DeepBO: Deep Neural-Network Bayesian Optimization of Polaritonic Metasurfaces in Continuous Space

Zihan Zhang
Zhiyuan College
School of Mathematical Science
Shanghai Jiao Tong University

Jintao Chen
Zhiyuan College
School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University

Kehang Cui
School of Material Science and Engineering
Shanghai Jiao Tong University
cuikehang@sjtu.edu.cn

Abstract

Thermophotovoltaics (TPVs) rely on selective thermal emitters to tailor the black-body radiation at high temperatures into band-matching emission for photovoltaic cells, resulting in power-conversion efficiencies surpassing the Shockley-Queisser limit. The selectivity of the thermal emitter must cover three orders of magnitude range of wavelengths, spreading from visible to the far infrared, which requires the superposition of multiple transformation theories of optics and degrees of freedom anisotropic geometries. It is extremely challenging to realize such high-dimensional complex metasurface design using conventional computational photonics. Here we develop a deep neural network-based Bayesian optimization (DeepBO) framework to screen a 16-dimensional design space of 10^{43} candidates, and realize a record-high spectral efficiency of 69% for the TPV emitter. We show that the neural network combined with Bayesian linear regression is an efficient and robust surrogate model which scales linearly with the size of data. We also reveal the underlying physical mechanisms of the geometric design of the TPV emitters using principal component analysis (PCA). We anticipate the DeepBO framework is a useful tool for data-intensive complex geometric design for photonics research community.

1 Introduction

Metasurfaces are sub-wavelength metal/dielectric structures that resonantly couple to the incoming electromagnetic fields [1]. The amplitudes, phases, and polarization states of the photons can be tailored in a customizable manner, enabling applications in quantum-optical processors, energy harvesting, sensing, communications, *etc*. The most critical feature of the metasurfaces is the high degrees of freedom in geometric inhomogeneity. However, the high-dimensional design space poses significant challenges for the structural design.

Integrating deep learning framework [2] with Maxwell solver is an effective method for the photonic design of the high-dimension metasurfaces. Intensive research has been done in this field. However, there are still two main issues. The first issue is the discretization of the design space. The design space is usually formulated into discretized unit pixels or voxels with fixed geometries so that they can be used as input (labels) for neural networks. This would degenerate the search space into a slice of the free-form continuous design space. The second issue is the lack of underlying physical mechanisms in deep neural networks. The generative neural network often produces fancy but unrealistic patterns which are not manufacturable. It is urgent to develop physics-informed framework for photonic design in continuous space.

Fourth Workshop on Machine Learning and the Physical Sciences (NeurIPS 2021).

Herein we realize a Deep neural network-based Bayesian Optimization (DeepBO) framework to design a thermal emitter in the continuous high-dimensional space to modulate the electromagnetic radiation for thermophotovoltaic (TPV) energy conversion. More importantly, through primary component analysis (PCA), we identify the underlying physics of the spectral characteristic control of the complex metasurfaces.

2 The DeepBO Framework

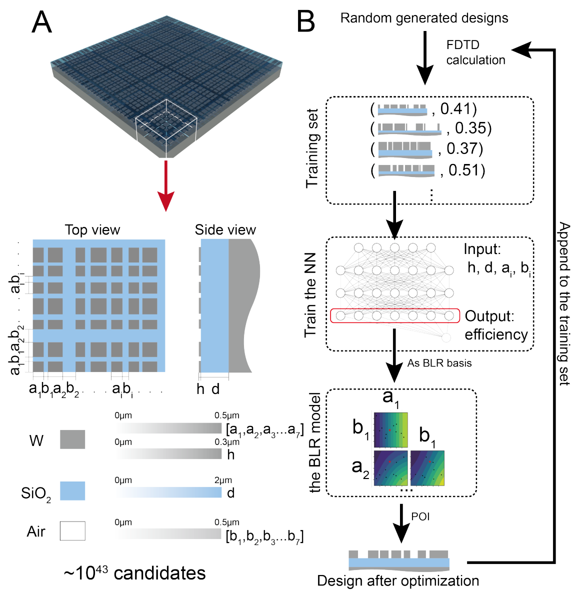

Figure 1: Design pipeline of the DeepBO. (A) The schematics of the thermal emitter and the 16-dimensional feature space. (B) Framework of DeepBO. The deep neural network is connected to a Bayesian linear regression, which gives the predicted distribution of the TPV efficiency in the feature space.

2.1 Design space

Our design of the thermal emitter, as illustrated in Figure 1A, is a three-layer periodic grating structure transferred from 3-d to 2-d and described by the feature vector $[h, d, a_1, b_1, a_2, ...]$, which lists all the parameters of each possible candidate in a 2-d unit cell. This simplifies a three-dimensional case to a two-dimensional one by using the symmetries of two orthogonal polarizations. Parameter $a_i (i = 1, 2, ..., 7)$, ranging from 0 μm to 0.5 μm continuously, denotes the width of each tungsten strip, while $b_i (i = 1, 2, ..., 7)$ with the same range is the length of the gap separating the neighboring tungsten strips, forming the first layer: a superposition of multi-fold periodic tungsten gratings, and the upper limit of the superposition is 7. The height of the tungsten strip is denoted by h, which ranges from 0 μm to 0.3 μm continuously. Parameter d denotes the thickness of SiO_2 layer ranging from 0 μm to 3 μm continuously, which composes the second layer under the grating structure. The bottom layer is an opaque tungsten substrate. If we regard 1 nm as the discrete unit of the ranges, a design space with $500^{14} * 300 * 2000 \approx 10^{43}$ candidates in total is formed.

2.2 Pipeline

The figure of merit (FOM) is defined as the TPV efficiency:

$$\eta = \frac{\int_0^{\lambda_{\text{band gap}}} bb(\lambda)\epsilon(\lambda)\lambda/\lambda_{\text{band gap}} d\lambda}{\int_0^{\infty} bb(\lambda) d\lambda},$$

where $\lambda_{\text{band gap}}$ is the wavelength corresponding to the band gap of the PV, $\epsilon(\lambda)$ is the emissivity of the design at the wavelength λ, and $bb(\lambda)$ is the intensity of the blackbody radiation at the working temperature of the TPV. In our case, $\lambda_{\text{band gap}} = 2.254$ nm, $\epsilon(\lambda)$ is calculated by FDTD simulation, and the working temperature is 1673K.

The DeepBO framework (Figure 1B) combines a deep neural network (NN) with the Bayesian linear regression (BLR) as the surrogate model. The NN, which takes the feature vector $[a_1, b_1, a_2, ..., h, d]$ of a design as input and the predicted TPV efficiency as output, is trained by a randomly probed set as *a priori*. The BLR performs a probabilistic linear regression based on the trained NN, and predicts *a posterior* distribution of the efficiency to the design space [3]. The POI algorithm is used to balance exploration (the highest uncertainty) and exploitation (the highest predicted value). We probe the predicted optimal design using FDTD Maxwell solver. The probed mapping between the design feature vector and FOM is appended to the training set *a priori*. The procedure loops until a satisfied design is eventually found.

3 Result

Figure 2A shows the history of the maximum efficiency with respect to the number of optimization loops. The spectral efficiency of the optimal structure is 0.69, which, to our best knowledge, is the highest spectral efficiency for the TPV emitters [4] [5]. The whole optimization process takes *ca.* 50 hours on a moderate desktop workstation with 56-Core 3.2 GHz CPU and 128 GB memory. We terminate the optimization process after 1400 loops as the maximum efficiency plateaus.

Figure 2B shows the comparison between the asymptotic time costs of DeepBO and Gaussian process (GP)-based BO from sklearn. It can be clearly observed that the time for each loop in DeepBO increases linearly with the training loop, while the time for each loop in GP-based BO increases cubically as GP has a computational complexity up to $O(n^3)$, where n is the number of known data points. The surrogate model in DeepBO consists of NN and BLR. The NN is used to perform the regression task, then the BLR is applied on the last hidden layer of the network to obtain a distribution of the regression result. As a reduced version of GP, the BLR is much less expressive than GP.

Moreover, the NN-BLR surrogate model in DeepBO offers higher accuracy and efficacy than GP. Figure 2C and 2D show the predicted mean value and distribution given by GP and NN-BLR along SiO_2 thickness, from the training set (Figure 2C) and the first 5 probed points (Figure 2D). The well trained NN in DeepBO fits the data more accurately, and even in the worst case the regression accuracy of NN in DeepBO will not lose out to GP. Compared with conventional deep neural network-based surrogate models, NN-BLR outputs the distribution of the regression results rather than a fixed value, which could provide more possibilities for exploration when sampling. The speed, accuracy and efficacy of NN-BLR makes DeepBO a powerful framwork for data-intensive geometric design.

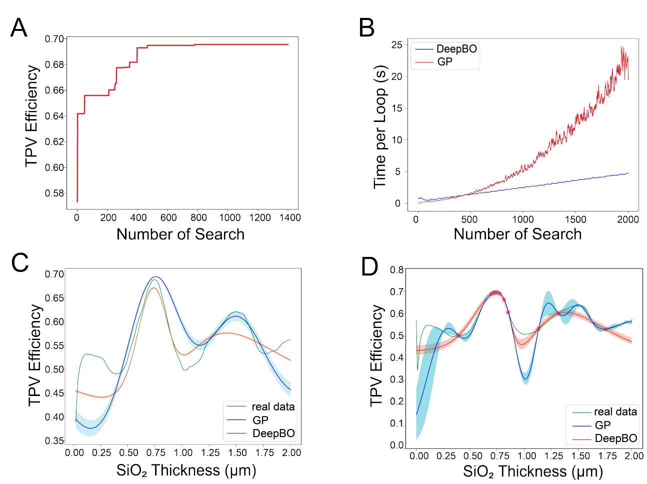

Figure 2: The performance of DeepBO. (A) Optimization history of the TPV efficiency. (B) Time per loop for DeepBO and GP-based BO. (C) Regression comparison of the trained surrogate model in GP-based BO and DeepBO. (D) The first five predicted probing points and their distribution given by GP-based BO and DeepBO.

4 Underlying Physical Mechanism

We further perform principal component analysis (PCA) on the 16-dimensional design space to reveal the physical mechanisms of the optimal structure which holds the record of the highest spectral efficiency of TPV emitters so far.

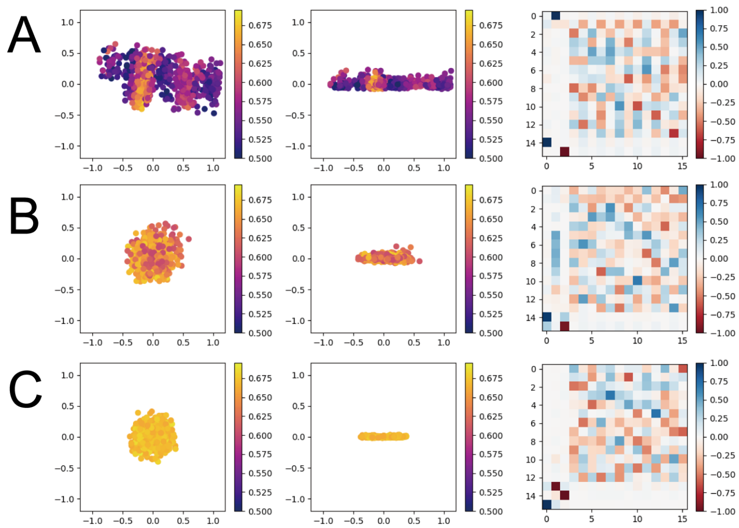

Figure 3: The PCA results on the search data set.

Since our model is periodic, a cyclic permutation on $\{a_i\}$ and $\{b_i\}$ gives exactly the same design. To make sure the parameters in PCA are physically meaningful, the feature vectors are cyclically permuted such that the largest a_i is in the position of a_1 before the PCA, while the positions of h, d remain the same. The regulated feature vector is denoted by $[h, d, a'_1, b'_1, a'_2, \ldots]$.

To explain the dynamic learning procedure, PCA is performed on the search set filtered by different thresholds of efficiency, as shown in Figure 3. The thresholds are: (A) efficiency > 0.50, 94% of training set, (B) efficiency > 0.60, 61% of training set, (C) efficiency > 0.66, 30% of training set. The first column in Figure 3 shows the projection of the filtered sets on the plane spanned by the first (x-axis) and second (y-axis) principal components, while the second column shows the projection on the plane spanned by the the first (x-axis) and least (y-axis) components. The third column shows the PCA components of each chosen data set, each row symbolizes a PCA component, the i-th row corresponding to the i-th principal component.

When the PCA is performed on the whole search set (Figure 3A), the data points distribute in a broad range, which demonstrates sufficient exploration over the entire design space. According to the range set of the parameters, if the data set is chosen uniformly, the variance in the d direction will be the largest, while the variance in a'_1 and b'_1 should be the least. The result of PCA shows the component with the largest variance is the direction of d, while the components with the least two variances is the direction of a'_1 and h. The order of variances calculated by the range setting agrees with the order of the principal components. The data set is composed with a wide range of low-efficiency points and a cluster of high-efficiency points, demonstrating a balance between exploration and exploitation.

When the efficiency threshold is higher, the variance in the d direction becomes lower. Meanwhile, the variance along a'_1 succeeds h to be the least component (Figure 3B). When the efficiency threshold is further increased to 0.66 (Figure 3C), the least three components are a'_1, h and d, *i.e.*, the distribution of high-efficiency points along this directions is narrow, indicating a'_1, h and d are key parameters of our model. Therefore, the performance of the TPV emitter is the most sensitive to the width and height of the largest tungsten strip, which corresponds to the surface plasmon polaritons, as well as the thickness of the SiO_2 layer, which corresponds to magnetic polaritons.

5 Conclusion

DeepBO innovatively combines the two state-of-the-art tools of black-box function fitting and optimization together, to achieve both accuracy and calculation efficiency in finding an optimal design of complicated material systems. The most suitable situation to use DeepBO is when the design space is extremely data-intensive. In this work, DeepBO proves its power by achieving a record design of TPV emitter, providing a new way of thinking to the global optimization of physical devices.

6 Acknowledgement

This work was financially supported by the Zhiyuan Scholar Program of Shanghai Jiao Tong University (SJTU) and Shanghai Pujiang Program. This work made use of shared facilities of the Zhiyuan Innovation Research Center (ZIRC) at SJTU.

References

[1] Wei Ting Chen, Alexander Y. Zhu, and Federico Capasso. Flat optics with dispersion-engineered metasurfaces. *Nature Reviews Materials*, 5:604–620, 2020.

[2] Wei Ma, Zhaocheng Liu, Zhaxylyk A. Kudyshev, Alexandra Boltasseva, Wenshan Cai, and Yongmin Liu. Deep learning for the design of photonic structures. *Nature Photonics*, 15:77–90, 2021.

[3] Jasper Snoek, Oren Rippel, Kevin Swersky, Ryan Kiros, Nadathur Satish, Narayanan Sundaram, Mostofa Patwary, Mr Prabhat, and Ryan Adams. Scalable bayesian optimization using deep neural networks. In *International Conference on Machine Learning*, pages 2171–2180. PMLR, 2015.

[4] Tobias Burger, Caroline Sempere, Bosun Roy-Layinde, and Andrej Lenert. Present efficiencies and future opportunities in thermophotovoltaics. *Joule*, 4(8):1660–1680, 8 2020.

[5] Reyu Sakakibara, Veronika Stelmakh, Walker R. Chan, Michael Ghebrebrhan, John D. Joannopoulos, Marin Soljacic, and Ivan elanovi. Practical emitters for thermophotovoltaics: a review. *Journal of Photonics for Energy*, 9(3):32713, 2019.