Inverse Photonic Design Using Deep Learning

Siddhant Gautam
Electrical Engineering, IIT Madras

December 24, 2019
About

**Title:** Deep Neural Network Inverse Design of Integrated Photonic Power Splitters

**Author:** Mohammad H. Tahersima, Keisuke Kojima, Toshiaki Koike-Akino, Devesh Jha, Bingnan Wang, Chungwei Lin & Kieran Parsons

**Publisher:** Nature, OSA

**Publishing Date:** 04 February 2019
Overview

**Key Points:**
- Using deep learning to predict optical response
- Inversely approximate designs for a targeted response
- Design power splitter with high efficiency and minimum reflection

**Features:**
- Minimum Reflection: 20 dB
- Transmission Efficiency: 90 %
Setting

Material
Artificially engineered subwavelength nanostructure

Applications
1. Optics
2. Integrated Photonics
3. Sensing, Metamaterials
4. Invisibility cloak

Purpose
Control EM fields into specific wavefronts

Drawbacks
Optimization of nanostructures - computationally intensive
E.g. FDTD simulation
Deep Learning based Inverse Design

Salient Features

- Models complex input-output relations
- Complex functions can be learnt in a data-driven fashion
- Uses: High energy physics, medical imaging, high energy physics

Usage

- Forward Design: Predict the optical response of a topology
- Inverse Design: Design a topology for a target optical response
Deep Learning for Forward Modeling to Predict Optical Response

Physics Involved

- Change in refractive index - reflection, refraction and scattering
- Scattering guides the beam to a target port
Training the Neural Network

- Generate labeled data for training using FDTD
- Feed DNN with numerical optical experimental data
- **Input**: Binary images of hole locations
- **Output**: Spectral response at T1&T2 and reflection from R

Pixel:
- 1 - Etched (silicon)
- 0 - Not Etched (silica)
Training Data for supervised learning

Optimization:

- Start with initial hole vector
- Using single stride binary search, solve
  \[
  \max FOM = \min(T_1) + \min(T_2) - \alpha \times \max(|R|)
  \]

1) An asymmetric optimization (\(\alpha=2\)) with a random initial vector
2) A symmetric search (\(\alpha=4\)) with a patterned initial vector
3) Histogram of all transmission train and test data label
4) Learning curve for training (lines) and test (dots) losses
<table>
<thead>
<tr>
<th>Design</th>
<th><strong>Forward Model</strong></th>
<th><strong>Inverse Model</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Inputs</td>
<td>20 × 20 HV arrays</td>
<td>63 x 1 SPEC vector</td>
</tr>
<tr>
<td>Type</td>
<td>Regression</td>
<td>Classification</td>
</tr>
<tr>
<td>Loss Function</td>
<td>Gaussian Log Likelihood</td>
<td>Bernoulli Log Likelihood</td>
</tr>
</tbody>
</table>

**Gaussian log-likelihood loss function:**

\[-\log P(Y|X,W) = \frac{1}{K} \sum_n \frac{1}{2} \log(2\pi\sigma^2) + \frac{1}{2\sigma^2}(y_n - w^T x_n)^2\]

where \(W\) denotes the model parameters, \(K\) is the number of training data.
Use lithography to generate nanostructures

Each pixel is a circle with 1 for etched & 0 for unetched

Change refractive index to determine propagation path

Generate forward data using FDTD

Spectral response

max FOM = min(T1) + min(T2) − α×max(|R|)

Randomly choose initial hole vectors

Optimize spectral transmission values

Collect labelled data for training

Compare with true solution

Forward Design Problem

Hole vector array (20 x 20)

Broadband Spectral data (63 x 1)
Forward Design Results

Spectrum approximation using deep ResNet; 80% input data for training and 20% of the total data for testing. a, b, and c represent predicted spectral response of three different power splitters.
Generate broadband transmission values for each port

Use SPEC as the input data batch and hole vectors as output labels

Predicted HVs using Bernoulli classifier (0 or 1)

Classification converges to 1 or 0 as loss reduces

Produced binary sequence fed back to solver

Run independent FDTD simulation to check validity of response

Inverse Design Problem

Broadband Transmission Values

Hole vectors (labels)
Inverse Design Results

Correlation coefficient. Fitting ResNet predicted transmission values vs true transmission values for (a) port 1 and (b) port 2 (b). Gray circle symbol size is proportional to gradient uncertainty.
Deep Residual Networks

- **Drawbacks of FCDNNs:**
  - Suffers from the problem of vanishing gradients
  - Increasing depth of doesn’t improve performance

- **Remedy - Using ResNet**
  - Improves the depth of training by using “identity shortcut connections”
  - Uses an additional identity function to allow smooth forward and backward propagation of gradients
DNN inverse design for devices with different splitting ratios. Electromagnetic energy density plots calculated using FDTD simulations.
Summary

- NN used for fast approximation of the optical response instead of computationally heavy numerical methods
- Inverse design: NN takes an optical response as input and provides an approximate solution nanostructure in a fraction of a second
- ResNet DNN architecture (8 layers) which allows smooth forward and backward propagation of gradients

**Applications:**
- a) A network that can approximate the spectral response of an arbitrary hole vector within this design space
- b) Use the inverse network to design a power splitter topology for any ratio