

Neural-Network-Based Bandgap Prediction for Photonic Crystal Waveguide Design

Hanyue Wang

Chongqing University, Chongqing , 400044, China

ABSTRACT

Photonic crystals (PCFs) are artificial microstructures characterised by a periodic refractive index distribution. The concept of photonic crystals was first proposed by Eli Yablonovitch in 1987, and subsequently introduced to the field of fibre optics by Philip Russell in 1992. The formation of a photonic bandgap is achieved by the introduction of structural defects within the core, thereby exerting control the propagation of light. In contrast to conventional optical fibres, PCFs employ photonic band gaps or total internal reflection to guide light, thereby ensuring low loss, high confinement and exceptional propagation performance. PCFs have a plethora of applications in the domain of optical communications and integrated photonic devices. However, extant photonic crystal structure designs principally rely on numerical calculation methods, including finite difference methods and transmission matrix methods, which incur high computational costs and protracted design cycles. This complicates the optimization process for complex structures. The proposed methodology integrates physical constraints with a neural network framework to facilitate rapid prediction and design evaluation. Initially, a data set is constructed through the utilization of numerical simulations, thereby facilitating the establishment of a correlation between structural parameters and bandgap properties. Subsequently, a feed-forward neural network is trained using these data to fit the nonlinear mapping between geometric parameters and intrinsic frequency response. Of particular significance is the incorporation of physical constraints into the design criteria, a process informed by the fundamental physical principle that "the working frequency of the waveguide must be within the photonic bandgap." This principle is integrated into the model, thereby ensuring the integrity and efficacy of the method. This approach enables a direct mapping between structural parameters and device usability. The findings of the study demonstrate that the proposed method exhibits a high degree of accuracy in prediction ($R^2 > 0.97$), thereby significantly reducing the computational expense when compared to conventional simulation methods. Additionally, the model effectively captures the non-linear relationship between structural parameters and electrical characteristics, facilitating efficient design of waveguides. This approach enhances the interpretive capabilities and practical applications of the model, providing a streamlined solution for the rapid design of complex photonic crystal components.

KEYWORDS

Photonic crystal fibers; Photonic bandgap prediction; Feed-forward neural network; Physical constraint; Waveguide usability evaluation; Rapid design; Structural parameter optimization

1 Introduction

Photonic crystal fiber (PCF) is a type of fiber optic cable with a periodic refractive index distribution. Unlike traditional fiber cables, which rely on total internal reflection to guide light, PCFs use equivalent refractive index modulation or the photonic bandgap effect ^[1]. This allows PCFs to precisely control light propagation by introducing periodic air holes in the cladding. This allows PCFs to offer significant advantages in single-mode transmission, large mode area, and dispersion control ^[2]. Consequently, PCFs have a wide range of applications in fields such as optical communications, high-power transmission, and the design of new photonic devices.

However, the structural design of photonic crystals typically relies on numerical methods such as the finite difference method and plane wave expansion, among which the time-domain finite difference method (FDTD) is the most widely used due to its relatively simple algorithm and high flexibility in terms of structure and materials. Using the time-domain finite-difference method, single-run simulations and Fourier transforms can be performed to obtain information across a wide frequency range. However, since the feature sizes of photonic crystals are typically on the same order of magnitude as the operating wavelength, or even smaller, the FDTD method requires small spatial and temporal steps during the discretization process, which leads to increased memory consumption and computation time ^[3]. Although these methods offer high accuracy, the high computational cost and long cycle times involved in parameter scanning and structural optimization make it difficult to meet the demands of rapid design and optimization for complex structures. Furthermore, existing research has largely focused on predicting physical quantities (such as spectra or bandgaps) and lacks the ability to directly assess the feasibility of device designs. Consequently, the predicted results still require subsequent simulation verification, which to some extent limits their practical value in engineering applications. For nanodevices with higher precision, as their complexity increases, hundreds or even thousands of simulations are often required to identify a reasonable design solution, and the computational cost of each simulation is extremely high ^[4]. Therefore, how to effectively reduce computational complexity while ensuring design accuracy remains an urgent challenge to be addressed.

In recent years, machine learning methods have garnered increasing attention in the field of photonic device design.

* **Corresponding Author:** , Hanyue Wang, 20234576@stu.cqu.edu.cn

Deep neural networks, as a versatile class of machine learning algorithms, capture and model highly nonlinear data relationships by serially stacking nonlinear processing layers. Although the introduction of machine learning into the field of nano-optics is relatively recent, significant breakthroughs have already been achieved, making data analysis and predictions that were once considered impossible now feasible^[5]. Neural networks are capable of simulating relationships based on nonlinear physics, such as the relationship between the geometry of photonic systems and their electromagnetic response, which offers a completely new perspective for solving both forward and inverse problems^[6]. In 2006, Hinton et al. first proposed that multi-hidden-layer artificial neural networks possess excellent feature learning capabilities and can effectively overcome the limitations of deep neural networks through layer-by-layer pre-training. Inverse design methods based on genetic algorithms can help identify non-intuitive, irregularly shaped photonic structures. In many applications, these structures outperform those designed empirically, such as silicon photonic devices, photonic crystals, and metamaterials. Deep learning modeling of photonic systems has only recently begun to be studied. In 2018, Ferreira et al. used a multilayer perceptron (MLP) to infer the dispersion relations of photonic crystals. The neural network they proposed can compute the dispersion relations of two-dimensional and three-dimensional photonic crystals within a few milliseconds, which is at least four orders of magnitude faster than photonic bandgaps^[7]. Since then, neural networks have been used to model the spectral responses of other plasmonic systems, including chiral nanostructures and lattice structures with custom color profiles, achieving end-to-end mapping by establishing structural design models^[8].

The aforementioned issues essentially stem from the complex, high-dimensional nonlinear mapping relationship between structural parameters and optical responses, which limits the efficiency of traditional numerical methods during parameter scanning and optimization. Therefore, to address this gap and simplify the analysis to highlight bandgap characteristics, this paper selects two-dimensional photonic crystal structures as the subject of study and proposes a method for rapidly predicting photonic crystal bandgaps that integrates neural networks. Leveraging the well-established application of neural networks in modeling nonlinear mappings, we constructed a mapping relationship between the parameters of the photonic crystal structure and its bandgap characteristics, and successfully achieved rapid prediction of key parameters for the first bandgap. Building on this, we further introduced the physical constraint that “the operating frequency of a photonic crystal waveguide is limited by the bandgap range,” elevating the bandgap prediction results from mere physical quantity outputs to engineering criteria for device design, thereby enabling direct assessment of the waveguide structure’s feasibility. Consequently, we established a closed-loop methodology comprising “parameter input—bandgap prediction—feasibility assessment—design screening,” transforming the structural optimization process from traditional iterative numerical validation into efficient, criterion-driven screening. Compared to existing data-driven bandgap prediction methods, our research introduces physically constrained feasibility criteria, enabling a direct mapping from structural parameters to device-level feasibility. While ensuring prediction accuracy, this method significantly reduces computational costs and substantially improves design efficiency. It establishes a direct mapping from physical response prediction to engineering design decisions, effectively overcoming the efficiency bottleneck of traditional numerical methods in complex structure optimization. This approach provides an efficient, scalable, and practically applicable technical pathway for the rapid engineering design of photonic crystal waveguides.

2 Materials and methods

2.1 Structure and Principle

Figure1 shows the structure of the photonic crystal we are studying. By introducing defect microcavities into the periodic array of silicon columns, the periodicity of the lattice can be disrupted. The periodically arranged high-refractive-index silicon columns produce Bragg scattering for light in specific frequency bands (as shown in Figure2 thereby forming a photonic bandgap and exhibiting excellent waveguide properties.

In non-magnetic media, Maxwell’s equations can be expressed as:

$$\nabla \cdot \mathbf{D} = 0, \nabla \cdot \mathbf{B} = 0, \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \quad (1)$$

Assume that the electric field is a periodic harmonic wave:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r})e^{-i\omega t}, \mathbf{H}(\mathbf{r}, t) = \mathbf{H}(\mathbf{r})e^{-i\omega t} \quad (2)$$

Substitute this into Maxwell's equations to obtain the time-domain form:

$$\nabla \times \mathbf{E}(\mathbf{r}) = i\omega\mu_0 \mathbf{H}(\mathbf{r}), \nabla \times \mathbf{H}(\mathbf{r}) = -i\omega\epsilon(\mathbf{r})\mathbf{E}(\mathbf{r}) \quad (3)$$

Eliminating the magnetic field $\mathbf{H}$ yields the governing equation for the electric field:

$$\nabla \times (\nabla \times \mathbf{E}(\mathbf{r})) = \omega^2 \mu_0 \epsilon(\mathbf{r}) \mathbf{E}(\mathbf{r}) \quad (4)$$

Periodic conditions through the medium:

$$\epsilon(\mathbf{r} + \mathbf{R}) = \epsilon(\mathbf{R}) \quad (5)$$

The final solution is the orthogonal normalization of the electric field:

$$\sum_{\mathbf{G}_l} |\mathbf{k} + \mathbf{G}_l| \times |\mathbf{k} + \mathbf{G}_l| E_k(\mathbf{G}_l) \tilde{\epsilon}^{-1}(\mathbf{G} - \mathbf{G}_l) = \frac{\omega^2}{c^2} E_k(\mathbf{G}) \quad (6)$$

where $\tilde{\epsilon}^{-1}(\mathbf{G} - \mathbf{G}_r)$ is the dielectric constant of the inverted vector.

This paper focuses on the geometric parameter R (the radius of the air hole) and uses the corresponding natural frequency ω to characterize the photonic bandgap properties. A neural network is trained using the relationship between ω and the hole radius R to enable rapid prediction of the frequency.

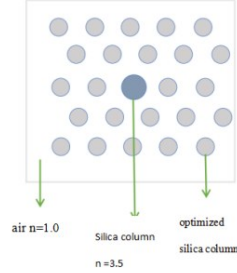


Figure 1 Photonic crystal structure

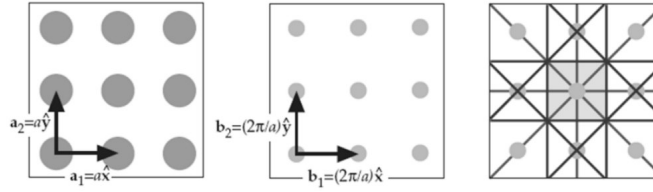


Figure 2 Upright Checkered Pattern and Inverted Checkered Pattern

2.2 Data Set Preparation

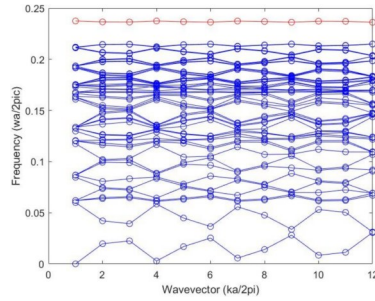


Figure 3 Energy band

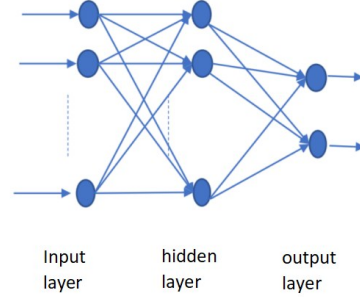


Figure 4 Structure of Neural Networks

In this study, the pore size R was selected as the input parameter, while the cutoff frequency ω of electromagnetic wave propagation in the photonic crystal was used as the output parameter. Using the open-source software package MPB, developed by Steven G. Johnson at the Massachusetts Institute of Technology (MIT) to calculate the band structure and electromagnetic modes of periodic dielectric structures, we obtained the band structures for different pore sizes and used them to train the neural network. The normalized air-to-pore radius R/a was selected as the primary structural parameter, and uniform sampling was performed within a preset parameter range, generating a total of 500 structural samples. For each set of parameters, the upper and lower cutoff frequencies of the first-photon bandgap were extracted as the model prediction targets; for each structural configuration, the corresponding band structure was calculated, and the cutoff frequency of the first-photon bandgap was extracted as the target output; to ensure the stability and convergence of the model, all input parameters were normalized prior to training.

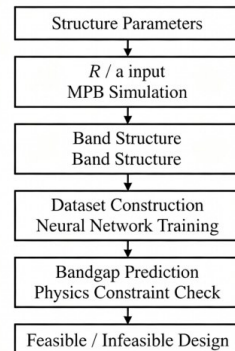


Figure 5 Overall Flowchart

2.3 Experimental Implementation

To ensure that the model has good generalization ability, the training set must contain a sufficiently large and representative number of samples. In this experiment, the training set comprises approximately 4/5 of the total samples (400 sets), while the remaining 1/5 (100 sets) serves as the test set, ensuring that the distribution patterns of the training and test sets are similar.

First, the number of input variables, the number of hidden layers, the number of neurons in each layer, and the number of output variables were determined to establish the network architecture required before creating the BP neural network (Figure 5).

After determining the network architecture, set the relevant training parameters (such as the number of training iterations and the learning rate), and then begin training the network. Once model training is complete, input the test set's input variables into the model; the model's output will serve as the corresponding prediction. The expression for this neural network is: $y = f(W_2 \cdot \sigma(W_1 x + b_1) + b_2)$

Finally, the error between the predicted values and the actual values in the test set is calculated using a loss function ($L = \frac{1}{N} \sum (y_{\text{pred}} - y_{\text{true}})^2$) to evaluate the model's generalization ability. Based on this, the number of hidden layer nodes is optimized to improve the model's generalization performance, and the predicted data is analyzed in conjunction with the physical constraints of the waveguide to achieve a preliminary prediction of the waveguide structure.

To enhance the model's applicability in engineering design, this paper introduces a physical constraint mechanism based on neural network predictions. In photonic crystal waveguides, guided modes can only exist stably when the operating frequency falls within the photonic bandgap.

Therefore, we define a feasibility criterion: if the operating frequency ω_{op} falls within the range $[\omega_{\text{min}}, \omega_{\text{max}}]$, the design is feasible; otherwise, it is infeasible. Based on this criterion, the model's output is extended from a single frequency prediction to a design feasibility assessment, enabling rapid screening of structural parameters. This method avoids the repetitive numerical verification of each candidate structure required in traditional design workflows, thereby significantly improving design efficiency.

3 Results and Discussions

We train the neural network (Figure4). The outputs of the hidden layer nodes use the sigmoid function as the activation function, while the outputs of the output layer nodes use a linear function as the activation function.

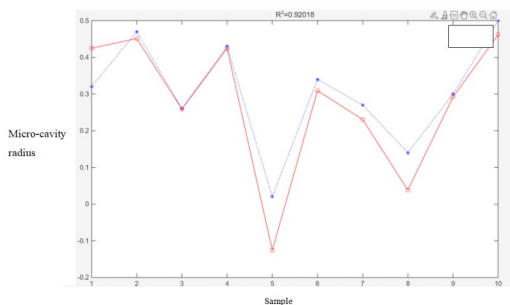


Figure 6 Predicted Results

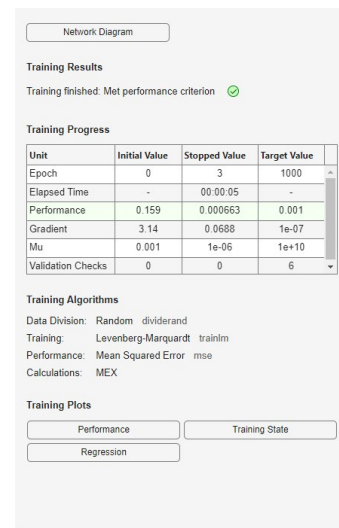


Figure 7 Performance

Figure7 compares the model-calculated values with the actual values for five test samples. It shows that the neural network generally provides a good fit for predicting the microcavity radius, with an R^2 as high as 0.970, indicating that the model exhibits high accuracy for most samples. However, for samples with extremely small radii (such as Sample 3), the prediction error is significant, indicating that the model's generalization ability in extreme ranges remains limited. Overall, the model performs excellently, but there is still room for improvement in extreme cases and with small sample sizes.

Furthermore, to validate the effectiveness of the proposed model, this paper compares it with traditional polynomial regression methods. The results show that, under the same dataset conditions, the neural network model outperforms traditional methods in both prediction accuracy and the ability to fit nonlinear relationships, further validating the effectiveness of the proposed method.

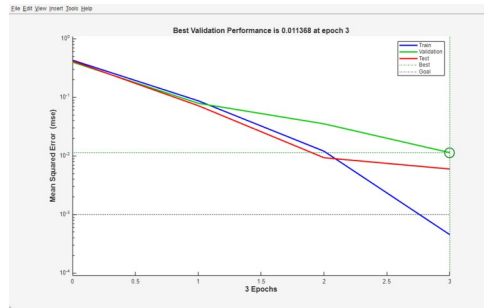


Figure 8 Root Mean Square Error

Figure 8 illustrates the evolution of the mean squared error over the number of training iterations. Blue represents the training set, green represents the validation set, and red represents the test set. The results show that the mean squared error is lowest in the fourth iteration. It should be noted that the overall fit does not increase linearly with the number of training iterations. Overall, the RMSE exhibits a general negative correlation with the number of training iterations and eventually converges to a minimum. This indicates that the training dataset is sufficiently large to effectively train a high-performance neural network.

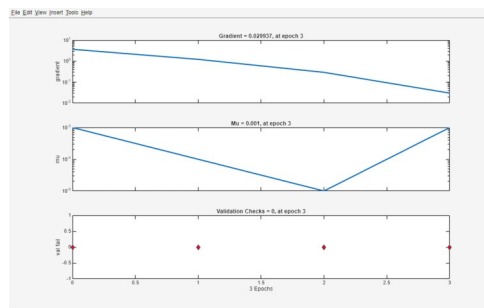


Figure 9 Training state

Figure 9 illustrates the training status. The training gradient has stabilized by the seventh training iteration, indicating that the training set is of high quality and that a good fit has been achieved. The validation error in the fourth training iteration is essentially zero, consistent with the optimal performance observed during training.

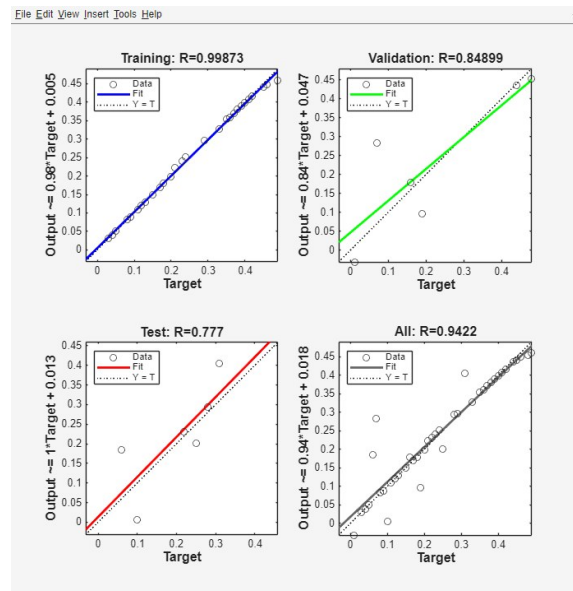


Figure 10 Regression

Figure 11 illustrates the regression performance of the neural network. The correlation coefficients for the training set, validation set, and test set reached $R = 0.99873$ and $R = 0.84899$, respectively, indicating that the model accurately fits the target during both the training and validation phases. The correlation coefficient for the test set is $R = 0.777$, which is slightly lower than those for the training and validation sets but still demonstrates strong generalization ability. The overall correlation coefficient for the entire dataset is $R = 0.9422$, indicating that the model exhibits high predictive accuracy across the entire dataset and can be effectively applied to the rapid prediction of microcavity radii.

Although the model has achieved a high degree of accuracy, some errors still remain. This dataset was generated by uniformly sampling the radii of air holes, but the corresponding bandgap response is inherently nonlinear. In particular, samples in extreme parameter ranges (such as those with extremely small radii) account for a small proportion of the dataset, resulting in insufficient learning of samples in these ranges during the training process. Furthermore, the current model employs a single-hidden-layer neural network with a limited number of neurons. Although it can theoretically approximate nonlinear functions, under-fitting occurs in specific regions during practical application, particularly in areas where physical properties change drastically, thereby leading to prediction errors.

To validate the effectiveness of the proposed neural network model, this paper introduces a polynomial regression method as a comparison model and uses a third-order polynomial to fit the relationship between structural parameters and bandgap frequency. The predictive performance of the two methods was evaluated under the same dataset conditions.

The results indicate that although the polynomial regression method can characterize the relationship between parameters and frequency to a certain extent, its prediction error increases significantly in regions with pronounced nonlinear variations. In contrast, the proposed neural network model can more accurately capture complex nonlinear mapping relationships. Benefiting from the neural network's advantages in approximating high-dimensional nonlinear functions, it outperforms the polynomial regression method in terms of both overall prediction accuracy and generalization ability.

Table 1 Comparison with Numerical Methods

Methods	R^2	MSE
Polynomial	0.92	1.2e-3
NN(article)	0.97	3.5e-4

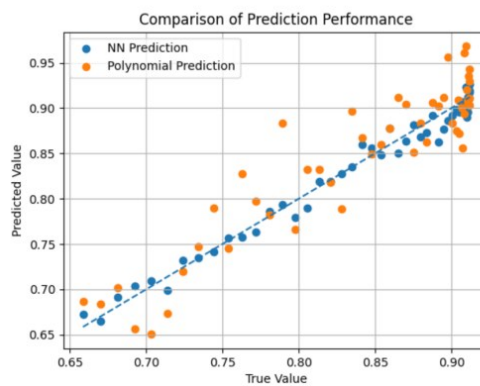
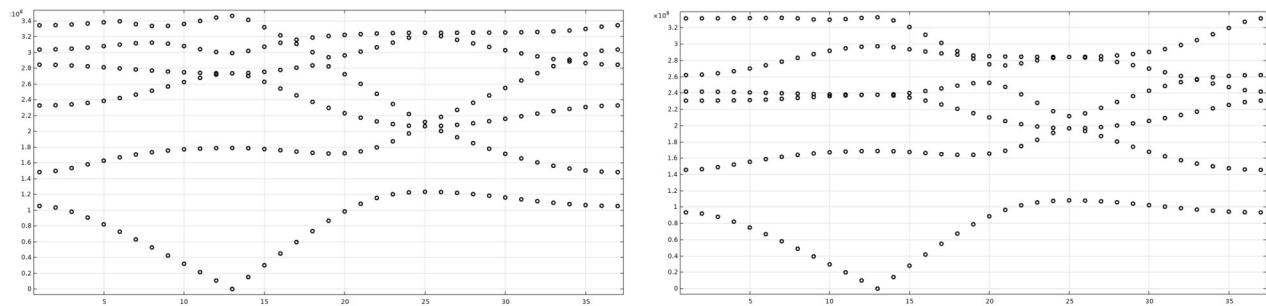


Figure 11 Comparison of prediction performance

These comparison results further validate the advantages of neural network methods in handling the complex nonlinear relationship between the structural parameters of photonic crystals and their electromagnetic response.

Furthermore, to further validate the reliability of the model, we performed numerical simulations of a typical structure using COMSOL Multiphysics to compare the band structure characteristics under different aperture conditions. We constructed the square unit cell mentioned above using silicon with a relative permittivity of 12 and a refractive index of 3.5, and introduced an air aperture at the center.

Through the software simulations, we varied the size of the central aperture to visually observe the relationship between the energy bands and the aperture size. Specifically, we set the length of the square to a and gradually increased the aperture size from $0.1a$ to $0.225a$ in $0.025a$ increments, obtaining six sets of data for comparison with the predicted energy bands.



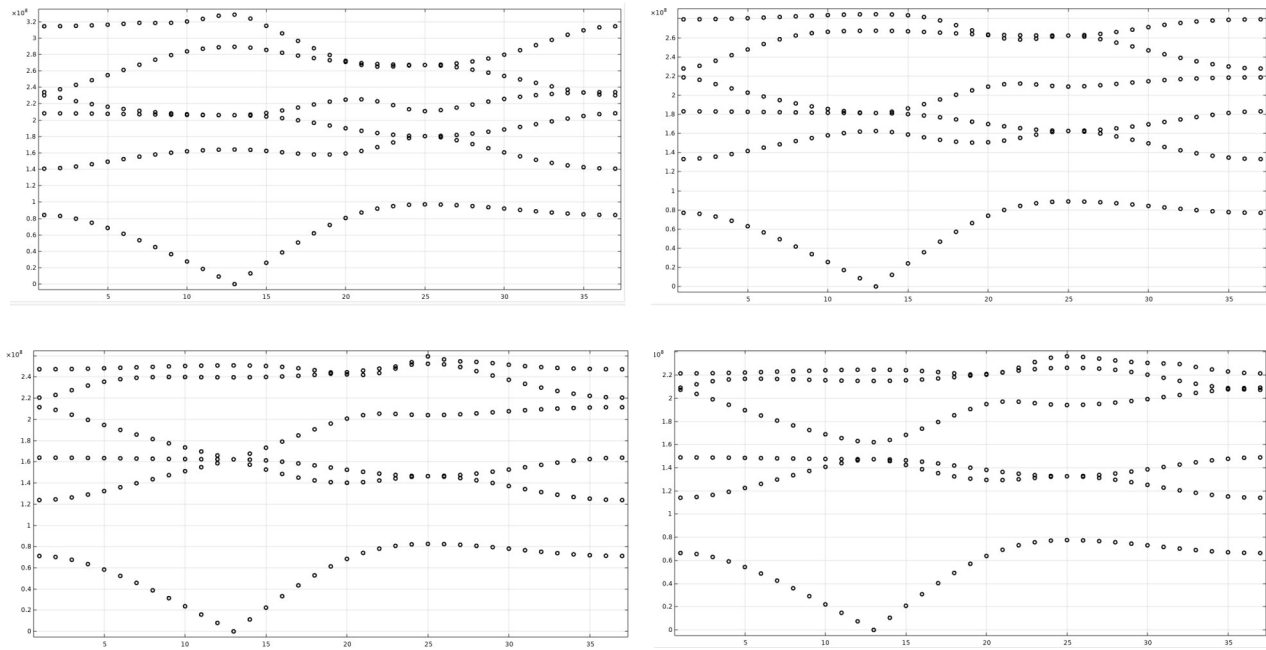


Figure 12 Energy band

The band structures of photonic crystals with different air-hole radii are shown in the figure, which are used to investigate the influence of geometric parameters on the band structure. The results indicate that the overall band distribution shifts significantly with changes in the hole radius, demonstrating that the natural frequency is highly dependent on the structural scale. Within a specific range of parameters, the opening and closing of the bandgap can be observed, indicating that the bandgap is highly sensitive to adjustments in the aperture size. Combined with previous neural network prediction results, it can be seen that the constructed model effectively captures the nonlinear mapping relationship between aperture changes and frequency response, thereby enabling rapid prediction and optimization of bandgap characteristics. This provides strong support for the efficient engineering design of photonic crystal devices (such as filters and waveguides).

4 Conclusion

Leveraging the advantages of neural networks in modeling nonlinear mappings, this paper constructs an efficient mapping model between the parameters of two-dimensional photonic crystal structures and their bandgap characteristics, thereby enabling rapid prediction of key parameters of the first bandgap. Building on this, we introduce the physical constraint that "the operating frequency of a photonic crystal waveguide is limited by the bandgap range," transforming traditional single-bandgap prediction results into usability criteria applicable to device design, thus allowing for direct evaluation of the waveguide structure's effectiveness. The results demonstrate that this method significantly reduces computational costs and effectively improves design efficiency while ensuring prediction accuracy, successfully achieving a direct mapping from physical response prediction to engineering design decisions. More importantly, by incorporating physical constraints, the model not only exhibits outstanding predictive capabilities but also offers clear engineering interpretability and practical application value. In summary, the model and methodology established in this paper not only validate the effectiveness of data-driven approaches in photonic crystal bandgap prediction but also enhance the model's engineering applicability through the introduction of physical constraints. Future research could further integrate higher-dimensional structural parameters and multi-objective optimization strategies, along with experimental validation, to advance the application and development of this method in practical photonic device design.

References

- [1] Russell, P. (2003). "Photonic Crystal Fibers." *Science* 299(5605): 358–362.
- [2] Birks, T. A., et al. (1997). "Endlessly single-mode photonic crystal fiber." *Optics Letters* 22(13): 961–963.
- [3] Pissourt, D., et al. (2004). "Comparative study of three methods for the simulation of two-dimensional photonic crystals." *Journal of the Optical Society of America A* 21(11): 2186–2195.
- [4] Liu, D., et al. (2018). "Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures." *ACS Photonics* 5(4): 1365–1369.
- [5] Wiecha, P. R., et al. (2021). "Deep learning in nano-photonics: inverse design and beyond." *Photonics Research* 9(5): B182–B200.
- [6] Jiang, J., et al. (2021). "Deep neural networks for the evaluation and design of photonic devices." *Nature Reviews Materials* 6(8): 679–700.
- [7] Yang, D., et al. (2023). "A hybrid method for photonic crystal fiber polarization filter based on artificial neural network and genetic algorithms." *Optical Fiber Technology* 80: 103426.
- [8] Ma, W., et al. (2021). "Deep learning for the design of photonic structures." *Nature Photonics* 15(2): 77–90.