Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)
16/1 (2023), 77-87. DOI: http://dx.doi.org/10.21609/jiki.v16i1.1147

Weld Defect Detection and Classification based on Deep Learning Method: A Review

Tito Wahyu Purnomo, Finkan Danitasari, Djati Handoko
[1]Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Kampus UI Depok, 16424, Indonesia

E-mail: djati.handoko@ui.ac.id

Abstract

The inspection of weld defects utilizing nondestructive testing techniques based on radiography is essential for ensuring the operability and safety of weld joints in metals or other materials. During the process of welding, weld defects such as cracks, cavity or porosity, lack of penetration, slag inclusion, and metallic inclusion may occur. Due to the limitations of manual interpretation and evaluation, recent research has focused on the automation of weld defect detection and classification from radiographic images. The application of deep learning algorithms enables automated inspection. The deep learning architectures for building weld defect classification models were discussed. This paper concludes with a discussion of the achievements of automation methods and a presentation of the research recommendations for the future.

Keywords: *weld defect, radiographic images, deep learning, convolutional neural network*

1. Introduction

It is critical to test metal or material-welded joints to verify that they meet design and service specifications and to ensure their safety and dependability. Cracks, porosity, gas pores, lack of penetration, slag inclusion, and metallic inclusion may occur on the weldment during the welding process. Poor edge quality, extreme stresses, and improper welding procedures cause cracks. The existence of moisture in the weld metals causes cavity or porosity defects in welded joints. A metallic inclusion defect occurs when metal, as well as tungsten, become entrapped in the weld seam. Lack of penetration is a dangerous defect that occurs when a joint has inadequate weld penetration. It has the potential to cause extensive damage to the weldment. Slag inclusion defects are caused by the presence of nonmetallic elements in welded metals.

Nondestructive testing (NDT) methods for detecting welding defects are typically divided into a number of categories, including visual or manual inspection, radiographic testing with an ionizing radiation source, including gamma rays or X-rays, eddy current testing, ultrasonic testing, dye penetrant testing, etc. In many industries, detecting defects in radiographic images is regarded as a basic requirement and is extensively utilized for controlling welding quality [27]. NDT experts have traditionally evaluated these radiographic images. Manual interpretation and evaluation of radiographic images are complex, subjective, inconsistent, time-consuming, and occasionally biased with respect to defects with similar features. As a result, it is advantageous to combine an automated artificial intelligence algorithm with a computer vision system for image evaluation and interpretation. In terms of working conditions, time required, digitalized documenting, objectivity, and increasing accuracy, automatic detection has significant advantages over manual evaluation.

In recent decades, researchers have concentrated on automated inspection techniques based on radiographic images. It has been demonstrated that radiographic images are typically difficult to examine due to their high noise, inconsistent gray distribution, and low contrast. Image quality has a significant impact on the detection of weld defects, particularly for small defects that are easily obscured by noise. Several methods for image pre-processing have been developed to eliminate or reduce these issues. At the beginning of research on the detection and classification of weld defects,

machine learning algorithms, including adaptive cascade boosting (AdaBoost) classifier, support vector machine (SVM), random forest (RF), neural network (NN), k-nearest neighbor (k-NN), and logistic regression (LR) were widely employed. Deep learning, on the other hand, is one of the machine learning implementation methods that utilizes an artificial neural network (ANN) to replicate the functions of the human brain. However, it is suspected that deep learning algorithms perform better at detecting and classifying weld defects in radiographic images. A deep neural network (DNN) is an ANN with multiple layers between the input and output layers. Convolutional neural network (CNN or ConvNet) is a form of ANN that is frequently employed in deep learning to analyze visual imaging. Deep belief network (DBN) was developed as a solution to the problems encountered when training deep layered networks with traditional neural networks. Transfer learning (TL) is another type of deep learning that involves applying a previously trained model to a new problem. Due to the fact that it can train DNNs with small datasets, it is currently very popular in deep learning.

The objective of this review is to examine methods for detecting and classifying weld defects from digital radiographic images. "Digital" may refer either to the direct recording of the radiographic image using a digital device or to the digitization of the radiographic image from conventional film. No distinction was made. This paper focused on digital radiographic image analysis techniques, including image preprocessing, defect detection, and defect classification. The most essential factor in the detection and classification of weld defects is the performance of the model based on deep learning.

2. Literature Review

We scoured searchable databases for relevant literature on the detection and classification of weld defects based on deep learning algorithms. We used the terms weld defects, radiographic images, deep learning, and CNN in our search. This is based on the object that will be used, which consists of radiographic images containing weld defects. For the detection and classification of weld defects, a CNN-based deep learning approach is used. After collecting the literature that aligned with the keywords, we considered that only journal articles and conference proceedings published within the last five years, or between 2017 and 2022, were considered. Finally, a total of 20 articles from 7 sources were obtained, with specific details depicted in Figure 1, which is a block diagram of this article's review stages.

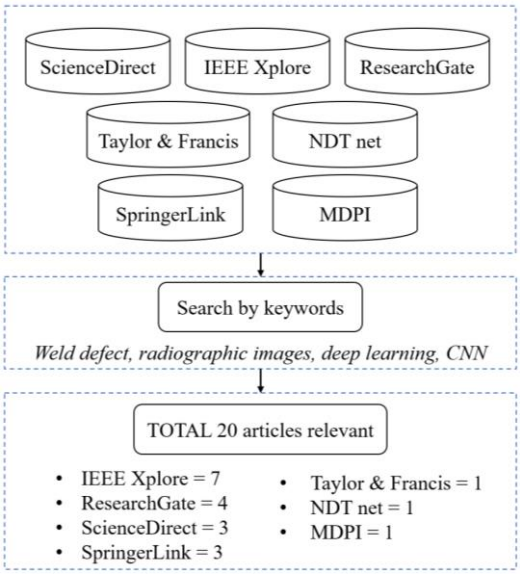


Fig. 1. Flowchart of the article's review stages.

After a comprehensive literature review, it is possible to conclude that the automatic detection system for welding defects is comprised primarily of the following technologies: image pre-processing, defect detection, and defect classification. In our paper, a summary of the literature is provided through a detailed analysis of each section. DNN, DBF, and CNN in various forms, including AlexNet, U-Net, VGG, ResNet, YOLO, MobileNet, Xception, EfficientNet, AF-RCNN, etc., are the deep learning architectures used to create models for detecting and classifying weld defects. Figure 2 represents the stages of a defect detection and classification system involving these aspects.

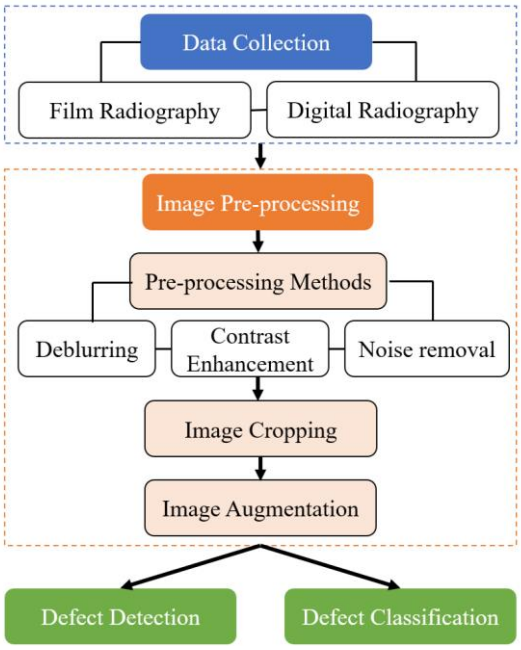


Fig. 2. Flowchart of the method for weld defect detection and classification.

After conducting an in-depth analysis of the collected literature, it is possible to conclude that an automatic welding defect detection and classification system involves the following stages: image collection, image pre-processing, defect detection, and defect classification. Each section of the literature is summarized and analyzed in our paper.

3. Data Collection

The publicly accessible GRIMA X-ray (GDXray) database served as the source for the radiographic image database used in multiple studies [3]. This database contains several categories of X-ray radiographic images, including castings, welds, baggage, nature, and settings. In our review, only the welds category was used as the dataset for the development of weld defect detection and classification models. This category includes 67 high-quality digitized radiographic images from a round-robin test on the detection of defects in welding seams. The radiographic images are stored as uncompressed TIFF files and have a 40.3 m pixel size. The experimental image datasets of [1], [6], [9], [12-14], [17-19], and [22] were retrieved from the GDXRay public database. Figure 3 depicts an example of a radiographic image obtained from the GDXray database.

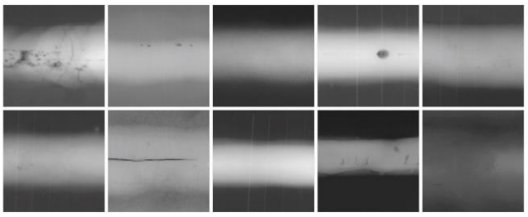


Fig. 3. A sample of radiographic images from the GDXRay database [3].

Furthermore, personal datasets that are not accessible to the public were utilized by [2], [5], [7], [8], [10], [11], [15], [16], [20], and [21]. For example, radiographic images obtained from real-time radiography utilizing a flat panel detector and digital detector array (DDA), as illustrated in Figure 4, are considered to be radiographic images [5, 16].

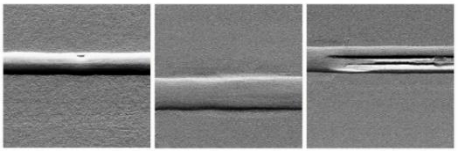


Fig. 4. A sample of radiographic images from DDA [16].

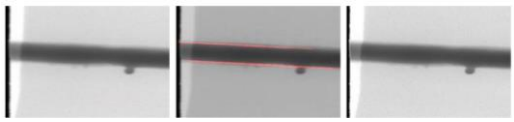
Meanwhile, the GDXray database was utilized by [2] to validate the experimental results obtained on the welding personal dataset. Each article utilized a distinct number of image datasets. The image dataset is obtained from a complete, high-resolution radiographic image, which may contain multiple types of defects.

4. Image Pre-processing

Typically, digital radiographic images have low contrast, noise, and an inconsistent grayscale distribution. However, image quality significantly affected the detection of weld defects, especially for small defects that are easily obscured by background noise. Various processing methods were employed to prevent or mitigate the occurrence of such problems.

4.1. Pre-processing Methods

Several articles that utilize personal databases, particularly real-time radiography, have performed several image pre-processing steps. Motion deblurring using the Hough Transform was utilized to eliminate motion blur in some images so that the weld area and defects could be seen explicitly [5]. In order to eliminate noise caused by defects in the base metal or film, noise removal was adopted to eliminate everything other than the welds and serves as training data alongside tagged information [15]. Contrast enhancement was performed to obtain a more distinct image of defects against the background [5]. Guo et al. [11] proposed a novel model known as the contrast enhancement conditional generative adversarial network (CECGAN), which is utilized as a resampling technique for improving radiographic image datasets. While addressing the weakness of feature extraction caused by low image contrast, the number of image samples is increased, and the data distribution within the images is balanced. Liu et al. [21] proposed image relief pre-processing to enhance the radiographic image's contrast to make it easier to detect defects. Other image pre-processing methods include RGB channel conversion [2], grayscale to binary conversion [17], histogram equalization [9, 16], and gradient processing [16]. Figures 5 and 6 illustrate an example of radiographic image pre-processing steps.



(a) (b) (c)

Fig. 5. The process of blind motion deblurring (a) Blurry image (b) Hough Transform implementation (c) Deblurred image [5].

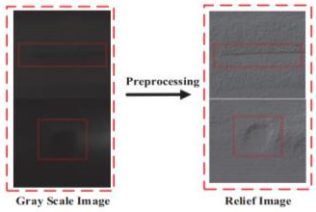


Fig. 6. Image relief preprocessing example [21].

A series of image pre-processing methods are necessary to ensure that the image is suitable to be utilized as a dataset.

4.2. Image Cropping

Following a series of image pre-processing steps, the random cropping method was used to obtain an image containing a single class of weld defect. Random cropping in the form of a square or rectangle can be performed manually or with a sliding-window approach. The pixel size of the randomly cropped image differs among the reviewed articles, including 32 x 32 [6,7,17,18], 64 x 64 [6,20], 71 x 71 [11], 96 x 96 [6], 128 x 128 [1,6,12], 160 x 160 [19], 224 x 224 [14], 227 x 227 [2], 240 x 480 [22], and 320 x 640 [9].

4.3. Image Augmentation

After obtaining an image patch with a single class of defect, the data augmentation step is performed to increase the number of datasets employed. This is intended to increase training accuracy and avoid overfitting. Several reviewed articles employ the image augmentation parameters, such as random rotation, horizontal and vertical flipping, random saturation, contrast adjustments, resizing, and clipping. Meanwhile, Zhang et al. [7] generated an artificial image patch with a defect using a data augmentation technique based on Wasserstein generative adversarial networks (WGAN). Figure 7 illustrates an example of the results of the image cropping and data augmentation method.

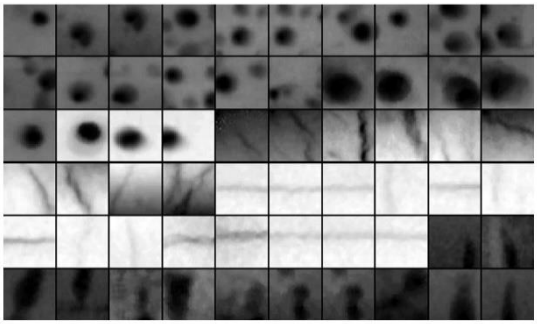


Fig. 7. The example of image cropping and data augmentation results [18].

After performing an image augmentation method, the number of image datasets for each type of defect is obtained. The dataset is then distributed for testing, validation, and training purposes. We have highlighted several distinct types of dataset distribution composition, as illustrated in Table 1.

Table 1. Composition of image datasets

Literature	Composition (%)		
	Training	Validation	Testing
[1]	60	20	20
[2], [6]-[9], [11]	80	-	20
[12]	70	15	15
[14]	75	-	25
[15]	94	6	-
[19]	40	30	30

Meanwhile, we were unable to find sufficient information on the dataset's composition from articles not listed in Table I.

5. Evaluation Indicator

The created image datasets serve as training, validation, and testing data for the proposed weld defect detection and classification model. A series of evaluation indicators is presented to verify the detection and classification results of weld defects and the performance of the proposed network model. For measuring training and testing performance, the recall, precision, F1-score, and accuracy are calculated using a confusion matrix. These indicators enable us to choose the most effective models and compare various models. The confusion matrix is depicted in Figure 8.

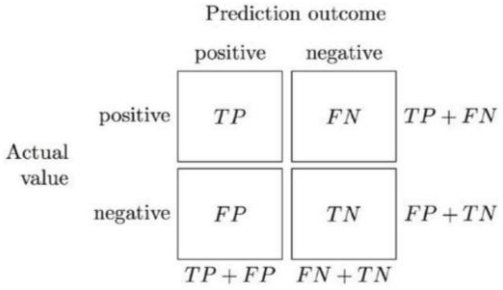


Fig. 8. Confusion matrix [16].

The confusion matrix, meanwhile, provides individual class-level insight into true and false classifications, as well as error types. The terms TP (true positive), TN (true negative), FP (false positive), and FN (false negative), which are resolved from the cells of the confusion matrix, define the precision (P) and recall (R) scores. The F1-score is the harmonic mean of P and R. Accuracy (Acc) is used to calculate the proportion

of correctly identified pixels. Yang et al. [9] also utilized additional evaluative metrics, including specificity (Sp), sensitivity (Se), area under the curve (AUC), which represents the area under the curve, such as the P-R Curve, which is between 0.1 and 1, and dice coefficient (Dice). The formulas for all of these metrics are as follows:

$$P = \frac{TP}{TP+FP} \quad (1)$$

$$S_e = R = \frac{TP}{TP+FN} \quad (2)$$

$$F1 = \frac{2PR}{P+R} \quad (3)$$

$$A_{CC} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$S_p = \frac{TN}{TN+FP} \quad (5)$$

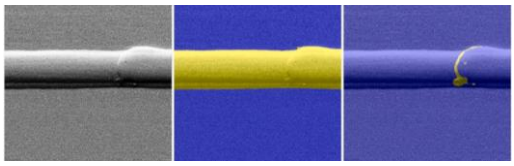
$$Dice = \frac{2|A \cap B|}{|A|+|B|} \quad (6)$$

where A and B represent two evaluation objects.

6. Defect Detection

The weld image contains not only the weld but also information about the background. The background is the portion of an image that does not require analysis. In welding defect detection, the area of the weld seam with defects becomes the target for identification. Subtracting the original image's background requires a segmentation method for identifying defects. Some articles include defect detection as part of defect classification. In this paper, however, we focused on the defect detection stage, i.e., determining the presence of defects in a radiographic image, also known as defect segmentation, excluding the classification of defect types. Yang et al. [9] proposed a defect detection method based on an Improved U-Net network to achieve high-precision automatic weld defect location. Created by the U-Net network, the proposed method can enhance radiographic image segmentation performance. In the presence of complex radiographic images with poor texture and contrast, a proposed model is required for the automatic and accurate localization of weld defects.

Sizyakin et al. [13] proposed an approach to detect weld defects using a combination of CNN and SVM. CNN is primarily used for defect and non-defect classifications. SVM is used to define defect boundaries accurately. To improve detection performance, the morphological filtration method was implemented. Golodov et al. [16] performed image segmentation in two stages: segmentation of the weld area from a whole image using the foreground segmentation network, or FgSegNet_v2 architecture, followed by segmentation of the weld defect using the FgSegNet_M architecture. Hou et al. [17] presented a three-stage algorithm for the automatic recognition of defects. Identifying a weld area using Otsu's method is the initial stage. Then, a sparse auto-encoder network (SAE) is trained and tested using radiograph image patches. Finally, a sliding-window approach is used to detect defects in a whole seam. Dong et al. [20] utilized a U-Net network by replacing its final softmax layer with RF to achieve better results. Figures 9 and 10 illustrate examples of weld defect detection results.



(a) (b) (c)

Fig. 9. Detection result using FgSegNet (a) Input image (b) Weld mask (c) Defect mask [16].

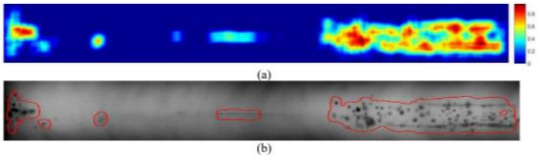


Fig. 10. Detection results using SAE (a) The probability map for defects (b) The detection result [17].

Figure 9(b) shows that the gray input image is segmented, resulting in a yellow welding area. Figure 9(c) also shows segmentation of weld defects in yellow. In comparison to Figure 9, Figure 10 shows defect segmentation using a probability map. Weld defects are bordered in red in the whole image. To determine the effectiveness of the proposed models, the performance evaluation of the proposed weld defect detection models is illustrated in Table 2.

Table 2. Weld defect detection model performance

Literature	Pre-processing	Architecture	Evaluation	Result
Yang et al., 2021 [9]	Similarity transformation; Gamma transformation; Linear transformation; Gaussian noise	Improved U-Net	Se	0.860
			Sp	0.999
			Acc	0.998
			AUC	0.884
			Dice	0.818

Literature	Pre-processing	Architecture	Evaluation	Result
Sizyakin et al., 2019 [13]	Morphological filtration; Contrast-limited adaptive histogram equalization	CNN – SVM	for MF + CNN + EB: R False alarm P F1-score	 0.4552 0.0278 0.6120 0.5221
Golodov et al., 2022 [16]	Histogram normalization; Gradient transformation	FgSegNet_v2 FgSegNet_M	F1-score (weld seg.) F1-score (defect seg.)	96.94% 91.19%
Hou et al., 2017 [17]	Binary to grayscale conversion; Thresholding using Otsu's method	SAE	Acc	91.84%
Dong et al., 2019 [20]	Shearing, skewing, flipping and elastic distortion operations	U-Net – RF	AUC	0.998

The proposed model for segmenting weld defects from radiographic images is highly effective. In terms of comparison with ground-truth values, the model proposed by Yang et al. [9] is preferred. While the model proposed by Golodov et al. [16] can detect not only weld defects but also the weld area relative to the base material, Therefore, these two models are more strongly recommended.

7. Defect Classification

ISO-6520-2:2013, "Welding and Allied Processes—Classification of Geometric Imperfections in Metallic Materials" [4] describes in detail the classes of weld defects. According to the articles that we have summarized, cracks, pores, lack of penetration, lack of fusion, solid inclusions, including slag and tungsten inclusions, undercuts, burn-through, excessive root penetration, concave root, overlap, and spatter are the most common classes of defects found in radiographic images. However, [9], [13], [17], and [20] did not specify the type of defect used, as the model only detects and classifies binary results, namely defects and non-defects. The choice of classifier is another crucial factor influencing defect classification performance. At the beginning of the study, defect classification was accomplished primarily through the use of traditional machine learning algorithm approaches. The classification results of these methods are always constrained by the quality of the designed features and the amount of image data. Deep learning technology can automatically extract deep representative features from a radiographic image, thereby eliminating the need to manually extract features. A further advantage of deep learning is that the network can utilize unlabeled data, like in the SAE [17] and DBN [22] architectures.

According to the articles we reviewed, the deep learning architectures are quite diverse, but they all refer to the CNN architecture, which is modified into several types based on the layer structure as shown in Table III. Thakkallapally [1] proposed a VGG-19-based CNN that was trained using transfer learning on a sample of 3000 radiographic images of size 128 x 128 pixels and belonging to three classes. Ajmi et al. [2] proposed a classification model based on the architecture of the pretrained network AlexNet, then compared the performance with deep convolutional activation features, or DCFA, GoogLeNet, VGG-16, VGG-19, ResNet50, and ResNet101. In the field of detecting steel pipe weld defects, Yang et al. [5] proposed object detection based on the YOLOv5 algorithm and compared it to the two-stage representative object detection algorithm Faster R-CNN. Pan et al. [6] proposed a classification model based on the TL-MobileNet structure by adding a full connection layer and a softmax classifier into the MobileNet architecture. Zhang et al. [7] proposed a WGAN-based method for image augmentation to address the limited number of datasets. Then, two DCNNs are trained on the augmented image sets using transfer learning techniques based on feature extraction. The two trained CNNs are combined in a multi-model ensemble framework to classify weld defects.

Yang et al. [8] proposed a three-stage defect classification model: define 11 welding defect features as inputs, construct a deep learning framework (unified DNN) in which multi-level features are included, and additionally, investigate pre-training and fine-tuning strategies to improve generalization performance with limited datasets. Naddaf et al. [10] developed and annotated over 100,000 radiographic images of various welds. On the basis of these data and annotations, an optimized CNN for defect classification is designed and trained. Guo et al. [11] utilized a generative adversarial network combined with transfer learning to address the data imbalance and improve accuracy. The Xception model is proposed as a feature extractor in the target network for transfer learning, while frozen–unfrozen training is used to fine-tune the classification model. Nazarov et al. [12] utilized the VGG-16 architecture for classification to solve the problem of transferring training due to the limited number of training images. To extract the

features of this dataset of weld defects, Kumaresan et al. [14] used transfer learning with pre-trained DCNNs, including VGG16 and ResNet50. On these extracted features, machine learning models such as LR, SVM, and RF were then trained. Oh et al. [15] proposed a faster R-CNN to classify welding defects automatically. To properly extract the radiographic testing image's features, two internal Faster R-CNN feature extractors were selected, compared, and their performances evaluated.

Golodov et al. [16] performed a classification of the weld area using EfficientNet-B3, followed by a classification of the weld defects using EfficientNet-B0. Hou et al. [18] created a model based on a DCNN for extracting deep features from radiographic images. In consideration of the imbalance in the number of image patches with different weld defects, three types of resampling techniques, namely ROS, RUS, and SMOTE, are employed to develop three balanced datasets. Using datasets, the classification abilities of five extracted types of features are compared using traditional and deep learning methods. Liu et al. [19] proposed the AF-RCNN algorithm to classify weld defects. ResNet and FPN serve as the network's backbone, while lightweight model channel attention and spatial attention mechanisms are utilized. Liu et al. [21] proposed a three-stage method for the classification of weld defects using a triplet DNN. The initial radiographic image is transformed into a relief image. Second, the feature vector is obtained by mapping the relief image. Finally, the SVM classifier detected the weld defect. Chang et al. [22] proposed a classification model based on DBN that classifies the weld feature curves extracted. A cylindrical projection method is proposed to increase the proportion of defect parts and address the problem of small defect loss. Finally, a SegNet-based system for classifying weld defects is proposed. As illustrated in Figures 11 and 12, the example results of defect classification are tested on whole radiographic images.

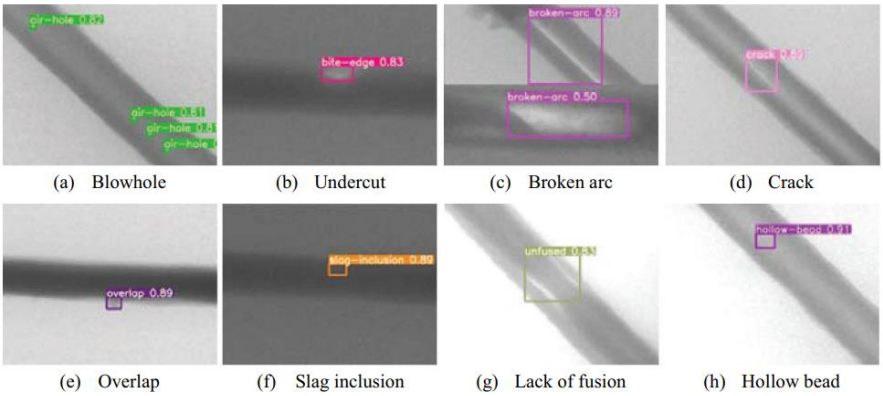


Fig. 11. Detection and classification results using YOLOv5 [5].

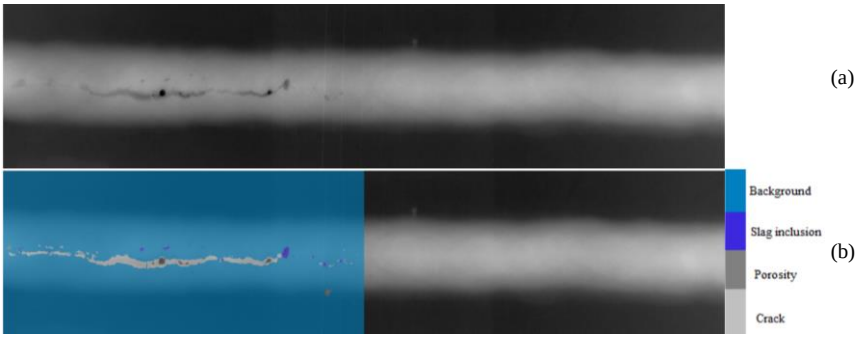


Fig. 12. Detection and classification result using WDC-SegNet (a) Original images (b) Defects recognition [22].

Figure 11 depicts the results of the detection of eight distinct images containing defects, indicated by colored boxes along with the defect's name and degree of accuracy. Figure 12 depicts a single image for three distinct types of defects. To determine the performance of the proposed models, Table 3 presents the results of the performance evaluation of the proposed classification models.

Table 3. Weld defect classification model performance

Literature	Pre-processing	Architecture	Number of Defect Classes	Evaluation	Result
Thakkallapally, 2019 [1]	-	VGG-19	3	Train Acc Val Acc Test Acc	93.17% 91.14% 91%
Ajmi et al., 2020 [2]	Conversion to RGB	AlexNet	2	Acc	100%
Yang et al., 2021 [5]	Motion deblurring; Light change; Gaussian noise; Saturation adjustment; Contrast and Sharpness	YOLOv5	8	mAP (mean of Average Precision)	98.7%
Pan et al., 2020 [6]	-	MobileNet	5	mAcc	97.69%
Zhang et al., 2019 [7]	Gaussian filter; Data augmentation with WGAN	Inception & MobileNet	4	Acc (normal) Acc (burn through) Acc (crack) Acc (porosity)	100% 94.77% 99.75% 99.67%
Yang et al., 2021 [8]	Image quality improvement	Unifed DNN	5	Train Acc Test Acc	97.95% 91.36%
Naddaf et al., 2020 [10]	-	VGG-16, VGG-19, AlexNet, ResNet	11	Acc	96%
Guo et al., 2021 [11]	Contrast enhancement (CECGAN)	Xception	5	F1-score Acc	0.909 92.5%
Nazarov et al., 2021 [12]	-	VGG-16	5	Acc	86%
Kumaresan et al., 2021 [14]	-	VGG-16 & ResNet-50	9 14	Acc (9 classes) Acc (14 classes)	99.4% 97.8%
Oh et al., 2020 [15]	Noise reduction; Contrast enhancement	Faster R-CNN, ResNet, Inception	2	mAP (ResNet with aug.)	0.532
Golodov et al., 2022 [16]	Histogram normalization; Gradient transformation	EfficientNet-B0	8	Acc (Top1) Acc (Top2)	82.73% 96.76%
Hou et al., 2019 [18]	Resampling (ROS, RUS, SMOTE); Reshape	DCNN	4	Acc (DCNN2): ROS Dataset RUS Dataset SMOTE Dataset	 96.3% 79.9% 97.2%
Liu et al., 2022 [19]	Image resizing	AF-RCNN	6	mAP	85.4%
Liu et al. 2020 [21]	Contrast enhancement; Image to relief conversion	Triplet DNN – SVM	5	Acc (multi-defects): Defect 1 Defect 2 Defect 3 Defect 4	 0.70 0.90 1.00 0.80
Chang et al., 2021 [22]	-	DBN, WDC-SegNet	4	Acc	98.6%

In terms of classification, weld defects that occur during the welding process can vary. Consequently, a model with the ability to classify more types of defects has better results. Based on the data in Table 3, a model that can classify eight or more defect classes has good performance for real applications, including the models proposed by Yang et al. [5], Naddaf et al. [10], Kumaresan et al. [14], and Golodov et al. [16]. However, it can be difficult to classify infrequent types of defects using these models. Thus, research into the development of a model capable of classifying more types of defects is continuous.

8. Discussion

The detection and classification of weld defects in radiographic images is one of the most essential and challenging welding inspection tasks. The limitations of manual or visual image evaluation and interpretation require an automated image inspection method utilizing computational technology. Researchers in the field of artificial intelligence have developed a deep learning-based method for inspecting weld defects in radiographic images. As a result, the deep learning architecture is able to become a model for detecting and classifying welding defects with good performance based on test results using evaluation indicators, particularly those that are most frequently associated with accuracy. The most fundamental aspect of this stage of detecting welding defects is that the model can distinguish the defective image from the background image. This becomes difficult when the evaluated radiographic images have poor quality, low contrast, noise, and an inconsistent grayscale

distribution. The result indicated that a series of image pre-processing methods, including motion deblurring [5], contrast enhancement [11], [15], [21], noise removal [15], image conversion [21], as well as image segmentation techniques, are capable of resolving this problem. In terms of defect classification, the model is able to recognize a diversity of defect types in a variety of classes, such as [5], [10], [14], and [16]. When the model is presented with a small number of datasets and there is an imbalance between classes, the classification of defects can become difficult. The limited number of datasets can be overcome by employing image augmentation techniques that vary the orientation and position of image defects artificially. The generative adversarial network method [11] and the resampling method [18] have been used to address the issue of imbalanced data. In summary, it can be determined that this classification model supports the evaluation and interpretation of weld defects in radiographic images by NDT specialists.

To get better detection and classification results, several recommendations can be considered. The stages of image preprocessing are determined according to the quality of the initial radiographic image. The pre-processing stages for radiographic images derived from conventional film digitization may differ from those derived from DDA or computed radiography (CR). An imaging plate records the radiographic image of CR, and during its use, artifacts in the form of scratches and spots can appear on the radiographic image [26]. The presence of these artifacts is disruptive to the image of the weld, as they can obscure defects that need to be inspected. Before an image is used as a dataset, the application of certain image filters is thus one of the most crucial image pre-processing parameters [27].

Important to the improvement of the defect classification model is the issue of the number of defect classes classified. The majority of studies have been able to classify the most common types of welding defects, including cracks, porosity, gas pores, lack of penetration, lack of fusion, slag inclusions, etc. According to ISO 6520-2:2013, there are approximately 60 types of defects, which are explained in detail and grouped into six major categories [4]. If the classification model can classify more types of defects, it will perform well. However, this will be problematic for types of defects that occur infrequently, as the amount of data available for these classes of defects will be extremely limited and imbalanced. In the future, it will be necessary to consider effective image augmentation methods in order to overcome this issue.

According to the articles we summarize, the weld defect dataset is comprised of radiographic images with a more elongated weld orientation than whole radiographic images. As a follow-up to the research conducted by [23-25] and [28], it will be necessary to conduct future studies on weld defect detection and classification based on radiographic image datasets with elliptical orientation or in radiographic techniques known as the double wall double image technique (DWDI). Figure 13 is an illustration of a radiographic image displaying an elliptical welding orientation.

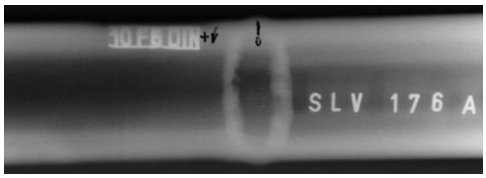


Fig. 13. A radiographic image sample using the DWDI technique [3].

Considering that the stages of detection and classification of weld defects are only a part of the entire radiographic image inspection, future research challenges are quite interesting to study. Following the detection and classification of weld defects, one of the subsequent advanced steps involves measuring the dimensions and depth of weld defects. Measuring the size of the defect is crucial because the size of a similar type of defect can differ from the others. In determining the performance limits of welded joints in materials, the size of the defect has to be considered. The size of the defect is one of the indicators that the welding process and quality control system require improvement. Therefore, it is highly probable that deep learning methods combined with computer vision can be utilized to develop a weld defect classification model that can automatically measure the dimensions and depth of defects in radiographic images.

In addition, it is necessary to conduct research on the detection and classification of welding defects in real time during the welding process in order to ensure the implementation of proper welding procedures. In actual conditions, welding defects occur due to errors in procedures, equipment, and environmental influences during the welding process. Because it needs to be understood that the model of detection and classification of weld defects cannot directly reduce or eliminate these defects, the results of the detection and classification of the types of weld defects that appear can be used as material for evaluating and improving procedures, equipment, and the environment during the welding process. For example, if there is a crack-type defect in a

welding area, this must be prevented in the future by preheating the metals, gradually cooling the weld joints, and maintaining acceptable weld joint gaps.

9. Conclusion

It is essential to test material-welded joints to ensure their safety and compliance with design and service specifications. Manual or visual evaluation and interpretation of conventional radiographic film are complex, time-consuming, subjective, inconsistent, and occasionally biased with respect to defects with similar features. Therefore, researchers have attempted to automate the evaluation and interpretation using deep learning-based computer assistance. As part of the evaluation and interpretation process, this article discussed an automation technique based on deep learning for detecting and classifying weld defects. This article discussed techniques for inspecting weld defects using digital radiographic images. This literature review provides information on the development of a model for the detection and classification of defects, whose stages involve data collection, image pre-processing, defect detection, and defect classification. First, the radiographic images used as datasets are categorized into two groups: public databases and private databases. Second, a series of image pre-processing methods are performed to obtain images that are adequate for use as training, validation, and testing datasets. Third, the defect detection and classification model is constructed using the obtained dataset, and its performance is evaluated using evaluation indicators. Concerning the detection of weld defects, the closeness of the segmentation results to the ground-truth value is of utmost importance. The model proposed by Yang et al. [9] and Golodov et al. [16] is therefore preferred. In terms of classification, models with the ability to classify more types of defects are preferred, including the model proposed by Yang et al. [5], Naddaf et al. [10], Kumaresan et al. [14], and Golodov et al. [16], which can classify eight or more classes of weld defects. Finally, the achievements of the defect detection and classification model are described, and recommendations are made to improve future research outcomes.

References

[1] B. C. Thakkallapally, “Defect classification from weld radiographic images using VGG-19 based convolutional neural network,” in *NDE 2019*, vol. 018, 2019.

[2] C. Ajmi, J. Zapata, S. Elferchichi, A. Zaafouri and K. Laabidi, "Deep learning technology for weld defects classification based on transfer learning and activation features," in *Advances in Materials Science and Engineering*, 2020, vol. 2020.

[3] D. Mery, V. Riffo, U. Zscherpel, G. Mondragón, I. Lillo, I. Zuccar, H. Lobel and M. Carrasco, “GDXray: the database of x-ray images for nondestructive testing,” in *Journal of Nondestructive Evaluation*, vol. 34.4, 2015, pp. 1-12.

[4] International Organization for Standardization. “Welding and Allied Processes – Classification of Geometric Imperfections in Metallic Materials – Part 2: Welding with Pressure (ISO 6520-2:2013).” Geneva, 2013.

[5] D. Yang, Y. Cui, Z. Yu and H. Yuan, “Deep learning based steel pipe weld defect detection,” in *Applied Artificial Intelligence*, vol. 35:15, 2021, pp. 1237-1249.

[6] H. Pan, Z. Pang, Y. Wang, Y. Wang and C. Lin, “A new image recognition and classification method combining transfer learning algorithm and MobileNet model for welding defect,“ in *IEEE Access*, vol. 8, 2020, pp. 119951-119960.

[7] H. Zhang, Z. Chen, C. Zhang, J. Xi and X. Le, “Weld defect detection based on deep learning method,” in *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, Vancouver, BC, Canada, 2019, pp. 1574-1579.

[8] L. Yang and H. Jiang, “Weld defect classification in radiographic images using unified deep neural network with multi-level features,” in *Journal of Intelligence Manufacture*, vol. 32, 2021, pp. 459–469.

[9] L. Yang, H. Wang, B. Huo, F. Li and Y. Liu, “An automatic welding defect location algorithm based on deep learning,” in *NDT & E International*, vol. 120, 2021, pp. 102435.

[10] M. Naddaf, S. Naddaf, H. Zargaradeh, S. M. Zahiri, M. Dalton, G. Elpers and A. R. Kashani, “Next-generation of weld quality assessment using deep learning and digital radiography,” in *Artificial Intelligence in Manufacturing, AAAI Spring Symposium Series*, California, USA, 2020, vol. 2020.

[11] R. Guo, H. Liu, G. Xie and Y. Zhang, “Weld defect detection from imbalanced radiographic images based on contrast enhancement conditional generative adversarial network and transfer learning,” in *IEEE Sensors Journal*, vol 21, no. 9, 2021, pp. 10844-10853.

[12] R. M. Nazarov, Z. M. Gizatullin and E. S. Konstantinov, “Classification of defects in welds using a convolution neural network,” in *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, Moscow, Russia, 2021, pp. 1641-1644.

[13] R. Sizyakin, V. V. Voronin, N. Gapon, A. Zelensky, and A. Pizurica, “Automatic detection of welding defects using the convolutional neural network,” in *Automated Visual Inspection and Machine Vision III*, Munich, Germany, 2019, vol. 11061.

[14] S. Kumaresan, K. S. J. Aultrin, S. S. Kumar and M. D. Anand, “Transfer learning with CNN for classification of weld defect,” in *IEEE Access*, vol. 9, 2021, pp. 95097-95108.

[15] S. Oh, M. Jung, C. Lim, and S. Shin, “Automatic detection of welding defects using faster R-CNN,” in *Applied Sciences*, 2020, vol. 10, no. 23, pp. 8629.

[16] V. A. Golodov and A. A. Maltseva, “Approach to weld segmentation and defect classification in radiographic images of pipe welds,” in *NDT & E International*, vol. 127, 2022, pp. 102597.

[17] W. Hou, Y. Wei, J. Guo, Y. Jin and C. Zhu, “Automatic detection of welding defects using deep neural

network," in *Journal of Physics: Conference Series*, Banda Aceh, Indonesia, 2017, vol. 933.

[18] W. Hou, Y. Wei, J. Guo, Y. Jin and C. Zhu, "Deep features based on a DCNN model for classifying imbalanced weld flaw types", in *Measurements*, vol. 131, 2019, pp. 482–489.

[19] W. Liu, S. Shan and H. Chen, "X-ray weld defect detection based on AF-RCNN," in Welding in the World, vol. 66, 2022, pp. 1165–1177.

[20] X. Dong, Taylor, C.J. Taylor and T.F. Cootes, "Small defect detection using convolutional neural network features and random forests," in *Computer Vision – ECCV 2018 Workshops*, Munich, Germany, 2019, vol. 11132.

[21] X. Liu, J. Liu, F. Qu, H. Zhu and D. Lu, "A weld defect detection method based on triplet deep neural network," in *2020 Chinese Control and Decision Conference (CCDC)*, Hefei, China, 2020, pp. 649-653.

[22] Y. Chang and W. Wang, "A deep learning-based weld defect classification method using radiographic images with a cylindrical projection," in *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021, pp. 1-11.

[23] C. C. B. Fioravanti, T. M. Centeno and M. R. De Biase Da Silva Delgado, "A deep artificial immune system to detect weld defects in DWDI radiographic images of petroleum pipes," in *IEEE Access*, vol. 7, 2019, pp. 180947-18096.

[24] F. M. Suyama, M. R. Delgado, R. D. da Silva and T. M. Centeno, "Welded joint detection of oil pipelines in radiographic images with Double Wall Double Image exposure," in *NDT & E International*, vol. 105, 2019, pp. 46-55.

[25] M. Kroetz et al., "Genetic algorithms to automatic weld bead detection in double wall double image digital radiographs," in *2012 IEEE Congress on Evolutionary Computation*, 2012, pp. 1-7.

[26] The American Society of Mechanical Engineers. "Radiographic Examination (ASME V Article 2)." New York City, 2017.

[27] International Atomic Energy Agency. "Guidelines on training, examination and certification in digital industrial radiology testing (RT-D)." Training Course Series No. 60, IAEA, Vienna, 2015.

[28] N. Boaretto and T. M. Centeno, "Automated detection of welding defects in pipelines from radiographic images DWDI," in *NDT & E International*, vol. 86, 2017, pp. 7-13.