Predicting particle properties in optical traps with machine learning

Hamilton, Lachlan H. W., McQueen, Lauren, Smee, Oscar, Nieminen, Timo, Rubinsztein-Dunlop, Halina, et al.


Event: SPIE Nanoscience + Engineering, 2020, Online Only
Predicting particle properties in optical traps with machine learning

Lachlan H.W. Hamiliton1,*,†, Lauren R. McQueen1,*†, Oscar W. Smee1,*†, Timo A. Nieminen1, Halina Rubinsztein-Dunlop1, and Isaac C.D. Lenton1

1School of Mathematics and Physics, The University of Queensland, Brisbane, QLD 4072, Australia
*Authors had equal contribution to the work and have been listed alphabetically

ABSTRACT

Identifying a particle in an optical trap can be a difficult task, especially for biological samples with low contrast. The relationship between particle properties, such as radius and refractive index, and measurable properties, such as the optical trap stiffness, tends to be both non-intuitive and non-linear. Here we aimed to estimate particle properties from easily measurable trap properties with machine learning. We demonstrate methods for real-time estimates of the radius and refractive index of particles trapped by optical tweezers. This is achieved by analysing a particle’s position and corresponding optical force with artificial neural networks. We designed and tested three networks, each using a different pre-conditioning method for the input data. We used trap stiffness measurements, position histograms and position-force time series data. Networks were initially trained and evaluated using synthetic data covering a range of particle refractive indices and radii. Each network successfully predicted the particle properties from the input data. However, initial results suggest that the trap stiffness and histogram network require a large amount of data, potentially complicating experimental realisation. On the other hand, the time series network required far less data, suggesting that it could easily be implemented experimentally. We subsequently demonstrated this by training a time series network for binary particle classification using experimental measurements of particle position and force. We found that with remarkably few force and position values we were able to accurately discriminate between different types of particles. This provides a proof of concept that real-time particle recognition is achievable with machine learning.

Keywords: optical trapping, machine learning, trap stiffness, time series, binary classification, particle classification, real time, convolutional neural networks

1. INTRODUCTION

Optical tweezers use the momentum of light to trap and manipulate particles. Particles are typically drawn close to regions of high intensity such as the focus of a tightly focused laser beam.1 Optical tweezers are able to exert femtonewton scale forces with high precision2, 3 making them useful for manipulating micrometer scale particles. Combined with various particle tracking methods3, 4 and methods for estimating the optical force acting on trapped particles,5–7 optical tweezers are an ideal tool for studying the mechanical and physical properties of a wide range of mechanical microsystems both biological and microrheological.

Many types of optical tweezers measurements rely on using particles with well characterised optical properties.8, 9 If we know the particle’s optical properties, we can easily calculate the expected optical forces and infer information about the surrounding environment from how the particle behaves in the optical trap. However, in many situations we don’t know the particle’s properties, for instance, if we are trapping an unknown particle inside a cell or micro-organism. In these situations, determining the particle’s optical properties is non-trivial: the relationship between optical force, position and particle properties is often both non-intuitive and non-linear.

Here we describe a solution involving training an artificial neural network to predict particle properties from properties that can be easily measured. More specifically, we use measurements of an optically trapped particle’s
position and force to estimate the particle's radius and refractive index. Although we focus on spherical particles, we don’t foresee any major difficulties in applying a similar technique to non-spherical particles. In Section 2, we describe three different networks: a trap stiffness network, a histogram network and a time series network; each network involving a different structure for the input data. To explore these different models, we used synthetic data generated with the five degree of freedom network described in Ref.:10 using this network allowed us to generate a large volume of data to train and evaluate the different networks. In Section 3 we present initial experimental verification of the time-series network for discriminating between two species of optically trapped particles with different size and refractive index. We find that once trained, our network is able to discriminate between the two different particles with a better than 98% accuracy with only 50 sample points at a 5kHz sample rate. Combined with fast hardware, such as by embedding our network on a field programmable gate array11 for real-time data processing, this method could be used for fast, real-time particle sorting.

2. PARTICLE PREDICTING NETWORKS

2.1 Overview of neural networks

A neural network can be treated as a transformation \( N(x_{in}) \rightarrow y_{out} \), from our input data \( x_{in} \) to our predicted outputs \( y_{out} \). It maps from the measured force and position \( x_{in} \) to the predicted radius and refractive index \( y_{out} \).

Training is facilitated by feeding the network a dataset with known outputs and using a learning rule to update its connections and maximise predictive accuracy. An unseen dataset is then used to validate that the network has learnt to predict the entire system, and not only the seen data. Each time this occurs is a training ‘epoch’. There are innumerable ways to structure the networks and many methods to process the input data to improve the learning rate and accuracy.

Our networks were trained on data simulated by the 5 degree of freedom (5DOF) neural network created by I.Lenton et al.10 The 5DOF network takes the position vector \((X, Y, Z)\), radius \( r \) and refractive index \( n_p \) of the particle we wish to simulate – hence 5 degrees of freedom. It then outputs the force in all dimensions. This network was trained by the Optical Tweezers Toolbox (OTT)12 with a circularly polarized beam of 800nm, in a medium of refractive index \( n_m = 1.33 \) and numerical aperture 1.02.10 The 5DOF network allows for significantly faster simulations when compared to analytical T-matrix methods such as the OTT.10 Creating a large amount of simulated data, for example when training our networks, is now possible in a short amount of time.

2.2 Trap stiffness network

The first network was trained on pre-processed data describing the linear trap stiffness along the beam axis, and along two orthogonal axes. When a particle is trapped in a Gaussian beam close to the equilibrium position, the optical force is approximately linear and related to the position by

\[
F(x) = -\kappa \cdot x
\]  

(1)

where \( \kappa \) is a vector describing the trap stiffness or spring constant of the trap.3,4,13–16 That is, it describes the force a particle will experience at some distance \( x \) from the equilibrium position.17 This force is generated by the gradient of the intensity, \( F \propto \nabla I \). In cylindrical coordinates, \((\rho, z) = (x^2 + y^2, z)\), we can describe the radial intensity of a circularly polarised Gaussian beam by

\[
I(\rho)_{\text{Gauss}} = I_0 \exp \left( -\frac{2\rho^2}{\omega_0^2} \right)
\]  

(2)

where \( \omega_0 \) is the beam waist and \( I_0 \) is related to the beam power. When the particle is near the trap center, i.e., when \( \rho \ll \omega_0 \), we observe

\[
I(\rho \ll \omega_0) = I_0 \left( 1 - \frac{2\rho^2}{\omega_0^2} \right) \implies F_\rho = -\kappa_\rho \rho
\]  

(3)
For a tightly focused linearly polarized trap, the intensity distribution around the focus will be asymmetric.\textsuperscript{16,18} In this case we can consider the components on the $x$ and $y$ axis separately,

\begin{equation}
F_x = \kappa_x x, \quad \langle E \rangle = \frac{1}{2}\kappa_x \langle x^2 \rangle.
\end{equation}

In an experiment, the trap stiffness can be measured by observing how the thermal motion of the particle is resisted by the optical forces.\textsuperscript{3,4,14,15} This can be found by equipartition theorem, allowing a Gaussian probability distribution to describe its motion. Note that the trap stiffness will be different in all dimensions for a linearly polarised beam.\textsuperscript{16,18}

\begin{equation}
\kappa_x = \frac{k_B T}{\langle x^2 \rangle}, \quad P(x) \propto \exp \left( -\frac{1}{2} \kappa_x x^2 \right)
\end{equation}

Fig. 1 demonstrates the calculation of $\kappa$, and further motivates the use of machine learning. Machine learning can be used to determine relationships between correlated variables. Complicated and often unpredictable variables can be linked, removing the need for theoretical analysis. The behaviour seen in the trapping landscape (Right) is non-intuitive and non-linear – yet it is decidedly non-random. This is an ideal system for machine learning.

Figure 1. Trap stiffness calculations, modelling and trapping landscape. The results from 5DOF simulations of particles in an optical trap after $10^5$ time steps. Left the linear regime where trap stiffness predicts the expected force on the particle. Middle a probability distribution compared with the measured histogram. This is used to determine the uncertainty in each measurement with $\chi^2$ tests. Right the trapping landscape of varying refractive index and radius

By pre-processing the position/force data we can extract properties such as trap stiffness. The “trap stiffness network” is a multiplayer perceptron network (MLP) trained on axial trap stiffness $\kappa$, calculated from the measured variance in position. We can effectively describe it as:

\begin{equation}
\mathcal{N}(\kappa_x, \kappa_y, \kappa_z) \rightarrow (r, n_p)
\end{equation}

To validate our model, we used numerical simulations. We used the 5DOF\textsuperscript{10} neural network trained with data from the Optical Tweezers Toolbox (OTT)\textsuperscript{12} to generate the forces. Fig. 1 shows an analysis of these simulations.

Preliminary results for the “trap stiffness network” are shown in Fig. 2. We used 10,000 radius-refractive index pairs from $1.36 \leq n_p \leq 1.75$ and $0.7\mu m \leq r \leq 1\mu m$, corresponding to the active area in Fig. 1-Right. We achieved a MAPE of 4.6%.
Figure 2. **Neural network results.** Results from the “trap stiffness network” trained on 3 simulated values of $\kappa_x, \kappa_y, \kappa_z$. **Top** schematic of the ANN. **Left** the network accuracy during each epoch. **Middle/Right** the network’s guess for the radius and refractive index values for each trap stiffness set.

Unfortunately, the high-accuracy state shown by Fig. 2 was unlikely compared to other local minima where the refractive index was incorrectly estimated. Fig. 3 shows the more likely state, and a frequency plot quantifying how unlikely the high accuracy state was (red circle).

Figure 3. **Local minima analysis.** Demonstrating the more likely and less optimal state the network finds when training. **Left** the frequency of each training accuracy value after 500 epochs. The network is initialized in a random initial state. Note that the best state (red circle) was only observed once after 600 trials. **Middle/Right** a typical sub-optimal radius and refractive index prediction that the network makes.

Current work is attempting to adjust the training parameters to force the state into the more accurate global minimum. However, by considering the trapping landscape (Fig. 4), we notice that there are regions of indistinguishable trap stiffness. That is, the values between neighbouring radii or refractive indices are very
similar. Furthermore, the information given by each axis is similar, meaning there may not be three inputs with unique information. Considering that we are trying to predict two outputs, it is likely that trap stiffness does not contain enough information for accurate and reliable particle predicting.

Figure 4. **Trapping landscape.** The axial trap stiffness in each direction, for each radius and refractive index pair used. Notice that the X and Y landscapes are almost identical, as the 5DOF network simulates a circularly polarized trap. Both the magnitude and the shape are similar, meaning there is little information gained when using both as an input.

### 2.3 Histcount network

\[ N(x_1, x_2, \ldots, x_n) \rightarrow (r, n) \] (7)

Another way to capture information about the dynamics of the trap is to pre-process the data into histograms or position probability distributions. The “histcount network” is a multilayer perceptron network (MLP) identifying refractive index and radius from the probability distribution of the positions. The relative frequency for a histogram of 50 bins, called histcounts, were used.

Using 5,000 simulations from \( 1.3 \leq n_p \leq 2 \) and \( 0.1 \leq r \leq 1\mu m \) we achieve a MAPE of 3.2\% using only the particle’s x-position measurements. This demonstrates a proof of concept for this approach. The next step will be to include the histcounts in 2 or 3 dimensions, or to include the force data.
2.4 Time series network

The third and final network we explored involved training the network using time series data. The aim for this network was to capture the particle dynamics by using ordered time series of the particle’s position and/or force. We chose the experimentally measurable properties of the particle’s locations and the forces acting on the particle as it moves about in the trap. The final goal is to predict the radius and refractive index of the particle from its dynamics.

To deal with the time series data, image classification models based on deep Convolutional Neural Networks (CNN) can be used. Modern deep CNN like ResNet are exceedingly powerful and are responsible for huge strides in the field of computer vision and hence could be promising for use in this problem. Time series can be thought of as a ‘one dimensional’ image, as such, we could imagine applying CNNs to time series *images* in much the same way as we apply them to regular images. The structure and order matter for both a time series...
and images. Therefore a CNN should be able to efficiently analyse the dynamic patterns in the time series much in the same way it could analyse important features in an image. By using ordered time series data, we preserve the dynamic information from the particle’s motion which is missing from shuffled data, preprocessed data and individual force/position measurements. The inputs and outputs of the network can be summarised as follows:

\[ N(F(t_1), F(t_2), \ldots, F(t_N)) \rightarrow (r, n_p) \]  

To start, the CNN was trained on only the force data. 30000 time series examples were generated for the train/test data. Each time series was composed of 1000 points (or 0.1 seconds of sample time) of the force in the \( x \), \( y \) and \( z \) directions. Each of the time series corresponded to a simulated particle with radii/refractive index in the space \( 1.5 \leq n_p \leq 1.7 \) and \( 0.4 \mu m \leq r \leq 0.6 \mu m \). The model was trained over several hours with a google colab GPU. With a training set of 27000 time series we achieved a mean absolute percentage error (MAPE) of 1.63% on the 3000-time series test set.

Fig. 6 breaks down the percentage error of the model on the unseen dataset by the true radii and refractive indices and shows that the error is actually concentrated, at least for prediction of radius, in certain locations such as the band of overestimation on the left and splotches of underestimation close to that band.

![Figure 6. Error heatmaps for “time series” network. The heatmap on the left shows the MAPE in the radius predictions, and the heatmap on the right shows the relative error in the refractive index predictions.](image)

There was a lot of overfitting occurring in the training of the model which could imply that some regularization/restructuring of the architecture could be a good avenue for improvement of the efficiency of the model. More data could be simulated to counteract the overfitting with diminishing returns at the cost of longer training and simulation time.

### 3. EXPERIMENTAL RESULTS

#### 3.1 Binary classification

We next took the time series network to real, experimental data in the case of a binary classification problem. A polystyrene (\( n = 1.57, \; r = 0.26 \pm 0.02 \mu m \)) and silica particle (\( n = 1.45, \; r = 0.83 \pm 0.05 \mu m \)) were placed in the trap and measurements of the positions and forces on the particle were taken.

The same structure of model was used as in the full regression model except the final layer was replaced with a classification output and both force and position time series were used but with only the \( x \) and \( y \) axes.
is to allow modest optical tweezer systems to be compatible with the network, as the z axis can be difficult to measure.

![Binary Classification: Training](image1)

Figure 7. **Binary classification training.** This plot shows the prediction accuracy in the training and validation (test) data compared with training epoch during binary classification. With a relatively short number of epochs, the network had accuracy above 95%.

The time series data collected from the trap was taken and cut up into chunked time series of 100 points in length (with a 5kHz sampling rate, this corresponds to 0.02 seconds in an experiment). Splitting up the time series in this way gave 940 series in the test set and 235 in the training. The model was trained over 200 epochs in several minutes with a Google Colab GPU (fig.7). Over 10 training runs mean accuracy was 99.9% on the unseen data - implying that the model can distinguish the particles with a high degree of accuracy.

The high accuracy with time series of only 0.02s made us question just how little data is needed. We tested this by splitting the time series into shorter chunks and retraining the model. The fig.8 shows how mean accuracy from 5 training runs of 100 epochs varies with the time series length. With 10 points per time series or just 0.002 seconds of sample time the model can classify the particles with a mean accuracy of 86.4% on the test set. With 200 points or 0.04 seconds, mean classification accuracy on the test set reached 100%.

![Time Series Length vs. Accuracy](image2)

Figure 8. **Binary classification accuracy vs. number of data points needed.** We tested the amount of data needed for an accurate binary classification, given that the network trains for the same number of epochs.
4. DISCUSSION

We have trained neural networks with different inputs and architectures, allowing the strengths and weaknesses for each method to be compared. We have provided three levels of data pre-processing: 3 single trap stiffness values, 50 position hiscounts and the entire force time series. We tested the ability of machine learning to generate predictions of the radius and refractive index for each case.

The trap stiffness network at first appeared to work, however Fig. 3 shows that when being trained, it is more likely to be stuck in a local minimum, where refractive index is inaccurately predicted. The local minima problem for the trap stiffness network may be solved by refining the optimisation parameters or exploring other network architectures. However, with the success of the time series network, and by considering the similarity of the trapping landscapes $\kappa_x(r, n_P)$ and $\kappa_y(r, n_P)$ in Fig. 4, it may be unsuitable for circularly polarized experimental systems. If the two landscapes are not unique, then they do not provide anymore information. It makes intuitive sense that with only one piece of information, it would be impossible to identify two properties for every radius/refractive index pair. For real systems, finding the Z-axial trap stiffness is difficult. In a circularly polarized trap, the symmetry of the $x$-$y$ plane means the trap stiffness will be the same. Therefore this method may not be suitable for real particle prediction in a circularly polarized beam.

The histcount network provides a powerful tool requiring only a single dimension of position data. This may be suited to optical tweezers systems which cannot measure the force of the particle in time. The high accuracy displayed in Fig. 5 motivates this network for experimental particle classification. The experimental verification of this network would allow particle classification to be achieved with only one dimension of positions. The drawback of this approach is an experiment must be run for enough time to accurately represent the motion with a histogram.

The time series network demonstrated a high accuracy for the 5DOF simulated seen in Fig. 6 across the radius/refractive index pairs in the range $1.5 \leq n_p \leq 1.7$ and $0.4\mu m \leq r \leq 0.6\mu m$. We then trained the time series network for one specific binary classification; however general classification in this entire range should still be possible. It may be possible to use this network without any pretraining on the experiment – provided the data is taken on an optical tweezers system with the same intensity, polarization, and beam width as the 5DOF network. Furthermore, it may not be a necessary to have accurate position or force data for estimating the particle properties. Instead of training a network using position and force data, it should be possible to train a time series network to use intensity of scattered light.

The ability to identify a particle in an optical trap is crucial in all experiments, and binary classification is only the beginning. Automated particle tracking, cell recognition and live classification should all be possible with the time series network. By requiring only 0.004s of data with a modest PSD-camera system, this approach is applicable to most optical trapping systems. The network can be trained with time series of any fixed length – meaning it is versatile enough to allow slower systems to provide enough information for particle classification.

5. CONCLUSION

We have demonstrated both with simulations and experiments that artificial neural networks can be used for estimating the optical properties of particles held in optical tweezers. With our work, we have demonstrated that live particle classification with 0.004s of data or less is possible. This can be utilized in the future for novel systems such as automatic particle tracking, sorting, cell recognition, and more.

The application of machine learning to optical tweezers has allowed for faster dynamic simulations. This was useful for generating the large amounts of data required to train and evaluate different network architectures. Overall, by applying machine learning to optical tweezers systems, we have presented new methods for predicting particle properties.

Funding

This work was funded by the Australian Government through the Australian Research Council’s Discovery Projects funding scheme (project DP180101002). L.H, L.M, O.S received support from The University of Queensland’s Summer Research Program scholarship. I.L. acknowledges support from an Australian Government RTP Scholarship.
Acknowledgement

We would like to acknowledge the computing resources of The University of Queensland’s getafix cluster.

Credit Authorship Contribution Statement


REFERENCES
