

Research Article

Open Access



Large language models enabled intelligent microstructure optimization and defects classification of welded titanium alloys

Suyang Zhang¹, William Yi Wang^{1,2,*} , Xinzhao Wang¹, Gaonan Li¹, Yong Ren¹, Xingyu Gao³ , Feng Sun¹, Bin Tang¹, Haifeng Song^{3,*}, Jinshan Li^{1,2,*} 

¹State Key Laboratory of Solidification Processing, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China.

²Innovation Center, NPU Chongqing, Chongqing 401135, China.

³Institute of Applied Physics and Computational Mathematics, Beijing 100088, China.

* **Correspondence to:** Prof. William Yi Wang and Prof. Jinshan Li, State Key Laboratory of Solidification Processing, Northwestern Polytechnical University, No. 127, Youyi West Road, Beilin District, Xi'an 710072, Shaanxi, China. E-mail: wywang@nwpu.edu.cn; ljsh@nwpu.edu.cn; Prof. Haifeng Song, Institute of Applied Physics and Computational Mathematics, No. 2, Fenghao East Road, Haidian District, Beijing 100088, China. E-mail: song_haifeng@iapcm.ac.cn

How to cite this article: Zhang S, Wang WY, Wang X, Li G, Ren Y, Gao X, Sun F, Tang B, Song H, Li J. Large language models enabled intelligent microstructure optimization and defects classification of welded titanium alloys. *J Mater Inf* 2024;4:34. <https://dx.doi.org/10.20517/jmi.2024.64>

Received: 28 Oct 2024 **First Decision:** 19 Nov 2024 **Revised:** 9 Dec 2024 **Accepted:** 16 Dec 2024 **Published:** 31 Dec 2024

Academic Editor: Chaolin Tan **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

The quick developments of artificial intelligence have brought tremendous attractive opportunities and changes to smart welding technology. In the present work, a novel model, ConvNeXt, which incorporates the advantages of convolutional neural networks (CNNs) and vision transformers (ViTs), has been designed to identify welding defects. The classification accuracy of the pre-trained ConvNeXt based on transfer learning method reaches as high as 99.52% after 500 iterations of training, while traditional CNNs of MobileNetV2 and ResNet34 achieve 85.94% and 93.41%, respectively. Moreover, the classification performance can be further improved through dataset optimization based on t-distributed stochastic neighbor embedding (t-SNE). In addition, arc geometrical features are added as input parameters for building a back propagation neural network to predict the formation of the weld seam, which has led to a reduction in the maximum prediction error for weld seam thickness from 0.8 to 0.6 mm. Furthermore, out of 28 sets of experimental parameters, only four sets result in errors exceeding 0.2 mm. It is worth noting that large language models (LLMs) are utilized to facilitate the automated programming for welding defect recognition, including ChatGPT 3.5, Bing Copilot, Claude3, and ERNIE Bot. LLM-aided automated



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



programming technology is applied to develop image stitching programs, achieving unsupervised automatic stitching of multiple welding tissue images and obtaining clear and wide-field weld ones. These case studies of deep learning technologies and automated programming based on LLMs set up a solidified building block for smart welding defect recognition during non-equilibrium solidification.

Keywords: Welding defect recognition, convolutional neural networks, back propagation neural network, large language models, automated programming

INTRODUCTION

Non-equilibrium solidification, characterized by exceptionally rapid cooling rates that obstruct thermodynamic equilibrium, assumes a pivotal role in contemporary manufacturing processes^[1,2]. Depending on the principles of non-equilibrium solidification, welding has become indispensable in producing materials with specific properties, especially in industries such as shipbuilding, navigation, and aerospace^[3-5]. However, manual welding is fraught with inefficiencies, high expenditure and inconsistent performance, which can compromise weld quality^[6]. Consequently, the advancement of intelligent welding monitoring systems has become imperative for enhancing performance and ensuring reliability in production^[7-9]. The systems encompass welding defect recognition, welding parameter-geometry relationship establishment, and automatic programming with the aid of large language models (LLMs)^[10,11].

On the one hand, deep learning algorithms have been considered as one kind of key component of intelligent welding monitoring systems, which are instrumental in recognizing welding defects^[12,13]. Defects including porosity, cracks, and insufficient fusion can significantly endanger weld quality without timely identification throughout the welding process^[14,15]. Continuous manual oversight of defects in the welding process is neither practical nor cost-effective^[16]. Fortunately, recent advancements in machine vision technology have endowed welding robots with the capability to autonomously identify defects^[17,18]. At the nucleus of machine vision resides the deep learning model, with convolutional neural networks (CNNs) excelling in image processing tasks^[19-21]. With multiple layers and deep architecture, CNNs can extract features from welding images to classify defects^[20,22]. Since the vision transformer (ViT) emerged in 2020, the potential of Transformer-based architectures in computer vision has gained widespread acknowledgment^[23,24]. However, ViT has notable drawbacks compared to CNNs including a large number of model parameters and high computational demands which pose challenges for achieving lightweight deployment^[25]. Meta AI attributes ViT's superior performance over CNNs to significant advancements in architectural design and optimization techniques, which inspired the creation of ConvNeXt^[26,27]. The innovation strikes a balance between recognition accuracy, storage requirements and computational efficiency, making ConvNeXt highly suitable for deployment in real-time welding monitoring systems without compromising production efficiency or imposing significant storage burdens^[28]. Meanwhile, the ConvNeXt architecture incorporates the exceptional ability of the Transformer framework to capture spatial and structural relationships in images, enabling it to deliver outstanding performance in welding defect image recognition tasks.

On the other hand, the intricate and nonlinear interactions among welding heat input, weld joint microstructures and the subsequent weldment performance present a formidable challenge^[29,30]. Precisely delineating the relationships, achieving accurate predictions of weld geometry and optimizing process parameters continue to pose substantial hurdles in the field^[31]. Despite the extensive accumulation of experimental data, the inherent complexity and scale of the information present significant barriers to uncovering the underlying principles through conventional analysis^[32,33]. This is where deep learning technology provides a transformative solution, leveraging the unparalleled capacity to model nonlinear

relationships^[34]. Integrating simulation techniques with deep learning models facilitates the automatic identification of correlations between welding parameters and weld seam microstructures, enabling accurate predictions of weldment performance^[35]. The approach not only optimizes process parameters and weld quality but also significantly reduces the reliance on costly and time-intensive experimental trials, pushing the boundaries of welding technology^[36].

Moreover, the development of the monitoring systems involves sophisticated image processing and neural network models, demanding considerable human effort and advanced programming expertise^[37-40]. In order to tackle the complexity of programming, the LLMs have emerged as powerful tools for automating programming tasks^[40-43]. Based on transformer architecture, LLMs utilize self-attention mechanisms to capture contextual dependencies in programming, thereby enabling the generation of coherent and efficient outputs^[44-46]. The models including GPT-3.5, BERT, and Copilot have demonstrated remarkable potential in automating code generation and optimization tasks^[47,48]. The capacity of programming explanation and generation renders the LLMs invaluable assets for programming automatic welding systems^[49,50].

To address the aforementioned challenges, a progressive monitoring system for the intelligent welding industry is presented in the research, harnessing wide-ranging image data amassed from tungsten inert gas (TIG) experiments. In Section “Weld state classification based on convolution neural network”, CNNs are employed for image data classification, followed by the evaluation of classification metrics and feature visualization to enhance the classifiers. In Section “Weld seam forming prediction on back propagation neural network”, backpropagation neural networks (BPNNs) are applied to establish a mapping between welding process parameters and welding formation geometry with model performance further ungraded by incorporating arc geometric features. In Section “Automatically programming based on LLMs”, LLMs are introduced to support programming tasks in image processing, with the code generation capabilities evaluated for arc contours extracting and welding image stitching.

MATERIALS AND METHODS

Artificial intelligence agents have been effectively integrated into the workflow of smart welding as assistants in the field of intelligent manufacturing^[51]. As illustrated in [Figure 1](#), LLMs are utilized to automate programming based on welding image data gathered from high-throughput experiments, with the objective of optimizing the defect detection process during welding of titanium alloys^[52]. The methodologies encompass titanium alloy tube-to-plate welding experiments, deep learning techniques and automatic programming of intelligent welding based on LLMs.

Tube-to-tube-sheet welding experiments for titanium alloy

Owing to the high chemical reactivity of titanium and the propensity to adsorb hydrogen, oxygen and nitrogen at elevated temperatures, conventional welding techniques such as manual metal arc welding, gas welding and CO₂ gas-shielded welding are deemed unsuitable^[53]. TIG welding is employed in the research, with the arc generated between the tungsten electrode and the workpiece to melt the metal, while the inert gas is introduced around the electrode to safeguard the metal and maintain arc stability^[54]. The schematic diagram of a TIG tube-to-tube-sheet welding robot system is shown in [Figure 2](#). The heat exchanger tubes, tube sheet and welding wire are all composed of TA2, which is an α -phase titanium alloy with excellent corrosion resistance and cold workability^[55]. The chemical composition of TA2 is presented in [Table 1](#). The heat exchanger tubes feature an outer diameter of 10 mm and a wall thickness of 1.5 mm. The tube sheet, measuring 100 mm in thickness, includes a 1 mm × 1 mm chamfer at a 45° angle. The diameter of the welding wire is 0.8 mm.

Table 1. The chemical composition table of titanium alloy TA2

Material	Fe	C	N	H	O	Ti
TA2	0.027	0.0063	0.0047	< 0.001	0.10	Margin

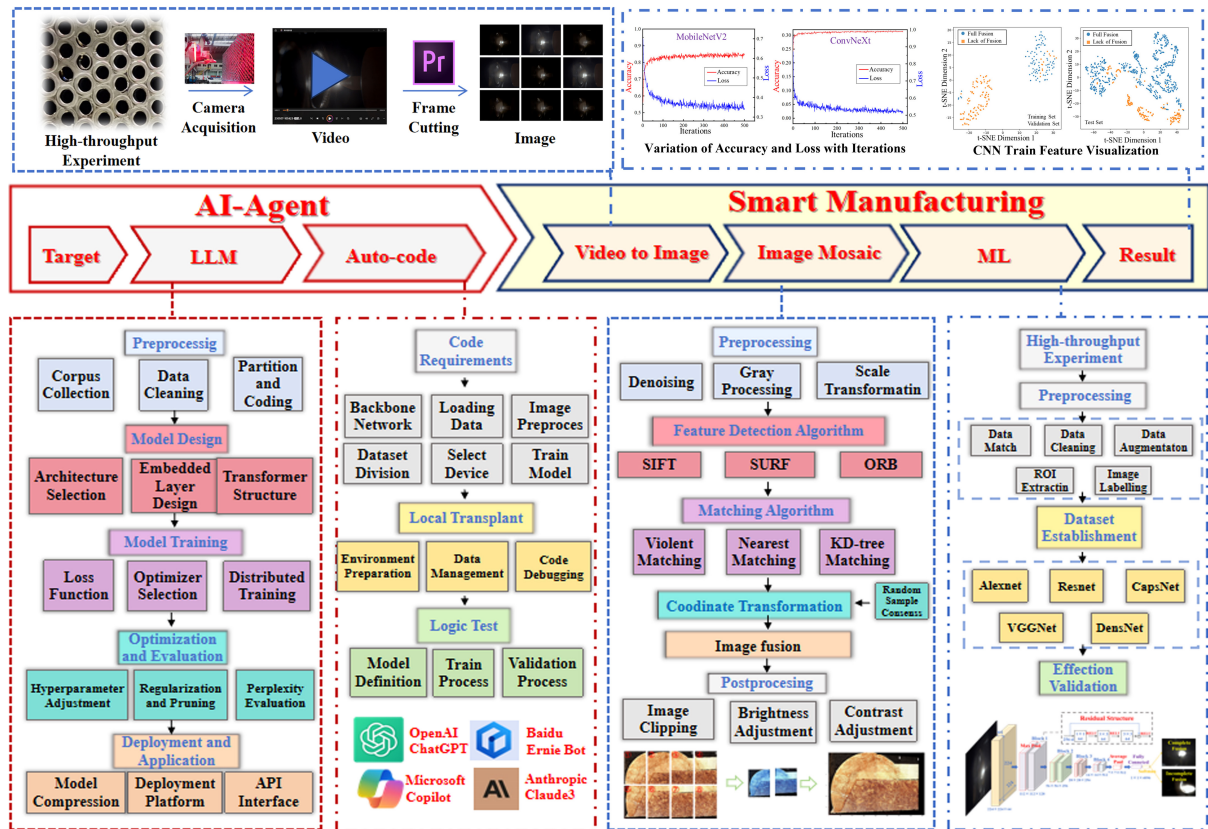


Figure 1. The technical roadmap of AI agent assistant smart welding consisted of five basic sections, including large language model, auto-coding via AI agent, image processing, image mosaic and machine learning for welding defect detection. AI: Artificial intelligence.

The intelligent system is equipped with the TPR 2000 welding machine, 500A programmable power supply, real-time data acquisition system and vision system^[56]. In the vision system, the Xiris XVC-1100 high dynamic range (HDR) camera is employed to observe the welding process with a maximum frame rate of 55 fps^[21]. High-speed imaging provides a direct approach to capturing the melting dynamics of the welding wire and the flow behavior of the molten pool^[57,58]. The camera position is refined through multiple experiments to capture high-quality images of the molten pool^[59]. The input/output channels of the camera are equipped with photoelectric isolation to protect against electromagnetic noise. As a full factorial experimental design requires testing all possible parameter combinations which is both time-intensive and costly, the orthogonal experimental method is employed to efficiently limit the number of tests while ensuring experimental validity^[60]. Table 2 shows the parameter values for each experimental sample, including pulse current (I_p), welding speed (V_s), pulse width (t_p), duty cycle ratio (δ) and weld seam thickness at the arc starting point (H).

Deep learning methods

BPNNs constitute a class of multi-layer feedforward neural networks trained through the error backpropagation algorithm^[61]. As depicted in Figure 3, the typical architecture of BPNNs comprises three

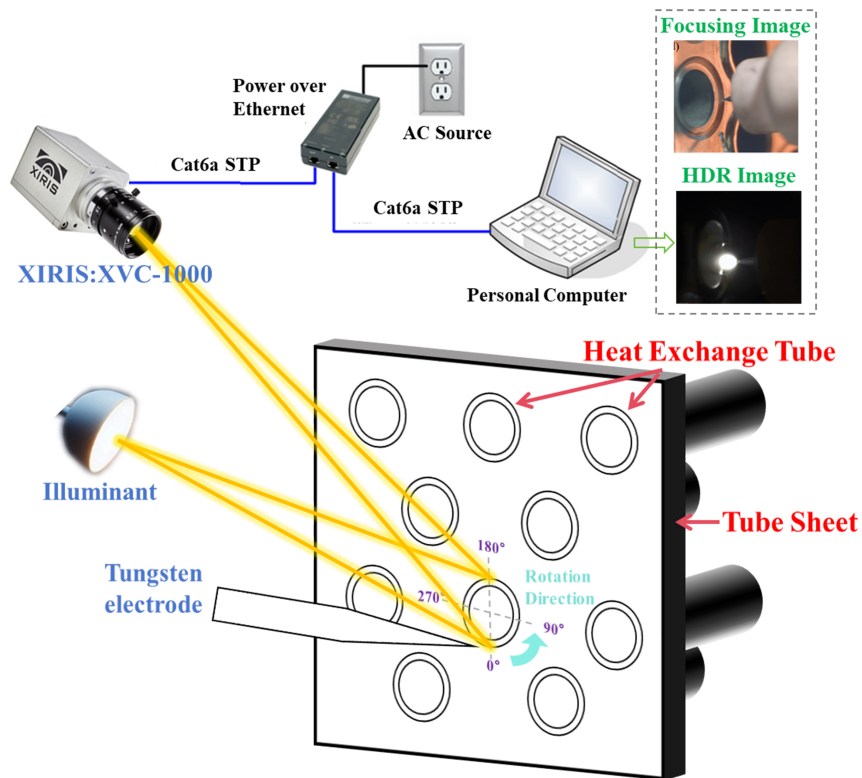


Figure 2. The schematic diagrams of intelligent robot system of TIG tube-to-tube-sheet welding with XVC-1100 HDR camera. TIG: Tungsten inert gas; HDR: High dynamic range.

primary layers: an input layer, one or more hidden layers, and an output layer^[62]. Neurons within these layers are interconnected by weights and no connections exist between neurons within the same layer or across non-adjacent layers. The input layer receives external input data and forwards the data to the hidden layer. As the core of the BPNNs, the hidden layer contains multiple neurons responsible for the nonlinear transformation and feature extraction of the input data. The number of hidden layers and neurons can be adjusted based on the complexity and specific requirements of the problem. The output layer receives processed information from the hidden layer and produces the final output of the BPNN.

CNNs represent a foundational class of deep learning models that have attained remarkable success in computer vision, especially in the domain of image recognition^[63,64]. CNNs excel at image perception due to the convolutional operations, which simulate the biological visual system by extracting features from input data through localized perception and weight sharing^[65]. The basic structure of CNNs consists of convolutional layers, pooling layers and fully connected layers. Convolutional and pooling layers are alternately stacked, followed by one or more fully connected layers that generate the final outputs^[66]. Numerous CNN models have been developed, with prominent examples including ResNet34, MobileNetV2 and ConvNeXt, which will be discussed in detail below.

ResNet34 architecture

In contrast to traditional machine learning, deep learning is distinguished by its intricate network architectures, which are essential for significantly enhancing performance. However, as the depth of the network increases, issues such as vanishing and exploding gradients may arise, negatively affecting the training process^[67]. As depicted in [Figure 4](#), ResNet mitigates this challenge by incorporating residual blocks,

Table 2. The technical parameters of pulsed TIG welding for each experimental group with the orthogonal experimental method

Group numbers	I_p (A)	V_s (mm/min)	t_p (s)	δ	H (mm)
1	125	120	0.04	0.67	1.917
2	75	60	0.08	0.50	1.955
3	95	90	0.10	0.33	1.525
4	85	100	0.08	0.67	1.789
5	105	120	0.10	0.50	1.61
6	125	80	0.04	0.33	0.915
7	85	110	0.10	0.67	1.713
8	105	60	0.04	0.50	1.748
9	125	90	0.08	0.33	1.963
10	75	90	0.08	0.67	2.211
11	95	110	0.10	0.50	1.268
12	115	60	0.04	0.33	0.981
13	105	80	0.08	0.67	1.829
14	125	100	0.10	0.50	1.657
15	85	120	0.04	0.33	1.664
16	95	100	0.04	0.67	1.597
17	115	120	0.08	0.50	1.536
18	75	80	0.10	0.33	1.720
19	115	110	0.04	0.67	1.718
20	85	80	0.08	0.50	1.623
21	105	100	0.10	0.33	1.452
22	115	80	0.10	0.67	1.940
23	75	100	0.04	0.50	1.634
24	95	120	0.08	0.33	1.643
25	95	60	0.08	0.67	1.863
26	115	90	0.10	0.50	1.447
27	75	110	0.04	0.33	1.806
28	125	60	0.10	0.67	1.856
29	85	90	0.04	0.50	1.133
30	105	110	0.08	0.33	1.803
31	75	120	0.10	0.67	1.683
32	95	80	0.04	0.50	1.943
33	115	100	0.08	0.33	1.616
34	105	90	0.04	0.67	1.543
35	125	110	0.08	0.50	1.479
36	85	60	0.10	0.33	1.640

TIG: Tungsten inert gas.

thereby enabling the network to learn incremental refinements through the innovative residual structure^[68]. Furthermore, the architecture substantially accelerates training and enhances convergence rates, particularly in the context of large-scale datasets. ResNet-34, a variant of the ResNet architecture featuring 34 layers, adeptly balances depth and computational efficiency, rendering it exceptionally effective for addressing complex challenges in the welding domain^[69].

MobileNetV2 architecture

MobileNetV2 is a streamlined neural network architecture optimized for efficient image classification on mobile devices and embedded systems^[70,71]. By employing depthwise separable convolutions, the

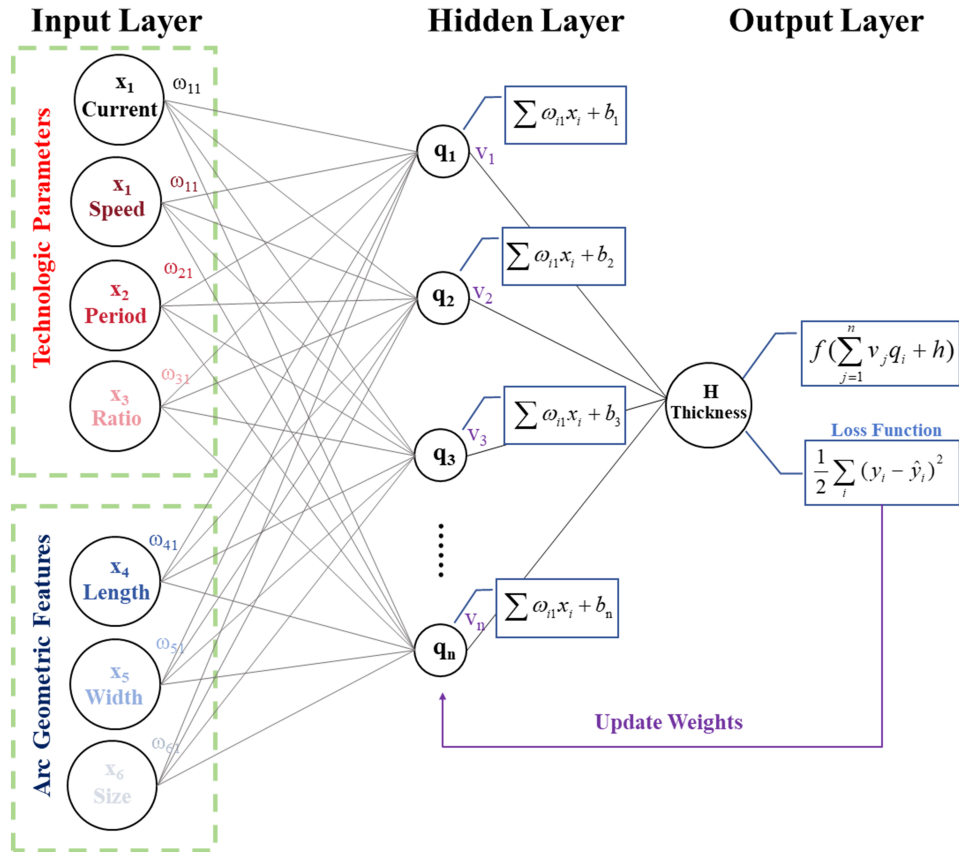


Figure 3. The structure of BPNN to predict weld seam thickness based on the input of technological parameters and arc geometric features. BPNN: Backpropagation neural network.

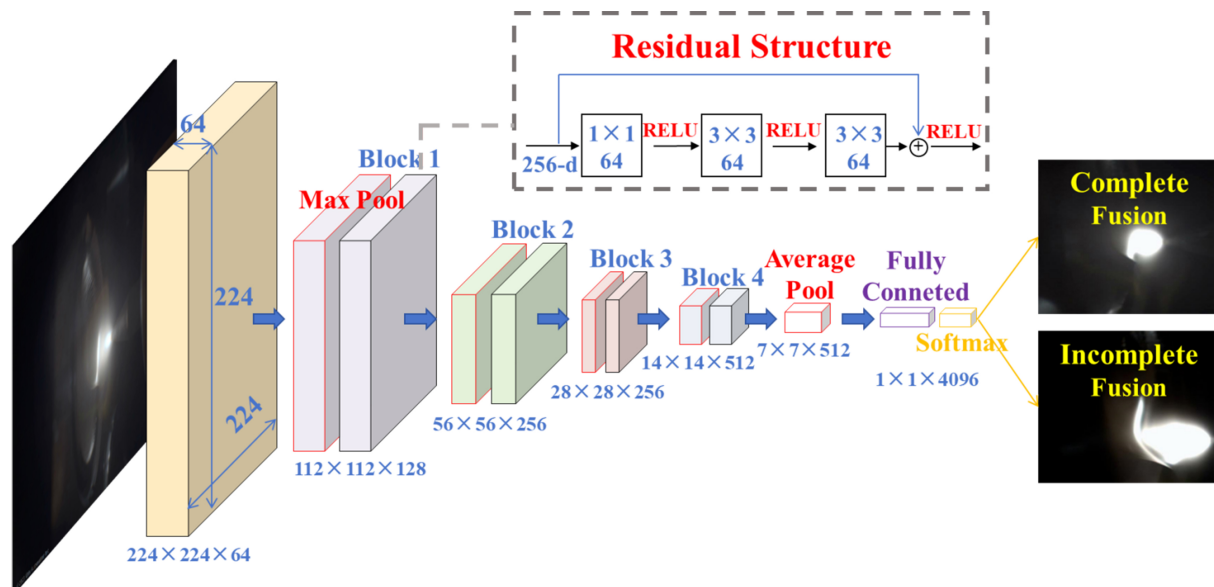


Figure 4. The structure of ResNet34 and the residual block structure for welding unfused defect classification.

architecture significantly reduces computational demands while preserving performance and minimizing

model size^[72]. The introduction of the inverted residual structure enhances gradient flow, mitigating vanishing gradients and improving overall network stability during training. The architecture also facilitates more efficient information transfer across layers, enabling deeper networks to capture data features more effectively. Compared to traditional residual connections, the inverted residual structure achieves greater computational efficiency with fewer parameters, optimizing both training and inference processes.

ConvNeXt architecture

ViT faces challenges of extensive parameters and substantial computational demands, which hinder the suitability for lightweight deployment^[73,74]. By addressing the limitations above, ConvNeXt leverages the design principles and optimization advancements inspired by ViT, combining with the efficiency of CNNs^[75,76]. The innovative architecture improves upon existing frameworks by substituting the commonly employed ReLU activation function with the Gaussian error linear unit (GELU):

$$\text{ReLU}(x) = \max(0, x) \quad (1)$$

$$\text{GeLU}(x) = 0.5x(1 + \tanh(\sqrt{\frac{2}{\pi}}(x + 0.044715x^3))) \quad (2)$$

Proposed by Hendrycks and Gimpel in 2016, GELU has gained attention for its smoother nonlinear characteristics, which improve model performance^[77]. The activation function integrates the advantages of both Sigmoid and ReLU, providing a continuous derivative that enhances gradient propagation during training and thereby mitigates the risk of gradient vanishing^[78]. Additionally, ConvNeXt enhances the performance using grouped convolution within the ConvNeXt Block, which partitions input feature maps into distinct groups for independent convolutions, thereby improving representational capacity and feature extraction efficiency^[79].

Automatic programming of intelligent welding based on LLMs

LLMs are advanced deep learning architectures trained on vast datasets, which can not only generate coherent natural language text but also deeply grasp the context and meaning^[80-82]. LLMs excel in various natural language processing tasks, such as text summarization, intelligent question-answering systems and machine translation, showcasing their versatility across multiple domains^[83]. The construction of LLMs involves several critical steps. The process begins with a requirements analysis to delineate the model's intended application, functionality and performance objectives. Relevant textual data is collected from diverse sources, including web pages, books and articles, followed by noise elimination^[84,85]. The model is trained guided by hyperparameters including learning rate, batch size and iteration count, while performance is validated based on metrics such as perplexity, F1 score, bilingual evaluation understudy (BLEU) and recall-oriented understudy for gisting evaluation (ROUGE). Additionally, LLMs undergo fine-tuning for specific tasks or domains, a process that generally requires fewer resources compared to the initial training phase^[86,87]. For the purpose of optimizing operational efficiency, compression is necessary through techniques including pruning and quantization. Finally, the model is deployed on suitable platforms equipped with a user-friendly application programming interface (API) for integration and subjected to real-time monitoring to ensure stability.

Figure 5 showcases the application of LLMs including ChatGPT-3.5, Copilot, Claude 3, and Ernie Bot in facilitating automated coding processes. In a specific test case focusing on the extraction of welding arc contours, all these models successfully generated executable programs to achieve the desired objectives. The input provided to the LLMs for this task was as follows: "We require processing welding arc images with three channels to extract the arc contours. It is known that the brightness of the arc area is significantly

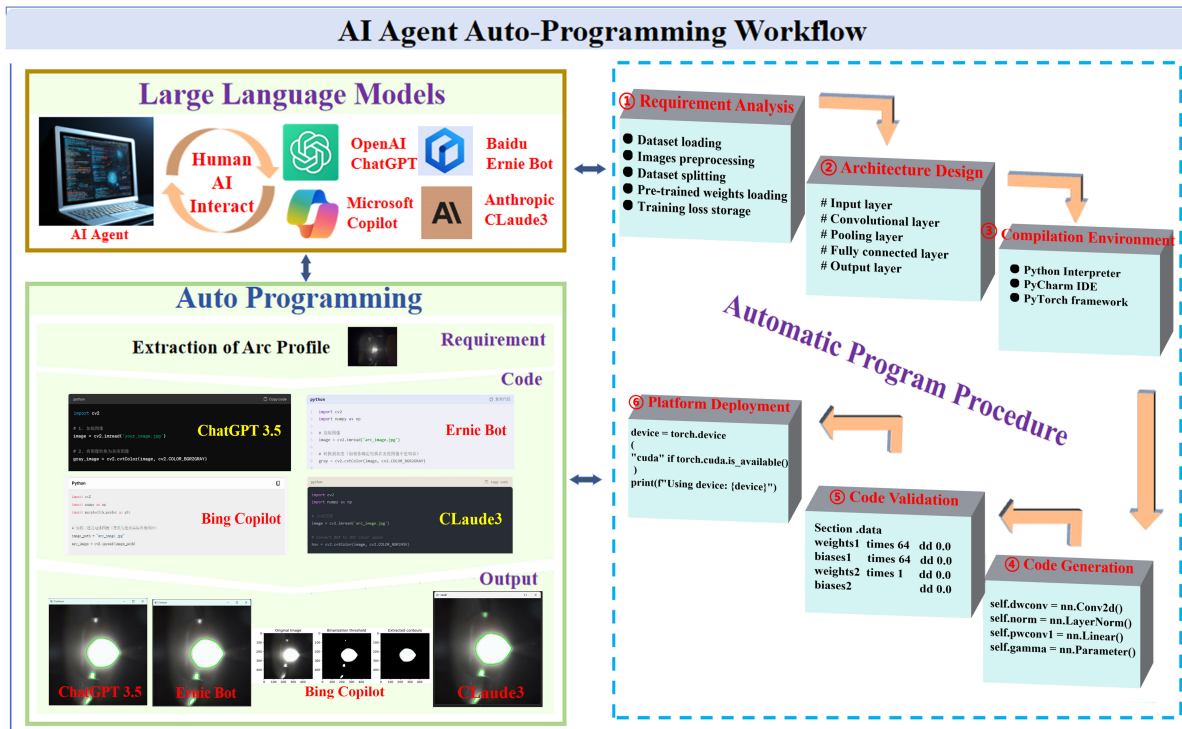


Figure 5. The process of automatic programming by a LLM via an AI agent based on ChatGPT-3.5, Copilot, Claude3 and Ernie Bot, comparing the user interface and model functions of different language models. LLM: Large language model; AI: Artificial intelligence.

higher than that of the non-arc area. Please generate executable Python programs to achieve this, along with necessary explanations”. The generated programs were seamlessly executed in the Python environment, producing welding arc contour images without requiring human intervention. This demonstrates the effectiveness of LLMs in automating complex image-processing tasks. However, the inherent stochasticity of LLMs can lead to different outputs for the same input, and the generated code might not fit the local environment. The limitations above can be mitigated by iterative refinement: users can re-input the generated programs into the LLMs with additional specifications to receive optimized versions. The iterative process can be repeated as needed until a fully functional program meeting all requirements is obtained.

RESULTS AND DISCUSSION

Weld state classification based on convolution neural network

In pulsed TIG welding, insufficient heating input may result in incomplete fusion of the weld seam, leading to defects that compromise weld strength and quality. CNNs excel in image-related tasks by directly processing raw images, thereby obviating the necessity for manually defined features and minimizing extraction errors. The capability renders the model less sensitive to image clarity, enhancing reliability and fault tolerance. In this section, CNNs will be applied to establish a classification model for unfused defects in TIG welding, which can be used for online monitoring of unfused defects during welding, preventing products with welding defects from flowing out of the production line and posing potential safety hazards during service.

Data augmentation and training parameters setting

Data augmentation is applied to experimental image data to improve the generalization and robustness of the CNNs, enabling the model to recognize images with different transformations and distortions^[88]. As

indicated in [Figure 6](#), the data augmentation methods encompass image flipping, random rotation, resizing, cropping and adjustment to lightness, saturation, contrast and color. Additionally, CNNs are heavily dependent on training samples, and changes in the target task or application routinely necessitate retraining and re-annotation, potentially reducing development efficiency. To address the problems above, the transfer learning method is introduced, which applies knowledge learned from one task to another related task^[89,90]. The models are pre-trained with the ImageNet dataset, which contains millions of labeled images commonly implemented for image recognition tasks^[91]. The pre-training enables the model to capture general image features and significantly accelerates the entire training efficiency.

Two strategies are typically adopted for updating weights in transfer learning models: partial layer freezing and full fine-tuning. Partial layer freezing fixes some layers after loading pre-trained weights, allowing only the remaining layers to be trained. Full fine-tuning utilizes pre-trained weights as initialization and updates the entire model with a lower learning rate, leveraging prior knowledge to speed up convergence. Determining the optimal freezing range necessitates extensive experimentation, which can be time-consuming and resource-intensive. Therefore, pre-trained weights are initially loaded during the construction of the defect recognition model, followed by comprehensive fine-tuning. Training parameters are shown in [Table 3](#). Due to the memory limitations of the RTX 3050 GPU, eight images per batch are used throughout local training, while 32 images per batch are used on the NVIDIA A100 platform. [Supplementary Materials](#) include the programs employed for the training and validation of CNNs in the research.

Performance evaluation of neural network models

Incomplete fusion of the weld seam during pulsed TIG welding can result in flaws that undermine the welding strength and overall quality. The CNN architectures including ResNet34, MobileNetV2, and ConvNeXt are placed into experiment to examine the unfused defect on pulsed TIG welding images. Two ResNet34 training experiments are conducted: one with randomly initialized weights, and the other with weights pre-trained on the ImageNet dataset. Both models are then trained on welding images from pulsed TIG experiments. As shown in [Figure 7](#), both versions of ResNet34 initially had a validation accuracy of around 0.55, indicating no classification capability. However, the classification accuracy of pre-trained ResNet34 swiftly ascends, reaching 93.41% after 500 iterations. In contrast, the accuracy of randomly initialized ResNet34 grows sluggishly, plateauing at 78.18%. Although further improvement is possible, it would require substantial computational resources, making it less cost-effective for defect classification. Furthermore, it is worth noting that the improvement of pre-trained ResNet34 significantly diminishes after the initial surge. The accuracy has increased marginally from 93.31% to 93.79% until 5000 epochs, indicating that the model's performance remained at a stable level after the initial rapid rise and was difficult to continue to improve.

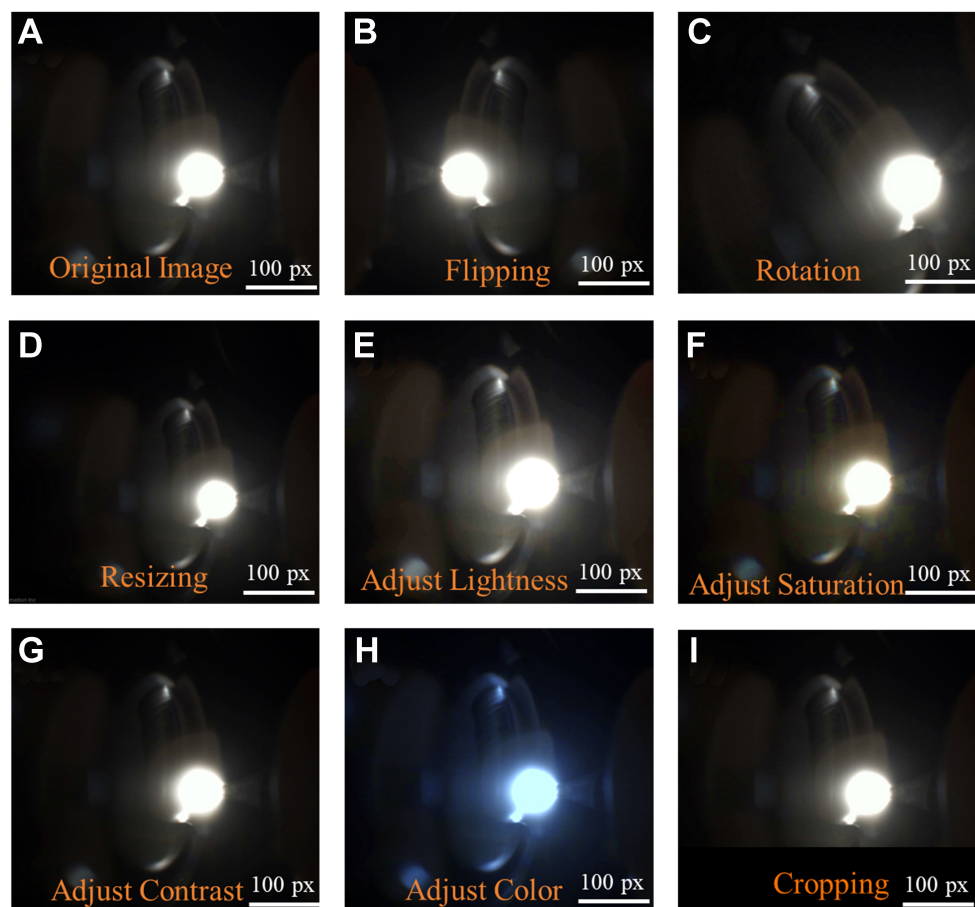
The confusion matrices provide a detailed evaluation of model performance, presenting precision and recall values for CNNs as shown in [Table 4](#). In the research, the fully fused state is defined as positive, while the unfused state is regarded as negative. Precision represents the proportion of true positives among all predicted positives; high precision ensures that samples predicted as fused are indeed fused. Recall measures the ratio of identified positive samples to all actual positive samples; high recall guarantees that no fused samples are overlooked by the classification model. In welding defect recognition, all positive samples are directly utilized in the workflow, making it essential to ensure that these samples are genuine positives. Therefore, precision takes precedence over recall. As shown in [Figure 8](#), ResNet34 achieved a precision of 96.52% and a recall of 92.5%, meeting industry requirements effectively. The defect classification model tends to classify well-fused images as non-fused, a trend that supports timely control measures during the

Table 3. Training parameter configuration of transfer learning model for welding unfused defect recognition with the strategy of loading pre-trained weights followed by performing entire fine-tuning

Training parameters	Parameters values
Image resize dimension	224 × 224
Batch	8 or 32
Learning rate	0.001
Epoch	5000
Loss function	Softmax cross entropy
Optimizer	Adam

Table 4. The accuracy precision and recall metrics of typical convolutional neural networks

Metrics	Accuracy (%)	Precision (%)	Recall (%)
Random initialization-ResNet34	78.18	81.23	80.56
500 Epochs trained-ResNet34	93.31	96.76	91.39
5000 Epochs trained-ResNet34	93.79	96.52	92.5
MobileNetV2	81.05	76.03	97.78
ConvNeXt	99.52	100	99.17

**Figure 6.** Data augmentation methods. (A) Original image; (B) flipping; (C) rotation; (D) resizing; (E-H) adjustment of lightness, saturation, contrast and color; (I) cropping.

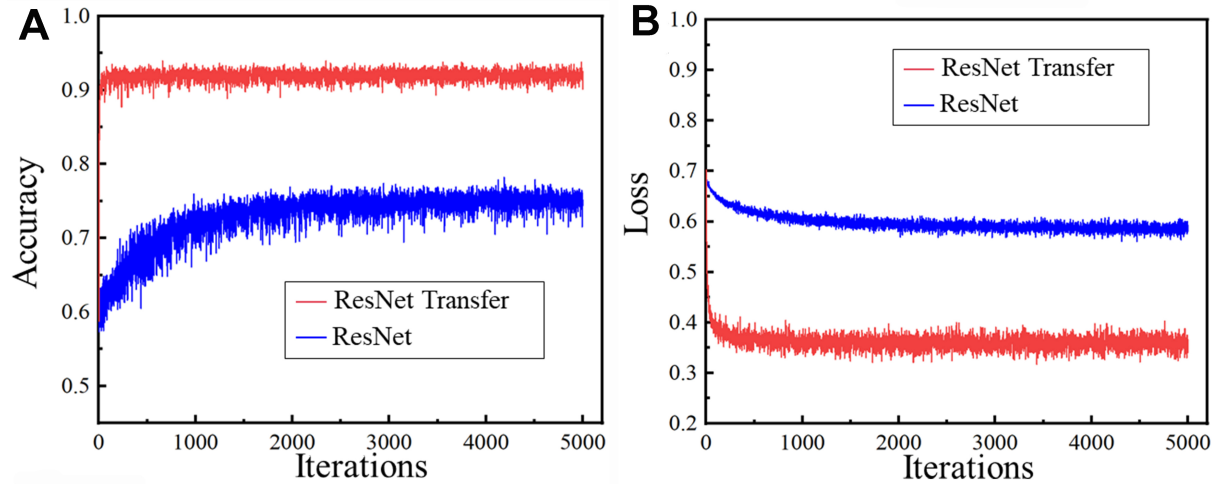


Figure 7. The performance of defect classification fluctuated over time based on the ResNet34 architecture, reflected in (A) training accuracy and (B) training loss.

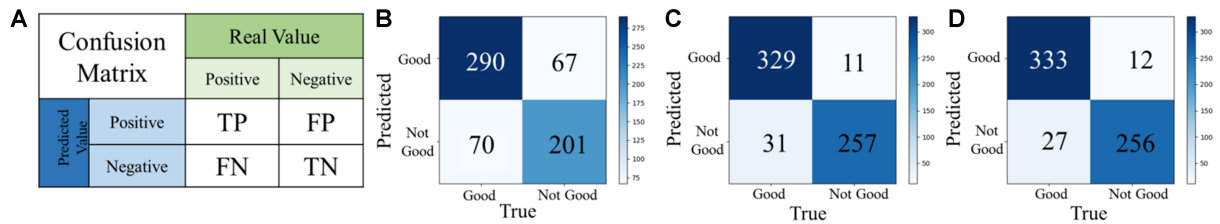


Figure 8. The confusion matrix evaluation of the ResNet34 architecture on defect classification task. (A) Confusion matrix diagram; (B) confusion matrix with randomly initialized weights; (C) loaded pre-trained weights and trained for 500 epochs; (D) loaded pre-trained weights and trained for 5000 epochs.

welding monitoring process to prevent defects. However, an excessively high defect response may hinder the efficiency of automated production.

The accuracy variation and confusion matrices for MobileNetV2 and ConvNeXt are presented in Figure 9. ConvNeXt achieved a remarkable classification accuracy of 99.52%, significantly surpassing MobileNetV2's 85.94% in defect detection. This disparity can be partly attributed to MobileNetV2 being a lightweight model that prioritizes computational efficiency, potentially sacrificing some performance. The confusion matrix indicates that MobileNetV2, with a precision of 76.03% and a recall of 97.78%, tends to classify unfused images as defect-free, posing a risk in the welding process. Based on the appraisal above, MobileNetV2 is considered unsuitable for welding defect recognition. In contrast, ConvNeXt demonstrates superior performance in welding image recognition, achieving higher accuracy than previous deep learning models^[12,21]. The welding monitoring system, incorporating the ConvNeXt model, achieves high-precision recognition of unfused defects without the necessity of human intervention. This capability holds significant potential for enhancing the quality of welding products and augmenting the productive efficiency within the automatic welding industry.

Performance optimization through t -distributed stochastic neighbor embedding

In supervised machine learning, the quality of a classifier is directly related to the quality of the data used for training. The presence of unwanted outliers in the data can significantly reduce the model's accuracy.

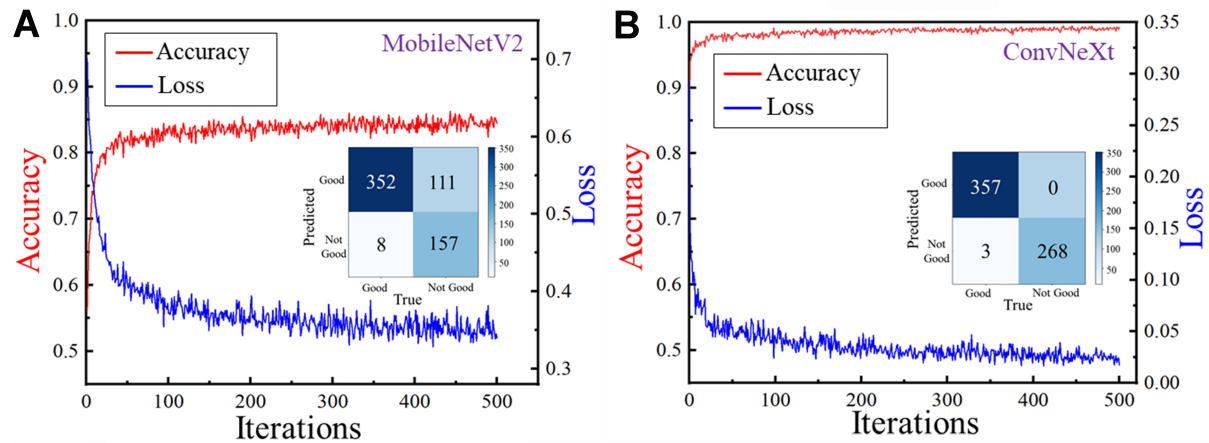


Figure 9. Model performance and confusion matrix loaded pre-trained weights and trained in the welding image dataset for 500 epochs. (A) MobileNetV2 architecture; (B) ConvNeXt architecture.

Therefore, identifying and eliminating these outliers is crucial for constructing a high-quality training dataset. The t-distributed stochastic neighbor embedding (t-SNE) is employed to model the original high-dimensional data into a low-dimensional embedding space using conditional probability distributions^[92]. The method optimizes the objective function based on the Kullback-Leibler divergence, applying gradient descent to find the most suitable embedding points in the low-dimensional space. By assigning each data point a position in a two-dimensional plot, t-SNE visualizes high-dimensional data with similar objects grouped closely together, while dissimilar objects are placed farther apart.

In the research, the dataset was randomly split into 70% for training, 20% for validation, and 10% for testing. As shown in Figure 10, the distribution of features extracted from the randomly initialized ResNet34 model does not exhibit typical clustering behavior in the t-SNE visualization. Through loading pre-trained weights, more distinct clustering was observed among samples with the same labels, suggesting that the pre-trained process can cluster features for weld seam images. However, because the general ImageNet dataset lacks weld pool images and related defect features, there is room for improvement in the model's performance on the self-constructed defect dataset. After 500 epochs of training, the output features of the trained model exhibit stronger intra-class cohesion and inter-class separation. Defect images are closely clustered in the two-dimensional space, while defect and non-defect images remain clearly separated. These visualization results indicate that transfer learning can equip the initial model with basic visual feature extraction capabilities, and the pre-trained model can effectively extract weld seam quality information from melt pool images.

Although t-SNE visualizations show the capability of CNN models to extract high-dimensional features for classification tasks, unscreened datasets may still fail to yield satisfactory results even after multiple training iterations. Therefore, t-SNE visualization is employed to classify and filter the dataset, aiming to improve model performance. In the context of pulse TIG welding, the most significant differences in melt pool images captured under identical welding parameters are in the arc features at peak and baseline moments. However, the features do not exhibit a strong correlation with the classification of unfused defects. To investigate whether differences in baseline and peak arc features affect model performance, additional experiments were conducted. The welding defect dataset was divided into categories as follows: the primary label indicating the presence of unfused defects and a sublabel indicating whether the image corresponds to the peak moment. Only primary labels were used for classification throughout training, while sub-labels

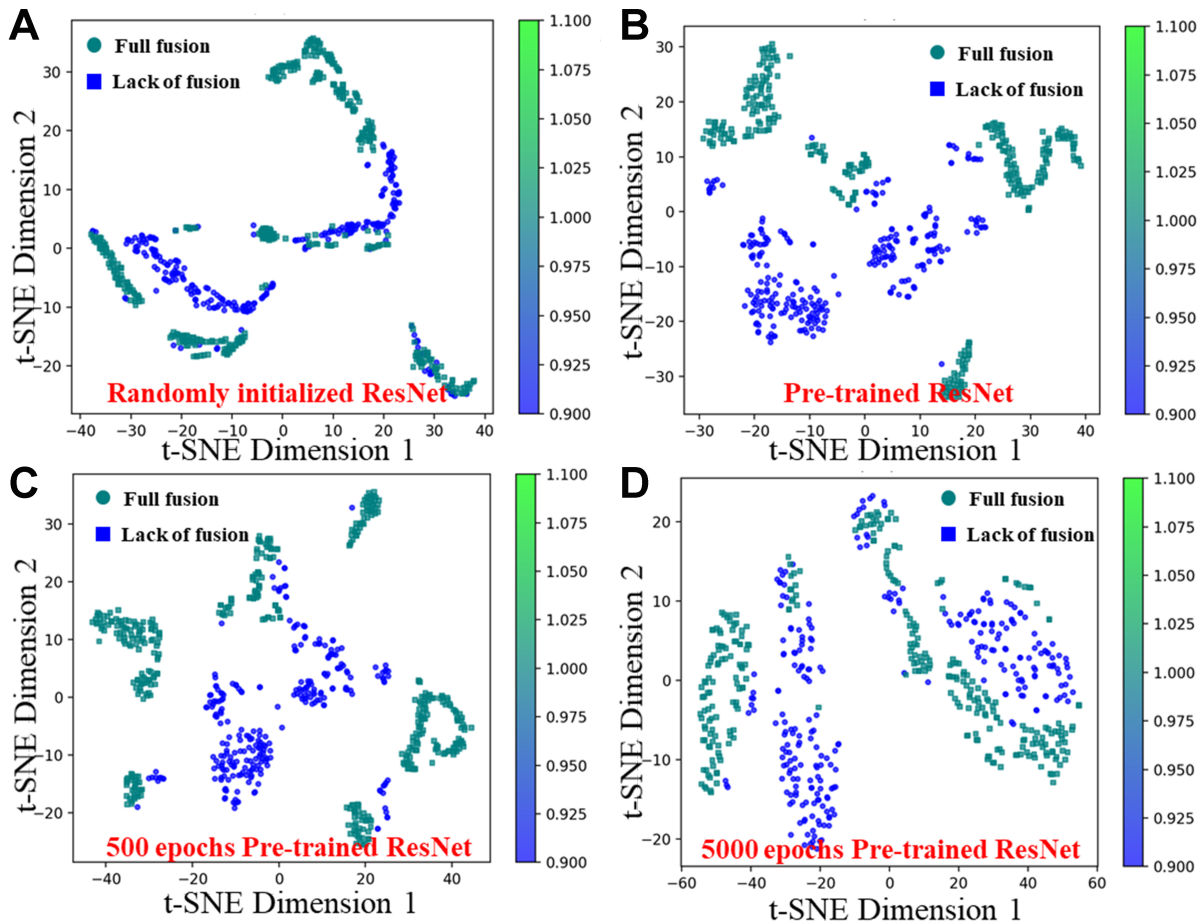


Figure 10. The t-SNE visualization of ResNet34 architecture. (A) With randomly initialized weights; (B) loaded pre-trained weights without being trained in the welding image dataset; (C) loaded pre-trained weights and trained for 500 epochs; (D) loaded pre-trained weights and trained for 5000 epochs. t-SNE: t-Distributed stochastic neighbor embedding.

were excluded from the training process. As shown in Figure 11, t-SNE visualizations reveal that even without feature alignment based on brightness during training, samples with similar brightness cluster noticeably. The phenomenon may be attributed to the simplicity of brightness as a feature, whereas unfused defects involve the senior features. The transfer learning-based architectures have already acquired the simple features during pre-training which interfered with defect classification. Accordingly, the dataset was refined by excluding all images associated with baseline currents, followed by retraining the models. Figure 12 illustrates a clearer distinction between defect and non-defect images in the t-SNE visualizations, demonstrating improved performance of defect recognition after separating peak moment images from baseline moment images. Therefore, optimizing data samples after the model's performance plateaus can further enhance the accuracy of deep learning models.

Weld seam forming prediction on back propagation neural network

The performance of weldments in operational conditions is significantly correlated with the morphology of the weld seams^[93]. The crucial function of intelligent welding systems is to establish a mapping relationship between welding parameters and weld appearance, thereby optimizing the welding conditions to produce superior weldments. This section proposes a method for predicting the thickness of weld seams with BPNNs. Additionally, by evaluating the correlation between arc length and welding voltage, the geometric

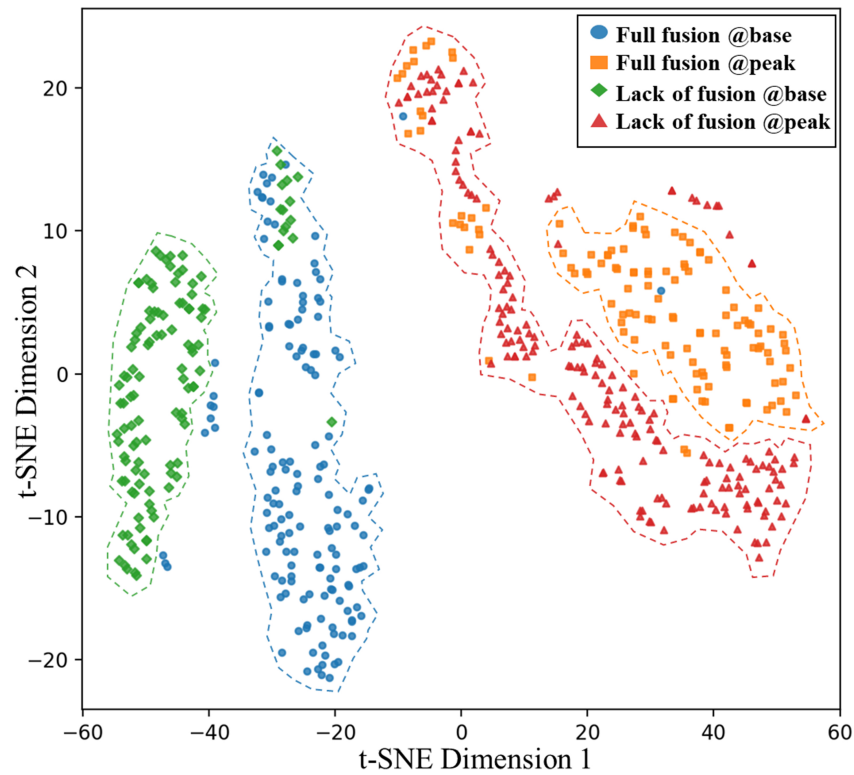


Figure 11. The t-SNE visualization on the dataset with primary labels of unfused defects and sublabels of lightness. t-SNE: t-Distributed stochastic neighbor embedding.

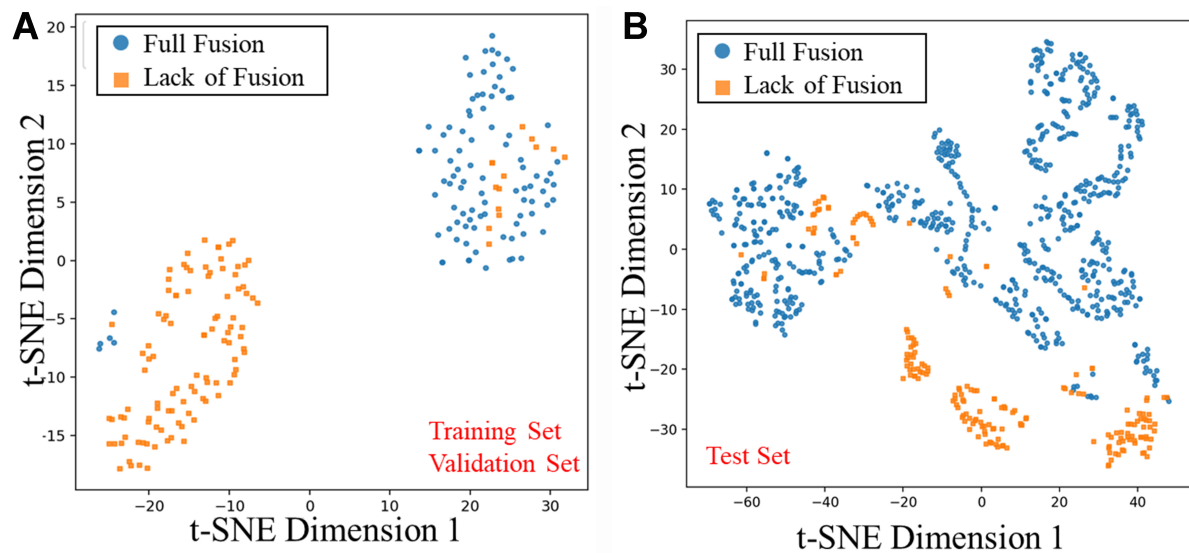


Figure 12. The t-SNE visualization on pruned dataset. (A) performance on training set and validation set; (B) performance on test set. t-SNE: t-Distributed stochastic neighbor embedding.

features of the weld are incorporated into the model to enhance prediction accuracy. Furthermore, various advanced image processing techniques are employed to efficiently identify and extract the geometric features of the welding arc. The [Supplementary Materials](#) provide access to the programs utilized for

welding contour extraction and BPNN training in the research.

Weld seam thickness prediction based on technologic parameters

In the context of predicting weld formation for tube-to-tube-sheet welding, the introduction of various interfering factors diminishes the uniformity of the weld formation process, thereby complicating the prediction of weld seam thickness^[94,95]. Traditional predictive models primarily depend on welding parameters, including welding current, V_s , welding period and δ which directly influence the welding outcome^[96]. The BPNNs are initially established in the research that relies exclusively on the welding technical parameters to evaluate the predictive efficacy and identify potential avenues for enhancement. To mitigate the adverse effects of incorrect labels and bolster the accuracy of BPNNs, the strategy for managing outliers is implemented. The data points exhibiting clear errors, primarily due to misalignment of the welding machine's core axis, were excluded. As illustrated in [Figure 13](#), although the predictions of weld seam thickness show marked improvement following data screening, considerable deviations between predicted and actual values persist. Consequently, the intricate interplay of multiple factors influencing weld formation makes achieving an optimal weld appearance extremely challenging through parameter adjustments alone. A deeper exploration of arc characteristics is essential to improve the predictive accuracy of BPNNs.

Correlation analysis between arc length and welding voltage

The arc contains crucial information about the welding process and can be used to predict welding performance^[97]. The arc shape directly reflects variations in welding process parameters and stability, closely related to the size and stress conditions of the molten pool which significantly affect weld quality. Therefore, leveraging arc geometry to predict welding performance is a practical approach. To evaluate the feasibility of the method, the relationship between arc length and key welding parameters specifically welding voltage is examined in the research. The 28th group is selected for the validation due to the slow V_s , which facilitates the collection of a larger volume of image data over the welding trajectory. Additionally, the high pulse duty cycle and prominent arc characteristics further enhance the data quality. The initial cycle of the welding torch travel takes 43.6 s, during which 1,244 molten pool images were captured. Through the thresholding technique, only the images recorded during the peak current phase are retained. Arc length variations over time are then extracted in batches with Python. As shown in [Figure 14](#), the fluctuations in arc length closely align with the monitored voltage values, confirming the correlation between arc length and welding voltage. The outcome demonstrates that the arc shape encapsulates valuable process information and can be applied to predict welding performance. The subsequent step is to extract the arc's geometric features from the welding images for further analysis.

Image processing for arc contour extraction

Efficient extraction of the arc contour requires specialized preprocessing techniques for the welding images, including three-dimensional grayscale distributions, region of interest (ROI) extraction and image enhancement. Firstly, the machine vision system processes images based on the distribution and gradient of the grayscale values. [Figure 15](#) presents the HDR images of pulsed TIG welding at both peak and base currents, along with their three-dimensional grayscale distributions and the grayscale variation along the direction of maximum arc length. The grayscale values approach saturation in the region illuminated by the arc, while the values drop sharply outside the arc area indicating a steep gradient between the arc and surrounding areas. The pronounced contrast allows for manageable extraction of the apparent arc contour by applying an appropriate grayscale threshold.

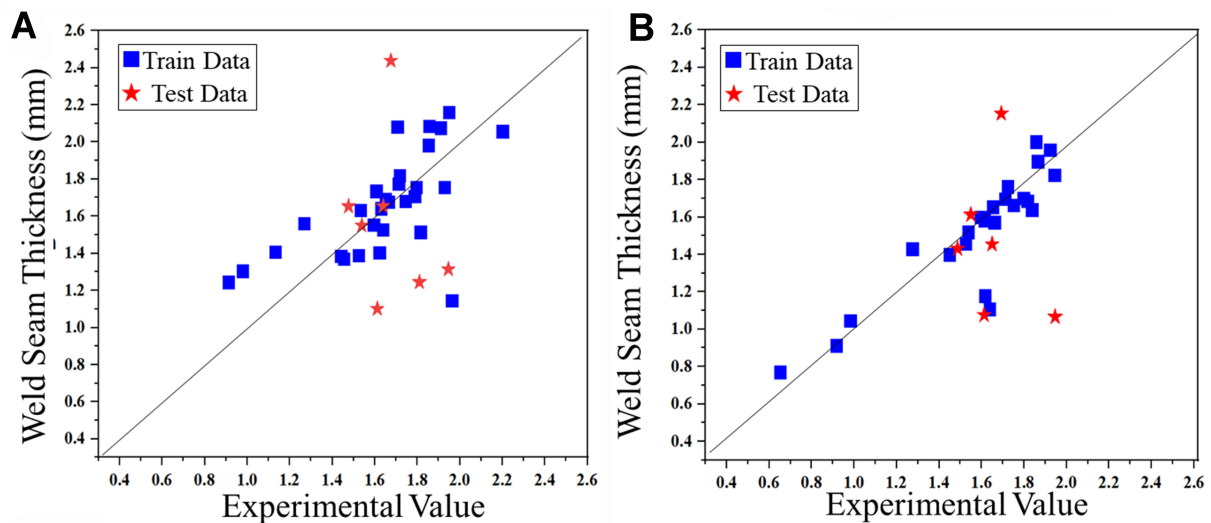


Figure 13. The weld seam thickness prediction only based on welding parameters. (A) Raw data; (B) data after screening to exclude obvious offset.

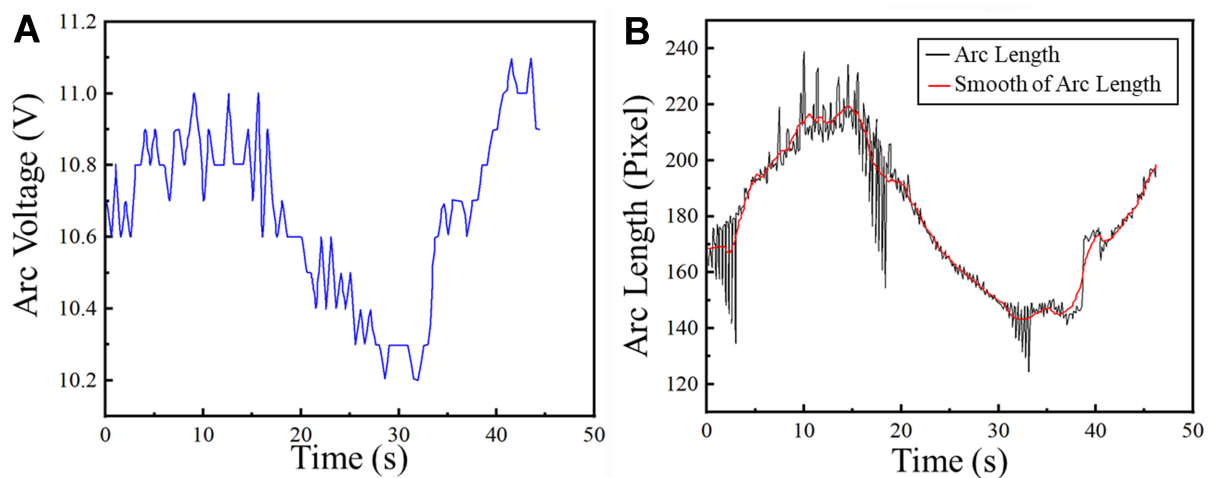


Figure 14. The correlation between voltage and arc length. (A) Welding voltage changes over time; (B) arc length varies with time.

Secondly, the original molten pool image size obtained by the information acquisition and processing system is $1,280 \times 1,024$. The exceptionally high local temperature of the molten pool generates intense thermal radiation, leading to a predominantly black background in the molten pool images. Nevertheless, the background area contains minimal information relevant to welding quality. To reduce redundancy and improve image processing efficiency, the image is cropped to extract the ROI corresponding to the molten pool. As depicted in Figure 16, considering possible camera motion during the process, the cropping window size is set to 500×500 , ensuring that the image captures only the molten pool region.

Thirdly, to further improve the visual effect of the arc image and achieve arc feature extraction, image enhancement technology is introduced. The original welding arc image, as shown in Figure 17A, lacks sufficient contrast for effective arc contour extraction. As shown in Figure 17B, adaptive histogram equalization is an improvement on histogram equalization, which not only effectively enhances the local contrast of the image but also avoids excessive contrast enhancement. In the case of different lighting

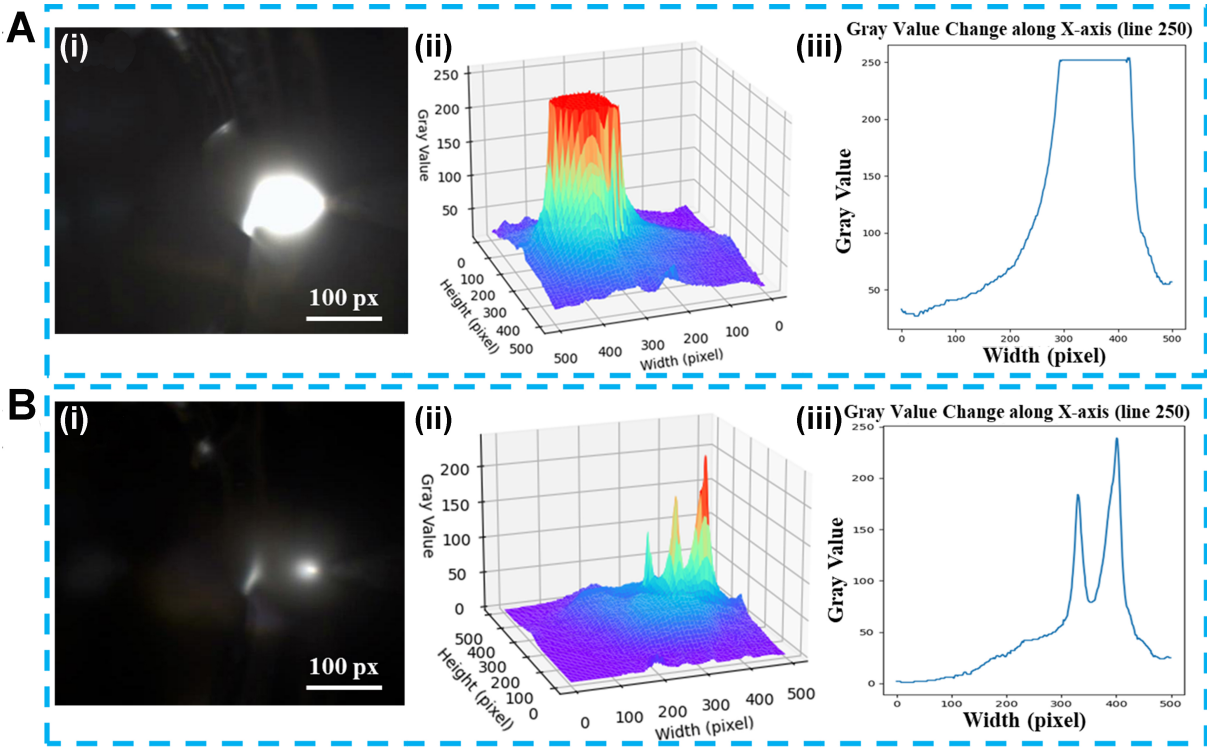


Figure 15. The grayscale distribution of pulsed TIG welding images. (A) At peak current; (B) at base current (i-iii) HDR images, three-dimensional grayscale distribution, grayscale variation along the direction of maximum arc length. TIG: Tungsten inert gas; HDR: High dynamic range.

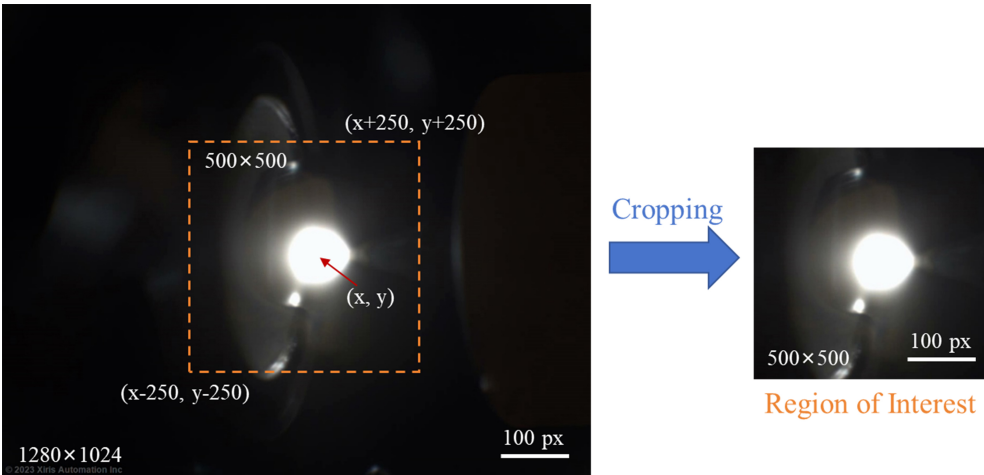


Figure 16. Extraction of ROI from melting pool images. ROI: Region of interest.

conditions in the image, the technique can retain more detailed information. As shown in [Figure 17C](#), gamma correction adjusts the brightness of the image by applying a power law transformation, and the degree of adjustment is determined by the gamma (γ) value. A γ value less than 1 will increase the contrast of the darker areas of the image, while a value greater than 1 will increase the contrast of the brighter areas. Applying gamma correction can adjust the overall contrast of the image. The texture information of the weld seam outside the molten pool in the original image is weak. Adjusting the Gamma value to 2.5 can

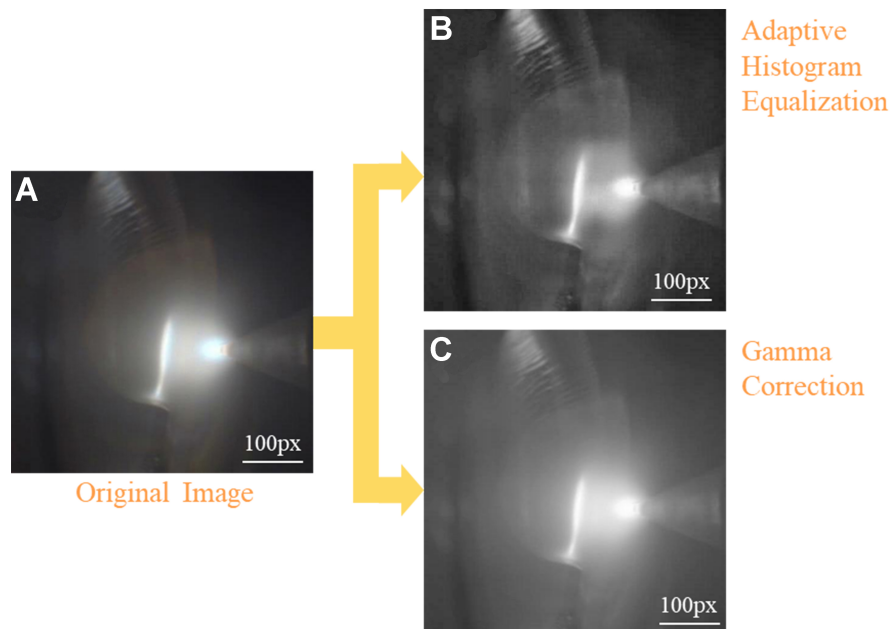


Figure 17. Image enhancement at base current moment. (A) Original image; (B) adaptive histogram equalization; (C) gamma correction.

enhance the morphological details of the “fish scale pattern” of the weld seam.

Weld seam thickness prediction incorporating arc geometric features

Utilizing straightforward programming derived from the OpenCV library enables the efficient extraction of two-dimensional morphological information regarding the arc in the molten pool image. As illustrated in [Figure 18](#), the specific steps to extract arc features are as follows. Firstly, the molten pool image is converted into a grayscale image, and the arc area is segmented from the image by setting the appropriate grayscale threshold. Secondly, canny edge detection algorithms are implemented to extract the arc length and arc width, which estimate the arc length and width by counting pixels along the arc contour. Finally, the arc area size is calculated according to the number of non-zero pixels in the binarized image, which corresponds to the pixels within the arc region. Meanwhile, significant changes in the shape and size of the molten pool only exist during the initial welding preheating stage and the arc closing insulation stage. The shape and size of the molten pool are relatively stable during the welding process with periodic fluctuations within a certain range. Therefore, ten molten pool images taken two seconds after the arc initiates are selected from each set of welding processes, and the average values are used as the representative input variables for the molten pool characteristics in that process. As shown in [Figure 19](#), the prediction values and prediction errors for the corner weld thickness of the BPNN added arc geometric features and the models only depending on process parameters for prediction are compared. After employing the arc features as input parameters, the maximum prediction error of the model for the corner weld thickness has decreased from the previous 0.8 to 0.6 mm. In 28 groups of experimental parameters, only four groups have an error exceeding 0.2 mm. Accordingly, the BPNN integrated with arc feature parameters has higher prediction accuracy and smaller fluctuations in prediction errors than the original process parameter BPNN.

Automatically programming based on LLMs

The AI agent is an intelligent entity driven by a LLM core, capable of perceiving the environment, making

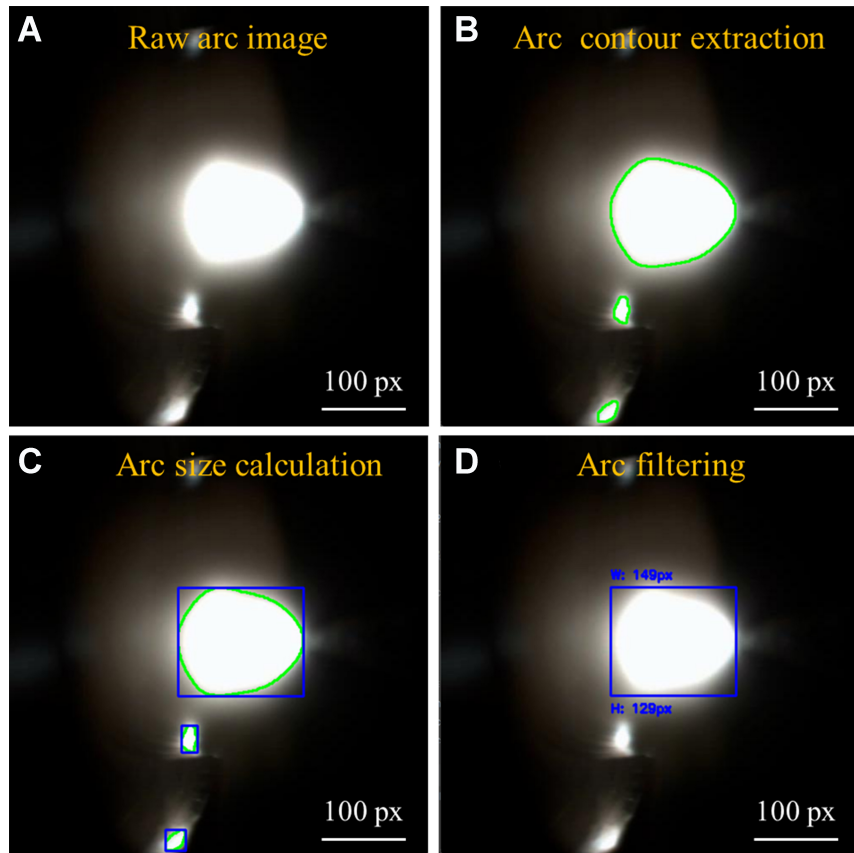


Figure 18. Extraction method of melting pool features. (A) Raw arc image; (B) arc light contour extraction; (C) arc size calculation; (D) arc filtering.

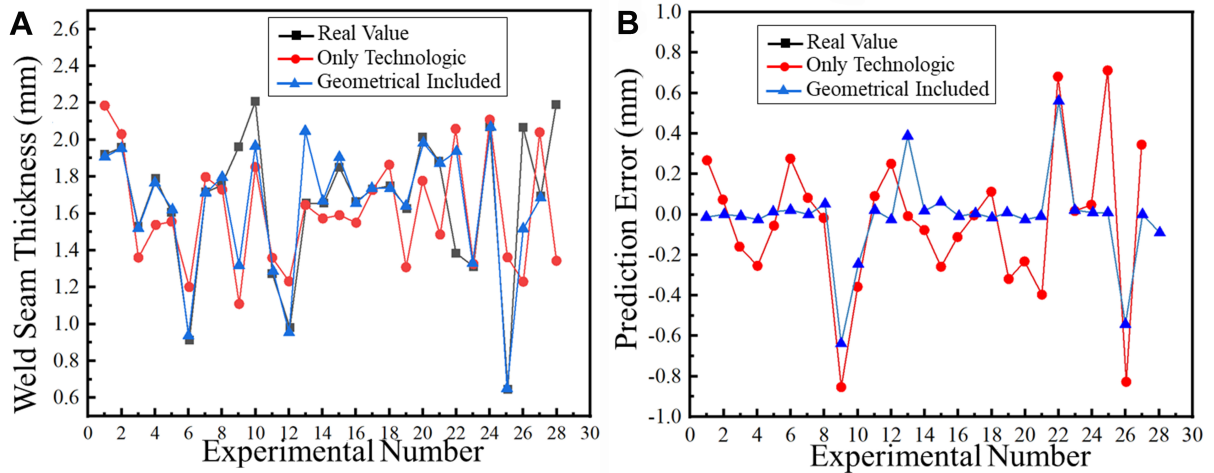


Figure 19. The comparison in the pulse TIG weld seam thickness prediction chart between the performance of the BPNN considering visual parameters and the BPNN relying solely on process parameters. (A) weld seam thickness value; (B) prediction error. TIG: Tungsten inert gas; BPNN: Backpropagation neural network.

decisions, and executing actions. The automatic programming module of the AI agent is an important application scenario of AI agent technology. In this section, the exploration of the programming generation

ability of AI agent and LLMs is conducted. The deep learning architecture for welding image defect recognition is generated and optimized by LLMs to explore the workflow to program assisted by artificial intelligence, which can be the pioneer for automatic programming in the welding automation field.

Welding defect recognition through automated programming technology

The procedure for using AI agents for automated programming in welding defect recognition tasks can follow the steps below. The first and most crucial step is to clarify the requirements that the programming is expected to fulfill. The automatic programming process is conducted by inputting the requirement into the LLMs and obtaining the outputting programs generated by LLMs. To make sure the LLMs can understand the requirement completely, clear steps for the programs are required for the input content to LLMs. In the input content, the program functions should be involved, including dataset loading, image preprocessing, dataset splitting, pre-trained weights loading, and training loss storage. Additionally, the development environment, including details such as the programming language (Python) and frameworks (PyTorch), should be clearly defined. It is worth noting that the outputs and user interfaces of four LLMs including ChatGPT-3.5, Claude, Microsoft Copilot, and Baidu ERNIE Bot were compared only in the simple task of extracting arc contours. For programming deep learning models aimed at welding defect recognition, ChatGPT was exclusively utilized, as it demonstrated the best ability to understand input requirements and generate highly applicable code.

When requirements are input into ChatGPT, the programs are generated automatically. However, automated programming cannot be accomplished in a single attempt and requires multiple interactions between humans and artificial intelligence. Occasionally, the code generated by ChatGPT may not be runnable in a local integrated development environment, resulting in compiler errors. These error messages can be fed back into ChatGPT, which will then generate solutions. There may also be new requirements for the programming. Developers can input the current programming along with the new requirements into the ChatGPT, which will then generate new programming with the added functionality automatically. Through the methodologies above, the programming can be finalized to ensure stable execution and successful fulfillment of the specified tasks. The ability of ChatGPT to iteratively refine outputs based on repeated inputs was most valued. Expecting a single attempt by ChatGPT to fully comprehend all requirements and generate flawless code was not the focus of this approach. Compared to human-written programs, those generated by ChatGPT tend to be more standardized, comprehensible and characterized by clear logic and comments.

Welding image stitching through automated programming technology

Another application of automated programming in welding defect recognition is the development of image stitching technology. The input for ChatGPT involves generating an executable program in a Python environment to stitch high-resolution welding tissue images. The program is automatically generated based on the following function modules: preprocessing, feature detection, feature matching, image fusion, and post-processing. The initial version of the program may not perfectly meet the stitching requirements. Adjustments can be made according to the specific function modules. For instance, various methods can be applied to the feature detection function, including scale-invariant feature transform (SIFT), speeded-up robust features (SURF), and oriented FAST and rotated BRIEF (ORB). In this research, LLMs are employed to generate programs for feature detection based on these methods, with each method tested within the full program to compare their performance in the image stitching task. After comparison, the best feature detector is selected, employing the SIFT method to handle the high-resolution characteristics inherent in welding images.

According to the automatic programming method, the complete programs for image splicing can be generated on the workflow as follows: Initially, the images undergo preprocessing which entails noise reduction, grayscaling and rescaling. Subsequently, the feature detector is established using the SIFT method to accommodate the high-resolution characteristics inherent in welding images. The feature detection phase extracts distinctive features from each image, resulting in the generation of descriptors. The feature-matching algorithms include brute-force matching, nearest neighbor matching and K -dimensional tree (KD-tree) matching. The KD-tree matching is utilized in the research to achieve an optimal balance between model efficiency and accuracy. The feature-matching algorithm discerns corresponding points between images based on the descriptors, thereby constructing a robust set of paired feature points. Furthermore, the random sample consensus (RANSAC) is employed to compute the homography matrix, which reflects the correspondence among feature points. Image fusion techniques including weighted averaging, Poisson blending and Laplacian pyramid blending facilitate seamless transitions in overlapping regions. The suitably sized blank canvas is prepared onto which the stitched images are rendered. Finally, post-processing enhances the quality of the stitched images through the cropping of extraneous areas and the adjustment of brightness to improve detail visibility. The image splicing programs applied in this research are available in the [Supplementary Materials](#). The technology abandons traditional manual marking methods and adopts the unsupervised stitching approach, enabling precise stitching of weld joint images through automated detection and matching of feature points^[98]. The technology exhibits remarkable adaptability, ensuring consistent performance across diverse scenarios, lighting conditions and sensor configurations. Meanwhile, by combining the clarity of high-magnification microscopy with the broad field of view offered by low-magnification microscopy, image stitching technology facilitates the acquisition of clear and complete microstructural images of weld joints.

The schematic illustration of the longitudinal cross-section of tube plate welding is shown in [Figure 20A](#) to depict the micrograph position of the weld seam vividly. Through the image-stitching process, lucid and expansive images of welding microstructures are obtained, facilitating the effortless observation of microstructural characteristics and the precise identification of welding defects. The majority of weld joint morphologies in the welding test samples exhibit a defect-free appearance, as depicted in [Figure 20B](#), indicating stable control of welding parameters and favorable welding conditions. As illustrated in [Figure 20C](#), the stitched image reveals porosity defects in the weld, which may be attributed to wet electrodes, moisture, oil, rust on the weldment or excessive V_s and currents. [Figure 21](#) showcases the morphology of the welding sample with crack defects in grayscale mode, which are caused by stresses generated during the cooling process. Notably, the center of the stitched image exhibits a black-striped area due to the absence of overlap between images captured by the camera, indicating that not all areas of the sample surface were captured.

CONCLUSIONS

In the present work, an in-depth investigation is conducted into the welding formation and microstructural characteristics of titanium alloys. First of all, the overall framework and specific structural design of the CNN-based defect detection model are elucidated. Image enhancement techniques are applied to augment the weld pool image dataset, and transfer learning methods are adopted to enhance model training effectiveness. The performance of ResNet34, MobileNetV2, and ConvNeXt on welding defect datasets is evaluated. The classification accuracy of the pre-trained ConvNeXt model, utilizing transfer learning, attains an impressive 99.52% following 500 training iterations. In contrast, the accuracies of MobileNetV2 and ResNet34 stand at 85.94% and 93.41%, respectively. Additionally, through the visualization technique of t-SNE, the operational dynamics of the deep learning model in defect detection tasks are thoroughly examined. Utilizing a systematic layer-by-layer feature extraction process, the model recalibrates the

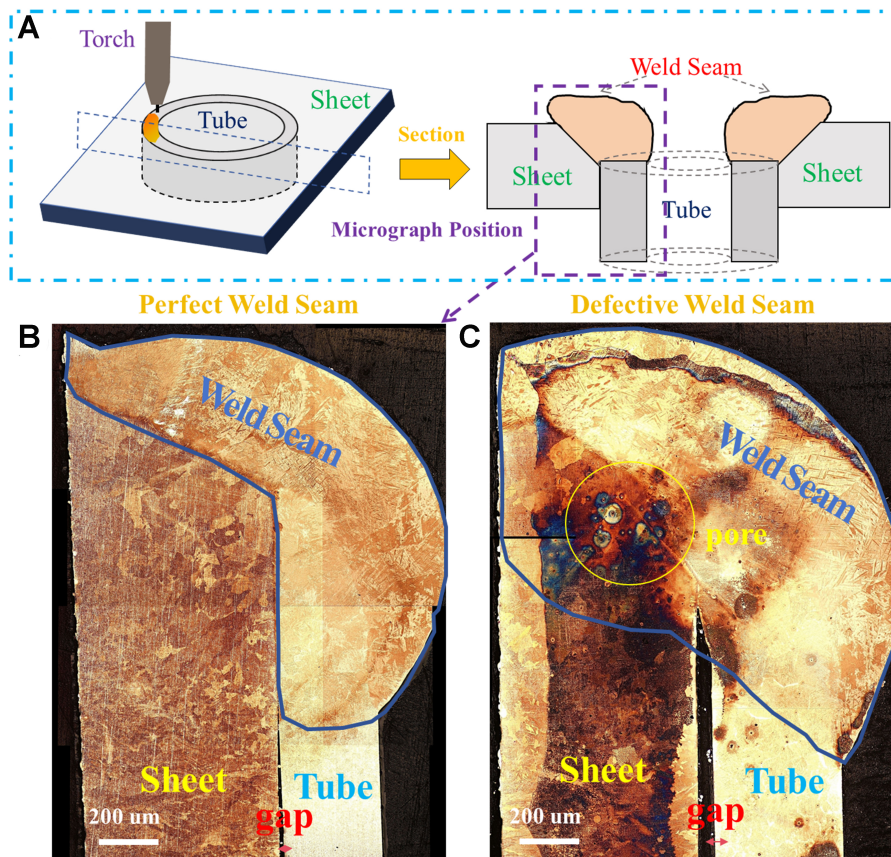


Figure 20. Schematic illustration of micrograph position of weld seam in the longitudinal cross-section of tube-to-tube-sheet welding and spliced images. (A) Micrograph position illustration; (B) spliced image of perfect weld seam; (C) spliced image of defective weld seam.

high-dimensional feature distribution of the data, thereby enhancing the efficacy of defect detection. In addition, a three-layer BPNN model is developed, incorporating both welding process parameters and weld feature quantities. The extracted arc length, arc width, and arc area are used as visual features for the dynamic prediction of weld seam thickness, improving model performance compared to models based solely on process parameters. The maximum prediction error for weld seam thickness decreases from 0.8 to 0.6 mm, and out of 28 sets of experimental parameters, only four sets result in an error exceeding 0.2 mm. Furthermore, an automated programming technique based on LLMs is developed to program deep learning models for welding defect recognition. Various LLMs including ChatGPT 3.5, Bing Copilot, Claude3 and ERNIE Bot are tested for their application in automated programming. Through the technology above, the program for image stitching is programmed, enabling unsupervised automatic stitching of multiple welding microstructure images. The technique results in clear and wide-field weld images, providing robust support for subsequent image recognition processes.

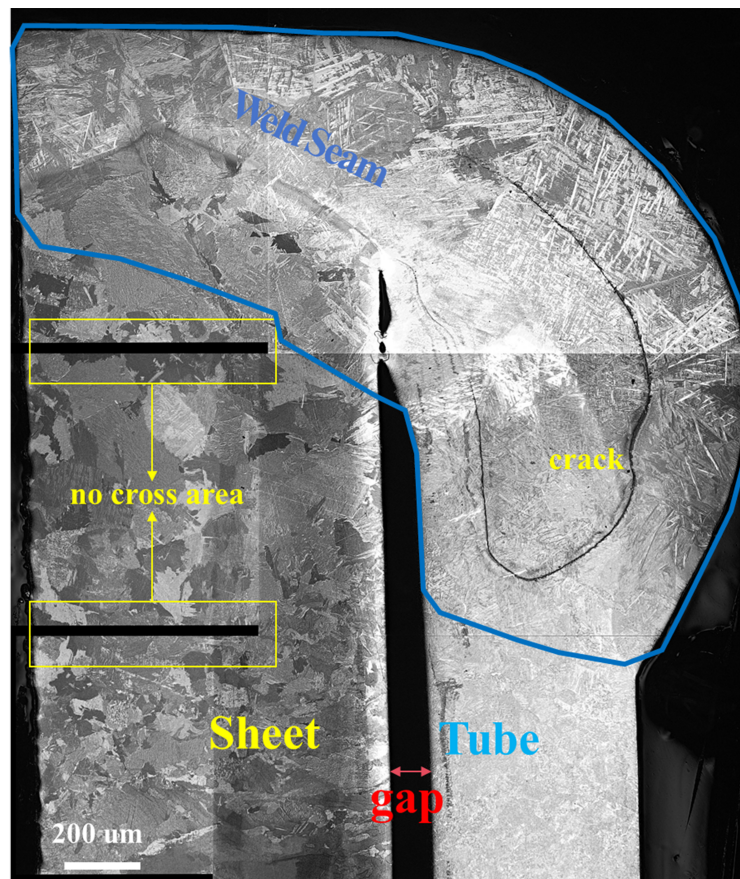


Figure 21. Spliced image of weld joint with crack defects in grayscale mode; meanwhile, the presence of black region in the image is due to the absence of cross area in the original photographs.

DECLARATIONS

Acknowledgments

This research was supported by the National Basic Scientific Research Project of China (Grant No. JCKY2020607B003). The authors gratefully acknowledge this funding, which was instrumental in the study's design, data collection, and analysis.

Authors' contributions

Writing - original draft preparation, conceptualization, methodology, data curation, investigation, formal analysis: Zhang S

Writing - original draft preparation, supervision, methodology, editing, validation, project administration, funding acquisition: Wang WY

Data curation, investigation, formal analysis, writing - original draft preparation: Wang X, Li G, Ren Y

Conceptualization, methodology, editing, methodology, project administration: Gao X

Conceptualization, supervision, methodology, editing, validation, project administration: Sun F, Tang B

Supervision, conceptualization, methodology, editing, methodology, project administration: Song H

Supervision, conceptualization, methodology, editing, methodology, project administration, funding acquisition: Li J

All authors have read and agreed to the published version of the manuscript.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Financial support and sponsorship

This research was supported by the National Basic Scientific Research Project of China (No. JCKY2020607B003).

Conflicts of interest

Wang WY is an Editor in the Junior Editorial Board of *Journal of Materials Informatics*. Wang WY was not involved in any steps of editorial processing, notably including the selection of reviewers, manuscript handling and decision-making. Li G, Ren Y and Sun F are affiliated with Western Superconducting Technologies Co., Ltd. The other authors declare that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2024.

REFERENCES

1. An Y, Xu X, Zhao Y, Hou H. Nonequilibrium solidification velocity, recalescence degree and grain refinement of highly undercooled Ni-based single-phase alloys. *J Alloys Compd* 2021;881:160658. DOI
2. Cheng Y, Wang G, Qiu Z, et al. Multi-physics simulation of non-equilibrium solidification in Ti-Nb alloy during selective laser melting. *Acta Mater* 2024;272:119923. DOI
3. Peng Z, Zhang X, Liu L, Xu G, Wang G, Zhao M. Effect of high-speed ultrasonic vibration cutting on the microstructure, surface integrity, and wear behavior of titanium alloy. *J Mater Res Technol* 2023;24:3870-88. DOI
4. Williams W. Development of structural titanium alloys for marine applications. *Ocean Eng* 1969;1:375-83. DOI
5. Auwal ST, Ramesh S, Yusof F, Manladan SM. A review on laser beam welding of titanium alloys. *Int J Adv Manuf Technol* 2018;97:1071-98. DOI
6. Su Y, Liang C, Wang D. Composition- and temperature-dependence of β to ω phase transformation in Ti-Nb alloys. *J Mater Inf* 2023;3:14. DOI
7. Li P, Zhang Y, Wang WY, et al. Coupling effects of high magnetic field and annealing on the microstructure evolution and mechanical properties of additive manufactured Ti-6Al-4V. *Mater Sci Eng A* 2021;824:141815. DOI
8. Park H, Rhee S. Estimation of weld bead size in CO₂ laser welding by using multiple regression and neural network. *J Laser Appl* 1999;11:143-50. DOI
9. Wang WY, Yin J, Chai Z, et al. Big data-assisted digital twins for the smart design and manufacturing of advanced materials: from atoms to products. *J Mater Inf* 2022. DOI
10. Vasan V, Sridharan NV, Balasundaram RJ, Vaithyanathan S. Ensemble-based deep learning model for welding defect detection and classification. *Eng Appl Artif Intell* 2024;136:108961. DOI
11. Wang B, Hu SJ, Sun L, Freiheit T. Intelligent welding system technologies: state-of-the-art review and perspectives. *J Manuf Syst* 2020;56:373-91. DOI
12. Feng Y, Chen Z, Wang D, Chen J, Feng Z. DeepWelding: A deep learning enhanced approach to GTAW using multisource sensing images. *IEEE Trans Ind Inf* 2020;16:465-74. DOI
13. Wang S, Zhang S, Wen S, Fernandez C. An accurate state-of-charge estimation of lithium-ion batteries based on improved particle swarm optimization-adaptive square root cubature kalman filter. *J Power Sources* 2024;624:235594. DOI
14. Ma D, Shu L, Zhou Q, Cao S, Jiang P. Online porosity defect detection based on convolutional neural network for Al alloy laser welding. *J Phys Conf Ser* 2021;1884:012008. DOI
15. Zhang Y, You D, Gao X, Katayama S. Online monitoring of welding status based on a DBN model during laser welding. *Engineering* 2019;5:671-8. DOI
16. Huang J, Zhang Z, Qin R, et al. Interpretable real-time monitoring of pipeline weld crack leakage based on wavelet multi-kernel

- network. *J Manuf Syst* 2024;72:93-103. DOI
17. Cheng Y, Yu R, Zhou Q, Chen H, Yuan W, Zhang Y. Real-time sensing of gas metal arc welding process - a literature review and analysis. *J Manuf Process* 2021;70:452-69. DOI
 18. Hossain R, Lewis J, Moore AL. In situ infrared temperature sensing for real-time defect detection in additive manufacturing. *Addit Manuf* 2021;47:102328. DOI
 19. Chen C, Lv N, Chen S. Welding penetration monitoring for pulsed GTAW using visual sensor based on AAM and random forests. *J Manuf Process* 2021;63:152-62. DOI
 20. Liu T, Wang J, Huang X, Lu Y, Bao J. 3DSMDA-Net: an improved 3DCNN with separable structure and multi-dimensional attention for welding status recognition. *J Manuf Syst* 2022;62:811-22. DOI
 21. Bacioiu D, Melton G, Papaalias M, Shaw R. Automated defect classification of aluminium 5083 TIG welding using HDR camera and neural networks. *J Manuf Process* 2019;45:603-13. DOI
 22. Yang J, Li S, Wang Z, Dong H, Wang J, Tang S. Using deep learning to detect defects in manufacturing: a comprehensive survey and current challenges. *Materials* 2020;13:5755. DOI PubMed PMC
 23. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: transformers for image recognition at scale. *arXiv* 2021, arXiv:2010.11929. Available online: <https://doi.org/10.48550/arXiv.2010.11929> (accessed 26 Dec 2024)
 24. Springenberg M, Frommholz A, Wenzel M, Weicken E, Ma J, Strodtz N. From modern CNNs to vision transformers: assessing the performance, robustness, and classification strategies of deep learning models in histopathology. *Med Image Anal* 2023;87:102809. DOI PubMed
 25. Mehta S, Rastegari M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer. *arXiv* 2021, arXiv:2110.02178. Available online: <https://doi.org/10.48550/arXiv.2110.02178> (accessed 26 Dec 2024)
 26. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A convnet for the 2020s. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2022 Jun 18-24; New Orleans, USA. IEEE; 2022. pp. 11976-86. DOI
 27. Hou Q, Lu CZ, Cheng MM, Feng J. Conv2Former: a simple transformer-style ConvNet for visual recognition. *IEEE Trans Pattern Anal Mach Intell* 2024;46:8274-83. DOI PubMed
 28. Lin M, Wu J, Meng J, Wang W, Wu J. Screening of retired batteries with gramian angular difference fields and ConvNeXt. *Eng Appl Artif Intell* 2023;123:106397. DOI
 29. Lei Z, Shen J, Wang Q, Chen Y. Real-time weld geometry prediction based on multi-information using neural network optimized by PCA and GA during thin-plate laser welding. *J Manuf Process* 2019;43:207-17. DOI
 30. Moon H, Na S. A neuro-fuzzy approach to select welding conditions for welding quality improvement in horizontal fillet welding. *J Manuf Syst* 1996;15:392-403. DOI
 31. Ai Y, Shao X, Jiang P, Li P, Liu Y, Yue C. Process modeling and parameter optimization using radial basis function neural network and genetic algorithm for laser welding of dissimilar materials. *Appl Phys A* 2015;121:555-69. DOI
 32. Yu R, Huang Y, Peng Y, Wang K. Monitoring of butt weld penetration based on infrared sensing and improved histograms of oriented gradients. *J Mater Res Technol* 2023;22:3280-93. DOI
 33. Yang L, Liu Y, Peng J, Liang Z. A novel system for off-line 3D seam extraction and path planning based on point cloud segmentation for arc welding robot. *Robot Cim Int Manuf* 2020;64:101929. DOI
 34. Zhang K, Yan M, Huang T, Zheng J, Li Z. 3D reconstruction of complex spatial weld seam for autonomous welding by laser structured light scanning. *J Manuf Process* 2019;39:200-7. DOI
 35. Liu T, Zheng P, Bao J. Deep learning-based welding image recognition: a comprehensive review. *J Manuf Syst* 2023;68:601-25. DOI
 36. Sahu PK, Pal S. Multi-response optimization of process parameters in friction stir welded AM20 magnesium alloy by Taguchi grey relational analysis. *J Magnes Alloy* 2015;3:36-46. DOI
 37. Kulal S, Pasupat P, Chandra K, et al. SPoC: search-based pseudocode to code. *arXiv* 2019, arXiv:1906.04908. Available online: <https://doi.org/10.48550/arXiv.1906.04908> (accessed 26 Dec 2024)
 38. Fedorenko E, Ivanova A, Dhamala R, Bers MU. The language of programming: a cognitive perspective. *Trends Cogn Sci* 2019;23:525-8. DOI PubMed
 39. Bobrow DG, Stefik MJ. Perspectives on artificial intelligence programming. *Science* 1986;231:951-7. DOI PubMed
 40. Ma J, Cao B, Dong S, et al. MLMD: a programming-free AI platform to predict and design materials. *npj Comput Mater* 2024;10:1243. DOI
 41. Nadeem M, Sohail SS, Javed L, Anwer F, Saudagar AKJ, Muhammad K. Vision-enabled large language and deep learning models for image-based emotion recognition. *Cogn Comput* 2024;16:2566-79. DOI
 42. Pei Z, Yin J, Neugebauer J, Jain A. Towards the holistic design of alloys with large language models. *Nat Rev Mater* 2024;9:840-1. DOI
 43. Mouret JB. Large language models help computer programs to evolve. *Nature* 2024;625:452-3. DOI PubMed
 44. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems; Long Beach, USA. Curran Associates Inc.; 2017, pp. 6000-10. DOI
 45. Wong MF, Guo S, Hang CN, Ho SW, Tan CW. Natural language generation and understanding of big code for AI-assisted programming: a review. *Entropy* 2023;25:888. DOI PubMed PMC
 46. Chiarello F, Giordano V, Spada I, Barandoni S, Fantoni G. Future applications of generative large language models: a data-driven case study on ChatGPT. *Technovation* 2024;133:103002. DOI

47. Fernandes LC. Programming computational electromagnetic applications assisted by large language models [Em Programmer's Notebook]. *IEEE Antennas Propag Mag* 2024;66:63-71. DOI
48. Peng D, Zheng L, Liu D, et al. Large-language models facilitate discovery of the molecular signatures regulating sleep and activity. *Nat Commun* 2024;15:3685. DOI PubMed PMC
49. Karnalim O, Toba H, Johan MC. Detecting AI assisted submissions in introductory programming via code anomaly. *Educ Inf Technol* 2024;29:16841-66. DOI
50. Zheng Y. Optimization of computer programming based on mathematical models of artificial intelligence algorithms. *Comput Electr Eng* 2023;110:108834. DOI
51. Wang WY, Zhang S, Li G, et al. Artificial intelligence enabled smart design and manufacturing of advanced materials: the endless Frontier in AI⁺ era. *MGE Advances* 2024;2:e56. DOI
52. Chen L, Chen Z, Yao X, et al. High-entropy alloy catalysts: high-throughput and machine learning-driven design. *J Mater Inf* 2022;2:19. DOI
53. Song G, Diao Z, Lv X, Liu L. TIG and laser-TIG hybrid filler wire welding of casting and wrought dissimilar magnesium alloy. *J Manuf Process* 2018;34:204-14. DOI
54. Weman K. 1 - Introduction to welding. In: *Welding processes handbook* (Second Edition). Woodhead Publishing; 2012. p.1-12. Available from: <https://shop.elsevier.com/books/welding-processes-handbook/weman/978-0-85709-510-7>. [Last accessed on 26 Dec 2024].
55. Zhang D, Liu Y, Liu R, et al. Characterization of corrosion behavior of TA2 titanium alloy welded joints in seawater environment. *Front Chem* 2022;10:950768. DOI PubMed PMC
56. Lei T, Wu C, Rong Y, Huang Y. The development of tube-to-tubesheet welding from automation to digitization. *Int J Adv Manuf Technol* 2021;116:779-802. DOI
57. Jin Z, Li H, Zhang C, Wang Q, Gao H. Online welding path detection in automatic tube-to-tubesheet welding using passive vision. *Int J Adv Manuf Technol* 2017;90:3075-84. DOI
58. Mu H, He F, Yuan L, Commins P, Ding D, Pan Z. A digital shadow approach for enhancing process monitoring in wire arc additive manufacturing using sensor fusion. *J Ind Inf Integr* 2024;40:100609. DOI
59. Pires J, Smith JS, Balfour C. Real-time top-face vision based control of weld pool size. *Ind Robot* 2005;32:334-40. DOI
60. Ugender S. Influence of tool pin profile and rotational speed on the formation of friction stir welding zone in AZ31 magnesium alloy. *J Magnes Alloy* 2018;6:205-13. DOI
61. Lin Q, Yang S, Yang R, Wu H. Transistor modeling based on LM-BPNN and CG-BPNN for the GaAs pHEMT. *Int J Numerl Model EL* 2024;37:e3268. DOI
62. Wang X, Sun L, Bai H, Yu K, Wang B. SCARA mechanical fault identification based on WPM-SE+BPNN method. *Meas Sci Technol* 2022;33:085007. DOI
63. Lee S, Kim H, Lieu QX, Lee J. CNN-based image recognition for topology optimization. *Knowl Based Syst* 2020;198:105887. DOI
64. Zhang Y, You D, Gao X, Zhang N, Gao PP. Welding defects detection based on deep learning with multiple optical sensors during disk laser welding of thick plates. *J Manuf Syst* 2019;51:87-94. DOI
65. Liu T, Zheng H, Zheng P, et al. An expert knowledge-empowered CNN approach for welding radiographic image recognition. *Adv Eng Inform* 2023;56:101963. DOI
66. Wang Z, Cao B, Liu J. Hyperspectral image classification via spatial shuffle-based convolutional neural network. *Remote Sens* 2023;15:3960. DOI
67. Dai F, Wen B, Hu Y, Gu X. A deep neural network potential model for transition metal diborides. *J Mater Inf* 2024;4:10. DOI
68. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27-30; Las Vegas, USA. IEEE; 2016. pp. 770-8. DOI
69. Zhu H, Sun M, Fu H, Du N, Zhang J. Training a seismogram discriminator based on ResNet. *IEEE Trans Geosci Remote Sens* 2021;59:pp.7076-85. DOI
70. Howard AG, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. *arXiv* 2017, arXiv.1704.04861. Available online: <https://doi.org/10.48550/arXiv.1704.04861> (accessed 26 Dec 2024)
71. Liu Z, Zhang X, Shen Z, Wei Y, Cheng KT, Sun J. Joint multi-dimension pruning via numerical gradient update. *IEEE Trans Image Process* 2021;30:8034-45. DOI
72. Ma S, Zhang Q, Li T, Song H. Basic motion behavior recognition of single dairy cow based on improved Rexnet 3D network. *Comput Electron Agr* 2022;194:106772. DOI
73. Han K, Wang Y, Chen H, et al. A survey on vision transformer. *IEEE Trans Pattern Anal Mach Intell* 2023;45:87-110. DOI
74. Sun W, Qin Z, Deng H, et al. Vicinity vision transformer. *IEEE Trans Pattern Anal Mach Intell* 2023;45:12635-49. DOI
75. Lv P, Xu H, Zhang Q, et al. An improved lightweight ConvNeXt for rice classification. *Alex Eng J* 2025;112:84-97. DOI
76. Zhu J, Feng Y, Liu Q, et al. An improved ConvNeXt with multimodal transformer for physiological signal classification. *IEEE Access* 2024. DOI
77. Hendrycks D, Gimpel K. Gaussian error linear units (GELUs). *arXiv* ;2016, arXiv:1606.08415. Available online: <https://doi.org/10.48550/arXiv.1606.08415> (accessed 26 Dec 2024).
78. Lee M, Wu Q. Mathematical analysis and performance evaluation of the GELU activation function in deep learning. *J Math* 2023;2023:1-13. DOI

79. Qiao Y, Zhang Q, Qi Y, Wan T, Yang L, Yu X. A waste classification model in low-illumination scenes based on ConvNeXt. *Resour Conserv Recy* 2023;199:107274. DOI
80. Jablonka KM, Ai Q, Al-Feghali A, et al. 14 examples of how LLMs can transform materials science and chemistry: a reflection on a large language model hackathon. *Digit Discov* 2023;2:1233-50. DOI
81. Choi J, Lee B. Accelerating materials language processing with large language models. *Commun Mater* 2024;5:449. DOI
82. Zang Y, Li W, Han J, Zhou K, Loy CC. Contextual object detection with multimodal large language models. *Int J Comput Vis* 2024. DOI
83. Mahowald K, Ivanova AA, Blank IA, Kanwisher N, Tenenbaum JB, Fedorenko E. Dissociating language and thought in large language models. *Trends Cogn Sci* 2024;28:517-40. DOI PubMed PMC
84. Wang H, Li J, Wu H, Hovy E, Sun Y. Pre-trained language models and their applications. *Engineering* 2023;25:51-65. DOI
85. Hu L, Liu Z, Zhao Z, Hou L, Nie L, Li J. A survey of knowledge enhanced pre-trained language models. *IEEE Trans Knowl Data Eng* 2024;36:1413-30. DOI
86. Lai Z, Wu T, Fei X, Ling Q. BERT4ST:: fine-tuning pre-trained large language model for wind power forecasting. *Energ Convers Manage* 2024;307:118331. DOI
87. Cook A, Karakuş O. LLM-commentator: novel fine-tuning strategies of large language models for automatic commentary generation using football event data. *Knowl Based Syst* 2024;300:112219. DOI
88. Zhang Z, Wen G, Chen S. Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding. *J Manuf Process* 2019;45:208-16. DOI
89. Yang X, Zhang Y, Lv W, Wang D. Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier. *Renew Energ* 2021;163:386-97. DOI
90. Balado J, Sousa R, Diaz-Vilariño L, Arias P. Transfer learning in urban object classification: online images to recognize point clouds. *Automat Constr* 2020;111:103058. DOI
91. Deng J, Dong W, Socher R, Li LJ, Li K, Li FF. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition; 2009 Jun 20-25; Miami, USA. IEEE; 2009. pp. 248-55. DOI
92. Husnain M, Missen MMS, Mumtaz S, Luqman MM, Coustaty M, Ogier J. Visualization of high-dimensional data by pairwise fusion matrices using t-SNE. *Symmetry* 2019;11:107. DOI
93. Moon H, Na S. Optimum design based on mathematical model and neural network to predict weld parameters for fillet joints. *J Manuf Syst* 1997;16:13-23. DOI
94. Gao Y, Hao K, Xu L, et al. Microstructure homogeneity and mechanical properties of laser-arc hybrid welded AZ31B magnesium alloy. *J Magnes Alloy* 2024;12:1986-95. DOI
95. Geng P, Ma H, Wang M, et al. Dissimilar linear friction welding of Ni-based superalloys. *Int J Mach Tool Manu* 2023;191:104062. DOI
96. Kim I, Son K, Yang Y, Yaragada P. Sensitivity analysis for process parameters in GMA welding processes using a factorial design method. *Int J Mach Tool Manu* 2003;43:763-9. DOI
97. Cook GE. Robotic arc welding: research in sensory feedback control. *IEEE Trans Ind Electron* 1983;IE-30:252-68. DOI
98. Nie L, Lin C, Liao K, Liu S, Zhao Y. Unsupervised deep image stitching: reconstructing stitched features to images. *IEEE Trans Image Process* 2021;30:6184-97. DOI