CS 229
Machine Learning
Spring 2016

# STRUCTURE PREDICTION OF OPTICAL FUNCTIONAL DEVICES WITH DEEP LEARNING

Kai Zhang[1], Shiyu Liu[2] and Qing Wang[3]

## 1 Optical device and simulation

### Simulation setup: optical device and electrical fields

***Simulation software***

- Rigorous Coupled-Wave Analysis (RCWA).

***Optical device***

- Silicon (Si) nanostructures represented by a binary vector.

***Input signal***

- Normal incident light.

***Output signal***

- Electric field distributed in a 2D space.

### Data set for learning

- 200000 training samples, 50000 test samples.
- Si structure $\boldsymbol{y} \in \{-1, +1\}^{20}$
- Output electrical field collected at near-field and far-field locations with $\boldsymbol{x} \in \mathbb{C}^{51}$

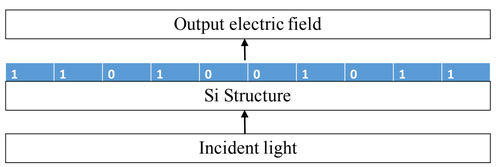


Figure 1 Schematic of 1D Si structure with simulation procedure.

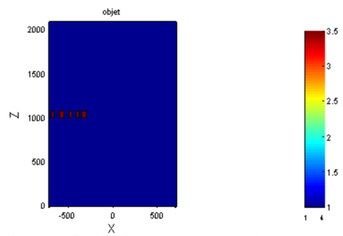


Figure 2 1D Si structure aligned along x-direction at height z = 1000.

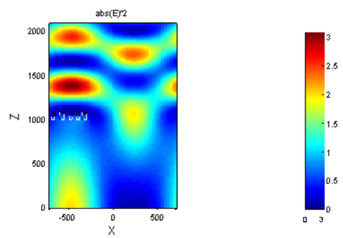


Figure 3 Output electrical field given light normal incident on Si structures shown in Figure 1.

## 2 Neural networks

### Neurons

***Activation function***

- Rectified Linear Unit (ReLU)

$$f(z) = \max(0, z)$$

### Fully connected neural networks

***Parameters at layer l***

- Neuron weights: $W^{(l)} \in \mathbb{R}^{H_l \times H_{l-1}}$
- Neuron bias: $b^{(l)} \in \mathbb{R}^{H_l}$

***Forward-feeding***

$$\boldsymbol{z}^{(l)} = f_j\left(W_l \boldsymbol{z}^{(l-1)} + \boldsymbol{b}_l\right)$$

- For all *j*, the activation function is ReLU in this project

***Prediction***

$$\hat{\boldsymbol{y}} = W_o \boldsymbol{z}^{(p)} + b_o$$

### Training process

***Cost function***

- Mean Square Error (MSE) function

$$J(W, b; \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{m}\sum_{i=1}^{m} \|\boldsymbol{y}^{(i)} - \hat{\boldsymbol{y}}^{(i)}\|_2^2$$

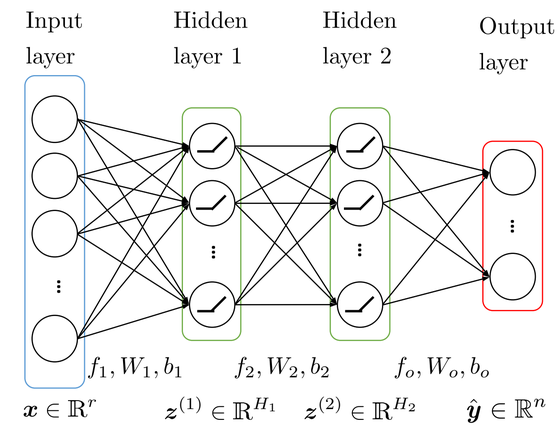


Figure 4 Neural networks with 2 hidden layers.

***Minimization procedure***

- Backward-propagation to find the derivatives of the cost function with respect to parameters $W^{(l)}, b^{(l)}$ between each layers
- Stochastic gradient descent (SGD) to find the optimal solution
- Constant learning rate $\alpha$ = 0.001

## 3 Training results and discussion

### Training configurations and definitions

***Neural networks configuration***

- The magnitude and phase of the transmitted electrical field are used as the input to the neural networks, i.e. $\boldsymbol{x} \in \mathbb{R}^{2\times 51 = 102}$
- The magnitude and phase are normalized before feeding into the neural network
- Two hidden layers with $H_1 = H_2 = 51$

***Error definitions***

- Training and testing error

$$\epsilon_{\text{train/test}} = \sqrt{\frac{\sum_{i=1}^{m} \|\hat{\boldsymbol{y}}^{(i)} - \boldsymbol{y}^{(i)}\|_2^2}{\sum_{i=1}^{m} \|\boldsymbol{y}^{(i)}\|_2^2}}$$

- Structure prediction error
  - Structure prediction

$$\hat{\bar{y}}_j = \begin{cases} +1, & \hat{y}_j > 0 \\ -1, & \hat{y}_j < 0. \end{cases}$$

  - error

$$\epsilon_{\text{prediction}} = \frac{1}{m}\sum_{i=1}^{m} \frac{\sum_{j=1}^{n} \mathbb{1}\{\hat{\bar{y}}_j \neq y_j\}}{n}$$

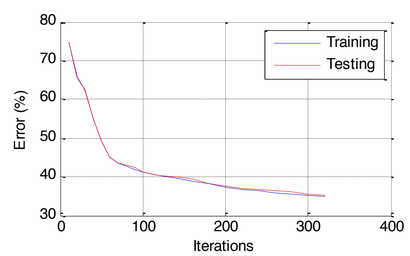


Figure 5 Learning curve for near-field electric field training using two-layer neural networks.

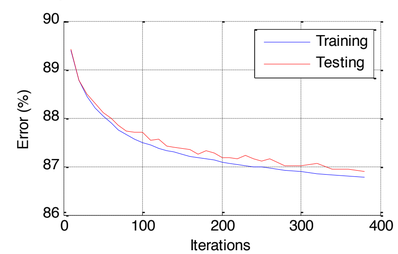


Figure 6 Learning curve for far-field electric field training using two-layer neural networks.

### Near-field electric field training

***Training & testing error***

- Train

$$\epsilon_{\text{train}} = 34.90\%$$

- Test

$$\epsilon_{\text{test}} = 35.40\%$$

***Structure prediction error***

- Train

$$\epsilon_{\text{prediction,train}} = 3.48\%$$

- Test

$$\epsilon_{\text{prediction,test}} = 3.59\%$$

### Far-field electric field training

***Training & testing error***

- Train

$$\epsilon_{\text{train}} = 86.75\%$$

- Test

$$\epsilon_{\text{test}} = 86.91\%$$

***Structure prediction error***

- Train

$$\epsilon_{\text{prediction,train}} = 29.47\%$$

- Test

$$\epsilon_{\text{prediction,test}} = 29.58\%$$

## 4 Application

### Physical implication

***Deviation efficiency***

We achieved 43.4% efficiency as a 45° beam deflector. The Si nanostructures were predicted by our deep learning algorithm using ideal 100% deflected electric fields as input in the test.

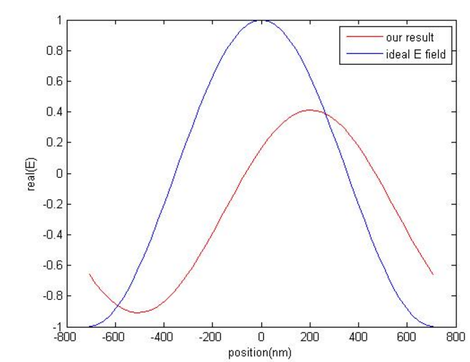


Figure 7 Comparison of 45° deflected transmitted light electric field of the predicted Si structure and the ideal 100% efficiency electric field

1. kzhang3@stanford.edu 2. shiyuliu@stanford.edu 3. wangqing@stanford.edu