*, 2078.
- (14) Herten, D. P.; Tinnefeld, P.; Sauer, M. *Appl. Phys. B—Lasers Opt.* **2000**, *71*, 765.
- (15) Sauer, M.; Angerer, B.; Han, K. T.; Zander, C. *Phys. Chem. Chem. Phys.* **1999**, *1*, 2471.
- (16) Sun, Y.; Whitehead, B. A.; Davis, L. M. *Proc. Adv. Fluorescence Sensing Technol. IV* **1999**, 3602, 379.
- (17) Enderlein, J.; Robbins, D. L.; Ambrose, W. P.; Keller, R. A. *J. Phys. Chem. A* **1998**, *102*, 6089.
- (18) Ambrose, W. P.; Goodwin, P. M.; Enderlein, J.; Semin, D. J.; Martin, J. C.; Keller, R. A. *Chem. Phys. Lett.* **1997**, *269*, 365.
- (19) Zander, C.; Drexhage, K. H.; Han, K. T.; Wolfrum, J.; Sauer, M. *Chem. Phys. Lett.* **1998**, *286*, 457.
- (20) Eggeling, C.; Fries, J. R.; Brand, L.; Gunther, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1556.