

Research Article

Multitask Learning by Multiwave Optical Diffractive Network

Jing Su,¹ Yafei Yuan ,² Chunmin Liu,¹ and Jing Li¹

¹Department of Optical Science and Engineering, Fudan University, Shanghai 200433, China

²Department of Electronic Engineering, Center for Intelligent Medical Electronics, Fudan University, Shanghai 200433, China

Correspondence should be addressed to Yafei Yuan; yyf@fudan.edu.cn

Received 26 March 2020; Revised 9 June 2020; Accepted 18 June 2020; Published 10 July 2020

Academic Editor: Ivan Giorgio

Copyright © 2020 Jing Su et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, there has been tremendous research studies in optical neural networks that could complete comparatively complex computation by optical characteristic with much more fewer dissipation than electrical networks. Existed neural networks based on the optical circuit are structured with an optical grating platform with different diffractive phases at different diffractive points (Chen and Zhu, 2019 and Mo et al., 2018). In this study, it proposed a multiwave deep diffractive network with approximately 10^6 synapses, and it is easy to make hardware implementation of neuromorphic networks. In the optical architecture, it can utilize optical diffractive characteristic and different wavelengths to perform different tasks. Different wavelengths and different tasks inputs are independent of each other. Moreover, we can utilize the characteristic of them to inference several tasks, simultaneously. The results of experiments were demonstrated that the network could get a comparable performance to single-wavelength single-task. Compared to the multinet network, single network can save the cost of fabrication with lithography. We train the network on MNIST and MNIST-FASHION which are two different datasets to perform classification of 32×32 inputs with 10 classes. Our method achieves competitive results across both of them. In particular, on the complex task MNIST_FASHION, our framework obtains an excellent accuracy improvement with 3.2%. In the meanwhile, MNSIT also has the improvement with 1.15%.

1. Introduction

Store and retrieve data units based on the von Neumann architecture are far more time-consuming and power-hungry than an optical device [1–6]. Different from modern computers, CIS integrated data computation, storage and fetch, is more effective, uses less power, has large storage capacity, and has higher integration level [7–9]. Besides, artificial neural network is [10] similar to the way in which human and animal store and process data, which is successful in a wide range of tasks such as image analysis [11], speech recognition [12], and language translation [13]. Artificial neural network can get comparable or even superior performance than the human with the increasing data volume, problem complexity, and structure depth. Most of the tasks cannot be migrated well in smart portable devices for its complexity and power. The less power, more efficiency, and faster speed are becoming more and more critical for deep learning implemented on the embedded device.

The neuromorphic computing seeks a brain-like processing, which overcomes the limitation from conventional computers. IBM [14] built a 5.4-billion transistor chip with 4096 neurosynaptic cores called True North, a fully functional digital chip. To provide the extreme complexity of the human cerebral cortex, Cianciulli et al. [15] combined complementary metal oxide semiconductors (CMOS) and two-terminal resistive device with electric circuits. Spin-transfer torque magnetic memory [16] (STTMRAM), with nonvolatility, high speed, and high endurance, is suitable as a stochastic memristive device, considering the functional implication of synaptic neuronal plasticity. Tait [17], inspired by spiking neural networks, integrated laser devices to explore highly interactive information at speeds with optical-electronic systems. This approach promises to incorporate photonic spike processing in the training system. Besides, Rios et al. dramatically improved storage capacity to implement all-photonic nonvolatile multilevel [18] where memory electric-photonic interconnect technologies bring out

only opportunities but also challenges to the unconventional circuits and systems. To overcome the wastage of optical-electric conversions coupling, all photic device can be performed with fast computational speed and lower power. On-chip nonvolatile photic device [19] would dramatically improve performance in existing brain-like neural networks [20] to eliminate electronic latency and reduce electronic consuming. The on-chip optical architecture is designed for network protocol computational element and waveguide medium to communicate among high-performance spiking neurons.

The architecture of fully optical network with Mach-Zehnder interferometer [21] promises the reduction of data-movement energy cost. All-optical diffractive deep neural network architecture (D2NN) [22] utilizes passive component and optical diffraction. D2NN can be easily scaled up and fabricated by 3D lithography [23] in a power-efficient manner.

In general, optical networks have more trainable parameters with complex-value modulation which provide phase and amplitude of each neuron rather than only amplitude in electric networks. Unfortunately, the optical device to form neural network has some problems. Firstly, all-optical neural network is designed for a single task, but multitasks [24] are significant and important. Secondly, learning rates for different tasks are important to the accuracy. It is nontrivial to balance the tasks and learning rates.

In this paper, to address the above two issues, we make most of the optical characteristics to express different tasks with different wavelengths. One is used as baseband and the other is used as a carrier frequency. Therefore, the base band wave can be set to a large learning rate and vice versa. Extensive experiments based on MNIST and MNIST-Fashion [25] are conducted to investigate the efficacy and properties of the proposed multiwavelength diffractive network (MWDN). In both tasks, the MWDN significantly performs the baselines even better in the same network.

2. Multiwavelength Diffractive Network

2.1. The Forward of the Network. Spatial domain implies per-wave in-plane propagation reasoning about diffractive in the particular phase and frequency, which can analyze and integrate different direction waves. It operates in the frequency domain. The wave distribution of the observation and aperture plane can be viewed as the linear combination with a great many monochromatic plane wave of different direction propagations. The amplitude and phase of each plane wave lies on the angular spectrum. The angular spectrum can be acquired by FFT analysis process [26]. The plane wave propagation is a complex task that takes into consideration many affecting factors, such as direct, phase, and amplitude.

As shown in Figure 1, in the top row of the picture, we first adjust optical grating parameters (height and the complex index of refraction), and then the height is altered by 3D printed, and the complex index of refraction is altered by laser light with different power in the last. Different power can alter different refraction of phase change materials. The

above figure shows the manufacturing process of the whole mask. The substrate changes the phase transition threshold (the phase transition changes the complex refractive index value) by calculating the film thickness, photolithography (adjusting the corresponding optical threshold), and photolithography power. After the whole system is made, in the middle row, different tasks are input, and the corresponding area is illuminated through the same network, which represents the label information. We use different wavelengths λ_1 and λ_2 to represent different categories. At the same time, the wavelengths are one long wave and one short wave, which have the effect of carrier wave, so that there is no impact between tasks.

We input images in MNIST and MNIST-FASHION simultaneously, the input optical wavelength of MNIST and MNIST-FASHION task is λ_1 and λ_2 , respectively. The diffractive network with a different task has the same optical parameters. The bottom of the figure is the optical carrier with different wavelengths λ_1 and λ_2 .

The top left of Figure 2 shows plan on the front view corresponding to the bottom of Figure 2, and the top right of Figure 2 is physical simulation system. The bottom of Figure 2 is the framework of diffractive network, and different colors denote different indexes of refractive on the side view. The task is input through the front mask, through the follow mask with different thicknesses and different phases.

Firstly, we convert image information to the phase and amplitude of optical information as the input of systems. Then, the optical grating is manufactured by 3D-printing device with different heights. In the following sections, we discuss that MWDN tackles the tasks predominately using the angular spectrum. MWDN by the 3D-printing would influence the amplitude and phase of the wave to $0\sim 1$ and $0\sim 2\pi$, for two tasks in the same network. For each layer of MWDN, we set the neuron size range $200\ \mu\text{m}$ to $700\ \mu\text{m}$, which is an effect tunable.

Following the Fresnel diffraction equation, we can consider the optical signal from the spatial domain to the frequency domain. The angular spectrum method of plane wave explains how wave propagates. It is the primary method of analyzing diffraction in the frequency domain. Based on the angular spectrum, the free space transfer function is to control free propagation. The wave plane can transfer angular spectrum by the FFT process, where diffractive data processing is more evident as follows:

$$\begin{aligned}
 A_0(f_x, f_y) &= \text{FFT}(U_0(x, y)) \\
 &= \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} U_0(x, y) e^{-j2\pi((f_x x/N) + (f_y y/M))}, \\
 A(f_x, f_y) &= A_0(f_x, f_y) \cdot H(f_x, f_y), \\
 U(x, y) &= \text{IFFT}(A(f_x, f_y)), \\
 H(f_x, f_y) &= \exp(jkz) \exp\left(j\pi\lambda z (f_x + f_y)^2\right),
 \end{aligned} \tag{1}$$

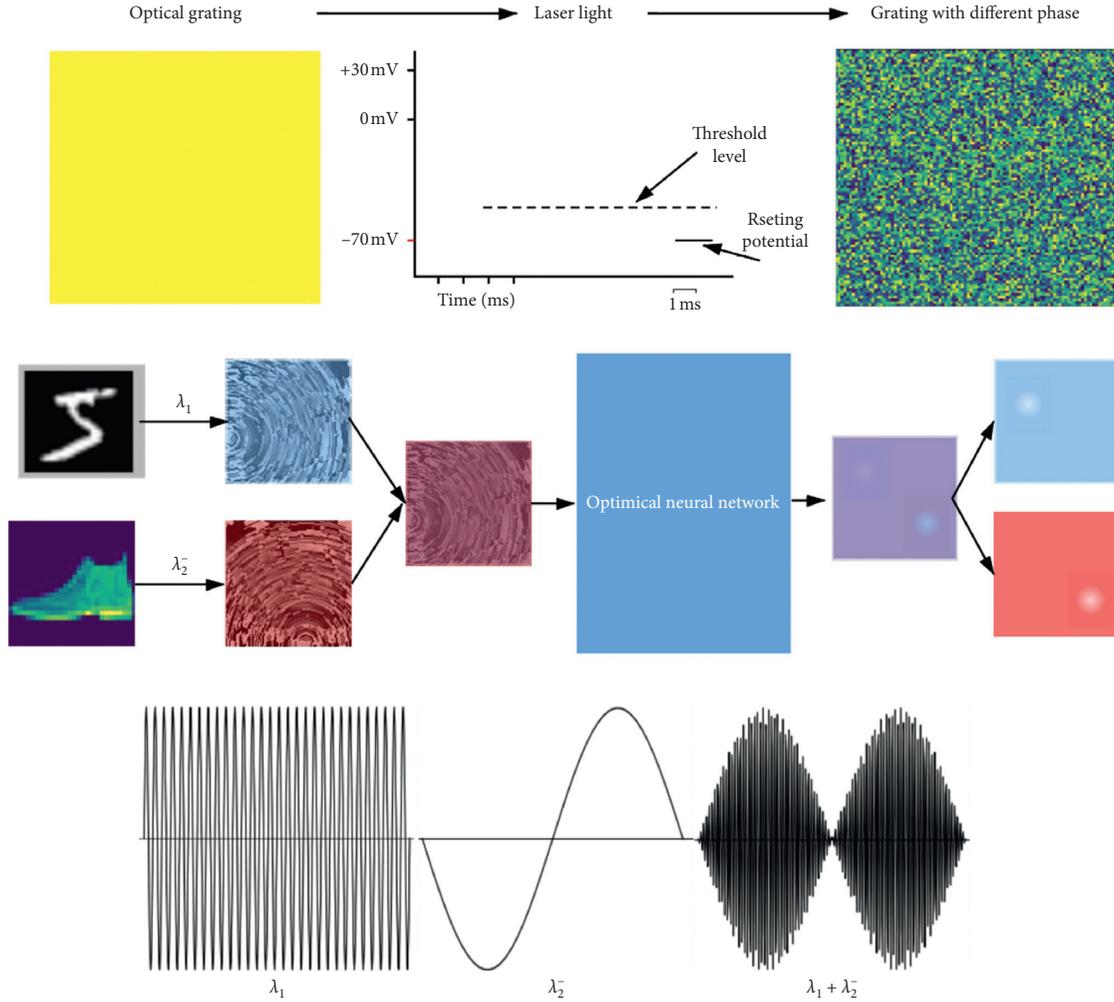


FIGURE 1: The architecture of multiwavelength diffractive network.

where f_x and f_y are space frequencies which correspond to x and y location ($f_x = 1/(x-x_0)$), $x-x_0$ is gaps in the optical map, N and M is the number of grooves on optical grating in height and width direction, $U_0(x, y)$ is the original field distribution, $U(x, y)$ is the field distribution after free-space transfer, $H(f_x, f_y)$ is transfer function, $A_0(f_x, f_y)$ is the original angular spectrum, and $A(f_x, f_y)$ is the angular spectrum after free-space transfer. The results of the inverse Fourier transform to transfer function are the impulse response function. The equation can be viewed as the Fourier transform:

$$\begin{aligned}
 h(x-x_0, y-y_0) &= F^{-1}(H(f_x, f_y)) \\
 &= \frac{1}{j\lambda z} \exp(jkz) \exp\left(j\frac{k}{2z}((x-x_0)^2 + (y-y_0)^2)\right).
 \end{aligned} \quad (2)$$

The output wave plane distribution propagates through 3D material and the field distribution is changed by the refractive index:

$$O(x, y) = U(x, y) \cdot (1 - \alpha \cdot \Delta d) \cdot \exp(j\phi),$$

$$\phi = \frac{2\pi\Delta n_1 \Delta d}{\lambda},$$

(3)

$$\Delta n_1 = n - n_0,$$

$$\alpha = \frac{4\pi k}{\lambda},$$

where α is extinction efficiency, n is the refractive index real part of 3D material, n_0 is the refractive index real part of the vacuum, k is the refractive index imagery part, λ is the wavelength, Δd is the height of material map, and ϕ is the phase difference. If we choose transparent material ($k \rightarrow 0$) ignoring the optical losses, the transmission coefficient of a neuron is composed of only phase term; if we select non-transparent material, the transmission coefficient of a neuron is composed of amplitude and phase in MWDN architecture.

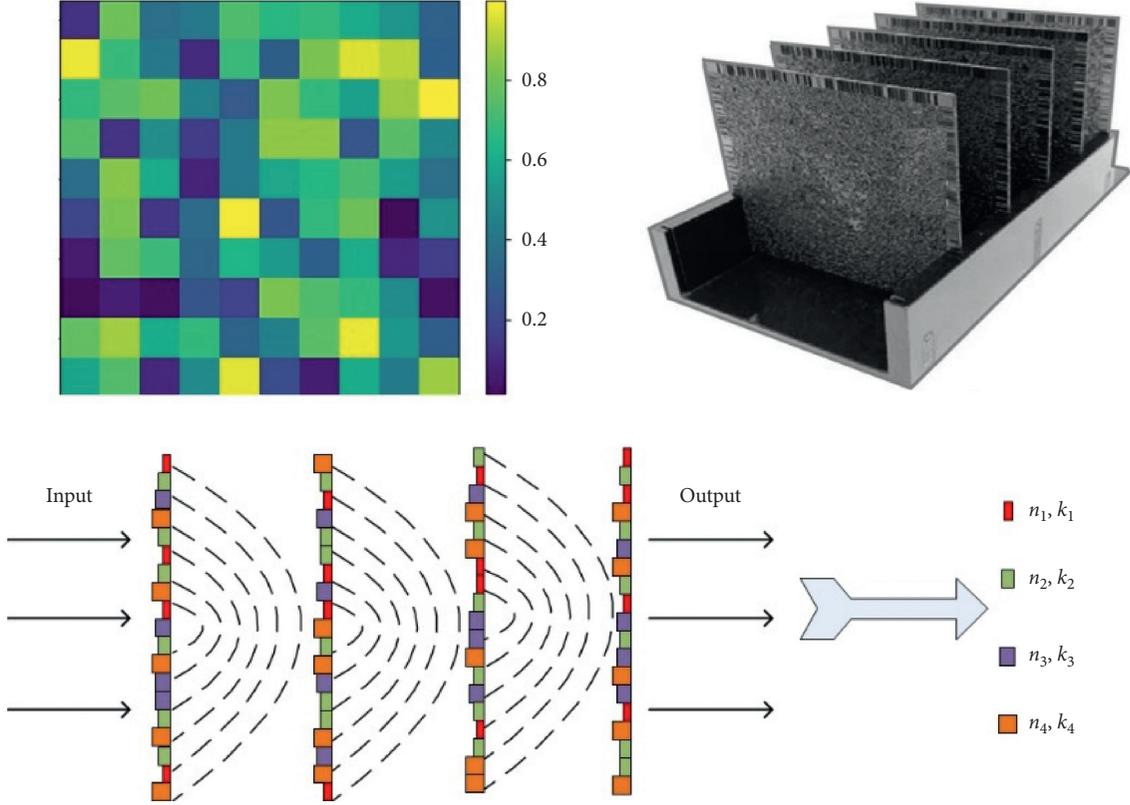


FIGURE 2: The framework of multiwavelength diffractive network.

According to the size of input data, an effective and flexible linear interpolation algorithm is to fit the diffractive input layer. The interconnection rate between adjacent layers relates to the distance and diffraction angle, which approach the critical value (1.0).

Furthermore, the number of the network layers and the axial distance are also tunable. The output layer can part into ten regions corresponding to ten classes, where the summation of light intensity can be detected in the wave plane region. Mean square error (MSE) uses to train MWDN parameters compared to the target. We aim to minimize a loss function, which increases target region wave intensity and decrease other regions. The training batch size is set to be 10 for the classifier.

2.2. The Backward of the Network. To train MWDN, we use the backpropagation algorithm with the Adam optimization method. We focus on the intensity of wave and define loss function with MSE between the output and target:

$$E(\Delta d) = \frac{1}{K} \sum_{k=1}^K (o_k - t_k)^2, \quad (4)$$

where K is the number of training data, o_k is the output of the MWDN, and t_k is the label of the corresponding input. The optimization problem can be written follows:

$$\min_{\Delta d_i^l} E(\Delta d_i^l), \quad (5)$$

where l is the layer and i is the l th layer location. The gradient of loss to all parameters can be calculated, which is used to update MWDN architecture parameters during the training process. Each batch of the training data is fed into MWDN, where each layer gradient can be calculated to update.

We use multiwavelength, we need to adjust a base value through one wavelength, and then fine-tune through another wavelength without affecting another task as follows:

$$\begin{aligned} \frac{\partial m_k^{M+1}}{\partial r} &= \left(\frac{\partial t_i^M}{\partial r} \right) m_i^M w_{i,k}^M + \left(\frac{\partial w_{i,k}^M}{\partial r} \right) t_i^M m_i^M \\ &= \sum_i \left(\frac{j2\pi}{\lambda_i} - \frac{2\Delta z}{r} - \frac{1}{2\pi r} \right) t_i^M m_i^M w_{i,k}^M. \end{aligned} \quad (6)$$

Our refractive index has a discrete value. Based on the principle of low-bit network, refractive index n is adjusted to be continuous in the backward and Bernoulli distribution to samples N_1 , N_2 , and N_m in the forward. There are m kinds of discrete N . The refractive index value depends on different phase values. In this paper, $m = 4$ is used, and the reasoning is as follows:

$$\frac{\partial m_k^{M+1}}{\partial n_i^l} = \frac{\partial m_k^{M+1}}{\partial \phi_i^l} \frac{\partial \phi_i^l}{\partial n_i^l} = \sum_l \frac{2\pi\Delta z}{\lambda_l} \frac{\partial m_k^{M+1}}{\partial \phi_i^l}. \quad (7)$$

For the forward step, we normalize n :

$$n_{\text{nor}} = \frac{(n_i^l - n_1)}{(n_2 - n_1)} Pn. \quad (8)$$

2.3. The Multiwavelength Setting. The optical diffractive network and deep neural network are markedly different. The function of the optical diffractive network is determined by wavelength and the parameters of the optical grating (height and complex refractive index). Multiwavelength diffractive network has a broad range of requirements that differ from the conventional network.

Different wavelengths have different effectiveness. We set different wavelengths for different tasks. Meanwhile, the network needs to ensure that different wavelengths do not affect each other. By setting one to baseband and the other to the carrier, the diffractive network is used to adjust optical plane wave independently. The algorithm can be considered as an efficient carrier algorithm. The ratio of baseband and carrier wavelengths is 1 : 30. The short wavelength is little influence to long wavelength and vice versa. If the phase difference of long wavelength is ϕ_1 and the phase difference of short wavelength is ϕ_2 , the corresponding relationship as follows:

$$\begin{aligned}\phi_1\lambda_1 &= \phi_2\lambda_2, \\ \frac{\lambda_1}{\lambda_2} &= 1 : 30.\end{aligned}\quad (9)$$

So, the equation can be as follows:

$$\begin{aligned}\frac{\phi_2}{\phi_1} &= 1 : 30, \quad \phi_2 \ll \phi_1, \\ \phi_1 &= 2\pi n + \phi'_1, \quad n = 0, 1, 2, \dots, \\ \phi_2 &= \frac{\pi n}{15} + \frac{\phi'_1}{30}, \quad \phi'_1 = 0 \sim 2\pi.\end{aligned}\quad (10)$$

The second term of ϕ_2 relative to the first term can be ignored, and the equation can be shown as follows:

$$\phi_2 \approx \frac{\pi n}{15}.\quad (11)$$

The multiwavelength diffractive network can be effective, and more powerful than deep neural network. Phase difference $\phi_i (i = 1, 2)$ can be obtained easily by adjusting the height on the diffractive network. Due to $\phi_2 \ll \phi_1$, then we adjust Δd for the minor learning rate for ϕ_2 , as well as the large learning rate for ϕ_1 , without one impacting the other:

$$\Delta d \approx \frac{\phi\lambda}{2\pi\Delta n}.\quad (12)$$

3. Experiment

In this work, we apply the proposed MWDN to implement on two different datasets MNIST and MNIST-Fashion.

3.1. Model Setup. By comparing to the state-of-the-art methods with accuracy and speed, MNIST and MNIST-FASION, in this method, achieve better performance. The size of the network is set to 200×200 , 500×500 , and 1500×1500 , each having a trainable height of the map. The

optical network possesses two types, one for phase modulation and the other for complex modulation. The MNIST and MNIST-FASION works with different optical wavelengths, and the input is altered by the optical grating mask.

Using the backward propagation, the model is trained with two task datasets alternately and validated its effectiveness. We train the network with different learning rates for different tasks, which overcome the drawbacks of local optimum to solve. As well as, all the parameters of the network are adjusted by the gradient descent algorithm to minimize the error.

3.2. Dataset. We evaluate the approach on two datasets and input information for neurons in the form of phase fed into the network. The two datasets have different data distributions, which are difficult to classify in the same network. The conventional networks require the input information to be independent and identical distribution. The task is to handle two different distribution data in the same network.

3.3. Experimental Analysis. For better performance, we set a different learning rate and different signal frequency to two datasets. The maximum half-cone diffraction angle is formulated as follows:

$$\varphi_{\max} = \sin^{-1}\left(\frac{\lambda}{2d_f}\right).\quad (13)$$

The light wavelength is 0.4 THz and 14.4 THz for MNIST and MNIST-Fashion. The neuron size is set to be $200 \mu\text{m}$. The height of the map and axial distance between two successive layers are trainable. As comparing the performance of MWDN and DN methods with the single task, the results are shown in Table 1. It is clear that the performance of MWDN would improve the accuracy of 1.15% and 3.2%, independently. To evaluate the multiwavelength for multitask, so we compare the multitask to a single task in Table 2. The multitask diffractive network enables consistent performance with a single task. The result can perform well in the same parameters. The experiment set of setting 1 is the same wavelength for comparison. Setting 2 is performed by a different wavelength. The DN-FASION and DN-MNIST are evaluated by independent diffractive network.

3.4. Convergence Analysis. Figures 3 and 4 demonstrate two classifiers for two datasets, where each dataset has ten target class. One classifier is set to classify in the same region and another is set to a different region. The different frequency setting to different datasets is effective and different regions in the MWDN have high accuracy and lower loss.

Finally, we report the performance on the validation data of MNIST and MNIST_FASION, a challenging task with different datasets. Using only a single network, two datasets classification can be accomplished simultaneously. We investigate the effects of various combinations of different datasets for MWDN. The results are shown in Figures 3 and 4. If the precision continues to increase and the loss decreases, the task is convergence. We find that we can

TABLE 1: The performance of MWDN and DN methods.

Method	MNIST (%)	λ (THz)	MNIST-Fashion (%)	λ (THz)
MWDN(PCM)	92.85	0.4	84.33	14.4
DN-MNIST	91.75	0.4	81.13	0.4

TABLE 2: The performance of MWDN and DN methods with multitask.

Method	MNIST (%)	MNIST-Fashion (%)	λ_1 (THz)	λ_2 (THz)
Setting 1	23.45	12.12	0.4	0.4
Setting 2	90.45	76.67	14.4	0.144
DN-Fashion	—	78.70	14.4	—
DN-MNIST	91.75	—	—	14.4

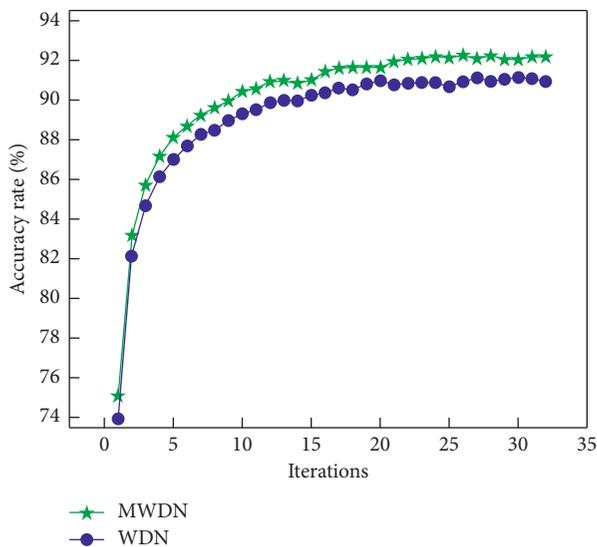


FIGURE 3: The accuracy in the training process of different methods.

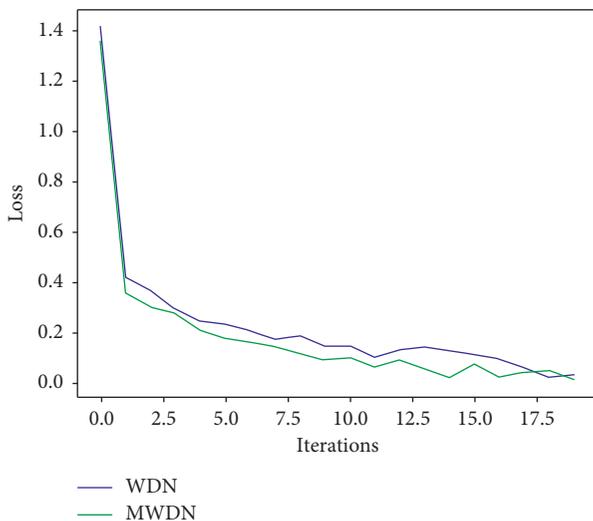


FIGURE 4: The loss in the training process of different methods.

implement two classes to the same network with MWDN algorithm. Compared to other approaches that use only single dataset as input, our approach even yields a boost.

4. Conclusion

In this paper, we propose a novel and multitasks optical network named as the multiwavelength diffractive network (MWDN). Based on plane wave propagation, our method can achieve comparable accuracy against the single-task network. We successfully apply MWDN to multitasks with different datasets distribution and provide a multiwavelength method with different model sizes. In the future, we aim to develop a more effective network to achieve complex tasks and reach better performance.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the financial support by National Key R&D Program of China (Grant no. 2017YFE0112000) and Shanghai Municipal Science and Technology Major Project (Grant no. 2017SHZDZX01). The authors wish to thank Doctor Y. C. Sun. Furthermore, the authors would like to express sincere thanks for light source and controller supported by 3 Lights and Light-ca Technology Corporation.

References

- [1] Y. Chen and J. Zhu, "An optical diffractive deep neural network with multiple frequency-channels," 2019, <https://arxiv.org/abs/1912.10730>.
- [2] W. Mo, C. L. Gutterman, Y. Li, S. Zhu, G. Zussman, and D. C. Kilper, "Deep-neural-network-based wavelength selection and switching in ROADMs systems," *Journal of Optical Communications and Networking*, vol. 10, no. 10, pp. D1–D11, 2018.
- [3] J. L. Jewell, "Regenerative pulsations from an intrinsic bistable optical device," *Applied Physics Letters*, vol. 40, no. 4, pp. 291–293, 1982.
- [4] J. Kraitl, H. Ewald, and H. Gehring, "An optical device to measure blood components by a photoplethysmographic method," *Journal of Optics A Pure & Applied Optics*, vol. 7, no. 6, pp. S318–S324, 2005.
- [5] Y. Liu, X. Lin, C. Wei, C. Zhang, and Y. S. Zhao, "3D-printed optical-electronic integrated devices," *Science China Chemistry*, vol. 62, no. 10, pp. 1398–1404, 2019.
- [6] S. Kim, S. G. Menabde, V. W. Brar, and M. S. Jang, "Functional midinfrared polaritons in van der Waals crystals," *Advanced Optical Materials*, vol. 8, no. 5, 2019.
- [7] M. Spencer, J. Eickholt, and J. Cheng, "A deep learning network approach to ab initio protein secondary structure prediction," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 12, no. 1, pp. 103–112, 2014.

- [8] X. Zhou, R. Takayama, S. Wang, T. Hara, and H. Fujita, "Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method," *Medical Physics*, vol. 44, no. 10, pp. 5221–5233, 2017.
- [9] C. Che, R. Lin, X. Zeng, E. Karim, G. John, and M. Xu, "Improved deep learning-based macromolecules structure classification from electron cryo-tomograms," *Machine Vision & Applications*, vol. 29, no. 8, pp. 1227–1236, 2018.
- [10] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5987–5995, IEEE, Honolulu, HI, USA, July 2017.
- [11] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," *Computer Vision—ECCV 2014*, Springer, Berlin, Germany, 2014.
- [12] L. Deng, "Deep learning: from speech recognition to language and multimodal processing," *Apsipa Transactions on Signal & Information Processing*, vol. 5, 2016.
- [13] S. P. Singh, A. Kumar, H. Darbari, L. Singh, and S. Jain, "Machine translation using deep learning: an overview," in *Proceedings of the 2017 International Conference on Computer, Communications and Electronics (Comptelix)*, IEEE, Jaipur, India, July 2017.
- [14] T. O. Kwok and Y. K. Kwok, "On the design, control, and use of a reconfigurable heterogeneous multi-core system-on-a-chip," in *Proceedings of the 2008 IEEE International Symposium on Parallel and Distributed Processing*, IEEE, Miami, FL, USA, April 2008.
- [15] T. F. Cianciulli, J. A. Lax, F. E. Cerruti et al., "Complementary value of transthoracic echocardiography and cinefluoroscopic evaluation of mechanical heart prosthetic valves," *Echocardiography*, vol. 21, no. 2, p. 211, 2010.
- [16] A. F. Vincent, J. Larroque, W. Zhao et al., "Spin-transfer torque magnetic memory as a stochastic memristive synapse," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 9, no. 2, pp. 166–174, 2015.
- [17] A. N. Tait, M. A. Nahmias, T. Yue, B. J. Shastri, and P. R. Prucnal, *Photonic Neuromorphic Signal Processing and Computing*, Springer, Berlin, Germany, 2013.
- [18] C. Ríos, M. Stegmaier, P. Hosseini et al., "Integrated all-photonic non-volatile multi-level memory," *Nature Photonics*, vol. 9, no. 11, pp. 725–732, 2015.
- [19] R. A. Hut, N. Mrosovsky, and S. Daan, "Nonphotic entrainment in a diurnal mammal, the European ground squirrel (*Spermophilus citellus*)," *Journal of Biological Rhythms*, vol. 14, no. 5, pp. 409–420, 1999.
- [20] K. Voutsas, G. Langner, J. Adamy, and M. Ochse, "A brain-like neural network for periodicity analysis," *IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics*, vol. 35, no. 1, pp. 12–22, 2005.
- [21] M. Dülk, S. Fischer, M. Bitter et al., "Ultrafast all-optical demultiplexer based on monolithic Mach–Zehnder interferometer with integrated semiconductor optical amplifiers," *Optical & Quantum Electronics*, vol. 33, no. 7–10, pp. 899–906, 2001.
- [22] X. Lin, R. Yair, N. T. Yardimci et al., "All-optical machine learning using diffractive deep neural networks," *Science*, vol. 361, no. 6406, pp. 1004–1008, 2018.
- [23] A. d. Campo and C. Greiner, "SU-8: a photoresist for high-aspect-ratio and 3D submicron lithography," *Journal of Micromechanics & Microengineering*, vol. 17, no. 6, pp. R81–R95, 2007.
- [24] Y. Tian, P. Luo, X. Wang, and X. Tang, "Pedestrian detection aided by deep learning semantic tasks," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Boston, MA, USA, June 2015.
- [25] L. Deng, "The MNIST database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.
- [26] C. F. V. Loan, *Computational Frameworks for the Fast Fourier Transform*, SIAM, Philadelphia, PA, USA, 1992.