WELDING RESEARCH

SUPPLEMENT TO THE *WELDING JOURNAL*, AUGUST 2023
Sponsored by the American Welding Society

Weld Geometry Prediction Based on Binocular Vision and Deep Learning

CNN structure was developed to predict weld size

BY S. Y. WANG, J. CHEN, C. Y. XIA, R. D. LI, W. B. ZHANG, AND C. S. WU

Abstract

To improve the level of welding automation in the industry, there are increasing requirements for highly intelligent and accurate inspections of the welding process in real time. This paper proposed a new method for predicting weld dimensions based on binocular imaging information and a deep learning system. The binocular imaging information was acquired by binocular vision equipment and an image processing algorithm. A convolutional neural network structure was developed by adding a fully connected block and loss function judgment. A new calculating procedure was proposed to extract and link the information of the processed weld pool image and the weld parameters effectively. With the help of 7394 training samples, the results of 1849 testing samples showed that the overall accuracy of the proposed model was higher than 93% for the prediction of weld dimensions, which could meet the requirements in practical applications.

Keywords

- Weld Geometry
- Deep Learning
- Binocular Vision
- Neural Network
- Gas Metal Arc Welding

Introduction

Due to its advantages of high efficiency, good reliability, and ease of automation, gas metal arc welding (GMAW) has been widely applied in various modern equipment manufacturing industries, such as ships, automobiles, pressure vessels, pipeline welding, etc. (Refs. 1, 2). It is often combined with robots for automatic welding in practical industrial production, but it is still difficult to achieve precise real-time regulation with machines during the complex welding process (Ref. 3). To improve the intelligent automation level of welding technology, scholars have paid a lot of attention to monitoring and detecting of the welding process.

In welding research, the geometry of the weld bead is a vital factor in evaluating the quality of the weld bead (Ref. 4). The method for predicting the status of the welding process by using the optical, electrical, and acoustic signals is widely accepted. The neural network models provide fairly accurate results for static modeling in a simple and fast way (Ref. 5). Zeng et al. trained multilayer neural networks to predict the optimal welding parameters of manual gas tungsten arc welding (GTAW) based on materials, weld position, and joint preparation (Ref. 6). Pernambuco et al. identified the types of discontinuities in the welding process based on an artificial neural network (ANN) and arc sound signal (Ref. 7). Cui et al. utilized a multisensor sensing system and support vector machine (SVM) to identify the joint penetration status (partial joint, complete joint, and excessive penetration) of Keyhole Tungsten Inert Gas welding from the signals of arc sound, welding current, and arc voltage (Ref. 8). Nagesh et al. tried to predict the value of reinforcement, weld width, and joint penetration by an ANN and welding parameters. However, the results showed that the maximum predicting error reached up to 22.39% (Ref. 9). That meant weld pool information should be provided to exactly predict the status of the welding process. Zhang et al. pointed out that the weld pool surface provided essential information for the understanding of the welding process, which helped human welders to make correct decisions to adjust robot parameters (Ref. 10). Therefore, the size of the weld pool surface is critical to the prediction of weld bead formation.

Lei et al. combined principal component analysis and a genetic algorithm to establish a neural network. They took

https://doi.org/10.29391/2023.102.014

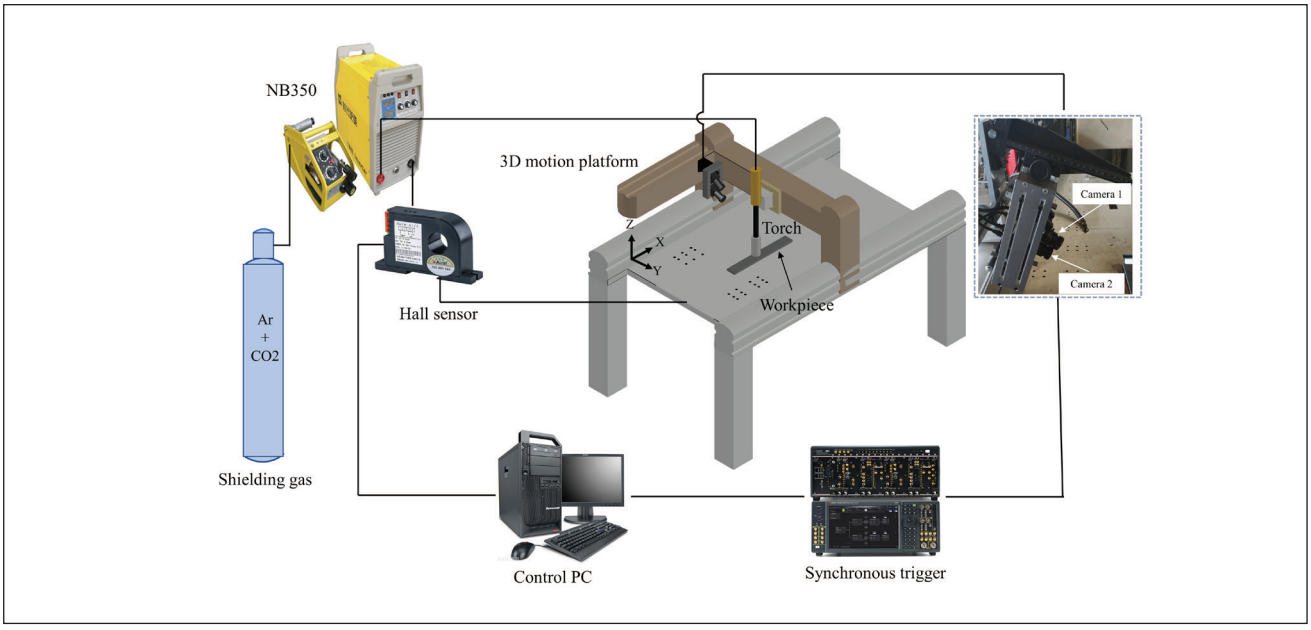


Fig. 1 — Schematic diagram of the hardware system for the experiment.

the weld pool surface's morphology and welding parameters as inputs to predict the weld waist width and weld back width (Ref. 11). Chang et al. developed a melt depth prediction model by optimizing a back propagation neural network (BPNN) with a genetic algorithm. The welding current, arc voltage, and weld pool characteristic parameters were used as inputs to predict joint penetration. Though the relative error of prediction was less than 4.6% (Ref. 12), the input data of weld pool width and length were easily affected by the capturing angle of a single camera. Martínez et al. established a welding geometry predicting framework for the GMAW process with a residual network (ResNet) model and support vector machine-kernel radial basis function. This predicting framework had a prediction error of ± 0.27 mm in the weld geometry (Ref. 13). Lu et al. proposed a molten pool online monitor model based on the video prediction framework (PredNet) and the convolutional classification network (LeNet-5). This model achieved 95% accuracy in predicting hump and penetration status, which provided a sufficient warning for the welding process (Ref. 14). For a single bevel GMAW with gap fluctuation, Nomura et al. built two types of convolutional neural network (CNN) models with the help of the top surface image of the weld pool to predict burn-through and joint penetration, respectively (Ref. 15). Comparing the results of actual and estimated penetration depth showed that the estimation accuracy still needed to be improved as the 11% value of absolute error in prediction was higher than 0.5 mm. Based on an active vision sensing system and the CNN, Cheng et al. proposed an innovative method to identify the penetration state of GTAW. The recognition accuracy could reach 98%. However, the laser reflection stripes in the above experiments were easily disturbed (Ref. 16).

Skilled welders can estimate the quality of welds based on the weld pool deformation they see. To reconstruct the three-dimensional (3D) information of weld pool deformation, Zhang et al. projected a structured light laser pattern lattice on the surface of the weld pool and imaged the specular reflection on an imaging plane to simulate the function of the welder's eyes (Refs. 17–19). Based on welding and shape parameters of the top surface of the weld pool, the dynamic neural network model was developed to predict the backside width of the weld pool during pulsed GTAW. Test data results indicated that the average prediction error of backside width was less than 6.3% (Ref. 20). However, this method required ideal experimental conditions, and it was difficult to obtain the laser spot, which was reflected from

Table 1 — Setup Parameters of the Two Cameras

Camera Parameters	Value
Exposure time	35 µs
Aperture	F5.6
Camera angle	30 ~ 38 deg
Working distance	200 ~ 300 mm
Capture frame rate	200 fps
Region of interest (ROI)	600 × 500 pixels

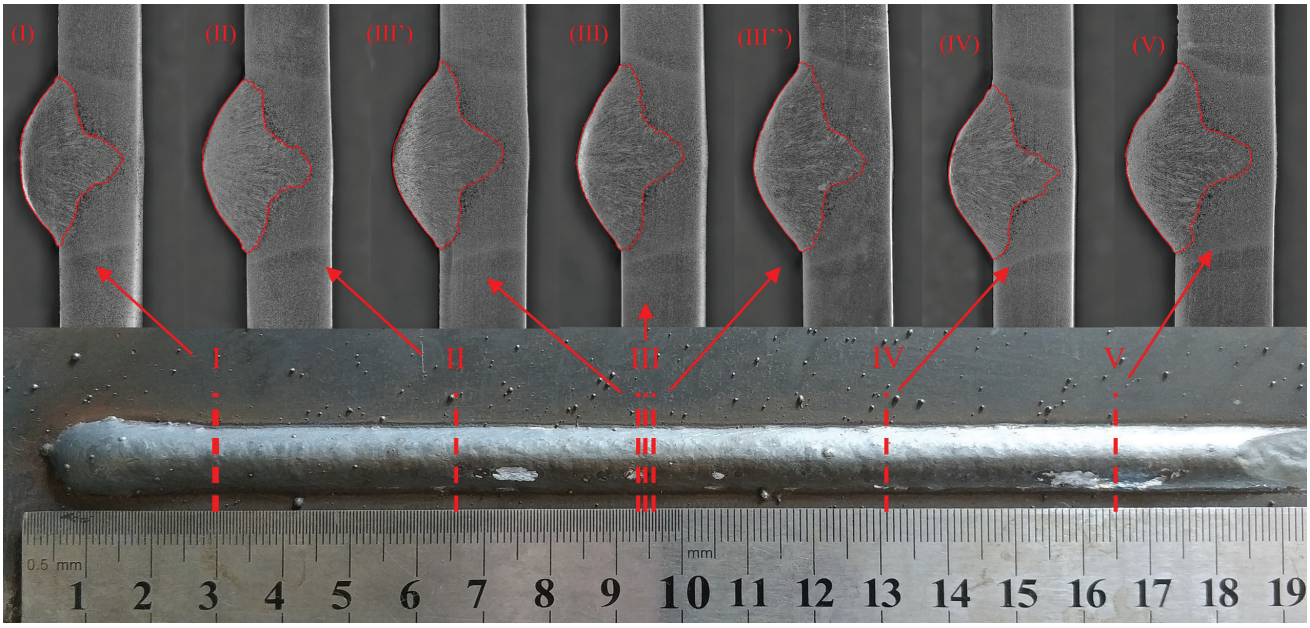


Fig. 2 — Metallographic images of weld sections for parameter 27.

Table 2 — Welding Parameters

Test Sample No.	Scheme of Test				Orthogonal Scheme
	A	B	C	D	
1	30	210	0.6	18	6, 2, 9, 9
2	25	270	0.85	12	1, 8, 8, 3
3	26	280	0.85	10	2, 9, 8, 1
4	33	210	0.75	16	9, 2, 6, 7
25	26	210	0.5	11	2, 2, 1, 2
26	32	270	0.8	17	8, 8, 1, 8
27	30	230	0.65	16	6, 4, 2, 7
80	26	270	0.8	18	2, 8, 7, 9
81	28	210	0.85	15	4, 2, 8, 6

the large deformed and fluctuated weld pool. Liang et al. constructed a prismatic stereo vision system and proposed a two-step matching algorithm based on the Harris algorithm, which efficiently reconstructed the weld pool surface in different penetration states for the pulsed gas metal arc welding process (Ref. 21). All the research mentioned above showed that there was an important correlation between the weld pool image and weld bead produced. In this paper, a prediction system of weld dimensions was proposed that included binocular vision equipment, an image processing system, and a deep learning system. The binocular vision equipment was utilized to capture the images of the weld pool from different shooting angles to avoid image distortion in the image processing process. The image processing system was applied to reconstruct the 3D weld pool image and transfer the 3D image into depth image. A further developed CNN structure was put forward to form the deep learning system and predict the weld dimensions based on the depth image of the weld pool and welding parameters. The model had high accuracy in predicting joint penetration, weld width, and reinforcement, which provides a foundation to monitor and control the welding process in real time in the future.

Experimental Design

Experimental Platform

As shown in Fig. 1, the entire hardware system of the experiment was mainly composed of three parts: GMAW system, Hall-sensor system, and binocular vision equipment. The GMAW system contained a GMAW welding machine (Times NB350), a shielding gas supply system, and a 3D motion platform, which carried the welding gun. The workpiece remained stationary during the welding process. The welding direction was set to the positive direction of the x coordinate axis as shown in Fig. 1. The binocular vision equipment consisted of two M3ST130 (M)-H CMOS cameras and ND 1000 neutral filters, which were placed in the front of the two camera lenses. The two cameras were arranged in the up-and-down direction and their clamping device could be precisely adjusted for the dip angle and position of the two cameras. In addition, the camera clamping device was fixed to the 3D motion platform to move synchronously with the welding gun, which meant the welding gun and CMOS cameras were relatively stationary. The two cameras captured the images of the weld pool from different perspectives simultaneously with the help of a synchronous trigger device. The setup parameters of the two cameras are shown in Table 1. The Hall-sensor system was composed of current and voltage sensors that collected electrical signals in real time at a frequency of 3000 Hz.

Experimental Design

Bead-on-plate welding was carried out in the experiments. The dimension of the Q235 mild steel plate was 250 x 50 x 5 mm, and the welding wire was H08Mn2Si solid wire with a diameter of 1.2 mm.

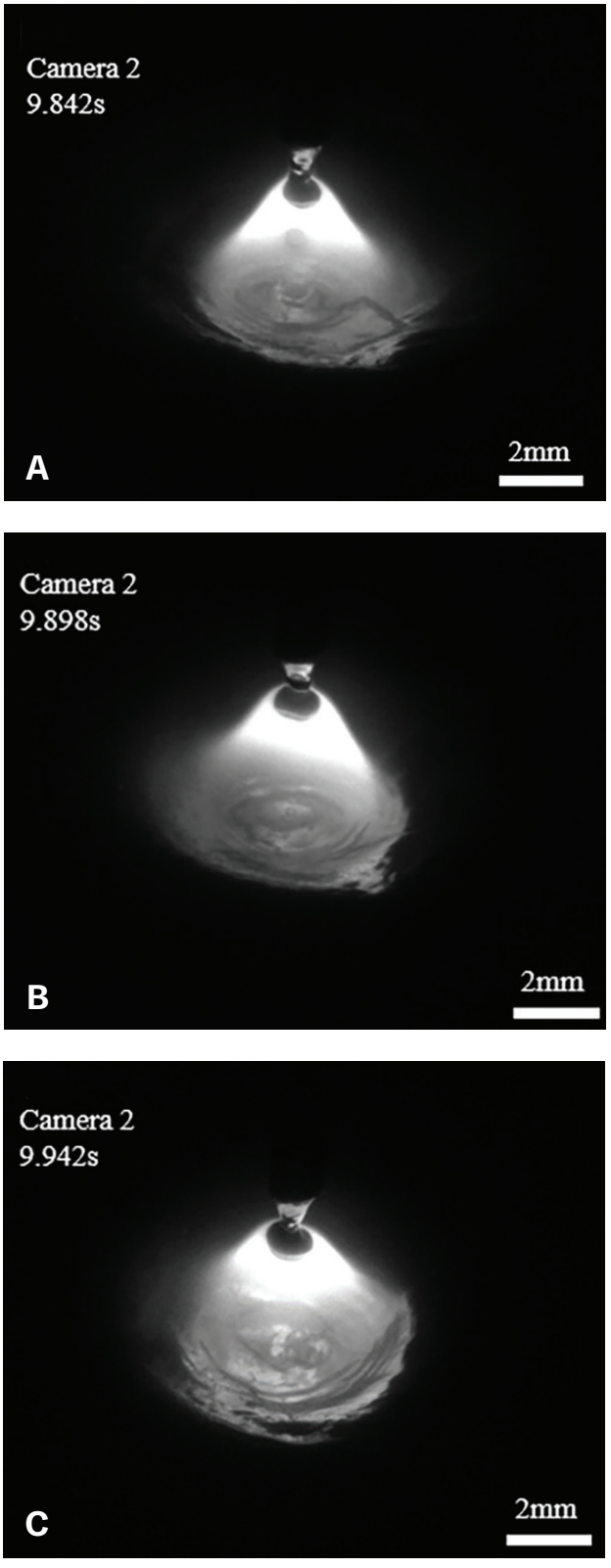


Fig. 3 — Weld pool image acquired within ± 0.05 s at the location of section III in Fig. 2. A — 9.842 s; B — 9.898 s; C — 9.942 s.

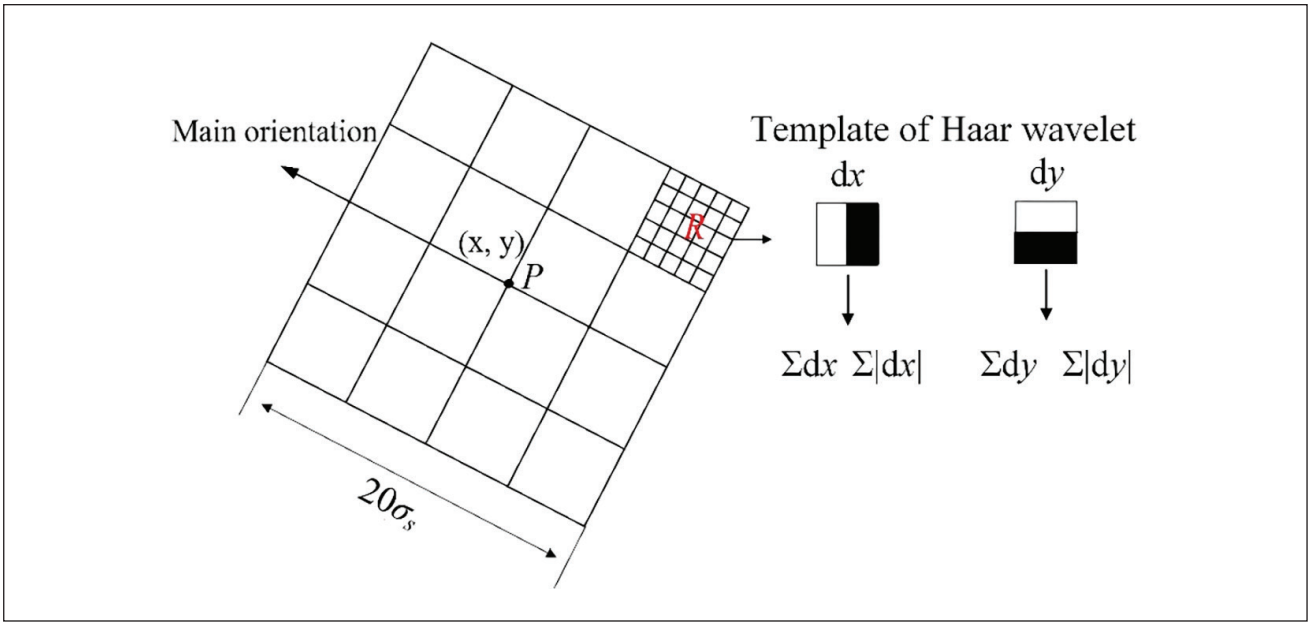


Fig. 4 — Construction of the SURF-BRISK-KAZE feature descriptor.

Table 3 — Feature Parameters of Weld Cross Section at Different Positions of Fig. 2 (Unit: mm)

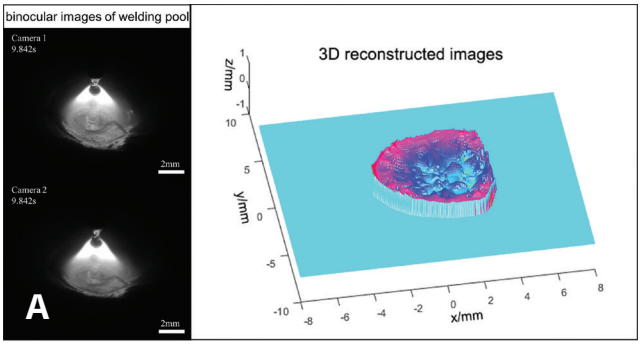
Position	I	II	III'	III	III''	IV	V
Joint penetration	3.3868	3.6126	3.3977	3.599	3.6099	3.5121	3.6371
Weld width	9.6246	9.7416	10.5386	10.5005	10.6719	9.7878	10.1904
Reinforcement	2.1601	2.3232	2.6768	2.6192	2.5517	2.4103	2.4456

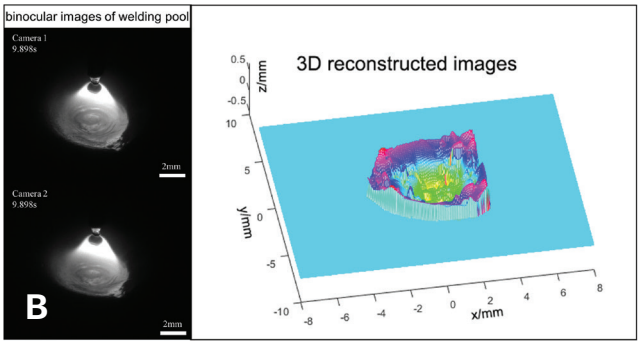
The experiments were designed by the orthogonal experimental design method and the factors consisted of welding current (200–280 A), arc voltage (25–33 V), nozzle height (10–18 mm), and welding speed (0.40–0.80 m/min). Each factor was divided equally into nine levels, and an L81 (94) orthogonal table was selected, as shown in Table 2. All the experiments were repeated two times.

Data Acquisition

The data acquired in the experiments included weld pool images, electrical parameters, weld width, joint penetration, and reinforcement. It was clear that the weld pool always oscillated during the whole GMAW process because of the variation of welding parameters and droplet transfer. Meanwhile, the heat dissipation and heat accumulation conditions were different when the welding arc continued moving on the different positions of the workpiece. As a result, there were small differences in the weld formation of different locations. For this paper, each weld joint was equally divided into six parts along its length (21 cm). Five different weld cross sections were obtained from each weld joint, as shown in Fig. 2, and their feature parameters are shown in Table 3.

The values in Table 3 proved that the weld width, joint penetration, and reinforcement of the weld fluctuated at different positions of the weld joint, especially the joint penetration. Though the weld pool always oscillated during welding processes, Fig. 2 shows the weld cross sections at position III' and position III'' were almost the same as the weld cross section at position III. It meant that each weld cross section corresponded to a series of fluctuating weld pool images that were acquired at the adjacent time, as shown in Fig. 3. In this study, the adjacent time was set to be ± 0.05 s of shooting time for each weld cross section. As a result, 20 pairs of continuous multiframe weld pool images were used to form 20 different 3D reconstructions of the weld pool





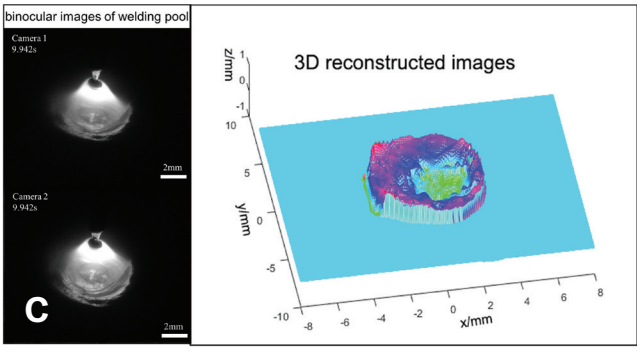


Fig. 5 — Three-dimensional reconstructed images at parameter 27. A — t = 9.842 s; B — t = 9.898 s; C — t = 9.942 s.

surface in the three-dimensional reconstruction section. All the above 3D reconstructions of the weld pool surface were corresponded to one weld cross section (many-to-one relationships).

Average welding electric parameters were calculated and utilized for the prediction of joint formation, which were expressed as:

$$I_{avg} = \frac{1}{n}\sum_{1}^{n} I_i \quad (1)$$

$$U_{avg} = \frac{1}{n}\sum_{1}^{n} U_i \quad (2)$$

where I_i and U_i are the instantaneous welding current and arc voltage detected by the Hall current and voltage sensors, respectively. n is the data quantity (300 in this study), which was counted from −0.05 to +0.05 s relative to the shooting time of each weld cross section.

The normalized value of welding speed, nozzle height, average welding current, and average arc voltage was calculated to solve conflicting measurement units in the prediction of joint formation, which was expressed as:

$$y' = \frac{y - y_{min}}{y_{max} - y_{min}} \quad (3)$$

Table 4 — Model Parameters

Parameter	Value
Kernel size	3 × 3
Initialization method	Kaiming initialization
Normalization	Batch normalization
Dropout	0.5
Optimizer	Adam
Loss function	Mean squared error
Activation function	Rectified linear unit

where y represents the real value of welding speed, nozzle height, average welding current or average arc voltage, and y_{min} and y_{max} are the minimum and maximum data, respectively, of the above parameters in the whole experiments.

Image Processing

Three-Dimensional Reconstruction

To extract key characteristics of weld pool geometry, the SURF-BRISK-KAZE algorithm was utilized in the 3D weld pool reconstruction system (Refs. 22–24), which supplied enough feature points of the weld pool image. This algorithm mainly extracted feature points from the ripples of the weld pool surface that were generated by droplet impingement. To match the feature points of the binocular images of the

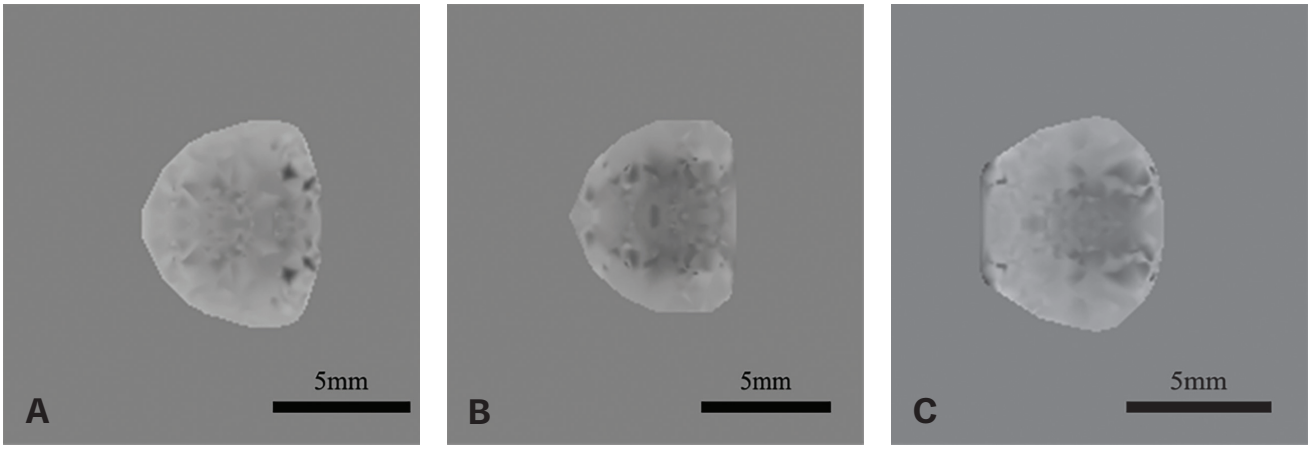


Fig. 6 — Nondimensionalization of 3D reconstructed images. A — 9.842 s; B — 9.898 s; C — 9.942 s.

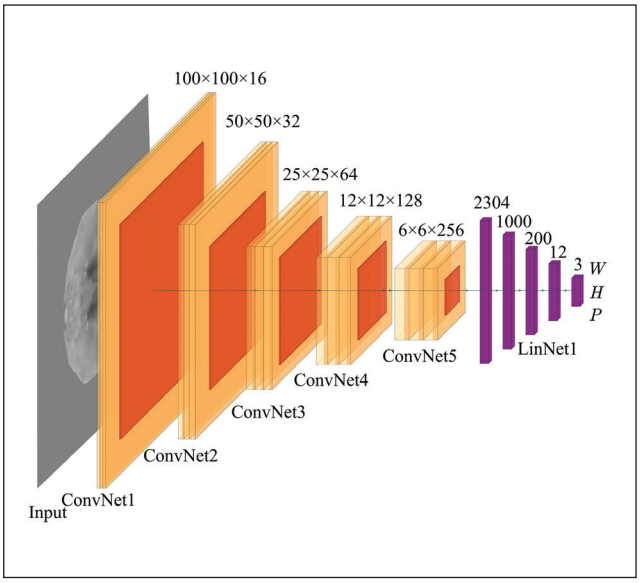


Fig. 7 — The architecture of the traditional CNN model.

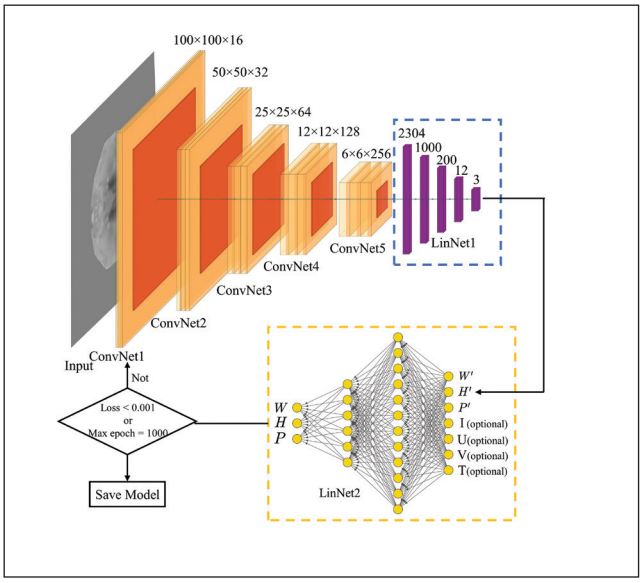


Fig. 8 — Developed CNN architecture for prediction of weld dimensions.

weld pool, the Euclidean distance between their descriptor was calculated as:

$$d_{m,n} = \sqrt{\sum_{i=1}^{64} (V_{1,m,i} - V_{2,n,i})^2} \quad (4)$$

where $V_{1,m}$ is the descriptor of feature point $P_{1,m}$ in the image of camera 1, and $V_{2,n}$ is the descriptor of feature point $P_{2,n}$ in the image of camera 2. *m* is the serial number of feature points in the image of camera 1, and *n* is the serial number of feature points in the image of camera 2. *i* is the serial number of dimensional vectors in the descriptor of the feature point. It was clear that there were *k* values of $d_{m,n}$ for feature point $P_{1,m}$ if the number of feature points in the image of camera 2 was *k*. The matching point in the binocular image of camera 2 corresponding to $P_{1,m}$ was detected if

$$\varepsilon_m = \frac{d_{m,min}}{d_{m,semin}} < 0.5 \quad (5)$$

where $d_{m,min}$ means the minimum value of $d_{m,n}$ for feature point $P_{1,m}$, and $d_{m,semin}$ means the second smallest value of $d_{m,n}$ for feature point $P_{1,m}$.

Table 5 — Global Training Parameters

Training Parameters	Parameter Value	Instructions
Image size	200 × 200	The size of the input image
Batch size	128	Number of samples captured in one training
Learning rate	0.001	Parameter update step

Table 6 — Predicting Accuracy of Traditional CNN Model

	Weld Width	Reinforcement	Joint Penetration
Accuracy (%)	87.63	85.84	75.38

The descriptor *V* was expressed as:

$$V = (V_1, V_2, \ldots, V_i), \quad i = 1 - 64 \tag{6}$$

($V_1 \ldots V_{64}$) was constructed by the SURF-BRISK-KAZE feature descriptor, as shown in Fig. 4. The feature point P(x, y) was considered as the center to generate the square region with a dimension of $20\sigma_s \times 20\sigma_s$. The main orientation of feature point P(x, y) pointed to a dominant vector of the summation of the Gaussian weighted Haar wavelet (Ref. 25). Then the square region was split into 4 × 4 subregions, and each subregion was divided into 5 × 5 small regions. Supposing the *y*-axis was along the main orientation of feature point P(x, y), there were four description vectors ($\Sigma\, dx$, $\Sigma\, |dx|$, $\Sigma\, dy$, $\Sigma\, |dy|$), which were calculated by the template of Haar wavelet and 25 small regions in the subregion of R. As a result, there were 64 description vectors in the 16 subregions of feature point P(x, y).

To improve the feature points matching the accuracy of the binocular images, an improved RANSAC algorithm was adopted to decrease the number of mismatching pairs (Ref. 26). The conversion relationship between the pixel coordinates of the matching point pairs and world coordinates was calculated by the preset coordinate system method to obtain the 3D point cloud of the weld pool. Then the LOWESS fitting algorithm was used to realize the 3D reconstruction of the weld pool surface, as shown in Fig. 5.

Nondimensionalization of the 3D Reconstructed Image

The results in Fig. 5 indicate that the point number was different in the reconstruction of different weld pool surfaces, which meant that the world coordinates of all matching points were hard to be at the input of the artificial intelligence network directly.

To facilitate convolution calculation of the artificial intelligence network, the length (*x*) and width (*y*) of the reconstructed image were converted to 200 × 200 pixels and the deformation (*z*) of the reconstructed image was converted to a gray value, as shown in equation 7.

$$\begin{cases} P_x = (n-1)\cdot\left(\dfrac{x - x_0}{a} + \dfrac{1}{2}\right) \\ P_y = (n-1)\cdot\left(\dfrac{y - y_0}{a} + \dfrac{1}{2}\right) \\ G(P_x, P_y) = 128\cdot\left(\dfrac{z}{b} + 1\right) \\ x = linespace\left(x_0 - \dfrac{a}{2}, x_0 + \dfrac{a}{2}, n\right) \\ y = linespace\left(y_0 - \dfrac{a}{2}, y_0 + \dfrac{a}{2}, n\right) \\ z = f(x, y) \end{cases} \tag{7}$$

where (x_0, y_0) are the center coordinates of the weld pool surface on *xoy* plane (unit mm), (P_x, P_y) are the pixel coordinates of weld pool surface on *xoy* plane, $G(P_x, P_y)$ is the nondimensionalization of weld pool deformation(z), and

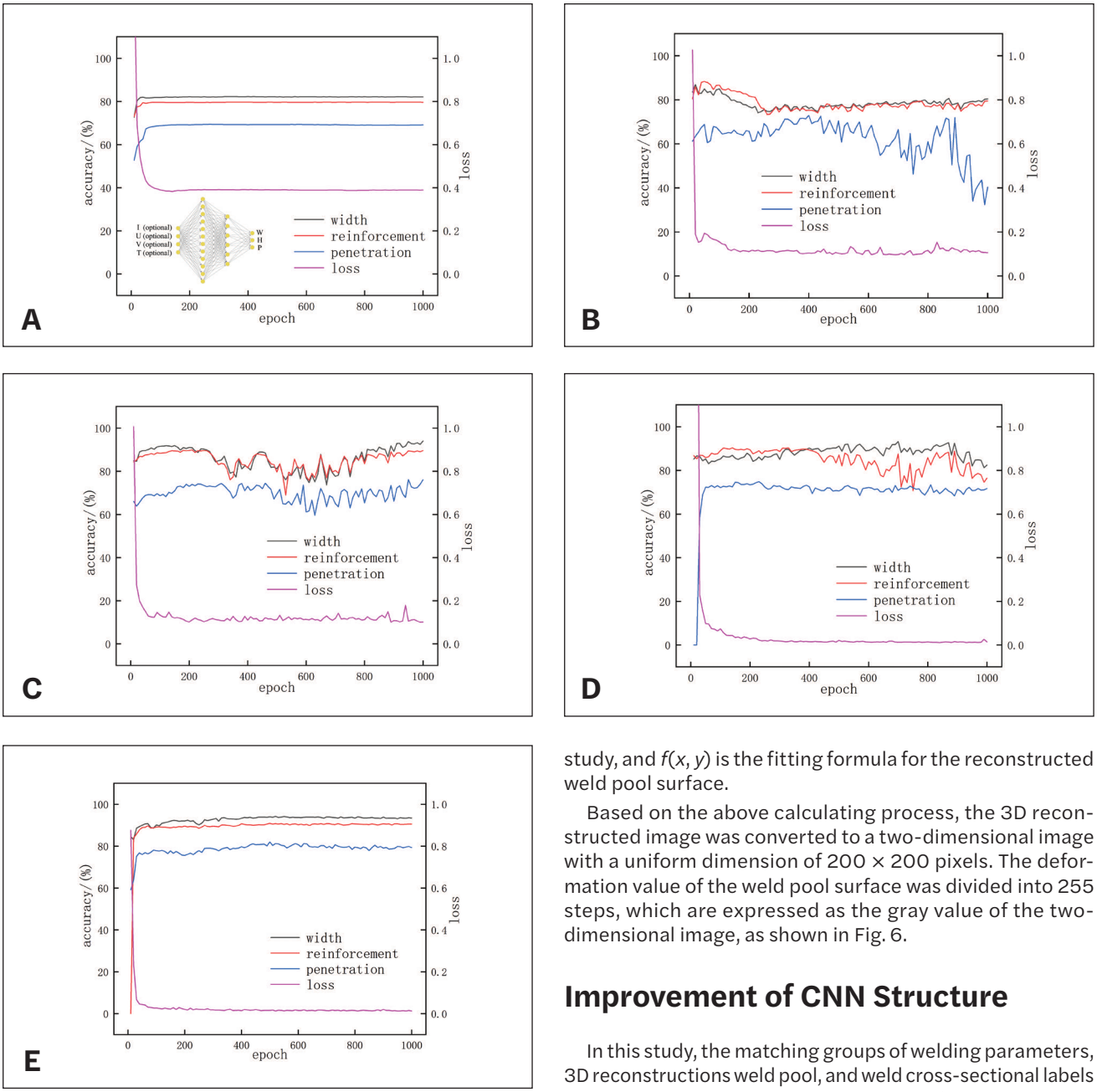


Fig. 9 — Developed CNN model trained by different combinations of welding parameters. A — BPNN model with I, U, V, *and* T*; B — developed CNN model with* I *and* U*; C — developed CNN model with* I, U, *and* V*; D — developed CNN model with* I, U, *and* T*; E — developed CNN model with* I, U, V, *and* T.

a and *b* are the constants obtained by the weld pool length and width, respectively. In this study, the statistical results of the reconstructed weld pool showed its width varied from 5 to 15 mm and its length varied from 5 to 14 mm in all the orthogonal tests, which indicated 16 mm was adequate for the value of *a*. Similarly, 3 mm was selected for constant *b* as all the concave depth of weld pool surfaces varied from −3 to 3 mm (relative to the *xoy* plane). Linespace (*x*, *y*, *n*) means that n values are generated from *x* to *y* evenly, *n* is 200 in this study, and *f*(*x*, *y*) is the fitting formula for the reconstructed weld pool surface.

Based on the above calculating process, the 3D reconstructed image was converted to a two-dimensional image with a uniform dimension of 200 × 200 pixels. The deformation value of the weld pool surface was divided into 255 steps, which are expressed as the gray value of the two-dimensional image, as shown in Fig. 6.

Improvement of CNN Structure

In this study, the matching groups of welding parameters, 3D reconstructions weld pool, and weld cross-sectional labels formed 9243 effective samples (data set). The number of weld cross-sectional labels was 463. At the same time, the samples of the dataset were randomly disrupted to facilitate the model learning's more-accurate features to prevent overfitting. Then the dataset was divided into a training dataset and a test dataset in the ratio of 8 (7394 groups):2 (1849 groups), which were used to train and test the predicting model, respectively.

The network architecture of the traditional CNN model, which consisted of five convolutional blocks (ConvNets 1–5) and a fully connected block (LinNet1), is shown in Fig. 7. Each convolutional block included three convolutional layers and a max pooling layer. The main function of the convolutional block was to extract the features of the simulated depth images. The fully connected block was used to predict the weld width *W*, reinforcement *H*, and joint penetration *P* based on the features of the simulated depth images.

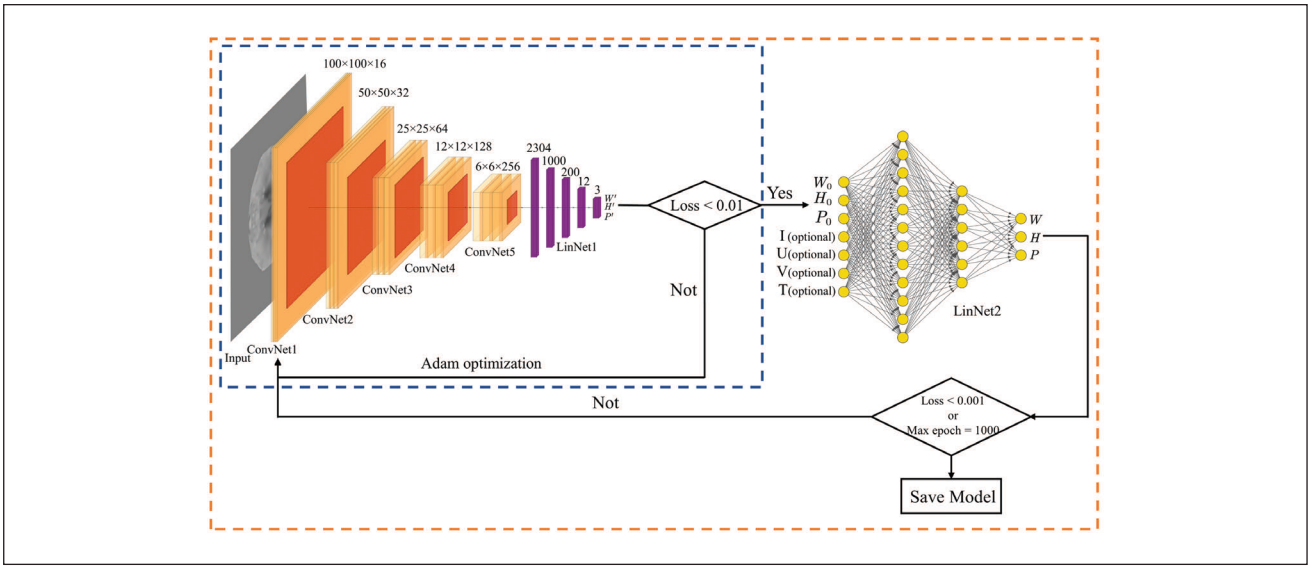


Fig. 10 — Further developed CNN architecture for prediction of weld dimensions.

Table 7 — Test Set Average Accuracy

Weld Shape	Weld Width	Reinforcement	Joint Penetration
Accuracy (%)	97.93	96.99	93.90

To improve the convergence speed and predicting accuracy of the model, the appropriate hyperparameters and corresponding optimization processes selected are shown in Table 4. The training parameters are shown in Table 5.

The predicting accuracy of the traditional CNN model was calculated by equation 8 and the results are shown in Table 6.

$$P = \left(1 - \frac{|y - y'|}{y}\right) \times 100\% \quad (8)$$

where P is the predicting accuracy, y is the value of experiments, and y' is the predicting value of the CNN model.

The low accuracy of predicting results shown in Table 6 indicated that the traditional CNN model needed to be improved as it only contained the information of 3D reconstructed images. The welding parameters, such as welding current, arc voltage, welding speed, nozzle height, and gas flow, should be considered and introduced into the CNN model. The developed network structure of the CNN model was proposed and is shown in Fig. 8. A new fully connected block (LinNet2) was added to the CNN model for the importing of welding parameters. W', H', and P' are the weld width, reinforcement, and joint penetration, respectively, of the weld pre-estimated by LinNet1. I, U, V, and T are the optional normalized value of the local average welding current (± 0.05 s as mentioned in the data acquisition section), local average arc voltage, welding speed, and nozzle height. The function of LinNet 2 was to predict W, H, and P based on the welding parameters and the results of LinNet1. The weight parameters of ConvNets(1-5), LinNet 1, and LinNet 2 were optimized repeatedly to increase the predicting accuracy of the developed CNN model until the value of loss function was lower than 0.001 or the epoch was higher than 1000. The model parameters and training parameters are shown in Tables 4 and 5, respectively. The influence of gas flow on weld bead formation was not considered in this study as it was a constant throughout the whole experiment (argon, 20 L/min).

Figure 9A shows the training results of the BPNN whose input was the electrical parameters (I, U, V, T), and the network architecture was the same as LinNet2. Though the training accuracy curve and loss curve were very smooth, the training accuracy predicted by (I, U, V, T) was very low as the training data were only 7394 groups in this study. Figures 9B–E show the loss curves and predicting accuracy curves of the developed CNN model trained by different combinations of welding parameters. Figure 9B shows the results of the developed CNN model with I and U. It shows that the training accuracies of weld width and reinforcement were less than 80%, and the training accuracy decreased heavily when the epoch was higher than 800. The stability and accuracy of the

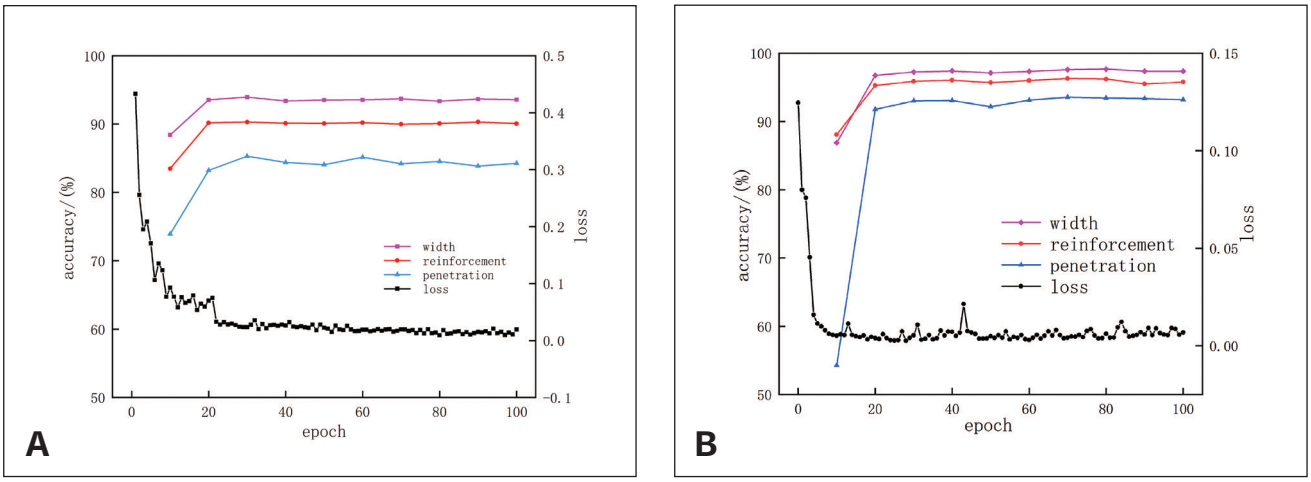


Fig. 11 — Network training process curve with and without binocular vision. A — With 2D image of weld pool surface; B — with binocular vision.

Table 8 — Model Prediction Performance Evaluation

Evaluation Indicator	Weld Width	Reinforcement	Joint Penetration
MAPE	2.09%	3.15%	5.55%
R^2 score	0.85	0.81	0.87

training results were enhanced when the welding speed and nozzle height, respectively, were introduced into the developed CNN model, as shown in Figs. 9C and D. The relative optimal results when all the welding parameters (*I*, *U*, *V*, *T*) are considered in the block of LinNet 2 are shown in Fig. 9E. However, the training accuracies of weld width, reinforcement, and joint penetration were still lower than 82.28%.

The developed CNN architecture in Fig. 8 shows that the loss function was not introduced into LinNet 1, which meant the pre-estimated results of *W'*, *H'*, and *P'* were only trained one time by ConvNets (1–5) and LinNet1 before they were introduced into LinNet2. As a result, the value of *W'*, *H'*, and *P'* with large deviation was utilized in LinNet 2 to predict *W*, *H*, and *P*, which reduced the predicting accuracy of the weld dimensions, though (*I*, *U*, *V*, *T*) were collected in high accuracy. Therefore, a further developed CNN architecture for the prediction of weld dimensions was proposed, as shown in Fig. 10. A loss judgment was inserted between LinNet 1 and LinNet 2 to increase the accuracy of *W'*, *H'*, and *P'* before they were used in LinNet 2. 'Loss 2 > = 0.01' meant the weight parameters of ConvNets (1–5) and LinNet 1 still needed to be optimized to increase the accuracy of *W'*, *H'*, and *P'*. W_0, H_0, and P_0 were the accurate values of *W'*, *H'*, and *P'*, respectively, when the loss was less than 0.01. Then this new model turned on the second fully connected block (LinNet2) to predict *W*, *H*, and *P* by W_0, H_0, P_0, and the welding parameters. The final results were output when loss 1 was less than 0.001 or the epoch was larger than 1000.

Results and Analysis

The further developed CNN model was trained and tested by Intel® Core™ i7-8700 and NVIDIA GeForce RTX 3080Ti. All the codes were compiled based on Python software (version 3.8). And the model was mainly built by PyTorch.

Training Results

Figure 11 shows the training accuracy curve and loss curve calculated by the further developed CNN model. It shows that the loss value, which had decreased to 0.01 when the epoch was only 10 times, declined rapidly. The loss value oscillated around 0.003 when the number of the epoch was higher than 20. However, the minimum value of the loss was 0.00265 (epoch = 29), which was larger than the desired value. The reasons were as follows: (1) There were only 7394 training samples used to optimize the weight parameters of ConvNets (1–5), LinNet 1, and LinNet 2; and (2) some welding parameters generated invalid data for training, as shown in Fig. 11, which meant the training samples distributed unevenly. As a result, more training data are needed to decrease loss value.

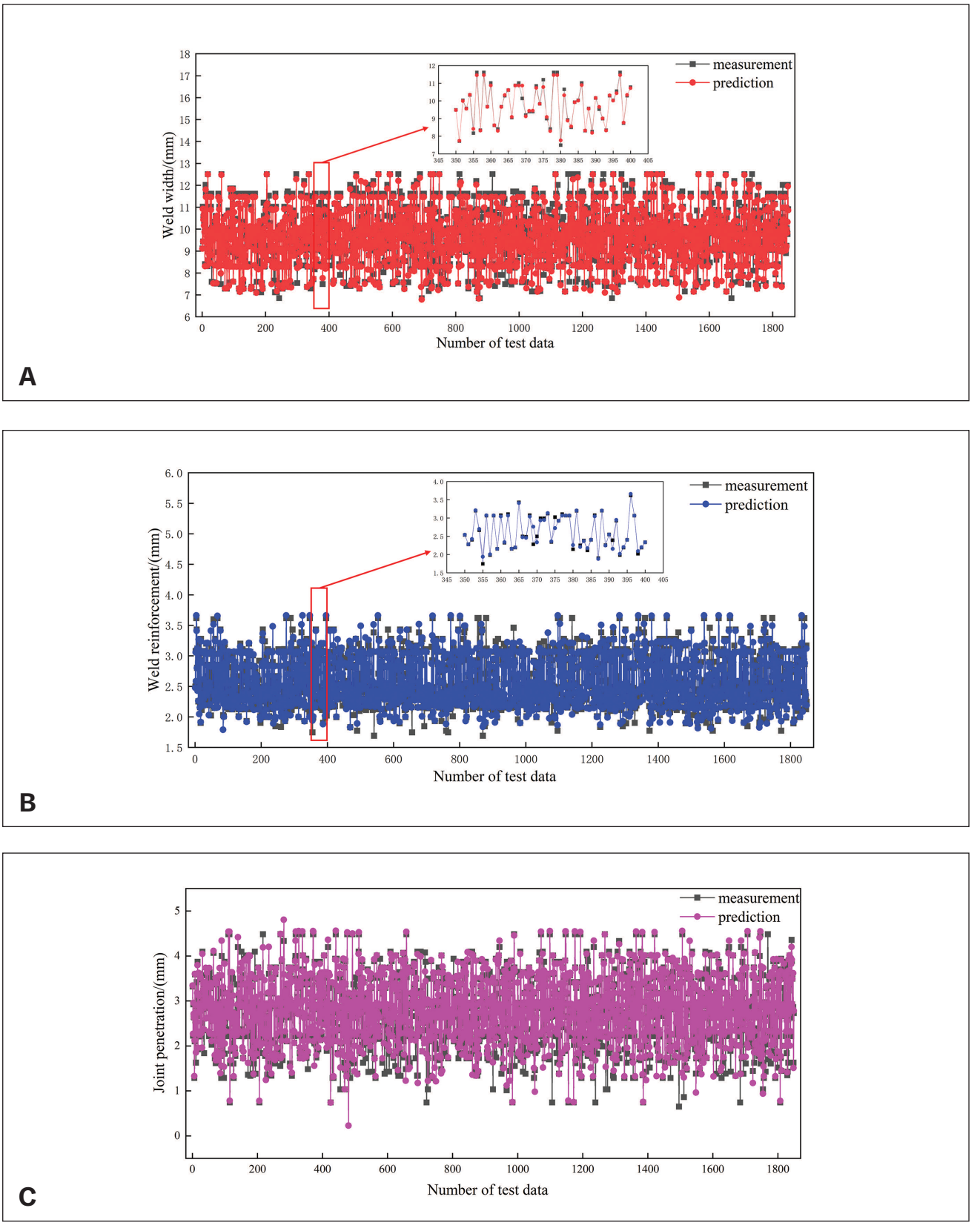


Fig. 12 — The comparison of predicted and measured values for testing data. A — Weld width; B — weld reinforcement; C — joint penetration.

The highest predicting accuracies of weld width, reinforcement, and joint penetration were 97.67, 96.31, and 93.55%, respectively, in the further developed CNN model. To verify the necessary 3D reconstruction of weld pool surface in the prediction of weld dimensions, a single 2D image of weld pool surface with welding parameters was also used to predict the weld dimensions under the same numbers of training samples and testing samples. The results showed that the highest predicting accuracies of weld width, reinforcement, and joint penetration were 95.3, 91.8, and 87.05%, respectively, and the minimum loss value was 0.0048. It is clear that the binocular vision method greatly improved the prediction capability of the further developed CNN model.

Model Performance Testing

Figure 12 shows the comparison of predicted and measured values for the weld width, reinforcement, and joint penetration of welds based on 1849 groups of testing data. Figures 12A and B show that most of the predicted and measured data dots for weld width and reinforcement coincided, which indicated good prediction accuracy. Figure 12C shows that a low coincidence rate appeared in the low value of joint penetration.

Table 7 shows that the average predicting accuracies of weld width, reinforcement, and joint penetration were 97.93, 96.99, and 93.90%, respectively. The prediction accuracy was calculated by Equation 8. The average processing time of each testing group was 17 ms. Further calculation showed that the average prediction accuracy of welding depth reached 97.78% when joint penetration was greater than 1.5 mm, and it decreased to 65.54% when joint penetration was lower than 1.5 mm. The reason was that 95% of data groups generated high joint penetration(> 1.5 mm) and only 5% of data groups generated low joint penetration(< 1.5 mm). The actual error between the predicted and measured value of joint penetration was less than 0.5 mm when joint penetration was lower than 1.5 mm. It means the training data for the low heat input should be increased in the future.

The network in this experiment was designed to solve the regression problem. The mean absolute percentage error (MAPE) and the coefficient of determination (R^2_score) of the predicted value and the real value were used to evaluate the network model. The expressions of MAPE and R^2_score were

$$MAPE = \sum_{i=1}^{n} \left| \frac{y_i - \tilde{y}_i}{y_i} \right| \times \frac{100}{n} \tag{9}$$

$$\bar{y}_i = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{10}$$

$$R^2 = 1 - \frac{\sum_i (\tilde{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \tag{11}$$

where y_i is the true value of i th data, $\tilde{y}_i$ is the predicted value of i th data, $\bar{y}_i$ is the true average value of test data, and n is the number of data in the test set.

Equations 9–11 indicate that a small MAPE value meant a low error of the model prediction value. The larger the R^2 value, the better the fitting effect of the model. The MAPE values and determination coefficient values of weld width, reinforcement, and joint penetration, respectively, of this model are shown in Table 8. The values of MAPE were 2.09, 3.15, and 5.55% for weld width, reinforcement, and joint penetration, respectively. The values of the R^2_score were 0.85, 0.81, and 0.87 for weld width, reinforcement, and joint penetration, respectively. It was shown that the network model had a good prediction performance for the weld dimensions.

Conclusion

This paper predicts the section geometry of GMAW a weld of low carbon steel based on binocular stereo vision and a convolution neural network.

1) A further developed CNN structure was designed. A training method was proposed to train the model by using the weld pool image first and then adding welding parameters to improve the accuracy of the model. The results showed that the loss curve under this method had stable fluctuation and rapid convergence, and the loss value could be reduced to 0.00265.

2) A method of transforming a 3D reconstructed weld pool image into a depth image was proposed, which facilitated the extraction of weld pool information by a convolution neural network.

3) The prediction accuracy of the developed model for weld width, reinforcement, and joint penetration reached 97.93, 96.99, and 93.90%, respectively. When joint penetration was greater than 1.5 mm (accounting for 95% of the test set), the average prediction accuracy of joint penetration was 97%. The prediction error of joint penetration was within an acceptable range. The average test time of one image was 17 ms, which met the requirements of welding real-time feedback control and lays a foundation for realizing the real-time feedback of weld section size.

Acknowledgments

The authors are grateful for the financial support for this project from the National Natural Science Foundation of China (Grant No. 52075296) and the key research and development program of Shandong Province (Grant No. 2021ZLGX01).

References

1. Hu, Z., Hua, L., Qin, X., Ni, M., Ji, F., and Wu, M. 2021. Molten pool behaviors and forming appearance of robotic GMAW on complex surface with various welding positions. *Journal of Manufacturing Processes* 64: 1359–1376. DOI: 10.1016/j.jmapro.2021.02.061

2. Lu, J., He, H., Shi, Y., Bai, L., Zhao, Z., and Han, J. 2021. Quantitative prediction for weld reinforcement in arc welding additive manufacturing based on molten pool image and deep residual network. *Additive Manufacturing* 41: 101980. DOI: 10.1016/j.addma.2021.101980

3. Kershaw, J., Yu, R., Zhang, Y. M., and Wang, P. 2021. Hybrid machine learning-enabled adaptive welding speed control. *Journal of Manufacturing Processes* 71: 374–383. DOI: 10.1016/j.jmapro.2021.09.023

4. Sen, M., Mukherjee, M., and Pal, T. K. 2015. Evaluation of correlations between DP-GMAW process parameters and bead geometry. *Welding Journal* 94(8): 265-s to 279-s.

5. Andersen, K., Cook, G. E., Karsai, G., and Ramaswamy, K. 1990. Artificial neural networks applied to arc welding process modeling and control. *IEEE Transactions on Industry Applications* 26(5): 824–830. DOI: 10.1109/IAS.1989.96968

6. Zeng, X. M., Lucas, J., and Fang, M. T. C. 1993. Use of neural networks for parameter prediction and quality inspection in TIG welding. *Transactions of the Institute of Measurement and Control* 15(2): 87–95. DOI: 10.1177/014233129301500204

7. Pernambuco, B. S. G., Steffens, C. R., Pereira, J. R., Werhli, A. V., Azzolin, R. Z., and Estrada, E. D. S. D. 2019. Online sound based arc-welding defect detection using artificial neural networks. In 2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR), and 2019 Workshop on Robotics in Education (WRE). *IEEE*: 263–268. DOI: 10.1109/LARS-SBR-WRE48964.2019.00053

8. Cui, Y., Shi, Y., Zhu, T., and Cui, S. 2020. Welding penetration recognition based on arc sound and electrical signals in K-TIG welding. *Measurement* 163: 107966. DOI: 10.1016/j.measurement.2020.107966

9. Nagesh, D. S., and Datta, G. L. 2002. Prediction of weld bead geometry and penetration in shielded metal-arc welding using artificial neural networks. *Journal of Materials Processing Technology* 123(2): 303–312. DOI: 10.1016/S0924-0136(02)00101-2

10. Zhang, Y., Wang, Q., and Liu, Y. 2021. Adaptive intelligent welding manufacturing. *Welding Journal* 100(1): 63-s to 83-s. DOI: 10.29391/2021.100.006

11. Lei, Z., Shen, J., Wang, Q., and Chen, Y. 2019. Real-time weld geometry prediction based on multi-information using neural network optimized by PCA and GA during thin-plate laser welding. *Journal of Manufacturing Processes* 43: 207–217. DOI: 10.1016/j.jmapro.2019.05.013

12. Chang, B., and Huang, J. 2020. Discrimination of molten pool penetration based on genetic algorithm optimization of BP neural network. *Journal of Physics: Conference Series* 1437(1): 012110. DOI: 10.1088/1742-6596/1437/1/012110

13. Martínez, R. T., Bestard, G. A., Silva, A. M. A., and Alfaro, S. C. A. 2021. Analysis of GMAW process with deep learning and machine learning techniques. *Journal of Manufacturing Processes* 62: 695–703. DOI: 10.1016/j.jmapro.2020.12.052

14. Lu, J., Xie, H., Chen, X., Han, J., Bai, L., and Zhao, Z. 2020. Online welding quality diagnosis based on molten pool behavior prediction. *Optics & Laser Technology* 126: 106126. DOI: 10.1016/j.optlastec.2020.106126

15. Nomura, K., Fukushima, K., Matsumura, T., and Asai, S. 2021. Burn-through prediction and weld depth estimation by deep learning model monitoring the molten pool in gas metal arc welding with gap fluctuation. *Journal of Manufacturing Processes* 61: 590–600. DOI: 10.1016/j.jmapro.2020.10.019

16. Cheng, Y. C., Wang, Q. Y., Jiao, W. H., Xiao, J., Chen, S. J., and Zhang, Y. M. 2021. Automated recognition of weld pool characteristics from active vision sensing. *Welding Journal* 100(5): 183-s to 192-s. DOI: 10.29391/2021.100.015

17. Zhang, W. J., Zhang, X., and Zhang, Y. M. 2015. Robust pattern recognition for measurement of three dimensional weld pool surface in GTAW. *Journal of Intelligent Manufacturing* 26(4): 659–676. DOI: 10.1007/s10845-013-0825-z

18. Song, H. S.,and Zhang, Y. M. 2007. Image processing for measurement of three-dimensional GTA weld pool surface. *Welding Journal* 86(10): 323-s to 330-s.

19. Liu, Y. K., Zhang, W. J., and Zhang, Y. M. 2015. Nonlinear modeling for 3D weld pool characteristic parameters in GTAW. *Welding Journal* 94(7): 231-s to 240-s.

20. Zhang, B., Shi, Y., Cui, Y., Wang, Z., and Hong, X. 2021. Prediction of keyhole TIG joint penetration based on high-dynamic range imaging. *Journal of Manufacturing Processes* 63: 179–190. DOI: 10.1016/j.jmapro.2020.03.053

21. Liang, Z., Chang, H., Wang, Q., Wang, D., and Zhang, Y. 2019. 3D reconstruction of weld pool surface in pulsed GMAW by passive biprism stereo vision. *IEEE Robotics and Automation Letters* 4(3): 3091–3097. DOI: 10.1109/LRA.2019.2924844

22. Khare, A., Mounika, B. R., and Khare, M. 2020. Keyframe extraction using binary robust invariant scalable keypoint features. *Twelfth International Conference on Machine Vision* (ICMV 2019). SPIE. Vol. 11433, 50–55. DOI: 10.1117/12.2559105

23. Ramkumar, B., Laber, R., Bojinov, H., and Hegde, R. S. 2020. GPU acceleration of the KAZE image feature extraction algorithm. *Journal of Real-Time Image Processing* 17(5): 1169–1182. DOI: 10.1007/s11554-019-00861-2

24. Shih, H. C., Ma, C. H., and Lu, C. L. 2019. An efficient fragment reconstruction using RANSAC algorithm. *IEEE 8th Global Conference on Consumer Electronics (GCCE)* 529, 530. DOI: 10.1109/GCCE46687.2019.9015254

25. Kang, T. K., Choi, I. H., and Lim, M. T. 2015. MDGHM-SURF: A robust local image descriptor based on modified discrete Gaussian-Hermite moment. *Pattern Recognition* 48(3): 670–684. DOI: 10.1016/j.patcog.2014.06.022

26. Gu, Z. N., Chen, J., and Wu, C. S. 2021. Three-dimensional reconstruction of weld pool surface by binocular vision. *Chinese Journal of Mechanical Engineering* 34(1): 1–13. DOI: 10.1186/s10033-021-00567-2

SONGYU WANG, **JI CHEN** (*chenji@sdu.edu.cn*), **CHUNYANG XIA**, **RUIDONG LI**, and **CHUANSONG WU** are with the Key Laboratory for Liquid-Solid Structural Evolution & Processing of Materials, Ministry of Education, Institute of Materials Joining, Shandong University, China. **WENBIN ZHANG** is with the China Construction Eighth Engineering Division.