Optimization of MZI-Based Photonic Neural Networks via Pseudo-Real Number Theory and GridNet Architecture

Li Wang[1*] Tao Wang[2]
[1]Luohe Food Engineering Vocational University, Luohe 462000, China
[2]Luohe Vocational Technology College, Luohe 462000, China
E-mail: wantlili@163.com, 3381062@qq.com
[*]Corresponding author

Keywords: pseudo-real number theory, neural network, MZI array, artificial intelligence, photon chip

Received: September 12, 2025

As information technology continues to evolve, the computing performance of existing electronic chips approaches its computational limit, and integrated photon technology has the advantages of fast computing speed and strong anti-interference ability. Therefore, a photon neural network algorithm based on pseudo real number theory is proposed in the study. This method employs a Mach-Zehnder interferometer array to construct a basic convolutional neural network, wherein the Mach-Zehnder interferometer splitter is replaced by a directional coupler to form the fundamental unit. The study employs singular value decomposition for matrix multiplication calculations, performs unitary matrix decomposition based on the rectangular decomposition principle, and utilizes pseudo-real number theory for weight training within the photonic neural network architecture. By extracting the real part of complex optical fields, this approach enables optical systems to compute arbitrary real-valued matrices. Experiments demonstrate that on the MNIST handwritten digit dataset, the photonic neural network converges to a loss function value of 0.08 after 20 iterations, achieving a minimum loss function value 0.02 and 0.09 lower than other algorithms respectively. Module U, housing the largest array of Mach-Zehnder interferometers, exhibits greater susceptibility to noise. The computational speed of the photonic neural network surpasses traditional methods by six orders of magnitude. Following optimization, its computational speed exceeds other approaches by 13 ns and 7 ns respectively. Ablation experiments reveals that the directional coupler module exerts the most significant influence on enhancing computational speed, while pseudo-real number theory has the greatest impact on recognition accuracy. From this, the approach proposed by the research can effectively improve the computing speed of neural networks and reduce the operating costs in actual production.

Povzetek: Raziskava kaže, da uporaba algoritma kačjega pastirja v brezžičnih senzorskih omrežjih izboljša energijsko učinkovitost, podaljša življenjsko dobo omrežja ter zagotavlja bolj prilagodljivo in zanesljivo usmerjanje v primerjavi s klasičnimi metodami, kot je PSO.

1 Introduction

Since the 21st century, computer technology, artificial intelligence technology, and big data technology have been rapidly advancing, and various algorithms and neural networks have become increasingly complex [1]. Traditional electronic computing systems face speed and energy consumption bottlenecks when processing large-scale data. Especially in the era of mobile Internet, various data emerge in endlessly, which has caused great pressure on traditional electronic computing systems [2-3]. Photon neural networks integrate photon and artificial neural networks, utilizing the high-speed propagation characteristics of light to complete numerous computational tasks within an extremely brief period, greatly improving computational efficiency [4]. At the same time, it also has low power consumption characteristics, with optical devices consuming much less power than electronic devices, making them more energy-efficient when handling complex computing tasks [5-6]. However, the neural network architecture for optical devices is not yet fully developed. Tian Y et al. raised a new activation approach to address the immature development of nonlinear activators in photon neural networks. This method used the Kramer Kronig relationship induction, which activated the relationship between light field amplitude and phase, and learned nonlinear features in relatively low dimensional convolutional spaces through linear features in high-dimensional convolutional spaces. Experiments showed that the learning ability of this method was significantly better than that of non-activated linear networks, comparable to networks with common activation functions [7]. Huang Y et al. proposed a PNN topology with Mach-Zehnder Interferometer (MZI) to enhance the scalability of photon neural networks. This method used a single tuned phase shifter to achieve vector matrix multiplication of any non-negative or real valued matrix, and designed and manufactures an 8×8 reconfigurable chip. Experiments showed that this method

achieved over 85% inference accuracy and lower optical path loss in handwritten digit recognition tasks [8]. Ouyang J et al. proposed a programmable photon neural network solver for solving the problem of quadratic unconstrained binary optimization in optoelectronic schemes. This method used photon chips to perform optical vector matrix multiplication to solve the cost function, and used an electronic processor to run heuristic algorithms to search for the optimal solution. The experiment showed that the calculation speed of this method was 1.66 TFLOP/s, which could effectively solve quadratic unconstrained binary optimization problems [9]. Pai S proposed a new integrated photon neural network to improve the performance of traditional neural networks. This method used a non-linear interleaved MZI mesh network to effectively convert optical encoded inputs, and used in-situ backpropagation to solve classification tasks. The backpropagation gradient of the phase shifter voltage was measured by interfering with the forward and backward propagation light. The experiment showed that the testing accuracy of this method was greater than 94%, and energy scaling analysis indicated a scalable approach for machine learning [10].

Zheng Z et al. raised a new dual adaptive training method to address the issue of significant accumulation of system errors in large-scale photon neural networks. This method introduced a system error prediction network with task similarity joint optimization function to achieve high similarity mapping between numerical models and physical systems, and performed high-precision gradient calculation during the dual backpropagation training process. Experiments showed that this method successfully trained photon neural networks with significant systematic errors and achieved high classification accuracy [11]. Ashtiani F et al. raised a new integrated end-to-end photon deep neural network to address the issues of on-chip optical nonlinearity and losses in photon devices. This network directly processed the light waves of the pixel array on the impact plate, performed sub-nanosecond image classification when the light waves propagated through the neuron layer, and performed linear calculations optically in the neurons. The experiment showed that the classification time of the network was less than 570ps, and the accuracy of two and four categories of handwritten letters was higher than 93.8% and 89.8%, respectively [12]. Kirtas M et al. raised a new electrical signal transmission method to address the issue of expensive modules for high-speed transmission of electrical signals in photon neural networks. This method focused on the quantization phenomenon in the model as an additional source of uncertainty, providing a framework that meets photon standards for training photon deep learning models with limited accuracy. Experiments demonstrated that this approach could effectively reduce the demand for expensive accessories while ensuring the computational speed and classification accuracy of photon neural networks [13].

To sum up, current research approaches have investigated the development of photon neural network structures from various angles and obtained specific outcomes. Nevertheless, there has been comparatively limited investigation into the linear computing level of photon neural networks, and there is no corresponding training method. Therefore, this research proposes a photonic neural network algorithm based on pseudo-real number theory. In terms of theoretical innovation, it pioneers the application of pseudo-real number theory in photonic neural network training. By employing optical projection computing units, it achieves an effective mapping from the complex domain to the real domain, thereby overcoming the limitation of conventional methods that can only process real-valued matrices. Architectural innovation involves replacing the beam splitter in conventional Mach–Zehnder Interferometers (MZI) with directional couplers. Combined with the GridNet architecture, this achieves an end-to-end trainable photonic neural network that significantly enhances computational speed while maintaining high precision. Algorithmically, a unitary matrix optimization method based on rectangular decomposition is proposed, working in tandem with singular value decomposition to provide a viable technical pathway for implementing large-scale photonic neural networks. The research aims to improve the computational speed of neural networks and reduce operational costs in practical production. A comparative analysis of the relevant work is presented in Table 1.

In Table 1, although existing research has achieved certain results in photonic neural networks, issues such as insufficient recognition accuracy and poor computational real-time performance persist. Concurrently, most approaches exhibit high hardware complexity, which constrains the further development of related technologies, while certain methods present significant limitations.

The research aims to address the following key questions: (1) Can pseudo-real number theory enhance the training convergence and recognition accuracy of MZI photonic neural networks without substantially increasing hardware complexity? (2) How does replacing conventional beam splitters with directional couplers affect the computational speed of photonic neural networks? What mechanisms underpin this performance improvement? (3) What are the relative contributions of each proposed module (singular value decomposition, rectangular decomposition, directional coupler, pseudo-real number theory, GridNet) to the overall system performance?

Research hypotheses: (1) Pseudo-real number theory can enhance performance through effective domain mapping; (2) Directional couplers can simplify system architecture and accelerate computation; (3) Synergistic effects exist among modules, collectively determining the system's final performance.

2 Methods and materials

2.1 Construction of photon neural network based on MZI

MZI is a measuring device based on optical interference phenomena, which can split a beam of light into two beams, propagate along two different paths, and then

recombine them to form an interference pattern. MZI has the characteristics of high sensitivity, modulatability, and integration, and can be applied in fields such as optical communication, optical measurement, optical computing, and sensors [14]. In photon neural networks, MZI is a key optical component that can perform linear operations such as matrix multiplication and weighted summation in neural networks [15]. In MZI, matrix multiplication is usually calculated using weight matrix singular value decomposition, and the singular value decomposition calculation of the matrix is in equation (1).

$$W = U\Sigma V^T \tag{1}$$

Table 1: Comparative analysis of relevant work

Method	Key technologies	Accuracy	Calculation speed	Hardware complexity	Principal limitations
Tian et al. [7]	Kramer-Kronig activation	95.6%	57ns	Medium	Non-linear activation implementation complexity
Huang et al. [8]	Single-tuned phase shifter	85.3%	31ns	Low	Restricted to non-negative matrices
Ouyang J et al. [9]	Optical vector matrix multiplication	87.5%	1.66s	High	Insufficient real-time capability
Pai et al. [10]	In-situ backpropagation	94.0%	69ns	Medium	Complex training process
Zheng et al. [11]	Dual adaptive training	82.4%	35ns	Medium	Focused on error compensation
Ashtiani et al. [12]	End-to-end photonic networks	93.8%	570ps	High	Limited scalability
This study	Pseudo-real number theory + GridNet	98.2%	12ns	Medium	/

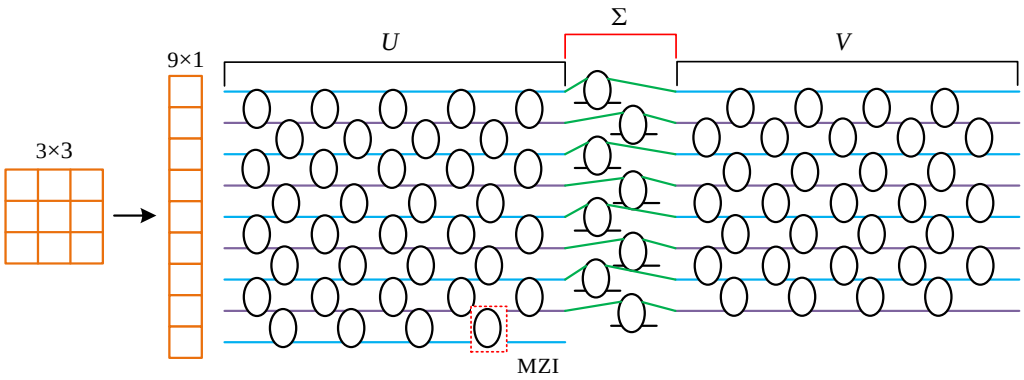

Figure 1: MZI array arrangement in CNN

In equation (1), W denotes a weight matrix of dimension $m \times n$, U denotes a unitary matrix of dimension $m \times m$, Σ denotes a diagonal matrix of dimension $m \times n$, V denotes a unitary matrix of dimension $n \times n$, and V^T denotes the complex conjugate of the unitary matrix V. The key to optical implementation lies in the fact that any unitary matrix can be physically realized through a MZI network, while a diagonal matrix can be achieved using a set of optical attenuators or amplifiers along a parallel waveguide. This study designs a linear matrix based on singular value decomposition theory, decomposing the transmission matrix (TM) into two unitary matrices and one diagonal matrix, and transferring them to an optical convolutional neural network (CNN) through mapping. This study is able to calculate any real valued matrix using this method, enabling data transmission in optical neural networks [16]. The MZI array arrangement in the CNN constructed in this study is in Figure 1.

In Figure 1, the convolutional kernels in the CNN are all unitary matrices using MZI, with a total of 8 kernels, each with a size of 3×3. The input of the convolutional layer is a 9×1 vector formed by expanding each unitary matrix. All unitary matrices are expanded to form a 9×8 matrix, which is then decomposed into two unitary matrices and one diagonal matrix using singular value decomposition, and connected using MZI. In the first unitary matrix area, there are a total of 36 MZI connection modules, in the diagonal matrix area, there are a total of 8 MZI modules, and in the second unitary matrix area, there are a total of 28 MZI modules. The MZI module in the unitary matrix region has two input terminals and two output terminals, while the MZI in the diagonal matrix region has only one input terminal and one output terminal. A global optical amplification device is added to the output terminals of all MZIs to compensate for attenuation

during optical signal transmission. The unitary matrix region specifically denotes a particular physical area on a photonic chip formed by multiple MZIs interconnected in a specific grid structure. The overall transmission matrix of this region is mathematically equivalent to a unitary matrix. The study adopts a grid arrangement to connect the MZIs in the convolutional layer, thereby achieving both matrix multiplication and addition simultaneously. MZI typically comprises an adjustable phase shifter and two beam splitters. The phase shifter imposes phase verification upon the optical mode, with its transmission matrix calculated as shown in equation (2) [17].

$$T_{PS} = \begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix} \tag{2}$$

In equation (2), T_{PS} represents the TM, ϕ represents the phase difference introduced by the phase shifter. The power ratio of reflected and transmitted light in the beam splitter is 50:50, resulting in the mixing of two input light signals. The transmission matrix of the beam splitter is calculated as shown in equation (3).

$$T_{BS} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \tag{3}$$

In equation (3), T_{BS} represents the TM of the beam splitter, i represents the imaginary unit, and the first column and first row of the TM of the beam splitter represent the amplitude coefficients of transmission and reflection, respectively. In the grid arrangement of the MZI, it is typically formed by cascading phase shifters, beam splitters, phase shifters, and beam splitters in sequence. This is represented as a 2×2 unit transmission matrix, derived by sequentially multiplying the transmission matrices. The overall transmission matrix of the MZI is shown in equation (4).

$$T_{MZI} = \begin{bmatrix} i\sin(\phi/2)e^{i\phi/2} & i\cos(\phi/2)e^{-i\phi/2} \\ \cos(\phi/2)e^{i\phi/2} & \sin(\phi/2)e^{-i\phi/2} \end{bmatrix} \tag{4}$$

In equation (4), T_{MZI} represents the overall TM of MZI. The study adopts the principle of rectangular decomposition for unitary matrix decomposition, and the unitary matrix obtained after 4 rounds of traversal is in equation (5).

$$U = DT_{1,2}T_{2,3}T_{3,4}T_{1,2}T_{3,4}^{-1}T_{2,3}^{-1} \tag{5}$$

In equation (5), D represents the diagonal matrix and T represents the rotation matrix. The study adopts spatial light theory to construct a non-destructive optical beam splitter, and uses two phase shifters to record the phase changes of two incident lights. The relationship between the input and output ports of the non-destructive optical beam splitter is in equation (6).

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = \begin{bmatrix} e^{i\varphi}\sin\omega & \cos\omega \\ e^{i\varphi}\cos\omega & -\sin\omega \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{6}$$

In equation (6), O_1 represents output optical signal 1, O_2 represents output optical signal 2, φ represents the phase value of the phase shifter, ω represents the transmission property parameters of the beam splitter, I_1 represents input optical signal 1, and I_2 represents input optical signal 2. The calculation of $\sin\omega$ and $\cos\omega$ for the beam splitter is in equation (7).

$$\begin{cases} \sin\omega = \sqrt{r} \\ \cos\omega = \sqrt{t} \end{cases} \tag{7}$$

In equation (7), r represents the reflectivity of the beam splitter and t represents the refractive index of the beam splitter. In the array composed of MZIs, the linear matrix obtained after inputting the optical signal is in equation (8).

$$\begin{bmatrix} W_{11} & \cdots & W_{1N} \\ \vdots & \ddots & \vdots \\ W_{N1} & \cdots & W_{NN} \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} O_1 \\ \vdots \\ O_N \end{bmatrix} \tag{8}$$

After obtaining the linear matrix, singular value decomposition is used to decompose the weight matrix, as shown in equation (9).

$$O = U\Sigma V^T \cdot I \tag{9}$$

By utilizing the principle of rectangular decomposition, the unitary matrices U and V^T are processed, and the decomposed MZI array is in Figure 2.

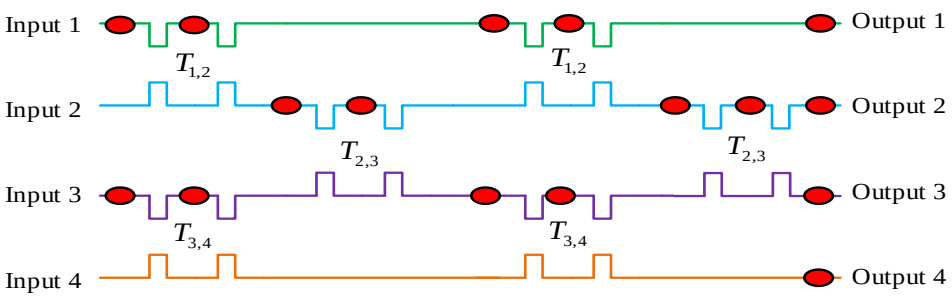

Figure 2: MZI array after unitary matrix decomposition

In Figure 2, there are four optical signal channels. After the optical signal is input into the MZI, it passes through $T_{1,2}$, $T_{3,4}$, $T_{2,3}$, $T_{1,2}$, $T_{3,4}$, and $T_{2,3}$ in sequence before being output. At the same time, to assess the functionality of the photon neural network constructed by MZI array, fidelity is used to measure the similarity between the target matrix and the matrix formed by MZI, as calculated in equation (10).

$$F = \left| \frac{\delta\left(U^{\dagger} U_g\right)}{N} \right|^2 \tag{10}$$

In equation (10), F represents the similarity between two matrices, δ represents the trace of the target matrix, $U^{\dagger}$ represents the conjugate transpose of the MZI matrix, U_g represents the target matrix, and N represents the size of the target matrix. The structure of the constructed photon neural network is in Figure 3.

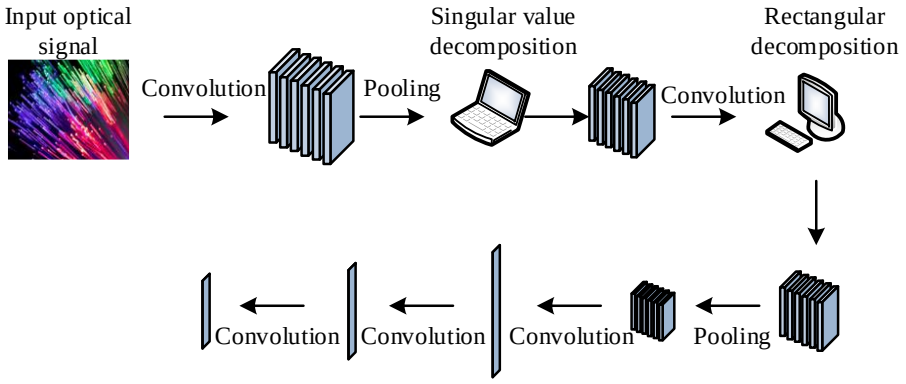

Figure 3: Specific structure of photon neural network

In Figure 3, the optical signal is input through two layers of 4×4 convolution kernels, and the transposed convolution kernels are processed using singular value decomposition. The decomposed matrix is then subjected to two max pooling layers, and the obtained unitary matrix is processed using the principle of rectangular decomposition. The processed information is sequentially processed through two convolutional layers, two max pooling layers, and three convolutional layers before being output.

2.2 Optimization of photon neural networks based on pseudo-real number theory

The TM of the photon neural network constructed using MZI can only represent the amplitude information of light in the real domain, while it can simultaneously represent amplitude and phase information in the complex domain, thus achieving more complex linear and nonlinear operations [18]. Therefore, the study adopts a training method based on pseudo-real number theory to optimize the TM of the photon neural network. Pseudo-real number theory constitutes a computational framework that maps optical signals from the complex domain to the real domain via optical projection computational units (such as directional couplers). This enables traditional real-number-based optimization algorithms to be directly applied for training photonic neural networks. Specifically, this theory leverages interference principles to extract real-part information from complex fields, thereby circumventing complex number operations while preserving critical phase information. The projection calculation unit constitutes a specific optical path structure formed by directional couplers. Its function is to interfere the complex optical field output from the MZI network with a reference light source. By measuring the intensity of the resulting interference pattern, it projects the real-part information of the complex optical field onto the measurable intensity domain. The study uses two directional couplers to replace the beam splitter in MZI. The improved MZI device is expressed using a two-dimensional unitary matrix, which exhibits tunable phase-shifting capability and wide adaptability. The output matrix of the improved MZI is in equation (11).

$$W' = \begin{bmatrix} e^{i\varphi}\sin\theta & e^{i\varphi}\cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix} \tag{11}$$

In equation (11), W' represents the output matrix of the improved MZI and θ represents the parameters of the internal phase shifter. According to the theory of pseudo-real numbers, any array composed of multiple different MZI devices can perform matrix multiplication operations to complete computational tasks. The input and output optical signals of multiple MZI networks are shown in equation (12) [19].

$$\begin{bmatrix} \sqrt{O_1}\exp\left(i\psi_1\right) \\ \sqrt{O_2}\exp\left(i\psi_2\right) \\ \sqrt{O3}\exp\left(i\psi_3\right) \end{bmatrix} = T_{MZI} \begin{bmatrix} \sqrt{I_1} \\ \sqrt{I_2} \\ \sqrt{I_2} \end{bmatrix} \tag{12}$$

In equation (12), ψ_1 represents the phase shift of the first MZI output optical signal, ψ_2 represents the phase shift of the second MZI output optical signal, and ψ_3 represents the phase shift of the third MZI output optical signal. Because commonly-used detection methods can only measure the intensity of light and cannot directly measure the phase value, only the real matrix calculated by the MZI array can be used. After introducing the theory of pseudo-real numbers, there is no longer a need for multiple MZI arrays to jointly calculate matrices. By using projection calculation units and any MZI network, all real matrices can be calculated. The study uses a directional coupler as the corresponding projection calculation unit, and the arbitrary output optical signal and initial input optical signal of MZI are used as the input terminals of the directional coupler. One of the output terminals of the directional coupler is calculated as shown in equation (13).

$$O_{z1} = |E_I|^2 + |E_O|^2 - 2|E_O||E_I|\cos\varphi' \quad (13)$$

In equation (13), O_{z1} represents the first output of the directional coupler, E_I represents the initial input light, E_O represents any output light of the MZI, φ' represents the phase difference between two input optical signals, the phase value of the initial input light is taken as zero, and the amplitude is taken as unit 1. Meanwhile, because the initial input light is directly fed into the directional coupler, while the output light undergoes attenuation after passing through the MZI network, the intensity of the input light is significantly higher than that of the output light. The output of the directional coupler is approximately equal to the real part of the amplitude of the output light plus one unit. The output calculation of the second directional coupler is in equation (14).

$$O_{z2} = |E_I|^2 + |E_O|^2 + 2|E_O||E_I|\cos\varphi' \quad (14)$$

In equation (14), O_{z2} represents the output of the second directional coupler. By using a detector to extract the output signal of the directional coupler and then processing it with high pass filtering, accurate output optical amplitude real part can be obtained. Therefore, based on the theory of pseudo-real numbers, a projection calculation unit is introduced to optimize the TM of MZI array and convert it into a unitary matrix. After the conversion is completed, any MZI can be used to calculate any real matrix. The study used GridNet network architecture and MZI array to build a photon neural network. The training process of GridNet photon neural network is in Figure 4.

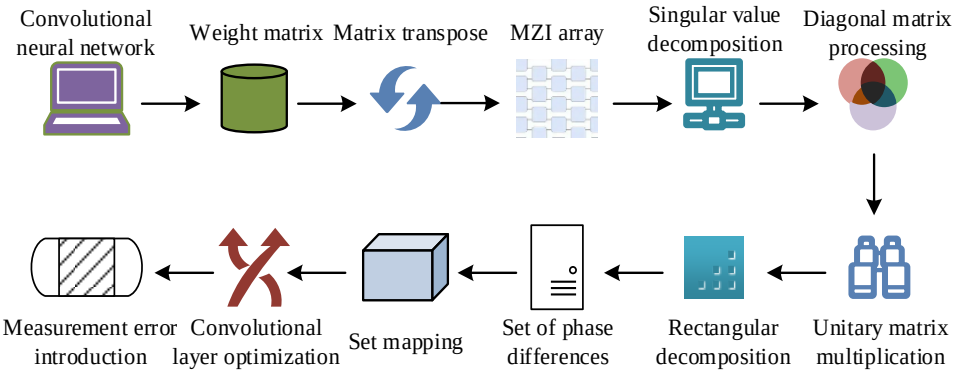

Figure 4: Specific training flow of GridNet photon neural network

In Figure 4, a CNN consisting of two convolutional layers is first set up, with a convolutional layer size of 4×4. The study calculates the weight matrix of the CNN, then transposes the matrix, and uses the transposed matrix as the TM of MZI. Singular value decomposition is performed on the TM to obtain a diagonal matrix and two unitary matrices, both with a scale of 4×4. The diagonal matrix is processed to ensure that each element is less than or equal to 1, thereby simulating the attenuation effect of MZI on light. The research adds imaginary parts to each element of a diagonal matrix to create a new unitary matrix. The three unitary matrices are multiplied to obtain a new TM. The transfer matrix is also a unitary matrix, which can be expressed using the Grid Wise Network (GridNet). Using the rectangular decomposition method to process the new TM, a set of MZI phase difference values is obtained, and then the set of phase difference values is mapped to the MZI array. Based on the mapped MZI array layout, convolutional layer optimization is performed by changing the weights during convolutional layer training to phase difference values, adjusting the training focus position, and introducing various measurement errors that may occur in MZI actual use, such as beam splitters, phase shifters, and detectors. GridNet's architecture comprises two convolutional layers, two max-pooling layers, two convolutional layers, two max-pooling layers, and three convolutional layers. The ReLU activation function is employed, utilizing the Adam optimizer with a batch size of 128. GridNet adopts L2 regularization with a learning rate of 0.001. The cross-entropy loss function (LF) of the photon neural network is in equation (15) [20-21].

$$L = -\frac{1}{x}\sum_{j=1}^{x}\sum_{k=1}^{K} p\left(s_j^k\right)\log q\left(s_j^k\right) \quad (15)$$

In equation (15), L represents the cross entropy LF, x represents the total number of data, K represents the total number of data categories, k represents a certain data category, j represents the number of data, j represents the true probability that data j belongs to category k, and $q\left(s_j^k\right)$ represents the predicted probability that data j belongs to category k. The hardware architecture of the constructed photon neural network is in Figure 5.

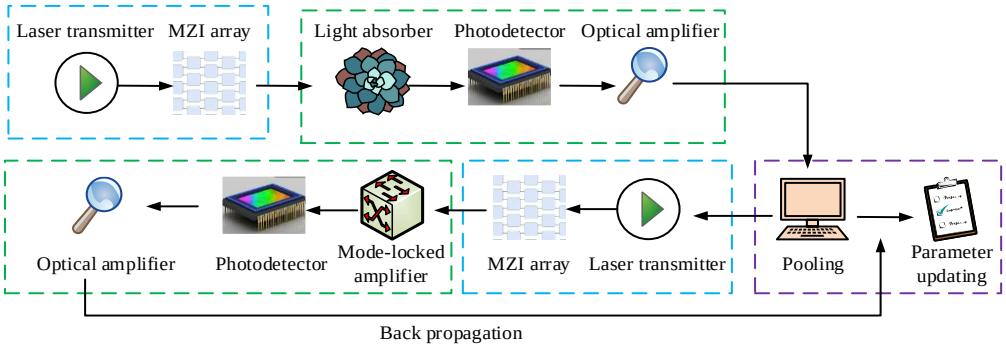

Figure 5: Hardware architecture of photon neural network

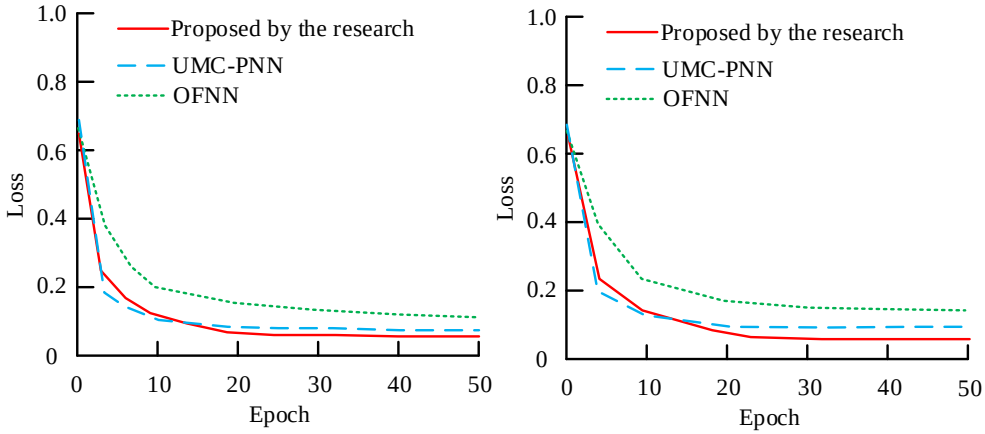

Figure 6: Comparison of LFs of different photon neural networks

In Figure 5, the photon neural network consists of two parts: optical and electrical. The optical part includes a laser emitter and an MZI array, while the nonlinear layer of the optical part consists of a light absorber, a detector, and a global optical amplification device. The optical signal is finally input into the field programmable gate array through an amplification device for pooling calculation. After the calculation is completed, it is then input into the fully connected layer composed of a laser emitter and an MZI array. Then, a lock-in amplifier is used for optical signal amplification to improve detection accuracy. The final output signal can be backpropagated to the field programmable gate array for parameter updating of the photon neural network. The evaluation metrics selected for this study encompassed not only the cross-entropy loss function but also accuracy, F1 score, and root mean square error (RMSE). Among these, accuracy served as the most intuitive indicator in classification tasks, reflecting the model's overall capability for correct classification. To address potential misleading interpretations of accuracy on imbalanced datasets, the study employed the F1 score, which represented the harmonic mean of precision and recall. This metric provided a more comprehensive assessment of the model's balanced performance across all categories, proving particularly suitable for verifying whether the model exhibited bias or blind spots towards specific categories. RMSE was employed to measure the discrepancy between the probability distribution of the network's predicted outputs and the true one-hot encoded label distribution. A lower RMSE indicated that the network's predicted probabilities were numerically closer to the true labels, reflecting not only classification accuracy but also the quality and calibration of the model's output confidence levels.

3 Results

3.1 Experimental analysis of photon neural network based on MZI

The hardware environment for running the photon neural network proposed by the research was Intel Core i5-12600H @ 2.60 GHz, GPU was NVIDIA GeForce RTX 3060Ti, memory was 16GB, operating system was Windows 11, and programming language was Python 3.8.0. During training, the learning rate of the neural network was set to 0.001, and the maximum number of iterations was set to 50. The experimental results were obtained from numerical simulation environments. Actual hardware implementation must account for non-ideal factors such as thermal noise and manufacturing tolerances, which will be further explored in future work. The study used the Modified National Institute of Standards and Technology database (MNIST) handwritten font recognition dataset for experimental analysis. The dataset contains 60000 training samples and

10000 test samples, each of which is a 28×28 grayscale image. The study utilized comparative approaches such as Optical Feedforward Neural Network (OFNN) and United Microelectronics Center Photon Neural Network (UMC-PNN). All experiments were conducted multiple times, with five randomized replicates, and the results were averaged. The comparison of LFs for different photon neural networks is in Figure 6.

In Figure 6 (a), when the iteration count was 5, the LFs of each method gradually converged. The proposed method tended to converge at 20 iterations, with a minimum LF value of 0.08, which was 0.02 and 0.09 lower than UMC-PNN and OFNN, respectively. UMC-PNN converged quickly in the early stages of iteration, but the minimum LF was lower than the method proposed in the study. In Figure 6 (b), when the iteration number was 0, that is, without training, the initial LFs of the three methods were relatively similar. In the test set, the minimum LF of the proposed method remained basically unchanged, and the initial convergence speed decreased slightly. The LFs of UMC-PNN and OFNN decreased by 0.03 and 0.02, respectively. The comparison of recognition accuracy between different photon neural networks is in Figure 7.

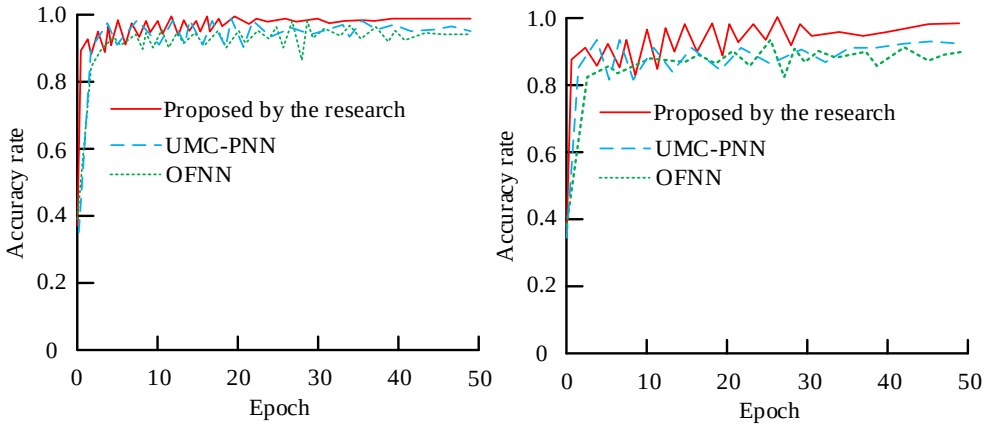

Figure 7: Comparison of recognition accuracy rates of different photon neural networks

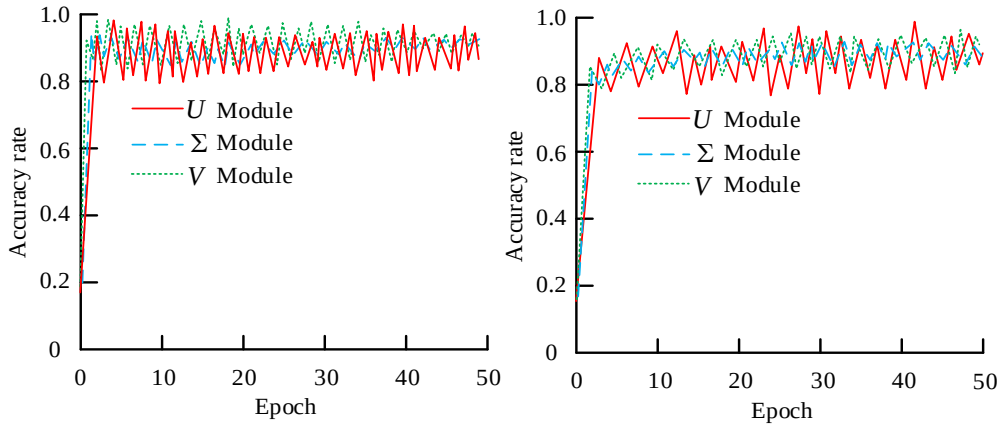

Figure 8: Comparison of recognition accuracy of different modules facing noise

In Figure 7 (a), the initial recognition accuracies of the proposed method, UMC-PNN, and OFNN were 0.374, 0.352, and 0.381, respectively. All three methods approached convergence after about 10 iterations, but there were significant fluctuations. After training, the maximum recognition accuracy of the proposed method was 0.982, which was 0.012 and 0.017 higher than UMC-PNN and OFNN, respectively. In Figure 7 (b), in the test set, the final recognition accuracy of the proposed method was basically the same, but the fluctuation amplitude increased significantly in the early stage of iteration. The final recognition accuracy of the proposed method was 0.065 and 0.073 higher than UMC-PNN and OFNN, respectively. In order to investigate the sensitivity of different modules in the MZI array to noise, random noise with a value of 0.2 was introduced into two unitary matrix modules and one diagonal matrix module. The comparison of the recognition accuracy of different modules in the face of noise is in Figure 8.

In Figure 8 (a), when noise was added to the unitary matrix U module, the model's recognition accuracy fluctuated the most because the U module had the most MZI arrays and was more affected by noise. When noise was added to the diagonal matrix Σ module, the model's fluctuation amplitude was the smallest. The average convergence values of the Σ module were 0.027 and 0.013 higher than those of the U and V modules,

respectively. In Figure 8 (b), due to the use of more MZI arrays in the UMC-PNN model, the fluctuation amplitude of its recognition accuracy increased. Similarly, when noise was added to the U module, the optimized model's recognition accuracy fluctuated the most. To further investigate the system's robustness to noise, the variation in recognition accuracy was tested under different noise levels (0.05, 0.1, 0.2, 0.3). The results are shown in Figure 9.

In Figure 9(a), when noise intensity was below 0.15, system performance showed no significant decline, with accuracy remaining above 0.90. However, performance deteriorated sharply when noise intensity exceeded 0.15. Results indicated the proposed method exhibited good tolerance to moderate noise levels. In Figure 9(b), the UMC-PNN model experienced a more rapid decline in accuracy under higher noise intensities.

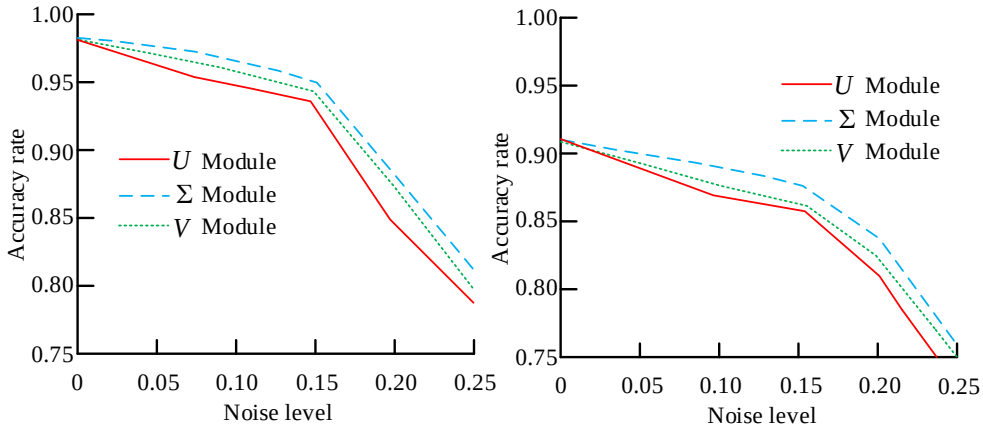

Figure 9: Robustness analysis of the model under different noise levels

Table 2: Compares the recognition performance of handwritten digits by different methods

Model	Calculation speed	Recognition accuracy	F1(%)	RMSE
Pseudo-real optimization	12ns	0.982	84.2	0.276
UMC-PNN	25ns	0.917	80.5	0.293
OFNN	19ns	0.909	78.7	0.307
CNN	24ms	0.914	82.6	0.284
SVM	15ms	0.896	79.5	0.295
KNN	7ms	0.857	76.3	0.302
DT	17ms	0.905	78.2	0.289
DBN	65ms	0.943	83.4	0.278

3.2 Experimental analysis of photon neural network optimized based on pseudo-real numbers

To investigate the recognition speed of handwritten digits using different methods, five comparative methods were added: CNN, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree (DT), and Deep Belief Network (DBN). The comparison of recognition performance of different methods for handwritten digits is in Table 2.

In Table 2, the research results were all theoretical calculations based on simulation models. The computational speed units for the three photonic neural networks were all ns, differing by six orders of magnitude from conventional convolutional neural networks and deep learning methods. Furthermore, after pseudo-real number optimization, the computational speeds of the photonic neural networks were 13 ns and 7 ns faster than UMC-PNN and OFNN respectively. The proposed method achieved the highest recognition accuracy among the evaluated approaches, alongside the optimal F1 score and RMSE. Photonic chips exhibited extremely low power consumption in static states, with primary energy expenditure arising from electro-optic conversion and phase modulators. Preliminary analysis indicated their energy efficiency potential significantly surpassed that of conventional electronic chips. To validate the method's generalization capability, experiments were conducted on the more complex Fashion-MNIST dataset. This dataset comprises 28×28 greyscale images of 10 categories of clothing items, totalling 70000 samples. The experimental results are presented in Table 3.

In Table 3, when confronted with the more complex Fashion-MNIST dataset, the performance of all algorithms declined. However, the computational speed of the proposed method decreased by merely 2 ns,

whereas the second-best OFNN algorithm saw a reduction of 6 ns. The proposed method achieved a recognition accuracy of 0.979, representing a decrease of 0.003 compared to the MNIST dataset, while the second-best UMC-PNN method saw a decline of 0.015. The proposed method maintained a relative advantage across all performance metrics, demonstrating high scalability. The proposed method exhibited lower standard deviations than the reference method across all performance metrics, with statistically significant differences observed between the results of each method (p<0.05). The LF and recognition accuracy of the optimized photon neural network using pseudo-real number theory are shown in Figure 10.

Table 3: Experimental results on the Fashion-MNIST dataset using different methods

Model	Calculation speed	Recognition accuracy	F1(%)	RMSE
Pseudo-real optimization	14±0.02ns	0.979±0.013	83.7±0.01	0.281±0.002
UMC-PNN	37±0.17ns	0.902±0.037	79.1±0.19	0.312±0.015
OFNN	25±0.35ns	0.895±0.029	75.4±0.32	0.325±0.027
CNN	43±1.24ms	0.897±0.041	78.2±0.58	0.304±0.042
t	12.716	11.563	14.092	11.864
p	0.001	0.005	0.000	0.004

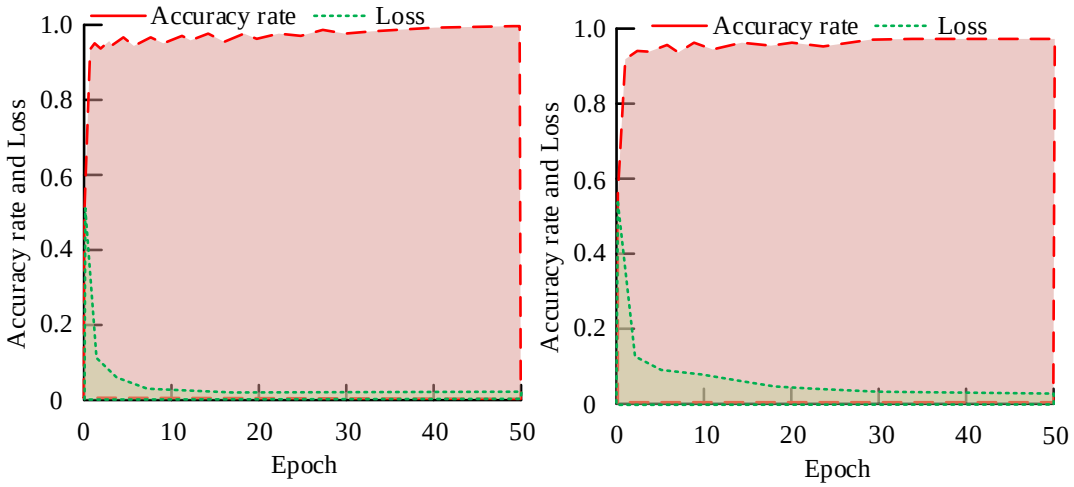

Figure 10: Changes in LF and recognition accuracy of photon neural network after optimization

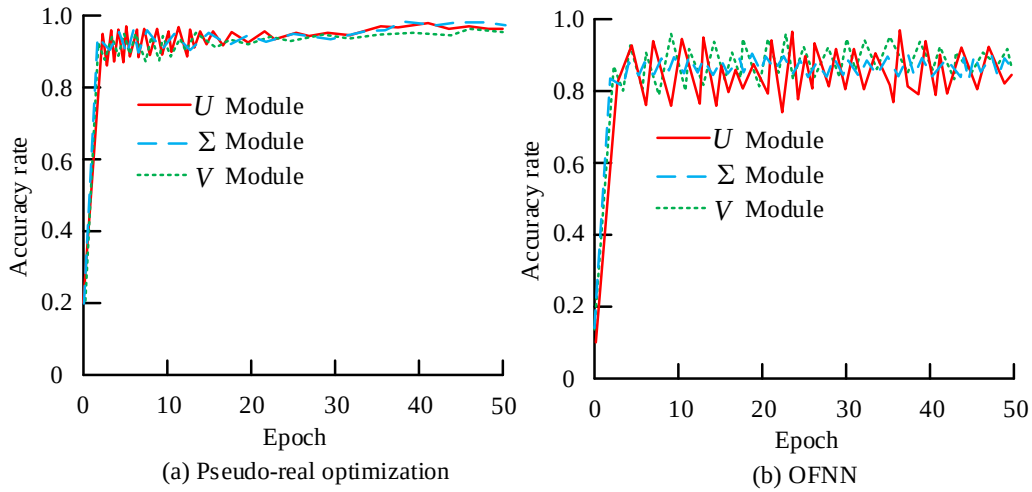

Figure 11: Comparison of recognition accuracy of different modules facing noise

In Figure 10 (a), after optimizing using pseudo-real number theory, the performance of the photon neural network was improved, with a maximum recognition accuracy and minimum LF of 0.987 and 0.05, respectively. In addition, the fluctuation amplitude of the photon neural network in the early stage of iteration was greatly reduced,

improving the stability of the network. In Figure 10 (b), in the test set, the maximum recognition accuracy of the improved network was 0.985. Although it decreased compared to the training set, it was still higher than before, with a minimum LF of 0.07. Random noise of 0.2 was introduced into the optimized network, and the comparison of recognition accuracy of different modules in the face of noise is in Figure 11.

In Figure 11 (a), the fluctuation amplitude of recognition accuracy of different modules of the optimized photon neural network was significantly reduced when facing random noise. All three modules could gradually converge in the later stage of iteration, and the maximum convergence value of the Σ module was 0.011 and 0.012 higher than that of the U module and V module, respectively. In Figure 11 (b), the OFNN model was greatly affected by random noise interference, and the model remained in a fluctuating state throughout the entire iteration cycle, unable to achieve convergence values. The ablation experiment analysis of the improved photon neural network is in Table 4 for comparison of experimental results.

Table 4: Results of improved photon neural network ablation experiments

Ablation module	Calculation speed	Recognition accuracy	F1(%)	RMSE
Singular value decomposition	18ns	0.954	81.5	0.295
Rectangular decomposition	15ns	0.936	83.2	0.284
Directional coupler	27ns	0.973	83.0	0.287
The theory of pseudo-real numbers	20ns	0.892	78.4	0.317
GridNet network architecture	14ns	0.903	79.5	0.309

Table 5: Comparison of the proposed method with existing advanced optical neural networks

Method	Accuracy	Loss	Speed	Noise Robustness	Scalability
Proposed	0.982	0.08	12ns	High	Medium
PCNN	0.917	0.10	25ns	Medium	High
UMFPNN	0.909	0.17	19ns	Low	Medium
Ashtiani et al. [12]	0.938	0.12	0.57ns	Medium	Low
Pai et al. [10]	0.940	0.11	/	High	Medium

In Table 4, the directional coupler module had the greatest impact on the computational speed of the improved photon neural network. After removing this module, the algorithm's calculation speed increased by 15ns and the recognition accuracy decreased by 0.009. The pseudo-real theory exerted the most significant influence on the accuracy of algorithm recognition. After removing this module, the recognition accuracy of the algorithm decreased by 0.090 and the RMSE increased by 0.041. The pseudo-real number theory preserved the phase information of complex optical fields through mathematical projection, mapping it onto real-valued quantities. This enabled the computation of arbitrary real matrices within a single MZI array. Conventional methods required multiple MZI arrays to process complex matrix operations, as standard detectors could not capture phase information, leading to information loss. Pseudo-real number theory efficiently integrated the amplitude and phase information of complex optical fields into real-number computations through directional couplers and approximate calculations, thereby preventing phase information loss. This highly efficient phase-preserving computation method enabled photonic neural networks to utilize the full information content of the optical field more comprehensively, consequently enhancing the recognition accuracy of the algorithms. The proposed method was compared with existing advanced optical neural networks as shown in Table 5.

In Table 5, the advanced comparison methods introduced in the experiment included the Photon Convolutional Neural Network (PCNN) and the Unitary Matrix Factorisation Photonic Neural Network (UMFPNN). The proposed method achieved the highest recognition accuracy and lowest loss value, with an accuracy 0.042 percentage points higher than the second-best method proposed by Pai et al. and a loss value 0.02 lower than PCNN. The proposed method exhibited high noise robustness and moderate scalability. The primary challenge faced by the proposed method in large-scale MZI arrays was noise accumulation. As illustrated in Figure 8, the unitary matrix module

exhibited the highest sensitivity to noise due to its largest number of MZI units. Compared to the dual adaptive training method proposed by Zheng et al. in [11], this research employed pseudo-real number theory to project phase information onto the real number domain, mitigating the accumulation of systematic errors to some extent. Nevertheless, when scaling the array further, strategies such as error prediction networks or quantization-aware training became necessary to ensure system stability. Regarding scalability, the GridNet architecture enabled efficient mapping of unitary matrices through rectangular decomposition, offering a viable pathway for realizing large-scale photonic neural networks

4 Discussion

To enhance the computational performance of photonic neural networks, this study proposes an improved photonic neural network based on pseudo-real number theory and the GridNet architecture. Compared to the UMC-PNN in Reference [8], the proposed method achieved a 6.5% improvement in recognition accuracy. The UMC-PNN relied on a single tunable phase shifter to perform vector multiplication on non-negative or real-valued matrices, thereby limiting its expressive capability for complex transformations. Conversely, the improved approach innovatively introduced projection computation units within a pseudo-real number framework, enabling a single MZI network to compute arbitrary real matrices. This substantially enhanced the network's feature extraction capabilities and flexibility. Compared to the end-to-end photonic network by Ashtiani et al. in [12], the improved method achieved higher recognition accuracy (0.982 vs 0.938) while maintaining high-speed computation through the MZI array, which provided precise and reconfigurable matrix weights. As illustrated in Figures 8 and 0, the improved method exhibited varying sensitivity to noise across different modules. The unitary matrix module, owing to its numerous MZIs, demonstrated the highest noise sensitivity, consistent with observations by Banerjee et al. [19]. Following optimization via pseudo-real theory, the system's noise robustness was significantly enhanced, indicating that this theory not only improved training efficacy but also fortified system stability. The study employed directional couplers as substitutes for beam splitters, a feasible approach from an integrated photonics manufacturing perspective. On silicon photonics chips, directional couplers served as core components for optical power splitting and combining, with manufacturing processes fully compatible with CMOS techniques. Compared to free-space beam splitters requiring separate fabrication and precision mechanical alignment, monolithic integrated directional couplers inherently offered advantages in size, stability, and scalability. Consequently, the research did not introduce an unconventional hardware modification but employed a more typical and mature component within integrated photonics to realize MZI functionality. Simultaneously, ablation experiments demonstrated the feasibility of incorporating directional couplers. Whilst simulation results were encouraging, practical hardware implementation still faces numerous challenges. These include manufacturing tolerances for directional couplers, the thermal sensitivity of the MZI, and crosstalk issues during large-scale integration, all requiring further resolution. Future work will focus on optimizing these practical issues to advance the methodology towards practical applications.

5 Conclusion

Aiming at the problem of traditional electronic chip computing performance approaching its limit, a photon neural network algorithm based on pseudo-real number theory was proposed for optical components. The experiment outcomes indicated that the LF of the photon neural network tended to converge after 20 iterations, with a minimum LF value of 0.08, which was 0.02 and 0.09 lower than UMC-PNN and OFNN, respectively. The photon neural network approached convergence after about 10 iterations, but the fluctuation amplitude was large. Its maximum recognition accuracy was 0.982, which was 0.012 and 0.017 higher than UMC-PNN and OFNN, respectively. The U module had the highest number of MZI arrays and was more affected by noise. The average convergence values of the Σ module were 0.027 and 0.013 higher than those of the U and V modules, respectively. The computing speed unit of photon neural networks was ns, which was 6 orders of magnitude lower than ordinary CNNs and deep learning methods. The maximum recognition accuracy and minimum LF of the optimized photon neural network were 0.987 and 0.05, respectively. When different modules of the optimized photon neural network faced random noise, the fluctuation amplitude of recognition accuracy was significantly reduced, and all three modules could gradually converge in the later stage of iteration. The directional coupler module had the greatest impact on improving the computational speed of photon neural networks. After removing this module, the algorithm's computation time increased by 15ns. The pseudo-real number theory had the greatest impact on recognition accuracy. After removing this module, the recognition accuracy decreased by 0.090. There are still some issues with this study, such as the proposed method mainly focusing on the application of photon neural networks at the linear level. In the future, nonlinear calculations of photon neural networks can be achieved based on the nonlinear characteristics of optics.

References

[1] Du Y, Su K, Yuan X, Li T, Liu K, Man H, Zou X. Implementation of optical neural network based on Mach–Zehnder interferometer array. IET

Optoelectronics. 2023, 17(1):1-11. https://doi.org/10.1049/ote2.12086

[2] Norris B R, Martinod M A, Tuthill P, Gross S, Cvetojevic N, Jovanovic N, et al. Machine-learning approach for optimal self-calibration and fringe tracking in photonic nulling interferometry. Journal of Astronomical Telescopes, Instruments, and Systems. 2023, 9(4):048005. https://doi.org/10.1117/1.JATIS.9.4.048005

[3] Yang F, Tian Q, N X, Pan Y, Wang F, Lá zaro J A, et al. Modulation format recognition scheme based on discriminant network in coherent optical communication system. Electronics. 2024, 13(19):3833-3845. https://doi.org/10.3390/electronics13193833

[4] Zhang L, Tian X, Jiang Y, Li X, Li Z, Li D, Zhang S. Adversarial network for multi-input image restoration under strong turbulence. Optics Express. 2023, 31(25):41518-41532. https://doi.org/10.1364/oe.503611

[5] Khonina S N, Kazanskiy N L, Skidanov R V, & Butt M A. Exploring types of photonic neural networks for imaging and computing—a review. Nanomaterials, 2024, 14(8): 697-709. https://doi.org/10.3390/nano14080697

[6] Huang C, Sorger V J, Miscuglio M, Al-Qadasi M, Mukherjee A, Lampe L, et al. Prospects and applications of photonic neural networks. Advances in Physics: X. 2022, 7(1):1981155-1981172. https://doi.org/10.1080/23746149.2021.1981155

[7] Tian Y, Zhao Y, Liu S, Li Q, Wang W, Feng J, Guo J. Scalable and compact photonic neural chip with low learning-capability-loss. Nanophotonics, 2022, 11(2):329-344. https://doi.org/10.1515/nanoph-2021-0521

[8] Huang Y, Yue H, Ma W, Zhang Y, Ao Y, Wang W, Chu T. Easily scalable photonic tensor core based on tunable units with single internal phase shifters. Laser & Photon Reviews. 2023, 17(10):2300001-2300015. https://doi.org/10.1002/lpor.202300001

[9] Ouyang J, Liu S, Yang Z, Wang W, Feng X, Li Y, Huang Y. 16-channel photonic solver for optimization problems on a silicon chip. Chip. 2025, 4(1):100117-100134. https://doi.org/10.1016/j.chip.2024.100117

[10] Pai S, Sun Z, Hughes T W, Park T, Bartlett B, Williamson I A, et al. Experimentally realized in situ backpropagation for deep learning in photonic neural networks. Science. 2023, 380(6643):398-404. https://doi.org/10.1126/science.ade8450

[11] Zheng Z, Duan Z, Chen H, Yang R, Gao S, Zhang H, Lin X. Dual adaptive training of photonic neural networks. Nature Machine Intelligence. 2023, 5(10):1119-1129. https://doi.org/10.1038/s42256-023-00723-4

[12] Ashtiani F, Geers A J, Aflatouni F. An on-chip photonic deep neural network for image classification. Nature. 2022, 606(7914):501-506. https://doi.org/10.1038/s41586-022-04714-0

[13] Kirtas M, Oikonomou A, Passalis N, Mourgias-Alexandris G, Moralis-Pegios M, Pleros N, Tefas A. Quantization-aware training for low precision photonic neural networks. Neural Networks. 2022, 155(4):561-573. https://doi.org/10.1016/j.neunet.2022.09.015

[14] Liao K, Dai T, Yan Q, Hu X, Gong Q. Integrated photonic neural networks: Opportunities and hallenges. ACS Photon. 2023, 10(7):2001-2010. https://doi.org/10.1021/acsphotonics.2c01516

[15] Wang T, Ma S Y, Wright L G, Onodera T, Richard B C, & McMahon P L. An optical neural network using less than 1 photon per multiplication. Nature Communications, 2022, 13(1), 123-135. https://doi.org/10.1038/s41467-021-27774-8

[16] Tsakyridis A, Moralis-Pegios M, Giamougiannis G, Kirtas M, Passalis N, Tefas A, Pleros N. Photonic neural networks and optics-informed deep learning fundamentals. APL Photon. 2024, 9(1): 011102. https://doi.org/10.1063/5.0169810

[17] Freire P, Manuylovich E, Prilepsky J E, Turitsyn S K. Artificial neural networks for photonic applications—from algorithms to implementation: tutorial. Advances in Optics and Photon. 2023, 15(3):739-834. https://doi.org/10.1364/AOP.484119

[18] Totovic A, Giamougiannis G, Tsakyridis A, Lazovsky D, Pleros N. Programmable photonic neural networks combining WDM with coherent linear optics. Scientific Reports. 2022, 12(1):5605-5617. https://doi.org/10.1038/s41598-022-09370-y

[19] Banerjee S, Nikdast M, Chakrabarty K. Characterizing coherent integrated photonic neural networks under imperfections. Journal of Lightwave Technology. 2022, 41(5):1464-1479. https://doi.org/10.1109/JLT.2022.3193658

[20] Ewaniuk J, Carolan J, Shastri B J, Rotenberg N. Imperfect quantum photonic neural networks. Advanced Quantum Technologies. 2023, 6(3):2200125-2200139. https://doi.org/10.1002/qute.202200125

[21] Paolini E, De Marinis L, Cococcioni M, Valcarenghi L, Maggiani L, Andriolli N. Photonic-aware neural networks. Neural Computing and Applications. 2022, 34(18):15589-15601. https://doi.org/10.1007/s00521-022-07243-z