OptoGPT: A Versatile Inverse Design Model for Optical Multilayer Thin Film Structures

Taigao Ma
Department of Physics
University of Michigan, Ann Arbor
taigaom@umich.edu

Haozhu Wang
Department of EECS
University of Michigan, Ann Arbor
hzwang@umich.edu

L. Jay Guo*
Department of EECS
University of Michigan, Ann Arbor
guo@umich.edu

Abstract

Optical multilayer thin film structures are widely used in various photonic applications. Inverse design is an important but difficult step to enable these applications, which seeks to find out the best structure (material & thickness arrangements) given a target optical response. Recently, deep learning-based methods have been developed to solve the inverse design efficiently. However, existing methods usually fix the material arrangements and only design the thickness, which is not versatile for a different material arrangement and may lead to sub-optimal performance. In this study, we resolve this issue by treating the structure as a sequence and using structure tokens to represent the material and thickness simultaneously. Later on, the inverse design problem can be formulated as a common sequence generation task conditioned on the input optical responses. Based on this, we propose **OptoGPT** to act as a versatile inverse design model that can design material and thickness simultaneously, significantly expanding the design capability. In addition, using probability resampling further provides a versatile method to satisfy fabrication and design requirements in practical applications.

1 Introduction

Optical multilayer thin film structures (shorten as "multilayer structures") are staked thin film layers (typical thickness around several hundred nanometers) with different materials at each layer. Because of the ease of fabrication, they are widely used in optical and energy applications, including structural colors[1, 2, 3], filters[4], photovoltaics[5, 6], display devices[7, 8], and radiative cooling[9, 10].

When light goes through a multilayer structure with different combinations of materials and thickness, photons will interact with these layered materials, leading to different optical responses. The forward prediction process from a given structure to its corresponding optical response is easy to calculate, either through accurate physics-based solvers[11, 12], or machine learning-based surrogate models[13, 14]. However, the inverse design process[15, 16, 17], which seeks to find out the best structure according to a desired optical response, is nontrivial as it is challenging to reveal the complicated inverse correlation.

Recently, researchers have started to use machine learning-based methods to facilitate the inverse design process, including tandem networks[18], GAN[19, 20], MDN[21], etc. After training, these neural networks are capable of capturing the intricate inverse relationship from the space of optical

NeurIPS 2023 Workshop on Deep Learning and Inverse Problems

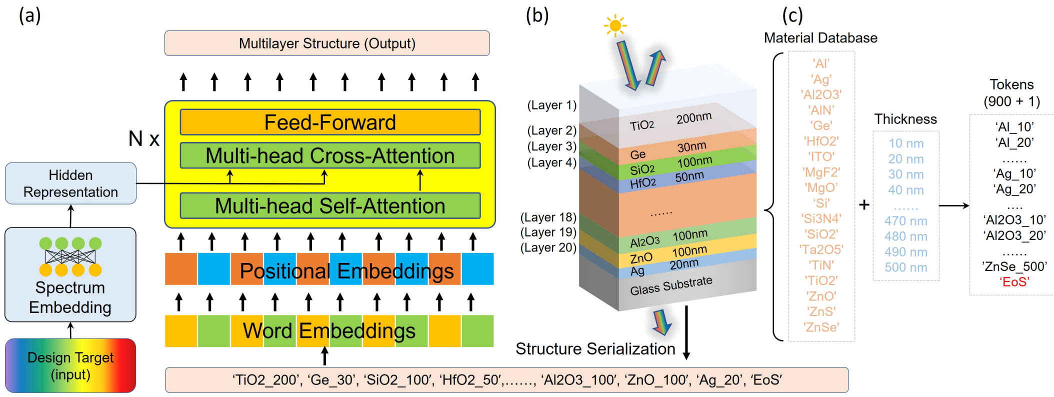

Figure 1: (a) Model Architecture of OptoGPT. (b) An example of the optical multilayer thin film structure as well as its structure serialization. (c) A diagram of forming structure tokens by concatenating the material and thickness at each layer.

responses to the space of structures and outputting designed structures directly when given a desired optical target. Although fast, existing inverse design models simplify the inverse design problem as thickness-only design while ignoring the material selections, e.g., the three-layer structure of $Ag/SiO_2/Ag$ in [20] and the six-layer structure of $MgF_2/SiO_2/Al_2O_2/TiO_2/Si/Ge$ in [22]. Therefore, these models only work when designing for specific types of multilayer structures, limiting their capabilities in versatile applications.

In this work, we propose the OptoGPT as a versatile inverse design model by leveraging the strong learning and generalization abilities of transformer. First, we reformulate the multilayer structure as a sequence and propose the structure token to denote the material and thickness at each layer. By doing so, the task of multilayer structure inverse design is equivalent to a sequence generation task conditioned on the input, e.g., the optical response. Our OptoGPT is a decoder-only transformer that auto-regressively generates the structure sequence when given a design target. A series of experiments demonstrate that our model can not only output versatile structures with different material combinations, but can also design structures that satisfy practical constraints during fabrication and manufacture.

2 Related Works

There have been many works that use deep learning-based models to tackle the inverse design of different types of photonic structures, including diffractive metagratings[23], metamaterials[24, 25], free-form metasurfaces[26, 27], etc. These models have demonstrated excellent inverse design performance in multiple photonic applications, including absorbers[22], reflectors[28], spectrometer[29], transmissive antenna[30], etc. In addition to inverse design, they have also been widely used as a fast surrogate simulator to speed up the time-consuming forward simulation and prediction process in scientific applications [31, 32, 33]. From another perspective, many research groups have already started to combine the transformer with their domain-specific scientific problems and proposed different ML models to solve complicated scientific discovery and engineering applications, including material design[34], chemistry[35], biology[36, 37, 38], etc. Our work serves as an exploration of using transformer to solve the inverse design in multilayer thin film structures.

3 Methods

3.1 Problem Set

For a given N-layer multilayer structure (see Fig. 1b), we denote the material arrangement as $m = [m_1, m_2, \ldots, m_N]$, and the thickness sequence as $t = [t_1, t_2, \ldots, t_N]$, where m_i, t_i refers to the material and thickness at the i_{th} layer, respectively. $m_i \in \mathbf{M}$ is a discrete variable that can take several distinct values from the material database $\mathbf{M}$. Then, we can denote a multilayer structure as

Table 1: Important parameters of OptoGPT

NAME	PARAMETER
EMBEDDING DIMENSION	1024
NUMBER OF DECODER BLOCKS (N)	6
NUMBER OF ATTENTION HEAD (A)	8
CONTEXT LENGTH (K)	21
NUMBER OF TOKENS	901
SPECTRA DIMENSION	142
BATCH SIZE	1000
OPTIMIZER	ADAM OPTIMIZER
TRAINING EPOCHS	$\sim$200
NUMBER OF PARAMETERS	58 M

$x = \{m, t\} \in \mathcal{X}$. Physics laws guarantee that there is a deterministic function $f(\cdot)$ that maps from the space of multilayer structure x to the d-dimensional optical properties $y = [y_1, y_2, \ldots, y_d] \in \mathcal{Y}$. This forward process can be solved using physical simulators[11, 12] or approximated by surrogate simulators[13, 14]. Differently, the inverse design problem focuses on solving the inverse mapping of $f^{-1}(\cdot)$, which seeks to identify a suitable x in the space of $\mathcal{X}$ such that its corresponding response $y = f(x)$ is close enough to the target response y^*. In this work, we consider designing the transmission and reflection spectra from 400 nm to 1100 nm. Both spectra are discretized by 10 nm, making d=2×71=142.

3.2 Structure Serialization

Existing inverse design models post strong restrictions on the design space of multilayer structures $\mathcal{X}$, e.g., setting N to a specific number and fixing the material arrangement m. Therefore, they only search for the thickness t in the limited design subspace $\mathcal{X}' \in \mathcal{X}$, which limits their applications. It would be more versatile if these models could also determine the total number of layers N and material combinations m automatically.

In order to achieve this, we start by reformulating the multilayer structure $x = \{m, t\} = \{[m_1, m_2, \ldots, m_N], [t_1, t_2, \ldots, t_N]\}$ as a sequence of $x = \{[m_1, t_1], [m_2, t_2], \ldots, [m_N, t_N]\}$. Then we propose a technique called structure serialization (see Fig. 1b-c), where structure tokens are used to represent the material and thickness information $[m_i, t_i]$ at each layer simultaneously, similar to how Natural Language Processing (NLP) researchers tokenize language sentences. By appending multiple tokens one-by-one, we can convert a multilayer structure into a sequence of tokens that a common sequence model can deal with. We also use a special token of 'EoS' (end of sequence) to enable the design of structures with different numbers of layers (we set the maximum to be 20 layers).

There are 18 different types of materials in our material database **M** (see Fig. 1c). In addition, the designed thickness ranges from 10 nm to 500 nm and we discretize it by 10 nm. Therefore, there are 18×50+1=901 tokens in our vocabulary. The total number of different types of material arrangements that our model can design expands to $18^{20} \sim 10^{25}$, significantly expanding our model's capability. This method is scalable and can be used to include other materials and extend to more layers.

3.3 Model Architecture

Since the multilayer structure is formulated as a sequence, we can treat the inverse design as a sequence generation task conditioned on the input of the design target. Therefore, we choose the decoder in the transformer as our model architecture (see Fig. 1a), similar to how GPT generates the language sequence. Here, spectrum embedding is a fully connected layer that maps the target optical response into the same embedding space of structure tokens. We generate a large training dataset with 10M samples, where each sample is a pair of randomly generated structures and the corresponding spectra simulated by Transfer Matrix Methods (TMM)[11]. The dataset generation takes $\sim$1200h using a single CPU. The dimension of the embedding is 1024 and the number of decoder blocks is 6, making the model size $\sim$58M. We train the model based on next-word prediction for 200 epochs by setting the batch size 1000, which takes about two weeks on a single NVIDIA 3090 GPU. Table 1 lists some important parameters.

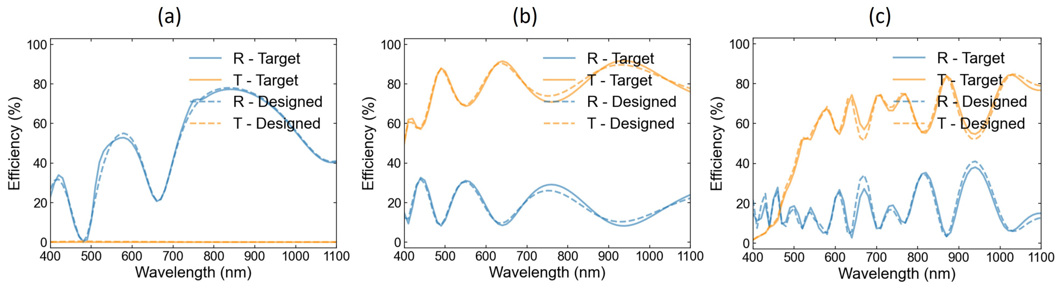

Figure 2: (a-c) Two examples of inverse design for transmission (T) and reflection (R) spectra. Real lines and dashed lines are the target spectra and simulated spectra of the designed structure, respectively.

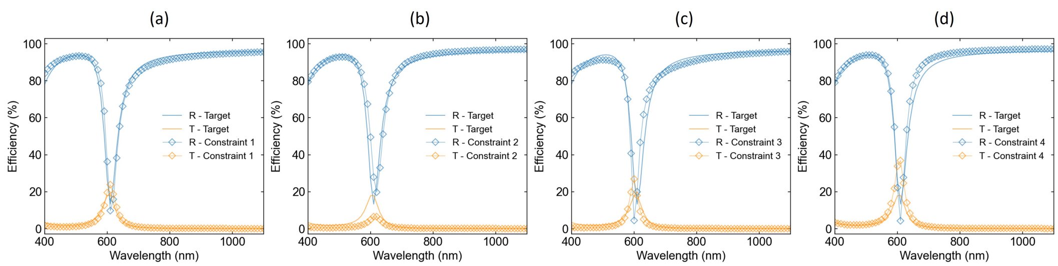

Figure 3: (a-d) Examples of inverse design reflection and transmission spectra with different constraints. Real lines and real lines with squares are the target spectra and simulated spectra of the designed structure, respectively.

4 Experiments

After training, we examine the inverse design performance by designing for transmission and reflection spectra that our model has never seen before. Once the model outputs a structure sequence, we convert it back into a multilayer structure and use TMM to simulate the spectra. The Mean Square Error (MSE) between the target spectra and the simulated spectra of the designed structure is used to evaluate the design accuracy. The averaged MSE on 1000 spectra is as low as 2×10^{-3}. In Fig. 2, we also give three examples to visualize the difference between the target spectra (real lines) and the designed spectra (dashed lines). We want to mention that the designed structures come from our model's output directly, although running an iterative optimization on thickness can further improve the accuracy.

To further demonstrate the versatility of our model, we use probability sampling to align the designed structures with practical constraints. For example, researchers or engineers may impose restrictions on the material selection or thickness range at any layers specific to the fabrication or design needs. This is done by removing these structure tokens that do not satisfy constraints from probability sampling during the auto-regressive generation process. As an example, we use OptoGPT to inverse design a Fabry-Parot resonator[39]. Here, the target spectra has a resonance absorption at 610 nm, leading to a dip in the reflection spectra (see Fig. 3). Now we consider adding four different constraints separately:

1. Fix the first layer to be 100 nm SiO_2
2. Remove Ag in the third layer
3. Limit the thickness of the first layer inside [10, 150] nm and remove Ag/Al in the first layer
4. Specify the material arrangement to be a three-layer $Ag/Si_3N_4/Ag$ structure and design the thickness only

The first constraint can be used when requiring a dielectric layer at the air interface to protect structures from oxidization when exposed to air, while the second constraint is practical when looking for an alternative to replace silver, considering silver is a precious metal. For the third constraint, we use it as a general example of adding thickness and material restrictions simultaneously. The fourth constraint specifies the number of layers as well as the material arrangements to be designed, which is a special case that converts our model into a thickness-only design model. Through probability resampling, our model can finish designs that satisfy these constraints. We simulate the spectra of these designed structures and compare them with the target spectra in Fig. 3a-d. We can find that the designed spectra are close to the target spectra, demonstrating that our model is versatile to different constraints required in practical applications, e.g., fabrication and manufacture.

5 Conclusion

In this work, we propose a new machine learning-based method to solve the inverse design problem in optical multilayer thin film structures. By formulating the multilayer structure as a sequence and using the structure token to represent the layered information, we convert the inverse design problem as a conditional sequence generation problem that can be well solved using transformer. Benefiting from this, our OptoGPT is capable of automatically determining the total number of layers and designing the material and thickness simultaneously. This differentiates our model from existing models, which usually post strong restrictions on the designed structures, and makes our model versatile to different applications, significantly expanding the model's capability in designing structures that satisfy practical constraints required in fabrication and manufacture. Our work expands the existing transformer applications from mainstream NLP and CV to optical science and demonstrates that transformer architecture is also an effective learner for photonic inverse design.

6 Acknowledgement

We thank the National Science Foundation (NSF PFI-008513 and NSF 2309403) for the support of this work.

References

[1] Zhengmei Yang, Chengang Ji, Dong Liu, and L Jay Guo. Enhancing the purity of reflective structural colors with ultrathin bilayer media as effective ideal absorbers. *Advanced Optical Materials*, 7(21):1900739, 2019.

[2] Danyan Wang, Zeyang Liu, Haozhu Wang, Moxin Li, L Jay Guo, and Cheng Zhang. Structural color generation: from layered thin films to optical metasurfaces. *Nanophotonics*, 12(6):1019–1081, 2023.

[3] Anwesha Saha, Taigao Ma, Haozhu Wang, and L Jay Guo. Environmentally sustainable and multifunctional chrome-like coatings having no chromium designed with reinforcement learning. *ACS Applied Materials & Interfaces*, 2023.

[4] Chengang Ji, Chenying Yang, Weidong Shen, Kyu-Tae Lee, Yueguang Zhang, Xu Liu, and L Jay Guo. Decorative near-infrared transmission filters featuring high-efficiency and angular-insensitivity employing 1d photonic crystals. *Nano Research*, 12:543–548, 2019.

[5] Mingzhen Liu, Michael B Johnston, and Henry J Snaith. Efficient planar heterojunction perovskite solar cells by vapour deposition. *Nature*, 501(7467):395–398, 2013.

[6] Chuantian Zuo, Henk J Bolink, Hongwei Han, Jinsong Huang, David Cahen, and Liming Ding. Advances in perovskite solar cells. *Advanced Science*, 3(7):1500324, 2016.

[7] Marc A Baldo, David F O'Brien, Y You, A Shoustikov, S Sibley, Mark E Thompson, and Stephen R Forrest. Highly efficient phosphorescent emission from organic electroluminescent devices. *Nature*, 395(6698):151–154, 1998.

[8] Hiroki Uoyama, Kenichi Goushi, Katsuyuki Shizu, Hiroko Nomura, and Chihaya Adachi. Highly efficient organic light-emitting diodes from delayed fluorescence. *Nature*, 492(7428):234–238, 2012.

[9] Aaswath P Raman, Marc Abou Anoma, Linxiao Zhu, Eden Rephaeli, and Shanhui Fan. Passive radiative cooling below ambient air temperature under direct sunlight. *Nature*, 515(7528):540–544, 2014.

[10] Shancheng Wang, Tengyao Jiang, Yun Meng, Ronggui Yang, Gang Tan, and Yi Long. Scalable thermochromic smart windows with passive radiative cooling regulation. *Science*, 374(6574):1501–1504, 2021.

[11] Steven J Byrnes. Multilayer optical calculations. *arXiv preprint arXiv:1603.02720*, 2016.

[12] Jean Paul Hugonin and Philippe Lalanne. Reticolo software for grating analysis. *arXiv preprint arXiv:2101.00901*, 2021.

[13] Taigao Ma, Haozhu Wang, and L Jay Guo. Ol-transformer: A fast and universal surrogate simulator for optical multilayer thin film structures. *arXiv preprint arXiv:2305.11984*, 2023.

[14] Yang Deng, Juncheng Dong, Simiao Ren, Omar Khatib, Mohammadreza Soltani, Vahid Tarokh, Willie Padilla, and Jordan Malof. Benchmarking data-driven surrogate simulators for artificial electromagnetic materials. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.

[15] Jiaqi Jiang, Mingkun Chen, and Jonathan A Fan. Deep neural networks for the evaluation and design of photonic devices. *Nature Reviews Materials*, 6(8):679–700, 2021.

[16] Taigao Ma, Mustafa Tobah, Haozhu Wang, and L Jay Guo. Benchmarking deep learning-based models on nanophotonic inverse design problems. *Opto-Electronic Science*, 1(1):210012–1, 2022.

[17] Simiao Ren, Ashwin Mahendra, Omar Khatib, Yang Deng, Willie J Padilla, and Jordan M Malof. Inverse deep learning methods and benchmarks for artificial electromagnetic material design. *Nanoscale*, 14(10):3958–3969, 2022.

[18] Dianjing Liu, Yixuan Tan, Erfan Khoram, and Zongfu Yu. Training deep neural networks for the inverse design of nanophotonic structures. *Acs Photonics*, 5(4):1365–1369, 2018.

[19] Sunae So and Junsuk Rho. Designing nanophotonic structures using conditional deep convolutional generative adversarial networks. *Nanophotonics*, 8(7):1255–1261, 2019.

[20] Peng Dai, Kai Sun, Xingzhao Yan, Otto L Muskens, CH de Groot, Xupeng Zhu, Yueqiang Hu, Huigao Duan, and Ruomeng Huang. Inverse design of structural color: finding multiple solutions via conditional generative adversarial networks. *Nanophotonics*, 11(13):3057–3069, 2022.

[21] Rohit Unni, Kan Yao, and Yuebing Zheng. Deep convolutional mixture density network for inverse design of layered photonic structures. *ACS Photonics*, 7(10):2703–2712, 2020.

[22] Wei Chen, Yuan Gao, Yuyang Li, Yiming Yan, Jun-Yu Ou, Wenzhuang Ma, and Jinfeng Zhu. Broadband solar metamaterial absorbers empowered by transformer-based deep learning. *Advanced Science*, 10(13):2206718, 2023.

[23] Jiaqi Jiang, David Sell, Stephan Hoyer, Jason Hickey, Jianji Yang, and Jonathan A Fan. Free-form diffractive metagrating design based on generative adversarial networks. *ACS nano*, 13(8):8872–8878, 2019.

[24] Wei Ma, Feng Cheng, and Yongmin Liu. Deep-learning-enabled on-demand design of chiral metamaterials. *ACS nano*, 12(6):6326–6334, 2018.

[25] Che Wang, Yuhao Fu, Ke Deng, and Chunlin Ji. Functional response conditional variational auto-encoders for inverse design of metamaterials. In *NeurIPS 2021 Workshop on Deep Learning and Inverse Problems*, 2021.

[26] Jiaqi Jiang and Jonathan A Fan. Simulator-based training of generative neural networks for the inverse design of metasurfaces. *Nanophotonics*, 9(5):1059–1069, 2019.

[27] Fufang Wen, Jiaqi Jiang, and Jonathan A Fan. Robust freeform metasurface design based on progressively growing generative networks. *ACS Photonics*, 7(8):2098–2104, 2020.

[28] Rohit Unni, Kan Yao, Xizewen Han, Mingyuan Zhou, and Yuebing Zheng. A mixture-density-based tandem optimization network for on-demand inverse design of thin-film high reflectors. *Nanophotonics*, 10(16):4057–4065, 2021.

[29] Junren Wen, Lingyun Hao, Cheng Gao, Hailan Wang, Kun Mo, Wenjia Yuan, Xiao Chen, Yusi Wang, Yueguang Zhang, Yuchuan Shao, et al. Deep learning-based miniaturized all-dielectric ultracompact film spectrometer. *ACS Photonics*, 10(1):225–233, 2022.

[30] Jaebum Noh, Yong-Hyun Nam, Sunae So, Chihun Lee, Sun-Gyu Lee, Yongjune Kim, Tae-Hyung Kim, Jeong-Hae Lee, and Junsuk Rho. Design of a transmissive metasurface antenna using deep neural networks. *Optical Materials Express*, 11(7):2310–2317, 2021.

[31] Johan Broberg, Maria Bånkestad, and Erik Ylipää. Pre-training transformers for molecular property prediction using reaction prediction. *arXiv preprint arXiv:2207.02724*, 2022.

[32] Nicholas Geneva and Nicholas Zabaras. Transformers for modeling physical systems. *Neural Networks*, 146:272–289, 2022.

[33] Pooya Khorrami, Olga Simek, Brian Cheung, Mark Veillette, Rumen Dangovski, Ileana Rugina, Marin Soljacic, and Pulkit Agrawal. Adapting deep learning models to new meteorological contexts using transfer learning. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 4169–4177. IEEE, 2021.

[34] Nihang Fu, Lai Wei, Yuqi Song, Qinyang Li, Rui Xin, Sadman Sadeed Omee, Rongzhi Dong, Edirisuriya M Dilanga Siriwardane, and Jianjun Hu. Material transformers: deep learning language models for generative materials design. *Machine Learning: Science and Technology*, 4(1):015001, 2023.

[35] Nathan Frey, Ryan Soklaski, Simon Axelrod, Siddharth Samsi, Rafael Gomez-Bombarelli, Connor Coley, and Vijay Gadepally. Neural scaling of deep chemical models. 2022.

[36] Renqian Luo, Liai Sun, Yingce Xia, Tao Qin, Sheng Zhang, Hoifung Poon, and Tie-Yan Liu. Biogpt: generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics*, 23(6):bbac409, 2022.

[37] Abel Chandra, Laura Tünnermann, Tommy Löfstedt, and Regina Gratz. Transformer-based deep learning for predicting protein properties in the life sciences. *Elife*, 12:e82819, 2023.

[38] Daria Grechishnikova. Transformer neural network for protein-specific de novo drug generation as a machine translation problem. *Scientific reports*, 11(1):321, 2021.

[39] Nur Ismail, Cristine Calil Kores, Dimitri Geskus, and Markus Pollnau. Fabry-pérot resonator: spectral line shapes, generic and related airy distributions, linewidths, finesses, and performance at low or frequency-dependent reflectivity. *Optics express*, 24(15):16366–16389, 2016.