Proceedings of the 7th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/2027.39024

Research on Inversion of Structural Parameters of Photonic Crystal Microcavities Based on Artificial Neural Networks

Jingxuan Xu

James Watt Engineering School, University of Glasgow, Glasgow, UK
599848643@qq.com

Abstract. Photonic crystal microcavity parameters directly determine their band characteristics and optical responses, while conventional forward design methods often involve repeated numerical simulations and high computational costs when solving parameter inversion problems. To address this issue, this paper proposes an artificial neural network-based method for the inversion of photonic crystal microcavity structural parameters, using photonic band features as the input and the microcavity radius as the output. The plane-wave expansion method is employed to generate photonic band data corresponding to different microcavity radii. Each sample consists of 50 TM bands at 12 k-points, forming a 600-dimensional input feature vector. A feedforward neural network trained with the Levenberg–Marquardt algorithm is then used to learn the nonlinear mapping between band features and microcavity radius. Numerical results show that the model achieves its best validation performance at the 6th epoch, with a validation MSE of 0.023263, and training is terminated at the 12th epoch due to the validation stopping criterion. The prediction results on independent test samples show that the network can effectively capture the variation trend of the microcavity radius. These results demonstrate the feasibility of using photonic band information for microcavity structural parameter inversion and provide a basis for data-driven inverse design of photonic crystal structures.

Keywords: photonic crystal, microcavity, artificial neural network, band structure, inverse design, plane-wave expansion

1. Introduction

Photonic crystals are artificial periodic dielectric structures capable of manipulating light propagation through their characteristic band structures, and have been widely studied in optical communication, sensing, and integrated photonics [1]. Traditional photonic crystal design usually relies on physical models, repeated parameter adjustment, and numerical simulation. As the structural degrees of freedom increase, this process becomes increasingly time-consuming and computationally expensive. Therefore, inverse design, which determines structural parameters from target optical responses, has attracted increasing attention. In recent years, data-driven methods such as neural networks have been introduced to learn the nonlinear relationship between structural parameters and optical responses, providing new approaches for rapid prediction and inverse design [1-6]. The plane-wave expansion method (PWE) is widely used to calculate the band structures of

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2027.39024

periodic photonic crystals by solving the corresponding electromagnetic eigenvalue problem [7, 8]. However, conventional forward design still requires repeated simulations when searching for structural parameters that produce desired band characteristics.

In this work, an artificial neural network is employed to establish an inverse mapping from photonic band features to the microcavity radius. The microcavity radius was varied from 0.01 to 0.50, generating 50 structural samples. For each sample, 50 TM bands at 12 k-points are calculated and expanded into a 600-dimensional feature vector. The corresponding microcavity radius was used as the prediction target. The proposed framework combines PWE-based data generation, feature construction, neural-network training, and independent testing. Its purpose is to evaluate whether high-dimensional band information can effectively encode structural characteristics and support microcavity parameter inversion, providing a basis for further studies involving larger datasets, multiple structural parameters, and more advanced network architectures.

2. Band data construction and feature extraction of photonic crystals

The periodic dielectric structure of the photonic crystal can form specific photonic energy bands. The introduction of microcavities or defect structures will change the local electromagnetic field distribution and the corresponding energy band response. Therefore, the energy band data corresponding to different microcavity radii contain information related to the structure size, providing a data basis for further establishing an inverse mapping relationship between "optical response - structural parameters".

The structure adopted in this study uses GaAs as the background material and sets a central microcavity in the periodic air hole array. The conventional air radius is denoted as r, and the radius of the central microcavity to be inverted is denoted as R. During the research process, the other structural conditions remain unchanged, and only the microcavity radius is changed to combine the energy band changes with the changes in the target structural parameters.

The calculation can adopt the plane wave expansion method (PWE). This method utilizes the translational symmetry of the periodic structure to expand the dielectric function and the electromagnetic field in the reciprocal lattice space as plane waves, and by solving the corresponding eigenvalue problem, obtaining the eigenfrequencies under different wave vectors. In this paper, PWE is first used to conduct forward numerical calculations on the photonic crystal structures corresponding to different microcavity radii, and then the obtained band frequencies are used as the input data for the neural network. Thus, PWE is responsible for generating "structure-energy band" samples with clear physical meaning, while the artificial neural network is responsible for learning the inverse relationship.

To construct the "band feature - microcavity radius" sample pairs that can be learned by neural networks, this paper conducts parameter scanning of the microcavity radius during the experimental process. The microcavity radius varies from 0.01 to 0.50, with a step size of 0.01, resulting in 50 sets of different structure samples. For each microcavity radius, 12 k points are selected along the predetermined Brillouin zone path, and the 50 TM bands at each k point are calculated. Therefore, each microcavity structure corresponds to a 12×50 band frequency matrix. In the data preprocessing stage, the TM band frequencies are extracted from the PWE output file and the k-point index, kx, ky, kz, and |k| are removed as auxiliary columns. Only 50 TM bands are retained. Then, the 12×50 band matrix of each structure sample is expanded into a 600×1 feature vector in a fixed order, and the corresponding microcavity radius is used as the target output. Finally, a 600-dimensional input feature and a 50-sample supervised learning dataset are formed. Before feature-vector construction, the extracted eigenfrequencies at each k-point are checked to ensure that the 50 TM bands are

DOI: 10.54254/2755-2721/2027.39024

arranged in ascending frequency order. After feature construction is completed, the program further checks the input dimensions, NaN, Inf, and completely repetitive samples, and saves the cleaned data for subsequent modeling. This process ensures that all samples have a unified data structure, while retaining the frequency change information on different k points and different bands, providing a foundation for neural networks to learn the nonlinear relationship between high-dimensional band features and microcavity radius.

3. Inverse model of artificial neural network microcavity radius

This paper employs a single-hidden-layer feedforward artificial neural network to establish an inverse mapping model from band characteristics to the radius of the microcavity. The input layer of the network corresponds to the 600-dimensional TM band characteristics of each sample, the hidden layer was used to complete feature combination and nonlinear transformation, and the output layer had one neuron that directly output the radius of the microcavity to be inverted.

All 50 sets of structural samples were randomly shuffled before modeling. Among them, 45 sets were used for network modeling, and the other 5 sets were reserved as independent test samples. The independent test samples did not participate in network parameter updates and were used after training to evaluate the predictive ability of the model on unseen data. During network training, the 45 modeling samples were further randomly divided into a training set, a validation set, and an internal test set. The training set was used to update the network weights, the validation set was used to monitor generalization performance and trigger early stopping, and the internal test set was used to assist in evaluating model fitting.

The network was trained using the Levenberg–Marquardt (trainlm) algorithm [9], with the performance function being the mean square error (MSE). The maximum number of training iterations was set to 10,000, the performance target was set to 1×10^{-12}, and the gradient threshold was set to 1×10^{-7}. Training would terminate when the performance target, the gradient stopping condition, or the validation early stopping condition was reached.

4. Model training results and performance of structural parameter inversion

The actual training results of the neural network are shown in Figure 1. The network trained for 48 epochs. At the end of training, the performance value decreased from 0.0287 to 2.23×10^{-21}, the gradient decreased to 3.66×10^{-10}, and the Mu value decreased to 1×10^{-9}. The validation-check count reached 6. The training interface reported that the performance criterion had been met, while the validation performance had not improved after the best-validation epoch.

The training and validation trends are further illustrated in Figure 2. The best validation performance is obtained at epoch 42, with a validation-set MSE of 1.15516×10^{-4}. After epoch 42, the training error continues to decrease, whereas the validation error no longer improves. This indicates that additional training mainly enhances the fitting of the training samples rather than improving validation performance. The training process therefore finishes at epoch 48.

DOI: 10.54254/2755-2721/2027.39024

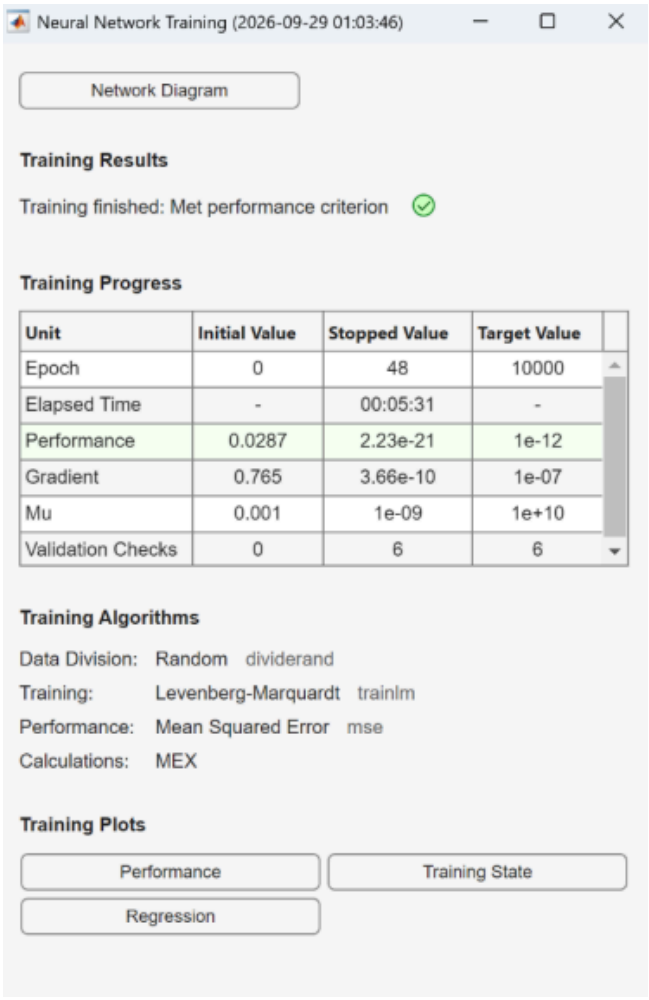


Figure 1. Summary of neural network training results

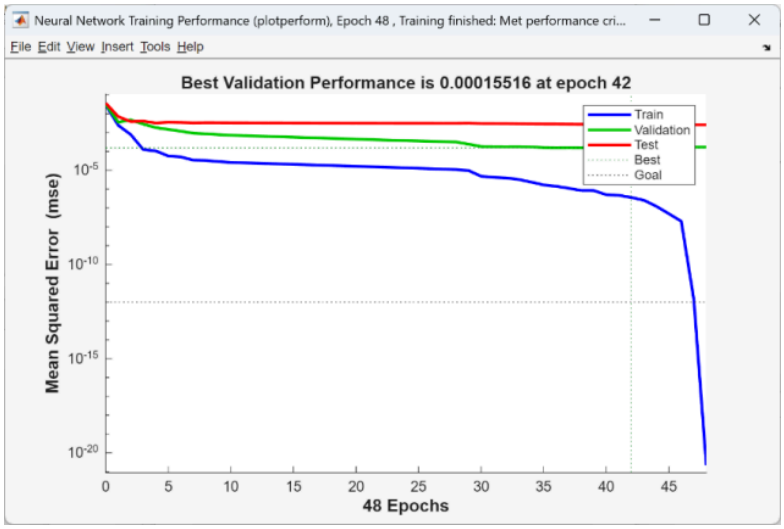


Figure 2. Neural network training performance curve (MSE)

Figure 3 shows the changes in the gradient, Mu value, and validation-check count during training. At epoch 48, the gradient decreases to 3.66×10^{-10} and the Mu value reaches 1×10^{-9}, while the

DOI: 10.54254/2755-2721/2027.39024

validation-check count reaches 6. Together with the performance curve, these results show that the network has reached a stable training state and that the validation set provides an effective safeguard against continued fitting after the best-validation epoch.

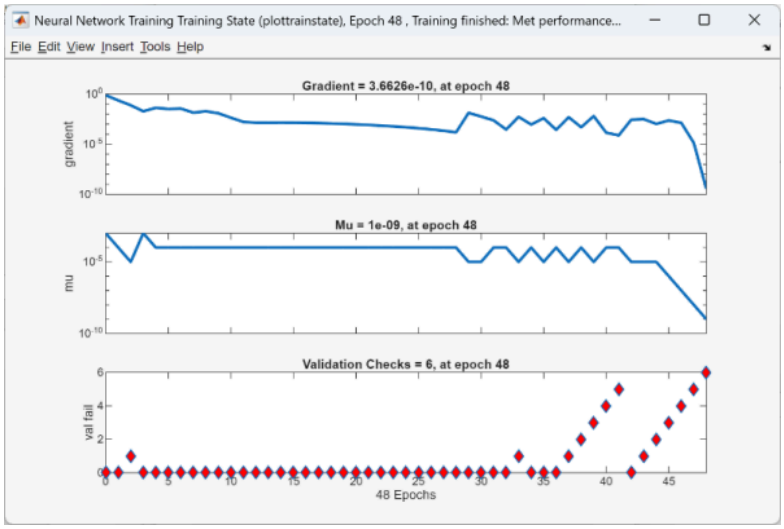


Figure 3. Training curve of neural network

After network training, the 5 reserved independent test samples were input into the model and the predicted microcavity radii were compared with the corresponding actual values. As shown in Figure 4, the predicted curve follows the overall variation of the actual microcavity radius, and most predictions are close to their target values, although small deviations remain for individual samples. The independent test result gives $R^2 = 0.99048$, indicating a strong agreement between the predicted and actual microcavity radii for this set of unseen samples.

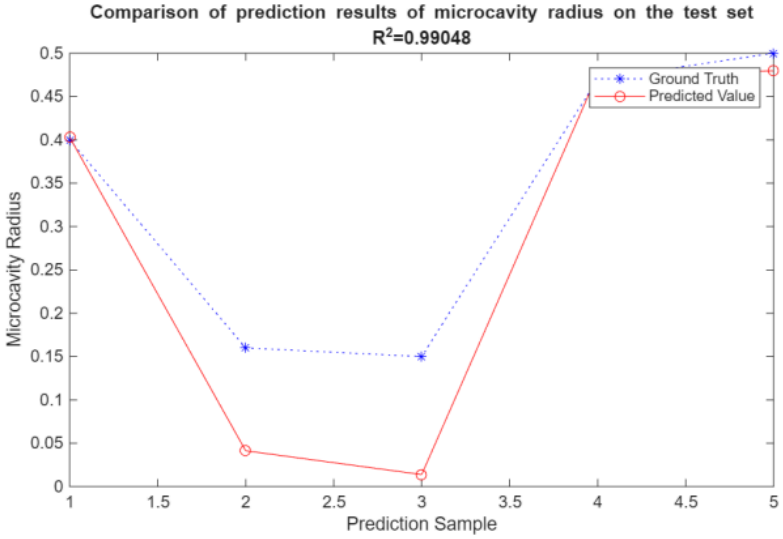


Figure 4. Comparison of prediction results for the radius of independent test set microcavities

The independent test results indicated that when using the 600-dimensional TM band frequencies as the input, the feedforward neural network could extract information related to the microcavity size from the high-dimensional band data and complete the regression prediction of the continuous microcavity radius. It should be noted that the independent test samples of this study are only 5

DOI: 10.54254/2755-2721/2027.39024

groups, so this result is mainly used to verify the feasibility of the structural parameter inversion scheme; for the prediction stability under a larger range of structural parameters and a more complex design space, further verification is still needed by expanding the sample size.

To further analyze the fitting performance across different data subsets, regression analysis was conducted on the training set, validation set, internal test set, and all modeling data, as shown in Figure 5. The corresponding correlation coefficients are R = 0.99999 for the training set, R = 0.99732 for the validation set, R = 0.9474 for the internal test set, and R = 0.98895 for all data. The training and validation results are both close to the ideal relationship, while the internal test set shows slightly larger deviations. Overall, the revised model exhibits substantially improved consistency across the data subsets and a reduced generalization gap compared with the previous experimental results.

Figure 5 presents the regression relationships between the network outputs and the target microcavity radii for the training, validation, internal test, and overall datasets. The regression points for the training and validation sets lie very close to the ideal Y = T line, corresponding to R values of 0.99999 and 0.99732, respectively. The internal test set achieves R = 0.9474, while the overall dataset reaches R = 0.98895. These results indicate that the network effectively captures the nonlinear relationship between the 600-dimensional TM-band features and the microcavity radius, although the slightly lower internal-test correlation suggests that additional independent samples are still desirable for a more robust assessment of generalization.

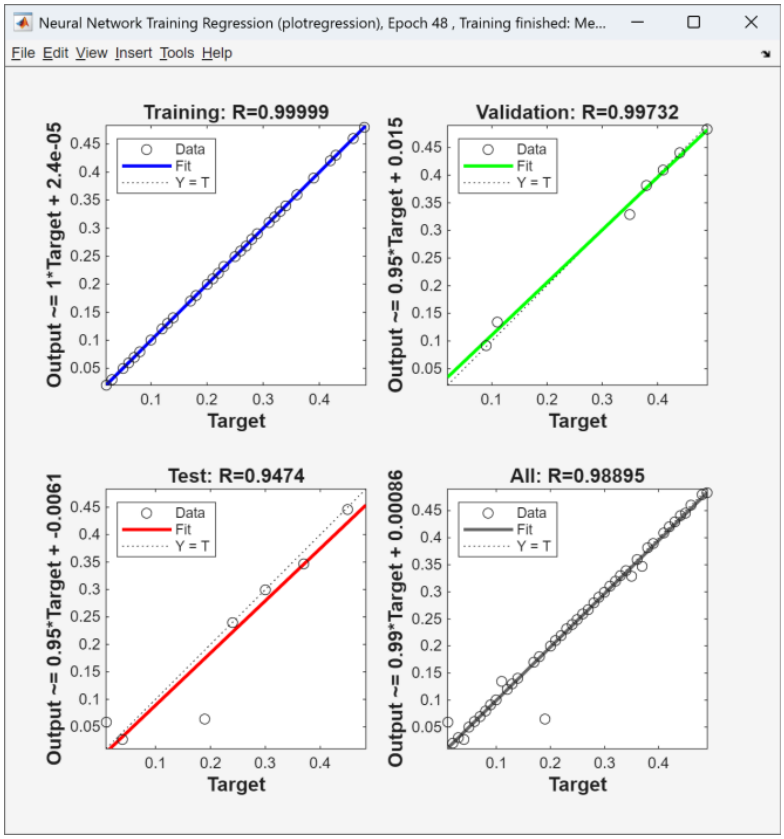


Figure 5. Neural network fitting regression analysis chart

Based on the MSE curve, regression analysis, and independent test results, the revised network has learned the nonlinear mapping relationship between photonic-crystal band characteristics and microcavity radius. The training and validation correlations are both close to 1, and the internal test

DOI: 10.54254/2755-2721/2027.39024

correlation remains above 0.94, showing that the generalization gap has been substantially reduced. Nevertheless, the dataset still contains only 50 structural samples, whereas each sample has 600-dimensional band characteristics, so the task remains a high-dimensional, small-sample regression problem. The current results therefore provide strong evidence for the feasibility of band-feature-based microcavity parameter inversion, but the stability of the model should still be evaluated using a larger dataset, repeated random splits, or cross-validation in future work.

5. Conclusion

This study developed a data-driven artificial-neural-network method for the inversion of photonic crystal microcavity parameters. Using the plane-wave expansion method, 50 structural samples were generated by varying the microcavity radius from 0.01 to 0.50 with a step of 0.01. For each structure, 50 TM bands at 12 k-points were calculated and expanded into a 600-dimensional feature vector, which was used as the network input with the corresponding microcavity radius as the prediction target.

A three-layer feedforward neural network with 15 hidden neurons was trained using the Levenberg–Marquardt algorithm. The best validation performance was obtained at epoch 42 with an MSE of 1.15516×10^{-4}, and training finished at epoch 48. The correlation coefficients of the training, validation, internal test, and overall datasets were 0.99999, 0.99732, 0.9474, and 0.98895, respectively. In addition, the 5 independent test samples achieved $R^2 = 0.99048$. These results demonstrate that the network can effectively learn the nonlinear inverse relationship between photonic band features and microcavity radius and provide improved consistency across training, validation, and test data.

Although the revised model shows markedly improved generalization performance, the dataset remains limited to 50 structural samples with 600-dimensional inputs. Therefore, the present results should be regarded as a feasibility demonstration rather than a complete assessment of robustness across a broad design space. Future work should expand the simulation dataset, use repeated random splits or cross-validation, systematically compare network architectures, and evaluate robustness under noise perturbations. The framework can also be extended from single-parameter radius inversion to joint inversion of multiple structural parameters, providing a broader basis for data-driven inverse design of photonic crystal structures.

References

[1] Deng, R., Liu, W., & Shi, L. (2024). Inverse design in photonic crystals. *Nanophotonics, 13*(8), 1219–1237. https://doi.org/10.1515/nanoph-2023-0750

[2] Christensen, T., Loh, C., Picek, S., et al. (2020). Predictive and generative machine learning models for photonic crystals. *Nanophotonics, 9*(13), 4183–4192. https://doi.org/10.1515/nanoph-2020-0197

[3] Nikulin, A., Zisman, I., Eich, M., et al. (2022). Machine learning models for photonic crystals band diagram prediction and gap optimisation. *Photonics and Nanostructures – Fundamentals and Applications, 52*, 101076. https://doi.org/10.1016/j.photonics.2022.101076

[4] Liu, D., Tan, Y., Khoram, E., & Yu, Z. (2018). Training deep neural networks for the inverse design of nanophotonic structures. *ACS Photonics, 5*(4), 1365–1369. https://doi.org/10.1021/acsphotonics.7b01377

[5] Ma, W., Liu, Z., Kudyshev, Z. A., et al. (2021). Deep learning for the design of photonic structures. *Nature Photonics, 15*, 77–90. https://doi.org/10.1038/s41566-020-0685-y

[6] Jiang, J., Chen, M., & Fan, J. A. (2021). Deep neural networks for the evaluation and design of photonic devices. *Nature Reviews Materials, 6*, 679–700. https://doi.org/10.1038/s41578-020-00260-1

[7] Johnson, S. G., & Joannopoulos, J. D. (2001). Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis. *Optics Express, 8*(3), 173–190. https://doi.org/10.1364/OE.8.000173

DOI: 10.54254/2755-2721/2027.39024

[8] Ho, K. M., Chan, C. T., & Soukoulis, C. M. (1990). Existence of a photonic gap in periodic dielectric structures. *Physical Review Letters, 65*(25), 3152–3155. https://doi.org/10.1103/PhysRevLett.65.3152
[9] Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks, 5*(6), 989–993. https://doi.org/10.1109/72.329697