photonics

MDPI

Article

Neural Network-Driven Transmission Characteristics Modeling and Manufacturing Error Detection for Photonic Lanterns

Zhuruixiang Sun [1], Xiang Li [1], Yao Lu [1,2,*], Tong Liu [3], Zilun Chen [4] and Zongfu Jiang [4]

[1] Space Engineering University, Beijing 101416, China; sunzhuruixiang@163.com (Z.S.); li_xiang_7756@163.com (X.L.)
[2] National Key Laboratory of Space Target Awareness, Beijing 101416, China
[3] National Department of Aerospace Science and Technology, Space Engineering University, Beijing 101416, China; liutong719@163.com
[4] National College of Advanced Interdisciplinary Studies, National University of Defense Technology, Changsha 410073, China; zilun2003@163.com (Z.C.); jiangzongfu7@163.com (Z.J.)
* Correspondence: lybenben1@163.com

Abstract

Traditional numerical simulation methods struggle to accurately characterize the transmission characteristics of finished photonic lanterns that contain manufacturing errors. This paper proposes a method for characterizing photonic lantern devices using neural networks. In an ideal 1×6 photonic lantern, the Mean Squared Error (MSE) for predicting intensity from the multimode to the single-mode end was reduced to 10^{-5}, and the neural network model can identify manufacturing error patterns, providing a new approach to addressing the precise characterization and product screening of novel irregular waveguides such as photonic lanterns.

Keywords: photonic lantern; neural network; transmission characteristics

1. Introduction

Photonics focuses on the macroscopic propagation and control of light in complex media and structures, with one of its core objectives being the realization of highly efficient, low-crosstalk multimode light field processing. Within this framework, photonic lanterns are fabricated by thermally isolating several single-mode cores within a single multimode core. They provide low-loss interfaces between single-mode and multimode systems [1]. To fully leverage the performance of photonic lanterns and guide their optimized design, precise experimental characterization and theoretical modeling of their transmission characteristics are crucial. However, manufacturing errors in photonic lanterns are difficult to measure precisely through computational simulations. Traditional physical methods—such as the Beam Propagation Method (BPM), Finite Difference Method (FDM), Eigenmode Expansion Method (EEM), or the solution of the photonic lantern transfer matrix—suffer from low simulation accuracy and are time-consuming when accounting for device manufacturing errors. In 2012, Nicolas K. Fontaine et al. at Bell Labs investigated the effects of different geometric arrangements of single-mode fiber bundles on the performance of photonic lanterns. They introduced the transfer matrix to describe the light transmission characteristics of photonic lanterns [2]. In 2014, Y. He and X. Chen pointed out that in numerical simulations, due to limited server capacity, it is impossible to include all modes in the calculations [3]. In 2019, M. Diab, S. Minardi, and colleagues demonstrated experimentally that symmetrically distributed output waveguides result in a transmission matrix with a

Check for updates

Received: 16 April 2026
Revised: 11 May 2026
Accepted: 13 May 2026
Published: 16 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

very poor condition number, allowing only the calculation of photonic lanterns under ideal conditions [4]. In 2021, S. G. Leon-Saval and his team used a nonlinear neural network to calculate the vector transmission matrix of photonic lanterns [5]. In 2025, Wu et al. systematically modeled and optimized the transmission characteristics of photonic lanterns in the O-band to a 2 μm wavelength range by establishing a numerical model based on adiabatic conditions, with the aim of achieving efficient mode conversion over a broad bandwidth [6]. In 2026, A. K. Taras et al. demonstrated that fabrication errors cause actual devices to lose the port symmetry present in ideal models, and that the measured relative mode dispersion differs significantly from ideal predictions, strongly confirming the limitations of relying solely on ideal model-based design [7]. To achieve high-precision calibration of photonic lanterns, this paper designs a neural network model of the photonic lantern's transmission characteristics, establishing a nonlinear mapping between the light-field data (intensity and phase) at the multimode and single-mode ends of the lantern. This approach achieves high-precision Mean Squared Error (MSE) for both ideal and manufacturing-error photonic lanterns, providing a new calibration path for optical devices such as photonic lanterns.

2. Neural Network Model of Photonic Lantern Transmission Characteristics

2.1. Photonic Lantern

The photonic lantern is an all-fiber mode converter that enables photonics-based manipulation of light. It is a specialized waveguide structure capable of achieving low-loss, adiabatic mode evolution between a multimode waveguide and an array of single-mode fibers. The function of the photonic lantern is to linearly decompose the multimode light field information at the focal plane into a set of stable, orthogonal basis modes, which manifest at the output as the intensity and phase information detected by the single-mode fiber array, as shown in Figure 1.

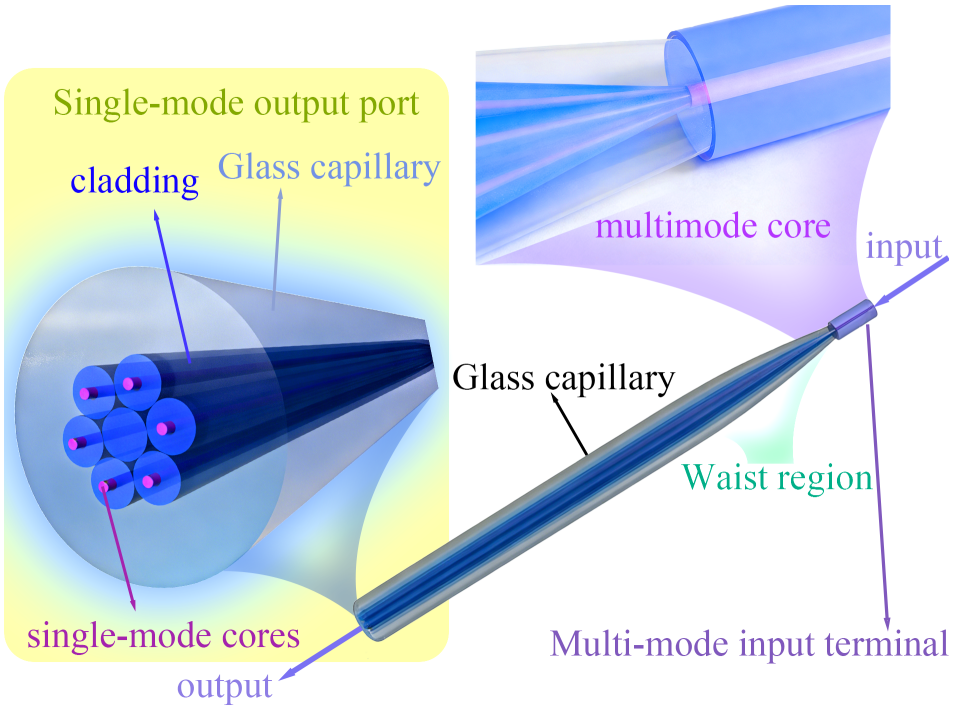


Figure 1. Structure of the photonic lantern.

As a regular optical waveguide (with a uniform longitudinal refractive index distribution), an optical fiber possesses specific boundary conditions. Under steady-state transmission conditions, a system of equations yields a finite number of discrete eigen

solutions. The specific transverse light-field distribution corresponding to each eigen solution is defined as a laser transverse mode [8–12], hereafter uniformly abbreviated as "mode". The main characteristics of modes are: their field distribution does not change with transmission distance; different modes within the same optical waveguide are orthogonal and can be ordered; and the total light field distribution in the fiber is a linear superposition of these modes.

Considering that common step-index fibers are circular homogeneous waveguides, the transverse distribution of the mode field is isotropic, and the refractive index varies little within each layer of the fiber (i.e., the weak-guide approximation holds), Gloge et al. proposed the scalar approximation [13], which simplifies the complex process of solving the Maxwell equations into solving the eigenvalue equations for a weak-guide step-index fiber:

$$\frac{UJ_{m+1}(U)}{J_m(U)} = \frac{WK_{m+1}(W)}{K_m(W)} \tag{1}$$

where J_m, K_m are the first-kind Bessel function and the second-kind (complex) Bessel function, respectively; U is the normalized transverse propagation constant in the core region: $U^2 = (k_0^2 n_1^2 - \beta^2)a^2$; W is the normalized transverse attenuation function in the cladding region: $W^2 = (\beta^2 - k_0^2 n_2^2)a^2$. Here, k_0 is the wave number in a vacuum, n_1 is the refractive index of the core, n_2 is the refractive index of the cladding, a is the core radius, and β is the propagation constant [14]. Defining V as the normalized frequency of the optical fiber [15]:

$$V^2 = k_0^2 a^2 \left(n_1^2 - n_2^2\right) = (k_0 a \mathrm{NA})^2 \tag{2}$$

it can be seen that the value of V is entirely determined by the optical fiber parameters.

The eigenvalue Equation (1) is a transcendental equation that requires numerical methods for solution. Once the geometric parameters of the fiber and the wavelength are determined, a numerical approximation of the solution can be obtained, revealing that the resulting propagation constant β takes discrete values between $k_0 n_2$ and $k_0 n_1$. For each integer m, there are n solutions, corresponding to different values of β_{mn}, U_{mn}, and W_{mn}. A spatial field distribution determined by each set of values is called a linear polarization mode [13], denoted by LP_{mn}. Assuming the polarization direction of the incident wave is parallel to the x axis, the transverse electric field of the LP_{mn} mode is represented by the scalar E_x, which is written in polar coordinates as:

$$E_x = Ae^{i\beta_{mn}z}\phi(m\theta)\begin{cases} \frac{J_m\left(\frac{U_{mn}}{a}r\right)}{J_m(U_{mn})}, 0 \le r \le a \\ \frac{K_m\left(\frac{W_{mn}}{a}r\right)}{K_m(W_{mn})}, r > a \end{cases} \tag{3}$$

where $\phi(m\theta)$ reflects the angular distribution of the light field. When $m > 0$, $\phi(m\theta)$ can take the form of $\cos(m\theta)$ or $\sin(m\theta)$, representing the two degenerate states of the LP_{mn} mode: the "even" mode and the "odd" mode, distinguished by the subscripts LP_{mne} and LP_{mno}. Figure 2 shows schematic diagrams of the near-field intensity and phase distributions for six typical LP modes in a step-index fiber. A 1×6 photonic lantern has exactly six modes, and the optical field information—including both intensity and phase—is input at the multimode end of the photonic lantern.

Numerical simulations of the evolution of a 1×6 photonic lantern have revealed a one-to-one correspondence between the input and output light fields of the photonic lantern. Drawing on the concept of the transmission matrix [16,17], this relationship between the input and output light fields of the photonic lantern can be concisely expressed as:

$$u = M \cdot v \tag{4}$$

where M is the transmission matrix of the photonic lantern, v is a column vector describing the optical field information at the single-mode end of the photonic lantern, $v^T = \left[v_1 e^{j\theta_1}, v_2 e^{j\theta_2}, v_3 e^{j\theta_3}\right]$ (where v_i and θ_i denote the amplitude and phase, respectively, of the fundamental mode optical field input from the i single-mode fiber), u is the column vector describing the optical field information at the multimode end of the photonic lantern, $u^T = \left[u_1 e^{j\phi_1}, u_2 e^{j\phi_2}, u_3 e^{j\phi_3}\right]$ (where u_i and ϕ_i denote the amplitude and phase of the i eigenmode in the output superposition field, respectively).

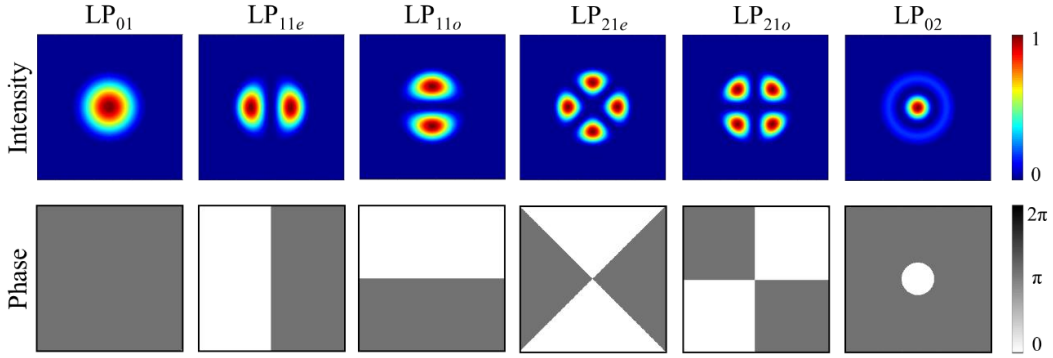


Figure 2. Schematic of near-field intensity and phase distributions for typical LP modes.

The physical structure of the photonic lantern exhibits inherent regularities, and its transmission characteristics can be expressed using nonlinear functions. Deep learning possesses strong capabilities for handling nonlinear problems and can approximate highly nonlinear functions with fewer parameters.

2.2. Neural Networks

A neural network is a computational model that mimics the connections between neurons in the human brain. It approximates complex functions by constructing multiple layers of nonlinear transformations. Its core lies in introducing nonlinear activation functions (LeakyReLU is used in this paper) to transform linear combinations into nonlinear mappings. The General Approximation Theorem states that, provided the hidden layer is sufficiently wide, this structure can approximate any continuous function on a compact domain with arbitrary precision, laying the theoretical foundation for neural networks to handle nonlinear regression problems [18].

From a mathematical perspective, the propagation process of a photonic lantern can be represented as a nonlinear operator T, which maps the input light field I_{in} into the output light field I_{out}:

$$I_{\text{out}} = T(I_{\text{in}}, \lambda) \tag{5}$$

where λ represents parameters such as wavelength. Neural networks can leverage their nonlinear mapping capabilities to learn an approximate representation of this operator. Through multiple layers of nonlinear transformations, any continuous function can be approximated, providing a mathematical guarantee of accurate modeling of the photonic lantern's transmission characteristics.

From a physical perspective, the photonic lantern is a multi-input, multi-output optical device whose transmission characteristics follow the laws of electromagnetic wave propagation described by Maxwell's equations. Although the underlying physical mechanisms are complex, this process is essentially a deterministic mapping: given specific input light-field parameters, a corresponding output light-field distribution is generated. Therefore, it can be abstracted as a deterministic nonlinear system, and neural networks are particularly adept at learning such complex input-output mapping relationships.

Consequently, if there exists an intrinsic, continuous functional relationship between the transmission characteristics of the photonic lantern and its input parameters, and if we can obtain a sufficiently large dataset that is well-distributed across the input parameter space, features high data quality, and has accurate labels, while simultaneously selecting an appropriate neural network architecture and conducting effective training (including measures to prevent overfitting), then the neural network has the potential to learn and establish a high-precision regression model for the transmission characteristics of the photonic lantern.

In regression tasks, the learning process is driven by minimizing a loss function (such as the MSE, used in this paper). Given the training data (x_i, y_i), the training objective is to minimize the loss function $L(\theta)$, as shown in Equation (6):

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - f(x_i, \theta))^2 \tag{6}$$

The core parameters are optimized using gradient descent, and the update rule is given by Equation (7):

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_\theta L(\theta^{(t)}) \tag{7}$$

where η is the learning rate. Gradient calculation relies on the backpropagation algorithm, which uses the chain rule to propagate the output error backward through each layer to the respective parameters, efficiently computing the gradient [19].

To prevent overfitting and improve generalization ability, regularization techniques are commonly employed. For instance, this paper uses Dropout regularization, which randomly deactivates a portion of neurons during training, to reduce cooperative adaptation among neurons. This enables the network to learn feature representations with greater redundancy and generalization, thereby enhancing the model's robustness. Additionally, this paper employs techniques such as learning rate decay to effectively regulate the training process. When training stalls, the search step size is adaptively refined, effectively balancing global optimization efficiency with local convergence accuracy to suppress overfitting and ensure the model possesses good generalizability.

3. Method

3.1. Light Field Dataset

Based on numerical simulations of traditional light-beam propagation algorithms (RSoft BeamPROP, BPM algorithm), this paper designs and implements a 1×6 photonic lantern for experiments, collecting large amounts of input light-field data (optical power and phase) and output light-field data (optical intensity and phase). This data is used to construct a light field dataset based on the physical principles of photonic lantern mode evolution.

The structure of the 1×6 photonic lantern described in this paper employs an adiabatic taper-coupling design between a multimode fiber (MMF) and a single-mode fiber array (SMF array). The core diameter is 10.0 μm, the cladding diameter is 125.0 μm, and the taper ratio is 13.25. The operating wavelength is 1.064 μm, the cladding refractive index is 1.435, and the refractive index difference is 0.01, meaning the core refractive index is 1.445. The numerical simulations were performed using RSoft BeamPROP with the following parameters: grid resolution of 0.8 μm in the X and Y directions (cross-section of the photonic lantern) and 2 μm in the Z direction (propagation direction); the simulation domain is a rectangular prism with a base side of 300 μm and a height of 26,800 μm, centered at the 1×6 multimode port of the photonic lantern; the propagation step size is set to 2 μm, consistent with the Z-axis resolution, and verification at 13,400 μm confirmed the accuracy

of the simulation. As shown in Figure 3, the input light field at the multimode end of the photonic lantern consists of a linear superposition of six orthogonal LP_{mn} modes. Each data set consists of randomly generated optical powers ([0–1]) and phase differences (in degrees, [0–360]) for the six fundamental modes. A total of N sets of input light field data are generated and saved as a .txt file. Through simulation, following a fixed order, the 6 pairs of average light intensities and phase differences (in degrees, [0–360]) detected at the 6 single-mode ports are recorded, yielding N sets of output light-field data, which are saved as text files. Each set of input light field data corresponds to a set of output light field data, forming a data pair, yielding a total of N pairs. The input and output light field data together form the input light field dataset. To optimize the neural network, the light-field data is normalized during training.

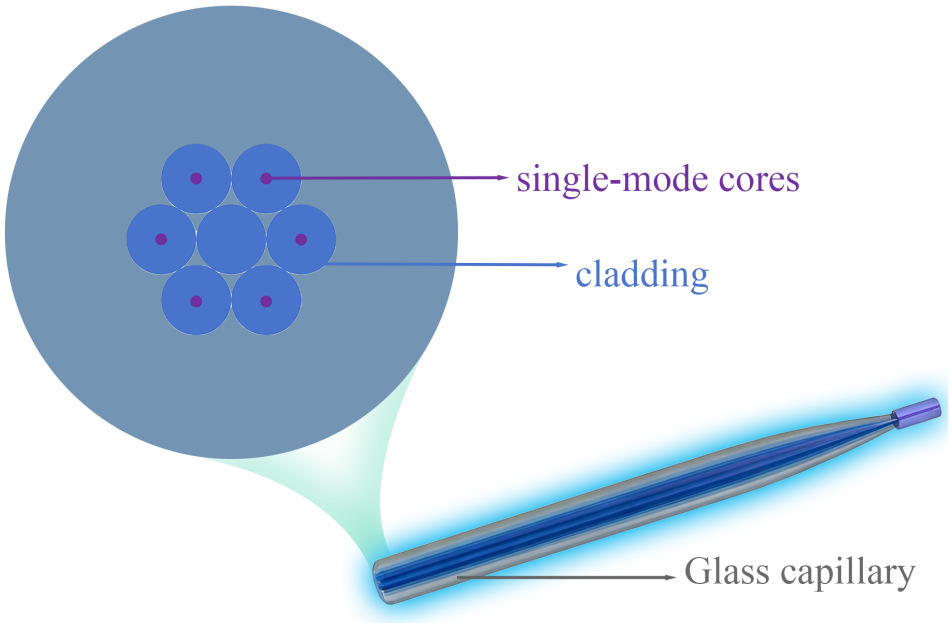


Figure 3. Cross-sectional view of the 1 × 6 photonic lantern.

As shown in Table 1, input and output data with the same numerical suffix correspond one to one. For example, Figure 4 shows the intensity and phase images for the input light field at the multimode end and the output at the single-mode end of an ideal 1 × 6 photonic lantern.

Table 1. Examples of Light Field Datasets.

Filename	Context
input0001	0.14 191 0.18 276 0.40 221 0.23 336 0.03 282 0.02 272
output0001	0.009012 6.1054 0.050812 186.04 0.02652 294.62 0.035892 245.89 0.076959 205.17 0.015572 241.77

This paper presents experiments using five model types: an ideal 1 × 6 photonic lantern, a 1 × 6 photonic lantern with misaligned fusion joints (two misalignment scenarios), and a 1 × 6 photonic lantern with non-uniform taper (non-uniform fiber volume and non-uniform arrangement).

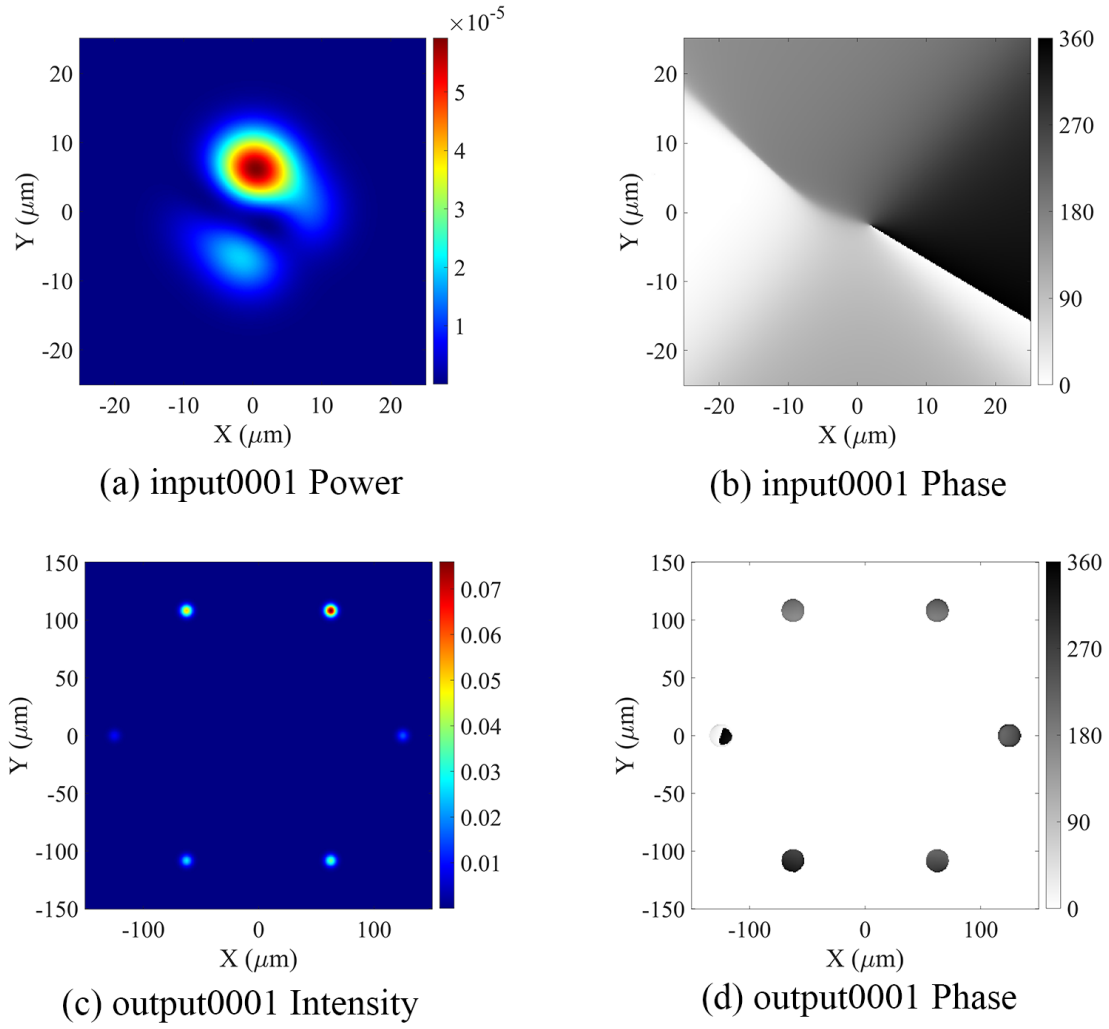


Figure 4. Intensity and phase images of the input light field at the multimode end and the output at the single-mode end of an ideal 1 × 6 photonic lantern.

3.1.1. Ideal 1 × 6 Photonic Lantern

An ideal photonic lantern enables low-loss, low-mode-dependent coupling between multiple single-mode fibers (SMFs) and a single multimode fiber (MMF) [1].

In an ideal photonic lantern, since the six input fundamental modes and the six single-mode fibers exhibit both vertical and horizontal symmetry, the data for the ideal 1 × 6 photonic lantern was expanded using vertical flipping and horizontal mirroring methods. As shown in Figure 2 above, based on the phase-symmetry relationships among the six fundamental modes, the light field dataset is expanded to 4 times the original number of light field data pairs, yielding $4N$ sets of light field data pairs. Specifically, LP_{01}, LP_{11e}, LP_{11o}, LP_{21e}, LP_{21o}, and LP_{02} correspond sequentially to the input light field data at the multimode end, while the single-mode fibers numbered 1 through 6 (from left to right and top to bottom) correspond to the output light field data. The phase distributions of LP_{01} and LP_{02} are fully symmetric and do not affect the phase evolution, whereas LP_{11e}, LP_{11o}, LP_{21e}, and LP_{21o} exhibit phase symmetry. When the phase of LP_{11o} and LP_{21o} is increased by 180°, the phase at single-mode outputs 3 and 4 increases by 180°; when the phase of LP_{11e} and LP_{21o} is increased by 180°, the phase at outputs 1, 2, and 3 increases by 180°; when the phase of LP_{11e}, LP_{11o}, and LP_{21o} is increased by 180°, the phase at outputs 1, 2, 3, 4, and 6 increases by 180°. Based on this principle, each original input-output pair is quadrupled: the first pair is an exact copy; the second pair increases the phase of the 3rd and 5th input modes by 180° and the 3rd and 4th output ports by 180°; the third pair increases the phase

of the 2nd, 3rd, and 5th input modes by 180° and the 1st, 2nd, 3rd, 4th, and 6th output ports by 180°. All phase values are kept within 0°–360° via modulo operation.

3.1.2. Lateral-Offset Splicing

As compared in Figures 5 and 6 below, lateral-offset splicing is a common and significant process defect in the manufacture of photonic lanterns, particularly when connecting single-mode fiber bundles to multimode fiber. This manufacturing error directly disrupts the ideal alignment of the optical path. It destroys the central symmetry of the light field, leading to unexpected coupling as the fundamental mode evolves into higher-order modes, thereby increasing insertion loss and being one of the primary sources of optical loss. Optical loss increases nonlinearly with increasing offset distance [20].

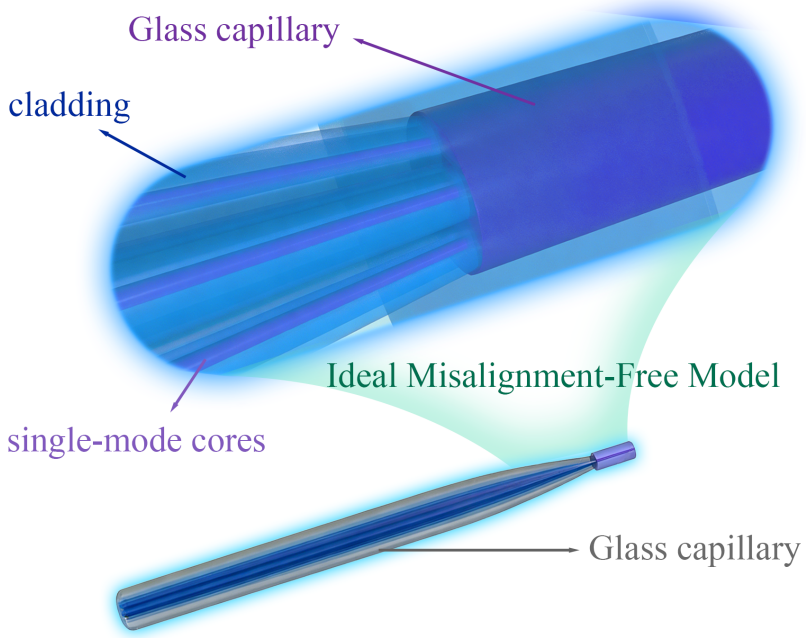


Figure 5. Structure of ideal photonic lantern.

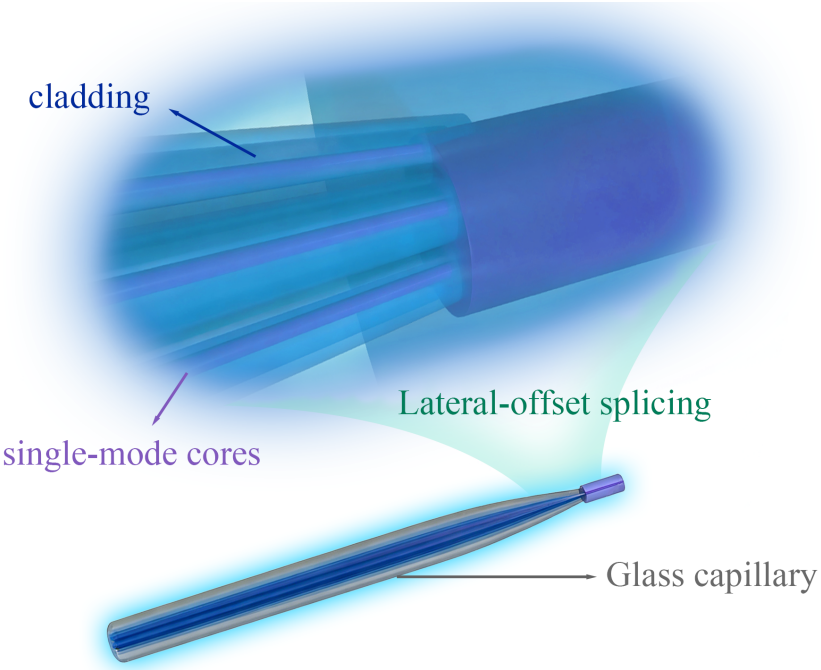


Figure 6. Structure of lateral-offset splicing.

In this paper, the end face of the waist region of the single-mode fiber was shifted relative to the matched multimode fiber during splicing, i.e., the entire structure was offset from the ideal optical axis, to simulate misaligned splicing. Two sets of optical field datasets (Unidirectional offset and Bidirectional offset) were obtained through simulation, with parameter configurations shown in Table 2.

Table 2. Parameters of the Offset Fusion Photonic Lantern Model.

Type	X Offset	Y Offset
Unidirectional offset	4 μm	0 μm
Bidirectional offset	−4 μm	4 μm

Since misaligned welding disrupts symmetry, the light field dataset here does not use symmetrically augmented data; instead, all training is performed using the original data. The following Figures 7–9 show the example of lateral-offset splicing light field data.

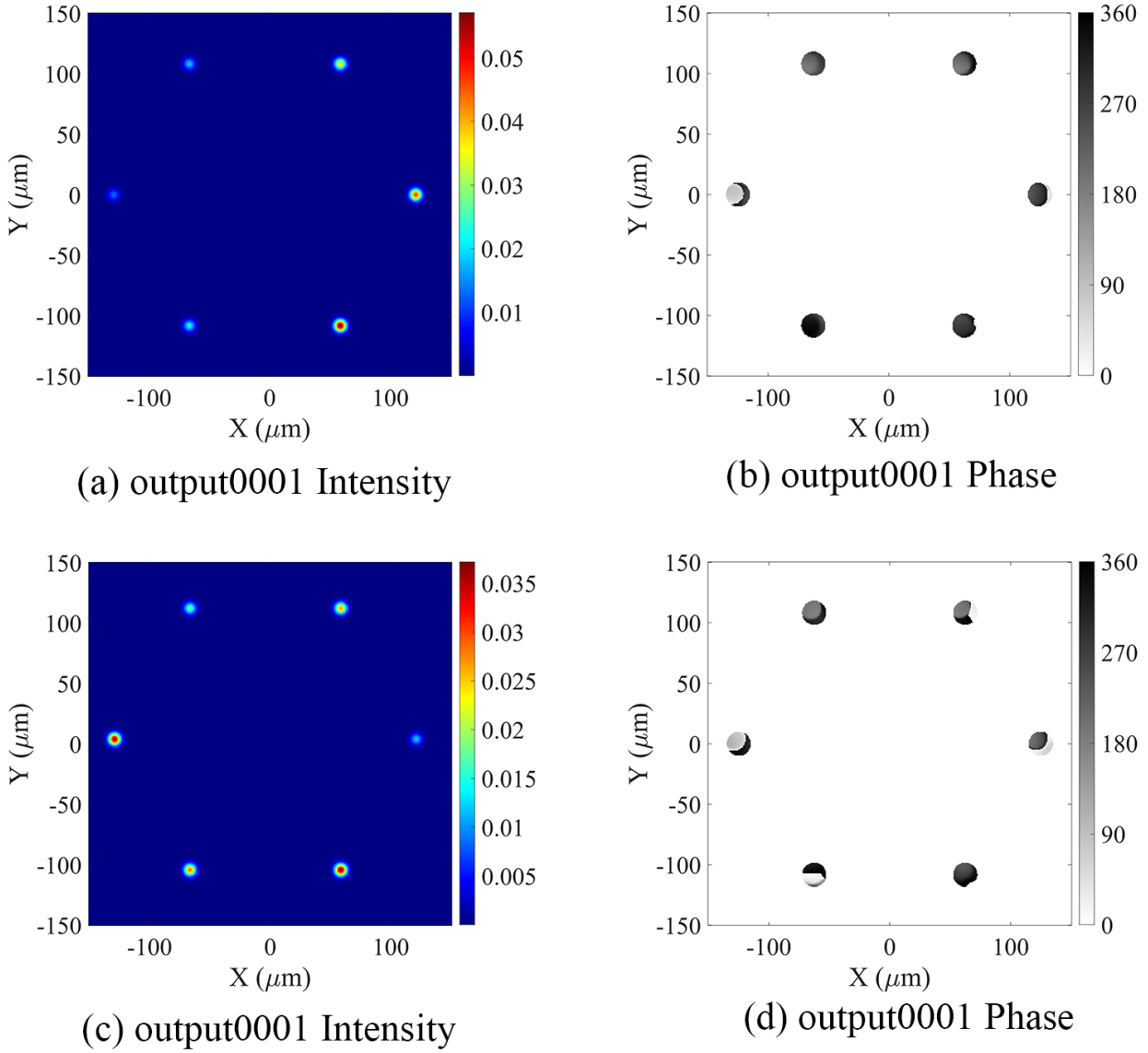


Figure 7. Intensity and phase images of the single-mode output from a 1 × 6 photonic lantern with offset splicing; (**a**,**b**) show the Unidirectional offset model, while (**c**,**d**) show the Bidirectional offset model.

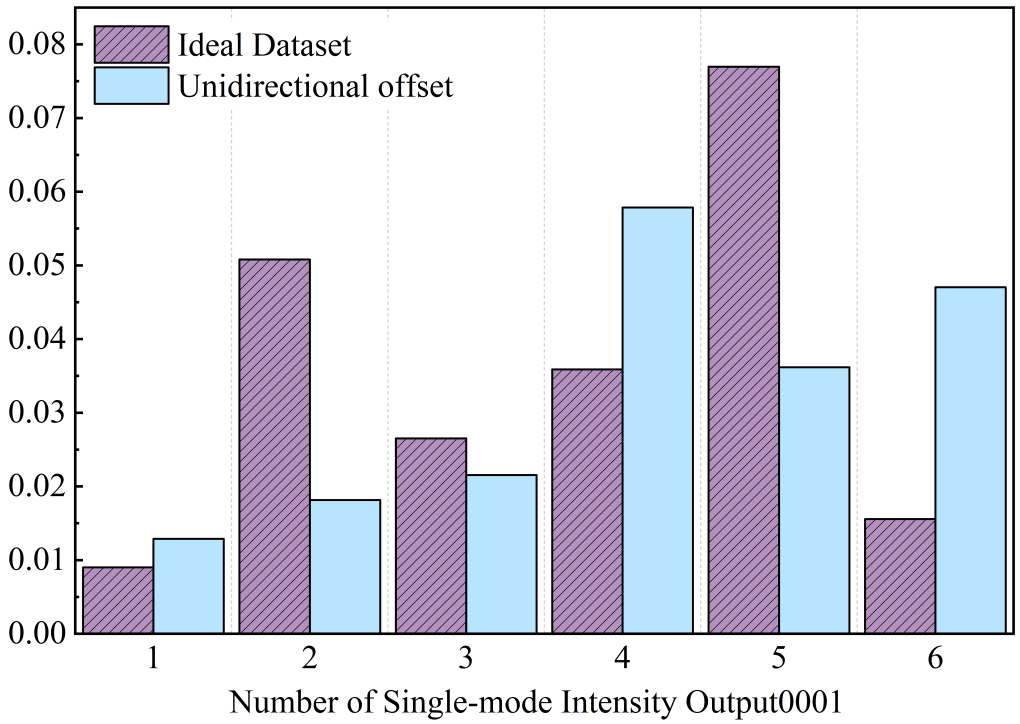


Figure 8. Intensity of the output from the single-mode end of a 1 × 6 photonic lantern with offset splicing (Unidirectional offset model).

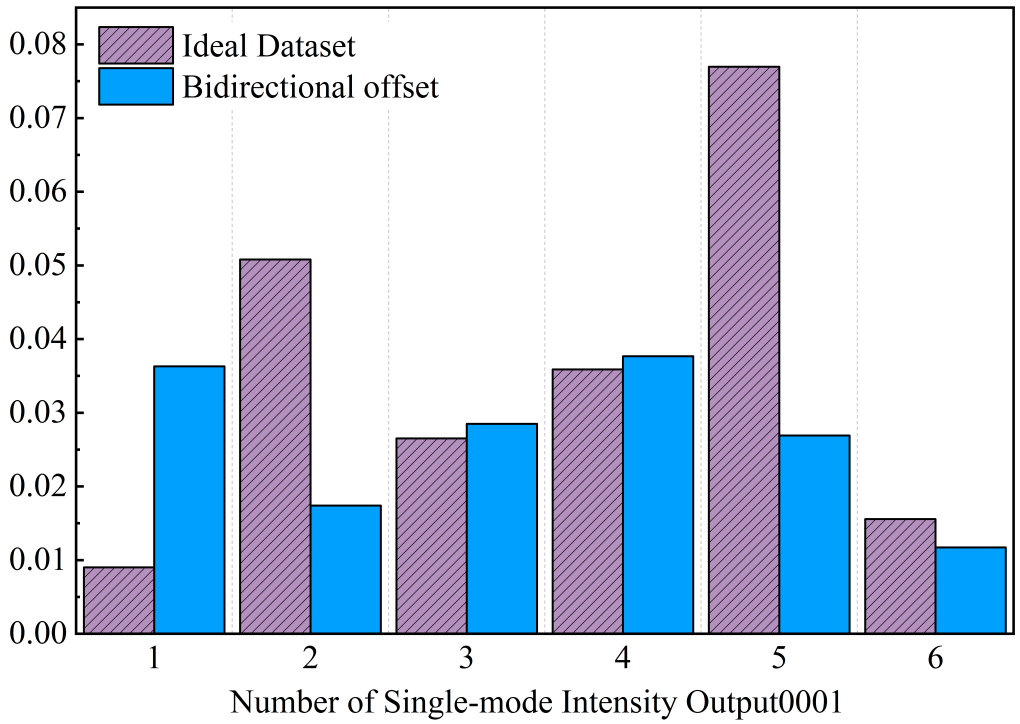


Figure 9. Intensity of the output from the single-mode end of a 1 × 6 photonic lantern with offset splicing (Bidirectional offset model).

3.1.3. Non-Uniform Tapering

Taper is a core process in the fabrication of photonic lanterns. It involves heating and stretching a fiber bundle to gradually reduce its cross-sectional dimensions, thereby achieving adiabatic mode evolution. The uniformity of the tapering process, particularly heating, is critical to the device's final performance [21]. Non-uniform tapering primarily results

from non-uniform distribution of the tensile force during electrode discharge, leading to non-uniform heating, as well as non-uniform internal arrangement of the fiber bundle.

Non-Uniform Heating

Photonic lanterns are typically manufactured using an optical fiber tapering process, which is extremely sensitive to the stability, shape, and position of the heat source (such as a flame) [20]. Flames are prone to turbulence and can fluctuate easily due to slight variations in airflow or residual air currents within the chamber; impure gases may also lead to fiber contamination [22]. A non-uniform static heat source typically produces a non-uniform taper waist diameter, leading to variations in the taper ratio across cores and disrupting phase-matching conditions between channels. This means that the cross-sectional dimensions of the taper zone vary along its length, and such variations can only be predicted through experimental calibration; they are difficult to control precisely using theoretical models [21].

As an example in Figure 10, errors caused by non-uniform heating reduced the coordinates and dimensions of certain core segments in four single-mode fibers within a 1×6 photonic lantern by 30%, resulting in these fibers being drawn thinner. Since non-uniform heating disrupts the symmetry of the photonic lantern, the light-field dataset here does not use symmetry-augmented data; instead, all training is performed on raw data. Example of light field data is shown in Figures 11 and 12.

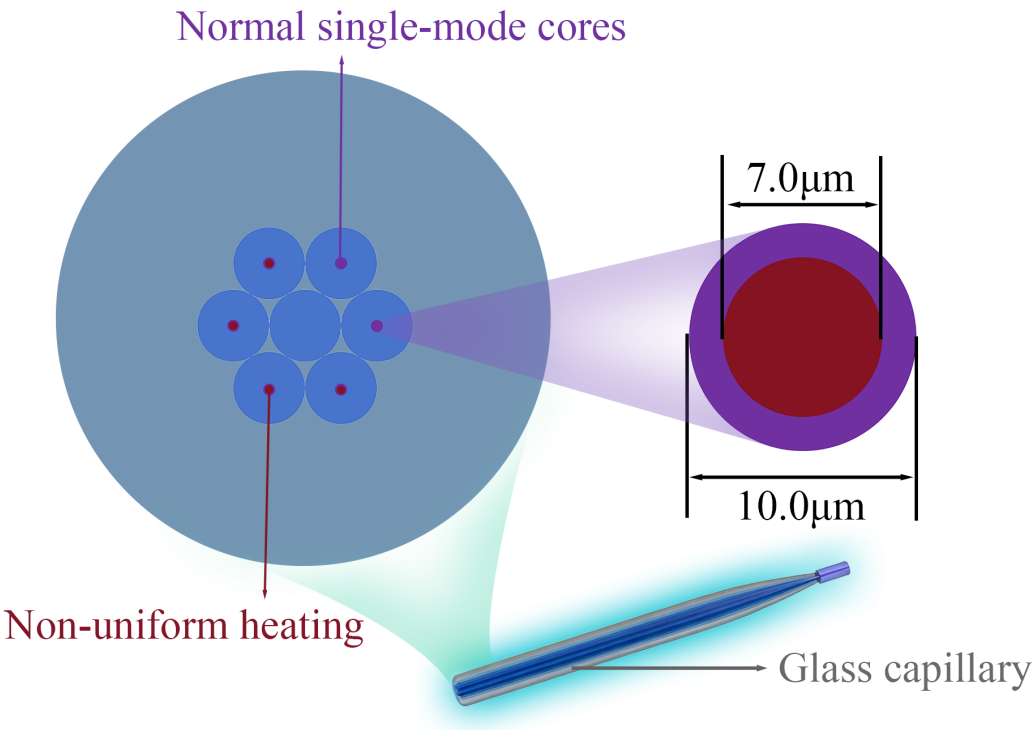


Figure 10. Non-uniform heating (when the end coordinates and dimensions of the core segments are multiplied by a factor of 0.7).

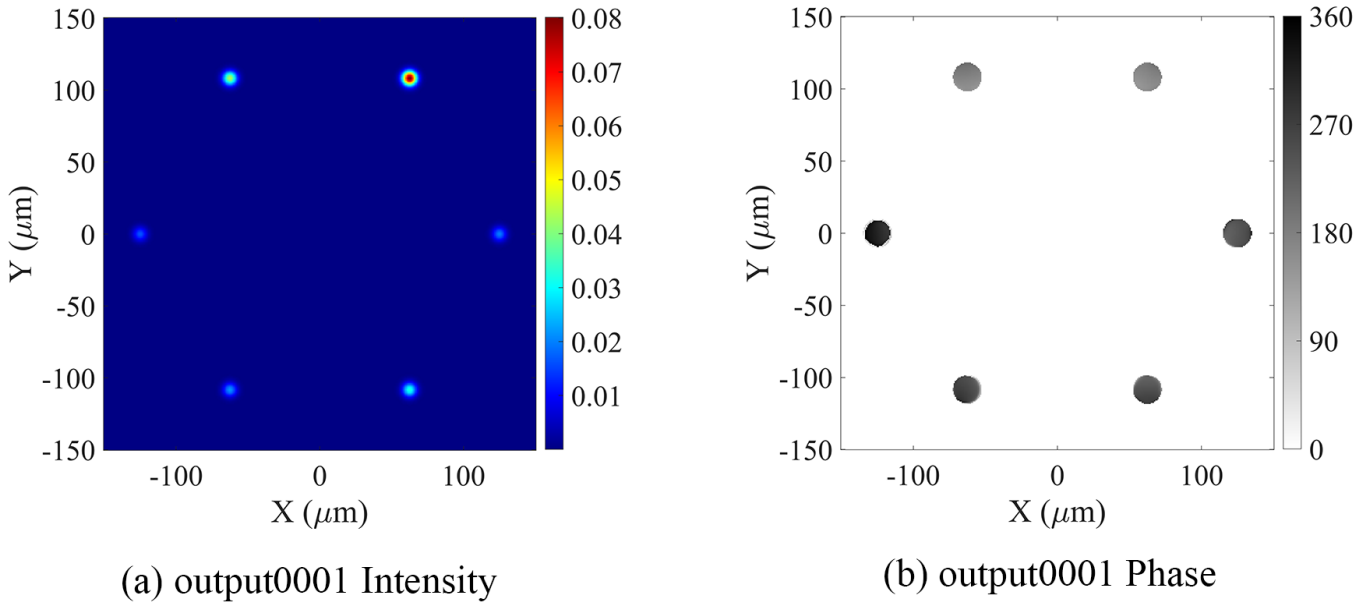


Figure 11. Intensity and phase images of the single-mode output from a 1 × 6 photonic lantern with non-uniform heating.

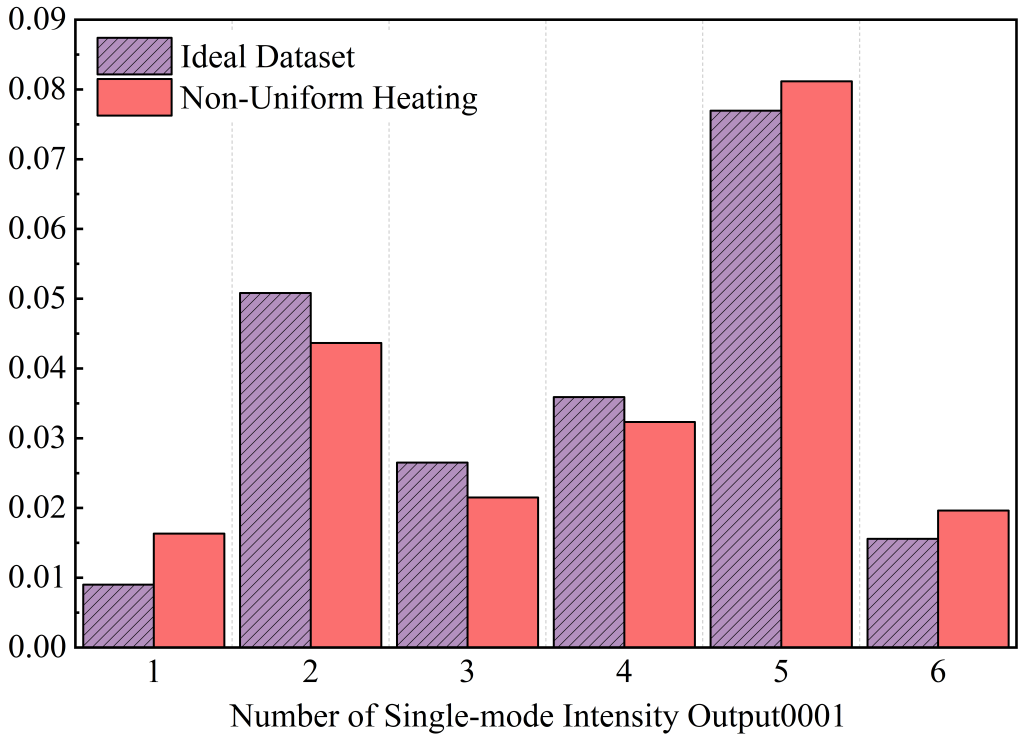


Figure 12. Intensity of the output from the single-mode end of a 1 × 6 photonic lantern with non-uniform heating.

Non-Uniform Arrangement

Ideally, the geometric arrangement of the cores should match the symmetry of the photonic lantern mode to minimize mode-dependent loss [23]. However, when bundling multiple single-mode fibers, if the core arrangement does not conform to the ideal geometric symmetry (e.g., a tight hexagonal arrangement) or if the core spacing is inconsistent, then during the drawing process, even with uniform heating, differences in the evolution paths and mutual coupling strengths of different cores will arise, equivalent to a form of structural

non-uniformity. This model simulates the geometric positional deviations caused by loose fiber packing within the capillary.

In a non-uniform arrangement shown in Figure 13, two symmetrically positioned cores are shifted inward by 10%, compressing the fiber arrangement and disrupting the ideal hexagonal close-packed structure. Example of light field data is shown in Figures 14 and 15.

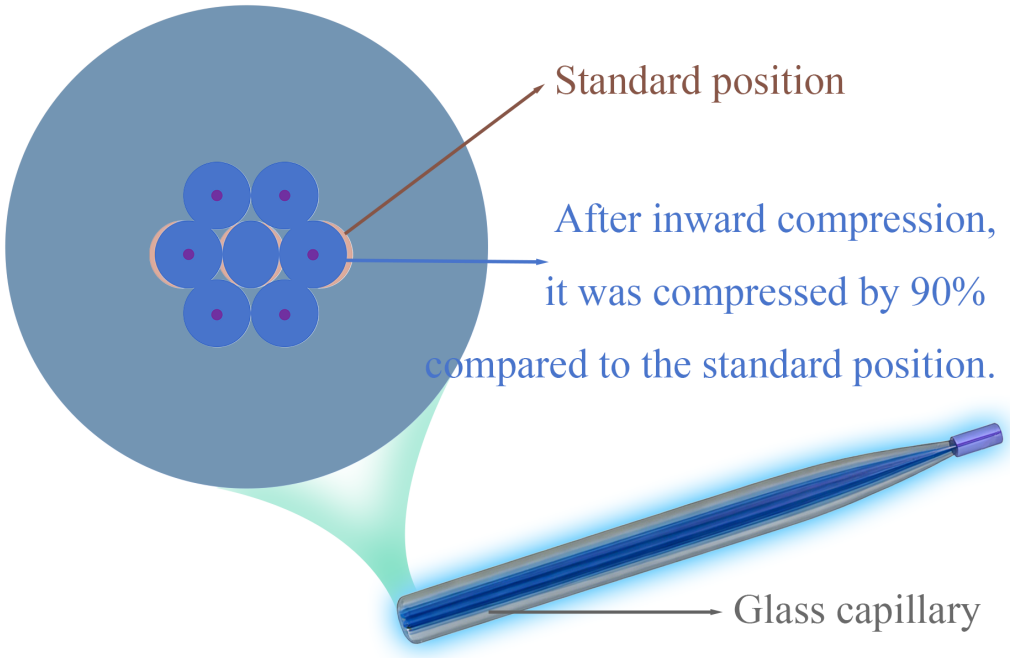


Figure 13. Non-uniform arrangement.

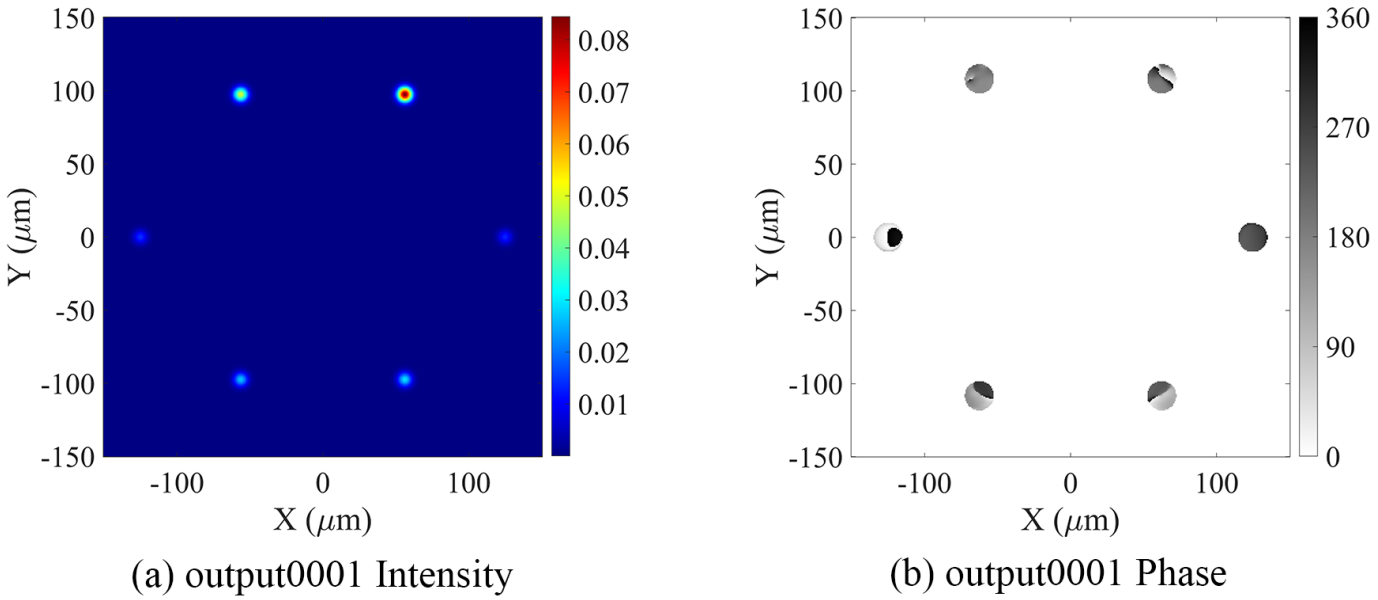


Figure 14. Intensity and phase images of the single-mode output from a 1 × 6 photonic lantern with non-uniform arrangement.

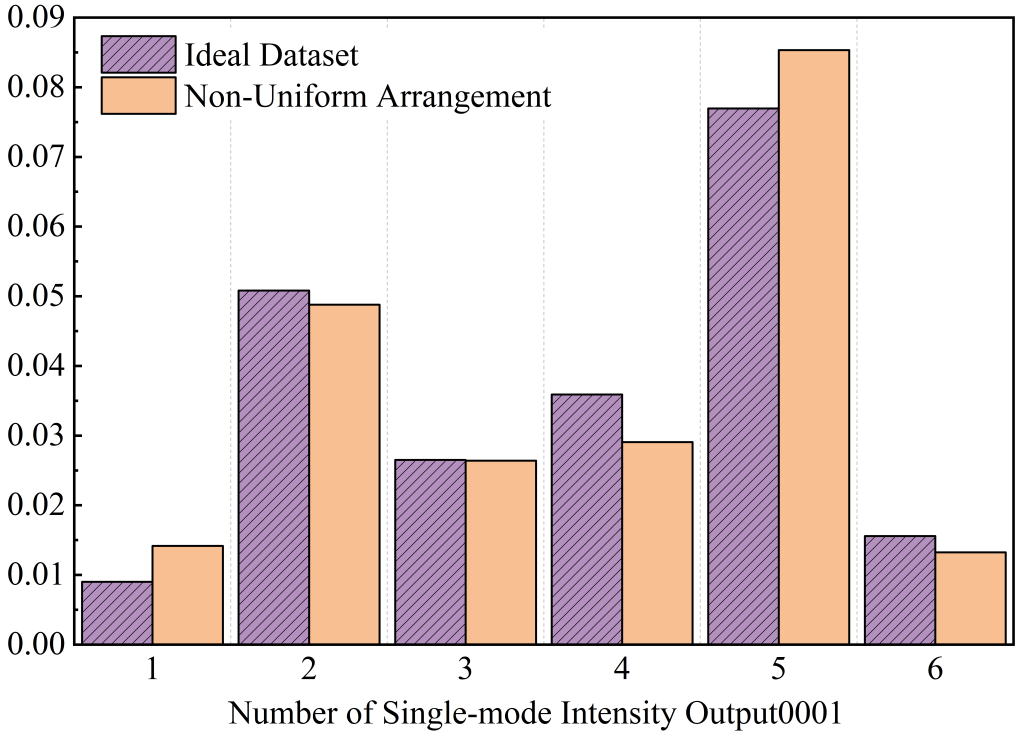


Figure 15. Intensity of the output from the single-mode end of a 1×6 photonic lantern with non-uniform arrangement.

3.2. Designing Neural Networks for Light Field Data Based on Light Field Datasets

This paper designs neural networks with different parameter settings based on a light-field dataset ($N > 1000$). The dataset is divided into training, validation, and test sets, ensuring that the validation and test sets are drawn from the original dataset, with a 90:5:5 split.

This paper adopts a Dropout network. A Dropout network is a model that systematically and strategically applies dropout or its variants within a deep neural network architecture. During the training phase, it uses a probability p as a hyperparameter to randomly remove neurons and their connections from the network for regularization, thereby preventing excessive co-adaptation among neurons and effectively mitigating overfitting [24]. For example, if the hyperparameter p is set to 0.2, each neuron has a 20% probability of being set to zero during each forward pass. Figures 16 and 17 illustrate a 3-layer Dropout network.

Derivation shows that for the activation output of layer l ($a^{(l)}$), during training, an element-wise multiplication is performed with a mask vector $m^{(l)}$, where the elements of $m^{(l)}$ are independently set to 0 with probability p and to 1 with probability $1 - p$, i.e., $\tilde{a}^{(l)} = a^{(l)} \odot m^{(l)}$, to perform random dropout.

Additionally, to avoid the need for random dropout during the testing phase while maintaining the expected value of the output, the retained activation values are scaled during training, i.e., $\tilde{a}^{(l)} = \frac{1}{1-p}(a^{(l)} \odot m^{(l)})$, which is referred to as Inverted Dropout.

Here, $\odot$ is an operator that performs element-wise multiplication followed by division by $1 - p$, ensuring that the expected values remain consistent between the training and testing phases.

In this paper, the model uses the MSE as the loss function to quantify the difference between the predicted light intensity and phase at the output and the true values. The MSE and the coefficient of determination R^2 are used to evaluate the performance of the regression model. The optimal model is obtained through parameter tuning.

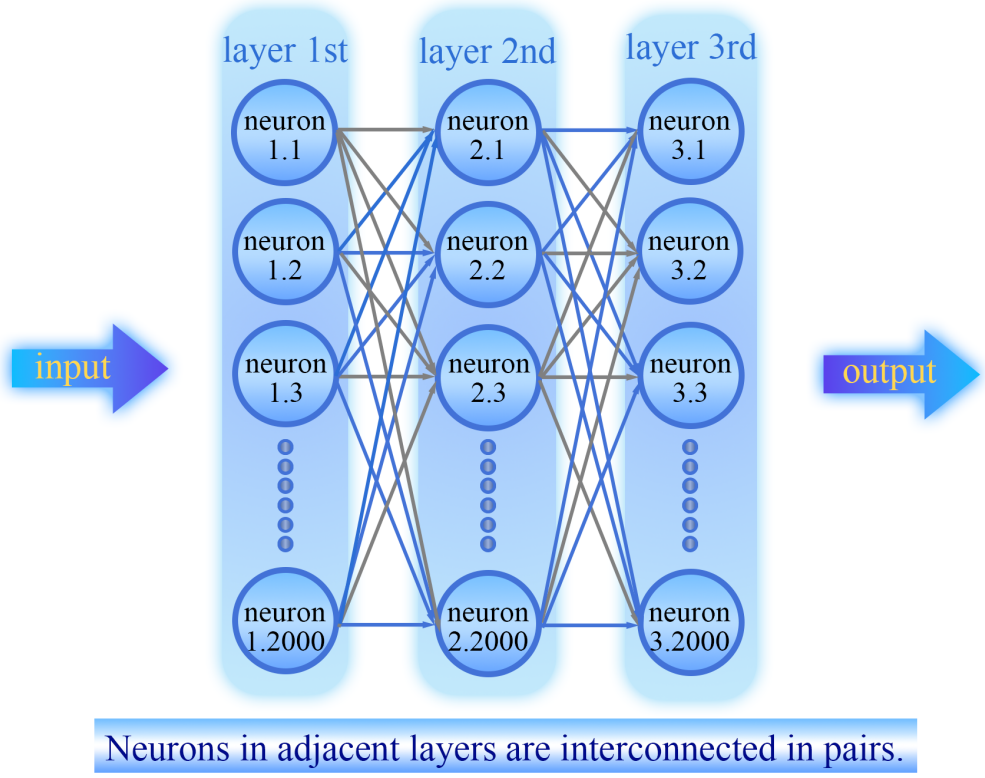


Figure 16. Overview of a Dropout network when $p = 0$.

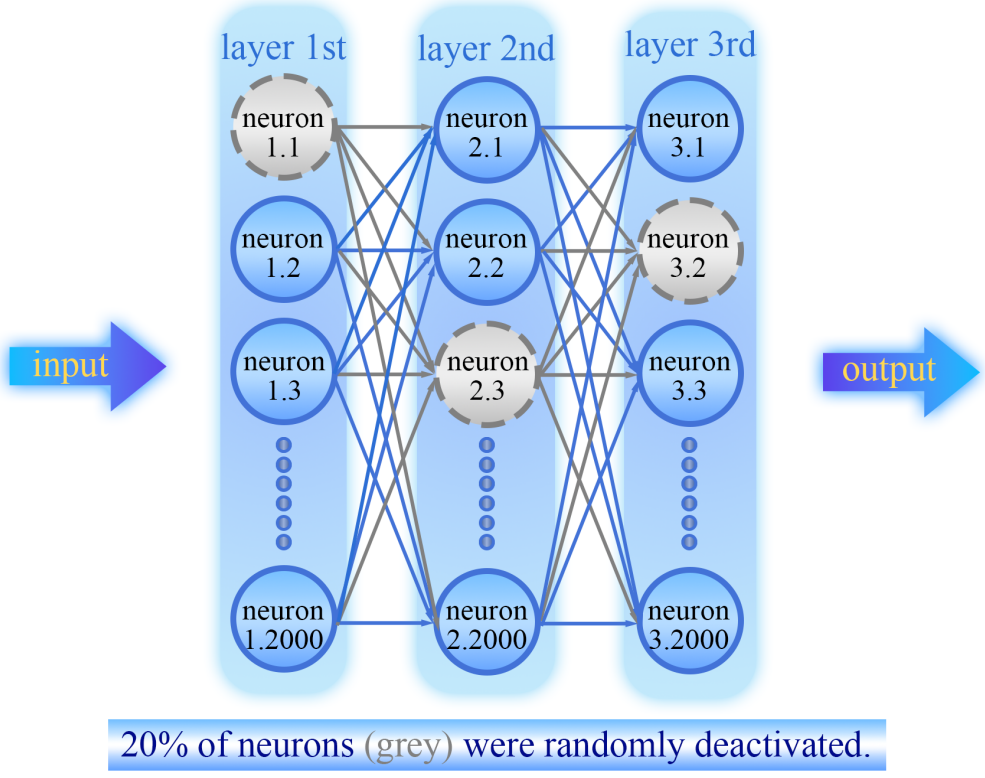


Figure 17. Dropout network when $p > 0$, using $p = 0.2$ as an example.

The dropout mechanism involves applying a random mask $m^{(l)}$ to the activation vector $a^{[l]}$ during the forward pass at layer l (where each element takes the value 0 with probability p and the value 1 with probability $1 - p$), resulting in:

$$\tilde{a}^{(l)} = \frac{1}{1-p}\left(m^{(l)} \odot a^{(l)}\right) \tag{8}$$

During testing, $m^{(l)} \equiv 1$ and the full $a^{[l]}$ is used directly.

For the light field intensity prediction task, let the network output be $\hat{I}_{\text{out}} = f_\theta(I_{\text{in}})$ and the true light field be I_{out}. The MSE loss is defined as:

$$L_{\text{MSE}}(\theta) = \frac{1}{N}\sum_{i=1}^{N} \|\hat{I}_{\text{out}}^{(i)} - I_{\text{out}}^{(i)}\|_2^2 \tag{9}$$

where N is the number of training samples.

After introducing dropout into each layer of the network, the overall objective remains to minimize MSE, but gradients are computed based on the "sparse" subnetworks at each iteration, thereby suppressing feature co-adaptation and improving generalization. The complete optimization problem can be written as:

$$\min_\theta E_{m^{[l]}}\left[\frac{1}{N}\sum_{i=1}^{N} \|f_\theta^{\text{drop}}(I_{\text{in}}^{(i)}) - I_{\text{out}}^{(i)}\|_2^2\right] \tag{10}$$

where f_θ^{drop} denotes the network with the dropout mask applied to the layer during forward propagation [25].

The network architecture in this paper employs LeakyReLU as the activation function across all layers, with a hyperparameter of 0.3, which is well-suited for continuous regression tasks in modeling the transmission characteristics of photonic lanterns. The Dropout design follows three principles: all hidden layers use the same Dropout rate to ensure consistent regularization effects; Dropout is applied after LeakyReLU to randomly deactivate activation values, thereby enhancing model robustness; the output layer does not use Dropout to ensure deterministic prediction results. Preliminary experiments show that setting the Dropout rate to 0.2 yields the lowest MSE on the validation set and relatively stronger generalization ability. With the Dropout rate kept constant, experiments were conducted with varying numbers of layers to identify the neural network model that achieves the best MSE on the validation set.

4. Results and Discussion

4.1. Performance of the Ideal Optical Sub-Lantern Transmission Characteristics Neural Network Model

This paper constructs a Dropout Network using dropout regularization, with all hyperparameters set to 0.2, meaning that in each training iteration, each neuron has a 20% probability of being dropped. The training batch size is set to 64, and the training runs 25,000 iterations. A strategy dynamically adjusts the learning rate; the system continuously monitors the validation set loss. If the improvement in the loss is less than 10^{-6} over 1000 consecutive iterations, the learning rate is reduced to 50% of its current value, with a lower limit set at 5×10^{-7}. Table 3 summarizes the performance of the transmission characteristic neural network. Figure 18 shows the training process. The data in Table 3 were obtained by validating the model on a test set of original light field data (unaugmented) that was not used in training, after 25,000 training iterations. The light-field data in this test set consisted of 300 pairs. Note that the numbers in parentheses represent the number of pairs in the light field dataset before augmentation.

Table 3. Performance Summary of the Dropout-Based Ideal Photonic Lantern Transmission Characteristics Neural Network.

Layers	Hidden Layer Size	Training Examples	Intensity MSE (Norm. Units)	Intensity R^2 (Norm. Units)	Phase MSE (Norm. Units)	Phase R^2 (Norm. Units)
3	2000	6000(1500)	9.793×10^{-5}	9.973×10^{-1}	4.063×10^{-3}	9.516×10^{-1}
5	2000	6000(1500)	1.167×10^{-4}	9.967×10^{-1}	3.814×10^{-3}	9.546×10^{-1}
9	2000	6000(1500)	1.196×10^{-4}	9.967×10^{-1}	4.529×10^{-3}	9.461×10^{-1}
15	2000	6000(1500)	1.219×10^{-4}	9.966×10^{-1}	4.267×10^{-3}	9.492×10^{-1}
21	2000	6000(1500)	1.203×10^{-4}	9.966×10^{-1}	3.772×10^{-3}	9.551×10^{-1}
27	2000	6000(1500)	1.230×10^{-4}	9.966×10^{-1}	4.640×10^{-3}	9.448×10^{-1}

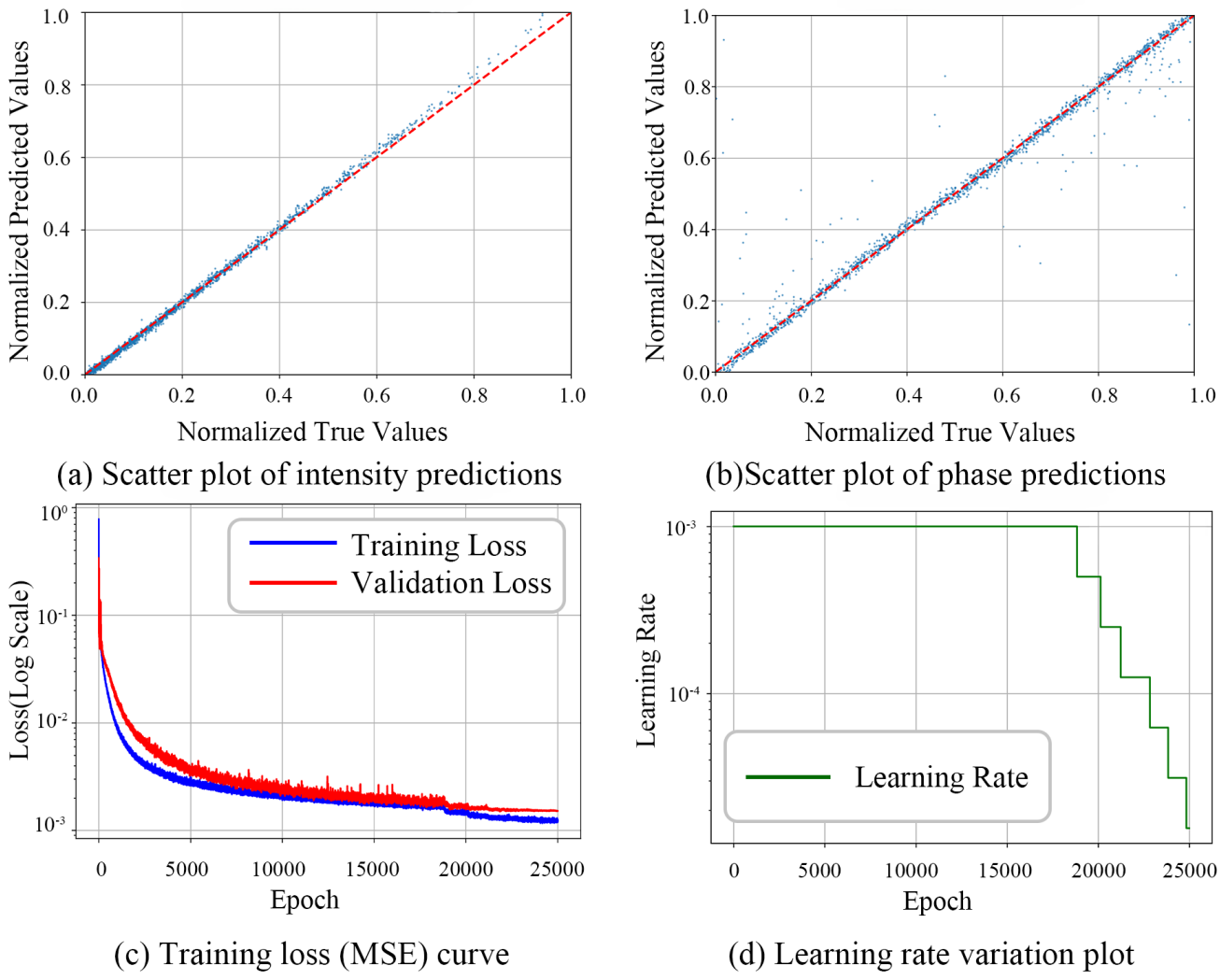


Figure 18. Training Process and Results of the Dropout Network.

Additionally, this paper uses the original light field dataset from an ideal 1×6 photonic lantern (1500 pairs in total, without $4\times$ data augmentation) to estimate the ideal transmission matrix via least-squares, serving as a control group for Dropout network training. As shown in Equation (4), this process can be expressed as the linear Equation $Y = T \times X$. Here, X is a column vector containing the complex amplitudes of each mode at the input; Y is also a column vector containing the complex amplitudes of each single-mode fiber at the output; and T is a complex matrix, i.e., the transmission matrix we seek, which fully describes the mapping relationship of the light field from the mode space to the spatial ports, including the coupling coefficients between each mode and each fiber as well as

the phase delay [26] during transmission. The least-squares method is used to obtain an optimal matrix T that minimizes the sum of squared errors between the predicted output $T \times X$ and the actual observed output Y. Figure 19 shows the prediction results of the ideal matrix for the input light field data; the test light field data aligns with the neural network, totaling 300 pairs.

It should be noted that, since the systematic nonlinear error in the transmission characteristics of photonic lanterns, the linear transmission matrix fails under a large number of random-state distributions, the transmission matrix calculated from a smaller dataset better aligns with the actual transmission characteristics. At the same time, neural networks can fit these nonlinear characteristics. Therefore, as the dataset size increases, the performance of the transmission matrix does not improve; instead, greater deviations occur. To control for variables and facilitate comparison with the neural network model, this paper used only 300 data pairs in calculating the transmission matrix.

As shown in Table 4 and Figure 19, the Dropout network achieves high accuracy in modeling the photonic lantern's transmission characteristics. The MSE for intensity can reach as low as 9.793×10^{-5}, and the MSE for phase can reach 4.063×10^{-3}, which is far superior to the MSE when the transfer matrix predicts the single-mode output. Since the transfer matrix of a photonic lantern is invertible, it follows that if a Dropout network is used to construct a reverse transfer characteristic neural network and trained in reverse using both the single-mode and multimode outputs, the reverse transfer characteristic neural network can also achieve high accuracy.

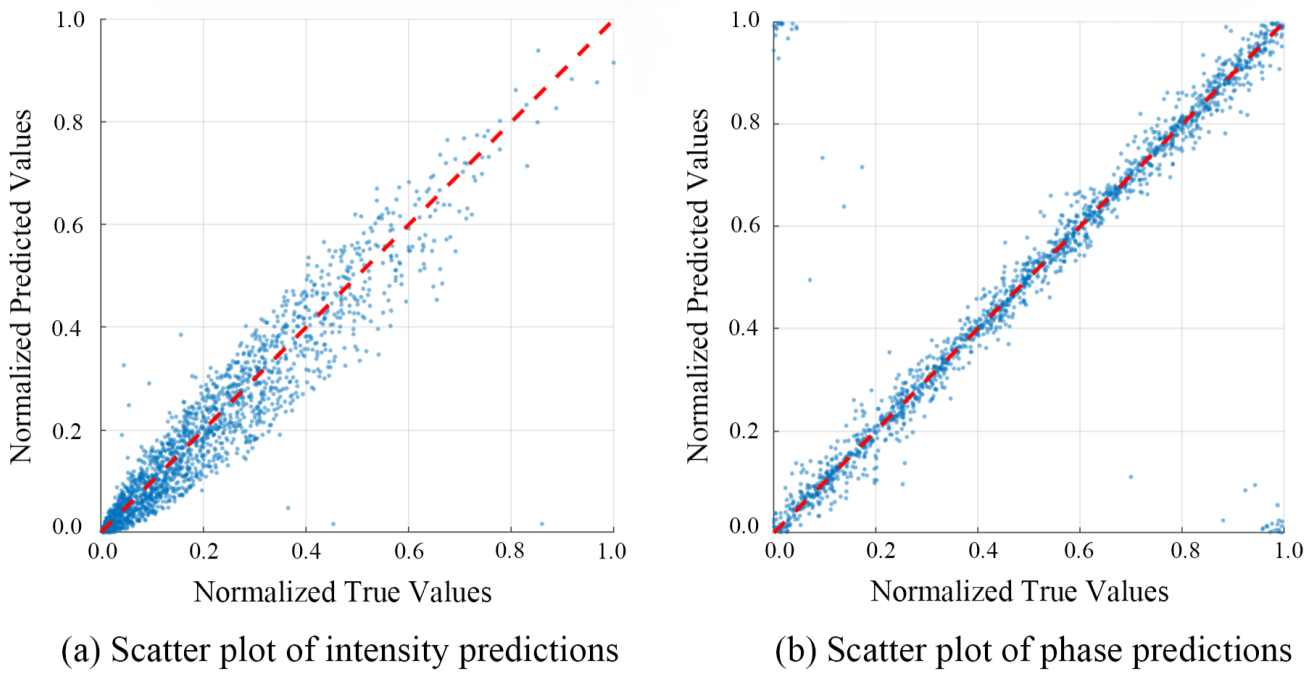


Figure 19. Scatter plot of transmission matrix prediction results.

Table 4. Performance of the Ideal Transmission Matrix Based on the Least Squares Method.

Type	Intensity MSE (Norm. Units)	Intensity R^2 (Norm. Units)	Phase MSE (Norm. Units)	Phase R^2 (Norm. Units)
Least Squares	3.175×10^{-3}	8.960×10^{-1}	1.989×10^{-2}	7.574×10^{-1}
Dropout Network	9.793×10^{-5}	9.973×10^{-1}	4.063×10^{-3}	9.516×10^{-1}

As shown in Figures 20 and 21, as well as Tables 5 and 6, taking the intensity and phase of Output 0001 detected by the single-mode port of an ideal 1×6 photonic lantern

as an example, for intensity prediction, the maximum error between the neural network model's predicted intensity and the ideal dataset is within 2.07%, whereas the maximum error in intensity prediction by the transfer matrix reaches 13.57%. For phase prediction, the maximum error between the neural network model's predicted phase and the ideal dataset is within 6.84%, whereas the maximum error for the transmission matrix's predicted phase reaches 24.43%.

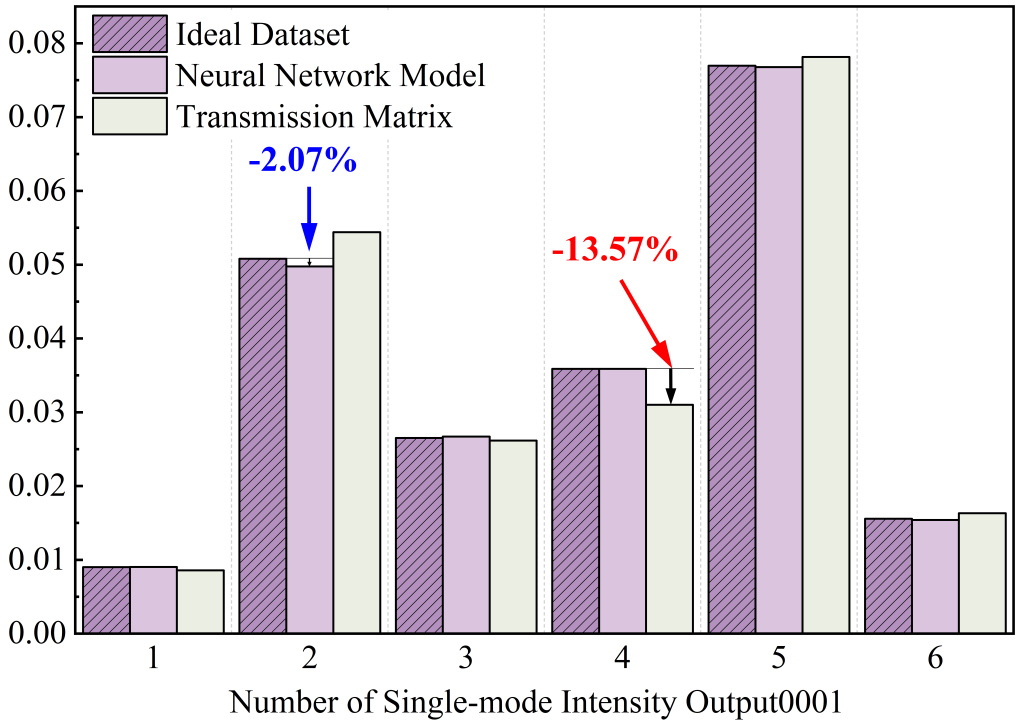


Figure 20. Intensity comparison at the single-mode output of the ideal dataset, the neural network model, and the transmission matrix.

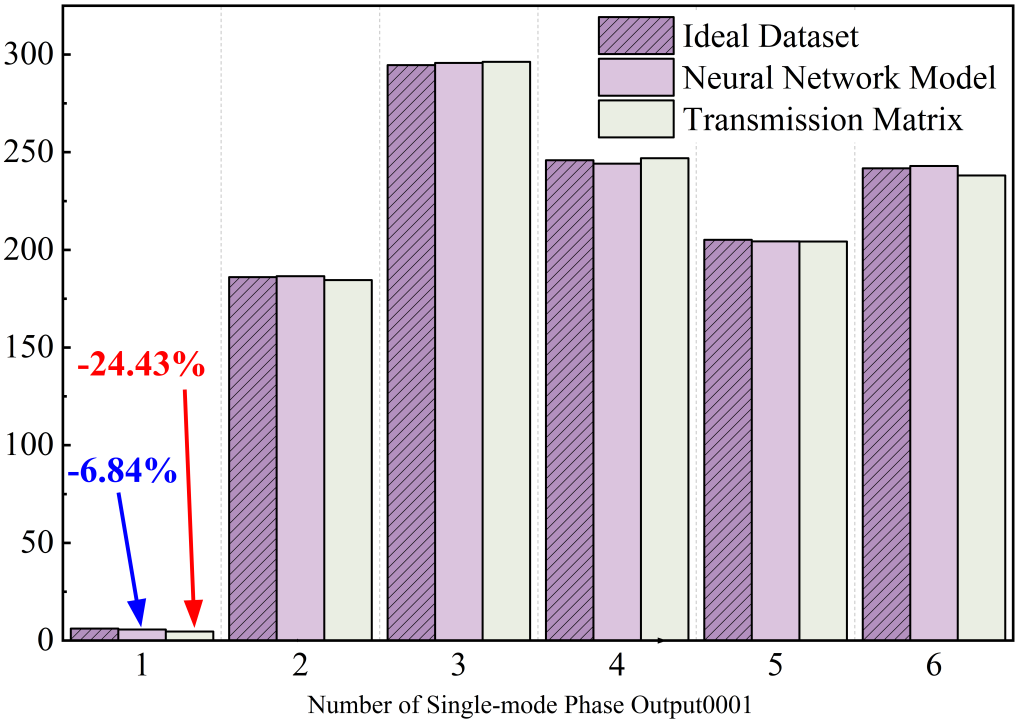


Figure 21. Phase comparison at the single-mode output of the ideal dataset, the neural network model, and the transmission matrix.

Table 5. Percentage error in optical intensity at the single-mode output of the neural network model and the transfer matrix.

Number of Single-Mode Output0001	Neural Network Model-Intensity	Transmission Matrix-Intensity
1	0.22%	−4.77%
2	−2.07%	7.09%
3	0.72%	−1.36%
4	−0.03%	−13.57%
5	−0.23%	1.55%
6	−1.09%	4.82%

Table 6. Percentage phase error of the single-mode output from the neural network model and transfer matrix.

Number of Single-Mode Output0001	Neural Network Model-Phase	Transmission Matrix-Phase
1	−6.84%	−24.43%
2	0.27%	−0.82%
3	0.38%	0.56%
4	−0.72%	0.43%
5	−0.39%	−0.43%
6	0.50%	−1.51%

During the experiment, this paper observed a strong positive correlation between dataset size and neural network accuracy: the larger the dataset, the lower the MSE, and the closer the coefficient of determination approaches 1. Currently, a Dropout network is used, with 6000 data points generated via 4-fold data augmentation from 1500 original data points for training. It can be inferred that if the data volume is increased, the accuracy of the transmission characteristic neural network can be improved while ensuring the model does not overfit during training.

When the current model predicts the single-mode output, the MSE for intensity is significantly lower than that for phase. Although the transmission characteristic neural network performs exceptionally well in handling energy distribution, phase information is more complex, and traditional Dropout networks still have certain limitations when processing phase relationships.

Based on our current research, we have considered two optimization schemes for phase feature learning: (1) sine-cosine encoding, which converts phase information from a scalar value to a sine-cosine pair, increasing the input light field dimension from $2N$ to $3N$, thereby preserving periodic information and eliminating boundary discontinuity issues; (2) an optimized complex-valued neural network (CVNN), which demonstrated superior performance to Dropout networks in phase modeling and exhibited greater physical consistency during preliminary experiments, but lagged behind in light intensity prediction accuracy and exhibited significant overfitting that was difficult to resolve, necessitating further optimization.

As shown in Figure 22, the transmission characteristic neural network also possesses high-efficiency processing capabilities. Once the photonic lantern transmission characteristic neural network model is trained, it can produce predictions at 53.30848 ms/sample, compared to simulation methods that take 30 to 60 min to obtain single-mode output light-field results, while maintaining high accuracy. The processing time is significantly faster than that of traditional simulation methods, providing a technical pathway for efficient batch processing of light-field data.

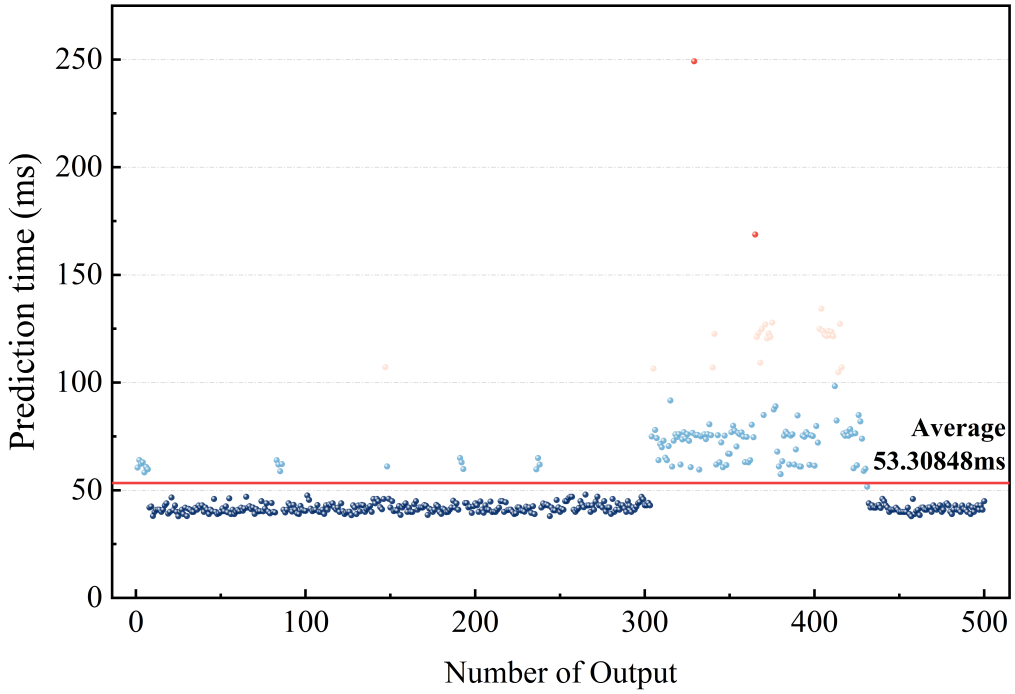


Figure 22. Prediction time of the Trained Neural Network.

4.2. Photonic Lantern Transmission Characteristics Neural Network Model

To verify the neural network model's calibration capability for the photonic lantern's transmission characteristics, a systematic cross-validation experiment was designed in this paper. The experiment uses MSE as the core metric for evaluating prediction accuracy, defined as shown in Equation (11):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{11}$$

where y_i is the true value, $\hat{y}_i$ is the model-predicted value, and N is the number of test samples. A smaller MSE value indicates higher prediction accuracy, while a larger MSE value indicates greater prediction error.

To analyze the accuracy of predictions using MSE, this paper defines the MSE Multiple Index called $m_{\alpha-\beta/\text{ideal}}$, whose value is the ratio of $\text{MSE}_{\alpha-\beta}$ to $\text{MSE}_{\text{ideal}}$, as shown in Equation (12). $\text{MSE}_{\alpha-\beta}$ is defined as the MSE of the α model's prediction of the β dataset, and $\text{MSE}_{\text{ideal}}$ is the MSE of the optimal ideal 1×6 photonic lantern neural network model's prediction of the ideal 1×6 photonic lantern dataset.

$$m_{\alpha-\beta/\text{ideal}} = \frac{\text{MSE}_{\alpha-\beta}}{\text{MSE}_{\text{ideal}}} \tag{12}$$

In this paper, only the neural network model for the optimal ideal 1×6 photonic lantern, trained on a large dataset (6000 pairs of extended data and 1500 pairs of original data), is used to predict different datasets. The neural network model for the manufacturing error photonic lantern, trained on a small dataset (500 pairs of original data), is used solely to predict the corresponding manufacturing error dataset.

Tables 7 and 8 present the normalized MSE and the MSE Multiple Index $m_{\alpha-\beta/\text{ideal}}$ for luminance across various models and test sets.

Analysis of the above data shows that the neural network model for the photonic lantern's transmission characteristics can effectively detect manufacturing errors.

Table 7. Intensity Norm.MSE Matrix.

Neural Networks Model	Photonic Lantern Dataset				
	Ideal	**Unidirectional Offset**	**Bidirectional Offset**	**Non-Uniform Heating**	**Non-Uniform Arrangement**
Ideal	2.418×10^{-3}	3.272×10^{-2}	5.149×10^{-2}	9.849×10^{-3}	6.339×10^{-3}
Unidirectional offset		9.660×10^{-3}			
Bidirectional offset			4.362×10^{-3}		
Non-Uniform Heating				9.415×10^{-3}	
Non-Uniform Arrangement					5.282×10^{-3}

Table 8. Intensity MSE Multiple Index $m_{\alpha-\beta/\text{ideal}}$ Matrix.

Neural Networks Model	Photonic Lantern Dataset				
	Ideal	**Unidirectional Offset**	**Bidirectional Offset**	**Non-Uniform Heating**	**Non-Uniform Arrangement**
Ideal	1	13.533	21.298	4.074	2.622
Unidirectional offset		3.995			
Bidirectional offset			1.804		
Non-Uniform Heating				3.894	
Non-Uniform Arrangement					2.185

Analysis of the first row of data in Table 8 reveals that, when using the ideal model to predict the manufacturing error dataset, the MSE for light intensity increased significantly, from 2.622 to 21.298. These data indicate a significant difference in the transmission characteristics between photonic lanterns with manufacturing errors and ideal photonic lanterns. Therefore, neural network models trained on ideal photonic lantern datasets and transmission matrices are not suitable for calculating the transmission characteristics of photonic lanterns with manufacturing errors. Furthermore, as concluded in Section 4.1, the MSE of the ideal photonic lantern neural network model is far superior to that of the transfer matrix when predicting the single-mode output. These data prove that the neural network model outperforms the transfer matrix in calibrating photonic lanterns with manufacturing errors. If the transfer matrix is used to predict the single-mode output of a photonic lantern with manufacturing errors, the error will be significantly greater than that of the neural network model.

By training a photonic lantern neural network model, we performed predictions on a test set of 500 pairs of similar photonic lanterns and obtained the model's MSE for light intensity along the diagonal. We found that the MSE was amplified by a factor of 1.804 to 3.995. Given that the neural network model for manufacturing error photonic lanterns was trained on only 500 pairs of original data, the model's prediction accuracy (MSE) is expected to be somewhat lower than that of the ideal photonic lantern model trained on a dataset of 6000 pairs (1500 pairs of original data). This further corroborates the inference in Section 4.1: there is a strong positive correlation between dataset size and neural network accuracy—the larger the dataset, the lower the MSE.

As shown in Table 9, compared to the light intensity prediction accuracy of the ideal photonic lantern neural network model on the manufacturing error training set, the accuracy of the Unidirectional Offset neural network model increased by 3.387 times, the Bidirectional Offset neural network model increased by 11.860 times, and the Non-Uniform Heating neural network model increased by 1.046 times. Furthermore, the Non-Uniform Arrangement neural network model saw a 1.200-fold improvement in accuracy. Analysis indicates that, compared to manufacturing errors caused by non-uniform heat

input and non-uniform arrangement, the neural network faces relatively greater difficulty in capturing the transmission characteristics of the latter two types of manufacturing errors during small-sample tuning, as shown below. Given that the MSE deviations for non-uniform heat input and non-uniform arrangement are relatively small in the ideal model, it can be inferred that the impact of taper non-uniformness-related errors on the transmission characteristics of the photonic lantern is relatively smaller compared to that of misalignment welding.

Table 9. Multiple Index m for Intensity MSE of Photonic Lanterns with Manufacturing Errors.

Type of Manufacturing Errors (Model/Dataset α)	$m_{\alpha-\alpha/\text{ideal}}$	$m_{\text{ideal}-\alpha/\text{ideal}}$	$\frac{m_{\text{ideal}-\alpha/\text{ideal}}}{m_{\alpha-\alpha/\text{ideal}}}$
Unidirectional offset	3.995	13.533	3.387
Bidirectional offset	1.804	21.298	11.860
Non-Uniform Heating	3.894	4.074	1.046
Non-Uniform Arrangement	2.185	2.622	1.200

Similarly, Tables 10 and 11 summarize the normalized MSE and the MSE Multiple Index for luminous intensity across different test sets for each model.

Table 10. Phase Norm.MSE Matrix.

Neural Networks Model	Photonic Lantern Dataset				
	Ideal	Unidirectional Offset	Bidirectional Offset	Non-Uniform Heating	Non-Uniform Arrangement
Ideal	3.160×10^{-3}	7.873×10^{-2}	5.149×10^{-2}	7.218×10^{-2}	6.773×10^{-2}
Unidirectional offset		1.939×10^{-2}			
Bidirectional offset			9.498×10^{-3}		
Non-Uniform Heating				2.576×10^{-2}	
Non-Uniform Arrangement					1.083×10^{-2}

Table 11. Phase MSE Multiple Index $m_{\alpha-\beta/\text{ideal}}$ Matrix.

Neural Networks Model	Photonic Lantern Dataset				
	Ideal	Unidirectional Offset	Bidirectional Offset	Non-Uniform Heating	Non-Uniform Arrangement
Ideal	1	24.914	36.172	22.84	21.432
Unidirectional offset		6.135			
Bidirectional offset			3.006		
Non-Uniform Heating				8.151	
Non-Uniform Arrangement					3.427

Analysis of the first row of data in Table 11 reveals that when using the ideal model to predict the manufacturing error dataset, the phase MSE increased significantly, from 21.432 to 36.172 times, a more pronounced increase than the variation in light intensity. These data indicate that changes in transmission characteristics caused by manufacturing errors have a particularly significant impact on phase, further corroborating the presence of significant differences between the transmission characteristics of photonic lanterns with manufacturing errors and those of ideal photonic lanterns.

Analysis of the phase MSE for the photonic lantern model with manufacturing errors along the diagonal shows that the MSE increased by a factor of 3.006 to 8.151. It can be inferred that, under small-sample conditions, the neural network model can still be further optimized to improve phase prediction accuracy. This also supports the conclusion in Section 4.1: traditional Dropout networks still have certain limitations when handling phase relationships.

As shown in Table 12, compared to the phase prediction accuracy of the ideal photonic lantern neural network model on the manufacturing error training set, the accuracy of the Unidirectional Offset neural network model improved by a factor of 4.061, the Bidirectional Offset neural network model by a factor of 12.033, and the Non-Uniform Heating neural network model by a factor of 2.802, and the Non-Uniform Arrangement neural network model achieved a 6.254-fold improvement.

Table 12. Multiple Index m for Phase MSE of Photonic Lanterns with Manufacturing Errors.

Type of Manufacturing Errors (Model/Dataset α)	$m_{\alpha-\alpha/\text{ideal}}$	$m_{\text{ideal}-\alpha/\text{ideal}}$	$\frac{m_{\text{ideal}-\alpha/\text{ideal}}}{m_{\alpha-\alpha/\text{ideal}}}$
Unidirectional offset	6.135	24.914	4.061
Bidirectional offset	3.006	36.172	12.033
Non-Uniform Heating	8.151	22.84	2.802
Non-Uniform Arrangement	3.427	21.432	6.254

As shown in Figure 23, since the phase transmission characteristics of photonic lanterns with manufacturing errors vary significantly compared to those of ideal photonic lanterns, and the variation in phase transmission characteristics is greater than that in light intensity transmission characteristics, the neural network models demonstrate improved accuracy when handling all four types of manufacturing errors compared to light intensity. This is particularly evident for non-uniform arrangement, where accuracy improved from 1.200 times that of light intensity to 6.254 times; and for non-uniform intensity distribution, accuracy improved from 1.046 times that of the intensity to 2.802 times.

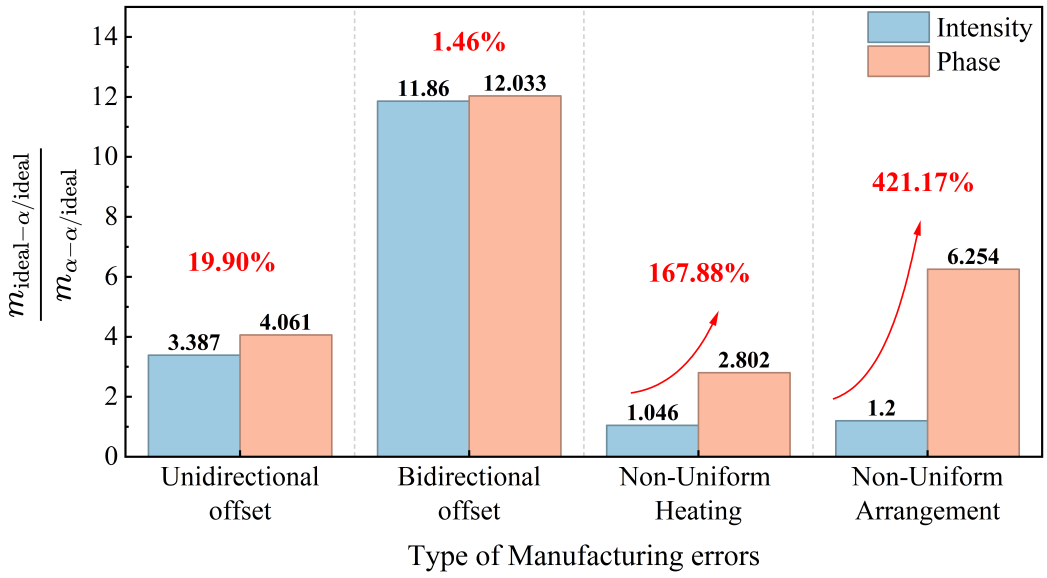


Figure 23. $m_{\text{ideal}-\alpha/\text{ideal}} / m_{\alpha-\alpha/\text{ideal}}$ of intensity versus phase.

5. Conclusions

This paper employs a neural network to characterize photonic lantern devices by training a model of their transmission characteristics. It achieved a predicted intensity MSE of 9.793×10^{-5} for an ideal 1×6 photonic lantern from the multimode end to the single-mode end, representing a 32.42-fold improvement over the intensity predictions MSE using the transmission matrix. The predicted phase MSE reached 3.814×10^{-3}, representing a 4.90-fold improvement over the MSE of phase predictions based on the transmission matrix. This study also found that the transmission characteristics of photonic lanterns with manufacturing errors differ significantly from those of ideal photonic lanterns; the neural network model for photonic lantern transmission characteristics achieves higher accuracy in modeling photonic lanterns with manufacturing errors than traditional modeling methods. When using the ideal model to predict the manufacturing error dataset, the intensity MSE increased by up to 21.298 times, and the phase MSE increased by up to 36.172 times, enabling the reliable detection of photonic lanterns with manufacturing errors. Compared to traditional simulation methods, the neural network model for photonic lantern transmission characteristics can compute light-field outputs more efficiently after training; the time required to process a single light field is reduced from minutes with traditional methods to milliseconds in this study. Additionally, this study found that the number of samples is significantly correlated with model accuracy. If the dataset is appropriately expanded, it is anticipated that accuracy will further improve, and the neural network model still holds considerable potential for calibrating the nonlinear relationships of photonic lanterns. However, using MSE as the sole evaluation metric can only determine the presence of manufacturing errors but cannot distinguish their specific types.

In future work, we will proceed to fabricate photonic lanterns and conduct further research on the transmission characteristic models—particularly those with manufacturing errors—by combining simulation models with experimental data. We will prepare a series of photonic lantern samples with known, controllable manufacturing errors (such as core offset and diameter deviation) and measure their transmission characteristics. By systematically comparing the predictions from our model (trained on simulation data) with actual measurement data, we will quantitatively evaluate the model's generalization capability and prediction accuracy under real-world conditions. Additionally, we will design and fabricate various $1 \times N$ photonic lantern models to further investigate how this neural network model can be applied to photonic lanterns with a wider range of configurations. In the next phase of research, we plan to introduce additional evaluation metrics and combine them with explainable artificial intelligence techniques to achieve not only detection but also classification of manufacturing error types.

Author Contributions: Conceptualization, Z.S., Y.L. and Z.J.; methodology, Z.S. and X.L.; software, Z.S. and T.L.; validation, Z.S., X.L. and T.L.; resources, Z.S., X.L., Y.L. and Z.C.; writing—original draft preparation, Z.S.; writing—review and editing, Y.L. and Z.J.; visualization, Z.S. and X.L. All authors have read and agreed to the published version of the manuscript.

Funding: National Natural Science Foundation of China (62405378); National Key Laboratory of Space Target Awareness (STA2025ZCA0404).

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Birks, T.A.; Gris-Sánchez, I.; Yerolatsitis, S.; Leon-Saval, S.G.; Thomson, R.R. The photonic lantern. *Adv. Opt. Photonics* **2015**, *7*, 107–167. [CrossRef]
2. Fontaine, N.K.; Ryf, R.; Bland-Hawthorn, J.; Leon-Saval, S.G. Geometric requirements for photonic lanterns in space division multiplexing. *Opt. Express* **2012**, *20*, 27123–27132. [CrossRef]
3. He, Y.; Chen, X. Optical Characteristic Research on Fiber Bragg Gratings Utilizing Finite Element and Eigenmode Expansion Methods. *Sensors* **2014**, *14*, 10876–10894. [CrossRef] [PubMed]
4. Diab, M.; Minardi, S. Modal analysis using photonic lanterns coupled to arrays of waveguides. *Opt. Lett.* **2019**, *44*, 1718–1721. [CrossRef] [PubMed]
5. Sweeney, D.; Norris, B.R.; Tuthill, P.; Scalzo, R.; Wei, J.; Betters, C.H.; Leon-Saval, S.G. Learning the lantern: Neural network applications to broadband photonic lantern modeling. *J. Astron. Telesc. Instrum. Syst.* **2021**, *7*, 028007. [CrossRef]
6. Wu, W.; Huang, Q.; Wang, C.; Huang, Y.; Li, Y.; Chen, L.; Zhao, J.; Ertman, S.; Halendy, M.; Xu, O.; et al. All-optical-fiber photonic lantern operates from O-band to 2 µm wavelength band and beyond. *Opt. Laser Technol.* **2025**, *192*, 113850. [CrossRef]
7. Taras, A.K.; Norris, B.R.; Betters, C.; Ross-Adams, A.; Tuthill, P.G.; Wei, J.; Leon-Saval, S. Illuminating the lantern: Coherent, spectro-polarimetric characterization of a multimode converter. *Opt. Express* **2026**, *34*, 1012–1025. [CrossRef] [PubMed]
8. Boyd, G.D.; Gordon, J.P. Confocal multimode resonator for millimeter through optical wavelength masers. *Bell Syst. Tech. J.* **1961**, *40*, 489–508. [CrossRef]
9. Kogelnik, H.; Li, T. Laser beams and resonators. *Appl. Opt.* **1966**, *5*, 1550–1567. [CrossRef]
10. Arnaud, J.A.; Kogelnik, H. Gaussian light beams with general astigmatism. *Appl. Opt.* **1969**, *8*, 1687–1693. [CrossRef]
11. Ganiel, U.; Hardy, A. Eigenmodes of optical resonators with mirrors having Gaussian reflectivity profiles. *Appl. Opt.* **1976**, *15*, 2145–2149. [CrossRef]
12. Casperson, L.W. Beam modes in complex lenslike media and resonators. *J. Opt. Soc. Am.* **1976**, *66*, 1373–1379. [CrossRef]
13. Gloge, D. Weakly guiding fibers. *Appl. Opt.* **1971**, *10*, 2252–2258. [CrossRef]
14. Paiano, G.; Pellicoro, M. Propagation constant of weakly guiding optical fibers: A new eigenvalue condition. *J. Light. Technol.* **2001**, *19*, 1592. [CrossRef]
15. Kogelnik, H.; Ramaswamy, V. Scaling rules for thin-film optical waveguides. *Appl. Opt.* **1974**, *13*, 1857–1862. [CrossRef] [PubMed]
16. Collins, S.A., Jr. Lens-system diffraction integral written in terms of matrix optics. *J. Opt. Soc. Am.* **1970**, *60*, 1168–1177. [CrossRef]
17. Streda, P.; Kucera, J.; MacDonald, A. Edge states, transmission matrices, and the Hall resistance. *Phys. Rev. Lett.* **1987**, *59*, 1973. [CrossRef]
18. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [CrossRef]
19. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
20. Mathew, N.M.; Grüner-Nielsen, L.; Castaneda, M.A.U.; Rottwitt, K. A Novel Fabrication Method for Photonic Lanterns. In *Proceedings of the 2018 Optical Fiber Communications Conference and Exposition (OFC)*; IEEE: Piscataway, NJ, USA, 2018. [CrossRef]
21. Leon-Saval, S.G.; Argyros, A.; Bland-Hawthorn, J. Photonic lanterns: A study of light propagation in multimode to single-mode converters. *Opt. Express* **2010**, *18*, 8430–8439. [CrossRef]
22. Brambilla, G.; Finazzi, V.; Richardson, D.J. Ultra-low-loss optical fiber nanotapers. *Opt. Express* **2004**, *12*, 2258–2263. [CrossRef]
23. Gross, S.; Withford, M.J. Ultrafast-laser-inscribed 3D integrated photonics: Challenges and emerging applications. *Nanophotonics* **2015**, *4*, 332–352. [CrossRef]
24. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
25. Han, X.; Papyan, V.; Donoho, D.L. Neural collapse under mse loss: Proximity to and dynamics on the central path. *arXiv* **2021**, arXiv:2106.02073.
26. Montoya, J.; Aleshire, C.; Hwang, C.; Fontaine, N.K.; Velázquez-Benítez, A.; Martz, D.H.; Fan, T.Y.; Ripin, D. Photonic lantern adaptive spatial mode control in LMA fiber amplifiers. *Opt. Express* **2016**, *24*, 3405–3413. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.