

Article

A Pre-Compensation Method for Post-Weld-Shift in Butterfly Semiconductor Laser Packaging

Jinwen Peng ¹, Lian Duan ^{1,2,*}  and Ji-An Duan ¹¹ State Key Laboratory of Precision Manufacturing for Extreme Service Performance, College of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China² College of Mechanical Engineering, Hunan University of Science and Technology, Xiangtan 411201, China

* Correspondence: lianduan@csu.edu.cn

Abstract

Post-weld shift (PWS) significantly degrades the coupling efficiency and consistency of butterfly semiconductor lasers during automated packaging. This study proposes a novel pre-compensation method based on a Particle Swarm Optimization–Backpropagation (PSO–BP) neural network to mitigate these effects. Finite element simulations and experimental analyses reveal that PWS consistently presented negative Y-direction displacement, driven by localized thermal contraction and residual stress. Results indicate that PWS magnitude is highly sensitive to laser energy and welding joint layout. By integrating PSO, the neural network achieved global convergence in just 25 iterations—15 fewer than the conventional BP model—with superior prediction accuracy. Implementation of this pre-compensation framework into an automated packaging system reduced the welding offset range from 0.5–4 μm to 0.5–2 μm , yielding a 26% average increase in optical coupling efficiency. This research provided a high-performance predictive framework suitable for the high-precision mass production of optoelectronic devices.

Keywords: post-weld shift; laser diode coupling; swarm optimization algorithm

1. Introduction

Butterfly laser diode modules are widely used in high-speed optoelectronic systems, including fiber optic communications, optical amplification and fiber sensing, owing to their high output power, excellent stability, and strong reliability [1–3]. However, in contrast to the stringent performance requirements of these devices, their packaging processes remain heavily reliant on manual expertise. This method results in low packaging efficiency, elevated manufacturing costs and inconsistent product quality, creating a significant bottleneck that impedes the scalable commercialization of advanced optoelectronics. During the packaging of butterfly semiconductor lasers, the fixation of the fiber pigtail predominantly utilizes laser welding. The laser welding process inputs high instantaneous energy, causing rapid localized melting and solidification at the joints [4]. This intense thermal cycling generates steep temperature gradients and non-uniform plastic deformation. Consequently, residual stresses propagate through the saddle clip and ferrule, displacing the optical fiber from its aligned position. This misalignment, termed post-weld shift (PWS), severely degrades the device's optical coupling efficiency [5,6].

The energy and direction of PWS are dictated by a highly coupled matrix of process parameters, including weld joint layout, laser pulse characteristics, initial clamping



Received: 25 May 2026

Revised: 14 June 2026

Accepted: 18 June 2026

Published: 30 June 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

configuration, welding sequence, and material geometry. Because these parameters interact synergistically, isolating individual variables for independent control and analysis is exceptionally difficult. For example, the heat input distribution across different weld joints alters the local temperature field and subsequent plastic shrinkage, leading to highly variable deformation vectors. Even a marginal fluctuation in laser energy (as small as 5–10%) can trigger a nonlinear transition in the residual stress field [7,8]. Furthermore, the clamping positions of the fixtures significantly alter the heat conduction paths, further complicating the final trajectory of the PWS. Consequently, PWS control is a complex, multi-physics prediction challenge rather than a linear problem solvable through single-parameter optimization. Currently, PWS mitigation strategies are broadly classified into two categories. The first focuses on the direct suppression of welding deformation through process optimization—such as reducing heat source energy, modifying weld joint layouts, and enhancing heat dissipation paths [9]. While these empirical adjustments can incrementally reduce PWS, the transient melt pool dynamics in laser welding exhibit pronounced nonlinearity and irreproducibility. Because minor variations in material properties or clamping conditions can render these optimized parameters entirely ineffective during mass production, relying solely on process parameter tuning cannot fundamentally resolve the PWS bottleneck.

To address the inherent limitations of process optimization, pre-welding compensation has emerged as a robust alternative widely adopted in high-precision industrial packaging. During the pre-welding phase, the ferrule position is intentionally adjusted slightly in the opposite direction of the expected offset. This deliberate pre-emptive alignment ensures that the inevitable post-welding displacement, resulting from thermal contraction and deformation, is precisely compensated. As a consequence, the optical fiber returns to the optimal coupling point upon completion of the welding process. The efficacy of the pre-compensation strategy is fundamentally dictated by the predictive accuracy of the underlying algorithm, which must map the complex relationships between input process parameters and the resulting displacement vector.

In recent years, the rapid advancement of artificial neural networks (ANNs) has revolutionized the field of optoelectronic manufacturing [10–12]. Due to their superior capacity for nonlinear mapping and high-dimensional data processing, NNs have become the focal point for modeling manufacturing processes, predicting residual stresses, and automating optical alignment. Despite the potential of data-driven approaches, the requirement for massive experimental datasets remains a prohibitive barrier to their implementation in precision laser packaging. Specifically, in butterfly laser coupling, the exorbitant cost of micron-level measurements and the inherent variability of material properties render large-scale data acquisition practically infeasible. To address the critical challenge of high-precision post-weld shift (PWS) prediction within data-scarce environments, this paper proposes an improved PSO-BP neural network. While conventional BP networks are highly sensitive to initial parameters and prone to local minima, the integration of the improved PSO algorithm globally optimized the network's initial weights and thresholds. Experimental results substantiate the superiority of the proposed model under limited sample conditions. Compared to traditional BP networks, the PSO-BP model achieved a remarkable reduction in prediction error. This robust predictive capability translates to a significant improvement in coupling efficiency, increasing from 68% to 82%, demonstrating its immense practical value in high-precision laser packaging.

2. Post-Weld-Shift Analysis

2.1. Optical Coupling Analysis

Prior to the laser welding packaging process, the optical coupling and alignment between the butterfly laser module and the single-mode fiber is typically performed. In edge-emitting laser-to-SMF coupling, the optical field at the fiber facet manifests as a hybrid profile rather than a pure Gaussian or uniform geometric beam. More precisely, the intensity distribution is modeled as a linear combination of a Gaussian distribution and the circular-aperture diffraction pattern, with each component assigned specific weighting coefficients. The expression is as follows [13]:

$$\begin{cases} \rho(x, y) = \sqrt{\left(\frac{x}{w_x}\right)^2 + \left(\frac{y}{w_y}\right)^2} \\ E(x, y) = E_0 \exp[-\rho^{2m}] \end{cases} \quad (1)$$

where m is the order of the super-Gaussian function, where a higher value results in a flatter central region and steeper edges.

The extent of the uniform area at the center of the light field can be modulated by adjusting the magnitude of m . In comparison with conventional Gaussian beams, flat-top super-Gaussian beams exhibit a more uniform intensity distribution at the center and a steeper transition at the edges. Such a profile more closely approximates the actual output optical field of optical fibers, thereby enhancing coupling efficiency. Figure 1a presents the numerical simulation results of the optical field at the optical fiber end face. The output field exhibits an approximately elliptical intensity distribution in the transverse cross-section, characterized by a relatively flat central region and a gradual roll-off at the edges, forming a structure analogous to an elliptical flat-top spot.

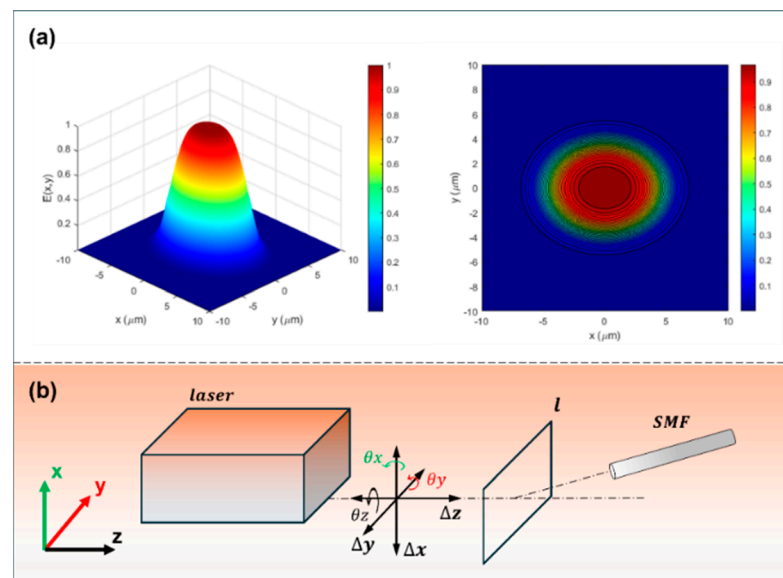


Figure 1. (a) Numerical simulation results of the optical field at the optical fiber end face; (b) the diagram of optical coupling between a semiconductor laser and an optical fiber.

Theoretically, fiber optic coupling requires precise adjustment of six degrees of freedom: three translational degrees along the x -, y - and z -axes, and three rotational degrees around the θ_x -, θ_y - and θ_z -axes. Given that the rotational displacement θ_z about the z -axis in single-mode fiber coupling is generally negligible in most scenarios, in this paper, we focused exclusively on the influence of the remaining five degrees of freedom on the fiber

coupling efficiency. Figure 1b shows the optical coupling between a semiconductor laser and an optical fiber.

The coupling efficiency of the laser and the single-mode fiber on plane l can be calculated by the following equation:

$$\left\{ \begin{array}{l} \eta_x = k_x \exp \left\{ -k_x \left[\frac{d_x^2}{2} \left(\frac{1}{r_{x0}^2} + \frac{1}{r_{f0}^2} \right) + \frac{\pi^2 \theta_x^2}{2\lambda^2} (r_x^2(z) + r_{f0}^2) - \frac{d_x \theta_x z}{r_{x0}^2} \right] \right\} \\ r_x^2(z) = r_x^2 \left(1 + \left(\frac{\lambda z}{\pi r_{x0}^2} \right)^2 \right) \\ k_x = 4r_{x0}^2 r_{f0}^2 / \left[(r_{x0}^2 + r_{f0}^2)^2 + \left(\frac{\lambda z}{\pi} \right)^2 \right] \end{array} \right. \quad (2)$$

$$\left\{ \begin{array}{l} \eta_y = k_y \exp \left\{ -k_y \left[\frac{d_y^2}{2} \left(\frac{1}{r_{y0}^2} + \frac{1}{r_{f0}^2} \right) + \frac{\pi^2 \theta_y^2}{2\lambda^2} (r_y^2(z) + r_{f0}^2) - \frac{d_y \theta_y z}{r_{y0}^2} \right] \right\} \\ r_y^2(z) = r_y^2 \left(1 + \left(\frac{\lambda z}{\pi r_{y0}^2} \right)^2 \right) \\ k_y = 4r_{y0}^2 r_{f0}^2 / \left[(r_{y0}^2 + r_{f0}^2)^2 + \left(\frac{\lambda z}{\pi} \right)^2 \right] \end{array} \right. \quad (3)$$

$$\eta_{ttl} = \eta_x \cdot \eta_y \quad (4)$$

where η_x and η_y are the coupling efficiencies in the x and y directions, respectively. r_{f0} is the radius of the mode field of the single-mode fiber (SMF), λ is the wavelength of the laser diode (LD), and z is the displacement error of the Z-axis. d_x (d_y) is the linear deviation of the center of the LD and SMF mode field in the vertical junction plane, θ_x (θ_y) represents the center of the LD mode field angle deviation from the center of the SMF mode field on the vertical junction plane, and r_{x0} (r_{y0}) is the mode field radius of the LD in the vertical junction plane.

To quantitatively analyze the alignment-dependent coupling efficiency of butterfly-type laser diodes and investigate the influence of various misalignments, including lateral, longitudinal and angular deviation, we designed a series of experiments wherein coupling efficiency is measured under varying misalignment conditions by adjusting the alignment errors between the laser and the optical fiber. The coupled system employed a step-scanning method, whereby the coupling platform drives the receiving optical fiber to scan along a specified direction while measuring the variation curve of received optical power along this axis. The ultra-precision motion stage achieves the repeat positioning accuracy of 0.5 μm and a resolution of 0.1 μm . In the experiment, a 1 V electrical signal was applied via an external power source, which theoretically translates into 1 mW of optical power through the butterfly laser. Another motion platform is used for holding single-mode optical fibers, and its motion direction is 5-DOF. An optical power meter with nanowatt-level resolution is connected to the terminal of the optical fiber to provide feedback on optical signal coupling alignment. Prior to scanning plane l , the laser is adjusted to its maximum power position along the z-axis.

Figure 2 shows the optical power distribution of XYZ directions. It can be seen that when the deviation in the X or Y direction exceeds 10 μm , the received power basically drops to 0 mW. Within the range of $\pm 1 \mu\text{m}$, the power ascent/descent gradient becomes significantly smaller, approximating a plateau. This phenomenon arises because the optical field before entering the fiber is generated by a laser with a beam quality factor $M^2 > 1$. Through the combined effects of collimating and focusing lenses, a weakly top-hat-shaped optical field, deviating from an ideal Gaussian profile, is formed [14]. This observation also corroborates the laser model we propose in Figure 1a.

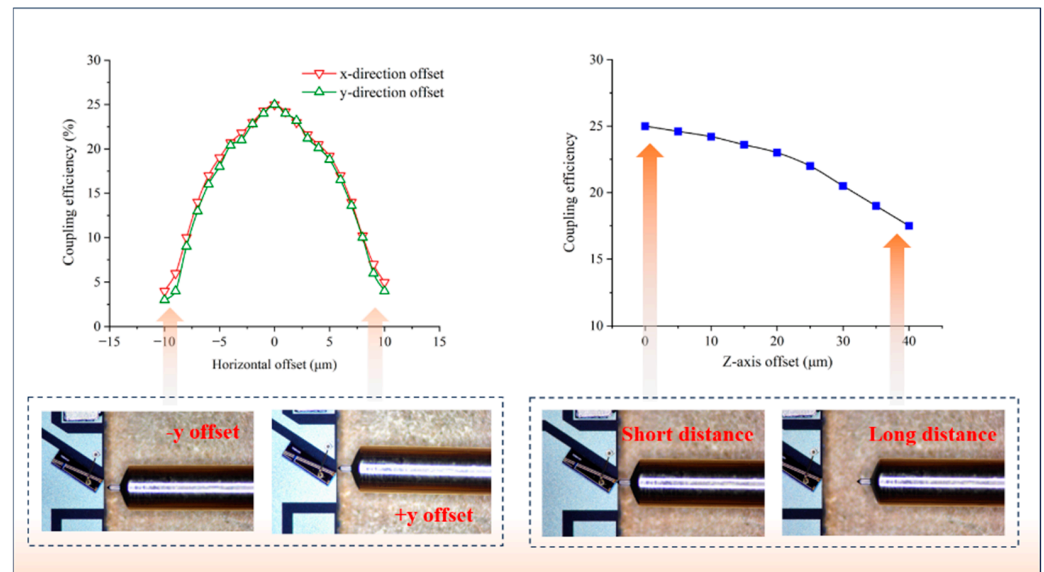


Figure 2. The influence of linear offset on coupling efficiency: x, y and z.

Figure 3 shows the optical power distribution of $\theta_x\theta_y$ directions. When the angle deviation exceeds 10° , the received optical power basically drops to 0 mW. By comparing the different directions, we found that the optimal coupling power in angle offset is lower than that in linear offset. This is because when the coupling platform is scanned along the angular direction, it introduces additional linear direction deviations, causing the originally well-coupled XYZ position to deviate from the optimal coupling position [15]. The analysis of the optical power distribution correlates the alignment offset with power loss. Through these experimental results, we obtained a comprehensive dataset of offset–power pairs. Then, we need to fix the angle of the angle-sensitive single-mode optical fiber to limit the minor disturbances during the packaging process.

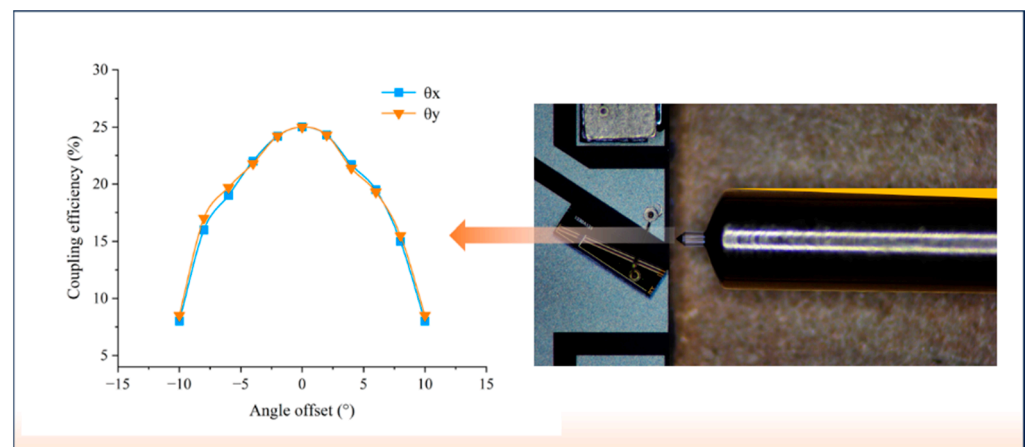


Figure 3. The influence of angular deviation on coupling efficiency: θ_x and θ_y .

2.2. Post-Weld-Offset Analysis

After the optical modules coupling, the components should be fixed. The butterfly laser module comprises a Kovar alloy housing, a thermoelectric cooler, a semiconductor laser chip, coupled optical fibers, and other fiber optic components (presented in Figure 4a). In order to facilitate the fixation of optical fibers, some mechanical components are also required, such as metal platforms, optical fiber tubes and saddle clamps. The saddle clamps are responsible for clamping and positioning the fiber optic sleeve, while constraining the

component displacement along the y direction. The laser welding procedure is as follows (presented in Figure 4b): first, the semiconductor laser chip is soldered onto a heat sink, followed by the fiber optic coupling process. Upon reaching the maximum power coupling point, the saddle clamp is welded onto the metal platform using a laser heat source, and finally, the fiber ferrule is welded onto the saddle clamp.

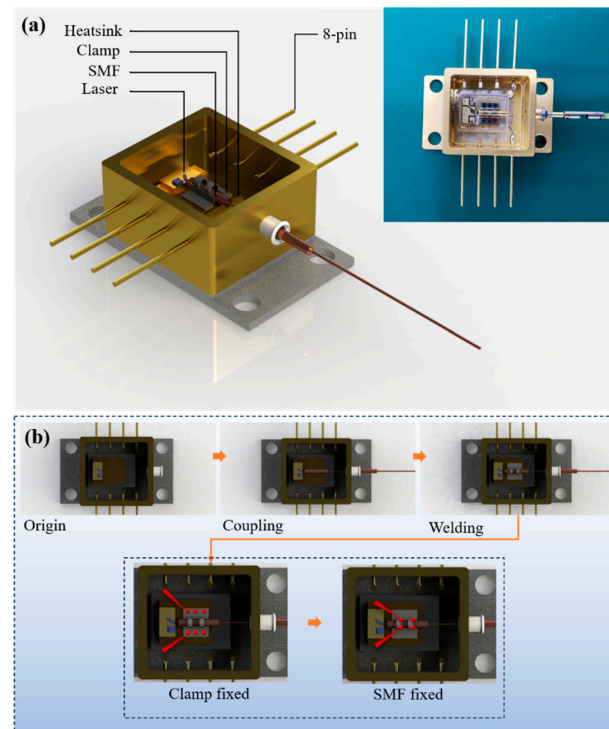


Figure 4. (a) The butterfly-shaped laser architecture and components; (b) the processes of butterfly-shaped laser module packaging.

Laser welding is a typical transient thermal process. During this transient process, the material will undergo solid-to-liquid, liquid-to-gas, and then reverse-solidification processes. Its temperature field satisfies the three-dimensional transient heat conduction equation [16]:

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q \quad (5)$$

where ρ is the density, c is the special heat of materials; k represents the thermal conductivity; and Q is the heat source, which is described by a Gaussian distribution model, and its surface heat flux density is expressed as

$$q(r_l) = \frac{3Q_0}{\pi r_{l0}^2} \exp\left(-3 \frac{r_l^2}{r_{l0}^2}\right) \quad (6)$$

where Q_0 is the efficient absorption of energy, and r_{l0} is the spot radius of the welding laser. The absorption of energy Q_0 is jointly determined by the pulse energy and the material absorption rate:

$$Q_0 = \eta_{weld} E_p \quad (7)$$

In the boundary condition setting, the initial welding temperature is set to the ambient temperature (25 °C). On the surface of the workpiece, there is intense heat exchange between the high-temperature metal and the surrounding environmental gases, including convective heat transfer and thermal radiation. Through a transient solution, the distribution characteristics of the temperature near the weld point over time can be obtained. The

element type for thermal analysis is selected as SOLID 90 (a 20-node 3D thermal analysis element with nonlinear thermal conduction, convection and material), and the element type for structural analysis is selected as SOLID 45 (an 8-node 3D structural analysis element with large plastic deformation and nonlinear material). All the nodes beneath the saddle clamp are assumed to be rigidly connected to the base, so their displacements and selected degrees of freedom are all constrained. Additionally, considering that the two laser beams are symmetrically and simultaneously welded from both sides during the welding process, it is approximately assumed that the post-welding stress-induced offset in the X direction is zero. In the simulation, only the post-welding offset in the Y direction is considered. In the simulation, we adopted the conduction method of thermal flow radiation for the treatment of the laser heat source. The thermal flow field adopts the conventional Gaussian distribution, so the laser heat source model in the loading conditions is

$$q(r) = \frac{3Q}{\pi r_0^2} \exp \left[-3 \left(\frac{r}{r_0} \right)^2 \right] \quad (8)$$

where r is the radius of the laser beam spot, which is equal to 500 μm . The radius of the focused spot r_0 is equal to 80 μm . Q is the energy transferred to the working surface. The laser pulse duration is 5 ms, and the material is 304 L stainless steel. During laser welding, not all the energy can be absorbed by the welding material, which means there is a problem of laser energy absorption efficiency. Here, the absorption efficiency is taken as 35%, and the energy Q transferred to the working surface is $1.25 \times 0.35 \text{ J}$ (the total energy generated by the laser beam $\times$ absorption efficiency). We took 100 s as the process time from the generation to the cooling to room temperature of a laser weld point and recorded the corresponding parameter change curves over time. If a second welding is to be carried out, it should also be done after 100 s when the weld point temperature has cooled to room temperature. Figure 5 shows the post-weld deformation distribution of the fiber optic sleeve and saddle clamp in the Y direction after the laser heat source was applied to the model.

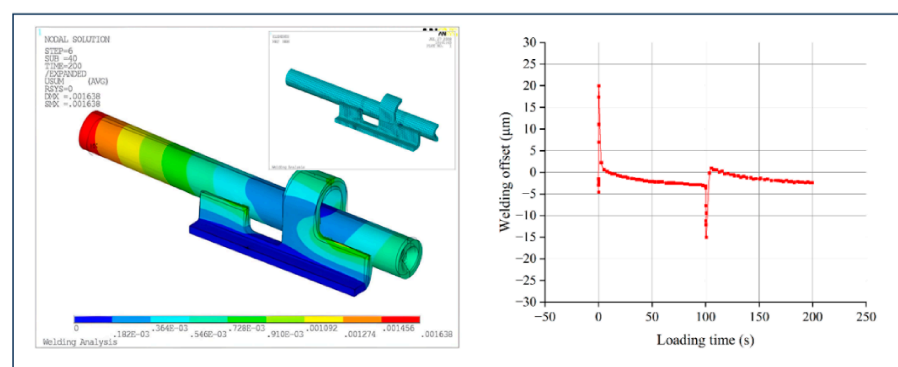


Figure 5. The finite element geometric model and analysis results: The relationship between laser loading time and welding offsets.

After it cooled down to room temperature, we loaded the laser heat source at the next weld point and recorded the post-weld shift generated at the front end of the optical fiber sleeve in the Y direction throughout the entire process. It can be clearly seen from Figure 5 that the front end of the optical fiber sleeve has a displacement along the negative Y-axis, which is consistent with the phenomenon of the sleeve sinking after welding in the actual welding process. When the laser irradiates the working surface, due to the instantaneous melting of the weld point, the fiber optic sleeve will produce a relatively large displacement along the positive Y-axis and then recover instantly. As the weld point gradually cools

down, the post-weld displacement will gradually decrease to a negative value. In the figure, when the weld point instantaneously melts, the post-weld displacement reaches $20.4\ \mu\text{m}$, and after cooling to room temperature, the post-weld displacement becomes $-3.16\ \mu\text{m}$. When the second weld point cools down completely, the final post-weld displacement becomes $-2.61\ \mu\text{m}$. The simulation results show that after the second welding is completed, the post-weld displacement is actually smaller than that after the first welding.

2.3. The Influence of Welding Parameters

The welding parameters consist of the welding position and laser energy. At the moment of welding, the front end of the sleeve initially jumps up briefly, and then, under the influence of cooling contraction and residual stress, it turns to negative displacement, ultimately sinking in the negative direction of the Y-axis. Each pair of weld points, upon formation, generates a contraction force during the transition from a molten to a solid state, which acts on the sleeve and induces downward displacement along the negative Y-axis; i.e., post-weld distortion consistently manifests in the negative Y direction. When the first pair of weld points is formed at the front end, the resulting contraction force lowers the front section of the sleeve while causing the rear end to tilt upward. Subsequently, when the second pair of weld points is created at the rear end, the contraction force sinks the rear section while elevating the front end. This process effectively compensates for the post-weld distortion induced by the preceding welding step.

We continued to use the finite element mesh model obtained previously, with all welding conditions remaining unchanged. First, we completed the first welding, and after the cooling process was over, we carried out the second welding. Five simulation calculations were conducted, with only the lateral distance d between the second pair of weld points and the first pair of weld points being changed each time. In the five calculations, d takes the values of 0.5 mm, 0.75 mm, 1 mm, 1.25 mm, 1.5 mm, 1.75 mm and 2.0 mm, respectively. Figure 6a shows the simulation results: as the distance between the two consecutive weld points increases, the compensation effect for post-weld shift does indeed decrease. When $d = 0.5\ \text{mm}$, the final post-weld shift is compensated to $-1.17\ \mu\text{m}$, while when $d = 2.0\ \text{mm}$, the final post-weld shift can only be compensated to $-2.41\ \mu\text{m}$.

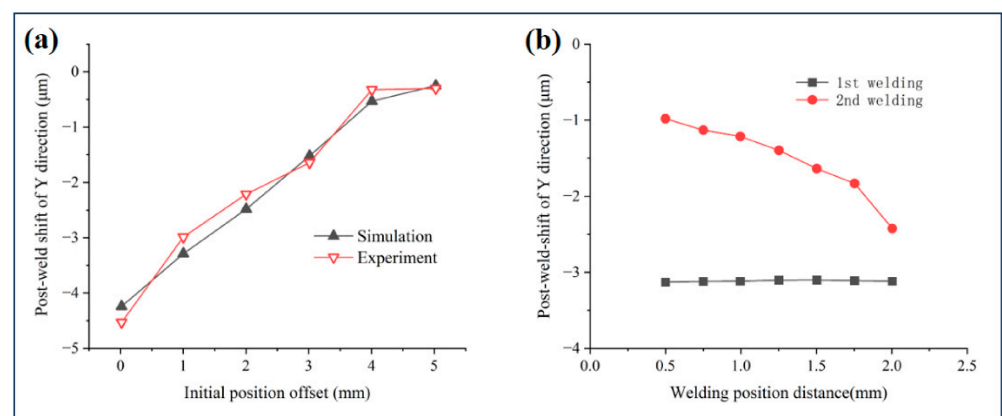


Figure 6. The influence of welding process parameters on the post-weld shifts of the butterfly laser. (a) The relationship between offsets and initial position offsets; (b) The relationship between offsets and Welding position distances.

The magnitude of the laser energy in welding will also directly affect the size of the post-welding offset. Using the same finite element mesh model obtained previously and keeping all other welding conditions unchanged, different laser energies were selected to weld the first pair of weld points. After the temperature cooled to room temperature,

the current post-weld shift was recorded. Since the Nd: YAG laser used for welding has six selectable energy levels, six different welding energies were accordingly set in the simulation. Figure 6b shows the results of the simulation. The simulation calculation results show that the post-weld shift significantly decreases as the welding energy reduces. Therefore, setting the welding laser energy reasonably can effectively suppress the post-weld shift.

The simulation results indicate that if the welding parameters and ambient conditions remain relatively stable, the resultant post-welding displacement will also exhibit consistency, implying a direct and stable mapping relationship between the welding parameters/environment and the displacement. Based on this, if the downward displacement of the fiber sleeve after welding, i.e., the post-welding offset, can be accurately predicted prior to the welding process using the current parameters, the desired position of the fiber sleeve can be pre-adjusted upward by the same magnitude. In this manner, the displacement induced by welding will precisely offset the pre-adjusted distance, thereby allowing the fiber to automatically return to its optimal coupling position.

3. Pre-Compensation Method of Post-Weld-Shift

3.1. Pre-Compensation System

The platform for conducting the post-welding offset measurement experiment of the laser welding machine was the LW4200 automatic laser welding machine of Shenzhen Lianying Laser Co. A high-precision laser displacement sensor was employed in the experiment to perform real-time measurement of the post-welding deviation in the Y direction in Figure 7. The laser displacement sensor adopted was the LK series semiconductor laser displacement sensor from Keyence Corporation of Japan. It consists of a light-emitting element and a position-sensitive detector (PSD) and uses triangulation to detect the target object. The effective measurement resolution of the Keyence LK series can reach 20 nm, which could meet the post-welding offset measurement requirements of this project.

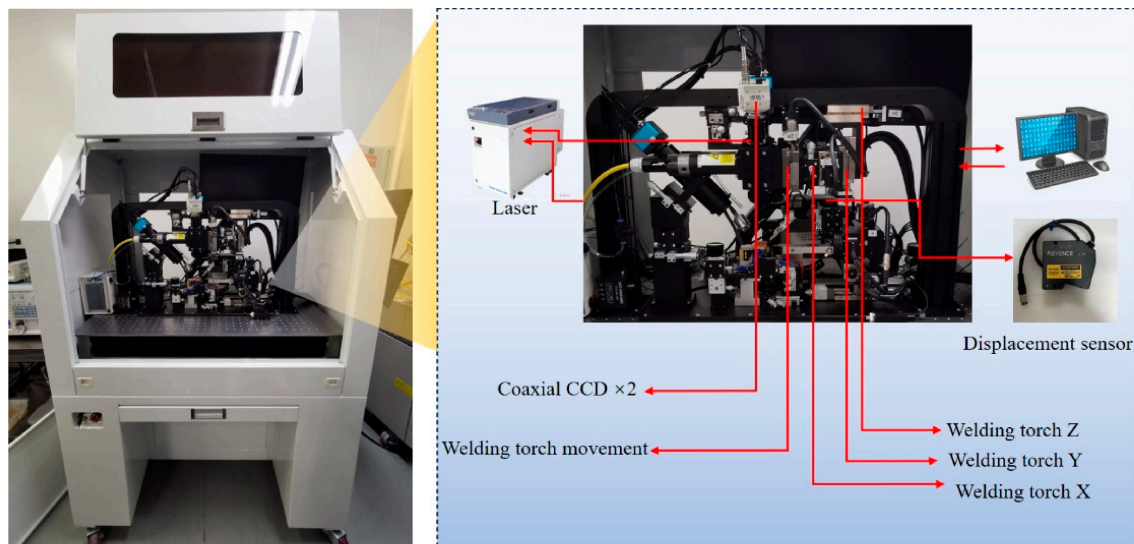


Figure 7. The influence of welding process parameters on the post-weld offset of the butterfly laser device. Butterfly laser automated packaging equipment and related motion platform.

With the aid of the real-time measurement system for post-weld shift, the pre-compensation experiment for post-weld shift was carried out in two steps: (1) We measured the post-weld shift after the first pair of weld points was completed. All adjustable parameters, including laser energy E , pre-offset of the sleeve y_1 , sleeve extension length l ,

and the position of the first pair of weld points z_1 , were assigned different initial values and combined in various ways. Then, the post-weld shift after welding was measured and recorded. (2) On the basis of the previous step, we welded the second pair of weld points again and measured the post-weld shift. The adjustable parameters included laser energy E , the current offset of the sleeve y_2 , the extension length of the sleeve l , the position of the first pair of weld points z_1 , and the position of the second pair of weld points z_2 . The range of parameter values is presented in Figure 8. The improved orthogonal table-based scheme divides the combinations of different process parameters into 36 groups. The experimental results obtained by arranging and combining different parameters are shown in Table 1.

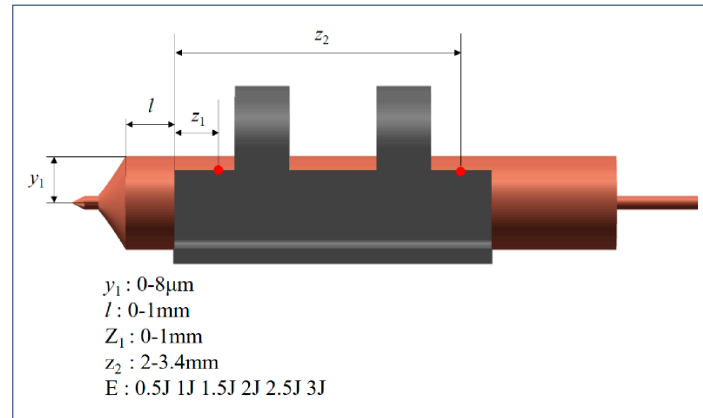


Figure 8. Schematic diagram of symbol representation for welding process parameters of butterfly laser.

Table 1. Experimental results under different parameter combinations.

Case	Laser Energy E (J)	Offset y_1 (μ m)	Extension l (mm)	Weld Spot z_1 (mm)	Weld Spot z_2 (mm)	Post-Weld-Shift d (μ m)
1	0.5	0	0.2	0.2	2.0	−2.14
2	0.5	2	0.4	0.4	2.5	1.05
3	0.5	4	0.6	0.6	3.0	−0.82
4	0.5	6	0.8	0.8	3.5	2.47
5	0.5	8	1.0	1.0	4.0	−1.36
6	0.5	0	0.2	0.4	3.0	0.59
7	1.0	0	0.4	0.6	3.5	−2.71
8	1.0	2	0.6	0.8	4.0	1.83
9	1.0	4	0.8	1.0	2.0	−0.15
10	1.0	6	1.0	0.2	2.5	2.92
11	1.0	8	0.2	0.4	3.0	−1.67
12	1.0	2	0.4	0.6	2.0	0.34
13	1.5	0	0.6	1.0	2.5	−2.48
14	1.5	2	0.8	0.2	3.0	1.21
15	1.5	4	1.0	0.4	3.5	−0.56
16	1.5	6	0.2	0.6	4.0	2.15
17	1.5	8	0.4	0.8	2.0	−0.93
18	1.5	4	0.6	0.8	2.5	0.77
19	2.0	0	0.8	0.4	4.0	−1.89
20	2.0	2	1.0	0.6	2.0	2.33
21	2.0	4	0.2	0.8	2.5	−0.42
22	2.0	6	0.4	1.0	3.0	1.56
23	2.0	8	0.6	0.2	3.5	−2.95
24	2.0	6	0.8	1.0	4.0	0.88
25	2.5	0	1.0	0.8	3.0	−1.12
26	2.5	2	0.2	1.0	3.5	2.64

Table 1. *Cont.*

Case	Laser Energy E (J)	Offset y_1 (μm)	Extension l (mm)	Weld Spot z_1 (mm)	Weld Spot z_2 (mm)	Post-Weld-Shift d (μm)
27	2.5	4	0.4	0.2	4.0	−0.37
28	2.5	6	0.6	0.4	2.0	1.91
29	2.5	8	0.8	0.6	2.5	−2.26
30	2.5	0	1.0	0.2	2.0	0.45
31	0.5	8	0.4	0.6	3.5	−1.54
32	1.0	6	0.2	1.0	4.0	2.78
33	1.5	4	1.0	0.8	3.0	−0.71
34	2.0	2	0.8	0.6	2.5	1.39
35	2.5	0	0.6	0.4	2.0	−2.03
36	1.5	0	0.2	0.2	2.0	0.26

It can be observed from the table that the post-welding offset, denoted as d , is governed by the intricate coupling between thermal input and structural constraints within the butterfly laser packaging process. Laser energy serves as the primary driver of displacement, where elevated energy levels intensify cooling contraction stresses within the melt pool. This deformation is further magnified by the extension length l through a leverage effect, where longer cantilever structures translate minor angular distortions at the weld site into significant terminal offsets. The pre-offset y_1 functions as a compensatory parameter, allowing for the identification of a structural equilibrium point to counteract inherent systemic shrinkage. Furthermore, the spatial configuration of the tandem weld spots (z_1 and z_2) defines the stress distribution; an optimized axial spacing facilitates torque balancing, thereby suppressing parasitic tilting of the ferrule. Consequently, minimizing the PWS requires a strategic trade-off: utilizing a refined low-energy regime while optimizing the y_1 compensation threshold and reducing the mechanical moment arm l to maintain sub-micron optical alignment.

3.2. PSO-BP Neural Network

The PSO-BP neural network is a hybrid machine learning model that deeply integrates swarm intelligence algorithms with traditional artificial neural networks [17–19]. Particle Swarm Optimization does not require gradient information. It encodes the weights and thresholds of the BP network as particles and conducts extensive searches throughout the entire sample space through information sharing of individual historical best and global historical best within the group, thereby providing the network with a set of high-quality initial parameters. After PSO finds an initial solution close to the global optimum, the BP network takes over the training and uses the gradient descent method for high-precision backpropagation fine-tuning to quickly approach the true optimal solution. Therefore, the post-weld shift pre-compensation method based on the PSO-BP neural network was suitably adopted for pre-compensation. The neural network architecture suitable for post-weld shift prediction in this study is shown in Figure 9a. It can be seen that after the PSO search is completed, it hands over the globally optimal position found, that is, the best set of weights and biases, to the BP neural network. The BP network is no longer randomly initialized but stands on the shoulders of giants. It uses the backpropagation algorithm to calculate the gradient and fine-tunes in the vicinity of this excellent initial point until it finds the precise local minimum (the optimal solution).

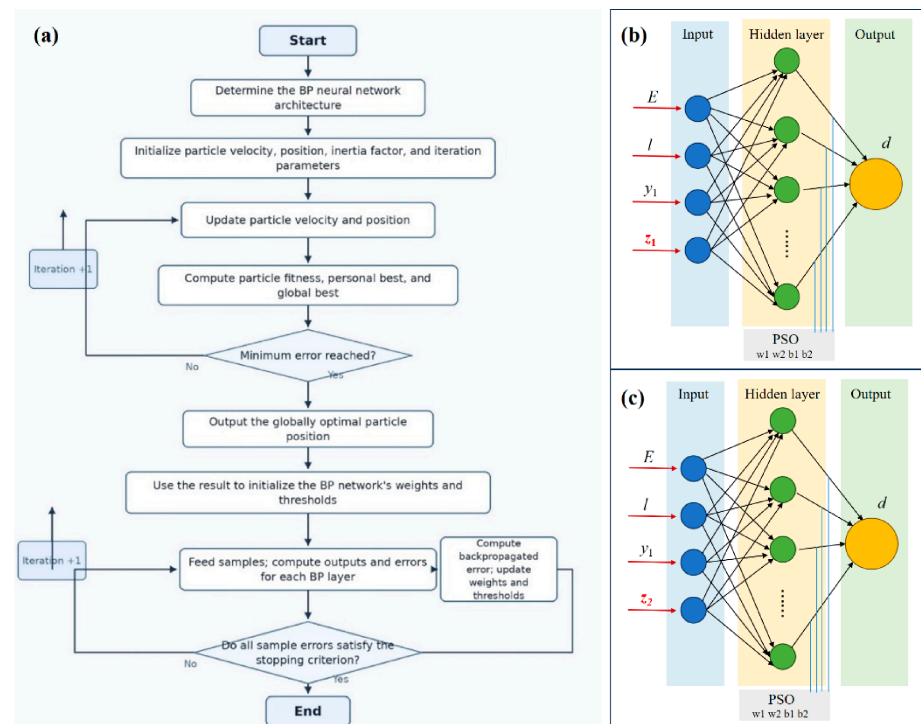


Figure 9. (a) The algorithm architecture for predicting post-weld offset using the PSO-BP network; (b,c) the BP neural network architectures of the first welding and the second welding.

Two neural network models were established to predict the post-weld shift of the butterfly-shaped semiconductor laser (shown in Figure 9b,c). Model 1 (Figure 9b) was used to predict the post-weld shift after the first pair of weld points was completed, and Model 2 (Figure 9c) was used to predict the post-weld shift after the second pair of weld points was completed. The BP neural network used in this study contains one hidden layer with 10 neurons. The number of hidden neurons was initially selected based on an empirical formula and then verified by trial training. The numerical values of the input samples of neural networks vary greatly. If the data are not processed, oscillation phenomena are likely to occur during network training. In this paper, all the data in the sample space were normalized; that is, all the input and output data are mapped to the same range. After the training was completed, they were remapped to the original data range. The hidden layer and output layer respectively adopt the hyperbolic tangent sigmoid function (tansig) and the linear function (purelin) as the transfer functions, and the training algorithm uses the momentum gradient descent method. The hyperparameter settings of the PSO-BP neural networks are shown in Table 2.

Table 2. Hyperparameter settings of the PSO-BP neural networks.

Parameter	Value	Parameter	Value
Model 1 inputs	E, y_1, l, z_1	Momentum coefficient	0.90
Model 2 inputs	E, y_1, l, z_1, z_2	Maximum epochs	1000
Output	d	MSE target	0.01
Number of hidden layers	1	Swarm size	20
Number of neurons in hidden layer	10	Maximum PSO iterations	60
Hidden-layer activation	tansig	Inertia weight w	0.4–0.9
Output-layer activation	purelin	Cognitive coefficient c_1	2
Learning rate	0.01	Social coefficient c_2	2

Figure 10a shows the training error curve of the PSO-BP neural network. The black line represents the mean square error, which is set to 1×10^{-2} . That is, the training stops when the mean square error of the network output reaches or is less than this value. The red curve represents the output of the mean square error after training. As a comparison method, the training curve of the traditional BP neural network is represented in green. It can be seen that basically, after 25 steps, the mean square error of the neural network model can be trained below the set value. This reduces the number of training iterations by 15 steps compared to the traditional BP method. Figure 10b shows the post-weld shift predicted by the trained BP neural network using all the adjustable welding parameters in the sample data as the network input. It can be seen that the trained PSO-BP neural network can simulate the change model of post-weld shift well, and the deviation between the predicted value and the true value of the model is very small.

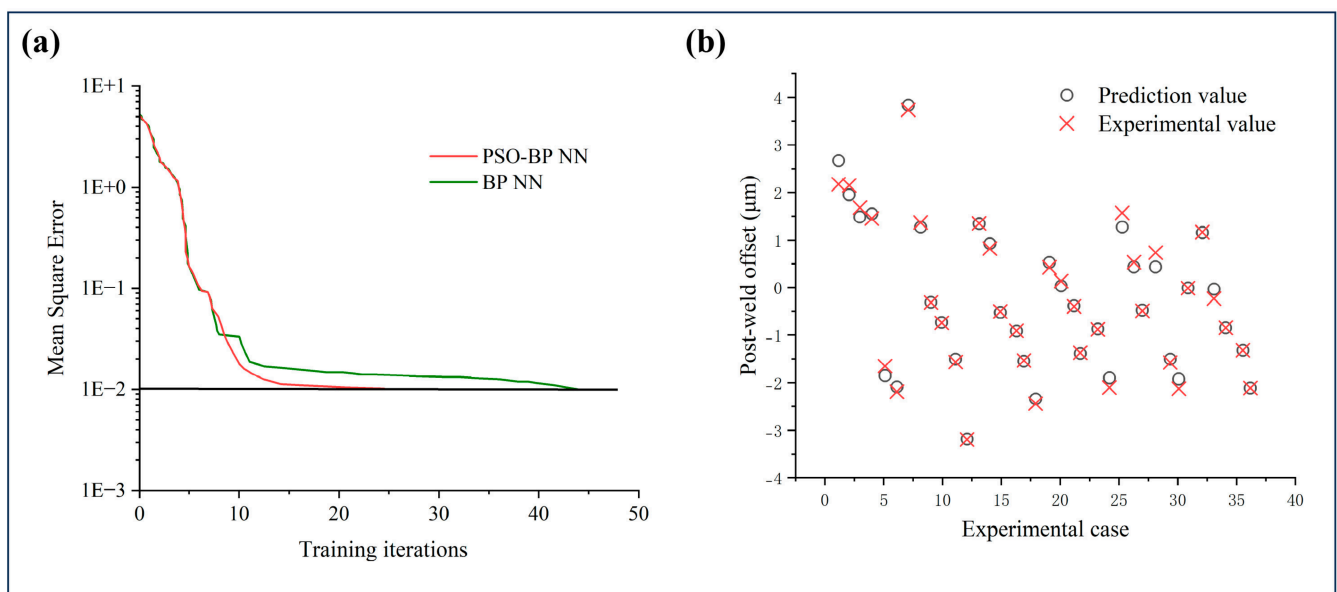


Figure 10. (a) Comparison of optimization performance between PSO-BP neural network and BP neural network; (b) comparison of experimental and predicted results of post-weld offset of butterfly laser diode based on PSO-BP neural network.

3.3. Experimental Results of PWS Compensation

This prediction method was introduced into the automatic packaging system to perform pre-compensation for the packaging of butterfly lasers. Figure 11a presents the statistical results of the post-weld offset for 50 different combinations of process parameters in the automated packaging system. It can be seen that the welding offset distance of the automated packaging system with a pre-compensation module is significantly smaller than that of the packaging system without a pre-compensation module. The offset range decreased from $0.5\text{--}4\text{ }\mu\text{m}$ to $0.5\text{--}2\text{ }\mu\text{m}$. Figure 11b illustrates the improvement in optical power achieved by the butterfly optical device after pre-compensation and automated packaging. The average increase in light coupling efficiency over 50 experiments (the difference between the coupled efficiency of the light device with pre-compensation and that of the device without pre-compensation) was 26%.

The detailed experimental data and lateral comparison results show that, compared with the automated packaging systems that rely on traditional fixed bias or do not use the pre-compensation method at all, the system introducing the PSO-BP pre-compensation algorithm successfully broke through the process bottleneck caused by complex welding thermal deformation. The spatial pose drift of the devices after laser welding thermal

shock was extremely effectively suppressed, and the transient drop amplitude of post-weld optical power was significantly reduced. Overall, not only did the post-weld optical performance of the devices achieve a significant and stable improvement, but the yield and batch-to-batch consistency in the mass production process also showed a qualitative leap. This fully confirms the huge application potential of this model in automated packaging production lines of high-end semiconductor lasers.

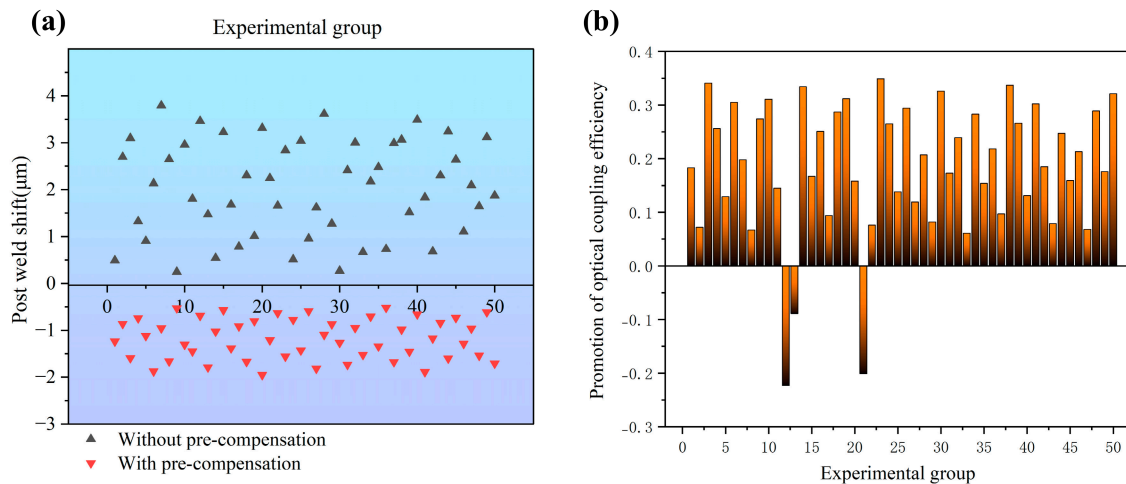


Figure 11. (a) The inhibitory effect of pre-compensation on post-weld offset in automated packaging; (b) the promotion of optical coupling efficiency via pre-compensation.

4. Conclusions

In this study, a novel pre-compensation method based on the PSO-BP neural network is presented to mitigate post-weld shift in butterfly semiconductor laser packaging. Through systematic investigation, the following conclusions are drawn:

- (1) Finite element simulation and experimental analysis reveal that PWS consistently manifests as a negative Y-direction displacement due to localized thermal contraction and residual stress.
- (2) The magnitude of the offset is highly sensitive to laser energy and welding joint layout. Specifically, increasing the axial spacing between tandem weld spots decreases the compensation effect, while reducing laser energy effectively suppresses the initial deformation.
- (3) The integration of Particle Swarm Optimization allows the BP neural network to achieve global convergence with limited experimental samples. The proposed PSO-BP model converged in 25 iterations, 15 steps fewer than the traditional BP method, with significantly lower mean square error.
- (4) The introduction of a pre-compensation module into the automated packaging process narrowed the welding offset to within 2 μm and achieved a 26% improvement in average coupling efficiency. These results confirm that the prediction-driven compensation strategy effectively mitigates packaging-induced losses and ensures high-consistency optical power output for butterfly lasers.

This paper offers a high-performance predictive framework that is highly applicable to the mass production of high-end optoelectronic devices.

Author Contributions: Conceptualization, L.D. and J.P.; methodology, J.P.; software, J.P.; validation, L.D. and J.-A.D.; investigation, J.P.; resources, J.-A.D.; data curation, J.P.; writing—original draft preparation, J.P.; writing—review and editing, J.P. and L.D.; funding acquisition, J.-A.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (NO. 2024YFF0504904) and the APC was funded by the National Natural Science Foundation of China (Grant No. U22A20200).

Data Availability Statement: The data can be made publicly available to the researchers who need it. Please contact the corresponding author to obtain it.

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

References

1. Song, J.H.; O'Brien, P.; Peters, F.H. Optimal laser welding assembly sequences for butterfly laser module packages. *Opt. Express* **2009**, *17*, 16406–16414. [[CrossRef](#)] [[PubMed](#)]
2. Hsu, Y.C.; Lin, S.H.; Tsai, H.H.; Hong, S.P. Experiment and Simulation on Reduction of Postweld Shift on Butterfly-Type Laser Module by Nd:YAG Laser Welding and Separated Dual Clips. *Fiber Integr. Opt.* **2014**, *33*, 315–324.
3. Zhang, Z.K.; Liu, Y.; Bai, J.H.; Yuan, H.Q.; Zhao, Z.P.; Liu, J.G.; Zhu, N.H. Ultra-wideband Butterfly Directly Modulated Semiconductor Lasers. *IEEE Photonics J.* **2017**, *9*, 1503309. [[CrossRef](#)]
4. Liu, H.N.; Wang, Y.J.; Zhang, J.Z.; Zhu, Z.Q.; Yang, S.L. Effect of laser spot offset on microstructure and properties of laser welded joint of medium Mn/DP980 dissimilar steel. *Mater. Lett.* **2024**, *363*, 13633. [[CrossRef](#)]
5. Duan, L.; Luo, D.H.; Zhong, S.S.; Zhang, F.; Zhou, H.B.; Duan, J.A. Post-weld-shift compensation strategy with improved U-net edge detection for reducing coupling loss of 25Gbps TOSA. *Opt. Express* **2025**, *33*, 12852–12867. [[PubMed](#)]
6. Wang, Z.Y.; Zhou, H.B.; Zhou, W.W.; Li, Z.Q.; Ju, X.; Peng, Y.C.; Duan, J.A. Experimental and numerical investigation of residual stress and post-weld-shift of coaxial laser diodes during the optoelectronic packaging process. *Weld. World* **2023**, *67*, 63–67.
7. Xiao, Z.X.; Qi, Y.; Chen, C.P.; Nie, X.J.; Wang, J.; Zhu, H.H. Analytical model of predicting residual stress-process parameter relationship for laser powder bed fusion of Ti6Al4V. *J. Mater. Res. Technol.* **2025**, *35*, 6372–6382. [[CrossRef](#)]
8. Ma, N.S.; Li, L.Q.; Huang, H.; Chang, S.; Murakawa, H. Residual stresses in laser-arc hybrid welded butt-joint with different energy ratios. *J. Mater. Process. Technol.* **2015**, *220*, 36–45.
9. Sheen, M.T.; Liu, C.C.; Ding, Z. Novel low-cost high-speed optic-electric laser diode pigtail module packaging technology. *Opt. Fiber Technol.* **2022**, *72*, 102976.
10. Yu, L.; Saida, K.; Hirano, S.; Chigusa, N.; Mochizuki, M.; Nishimoto, K. Application of neural network-based hardness prediction method to HAZ of A533B steel produced by laser temper bead welding. *Weld. World* **2017**, *61*, 483–498.
11. Huang, W.W.; Gao, X.D.; Huang, Y.H.; Zhang, Y.X. Improved Convolutional Neural Network for Laser Welding Defect Prediction. *Int. J. Precis. Eng. Manuf.* **2023**, *24*, 33–41.
12. Li, C.K.; Wang, J.X.; Dai, Y.; Shi, Y. Experimental validation of saliency maps for understanding deep neural networks for weld penetration prediction. *J. Manuf. Process* **2023**, *88*, 22–33. [[CrossRef](#)]
13. Duan, L.; Xu, C.; Zhong, S.S.; Zhou, H.B.; Duan, J.-A. Optical power auto-alignment method with eugenics sorting for enhancing the alignment speed and robustness of fiber-grating couplers. *Opt. Express* **2022**, *30*, 39544–39560. [[PubMed](#)]
14. Li, J.Z.; Li, K.Y.; Zhang, G.; Li, Z.; Chen, Z.Q. Restorable, enhanced, and multifaceted tunable optical properties in a coaxial quantum well driven by the electric field and intense laser field. *Opt. Express* **2022**, *30*, 3474–3487. [[CrossRef](#)] [[PubMed](#)]
15. Liu, C.; Zhao, C.Y.; Wen, B.C. Investigation on coupled vibration of machine tool table system with position deviations. *Int. J. Adv. Manuf. Technol.* **2021**, *114*, 2321–2327. [[CrossRef](#)]
16. Wang, L.; Wei, Y.H.; Chen, J.C.; Zhao, W.Y. Macro-micro modeling and simulation on columnar grains growth in the laser welding pool of aluminum alloy. *Int. J. Heat. Mass. Tran.* **2018**, *123*, 826–838. [[CrossRef](#)]
17. Xu, J.J.; Rong, Y.M. Improving the prediction accuracy of thermal finite element analysis for laser welding through an automated optimization method. *Int. J. Adv. Manuf. Tech.* **2022**, *123*, 1657–1668. [[CrossRef](#)]
18. Han, R.; Fang, X.; Zhang, H.; Ren, D.; Liu, L.; Song, G. Adaptive Control for Laser-GMAW Hybrid All-Position Root Pass Welding. *Weld. J.* **2026**, *105*, 5.
19. Zhang, H.T.; Bai, X.W.; Dong, H.H.; Zhang, H.O. Modelling and Prediction of Process Parameters with Low Energy Consumption in Wire Arc Additive Manufacturing Based on Machine Learning. *Metals* **2024**, *14*, 567. [[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.