Journal of Manufacturing Processes 45 (2019) 208–216

Contents lists available at ScienceDirect

Journal of Manufacturing Processes

journal homepage: www.elsevier.com/locate/manpro

Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding

Zhifen Zhang[a], Guangrui Wen[a,*], Shanben Chen[b]

[a] School of Mechanical Engineering, Xi'an Jiao Tong University, Xi'an, 710049 China
[b] School of Material Sciences and Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

ARTICLE INFO

Keywords:
Deep learning
Defects detection
Al alloy
Robotic arc welding
Convolutional neural networks
Weld images
Feature visualization

ABSTRACT

Accurate on-line weld defects detection is still challenging for robotic welding manufacturing due to the complexity of weld defects. This paper studied deep learning–based on-line defects detection for aluminum alloy in robotic arc welding using Convolutional Neural Networks (CNN) and weld images. Firstly, an image acquisition system was developed to simultaneously collect weld images, which can provide more information of the real-time weld images from different angles including top front, top back and back seam. Then, a new CNN classification model with 11 layers based on weld image was designed to identify weld penetration defects. In order to improve the robustness and generalization ability of the CNN model, weld images from different welding current and feeding speed were captured for the CNN model. Based on the actual industry challenges such as the instability of welding arc, the complexity of the welding environment and the random changing of plate gap condition, two kinds of data augmentation including noise adding and image rotation were used to boost the CNN dataset while parameters optimization was carried out. Finally, non-zero pixel method was proposed to quantitatively evaluate and visualize the deep learning features. Furthermore, their physical meaning were clearly explained. Instead of decreasing the interference from arc light as in traditional way, the CNN model has taken full use of those arc lights by combining them in a various way to form the complementary features. Test results shows that the CNN model has better performance than our previous work with the mean classification accuracy of 99.38%. This paper can provide some guidance for on-line detection of manufacturing quality in metal additive manufacturing (AM) and laser welding.

1. Introduction

Aluminum alloy is one of the key material for light weighing manufacturing in various industry applications, such as aerospace, automotive and so on. Arc welding manufacturing of aluminum alloy components, for instance gas tungsten arc welding (GTAW) has always been the main technology for aerospace industry due to its high precision and good stability. Therefore, guaranteeing the welding quality of aluminum alloy component with zero-defect is of great importance. At present, more and more robotics are gradually replacing human welder, which can greatly improve the production efficiency. However, the down side is that there is no more human intelligence interference to control or stop the weld defect during the dynamic welding process. Furthermore, due to the complexity of welding process and randomness of process interference, weld defects are usually inevitable. Various technologies of Non Destructive Testing (NDT) are applied post welding and performed by experienced workers, which are costly and low efficient. Therefore, achieving real-time monitoring and detecting of weld defects during the robotic welding process is quite urgent in need.

Weld surface defects including under penetration and burning through can greatly weaken the strength of joints. Even if precautionary measures are well made before welding, the defects still cannot be effectively controlled because of its high randomness and complex influence factors, such as assembly form, welding structure, welding process parameters and heat accumulation effect. This kind of defects are not allowed in the welding products based on ISO10042 inspection standards. So, using real-time monitoring and detection technology of welding seam quality might timely detecting seam defects, improving the stability of welding quality and manufacturing efficiency; moreover, promoting the intelligent degree of robotic welding manufacturing.

Sensing technology is the key to achieve on-line defects detection. By accurately sensing certain type of information, the welding robot might be "smart" enough to identify the seam quality of aluminum alloy in real-time [1]. During the dynamic welding process, multiple types of information are releasing including welding audible sound, arc light, welding current and voltage. Arc audible sound information is reported

* Corresponding author.
E-mail addresses: zzf919@xjtu.edu.cn (Z. Zhang), grwen@xjtu.edu.cn (G. Wen), sbchen@sjtu.edu.cn (S. Chen).

https://doi.org/10.1016/j.jmapro.2019.06.023
Received 25 March 2019; Received in revised form 11 June 2019; Accepted 24 June 2019
Available online 11 July 2019

to be related to seam penetration [2,3]. Zhang [3] found that the polarity exchange of AC welding power is the key source of arc sound pressure and signal frequency energy. Seam defects of under penetration and burning through were identified using the proposed classification model. However, environment noise in welding workshop can be loud and uncertain, which normally requires complex algorithm and is difficult for industry application. Arc spectrum emission is also reported to be able to detect certain macro weld defects by means of non-contact spectrometer since it contains abundant information about emission intensity of various chemical elements. Mirapeixa [4] quantitatively researched the seam quality of Usibor blanks in the laser welding process while the ratio of line-to-continuum for the Al I emission line at 396.15 nm was extracted as the monitoring feature. Song et al. [5] predicted the Al concentration during a laser additive manufacturing process by means of Al/Ti line-intensity ratio and support vector regression. Zhang [6] proposed several spectrum bands of interest and extracted their statistical features to detect surface oxidation defects. Furthermore, Huang et al. [7] utilized a k-medoids algorithm to select spectral lines with more sensitivity to seam porosity. However, arc optical spectrum is still researched on laboratory-level. More important, it can be quite sensitive to the setup position and material component. Its reliability and generality still need to improve. Besides, its results are highly dependent on equipment resolution of the spectrometer, which leads to high cost and is unacceptable for industry application.

Accurate detection of multiple seam defects, either surface one or inner one is still difficult due to the complex generation mechanism of defects and their own limitations. In addition, those defects can happen in short time and be probably unseen inside the weld pool and seam. Multisensory fusion technology [8–11] might provide a possible solution for this problem, but it will bring much more burden to the monitoring system with all kinds of sensors, different information processing methods and complex algorithms. With all the high cost, it is not easy for industry application at present. Welding pool vision and X ray imaging are more direct in terms of evaluating the welding process and seam quality. Liu et al. [12] acquired the welding current and weld pool vision signal during the GTAW process to control the automated process using a data-driven method. Furthermore, the fusion [13] between the machine algorithm and human intelligence was investigated for the next generation of intelligent welding robots. Liu and Zhang [14] studied to extract the human welder's response against 3-D weld pool surface and transfer to the welding robots to perform automated welding tasks. He et al. [15] collected images of a segment gap and welding pool using a vision sensor that integrated the laser structure light with a CCD image. Calta et al. [16] developed a synchrotron X-ray imaging system to collect the dynamic information of a melt pool in a laser powder bed fusion process. The formation of keyhole pores was revealed for Ti-6Al-4 V. X-ray imaging can collect more inner information of the weldment, but higher resolution comes with higher radiation and damage to workers. Its acquiring equipment can be very complex and still on laboratory research stage.

Weld pool vision can acquire more macro level and rich information, including wire filling, shape of weld pool and arc shape with high resolution and low cost since CCD technology is more mature and has been widely used in industry. To overcome these problems, traditional methods includes two steps including feature extraction based on image processing and classification modelling, as shown in Fig. 1. Its real-time performance, stability and robustness of image system is still a huge challenge not to mention the complexity, efficiency and huge cost of algorithm.

The convolutional neural network, as one of the well-known deep learning architectures, has powerful ability of feature learning owning to multiple convolution and pooling layers. The greatest advantage of CNN is that the features of each hidden layer are not designed manually, but are learned from the input data automatically. With decades developing, many various CNN models have been proposed and applied in diverse applications, such as cracks detection in civil infrastructures [17], myocardial infarction detection using ECG signals [18], fault classification of machinery [19], detect cerebral micro bleeds from magnetic resonance images [20] and so on. CNN has been applied to weld defects detection both for on line [21,22] and off line [23]. Literature [21] combines CNNs with long short-term memory networks to extract more comprehensive and fusion features while welding burning through and welding deviation were detected for CO2 welding. S.A. Shevchik et.al [22] studied the on-line detection of porosity with different types and concentrations in selective laser melting of stainless steel using acoustic emission and spectral convolutional neural network.

However, to get the clear information of defects from real-time weld images often encounter several challenges for CNN as follows:

1) Difficult to get a clear vision of the welding pool under the powerful exposure of arc light; hard to be characterized using simple features;
2) Dynamic changing of arc light intensity under different welding power; Arc drifting and unstable arc plasma because of constantly searching for the minimum ionization energy path, leading to dynamic weld images with quite different features.
3) Due to the high diversity of welding condition in the application of robotic rapid welding, the model generalization ability needs to be improved using technologies such as data augmentation, which has not been reported yet as well as the visualization of CNN features.

In this paper, on-line detection of weld defects for Al alloy in pulsed GTAW was researched using convolutional neural networks and weld image. Contributions can be included as follows:

1) A weld image-based CNN classification model was proposed to identify weld seam penetration defects in real-time for Al alloy in pulsed GTAW based on the three-way pool images, which can provide more information of the real-time weld images from different angles including top front, top back and back seam.
2) Data augmentation including noise addition and image rotation was utilized to boost the CNN dataset and has improved the classification accuracy about 3.88% based on the actual industry challenges analyzed in Fig.3 such as the instability of welding arc, the complexity of the welding environment and the random changing of plate gap condition. In order to improve the robustness and generalization ability of the CNN model, weld images from different welding current and feeding speed were captured for the CNN model.
3) The deep learning image features of the CNN model were visualized and analyzed using the proposed Non-zero Pixel (NOP) method. Their physical meaning were clearly explained. Instead of decreasing the interference from arc light as in traditional way, the CNN model has taken full use of those arc lights by combining them in a various way to form the complementary features.

2. Theory of CNN

Convolutional neural networks [24,25] are a specialized kind of neural network that use convolution in place of general matrix multiplication in some of their layers. Typical CNN consists convolutional layers, pool layers and full connection layers.

Convolutional layer introduces two important ideas: local receptive fields and shared weights. Within the convolutional layer's feature map, each neuron is connected to previous layer's maps through a learnable filter (convolutional kernel). Neurons within the same feature map share the same filter which has a small receptive field, but extends through the full corresponding previous layer's map. Multiple maps presenting different features are produced by using distinct filters. Mathematically, the convolutional layer is defined as follows:

$$x_j^l = f\left(\sum_{i\in M_j} x_i^{(l-1)} * k_{ij}^l + b_j^l\right) \tag{1}$$

Where * represents the convolution operator; x_j^l represents the jth feature map in the lth layer; M_j represents the input subset of feature

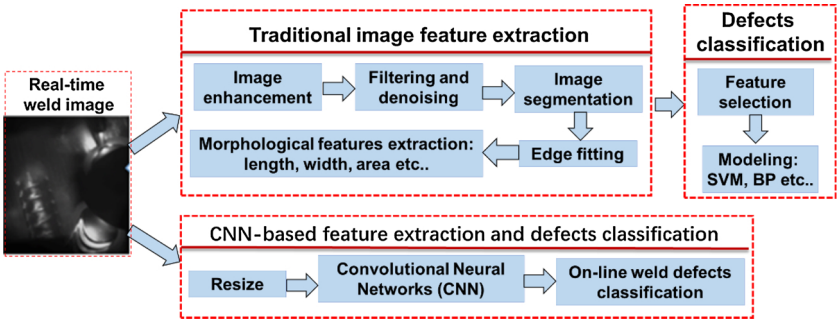


Fig. 1. Comparison of traditional method and proposed method in this paper.
a) Schematic of the image sensing system b) The photograph of visual sensor. c) Real-time three-way weld image.

maps that are used to calculate $x_j^l k_{ij}^l$is the weight matrix of the kernel connecting the ith feature map in the l-1th layer with the jth feature map in the lth layer; b_j^l is bias; $f(\cdot)$represents the activation function, which enables the network to acquire a nonlinear expression of the input signal to enhance the representation ability.

In this paper, Leaky Rectified Linear Units (L-ReLU) activation function is used in convolution layers and full connection layers for its multiple advantages, such as making the weights more trainable, better gradient propagation, scale-invariant and avoiding neurons death. The formula of L-ReLU is described as follows:

$$f(x) = \begin{cases} 0.01x, x \le 0 \\ x, \quad x \ge 0 \end{cases} \tag{2}$$

After convolution, pooling operations are used to replaces the output of a certain location with a summary statistic of nearby.

$$x_j^l = pooling(x_j^{l-1}) + b_j^l \tag{3}$$

Where, $pooling(\cdot)$ denotes the pooling operation, which is a form of non-linear down-sampling. Max pooling is the most common pooling functions. Pooling helps to make the features become approximately invariant to small translations of the input and improve statistical and computational efficiency further. It can be related to a feature selection. Finally, the extracted features are input into the full connection layers, which is similar with multiplayer perception (MLP) network. All layers are trained simultaneously using a version of the back-propagation learning algorithm. During the training progress, a CNN with parameters, $\theta = \{k_{ij}, w_{ij}, b_j\}$, is trained and optimized by minimizing its loss function, where cross entropy is used and can be described:

$$H(p, q) = -\sum_x p(x)\log q(x) \tag{4}$$

Where, H is the cross entropy, which can determine how close the actual output is to the expected output. p is the expected output of sample x, q is the actual output of sample x.

2.1. Experiment setup

The experimental system includes multiple signal acquisition system, welding system and process controlling and motion system. The experimental system diagram and welding parameters can be found in literature [3]. Both arc sound pressure and voltage signal were synchronously acquired with the sampling rate of 40 kHz by a data acquisition card. The welding system includes the TIG welding power with OTC INVERTER ELESON 500P-type, CM-271 type wire feeder, water cold tank, welding torch and 99.99% pure Argon as the shielding gas. For the control and motion system, a host computer using VC++ multithreading application was utilized to control the welding parameters such as feed speed, time and so on. It can also drive work piece to move with multiple freedoms while the torch and the designed multisensory system remained quiescent.

For the weld image acquiring, a three-way image acquisition system was developed to simultaneously collect the vision information from three directions including top center of the welding pool, front of the welding pool as well as back of the welding seam. The schematic diagram of the image sensing system and the photograph of visual sensor were given in Fig.2-a and b. It includes a CCD camera (DH-SV1420FM), mirror system, filter system and "L" type aluminum alloy holder. The USB digital transmission method was adopted to improve the transmission speed and anti-interference ability. Multiple mirror reflection was designed in order to capture the weld image from the three directions. The adjustable secondary mirror is fixed at a 45 ° angle, which facilitates the adjustment of the reflected light path. The filter system consists of proper filtering and dimming in order to decrease the interference from arc light. In addition, considering the pulsed welding current of the pulsed GTAW method, the three-way images were acquired within the period of pulse base current, which might have less arc light. Its specific time was set to 36 ms after the peak current drops behind the edge. Its sampling rate is the same with pulse frequency, *e.g.*, 1 Hz, with the pixel size of 1392*1040. Fig.2-b displays the acquired real-time three-way weld image, from which the front edge of weld pool and wire filler can be seen from the top front image, while the clear view of the weld pool and arc can be obtain from the top back image. Moreover, the back view of the weld seam gives the real-time information about the forming quality.

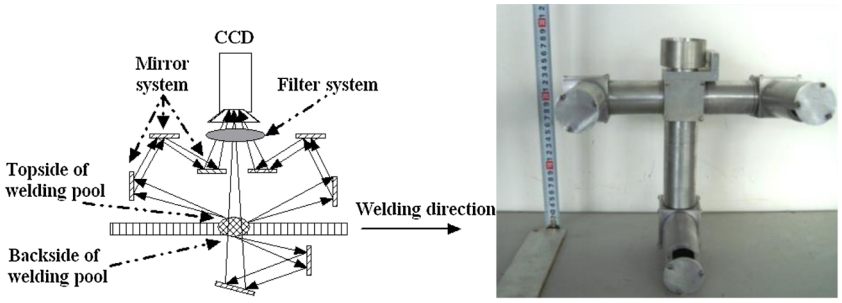
For precise defects detection based on weld images during the dynamic welding process of Al alloy, several challenges are encountered and are demonstrated in Fig.3.

1) Due to the instability of welding arc, arc light is dynamically changing even the acquisition time is settled. Therefore, arc light might be quite strong and useful information about weld quality and defects can be submerged by the strong light interference, as shown in Fig.3-a.
2) Due to the complexity of the welding environment, it is difficult to keep the CCD camera and vision acquisition system in a constant position. Therefore, the captured images might have different characteristics as shown in Fig.3-b, even miss certain information as shown in Fig.3-d.
3) Gap between the weld plate might randomly change due to the assemble error or plate transformation. Thus, it leads to arc light leaking from the gap and disturbs the real-time weld image, as shown in Fig. 3-c.
4) Reflection of arc light is another huge problem when welding Al alloy, which can be found in Fig.3-b–d. The reflection obviously increase the complexity of detection samples.

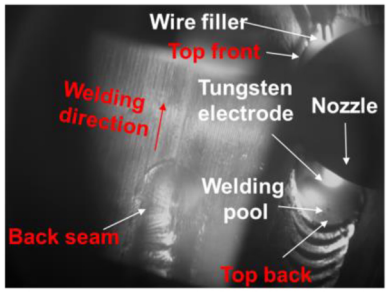
3. Method and results

3.1. Datasets prepare

In this section, datasets of weld images used for training and testing the CNN classification model was carefully introduced. As listed in Table 1, three types of weld seam quality including under penetration



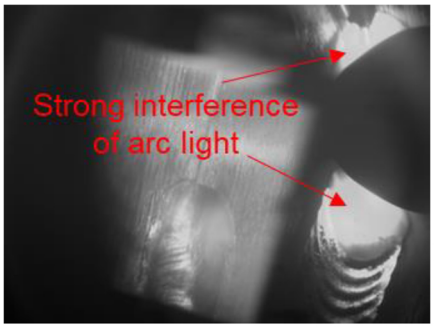
a) Schematic of the image sensing system b) The photograph of visual sensor



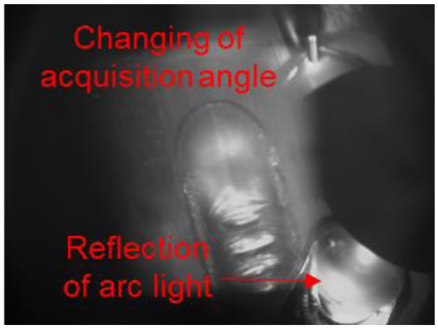
c) Real-time three-way weld image

Fig. 2. Weld image acquiring for Al alloy in GTAW.
a)Full penetration b) Changing of acquisition angle.
c) Arc light leaking from gap d) Under penetration.

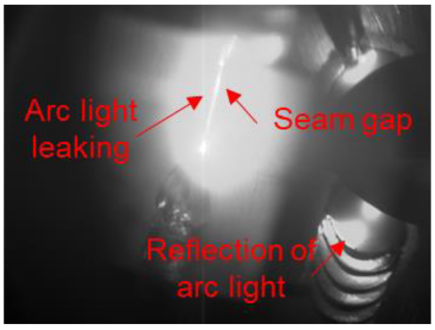
defects, normal penetration and burning through defects were obtained by designing the rectangle-shaped, dumbbell-shaped and ladder-shaped weld plates, which can be seen in Fig. 4. Besides, considering the obvious heat accumulation effect of Al alloy thin plate, there is no preheating of plates, which caused the under penetration at the beginning of the welding, while presetting gap can generate the burning through defects. Moreover, different weld parameters were used such as welding current, feeding speed in order to cover more welding conditions, increase the diversity of the datasets and then improve the generation ability of classification model.



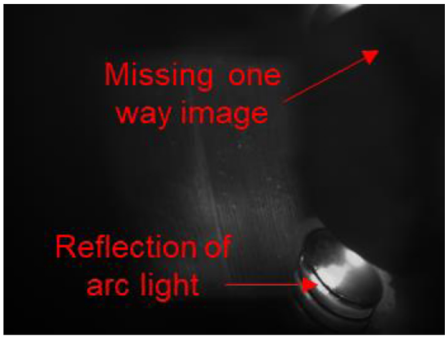
a)Full penetration



b) Changing of acquisition angle



c) Arc light leaking from gap



d) Under penetration

Fig. 3. Different weld images acquired for Al alloy in GTAW.
a) Rectangle-shaped plate b) dumbbell-shaped plate c) ladder-shaped plate.

Table 1
Data preparing of different seam penetration.

Shape of plate	Welding current	Feeding speed	Weld seam penetration
Rectangle shape Dumbbell-shaped Ladder-shape	240A, 260A 195A, 205A, 230A,	10 mm/s,7 mm/s 10 mm/s,7 mm/s	Under penetration Normal penetration Burning through

3.2. CNN based on weld image

In this section, a CNN defects classification model for Al alloy in GTAW was developed based on CNN and weld images. Its framework was carefully designed based on the feature of weld image as displayed in Fig.5.

3.2.1. Convolution and pooling

The original size of the weld image was 1392*1040 pixel, which needs to be resized into 100*100 before inputting into the CNN model, which can be seen in Fig. 6. The original image was resized by performing interpolation to down-size the images. For down-sampling, the arithmetic mean was applied. The data in the three channels have been analyzed and shows little difference. Therefore, the original images have been transferred into grey-scale and used the kernel size with 5*5 other than 5*5*3. As listed in Table 2, the kernel size for the first convolution layer (L1-C1) is 5×5 with 1stides, which means it will convolve by skipping one pixel in x-direction and y-direction. The second convolution layer (L3-C2) is the same with the first one, while the third (L5-C3) and fourth(L7-C5) convolution layer have the kernel size of 3×3 in 3 dimension with 1stides. Each convolution layer is followed by a pooling layer with the size of 2×2 in 3 dimension with 2 strides as in L2-P1, L4-P2, L6-P3 and L8-P4. In this paper, max pooling performs the local max operation over the input features from the last layer, which can reduce the data size and obtain location-invariant features. That means the pooling layer can reduce the size of the weld image, moreover, it can extract the invariance feature to shift, scale and distortion. Thus, it is proper for the complex welding environment and randomly changing weld images. The number of filters for each convolution layer are listed in Table 2, which increased with the layer number in order to extract deeper feature of the input weld image.

3.2.2. Fully connected network

The fully connected network is similar with the neural network, which can defined as follow.

$$x_j^l = f\left(\sum_{i\in W_j} x_i^{l-1} * W_{ij}^l + b_j^l\right) \tag{5}$$

Where, W^{ij}represents the weight connecting the ith neuron in l-1th layer with the jth neuron of in lth layer. $W_j$represents the number of neurons in l-1layer. After the last layer of pooling, the original size of 100*100 is down to 6*6*128. 3-D maps is rasterized to 1-D vectors and connected with each units in the following full connection layers L9 –F1 and L10 –F2 with the neural number of 1024 and 512. The last layer is a logistic-regression layer containing three outputs corresponding to the three class of welding defects.

3.3. Activation and optimization

In this paper, Rectified Linear Unit (ReLU) function is used as the activation function for the CNN model both in convolution layer and full connection layer. It enables the network to acquire a nonlinear expression of the input signal to enhance the representation ability and make the learned features more dividable. ReLU makes the weights in the shallow layer more trainable when using back-propagation learning method to optimize the parameters. Here, Adam (Adaptive Moment Estimation) algorithm is used for the optimization in this paper. Its can be described as below:

$$\begin{aligned} m_t &= \mu * m_{t-1} + (1-\mu) * g_t, \\ n_t &= \upsilon * n_{t-1} + (1-\upsilon) * (g_t)^2, \\ \hat{m}_t &= \frac{m_t}{1-\mu_t}, \\ \hat{n}_t &= \frac{n_t}{1-\upsilon_t}, \\ \triangle\theta_t &= -\frac{\hat{m}_t}{\sqrt{\hat{n}_t+\varepsilon}} * \eta \end{aligned} \tag{6}$$

$m_t$and $n_t$are the first order moment estimation and the second order moment estimation of the gradient; $-\frac{\hat{m}_t}{\sqrt{\hat{n}_t+\varepsilon}}$is an dynamic regularizer to the learning rate; $\mu$is the momentum factor, $\eta$is a global learning rate, $\varepsilon$is a constant to make sure the denominator is not zero.

The output layer will use Softmax to calculate the probability of the sample belonging to each class. All the layers are trained simultaneously using a version of the back-propagation learning algorithm.

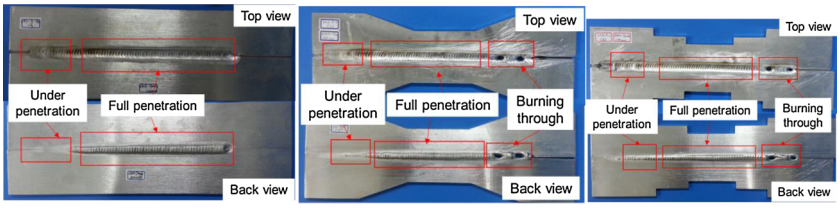
During the training progress, a CNN with parameters, $\theta = \{k_{ij}, w_{ij}, b_j\}$, is trained by minimizing the cross-entropy loss function. Specially, the loss function is defined as follow:

$$\begin{aligned} Li &= -\log\left(\frac{e^{f_{yi}}}{\sum_j e^{fj}}\right) \\ f_i(z) &= \frac{e^{zj}}{\sum_k e^{zk}} \end{aligned} \tag{7}$$

Where f_i is the softmax function and its output changes from 0 to 1. To minimize the loss function, those weight tensors are first assigned as random numbers near to zero.

For the initialization of the training model, normal distribution was applied to generate the random number for model parameters.

And then the forward propagation is executed. Cross entropy is used to determine how close the actual output is to the desired output, which refers to under penetration, normal penetration and burning through.



a) Rectangle-shaped plate b) dumbbell-shaped plate c) ladder-shaped plate

Fig. 4. Welding seams with different penetration defects of Al alloy in GTAW.

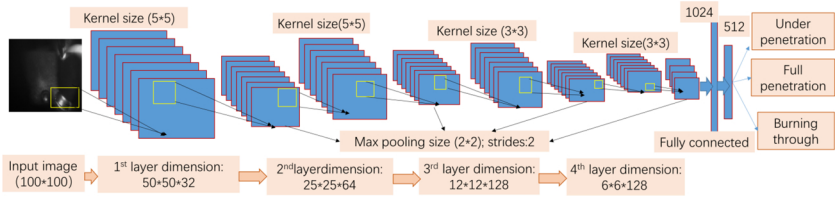


Fig. 5. Architecture of CNN model of weld defects classification.

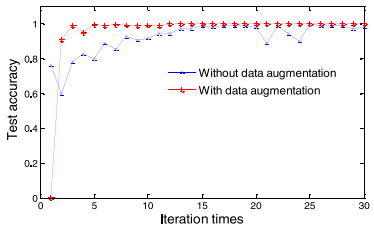


Fig. 6. Comparison between the datasets with and without data augmentation.

Table 2
Parameters for the CNN model of weld defects classification.

Layer	Kernel size	Stride	Operator	Number or method
L1-C1	5 × 5	1	convolution	Filters number 32, strides(1,1)
L2-P1	2 × 2	2	pooling	Max
L3-C2	5 × 5	1	convolution	Filters number:64 Padding:same
L4-P2	2 × 2	2	pooling	Max
L5-C3	3 × 3	1	convolution	Filters number:128 Padding:same
L6-P3	2 × 2	2	pooling	Max
L7-C4	3 × 3	1	convolution	Filters number:128 Padding:same
L8-P4	2 × 2	2	pooling	Max
L9 –F1	–	–	full connected	1024
L10–F2	–	–	full connected	512
–	–	–	activation	ReLu
–	–	–	loss	softmax_cross_entropy
Output	–	–	–	0 1 2

3.4. Data augmentation

In this section, two kinds of data augmentation were applied to boost the original dataset including noise addition and rotation. For salt and pepper noise, the noise density is considered as the percentage of the image noise area, which is set to 0.05. Gaussian white noise of mean 0 and variance 0.01 was added to weld images. For normal penetration sample, salt-pepper noise was added to each weld image. For the under penetration and burning through defects, both two types of noise were added in order to boost the defect samples. Besides, the images were randomly rotated from 15 degrees to 30 degrees. Furthermore, the data in the three channels of the weld image have been analyzed. The value from the three channels shows little difference. Therefore, the original images have been transferred into grey-scale. The state of weld penetration was manually offline identified based on the state of back seam as seen in Fig.4. There is no formation of back seam for under penetration while the back seam is burned through for burning through, which both weaken the strength of weld seam. The final datasets for CNN defects classification model was given in Table 3, in which the whole data was divided into two parts with certain proportion before training. Specifically, 80% of the total data was used for training of the CNN model while the rest is for model testing.

3.5. Optimization of the CNN model

3.5.1. Data augmentation

Based on the framework of CNN in last section, the data augmentation of noise addition was investigated regarding its influence to classification accuracy of the model. In addition, the learning rate was set to 0.001 with maxpooling, resized to 100*100, 30 of iteration times and batch size of 10. In order to stop iterating of the model, iteration number is decided when the changing curve of classification rate is stable.Fig.6 displayed the curves of test accuracy along with the iteration times for datasets with and without data augmentation. It can be clearly seen from Fig. 7-a and b that the model has shown higher accuracy. By calculating its mean value, the average test accuracy without DA is 92.3% while it is 96.18% for data with DA, which means the average accuracy was improved by 3.88% by means of the data augmentation technology.

Table 3
Datasets after data augmentation.

Datasets	Under penetration	Full penetration	Burning through	Total number	Data augmentation
Total	1200	2148	1416	4764	Noise addition
Training	960	1718	1132	3810	Rotation
Testing	240	429	283	952	

3.5.2. Resize of input image

In this section, the size of the input image was investigated by resizing to 100*100, 200*200 and 300*300 from original size of 1390*1040 in terms of the test accuracy and test time of the CNN model. The learning rate was set to 0.0015, while 20 of iteration times and batch size of 10 were used with maxpooling. For resizing image of 200*200, the feature dimension was 100*100*32, 50*50*64, 25*25*128 and 12*12*128 for the four layers respectively. For resizing image of 300*300, the feature dimension was 150*150*32, 75*75*64, 37*37*128 and 18*18*128 for the four layers respectively. Their test accuracy and testing time for each sample were compared and displayed in Fig.7.

Test time was also calculated under the computer environment of Inter (R) Core (TM) i7-8700 CPU@3.20 GHz and 32 G memory in 64bits of Windows 10 while the programming environment is Python.

Fig. 7-a and b display the test accuracy and time for different size of input image, from which it can be seen that resizing to 100*100 shows the best performance in terms of the test classification accuracy and time efficiency. In comparison, the size of the input image might have more influence on the test time efficiency than the classification accuracy. Both factors should be weighted because for the industry of robotic welding, especially the rapid manufacturing one, on-line detection efficiency is quite important as well as the inspection accuracy.

3.5.3. Filters of convolution layer

In this section, the filters number of convolution layer was carefully investigated. Four different structures and number of convolution filters were designed and compared regarding the test accuracy and test time of the CNN model. Besides the number change of number of convolution filters, other parameters are the same with that in Table 2.

Fig. 8-a shows the mean statistics of the test accuracy for the four different number of filters in each convolution layers. It can be seen that structure of 32-64-128-128 has shown the best performance with the highest mean test accuracy of 94% and lower standard deviation. For the bigger number of filters of 64-128-256-256, although the test accuracy and std showed the similar value with that of 32-64-128-128, the feature number might be excessive and redundant feature can decrease the performance the model, especially the testing time and efficiency. Judging from Fig. 8-b, the testing time for 10 samples increased with the number of filter, which means that the test efficiency is greatly affected

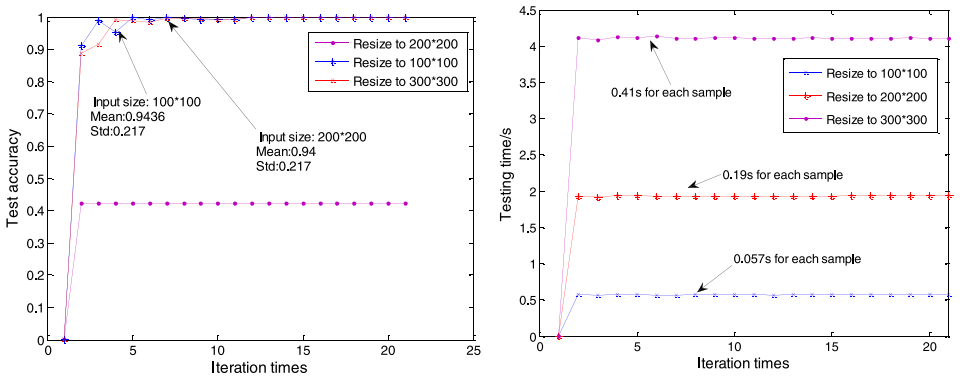


Fig. 7. Comparison of input image size for the CNN model.
a) Mean test accuracy b) Testing time.

by the filters number in convolution layer. In addition, proper filter number is important for the accurate classification in order to extract appropriate number of image features for the specific classification purpose. For comprehensive consideration, 32-64-128-128 contributes the best performance of CNN in terms of test accuracy and testing time.

3.6. Visualization of deep learning features

Number In this section, the deep learning image features from each convolution-pooling layer were analyzed and visualized using the proposed Non-zero Pixel (NOP) method in order to reveal the extraction pattern of the CNN model and explain the physical meaning of the deep learning features. Different number of filters in convolution layer were quantitatively compared and the extracted image features in each layer were visualized based on NOP.

3.7. NOP

NOP was proposed to quantitatively evaluate the features extracted in each convolution –pooling layer and then visualize the images with learned features from the CNN classification model. Bellowing is the details of the NOP method.

The key of NOP method is to find the images with non-zero pixels from all the output data in each layer. First, the original output data needs to be reshaped into 2-dimension image with the size of w*l. Then, search every data point from the 2D image to find out if there is any pixel whose value is zero and count the number of images with non-zero pixels. Based on the careful reviewing all the feature data, it is found that most of the feature value is zero providing little information of the original image, which means that nothing has been learned from the image using the convolution-pooling algorithm. Therefore, the key of NOP is to find the images with non-zero pixels. Fig. 9 gives the example of NOP for image features extraction for the CNN model designed in this paper. The output dimension from first convolution-pooling layer is 100*100*32, therefore, when doing NOP its output data needs to be reshaped into 100*100 with the number of 32 for 2D weld feature images. For the fourth convolution-pooling layer, the number of 2D weld feature image will be 128, NOP can determine how much of them have the effective image feature and map the extracted feature back to the original weld images.

4. Quantitative comparing based on NOP

Table 4 lists the numbers of NOP feature for different framework of CNN model in each convolution-pooling layer, from which it can be seen that the structure of 32-64-128-128 owns the highest number of NOP feature comparing with other network structures. This might contribute the highest classification accuracy (Fig. 8a) of CNN model due to more features have been deeply extracted. It also means that the more the filters is, the deeper the CNN can dig, the more the effective feature is. Besides, the NOP features for 64-128-256-256 is much smaller even through its classification accuracy is close to the former one, which limits its performance because of the limited extracted feature. The other thing is that there are only few images owning useful information based on the NOP method as seen in Table 4. We have tried to use less or more filters to change and compare the model structure as displayed in Fig. 8-a, from which the mean test accuracy and standard deviation have been compared. One thing is clear that the performance did not improved when less or more filters are used. In future, it might be worth to investigate other CNN model structures such as RCNN to focus on the main area of the image to increase the efficiency of CNN.

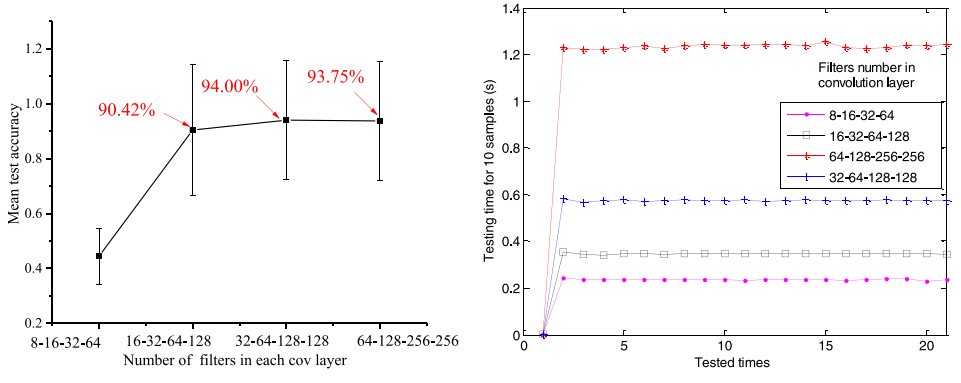


Fig. 8. Comparison between different filters numbers of convolution layer for the CNN model.

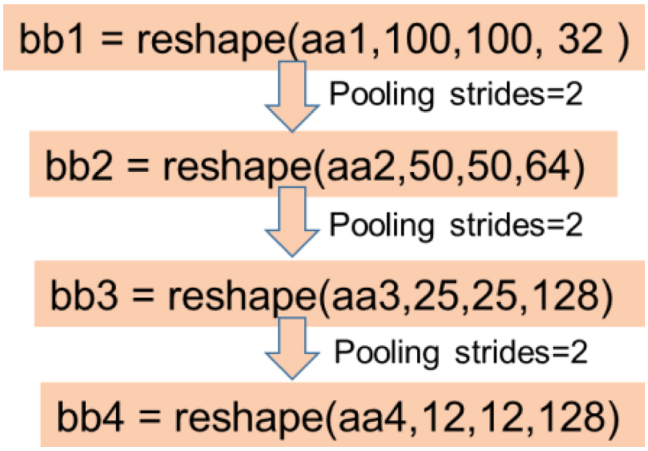


Fig. 9. Extraction of image features based on NOP for the CNN model.

Table 4
Numbers of NOP features for different framework of CNN and pooling method.

Convolution layer	8-16-32-64	16-32-64-128	32-64-128-256	64-128-256-256
1stlayer	5	3	14	7
2nd layer	5	3	3	8
3rd layer	0	8	79	12
4th layer	0	30	35	55

4.1. Visualization of DL features based on NOP

Fig. 10 demonstrated the learned deep image features from the CNN model with the convolution framework of 16-32-64-128 using max-pooling for a full penetration weld image. NOP was utilized to select these effective features with non-zero pixels.

After convolution-pooling processing in the first layer in Fig. 10, it can be seen that the image has been enforced by denosing and contrast improving, especially for the back view of weld seam, which has shown the higher brightness. Then, the second layer of convolution reserved only one effective image feature, in which the weld seam, arc light and wire filler have been blurred due to the further convolution. However, the different degree of arc light has been extracted from the top weld pool, wire filler and back seam. In the third layer, the extracted features were quite similar with each other and cannot be clearly distinguished in naked eyes.

Finally, more deep features have been extracted with different characteristics, which were marked in red rectangle in the fourth layer. First, several bright areas with different degree of brightness was extracted as deep features, which were originated from the arc light or its refection. Second, those bright areas either solely appears or combines with each other, which demonstrates the complementary features. For instance, the bright area from top weld pool and back seam were combined into one feature while the top weld pool and wire filler were also combined to form the new feature. Therefore, by means of visualization of the deep learning features, those image features can be well explained in terms of its physical meaning and extraction pattern.

5. Discussions and conclusions

In this paper, a new on-line detection of weld defects for Al alloy in pulsed GTAW was proposed based on convolutional neural networks and weld images. The image acquisition system developed in this paper can capture the images of weld pool from different angles and provide more abundant information for accurate defects classification. However, it contains multiple mirror reflection and might be difficult to use in practical industry application. In future, the image acquisition system will be more simplified to improve its practicability.

Two kinds of data augmentation methods were applied to boost the original dataset including noise addition and image rotation and has improved the accuracy about 3.88%. Moreover, several key parameters including filters of convolution layer and size of input image were optimized for the CNN model. Then, a new CNN classification model based on weld image with 11 layers was designed to identify weld seam penetration defects.

Comparing with our previous work [3], audible arc sound was utilized to identify the defects of under penetration and burning through based on the proposed classification model SVM-GSCV, by means of which the test accuracy was changing from 81.52% to 98.46% and the average accuracy of classification is 87.16%. In this paper the

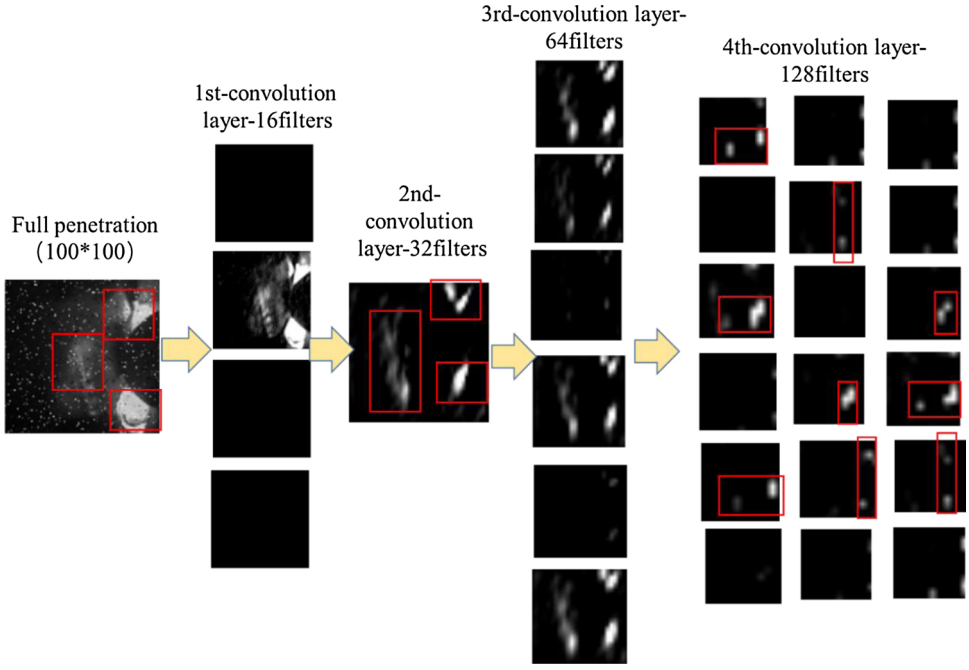


Fig. 10. Visualization of deep learning features in the CNN model based on NOP.

same batch of experiment dataset were researched for the same detection purpose using the designed CNN model based on weld images, which has shown the better performance with the mean accuracy of 99.38% with the std value of 0.018.

Finally, visualization of the deep learning features in each convolution layer was carried out using the proposed NOP method to explain the physical meaning of the deep learning features. It was found that the deep learning features were mainly dependent on the brightness of arc light due to the strong arc light emission. Instead of decreasing or eliminating the interference from arc light as in traditional research way, the CNN model has taken full use of those arc lights by combining them in a various way to form the complementary features, which provides us a new and innovative solution for real-time monitoring information processing and defect detection.

In future, more macro weld defects should be included and investigated to expand the datasets and improve the robustness of the CNN model. It is worth to mention that the proposed defects classification model might be limited to surface defects and can not detect the inner defects like porosity, cracks due to the limitation of the vision sensing. Although, other types of sensing technologies, such as arc optical spectrum, thermal infrared temperature might provide better solution to achieve on-line detection of inner defects. In addition, more effective optimization method is needed for multiple parameters of the CNN model. Other advanced method including R-CNN or Fast R-CNN might be considering to selectively search regions of interests of the image.

Acknowledgement

The work was supported by National Natural Science Foundation of China, No.51605372, China postdoctoral science foundation funding, No.2018T111052, No.2016M602805, National Natural Science Foundation of China, No. 51775409, the Program for New Century Excellent Talents in University (NCET-13-0461).

References

[1] Chen SB, Wu J. Intelligentized methodology for arc welding dynamical processes: visual information acquiring, knowledge modeling and intelligent control vol. 29. Springer Verlag; 2008.
[2] Lv N, Xu Y, Li S, Yu X, Chen S. Automated control of welding penetration based on audio sensing technology. J Mater Process Technol 2017;250:81–98.
[3] Zhang Z, Wen G, Chen S. Audible Sound-based Intelligent Evaluation for Aluminum Alloy in Robotic Pulsed GTAW: mechanism, feature selection and defect detection. IEEE Trans Industr Inform 2017.
[4] Mirapeix J, Vila E, Valdiande JJ, Riquelme A, Garcia M, Cobo A. Real-time detection of the aluminium contribution during laser welding of Usibor1500 tailor-welded blanks. J Mater Process Technol 2016;235:106–13.
[5] Song L, Huang W, Han X, Mazumder J. Real-time composition monitoring using support vector regression of laser-induced plasma for laser additive manufacturing. Ieee Trans Ind Electron 2017;64:633–42.
[6] Zhang Z, Yu H, Lv N, Chen S, Zhang Z, Yu H, et al. Real-time defect detection in pulsed GTAW of Al alloys through on-line spectroscopy. J Mater Process Technol 2013;213:1146–56.
[7] Huang Y, Wu D, Lv N, Chen H, Chen S. Investigation of porosity in pulsed GTAW of aluminum alloys based on spectral and X-ray image analyses. J Mater Process Technol 2017;243:365–73.
[8] Chen B, Wang J, Chen S. A study on application of multi-sensor information fusion in pulsed GTAW. Ind Rob 2010;37:168–76.
[9] Wu D, Huang Y, Chen H, He Y, Chen S. VPPAW penetration monitoring based on fusion of visual and acoustic signals using t-SNE and DBN model. Mater Des 2017;123:1–14.
[10] Zhang Z, Chen S. Real-time seam penetration identification in arc welding based on fusion of sound, voltage and spectrum signals. J Intell Manuf 2017;28(January):207–18.
[11] You D, Gao X, Katayama S. Multisensor fusion system for monitoring high-power disk laser welding using support vector machine. IEEE Trans Industr Inform 2014;10(May):1285–95.
[12] Liu Y, Zhang Y. Iterative local ANFIS-Based human welder intelligence modeling and control in pipe GTAW process: a data-driven approach. Ieee/asme Trans Mechatron 2015;20(June):1079–88.
[13] Liu YK, Zhang YM. Fusing machine algorithm with welder intelligence for adaptive welding robots. J Manuf Process 2017;27(June):18–25.
[14] Liu YK, Zhang YM. Supervised learning of human welder behaviors for intelligent robotic welding. Ieee Trans Autom Sci Eng 2017;14(July):1532–41.
[15] He Y, Chen Y, Xu Y, Huang Y, Chen S. Autonomous detection of weld seam profiles via a model of saliency-based visual attention for robotic arc welding. J Intell Robot Syst 2016;81(March):395–406.
[16] Calta NP, Wang J, Kiss AM, Martin AA, Depond PJ, Guss GM, et al. An instrument for in situ time-resolved X-ray imaging and diffraction of laser powder bed fusion additive manufacturing processes. Rev Sci Instrum 2018;89:055101.
[17] Cha Y-J, Choi W, Buyukozturk O. Deep learning-based crack damage detection using convolutional neural networks. Comput Civ Infrastruct Eng 2017;32(May):361–78.
[18] Acharya UR, Fujita H, Oh SL, Hagiwara Y, Tan JH, Adam M. Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. Inf Sci (Ny) 2017;415:190–8.
[19] Jia F, Lei Y, Lu N, Xing S. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. Mech Syst Signal Process 2018;110:349–67.
[20] Dou Q, Chen H, Yu L, Zhao L, Qin J, Wang D, et al. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans Med Imaging 2016;35:1182–95.
[21] Liu T, Bao J, Wang J, Zhang Y. A hybrid CNN–LSTM algorithm for online defect recognition of CO2 welding. Sensors 2018;18:4369.
[22] Shevchik SA, Kenel C, Leinenbach C, Wasmer K. Acoustic emission for in situ quality monitoring in additive manufacturing using spectral convolutional neural networks. Addit Manuf 2018;21(May):598–604.
[23] Zhenhua C, Guoliang H, Chao L, Guo C. Automatic recognition of weld defects in TOFD D-Scan images based on faster R-CNN. J Test Eval 2018;48:14. pp.-14 pp., 2018.
[24] Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 1980;36:193–202.
[25] Bengio Y. Learning deep architectures for AI. Found Trendsâ® Mach Learn 2009;2:1–127.