Deep Learning for the Inverse Design of Nanophotonic structures

ADEPU VISHAL VARDHAN (170108003) and ADITYA UPPALA (170108039)

Under the guidance of

Dr. Debabrata Sikdar

Abstract

Data Inconsistency leads to slower training process and inaccurate results when Neural networks are trained for predicting the design parameters of photonic devices. These issues arise due to the fundamental property of non-uniqueness in all scattering problems. Tandem network-based models are an effective solution to counteract these issues.

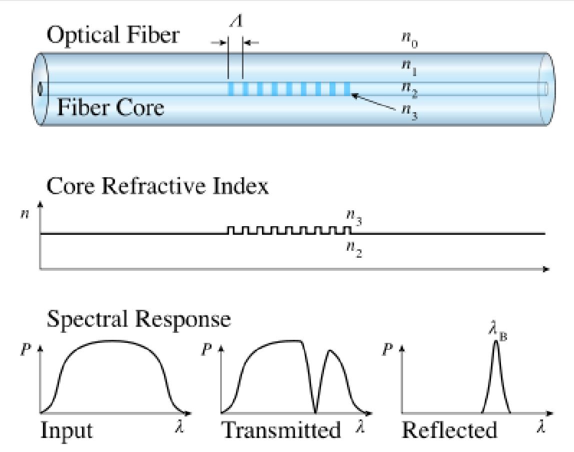
In this project, we present the performance of tandem architecture-based models over direct inverse prediction models for a Fiber Bragg grating design parameter estimation task.

Introduction

Data inconsistency is a major issue while training deep neural networks. When there is an inconsistency with the training data, the training process slows down, and the trained model is not accurate. Data can be inconsistent due to various reasons, one of them being the property of non-uniqueness.

Tandem architecture is an effective solution to overcome the issue of non-uniqueness property and thereby train the inverse models with large amounts of data containing non-unique instances. Tandem means "to connect in series". Tandem architecture cleverly employs forward and backward models in series to mitigate the one-to-many instances.

In this project, we apply this concept of Tandem network to a design parameter estimation of Fiber Bragg grating given the reflection power spectra as input. Fiber Bragg Grating is a type of distributed Bragg reflector built in the core of an optical fiber to selectively reflect certain wavelengths of light and transmit the rest. This effect is achieved by periodically varying the refractive index along the segment as shown in the figure below.



Full equation for the reflected power ($P_B(\lambda)$) is

$$P_B(\lambda) = \frac{\sinh^2\left[\eta(V)\delta n_0 \sqrt{1-\Gamma^2}\,\frac{N\Lambda}{\lambda}\right]}{\cosh^2\left[\eta(V)\delta n_0 \sqrt{1-\Gamma^2}\,\frac{N\Lambda}{\lambda}\right] - \Gamma^2}$$

where,

$$\Gamma(\lambda) = \frac{1}{\eta(V)\delta n_0}\left[\frac{\lambda}{\lambda_B} - 1\right]$$

Here ,

η is the fraction of power in the core,

δn_0 is the variation in refractive index (n_3-n_2),

N is the number of grating,

Λ is grating period,

λ_B is the reflected wavelength,

n_e is the effective refractive index of the grating in the fiber core.

Model Architecture and Training

Here the parameters η, δn_0, Λ , λ_B , N from the reflected power relation act as design variables and reflected power function sampled at 10 nm acts as Target response. We train models to predict the design variables given target response. Tandem network is trained in 2 stages. (i) first train the forward neural network and (ii) train the backward Neural network by stacking the freezed forward model at the output layer as shown in the Figure on the right.

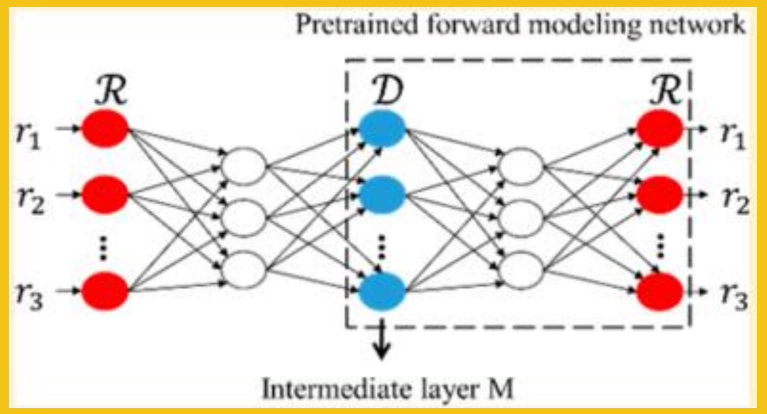


Figure showing the architecture of the tandem model

Tandem Forward model architecture

Input units : 5
Hidden layer 1 : 256 dense units with relu activation + batch normalization
Hidden layer 2 : 256 dense units with relu activation + batch normalization
Hidden layer 3 : 256 dense units with relu activation + batch normalization
Hidden layer 4 : 256 dense units with relu activation + batch normalization
Output layer : 300 dense units with sigmoid activation

Tandem Backward model architecture

Input units : 300
Hidden layer 1 : 64 dense units with relu activation + batch normalization
Hidden layer 2 : 64 dense units with relu activation + batch normalization
Hidden layer 3 : 64 dense units with relu activation along with batch normalization
Hidden layer 4 : 32 dense units with relu activation along with batch normalization
Output layer : 5 dense units with sigmoid activation

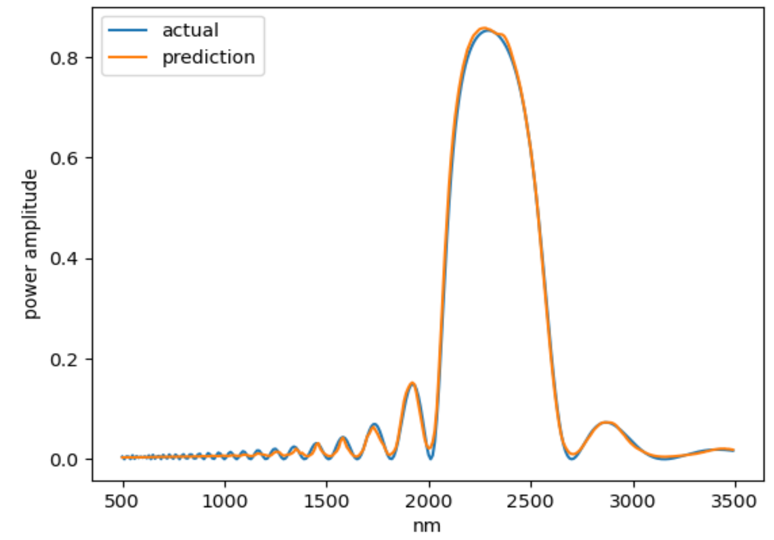
Direct inverse model architecture

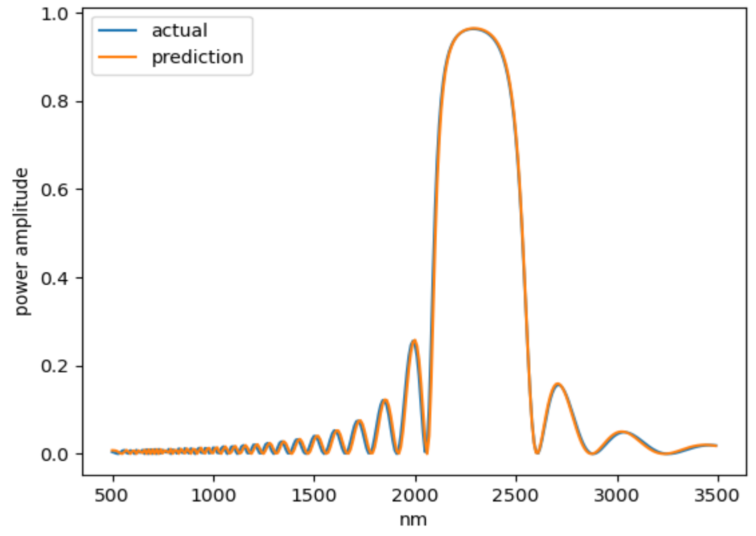
Input units : 300
Hidden layer 1 : 64 dense units with relu activation + batch normalization
Hidden layer 2 : 64 dense units with relu activation + batch normalization
Hidden layer 3 : 64 dense units with relu activation along with batch normalization
Hidden layer 4 : 32 dense units with relu activation along with batch normalization
Output layer : 5 dense units with sigmoid activation

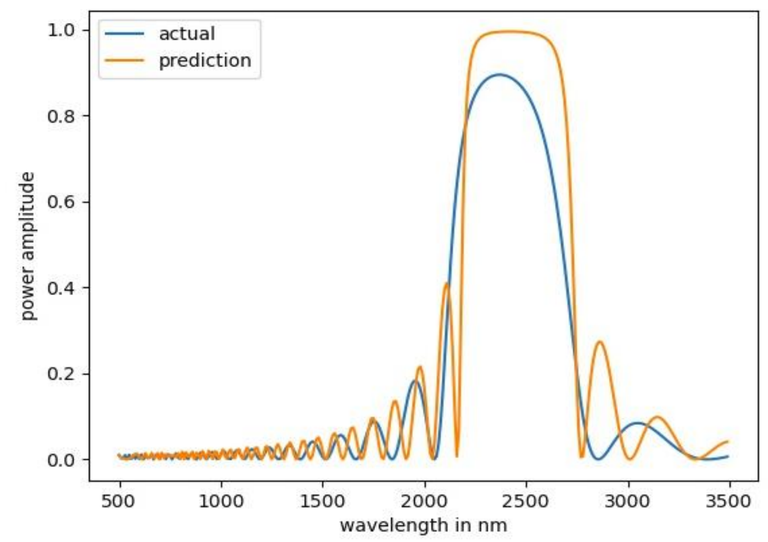
Here the forward model is trained until a reasonable test set accuracy is obtained, later the tandem model is built by stacking the untrained backward model and trained forward model. Tandem model is trained on the inverse dataset and with forward model weights set to non-trainable. To compare the model accuracy with the direct inverse prediction models we build a direct inverse model with the architecture same as backward model of tandem and train it with the same data, once trained we perform vs prediction on the test set and estimate the performance.

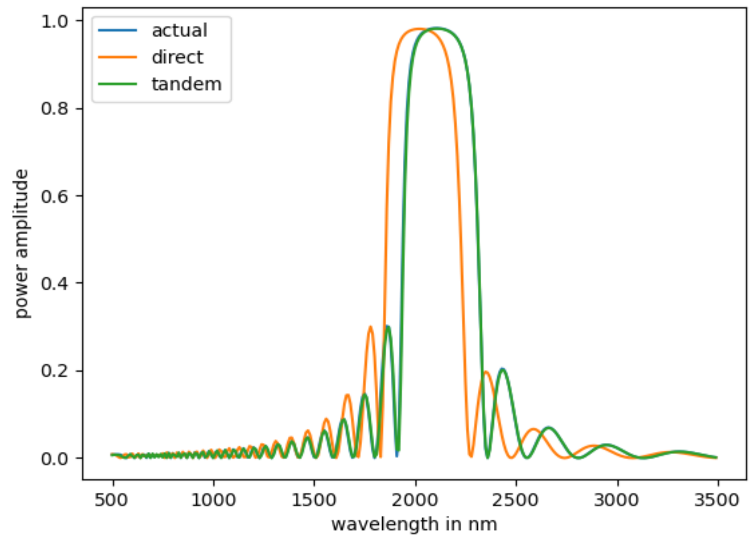
Results and findings

- By comparing the test set performances of the tandem architecture based model and direct inverse model, we find that the tandem based models perform very well compared to the direct models. In this task, we achieved a 100 times greater accuracy with the tandem model compared to the direct model.
- Below are plots of test set performances of forward model (top left), backward model (top right), direct model (bottom left), tandem architecture-based model vs direct model (bottom right).









Conclusion

- As discussed via Fiber Bragg grating modelling task in this project, tandem architecture-based models can be used to overcome the non-uniqueness problem in the neural networks and often give a good boost in accuracy for the inverse prediction tasks.
- The above model utilizing the tandem architecture achieves at least a 100 times less loss compared to the model without tandem architecture.

References

[1] **Fathy M. Mustafa, Mofreh Toba and Tamer M. Barakat** , "New Simulation of Fiber Bragg Grating: Maximum Reflectivity and Narrow Bandwidth without Side Lobes", International Journal of Applied Engineering Research, vol. 14, Number 11 (2019) pp. 2667-2674

[2] **Silvia Ferrari and Robert F. Stengel,** "Smooth Function Approximation using Neural Networks", IEEE Transactions on Neural Networks, vol. 16, no. 1, Jan 2005

[3] **Dianjing Liu, Yixuan Tan, Erfan Khoram, and Zongfu Yu***, "Training deep neural networks for the inverse design of nanophotonic structures", ACS Photonics 2018, S, 1365-1369