

## Article

# Sizing-Based Flaw Acceptability in Weldments Using Phased Array Ultrasonic Testing and Neural Networks

Seung-Eun Lee <sup>1</sup>, Jinhyun Park <sup>1</sup>, Yun-Taek Yeom <sup>2</sup> , Hak-Joon Kim <sup>1</sup> and Sung-Jin Song <sup>1,\*</sup>

<sup>1</sup> Department of Mechanical Engineering, College of Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea

<sup>2</sup> Department of Smart Mechanical Components and Materials, Dongyang University, Yeongju 36040, Republic of Korea

\* Correspondence: sjsong@skku.edu; Tel.: +82-31-299-6625

**Abstract:** Liquefied Natural Gas (LNG) is one of the major renewable energy sources and is stored and carried in a storage tank that is designed following international standards. Since LNG becomes highly unstable when it encounters oxygen in the air, a leakage from an LNG storage tank can cause a catastrophic industrial accident. Thus, the inspection of LNG storage tanks is one of the priorities to be completed before LNG is stored in a storage tank. Recently, the usage of Phased Array Ultrasonic Testing (PAUT) has been gradually increasing as the risks of RT emerge. PAUT has some obstacles to overcome in order to substitute RT, such as efficiency and accuracy. Specifically, the cost issue must be addressed. Therefore, many attempts to combine PAUT with Artificial Neural Networks (ANN) have been made. PAUT provides many types of 2D images of the inspected weldment. The S-scan is one of the 2D images provided by PAUT, and it displays the cross-sectional view of the specimen with a single transducer. The inspectors examine the S-scan image and other provided images of PAUT to detect, classify and size the flaw that exists in the weldment so that the decision of whether the inspected weldment with the flaw is acceptable can be made. Nowadays, most of the previous research on PAUT and ANN focuses on detecting and classifying the flaws in B-scan or S-scan images. However, the last step to determine the flaws' acceptability is not yet covered. In this study, the flaw acceptance criteria of PAUT in various international standards are listed. EXTENDE CIVA is used to create the PAUT S-scan images. The S-scan images are labeled with the listed acceptance criteria. Then, they are used in Mask R-CNN training. After the training, some new S-scan images with flaws are used to test the performance, and this showed 96% precision and 87% recall. With the algorithm, the acceptability of a flaw in a weldment can be determined efficiently and it will reduce the burden of PAUT usage and reduce the time required for a full-length inspection.

**Keywords:** Phased Array Ultrasonic Testing; Mask R-CNN; flaw acceptance criteria; Artificial Neural Network



**Citation:** Lee, S.-E.; Park, J.; Yeom, Y.-T.; Kim, H.-J.; Song, S.-J. Sizing-Based Flaw Acceptability in Weldments Using Phased Array Ultrasonic Testing and Neural Networks. *Appl. Sci.* **2023**, *13*, 3204. <https://doi.org/10.3390/app13053204>

Academic Editor: Marco Scalerandi

Received: 19 January 2023

Revised: 28 February 2023

Accepted: 28 February 2023

Published: 2 March 2023



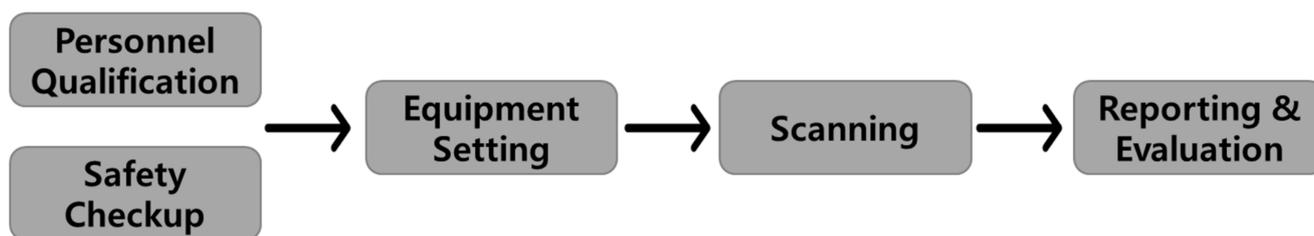
**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Liquefied Natural Gas (LNG) is a liquefied form of natural gas, typically methane, which is developed for convenience of storage, where the liquid state occupies 600 times less space than the gaseous state. LNG was first introduced in the 1910s and was used in the 1940s to substitute fossil fuels such as coal and petroleum [1]. Now, LNG has become one of the major renewable energy sources and is employed around the world. However, the storage of LNG is a major task, since, although the liquid state is more stable than the gaseous state, when exposed to the air, LNG can vaporize and return to a gaseous state, and become extremely unstable. For the safe storage, carrying, and receiving of LNG, a massive storage tank composed of 9% nickel steel or stainless steel should be used, since austenitic metal alloys such as 9% nickel steel do not have a ductile and brittle transition temperature, hence being more stable under brittle fracture. The storage tanks, after being

constructed, must be inspected via Nondestructive Evaluation (NDE) methods so that the possible risk of leakage is prevented. If not, there is a huge risk of catastrophic industrial accidents. The world has already witnessed serious LNG leakage accidents, such as the disaster caused by the leakage of LNG in 1944 [2].

Radiographic Testing (RT) is one of the main NDE methods used to inspect and evaluate LNG storage tanks. RT is a more efficient and practical NDE method that brings more precise and practical results than other NDE methods. However, the danger of radiation exposure when inspecting with RT has become an emerging issue recently. Phased Array Ultrasonic Testing (PAUT) is one of the advanced derivatives of the Ultrasonic Testing (UT) methods in NDE. The market share of UT has dramatically increased to 31.10% [3] in lieu of RT to avoid the risks that have emerged lately. Moreover, PAUT is an image-based testing method, similar RT. However, the accuracy and efficiency of PAUT still does not reach that of RT, and these remain as obstacles to the substitution of RT by PAUT. Compared with RT, PAUT needs more skilled and experienced operators due to the complexity in the procedure of evaluation. Using steering and focusing techniques, the ultrasonic beam can be controlled and provide many other types of image-based outcomes. However, PAUT still uses an ultrasonic beam, and ultrasonic beams are reflected, deflected, and attenuated by various geometrical aspects. These aspects make the outcome of PAUT extremely difficult to evaluate, even for skilled experts. To overcome these obstacles, many attempts to combine PAUT with Artificial Neural Networks (ANN) have been made. Figure 1 below briefly introduces the overall procedure of PAUT [4].



**Figure 1.** Overall procedure of PAUT.

PAUT uses the steering and focusing of an ultrasonic beam produced by the phased arrayed elements in a single transducer to penetrate the specimen. Nowadays, the PAUT equipment usually saves the scanned data so that they can be evaluated after the scanning is done. The evaluation requires much experience, since it must correctly follow the internationally implemented standards, and has a high cost, so attempts to automate this procedure with ANN or other Artificial Intelligence (AI) techniques have been previously made. Song et al. [5] used probabilistic neural networks for ultrasonic weldment flaw signals. Park et al. [6] used the system invariant method and residual neural network for ultrasonic flaw classification in weldments. Munir et al. [7] used the dropout technique for ultrasonic flaw classification in weldments. Munir et al. [8] also used a Convolutional Neural Network (CNN) for ultrasonic flaw classification in noisy conditions. Cruz et al. [9] used neural networks for feature selection in metal weldment UT. Sambath et al. [10] also used ANN for feature selection in an ultrasonic A-scan. Virupakshappa et al. [11] used a CNN for ultrasonic A-scan flaw detection. Medak et al. [12] used the EfficientDet Network for flaw detection in PAUT B-scan images. Siljama et al. [13] used a CNN for flaw detection in PAUT B-scan images. Ho et al. [14] used a two-stage ANN for flaw classification in PAUT S-scan images.

Most of the previous research, except for [5], focused on the detection of the flaw or classification of the flaw using neural networks. Detection of the flaw mainly focuses on determining whether the flaw exists in various scanning methods. Classification of the flaw seeks to distinguish the type of the flaw among various types of existing flaws. However, these attempts cannot practically evaluate whether the flaw is dangerous, because they lack the procedure of sizing of the flaws detected by PAUT. The sizing of PAUT outcomes

according to API 620 [15] and ISO 11666 [16] aims to set up a threshold for similar types of flaws so that the flaws can be determined acceptable or not. When the sizing process of flaw images obtained by PAUT can determine the flaw's acceptability following the size of the flaw, then the class of the flaw can be divided into two variables, acceptable or not acceptable. Until now, the sizing of flaw data obtained by PAUT was not considered important, but the sizing procedure must be handled and automated with the ANN so that the efficiency of PAUT can be enhanced.

Mask R-CNN is a two-stage object detection model that is trained with images that are masked and annotated by classes. If the classes when training Mask R-CNN are labeled according to the sizes of flaws, Mask R-CNN will be able to determine the flaws' acceptability.

In this paper, the flaw will be determined according to the reasonable acceptance criteria with a two-stage object detection ANN model, Mask R-CNN. The data will be designed and obtained using the renowned ultrasonic simulation tool, EXTENDE CIVA, following the acceptance criteria that are described in this paper below, referring to API 620 Annex U. Since the ultrasonic S-scan image completely volumetrically presents both vertical and horizontal directions of the weldment and the flaw, a simulated ultrasonic S-scan will be a useful data tool for the training of Mask R-CNN. The data will then be masked and annotated with an image annotation tool called VGG Image Annotator. The class of the flaw data obtained with PAUT must be divided into two classes, acceptable or not acceptable (NA), according to the size of the flaw. After the masking process, the data will be used to train the Mask R-CNN model and the performance of the trained model will be tested.

## 2. International Standards for the Sizing of PAUT

Table 1 below shows the international standards that are used in the flaw sizing of PAUT in weldments. API Standard 620 is referred to in this study because the standard indicates the methods used to design and construct LNG storage tanks.

**Table 1.** The international standards of PAUT flaw sizing technique.

Standards	Title	Edition
API Standard 620 [15]	Design and Construction of Large, Welded, Low-Pressure Storage Tanks	12th Ed., 2013, Addendum 2, 2018
ISO 13588 [4]	Non-destructive testing of welds—Ultrasonic testing—Use of automated phased array technology	2nd Ed., 2019-02
ISO 22825 [17]	Non-destructive testing of welds—Ultrasonic testing—Testing of welds in austenitic steels and nickel-based alloys	3rd Ed., 2017-09
ISO 17640 [18]	Non-destructive testing of welds—Ultrasonic testing—Techniques, testing levels, and assessment	4th Ed., 2018-10
ISO 19285 [19]	Non-destructive testing of welds—Phased Array Ultrasonic Testing (PAUT)—Acceptance levels	1st Ed., 2017-08
ISO 11666 [16]	Non-destructive testing of welds—Ultrasonic testing—Acceptance levels	2nd Ed., 2018-01

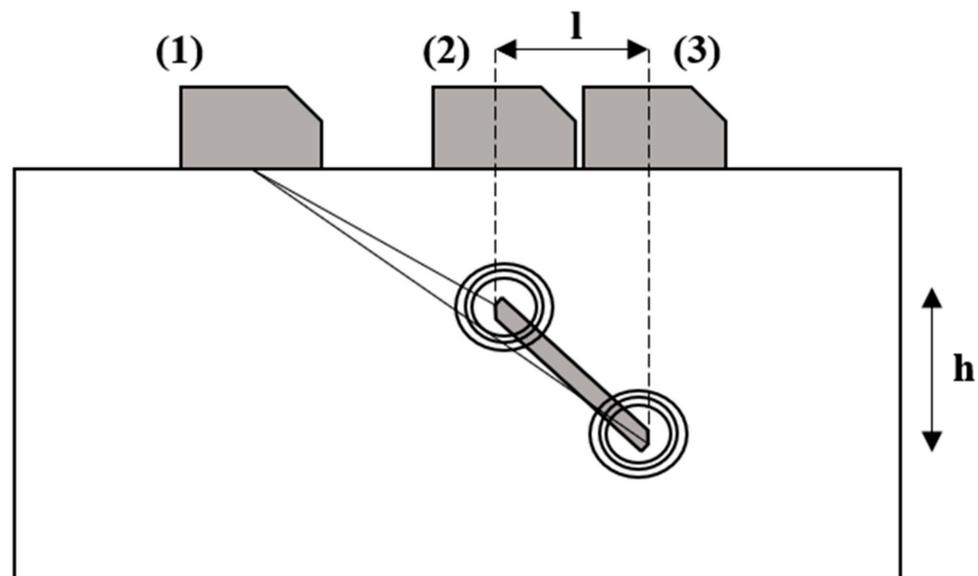
Amongst the standards, API 620 Annex U presents the acceptance criteria for the flaws in weldments. The acceptability of the flaws can be determined according to the acceptance criteria presented in API 620 Annex U, since the acceptance criteria are the output of fracture mechanical experiments. ISO 13588 is referred to because it is an international standard for the use of automated PAUT. ISO 22825 is referred to because it contains the PAUT procedure for austenitic steels and nickel-based alloys. Table 2 presents the acceptance criteria listed in API 620 Annex U. Moreover, the sizing methodology of the flaw image obtained by PAUT must follow the flaw sizing methodologies introduced in API 620 Annex U and ISO 11666.

**Table 2.** Maximum flaw acceptance criteria according to the thickness of weldments, Adapted with permission from Ref. [15]. Copyright 2014, American Petroleum Institute.

Thickness at Weld, (t) mm	Maximum Acceptance Flaw Lengths, (l) mm							
	For Surface Flaw with Height (h) mm No Greater Than			For Sub-Surface Flaw with Height (h) mm No Greater Than				
	h = 2	h = 2.5	h = 3	h = 2	h = 3	h = 4	h = 5	h = 6
6 to <10	8	4	3	5	3	Not allowed	Not allowed	Not allowed
10 to <13	8	8	4	41	5	4	Not allowed	Not allowed
13 to <19	8	8	4	38	8	5	4	3
19 to <25	8	8	4	75	13	8	6	5
25 to <32	9	8	4	100	20	9	8	6
32 to <38	9	8	4	125	30	10	8	8
38 to <44	9	8	4	150	38	10	9	8
44 to <50	9	8	4	175	60	11	9	8

### 3. Flaw Sizing Methods

Figure 2 demonstrates the sizing methods of PAUT according to ISO 11666 [16] and ISO 13588 [17]. The sizing method introduced in ISO 11666 and ISO 13588 is practically the same method that is described in API 620 Annex U. However, API 620 Annex U only describes the sizing method by UT. Therefore, the sizing methods of ISO 11666 and ISO 13588 are used for this study. Below, (1), (2), and (3) denote the scan direction. When the transducer is at position (1) of Figure 2, the height sizing is available. The ultrasonic beam is reflected by the tip and bottom of the flaw, so they are called the tip echo and bottom echo. The distance between the tip and bottom echoes is the height of the flaw. Moreover, the distance of position (2) and (3) is the length of the flaw embedded. After determining the actual size of the flaw, the flaw can be determined as acceptable or not with the acceptance criteria in Table 2.



**Figure 2.** Flaw sizing method of PAUT; (1), (2), (3) show the scan direction accordingly.

### 4. PAUT Simulation

#### 4.1. PAUT Simulation Settings

EXTENDE CIVA is a well-known NDE simulation tool used to simulate a variety of NDE methods. By using EXTENDE CIVA, various types of NDE simulations, including UT, Guided Wave Testing, and Eddy-Current Testing (ECT), are possible. PAUT also is one of the possible simulation options of EXTENDE CIVA [20].

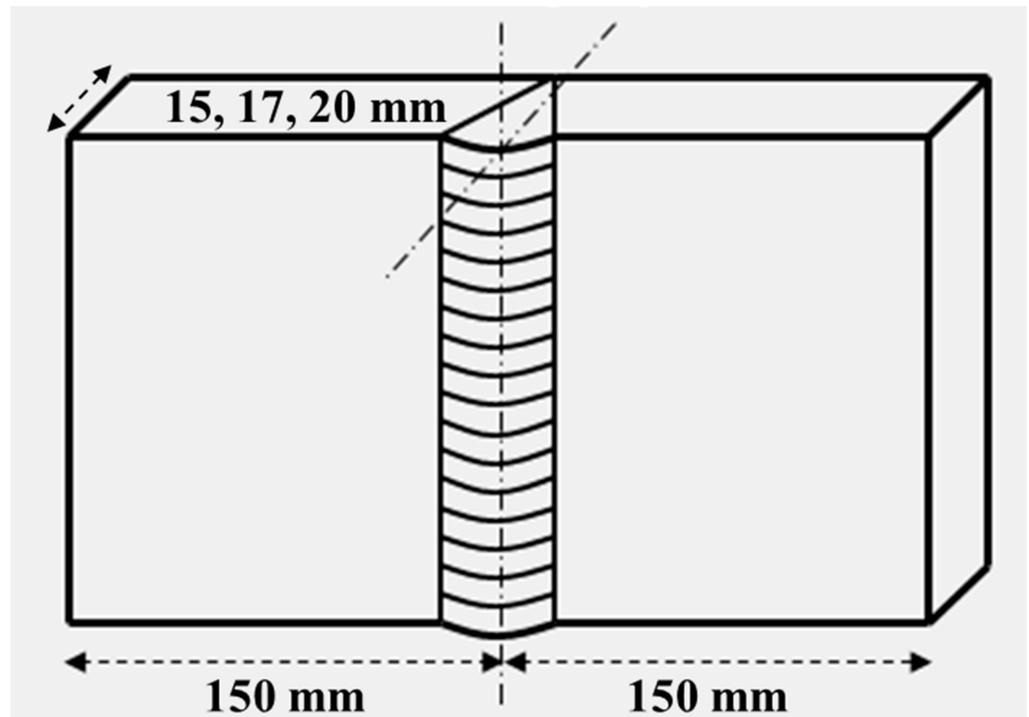
For a reasonable simulation, the specimens are designed using the EXTENDE CIVA weldment geometry setting. In this research, 10 different butt-welded specimens with 3 different heights, 300 mm × 300 mm × 15 mm ( $W \times L \times H$ ), 300 mm × 300 mm × 17 mm ( $W \times L \times H$ ), 300 mm × 300 mm × 20 mm ( $W \times L \times H$ ), are designed. The specimens are inspired by an artificial flaw specimen, shown in Figure 3, with 300 mm × 300 mm × 15 mm ( $W \times L \times H$ ). In the specimens, the weldments and the flaws are embedded. The parent material of specimens is stainless steel (SUS 304) and the weld material is Inconel. For the weld geometry, a double-V geometry is used. Moreover, the heat-affected zone (HAZ) is set to 5 mm height along the weld zone. The contact transducer for PAUT simulation also must be designed in the simulation setting. The wave type, incidence angle, and frequency are properly updated for precise simulation. The flaws and flaw sizes are designed referring to the acceptance criteria in Table 2, where there exist two types of flaws, the surface flaw and the sub-surface flaw. The set-up data are presented in Table 3. Meanwhile, Figure 4 shows the weldment geometry setting for EXTENDE CIVA.



**Figure 3.** Sample of an artificial flaw specimen.

**Table 3.** EXTENDE CIVA simulation settings.

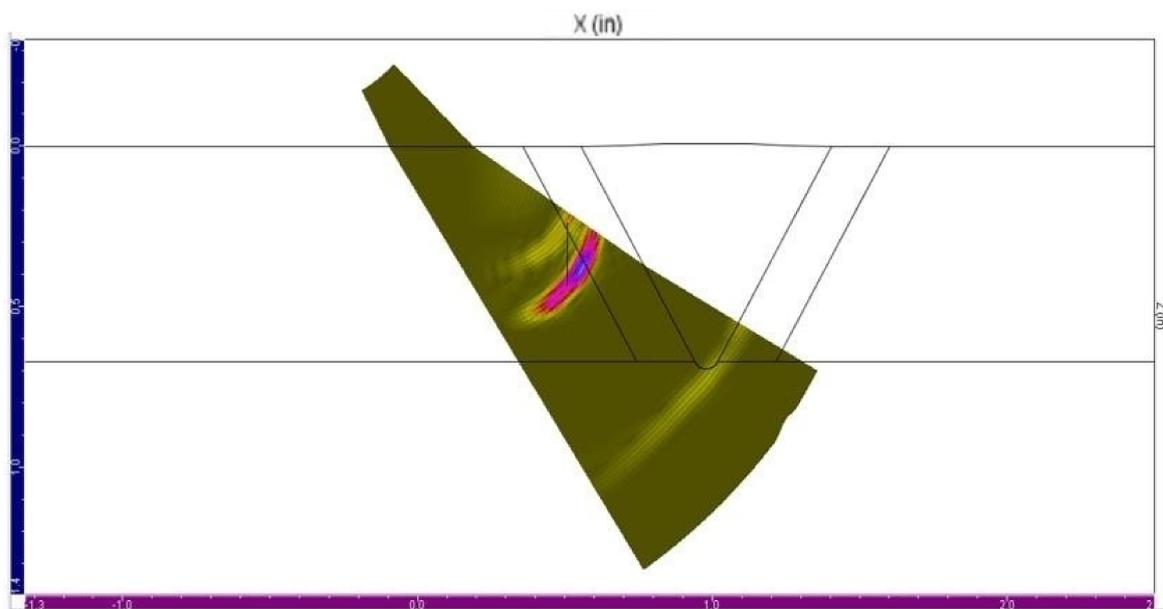
EXTENDE CIVA Simulation Settings	
Parent Material	Stainless Steel (SUS 304)
Weld Material	Inconel
Specimen Geometry	300 mm × 300 mm × 15 mm ( $W \times L \times H$ ), 300 mm × 300 mm × 17 mm ( $W \times L \times H$ ), 300 mm × 300 mm × 20 mm ( $W \times L \times H$ ).
Weld Geometry	Double-V
Heat-Affected Zone	5 mm
Wave Type	Transverse
Incidence Angle	45°
Wave Frequency	2.25 MHz
Flaw Type	Surface Flaw Sub-Surface Flaw
Flaw Geometry	Rectangle
Scan Distance	30 mm
Scanning Mode	Raster Scan
Scanning Step	1 mm



**Figure 4.** EXTENDE CIVA simulation weldment geometry settings.

#### 4.2. PAUT Simulation Results

Figure 5 is one of the S-scan results of the CIVA PAUT simulation with specimen #2, flaw #2. The result page displays many ultrasonic scan methods, such as an ultrasonic C-scan, D-scan, and S-scan, together. Among the scan methods in the result page, the S-scan seems to convey the height and length information of the flaw. Therefore, in this research, the S-scan images were the training data. We obtained a total of 498 S-scan images.



**Figure 5.** S-scan data sample used in training of algorithm acquired from EXTENDE CIVA simulation results, defect 2, specimen #2.

## 5. Database Construction

### 5.1. Database Classification

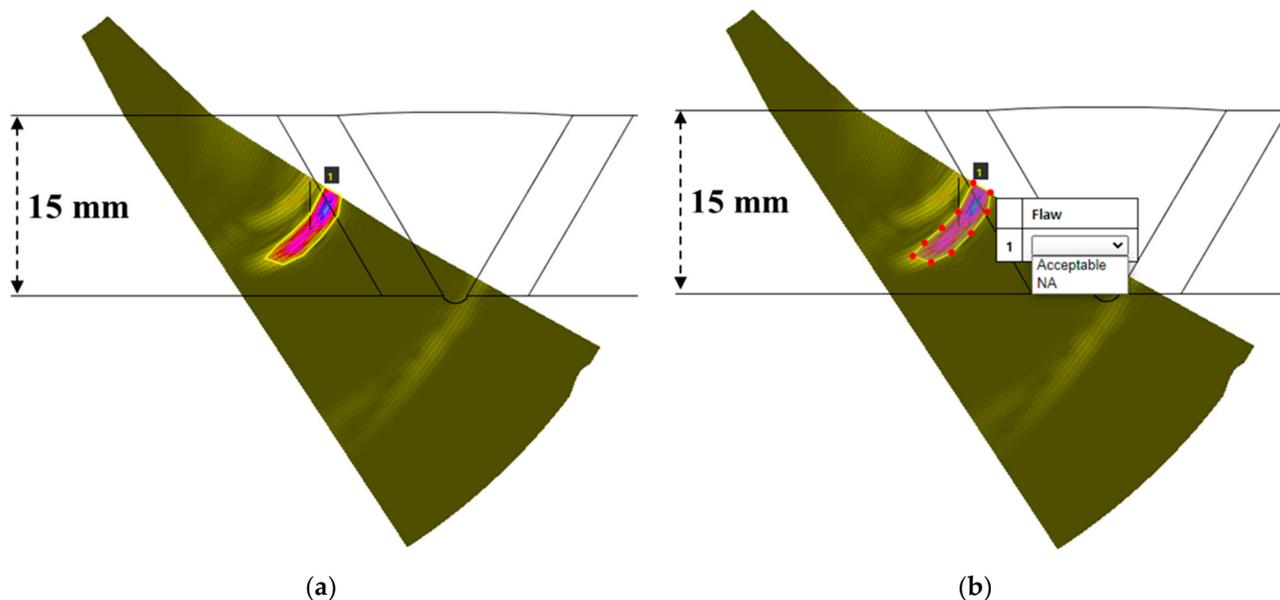
After the simulation, a total of 498 flaw images in the S-scan are obtained. Before the masking or labeling of the images, the images must be classified into training data and validation data, since Mask R-CNN requires a training dataset and validation dataset in training. Sixty randomly selected images out of 498 images are classified to be used for the validation process. Table 4 shows how the data are classified.

**Table 4.** Database and data classification.

	Number of Data
S-Scan Image	498
Training Data	438
Validation Data	60

### 5.2. Data Masking

For the masking of datasets, VGG Image Annotator (VIA) is used [21]. VIA is a manual image annotation tool developed by the Visual Geometry Group (VGG). The application is very light in storage and does not require internet access. Since the region should be masked in a certain format, the region shape must be selected among the 5 shape options. In this study, the polygon shape is used for the masking of regions so that the shape of the flaw can be precisely masked, as shown in Figure 6.



**Figure 6.** Data masking and annotation process with VIA, defect 2, specimen #2, 1 in the figure means that it is labeled as first category which is “Acceptable” in this study: (a) masking of the flaw (region shape: polygon); (b) annotation of the flaw.

All 438 images for the training and 60 images for the validation of Mask R-CNN are masked with the polygon, a region shape option. Then, the masked regions are annotated in the region attribute tab of VIA. The regions are annotated into two classes, acceptable and not acceptable (NA).

## 6. Training

The training and validation datasets are used in the training of Mask R-CNN. Figure 7 schematically presents the training process of Mask R-CNN, described in [22], together with the dataset obtained in this study. First, using binary interpolation, the input images are

resized. Then, the input size is padded to enter the backbone network. Using ResNet-101, each layer produces a feature map. Mask R-CNN uses the Feature Pyramid Network (FPN) before the Region Proposal Network (RPN) to become flexibly applicable to different sizes of object images. RPN is applied to the feature map acquired by the FPN and the classification and bounding box regression outcomes are acquired. Then, an anchor box is made by projecting the bounding box regression outcome onto the original image. Using non-max suppression, the anchor boxes with lower scores are deleted. The ROI aligner uniformizes the sizes of the remaining anchor boxes. Moreover, with Mask R-CNN, the actual boundary of the target object can be trained with a mask branch. Using Mask R-CNN, the classification of flaws according to their size is possible [23].

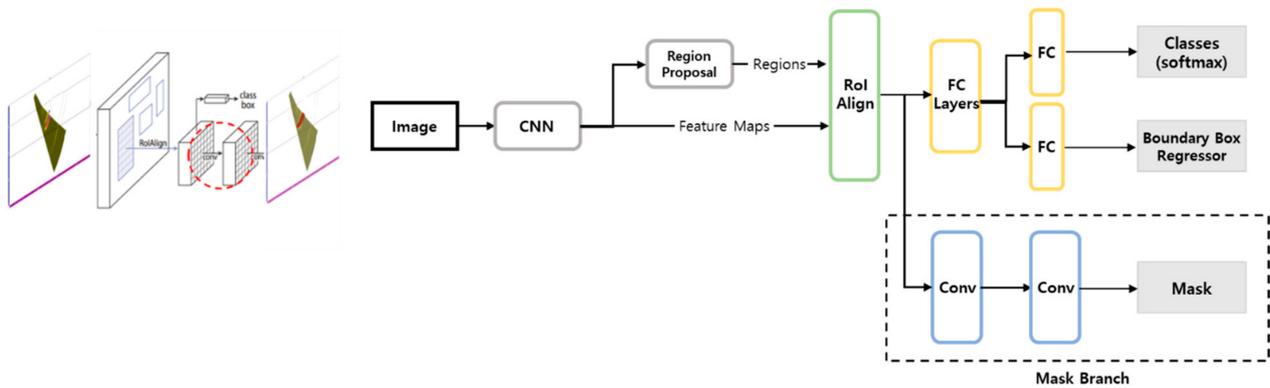


Figure 7. Mask R-CNN framework for instance segmentation and the mask branch.

Table 5 contains the crucial parameter settings for the training of Mask R-CNN. “Image per GPU” is the number of images processed simultaneously by the GPU. The better the performance of the GPU for training, the more images can be processed. “Steps per Epoch” refers to the number of extracted images per iteration. The larger the number filled in, the longer the training time. Filling in a large number can improve the accuracy, but it can lead to the overfitting of the model. “Validation Steps” is used to verify the images after each iteration cycle. Only a part of the training set is set to be used because it mainly affects the time of training. “RPN Anchor Scales” refers to the size of windows selected through RPN. “RPN Train Anchors per Images” is the number of windows selected per image, and “RPN Anchors Ratios” is the ratio corresponding to the selected window. “Train ROI per Image” is obtained according to the number set through iterative learning debugging. The larger the number of objects to be detected in the image, the larger the value should be set. If there are many “Number Classes”, it should be set high. In this research, there are a total of three classes: background, acceptable, and NA. “RPN NMS Threshold” deletes a box with an overlap rate higher than the threshold during learning. If the value of “Detection NMS Threshold” is large, the result may have more labeling boxes or overlapping of the boxes. “Detection Min Confidence” is a value indicating the probability that the detected object is similar to the annotated category, and if the probability is lower than the threshold, it does not detect the object during learning and is not displayed in the prediction.

Table 5. Crucial parameters used in the training.

Variables	Settings
Images per GPU	2
Steps per Epoch	20
Validation Steps	50
RPN Anchor Scales	(32, 64, 128, 256, 512)
RPN Train Anchors per Image	256
RPN Anchors Ratios	[0.5, 1.2]
Train ROIs per Image	200

**Table 5.** *Cont.*

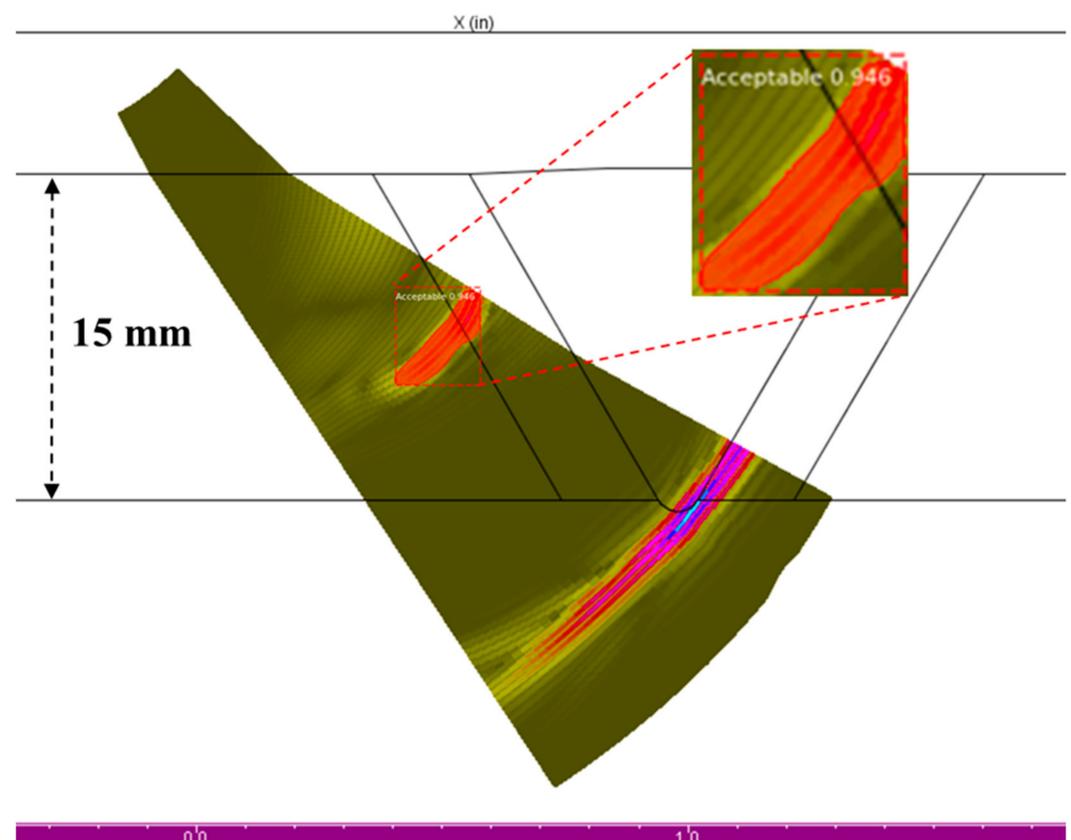
Variables	Settings
Learning Rate	0.001
Learning Momentum	0.9
Weight Decay	0.0001
Number Classes	1 + 2 (Background + Classes)
RPN NMS Threshold	0.7
Detection NMS Threshold	0.3
Detection Min Confidence	0.9

## 7. Results and Discussion

In this study, a total of 60 PAUT simulation S-scan flow images were used for the testing, of which 29 images for acceptable flaws and 31 images for not acceptable flaws were randomly selected. The simulation S-scans contained 1 mm, 2 mm, 4 mm, and 6 mm lengths of flaws. The 1 mm and 2 mm flaws were annotated as acceptable and the 4 mm and 6 mm flaws were annotated as not acceptable (NA).

### 7.1. Acceptable Flaws

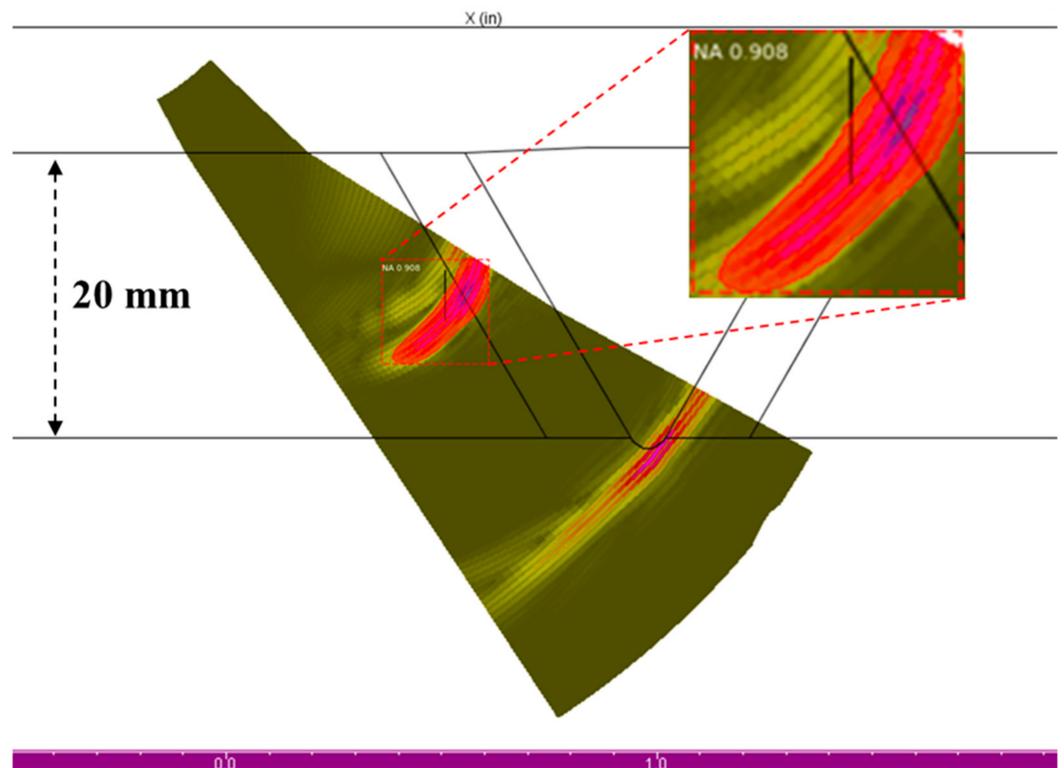
Figure 8 shows the testing result of the acceptable flaw that has the length of 2 mm, on a 15-mm-thick specimen. The 2 mm flaw is acceptable according to the acceptance criteria, so the model should determine it as acceptable. The trained model gave the result correctly.



**Figure 8.** Acceptable flaw result, flaw length: 2 mm.

### 7.2. Not Acceptable Flaws

Figure 9 shows the testing result of a not acceptable flaw that has the length of 4 mm, on a specimen with a thickness of 20 mm. The 4 mm flaw also is correctly determined according to the acceptance criteria.



**Figure 9.** Not acceptable (NA) flaw result, flaw length: 4 mm.

### 7.3. Confusion Matrix

Figure 10 is a confusion matrix of the Mask R-CNN detection model [23]. The confusion matrix is basically the relationship of the actual answer and the prediction of the model. When the answer can be either true or false, the prediction of the model can be also true or false. In the confusion matrix, there are four possibilities: true positive (TP), where the model predicts true while the answer is true; false positive (FP), where the model predicts true while the answer is false; false negative (FN), where the model predicts false while the answer is true, and true negative (TN), where the model predicts false while the answer is false. The confusion matrix gives precision and recall as the outputs. Precision is the ratio of what the model classifies as true to what is true. Recall is the fraction of what the model predicts to be true among what is true. To evaluate the performance of an algorithm, both precision and recall must be considered. The precision and recall can be expressed in the following way.

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

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

Class A, located in the first row and first column of the confusion matrix, recorded 0.00%, because Class A is the background. In the case of acceptable, Class B, 29 out of 29 flaw images were correctly detected and the recall recorded 100%, but the actual precision was 82.86%, as the acceptance flaws were incorrectly detected six times in the background. In the case of Class C, 27 flaws were correctly detected in 31 defect images, but the recall recorded 87.10% as it failed to detect four flaw images in the background. On the other hand, in the case of precision, it was detected incorrectly once in the background and recorded as 96.43%. Table 6 below shows the results of the confusion matrix in the table.

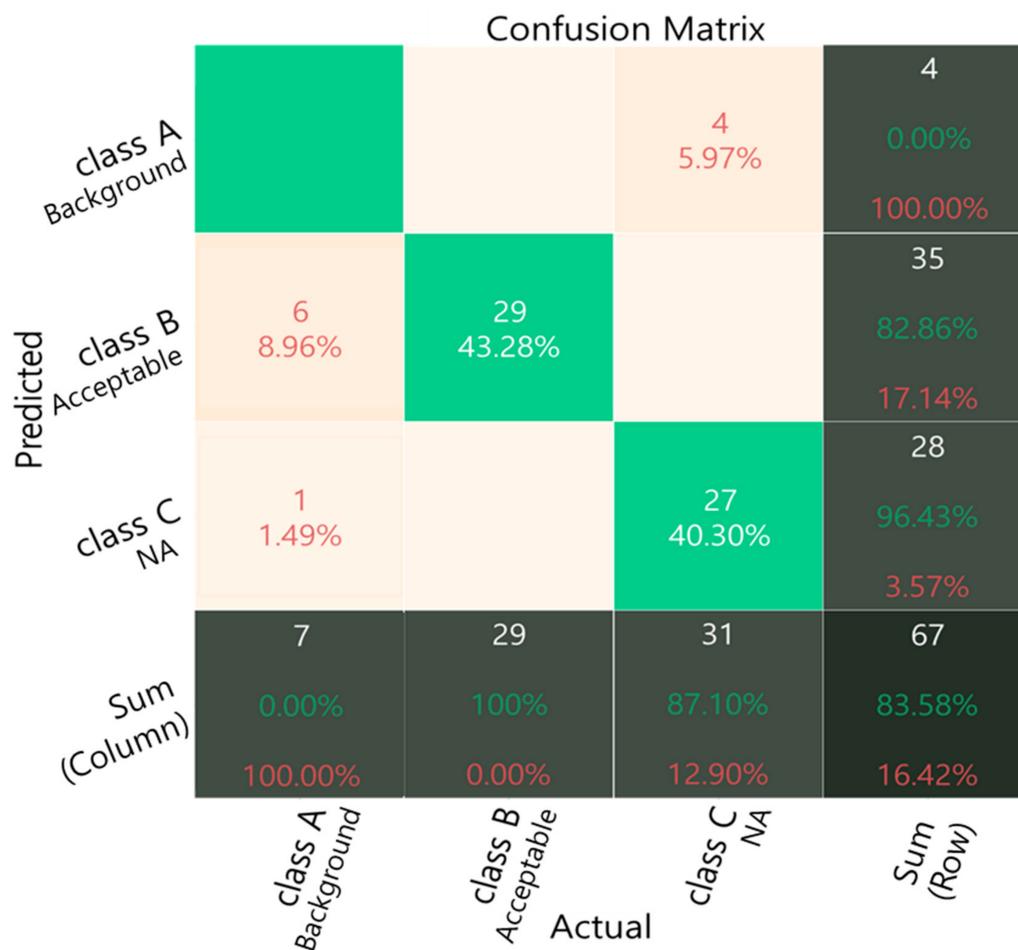


Figure 10. Confusion matrix of the trained model.

Table 6. Precision and recall of the trained model.

Class	Precision	Recall
Class B: Acceptable Flaws (Acceptable)	82.86%	100%
Class C: Not Acceptable Flaws (NA)	96.43%	87.10%

### 8. Conclusions

In this study, to enhance and speed up the Phased Array Ultrasonic Testing (PAUT) evaluation process, an algorithm to size and to determine the acceptance of flaws within a weldment is developed. A weldment wherein a flaw is embedded is evaluated by the size of the flaw. When the flaw does not meet the acceptance criteria described in an adequate international standard, the weldment is determined unacceptable. Therefore, an algorithm that determines a flaw in a weldment referring to the acceptance criteria should be developed.

For this research, the crucial international standards when evaluating Liquefied Natural Gas (LNG) storage tanks with PAUT are listed. API 620 Annex U contains the acceptance criteria for flaws in weldments. The sizing of a flaw in an LNG storage tank PAUT evaluation must follow the sizing methods introduced in ISO 13588 and ISO 11666.

Referring to the international standards, a total of 498 S-scan PAUT simulation images were designed and made. All the PAUT S-scan images were designed and made by PAUT simulation with EXTENDE CIVA. The flaws embedded in the simulation had the same rectangular shape, with different sizes. The image dataset was then divided into two different datasets, one for training and the other for validation. Then, the images were masked and annotated using VGG Image Annotator (VIA). The developed algorithm

using Mask R-CNN was trained with two different classes of S-scan, acceptable and not acceptable (NA).

After the training, 60 S-scan images, 29 acceptable flaw images, and 31 not acceptable flaw images were used for the performance testing of the trained model. For Class B, which was acceptable flaws, the precision was 82.86% and the recall was 100%. For Class C, which was unacceptable flaws, the precision was 96.43% and the recall was 87.10%. The testing result shows that Mask R-CNN was properly trained.

The result of this research proves the applicability of ANN in PAUT S-scan data sizing. Although the model showed good performance, it still needs some improvements before its application in the field. Most importantly, the dataset was collected from the PAUT simulation. For the application of the model, the model must be trained with actual PAUT data. With actual PAUT data, the model will generate more practical results and become part of a more efficient inspection system.

**Author Contributions:** Conceptualization, S.-J.S., S.-E.L. and J.P.; methodology, S.-J.S. and S.-E.L.; software, H.-J.K., S.-E.L. and J.P.; validation, S.-J.S., S.-E.L., J.P., H.-J.K. and Y.-T.Y.; formal analysis, S.-E.L.; investigation, S.-E.L.; resources, S.-E.L. and J.P.; data curation, S.-E.L.; writing—original draft preparation, S.-E.L.; writing—review and editing, S.-E.L. and S.-J.S.; visualization, S.-E.L.; supervision, S.-J.S. and H.-J.K.; project administration, Y.-T.Y.; funding acquisition, S.-J.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2020R1I1A1A01073812).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

LNG	Liquefied Natural Gas
PAUT	Phased Array Ultrasonic Testing
RT	Radiographic Testing
ANN	Artificial Neural Network
AI	Artificial Intelligence
CNN	Convolutional Neural Network
UT	Ultrasonic Testing
NDE	Nondestructive Evaluation
ECT	Eddy-Current Testing
HAZ	Heat-Affected Zone
VIA	VGG Image Annotator
VGG	Visual Geometry Group
FPN	Feature Pyramid Network
RPN	Region Proposal Network
NA	Not Acceptable
RT	Radiographic Testing

## References

1. Lun, H.; Fillippone, F.; Roger, D.C.; Poser, M. Design and construction aspects of post-tensioned LNG storage tank in Europe and Australasia. In Proceedings of the New Zealand Concrete Industry Conference, Christchurch, New Zealand, 29 September–1 October 2006.
2. PetroSkills. LNG Containment Tanks: Why Is the Internal Liner 9% Ni and Not 7% Ni? 2021. Available online: <http://www.jmcampbell.com/tip-of-the-month/2021/05/lng-containment-tanks-why-is-the-internal-liner-9-ni-and-not-7-ni/> (accessed on 5 January 2023).

3. Innopolis Foundation. Non-Destructive Testing and Inspection Market. 2019. Available online: <https://innopolis.or.kr> (accessed on 6 January 2023).
4. ISO 13588:2019; Non-Destructive Testing of Welds—Ultrasonic Testing—Use of Automated Phased Array Technology. International Organization for Standardization: London, UK, 2019. Available online: <https://www.iso.org/standard/72747.html> (accessed on 1 April 2021).
5. Song, S.J.; Schmerr, L.W. Ultrasonic flaw classification in weldments using probabilistic neural networks. *J. Nondestruct. Eval.* **1992**, *11*, 69–77. [[CrossRef](#)]
6. Park, J.; Lee, S.E.; Kim, H.J.; Song, S.J.; Kang, S.S. System Invariant Method for Ultrasonic Flaw Classification in Weldments Using Residual Neural Network. *Appl. Sci.* **2022**, *12*, 1477. [[CrossRef](#)]
7. Munir, N.; Kim, H.J.; Song, S.J.; Kang, S.S. Investigation of deep neural network with drop out for ultrasonic flaw classification in weldments. *J. Mech. Sci. Technol.* **2018**, *32*, 3073–3080. [[CrossRef](#)]
8. Munir, N.; Kim, H.J.; Park, J.; Song, S.J.; Kang, S.S. Convolutional neural network for ultrasonic weldment flaw classification in noisy conditions. *Ultrasonics* **2019**, *94*, 74–81. [[CrossRef](#)] [[PubMed](#)]
9. Cruz, F.C.; Simas Filho, E.F.; Albuquerque, M.C.S.; Silva, I.C.; Farias, C.T.T.; Gouvêa, L.L. Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing. *Ultrasonics* **2017**, *73*, 1–8. [[CrossRef](#)] [[PubMed](#)]
10. Sambath, S.; Nagaraj, P.; Selvakumar, N. Automatic defect classification in ultrasonic NDT using artificial intelligence. *J. Nondestruct. Eval.* **2011**, *30*, 20–28. [[CrossRef](#)]
11. Virupakshappa, K.; Marino, M.; Oruklu, E. A multi-resolution convolutional neural network architecture for ultrasonic flaw detection. In Proceedings of the 2018 IEEE International Ultrasonics Symposium (IUS), Kobe, Japan, 22–25 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–4. [[CrossRef](#)]
12. Medak, D.; Posilović, L.; Subašić, M.; Budimir, M.; Lončarić, S. Automated defect detection from ultrasonic images using deep learning. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2021**, *68*, 3126–3134. [[CrossRef](#)] [[PubMed](#)]
13. Siljama, O.; Koskinen, T.; Jessen-Juhler, O.; Virkkunen, I. Automated flaw detection in multi-channel phased array ultrasonic data using machine learning. *J. Nondestruct. Eval.* **2021**, *40*, 67. [[CrossRef](#)]
14. Ho, M.P.; Ngai, W.K.; Chan, T.W.; Wai, H.W. An artificial neural network approach for parametric study on welding defect classification. *Int. J. Adv. Manuf. Technol.* **2022**, *120*, 527–535. [[CrossRef](#)]
15. API 620; Design and Construction of Large, Welded, Low-Pressure Storage Tanks. American Petroleum Institute: Washington, DC, USA, 2018.
16. ISO 11666:2018; Non-Destructive Testing of Welds—Ultrasonic Testing—Acceptance Levels. International Organization for Standardization: London, UK, 2018. Available online: <https://www.iso.org/standard/69610.html> (accessed on 1 April 2021).
17. ISO 22825:2017; Non-Destructive Testing of Welds—Ultrasonic Testing—Testing of Welds in Austenitic Steels and Nickel-Based Alloys. International Organization for Standardization: London, UK, 2017. Available online: <https://www.iso.org/standard/69428.html> (accessed on 2 June 2021).
18. ISO 17640:2018; Non-Destructive Testing of Welds—Ultrasonic Testing—Techniques, Testing Levels, and Assessment. International Organization for Standardization: London, UK, 2018. Available online: <https://www.iso.org/standard/75737.html> (accessed on 3 May 2021).
19. ISO 19285:2017; Non-Destructive Testing of Welds—Phased Array Ultrasonic Testing (PAUT)—Acceptance Levels. International Organization for Standardization: London, UK, 2017. Available online: <https://www.iso.org/standard/64266.html> (accessed on 1 April 2021).
20. Extende. Ultrasonic Testing with CIVA. 2022. Available online: <https://www.extende.com/ultrasonic-testing-with-civa> (accessed on 6 January 2023).
21. Dutta, A.; Gupta, A.; Zissermann, A. VGG Image Annotator (VIA). 2016. Available online: <http://www.robots.ox.ac.uk/vgg/software/via> (accessed on 6 January 2023).
22. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
23. Ait Jeddi, Y. Confusion Matrix for Matterport Implementation of Mask R-CNN. 2021. Available online: <https://github.com/Altimis/Confusion-matrix-for-Mask-R-CNN> (accessed on 12 January 2023).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.