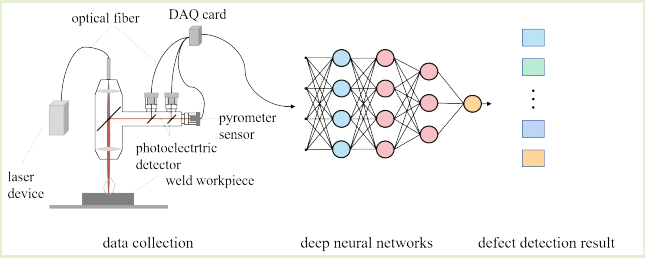
Real-Time Defect Detection Scheme Based on Deep Learning for Laser Welding System

Peng Peng, Kui Fan, Xueqiang Fan, Hongping Zhou, and Zhongyi Guo

***Abstract*—Laser welding, as an important material processing technology, has been widely used in various fields of industry. In most industrial welding production and processing, high precision is required for welding parameters and fixed work pieces. However, in the process of laser welding, serious heat transfer effect will bring unpredictable welding deviations, and even a small deviation will lead to serious welding defects, which will affect the quality of the welded products. Traditional nondestructive testing methods have been widely used, but they have been proved to have some limitations. Existing laser welding defect detection schemes are mainly focused on the detection of postweld defects, which requires a large amount of data, and the real-time detection cannot be guaranteed. In this article, we propose a data acquisition system for collecting changes in physical characteristics during laser welding with the aids of multiple sensors. Based on the data originating from sensors' system, an efficient laser welding defect detection model has been designed and investigated based on the multiscale convolutional neural network (MSCNN), bidirectional long short-term memory (BiLSTM), and attention mechanism (AM). The final proposed MSCNN-BiLSTM-AM fusion detection model can achieve 99.38% detection accuracy, which make the laser welding system more efficient and more suitable.**



***Index Terms*— Attention mechanism (AM), data enhancement, defect detection, integrated learning, laser welding.**

I. Introduction

METALS are indispensable and important materials in modern industrial production and manufacturing. Metal materials are indispensable for everyday household appliances, transportation, cargo-laden ships and airplanes, corrosion-resistant and pressure-resistant chemical equipment, and aerospace and national defense industries. In the manufacturing process of these industrial products, various parts need to be connected to make products according to the design requirements. Laser welding is an ideal processing method to connect these parts [1], [2], [3] and widely used in industrial production and manufacturing process, which is an improvement and innovation of the traditional welding methods [4], [5], [6], [7], [8]. Compared with the traditional methods, laser welding has the characteristics of higher power density, fast speed, large depth, and little influence on the thermal distortion of the welded parts. In addition, laser welding can weld refractory materials, such as titanium, quartz, and so on, and can weld anisotropic materials with good results [9], [10], [11]. Because of this, laser welding technology is gradually replacing the traditional welding methods and has been widely used in civil production, military engineering technology, aerospace technology, and other fields [12], [13], [14].

As an important means of welding, laser welding can realize the automation and flexible processing of the welding process, in which the high speed of laser welding meets the inherent requirements of the development of modern industry, but like the traditional welding methods, there will be some welding defects caused by aging of equipment, parameter differences, welding materials, and so on. How to effectively and timely discover the defects in the laser welding process is particularly important. There are many external and internal causes that cause the defects of laser welding, and there will be cracks, breakdown, defocusing, distortion, lack of welding, and other welding defects in the process of laser welding. These welding defects will reduce the quality of component products and even invalidate their functions. A small welding defect may cause enormous economic and even personal injury. Therefore, defect detection of laser welding has become a very important

Manuscript received 8 May 2023; accepted 13 May 2023. Date of publication 23 May 2023; date of current version 1 August 2023. This work was supported by the National Natural Science Foundation of China under Grant 61775050. The associate editor coordinating the review of this article and approving it for publication was Dr. Daniele Tosi. *(Corresponding author: Zhongyi Guo.)*

The authors are with the School of Computer and Information, Hefei University of Technology, Hefei 230009, China (e-mail: pengpeng2021m@163.com; fankuihfut@163.com; fanxueqiang2022B@163.com; ciangela@hfut.edu.cn; guozhongyi@hfut.edu.cn).

Digital Object Identifier 10.1109/JSEN.2023.3277732

inspection link. Now, with the development of miniaturization and lightweight industrial equipment, the traditional welding defect detection scheme appears to be somewhat inadequate and cannot be applied to more complex and variable welding scenarios.

With the rapid development and maturity of the theory of deep learning, the performance of deep learning in various fields is brilliant [15], [16]. Intelligent fault diagnosis technology relying on in-depth learning has become a research hot spot and development direction in the field of artificial intelligence. At present, the main intelligent fault diagnosis algorithms are artificial neural network (ANN), support vector machine (SVM), random forest, and so on [17], [18]. They are developing from single strategy classification and prediction to multistrategy fusion. Now, the in-depth learning technologies, such as machine vision, biometric recognition, natural language processing, speech recognition, and intelligent decision automation control, are integrated into the production, operation, and defect detection of the equipment, which enables the equipment to have the ability of self-perception, self-adaptive, self-control, self-learning, and so on, thus greatly improving the overall operation stability and reliability of the industrial system [19].

Based on the powerful feature expression ability of deep learning [20], an automatic welding defect location method has been proposed based on U-net network, which includes data enhancement and welding defect location. Kan and Kalkan [21] present an automatic welding defect detection method based on image processing and machine learning, in which the welding material is divided into several subimages, and each sub image is detected and classified. Deng et al. [22] study these data based on industrial image processing algorithm and deep learning algorithm. In this study, the median filter is used to remove the noise in the weld image. Image enhancement technology is used to improve the image contrast of different regions. Deep convolution neural network (CNN) is used for feature extraction, and the activation function and adaptive pool method can be improved. A multisensor data fusion network based on CNN is presented for in-process defect detection, called IDDNet [23]. The experimental results show that IDDNet achieves better classification results than SVM with an overall accuracy of 97.57%. An integrated deep learning architecture [15] has been proposed and investigated based on CNN, gated recursive unit (GRU), and high-performance classification algorithms, such as k-nearest neighbor (KNN) and SVM, in which a synthesis scheme based on classical machine learning method is proposed, and the optimal hyperparameters of each algorithm are determined by extensive grid search. Miao et al. [24] present a comprehensive identification method of weld defects based on CNN combined with eddy current testing and 3-D laser scanning. The detection principle and equipment of the two detection methods are introduced. In order to realize the real-time detection of narrow lap welding, a two-stage defect identification model is established, which greatly improves the efficiency of weld defect identification. Aiming at the problem of low measurement accuracy of the existing detection methods, a welding surface defect detection system based on the principle of optical coherent distance measurement has been designed [25]. Based on the principle of Michelson interferometer, a point on the welding surface is scanned by broadband laser to collect the interference light intensity of the reference arm and the measuring arm.

The main contributions of this article can be summarized as follows.

First, we propose an intelligent laser welding defect detection scheme [multiscale convolutional neural network (MSCNN)–bidirectional long short-term memory (BiLSTM)–attention mechanism (AM)] based on MSCNN, BiLSTM and AM. Because the multimodel integration and fusion can take into account the advantages of different models, the proposed fusion model can not only give full play to the spatial feature extraction ability of CNN but also extract the temporal correlation information of features by using BiLSTM. In addition, AM is introduced in the appropriate position of the network to reduce or filter the impact of "noise" features on the final classification results and improve the recognition performance of the overall detection scheme.

Second, when designing CNN, we tried the multichannel and multiscale cascading mode of CNN and BiLSTM, which fused the feature information within the range of different receptive fields, greatly enriched the diversity of features learned by the model, and made it possible to train a better depth model. Finally, through a large number of comparative experiments and ablation experiments, we verified the superiority of the improved scheme proposed in this chapter and confirmed that the advantage of introducing subalgorithm model into the theoretical design was reflected in the integrated model.

The rest parts are organized as follows. Section II briefly introduces the basic theory of deep learning model for defect detection. Section III describes the experimental device of data acquisition. In Section IV, the intelligent detection model proposed in this article is introduced in detail, and the effectiveness and superiority of the proposed method are studied and proved. Finally, Section V draws a conclusion.

II. Principle of Deep Learning Defect Detection

As an improved network of RNN, the LSTM is widely used in the classification of time-series data, especially in text classification, because it can capture the time-series characteristics well and handle the data with front–back dependence well. Although the LSTM model captures long-distance dependencies and successfully solves the problems of gradient disappearance and loss of historical information in the RNN model, the LSTM cannot extract back-to-front information, that is, the ability of LSTM to fuse front-to-back information is poor. Therefore, the BiLSTM model is introduced in order to better capture the correlation characteristics before and after the data. However, not every feature information contributes the same to the classification of final laser welding quality status in all fused information. Inspired by the fact that the human brain pays attention to some key information when it remembers something, different attentions are allocated to different levels of characteristic information. This attention allocation mechanism allows the in-depth learning model to

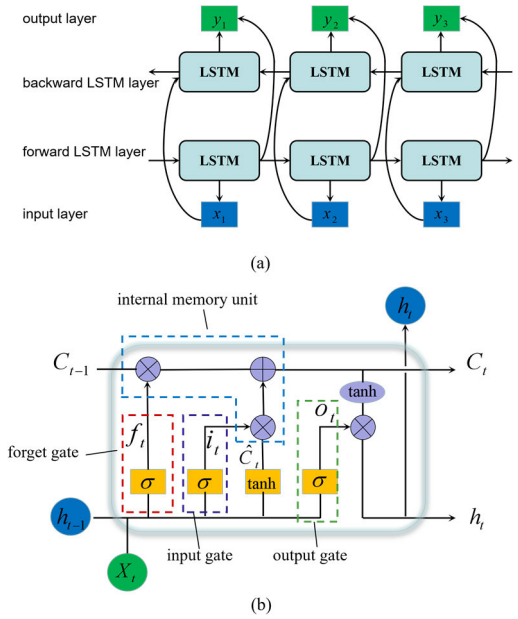


Fig. 1. (a) Network structure diagram of the BiLSTM. (b) Basic building blocks of the LSTM.

filter out more detailed information that is critical to the outcome of the target from a large amount of information, thus improving the expected performance level of the in-depth learning model.

A. Basic Mechanism and Principle of BiLSTM Network

The BiLSTM is a combination of forward LSTM and backward LSTM, and the network structure is shown in Fig. 1(a). BiLSTM adds a backward LSTM layer relative to LSTM. Such a structure allows BiLSTM to simultaneously handle both-way temporal characteristics, and each network layer can capture both past (forward) and future (backward) feature information at a specific time step. After processing the forward and backward LSTMs, the output of the two LSTMs is stitched together, so the network as a whole can fuse forward and backward temporal correlation information. BiLSTM approximates the bidirectional loop neural network in its overall structure, and the structural unit of the LSTM network is still used in the main part of the computational function. The structure unit of LSTM mainly consists of four components: input gate, output gate, forget gate, and internal memory unit, which are separated as shown in Fig. 1(b).

In Fig. 1(b), h_{t-1} is the final output value of the LSTM neuron unit at the last moment, C_{t-1} is the cell state of the smart grid unit at the previous moment, x_t is the input at the current moment, σ is the activation function, f_t is the output of the forgotten gate at the current moment, i_t is the output of the input gate at the current moment, $\hat{C}_t$ is the candidate cell state at the current moment, o_t is the output value of the output gate, C_t is the cell state at the current moment, and h_t is the output of the current moment. Gates in LSTM are a selective path for information to enter. It consists of a layer of sigmoid function calculations and point multiplication operations. Input–output and forgetting gates work together to protect and control the state of smart grid cells. The sigmoid layer outputs values between 0 and 1 describing how much information can be passed in each part of the information, 0 for "no amount allowed" and 1 for "any amount allowed." The LSTM calculation process is shown as follows.

First, the current output of the amnesia gate is calculated from the output value h_{t-1} of the LSTM neuron unit at the previous moment and the current input data x_t, and the information to be forgotten is selected, as shown in (1). The values of f_t in the formula range from 0 to 1, W_f is the weight matrix of the gate, and L is the offset of the gate

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right). \tag{1}$$

Second, according to the output value h_{t-1} of the LSTM neuron unit at the previous moment, the input gate output i_t at the current moment, the input data x_t at the current moment, and the candidate cell state $\hat{C}_t$ at the current moment, the information to be memorized is calculated and selected, that is, to determine what new information is stored in the smart grid cell state. The calculation process is shown in the following equations:

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

$$\hat{C}_t = \tanh\left(W_C\left[h_{t-1}, x_t\right] + b_C\right). \tag{3}$$

The values of i_t range from 0 to 1, W_i is the weight matrix of the input gates, b_i is the input bias, W_C is the weight of the candidate input gates, and b_C is the bias of the candidate input gates.

Then, the state C_t of the current smart grid cell is calculated from the internal memory cell. As shown in (4), the value of C_t ranges from 0 to 1

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t. \tag{4}$$

Finally, the hidden layer state values of the output gate and the current moment are calculated as shown in the following equations:

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{5}$$

$$h_t = o_t * \tanh\left(C_t\right). \tag{6}$$

Through the above calculation process, we can get a sequence of hidden layer states with the same length as the original data. When the BiLSTM is used in the classification of time series $\{h_0, h_1, h_2, \ldots, h_{n-1}\}$, the forward extracted hidden layer state sequence and the backward extracted hidden layer state sequence are stitched together to form a time-series two-way feature fusion sequence, which can be used as an input for subsequent classification network layer. However, not all the extracted features have the same contribution to the classification task, some of them have commonalities or interfere with the classification results. Therefore, how to enlarge the role of contributing features in classification tasks and reduce the impact of invalid and irrelevant features on classification results is particularly important. The introduction of AM is an excellent solution to this problem.

B. Principles of AM

Attention originates from the human visual system. When people look at things, they remember some key information

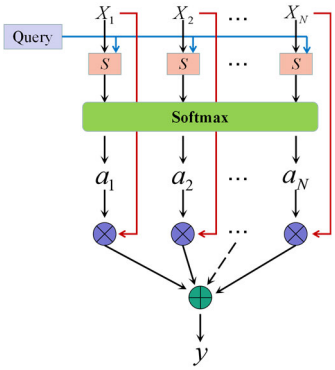


Fig. 2. Basic structure of the AM.

Fig. 3. Experimental system for laser welding.

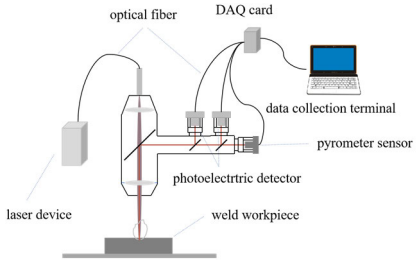


Fig. 4. Diagram of data collection design for laser welding process.

for some purpose or instinctive purpose, such as shape, color, size, and so on. They also ignore some visible information. When observing the outside world, human beings use this physiological mechanism to allocate their attention through the brain to obtain more important target information. This allows people to use limited attention resources to quickly get useful information about their goals from a large amount of information. In deep learning, neural network is a kind of brain-like computing model, so the AM in neural network is different from that of human visual nervous system. As a result, AMs excel in many scenarios in the field of deep learning. Attention-based neural networks have attracted great interest from researchers all over the world in the study of in-depth learning.

AMs can be divided into hard-AM and soft-AM according to their differentiability. Hard AMs select one of the input vectors as the output of the mechanism based on the distribution of attention, which results in an area being either fully concerned or not at all. Therefore, the hard AM is an indispensable attention allocation scheme, which makes the functional relationship between the overall loss function and the attention distribution of the model indistinguishable and makes it impossible to train the model using the back-propagation algorithm. In the research and use of in-depth learning models and AMs, most of them are calculated in the form of soft attention. Soft AM uses attention distribution to weight sum and fuse input vectors and uses different weighting coefficients from 0 to 1 to indicate the degree of attention each area receives. The larger the value assigned by the weight factor, the more attention allocated, that is, the more important the corresponding information is.

The calculation of AM is essentially to assign attention weights to all input information. The mathematical calculation process of the mechanism can be expressed as (7) and (8). W_a is the AM's weight matrix, which represents the degree to which different location information needs to be emphasized. e_n is the result of the first weighting calculation, b indicates a bias in the AM, $X = [X_1, X_2, \ldots, X_N]$ is the AM's input, and a_n represents the attention weight obtained by entering $X = [X_1, X_2, \ldots, X_N]$. Generally, there are two steps in the AM: 1) calculating the distribution of attention and 2) calculate a weighted average of the input information based on the attention distribution. Fig. 2 gives a schematic of the AM's basic structure, in which $X = [X_1, X_2, \ldots, X_N]$ represents N items of input information, X is called source in the industry, and query represents the corresponding query space after partitioning the input information. First, the input information X and query are passed to the attention score function s, which can be designed and selected in two ways: additive attention and zooming point product attention. Subsequently, the score results are entered into the Softmax layer to obtain the attention weight factor $[a_1, a_2, \ldots, a_N]$ of the input information. Finally, the input information is used to weighted average the attention weight factor vector to get the final result

$$e_n = \tanh\left(W_a\left[X_1, X_2, \ldots, X_N\right] + b\right) \tag{7}$$

$$a_n = \frac{\exp\left(e_n\right)}{\sum_{k=1}^{N} \exp\left(e_k\right)}. \tag{8}$$

III. Design of Data Acquisition System for Laser Welding Process Based on Sensor

According to the law found in the experiment, the designed laser welding process data acquisition system mainly includes a Southampton Photonics Inc. (SPI) pulse fiber laser, a galvanometer scanner, a flat field lens, two vibration mirrors, and a sensor-based front-end data acquisition system. Fig. 3 demonstrates the laser welding experimental system. By turning the vibration mirror to change the laser path, the galvanometer scanner can reflect the laser to the desired position.

Front-end signal acquisition consists of two photoelectric detectors, a pyrometer sensor, and some optical elements. The overall schematic of the data collection system is shown in Fig. 4. The acquisition equipment and laser welding equipment were coaxial installed. The PDA10A2 photodetector in wavelength range of 200–1100 nm is used to collect plasma and light intensity signals. The pyrometer sensor uses a PT2400 high-temperature optical fiber sensor in the range of

−40 °C to 750 °C to monitor temperature changes during the welding. The data acquisition card is used to collect the electrical signal generated by the sensors, and the collected information is analyzed and processed on the computer.

The standard welded parts are made of SUS304 stainless steel plate with a thickness of about 0.3 mm and a radius of 22 mm. A continuous wave optical fiber laser was used in the experiment. The laser output power was 80 W and the welding speed was 50 mm/s. During the data collection experiment, 130 sampling time points are set for each standard workpiece during the welding process, and plasma strength, light intensity, and temperature values are collected at each sampling time.

According to the experimental device and data collection scheme, we have collected and constructed a dataset of 6467 samples of laser welding quality, including nine welding defect states and one welding qualification state. For the collected welding process data, standardization is required in order to eliminate errors caused by different dimensions or large differences in values. We converted the welding quality data into a standard dataset with a mean value of 0 and a variance of 1. The ten welding quality states in the dataset are "qualified," "defocus 3 mm," "defocus-3 mm," "deformation," "cracks," "repetition," "lack of weld," "drift," "tilt, " and "watermarks, " respectively.

IV. MSCNN and BiLSTM Fusion Model for Laser Welding Defect Detection

A. Model Design

For each sample, the monitoring data of the welding process have time-series characteristics. Although the CNN can extract the spatial characteristics of the laser welding quality data, only relying on the local spatial features of CNN will ignore the time-series characteristics contained in the data. The LSTM is suitable for processing time-series data. Relative to RNN, the LSTM considers the dependence and correlation between long-distance time-series information. However, because it only considers one-way data processing, the extracted feature information is insufficient. The BiLSTM is composed of forward LSTM and backward LSTM, which essentially solves the problem of LSTM. The forward LSTM is used to capture the above information, while the backward LSTM is used to capture the context information. Finally, the context data features are fused to provide a more comprehensive understanding of the overall time-series characteristics of the data. For complex features and large amounts of data, the application of AM can highlight important feature information that contributes a lot to the final classification result after feature fusion and improve the upper limit of model performance. Combining various theories and practices, we propose to cascade CNN and BiLSTM and introduce AM. The CNN-BiLSTM-AM integrated model will take into account the advantages of both CNN and BiLSTM, which can extract the space-time characteristics of the laser welding quality data and solve the problem of long-term dependence on missing information. Moreover, the added AM can filter out the insignificant noise information in the fused feature information, reduce or even eliminate its impact on the final laser welding quality status

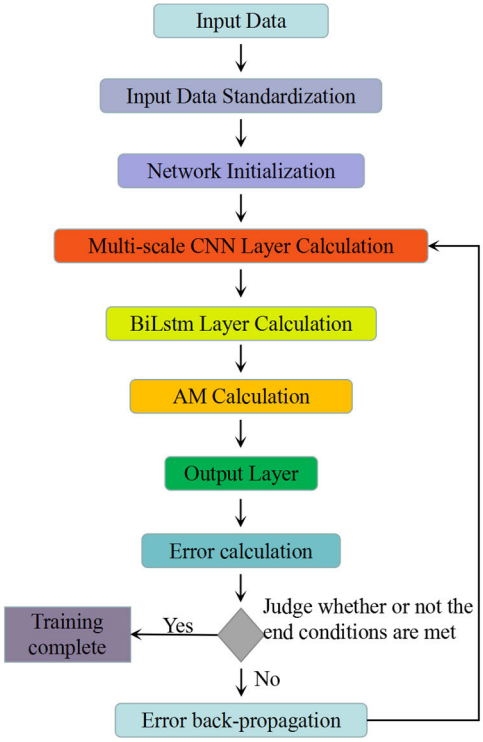


Fig. 5. Diagram of MSCNN-BiLSTM-AM training process.

recognition, and highlight the role of useful information in the fused feature information. Here, the AM used in the fusion model is self-attention.

In the fields of classification and recognition, the richer the diversity of features mastered by general models, the better the model performances. Single-scale convolution nuclei can only extract local regional features within the convolution nucleus's receptor field, so some potential information under larger or smaller receptor fields may be lost. After theoretical demonstration and experimentation, we finally propose to use multiscale convolution kernel to extract characteristic information under different field conditions. Extracting features using MSCNN needs moderation, because using different convolution cores will increase the order of the overall network parameters and the overall training cost of the network. Fig. 5 shows the MSCNN-BiLSTM-AM integrated model training flow diagram. In terms of data preprocessing, standardized processing strategies are still used to eliminate the impact of data dimension and magnitude differences. For the design of CNN, BiLSTM, and AM core functional layers, we use some experimental experience and design strategies, by taking into account the overall lightweight model and giving full play to the performance advantages of each submodule.

For the integration model of MSCNN-BiLSTM-AM fusion, we have made sufficient experimental comparisons and design considerations. In order to extract different types of features, we propose a multiresolution 1-D convolution structure for CNN design. We implement two convolution layer branches with different convolution core sizes, which are designed to select time step ranges with different kernel sizes to extract different spatial features from the original data. Each branch contains two convolution layers, in which each contains a batch normalization layer and an ReLU activation function.

TABLE I
NETWORK STRUCTURE OF MSCNN–BiLSTM–AM MODEL

Operation	Kernel	Strides	Feature Maps	BN?	Dropout	Nonlinearity
input (1 × 3 × 130)						
Sequential(Branch One)						
Convolution	8 × 1	1 × 1	1 × 16 × 125	√	0.0	ReLU
Convolution	8 × 1	1 × 1	1 × 32 × 120	√	0.0	ReLU
Bi-LSTM	N/A	N/A	1 ×120 × 128, 2 × 1 × 64	N/A	0.0	N/A
Attention	N/A	N/A	1 × 128	N/A	0.0	N/A
Sequential(Branch Two)						
Convolution	32 × 1	1 × 1	1 × 16 × 101	√	0.0	ReLU
Convolution	32 × 1	1 × 1	1 × 32 × 72	√	0.0	ReLU
Bi-LSTM	N/A	N/A	1 × 72 × 128, 2 × 1 × 64	N/A	0.0	N/A
Attention	N/A	N/A	1 × 128	N/A	0.0	N/A
Concat(Feature Integration)						
Concat	N/A	N/A	1 × 256	×	0.0	N/A
Linear	N/A	N/A	1 × 128	×	0.0	ReLU
Linear	N/A	N/A	1 × 10	×	0.0	N/A

In designing LSTM, we use a BiLSTM layer on each branch to extract time-series features. BiLSTM adds sequence reverse information to LSTM, which extracts more time-series features than one-way LSTM and generally outperforms one-way LSTM. Next, we fuse and stitch the features extracted from the two branches and then input the feature information of the two branches into two linear models. Finally, the ReLU activation function can be used to get the classified output. Table I shows the network structure of the MSCNN-BiLSTM-AM model.

B. Experimental Discussion

In this study, the performance of weld defects identification is assessed by using the following three classical evaluation indexes of classification problem, i.e., recall, precision, and $F1$-score ($F1$):

$$\text{Precision} = \left(1 - \frac{N_+^-}{N^+ - N_-^+ + N_+^-}\right) \tag{9}$$

$$\text{Recall} = \left(1 - \frac{N_-^+}{N^+}\right) \tag{10}$$

$$F_1 = 2 \cdot \frac{\text{Pre} \cdot \text{Rec}}{\text{Pre} + \text{Rec}} \tag{11}$$

where N^+ is the total number of weld defect, N_-^+ is the number of weld defect incorrectly predicted as nonweld defect, N^- is the total number of nonweld defect, and N_+^- is the number of nonweld defect incorrectly predicted as weld defect.

This part uses the neural network model to learn the time-space characteristics of plasma strength, light intensity, and temperature change data in laser welding process and combines AM to improve the upper limit of the integrated model to identify and classify the quality of laser welding as a whole. The performance changes of the models with and without AM have been compared in this study. Under the condition

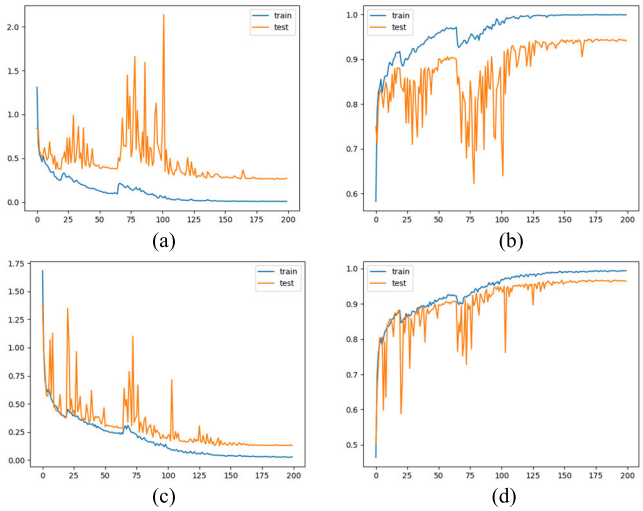


Fig. 6. Comparisons of (a) and (c) loss function and (b) and (d) recognition accuracy changes on training and test sets between (a) and (b) MSCNN–BiLSTM and (c) and (d) MSCNN–BiLSTM–AM models without data enhancement conditions.

without and with data enhancement, the MSCNN-BiLSTM-AM model demonstrates high-quality detection performance of laser welding, in which no data enhancement condition refers to that the model only uses 6467 original laser welding quality datasets (5174 training data and 1293 test data), while using data enhancement condition refers to training the designed model by expanding the training set of original data with 1147 enhanced quality data generated in previous research papers [26].

In the absence of data enhancement, Fig. 6(a) and (b) shows the change curves of MSCNN-BiLSTM loss function and recognition accuracy on training and test sets, respectively, while Fig. 6(c) and (d) shows the change curves of MSCNN-BiLSTM-AM loss function and recognition accuracy

TABLE II
PERFORMANCE METRICS OF MSCNN–BiLSTM INTEGRATED MODEL IN EACH CATEGORY WITH DATA ENHANCEMENT

Class / Metrics	Qua (%)	Def3 (%)	Def-3 (%)	Defor (%)	Cra (%)	Rep (%)	LoW (%)	Dri (%)	Tilt (%)	W (%)
Precision	100	98.36	99.18	98.44	99.18	99.26	98.29	100	100	93.65
Recall	100	99.17	100	94.74	100	99.26	99.14	98.44	97.67	98.33
F1	100	98.76	99.59	96.55	99.59	99.26	98.71	99.21	98.82	95.93

TABLE III
PERFORMANCE METRICS OF MSCNN–BiLSTM–AM INTEGRATED MODEL IN EACH CATEGORY WITH DATA ENHANCEMENT

Class / Metrics	Qua (%)	Def3 (%)	Def-3 (%)	Defor (%)	Cra (%)	Rep (%)	LoW (%)	Dri (%)	Tilt (%)	W (%)
Precision	100	99.17	99.18	100	99.18	99.27	100	100	100	96.74
Recall	100	99.17	100	97.74	100	100	99.14	99.22	99.22	99.17
F1	100	99.17	99.59	98.86	99.59	99.63	99.57	99.61	99.61	97.94

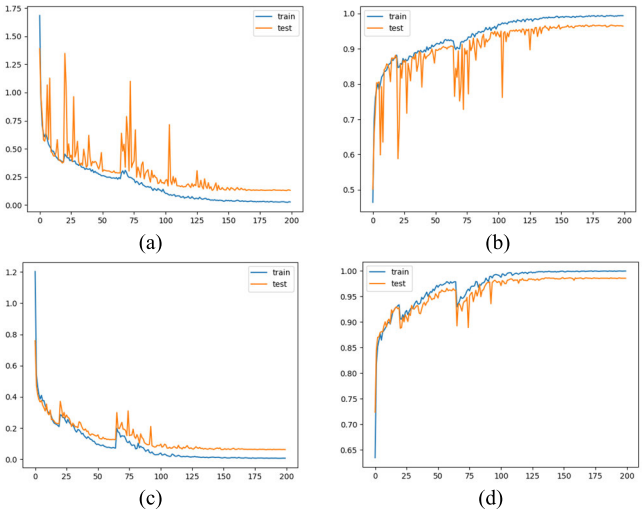


Fig. 7. Comparison of (a) and (c) loss function and (b) and (d) recognition accuracy changes on training and test sets between (a) and (b) MSCNN–BiLSTM and (c) and (d) MSCNN–BiLSTM–AM models with data enhancement conditions.

TABLE IV
CLASSIFICATION ACCURACY OF MSCNN–BiLSTM–AM AND DIFFERENT MODELS UNDER DIFFERENT AMOUNTS OF TRAINING DATA

	No data enhancement	Data Enhancement
MSCNN-BiLSTM-AM	98.49%	99.38%
MSCNN-BiLSTM	98.37%	98.68%
CNN-BiLSTM-AM	97.86%	97.33%
CNN-AM	96.91%	97.15%
ACGAN-SVM-CNN	\	98.37%
CNN	96.83%	96.88%
SVM	83.35%	81.47%
Naive Bayes	75.87%	75.89%
Random Forest	86.43%	86.96%

on training and test sets, respectively. It can be seen that the network can converge after 200 epochs. After the 150th epoch, the performance of the model is stable on both the training set and the test set, but whether there is or not AM, the model still shows a large fluctuation on the test set during the pretraining period. Under these conditions, the MSCNN-BiLSTM model achieves maximum classification accuracy of 98.37% on the test set, while the MSCNN-BiLSTM-AM model achieves maximum classification accuracy of 98.49% on the test set.

With data enhancement, Fig. 7(a) and (b) shows the change curves of MSCNN-BiLSTM loss function and recognition accuracy on training set and test set, respectively, and Fig. 7(c) and (d) shows the change curves of MSCNN-BiLSTM-AM loss function and recognition accuracy on training set and test set, respectively. It can be seen that both MSCNN-BiLSTM and MSCNN-BiLSTM-AM models converge more quickly with sufficient sample diversity, and the overall loss function and model recognition accuracy fluctuate less. Especially from Fig. 7(d), we can see that the performance of MSCNN-BiLSTM-AM model is more stable and superior on the test set, and the classification recognition accuracy can reach to 95% at the 50th epoch. On the one hand, the real improvement of data diversity enhances the model's understanding of the potential rules of empirical data. On the other hand, the AM plays a positive role in promoting the model recognition ability when the data richness is sufficient. Under this condition, the highest classification accuracy of the MSCNN-BiLSTM model on the test set reaches to 98.68%, while the highest classification accuracy of the MSCNN-BiLSTM-AM model is 99.38%.

With data enhancement, Table II is the ability of the MSCNN-BiLSTM fusion algorithm model to identify each laser welding quality category. Table III is the classification performance index of the MSCNN-BiLSTM-AM fusion algorithm model in each laser welding quality category. The abbreviations "Qua," "Def3," "Def-3," "Defor," "Cra," "Rep," "LoW," "Dri," "Tilt," and "W" in the table represent "qualified," "defocus 3 mm," "defocus-3 mm," "defocus-3 mm," "deformation," "cracks," "repetition," "lack of weld," and "defocus of weld," respectively. There are ten welding quality states, drift, tilt, and watermarks. The results show that the designed detection model maintains high sensitivity to the "qualified" quality state and greatly improves the

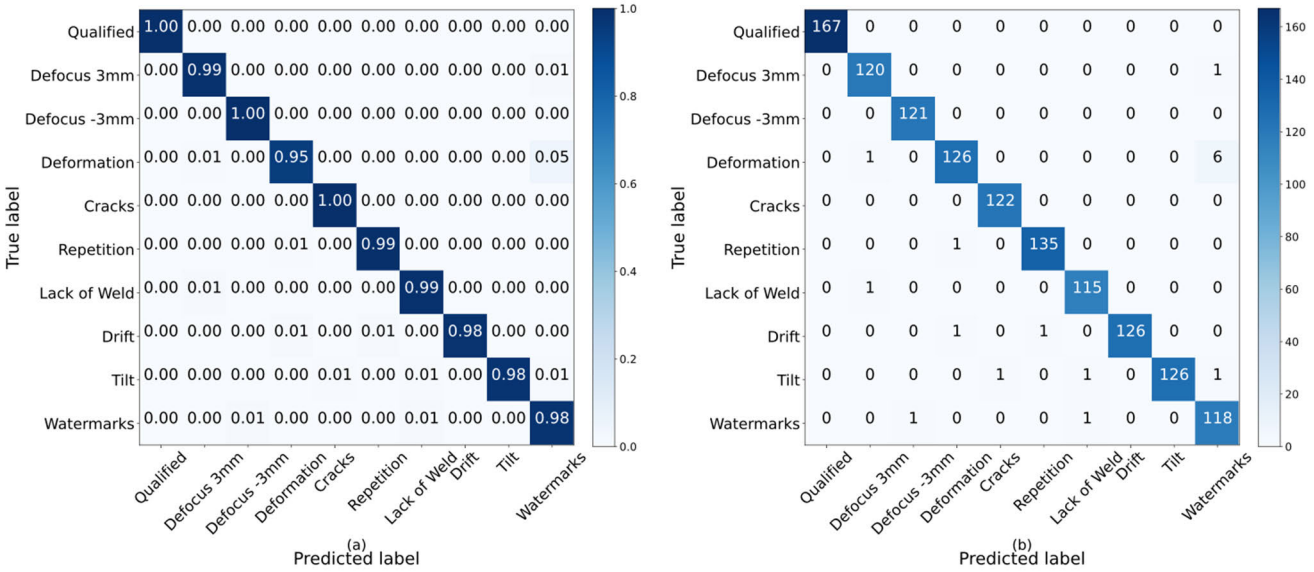


Fig. 8. Confusion matrix of MSCNN–BiLSTM integrated model for welding defect identification. (a) Confusion matrix of the classification result. (b) Confusion probability matrix of the classification result.

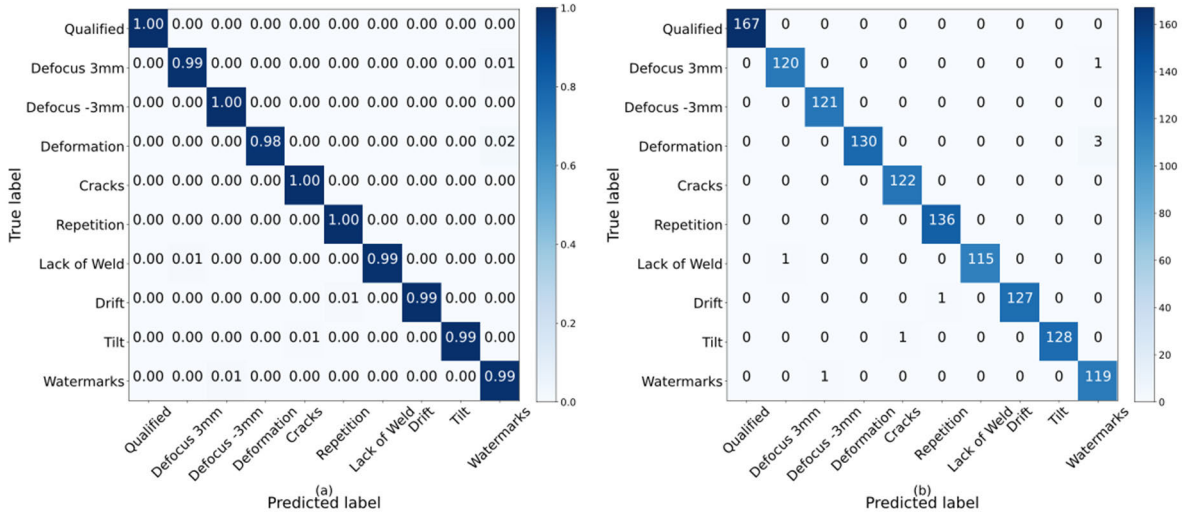


Fig. 9. Confusion matrix of MSCNN–BiLSTM–AM integrated model for welding defect identification. (a) Confusion matrix of the classification result. (b) Confusion probability matrix of the classification result.

detection capability of the defect category "defocus 3 mm." Figs. 8 and 9 show the confusion matrix diagrams obtained from MSCNN-BiLSTM and MSCNN-BiLSTM-AM cross-validation on the enhanced dataset, respectively.

To verify the superiority of the overall performance of MSCNN-BiLSTM-AM, the proposed fusion algorithm model is compared with single-scale CNN, SVM [15], Naive Bayes [27], random forest, and gradient boosting decision tree (GBDT) [28] under different data volume conditions. It can be seen that the performance of the proposed improved detection model is better than those of the previous schemes.

Ablation experiments have been performed on the proposed MSCNN-BiLSTM-AM model to compare the importance of different functional submodules in MSCNN-BiLSTM-AM for overall performance. The ablation experiment is similar to the control variable method in that it deletes part of the network in the fusion model and studies the performance changes of the network after the deletion to understand the role and role of the deleted part in the overall model. In Table IV, by comparing the first five rows of data, MSCNN-BiLSTM-AM has the best ability to identify welding defects. Adding BiLSTM and AM can improve the overall recognition performance of the model. Further comparisons show that the introduction of MSCNN has a more significant effect on model promotion than BiLSTM and AM. The experimental results verify the feasibility of our designed method. In addition, even with increasing sample diversity, some traditional machine learning methods (such as SVM and Naive Bayes model in Table IV) still fail to learn more potential regularity information from the expanded samples, which confirms that some traditional machine learning models are unable to handle complex industrial application scenarios, while deep learning models are

superior. The multimodel fusion strategy also makes the fusion model more suitable for the analysis and processing of specific scenarios and data by taking advantage of the best practices.

V. Conclusion

The main content of this article is to propose an MSCNN-, BiLSTM-, and AM-based intelligent detection scheme for laser welding defects (MSCNN–BiLSTM–AM). Because the integration of multiple models can take into account the advantages of different models, the fusion model proposed in this article can not only extract the spatial feature extraction ability of CNN but also extract the time-series association information of features using BiLSTM. Meanwhile, the introduction of AM in the appropriate location of the network reduces or filters the impact of "noise" features on the final classification results and improves the recognition performance of the overall detection scheme. In addition, when designing the CNN, we tried the multichannel and MSCNN and BiLSTM cascade, which fused the feature information in different fields, greatly enriched the diversity of features learned by the model, and made it possible to train a better depth model. Finally, we verify the superiority of the improved scheme proposed in this article through a large number of comparison and ablation experiments and verify that the advantages of introducing the subalgorithm model in the theoretical design are reflected in the integrated model.

References

[1] K. I. Yaakob, M. Ishak, M. M. Quazi, and M. N. M. Salleh, "Optimizing the pulse wave mode low power fibre laser welding parameters of 22Mnb5 boron steel using response surface methodology," *Measurement*, vol. 135, pp. 452–466, Mar. 2019.

[2] K. He, Q. Li, and J. Chen, "An arc stability evaluation approach for SW AC SAW based on Lyapunov exponent of welding current," *Measurement*, vol. 46, no. 1, pp. 272–278, Jan. 2013.

[3] Y. Y. Qian, X. K. Sun, S. Zhou, and Y. M. Wang, "Research on vibration suppression for rail welding head straightness detecting system," *Meas. Control Technol.*, vol. 36, no. 2, pp. 75–79, 2017, doi: 10.3969/j.issn.1000-8829.2017.02.018.

[4] H. Shen, J. Jin, B. Liu, and Z. Zhou, "Measurement and evaluation of laser-scanned 3D profiles in wire arc hybrid manufacturing processes," *Measurement*, vol. 176, May 2021, Art. no. 109089.

[5] Z. Zhang, Y. Huang, R. Qin, Z. Lei, and G. Wen, "Real-time measurement of seam strength using optical spectroscopy for Al–Li alloy in laser beam welding," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[6] Y. Zou and R. Lan, "An end-to-end calibration method for welding robot laser vision systems with deep reinforcement learning," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4270–4280, Jul. 2020.

[7] Y. Ma et al., "A fast and robust seam tracking method for spatial circular weld based on laser visual sensor," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.

[8] W. Su, W. Luo, X. Dong, and R. Liu, "Residual stress analysis of laser welding between output needle and helix," *J. Phys., Conf. Ser.*, vol. 1802, no. 2, Mar. 2021, Art. no. 022088.

[9] A. Tsukui, "Laser welding method and laser welding apparatus," US20 210 213 563 A1 Jul. 15, 2021.

[10] F. Maiwald, S. Englmaier, and S. Hierl, "Online pyrometry for weld seam localization in absorber-free laser transmission welding of transparent polymers," *J. Laser Micro/Nanoeng.*, vol. 16, no. 1, pp. 1–6, 2021.

[11] J. Xue, M. H. R. Dehkordi, A. Abdelahi, A. Abdelahi, E. Rasti, and Z. Li, "Experimental investigation of temperature field, defects, and mechanical strength in dissimilar laser bonding of Ti6Al4 V and polyethylene terephthalate," *J. Laser Appl.*, vol. 33, no. 1, Feb. 2021, Art. no. 012038.

[12] N. Kumar, N. Kumar, and A. Bandyopadhyay, "A state-of-the-art review of laser welding of polymers—Part I: Welding parameters," *Weld. J.*, vol. 100, no. 7, pp. 221–228, Jul. 2021.

[13] M. Oikawa, K. Atsumi, Y. Otsuka, and N. Kawada, "Development of condition monitoring system for electric resistance spot welding used to manufacture railway car bodies," *J. Robot. Mechatronics*, vol. 33, no. 2, pp. 421–431, Apr. 2021.

[14] E. W. Teichmann, J. Kelbassa, A. Gasser, S. Tarner, and J. H. Schleifenbaum, "Effect of wire feeder force control on laser metal deposition process using coaxial laser head," *J. Laser Appl.*, vol. 33, no. 1, Feb. 2021, Art. no. 012041.

[15] C. Knaak, J. von Eßen, M. Kröger, F. Schulze, P. Abels, and A. Gillner, "A spatio-temporal ensemble deep learning architecture for real-time defect detection during laser welding on low power embedded computing boards," *Sensors*, vol. 21, no. 12, p. 4205, Jun. 2021.

[16] M. M. Naddaf-Sh et al., "Defect detection and classification in welding using deep learning and digital radiography," in *Fault Diagnosis and Prognosis Techniques for Complex Engineering Systems*, 2021, pp. 327–352.

[17] D. Medak, L. Posilovic, M. Subašic, M. Budimir, and S. Loncaric, "Deep learning-based defect detection from sequences of ultrasonic B-scans," *IEEE Sensors J.*, vol. 22, no. 3, pp. 2456–2463, Feb. 2022.

[18] Y. Cheng, H. G. Deng, Y. X. Feng, and J. J. Xiang, "Weld defect detection and image defect recognition using deep learning technology," Tech. Rep., 2021, doi: 10.21203/rs.3.rs-149365/v1.

[19] W. Zhao, F. Chen, H. Huang, D. Li, and W. Cheng, "A new steel defect detection algorithm based on deep learning," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–13, Mar. 2021.

[20] L. Yang, H. Wang, B. Huo, F. Li, and Y. Liu, "An automatic welding defect location algorithm based on deep learning," *NDT E Int.*, vol. 120, Jun. 2021, Art. no. 102435.

[21] Y. C. Kan and H. Kalkan, "Automatic detection and classification of laser welding defects," in *Proc. Innov. Intell. Syst. Appl. Conf. (ASYU)*, Oct. 2021, pp. 1–5.

[22] H. Deng, Y. Cheng, Y. Feng, and J. Xiang, "Industrial laser welding defect detection and image defect recognition based on deep learning model developed," *Symmetry*, vol. 13, no. 9, p. 1731, Sep. 2021.

[23] F. Deng et al., "A multi-sensor data fusion system for laser welding process monitoring," *IEEE Access*, vol. 8, pp. 147349–147357, 2020.

[24] R. Miao et al., "Real-time defect identification of narrow overlap welds and application based on convolutional neural networks," *J. Manuf. Syst.*, vol. 62, pp. 800–810, Jan. 2022.

[25] S. Lü, D. Hongchang, X. Yang, and L. Yongkun, "Surface welding defect detection based on Michelson interferometer," *Laser Optoelectron. Prog.*, vol. 58, no. 19, 2021, Art. no. 1912005, doi: 10.3788/LOP202158.1912005.

[26] K. Fan, P. Peng, H. Zhou, L. Wang, and Z. Guo, "Real-time high-performance laser welding defect detection by combining ACGAN-based data enhancement and multi-model fusion," *Sensors*, vol. 21, no. 21, p. 7304, Nov. 2021.

[27] Y. Zhu et al., "A lightweight multiscale attention semantic segmentation algorithm for detecting laser welding defects on safety vent of power battery," *IEEE Access*, vol. 9, pp. 39245–39254, 2021.

[28] Z. Xiao-Yu, Y. Yun-Yang, L. Yi-An, and G. Shang-Bing, "Flame detection based on GBDT feature for building," in *Proc. Int. Smart Cities Conf. (ISC)*, Sep. 2017, pp. 1–5.