Research Article

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

Abstract: Deep learning (DL) has emerged as a promising tool for photonic inverse design. Nevertheless, despite the initial success in retrieving spectra of modest complexity with nearly instantaneous readout, DL-assisted design methods often underperform in accuracy compared with advanced optimization techniques and have not proven competitive in handling spectra of practical usefulness. Here, we introduce a tandem optimization model that combines a mixture density network (MDN) and a fully connected (FC) network to inversely design practical thin-film high reflectors. The multimodal nature of the MDN gives access to infinite candidate designs described by probability distributions, which are iteratively sampled and evaluated by the FC network to allow for rapid optimization. We show that the proposed model can retrieve the reflectance spectra of 20-layer thin-film structures. More interestingly, it reproduces with high precision the periodic structures of high reflectors derived from physical principles, even though no such information is included in the training data. Improved designs with extended high-reflectance zones are also demonstrated. Our approach combines the high-efficiency advantage of DL with the optimization-enabled performance

improvement, enabling efficient and on-demand inverse design for practical applications.

1 Introduction

Deep learning (DL) has emerged as a promising and powerful tool for many tasks in optics and photonics, such as image reconstruction and enhancement in microscopy [1], feature recognition in spectroscopy [2], and inverse design of photonic devices [3–8], to name a few. Conceptually, DL uses many-layered neural networks (NNs), where a series of parameters are adjusted upon exposure to a repository of training data, to learn an abstract and generalizable mapping from the inputs to the outputs of NNs. In the case of inverse design, the inputs are the desired optical responses, and the outputs are the design variables of a coinciding structure. Owing to their remarkable ability to unearth unintuitive and hidden relations within the data, DL-based methods have been applied to designing various photonic structures, including metasurfaces [9–11], metagratings [12, 13], and multilayer nanoparticles and thin films [14–21], etc. The choice of the assisting NNs also diversifies quickly from fully connected (FC) networks to many advanced models (discriminative or generative) and their ensembles [22–25]. While these efforts have shown encouraging ability to reproduce spectra based on fully random designs or simple physical processes such as a Lorentzian resonance, little has been reported at the level of solving practical problems, for which some rules of design have been derived from physical principles. Thus, DL has yet to convincingly show a capacity to match or replace traditional physics-based design methods. In addition, in comparing with optimization techniques for inverse design [26–28], DL offers an unparalleled advantage of high speed after training but is less favorable in accuracy, especially when the target optical responses contain steep features resulting from

*Corresponding author: Yuebing Zheng, Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712, USA; and Texas Materials Institute, The University of Texas at Austin, Austin, TX 78712, USA E-mail: zheng@austin.utexas.edu. https://orcid.org/0000-0002-9168-9477 (Y. Zheng)
Rohit Unni and Kan Yao, Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712, USA; and Texas Materials Institute, The University of Texas at Austin, Austin, TX 78712, USA. https://orcid.org/0000-0001-7905-8317 (R. Unni)
Xizewen Han, Department of Statistics and Data Science, The University of Texas at Austin, Austin, TX 78712, USA
Mingyuan Zhou, Department of Statistics and Data Science, The University of Texas at Austin, Austin, TX 78712, USA; and McCombs School of Business, The University of Texas at Austin, Austin, TX 78712, USA

sophisticated physical processes. Recent works have shown the potential of incorporating both DL and optimization in inverse design [10, 14, 29], which could combine the advantages of high speed and scalability from DL with the higher maximum performance afforded by optimization techniques. This attempt raises the possibility that the most fruitful avenue for application of DL to inverse design may be in concert with existing approaches.

In this work, we report a tandem model that combines two NNs with an optimization process to inversely design a type of multilayer thin-film structure with wide applications, which are known as the high reflectors. Among the structures to which DL-based design methods have been applied, multilayer stacks are of particular interest [15–18, 30]. On one hand, thorough understanding of the properties of multilayer structures sets the basis of thin-film optics. Albeit the seemingly simple geometry, multilayer thin films can provide a variety of optical properties (see Figure 1A), depending on the choice of materials and their arrangements. The associated devices span beam splitters [31], high reflectors [32], antireflection coatings [33], and numerous optical filters [34]. In different application scenarios, the exact line shape of the desired spectra is case-specific, each corresponding to a structure to be inversely solved. Practical thin-film devices can have tens or hundreds of layers, making the task challenging enough to stimulate innovative design methods. On the other hand, the development of thin-film optics has reaped many classical designs based on physical principles or rules of thumb [34]. Retrieving and outperforming these devices can be good gauges to assess the performance of numerical design methods.

Our two NNs include a mixture density network (MDN), which solves the inverse problem to suggest probability distributions of "raw" designs [35], and a standard FC network as a forward simulator to instantly evaluate any updated design during iterative optimization. The applicability of the proposed model is beyond retrieval of random thin-film structures. We show that the classical design of high reflectors, derived from the principle of interference of light, can be reconstructed with high precision, even though no information about the periodicity or quarter-wavelength layer thickness is introduced into the training data. Improved designs with extended high-reflectance regions, which are not easily obtainable by physics-based design rules, are also demonstrated. We emphasize that the use of two NNs in this work differs from previous implementations of a tandem architecture [15, 36], which feature the inverse and forward networks directly connected to relieve non-uniqueness. Here, the non-uniqueness is relieved by the
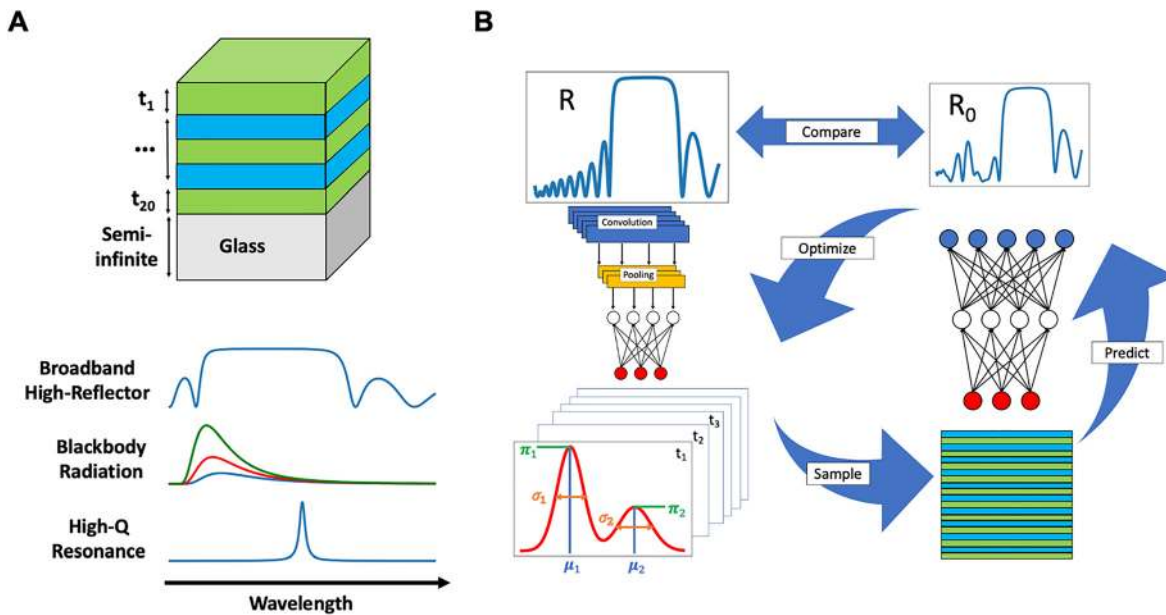

Figure 1: (A) Top: diagram of a 20-layer thin-film structure consisting of two alternating materials placed on a semi-infinite glass substrate. Once the materials are chosen, the thicknesses of layers $t_1 - t_{20}$ constitute the design parameters for any requested optical responses. Bottom: examples of optical spectra of the multilayer thin-film structure featuring selected line shapes for various applications, including broadband high reflection, black body radiation, and narrow peaks from resonances. (B) Schematic of the tandem model with adjoining optimization method. The requested or target spectrum R (navy curve) fed into the MDN (left column) produces probability distributions of the design parameters (red curve), which are repeatedly sampled and evaluated with the forward model (right column) to optimize the designs until the predicted optical response R_0 (blue curve) is close enough to the target spectrum R.

MDN alone, and the forward model is indirectly connected via the optimization method to iteratively improve designs after initial readout. The model is also different from recent works that use NNs only to accelerate the adjoint simulations in an optimization framework [30, 37]. In fact, apart from resampling, our whole model is DL-based. The unique ability of MDN to predict multiple optima of the objective function can potentially aid in the search of a global optimum rather than getting stuck at a local one. Therefore, our approach combines the advantage of DL in high speed with the optimization-enabled performance improvement, enabling efficient and on-demand inverse design of structures and devices for practical applications.

2 Materials and methods

As illustrated in Figure 1B, we construct a tandem model consisting of two independent deep neural networks. The first network is an MDN, designated to solve the inverse problem, i.e., seeking a design that will produce the desired optical properties. The unique feature of MDNs is that the outputs are not deterministic discrete design variables but their probability distributions over the possible ranges. These distributions are built as a mixture of M (we take $M = 16$ in this work) Gaussian distributions, each parametrized by a mean μ, variance σ and a mixing weight π, with the mixing weights shared across all design mixtures. These parameters are what the network outputs in lieu of directly outputting design values. Candidate designs are then generated by sampling the distributions of the individual variables. To train the model, desired outputs are measured for goodness of fit with the model-produced distributions using a metric such as log-likelihood, and this allows for multiple correct outputs to be modeled as different peaks in the distribution. This difference from ordinary NNs alleviates the non-uniqueness problem in inverse design, and more importantly, opens the opportunity of subsequent optimization for performance improvement and better search of a global optimum in the design space. The second network deals with the forward problem, functioning as a simulator to predict the optical response of a given structure. This task is relatively simple and can thus be carried out by a standard FC network. A full description of the model architectures, hyperparameters, and training curves is given in Section 1 of the Supplementary materials. Upon completion of training, both the suggestion of candidate designs for a given spectrum and the prediction of optical properties for a given design are almost instantaneous.

To work jointly as a tandem, the two NNs are connected through an optimization procedure, which essentially post-processes the output of the MDN by iteratively evaluating and updating sampled candidate designs with the assistance of the forward network. The full model works as follows for a design task involving N variables. First, a desired optical spectrum R is fed into the MDN, which produces at the output N probability distributions, each corresponding to one design variable. Without complex sampling strategies, the initial guess of the design is made by assigning every variable the value at the most prominent peak of its own distribution. Next, the initial design suggestion is sent to the forward network for a prediction of the optical response, denoted by R_M and also labeled as R_0 to indicate that it is the predicted response of the active candidate design. The performance of this design is evaluated by comparing R_0 with the ground truth R based on some metrics, including root mean square error (RMSE). Then, the optimization process starts. Of the N variables of the active candidate design, a randomly chosen one is resampled for a specified number of times based on its probability distribution for new guesses, while the remaining $N - 1$ variables are fixed. Once the predicted response of a new guess R_i is closer to R than R_0, that guess becomes the new active candidate design, and R_0 is updated accordingly. The resampling and evaluation repeat for all the N variables in a random order, and the process of cycling through all the variables can also be repeated for any number of times. A more detailed account on the numbers chosen for optimization and their relation to improvement in the design is provided in Section 2 of the Supplementary materials. If the forward model is accurate enough, the prediction of the design guesses' error relative to the ground truth will closely approximate the true error, and the design will improve over time. Lastly, after the optimization is complete, the optical response of the finalized design R' is simulated by plugging the design variables in an electromagnetic solver, which computes the real properties of the final design. It should be emphasized that through the entire design process, simulation is only used once at the very end for verification.

We apply our model to inverse design based on the reflectance from a stack of dielectric thin films of alternating high and low refractive indices. The whole structure consists of 20 layers of magnesium fluoride (MgF_2) and tantalum pentoxide (Ta_2O_5) placed on a glass substrate and with an air cladding, as shown in Figure 1A. For the sake of comparison to physics-based design rules, only normal incidence is considered. This reduces the design variables to the thicknesses of each of the 20 layers, forming a 20-dimensional vector. However, inclusion of the angle of incidence in the design variables has no intrinsic difficulty [16]. The reflectance is calculated for the wavelength interval of 400–1000 nm using analytical formulae based on the Fresnel equations [38], and the spectrum is discretized into 300 points. We limit the ranges of possible physical thicknesses to between 50 and 150 nm for MgF_2 and between 30 and 120 nm for Ta_2O_5, both of which cover the optical thickness of quarter-wavelength for a large portion of the wavelengths of interest. The dielectric functions of MgF_2 and Ta_2O_5 are taken from Refs. [39, 40] unless otherwise specified. To train the two separate models, we use the same dataset of 828,000 samples, split into 70% for training and the remaining 30% for testing. In sample generation, 50% of the data uniformly samples the entire thickness range for all design variables, 25% of the data uniformly samples the upper half of the thickness range for each, and the last 25% uniformly samples the lower half. The uniform sampling means all values within the chosen interval are equally likely to be chosen. Other sampling techniques such as low-discrepancy sequences [41] are also tested and produce qualitatively similar results, as discussed in the Supplementary materials, Section 3. Compared with sampling the full thickness ranges all at once, this strategy ensures that structures with relatively balanced thicknesses are less likely underrepresented, given the enormous design space spanned by 20 independent variables.

3 Results

Both models are trained via gradient descent for 500 epochs on the Stampede2 computer of Texas Advanced

Computing Center (TACC). A single compute node is used for training, which takes a total of approximately 23 h to complete, in addition to about 1 h for data generation. The forward model uses RMSE in the optical response as its loss function and converges to a final RMSE of 0.02 for both the training and test datasets. The MDN uses the negative log-likelihood metric and converges to a value of –18 for both datasets. The optimization process is set to run four cycles of resampling and evaluation through all the 20 variables, and each variable is sampled 50 times in one cycle. With these settings, the optimization of a single design takes approximately 1 min on the same TACC compute node or 6 min on a commercial laptop.

3.1 Retrieval of random stacks

We first examine the tandem optimization model on the spectra produced by random structures in the test dataset. The model can retrieve those designs fairly well even with the MDN alone, and progressive improvements can be seen during the optimization process. Figure 2A presents an example randomly chosen from the test dataset. Three spectra taken at different steps of the design, namely the predicted response of MDN's initial guess R_M (black dash curve), the predicted response of the final design R_0 (green dash curve), and the simulated response of the final design R' (blue curve), are compared with the desired spectrum as ground truth R (red curve). Noticeably, despite the complex

line shape of the target spectrum, the initial guess that simply takes the peak values of the outputs of MDN can already resemble the ground truth design closely at most wavelengths. The deviations between the initial guess and the ground truth are diminished by optimization. Both the predicted and simulated responses of the final design overlap the target spectrum R almost perfectly, confirming the effectiveness of the optimization procedure and the accuracy of the forward network. To get a quantitative idea of the performance of MDN and the improvement by optimization, we conduct a statistical study over 500 samples randomly chosen from the test dataset. Figure 2B shows the histograms of RMSE between the requested spectra and the responses of the initial guesses by MDN and final designs after optimization. The initial designs, sampled at the most prominent probability peaks predicted by the MDN, give an average response RMSE of 0.09. In comparison, after optimization, the model produces final designs with an average RMSE of 0.04, improved by more than a factor of two. This improvement highlights the unique benefits of the probabilistic nature of the MDN. By resampling and using the forward model to evaluate samples without additional simulations, designs can be continuously optimized after an initial sample. The optimization can run for any arbitrary number of rounds and samples, with the processing time scaling linearly with each factor, however the improvements in response RMSE see diminishing returns. A more detailed description is provided in Section 2 of the Supplementary materials.
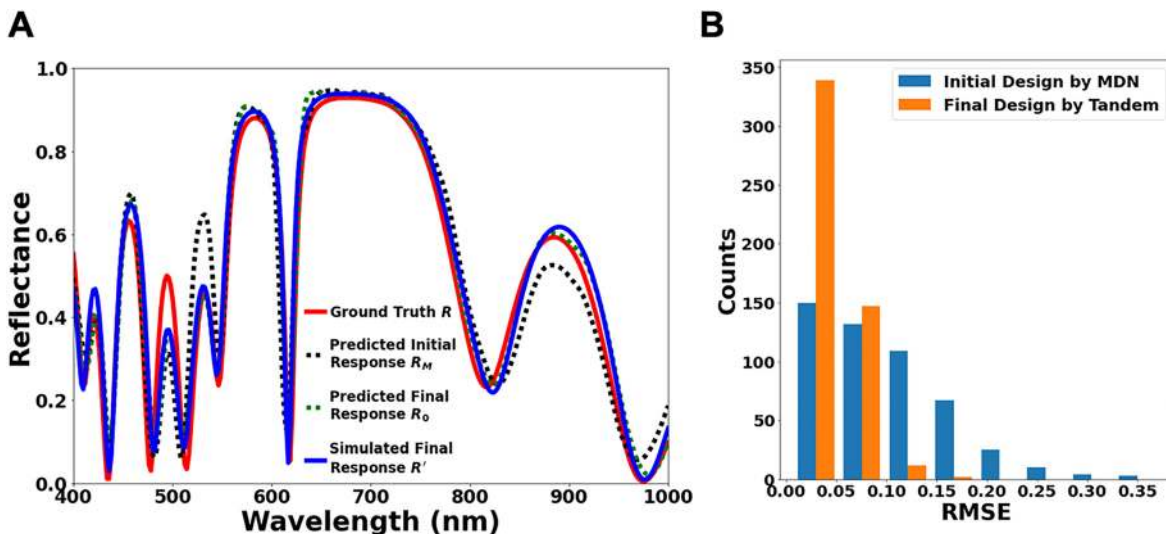

Figure 2: (A) Comparison of a requested spectrum R (red curve) with the simulated response R' of the final design retrieved by the model (blue curve) for a random case from the test dataset. Forward model predictions of the response of the initial MDN design R_M (black dash curve) and of the final design R_0 (green dash curve) are also shown. (B) Histogram of RMSE matching between the requested spectrum and the response of the model-suggested design for 500 randomly chosen test dataset samples before optimization (initial design from the MDN) and after optimization (final design by the tandem optimization model).

3.2 Retrieval of classical high reflectors

Practical applications of the multilayer thin films usually require optical responses that are much more sophisticated than a series of slowly varying peaks. Typical features like a constant response over a certain wavelength range and steep edges cannot be produced by randomly assembled layers as building blocks. Therefore, the ability to retrieve spectra from random structures is no guarantee of practical applications. To prove that the proposed model can be a competitive tool to tackle real-world applications, we further test it with a classical type of thin-film high reflectors, the distributed Bragg reflectors (DBRs) [42, 43], which feature a flat high-reflectance band and are originally accessed by physics-based approaches. The most popular design of DBRs consists of a stack of quarter-wave dielectric films with alternating high and low refractive indices. In other words, a DBR has a periodic structure, and

with a given pair of materials, the thicknesses of neighboring layers are determined by the central wavelength of operation. We compute the reflectance of several DBRs targeting at different wavelengths and feed the spectra into the MDN as the desired optical properties. The proposed model performs very well in retrieving these DBR designs, as shown in Figure 3. In both cases at shorter and longer wavelengths, the predicted and simulated responses of the final designs are highly coincident with the ground truth spectra. More impressively, the model manages to "discover" periodicity in the appearance of two materials, and the suggested thicknesses only deviate slightly from the quarter-wavelengths. With a total of 20 independent design variables, the size of our dataset gives an extremely low chance of any individual training samples being periodic to supply this knowledge to the model. Therefore, the successful retrieval of DBRs, despite their significant divergence from the random structures used for training,
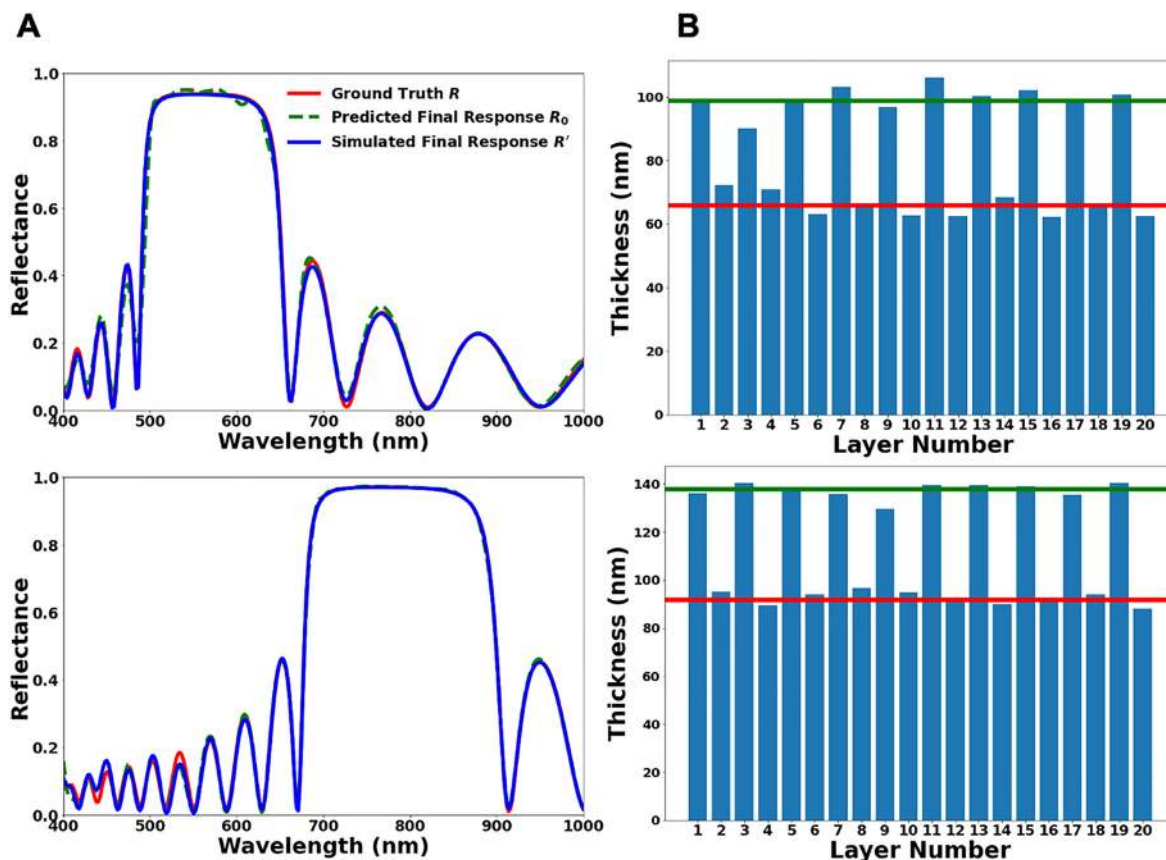

Figure 3: Inverse design of DBRs consisting of alternating high- and low-index films of quarter-wavelength thicknesses stacked in a periodic fashion.
(A) Comparisons of the requested spectrum R from a DBR (red curve) with the simulated response R' of the final design retrieved by our model (blue curve). Also shown is the predicted response R_0 of the final design (green dash curve). Two examples are selected for shorter (top) and longer (bottom) central wavelengths. (B) Visualization of the 20-layer thicknesses for the corresponding designs (top: shorter central wavelength; bottom: longer central wavelength) retrieved by the tandem optimization model for both requested spectra. Green and red lines mark the alternating thicknesses of MgF_2 and Ta_2O_5, respectively, in the ground truth designs of DBRs.

suggests that the model has a strong ability to intuit the design-response mapping on a fundamental level.

The working principle of DBRs is the constructive interference of light fields reflected from each interface of the multilayer thin films, while a random arrangement of thicknesses cannot lead to such an effect. We note two interesting phenomena regarding this difference. One is that DBRs are tolerant of small disturbances to the ideal design. In Figure 3B, the deviations of the retrieved thicknesses from ground truths are largely attributed to this tolerance. The other is that (quasi-)periodic structures may apply implicit constraints that suppress multi-modality in the design. The final thicknesses after optimization in Figure 3 are close to those taken from the MDN's initial suggestion. In contrast, randomly sampled designs have a much greater chance to contain at least one layer that exhibits secondary peaks in the corresponding probability distribution. A brief discussion of multi-modality is presented in Section 3 of the Supplementary materials.
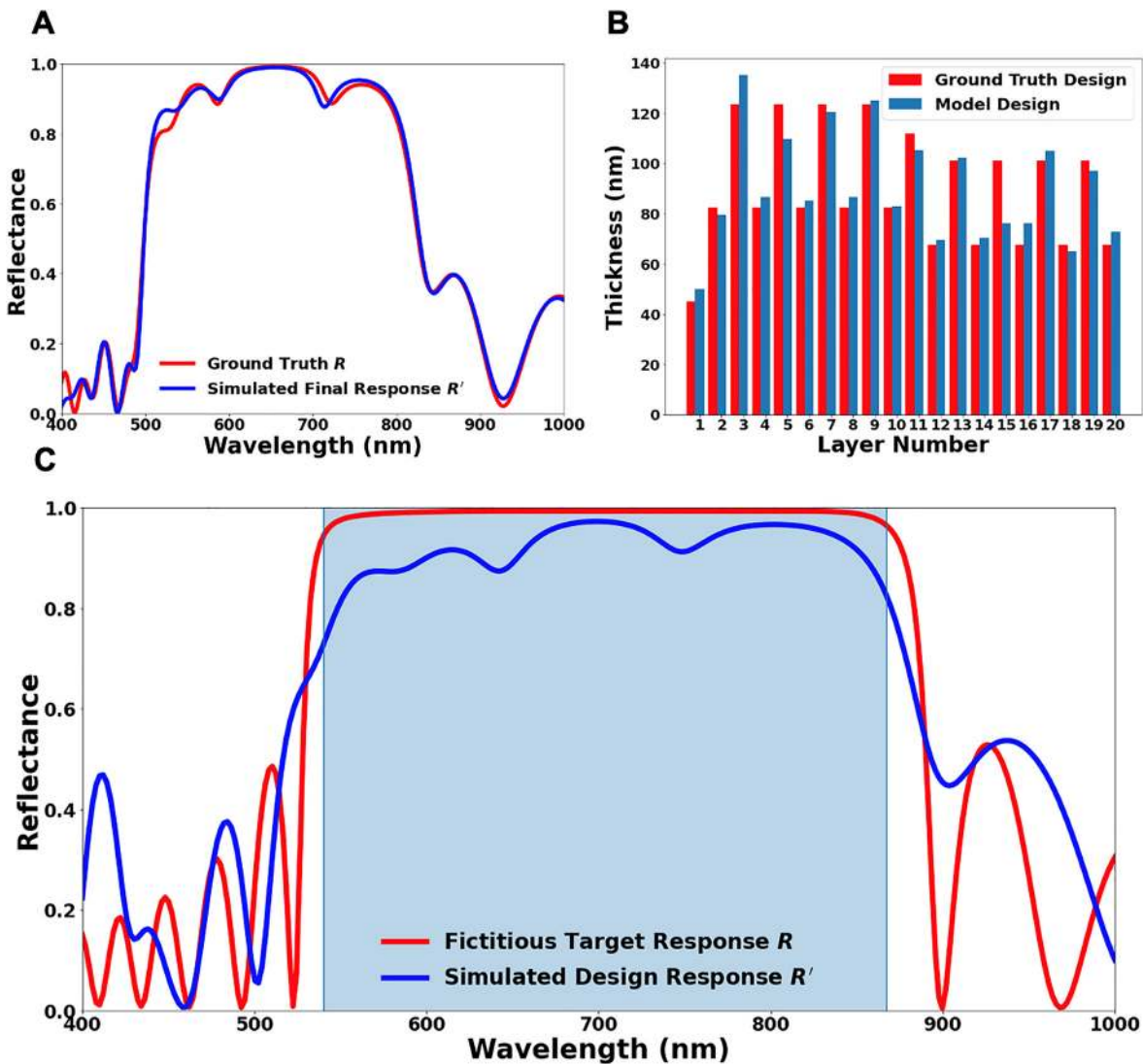

Figure 4: On-demand inverse design of thin-film high reflectors with extended high-reflectance zones.
(A) Comparison of a requested spectrum R and the simulated response R' of the design retrieved by the model. The target design is composed of two 9-layer periodic sub-stacks with a spacer layer in the middle and a cladding layer on the top, obtained with a known optimization strategy based on the principle of interference. (B) Comparison of design variables between the ground truth design and model-produced design. (C) Inverse design of a high reflector with ultrabroad high-reflectance zone. The requested spectrum R (red curve) is fictitious, taken from a 20-layer DBR and artificially extended in the high-reflectance region for a bandwidth not achievable by known design methods based on physical principles. The blue curve shows the simulated response R' of the design suggested by the tandem optimization model. Note that in this case, optimization is applied to the high-reflectance region only, as highlighted by the shaded area.

3.3 Inverse design of reflectors with extended high-reflectance zones

Our next task is set to be more ambitious. Standard DBRs have a limited bandwidth, determined by the choice of materials. In the frequency domain, the width of the high-reflectance zone Δf is given by [34]

$$\frac{\Delta f}{f_0} = \frac{4}{\pi} \arcsin\left(\frac{n_H - n_L}{n_H + n_L}\right), \qquad (1)$$

where f_0 is the central frequency, and n_H and n_L are the refractive indices of the high and low index layers, respectively. Extending the width of the high-reflectance zone is a task of practical concern, and different attempts have been made based on physical considerations or optimization algorithms [34]. We first challenge our model with a design manually optimized with the guidance of interference conditions, which requests stacking two 9-layer periodic structures covering different wavelengths with a spacing layer in the middle and a cladding layer on the top. Again, the model retrieves the design with high precision. To minimize the influence of the dielectric loss, we train the model with a different dataset generated using the dielectric function of Ta_2O_5 measured by Gao et al. [44]. Figure 4A compares the performance of the final design and the ground truth, showing an excellent agreement. The small dips in the extended high-reflectance zone are unavoidable mainly because of the limited number of layers. Figure 4B shows the comparison of design variables. The two distinct sub-stacks can be clearly recognized, despite that no such information is provided in the training data. With only nine layers in each sub-stack, the extraction of strict periodicity is more difficult than in the standard DBRs.

Finally, we solve the ultimate task to search for high-reflector designs that can outperform the examples derived from known principles. Although any fictious spectra can be used as the input to the model, we choose to artificially extend the high-reflectance zone of a DBR to be wider than what is achievable by the previous optimization method [34]. This manipulation creates a spectrum in Figure 4C (red curve), which does not coincide with any known structure and may not be physically possible with the constraints of the design space. However, the model successfully finds a design with consistent high reflectance across the desired wavelength range (blue curve). In this specific task, the optimization process is customized to evaluate the candidate designs only in the target high-reflectance zone. The flexibility of tuning the objective function of optimization further enhances the search for the closest possible matching design to any input spectra.

4 Conclusions and future study

We propose a framework for inverse design based on two NNs combined through an optimization process and demonstrate its ability to outperform physics-based methods in designing thin-film high reflectors. The first NN solving the inverse problem is the MDN. Its unique probabilistic nature allows generation of candidate designs with information of uncertainty, which enables progressive performance improvement through iterative evaluation and resampling during the optimization process. The second NN has a standard FC architecture to solve the forward problem. It serves as a simulator to make instant and accurate predictions of the response of candidate designs from the MDN. We apply this tandem optimization model to on-demand inverse design of high reflectors based on 20-layer thin films. In addition to passing the ordinary test of reproducing the reflectance spectra of random structures, the model successfully retrieves a series of high reflector designs that are derived by physics-based methods. In particular, DBRs with periodic quarter-wave layers are obtained with high accuracy, even though no prior knowledge of periodicity is in the training data. We further demonstrate designs with extended high-reflectance zones, originally accessible only by physical principles or other optimization techniques. Lastly, the model shows a strong ability to search for the closest possible solution to unrealistic reflectors with an artificially widened high-reflectance zone. The proposed model can both quickly and repeatedly produce high-performance designs that are competitive with those obtained by other time-consuming inverse design approaches, making it a promising tool for on-demand industrial applications.

It is worth mentioning that the individual components of the proposed model are highly adaptable. For example, the MDN can be replaced by generative NNs to tackle pattern-based designs, and the optimization would then take samples in the latent space instead of the probability distributions [10]. The success in retrieving periodic structures of DBRs is also encouraging for other uses beyond inverse design, such as knowledge discovery [45–49]. We envision that the combination of DL algorithms and optimization techniques will provide new opportunities for advancing both optical physics and engineering applications.

References

[1] K. de Haan, Y. Rivenson, Y. Wu, and A. Ozcan, "Deep-learning-based image reconstruction and enhancement in optical microscopy," *Proc. IEEE*, vol. 108, no. 1, pp. 30–50, 2020.

[2] C.-S. Ho, N. Jean, C. A. Hogan, L. Blackmon, S. S. Jeffrey, M. Holodniy, et al., "Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning," *Nat. Commun.*, vol. 10, no. 1, p. 4927, 2019.

[3] K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics*, vol. 8, no. 3, pp. 339–366, 2019.

[4] R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nanoscale Adv.*, vol. 2, no. 3, pp. 1007–1023, 2020.

[5] J. Jiang, M. Chen, and J. A. Fan, "Deep neural networks for the evaluation and design of photonic devices," *Nat. Rev. Mater.*, vol. 6, no. 8, pp. 679–700, 2021.

[6] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics*, vol. 15, no. 2, pp. 77–90, 2021.

[7] P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," *Photonics Res.*, vol. 9, no. 5, pp. B182–B200, 2021.

[8] S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics*, vol. 9, no. 5, pp. 1041–1057, 2020.

[9] Z. Liu, D. Zhu, S. P. Rodrigues, K. T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.*, vol. 18, no. 10, pp. 6560–6576, 2018.

[10] Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine-learning-assisted metasurface design for high-efficiency thermal emitter optimization," *Appl. Phys. Rev.*, vol. 7, no. 2, p. 021407, 2020.

[11] C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, "Deep learning for accelerated all-dielectric metasurface design," *Opt. Express*, vol. 27, no. 20, pp. 27523–27535, 2019.

[12] J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, "Free-form diffractive metagrating design based on generative adversarial networks," *ACS Nano*, vol. 13, no. 8, pp. 8872–8878, 2019.

[13] S. Inampudi and H. Mosallaei, "Neural network based design of metagratings," *Appl. Phys. Lett.*, vol. 112, p. 241102, 2018.

[14] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, et al., "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.*, vol. 4, no. 6, p. eaar4206, 2018.

[15] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics*, vol. 5, no. 4, pp. 1365–1369, 2018.

[16] R. Unni, K. Yao, and Y. Zheng, "Deep convolutional mixture density network for inverse design of layered photonic structures," *ACS Photonics*, vol. 7, no. 10, pp. 2703–2712, 2020.

[17] R. Hegde, "Sample-efficient deep learning for accelerating photonic inverse design," *OSA Contin.*, vol. 4, no. 3, pp. 1019–1033, 2021.

[18] J. Jiang and J. A. Fan, "Multiobjective and categorical global optimization of photonic structures based on ResNet generative neural networks," *Nanophotonics*, vol. 10, no. 1, pp. 361–369, 2021.

[19] S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core–shell nanoparticles," *ACS Appl. Mater. Interfaces*, vol. 11, no. 27, pp. 24264–24268, 2019.

[20] Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljacic, "Migrating knowledge between physical scenarios based on artificial neural networks," *ACS Photonics*, vol. 6, no. 5, pp. 1168–1174, 2019.

[21] S. So, D. Lee, T. Badloe, and J. Rho, "Inverse design of ultra-narrowband selective thermal emitters designed by artificial neural networks," *Opt. Mater. Express*, vol. 11, no. 7, pp. 1863–1873, 2021.

[22] W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano*, vol. 12, no. 6, pp. 6326–6334, 2018.

[23] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Adv. Mater.*, vol. 31, no. 35, p. 1901111, 2019.

[24] Z. Liu, D. Zhu, K. T. Lee, A. S. Kim, L. Raju, and W. Cai, "Compounding meta-atoms into metamolecules with hybrid artificial intelligence techniques," *Adv. Mater.*, vol. 32, no. 6, p. 1904790, 2020.

[25] I. Sajedian, J. Kim, and J. Rho, "Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks," *Microsyst. Nanoeng.*, vol. 5, no. 1, p. 27, 2019.

[26] J. S. Jensen and O. Sigmund, "Topology optimization for nano-photonics," *Laser Photonics Rev.*, vol. 5, no. 2, pp. 308–321, 2011.

[27] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics*, vol. 12, no. 11, pp. 659–670, 2018.

[28] J. A. Fan, "Freeform metasurface design based on topology optimization," *MRS Bull.*, vol. 45, no. 3, pp. 196–201, 2020.

[29] F. Wen, J. Jiang, and J. A. Fan, "Robust freeform metasurface design based on progressively growing generative networks," *ACS Photonics*, vol. 7, no. 8, pp. 2098–2104, 2020.

[30] D. Zhang, Q. Bao, W. Chen, Z. Liu, G. Wei, and J. J. Xiao, "Inverse design of an optical film filter by a recurrent neural adjoint method: an example for a solar simulator," *J. Opt. Soc. Am. B*, vol. 38, no. 6, pp. 1814–1821, 2021.

[31] L. Li and J. A. Dobrowolski, "High-performance thin-film polarizing beam splitter operating at angles greater than the critical angle," *Appl. Opt.*, vol. 39, no. 16, pp. 2754–2771, 2000.

[32] R. R. Willey, *Practical Design and Production of Optical Thin Films*, New York, CRC Press, 2002.

[33] M. Keshavarz Hedayati and M. Elbahri, "Antireflective coatings: conventional stacking layers and ultrathin plasmonic metasurfaces. A mini review," *Materials*, vol. 9, no. 6, p. 497, 2016.

[34] H. A. Macleod, *Thin-Film Optical Filters*, Boca Raton, CRC Press, 2010.

[35] C. M. Bishop, *Mixture Density Networks*, Birmingham, Aston University: Neural Computing Research Group, 1994.

[36] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light Sci. Appl.*, vol. 7, no. 1, p. 60, 2018.

[37] Y. Deng, S. Ren, K. Fan, J. M. Malof, and W. J. Padilla, "Neural-adjoint method for the inverse design of all-dielectric metasurfaces," *Opt. Express*, vol. 29, no. 5, pp. 7526–7534, 2021.

[38] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Cambridge, Cambridge University Press, 1999.

[39] L. V. Rodríguez-de Marcos, J. I. Larruquert, J. A. Méndez, and J. A. Aznárez, "Self-consistent optical constants of MgF_2, LaF_3, and CeF_3 films," *Opt. Mater. Express*, vol. 7, no. 3, pp. 989–1006, 2017.

[40] L. V. Rodríguez-de Marcos, J. I. Larruquert, J. A. Méndez, and J. A. Aznárez, "Self-consistent optical constants of SiO_2 and Ta_2O_5 films," *Opt. Mater. Express*, vol. 6, no. 11, pp. 3622–3637, 2016.

[41] H. Niederreiter, "Low-discrepancy and low-dispersion sequences," *J. Number Theor.*, vol. 30, no. 1, pp. 51–70, 1988.

[42] H. Kogelnik and C. V. Shank, "Stimulated emission in a periodic structure," *Appl. Phys. Lett.*, vol. 18, no. 4, pp. 152–154, 1971.

[43] W. Shyh, "Principles of distributed feedback and distributed Bragg-reflector lasers," *IEEE J. Quant. Electron.*, vol. 10, no. 4, pp. 413–427, 1974.

[44] L. Gao, F. Lemarchand, and M. Lequime, "Exploitation of multiple incidences spectrometric measurements for thin film reverse engineering," *Opt. Express*, vol. 20, no. 14, pp. 15734–15751, 2012.

[45] R. Iten, T. Metger, H. Wilming, L. Del Rio, and R. Renner, "Discovering physical concepts with neural networks," *Phys. Rev. Lett.*, vol. 124, no. 1, p. 010508, 2020.

[46] Z. Liu, L. Raju, D. Zhu, and W. Cai, "A hybrid strategy for the discovery and design of photonic structures," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 10, no. 1, pp. 126–135, 2020.

[47] Y. Kiarashinejad, M. Zandehshahvar, S. Abdollahramezani, O. Hemmatyar, R. Pourabolghasem, and A. Adibi, "Knowledge discovery in nanophotonics using geometric deep learning," *Adv. Intell. Syst.*, vol. 2, no. 2, p. 1900132, 2020.

[48] Y. Kiarashinejad, S. Abdollahramezani, M. Zandehshahvar, O. Hemmatyar, and A. Adibi, "Deep learning reveals underlying physics of light–matter interactions in nanophotonic devices," *Adv. Theory Simul.*, vol. 2, no. 9, p. 1900088, 2019.

[49] C. Yeung, J. M. Tsai, B. King, Y. Kawagoe, D. Ho, M. W. Knight, et al., "Elucidating the behavior of nanophotonic structures through explainable machine learning algorithms," *ACS Photonics*, vol. 7, no. 8, pp. 2309–2318, 2020.