Applied Physics Letters ARTICLE pubs.aip.org/aip/apl

Improved spectral filtering of broadband diffractive neural network by loss function engineering

Cite as: Appl. Phys. Lett. **126**, 081107 (2025); doi: 10.1063/5.0257384
Submitted: 9 January 2025 · Accepted: 13 February 2025 ·
Published Online: 26 February 2025

View Online Export Citation CrossMark

Bolin Li,[1,2] Guangrui Luan,[1,2] Yinfei Zhu,[2] Jinlei Fei,[2] Min Gu,[1,2] and Jian Lin[1,2,a)]

AFFILIATIONS

[1]School of Artificial Intelligence Science and Technology, University of Shanghai for Science and Technology, Shanghai, China
[2]Institute of Photonic Chips, University of Shanghai for Science and Technology, Shanghai, China

[a)]**Author to whom correspondence should be addressed:** jianlin@usst.edu.cn. **Tel.**: +65-65168856. **Fax**: +65-68723069

ABSTRACT

We engineer the loss function by removing the conventional physics-based energy constraint during the training of broadband diffractive neural networks (DNNs) to enhance their spectral filtering capabilities of supercontinuum light. Simulations show that compared to DNNs trained with conventional loss function, the suppression of out-of-band spectral intensities can be improved by three orders of magnitude, resulting in an extinction coefficient of 10^{-6}. Additionally, the spectral resolution can be enhanced by over 50% with a 6.6% improvement of energy efficiency. These findings are corroborated by experiments conducted with a two-layer DNN. The proposed method holds promise for enhancing the performance of broadband DNNs across various applications, including spectral reconstruction, spectrum classification, and color image processing, among others.

Artificial neural networks (ANNs), as a subset of machine learning models, have attracted considerable attention due to their rapid development and widespread adoption across diverse fields,[1] such as image classification and processing,[2,3] nature language processing,[4] medical diagnosis,[5] as well as computer vision and autonomous driving.[6] Most ANNs are implemented via electronic computing; however, their performance currently faces significant challenges posed by the conventional von Neumann architecture,[7,8] notably the rapidly increasing energy consumption.[9]

In this context, optical neural networks (ONNs) exhibit the potential to surpass electronic counterparts by enabling high-speed parallel inference in the optical domain. Several ONN structures have been proposed, including photonic reservoir computing,[10] optical interference neural network,[11] photonic spiking neural network,[12] optical recurrent neural network,[13] and diffractive optical neural network (DNN).[14–25] Among them, DNNs possess advantages in terms of high parallelism and connectivity, and they have been applied in various tasks, such as image classification,[14,24] phase retrieval,[16] data encryption,[17] spectrum manipulation,[19,22,25] and image reconstruction.[18,20,23]

DNNs perform computations through the spontaneous interference of light at each layer following propagation, alongside controlled complex modulations of light phase and/or amplitude imposed by each pixel (referred to as neurons in DNNs) of the diffractive layers. The modulation depths of these pixels in every layer are trained using an error-backpropagation algorithm with a suitably designed loss function. By tuning the loss function, the performance of a DNN can be optimized.[19] Physical constraints are typically incorporated into the design of loss functions, such as the stipulation that the upper limit of a DNN's energy efficiency is 100%.[19,26] However, these conventional loss functions may restrict the overall performance of the DNN.

In this study, we design the loss function without imposing any physics-based energy constraint (PBEC) on the output signal during the training of a DNN to improve its spectral filtering capabilities for supercontinuum light. The loss function incorporates specific parameters to govern the intensities of in-band and out-of-band signals, allowing for their ratio to be adjusted to optimize the DNN's output. Simulation shows that compared to DNNs trained with conventional loss function, the out-of-band spectral intensities can be suppressed by up to 3 orders, resulting in an extinction coefficient of 10^{-6} with energy efficiency exceeding 70%. The spectral resolution can also be improved by more than 50%. A two-layer DNN constructed with a single SLM is employed to validate these results. The proposed loss function may prove beneficial in enhancing the performance of

broadband DNNs in various tasks including spectral reconstruction, spectrum classification, color image processing, and so on.

The schematic of the DNN model in this work is shown in Fig. 1(a). It contains two phase-only diffractive layers and one detector. Each layer has 512×512 square pixels with a pixel size of $8\,\mu$m. The detector is a square with 6×6 pixels at the center of detector plane. The network is trained by optimizing the phase values with respect to the loss function. Light modulated by the lth layer and propagating to the $(l+1)$th layer is calculated using the angular spectrum theory as follows:[27]

$$E_{l+1}(x,y,z,\lambda) = F^{-1}\{F\{E_l(x,y,0,\lambda)\cdot w_l(x,y,\lambda)\}\cdot H(u,v,z,\lambda)\}, \tag{1}$$

$$w_l(x,y,\lambda) = \exp\,[j\Phi_l(x,y,\lambda)], \tag{2}$$

$$H(u,v,z,\lambda) = \exp[jkz\sqrt{1-\lambda^2(u^2+v^2)}], \tag{3}$$

where F and F^{-1} represent the Fourier and inverse Fourier transforms; w_l is the modulation factor; Φ_l is the phase delay induced by the lth layer; H is a phase shifting factor describing the propagation of light for a distance of z. Coordinates u and v are spatial frequencies with $u = \frac{\cos\alpha}{\lambda}$ and $v = \frac{\cos\beta}{\lambda}$, in which α, β are the angles between wave vectors and x, y axis. The input spectrum ranges from 450 to 850 nm, which is discretized into 2035 uniformly distributed sampling points for training. In each iteration, 50 wavelengths are randomly selected and input into the neural network. The label of the training, i.e., the target spectrum $g(\lambda_i)$ centered at λ_0 with spectral width of $\frac{\Delta\lambda}{2}$ is designed as follows:

$$g(\lambda_i) = \begin{cases} -b\dfrac{(\lambda_{max}-\lambda_{min})}{\left(\lambda_0-\dfrac{\Delta\lambda}{2}-\lambda_{min}\right)}, & \lambda_i < \lambda_0-\dfrac{\Delta\lambda}{2}, \\ a\dfrac{(\lambda_{max}-\lambda_{min})}{\Delta\lambda}, & \lambda_0-\dfrac{\Delta\lambda}{2} < \lambda_i < \lambda_0+\dfrac{\Delta\lambda}{2}, \\ -b\dfrac{(\lambda_{max}-\lambda_{min})}{\left(\lambda_{max}-\lambda_0-\dfrac{\Delta\lambda}{2}\right)}, & \lambda_i > \lambda_0+\dfrac{\Delta\lambda}{2}, \end{cases} \tag{4}$$

where b and a are the inhibition and gain coefficients. The ratio of b:a is defined as the inhibition ratio. Then, the loss function is calculated by

$$L_1 = \sum_{i\in T} -g(\lambda_i)I_{out} \tag{5}$$

and the conventional loss function with PBEC can be expressed as

$$L_c = \sum_{i\in T} g(\lambda_i)(I_{in}(\lambda_i) - I_{out}(\lambda_i)), \tag{6}$$

where I_{out} is the power detected within the detector area; I_{in} is total power input into the DNN; T represents sequence numbers of the 50 discrete wavelengths randomly selected in each training batch. The resulting error is backpropagated to update the modulation factor [Eq. (2)] for the next iteration. The energy efficiency of wavelength λ_i with a certain bandwidth is evaluated by the following equation:

$$\eta = \sum_{i\in B}\frac{I_{out}(\lambda_i)}{I_{in}(\lambda_i)}, \tag{7}$$

where B represents the sequence numbers of discrete wavelengths within this band.

The schematic of the experimental setup is shown in Fig. 1(b). The light source is from a photonic crystal fiber (PCF; femtoWHITE 800, NKT Photonics Inc.) pumped by a femtosecond laser (Chameleon, Coherent Inc.) tuned to 800 nm. A half-wave plate and a polarization beam splitter (PBS) are used to adjust the pump power. The spectral shape of the generated supercontinuum can be adjusted by the half-wave plate and PBS after the PCF. A spatial light modulator (SLM; 1920×1200 pixels with an 8-μm pixel size; EXULUS-HD2, Thorlabs Inc.) and a mirror are used to construct the two phase-modulation layers. Each layer contains 512×512 neurons (pixels). Before shining onto the SLM, the collimated output of the PCF is expanded by a 4-f system to about 5 mm to match the sizes of the two layers. The light path from layer 1 to layer 2 is set as 10 cm by adjusting the position of the mirror. This distance is determined under two constrains: the light incident angle should be close to normal with a deviation less than 10° to ensure the operation of the SLM as product design; and the neurons of the first layer can be fully connected to each other on layer 2. The modulated light is then collected at the center of the detector plane by a multi-mode fiber with 50-μm core diameter connected to a spectrometer (USB4000, Ocean Optics Inc.) for data readout.

Figure 2 shows the simulated single-passband filtering performance of DNNs trained with and without PBEC incorporated into the loss function. The input spectrum has uniform intensity across all wavelengths. The target wavelengths for filtering are 550, 600, 650, 700, and 750 nm, each with a bandwidth of 16 nm. It can be seen that when conventional loss function is used, the filtered spectra are centered around the targeted wavelengths, but noticeable out-of-band signals are evident on both sides of the target peak across the spectral range. In contrast, when the PBEC is removed from the loss function in the DNN training process, the out-of-band signals are significantly suppressed, and the peak widths at all target wavelengths are noticeably reduced with limited intensity attenuation. These results demonstrate the effectiveness of the proposed DNN architecture for spectral filtering across a broad spectral range and emphasize the importance of removing the PBEC in the loss function to enhance the spectral filtering performance.

To further investigate the impact of removing the PBEC at varying inhibition ratios (0:1, 1:10, 1:5, 1:1, 5:1, and 10:1) on DNN spectral filtering performance, we calculate the filtered spectra targeting 600 nm, as shown in Fig. 3. Figure 3(a) shows that the inhibition ratio has minimal impact on the DNN output when PBEC is included in the loss function. However, after PBEC is removed, the intensity and spectral width of the output spectra decrease significantly with increasing inhibition ratio, particularly beyond 1:1. Figure 3(c) presents the energy efficiency at different inhibition ratios. When PBEC is included in the loss function, the energy efficiency remains relatively insensitive to changes in the inhibition ratio. After PBEC removal, energy efficiency increases by approximately 7% at low inhibition ratios (0:1, 1:10, 1:5) and then decreases as the inhibition ratio increases. Figures 3(d) and 3(e) display the same data as Figs. 3(a) and 3(b), respectively,

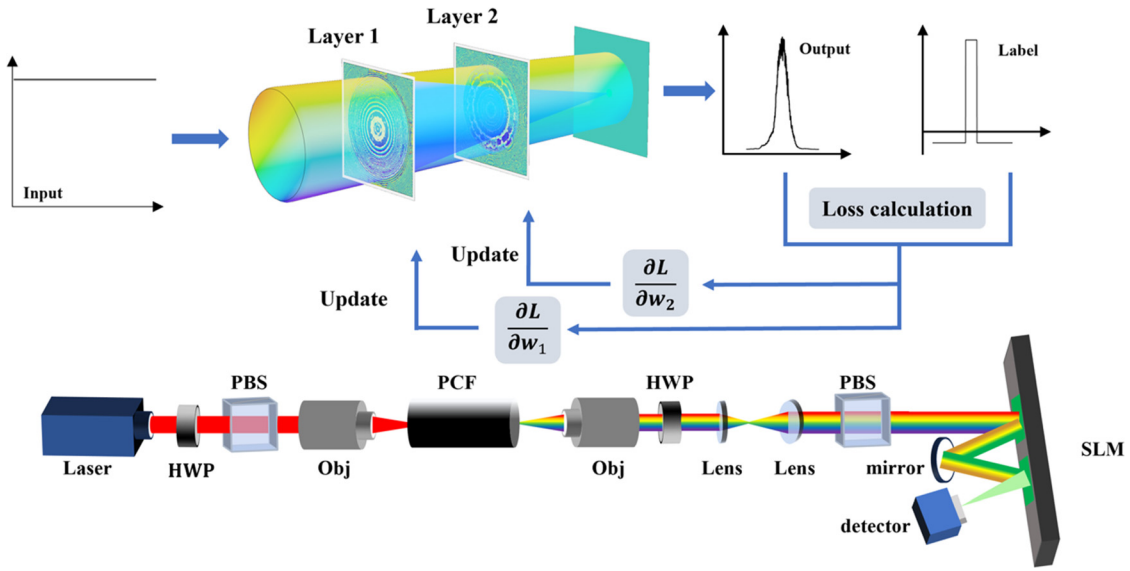

FIG. 1. (a) Illustration of the DNN model in this work. (b) Schematic of the experimental setup. HWP: half wave plate; PBS: polarization beam splitter; Obj: objective; PCF: photonic crystal fiber; SLM: spatial light modulator.

on a logarithmic scale. The results show that when PBEC is included, the normalized out-of-band signal intensity is approximately 10^{-3}–10^{-4}. However, after PBEC removal, this intensity is significantly reduced to around 10^{-6}–10^{-7} at moderate and high suppression ratios (1:1, 5:1, 10:1). These findings highlight the effectiveness of removing PBEC in suppressing out-of-band signals during the spectral filtering process of broadband DNNs.

Next, we investigate the performance of DNNs at various target spectral widths, as shown in Fig. 4. Figure 4(a) illustrates that when PBEC is included in the loss function, the output spectra become

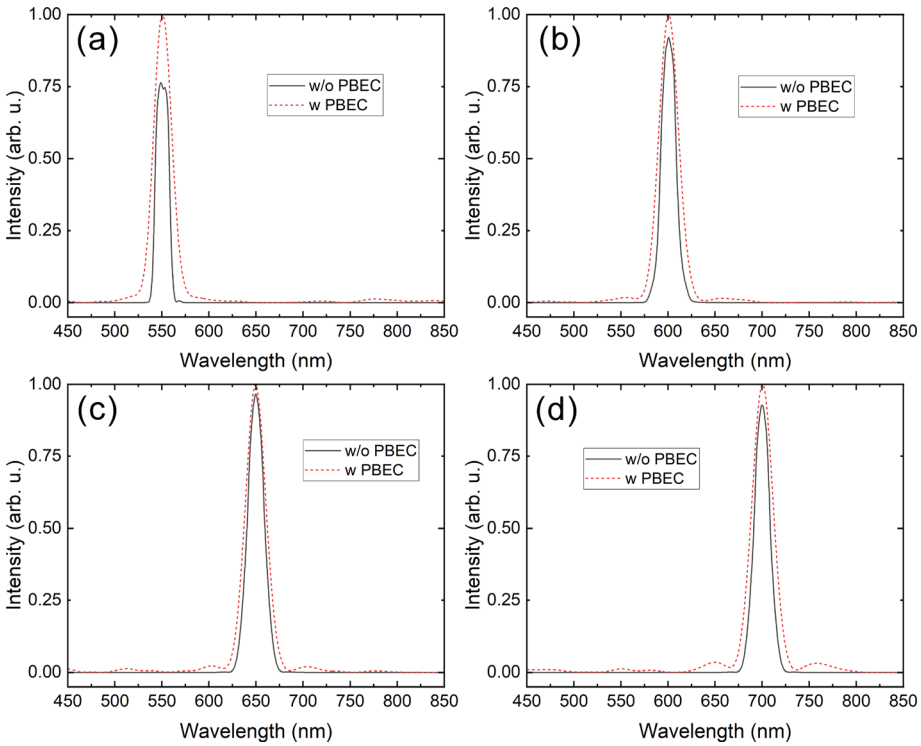

FIG. 2. Simulated spectra of single passband spectral filtering from DNN trained by loss function with and without PBEC. The target wavelengths are (a) 550, (b) 600, (c) 650, (d) 700 nm. The inhibition ratio is 1:1.

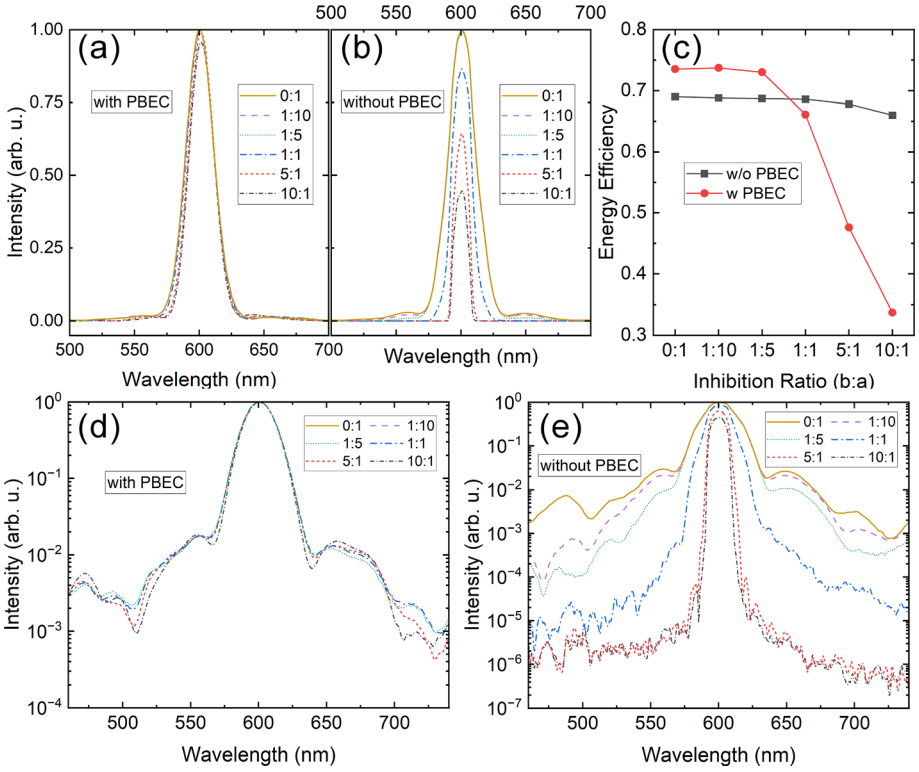

FIG. 3. Simulated spectra of single passband spectral filtering by DNN with various inhibition ratios displayed in (a) and (b) linear and (d) and (e) logarithmic scales and (c) the corresponding energy efficiencies. (d) Simulated FWHMs of DNN output spectra with and without ILF.

narrower as the target spectral width decreases. For target spectral widths smaller than 4 nm, the spectral shape becomes irregular, and at the extreme narrow target width of 0.16 nm, the spectral intensity nearly vanishes. In contrast, as shown in Fig. 4(b), when PBEC is removed, the spectra maintain a regular shape, and the spectral intensity per nanometer increases as the target spectral width decreases, with a slight decrease in intensity at a target width of 0.16 nm. Figure 4(c) quantitatively presents the full width at half maximum (FWHM) of the DNN output spectra. The results show that after PBEC removal, the spectral FWHM is significantly reduced, with more than a 50% reduction observed at the target width of 8 nm. Additionally, the FWHM values are approximately 2 nm larger than the target values, demonstrating an almost linear relationship with the target values. These results reveal that the spectral resolution can be substantially improved by removing the PBEC from the loss function.

To validate the simulated results, we build a two-layer DNN setup depicted in Fig. 1(b) for the processing of the supercontinuum generated by a PCF. As shown in Fig. 5(a), the supercontinuum covers a broad range of wavelengths from 450 to 850 nm and has relatively high intensities starting from 500 nm. Figures 5(b)–5(e) display the spectra filtered by DNNs both with and without PBEC in the loss function, targeting at 550, 600, 650, 700 nm, each with a bandwidth of 16 nm. It is evident that when PBEC is included in the loss function, filtered spectra exhibit either broader widths or noticeable side peaks due to insufficient attenuation of out-of-band signals. Conversely, after PBEC removal, out-of-band signals are effectively suppressed, closely

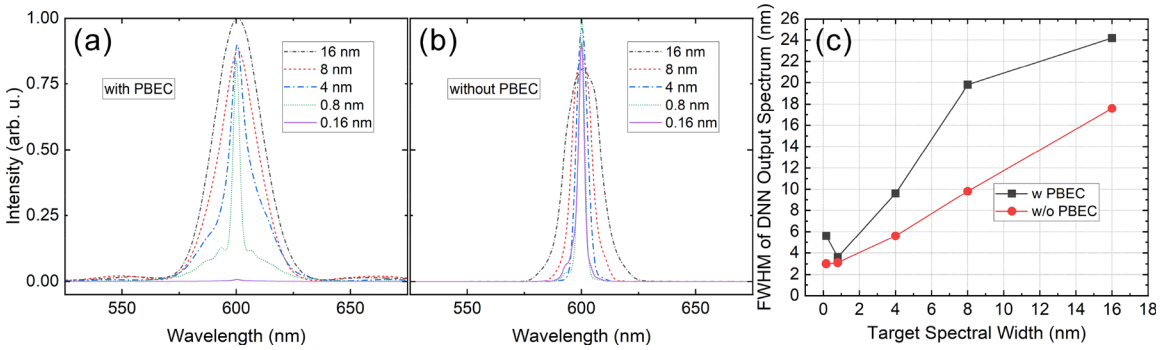

FIG. 4. Simulated output spectra of DNNs trained (a) with and (b) without PBEC at different target spectral widths (b:a = 1:1) and (c) the corresponding FWHMs.

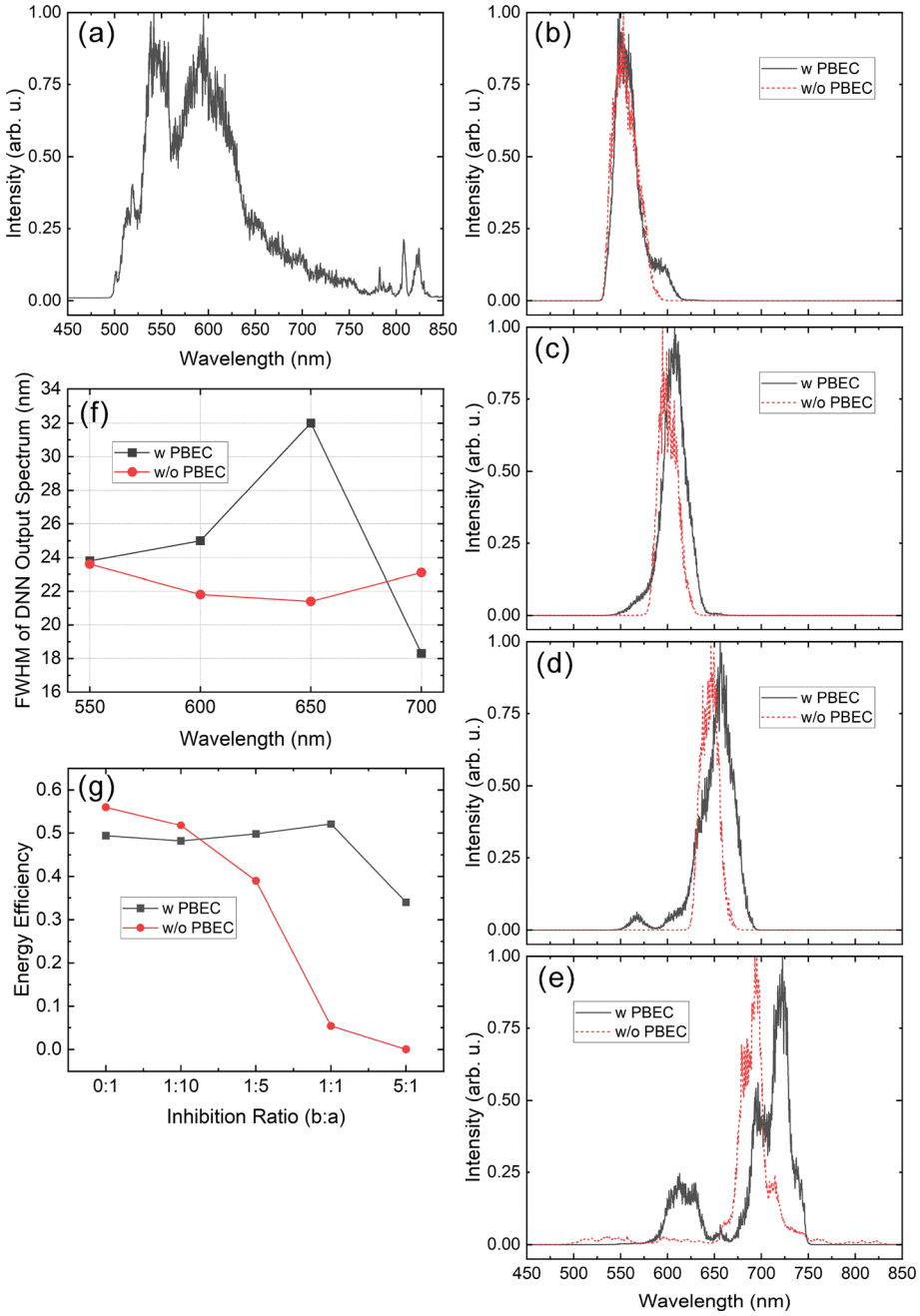

FIG. 5. Experimental results of spectral filtering of DNN. (a) Spectrum of supercontinuum measured before the SLM. (b)–(e) Filtered spectra targeted at wavelengths of 550, 600, 650, 700 nm with and without PBEC, and corresponding (h) FWHMs and (g) energy efficiencies.

resembling the simulation results, and peak wavelengths align more closely with the target values. The residual signals observed between 550 and 650 nm in Fig. 5(e) could be attributed to the relatively high intensity of the input light within this spectral region. Furthermore, the shapes of the filtered spectra are more symmetric around the central wavelengths, demonstrating narrower FWHM, particularly at longer wavelengths [Fig. 5(f)]. For instance, the FWHM at the target wavelength of 650 nm decreases from 44.8 to 23 nm, representing a reduction of 48.7%. Figure 5(g) presents energy efficiencies at different inhibition ratios, revealing that energy efficiency is less sensitive to the inhibition ratio changes when PBEC is included in the loss function; while after PBEC removal, the energy efficiency becomes higher at low inhibition ratios (0:1, 1:10) and decreases with increase in the inhibition ratio, similar to the simulation results, although the observed efficiencies are lower. This discrepancy may be attributed to the limited accuracy of spatial alignment between the two diffractive layers, constrained by the 8 μm pixel size of the SLM. These results validate

the practical effectiveness of PBEC removal in enhancing the spectral filtering capabilities of broadband DNNs.

It is noteworthy that the pixel size of the diffractive layer is 8 μm, which is an order of magnitude larger than the working wavelengths. Since the spatial chromatic dispersion of scattered light is inversely related to the scatter size, the spectral resolution of the DNN could be further enhanced by reducing the pixel (neuron) size. This can be achieved by employing spatial light modulators (SLMs) with smaller pixel sizes or utilizing various advanced fabrication techniques. For instance, the feature sizes produced by two-photon direct laser writing can reach sub-wavelength scales in the visible wavelength range. Additionally, the spatial alignment of the diffractive layers could be more precise compared to that of free-space setups.

In this study, we present a practical approach to enhance the spectral filtering performance of broadband DNNs by eliminating the PBEC from the loss function during the training phase. Simulation results demonstrate that this method leads to a substantial reduction in out-of-band signal intensity, achieving suppression levels of up to three orders of magnitude and an extinction coefficient of 10^{-6}. Additionally, the spectral resolution is improved by more than 50%, accompanied by a 6.6% increase in energy efficiency. These results are further validated through experiments conducted with a two-layer DNN. In conclusion, the removal of PBEC proves to be an effective strategy for enhancing the performance of broadband DNNs, with potential applications in spectral reconstruction, spectrum classification, and color image processing, among other areas.

This study was financially supported by the Science and Technology Commission of Shanghai Municipality (Grant No. 21DZ1100500), the Shanghai Municipal Science and Technology Major Project, the Shanghai Frontiers Science Center Program (2021-2025 No. 20), and the National Natural Science Foundation of China (Grant No. 11874267).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Bolin Li: Data curation (equal); Formal analysis (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal). **Guangrui Luan:** Data curation (equal); Investigation (equal); Software (equal). **Yinfei Zhu:** Software (equal). **Jinlei Fei:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Software (equal). **Min Gu:** Funding acquisition (equal); Resources (equal). **Jian Lin:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**(7553), 436–444 (2015).
[2]A. Krizhevsky, I. Sutskever, and G. E. Hinton, Commun. ACM **60**(6), 84–90 (2017).
[3]G. Barbastathis, A. Ozcan, and G. Situ, Optica **6**(8), 921–943 (2019).
[4]G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, IEEE Signal Process Mag. **29**(6), 82–97 (2012).
[5]A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, Nat. Med. **25**(1), 24–29 (2019).
[6]S. P. Rodrigues, Z. Q. Yu, P. Schmalenberg, J. Lee, H. Iizuka, and E. M. Dede, Nat. Photonics **15**(2), 66–67 (2021).
[7]P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura, B. Brezzo, I. Vo, S. K. Esser, R. Appuswamy, B. Taba, A. Amir, M. D. Flickner, W. P. Risk, R. Manohar, and D. S. Modha, Science **345**(6197), 668–673 (2014).
[8]J. Pei, L. Deng, S. Song, M. G. Zhao, Y. H. Zhang, S. Wu, G. R. Wang, Z. Zou, Z. Z. Wu, W. He, F. Chen, N. Deng, S. Wu, Y. Wang, Y. J. Wu, Z. Y. Yang, C. Ma, G. Q. Li, W. T. Han, H. L. Li, H. Q. Wu, R. Zhao, Y. Xie, and L. P. Shi, Nature **572**(7767), 106–111 (2019).
[9]Q. M. Zhang, H. Y. Yu, M. Barbiero, B. K. Wang, and M. Gu, Light **8**, 42 (2019).
[10]M. Rafayelyan, J. Dong, Y. Q. Tan, F. Krzakala, and S. Gigane, Phys. Rev. X **10**(4), 041037 (2020).
[11]Y. C. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. J. Zhao, H. Larochelle, D. Englund, and M. Soljacic, Nat. Photonics **11**(7), 441–447 (2017).
[12]J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, Nature **569**(7755), 208–214 (2019).
[13]J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner, Optica **5**(6), 756–760 (2018).
[14]X. Lin, Y. Rivenson, N. T. Yardimei, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, Science **361**(6406), 1004–1008 (2018).
[15]T. Yan, J. M. Wu, T. K. Zhou, H. Xie, F. Xu, J. T. Fan, L. Fang, X. Lin, and Q. H. Dai, Phys. Rev. Lett. **123**(2), 023901 (2019).
[16]E. Goi, S. Schoenhardt, and M. Gu, Nat. Commun. **13**(1), 7531 (2022).
[17]E. Goi, X. Chen, Q. M. Zhang, B. P. Cumming, S. Schoenhardt, H. T. Luan, and M. Gu, Light **10**(1), 40 (2021).
[18]Y. C. Zhang, Q. M. Zhang, H. Y. Yu, Y. N. Zhang, H. T. Luan, and M. Gu, Sci. Adv. **10**(24), eadn2205 (2024).
[19]Y. Luo, D. Mengu, N. T. Yardimci, Y. Rivenson, M. Veli, M. Jarrahi, and A. Ozcan, Light **8**, 112 (2019).
[20]J. X. Li, D. Mengu, N. T. Yardimci, Y. Luo, X. R. Li, M. Veli, Y. Rivenson, M. Jarrahi, and A. Ozcan, Sci. Adv. **7**(13), eabd7690 (2021).
[21]T. K. Zhou, L. Fang, T. Yan, J. M. Wu, Y. P. Li, J. T. Fan, H. Q. Wu, X. Lin, and Q. H. Dai, Photonics Res. **8**(6), 940–953 (2020).
[22]Y. Z. Cheong, L. Thekkekara, M. Bhaskaran, B. del Rosal, and S. Sriram, Adv. Photonics Res. **5**(6), 2300310 (2024).
[23]D. Mengu, A. Tabassum, M. Jarrahi, and A. Ozcan, Light **12**(1), 86 (2023).
[24]Z. Y. Duan, H. Chen, and X. Lin, Nanophotonics **12**(5), 893–903 (2023).
[25]Z. Wang, H. Chen, J. A. Li, T. F. Xu, Z. J. Zhao, Z. Y. Duan, S. Gao, and X. Lin, Nanophotonics **13**(20), 3883–3893 (2024).
[26]Y. L. Zhu, Y. Y. Chen, and L. Dal Negro, Opt. Lett. **47**(24), 6309–6312 (2022).
[27]M. Gu, *Advanced Optical Imaging Theory* (Springer, Berlin, New York, 2000).