



Deep learning improves performance of topological bending waveguides

ITSUKI SAKAMOTO,^{1,3} SHO OKADA,¹ NOBUHIKO NISHIYAMA,¹
XIAO HU,² AND TOMOHIRO AMEMIYA^{1,4} 

¹Department of Electrical and Electronic Engineering, Tokyo Institute of Technology, Tokyo 152-8552, Japan

²Research Center for Materials Nanoarchitectonics (MANA), National Institute for Materials Science (NIMS), Ibaraki 305-0044, Japan

³sakamoto.i.ab@m.titech.ac.jp

⁴amemiya.t.ab@m.titech.ac.jp

Abstract: This study introduced design informatics using deep learning in a topological photonics system and applied it to a topological waveguide with a sharp bending structure to further reduce propagation loss. The sharp bend in the topological waveguide composed of two photonic crystals wherein dielectrics having C_{6v} symmetry were arranged in triangle lattices of hexagons, and the designing of parameters individually for 6×6 unit cells near the bending region using deep learning resulted in an output improvement of 60% compared to the initial structure. The proposed structural design method has high versatility and applicability for various topological photonic structures.

© 2024 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

The research field focused on tracing the topology of electron systems in topological insulators and Weyl semimetals to photon systems is referred to as topological photonics, which has been progressing rapidly in recent years [1–4]. The greatest feature of topological photonics is the creation of photonic structures with mathematically different properties (i.e., different topologies) by simultaneously controlling interactions within and between unit cells in nonperiodic structures. Their benefit is that their local behavior is determined by a global parameter (e.g. spin Chern number), which is broadly insensitive to local perturbations. This facilitates more flexible design of devices compared to photonic crystals [5–8], which primarily interact between unit cells, and metamaterials [9–15], which primarily interact within unit cells.

One of the most well-known phenomena in topological photonics is the topological edge state that occurs at the interface of two photonic crystals distinct in topology. This allows the propagation of light with a specific orbital angular momentum and spin [16,17]. Consequently, unprecedented novel optical functions can be realized on optical circuits, such as light propagation resistant to sharp bends and unidirectional propagation dependent on circular polarization [18–23].

Figure 1 shows a series of elements that are responsible for propagation, input/output, and branching in a topological photonic system formed on an optical circuit; however, these are realized by combining two photonic crystals distinct in topology [16,17,24–26]. Here, if the parameters could be individually optimized for each unit cell constituting the crystal in the vicinity of each device (for example, the yellow region in Fig. 1), the device performance may be improved. However, in reality, the number of parameters increases exponentially, thus, this is not realistic in terms of calculation costs.

Deep learning is a machine learning method that uses neural networks (particularly multi-layered neural networks) that reproduce the mechanism of human neurons. It has been introduced in various fields such as image recognition, speech recognition, and language translation [27–29]. Under such circumstances, the introduction of deep learning has begun in the design of various

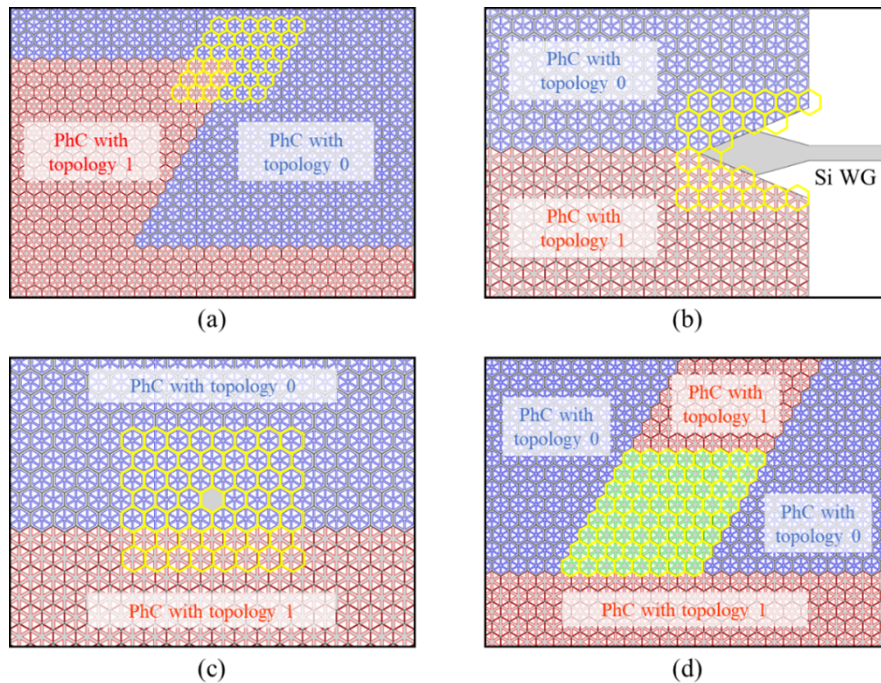


Fig. 1. Series of elements responsible for basic operations in a topological photonics system formed on an optical circuit. (a) Sharp bending propagation, (b) horizontal coupling with silicon waveguides, (c) vertical coupling with free space, and (d) branching at a suitable intensity ratio. In the vicinity of each device (yellow region), if it is possible to optimize the parameters individually for each unit cell constituting the crystal, there is a possibility that device performance can be improved.

photonic structures [30,31]. In this study, we introduced design informatics using deep learning for the first time in a topological photonics system. Adapting the above method to a topological waveguide with a sharp bending structure, the propagation loss was reduced. This facilitated the optimization for each unit cell of the topological photonic crystal constituting the device, thereby enabling more flexible device design.

2. Base element structure

The topological photonic crystals used in this study adopted a typical structure wherein dielectrics having C_{6v} symmetry were arranged in triangle lattices of hexagon [3]. The unit cell is shown in Fig. 2(a). The structure within the unit cell was determined by three design parameters: the length L of one side of the equilateral triangular hole, distance R from the center of the unit cell to the center of gravity of the equilateral triangular hole, and period a of the unit cell. Here, the period a was fixed at 730 nm, and the design parameters (L , R) of the two reference photonic crystals distinct in topology were set to (281, 213) and (284, 264) nm, respectively, to facilitate operations in the 1.55 μm range, which is the optical communication band (also see Appendix for band structures of the two reference photonic crystals).

In this study, we considered the formation of a sharp bend in a topological waveguide composed of two photonic crystals, and improved the bending loss by designing the parameters individually for 6×6 unit cells in the vicinity of the bending region, as shown in Fig. 2(b) (The range of 6×6 was determined to be appropriate based on the mode distribution profiles from the propagation analysis). Here, when two parameters L and R were considered in 6×6 unit cells, a total of 72

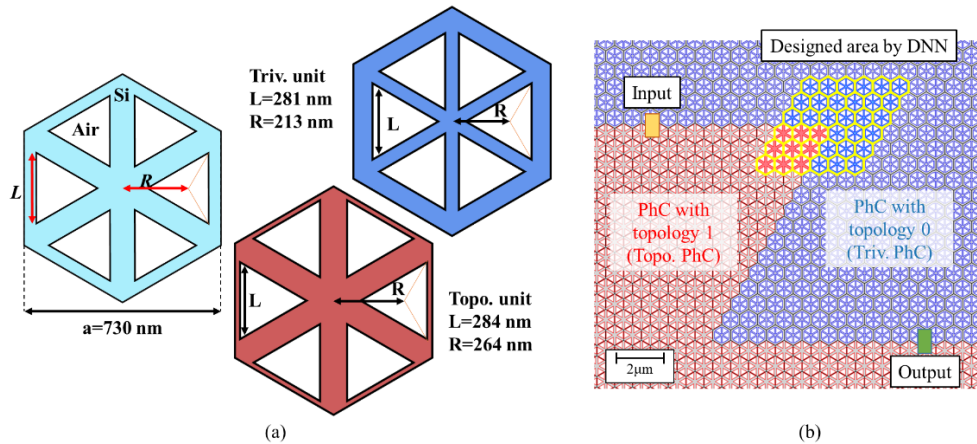


Fig. 2. (a) Unit cell of the topological photonic crystal used in this study. A typical structure in which dielectrics having C_{6v} symmetry were arranged in triangle lattices of hexagon. (b) A topological waveguide with a sharp bend used in the analysis. We consider improving the bending loss by designing the parameters individually for 6×6 unit cells near the bending region.

parameters were obtained. Owing to the impossibility of maximizing the coupling efficiency while changing all these parameters deep learning was considered to be effective from such a perspective.

In this design informatics, we used convolutional neural networks (CNN), which are widely used during image recognition. When using CNN for image recognition, multiple channels (R, G, B) can be considered to specify the color of each pixel. Further, we can refer to pixel position-dependent features that are not considered in fully coupled networks. The above features are similar to the fact that each unit cell of a topological photonic crystal has each design parameter (L , R) and that the entire device has acquired a two-dimensional structure combining unit cells. Thus, CNN is suitable for design informatics in topological photonics.

The overall flow of structural design by deep learning is as follows. First, the datasets to be used for training were converted (Section 3). The network was then configured and optimized by learning these datasets (Section 4). Subsequently, the optimized network was used to determine the structure with high coupling efficiency according to the search algorithm (Section 5).

3. Conversion of datasets used for learning

In the element structure shown in Fig. 2(b), a dataset for deep learning was acquired. Among the design parameters (L , R) in each unit cell inside the 6×6 region near the bend, only R was used as a parameter for deep learning considering the errors when actually fabricating the device. Thus, R_i ($i = 1-36$) of each unit cell in the 6×6 region, and the output value P of the device at that time were used as a dataset.

First, R_i was randomly changed according to a Gaussian distribution centering on the design parameter $R = 231$ and 264 nm of the two reference photonic crystals in order not to break the topological protection in the waveguide (see Fig. 3(a)). Herein, L was fixed at 281 and 284 nm for the two reference photonic crystals distinct in topology, respectively. For the output value P , propagation analysis was conducted employing the finite difference time domain (FDTD) method, and 2000 trials were conducted to read the output intensity for R_i that was randomly. Here, Synopsys Rsoft Photonic Device Tools was used for this simulation. In simulation, Rectangle was used as the light source, and PML was used as the boundary condition. The refractive

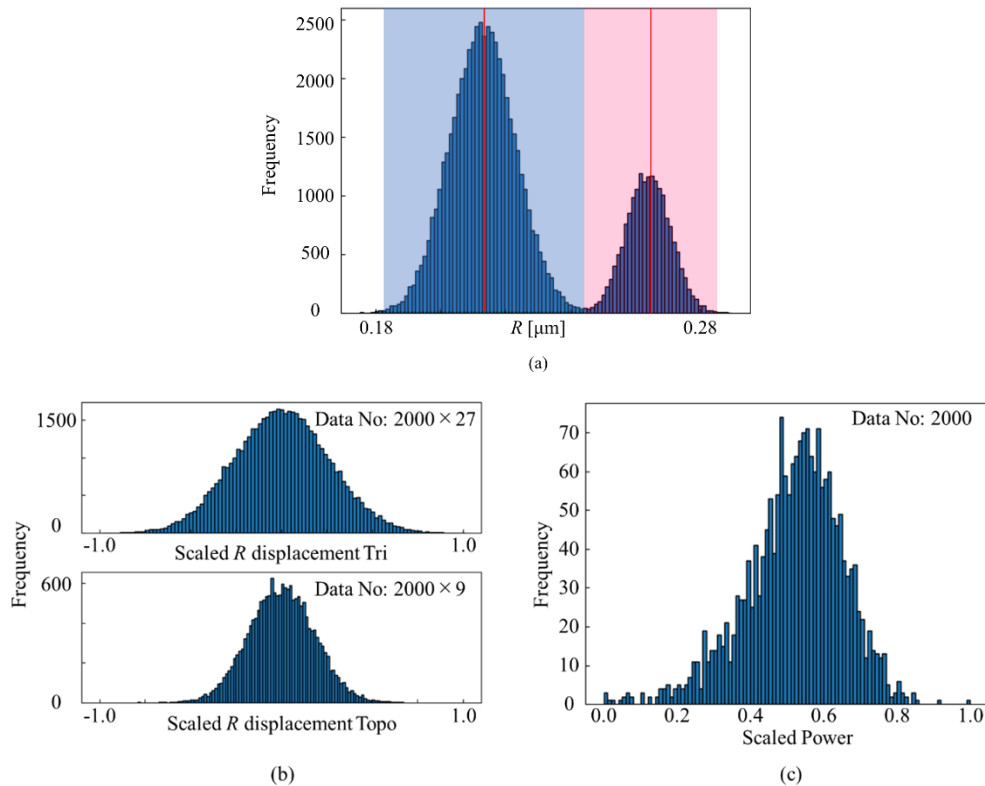


Fig. 3. (a) The Gaussian distribution of the parameter R_i centering on the design parameter $R = 231$ and 264 nm of the two reference photonic crystals. Histogram of rescaled input data. (b) Design parameter R_i of each unit cell: Displacement value from the R value of the reference photonic crystal. (c) Output value P : Output intensity value after propagation.

index of silicon was set to 3.4764, and the output intensity P was set to monitor intensity/input intensity (a.u.). Under specific computing resources (Intel Xeon Gold 6238R @ 2.20GH, Core: 28, Memory: 768 GiB), the time required per simulation was about 20 min. Consequently, 2000 combination datasets of the variable R_i and output value P for 36 unit cells were obtained.

Next, the obtained datasets were normalized to render them suitable for learning. Figure 3(b) and 3(c) shows the histogram of the datasets after normalization. At this time, the variable R_i was converted into a displacement value from the value of R of the reference photonic crystal, and rescaled such that the displacement value range was -1 to 1 to further improve the prediction accuracy. In addition, the output value P was rescaled such that the range of value was 0 – 1 .

4. Network configuration and optimization

Figure 4 shows the configuration diagram of the network used in this study. In this configuration, padding was performed along the vertical and horizontal directions to provide parameters of the peripheral structure (the parameter given by padding was 0). For a matrix of 10×10 after padding, eight types of feature maps were created by performing convolution using a 3×3 kernel (at this time, stride was 1 in both directions). Similarly, in the second layer, convolution was performed using a 3×3 filter, and 16 types of feature maps were generated. Thereafter, the third layer was connected to 576 neurons via the ReLU activation function, the fourth layer to 96

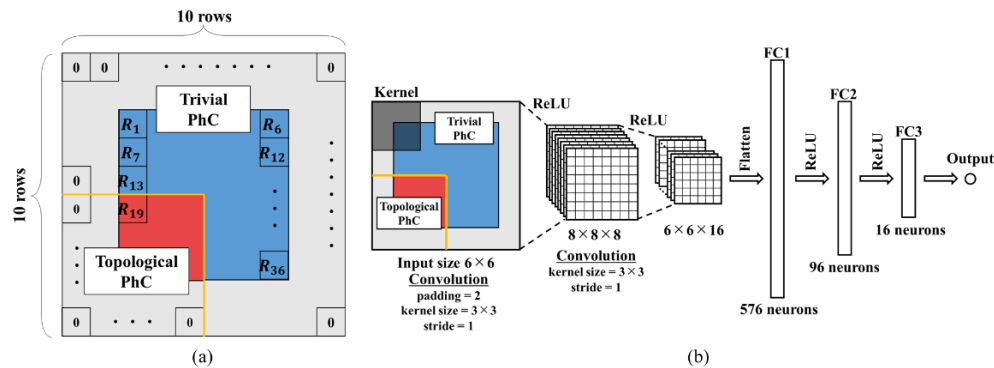


Fig. 4. Configuration diagram of the network used in this study. By convolving a 10×10 matrix after padding using a 3×3 kernel, eight types of feature maps were generated in the first layer and 16 types in the second layer. The third layer was connected to 576 neurons via the ReLU activation function, the fourth layer to 96 neurons via ReLU, and the fifth layer to 16 neurons via ReLU by full coupling.

neurons via ReLU, and the fifth layer to 16 neurons via ReLU by full coupling. Finally, they were combined to obtain one output value.

The above network was trained using the prepared datasets. Of the 2000 datasets prepared in Section 3, 1900 were used as training data for the network, and the remaining 100 were used as test datasets. The loss function in the learning was set to mean squared error (MSE), the batch size was set to 100, and the learning rate was set to 0.001. The learning was repeated until the loss of the test datasets converged using Adam as the optimization algorithm. MSE is a loss function used in regression problems and is sensitive to noise in the training data. Since the simulation data in this study did not contain noise, we adopted MSE. Adam is now widely used and is a common method [32]. Figure 5(a) shows the number of times of learning vs. prediction error. As evident, the internal parameters of the network were changed in an appropriate direction.

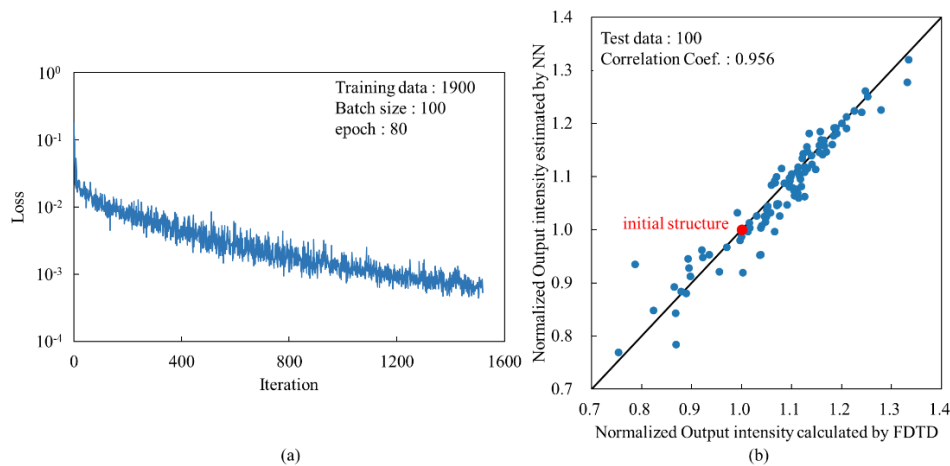


Fig. 5. (a) Number of times of learning vs. prediction error for 1900 training data. (b) Accuracy verification results of the output value obtained by propagation analysis by the FDTD method for 100 test data and the predicted output value by the learned network.

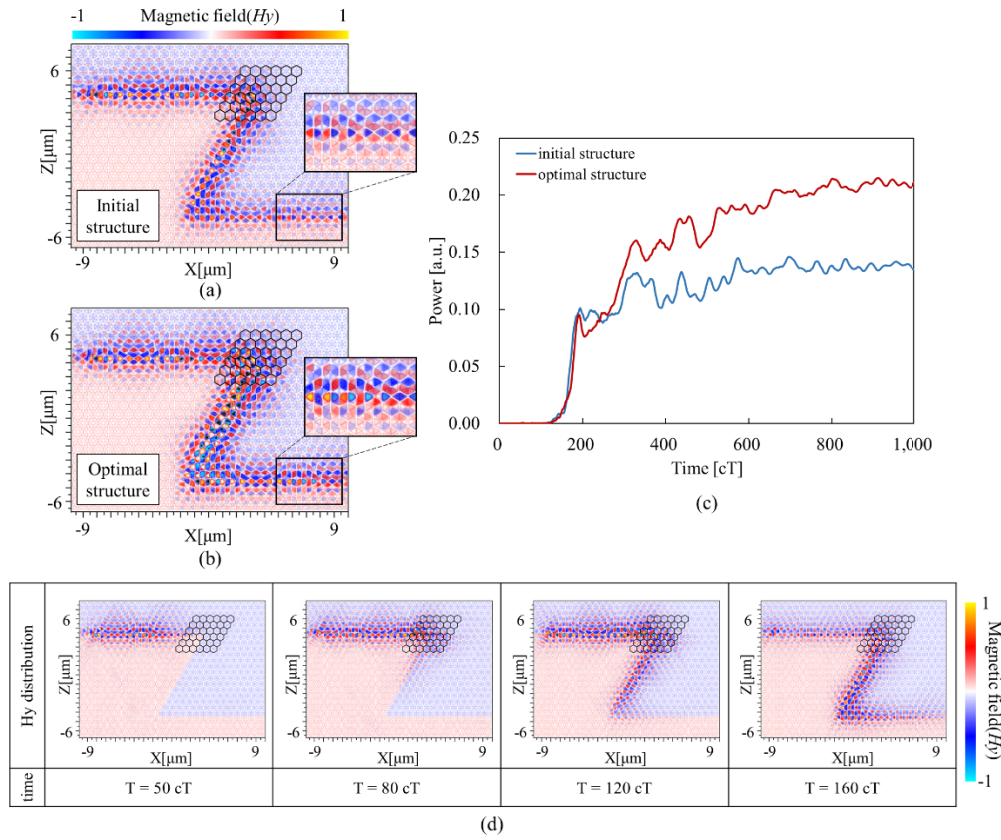


Fig. 6. (a) Propagation mode distribution in the initial structure without deep learning process. (b) Propagation mode distribution in the optimal structures obtained by deep learning process. (c) Time dependence of the output propagation intensity for the initial and optimal structures. (d) Time-dependent propagation mode distributions in the optimal structure obtained by deep learning ($1 \text{ cT} = 1/3 \times 10^{-15} \text{ s}$).

Figure 5(b) shows the accuracy verification results of the output value P obtained using the propagation analysis by the FDTD method for 100 test data and the predicted output value P by the learned network. The correlation coefficient at this time was 0.956, and it was observed that the propagation characteristics of the topological waveguide whose parameters were designed individually for 36 unit cells (6×6 unit cells) near the bending region were predictable by using this learned network.

5. Structural optimization using learned network

Finally, the 36 unit cells (6×6 unit cells) near the bending region were optimized using the learned network to minimize the bending loss of the topological waveguide. In the learned network, the predicted output value P is expressed as:

$$P = f_{NN}(R_1, R_2, R_3, R_4, \dots, R_{35}, R_{36}) \quad (1)$$

At this time, the following mean squared error is expressed as the loss function.

$$Loss = (P - \hat{P})^2 \quad (2)$$

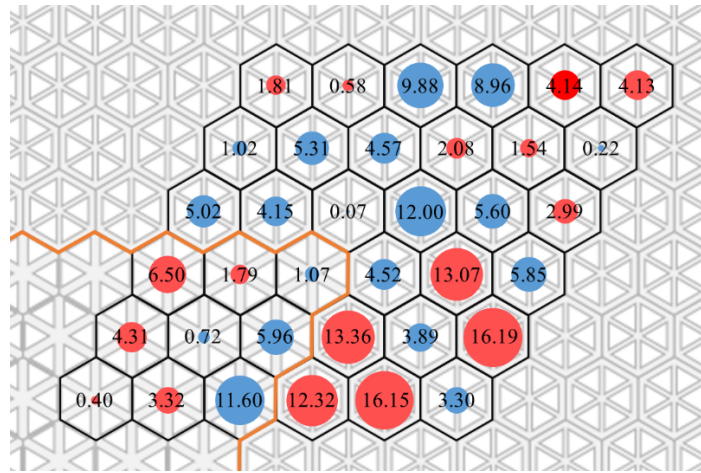


Fig. 7. Amount of shift from the initial structure of each unit cell in the optimal structure obtained by deep learning. Red circles indicate the increment (%) of R from the initial structure, and blue circles indicate the decrement (%) of R from the initial structure.

where \hat{P} represents the target output value.

First, using the structure R_i ($i = 1 \sim 36$) that yielded initial value, R_i was updated according to the following equation based on the gradient descent method.

$$R_i \leftarrow R_i + \eta \frac{\partial Loss}{\partial R_i} \quad (3)$$

where $\eta = 0.001$ was set, and the target output value \hat{P} was set to a value slightly higher than the output value P achieved by the structure. After updating the R_i until Eq. (3) converged, the target output value \hat{P} was newly reset, and the above calculation was repeated.

Figures 6(a) and 6(b) shows the propagation mode distributions in the initial structure without deep learning process (i.e. the structure composed only of two photonic crystals distinct in topology) and the optimal structure obtained by deep learning process, respectively. From the mode intensities after passing through the bending region, it can be seen that the bending loss is suppressed in the optimal structure, and the mode shapes are also maintained as in the initial structure. Figure 6(c) shows the time dependence of the output intensity for the initial and optimal structures (the time dependence of the mode distribution in the optimal structure is also shown in Fig. 6(d)). In the calculations, no changes were made to any structures except for the bending. Therefore, it can be concluded that the output improvement seen by comparing these results is due only to the bending structure, which was designed by deep learning. In the optimal structure, the stable output value P at a certain time after the propagating light reached the output port was 0.211. Compared with the output value 0.135 of the initial structure, an output improvement of approximately 1.6 times (60%) was confirmed. The best output value P for a total of 2000 data sets was 0.196, confirming that the best value for the data sets was exceeded.

Topologically protected waveguides are inherently tolerant to steep bending and exhibit high propagation efficiency. Therefore, in order not to break the topological protection in the waveguide, as shown in Fig. 3, the deep learning design was also performed with parameter distributions that do not deviate significantly from the two reference photonic crystals. While this ensures that topological protection is not broken even in the vicinity of optimized steep bending structures, the modes of light are appropriately broadened due to the change in photonic bands in the vicinity of the bending, thereby improving the propagation efficiency by pseudo-relaxation of the sharp bending.

Figure 7 shows the amount of shift from the initial structure of each unit cell in the optimal structure. Here, the red circles indicate the magnitude of the increase rate (%) of R from the initial structure, and the blue circles indicate the magnitude of the decrease rate (%) of R from the initial structure. Thus, there is no clear regularity in the structure distribution of unit cells, and it is useful to adopt the method using deep learning.

6. Conclusion

This study proposed a computational method based on deep learning to improve the propagation characteristics for sharp bending structures in topological photonic crystals. The structural design method used in this study has high versatility for topological photonic structures. Thus, this method is expected to be applied to various topological photonic systems.

Appendix

Figure 8 shows the photonic band structures of the two reference topological photonic crystals used in this study.

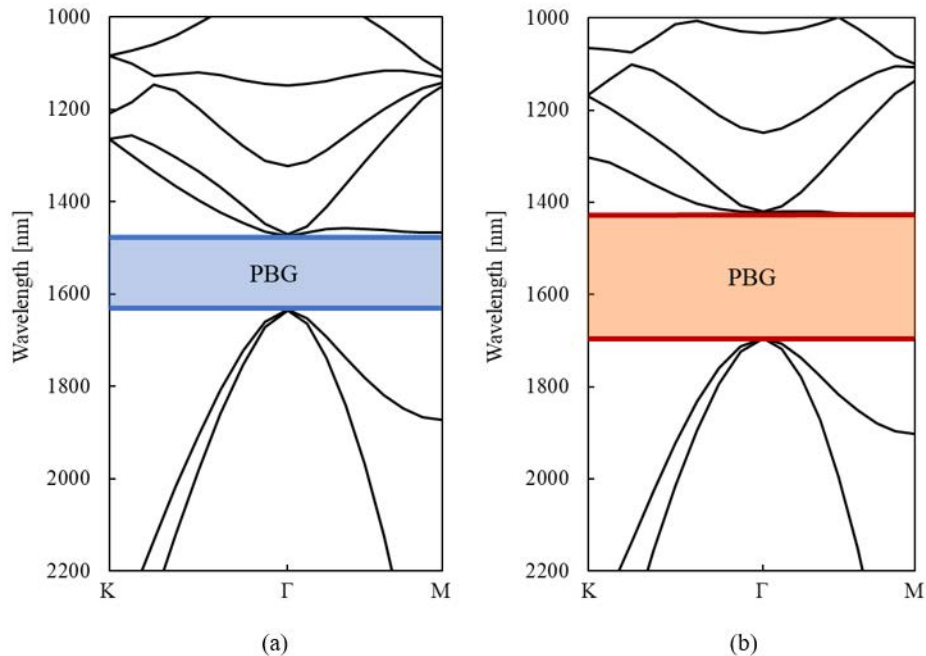


Fig. 8. Band structures of the two reference photonic crystals : (a) (L, R) = (281, 213), (b) (L, R) = (284, 264) nm.

Funding. Japan Society for the Promotion of Science (#22H01520); Adaptable and Seamless Technology Transfer Program through Target-Driven R and D (JPMJTR22RG); Core Research for Evolutional Science and Technology (JPMJCR18T4).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. F. D. M. Haldane and S. Raghu, "Possible realization of directional optical waveguides in photonic crystals with broken time-reversal symmetry," *Phys. Rev. Lett.* **100**(1), 013904 (2008).
2. Z. Wang, Y. Chong, J. D. Joannopoulos, *et al.*, "Observation of unidirectional backscattering-immune topological electromagnetic states," *Nature* **461**(7265), 772–775 (2009).
3. L.-H. Wu and X. Hu, "Scheme for achieving a topological photonic crystal by using dielectric material," *Phys. Rev. Lett.* **114**(22), 223901 (2015).
4. T. Ozawa, H. M. Price, A. Amo, *et al.*, "Topological photonics," *Rev. Mod. Phys.* **91**(1), 015006 (2019).
5. J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, "Photonic crystals: putting a new twist on light," *Nature* **386**(6621), 143–149 (1997).
6. T. Baba, "Slow light in photonic crystals," *Nat. Photonics* **2**(8), 465–473 (2008).
7. K. Kondo, T. Tatebe, S. Hachuda, *et al.*, "Fan-beam steering device using a photonic crystal slow-light waveguide with surface diffraction grating," *Opt. Lett.* **42**(23), 4990–4993 (2017).
8. T. Asano, Y. Ochi, Y. Takahashi, *et al.*, "Photonic crystal nanocavity with a Q factor exceeding eleven million," *Opt. Express* **25**(3), 1769–1777 (2017).
9. J. B. Pendry, A. J. Holden, D. J. Robbins, *et al.*, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Trans. Microwave Theory Tech.* **47**(11), 2075–2084 (1999).
10. R. A. Shelby, D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science* **292**(5514), 77–79 (2001).
11. T. Amemiya, T. Shindo, D. Takahashi, *et al.*, "Nonunity permeability in metamaterial-based GaInAsP/InP multimode interferometers," *Opt. Lett.* **36**(12), 2327–2329 (2011).
12. N. I. Zheludev and Y. S. Kivshar, "From metamaterials to metadevices," *Nat. Mater.* **11**(11), 917–924 (2012).
13. T. Amemiya, T. Kanazawa, S. Yamasaki, *et al.*, "Metamaterial waveguide devices for integrated optics," *Materials* **10**(9), 1037 (2017).
14. T. Amemiya, S. Yamasaki, M. Tanaka, *et al.*, "Demonstration of slow-light effect in silicon-wire waveguides combined with metamaterials," *Opt. Express* **27**(10), 15007–15017 (2019).
15. M. Tanaka, T. Amemiya, H. Kagami, *et al.*, "Control of slow-light effect in a metamaterial-loaded Si waveguide," *Opt. Express* **28**(16), 23198–23208 (2020).
16. M. Kim, Y. Kim, and J. Rho, "Spin-valley locked topological edge states in a staggered chiral photonic crystal," *New J. Phys.* **22**(11), 113022 (2020).
17. N. Parappurath, F. Alpeggiani, L. Kuipers, *et al.*, "Direct observation of topological edge states in silicon photonic crystals: Spin, dispersion, and chiral routing," *Sci. Adv.* **6**(10), eaaw4137 (2020).
18. S. Peng, N. J. Schilder, X. Ni, *et al.*, "Probing the band structure of topological silicon photonic lattices in the visible spectrum," *Phys. Rev. Lett.* **122**(11), 117401 (2019).
19. D. Smirnova, S. Kruk, D. Leykam, *et al.*, "Third-harmonic generation in photonic topological metasurfaces," *Phys. Rev. Lett.* **123**(10), 103901 (2019).
20. Z. K. Shao, H. Z. Chen, S. Wang, *et al.*, "A high-performance topological bulk laser based on band-inversion-induced reflection," *Nat. Nanotechnol.* **15**(1), 67–72 (2020).
21. Z. Q. Yang, Z. K. Shao, H. Z. Chen, *et al.*, "Spin-momentum-locked edge mode for topological vortex lasing," *Phys. Rev. Lett.* **125**(1), 013903 (2020).
22. H. Yoshimi, T. Yamaguchi, Y. Ota, *et al.*, "Slow light waveguides in topological valley photonic crystals," *Opt. Lett.* **45**(9), 2648–2651 (2020).
23. Y. Ota, F. Liu, R. Katsumi, *et al.*, "Photonic crystal nanocavity based on a topological corner state," *Optica* **6**(6), 786–789 (2019).
24. H. Kagami, T. Amemiya, S. Okada, *et al.*, "Topological converter for high-efficiency coupling between Si wire waveguide and topological waveguide," *Opt. Express* **28**(22), 33619–33631 (2020).
25. H. Kagami, T. Amemiya, S. Okada, *et al.*, "Highly efficient vertical coupling to a topological waveguide with defect structure," *Opt. Express* **29**(21), 32755–32763 (2021).
26. T. Amemiya, S. Okada, H. Kagami, *et al.*, "Splitter of topological photonic waveguide in semiconductor platform," *Research Square*, PPR682421 (2023).
27. A. Radford, J. W. Kim, C. Hallacy, *et al.*, "Learning transferable visual models from natural language supervision. In: International conference on machine learning," In: *Proceedings of Machine Learning Research*, 2021. p. 8748–8763.
28. R. Rombach, A. Blattmann, D. Lorenz, *et al.*, "High-resolution image synthesis with latent diffusion models," In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022. p. 10684–10695.
29. A. Vaswani, N. Shazeer, and N. Parmar, *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems*, **30**, 2017.
30. T. Asano and S. Noda, "Optimization of photonic crystal nanocavities based on deep learning," *Opt. Express* **26**(25), 32704–32717 (2018).
31. R. Li, X. Gu, K. Li, *et al.*, "Deep learning-based modeling of photonic crystal nanocavities," *Opt. Mater. Express* **11**(7), 2122–2133 (2021).
32. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv*, arXiv:1412.6980 (2014).