Opto-Electronic Science

Supplementary information

January 2023, Vol. 2, No. 1

DOI: 10.29026/oes.2023.220019

Chiral detection of biomolecules based on reinforcement learning

Yuxiang Chen[1†], Fengyu Zhang[2,4†], Zhibo Dang[1], Xiao He[1], Chunxiong Luo[2,4], Zhengchang Liu[3], Pu Peng[1], Yucheng Dai[3], Yijing Huang[1], Yu Li[3] and Zheyu Fang[1,3*]

[1]School of Physics, Peking University, Beijing 100871, China; [2]The State Key Laboratory for Artificial Microstructures and Mesoscopic Physics, School of Physics & Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China; [3]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China; [4]Wenzhou Institute University of Chinese Academy of Sciences, Wenzhou 325001, China.

[†]These authors contributed equally to this work.

*Correspondence: ZY Fang, E-mail: zhyfang@pku.edu.cn

This file includes:

Supplementary information for this paper is available at https://doi.org/10.29026/oes.2023.220019

Section 1: Finite-difference time domain (FDTD) simulations

We used 3D FDTD simulations (Lumerical Solutions Ltd.) to determine the complex reflectivity of metal nanostructures. According to nanostructures proposed by artificial intelligence algorithm, simulations were completed under periodic boundary condition (horizontal direction) and perfectly matched layer absorbing boundary condition (vertical direction). In order to decrease simulation time and to ensure enough accuracy of simulation, the mesh of the region near nanostructures was set to 2 nm and the overall simulation time as 1000 fs.

Section 2: Data enhancement

The significance of data enhancement. It is widely believed that the size of database plays a crucial role in artificial intelligence algorithm. In this project, the generation of training dataset depends on electromagnetic simulations which demand extensive computational resources and time. Data enhancement utilizes the symmetry of the system to multiply the dataset without any extra simulations. This technology enlarges significantly training data and improves the performance of neural networks.

The implementation of data enhancement.[S1] Data enhancements that we use depend on the rotational symmetry and mirror symmetry. R_{RCP} and R_{LCP} are defined as the reflectivity with a RCP/LCP input. When a nanostructure called A rotates 90 degrees, corresponding R_{RCP} and R_{LCP} do not vary. On the other hand, for nanostructure A, mirror symmetry transformation exchanges R_{RCP} and R_{LCP}. The formulas are as follows:[S1]

$$\begin{cases} R_{\mathrm{RCP}}\left(\mathrm{rot90}\left(\mathrm{A}\right)\right) = R_{\mathrm{RCP}}\left(\mathrm{A}\right) \\ R_{\mathrm{LCP}}\left(\mathrm{rot90}\left(\mathrm{A}\right)\right) = R_{\mathrm{LCP}}\left(\mathrm{A}\right) \end{cases},$$

$$\begin{cases} R_{\mathrm{RCP}}\left(\mathrm{flip}\left(\mathrm{A}\right)\right) = R_{\mathrm{LCP}}\left(\mathrm{A}\right) \\ R_{\mathrm{LCP}}\left(\mathrm{flip}\left(\mathrm{A}\right)\right) = R_{\mathrm{RCP}}\left(\mathrm{A}\right) \end{cases}.$$

It is obvious to obtain a D_4 point group based on mirror symmetry transformation and 90-degree rotation. After anyone of eight transformations in the D_4 point group, the new reflectivity of a given nanostructure has a simple relationship with the original one. Therefore, the dataset is enlarged to 8 times through the eight transformations.

Section 3: Convolutional neural networks (CNN)

The structures of CNNs. As the expressions in the text, there are several different convolutional neural networks involved in the project. As shown in Fig. S1, all of the CNNs consist of convolutional layers, pooling layers, activation layers and fully connected layers. The kernel size of all convolutional layers is 3×3, and pooling layers are set as 2×2 max pooling. We choose Rectified Linear Unit (ReLU) as activation function except the output layer. The final activation layer is sigmoid function which transforms the output to interval [0,1]. In addition, batch normalization method is involved to accelerate the convergence of models.

Section 4: Bayesian optimization (BO) based on Gaussian process priors[S1]

BO is an effective optimization strategy for a complicated function $f(x)$. An acquisition function which determines the new nanostructure is the core of the strategy. In this project, as an estimate, Gaussian process is supposed to be a prior distribution of $f(x)$. Define $\{(x_i, y_i)\,|i = 1,2,3,\cdots n\}$ as the available data with a capacity of n. According to Gaussian distribution, $y_i \sim \mathcal{N}\left(f(x_i), \sigma^2(x_i)\right)$, where $\sigma^2(x_i)$ denotes the variance noise. Analyzation of the available data predicts a mean function $\mu(x)$ and corresponding variance function $\sigma^2(x)$. $\mu(x)$ and $\sigma^2(x)$ describe an interval of $f(x)$, so it is easy to find the best one among all possible values of $f(x)$, and the corresponding x is exactly the result of the acquisition function.

Section 5: Comparison of reinforcement learning and classic supervised learning

It is a fact that the most time-consuming process in the design is exactly electromagnetic simulations, so we aim at decreasing the number of simulations. The algorithm we propose is an exploration of parameter space, which trains ANNs and searches for new structures at the same time. On the contrary, classic supervised learning firstly completes training process, then it uses the trained model to find structures. The main difference is that our algorithm eliminates numer-

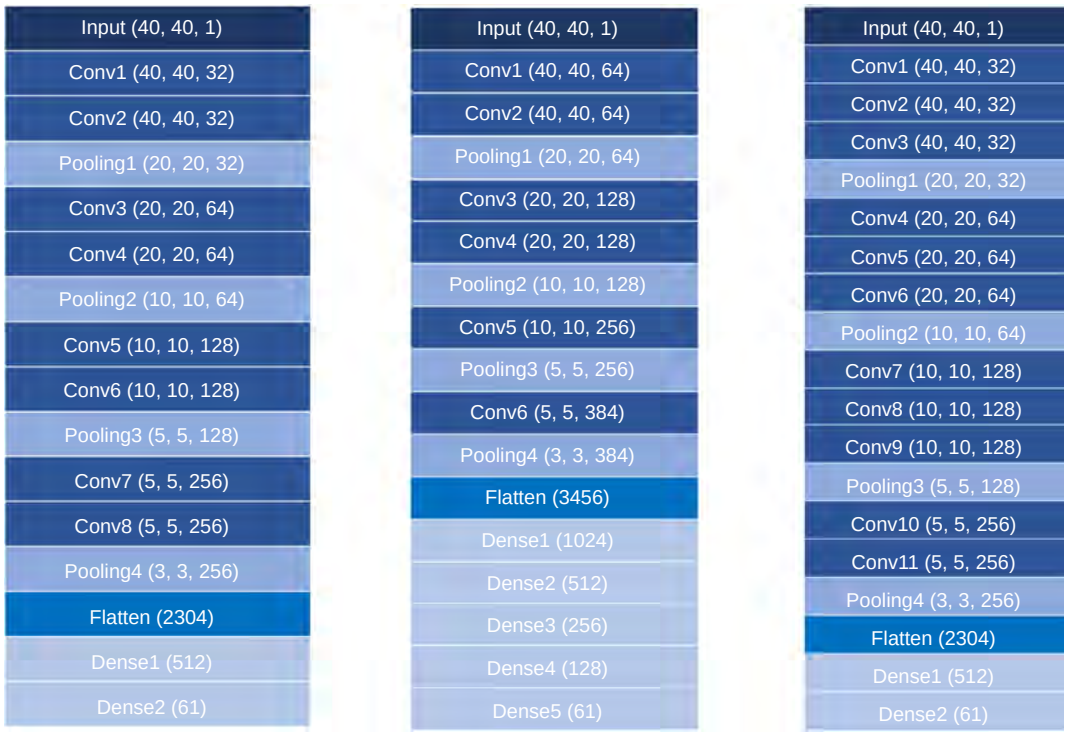

Fig. S1 | Convolutional neural networks with different structures. Input represents the input layer. Conv represents convolutional layers. Pooling represents max pooling layers. Dense represents fully connected layers. Flatten represents flatten layers. Numbers in brackets denote the output size of corresponding network layers. There is an activation layer which is not shown in the figure after every convolutional layer and fully connected layer.

ous undesirable structures. Although the number of structures with significant chirality is appreciable, most structures possess weak chirality. Classic supervised learning needs to simulate the optical response of diverse structures including a great quantity of weak-chiral structures, which caused a huge waste of time. For example, our algorithm does 8000 simulations to obtain the optimized structures, and three thousand of them exhibit CD values over 0.4. According to the data distribution, weak-chiral structures occupy about 99% of the whole parameter space. As an estimate, simple supervised machine learning needs a dataset consist of 300,000 structures to achieve similar effects, which costs several months in simulations. Our algorithm takes merely 2 weeks. Therefore, the economization on calculation resource is very significant.

Section 6: Microfluidic chips

Figure S2 illustrates the photos of microfluidic chips.

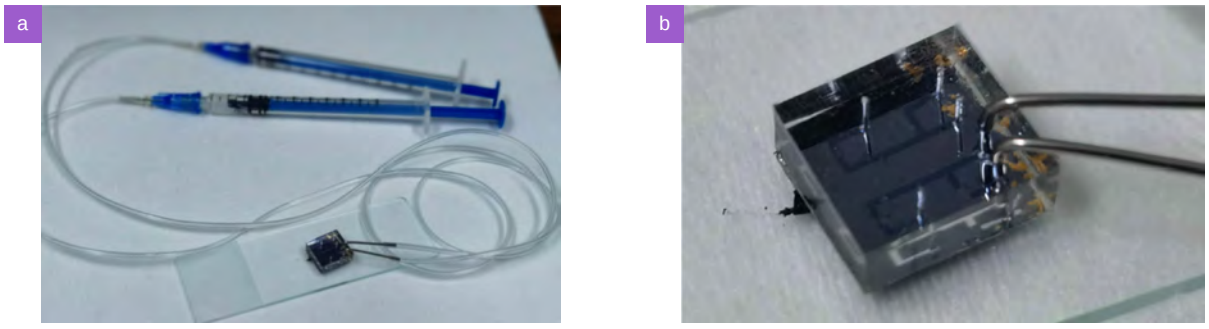

Fig. S2 | Photos of microfluidic chips.

Section 7: Experimental details

Fabrication of microfluidic chips. The first step is the fabrication of metal nanostructures. Two layers of different positive resists (PMMA A2 950 and MMA) were spun on the Si/SiO_2 substrate by spin-coating method. A scanning electron microscope (SEM, FEI Quanta 450 FEG) with a Nano Pattern Generation System (NPGS) module patterned the nanostructures proposed by the artificial intelligence algorithm. A gold layer with thickness of 30 nm was deposited by an electron beam evaporator (DE400DHL, DE Technology). Finally, metal nanostructures on the substrate were completed after lift-off process. Microfluidic structures made of Polydimethylsiloxane (PDMS) were fabricated on the substrate with metal nanostructures. PDMS can adhere to SiO_2 layer tightly to avoid leakages.

Chirality measurements. We used a dark-filed microscope (HSI V3, CytoViva Co.) to measure all of the spectra. A linear polarizer and quarter-wave plates fixed additionally in the microscope distinguish LCP and RCP light, and the system can gauge the reflectivity with a LCP/RCP input. All the spectra were normalized according to the reflective spectra of an Ag mirror.

CL measurements. The CL detection system (Gatan MonoCL4 Plus) is equipped on SEM (FEI Quanta 450 FEG). The detector included linear polarizers and quarter-wave plates in order to distinguish LCP/RCP components. Metal nanostructures were excited by a 30 kV electron beam, and photomultipliers amplified chiral CL emissions collected by a parabolic mirror. All the CL signals were modulated by a 650 nm bandpass filter.

References

S1. Li Y, Xu YJ, Jiang ML, Li BW, Han TY et al. Self-learning perfect optical chirality via a deep neural network. *Phys Rev Lett* **123**, 213902 (2019).