PHOTONICS Research

High-speed and versatile ONN through parametric-based nonlinear computation

Xin Dong,[1,†] Yuanjia Wang,[1,†] Xiaoxiao Wen,[1] Yi Zhou,[1,2,3] and Kenneth K. Y. Wong[1,2,4]

[1]Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China
[2]Advanced Biomedical Instrumentation Centre, Hong Kong, China
[3]e-mail: zhouyi08@hku.hk
[4]e-mail: kywong@eee.hku.hk
[†]These authors contributed equally to this work.

Neural networks (NNs), especially electronic-based NNs, have been rapidly developed in the past few decades. However, the electronic-based NNs rely more on highly advanced and heavy power-consuming hardware, facing its bottleneck due to the slowdown of Moore's law. Optical neural networks (ONNs), in which NNs are realized via optical components with information carried by photons at the speed of light, are drawing more attention nowadays. Despite the advantages of higher processing speed and lower system power consumption, one major challenge is to realize reliable and reusable algorithms in physical approaches, particularly nonlinear functions, for higher accuracy. In this paper, a versatile parametric-process-based ONN is demonstrated with its adaptable nonlinear computation realized using the highly nonlinear fiber (HNLF). With the specially designed mode-locked laser (MLL) and dispersive Fourier transform (DFT) algorithm, the overall computation frame rate can reach up to 40 MHz. Compared to ONNs using only linear computations, this system is able to improve the classification accuracies from 81.8% to 88.8% for the MNIST-digit dataset, and from 80.3% to 97.6% for the Vowel spoken audio dataset, without any hardware modifications.   © 2025 Chinese Laser Press

https://doi.org/10.1364/PRJ.553388

1. INTRODUCTION

In the past decades, with the help of powerful electronic hardware, neural networks (NNs) have become an important part of both scientific applications and daily life areas [1–4]. However, as the exponential scaling of electronic transistors approaches its physical limit, the development of electronic hardware ruled by Moore's law has seen a tendency to slow down. This becomes a bottleneck for NN computation performances [5,6]. Thus, it is of great significance that a new generation of NN computation platforms is developed. Optical systems, which can also provide computation solutions as information carriers, have great potential in altering traditional NN structures into optical neural networks (ONNs), with the advantages of high energy efficiency, low crosstalk, light-speed process, and massive parallelism computation [7–9].

Though the information in ONNs is processed at the speed of light, most exciting research on ONNs usually involves direct spatial manipulations and observations with the help of other electronic devices, such as digital micromirror devices, spatial light modulators, and cameras. The frame rate of these items, to some extent, is also determined by electronic hardware and may only reach several thousand Hz, limiting the computation speed of ONNs [10,11]. Therefore, it

is desired that superior optical manipulation and observation methods are implemented for higher-speed ONNs. On the one hand, mode-locked lasers (MLLs), which can produce pulses of ultrashort durations, are able to generate ultra-wide-bandwidth carrier lasers. In this way, more efficient spatial encoding with waveshapers can be realized [12,13]. Further, by applying temporal dispersion, wavelength-to-time mapping can be achieved. Utilizing the dispersive Fourier transform (DFT) technique, spectrum observations can be conducted with a higher frame rate, as proven to be effective in microscopy and soliton dynamics observation applications [14].

Another challenge of ONNs is the realization of effective and flexible nonlinear computations. In NN systems, nonlinear computations (i.e., activation functions) are required for closer fittings and higher accuracies [15]. Though efforts have been seen for specific activation functions, with multimode fibers (MMFs) [16] or diffractive metasurface [17,18], these modules are usually designed aiming at one specific task function, and thus lack the adaptability for different tasks. One of the potential solutions to this challenge is to apply the nonlinear Schrödinger equation in the system [19,20], specifically with the highly nonlinear fiber (HNLF). The parametric process in HNLF can project the input vector to a separable

vector space, while maintaining certain flexibilities of presented nonlinear functions by optimizing the equation variables [21]. Hence, it is promising to implement HNLF for nonlinear computations in ONNs. Zhou *et al.* [22] designed a nonlinear Schrödinger kernel-based hardware accelerator that can nonlinearly project data into a linear, predictable space. Although the system has proved to be effective for different types of tasks, the optical kernels applied are data-dependent and not trained; therefore, they are still not flexible enough when it comes to reusable system implementation.

Based on the above discussions, we present here a novel, high-speed, and versatile optical neural network. An adaptive nonlinear computation module via HNLF and its parametric processes are implemented in this system. Aiming at a versatile system for different types of tasks, we specifically researched how the gains of the nonlinear Schrödinger process affect the system accuracy and thus realize the adaptive nonlinear computation by optimizing the process variables. In this way, the system can be achieved with versatility for different types of tasks. Meanwhile, this ONN system can further achieve an overall computation frame rate of up to 40 MHz with the specially designed MLL source combined with DFT processes. Evaluations are conducted on both handwritten digit datasets and spoken audio datasets with one system setup without any hardware modification. After applying nonlinear computations, the classification accuracy is increased from 81.8% to 88.8% for the MNIST-digit image dataset, and from 80.3% to 97.6% for the Vowel spoken audio dataset.

2. METHODS

The novel ONN system design is illustrated in Fig. 1. The laser source is an inertial-free swept MLL source designed based on our previous research (102 nm, 10 dB optical bandwidth centered at 1.55 μm, 22 mW direct output average power, and 1.6 ps pulse width) with a repetition rate of 40 MHz [13]. The input data will be converted into a spectral form and a waveshaper is then leveraged to directly modulate the input data onto the MLL pulses. Since the parametric process is polarization-dependent, a polarization controller is applied at the output of the waveshaper to optimize the parametric process realized in the HNLF.

The parametric process is regarded as the key to the nonlinear computation part of the system. Specifically, HNLF is used here to realize optical computation in the spectrum domain, in which the pulses propagate following the nonlinear Schrödinger equation expressed as Eq. (1):

$$\frac{\partial U}{\partial z} = -\frac{\alpha}{2}U - \frac{i\beta_2}{2}\frac{\partial^2 U}{\partial T^2} + i\gamma P_0 |U|^2 U, \tag{1}$$

where U is the initial field; α is the fiber loss coefficient; $T = t - z/v_g$, with t being the physical time and v_g being the group velocity; and γ is the nonlinear coefficient. It can be seen that Eq. (1) is a combination of linear attenuation, second-order dispersion, and the Kerr effect, which plays an important role in the nonlinear parametric process [19,20].
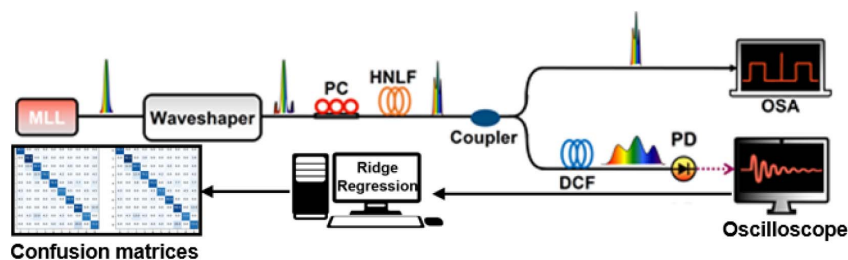
It is worth noting that, apart from the advantages of light-speed computations and low power consumption, by modifying the modulation signal and optimizing the gains of the nonlinear Schrödinger process in HNLF, our network features can be adapted to handle different types of tasks without reconstructing the system. This will be further demonstrated in the evaluation sections. Hence, our system provides great potential for a variety of advanced applications.

Since this computation is fully realized in a physical approach at the speed of light, the system latency is greatly shortened, and the computation capacity is no longer limited by conventional hardware. Moreover, the MLL pulse has a large peak power with low average power because of the low duty ratio; the parametric process can then be carried out by consuming only a few milliwatts with the help of the HNLF, which can introduce a strong nonlinear phase shift on a low power level [23]. Thus, this nonlinear computation process offers great energy conservation.
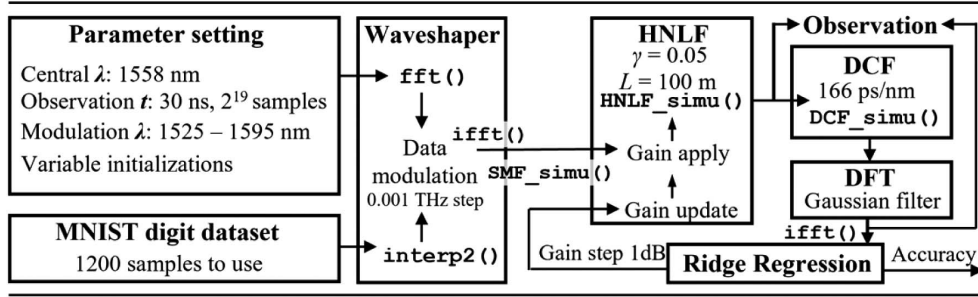
After the nonlinear computations, the proceeding pulse is divided by a 90/10 coupler. Ten percent of the pulse is observed directly by the optical spectrum analyzer (OSA) for spectrum references, and 90% is stretched by a dispersion compensation fiber (DCF) with a dispersion of 166 ps/nm to execute the DFT process for high-frame-rate observations. Based on the temporal far-field diffraction with femtosecond laser pulses (MLL) and dispersive components (DCF), the transfer function of the temporal dispersion is expressed as Eq. (2):

$$H(\omega) = \exp\left(\frac{i\beta_2 z\omega^2}{2}\right), \tag{2}$$

where β_2 is the second-order dispersion parameter, z is the propagation distance, and $\Phi = \beta_2 z$ is the group delay dispersion. After the dispersion module, the wavelength

Fig. 1. Experimental setup of parametric process-based ONN. DCF, dispersion compensation fiber; HNLF, highly nonlinear fiber; MLL, mode-locked laser; OSA, optical spectrum analyzer; PC, polarization controller; PD, photodetector.

Simulation block diagram:

Fig. 2.　Simulation coding block diagram. Self-implemented simulation functions: SMF_simu( ), HNLF_simu( ), and DCF_simu( ).

information region is mapped to the temporal information region [14,24,25]. In this way, the encoded spectrum is recorded by the oscilloscope in the time domain. Because the spectrum is acquired by DFT in real-time, the frame rate of the ONN is equal to the 40 MHz repetition rate of the MLL. Finally, the observed spectrum is used as the input to a digital decision layer (ridge regression) to complete the classification task.

The feasibility of the proposed system is initially assessed through simulations. The ONN structure depicted in Fig. 1 is implemented using MATLAB. The coding block diagram and related parameter configurations are presented in Fig. 2, where functions such as SMF_simu( ), HNLF_simu( ), and DCF_simu( ) are custom-coded for simulating the respective components in the system. The testing dataset is the MNIST-digit dataset, which is a collection of handwritten digit images in different fonts. In our evaluation, 1200 samples are applied for classification tasks, with 1000 samples for training and 200 samples for testing. These 2D digit images are first resized into 1D arrays by pixels, and then converted into spectral form and modulated into the laser source signals by a waveshaper. The regression penalty term, denoted as alpha, is configured to be 100. A comparison of flattened data sample "1" both before and after the HNLF is illustrated in Fig. 3(a), focusing on an observation wavelength range of 1520–1600 nm. It can be observed that a significant nonlinear transformation is applied to the original data input. Meanwhile, the relationship between the decay of input power gain and accuracy is given in Fig. 3(b). By optimizing the input gain value, an improved accuracy from 77.17% to 83.78% is achieved. This indicates that in practical applications, by properly adjusting system variables, our system can be adapted to completely different tasks without hardware modifications.
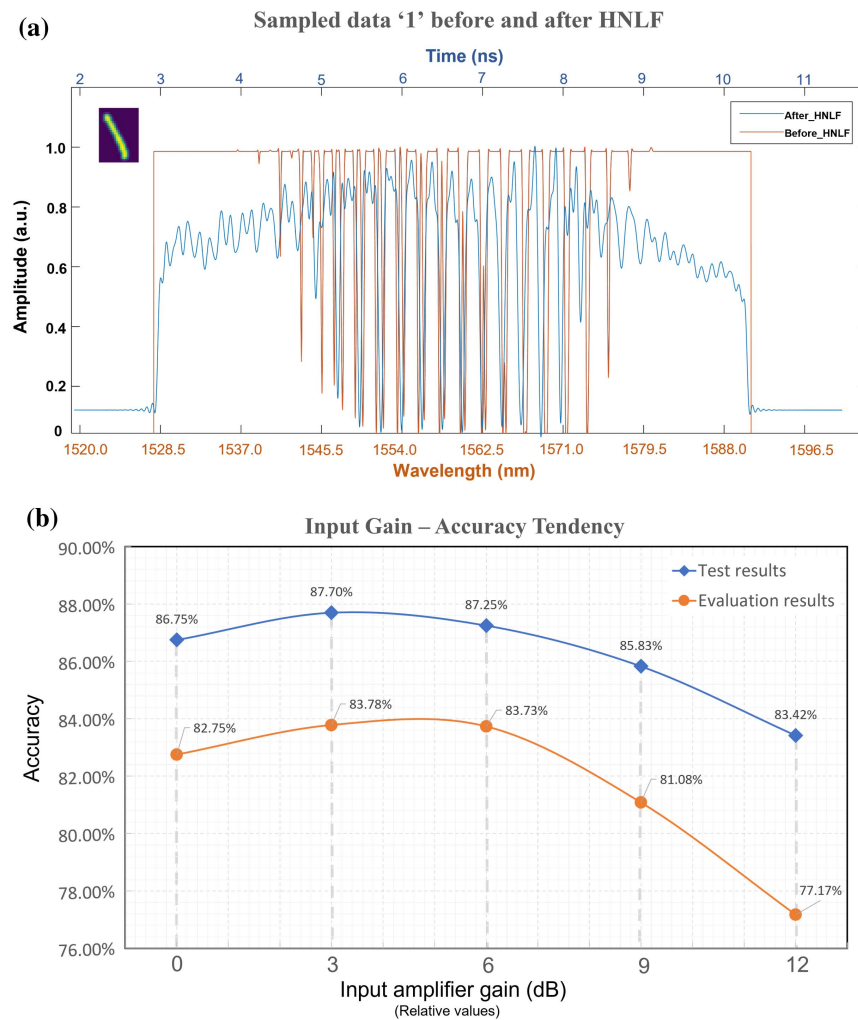
3. EXPERIMENTAL RESULTS

Based on the above results, the system is physically implemented with the proposed structure and the same parameter settings as in the simulation. The utilized waveshaper is the Finisar Waveshaper 4000B. In order to further investigate the performances regarding the versatility and accuracy improvement that benefited from the parametric process on different tasks, evaluations on both digit classifications and audio recognitions are carried out with the MNIST-digit and the Vowel audio datasets,

respectively. By utilizing the DFT method, the input data is processed at an overall computation frame rate of 40 MHz during the evaluations conducted with our specially designed laser source. With a network structure consisting of one full-connect (1-FC) layer, the system is capable of achieving a calculation speed exceeding 620 GFLOPs.

For the MNIST dataset, it can be observed that with the physical system, there is a distinct difference between the encoded spectrum structures before and after HNLF, as illustrated in Fig. 4(a), thereby resulting in a significant distinction between features and an increase in classification accuracy. The confusion matrix with and without nonlinearity interaction is shown in Fig. 4(b). Without nonlinear computations, tests on MNIST-digit data only provide an accuracy of 81.5%, whereas we obtain 87.7% accuracy on the test set with the additional nonlinear computations. Benefiting from additional complex physical effects, the physical system even achieves a further accuracy improvement compared to the simulation result, also as indicated in Fig. 3(b).

Since ridge regression is a linear model, the improvement in classification performance confirms that it originates from nonlinear computation. In order to compare our ONN system with conventional NNs, we applied various digital linear approaches [ridge regression, 1-FC, logistic regression, and linear discriminant analysis (LDA)], as well as digital activation functions (ReLU, GeLU, and leaky-ReLU), to the same dataset. The classification report, presented in Table 1, includes results for both optical input data (referring to the output of the optical system before the HNLF) and digital input data (referring to the original MNIST-digit data). It is important to highlight that our framework achieves a comparable classification accuracy when compared to these digital neural network models. Particularly noteworthy is that our system achieves competitive accuracy with the nonlinearity introduced by the parametric process using HNLF, even when compared to the nonlinearity provided by commonly used digital activation functions. This further validates the effectiveness of the novel method we propose.

Next, we further increase the number of digit samples to 6000, with 5000 samples for training and 1000 samples for testing. The confusion matrix for both tests with and without nonlinear computation is shown in Fig. 4(c). The

(a)

(b)

Fig. 3. (a) Simulated data sample (flattened, "1") with MNIST-digit database before nonlinear interaction in HNLF (red, frequency domain) and after HNLF (blue, time domain); (b) tendency of accuracy in relation of input power decay.
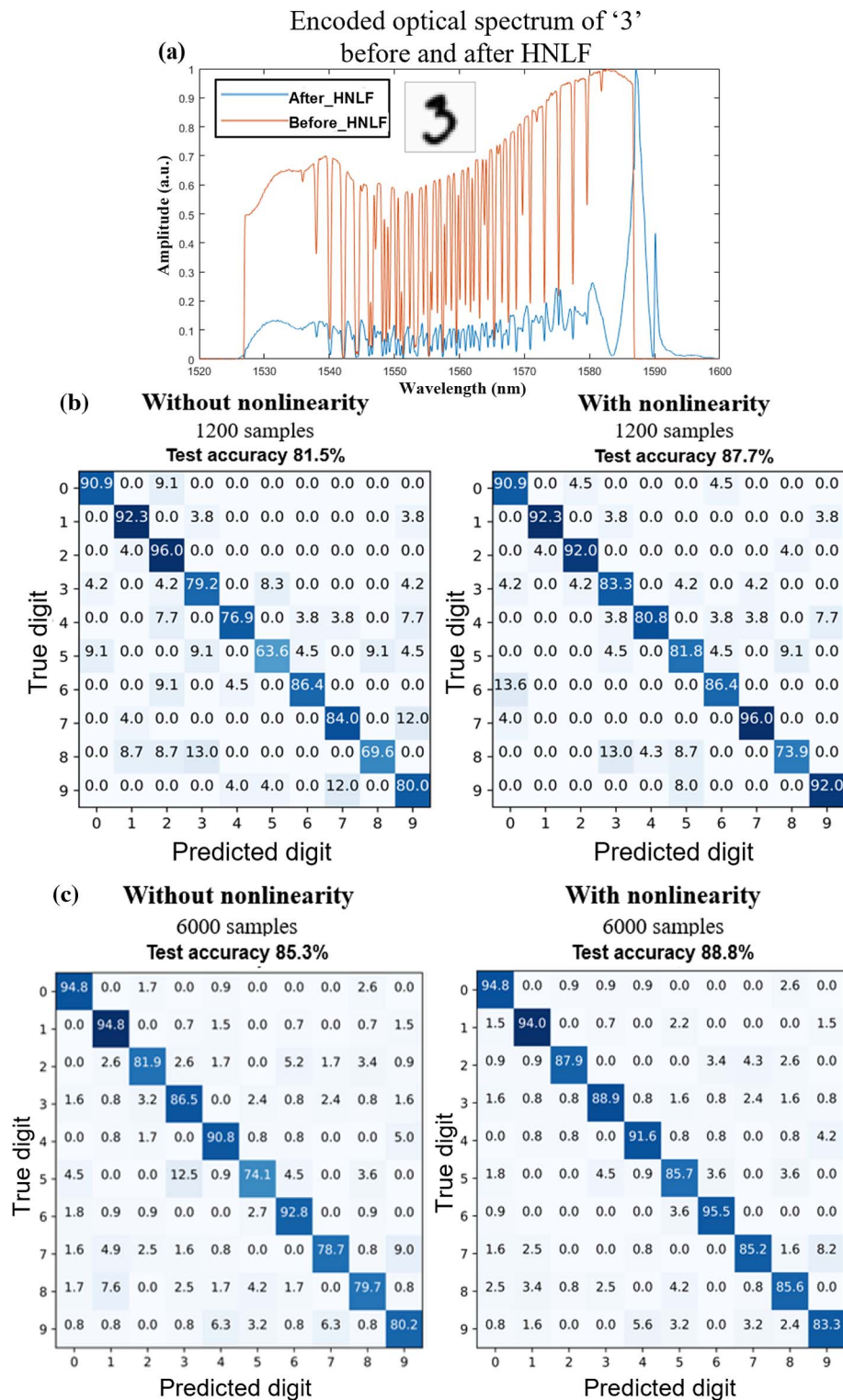
improvement in accuracy observed in both tasks is logical, as increasing the quantity of training samples is a well-known method for enhancing accuracy in learning outcomes, as mentioned in Ref. [26]. Specifically, it is noted that for tasks without nonlinear computations, the accuracy shows a notable change from 81.5% to 85.5%. On the other hand, in tasks involving nonlinear computations, the accuracy also increases, albeit by a smaller margin, from 87.7% to 88.8%. This minor increase can be attributed to the already competitive results achieved with smaller datasets. These findings suggest that our system is adept at extracting additional data features even when provided with limited data samples, thereby offering a potential solution to address the challenges associated with higher power and memory consumption that typically arise with larger datasets in order to achieve improved accuracy. The comparison in Table 1 can also confirm the same conclusion: with the nonlinearity via parametric process, the training parameter for similar accuracy is reduced from 50,890 to 784, which is a 98% decrease. With fewer data samples to process, the overall speed and efficiency of

the system, to some extent, are correspondingly further improved.

For audio recognition tasks, an audio dataset with 259 samples of different pronunciations, including "ae," "ah," "aw," "uw," "er," "iy," and "ih," from the Vowel dataset is used. Similarly, with the same system setup, these sampled audio tracks are then modulated into the source signals in spectral forms by the waveshaper. For ridge regressions, the ratio between the training set and the testing set is also 5:1, with the regression penalty term set to 100.

By optimizing the input power gain of the HNLF, a significantly increased accuracy from 80.3% to 97.6% is observed after HNLF, as shown by the confusion matrices with and without nonlinearity presented in Fig. 5(a). In order to intuitively show the improvements after HNLF, the linear discriminant analysis is illustrated in Fig. 5(b). Before HNLF, the feature components of different pronunciations are heavily coupled with each other, especially for "er" (brown dots), which is highly interfered by "iy" (red dots). This results in a 0% accuracy when recognizing "er." After HNLF, it can be seen that
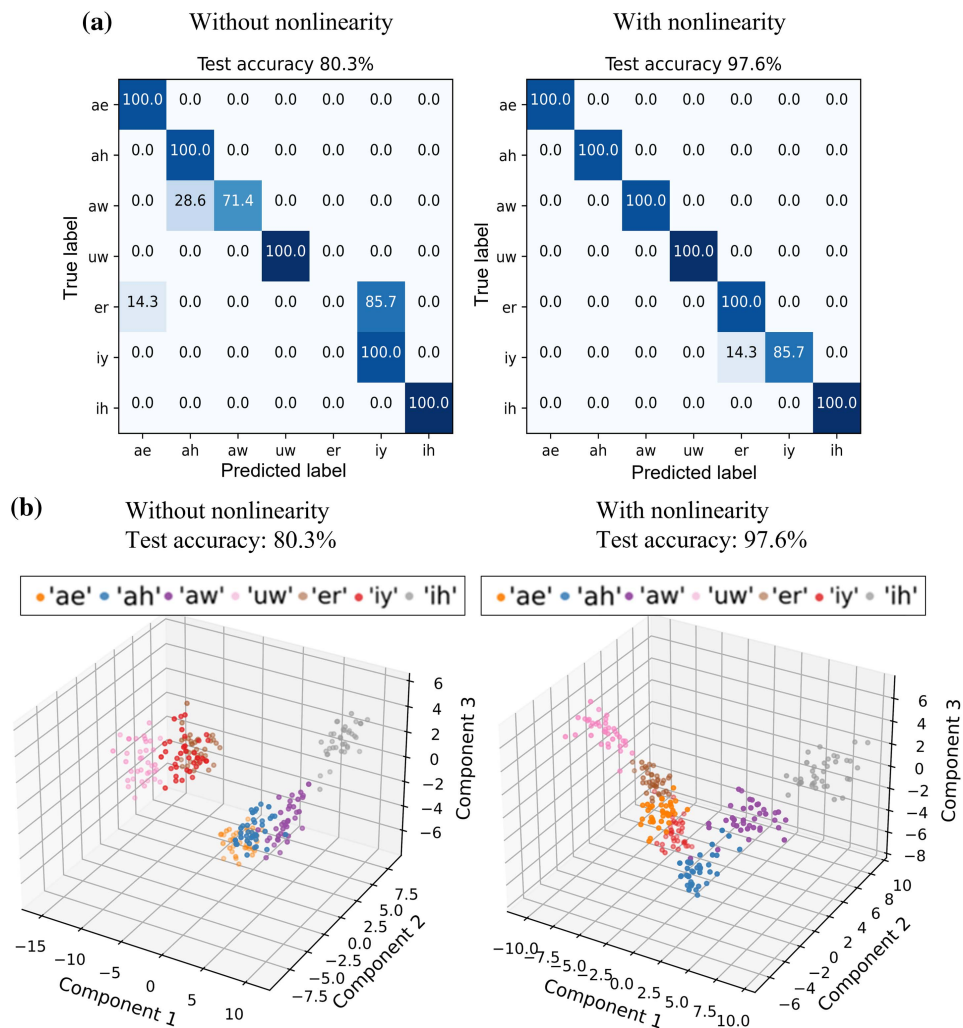
Fig. 4. (a) Encoded optical spectrum with MNIST-digit database before and after nonlinear interaction in HNLF; (b) confusion matrices with and without nonlinearity, 1200 samples; (c) confusion matrices with and without nonlinearity, 6000 samples.

the feature components are better extracted for the pronunciations, hence obtaining a more distinct distribution and a much higher recognition accuracy. Thus, it is proved that with the adaptive parametric process nonlinearity, our system can also be modified for a more complex audio recognition task in addition to relatively simpler digit classifications, and implies potential for more pragmatical types of applications, such as motion recognition and machine visioning.

Table 1. Classification Reports of MNIST-Digit Dataset[a]

	With Nonlinearity				
Input data	Optical	Optical	Optical	Optical	Optical
Architecture	Ridge regression HNLF	1-FC layer HNLF	1-FC layer ReLU	1-FC layer GeLU	1-FC layer leaky-ReLU
Accuracy	87.8%	87.3%	88.5%	88.2%	89.4%
Parameters	784	784	784	784	784
	Without Nonlinearity				
Input data	Optical	Digital	Digital	Digital	Digital
Architecture	Ridge regression	1-FC layer	Logistic regression	LDA (1200 samples) LDA (6000 samples)	1-FC layer
Accuracy	81.5%	80.3%	81.4%	65.8% 82.6%	87.5%
Parameters	784	784	784	784	50,890

[a]1200 samples, 0.01 learning rate (if applicable).

Fig. 5. (a) Confusion matrices with and without nonlinearity, 1200 samples; (b) linear discriminant analysis with and without nonlinearity, 1200 samples.

4. CONCLUSION

In summary, a novel high-speed and versatile parametric optical neural network that has high potential for applications in multiple areas is demonstrated in this paper. Aiming at the common challenge of flexible nonlinear activation functions in ONNs, we propose a parametric process-based realization that can be adapted for different tasks without modifying the hardware system. By implementing a specially designed MLL and DFT

process, the overall data processing speed can reach a high speed up to 40 MHz with a power consumption level of only a few milliwatts. Evaluations are carried out on both digit classification and audio recognition tasks. The accuracy is improved from 81.5% without nonlinearity to 88.8% after HNLF for the MNIST-digit dataset, and a satisfying increase in audio recognition from 80.3% to 97.6% is achieved for the Vowel dataset, which proves the versatility and potential of our system for a wide range of applications. Meanwhile, an insight into the comparison with the conventional ONN suggests that our nonlinear computation is able to effectively extract previously hidden data features with a 98% reduced scale of parameter size, which further increases the calculation capacity and efficiency of the system. Furthermore, with a consistent ultrafast operation time of 25 ns for each input image manipulation, the overall computing speed can be significantly enhanced to hundreds of TFLOPs by incorporating more complex network structures. If an FC layer with a 1:1000 pixel-node mapping structure is implemented, a computation speed exceeding 600 TFLOPs can be attained. This level of computational performance approaches the speed capabilities of a cutting-edge cluster of GPUs. Admittedly, the information-imparting process speed is still limited by the waveshaper. This challenge can be potentially overcome with temporal modulation methods, which will be a worthy direction in our future work.

Disclosures. The authors declare no conflicts of interest.

Data Availability. All data and codes are available from the authors upon reasonable request.

REFERENCES

1. M. Baek, F. DiMaio, I. Anishchenko, *et al.*, "Accurate prediction of protein structures and interactions using a three-track neural network," Science **373**, 871–876 (2021).
2. J. Jumper, R. Evans, A. Pritzel, *et al.*, "Highly accurate protein structure prediction with AlphaFold," Nature **596**, 583–589 (2021).
3. M. Partridge, "Spectra evolution during coating," https://figshare.com/articles/dataset/Spectra_evolution_during_coating/1004612?file=3177041 (2019).
4. Y. Rivenson, Z. Göröcs, H. Günaydin, *et al.*, "Deep learning microscopy," Optica **4**, 1437–1443 (2017).
5. D. Patterson, J. Gonzalez, Q. Le, *et al.*, "Carbon emissions and large neural network training," arXiv, arXiv:2104.10350 (2021).
6. M. Asghari and A. V. Krishnamoorthy, "Energy-efficient communication," Nat. Photonics **5**, 268–270 (2011).
7. A. Momeni and R. Fleury, "Electromagnetic wave-based extreme deep learning with nonlinear time-Floquet entanglement," Nat. Commun. **13**, 2651 (2022).
8. S. Bavikadi, A. Dhavlle, A. Ganguly, *et al.*, "A survey on machine learning accelerators and evolutionary hardware platforms," IEEE Des. Test **39**, 91–116 (2022).
9. G. Wetzstein, A. Ozcan, S. Gigan, *et al.*, "Inference in artificial intelligence with deep optics and photonics," Nature **588**, 39–47 (2020).
10. J. Spall, X. Guo, and A. I. Lvovsky, "Hybrid training of optical neural networks," Optica **9**, 803–811 (2022).
11. T. Zhou, L. Fang, T. Yan, *et al.*, "*In situ* optical backpropagation training of diffractive optical neural networks," Photonics Res. **8**, 940–953 (2020).
12. D. Huang, Y. Shi, F. Li, *et al.*, "Fourier domain mode locked laser and its applications," Sensors **22**, 3145 (2022).
13. J. Kang, P. Feng, X. Wei, *et al.*, "102-nm, 44.5-MHz inertial-free swept source by mode-locked fiber laser and time stretch technique for optical coherence tomography," Opt. Express **26**, 4370–4381 (2018).
14. A. Mahjoubfar, D. V. Churkin, S. Barland, *et al.*, "Time stretch and its applications," Nat. Photonics **11**, 341–351 (2017).
15. S. R. Dubey, S. K. Singh, and B. B. Chaudhuri, "Activation functions in deep learning: a comprehensive survey and benchmark," Neurocomputing **503**, 92–108 (2022).
16. U. Teğin, M. Yıldırım, İ. Oğuz, *et al.*, "Scalable optical learning operator," Nat. Comput. Sci. **1**, 542–549 (2021).
17. J. Shi, L. Zhou, T. Liu, *et al.*, "Multiple-view D²NNs array: realizing robust 3D object recognition," Opt. Lett. **46**, 3388–3391 (2021).
18. H. Dou, Y. Deng, T. Yan, *et al.*, "Residual D²NN: training diffractive deep neural networks via learnable light shortcuts," Opt. Lett. **45**, 2688–2691 (2020).
19. P. L. McMahon, A. Marandi, Y. Haribara, *et al.*, "A fully programmable 100-spin coherent Ising machine with all-to-all connections," Science **354**, 614–617 (2016).
20. T. Zhou, F. Scalzo, and B. Jalali, "Nonlinear Schrödinger kernel for hardware acceleration of machine learning," J. Lightwave Technol. **40**, 1308–1319 (2022).
21. X. Jiang, D. Wang, X. Chen, *et al.*, "Physics-informed neural network for optical fiber parameter estimation from the nonlinear Schrödinger equation," J. Lightwave Technol. **40**, 7095–7105 (2022).
22. M. Nakajima, K. Tanaka, and T. Hashimoto, "Neural Schrödinger equation: physical law as deep neural network," IEEE Trans. Neural Netw. Learn. Syst. **33**, 2686–2700 (2021).
23. C. S. Brès, A. Della Torre, D. Grassani, *et al.*, "Supercontinuum in integrated photonics: generation, applications, challenges, and perspectives," Nanophotonics **12**, 1199–1244 (2023).
24. G. Herink, B. Jalali, C. Ropers, *et al.*, "Resolving the build-up of femtosecond mode-locking with single-shot spectroscopy at 90 MHz frame rate," Nat. Photonics **10**, 321–326 (2016).
25. K. Goda, K. K. Tsia, and B. Jalali, "Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena," Nature **458**, 1145–1149 (2009).
26. A. Alwosheel, S. van Cranenburgh, and C. G. Chorus, "Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis," J. Choice Modelling **28**, 167–182 (2018).