Parameter Selection of Direct Modulation Semiconductor Laser for Shaping Current Based on Convolutional Neural Network

Qing-An Ding [ID], Huixin Liu, Xiaojuan Wang, Gaoyang Zhu [ID], Liuge Du [ID], Xudong Cheng, Li Zheng, Junkai Li, and Qunying Yang

Abstract—The shaping current technology can efficiently and low-costly suppress the relaxation oscillations (ROs) of the direct modulation semiconductor laser (DML) for the high-performance optic system. The parameter selection is the key problem to precisely constructing the injection current form to obtain the desired output waveform. A novel framework based on convolutional neural network (CNN) is proposed to predict the shaped current parameters avoiding the time-consuming and computationally complex problems with analytical solutions. In the network training, batch and min-max normalizations are adopted to optimize neural networks, which aim to accelerate the convergence and improve their approximation ability. The trained inverse CNN named by feeding into the desired data samples from DML output waveform is used to achieve parameter selection for constructing the injection current. Also, the trained forward CNN would verify the validity of selected parameters responding to the output waveform, and get the unique corresponding relationship between them. Simulation results own high agreement with the theoretical values and show that the CNN models provide a powerful tool to select parameters of shaped current with accurate and fast capabilities.

Index Terms—Semiconductor lasers, relaxation oscillations, parameter selection, convolutional neural network.

I. INTRODUCTION

SEMICONDUCTOR laser is a key element in optical communication transmission systems whose performance can determine the quality of the whole system [1], [2]. In order to transmit optical communication information, semiconductor lasers usually adopt either external modulation (EM) or direct modulation (DM). EM refers to the method that modulates the output intensity of semiconductor laser with the external modulator based on the external signal. This method has the advantages of high transmission bit rate and long transmission distance, but it needs expensive external equipment [3]. DM alters the injected current of the laser to generate the time-dependent output in the optical intensity, which is simple and low cost. But, the relaxation oscillations (ROs) produced by this method would cause an increasing bit error rate in optical communication [4]. To overcome the ROs, many techniques have been presented, including external electrical circuit [5], outside optical feedback [6], light injection[7], and the modification of the physical laser structure [8]. Compared with these methods, shaping current is easier to achieve and cheaper. The ROs are suppressed by injecting the switching shaped current of direct modulation laser (DML) [3], [9], [10]. However, the switching shaped current has the characteristic of discontinuity.

To overcome the discontinuity, our team has presented an improved continuous shaping current [11], which is a more attractive solution with the characteristic of flexibility. The ROs in the desired output light pulse are determined by the parameters of shaped current, which are difficult to obtain. For the traditional numerical method, a random set of parameters is given according to the parameter ranges to calculate the output waveform, and it is compared with the target waveform to compute the mean square error. If the target precision or the maximum number of iterations is achieved, a set of parameters is output. This method is limited by iterations, so it tends to get the problem of local optimum and huge calculation. Particularly for the semiconductor laser that involves nonlinear processes and multi-physics interactions, the traditional numerical method is very time-consuming.

Because of the flexibility and the learning ability of artificial neural network (ANN), it has been applied in the inverse design [12], [13] and tackles the non-uniqueness problem in the training. For the inverse design of semiconductor laser, P. Feng and Z. Ma have proposed a method that combines

Qing-An Ding is with the School of Electronics and Information Engineering, Shandong University of Science and Technology, Qingdao 266510, China (e-mail: aqd6677@126.com).

Huixin Liu, Xiaojuan Wang, Xudong Cheng, Li Zheng, Junkai Li, and Qunying Yang are with the Microwave and Optical Communication Studio, Shandong University of Science and Technology, Qingdao 266510, China, and also with the College of Electronic Information Engineering, Shandong University of Science and Technology, Qingdao 266510, China (e-mail: 2895187400@qq.com; 1023710441@qq.com; 480346326@qq.com; 3449950944@qq.com; 2261236194@qq.com; 642864604@qq.com).

Gaoyang Zhu is with the School of Electronics and Information Engineering, Shandong University of Science and Technology, Qingdao 266510, China (e-mail: gaoyangzhu@sdust.edu.cn).

Liuge Du is with the School of Information Science and Engineering, Shandong University, Qingdao 266237, China (e-mail: lgdu@sdu.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2022.3221118

ANN with particle swarm optimization to improve the accuracy and speed compared with traditional method [14], [15]. As a special neural network, convolutional neural network (CNN) introduces convolution and pooling operations to generate deep features, thereby enhancing the ability of pattern recognition [16]. CNN has been used in signal extraction [17] and analysis [18] in photonics, which is reliable and effective. Under this background, we propose a scheme to select the parameter for constructing the shaped current waveforms based on CNN.

In this tutorial, the novelty of the proposed solution introduces the CNN model instead of the DML to construct the complicated relationship between the injection current and output waveform by forward and inverse intelligently artificial training. Initially, the theoretical injection current and desired output waveform are derived from the rate equations to get training samples. The batch and min-max normalizations are embedded in neural networks to get better performance. Afterward, CNN models are trained to select and verify the parameters, while the loss function is used to evaluate the performance of CNNs. By this way, the CNN overcomes the problem of time-consuming and huge calculation of traditional numerical method. Moreover, a series of input tests for CNNs indicate that our proposed scheme is useful for the shaping current technology and has the advantages of low cost, high-efficient and accuracy.

II. SHAPING CURRENT TECHNOLOGY

A. Dimensionless Rate Equations

The rate equations are an important theory for the study of semiconductor lasers. The semiconductor laser is assumed to be a single-mode laser. In this case, the laser dynamics can be expressed by a system of two dimensional ordinary differential equations driven by carrier density and photon density [4] as follows:

$$\frac{dN}{dt} = -\gamma_s N - G(N, S)S + \frac{J_0 + J(t)}{ed}, \quad (1)$$

$$\frac{dS}{dt} = -\gamma_c S + \Gamma G(N, S)S, \quad (2)$$

where N and S represent the carrier density and photon density, J_0 is the bias current, $J(t)$ is the modulation around the bias, $G(N(t), S(t))$ is the optical gain coefficient including nonlinear effects, Γ is the restriction factor, d is the active layer thickness of the laser, γ_c and γ_s is the decay rate of photon and spontaneous carrier.

It is necessary to use dimensionless rate equations for achieving the purposes of numerical calculation and analysis. In the absence of modulation ($J(t) = 0$), the dimensionless quality is carried out at the fixed point (S_0, N_0) of the laser. The nonlinear $G(N(t), S(t))$ around this point is expanded to the first order as:

$$G(N, S) = G_0 + G_N(N - N_0) + G_P(S - S_0), \quad (3)$$

where G_0 is the optical gain constant, $G_n = \partial G / \partial N$, and $G_p = \partial G / \partial S$.

TABLE I
THE VALUE AND MEANING OF PARAMETERS IN EQUATION

Symbol	Value	Description
J	2/3	bias current at fixed point
γ_s	$1.458 \times 10^9 \, s^{-1}$	spontaneous carrier decay rate
γ_n / J	$2.0 \times 10^9 \, s^{-1}$	gain variation with carrier density
γ_p / J	$3.6 \times 10^9 \, s^{-1}$	gain variation with photon density
γ_c	$3.6 \times 10^{11} \, s^{-1}$	photon decay rate
ω_R	$\sqrt{\gamma_c \gamma_n + \gamma_s \gamma_p} \sim 1/45 ps$	relaxation oscillation frequency
T_0	1.8	rise time

At the time of $dS/dt = dN/dt = 0$, (S_0, N_0) is submitted for (1) and (2) and then we can obtain the following equations

$$G_0 S_0 = \frac{J_0}{ed} - \gamma_s N_0, \quad (4)$$

$$\Gamma G_0 = \gamma_C. \quad (5)$$

The dimensionless photon density, the dimensionless carrier density, and the current can be defined by:

$$s(t) = \frac{S(t)}{S_0}, \quad (6)$$

$$n(t) = \frac{N(t)}{N_0}, \quad (7)$$

$$J + J_M = \frac{J_0 + J(t)}{\gamma_s N_0 ed} - 1. \quad (8)$$

And the dimensionless gain $g = G(t)/G_0$ is equivalent to

$$g(n, s) = 1 + \frac{\gamma_n}{J\gamma_s} n - \frac{\gamma_p}{\gamma_c}(s - 1), \quad (9)$$

where $\gamma_n = G_n S_0$ and $\gamma_p = -\Gamma G_p S_0$. The angular frequency of relaxation oscillation is $\omega_R = \sqrt{\gamma_c \gamma_n + \gamma_s \gamma_p}$ and the dimensionless time is $\tau = t\omega_R$, then the dimensionless photon density and dimensionless current density can be expressed in terms of

$$\frac{ds}{d\tau} = \frac{\gamma_c}{\omega_R}([g(n, s) - 1]s), \quad (10)$$

$$\frac{dn}{d\tau} = \frac{\gamma_s}{\omega_R}(J + J_M - n - jg(n, s)s). \quad (11)$$

To obtain the input current and corresponding output light pulse of DML, it is necessary to derive the dimensionless rate equations. The dimensionless qualities and parameters are shown in Table I, which are based on the works of Lucas Illing [4].

B. Shaping Current Technology Based on Dimensionless Rate Equations

The laser is equated to a nonlinear system. $s = e^{y(\tau)}$ is introduced to get the output waveform function (s), then the differential equation can be derived as $\frac{ds}{d\tau} = e^{y(\tau)}y'(\tau)$ and combine it with $ds/d\tau$ in the formula (10). The $n(\tau)$ can be

derived as the following expression:

$$n(\tau) = \frac{J\gamma_s}{\gamma_n}\left[\frac{e^{y(\tau)}y'(\tau)}{\gamma_C s}\omega_R + \frac{\gamma_P}{\gamma_C}(s-1)\right], \quad (12)$$

then $\frac{dn}{d\tau} = \frac{J\gamma_S}{\gamma_n\gamma_C}\left[y''(\tau)\omega_R + \gamma_P e^{y(\tau)}y'(\tau)\right]$ can be deduced, after that we solve it and $ds/d\tau$ in the (10) simultaneously to obtain the driving current (J_M). The relationship between driving current and output waveform can be expressed as:

$$J_M = \frac{J\omega_R^2}{\gamma_n\gamma_c}\left[y''(\tau) + \left(\frac{\gamma_p+\gamma_n}{\omega_R} + \frac{\gamma_s}{\omega_R}\right)y'(\tau) + (s-1)\right]. \quad (13)$$

This formula relates a given waveform s, via $y(\tau)$, to the injection current ($J_S = 3 + J + J_M \geq 0$, the relaxation oscillations in the desired output waveform can be suppressed by injecting J_S) that causes it, which involves complex nonlinear problems. The desired output waveform (s) is selected as an approximate rectangular pulse due to the symmetry of rectangular wave. Then, s is divided into two symmetrical parts, i.e., rising edge (S_{on}) and failing edge (S_{off}).

To realize the effect of steady approximation to rectangular pulse, we use the high-order Fourier series $y(\tau)$ to construct s and $y(\tau)$ can be defined as follows [11]:

$$y(\tau) = P\left(\frac{\tau}{T}\right) + \left[\sum_{k=1}^{N} a_k \sin\left(2\pi f\left(\frac{\tau}{T}\right)\right)\right]^4, \quad (14)$$

where $N = 2$. a_k is the constant parameter that determines and resolves whether s approximates to an ideal rectangular pulse. $P(x)$ is a seventh-order polynomial, which can be expressed in the rising edge as:

$$P_{on}(x) = (-20x^7 + 70x^6 - 84x^5 + 35x^4)\frac{\ln s(0)}{\ln s(T_0)} + \ln[s(T_0)]. \quad (15)$$

Here $x = \frac{\tau}{T}$ and T is the dimensionless transition time. To minimize the transition time between the rising and falling in s, the f-function is given by

$$f(x) = \frac{(a_0+1)x}{a_0x+1}, \quad (16)$$

where the value a_0 is 2. In the paper, we only examine the S_{on}.

The parameters (a_1, a_2) are the key to constructing injection current since the current can differ greatly even with small changes of them. The selection of parameters in injection current, according to the desired output waveform, is a problem of multi-parameter non-linearization [19]. The parameter selection and obtaining J_S with numerical methods is still challenging, as in time-cost and vast calculation. Therefore, we propose a method based on CNN to acquire the parameters and injection current with fast speed and accuracy. Also, the forward neural networks are trained in advance to verify the selected parameters.
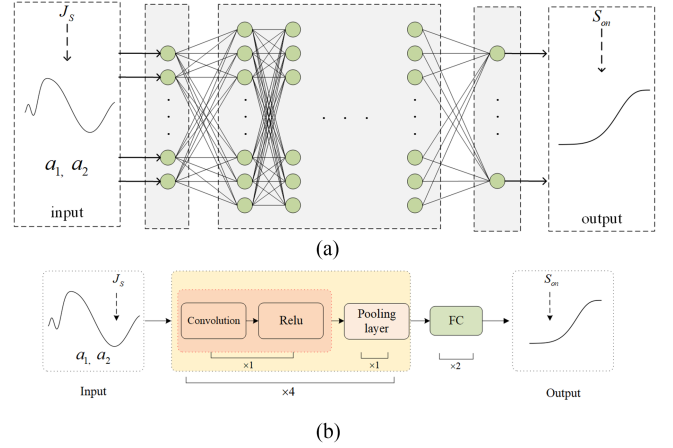

Fig. 1. (a) a DNN architecture for parameter verification. (b) a CNN architecture for parameter verification.

TABLE II
THE PARAMETER SETTINGS OF DNN

Parameter	Value
Learning rate	0.0008
Hidden layers	20
Hidden units	100
Activation function	Tan Hyperbolic
Iteration	15000
Optimizer	Adam

III. SHAPING CURRENT BASED ON ARTIFICIAL NEURAL NETWORK

A. Artificial Neural Network Architectures for Parameter Verification

To obtain the ANN that can predict the S_{on} corresponding to parameters (a_1, a_2) and J_S, a CNN and a DNN are constructed as shown in Fig. 1. The two architectures contain 203 input units, which are used to receive the parameters (a_1, a_2) and the sampling points for J_S.

The values of the parameters are randomly selected in the specified ranges of $a_1 = [0.13, 0.33]$ and $a_2 = [0.1, 0.23]$, which are obtained from the shaping current derivation based on single mode rate equations and a large number of scanning using the calculation software.

The DNN architecture consists of an input layer, 20 hidden layers, and an output layer, and its hyper-parameter setting is shown in Table II. The DNN based on the back-propagation algorithm is trained by sampled data [20], [21]. The hidden layers and output layer process the sign, and the ultimate result (S_{on}) is exported by the output layer neurons [21].

The CNN architecture is composed of convolutional, pooling, and fully connected layers and is a kind of neural network with local connection and weight sharing [22]. The filter in the convolutional layer is used to scan the input data to seek features, and the size we employed is 1×3 [18]. After the convolutional layer, the maximum pooling is implemented by 2×2 kernel, and the size of the object is reduced by a factor

TABLE III
THE HYPER-PARAMETER FOR CNN1

| Parameter	Value
Learning rate	0.001
Batch size	2048
Activation function	Rectified Linear Unit
Optimizer	Adam

Fig. 2. Training and validation loss curves for forward models of the semiconductor laser. (a) the comparison of training and validation loss curves for DNN. (b) the comparison of training and validation loss curves for CNN1.

of two [18], [23]. Before training, the input signal containing parameters (a_1, a_2) and J_S is converted to the form of 1×203. Therefore, the filter and maximum pooling used in the CNN are two dimensional. In addition, zero padding is applied to all the convolutional and pooling layers whose strides are set to 1 and 2 respectively. Table III presents the hyper-parameter setting of CNN architecture.

The two models are trained to map the semiconductor laser behavior, which produce S_{on} in response to given parameters (a_1, a_2) and injection current. The CNN model is defined as CNN1. During the process of training, the loss function convergences faster and stabilizes at smaller value when the DNN and CNN adopt the Tan Hyperbolic and Rectified Linear Unit activation functions, respectively. Therefore, to better realize the forward modeling of semiconductor laser based on DNN and CNN, the two models employ different activation functions.

B. The Training of Neural Networks

49999 data samples are generated before the training of neural networks. 44999 samples are used for training, while the validation and test sets contain 2500 samples respectively. In this paper, the loss function adopts the form of mean squared error (MSE) which can be written by:

$$MSE = \frac{\sum_{i=1}^{n} (y_i - y'_i)^2}{n}, \tag{17}$$

where y_i is the ith theoretical value, y'_i is the ith predicted figure, and n is the sample size. In the training process, the trainable parameters are continuously updated with the reference target of achieving the smallest possible MSE. The dropout method is adopted to randomly abandon neurons which aims to avoid overfitting and improve the generalization ability of the networks. Each neuron is reserved with a fixed probability 0.7 in the hidden layer of DNN and fully connected layer of CNN. The success of the neural network can be testified by comparing the neural network predicted values with the theoretical values of shaping current technology.

C. Neural Network Optimization for Parameter Verification

Since DNN model includes many hidden layers, the data distribution will be changed in the training process and the convergence speed of the model will be affected to some extent. Thus, batch normalization (BN) layer as a special layer is added before the input layer and each hidden layer of DNN. The input data of these layers can be normalized to the standard normal distribution so that the values are transformed into the input-sensitive region. The net input values processed by BN
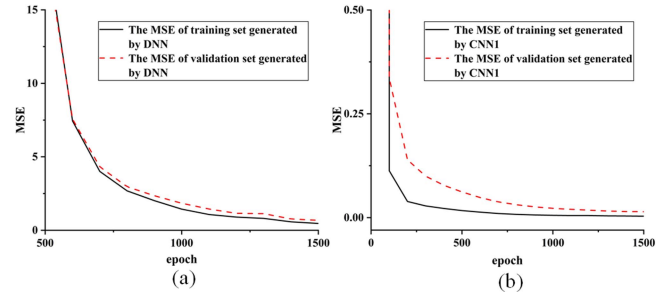
can be defined as

$$\widehat{z}^{(l)} = \frac{z^{(l)} - E\left[z^{(l)}\right]}{\sqrt{\text{var}(z^{(l)}) + \varepsilon}}, \tag{18}$$

where $z^{(l)}$ is the net input value of layer l, $E[z^{(l)}]$ is the mean of $z^{(l)}$, $\text{var}(z^{(l)})$ is the variance of $z^{(l)}$, and ε is a constant to prevent the denominator equal 0.

At the same time, in order not to cause negative effects on the neural network, scaling and shifting are added to change the value range.

$$\hat{z}^{(l)} = \frac{z^{(l)} - \mu}{\sqrt{\sigma^2 + \varepsilon}} \times \gamma + \beta, \tag{19}$$

where γ is scaling parameter and β is shifting parameter. They can be adjusted continuously during the training of the neural network.

To make the training easier, min-max normalization is adopted to normalize the input data of inverse CNN into the same scale. Since most of the gradient direction is approximate to the optimal search direction, the training efficiency is greatly improved [22]. The min-max normalization converts the inputs of the input layer to the range of [01], and the transformation formula is defined as

$$x^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \tag{20}$$

where x_i is the input data, $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of features x_i in all data samples.

Fig. 2 shows the training and validation loss curves comparison between DNN and CNN1, the convergence of MSE generated by CNN1 is much faster and the value is much smaller. For better achieving the forward modeling, the MSE between theoretical and predicted S_{on} should be less than 0.001. It can be seen that the training curve of DNN drops drastically and gradually stabilizes around 10^{-4}, while the training curve of CNN1 tends to be steady and finally stabilizes around 10^{-6}. In a word, CNN1 should be more accurate to predict the S_{on} corresponding to each combination parameter and J_S.

To further demonstrate the consistency between the theoretical values of S_{on} and the predicted values of the neural networks, a sample is selected randomly from the test dataset. DNN and CNN1 are injected the same J_S, a_1 and a_2, but Fig. 3(a) and (b) illustrate the diverse output. Fig. 3 indicates the
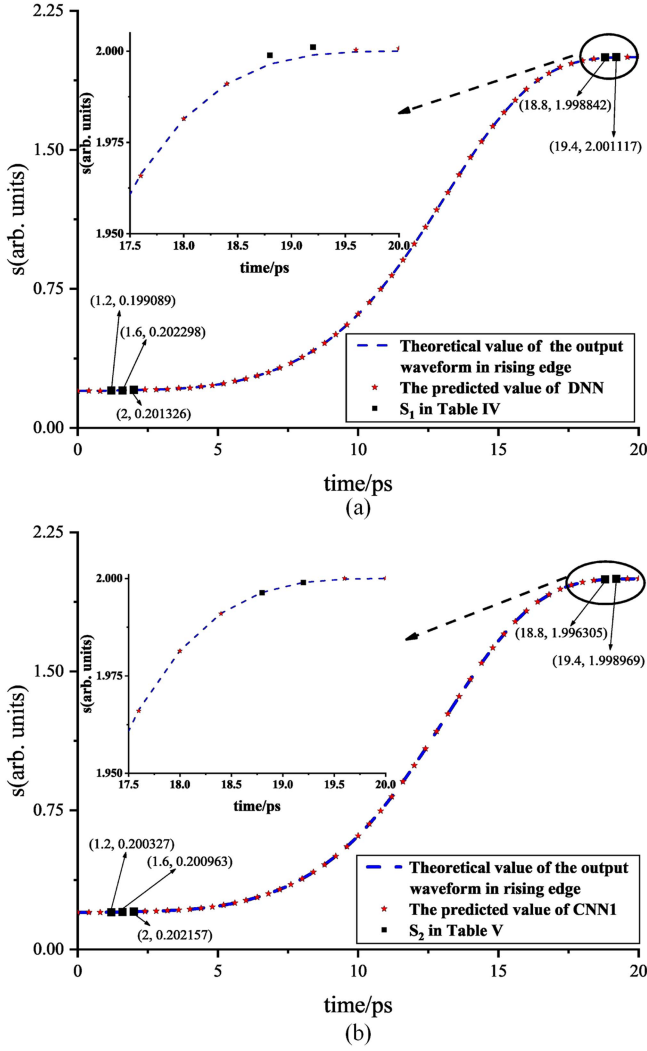
Fig. 3. The comparison between the theoretical values of the output waveform and the approximate values of neural network models. (a) the comparison between theoretical values and the DNN approximate values. (b) the comparison between theoretical values and the CNN1 approximate values.

TABLE IV
THE COMPARISON BETWEEN S_0 AND S_1

Parameter	Value				
J_S	3.96472	3.283423	3.95601	3.408711	4.759021
S_0	0.200371	0.202702	0.201294	1.996452	1.998974
S_1	0.199089	0.201326	0.202298	1.998842	2.001117
E_1	0.001282	0.001376	0.001004	0.00239	0.002143
RE_1	0.64%	0.68%	0.5%	0.12%	0.11%

TABLE V
THE COMPARISON BETWEEN S_0 AND S_1

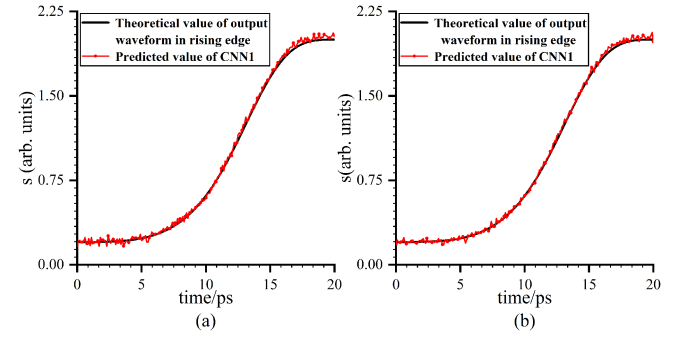| Parameter	Value				
J_S	3.96472	3.283423	3.95601	3.408711	4.759021
S_0	0.200371	0.202702	0.201294	1.996452	1.998974
S_2	0.200327	0.202157	0.200963	1.996305	1.998969
E_2	0.000044	0.000545	0.000331	0.000147	0.000005
RE_2	0.022%	0.27%	0.16%	0.007%	0.0003%

Fig. 4. The comparison between the predicted output waveform resulted from noisy input and theoretical output waveform. (a) Sample 1. (b) Sample 2.

favorable approximation ability of DNN and CNN1, meantime the ROs are completely suppressed in S_{on}. In Fig. 3(b), the stars are highly coincident with the blue dotted line. However, in Fig. 3(a) the stars fluctuate around the theoretical dotted line. The MSE (calculated by (17)) between the predicted values of neural networks and theoretical values in Fig. 3(a) and (b) are 2.17×10^{-4} and 6.44×10^{-6} respectively.

Tables IV and V show partial values of Fig. 3. S_0 represents the theoretical value of the output waveform in the rising edge in Fig. 3. S_1 and S_2 are output waveform values predicted by DNN and CNN1, which are indicated by black squares in Fig. 3(a) and (b). E_1 and E_2 are the absolute errors between the predicted value and theoretical value. RE_1 and RE_2 are the relative errors between the theoretical and the predicted values.

As shown in Tables IV and V, the value of E_2 is much smaller than the value of E_1, and also the RE_2 is much smaller than RE_1. The prediction time of a sample by DNN is 0.0473 s, which is 5.5 times that of CNN1. In comparison, the forward CNN

achieves better performance than DNN in terms of speed and accuracy with fewer optimization algorithms. Therefore, CNN1 is adopted to verify the parameters selected by inverse CNN.

To demonstrate the applicability of CNN1, 100 sets of parameters and J_S with Gaussian noise are injected into it. The MSE between the predicted and theoretical S_{on} is 2.7×10^{-4}, which still meets the quantitative evaluation. Fig. 4 shows the comparison between the theoretical and predicted values of semiconductor laser output waveform. The trends of red and black lines are similar, which illustrate the robustness of CNN1.

IV. PARAMETER SELECTION AND VERIFICATION OF SHAPING CURRENT TECHNOLOGY

A. Parameter Selection of Shaping Current Technology

An inverse CNN model is utilized to realize parameter selection (produces a_1, a_2 and the sampling points for J_S response to S_{on}) to construct shaped current of semiconductor lasers, as shown in Fig. 5. The CNN model is designed to have 201 input units. To address the problem of difficult training, the min-max normalization is adopted to process the input data of inverse
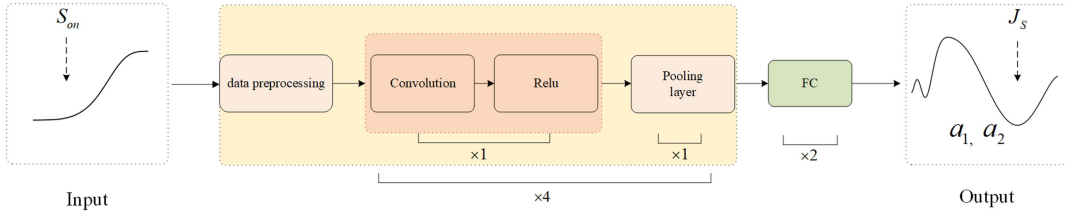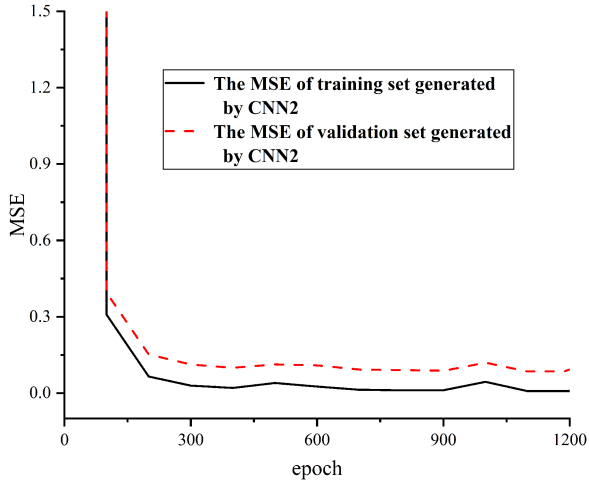
Fig. 5. The CNN schematic for parameter selection.

Fig. 6. The variations of MSEs during training and validating for CNN2.

TABLE VI
THE COMPARISON BETWEEN a_1 AND a_2

Parameters	a_1	a_2
Theoretical value	0.198	0.21245
Predicted value	0.197821	0.2127948
Absolute error	0.000179	0.0003448
Relative error	0.09%	0.16%

exhibit a good approximation ability of CNN2 from the inverse modeling of semiconductor laser. The absolute and relative errors in Table VI are calculated from sample 1, which concretely show the difference between the theoretical and predicted parameters.

As shown in Fig. 7 and Table VI, the parameters and J_S predicted by CNN2 are highly coincident with the theoretical values, so we conclude that CNN2 can overcome the challenges in the parameter selection of shaping current technology.

To confirm the robustness of inverse CNN, the relative Gaussian noise is added to the S_{on}. 100 samples S_{on} are randomly selected from the test set with noise and put into the trained inverse CNN. The MSE results between the theoretical and predicted values of the parameters (a_1, a_2) and J_S are 0.00086, 0.0011, and 0.0151, which still satisfy the evaluation. Fig. 8 compares the J_S resulting from the S_{on} with noise, which exhibits a good prediction ability of inverse CNN.

B. The Validity Verification of the Selected Parameters

Since the generalization capacity of CNN models, they are combined to select and verify the parameters of shaping current. A set of desired values of S_{on} (different from train, validation, and test sets) is inputted into CNN2. Then the predicted parameters and J_S are put into CNN1.

Fig. 9 depicts the J_S predicted by CNN2, and the predicted a_1 and a_2 are 0.3296627 and 0.1970667. Then they are inputted into CNN1 to acquire a sample of S_{on}. As shown in Fig. 10, the predicted values of CNN1 and desired S_{on} are highly fitted, and the MSE (calculated from (17) between them is 1.68×10^{-6}. Also, it proves the high approximation capabilities of the two CNN models.

Table VII demonstrates some of the data in Fig. 10. S_3 is the desired output waveform value. S_4 represents S_{on} value obtained by inputting predicted J_S and parameters into trained CNN1. E_3 and RE_3 (calculated from the difference between S_3 and S_4) are the absolute and relative errors, respectively. As can be seen

CNN, which is defined as CNN2. The hyper-parameter setting of CNN2 is the same as CNN1, as shown in Table III, also the filter size, pooling method, padding, and stride setting. For CNN2, the input signal (S_{on}) is transformed to the form of 1×201.

Fig. 6 demonstrates the MSEs (the difference between the theoretical and predicted ones of J_S and parameters) of CNN2 during the training and validating, the tendencies of them are similar. It can be seen that the MSEs decline sharply in the first few epochs and finally converge around 10^{-3} and 10^{-2}. We can conclude that CNN2 is no overfitting problem since the gap of MSEs between the training and validation sets is small. To ensure the validity of CNN2, the model should meet the quantitative evaluation of MSE less than 0.04.

Moreover, to prove the effectiveness and stability of CNN2, 100 desired S_{on} randomly selected from the test set are put to it. The MSE (calculated by (17) of the predicted parameters (a_1, a_2) and the theoretical values derived from these 100 samples are 4.12×10^{-5}, 2.21×10^{-5}, respectively. The MSE between the theoretical and predicted J_S derived from the 100 samples is 0.003, which illustrates the approximate ability of CNN2.

To more intuitively indicate the reliability of CNN2, four S_{on} samples are selected from the test set, as well as the predicted and theoretical values are presented. Fig. 7 demonstrates the comparison between the predicted J_S generated by CNN2 with the theoretical values of the shaping current technology. The MSEs (calculated by (17) between the theoretical and predicted values of J_S are 0.00268, 0.00216, 0.0045, and 0.005, which
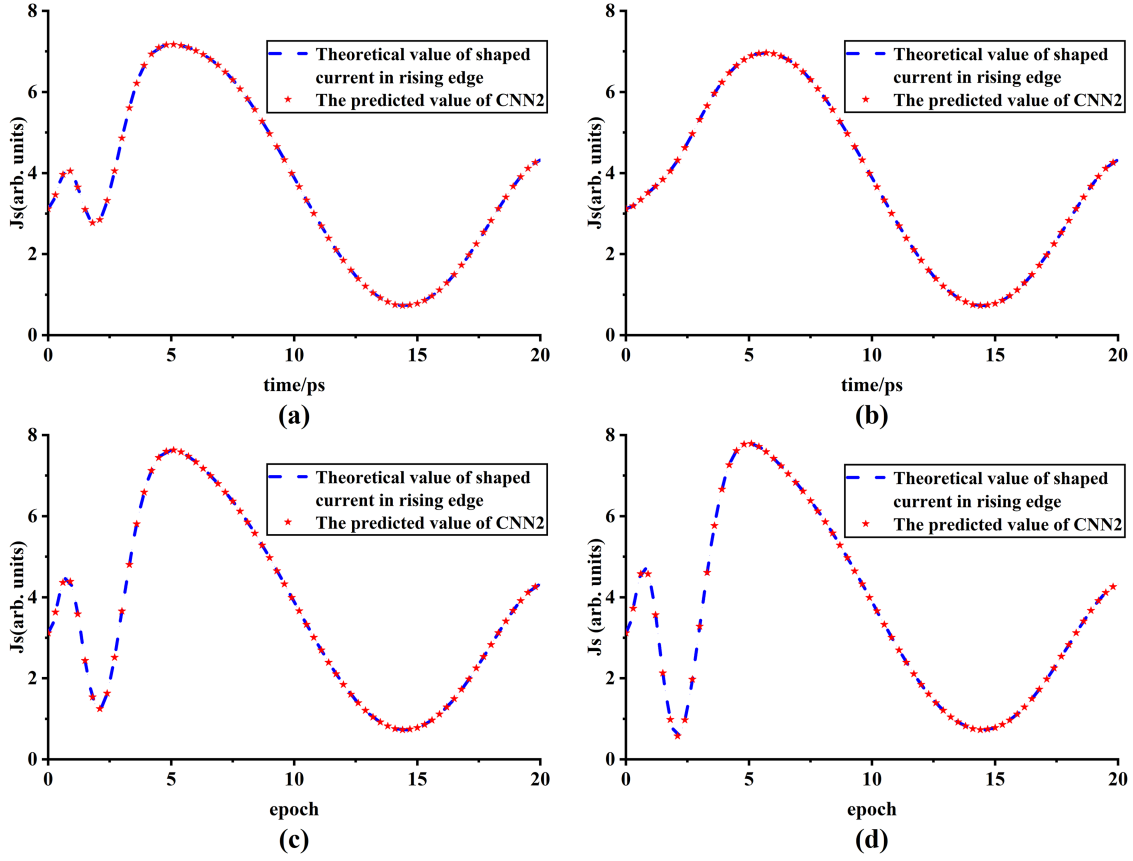
Fig. 7. The comparison between theoretical values of the J_S and CNN2 approximate values. (a) Sample 1. (b) Sample 2. (c) Sample 3. (4) Sample 4.
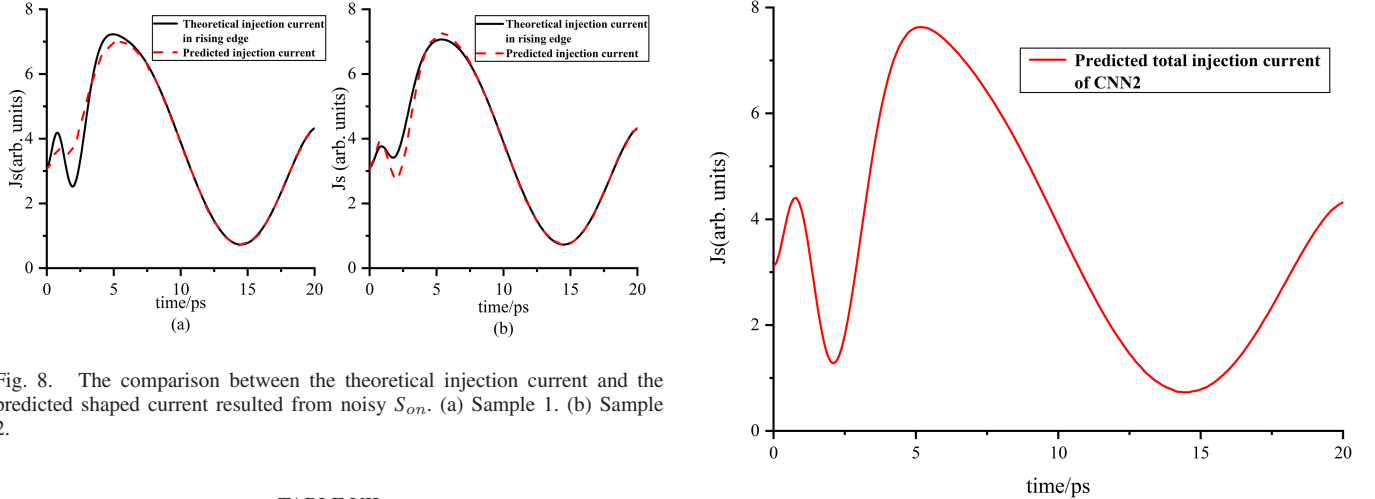

Fig. 8. The comparison between the theoretical injection current and the predicted shaped current resulted from noisy S_{on}. (a) Sample 1. (b) Sample 2.

Fig. 9. The predicted injection current of CNN2.

from Table VII, the values of E_3 and RE_3 are small, then the approximation and generalization abilities of CNN models are further proved.

Comparison among the several groups of experimental results, we can conclude that the CNNs can accurately realize the parameter selection of shaping current technology with fewer resources.

TABLE VII
THE COMPARISON BETWEEN S_3 AND S_4

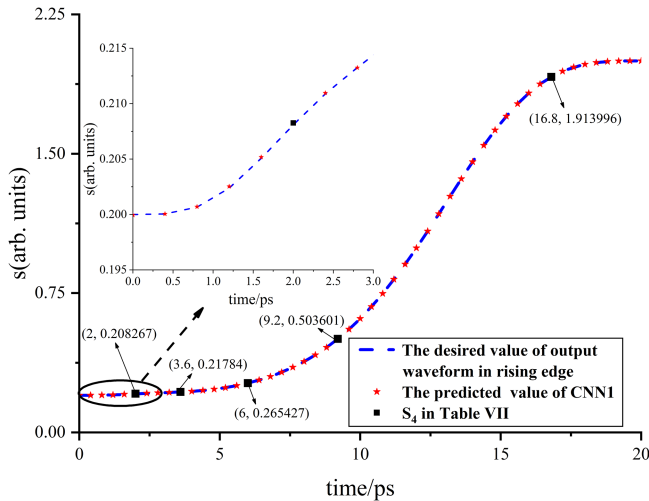| Parameter	Value				
J_S	1.327752	5.591380	7.398908	4.768610	1.726744
S_3	0.208098	0.218001	0.265504	0.503436	1.914127
S_4	0.208267	0.21784	0.265427	0.503601	1.913996
E_3	0.000169	0.000161	0.000077	0.000165	0.000131
RE_3	0.081%	0.074%	0.03%	0.033%	0.007%

Fig. 10. The comparison between desired values of the output waveform and CNN1 approximate values.

V. CONCLUSION

In the construction of shaped current, the novel parameter selection scheme based on CNN trained by data samples from dimensionless rate equations is suggested. The BN and min-max normalizations successfully solve the difficult training problems of convergence and reduce the size of neural networks. With normalizing data, the proposed inverse CNN works well in realizing the parameter selection for shaping current technology. Moreover, the trained forward CNN is efficiently utilized to verify the parameters corresponding to S_{on}, which is faster and more accurate than DNN. By combining the results of two CNNs, the parameters selected are very suitable for the theoretical values, and also the generalization ability of the CNNs is demonstrated. The simulation results indicate that the trained CNNs have strongly computing ability and high accuracy to realize the parameter selection and verification.

The proposed scheme would give insights to the development of shaping current technology and can be used as practical design guidelines for suppressing relaxation oscillations of DML. A potential future study is using the CNN predicted injection current to modulate a real semiconductor laser, and analyzing the output waveform.

REFERENCES

[1] B. Z. L. Jiusheng, "Artificial neural network modeling of the modulation characteristics of semiconductor lase," *Chin. J. Laers*, vol. 31, no. 11, pp. 1301–1304, 2004.

[2] L. Li, Y. Shi, Y. Zhang, H. Zou, J. Shen, and X. Chen, "Study on a DFB laser diode based on sampled grating technique for suppression of the zeroth order resonance," *IEEE Photon. J.*, vol. 9, no. 2, Apr. 2017, Art. no. 1501909.

[3] B. Nie, Q.-A. Ding, C. Liu, X. Cheng, X. Wang, and L. Zhang, "Suppression of relaxation oscillations in DFB lasers through shaping injection current waveforms," in *Proc. IEEE 20th Int. Conf. Commun. Technol.*, 2020, pp. 975–979.

[4] L. Illing and M. B. Kennel, "Shaping current waveforms for direct modulation of semiconductor lasers," *IEEE J. Quantum Electron.*, vol. 40, no. 5, pp. 445–452, May 2004.

[5] Y. Suematsu and T.-H. Hong, "Suppression of relaxation oscillation in light output of injection lasers by electrical resonance circuit," *IEEE J. Quantum Electron.*, vol. 13, no. 9, pp. 756–762, Sep. 1977.

[6] D. Zhao, S. Andreou, W. Yao, D. Lenstra, K. Williams, and X. Leijtens, "Monolithically integrated multiwavelength laser with optical feedback: Damped relaxation oscillation dynamics and narrowed linewidth," *IEEE Photon. J.*, vol. 10, no. 6, Dec. 2018, Art. no. 6602108.

[7] R. Lang and K. Kobayashi, "Suppression of the relaxation oscillation in the modulated output of semiconductor lasers," *IEEE J. Quantum Electron.*, vol. 12, no. 3, pp. 194–199, Mar. 1976.

[8] R. S. Tucker, "High-speed modulation of semiconductor lasers," *IEEE Trans. Electron Devices*, vol. 32, no. 12, pp. 2572–2584, Dec. 1985.

[9] Q. Junqiu, *Modulation Characteristics and Optimizations of Directly Modulated Distributed Feedback Laser*. Hubei, Wuhan, China: Huazhong Univ. Science and Technology, 2015.

[10] N. Dokhane and G. Lippi, "Improved direct modulation technique for faster switching of diode lasers," *IEE Proc.-Optoelectron.*, vol. 149, no. 1, pp. 7–16, 2002.

[11] Q. Ding et al., "Continuous current-injected waveforms shaping for suppressing relaxation oscillations of direct modulation based on equivalent circuit model," *Opt. Exp.*, vol. 30, no. 11, pp. 19273–19287, 2022.

[12] Y. Long, J. Ren, Y. Li, and H. Chen, "Inverse design of photonic topological state via machine learning," *Appl. Phys. Lett.*, vol. 114, no. 18, 2019, Art. no. 181105.

[13] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *Acs Photon.*, vol. 5, no. 4, pp. 1365–1369, 2018.

[14] P. Feng and Y. Li, "Semiconductor laser parameter inverse design method based on artificial neural network and particle swarm optimization," *Chin. J. Lasers*, vol. 46, no. 7, 2019, Art. no. 0701001.

[15] Z. Ma and Y. Li, "Parameter extraction and inverse design of semiconductor lasers based on the deep learning and particle swarm optimization method," *Opt. Exp.*, vol. 28, no. 15, pp. 21971–21981, 2020.

[16] S. Liu, H. Ji, and M. C. Wang, "Nonpooling convolutional neural network forecasting for seasonal time series with trends," *IEEE Trans. Neural Netw.*, vol. 31, no. 8, pp. 2879–2888, Aug. 2020.

[17] C. Liu, L. Li, X. Huang, X. Wang, and C. Wang, "Audio signal extraction and enhancement based on CNN from laser speckles," *IEEE Photon. J.*, vol. 14, no. 1, Feb. 2022, Art. no. 6807205.

[18] S. Guo et al., "Automatic quantification of subsurface defects by analyzing laser ultrasonic signals using convolutional neural networks and wavelet transform," *IEEE Trans. Ultrasonics, Ferroelect., Freq. Control*, vol. 68, no. 10, pp. 3216–3225, Oct. 2021.

[19] X. Guangzong, L. Xingwu, and Z. Bin, "A novel global optimization method to design laser parameters with artificial neural network and genetic algorithm," *Chin. J. Lasers*, vol. 37, no. 5, pp. 1203–1208, 2010.

[20] S. Boscolo and C. Finot, "Artificial neural networks for nonlinear pulse shaping in optical fibers," *Opt. Laser Technol.*, vol. 131, 2020, Art. no. 106439.

[21] J.-S. Li and Z.-W. Bao, "Artificial neural network for the noise characteristics of laser modeling," in *Proc. IEEE Int. Conf. Mach. Learn. Cybern.*, 2004, vol. 5, pp. 3219–3222.

[22] Q. Xipeng, *Neural Network and Deep Learning*. Beijing, China: China Machine Press, 2020.

[23] V. Bangari et al., "Digital electronics and analog photonics for convolutional neural networks (DEAP-CNNs)," *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 1, Jan./Feb. 2020, Art. no. 7701213.