

Review

Deep learning-based welding image recognition: A comprehensive review

Tianyuan Liu^a, Pai Zheng^{a,*}, Jinsong Bao^{b,*}^a Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, 999077, Hong Kong, China^b College of Mechanical Engineering, Donghua University, Shanghai 201620, China

ARTICLE INFO

Keywords:

Welding system
Computer vision
Image recognition
Deep learning
Convolutional neural network
Explainable AI

ABSTRACT

The reliability and accuracy of welding image recognition (WIR) is critical, which can largely improve domain experts' insight of the welding system. To ensure its performance, deep learning (DL), as the cutting-edge artificial intelligence technique, has been prevailingly studied and adopted to empower intelligent WIR in various industry implementations. However, to date, there still lacks a comprehensive review of the DL-based WIR (DLBWIR) in literature. Aiming to address this issue, and to better understand its development and application, this paper undertakes a state-of-the-art survey of the existing DLBWIR research holistically, including the key technologies, the main applications and tasks, and the public datasets. Moreover, possible research directions are also highlighted at last, to offer insightful knowledge to both academics and industrial practitioners in their research and development work in WIR.

1. Introduction

Machine vision plays a critical role in assuring weld quality through seam tracking [1] and online monitoring [2] and NDT [3] post welding. Its core, WIR, has gone through two main stages, per the use of feature engineering or DL models [4]. With rapid developments of third wave AI innovation and high-performance computing infrastructure [5,6], DLBWIR has become a dominant research hotspot [7,8].

In literature, the concept of DL was first introduced by Hinton et al. [9] in 2006. Then, the overwhelming dominance of AlexNet in 2012 led to the explosion of DL in the CV domain [10]. Not until 2016, when AlphaGo defeated the world Go champion Lee Sedol, DL has started to attract ever-increasing attention from industry [11]. Wang et al. [12] summarized the commonly used DL structures in manufacturing, namely CNN, RNN, AE and DBN. Nevertheless, to the authors' best knowledge, there still lacks comprehensively reviews on the DLBWIR as the fundamental basis for researchers and industrial practitioners in this domain. Motivated by it, this paper presents a holistic review of recently published articles on DLBWIR, selected by the Web of Science database from January 2016 to June 2022, as shown in Fig. 1.

As DL came to the forefront of industry in 2016, Günther et al. [13] conducted initial explorations and mapped out a DL-based intelligent welding framework. After a year of precipitation, more scholars began to publish technical articles ever since 2018. During the outbreak phase (last three years, 2019–2021), the average annual growth rate of the

number of publications is 123 % $((24-7)/7 + (46-24)/24 + (62-46)/46)/3$. Furthermore, as of the end of June, 52 articles have been published in 2022. This indicates that the research on DLBWIR is still at a high rate of growth. Fig. 2. plots the unit locations of the article's participants, from which it can be seen that the research team came from 27 countries in Asia, North America, Europe, South America, Oceania and Africa, indicating that the topic has attracted global interest (Since an article may involve research institutions from more than one country, thus $\sum_{location_i} number_i > \sum_{year_j} number_j$, as shown later). In response to this statistic, the authors argue that this is driven by smart manufacturing strategic plans such as the American Industrial Internet, German Industry 4.0 and Made in China 2025. AI technology with DL as the core is the supporting element of these strategic planning. As a result, research related to DLBWIR as one of the instantiated scenarios for smart manufacturing also shows the trends in Fig. 1 and the distribution of research in Fig. 2.

Completeness is the fundamental for review work. This paper covers all relevant journal articles published from January 2016, when DL first entered the field of WIR. Table 1 summarizes the search topics, types of reference, source of databases and reference screening methods in this review. As a result, 230 articles were derived. Next, the authors assessed the relevance of each paper to the review topic based on the content of each article's abstract, introduction, method and conclusion. For example, some articles contain the keywords "DL" or "welding image", but they do not include DL and welding image as a whole, which were therefore excluded from further review. Eventually, 192 articles

* Corresponding authors.

E-mail addresses: pai.zheng@polyu.edu.hk (P. Zheng), bao@dhu.edu.cn (J. Bao).

Nomenclature	
WIR	Welding image recognition
DL	Deep learning
DLBWR	DL-based WIR
AI	Artificial intelligence
CV	Computer vision
CNN	Convolutional neural network
RNN	Recurrent neural network
AE	Auto encoder
DBN	Deep belief network
PRW	Pre-welding
IW	In-welding
POW	Post-welding
CCD	Charge coupled device
TFT	Time-frequency transform
OM	Optical microscope
ROI	Region of interest
PreP	Preprocessing
NormL	Normalization
DA	Data augmentation
TL	Transfer learning
MML	Multi-modal learning
GAN	Generative adversarial network
DSC	Depthwise separable convolution
NDT	Non-destructive testing
ME	Model ensemble
AM	Attention mechanism
CAM	Class activation mapping
GradCAM	Gradient-weighted class activation mapping
BN	Batch normalization
KCV	K-fold cross-validation
LSTM	Long short-term memory
Acc	Accuracy
MSE	Mean squared error
DE	Detection error
MAP	Mean average precision
GOF	Goodness of fit
Rec	Recall
IOU	Intersection over union
TE	Tracking error
PRMSE	Pixel root mean square error
SS	Structural similarity
DWIQI	Duplex wire-type image quality indicator
KG	Knowledge graph
WDKG	Welding domain knowledge graph

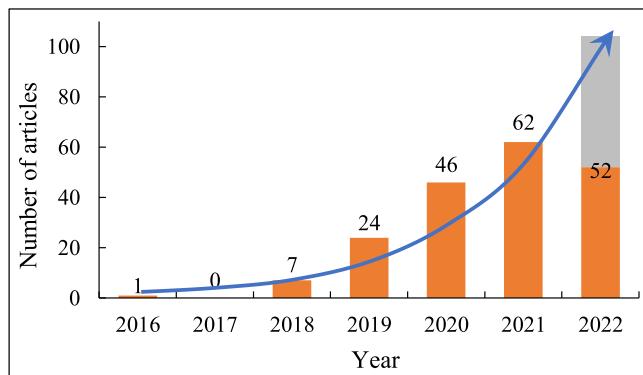


Fig. 1. Statistics on the number of articles published each year.

published in IEEE, Elsevier, Springer, Taylor & Francis, MDPI, ASME, AWS, SAGE, and Hindawi were selected.

2. Connotations and motivations

2.1. Welding image

2.1.1. What is welding image

Welding is a manufacturing process and technique for joining metals or other thermoplastic materials such as plastics by means of heat, high temperature or pressure. There are abundant visual information sources, either being related to or reflecting the quality of the weld produced, from all three stages: PRW, IW, and POW. Such information helps assure the welding process operate appropriately to produce desired welds and decide if produced welds meet quality requirements. Per literature, welding image can be defined as “grid-type data of single or multiple channels obtained by direct or indirect methods in the three stages of PRW, IW and POW, which can reflect the characteristics of weld seam and welding quality and promote welding process improvement” [7,14]. Therefore, its core is to reflect welding quality information, which is characterized by grid-type data from any of the welding stages. From the source point of

view, welding images can be obtained directly through visual sensing, or indirectly through other sensing methods. The division of welding stages and typical welding images corresponding to each stage are shown in Fig. 3. In general, the WIR of PRW, IW, and POW is used to solve the welding tracking, welding process monitoring and defect detection tasks, respectively.

2.1.2. Why image-based

One of the main characteristics of image compared to other data types is its spatial nature so that it can contain comprehensive features such as position, shape, size and texture that reflect the state of the welding process producing the welds or directly depict the welds. However, acquiring an image that sufficiently reflects the target welding information is often challenging and such challenge varies per requirement, welding state and manufacturing environment. In response to this problem, scholars have used different approaches to acquire welding images to ensure specific engineering requirements be met. Du et al. [15] scanned the weld joint with a laser strip and imaged it with a CCD camera to obtain the position of the weld seam. Liu et al. [16] imaged molten pool and keyhole from the front by a CCD camera to identify different penetration levels in laser welding. Jiao et al. [17] simultaneously collected the front and back information of the molten pool to observe the penetration state. Li et al. [18] collected both the back image and the front laser reflection image of the molten pool to obtain sufficient penetration information. Lu et al. [19] used a CCD camera to image workers’ behavior to judge whether they conform to welding specifications. Knaak et al. [20] observed the surface defects of the molten pool by mid-wave infrared imaging technology. Jin et al. [21] converted the current signal into a two-dimensional image by means of TFT, and used it to visually display the information of the forming quality of the backside of the weld. For the problem of welding fault identification, Lee et al. [22] designed an interpretable TFT method. Li et al. [23] reflected the molten pool size information by simulated laser point reflection image. Yang et al. [24] photographed the appearance of the weld by CCD to observe its surface forming quality. In the references [25–27], the internal defect information of the weld was obtained by imaging methods of X-ray, microscope and TFT, respectively. Table 2 summarizes the different ways in which welding image has been acquired in literature and their contents.

2.1.3. How to obtain welding image

It can be seen from Fig. 4 that the current research is mainly focused on obtaining image data in the stage of IW and POW. This is mainly because the technology of seam tracking for PRW is currently at a more mature stage [38], while online closed-loop monitoring for IW and intelligent NDT for POW are still challenging tasks. Fig. 4 provides statistics on the direct and indirect ways of acquiring welding images during the different stages. One can find that the CCD dominates welding image acquisition technology and has been widely used in PRW, IW and POW. The authors attribute this statistic to the fact that CCD imaging is low cost, rich information, no contact with the workpiece, and wide application scenarios [39,40].

In PRW stage, the current research mainly uses laser stripes to scan the weld bead and groove, and then uses CCD to shoot the laser stripes to reflect the welding quality information.

In IW stage, CCD is used to directly photograph the melting area to observe the dynamic process of this area from the outside. In addition, other sensing methods such as sound and electricity are used to obtain one-dimensional signals in the welding process, and then two-dimensional welding images are obtained through TFT. It also simulates the image during the welding process based on the theoretical model.

In POW stage, the CCD is mainly used to take images of the appearance of the weld seam in order to observe the quality of the surface formation. However, this welding image does not show the internal defects of the welding seam. Based on the OM, internal defects can be seen from the weld profile in a destructive way, while the internal defects can be seen from the image in a non-destructive way based on the X-ray scanning method. In addition, one-dimensional signals of weld formation can be obtained through sound or ultrasound, and then two-dimensional welding image can be obtained through TFT.

Table 1

Systematic method on screening relevant articles.

Searching Index	Specific Content
Search topics	TS = ((weld image OR molten pool OR weld penetration OR weld globule transition OR weld seam OR weld radiographic OR weld defect) AND (deep learning OR DL OR CNN OR convolutional neural network))
Article type	Scientific/technical articles published in peer-reviewed journals and published online
Search period	From January 2016 to June 2022
Database	Science citation index expanded (SCIE) and Emerging sources citations index (ESCI)
Screening procedure	The relevance with the research topic as judged by the contents of abstract, introduction, method and conclusion of every papers.
Classification scheme	Key technologies of current development of DLBWIR (as shown in Section 3), applications of DLBWIR (as shown in Section 4) and available public datasets (as shown in Section 5)

2.2. Welding image recognition

After obtaining the welding image, the abstract feature information reflecting the welding quality such as size, shape and category, can be recognized from the direct information including brightness, color and pixel distribution contained in the welding image by the WIR technology. According to the actual task requirements, this paper divides the WIR task into four levels as shown in Fig. 5: 1) *Classification*. It aims to discover the object category in the welding image, which is an image-level recognition task (*What?*). 2) *Detection*. It aims to find the position of the object in the welding image, which is a target-level recognition task (*Where?*). 3) *Segmentation*. It aims to calculate the size of the object in the welding image, which is a pixel-level recognition task (*How much?*). 4) *Tracking*. It aims to solve when and where the target is in the welding image, which is a target-level recognition task for dynamic images (*When and where?*). It is worth noting that the above four types of

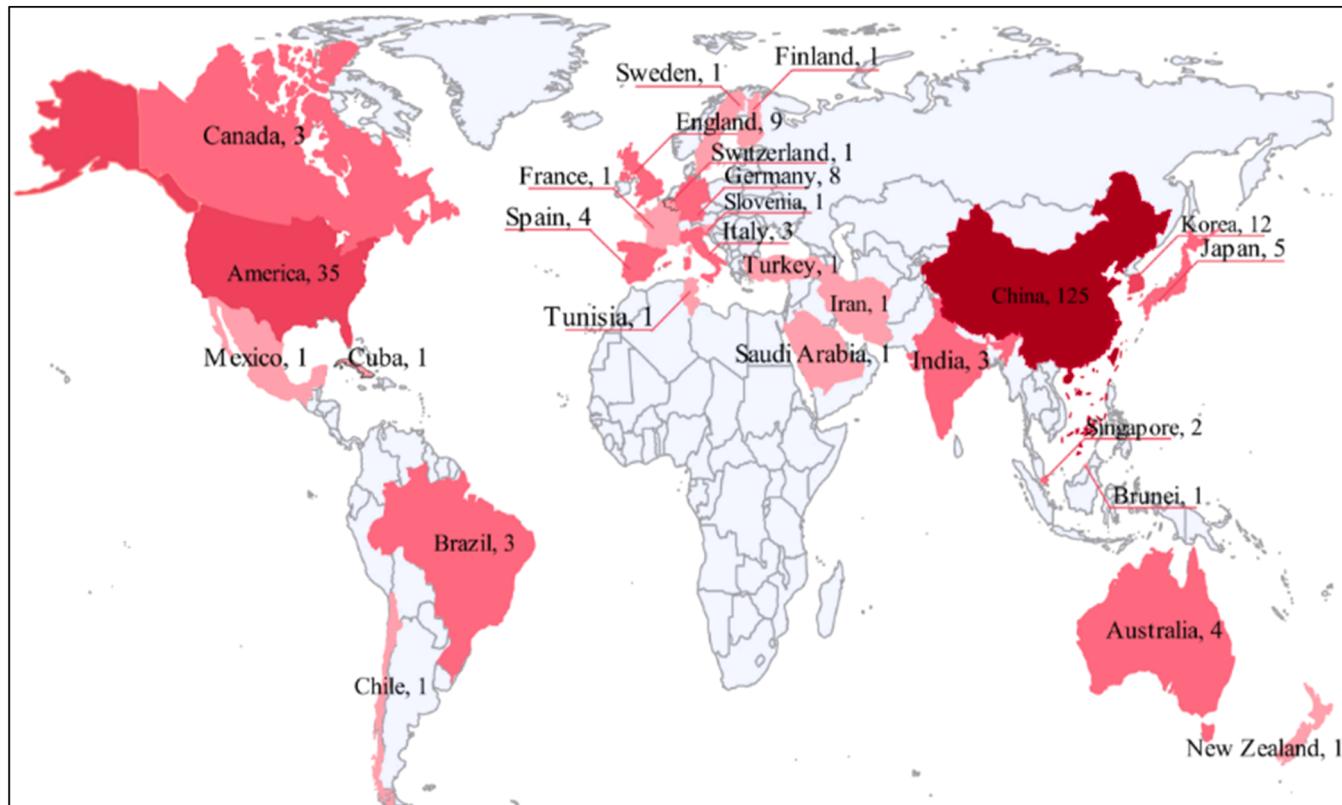


Fig. 2. Geographical distribution of published works in DLBWIR worldwide.

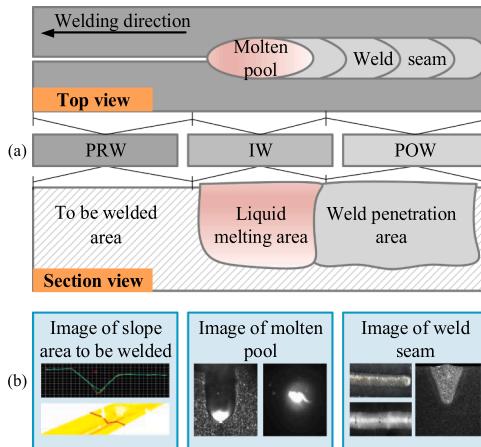


Fig. 3. Conceptual scope of welding image: (a) division of welding stages, and (b) typical welding image.

tasks are not completely independent. For example, the tasks of the second and third levels generally also include the tasks of the first level. Therefore, welding image classification is the most fundamental recognition task, and there are relatively many studies on this level.

2.3. Deep learning

2.3.1. What is deep learning

Table 3 lists several typical definitions of DL. Although scholars have defined DL from multiple perspectives such as structure, process, purpose, scope, function, and form, *they lack a description of DL from an implementation perspective*. Therefore, in conjunction with engineering practice, this paper refers to the data, model, training and platform technologies involved in the DLBWR process together with the basic connotation of DL mentioned in Table 3 below.

2.3.2. Why deep learning-based

Assuming that the welding image is x and the welding quality information is Q , the process of WIR is to seek f so that $f(x) = Q$. In general, the development of WIR has gone through two stages: feature engineering-based and DL-based. In the feature engineering-based method, it generally follows the technical route of image denoising, image enhancement, object segmentation, feature extraction, feature selection, and feature recognition, that is, $f(x) = f_{\text{Rec}}(f_{\text{Sel}}(f_{\text{Ext}}(f_{\text{Seg}}(f_{\text{Enh}}(f_{\text{Den}}(x))))))$. This idea decomposes the process of finding f into several sub-tasks. Although it has strong interpretability, it has the following shortcomings. First, the process involves too many intermediate subtasks. This not only spreads the optimization goal, but also fails to achieve the global optimum. Second, the process relies too much on expert empirical knowledge. This not only makes it difficult to give full play to the value of data, but also requires higher multidisciplinary knowledge collaboration. The DL-based method can directly establish the end-to-end mapping from x to Q , that is, $f(x)$ can be directly derived from $\text{DL}(x)$. This modelling approach requires only one objective function to be optimized and eliminates intermediate steps and cumulative errors. In addition, the process can be data-driven to adaptively learn and extract abstract visual features of the object, avoiding the need for a manual experience-driven feature extraction process. Therefore, this idea can not only achieve global optimization, but also give full play to the value of data and improve the efficiency of model development.

2.3.3. How to perform DLBWR

As DLBWR is based on a DL approach to solving WIR tasks, DLBWR is largely in line with the technical framework followed by DL in other areas of image recognition. Although unsupervised learning is also an

effective method for image recognition, it has been studied very little in the WIR field. Therefore, based on the definition of DL, the paradigm of supervised DLBWR can be summarized in Fig. 6. In practice, three major ways of obtaining welding image labels are: expert experience-based manual labelling [46], welding process-based automatic labelling [47] and the high precision inspection-based automatic labelling [48]. After acquiring the welding image data and the labels corresponding to the different tasks, a certain hypothesis function is learned from the hypothesis space defined by the structure of the model for the different tasks as the final model according to the training method. The process is supported by the platform provided by the hardware and software environment. This well-trained model is an approximation of the real model.

2.4. Related surveys

In view of the important role of WIR in welding systems, scholars have reviewed the literature in the field from different perspectives. Table 4 summarizes the scope, taxonomy and relevance to DLBWR of different review articles.

The following conclusions can be drawn from a comparative analysis of recently published review articles. The academic community generally believes that DL, especially CNN, has significant advantages in WIR and is the research direction to solve the WIR problem. Despite the significant contributions of each reviewed survey, *these review articles either present DL as a future perspective for solving WIR problems or analyze some of the DLBWR literature as it relates to individual welding stages*.

Accordingly, this paper will provide a comprehensive review of the DLBWR literature involved in the three stages of PRW, IW and POW. Although people are very curious to know "what are the difficulties and challenges in DLBWR?", actually, this question is not easy to answer and may even be over-generalized. Since the welding image at different stages and the WIR tasks in different scenarios have totally different objectives and constraints, their difficulties may vary from each other. As such, it is difficult to generalize existing research from the perspective of unifying issues and challenges. Therefore, this paper will summarize the key technical components of DLBWR from an overall perspective in Section 3, as well as the challenges and corresponding solutions brought by the characteristics of welding image to the application of DL. In Section 4, the engineering requirements, vision tasks, difficulties, solutions and results of DLBWR tasks in different welding stages will be dissected. Given the key position of data in DLBWR, this paper will summarize the public datasets in the field and analyze their features in Section 5. Finally, an outlook on future directions, both in terms of literature and authors, will be presented in Section 6.

3. Key technologies of DLBWR

Based on the paradigm of DLBWR, this section summarizes the key technologies involved in DLBWR from four aspects: *data, model, training and platform*.

3.1. Data technology

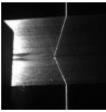
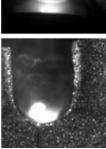
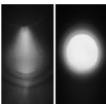
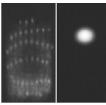
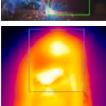
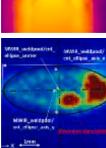
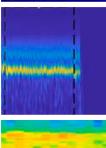
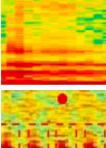
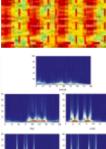
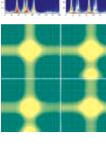
3.1.1. Welding image itself

As shown in Fig. 7, the key technologies for the welding image itself in the literature mainly include ROI, PreP, and NormL.

Table 5 summarizes the specific ROI, PreP, and NormL methods involved in the existing literature. Welding image generally has the characteristics of high pixel, small object, and large background. ROI refers to finding out the region of interest in the original image and identifying it. Therefore, the ROI operation on the original welding image can make the model focus on learning the characteristics of the key area while reducing the amount of calculation. Furthermore, as the input image size is fixed for some classical network structures and pre-trained models, ROI methods become one of the most commonly used

Table 2

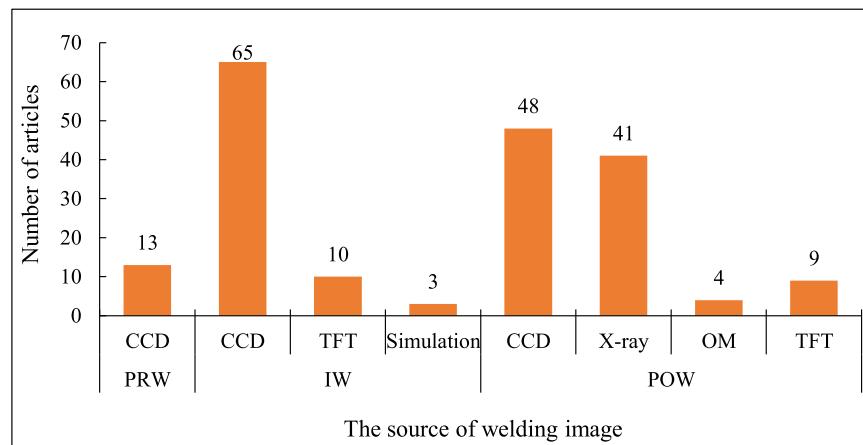
Comparison of typical welding image at different stages.

Ref.	Year	Stage	Method	Sketch Map	Image content	Welding information
[15]	2019	PRW	CCD		Single-stripe laser, a single weld joint and groove.	Weld joint position
[28]	2021	IW	CCD		Front of keyhole	Penetration state
[16]	2021				Front of molten pool and keyhole	Penetration state
[17]	2020				Front and back of molten pool	Penetration state
[18]	2020				The back side of the melt pool, and the front laser stripe.	Penetration state
[29]	2019				Weld seams are captured in three directions by multiple specular reflections.	Penetration state
[19]	2020				Protective products for manual welding	Work safety and welding processes
[30]	2022				Infrared imaging of the molten pool	Penetration state
[20]	2021				Mid-wave infrared imaging of the molten pool	Surface defects
[21]	2020	TFT			Current signal	Weld back forming
[31]	2021				Vibration signal	Internal defects in welded joint
[32]	2020				Sound signal	Penetration state
[33]	2019				Eddy current signal	Internal defects in weld seam
[22]	2022				Acceleration, displacement, current and voltage.	Interpretable transformation to reflect welding failure mechanisms

(continued on next page)

Table 2 (continued)

Ref.	Year	Stage	Method	Sketch Map	Image content	Welding information
[34]	2021		Simulation		Molten pool	Internal defects in weld seam
[23]	2021				Laser reflection point	Size of molten pool
[24]	2021	POW	CCD		Weld appearance	Surface defects in weld seam
[35]	2020				Multi-stripe laser scan multiple weld bead and groove	Surface defects in weld seam
[25]	2021		X-ray		Inside the weld seam	Internal defects in weld seam
[26]	2020		OM		Micro-organization	Internal defects in weld seam
[36]	2021		TFT		Sound signal	Internal defects in weld seam
[27]	2020				Ultrasound signal	Internal defects in weld seam
[37]	2022				Time-of-flight diffraction	Internal defects in weld seam

**Fig. 4.** The way of acquiring welding image in different welding stages.

techniques for welding image data itself.

Due to the interference of arc light, spatter, and vibration in the welding environment, PreP methods represented by filtering, denoising, and enhancement have also been extensively researched. The PreP results in high quality welding image that facilitates the learning of feature information.

Meanwhile, the purpose of welding image NormL is to compress the grey scale range, thus reducing the search space and speeding up the learning and convergence of the DL model.

3.1.2. Welding image dataset

In addition to the problems of the welding image itself, there is also a widespread problem of small samples at the dataset level, which brings a huge obstacle to the in-depth application of the DL methods. From the data point of view, scholars have mainly addressed this problem through DA, TL, and GAN. **Fig. 8** plots the number of occurrences of the three methods in the literature.

Table 6 summarizes the specific DA, TL, and GAN methods involved in the existing studies.

The idea of DA is to artificially transform the original image in the spatial, greyscale and color domains to obtain new samples of welding

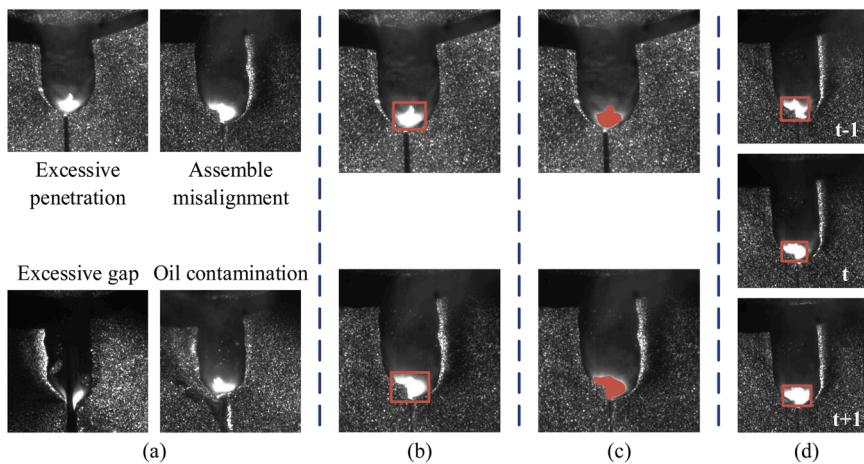


Fig. 5. Four basic tasks of WIR: (a) classification, (b) detection, (c) segmentation, and (d) tracking.

Table 3

Typical definitions of DL.

Ref.	Year	Definition	Keywords	Highlights
[41]	2009	DL is about learning feature hierarchies with features from higher levels of the hierarchy formed by the composition of lower level features.	Feature hierarchies, composition	Structure, process
[42]	2014	DL is a class of machine learning algorithms that uses multiple layers to progressively extract higher-level features from the raw input.	Machine learning, multiple layers	Scope, process
[43]	2015	DL (also known as deep structured learning or hierarchical learning) is learning data representations, as opposed to task-specific algorithms. Learning can be supervised, semi-supervised or unsupervised.	Hierarchical learning, data representations	Function, form
[44]	2015	DL is a multi-level feature learning method that uses simple but non-linear components to gradually abstract the features of each layer (starting from the original data) into higher-level features.	Multi-level, feature learning	Structure, process
[45]	2018	DL is not only the process of learning the relationship between two or more variables, but also the process of learning the knowledge that controls the relationship and making the knowledge meaningful.	Learning relationship, learning knowledge	Purpose

image to expand the dataset [70]. DA is the simplest manner to implement and the process of learning the various transformations greatly improves the inductive bias of the model, making it the most commonly used method for inadequate welding image data. In practice, there are both offline and online ways to achieve DA.

Meanwhile, the idea of TL is to use the knowledge learned by the model to solve the source task based on the source dataset to improve the target task based on the target dataset [71]. Essentially, the model structure is first pre-trained based on the source dataset and then fine-tuned based on the target dataset. Where a source dataset is a dataset with a rich sample size and label information that is different from the target dataset. The target dataset is a dataset that currently holds insufficient data volume and label information to solve the target task. Source datasets can be classified into out-of-domain and in-domain

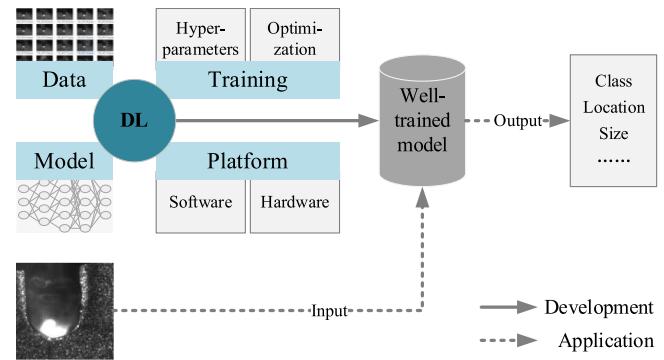


Fig. 6. The overall paradigm of DLBWIR.

datasets from the perspective of relevance to the target dataset.

Moreover, GAN [72] is to learn the distribution of existing data through generative network, and to expand the dataset by generating new welding image samples based on the learned distribution. The commonly used GAN structures in DLBWIR are DCGAN [73], CGAN [74], CycleGAN [75], BAGAN [76] and WGAN [77].

3.2. Model technology

3.2.1. Backbone

After acquiring the welding image, a suitable model structure needs to be selected to learn the feature patterns embedded in the welding image. Fig. 9 summarizes the backbone structures that have been used for different DLBWIR tasks. It is worth noting that some studies have optimized and improved some hyperparameters of the backbone according to the specific problem (e.g., changing the number of layers and the number of convolutional kernels, etc.), but the core of the network still belongs to that backbone, so it is counted as such for statistical purposes.

Among all backbone structures, Lenet5 [96] has received the most applications, mainly because the structure is the simplest and contains the fewest parameters to be trained, which is more suitable for WIR with small samples. It also contains fewer key operations and techniques and is therefore more scalable. Resnet [97] introduces the residual structure to improve the trainability of deeper DL models. VGG [98] maintains the size of the receptive field by reducing the size of the convolution kernel and increasing the depth of the network. Alexnet [10] not only implements an efficient parallel computing structure, but also proposes classical training techniques to prevent overfitting of the model. Mobilenet [99] decomposes the classic convolution operation into two stages of

Table 4
Review articles related to DLBWIR.

Ref.	Year	Scope	Taxonomy	Relation to DLBWIR
[38]	2019	Weld seam tracking task in the PRW stage	Sensing technology, feature extraction technology and application scenarios for seam tracking task.	The authors found that vision is the most widely used sensing method for seam tracking task, and believe that AI technology is the future direction of quality monitoring and process optimization.
[4]	2021	Monitoring task in the IW stage	Sensing, preprocessing, segmentation, feature definition, extraction, and recognition techniques for WIR.	The authors provide a brief analysis of the CNN-based WIR literature in the feature recognition stage and argue that DL techniques are the future direction of WIR.
[7]	2020	Laser welding monitoring task in the IW stage	Visual sensing technology and visual monitoring application scenarios.	The authors consider AI and in particular DL-based quality monitoring and evaluation system to be the most interesting and challenging future direction.
[49]	2022	Laser welding monitoring task in the IW stage	Visual sensing, feature analysis and modelling techniques.	A summary of modeling methods based on DL and CNN in the laser welding literature is presented.
[50]	2020	NDT task in the POW stage	Preprocessing, segmentation, feature extraction/selection and recognition techniques.	The application of DL, especially CNN methods, in the feature recognition stage is briefly summarized.
[51]	2021	Vision sensing task in the PRW, IW and POW stages	Vision sensing application scenarios.	The authors believe that visual analysis modules can significantly reduce the dependence of welding systems on the operator. AI and especially CNN are used as future research directions to fully exploit welding features.
[8]	2020	Intelligent welding system	Enabling technologies, platform technologies and typical application scenarios.	The authors identify AI and DL as key enabling technologies for intelligent welding system.
[14]	2020	Intelligent welding system	The welding manufacturing process is broken down into three stages: process pre-design, optimized design and in-line monitoring, and the challenges and possible solutions for each stage are condensed through formal descriptions.	The authors believe that DL is an effective approach to solving complex modelling problems in the welding manufacturing.

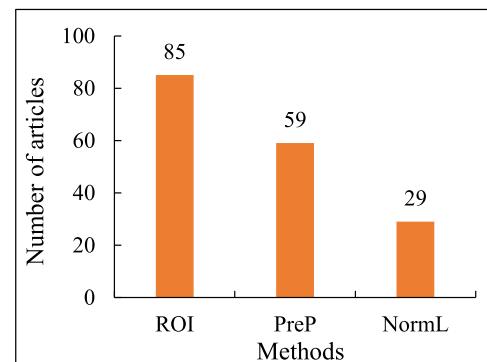


Fig. 7. Key technologies for welding image itself.

Table 5
The specific technologies for welding image itself and their comparison.

Ref.	Year	Key elements	Method	Pros	Cons
[52]	2021	Crop	ROI	Improve the processing efficiency of DL and promote DL to learn key features of welding images.	Loss of semantic information.
[53]	2019	Estimating ROI from center point and arc of the transmission			
[54]	2020	Resize			
[55]	2019	Crop and down-sampling			
[56]	2021	Crop and resize			
[57]	2020	Crop, resize and binarization			
[58]	2019	Filtering, enhancement, and segmentation	PreP	Improve the quality of welding images.	This is a non-end-to-end operation and lacks interaction with DL-based welding image feature learning process.
[59]	2021	Denoising and contrast enhancement			
[60]	2020	Histogram equalization and gray scale stretch			
[61]	2021	Deblur			
[62]	2020	Two-channel synthesis			
[63]	2020	Image alignment			
[64]	2021	Gauss low-pass filter and median filter			
[65]	2021	Spatial domain filtering and gradient edge extraction			
[66]	2019	The gray value range is (0, 1)	NormL	The model solution space is reduced to speed up model learning and convergence.	Lack of consideration for adaptability to activation functions.
[34]	2021	The gray value range is (-0.5, +0.5)			
[67]	2020	Z-NormL			
[68]	2020	Zero-center NormL			
[69]	2022	Normalize along the time axis			

depthwise convolution and pointwise convolution to reduce the amount of parameters and calculations, making it possible to deploy the DL model on the mobile terminal.

On one hand, R-CNN [100–103] and SSD [104] are the three most commonly used families of DL structures for object detection. The R-CNN series is two-stage detector. Although it is slower, it has accuracy

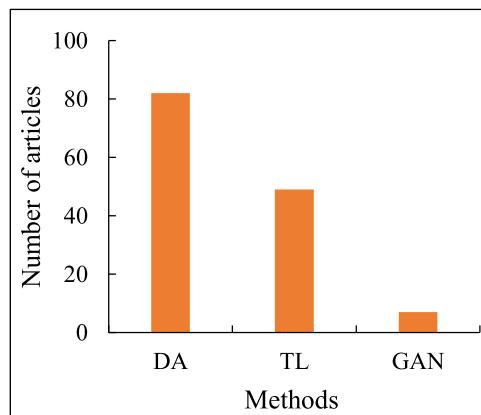


Fig. 8. Key technologies for welding image dataset.

advantages. Meanwhile, YOLO and SSD series are one-stage detectors with the advantage of fast speed. The YOLO series offers a significant increase in computing speed with less loss of accuracy, and the models in this series are always being maintained and updated, making them the most widely used in the DLBWIR field.

On the other hand, FCN [105] and U-Net [106] are two classical image segmentation structures, both consisting entirely of convolutional layers. FCN uses point-by-point summation to fuse multi-scale features, and U-Net uses concatenate to fuse multi-scale features [107]. Also, Mask R-CNN [108] is a flexible structure that can be oriented to classification, detection and segmentation. U-Net has been the most studied application for its ability to maintain high segmentation accuracy even with small samples. However, the current solution for the tracking task is to convert it into a temporal object detection or segmentation task, and there is still a lack of application of network structures such as HCF [109], MDNet [110], ECO [111], etc. that are directly modelled for the tracking task.

3.2.2. Improved structure

The improved structure based on backbone have also been widely studied and applied in DLBWIR. We summarize the improved structure that appear in the literature into four categories: MML, ME, AM, and explainability.

Although visual sensing has the advantage of being rich in information, there are also some characteristic information that cannot be obtained by visual sensing, such as the internal quality information of the weld during the welding process. In response to this problem, MML can utilize the complementarity between multiple modalities to learn a more comprehensive feature representation [112]. The shared representation between multiple modalities allows the DL model to have a stronger inductive preference for certain hypotheses and can therefore serve to limit the hypothesis space.

The welding image is formed by a combination of strongly coupled factors and the features learned by a single DL model may not be able to describe the complex patterns embedded in the welding image. The main idea of ME is to train multiple DL models simultaneously and then have multiple models make voting decisions, thus improving the overall decision accuracy.

In addition to key objects, non-main information such as background and interference widely exist in welding image, which presents an obstacle to the DL model learning the main features. The AM allows for more weighting of important feature information, thus preventing important information from being masked by noise. As such, the AM not only helps the model to decide which features in the welding image need more attention, but also to allocate the limited computational resources to the more important parts. The imposed domain of attention in DLBWIR involves spatial [113], channel [114], spatial channel mixing [115] and text semantic [16].

Table 6

The specific technologies for welding image dataset and their comparison.

Ref.	Year	Key elements	Method	Pros	Cons
[78]	2022	Stick	DA	Simple, intuitive and easy to operate	Insufficient diversity of samples
[79]	2021	Brightness and cropping			
[18]	2020	Translation and rotation			
[20]	2021	Mirroring and rotation			
[80]	2021	Mirroring and adding noise			
[61]	2021	Lighting, rotation, cropping, adding noise, flipping, saturation, contrast, resize, and cutting			
[81]	2021	Flipping, color jittering and affine transformation			
[82]	2020	Rotation, flipping, and the aspect ratio			
[24]	2021	Affine transformation, gamma transformation, linear transformation, resize, and adding noise			
[67]	2020	Shearing, skewing, flipping and elastic distortion			
[83]	2020	Rotation, shearing, shifting, zooming, and channel substitution			
[84]	2021	Rotation, flipping, color jittering, shifting, four images are cropped and stitched into one			
[85]	2021	Flipping, cutmix, and mosaic			
[86]	2019	Transferred from the out-of-domain dataset without fine turning	TL	Improving the initialization performance of model	The network structure for the target task needs to be the same as the pretrained network structure.
[87]	2021	Transferred from the dataset in the domain			
[88]	2020	Transferred from the dataset with 1,600,000 images in the domain.			
[89]	2019	Transferred from the dataset in the domain, and conduct self-supervised pre-training.			
[90]	2021	Transferred from the dataset in the domain, and conduct unsupervised pre-training.			
[91]	2022	Comparing the effect of TL based on different backbone			

(continued on next page)

Table 6 (continued)

Ref.	Year	Key elements	Method	Pros	Cons
[92]	2020	Convolutional layers replace perceptrons	GAN	The new samples are more diverse	It is difficult to learn a reliable data distribution while the original welding image data is insufficient.
[26]	2020	Add conditional information to the generator and discriminator			
[93]	2022	No image pairs required			
[94]	2022	As an enhancement tool to restore the balance of the dataset			
[95]	2020	Solve the problem of collapse mode			

The current DL is generally regarded as a black box model with poor explainability. The literature [116] argues that a key component of AI systems is the ability to explain decisions, recommendations, predictions, actions and the process. DLBWIR mainly adopts the methods of feature map visualization and class activation mapping [117,118] respectively trying to explain what the model has learned and what is the basis for the model's decision-making for individuals. Table 7 summarizes the commonly used model improvement methods in DLBWIR.

3.3. Training technology

After acquiring the dataset and determining the hypothesis space from the model structure, the model needs to be properly trained in order to find a hypothesis function that performs well. Fig. 10

summarizes the optimizer and training techniques involved in the DLBWIR model training process.

The optimizer guides the model parameters to update the appropriate size in the correct direction in the process of DL backpropagation, so that the updated parameters make the loss function value continue to approach the global minimum, which plays an important role in the model training process. It can be seen from the definition that the core of the optimizer includes two aspects, namely the optimization direction and the update step size. Different optimizers appearing in the literature are optimized and improved for these two. It can be seen from Fig. 10 that Adam and SGD are more commonly used at present. This is because Adam has the advantage of fast convergence speed, SGD has the advantage of strong generalization ability, and each DL framework has better support for these two optimizers. A detailed comparison of DL optimizers can be found in the reference [140]. For the problem of weld appearance detection, Li et al. [141] conducted a comparative study on the effects of different optimizers. For the TIG welding defect identification problem, Sekhar et al. [142] compared four optimizers. However, there is a lack in DLBWIR to analyze the characteristics of the welding image data and select the optimizer accordingly.

The use of batch training can improve memory utilization and accelerate the training process of the model. It is the most widely used training technique under the condition of rapid development of hardware technology. Compared with training based on single sample, training based on batch can lock the direction of gradient descent faster and cause smaller training shock. Batch training has applications in multiple tasks such as classification, detection, segmentation, and tracking. However, the selection of batch size in the current research is based on experience, and there is no clear basis to select a batch size suitable for specific DLBWIR tasks.

BN first calculates the mean and variance of the current batch, then normalizes the current batch, and finally introduces a scale factor and a

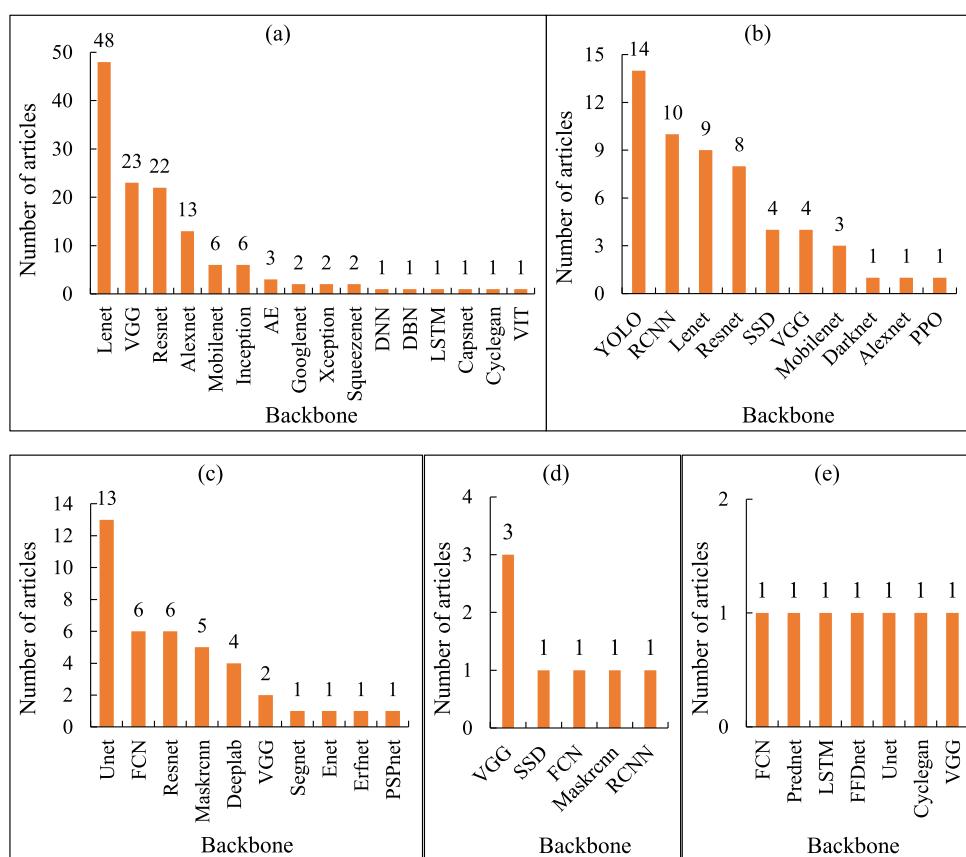
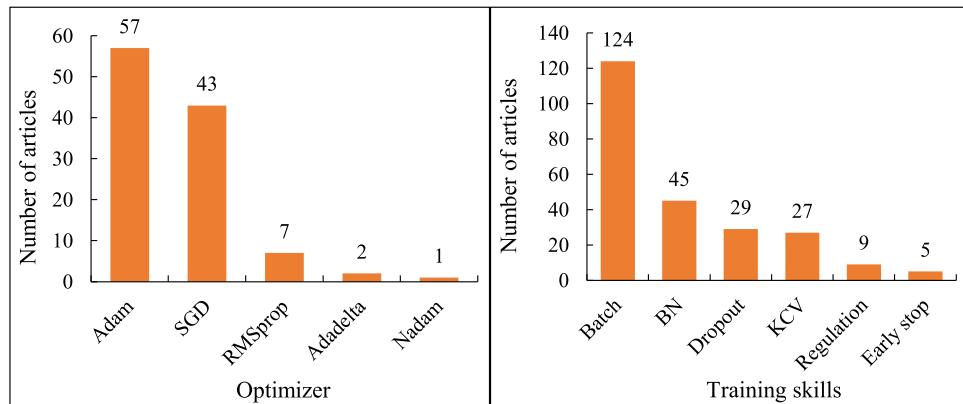


Fig. 9. Backbone commonly used in DLBWIR: (a) classification, (b) detection, (c) segmentation, (d) tracking, and (e) others.

Table 7

The specific technologies for improved structure and their comparison.

Ref.	Year	Key elements	Method	Pros	Cons
[119]	2021	Infrared thermal and visual image	MML	Complementary information,	The primary and secondary relationship
[120]	2022	High-speed camera and coherent light imaging		Limiting the hypothesis space	between modalities is not clear.
[121]	2022	Image and temperature			
[122]	2020	Image and sound			
[123]	2022	Eddy current detection and 3D laser scanning			
[124]	2020	Image, sound, voltage and current			
[16]	2021	Image and text			
[125]	2022	Handcrafted and learned features			
[95]	2020	Inception and MobileNet	ME	Voting decisions with high robustness	The training process is difficult and inefficient.
[126]	2021	Lenet5 and MLP			
[127]	2019	Alexnet, Densenet, Resnet, VGG, and Capsnet			
[128]	2018	Lenet5 and Lenet5			
[20]	2021	CNN, GRU, SVM and KNN			
[16]	2021	Label text	AM	Pay attention to important information and optimize the allocation of resources.	There is a lack of research on task justification and where to add it.
[129]	2019	Spatial			
[130]	2020	Channel			
[56]	2021	Spatial channel mixing			
[131]	2022	Spatiotemporal and channel attention			
[132]	2021	Spatial correlation and Senet			
[133]	2022	Bilinear attention			
[134]	2022	Two modal cross-attention			
[135]	2022	Four modal cross-attention			
[136]	2022	Self-attention			
[137]	2018	Feature map visualisation	Explainability	Visualization of the model's learned knowledge and basis for decision making.	Lack of semantic-level explanations.
[136]	2022	Attention visualization			
[138]	2020	GradCAM			
[82]	2020	CAM and Guided GradCAM			
[139]	2022	Multi-scale fusion features-based CAM			

**Fig. 10.** Key technologies of the model training process.

translation factor to retain the learned features [143]. This method makes the distribution of each layer's input data relatively stable, reduces the model's dependence on network initialization, accelerates the model's convergence speed, and reduces the risk of model overfitting.

Dropout refers to making some neurons stop working with a certain probability when the signal propagates forward [144]. This can make the model not too dependent on some local features, thereby improving the generalization ability. However, dropout is designed for a more generalized neural network, ignoring the correlation of parameters in the convolution kernel. Therefore, Pan et al. [145] used a more targeted Dropblock [146] method.

KCV generally divides the original dataset into k sub-datasets, a single sub-dataset is used for the verification, and other sub-datasets are used for training model. KCV is repeated k times, and the result of averaging k times can get a single estimate [147–152]. In addition,

Wang et al. [153] used a Monte Carlo cross-validation method. The advantage of this method is that the limited sample space can be utilized as much as possible and data interference can be eliminated to the greatest extent. However, the current literature does not consider the original data distribution when dividing the sub-datasets. Therefore, it is necessary to consider a hierarchical cross-validation method that combines data distribution in future applications.

The essence of parameter regularization is to limit the parameter set to be optimized, thereby standardizing the training direction of the model, and limiting the scale of the hypothesis space to prevent the model from overfitting. L1 [154] and L2 [155] parameter regularization methods are two commonly used in the literature. However, the parameters in the convolution kernel have a local correlation, and the L1 and L2 regularization methods lack consideration of this correlation. Therefore, Liu et al. [156] took into account the characteristics of the

local correlation of the parameters in the convolution kernel, and designs a coarse-grained parameter regularization method for the classification of molten pool image.

The basic idea of the early-stopping method is to stop training when the model's performance on the validation set begins to decline [157]. In this way, a model with better generalization performance can be obtained. From the perspective of the limitation of the hypothesis space, it can be considered that this method limits the parameter space of the optimization process to a small neighborhood of the initial parameters, so it can achieve a similar effect to L2 parameter regularization. The disadvantage of early-stopping is that different methods are not taken to solve the two problems of optimizing the loss function and reducing the verification loss, which may face the problem of insufficient training loss.

3.4. Platform technology

An important reason for the popularity and application of DL technology in the industry is that a large number of excellent DL frameworks have been open sourced, which has brought convenience to the industry in terms of implementation. Fig. 11 summarizes the DL framework used in DLBWIR related literature. Tensorflow [158] has a professional visualization platform for the training process, and its symbolic programming features are more compatible with the underlying computer code than other DL frameworks and easier to deploy in industrial scenarios, hence it is currently the most widely used in DLBWIR [159]. Keras [160] and Pytorch [161] have also been widely used due to the convenience brought by their highly modular packaging. In view of the fact that most of the personnel who carry out WIR task come from the engineering field, the Matlab framework [162], which is more popular in the engineering field, has also been used in several ways.

Through the development of DL, it can be found that computing power has always been a key factor hindering the development and landing of DL. Therefore, it can be seen from Fig. 12 that most of the studies in DLBWIR clearly indicate that there is a GPU accelerated environment. Further, nearly 1/2 of the literature that specifies an operating system chooses Linux due to its high stability, scalability and open source. In the absence of a local GPU acceleration environment, the solution in reference [82,163] is to implement GPU acceleration based on Google's Colab cloud server.

3.5. Brief summary

This section provides an overview of the key technologies involved in DLBWIR based on the DLBWIR research paradigm.

For sensing technology, existing research covers the three stages of PRW, IW and POW, which can be used to acquire welding image in both direct and indirect ways. Two types of images reflecting both internal and external information about the weld can be acquired. However, there are few studies on multi-source sensing to achieve information

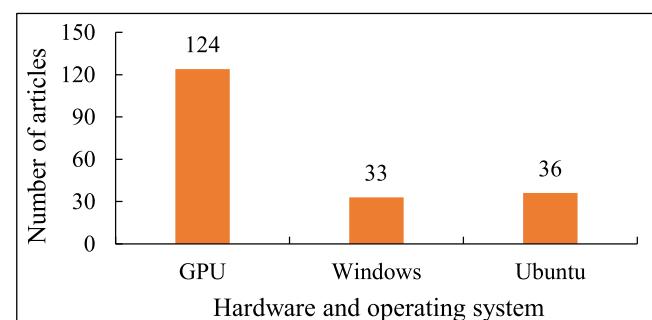


Fig. 12. Realization environment of DLBWIR research.

complementarity at the same stage in the existing literature, and there is no research on multi-source asynchronous sensing to achieve cross-stage information reasoning.

For data technology, the existing research includes two levels of key technologies for the welding image itself and the welding image dataset. The widespread problems of pixel redundancy, visual interference, and drastic grey-scale changes in welding image have been addressed by existing research through three types of methods: ROI, Prep, and NormL, respectively. In response to the small size of the welding image dataset, existing data technology has reduced the impact of small samples through DA, TL, and GAN. However, the ROI, Prep, and NormL techniques for the welding image itself lack interaction with the DL training process. The DA, TL, and GAN techniques for the welding image dataset are lacking in combination with the characteristics of welding image. For example, Zhang et al. [29] added different noise types for different penetration states and used a rotated-based DA method combining the image characteristics of the three views.

For model technology, existing research has been carried out mainly on the basis of backbone structure and improved structure in the DL field. The Lenet5 with simple structure and strong scalability is the most widely used in DLBWIR. In order to further improve the advantages of DLBWIR, three improvement ideas of MML, AM and ME are respectively used to supplement feature information sources, strengthen the learning of key features in welding image, and improve the basis for decision-making.

In order to obtain a DLBWIR model with strong generalization ability, existing research includes the main techniques for the training process of DL models. However, the relevant training techniques lack integration with the characteristics of the welding image and the optimal choice of hyperparameters.

For platform technology, open source frameworks and hardware acceleration environments provide platform support for the development of DLBWIR.

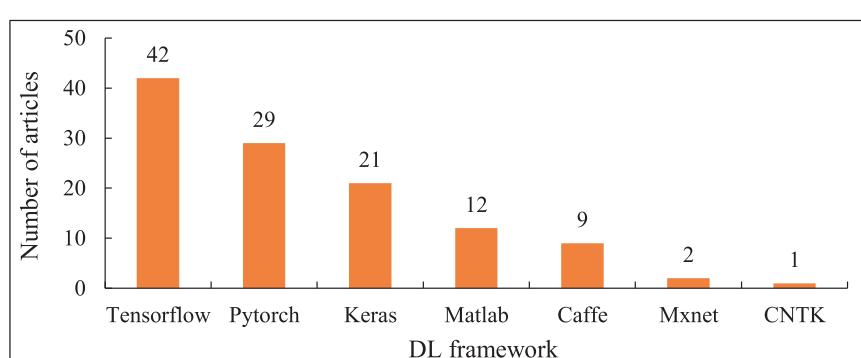


Fig. 11. The DL framework used in DLBWIR.

4. Applications of DLBWIR

This section provides an overview of the current status of DLBWIR applications from a WIR task type perspective. Fig. 13 summarizes the amount of DLBWIR posts in different task applications. The statistics were based on 192 papers that had been searched and filtered by the rules described in the Introduction. The different tasks were distinguished on the basis of the supervisory information and task scenarios as described in Section 2.2. As the classification task is the basis and prerequisite for several other tasks, the development of technology oriented towards the classification task generally precedes the other tasks. And the cost of image-level labeling for classified task is lower than other tasks. Therefore, the applications oriented to classification task account for the vast majority, and applications oriented to object detection, image segmentation, target tracking, and other tasks decrease in turn.

4.1. Welding image classification

Welding image classification is to discover the category of objects contained in the welding image, which is a coarse-grained WIR task.

The classification task for PRW is mainly the recognition of the bevel shape and then the adjustment of the welding process based on the recognition results. Tian et al. [164] first used different laser sources to obtain weld bevel profile features and then designed a 10-layer CNN to classify the bevel types. In the reference [15], a 6-layer CNN was designed to identify the bevel type in response to strong noise interference in the field of view, and to narrow the parameter search for subsequent tasks accordingly. Cruz et al. [57] designed a 9-layer CNN to identify whether the parts to be welded are aligned.

The classification task for IW is mainly to identify the state of the melting zone, and the purpose is to provide a basis for the traceability and online control of welding quality. To better predict the melt depth, Jiao et al. [17] synthesised the current image with the previous 1/6 and 2/6 s images to reflect the dynamic welding phenomenon and built a 9-layer CNN to identify the image features. Zhang et al. [29] used different data augmentation methods for different welding defects and designed a 10-layer CNN to achieve classification of the three penetration states. Miao et al. [33] designed a two-stage detection strategy to improve the efficiency of CNN-based defect recognition for welding image acquired by TFT. The references [124,131] introduced temporal information in the identification of weld states and convert disturbing information for a single image into useful information to support model decisions. Liu et al. [137] found that the molten pool features extracted by the convolution kernel were redundant through feature map visualization, so it was integrated into the LSTM network to adaptively fuse the redundant features. In order to learn richer molten pool features, Liu et al. [156] proposed a coarse-grained parameter regularization method, taking into account that the weight parameters in the convolution kernel have local correlations, which facilitated the convolutional kernel to be trained in the direction of increasing variance. Liu et al. [16] addresses

the problem of small visual differences between different penetration states by designing a label semantic attention that uses discriminative inter-label texts to guide the model in learning discriminative visual features. In the reference [165,351]-dimensional features were first extracted manually, followed by a stacked sparse autoencoder to establish a mapping relationship between the weld state and the input features. Xue et al. [81] used Mobilenet as the skeleton network, which greatly reduces the model size and offers the possibility of practical deployment of the model. The references [138,166] have improved the explainability of models through model decision basis visualization and feature map visualization respectively.

The classification task for POW focuses on the identification of external or internal defects in the weld seam to enable intelligent quality inspection task. In the reference [167], three resampling methods were designed and compared for the problem of defect sample imbalance in weld radiographic image recognition. Jiang et al. [168] designed four pooling methods for different weld defect types and different regions in the image. Kumaresan et al. [169] used a TL strategy to address the problem of a small sample of weld seam radiographic image. Miao et al. [170] devised a two-stage identification method based on the fact that defects in weld appearance do not occur frequently, i.e. identifying the presence or absence of defects and then identifying the specific type of defect. In the reference [36], a one-dimensional signal was first acquired using the acoustic emission technique, then a two-dimensional image was obtained by TFT, and then a CNN was built to identify whether the welded steel pipe was subject to a natural gas leak. Zhang et al. [171] designed a 7-layer CNN and classified the weld microstructure images corresponding to the three defects. Feng et al. [132] firstly learned the difference between normal and defect samples through comparative learning, and secondly filtered the same features and extracts the different features of the image pairs by means of the spatial correlation AM, which provides an effective attention for learning the differences between images of different armature welding surfaces. Hu et al. [172] adopted ELU activation function and grayscale adaptive pooling strategy in the task of X-ray defect recognition. To address the small sample problem faced by weld radiographic defect detection tasks, Di and Li [173] investigated the effect of freezing different layers on transfer learning methods. Zhang et al. [174] weighted the loss function for the problem of imbalanced defect categories in weld radiographic images. To address the problem of low quality of X-ray images, Liu et al. [175] firstly embed defect images into local and global feature spaces, and secondly maintain the consistency of the feature space based on a new consistency strategy.

In summary, the main application scenarios of welding image classification include the identification of bevel types, melt zone states and weld defects. The current research mainly focuses on issues such as noise interference, lack of welding feature information, insufficient data volume, small visual differences, high calculation costs, and poor explainability. However, the current research has been modeling DLBWIR as a single-label classification, and the actual welding image often contains a

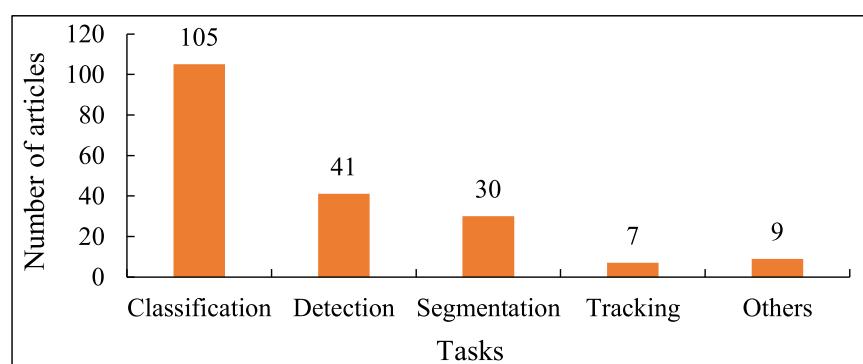


Fig. 13. Posting volume for different DLBWIR tasks.

variety of label categories. Therefore, the authors believe that the accuracy and interpretability of the classification model can be improved by modeling DLBWIR as a multi-label classification task.

4.2. Welding object detection

Object detection is to determine the category and location of objects

Table 8
Application of DL in welding image object detection.

Ref.	Year	Stage	Issue and challenge			Method and result			Data sources, volumes and metrics
			Engineering requirement	Visual task	Difficulty and challenge	Backbone	Key elements		
[176]	2018	IW	Welding quality online monitoring	Laser spot detection	Incomplete laser spot image, large field of view and complex calculation.	Not mentioned	Two-stage detection of first coarse and then fine	Laboratory, 400 and Acc of 97.96 %	
[177]	2021		Prediction of weld forming quality	Back-bead width prediction	Weld width is difficult to measure directly	Lenet5	Simultaneous acquisition of front and back images of the molten pool. Define the width detection on the back of the weld as a regression problem.	Laboratory, 28494 and MSE of 0.096	
[178]	2021		Welding quality online monitoring	The monitoring of weld width and reinforcement	Global information is difficult to extract	Resnet	Non-local attention	Laboratory, DE of 0.1805 mm	
[34]	2021		Welding quality online monitoring	Weld defect detection	Lack of dataset	YOLO	Multiple dataset construction methods	Simulation, 3892 and MAP of 85 %	
[179]	2021		Laser beam absorbance monitoring	Top and bottom diameter detection of keyhole	The laser beam absorptance inside a keyhole varies rapidly	YOLO	Dual vision sensing	Laboratory, 240 and GOF of 99.76 %	
[119]	2021		Welding quality online monitoring	Penetration status recognition	Arc interference and small sample size	Faster R-CNN	Dual vision sensing and TL	Laboratory, 2832 and Acc of 95.58 %	
[19]	2020		Safety monitoring of welding work	Welding protective equipment detection	The object scale is different and the model is complex	SSD	Multi-scale feature fusion and DSC	Production, 10000 and MAP of 87.45 %	
[180]	2022		Welding quality online monitoring	Penetration status recognition	The penetration state changes dynamically and is difficult to measure directly	Lenet5-LSTM	Fusion of timing information	Laboratory, 28560 and DE of 3 mm	
[35]	2020	POW	Automatic NDT	Weld line detection	Strong noise and complex model	SSD	Parallel structured light and DSC	Laboratory, 10000 and DE of 1 mm	
[181]	2021		Connector quality detection	Solder joint detection	Small sample size, small object and the model is complex	Faster R-CNN	DA, TL and the region clustering-based anchor box generation	Laboratory, 1654 and MAP of 94.1 %	
[84]	2021		Body welding quality control	Solder joint detection	Strong interference, small object, small sample size and the model is complex	YOLO	Feature pyramid, DA and DSC	Production, 400 and MAP of 89.32 %	
[182]	2022		Automatic NDT	Weld defect detection	Low contrast and strong noise	YOLO	Contrast-limited adaptive histogram equalization, non-local mean denoising and GIOU loss	Production, 256 and MAP of 51.2 %	
[91]	2022		Weld seam polishing	Weld seam detection	Small sample size	Faster R-CNN	DA and TL	Laboratory, 5000 and Acc of 91.68 %	
[78]	2022		Automatic NDT	Weld defect detection	Small sample size	YOLO	Stick-based DA	Public dataset, 3000 and MAP of 79.33 %	
[183]	2022		Automatic NDT	Weld defect detection	Small object	Faster R-CNN	Bidirectional feature pyramid	Laboratory, MAP of 85.5 %	
[184]	2022		Quality inspection of body welded joints	Weld defect detection	Fine-grained	YOLO	Label smoothing	Laboratory, 1100 and MAP of 90.1 %	
[185]	2021		Rail quality detection	Flash butt weld detection	Small object	YOLO	Multi-scale feature fusion	Production, 4529 and MAP of 87.41 %	
[61]	2021		Steel pipe quality detection	Weld defect detection	Small sample size and real-time	YOLO	Multiple DA methods and single stage detection	Production, 30672 and MAP of 98.7 %	
[186]	2019		PCB board quality detection	Solder joint detection	Simultaneous multiple recognition tasks with small sample size	Mask R-CNN	Mask R-CNN and TL	Laboratory, 88 and MAP of 97.4 %	

in the welding image at the same time, which is a fine-grained WIR task. In view of the fact that the regression task can obtain specific values, this section considers the regression problem as an object detection task and summarizes it.

Lin et al. [176] first used a CNN to initially localize the laser spot area, then increased the optical magnification of the camera and used template matching for high precision localization. Wang et al. [177] first

used CNN to extract the visual features of the molten pool, and then established the correspondence between the visual features and the width of the back weld to achieve the key feature detection of the molten pool object. In the reference [178], after CNN extracted the visual features of the molten pool, a regression method was used to simultaneously detect the width and reinforcement of the deposited layer. In the reference [34], six auxiliary datasets were created for the small sample problem, and both porosity and slag defects were effectively detected based on YOLOv4. Kim et al. [179] collected the images of the top keyhole and the bottom aperture simultaneously, and then measured the aperture size based on YOLOv4, and finally established the relationship between the aperture size and the absorption rate of the laser beam. A dual-input Faster R-CNN model with infrared thermal and visual images as input, including a simultaneous feature extraction module, a convolutional descriptor selection module and a synthetic feature-based recognition module, was developed in the reference [119]. Lu et al. [19] used an idea based on multi-scale feature fusion to improve the SSD structure and apply it to welders' behaviour detection. Yu et al. [180] developed a CNN-LSTM model and extracted sufficient penetration information from the molten pool image sequence.

Dong et al. [35] combined Mobilenet and SSD structures to achieve lightweight detection of weld lines. In the reference [181], a clustering-based anchor generation method was proposed and combined with Faster R-CNN to achieve connector weld joint defect detection. Dai et al. [84] not only used a lightweight Mobilenet as the backbone of YOLOv3, but also designed a feature pyramid with cross-scale connections and employed a complete IOU loss to speed up model convergence and applied it to quality detection of resistance spot welds. Aiming at the non-derivable problem of the IOU loss function, Yun et al. [182] first enhanced the X-ray image through contrast-limited adaptive histogram equalization, and secondly used GIOU as the loss function for training the YOLO model. Liu et al. [91] used DA and TL methods to solve the problem of insufficient weld appearance image data, and compared the effects of TL based on different backbone. Mery et al. [78] designed a DA method called Stick. For the problem of small defect targets in X-ray images, Chen et al. [183] adopted a bidirectional fusion feature pyramid structure. Li et al. [184] adopted the label smoothing technique for the problem of small variation in the features of appearance defective. In the reference [185], a minimum scale prediction network was designed to improve the detection accuracy of the model for small-scale objects and applied to the detection of welding defects in railway tracks. In order to identify multiple defects in the weld seam of steel pipes, Yang et al. [61] used a single-stage inspection structure YOLOv5 to detect 8 types of defects. Wu et al. [186] first pre-trained the skeleton network Resnet101 on the COCO dataset, and then detected the location of solder joint defects based on Mask R-CNN. Table 8 summarizes the key elements of DL in the welding image object detection tasks. The identification of key elements is based on the targeted approach used by the literature to address the difficulty and challenge, the innovation point in the literature, and the associated point with the theme of this paper.

In summary, the main application scenarios of welding object detection include the detection of process quality, forming quality, weld position and defect location. In addition to the common problems similar to classification task, the current research also includes small object and cross-scale issues. However, although the current object detection tasks are modeling and positioning the labels and positions of multiple objects in the welding image, the relationship between objects is not considered. Therefore, the authors believe that the reliability of DLWIR can be improved through modeling relationship between multiple objects.

4.3. Welding image segmentation

Segmentation is the most refined WIR task, and its essence is a pixel-by-pixel classification task. Due to the pixel-by-pixel classification characteristic of segmentation task, the model inference speed is slow

while obtaining refined recognition results. Therefore, segmentation task is often used in quantitative detection of objects to be identified in POW images.

Yang et al. [187] designed an attention dense convolutional block and used pixel segmentation to extract accurate laser stripes in a strongly disturbed environment. Wang et al. [153] designed a segmentation network with a channel AM and a separable structure for extracting laser stripes.

Mi et al. [188] used dilated convolution to increase the receptive field and designed a semantic segmentation network to segment the molten pool and the splash. In the literature [52], a segmentation network Epnet was designed to calculate the molten pool width and then predict the forming width of the weld. Tan et al. [189] first divided the welding image into small blocks, and then designed CNN and threshold neural network to segment each small block.

Yang et al. [190] proposed a branching structure for encoding spatial and contextual features, respectively, for the small target and cross-scale problems of appearance defects. Ling et al. [191] designed a parallel network to obtain semantic differences for the problem of small and irregular targets of weld defects, and used the Lovasz-softmax loss function in the training process. Li and Li [192] proposed a high-frequency feature enhancement network for the problem of insignificant pixel variation and small area of defects on the welded steel surface. Pandiyan et al. [193] designed an encoding-decoding type CNN to realize the segmentation of four types of welds, which provides a basis for the optimization of robot grinding parameters. Liu et al. [194] proposed a fusion block to extract the most important features in the three dimensions of the normal map, and then constructed a CNN to segment the key points of the welded stud from the normal map. Oh and Ki [26] used the first generator to generate the weld segmentation map, and the second generator converted it into an OM image, which realized the direct prediction of the cross-sectional welding image from the laser welding process parameters. Due to the lack of image data of weld appearance, Yang et al. [24] segmented the weld area based on U-Net, and then merged with the workpiece image to achieve data augmentation. In order to segment laser welding defects of different sizes and shapes, Zhu et al. [195] proposed a high-precision and high-efficiency lightweight segmentation algorithm, which mainly includes a feature extraction module, an attention module, a localisation module and a boundary anti-aliasing module. Jang et al. [196] regarded Resnet as the backbone of the segmentation network, and realized the measurement of acicular ferrite content that had a significant impact on the mechanical properties of carbon steel. In the reference [197], a Mask R-CNN network for defect segmentation of weld X-ray images was constructed based on transfer learning for the problem of small sample size of weld X-ray images. Zhang et al. [198] firstly proposed a compressed U-Net to segment the weld pad and the weld joint, secondly a template based approach to confirm the validity of each weld joint and finally a heuristic algorithm to determine the defect type. In order to obtain more precise details of defects, Yang et al. [25] added cross-scale skip connections to U-Net. Wang and Shen [199] proposed a boundary-aware semantic segmentation method, which combined spatial pyramid pooling and spatial AM to improve the segmentation accuracy and robustness of the welding area. For the problem of complex background and low contrast of X-ray images, Yang et al. [200,201] proposed an attention fusion module and a cross-scale connection method. Dong et al. [90] proposed a novel unsupervised local deep feature learning method based on image segmentation to build a network which can extract useful features from an image. Nowroth et al. [202] proposed an improved method combining dilated convolution and pyramidal pooling for segmenting OM images. Table 9 summarizes the key elements of DL in the application of welding image segmentation task.

In summary, the main application scenarios of welding object segmentation include segmentation of weld key points, melt zone objects, and defect objects. In addition to facing common challenges similar to

Table 9

Application of DL in welding image segmentation.

Ref.	Year	Stage	Issue and challenge			Method and result		
			Engineering requirement	Visual task	Difficulty and challenge	Backbone	Key elements	Data sources, volumes and metrics
[187]	2022	PRW	Weld seam tracking	Laser stripe segmentation	Strong interference, across scales and unbalance	U-Net	Pixel segmentation task, BiConvLSTM module and weighted loss	Laboratory, 192 and Acc of 99.4 %
[153]	2022		Weld seam tracking	Laser stripe segmentation	Strong interference and real-time	Erfnet	Channel attention and DSC	Laboratory, 737 and segmentation success rate of 96 %
[188]	2021	IW	Directed energy deposition monitoring	Segmentation of molten pool and spatter	Strong interference and real-time	PSPnet	Dilated convolution and pyramid pooling	Laboratory, 700 and Acc of 94.71 %
[52]	2021		Width monitoring of cladding layer in arc additive manufacturing	Segmentation of molten pool	Small sample size, real-time and difficulty in contour extraction	Enet	DA, DSC and pyramid pooling	Laboratory, 600 and Acc of 94.18 %
[189]	2020		Powder bed laser melting process monitoring	Segmentation of spatter	Strong interference and small sample size	VGG	Threshold neural network and processing in blocks	Laboratory, 210 and Rec of 80.48 %
[190]	2022	POW	Safety valve welding quality inspection	Segmentation of weld seam	Small sample size and across scales	Proposed	Spatial features, contextual features and feature fusion	Laboratory, 4708 and Mean IOU of 86.704 %
[191]	2022		PCB solder joint quality inspection	Solder joint defect segmentation	Small sample size, small object and irregular shapes	Deeplab	Parallel network and Lovasz-softmax	Laboratory, 340 and Acc of 96.78 %
[192]	2022		Surface defect detection of welded rebar	Segmentation of welding seam defect	Defects are insignificant and small in size	U-Net	High-frequency components of low-level features	Production, 1580 and Mean IOU of 86.2 %
[193]	2019		Robot polishing endpoint detection	Segmentation of weld seam	Strong interference, unbalance and robustness	Segnet	DA, class-weighted loss and regularization	Laboratory, 2000 and Acc of 99.35 %
[194]	2020		Welding stud detection	Segmentation of stud key point	There are noise spots, missing spots, and the need to measure multiple studs simultaneously.	FCN	Channel separation and feature fusion	Laboratory, 1427 and Rec of 0.98 %
[26]	2020		Weld microstructure detection	Segmentation of weld section	Small sample size and complex OM image content	FCN	CGAN and two-stage image generation	Laboratory, 1638 and GOF of 93.6 %
[24]	2020		Weld position detection	Segmentation of weld seam	Complex and diverse industrial environment	U-Net	DA based on segmented image fusion	Laboratory, 1940 and Acc of 96.2 %
[195]	2021		Welding defect detection of battery cover safety hole	Segmentation of welding defect	Defect size and shape are variable and require real-time	U-Net	Multi-scale AM and DSC	Production, 7263 and Mean IOU of 84.67 %
[196]	2020		Acicular ferrite content measurement	Segmentation of weld microstructure	Small sample size and imbalance	FCN	DA and category sensitive loss	Laboratory, 144 and Mean IOU of 86.2 %
[197]	2018		Automatic NDT	Segmentation of welding seam defect	Simultaneous multiple recognition tasks with small sample size	Mask R-CNN	Mask R-CNN and TL	Public dataset, 88 and MAP of 85 %
[198]	2020		Quality inspection of laser solder joints in battery production	Segmentation of solder joint and solder pad	Small sample size and real-time	U-Net	DA and network compression	Laboratory, 37,877 and Acc of 99.85 %
[25]	2021		Automatic NDT	Segmentation of welding seam defect	Small sample size and across scales	U-Net	DA and skip connections	Public dataset, 20 and Acc of 99.8 %
[199]	2020		Automatic NDT	Segmentation of welding seam defect	Small sample size, strong interference and variable object shape	U-Net	DA, AM and atrous spatial pyramid pooling	Production, 300 and Acc of 93.7 %
[200]	2022		Automatic NDT	Segmentation of welding seam defect	Complex backgrounds, poor contrast, weak texture, and class imbalance	U-Net	Attention fusion block and cross-scale connection	Public dataset, 20 and Acc of 99.8 %
[90]	2021		Automatic NDT	Segmentation of welding seam defect	Difficulty in labeling data	U-Net	Unsupervised	Production and simulation, 24,111 and Acc of 89.55 %
[202]	2022		Weld microstructure detection	Segmentation of weld section	Small sample size and across scales	Deeplab	TL, dilated convolution and pyramid pooling	Laboratory, 282 and Mean IOU of 76.88 %

classification and detection tasks, this application is particularly affected by the difficulty of pixel-level labelling and the difficulty of segmenting pixels near boundaries. Therefore, the authors suggest that weakly supervised welding image segmentation based on image-level labelling information could be investigated in the future. In addition, the hard-to-segment samples can be mined explicitly by modelling the contextual relationships of the boundary pixels.

4.4. Welding object tracking

In the object tracking task of the welding image, the current

approach is to first detect the area of the target, and then calculate the object position to achieve the tracking task [203]. The end-to-end DL structure has not been applied, such as [204,205]. Furthermore, the current research focuses on the seam tracking task of PRW.

Zou et al. [206] first input the welding image into the convolutional long and short-term memory network, and then detected the welding feature points through the similarity matching between multiple feature layers. Zou and Zhou [86] firstly used CNN to learn welding image features, and secondly proposed a weld feature point detection method based on multi-correlation filter collaboration considering the continuity of feature points in adjacent frames. Zou et al. [87] used the SSD

model to learn global and local weld features and combined it with RNN to learn contextual information of visual features to achieve weld tracking under strong noise interference. Xiao et al. [207] used RCNN to identify the weld type and detect the laser stripe ROI, then designed different centre extraction algorithms for different types of welds and proposed a priori model to ensure the stability of the algorithm. In order to identify laser stripes and track welds in a strong interference environment, Zhang et al. [208] designed a full convolutional network based on the idea of segmentation. Li et al. [209] first used Mask R-CNN to detect the weld position and then extracted the weld line to assist the robot in its inspection. Table 10 summarizes the key elements of DL in the welding object tracking task.

In summary, the main application scenario for welding object tracking is the weld seam tracking task by capturing key points/objects in the area to be welded, and the current research focuses on the problems of strong arc light interference and small samples. As far as application scenarios are concerned, the authors believe that the task can be applied to the object tracking scenario of the IW stage. In terms of research points, the authors believe that tracking multiple objects in welding image will be an interesting direction.

4.5. Others

Due to the difficulty of dataset construction, the DL method is less used in subdivided WIR tasks.

Wang et al. [92] designed a multi-scale feature fusion semantic segmentation network Res-Seg, which detected the edge of the molten pool with good continuity, and laid the foundation for the calculation of width for molten pool and weld seam. Wang et al. [130] designed a generative network to predict the future molten pool image, and then input it into the regression network to guide weld reinforcement information. In order to avoid extensive analytical calculations, Li et al. [23] used a LSTM network to establish a mapping mechanism from 2D imaging points in the imaging plane to 3D reflection points on the molten pool surface, enabling data-driven 3D reconstruction of the molten pool surface. In the reference [31], a transfer learning method based on multi-objective instance weighting was designed to detect welding quality from the perspective of equipment fault diagnosis and was validated with robotic spot welding tests. In order to know the crystallographic orientation during welding, Singh et al. [210] established a generative model from ultrasonic wave travel times to crystallographic orientations. Mirzapour et al. [211] proposed a fast and flexible denoising CNN for the denoising problem of weld radiographic images. In order to perform NDT with a low-cost method, Rohkohl et al. [212]

generate computed tomography results from eddy current testing results. Aiming at the problem of industrial fakes caused by repeated imaging in the field of NDT, Gao et al. [213] proposed an X-ray image similarity comparison algorithm based on Siamese neural network. Table 11 summarizes the scenarios and key elements of DL in other subdivided WIR tasks.

4.6. Brief summary

The application of DL in welding image classification, object detection, image segmentation and target tracking has realized the instantiation of DLBWR. However, different task types have different applicable scenarios. Table 12 summarizes the applicable scenarios for different task types and the pros and cons of the application process. In general, DL mainly faces challenges at three levels of data (small sample, imbalance), feature (strong interference, fine-grained) and model (real-time, explainability) in various WIR tasks, and related solutions are developed around the key technologies mentioned in Section 3. However, related research focuses on the application of DL methods in WIR, which lacks the guidance of welding domain knowledge.

5. Public datasets for DLBWR

Welding image datasets are the basis for DLBWR research. The publication of large datasets such as ImageNet [214] for classification tasks, COCO [215] or object detection tasks, SBD [216] for segmentation tasks, and LaSOT [217] for tracking tasks is an important reason for the rapid development of DL in the CV field. Table 13 summarizes the public datasets that exist in the field of DLBWR. However, the public datasets of DLBWR still have the following shortcomings: (a) small amount of data; (b) few label categories; (c) unbalanced data; (d) lack of datasets for subdivided tasks.

6. Future perspectives

6.1. Perspectives from the literature

Approximately 50 % of the examined 192 articles gave a clear research outlook. As shown in Fig. 14, this section summarizes their relevant future perspectives in terms of the DLBWR paradigm as well as system integration in the following six areas.

6.1.1. Sensing

Visual sensing has the advantage of containing a wealth of

Table 10
Application of DL in welding image object tracking.

Ref.	Year	Stage	Issue and challenge			Method and result		
			Engineering requirement	Visual task	Difficulty and challenge	Backbone	Key elements	Data sources, volumes and metrics
[206]	2020	PRW	Weld seam tracking	Welding line feature point detection	Strong interference	Faster R-CNN	Fusion of timing information and periodically initialize network	Laboratory, 46819 and TE of 0.5 mm
[86]	2019		Weld seam tracking	Welding line feature point detection	Strong interference and small sample size	VGG	Fusion of timing information, correlative filtering and TL	Laboratory, 0 and TE of 0.25 mm
[87]	2021		Weld seam tracking	Welding line feature point detection	Strong interference and small sample size	SSD	Combine timing information, multi-scale features, DA and TL	Laboratory, 40000 and TE of 0.5 mm
[207]	2019		Weld seam tracking	Weld line center point detection	Strong interference and robustness	Faster R-CNN	Two-step extraction and type-related weld centre extraction method.	Laboratory, 4000 and TE of 0.28 mm
[208]	2019		Weld seam tracking	Weld line position detection	Strong interference and small sample size	FCN	TL and multi-scale feature fusion	Laboratory, 1000 and PRMSE of 3.87
[209]	2022	POW	Robot inspection	Weld seam position detection	Complex structure and non-salient objects	Mask R-CNN	Multi-scale feature fusion	Laboratory and production, 1500 and TE of 20 mm

Table 11

Application of DL in other WIR tasks.

Ref.	Year	Stage	Issue and challenge			Method and result		
			Engineering requirement	Visual task	Difficulty and challenge	Backbone	Key elements	Data sources, volumes and metrics
[92]	2020	IW	Weld width prediction	Edge detection of molten pool	Small sample size and arc interference	FCN	DA and multi-scale feature fusion	Laboratory, 1100 and TE of 0.2 mm
[130]	2020		Weld reinforcement prediction	Generation of molten pool image	Non-salient features	Prednet	Combine timing information and channel attention	Laboratory, 25830 and SS of 93.27 %
[23]	2021		3D reconstruction of molten pool	2D imaging points are mapped to 3D reflection points	Strong interference, small sample size and difficult to measure directly.	LSTM	Combine timing information and train the model based on the simulation data	Laboratory and simulation, 5000 and MSE of 0.002
[31]	2021		Welding product fault diagnosis	Robot spot welding defect diagnosis	Small sample size and transferability issue	Resnet	Case-based TL	Laboratory, 2160 and Acc of 98.33 %
[210]	2022		Prediction of weld tissue structure	Image generation	Determining the microstructure of weld seam by NDT methods	DNN	Build a generative model from ultrasonic wave travel times data to crystallographic orientation.	Simulation, 10000 and within 3°
[211]	2021	POW	Weld radiographic image recognition	Image denoising	Image is blurry and details are not noticeable	FFDnet	Selective image enhancement	Public dataset and production, DWIQL of 96 %
[212]	2022		Automatic NDT	Image generation	Comprehensive and inexpensive online inspection	U-Net	Generate high-cost test results from low-cost test results	Laboratory, 294 and Acc of 93.7 %
[213]	2022		Fake film identification	Image similarity comparison	Small sample size and larger image resolution	VGG	Comparisons by block	Production, 342 and Acc of 93.458 %

Table 12

Comparison of applications of DLBWIR in different tasks.

Task	Application scenario	Pros	Cons
Classification	Recognition of groove type, penetration state and weld defect type.	Welding image annotation is simple and wide range of applications.	Unable to achieve quantitative recognition and poor explainability.
Detection	Recognition of the object location of melting zone and welding defects.	Preliminary quantitative recognition can be achieved to meet the needs of automated detection.	Welding image is difficult to label by object and the number of objects and backgrounds is not balanced.
Segmentation	Quantitative recognition of melting zone objects and weld defect objects.	It can realize refined recognition and provide quantitative basis for welding process optimization. It has strong explainability.	Pixel-by-pixel annotation of welding image is difficult. Model inference time is long.
Tracking	Bevel and weld seam key point tracking.	It can realize the recognition of moving object.	Existing research achieves object tracking through frame-by-frame detection, which is computationally expensive. The scope of application is limited.
Others	Welding image denoising, generation, and edge detection.	It can solve the subtask requirements in the system task.	Welding image annotation is difficult and has limited application.

information, however, there are some features that cannot be obtained by visual sensing, such as information about defects inside the weld seam during the welding process. Although the references [32,223] launched a preliminary study on multi-modal sensing, it had only been a simple fusion of multi-modal information at the classification stage and lacked an in-depth development of MML. Therefore, Jiang et al. [119] decomposed this problem into three sub-problems: 1) selecting

multi-modal data types, 2) simultaneous multi-modal data acquisition, and 3) multi-modal data alignment.

6.1.2. Dataset

DL models have more parameters to be trained than traditional machine learning models. Increasing the amount of data in the welding image not only prevents overfitting of the DL model, but also improves the generalization ability of the model. The references [17,164,224,225] suggest that future research will require the creation of large-scale, high-quality welding image datasets. This includes extending the label categories of the welding image data, extending the test conditions to obtain comprehensive welding image data, and extending the size of the dataset based on both. In addition, Dai et al. [94] identified few-shot learning and zero-shot learning as future research directions for DLBWIR.

6.1.3. Model structure

A typical intelligent welding systems contains multiple modules and has very limited storage and computing resources that can be allocated to the vision module. Therefore, the references [195,219,226] consider small and fast DLBWIR models as the way forward, resulting in easy deployment at the edge and on mobile as well as real-time applications.

In addition, most of the current DLBWIR is a data-driven model and lacks the participation of expert knowledge. This not only causes a waste of expert knowledge, but also makes it difficult to move towards strong intelligent welding [227]. Therefore, the related literature believes that the model structure driven by data and knowledge will be an important research direction, such as adding manual features [28], physical and mathematical principles [228,229], and welding mechanism [230].

6.1.4. Training

Hyperparameters are crucial for the training, convergence and recognition ability of the model. However, most of the current literature is based on empirical selection of hyperparameters. The references [131, 167] believe that the hyperparameters involved in the DLBWIR model (such as learning rate, batch size, optimizer, etc.) need to be compared and optimized to obtain better recognition results.

In addition, the references [20,230] believe that the problem of imbalance training is often faced in WIR tasks. To solve this problem, it is necessary to design a class-related loss function to balance the problem of drift in the learning process caused by the sample size.

Table 13
Public datasets for DLBWIR.

Ref.	Year	Stage	Amount	Label	Task	URL
[187]	2022	PRW	192	Laser stripe	Segmentation	https://aistudio.baidu.com/aistudio/datasetdetail/106021/
[218]	2019	IW	33,254	Good weld, burn through, contamination, lack of fusion, misalignment and lack of penetration	Classification	https://www.kaggle.com/danielbacioiu/tig-aluminium-5083
[55]	2019		45,058	Good weld, burn through, contamination, lack of fusion, lack of shielding gas and high travel speed	Classification	https://www.kaggle.com/danielbacioiu/tig-stainless-steel-304
[219]	2021		4163	Short circuit and no short circuit	Classification	https://doi.org/10.17632/2nyjp89bf.1
[220]	2015	POW	88	Non-defective, lack of penetration, porosity, slag inclusion and crack	Detection and Segmentation	https://domingomery.ing.puc.cl/material/gdxray/
[171]	2019		162	Under-melt, beautiful-weld and over-melt	Classification	https://buffalo.box.com/s/13ccf2flyaqlzfcx0egy2vfwm0ym3sv
[193]	2019		2000	Four weld removal states and background	Classification and Segmentation	https://data.mendeley.com/datasets/2pcnt8kpw9/1
[198]	2020		37,877	Four tasks for welding pad and welding spot segmentation, three tasks for qualifying battery, obliquely placed welding pad, welded through welding spot and placed-over-highly electrode tab classification.	Segmentation and Classification	https://pan.baidu.com/s/1c82KVsjTieMp-pYCCn8l5w
[221]	2021		37	Good and defect	Classification	https://www.mdpi.com/2075-4701/11/4/535/s1
[222]	2021		297	Spatter, lack of fusion and normal	Segmentation	https://www.kaggle.com/ruthka/maskrcnn-mswellddefect
[61]	2021		3408	Blowhole, undercut, broken arc, crack, overlap, slag inclusion, lack of fusion, and hollow bead	Classification and Detection	https://github.com/huangyebiaoke/steel-pipe-weld-defect-detection/releases/tag/1.0
[191]	2022		340	Missing solder, missing proteus, solder short, overturned direction and solder ball.	Segmentation	https://pan.baidu.com/share/init?surl=_xLYqL8HwxcsBO3iowsyQ

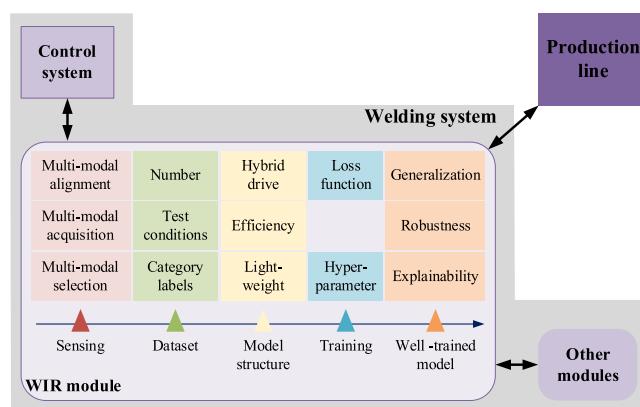


Fig. 14. Summary of research outlooks addressed in the literature.

6.1.5. Well-trained model

In the current research, most of the welding images are collected based on small-scale welding processes. Therefore, Nomura et al. [231] indicate that improving the generalization of the final model will be a research direction, that is, to ensure that the model has the generalization ability under variable operating conditions. Fioravanti et al. [232] believe that online learning methods should be explored to improve the generalization ability of the model. In addition, Yang et al. [61] argue that it is necessary to carry out few-shot learning research for the current situation of small samples of welding image, and automatically update the system as new samples are added to the design model.

Liu et al. [156] believe that many simplifications and idealization assumptions have been made in the current DLBWIR research, such as welding image data from the same distribution, and the model will not be disturbed and attacked. This will lead to insufficient flexibility and robustness of our algorithm [233,234]. In the future, we should model more general situations and conduct attack tests.

Liu et al. [131] argue that the current DLBWIR is a black box model, and people have no way of knowing the decision basis and reliability of the model, which limits the application of DLBWIR in safety-sensitive tasks such as aerospace and autonomous driving. Therefore, improving the explainability of the model is essential for developers to optimize the model and increase the confidence of users.

6.1.6. Integration

The WIR module does not exist independently, but exists in the robot welding system, the robot welding production cell, and the manufacturing system. Therefore, Nogay and Akinci [235] consider the DL model as one of the modules in the robot welding cell, and emphasize its horizontal integration with other modules such as motion modules. Wang et al. [236] emphasize the vertical integration between DLBWIR module and control system. Cruz et al. [57] and Chu et al. [237] emphasize the end-to-end integration of existing models with production lines and even information systems through internet of things to achieve interconnectivity.

6.2. Perspectives from the authors

6.2.1. Public datasets and pre-trained models

Data is the source of power for DL. While the importance of datasets is also addressed in the DLBWIR literature and the creation of large, high-quality datasets is identified as future work, there is a lack of emphasis on the openness of datasets. From the history of DL in the field of natural images, one can easily see the important role played by public datasets in driving the development. Therefore, it is believed that it will be an important task to establish an open welding image data set. Furthermore, in order to build a reliable welding image dataset, the authors believe that three aspects should be emphasized. In the dataset design stage, both its practical values in solving domain problems and its academic merits in advancing research have to be considered holistically. In the image acquisition stage, raw images containing rich information about the quality of the weld should be acquired through the use of auxiliary light, filters and other ancillary equipment. In the image labelling stage, several experts should simultaneously label and vote on the final label to improve the accuracy and consistency of the label. A high-quality dataset should be characterized by a large amount of data, a balanced sample size for each category, a large number of label categories, and a high label confidence, to satisfy the task of refined recognition.

It is extremely difficult to establish a large data set for each subdivided scenario. The idea in the CV field for this problem is to pre-train the model structure on a large public dataset, and then fine-tune the weights learned by the model based on the data in the target domain. The pre-training model has played an important role in promoting image recognition tasks in the subdivision field. Therefore, it is not only

necessary to establish a public dataset in the field of welding image, but also to pre-train a batch of classic models for WIR tasks in subdivision fields and teams that lack large-scale test conditions.

6.2.2. Semantic explainability for DLBWIR

The welding process is carried out under clear process guidance. The welding image contains rich welding semantic information, which is a carrier that reflects the welding state. However, DLBWIR is an end-to-end way to establish the association relationship between the welding image and the welding state. This black box model not only lacks the reflection of the decision-making basis of the model, but also lacks an explicit expression of the causal relationship between the welding image and the welding state. Although the references [82,137] used feature map visualization and class activation mapping methods to show users the features learned by DL and the basis for making decisions for specific examples, this explanation is a coarse-grained explanation oriented by human visual perception. This explanation neither include the job semantics in the welding process, nor the reasoning and revealing of the causal relationship of the welding status information [139]. Therefore, people still cannot intuitively capture the decision-making basis of the model through this visual explanation and optimize the welding process accordingly.

As a highly readable external knowledge carrier, the KG provides a way for the semantic expression of the welding manufacturing process [133]. Through KG, users can infer the associated path as the semantic explanation of the model decision. This type of association path not only expresses the semantic information of the entities and relationships in

the KG, but also helps developers discover the causal relationship between the welding image and the welding state. Thus, giving the DL model semantic explainability and causality reasoning capabilities, it provides a basis for developers to improve the model and optimize the welding process.

This part takes the radiographic image detection of welding seams of aerospace structural parts as an example to illustrate the semantic explainability and causal reasoning framework of DLBWIR based on KG. As shown in Fig. 15, it can be divided into three layers: *data layer*, *knowledge layer*, and *explanation layer*. The *data layer* mainly includes welding context data and multi-modal data. The multi-modal data in the process design stage mainly includes process cards and text documents. The multi-modal data in the simulation stage includes temperature field simulation data, stress field simulation data, and molten pool flow simulation data. The PRW stage mainly includes point cloud and image data. The IW stage includes weld pool image data, instantaneous current data, welding sound data, and so on. The multi-modal data in the POW stage includes weld radiographic image data, weld actual width data, weld metallographic image data, etc. After acquiring welding data, in the *knowledge layer*, knowledge extraction, knowledge representation, semantic alignment, knowledge fusion, and knowledge storage operations are performed on each modal data to form knowledge modules of each welding stage and synthesize to form the WDKG. In the *explanation layer*, there are two parts that can be explained by semantics and the basis for decision-making. Through causal reasoning technology, various path dependencies that produce various defects can be found, thereby forming semantic explainability.

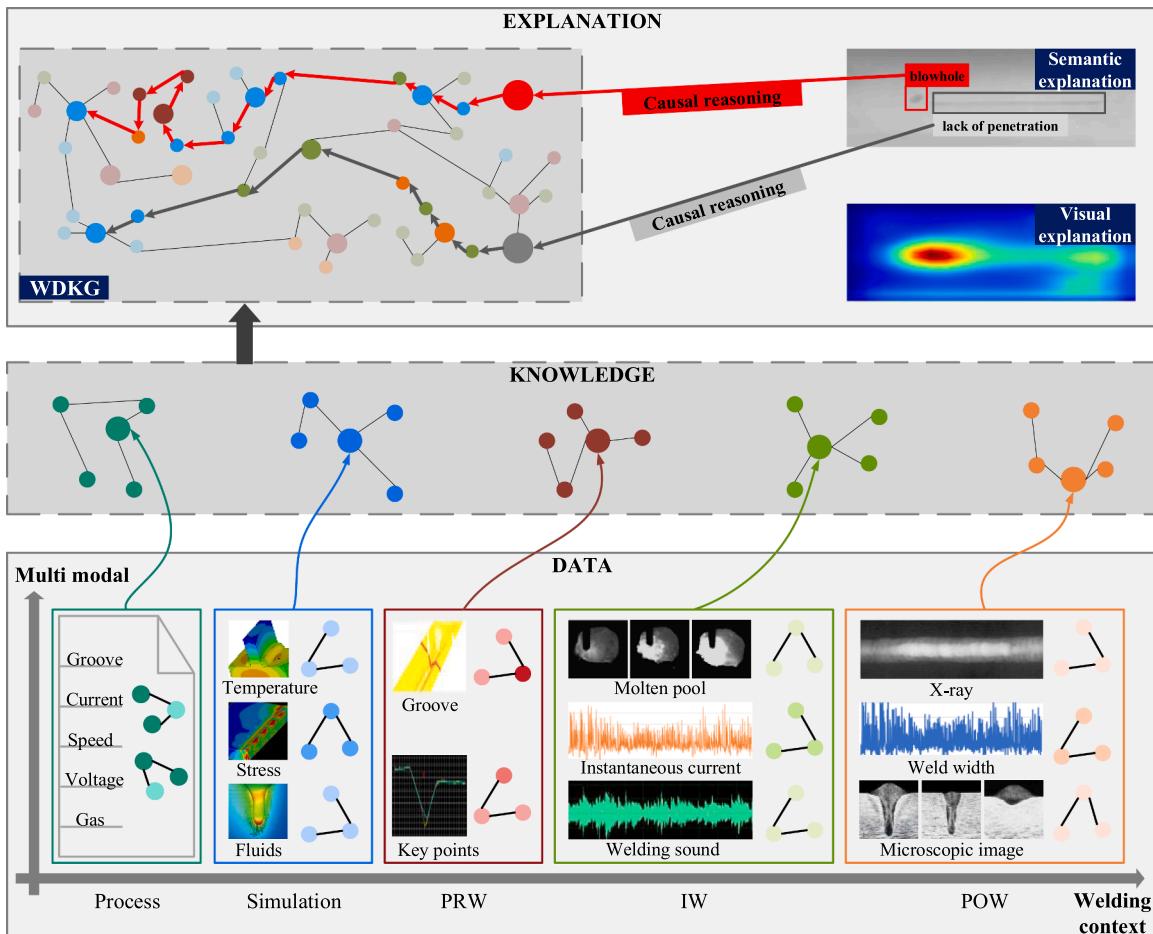


Fig. 15. KG-empowered semantic explanation for DLBWIR. The multimodal data involved in the welding context is integrated in the data layer. In the knowledge layer, the sub-KG corresponding to each welding stage is constructed. Each sub-KG is fused at the explanation layer, and the basis for explanation reasoning of welding defects is provided.

As shown in Fig. 15, for a weld seam, visual explanation can provide the basis for decisions made by the DL model. However, this description is imprecise and does not contain engineering semantics. The fusion of knowledge modules formed by welding contexts can lead to semantic explanation. In terms of semantic explanation, if the DL-based target detection algorithm detects porosity defects, it can correspond to semantic knowledge such as stress concentration, impure shielding gas, too much welding current, etc. Suppose the DL-based target detection algorithm detects an unperforated defect. In that case, it can correspond to the semantic knowledge of low molten pool temperature, too small bevel angle, too small welding current, too large welding speed, etc. The explanation formed by semantic association of the causes of welding defects is for front-line operators, and its essence is the semantic knowledge formed by in-depth annotation of multi-stage and multi-modal information.

7. Conclusion

DLBWIR is becoming a core technology to provide forecasting and decision making in intelligent welding systems. This paper reviewed a total of 192 previous publications to summarize the state-of-the-art of the DLBWIR research and application. The main contributions of this work are summarized as follows:

- 1) The key enabling technologies for the DLBWIR, such as sensing, data, model, training, and platform were outlined. Moreover, the characteristics and scope of application of the relevant technologies were summarized.
- 2) The current applications of DL in different WIR tasks are summarized, based on which, the most challenging is the lack of guidance from welding domain knowledge.
- 3) Two promising research works, namely opening datasets and semantic explainability are highlighted, which are the key to promote the in-depth research and practical application of DLBWIR.

Despite the rapid growth, DLBWIR remains a rapidly evolving field. Many pressing issues should be addressed to enhance its viability in practice. In that regard, we hope this paper can provide insightful knowledge and guide to attract more researchers in addressing the future directions of the DLBWIR research and application.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research is partially funded by the Mainland-Hong Kong Joint Funding Scheme (MHX/001/20), Innovation and Technology Commission (ITC), Hong Kong Special Administration Region of the People's Republic of China, National Key R&D Programs of Cooperation on Science and Technology Innovation with Hong Kong, Macao and Taiwan (SQ2020YFE020182), Ministry of Science and Technology (MOST) of the People's Republic of China, and the Centrally Funded Postdoctoral Fellowship Scheme (1-YXBM), The Hong Kong Polytechnic University.

References

- [1] Luo Z, Dai JS, Wang C, Wang F, Tian Y, Zhao M. Predictive seam tracking with iteratively learned feedforward compensation for high-precision robotic laser welding. *J Manuf Syst* 2012;31(1):2–7. <https://doi.org/10.1016/j.jmsy.2011.03.005>.
- [2] Zhang Y, Zhao Z, Zhang Y, Bai L, Wang K, Han J. Online weld pool contour extraction and seam width prediction based on mixing spectral vision. *Opt Rev* 2019;26(1):65–76. <https://doi.org/10.1007/s10043-018-0479-3>.
- [3] Yan Z, Xu H, Huang P. Multi-scale multi-intensity defect detection in ray image of weld bead. *NDT E Int* 2020;116:102342. <https://doi.org/10.1016/j.ndteint.2020.102342>.
- [4] Fan X, Gao X, Liu G, Ma N, Zhang Y. Research and prospect of welding monitoring technology based on machine vision. *Int J Adv Manuf Tech* 2021;115: 3365–91. <https://doi.org/10.1007/s00170-021-07398-4>.
- [5] Sahoo S, Lo CY. Smart manufacturing powered by recent technological advancements: a review. *J Manuf Syst* 2022;64:236–50. <https://doi.org/10.1016/j.jmsy.2022.06.008>.
- [6] Wang J, Xu C, Zhang J, Zhong R. Big data analytics for intelligent manufacturing systems: a review. *J Manuf Syst* 2022;62:738–52. <https://doi.org/10.1016/j.jmsy.2021.03.005>.
- [7] Cai W, Wang J, Jiang P, Cao L, Mi G, Zhou Q. Application of sensing techniques and artificial intelligence-based methods to laser welding real-time monitoring: A critical review of recent literature. *J Manuf Syst* 2020;57:1–18. <https://doi.org/10.1016/j.jmsy.2020.07.021>.
- [8] 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. <https://doi.org/10.1016/j.jmsy.2020.06.020>.
- [9] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput* 2006;18(7):1527–54. <https://doi.org/10.1162/neco.2006.18.7.1527>.
- [10] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Commun ACM* 2017;60(6):84–90. <https://doi.org/10.1145/3065386>.
- [11] Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Driessche GVD, et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016;529 (7587):484–9. <https://doi.org/10.1038/nature16961>.
- [12] Wang J, Ma Y, Zhang L, Gao RX, Wu D. Deep learning for smart manufacturing: methods and applications. *J Manuf Syst*, 48(C); 2018. p. 144–56. <https://doi.org/10.1016/j.jmsy.2018.01.003>.
- [13] Günther J, Pilarski PM, Helfrich G, Shen H, Diepold K. Intelligent laser welding through representation, prediction, and control learning: an architecture with deep neural networks and reinforcement learning. *Mechatronics* 2016;34:1–11. <https://doi.org/10.1016/j.mechatronics.2015.09.004>.
- [14] Zhang Y, Yang YP, Zhang W, Na SJ. Advanced welding manufacturing: a brief analysis and review of challenges and solutions. *J Manuf Sci Eng* 2020;142(11): 110816. <https://doi.org/10.1115/1.4047947>.
- [15] Du R, Xu Y, Hou Z, Shu J, Chen S. Strong noise image processing for vision-based seam tracking in robotic gas metal arc welding. *Int J Adv Manuf Tech* 2019;101 (5):2135–49. <https://doi.org/10.1007/s00170-018-3115-2>.
- [16] Liu T, Bao J, Zheng H, Wang J, Yang C, Gu J. Learning semantic-specific visual representation for laser welding penetration status recognition. *Sci China Technol Sc* 2021;1–14. <https://doi.org/10.1007/s11431-021-1848-7>.
- [17] Jiao W, Wang Q, Cheng Y, Yu R, Zhang Y. Prediction of weld penetration using dynamic weld pool arc images. *Weld J* 2020;99(11):295–302. <https://doi.org/10.29391/2020.99.027>.
- [18] Li C, Wang Q, Jiao W, Johnson, Zhang Y. Deep learning-based detection of penetration from weld pool reflection images. *Weld J* 2020;99(9):239–45. <https://doi.org/10.29391/2020.99.022>.
- [19] Lu H, Li C, Chen W, Jiang Z. A single shot multibox detector based on welding operation method for biometrics recognition in smart cities. *Pattern Recogn Lett* 2020;140:295–302. <https://doi.org/10.1016/j.patrec.2020.10.016>.
- [20] Knaak C, von Eßen J, Kröger M, Schulze F, Abels P, Gillner A. A spatio-temporal ensemble deep learning architecture for real-time defect detection during laser welding on low power embedded computing boards. *Sens Basel* 2021;21(12): 4205. <https://doi.org/10.3390/s21124205>.
- [21] Jin C, Shin S, Yu J, Rhee S. Prediction model for back-bead monitoring during gas metal arc welding using supervised deep learning. *IEEE Access* 2020;8: 224044–58. <https://doi.org/10.1109/ACCESS.2020.3041274>.
- [22] Lee J, Noh I, Lee J, Lee SW. Development of an explainable fault diagnosis framework based on sensor data imagination: a case study of the robotic spot-welding process. *IEEE Trans Ind Inf* 2022;18(10):6895–904. <https://doi.org/10.1109/TII.2021.3134250>.
- [23] Li L, Cheng F, Wu S. An LSTM-based measurement method of 3D weld pool surface in GTAW. *Measurement* 2021;171:108809. <https://doi.org/10.1016/j.measurement.2020.108809>.
- [24] Yang L, Fan J, Liu Y, Li E, Peng J, Liang Z. Automatic detection and location of weld beads with deep convolutional neural networks. *IEEE Trans Instrum Meas* 2021;70:1–12. <https://doi.org/10.1109/TIM.2020.3026514>.
- [25] Yang L, Wang H, Huo B, Li F, Liu Y. An automatic welding defect location algorithm based on deep learning. *NDT E Int* 2021;120:102435. <https://doi.org/10.1016/j.ndteint.2021.102435>.
- [26] Oh S, Ki H. Cross-section bead image prediction in laser keyhole welding of AISI 1020 steel using deep learning architectures. *IEEE Access* 2020;8:73359–72. <https://doi.org/10.1109/ACCESS.2020.2987858>.
- [27] Yan Y, Liu D, Gao B, Tian G, Cai Z. A deep learning-based ultrasonic pattern recognition method for inspecting girth weld cracking of gas pipeline. *IEEE Sens J* 2020;20(14):7997–8006. <https://doi.org/10.1109/JSEN.2020.2982680>.
- [28] Wu D, Hu M, Huang Y, Zhang P, Yu Z. In situ monitoring and penetration prediction of plasma arc welding based on welder intelligence-enhanced deep random forest fusion. *J Manuf Process* 2021;66:153–65. <https://doi.org/10.1016/j.jmapro.2021.04.007>.
- [29] 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. <https://doi.org/10.1016/j.jmapro.2019.06.023>.
- [30] Yu R, He H, Han J, Bai L, Zhao Z, Lu J. Monitoring of back bead penetration based on temperature sensing and deep learning. Measurement 2022;188:110410. <https://doi.org/10.1016/j.measurement.2021.110410>.
- [31] Lee K, Han S, Pham VH, Cho S, Choi HJ, Lee J, et al. Multi-objective instance weighting-based deep transfer learning network for intelligent fault diagnosis. Appl Sci-Basel 2021;11(5):2370. <https://doi.org/10.3390/app11052370>.
- [32] Ren W, Wen G, Xu B, Zhang Z. A novel convolutional neural network based on time-frequency spectrogram of arc sound and its application on GTAW penetration classification. IEEE T Ind Inf 2020;17(2):809–19. <https://doi.org/10.1109/TII.2020.2978114>.
- [33] Miao R, Gao Y, Ge L, Jiang Z, Zhang J. Online defect recognition of narrow overlap weld based on two-stage recognition model combining continuous wavelet transform and convolutional neural network. Comput Ind 2019;112:103115. <https://doi.org/10.1016/j.compind.2019.07.005>.
- [34] Gantala T, Balasubramanian K. Automated defect recognition for welds using simulation assisted TFM imaging with artificial intelligence. J Nondestruct Eval 2021;40(1):1–24. <https://doi.org/10.1007/s10921-021-00761-1>.
- [35] Dong Z, Mai Z, Yin S, Wang J, Yuan J, Fei Y. A weld line detection robot based on structure light for automatic NDT. Int J Adv Manuf Tech 2020;111(7):1831–45. <https://doi.org/10.1007/s00170-020-05964-w>.
- [36] Song Y, Li S. Gas leak detection in galvanised steel pipe with internal flow noise using convolutional neural network. Process Saf Environ 2021;146:736–44. <https://doi.org/10.1016/j.psep.2020.11.053>.
- [37] Zhi Z, Jiang H, Yang D, Gao J, Wang Q, Wang X, et al. An end-to-end welding defect detection approach based on titanium alloy time-of-flight diffraction images. J Intell Manuf 2022. <https://doi.org/10.1007/s10845-021-01905-w>.
- [38] Rout A, Deepak BBV, Biswas BB. Advances in weld seam tracking techniques for robotic welding: a review. Robot Cim-Int Manuf 2019;56:12–37. <https://doi.org/10.1016/j.rcim.2018.08.003>.
- [39] Lei T, Rong Y, Wang H, Huang Y, Li M. A review of vision-aided robotic welding. Comput Ind 2020;123:10326. <https://doi.org/10.1016/j.compind.2020.103326>.
- [40] Zhang Y, Wang Q, Liu Y. Adaptive intelligent welding manufacturing. Weld J 2021;100(1):63–83. <https://doi.org/10.29391/2021.100.006>.
- [41] Bengio Y, Montreal U. Learning deep architectures for AI. Found Trends Mach Le 2009;2(1):1–87. <https://doi.org/10.1561/2200000006>.
- [42] Deng L, Yu D. Deep learning: methods and applications. Found Trends Signal 2014;7(3–4):197–387. <https://doi.org/10.1561/2000000039>.
- [43] Schmidhuber J. Deep learning in neural networks: an overview. Neural Netw 2015;61:85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>.
- [44] Lecun Y, Bengio Y, Hinton GE. Deep learning. Nature 2015;521(7553):436–44. <https://doi.org/10.1038/nature14539>.
- [45] Zhang W, Yang G, Lin Y, Ji C, Gupta MM. On definition of deep learning. In: IEEE World Automation Congress; 2018. p. 232–6. (<https://doi.org/10.23919/WAC.2018.8430387>).
- [46] Oh S, Jung M, Lim C, Shin S. Automatic detection of welding defects using Faster R-CNN. Appl Sci-Basel 2020;10(23):8629. <https://doi.org/10.3390/app10238629>.
- [47] Zhang Z, Li B, Zhang W, Lu R, Wada S, Zhang Y. Real-time penetration state monitoring using convolutional neural network for laser welding of tailor rolled blanks. J Manuf Syst 2020;54:348–60. <https://doi.org/10.1016/j.jmsy.2020.01.006>.
- [48] Shevchik S, Quang T, Meylan B, Farahani F, Olbinado M, Back A, et al. Supervised deep learning for real-time quality monitoring of laser welding with X-ray radiographic guidance. Sci-Rep-UK 2020;10:3389. <https://doi.org/10.1038/s41598-020-60294-x>.
- [49] Wu D, Zhang P, Yu Z, Gao Y, Zhang H, Chen H, et al. Progress and perspectives of in-situ optical monitoring in laser beam welding: sensing, characterization and modeling. J Manuf Process 2022;75:767–91. <https://doi.org/10.1016/j.jmapro.2022.01.044>.
- [50] Hou W, Zhang D, Wei Y, Guo J, Zhang X. Review on computer aided weld defect detection from radiography images. Appl Sci-Basel 2020;10(5):1878. <https://doi.org/10.3390/app10051878>.
- [51] Xu Y, Wang Z. Visual sensing technologies in robotic welding: recent research developments and future interests. Sens Actuat A-Phys 2021;320:112551. <https://doi.org/10.1016/j.sna.2021.112551>.
- [52] Wang Y, Lu J, Zhao Z, Deng W, Han J, Bai L, et al. Active disturbance rejection control of layer width in wire arc additive manufacturing based on deep learning. J Manuf Process 2021;67:364–75. <https://doi.org/10.1016/j.jmapro.2021.05.005>.
- [53] Park JK, An WH, Kang DJ. Convolutional neural network based surface inspection system for non-patterned welding defects. Int J Precis Eng Man 2019;20(3):363–74. <https://doi.org/10.1007/s12541-019-00074-4>.
- [54] Yang Y, Pan L, Ma J, Yang R, Zhu Y, Yang Y, et al. A high-performance deep learning algorithm for the automated optical inspection of laser welding. Appl Sci-Basel 2020;10(3):933. <https://doi.org/10.3390/app10030933>.
- [55] Bacioiu D, Melton G, Papaelias M, Shaw R. Automated defect classification of SS304 TIG welding process using visible spectrum camera and machine learning. NDT E Int 2019;107:102139. <https://doi.org/10.1016/j.ndteint.2019.102139>.
- [56] Ma G, Yu L, Yuan H, Xiao W, He Y. A vision-based method for lap weld defects monitoring of galvanized steel sheets using convolutional neural network. J Manuf Process 2021;64:130–9. <https://doi.org/10.1016/j.jmapro.2020.12.067>.
- [57] Cruz YJ, Rivas M, Quiza R, Beruvides G, Haber RE. Computer vision system for welding inspection of liquefied petroleum gas pressure vessels based on combined digital image processing and deep learning techniques. Sens-Basel 2020;20(16):4505. <https://doi.org/10.3390/s20164505>.
- [58] Zhu H, Ge W, Liu Z. Deep Learning-based classification of weld surface defects. Appl Sci-Basel 2019;9(16):3312. <https://doi.org/10.3390/app9163312>.
- [59] Deng H, Cheng Y, Feng Y, Xiang J. Industrial laser welding defect detection and image defect recognition based on deep learning model developed. Symmetry-Basel 2021;13(9):1731. <https://doi.org/10.3390/sym13091731>.
- [60] Li Y, Hu M, Wang T. Weld image recognition algorithm based on deep learning. Int J Pattern Recogn 2020;34(08):2052004. <https://doi.org/10.1142/S0218001420520047>.
- [61] Yang D, Cui Y, Yu Z, Yuan H. Deep learning based steel pipe weld defect detection. Appl Artif Intell 2021;35(15):1237–49. <https://doi.org/10.1080/08839514.2021.1975391>.
- [62] Cheng Y, Wang Q, Jiao W, Yu R, Chen S, Zhang Y, et al. Detecting dynamic development of weld pool using machine learning from innovative composite images for adaptive welding. J Manuf Process 2020;56:908–15. <https://doi.org/10.1016/j.jmapro.2020.04.059>.
- [63] Wang D, Dong Y, Pan Y, Ma R. Machine vision-based monitoring methodology for the fatigue cracks in U-rib-to-deck weld seams. IEEE Access 2020;8:94204–19. <https://doi.org/10.1109/ACCESS.2020.2995276>.
- [64] Yang L, Fan J, Huo B, Liu Y. Inspection of welding defect based on multi-feature fusion and a convolutional network. J Nondestruct Eval 2021;40(4):1–11. <https://doi.org/10.1007/s10921-021-00823-4>.
- [65] Tao X, Peng L, Tao Y, Ye C. Inspection of defects in weld using differential array ECT probe and deep learning algorithm. IEEE Trans Instrum Meas 2021;70:1–9. <https://doi.org/10.1109/TIM.2021.3099566>.
- [66] Choi SG, Hwang I, Kim YM, Kang B, Kang M. Prediction of the weld qualities using surface appearance image in resistance spot welding. Met-Basel 2019;9(8):831. <https://doi.org/10.3390/met9080831>.
- [67] Dong X, Taylor CJ, Cootes TF. Defect detection and classification by training a generic convolutional neural network encoder. IEEE Trans Signal Proces 2020;68:6055–69. <https://doi.org/10.1109/TSP.2020.3031188>.
- [68] Ajmi C, Zapata J, Elferchichi S, Zaafouri A, Laabidi K. Deep learning technology for weld defects classification based on transfer learning and activation features. Adv Mater Sci Eng 2020:1574350. <https://doi.org/10.1155/2020/1574350>.
- [69] Guo S, Wang D, Chen J, Feng Z, Guo W. Predicting nugget size of resistance spot welds using infrared thermal videos with image segmentation and convolutional neural network. J Manuf Sci Eng 2022;144(2):021009. <https://doi.org/10.1115/1.4051829>.
- [70] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019;6(1):1–48. <https://doi.org/10.1186/s40537-019-0197-0>.
- [71] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data En 2009;22(10):1345–59. <https://doi.org/10.1109/TKDE.2009.191>.
- [72] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. Commun ACM 2020;63(11):139–44. <https://doi.org/10.1145/3422622>.
- [73] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434; 2015.
- [74] Mirza M, Osindero S. Conditional generative adversarial nets. arXiv Prepr arXiv 2014;1411:1784.
- [75] Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: IEEE International Conference on Computer Vision; 2017. p. 2242–51. (<https://doi.org/10.1109/ICCV.2017.244>).
- [76] Mariani G, Scheidegger F, Istrate R, Bekas C, Malossi C. BAGAN: data augmentation with balancing gan. arXiv preprint arXiv; 2018. p. 1803.09655.
- [77] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: Proceedings of the 34th international conference on machine learning; 2017. p. 214–23.
- [78] Mery D, Kaminetzky A, Golborne L, Figueiroa S, Saavedra D. Target detection by target simulation in X-ray testing. J Nondestruct Eval 2022;41:21. <https://doi.org/10.1007/s10921-022-00851-8>.
- [79] Cai W, Jiang P, Shu LS, Geng SN, Zhou Q. Real-time laser keyhole welding penetration state monitoring based on adaptive fusion images using convolutional neural networks. J Intell Manuf 2021. <https://doi.org/10.1007/s10845-021-01848-2>.
- [80] Cheng YC, Wang QY, Jiao WH, Xiao J, Chen SJ, Zhang YM. Automated recognition of weld pool characteristics from active vision sensing. Weld J 2021;100(5):183–92. <https://doi.org/10.29391/2021.100.015>.
- [81] Xue B, Chang B, Du D. Multi-output monitoring of high-speed laser welding state based on deep learning. Sens-Basel 2021;21(5):1626. <https://doi.org/10.3390/s21051626>.
- [82] Xia C, Pan Z, Fei Z, Zhang S, Li H. Vision based defects detection for keyhole TIG welding using deep learning with visual explanation. J Manuf Process 2020;56:845–55. <https://doi.org/10.1016/j.jmapro.2020.05.033>.
- [83] Ajmi C, Zapata J, Martínez-Álvarez JJ, Doménech G, Ruiz R. Using deep learning for defect classification on a small weld X-ray image dataset. J Nondestruct Eval 2020;39(3):1–13. <https://doi.org/10.1007/s10921-020-00719-9>.
- [84] Dai W, Li D, Tang D, Jiang Q, Wang D, Wang H, et al. Deep learning assisted vision inspection of resistance spot welds. J Manuf Process 2021;62:262–74. <https://doi.org/10.1016/j.jmapro.2020.12.015>.
- [85] Zhou C, Li D, Wang P, Sun J, Huang Y, Li W. ACR-Net: attention integrated and cross-spacial feature fused rotation network for tubular solder joint detection. IEEE Trans Instrum Meas 2021;70:1–12. <https://doi.org/10.1109/TIM.2021.3094837>.

- [86] Zou Y, Zhou W. Automatic seam detection and tracking system for robots based on laser vision. *Mechatronics* 2019;63:102261. <https://doi.org/10.1016/j.mechatronics.2019.102261>.
- [87] Zou Y, Zhu M, Chen X. A robust detector for automated welding seam tracking system. *J Dyn Sys Meas Control* 2021;143(7):071001. <https://doi.org/10.1115/1.4049547>.
- [88] Gonzalez-Val C, Pallas A, Panadeiro V, Rodriguez A. A convolutional approach to quality monitoring for laser manufacturing. *J Intell Manuf* 2020;31(3):789–95. <https://doi.org/10.1007/s10845-019-01495-8>.
- [89] Yang L, Wang Z, Gao S, Shi M, Liu B. Magnetic flux leakage image classification method for pipeline weld based on optimized convolution kernel. *Neurocomputing* 2019;365:229–38. <https://doi.org/10.1016/j.neucom.2019.07.083>.
- [90] Dong X, Taylor CJ, Cootes TF. Automatic aerospace weld inspection using unsupervised local deep feature learning. *Knowl-Based Syst* 2021;221:106892. <https://doi.org/10.1016/j.knosys.2021.106892>.
- [91] Liu C, Chen S, Huang J. Machine vision-based object detection strategy for weld area. *SCI Program-Neth* 2022;2022:1188974. <https://doi.org/10.1155/2022/1188974>.
- [92] Wang Y, Han J, Lu J, Bai L, Zhao Z. TIG stainless steel molten pool contour detection and weld width prediction based on res-seg. *Met-Basel* 2020;10(11):1495. <https://doi.org/10.3390/met10111495>.
- [93] Ai J, Tian L, Bai L, Zhang J. Data augmentation by a CycleGAN-based extra-supervised model for nondestructive testing. *Meas Sci Technol* 2022;33:045017. <https://doi.org/10.1088/1361-6501/ac5ec3>.
- [94] Dai W, Li D, Tang D, Wang H, Peng Y. Deep learning approach for defective spot welds classification using small and class-imbalanced datasets. *Neurocomputing* 2022;477:46–60. <https://doi.org/10.1016/j.neucom.2022.01.004>.
- [95] Le X, Mei J, Zhang H, Zhou B, Xi J. A learning-based approach for surface defect detection using small image datasets. *Neurocomputing* 2020;408:112–20. <https://doi.org/10.1016/j.neucom.2019.09.107>.
- [96] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Prpc IEEE* 86(11); 1998. p. 2278–324. (<https://doi.org/10.1109/5.726791>).
- [97] He K, Zhang X, Ren SQ, Sun J. Deep residual learning for image recognition. In: *IEEE conference on computer vision and pattern recognition*; 2016. p. 770–778. (<https://doi.org/10.1109/CVPR.2016.90>).
- [98] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv Prepr arXiv* 2015;1409:1556.
- [99] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang WJ, Weyand T, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv: 1704.04861*; 2017.
- [100] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2014. p. 580–7. (<https://doi.org/10.1109/CVPR.2014.81>).
- [101] Girshick R, Farst R-CNN. In: *IEEE International Conference on Computer Vision*. 2015: 1440–1448. (<https://doi.org/10.1109/ICCV.2015.169>).
- [102] Ren SQ, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal* 2017;39(6). <https://doi.org/10.1109/TPAMI.2016.2577031>. 1137–49.
- [103] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2016. p. 779–88. (<https://doi.org/10.1109/CVPR.2016.91>).
- [104] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, et al. SSD: single shot multibox detector. In: *European Conference on Computer Vision*; 2016. p. 21–37. (https://doi.org/10.1007/978-3-319-46448-0_2).
- [105] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal* 2017;39(4):640–51. <https://doi.org/10.1109/TPAMI.2016.2572683>.
- [106] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: *International conference on medical image computing and computer-assisted intervention*; 2015. p. 234–41. (https://doi.org/10.1007/978-3-319-24574-4_28).
- [107] Yu R, Kershaw J, Wang P, Zhang Y. Real-time recognition of arc weld pool using image segmentation network. *J Manuf Process* 2021;72:159–67. <https://doi.org/10.1016/j.jmapro.2021.10.019>.
- [108] He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: *IEEE International Conference on Computer Vision*; 2017. p. 2961–9. (<https://doi.org/10.1109/ICCV.2017.322>).
- [109] Ma C, Huang J, Yang X, Yang M. Hierarchical convolutional features for visual tracking. In: *IEEE International Conference on Computer Vision*; 2015. p. 3074–3082. (<https://doi.org/10.1109/ICCV.2015.352>).
- [110] Nam H, Han B. Learning multi-domain convolutional neural networks for visual tracking. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2016. p. 4293–4302. (<https://doi.org/10.1109/CVPR.2016.465>).
- [111] Danelljan M, Bhat G, Khan F, Felsberg MECO.: efficient convolution operators for tracking. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2017. p. 6931–6939. (<https://doi.org/10.1109/CVPR.2017.733>).
- [112] Nazir Q, Shao C. Online tool condition monitoring for ultrasonic metal welding via sensor fusion and machine learning. *J Manuf Process* 2021;62:806–16. <https://doi.org/10.1016/j.jmapro.2020.12.050>.
- [113] Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K Spatial transformer networks. In: *International Conference on Neural Information Processing Systems*; 2015. p. 2017–25.
- [114] Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-excitation networks. *IEEE Trans Pattern Anal* 2020;42(8):2011–23. <https://doi.org/10.1109/TPAMI.2019.2913372>.
- [115] Woo S, Park J, Lee JY, Kweon IS. CBAM: convolutional block attention module. In: *European Conference on Computer Vision*; 2018. p. 3–19.
- [116] Biran O, Cotton C. Explanation and justification in machine learning: a survey. In: *International Joint Conference on Artificial Intelligence*; 2017. p. 8–13.
- [117] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D.. Grad-cam: visual explanations from deep networks via gradient-based localization. In: *IEEE International Conference on Computer Vision*; 2017. p. 618–626. (<https://doi.org/10.1109/ICCV.2017.74>).
- [118] Zhou BL, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2016. p. 27–30. (<https://doi.org/10.1109/CVPR.2016.319>).
- [119] Jiang R, Xiao R, Chen S. Prediction of penetration based on infrared thermal and visual images during pulsed GTAW process. *J Manuf Process* 2021;69:261–72. <https://doi.org/10.1016/j.jmapro.2021.07.046>.
- [120] Ma D, Jiang P, Shu L, Geng S. Multi-sensing signals diagnosis and CNN-based detection of porosity defect during Al alloys laser welding. *J Manuf Syst* 2022;62: 334–46. <https://doi.org/10.1016/j.jmsy.2021.12.004>.
- [121] Jamnikar ND, Liu S, Brice C, Zhang X. In-process comprehensive prediction of bead geometry for laser wire-feed DED system using molten pool sensing data and multi-modality CNN. *Int J Adv Manuf Tech* 2022;121:903–17. <https://doi.org/10.1007/s00170-022-09248-3>.
- [122] Wu D, Huang Y, Zhang P, Yu Z, Chen H, Chen S. Visual-acoustic penetration recognition in variable polarity plasma arc welding process using hybrid deep learning approach. *IEEE Access* 2020;8:120417–28. <https://doi.org/10.1109/ACCESS.2020.3005822>.
- [123] Miao R, Shan Z, Zhou Q, Wu Y, Ge L, Zhang J, et al. Real-time defect identification of narrow overlap welds and application based on convolutional neural networks. *J Manuf Syst* 2022;62:800–10. <https://doi.org/10.1016/j.jmsy.2021.01.012>.
- [124] Chen C, Xiao R, Chen H, Lv N, Chen S. Prediction of welding quality characteristics during pulsed GTAW process of aluminum alloy by multisensory fusion and hybrid network model. *J Manuf Processes*, 68(B); 2020. p. 209–24. (<https://doi.org/10.1016/j.jmapro.2020.08.028>).
- [125] Hou W, Rao L, Zhu A, Zhang D. Feature fusion for weld defect classification with small dataset. *J Sens* 2022;2022:8088202. <https://doi.org/10.1155/2022/8088202>.
- [126] Kershaw J, Yu R, Zhang YM, Wang P. Hybrid machine learning-enabled adaptive welding speed control. *J Manuf Process* 2021;71:374–83. <https://doi.org/10.1016/j.jmapro.2021.09.023>.
- [127] 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* 2019;16 (1):465–74. <https://doi.org/10.1109/TII.2019.2937563>.
- [128] Cai N, Cen G, Wu J, Li F, Wang H, Chen X. SMT solder joint inspection via a novel cascaded convolutional neural network. *IEEE Trans Comp Pack Man* 2018;8(4): 670–7. <https://doi.org/10.1109/TCMP.2018.2789453>.
- [129] Yang L, Wang Z, Gao S. Pipeline magnetic flux leakage image detection algorithm based on multiscale SSD network. *IEEE Trans Ind Inf* 2019;16(1):501–9. <https://doi.org/10.1109/TII.2019.2926283>.
- [130] Wang Y, Zhang C, Lu J, Bai L, Zhao Z, Han J. Weld reinforcement analysis based on long-term prediction of molten pool image in additive manufacturing. *IEEE Access* 2020;8:69908–18. <https://doi.org/10.1109/ACCESS.2020.2986130>.
- [131] 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. <https://doi.org/10.1016/j.jmsy.2021.01.017>.
- [132] Feng T, Huang S, Liu J, Wang J, Fang X. Welding surface inspection of armatures via CNN and image comparison. *IEEE Sens J* 2021;21(19):21696–704. <https://doi.org/10.1109/JSEN.2021.3079334>.
- [133] Li Q, Yang B, Wang S, Zhang Z, Tang X, Zhao C. A fine-grained flexible graph convolution network for visual inspection of resistance spot welds using cross-domain features. *J Manuf Process* 2022;78:319–29. <https://doi.org/10.1016/j.jmapro.2022.04.025>.
- [134] Li W, Qi J, Sun H. TAF2-Net: triple-branch attentive feature fusion network for ultrasonic flaw detection. *IEEE Trans Instrum Meas* 2022;71:1–12. <https://doi.org/10.1109/TIM.2022.3150592>.
- [135] Li Z, Chen H, Ma X, Chen H, Ma Z. Triple pseudo-siamese network with hybrid attention mechanism for welding defect detection. *Mater Des* 2022;217:110645. <https://doi.org/10.1016/j.matdes.2022.110645>.
- [136] Wang Z, Chen H, Zhong Q, Lin S, Wu J, Xu M, et al. Recognition of penetration state in GTAW based on vision transformer using weld pool image. *Int J Adv Manuf Tech* 2022;119:5439–52. <https://doi.org/10.1007/s00170-021-08538-6>.
- [137] Liu T, Bao J, Wang J, Zhang Y. A hybrid CNN–LSTM algorithm for online defect recognition of CO₂ welding. *Sens-Basel* 2018;18(12):4369. <https://doi.org/10.3390/s18124369>.
- [138] Božič A, Kos M, Jezeršek M. Power control during remote laser welding using a convolutional neural network. *Sens-Basel* 2020;20(22):6658. <https://doi.org/10.3390/s20226658>.
- [139] Liu T, Zheng H, Bao J, Zheng P, Wang J, Yang C, et al. An explainable laser welding defect recognition method based on multi-scale class activation mapping. *IEEE Trans Instrum Meas* 2022;71:1–12. <https://doi.org/10.1109/TIM.2022.3148739>.
- [140] Ruder S. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*; 2016.

- [141] Li Y, Mu H, Polden H, Li H, Wang L, Xia C, Pan Z. Towards intelligent monitoring system in wire arc additive manufacturing: a surface anomaly detector on a small dataset. *Int J Adv Manuf Tech* 2022;120:5225–42. <https://doi.org/10.1007/s00170-022-09076-5>.
- [142] Sekhar R, Sharma D, Shah P. Intelligent classification of tungsten inert gas welding defects: a transfer learning approach. *Front Mech Eng* 2022;8:824038. <https://doi.org/10.3389/fmech.2022.824038>.
- [143] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International Conference on Machine Learning; 2015. p. 448–56.
- [144] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 2014; 15(1):1929–58.
- [145] Pan H, Pang Z, Wang Y, Wang Y, Chen L. A new image recognition and classification method combining transfer learning algorithm and mobilenet model for welding defects. *IEEE Access* 2020;8:119951–60. <https://doi.org/10.1109/ACCESS.2020.3005450>.
- [146] Ghiasi G, Lin TY, Le Q.V. Dropblock: a regularization method for convolutional networks. *arXiv preprint arXiv:1810.12890*; 2018.
- [147] Yang L, Jiang H. Weld defect classification in radiographic images using unified deep neural network with multi-level features. *J Intell Manuf* 2020;32(2):459–69. <https://doi.org/10.1007/s10845-020-01581-2>.
- [148] Chen Y, Su S, Yang H. Convolutional neural network analysis of recurrence plots for anomaly detection. *Int J Bifurcat Chaos* 2020;30(01):2050002. <https://doi.org/10.1142/S0218127420500029>.
- [149] Zhang B, Liu S, Shin YC. In-process monitoring of porosity during laser additive manufacturing process. *Addit Manuf* 2019;28:497–505. <https://doi.org/10.1016/j.addma.2019.05.030>.
- [150] Liu J, Guo F, Zhang Y, Hou B, Zhou H. Defect classification on limited labeled samples with multiscale feature fusion and semi-supervised learning. *Appl Intell* 2021;1–16. <https://doi.org/10.1007/s10489-021-02917-y>.
- [151] Li S, Gao J, Zhou E, Pan Q, Wang X. Deep learning-based fusion hole state recognition and width extraction for thin plate TIG welding. *Weld World* 2022;66: 1329–47. <https://doi.org/10.1007/s0194-022-01287-4>.
- [152] Xiao M, Yang B, Wang S, Chang Y, Li S, Yi G. Research on recognition methods of spot-welding surface appearances based on transfer learning and a lightweight high-precision convolutional neural network. *J Intell Manuf* 2022. <https://doi.org/10.1007/s10845-022-01909-0>.
- [153] Wang B, Li F, Lu R, Ni X, Zhu W. Weld feature extraction based on semantic segmentation network. *Sens-Basel* 2022;22(11):4130. <https://doi.org/10.3390/s22114130>.
- [154] Park MY, Hastie T. L1-regularization path algorithm for generalized linear models. *J R Stat Soc B* 2007;69(4):659–77. <https://doi.org/10.1111/j.1467-9868.2007.00607.x>.
- [155] Cortes C, Mohri M, Rostamizadeh A. L2 regularization for learning kernels. *arXiv Prepr arXiv* 2012;1205:2653.
- [156] Liu T, Bao J, Wang J, Zhang Y. A coarse-grained regularization method of convolutional kernel for molten pool defect identification. *J Comput Inf Sci Eng* 2020;20(2):021005. <https://doi.org/10.1115/1.4045294>.
- [157] Prechelt L. Early stopping-but when? *Neural Netw: Tricks Trade* 1998;55–69. <https://doi.org/10.1007/3-540-49430-8-3>.
- [158] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J., et al. Tensorflow: a system for large-scale machine learning. In: Operating Systems Design and Implementation; 2016. p. 265–83.
- [159] Parvat A., Chavan J., Kadam S., Dev S., Pathak V. A survey of deep-learning frameworks. In: International Conference on Inventive Systems and Control; 2017. p. 1–7. (<https://doi.org/10.1109/ICISC.2017.8068684>).
- [160] Gulli A, Pal S. Deep learning with Keras. Packt Publishing Ltd; 2017.
- [161] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: an imperative style, high-performance deep learning library. *Adv Neural Inf Process Syst* 2019;32:8026–37.
- [162] Han CM, Kim TW, Choi HW. Assessment of laser joining quality by visual inspection, computer Simulation, and deep learning. *Appl Sci-Basel* 2021;11(2): 642. <https://doi.org/10.3390/app11020642>.
- [163] Mishra A, Patti A. Deep convolutional neural network modeling and laplace transformation algorithm for the analysis of surface quality of friction stir welded joints. *ADCAIJ-Adv Distrib C*, 10(3); 2021. 3–7–320. (<http://hdl.handle.net/10366/147243>).
- [164] Tian Y, Liu H, Li L, Yuan G, Feng J, Chen Y, et al. Automatic identification of multi-type weld seam based on vision sensor with silhouette-mapping. *IEEE Sens J* 2020;21(4):5402–12. <https://doi.org/10.1109/JSEN.2020.3034382>.
- [165] Zhang Y, You D, Gao X, Wang C, Li Y, Gao PP. Real-time monitoring of high-power disk laser welding statuses based on deep learning framework. *J Intell Manuf* 2019;31:799–814. <https://doi.org/10.1007/s10845-019-01477-w>.
- [166] Jia CB, Liu XF, Zhang GK, Zhang Y, Yu CH, Wu CS. Penetration/keyhole status prediction and model visualization based on deep learning algorithm in plasma arc welding. *Int J Adv Manuf Tech* 2021;117(11):3577–97. <https://doi.org/10.1007/s00170-021-07903-9>.
- [167] Hou W, Wei Y, Jin Y, Zhu C. Deep features based on a DCNN model for classifying imbalanced weld flaw types. *Measurement* 2019;131:482–9. <https://doi.org/10.1016/j.measurement.2018.09.011>.
- [168] Jiang H, Hu Q, Zhi Z, Gao J, Gao Z, Wang R, et al. Convolution neural network model with improved pooling strategy and feature selection for weld defect recognition. *Weld World* 2021;65(4):731–44. <https://doi.org/10.1007/s40194-020-01027-6>.
- [169] Kumaresan S, Aultrin KSJ, Kumar SS, Anand MD. Transfer learning with CNN for classification of weld defect. *IEEE Access* 2021;9:95097–108. <https://doi.org/10.1109/ACCESS.2021.3093487>.
- [170] Miao R, Jiang Z, Zhou Q, Wu Y, Gao Y, Zhang J, et al. Online inspection of narrow overlap weld quality using two-stage convolution neural network image recognition. *Mach Vis Appl* 2021;32(1):27. <https://doi.org/10.1007/s00138-020-01158-2>.
- [171] Zhang B, Jaiswal P, Rai R, Guerrier P, Baggs G. Convolutional neural network-based inspection of metal additive manufacturing parts. *Rapid Prototyp J* 2019;25 (3):530–40. <https://doi.org/10.1108/RPJ-04-2018-0096>.
- [172] Hu A, Wu L, Huang J, Fan D, Xu Z. Recognition of weld defects from X-ray images based on improved convolutional neural network. *Multimed Tools Appl* 2022;81: 15085–102. <https://doi.org/10.1007/s11042-022-12546-3>.
- [173] Di X, Li L. Image defect recognition method based on deep learning network. *Mob Inf Syst* 2022;2022:5696334. <https://doi.org/10.1155/2022/5696334>.
- [174] Zhang Z, Liu X, Sun X. Image recognition of limited and imbalanced samples based on transfer learning methods for defects in welds. *Proc Inst Mech Eng B-J Eng* 2022. <https://doi.org/10.1177/09544054221082779>.
- [175] Liu J, Liu X, Qu F, Zhang H, Zhang L. A defect recognition method for low-quality weld image based on consistent multiscale feature mapping. *IEEE Trans Instrum Meas* 2022;71:1–11. <https://doi.org/10.1109/TIM.2022.3171609>.
- [176] Lin CS, Huang YC, Chen SH, Hsu YL, Lin YC. The application of deep learning and image processing technology in laser positioning. *Appl Sci-Basel* 2018;8(9):1542. <https://doi.org/10.3390/app8091542>.
- [177] Wang Q, Jiao W, Wang P, Zhang Y. A tutorial on deep learning-based data analytics in manufacturing through a welding case study. *J Manuf Process* 2021; 63:2–13. <https://doi.org/10.1016/j.jmapro.2020.04.044>.
- [178] Wang Y, Xu X, Zhao D, Deng W, Han J, Bai L, et al. Coordinated monitoring and control method of deposited layer width and reinforcement in WAAM process. *J Manuf Process* 2021;71:306–16. <https://doi.org/10.1016/j.jmapro.2021.09.033>.
- [179] Kim H, Nam K, Oh S, Ki H. Deep-learning-based real-time monitoring of full-penetration laser keyhole welding by using the synchronized coaxial observation method. *J Manuf Process* 2021;68:1018–30. <https://doi.org/10.1016/j.jmapro.2021.06.029>.
- [180] Yu R, Kershaw J, Wang P, Zhang Y. How to accurately monitor the weld penetration from dynamic weld pool serial images using CNN-LSTM deep learning model? *IEEE Robot Autom Let*, 7(3); 2022. p. 6519–25. (<https://doi.org/10.1109/LRA.2022.3173659>).
- [181] Zhang K, Shen H. Solder joint defect detection in the connectors using improved Faster-R-CNN algorithm. *Appl Sci-Basel* 2021;11(2):576. <https://doi.org/10.3390/app11020576>.
- [182] Yun GH, Oh SJ, Shin SC. Image preprocessing method in radiographic inspection for automatic detection of ship welding defects. *Appl Sci-Basel* 2022;12(1):123. <https://doi.org/10.3390/app12010123>.
- [183] Chen Y, Wang J, Wang G. Intelligent welding defect detection model on improved R-CNN. *IETE J Res* 2022;1–10. <https://doi.org/10.1080/03772063.2022.2040387>.
- [184] Li H, Li B, Wen X, Wang H, Wang X, Zhang S, et al. A detection and configuration method for welding completeness in the automotive body-in-white panel based on digital twin. *Sci Rep-UK* 2022;12:7929. <https://doi.org/10.1038/s41598-022-11440-0>.
- [185] Chen Z, Wang Q, Yang K, Yu T, Yao J, Liu Y, et al. Deep learning for the detection and recognition of rail defects in ultrasound B-scan images. *Transp Res Rec* 2021; 2675(11):888–901. <https://doi.org/10.1177/03611981211021547>.
- [186] Wu H, Gao W, Xu X. Solder joint recognition using mask R-CNN method. *IEEE Trans Comp Pack Man* 2019;10(3):525–30. <https://doi.org/10.1109/TCMP.2019.2952393>.
- [187] Yang L, Fan J, Huo B, Li E, Liu Y. Image denoising of seam images with deep learning for laser vision seam tracking. *IEEE Sens J* 2022;22(6):6098–107. <https://doi.org/10.1109/JSEN.2022.3147489>.
- [188] Mi J, Zhang Y, Li H, Shen S, Yang Y, Song C, et al. In-situ monitoring laser based directed energy deposition process with deep convolutional neural network. *J Intell Manuf* 2021;1–11. <https://doi.org/10.1007/s10845-021-01820-0>.
- [189] Tan Z, Fang Q, Li H, Liu S, Zhu W, Yang D. Neural network based image segmentation for spatter extraction during laser-based powder bed fusion processing. *Opt Laser Technol* 2020;130:106347. <https://doi.org/10.1016/j.optlastec.2020.106347>.
- [190] Yang Y, He Y, Guo H, Chen Z, Zhang L. Semantic segmentation supervised deep-learning algorithm for welding-defect detection of new energy batteries. *Neural Comput Appl* 2022. <https://doi.org/10.1007/s00521-022-07474-0>.
- [191] Ling Z, Zhang A, Ma D, Shi Y, Wen H. Deep siamese semantic segmentation network for PCB welding defect detection. *IEEE Trans Instrum Meas* 2022;71: 1–11. <https://doi.org/10.1109/TIM.2022.3154814>.
- [192] Li Z, Li D. A high-frequency feature enhancement network for the surface defect detection of welded rebar. *Struct Control Health* 2022;29(8):e2983. <https://doi.org/10.1002/stc.2983>.
- [193] Pandiyan V, Murugan P, Tjahjowidodo T, Caesarendra W, Manyar OM, Then DJH. In-process virtual verification of weld seam removal in robotic abrasive belt grinding process using deep learning. *Robot Cim-Int Manuf* 2019;57:477–87. <https://doi.org/10.1016/j.rcim.2019.01.006>.
- [194] Liu H, Yan Y, Song K, Chen H, Yu H. Efficient optical measurement of welding studs with normal maps and convolutional neural network. *IEEE Trans Instrum Meas* 2020;70:1–14. <https://doi.org/10.1109/TIM.2020.3024389>.
- [195] Zhu Y, Yang R, He Y, Ma J, Guo H, Yang Y, et al. A lightweight multiscale attention semantic segmentation algorithm for detecting laser welding defects on

- safety vent of power battery. *IEEE Access* 2021;9:39245–54. <https://doi.org/10.1109/ACCESS.2021.3064180>.
- [196] Jang J, Van D, Jang H, Baik DH, Yoo SD, Park J, et al. Residual neural network-based fully convolutional network for microstructure segmentation. *Sci Technol Weld Joi* 2020;25(4):282–9. <https://doi.org/10.1080/13621718.2019.1687635>.
- [197] Ferguson MK, Ronay A, Lee YTT, Law KH. Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning. *Smart Sustain Manuf* 2018;2:1–43. <https://doi.org/10.1520/SSMS20180033>.
- [198] Zhang H, Di X, Zhang Y. Real-time CU-Net-based welding quality inspection algorithm in battery production. *IEEE Trans Ind Electron* 2020;67(12):10942–50. <https://doi.org/10.1109/TIE.2019.2962421>.
- [199] Wang L, Shen Q. Visual inspection of welding zone by boundary-aware semantic segmentation algorithm. *IEEE T Instrum Meas* 2020;70:1–9. <https://doi.org/10.1109/TIM.2020.3010665>.
- [200] Yang L, Fan J, Huo B, Li E, Liu Y. A nondestructive automatic defect detection method with pixelwise segmentation. *Knowl-Based Syst* 2022;242:108338. <https://doi.org/10.1016/j.knosys.2022.108338>.
- [201] Yang L, Song S, Fan J, Huo B, Li E, Liu Y. An automatic deep segmentation network for pixel-level welding defect detection. *IEEE Trans Instrum Meas* 2022;71:1–10. <https://doi.org/10.1109/TIM.2021.3127645>.
- [202] Nowroth C, Gu T, Grajczak J, Nothdrift S, Twiefel J, Hermsdorf J, et al. Deep learning-based weld contour and defect detection from micrographs of laser beam welded semi-finished products. *Appl Sci-Basel* 2022;12(9):4645. <https://doi.org/10.3390/app12094645>.
- [203] Zou Y, Li J, Chen X, Lan R. Learning siamese networks for laser vision seam tracking. *J Opt Soc Am A Opt Image Sci Vis* 2018;35(11):1805–13. <https://doi.org/10.1364/JOSAA.35.001805>.
- [204] Ma C, Huang JB, Yang X, Yang MH. Hierarchical convolutional features for visual tracking. In: *IEEE International Conference on Computer Vision*; 2015. p. 3074–82. <https://doi.org/10.1109/ICCV.2015.352>.
- [205] Nam H, Han B. Learning multi-domain convolutional neural networks for visual tracking. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2016. p. 4293–302. <https://doi.org/10.1109/CVPR.2016.465>.
- [206] Zou Y, Lan R, Wei X, Chen J. Robust seam tracking via a deep learning framework combining tracking and detection. *Appl Opt* 2020;59(14):4321–31. <https://doi.org/10.1364/AO.389730>.
- [207] Xiao R, Xu Y, Hou Z, Chen C, Chen S. An adaptive feature extraction algorithm for multiple typical seam tracking based on vision sensor in robotic arc welding. *Sens Actuat A-Phys* 2019;297:111533. <https://doi.org/10.1016/j.sna.2019.111533>.
- [208] Zhang Y, Yang G, Wang Q, Ma L, Wang Y. Weld feature extraction based on fully convolutional networks. *Chin J Lasers* 2019;46(3):0302002. <https://doi.org/10.3788/CJL201946.0302002>.
- [209] Li J, Jin S, Wang C, Xue J, Wang X. Weld line recognition and path planning with spherical tank inspection robots. *J Field Robot* 2022;39(2):131–52. <https://doi.org/10.1002/rob.22042>.
- [210] Singh J, Tant K, Mulholland A, MacLeod C. Deep learning based inversion of locally anisotropic weld properties from ultrasonic array data. *Appl Sci-Basel* 2022;12(2):532. <https://doi.org/10.3390/app12020532>.
- [211] Mirzapour M, Movafeghi A, Yahaghi E. Quantitative weld defect sizing using convolutional neural network-aided processing of RT images. *Insight* 2021;63(3):141–5. <https://doi.org/10.1784/insi.2021.63.3.141>.
- [212] Rohkohl E, Kraken M, Schönenmann M, Breuer A, Herrmann C. How to characterize a NDT method for weld inspection in battery cell manufacturing using deep learning. *Int J Adv Manuf Tech* 2022;119. <https://doi.org/10.1007/s00170-021-08553-7>. 4829–2843.
- [213] Gao W, Li C, Yang T, Li L, Li X. A method for detection of rare industrial fake film based on deep learning. *Secur Commun Netw* 2022; 2022. 3253190. <https://doi.org/10.1155/2022/3253190>.
- [214] Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2009. p. 248–55. <https://doi.org/10.1109/CVPR.2009.5206848>.
- [215] Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan, et al. Microsoft COCO: common objects in context. In: *European Conference on Computer Vision*; 2014. p. 740–55. https://doi.org/10.1007/978-3-319-10602-1_48.
- [216] Gould S, Fulton R, Koller D. Decomposing a scene into geometric and semantically consistent regions. In: *International Conference on Computer Vision*; 2009. p. 1–8. <https://doi.org/10.1109/ICCV.2009.5459211>.
- [217] Fan H, Lin L, Yang F, Chu P, Deng G, Yu S, et al. Lasot: a high-quality benchmark for large-scale single object tracking. In: *IEEE Conference on Computer Vision and Pattern Recognition*; 2019. p. 5374–83. <https://doi.org/10.1109/CVPR.2019.900552>.
- [218] Baciuoi D, Melton G, Papaelias M, Shaw R. Automated defect classification of aluminium 5083 TIG welding using HDR camera and neural networks. *J Manuf Process* 2019;45:603–13. <https://doi.org/10.1016/j.jmapro.2019.07.020>.
- [219] Martínez RT, Bestard GA, Silva AMA, Alfaro SCA. Analysis of GMAW process with deep learning and machine learning techniques. *J Manuf Process* 2021;62:695–703. <https://doi.org/10.1016/j.jmapro.2020.12.052>.
- [220] Mery D, Riffó V, Zscherpel U, Mondragón G, Lillo I, Zuccar I, et al. GDxRay: the database of X-ray images for nondestructive testing. *J Nondestruct Eval* 2015;34(4):42. <https://doi.org/10.1007/s10921-015-0315-7>.
- [221] Hartl R, Bachmann A, Habedank JB, Semm T, Zaeh MF. Process monitoring in friction stir welding using convolutional neural networks. *Met-Basel* 2021;11(4):535. <https://doi.org/10.3390/met11040535>.
- [222] Kandala RR, Abhinay A, Aechana S; 2021. (<https://www.kaggle.com/ruthka/maskrcnn-msweldefect>).
- [223] Zhang Y, You D, Gao X, Zhang N, Gap 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. <https://doi.org/10.1016/j.jmsy.2019.02.004>.
- [224] Cheng Y, Chen S, Xiao J, Zhang Y. Dynamic estimation of joint penetration by deep learning from weld pool image. *Sci Technol Weld Joi*, 26(4); 2021. p. 279–85. <https://doi.org/10.1080/13621718.2021.1896141>.
- [225] Shang J, An W, Liu Y, Han B, Guo Y. Oil pipeline weld defect identification system based on convolutional neural network. *KSII T Internet Inf*, 14(3); 2020. p. 1086–103. <https://doi.org/10.3837/tiiis.2020.03.010>.
- [226] Yang Y, Yang R, Pan L, Ma J, Zhu Y, Diao T, et al. A lightweight deep learning algorithm for inspection of laser welding defects on safety vent of power battery. *Comput Ind* 2020;123:103306. <https://doi.org/10.1016/j.compind.2020.103306>.
- [227] Zhang B, Zhu J, Su H. Toward the third generation of artificial intelligence. *Sci Sin Inf* 2020;50(9):1281–302. <https://doi.org/10.1360/SSI-2020-0204>.
- [228] Kästner L, Ahmadi S, Jonietz F, Jung P, Caire G, Ziegler M, et al. Classification of spot-welded joints in laser thermography data using convolutional neural networks. *IEEE Access* 2021;9:48303–12. <https://doi.org/10.1109/ACCESS.2021.3063672>.
- [229] Maldonado-Ramirez A, Rios-Cabrera R, Lopez-Juarez I. A visual path-following learning approach for industrial robots using DRL. *Robot Cim-Int Manuf*, 71; 2021. p. 102130. <https://doi.org/10.1016/j.rcim.2021.102130>.
- [230] Guo R, Liu H, Xie G, Zhang Y. Weld defect detection from imbalanced radiographic images based on contrast enhancement conditional generative adversarial network and transfer learning. *IEEE Sens J* 2021;21(9):10844–53. <https://doi.org/10.1109/JSEN.2021.3059860>.
- [231] Nomura K, Fukushima K, Matsumura T, Asai S. Burn-through prediction and weld depth estimation by deep learning model monitoring the molten pool in gas metal arc welding with gap fluctuation. *J Manuf Process* 2021;61:590–600. <https://doi.org/10.1016/j.jmapro.2020.10.019>.
- [232] Fioravanti CCB, Centeno TM, Delgado MRDBDS. A deep artificial immune system to detect weld defects in DWI radiographic images of petroleum pipes. *IEEE Access* 2019;7:180947–64. <https://doi.org/10.1109/ACCESS.2019.2959810>.
- [233] Dung CV, Sekiya H, Hirano S, Okatani T, Miki C. A vision-based method for crack detection in gusset plate welded joints of steel bridges using deep convolutional neural networks. *Autom Constr* 2019;102:217–29. <https://doi.org/10.1016/j.autcon.2019.02.013>.
- [234] Sassi P, Tripicchio P, Avizzano CA. A smart monitoring system for automatic welding defect detection. *IEEE T Ind Electron* 2019;66(12):9641–50. <https://doi.org/10.1109/TIE.2019.2896165>.
- [235] Nogay HS, Akinci TC. Classification of operation cases in electric arc welding machine by using deep convolutional neural networks. *Neural Comput Appl* 2021; 33(12):6657–70. <https://doi.org/10.1007/s00521-020-05436-y>.
- [236] Wang Q, Jiao W, Zhang YM. Deep learning-powered digital twin for visualized weld joint growth monitoring and penetration control. *J Manuf Syst* 2020;57:429–39. <https://doi.org/10.1016/j.jmsy.2020.10.002>.
- [237] Chu Y, Feng D, Liu Z, Zhao Z, Wang Z, Xia X, et al. Hybrid-learning-based operational visual quality inspection for edge-computing-enabled IoT system. *IEEE Internet Things* 2022;9(7):4958–72. <https://doi.org/10.1109/JIOT.2021.3107902>.