IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 September 2023, accepted 8 October 2023, date of publication 16 October 2023, date of current version 20 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3324717

RESEARCH ARTICLE

Multi-Branch Deep Fusion Network-Based Automatic Detection of Weld Defects Using Non-Destructive Ultrasonic Test

KYEEUN KIM, KEO SIK KIM, AND HYOUNG-JUN PARK
Electronics and Telecommunications Research Institute, Daejeon 34129, South Korea

Corresponding author: Keo Sik Kim (keosikis@etri.re.kr)

This work was supported by the Electronics and Telecommunications Research Institute (ETRI) funded by the Korean Government through the Honam Region Regional Industry-Based ICT Convergence Technology Advancement Support Project 23ZK1100.

ABSTRACT This study introduces a deep learning engine designed for the non-destructive automatic detection of defects within weld beads. A 1D waveform ultrasound signal was collected using an A-scan pulser receiver to gather defect signals from inside the weld bead. We established 5,108 training datasets and 500 test datasets for five pass/fail labels in this study. We developed a multi-branch deep fusion network (MBDFN) model that independently trains 1D-CNN for local pattern learning within a sequence and 2D-CNN for spatial feature extraction and then combines them in an ensemble method, achieving a classification accuracy of 92.2%. The resulting deep learning engine has potential applications in automatic welding robots or welding inspection systems, allowing for rapid determination of internal defects without compromising the integrity of the finished product.

INDEX TERMS Deep learning, quality management, welding, automatic testing.

I. INTRODUCTION

Non-destructive testing (NDT) techniques such as scanning electron microscopy (SEM) and scanning helium ion microscopy (SHIM) play an important role in inspecting secondary electron (SE) signals in semiconductor chips without compromising the integrity of the material [1], [2], [3], [4], [5]. Similarly, in welding applications used in automotive and ship manufacturing, NDT plays a key role in inspecting welded joints for defects [6], [7], [8], [9], [10]. These tests are particularly employed to identify various defects within the bead, such as cracks, incomplete penetration, lack of fusion (LF), and porosity [11]. Among NDT methods applicable to bead defect detection, magnetic testing (MT), penetration testing (PT), radiographic testing (RT), and ultrasonic testing (UT) are commonly utilized [7], [12], [13]. However, magnetic testing is most suitable for examining magnetic materials, penetrant testing is primarily used to observe surface defects, and radiographic testing faces challenges in detecting internal microscopic defects and localizing them accurately [10], [13]. In our study, the focus is on employing ultrasonic testing to detect defects within the beads.

Ultrasonic testing utilizes a piezoelectric probe to transmit specific-frequency ultrasonic waves, into the material's interior [7], [15]. These waves travel in a straight line and interact with different media, such as defects or other materials, causing reflection or refraction based on the medium's physical properties. Subsequently, the probe receives the reflected or refracted ultrasonic signal and presents it as a pulse signal to the user [7], [15]. By analyzing the shape of the received pulse signal, the presence and type of defect can be determined, as the signal's shape varies according to the internal defect's characteristics [7], [15], [16], [17].

The ultrasonic detection methods can be classified into one-dimensional UT and phased-array ultrasonic testing (PAUT) based on their signaling approach [18], [19]. In one-dimensional UT, a single sound wave with fixed angle and frequency is transmitted and received through a probe [7]. PAUT allows transmission and reception of ultrasonic waves in an array, enabling adjustment of angle and frequency, and presenting the received sound waves as an image [18], [19]. This user-friendly feature allows easy defect detection and

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy.

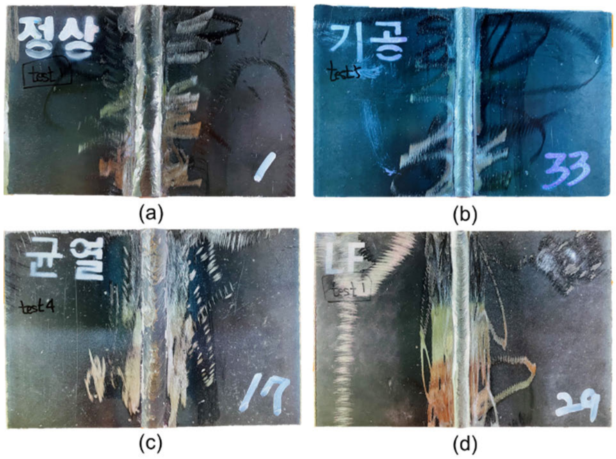

FIGURE 1. Photographs of internal defect samples; (a) pass, (b) porosity, (c) crack, and (d) lack-of-fusion.

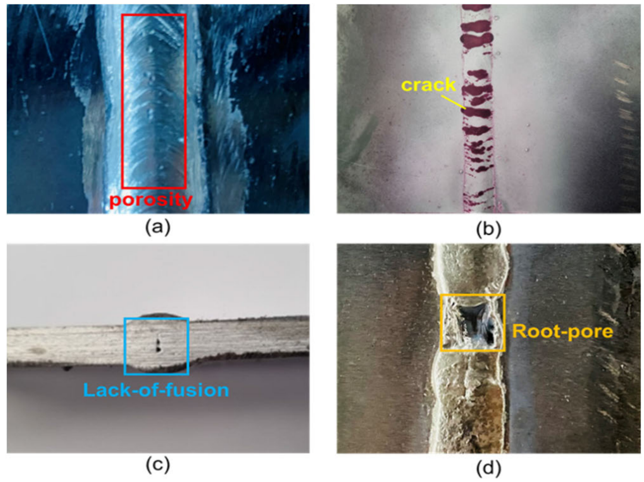

FIGURE 2. Internal defect examples; (a) porosity, (b) crack, (c) lack-of-fusion, and (d) root-pore.

classification [7], [19], though it comes with the drawback of high equipment cost. Conversely, UT is cost-effective but relies on a single pulse signal for defect identification, leading to signal distortions depending on inspection (e.g., probe-workpiece contact, probe position, orientation) and requiring expert-level discernment. Consequently, inspection results might be subjective and vary based on the inspector [13], [19].

To overcome these limitations, this study aims to enhance the classification performance of non-destructive UT as applying deep learning algorithms and signal processing techniques [11], [18], [19]. We propose a novel deep learning-based UT engine that consists of UT device to acquire a single ultrasonic signal and convolutional neural network (CNN)-based classifier capable of distinguishing four different types of defects within a weld bead [18], [19], [20], [21], [22].

II. EXPERIMENTAL DETAILS

A. SAMPLE PRODUCTION

To create an ultrasonic dataset of internal weld bead defects, we prepared the four types of weld samples: one ('pass') without defects and the other with internal defects including pores, cracks, lack-of-fusion. The samples were produced by a company that manufactures actual I-shaped butt CO_2 welds and were designed to present the types found in the 'standard ultrasonic kit (FLAWTECH co., LTD.)' used as a reference for ultrasonic testing.

Pores are caused by moisture from the surrounding air or entrapment of the surface oxide film during melting [23], [24], and depending on the cause, they can appear as porosity, root-pore, etc. In this study, 'root-pore' and 'porosity' were simultaneously present in the pore sample, and data of both defect patterns were obtained from the porous sample as shown in Figure 1(b). The remaining data were obtained from the corresponding samples, such as 'pass' / 'crack' / 'lack-of-fusion' samples, respectively, as shown in Fig. 1(a), (c) and (d). The weld base material was SAPH440, a carbon steel, with a thickness of 7t (7 mm), subject to a 10% thickness error. We used a KC-28 carbon steel welding rod and performed I-shaped butt CO_2 welding between the materials. Fig. 1 displays a photograph of the manufactured sample. The total number of samples for the four types of normal and defective welds was 146.

In Fig. 2, we present defect samples exhibiting porosity, cracks, lack-of-fusion, and root pores. These defects were randomly distributed within the sample. To identify the defect locations, we employed MT, and only the data collected at those identified locations were used as defect data. Fig. 2(a) displays a porous sample, where porosity is observed randomly distributed inside the bead. In Fig. 2(b), we see a cracked sample, with the identified cracks highlighted in red through magnetic particle detection. Fig. 2(c) showcases a lack-of-fusion sample, revealing a defect situated in the middle of the bead. Lastly, Fig. 2(d) exhibits a root pore sample, indicating the presence of a root pore on the back of the bead.

B. DATA ACQUSITION

For this study, all ultrasonic signals were collected using a 5MHz-frequency 70-degree probe (KN5-70, Kyungdo Enterprise Co. Ltd., Republic of Korea) and a pulser-receiver (OPBOX 2.0, Optel, Poland) capable of transmitting and receiving ultrasonic signals from a single probe. The OPBOX is equipped with a built-in 10-bit A/D converter, which efficiently converts ultrasonic analog signals into digital data, eliminating the need for a separate signal converter like an oscilloscope.

The ultrasonic settings were configured as follows, taking into account the speed of ultrasonic propagation in the welding material: velocity 3,420 m/s, pulse width 0.1 μs, sampling frequency 50 MHz, analog filters 4 - 10 MHz, gain (constant) 50 dB, and binning of 64. The OPBOX software allowed easy ultrasound setup and signal verification.

Due to the dead zone at the beginning of the ultrasonic signal caused by the initial pulse, the probe's frequency needed

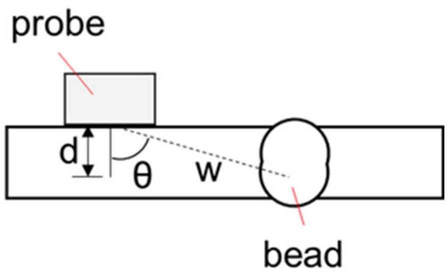

FIGURE 3. Schematic of ultrasonic testing.

careful consideration [7], [17]. As the weld sample in this study was relatively thin at 7t (7 mm), a 5 MHz probe with a short dead zone of less than 20 μs was chosen. The data before 20 μs was considered as the dead zone, excluding it from model training due to noise interference [17]. Additionally, analysis revealed a peak beyond 65 μs on the time axis, requiring calculation to interpret the signal from a thickness of several millimeters. Fig. 3 depicts a schematic diagram illustrating the correlation between ultrasonic propagation distance, detection angle, and defect depth during ultrasonic testing.

To calculate the defect depth d, we use the following equation:

$$d = W \times \cos\theta \tag{1}$$

where d is the defect depth and, W is the propagation distance of the ultrasonic wave. The propagation distance of the ultrasonic wave W was calculated using the following equation:

$$W = v \times t \tag{2}$$

where v is the ultrasonic velocity, and t is the propagation time. Using Eq. (1) and (2), we calculated the point on the time axis at 65 μs to have a depth of approximately 7.6 mm, confirming the signal originates from the floor surface.

In Fig. 4, the 1D-waveform of ultrasonic signals for normal and defect types is presented. Fig. 4(a) displays a 1D waveform obtained from a pass sample, where no defect peak is observed beyond 20 μs. However, a peak signal reflected from the back of the weld sample appears after the time axis of approximately 65 μs. Fig. 4(b) shows the waveform from the porosity sample, revealing a peak at around 50 μs (approximately 5.8 mm depth) caused by the defect and another peak beyond approximately 65 μs obtained from the back of the sample. Fig. 4(c) presents the signal from a root-pore sample, demonstrating a defect-induced peak at about 60 μs (around 7 mm deep) due to pores located at the bottom of the sample. For cracks, Fig. 4(d) shows distributed peaks starting at 50 μs (around 5.8 mm deep), while Fig. 4(e) illustrates a large peak attributed to lack-of-fusion around 55 μs (approximately 6.4 mm deep).

Table 1 presents the organization of the collected ultrasound training datasets, classified by labels. Out of the 146 samples, we acquired 5,108 ultrasound 1D waveform data from 126 samples. Approximately 80% of these data

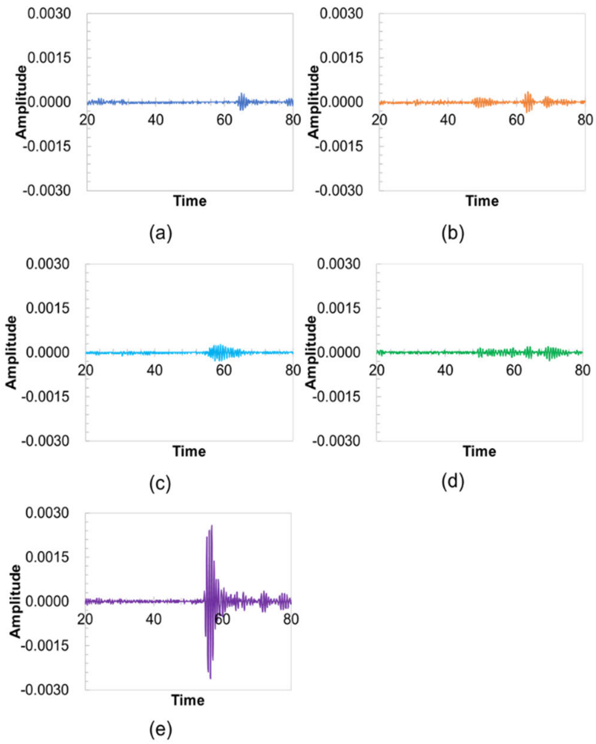

FIGURE 4. 1D waveform; (a) pass, (b) porosity, (c) root-pore, (d) crack, and (e) lack-of-fusion.

(4,086) were used for training the AI model, while the remaining 20% (1,022) formed the validation dataset. For the test dataset, we collected 100 data points per defect from the remaining 20 samples, resulting in a total of 500 data points.

The different amount of data per label can be attributed to the different features and patterns inherent in each label type. All defect signals were verified by two experienced experts with more than 10 years of experience in manufacturing using welding. Lack-of-fusion was only able to obtain signals similar to what can be seen in Figure 4(e), and was determined to be due to the presence of only one defect pattern. Similarly, for the pass, porosity, and root-pore labels, which do not have significantly different defect patterns, we were able to collect about a thousand data each. However, since cracks can produce transverse and longitudinal cracks depending on the direction of growth and have different shapes of 1D waveforms [25], [26], we collected about two thousand data for training the algorithm, which is about twice as many as the other labels.

C. DEEP LEARNING-BASED CLASSIFIER

We developed a deep learning-based algorithm to classify the UT signal into normal and one of the four defect patterns present in the weld bead. External inclusions like pores or slag along the ultrasonic path, as well as improper welding conditions resulting in incomplete structures, can alter the phase and magnitude of the reflected ultrasonic signal [8]. These changes manifest as peaks in the 1D UT waveform, and

TABLE 1. UT 1D waveform dataset.

	Label	Train	Validation	Test	Total
Pass	0	703	176	100	979
Porosity	1	734	195	100	1,029
Crack	2	1,522	368	100	1,990
Lack-of-fusion	3	386	94	100	580
Root-pore	4	741	189	100	1,030
Total		4,086	1,022	500	5,608

their locations and shapes serve as crucial features to detect internal defects. Therefore, we devised a deep learning-based classifier that can simultaneously learn various temporal and spatial features of the above-mentioned peaks, and this network is called multi-branch deep fusion network (MBDFN). It consists of a parallel network structure of 1D- or 2D convolution branches, named as multi-branch, and each branch has consecutive convolution modules, named as deep fusion network.

1D-CNN is useful for extracting temporal local features. For example, time-varying features such as oscillations and sparkles can be extracted from one-dimensional UT waveform data. Since the order of this kind of data is usually important, 1D-CNN is effective at recognizing and learning local patterns within sequences. On the other hand, 2D-CNN has advantages in spatial feature extraction. For example, spatial information such as location and peak spacing can be extracted from two-dimensional data, such as a one-dimensional UT waveform converted to a gray-scale image. Since spatial structure is important in two-dimensional data such as images, 2D-CNN is useful for recognizing and learning patterns from different parts of an image [27], [28]. The detailed structure of each CNN is as follows.

Our 1D-CNN structure, as depicted in Fig. 5, involves a 1D convolutional layer for pattern classification using the temporal features of significant peaks along the time axis (or spatial axis) of the UT 1D waveform. To enhance the reliability and efficiency of model training and prediction, as well as to improve the model's performance and generalization capability, we normalized the 1D UT waveforms to a range of 0 to 1. Additionally, we introduced random noise (sigma=0.03) to further normalize the signal. The signal then underwent sequentially repeated convolutional modules, including the 1D convolutional layer, activation layer, pooling layer, and drop-out layer. The output passed through repeated dense layers, and the final decision was made using the softmax function. To address the gradient banishing issue in deeper layers, we implemented a residual connection between the input and output of the 1D convolution module [29]. Optimal hyperparameters for the sequential convolution modules and dense layers, such as the number and size of 1D convolution

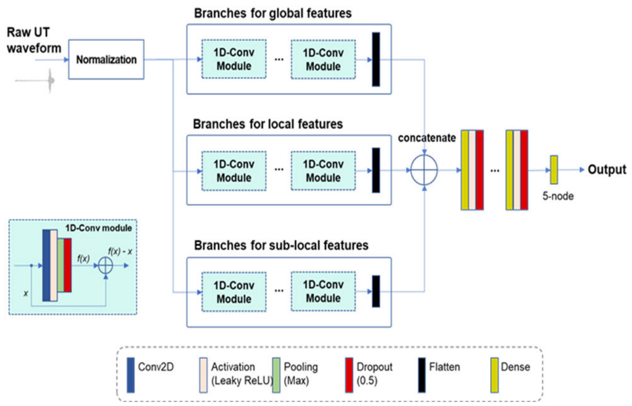

FIGURE 5. Structure of 1D-CNN of UT waveform signal.

TABLE 2. Hyperparameters of 1D-CNN.

Hyperparameter		Value
Optimizer		Leaky ReLU
Learning rate		[0.01 ~ 0.0001]
Epoch		10,000
Early stopping		Yes (patient:50 epochs)
Branches	Number	Various (3~5)
1D Conv. module	Repetition	Various (7~10)
Conv. 1D	Filters	Various (16)
	Kernel size	Various ([3, 5, 7])

filters and the number of hidden layers, were obtained via grid search under various training conditions. Table 2 presents the hyperparameters of the 1D-CNN.

2D-CNN structure is depicted in Fig. 6. It comprises parallel 2D convolution modules, including a 2D convolutional layer, activation layer, pooling layer, and drop-out layer, similar to 1D-CNN. All modules are concatenated and fed into repeated dense layers. The final decision is obtained using the softmax function. Additionally, similar to the 1D convolution module in Fig. 5, the 2D convolution module also incorporates residual connections between its input and output. To determine optimal hyperparameters for the sequential 2D convolution modules, we conducted a grid search under various training conditions. The filter sizes in the modules for global features were set to be larger than those for local features, as larger filter sizes provide more global features. Table 3 outlines the hyperparameters of the 2D-CNN.

For the final classification of UT signals, we selected 1D-CNN and 2D-CNN candidates with the highest macro-average recall. By ensembling these candidates, we identified the optimal structure based on the best macro-average recall. The macro-average recall, which represents the average detection performance (recall) per class, was considered as a suitable performance indicator for comparing classification performance among models.

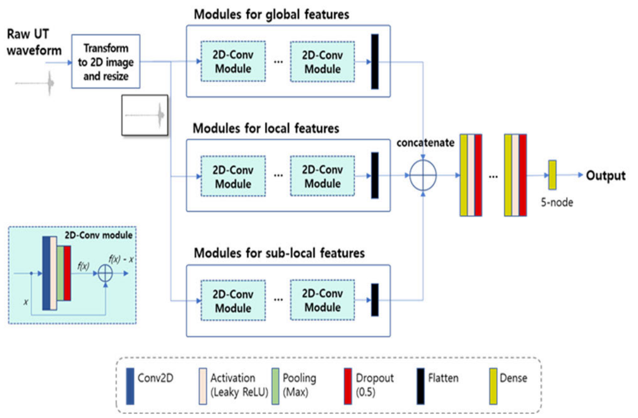

FIGURE 6. Structure of 2D-CNN of UT image.

TABLE 3. Hyperparameters of 2D-CNN.

Hyperparameter		Value
Optimizer		Leaky ReLU
Learning rate		[0.002 ~ 0.00001]
Epoch		200
Early stopping		Yes (patient:20 epochs)
Branches	Number	Various (3~5)
2D Conv. module	Repetition	Various (2~9)
Conv. 2D	Filters	Various (16~64)
	Kernel size	Various ((3x3) ~ (11x11))

To obtain the final decision, we utilized soft voting of the logit from the combined models. Subsequently, the node with the highest value was classified as the final class using softmax. A visual representation of the MBDFN's overall configuration is referred to Fig. 7.

III. RESULTS AND DISCUSSION

To determine the optimal hyperparameters of the 1D- and 2D-CNN, we evaluated models with a macro-average recall exceeding 0.87. The final classification model was chosen by selecting the model combination with the highest macro-average recall among multiple combinations.

For the hyperparameters of the 1D-CNN, after fixing the learning rate and random seed (as per Table 2), we conducted a grid search. The resulting model had three branches in the CNN module, with a 1D convolutional layer and its filter size of 10 in each branch. The kernel sizes were 7 × 7, 5 × 5, and 3 × 3 for the respective branches, with the activation function being Leaky ReLU. Additionally, the MaxPooling 1D layer had a pool size of 2 × 2, and the dropout ratio was set to 0.2. Table 4 displays the classification performance index of the top five models with macro-average recall, obtained by varying the learning rate from 0.001 to 0.00005 for the 1D-CNN with the optimal hyperparameters. The recall and

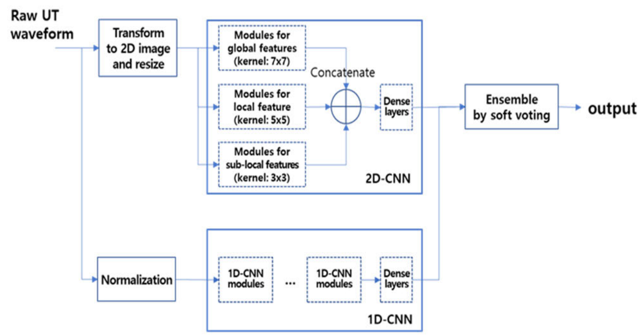

FIGURE 7. Overall structure of MBDFN for final classification of defective UT.

TABLE 4. Training results of the 1D-CNN model.

Model number	Precision	Recall	F1-score	Accuracy
1	0.893	0.890	0.890	0.890
2	0.890	0.886	0.885	0.886
3	0.891	0.884	0.884	0.884
4	0.886	0.880	0.879	0.880
5	0.878	0.874	0.873	0.874

F1-score of the top five models ranged from 0.873 to 0.89, with a mean recall of 88.2% and a standard deviation of 0.7%.

Table 5 displays the training results of 1D-CNN Model-1, which achieved the highest macro-average recall. The model's performance metrics are as follows: F1-score of 89.0%, precision of 89.3%, recall of 89.0%, and accuracy of 89.0%.

Fig. 8 illustrates the confusion matrix of the 1D-CNN model mentioned above, featuring five labels: pass (0), porosity (1), crack (2), lack of fusion (3), and root pore (4). According to the matrix, the classification accuracy for the four labels (pass, porosity, lack of fusion, and root pore) is nearly 90%. However, the accuracy for the crack label is 75%, particularly when misclassified as porosity (18%), which shares a similar 1D waveform. This discrepancy is attributed to the various forms of crack generation, such as transverse and longitudinal cracks, during the production of internal defect samples. Additionally, ultrasonic signals are collected differently based on the crack's shape [25], [26]. When the crack progression direction aligns with the ultrasonic progression direction, the crack may be recognized as a point, generating a waveform resembling porosity. On the other hand, when the crack progression direction and the ultrasonic progression direction do not align, a disconnected progression surface, such as lack of fusion or root pore, can generate a waveform. Although we considered different crack types and trained with about twice as much crack data as the other labels, we did not get good results in crack classification. To enhance performance, it may be beneficial to differentiate cracks into separate labels, such as transverse and longitudinal cracks, based on their generation direction.

TABLE 5. Results of 1D-CNN best performance model.

	Precision	Recall	F1-score	Support
Pass	0.941	0.960	0.955	100
Porosity	0.795	0.930	0.857	100
Crack	0.872	0.750	0.806	100
Lack-of-fusion	0.969	0.940	0.954	100
Root-pore	0.879	0.870	0.874	100
Accuracy			0.890	500
Macro avg	0.893	0.890	0.890	500
Weighted avg	0.893	0.890	0.890	500

After conducting a grid search with fixed learning rate and random seed for the 2D-CNN (as per Table 3), the hyperparameters of the model with the best macro-average recall performance are as follows. Similar to the 1D-CNN, the CNN module includes three branches: each branch contains a 2D convolutional layer with 16 filters. The kernel sizes for the branches are 7 × 7, 5 × 5, and 3 × 3, with the activation function being Leaky ReLU. The MaxPooling 2D layer has a pool size of 2 × 2, and the dropout ratio is 0.2.

Table 6 presents the classification performance indices of the top five models with macro-average recall values obtained by varying the learning rate from 0.001 to 0.00005 for the 2D-CNN with the optimal hyperparameters. All eight models achieved macro-average recall values exceeding 87%, and their accuracies were also above 87%.

Table 7 displays the training results of Model-1, which exhibited the best performance among the 2D-CNNs. The model demonstrated precision, recall, F1-score and accuracy rates of 88.7%, 88.6%, 88.6%, and 88.6%, respectively.

Fig. 9 presents the confusion matrix of the 2D-CNN model. According to the confusion matrix, three labels, pass (0), porosity (1), lack of fusion (3), achieved a classification accuracy of over 90%. However, similar to the 1D-CNN results, the crack (2) and root-pore (4) labels had classification accuracies of 79% and 82%, respectively. Additionally, cracks were frequently misclassified by other labels.

Table 7 presents the performance of the MBDFN achieved through the optimal combination of the 1D-CNN and 2D-CNN models. MBDFN combines the eight 1D-CNN models and four 2D-CNN models mentioned earlier. Its performance was confirmed to be 92.4% precision, 92.1% F1-score, 92.2% recall, and 92.2% accuracy.

Fig. 10 illustrates the confusion matrix of the MBDFN model. Remarkably, all four labels achieved a classification accuracy of over 90%, representing a notable improvement compared to the individual models. Nevertheless, in the MBDFN network, the recall for cracks (2) remained at 78%, indicating that cracks are still frequently misclassified as porous defects such as porosity (1) and root-pore (4).

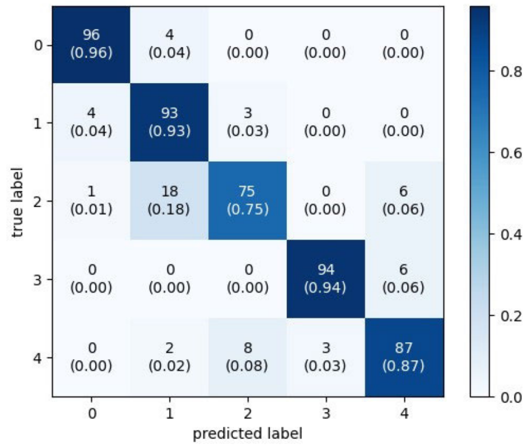

FIGURE 8. Confusion matrix of 1D-CNN model.

TABLE 6. Training results of the 2D-CNN model.

Model number	Precision	Recall	F1-score	Accuracy
1	0.889	0.888	0.888	0.888
2	0.890	0.886	0.886	0.886
3	0.882	0.884	0.882	0.884
4	0.881	0.882	0.880	0.882
5	0.881	0.880	0.880	0.880

TABLE 7. Results of 2D-CNN best performance model.

	Precision	Recall	F1-score	Support
Pass	0.941	0.960	0.950	100
Porosity	0.826	0.900	0.861	100
Crack	0.806	0.790	0.798	100
Lack-of-fusion	0.980	0.960	0.970	100
Root-pore	0.882	0.820	0.850	100
Accuracy			0.886	500
Macro avg	0.887	0.886	0.886	500
Weighted avg	0.887	0.886	0.886	500

Fig. 11 depicts the receiver operating characteristic (ROC) curve of the MBDFN model. The overall area under the curve (AUC) was 0.99, with individual AUC values for each label being 1.00 for pass, 0.99 for porosity, 0.98 for crack, 1.00 for lack-of-fusion, and 0.99 for root-pore.

MBDFN were developed to provide standardized defect detection and classification performance regardless of the operator's level of expertise. However, due to the structure of MBDFN, which consists of multiple models, detecting defects in real-time is challenging. While the inference time of a single 1D CNN model comprising an MBDFN averages 2.53±1.07 seconds with the mAP of 89%, the inference time of an MBDFN takes about 8 seconds longer and the mAP

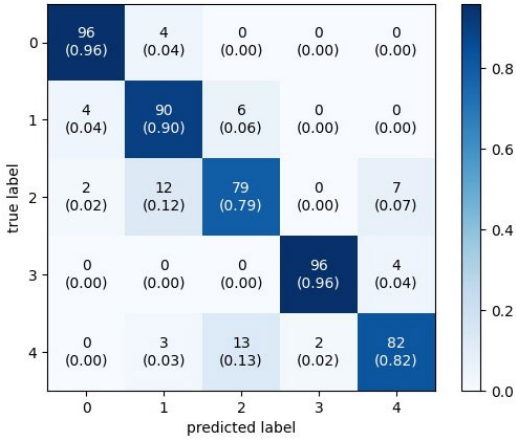

FIGURE 9. Confusion matrix of 2D-CNN model.

TABLE 8. Results of MBDFN.

	Precision	Recall	F1-score	Support
Pass	0.980	0.990	0.985	100
Porosity	0.836	0.970	0.898	100
Crack	0.907	0.780	0.839	100
Lack-of-fusion	0.980	0.980	0.980	100
Root-pore	0.918	0.890	0.904	100
Accuracy			0.922	500
Macro avg	0.924	0.922	0.921	500
Weighted avg	0.924	0.922	0.921	500

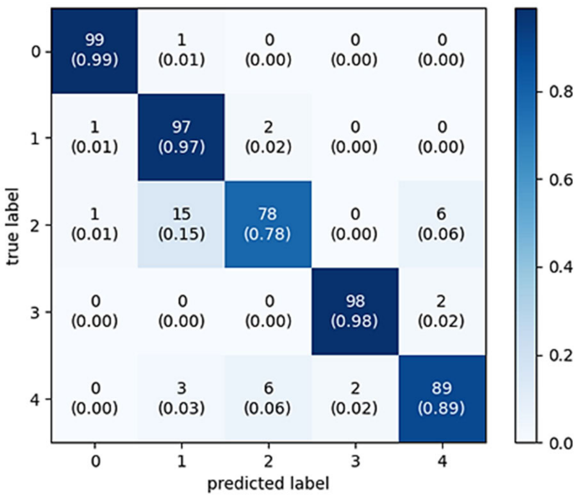

FIGURE 10. Confusion matrix of MBDFN model.

improves from 3.6% to 92.2%. This is mainly due to the model loading time and iteration of the 1D and 2D CNNs, while the inference time for unit data is expected to be within a few milliseconds. In the future, we expect that applying model optimization engines such as TensorRT and ONNX will further reduce model loading time, enabling real-time processing.

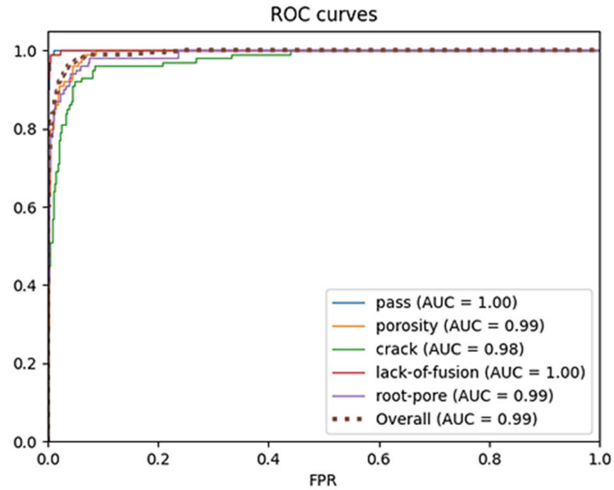

FIGURE 11. ROC curve of ensemble model.

In addition, the defect samples used in this study were produced by a company that manufactures I-shaped butt welds using CO_2 welding. We asked them to produce samples with defects that are common in real-world welding processes. We assumed that the types of defects would be similar in other applications or welding methods. However, the quality of the weld can change the appearance of the ultrasonic signal. For example, if the backside of the weld bead is not smooth, or if the bond between the base metal and the bead is poor, multiple ultrasonic peaks may occur on the backside, giving a different result than the actual defect. This issue can be addressed with additional data training, and coverage can be increased in the future by adding data collected from different environments.

IV. CONCLUSION

In this study, we proposed a deep-learning engine for automatically and non-destructively classifying defects in weld beads from 1D UT signals. Using an A-scan pulser-receiver and angle beam ultrasonic probe, we obtained 1D waveform ultrasound signals, which consists of 5,108 UT signals used for train dataset and 500 UT signals used for test dataset, from the weld bead. For classifying UT signal to the corresponding labels, such as pass and four types of defects, MBDFN model was developed through optimal combination of both the 1D-CNN models and 2D-CNNs. The confirmed classification accuracy for normal and four types of internal defects was 92.2%. We expect that the deep learning model presented in this paper could find practical applications in automatic welding robots or welding inspection systems to swiftly determine the presence of defects in finished products without the need for destructive testing.

REFERENCES

[1] A. K. W. Chee, R. F. Broom, C. J. Humphreys, and E. G. T. Bosch, "A quantitative model for doping contrast in the scanning electron microscope using calculated potential distributions and Monte Carlo simulations," *J. Appl. Phys.*, vol. 109, no. 1, Jan. 2011, Art. no. 013109, doi: 10.1063/1.3524186.

[2] A. K. W. Chee, "Fermi level pinning characterisation on ammonium fluoride-treated surfaces of silicon by energy-filtered doping contrast in the scanning electron microscope," *Sci. Rep.*, vol. 6, no. 1, p. 32003, Aug. 2016, doi: 10.1038/srep32003.

[3] H. Kwak, S. Ryu, S. Cho, J. Kim, Y. Yang, and J. Kim, "Non-destructive thickness characterisation of 3D multilayer semiconductor devices using optical spectral measurements and machine learning," *Light, Adv. Manuf.*, vol. 2, pp. 9–19, Mar. 2021, doi: 10.37188/lam.2021.001.

[4] D. Ukolov, A. Baluev, P. Gromova, A. Pechenkin, and R. Mozhaev, "Laser scanning confocal IR microscopy for non-destructive testing of semiconductors," in *Proc. Moscow Workshop Electron. Netw. Technol. (MWENT)*, Jun. 2022, pp. 1–3, doi: 10.1109/MWENT55238.2022.9802329.

[5] R. Ong and K. Y. Cheong, "Non-destructive electrical test detection on copper wire micro-crack weld defect in semiconductor device," in *Proc. IEEE 37th Int. Electron. Manuf. Technol. (IEMT) 18th Electron. Mater. Packag. (EMAP) Conf.*, Sep. 2016, pp. 1–6, doi: 10.1109/IEMT.2016.7761967.

[6] J. R. Deepak, V. K. Bupesh Raja, D. Srikanth, H. Surendran, and M. M. Nickolas, "Non-destructive testing (NDT) techniques for low carbon steel welded joints: A review and experimental study," *Mater. Today, Proc.*, vol. 44, pp. 3732–3737, Jan. 2021, doi: 10.1016/j.matpr.2020.11.578.

[7] J. Jodhani, A. Handa, A. Gautam, and R. Rana, "Ultrasonic non-destructive evaluation of composites: A review," *Mater. Today, Proc.*, vol. 78, pp. 627–632, Jan. 2023, doi: 10.1016/j.matpr.2022.12.055.

[8] M. Shaloo, M. Schnall, T. Klein, N. Huber, and B. Reitinger, "A review of non-destructive testing (NDT) techniques for defect detection: Application to fusion welding and future wire arc additive manufacturing processes," *Materials*, vol. 15, no. 10, p. 3697, May 2022, doi: 10.3390/ma15103697.

[9] K. A. Reddy, "Non-destructive testing, evaluation of stainless steel materials," *Mater. Today, Proc.*, vol. 4, no. 8, pp. 7302–7312, 2017, doi: 10.1016/j.matpr.2017.07.060.

[10] L. S. Rosado, T. G. Santos, M. Piedade, P. M. Ramos, and P. Vilaça, "Advanced technique for non-destructive testing of friction stir welding of metals," *Measurement*, vol. 43, no. 8, pp. 1021–1030, Oct. 2010, doi: 10.1016/j.measurement.2010.02.006.

[11] H. Sun, P. Ramuhalli, and R. E. Jacob, "Machine learning for ultrasonic nondestructive examination of welding defects: A systematic review," *Ultrasonics*, vol. 127, Jan. 2023, Art. no. 106854, doi: 10.1016/j.ultras.2022.106854.

[12] J. Blitz, *Electrical and Magnetic Methods of Non-Destructive Testing*. Dordrecht, The Netherlands: Springer, 1997.

[13] J. P. Yun, S. Choi, J.-W. Kim, and S. W. Kim, "Automatic detection of cracks in raw steel block using Gabor filter optimized by univariate dynamic encoding algorithm for searches (uDEAS)," *NDT & E Int.*, vol. 42, no. 5, pp. 389–397, Jul. 2009, doi: 10.1016/j.ndteint.2009.01.007.

[14] D. Yang, Y. Cui, Z. Yu, and H. Yuan, "Deep learning based steel pipe weld defect detection," *Appl. Artif. Intell.*, vol. 35, no. 15, pp. 1237–1249, Dec. 2021, doi: 10.1080/08839514.2021.1975391.

[15] Y. Guo, Z. Xiao, L. Geng, J. Wu, F. Zhang, Y. Liu, and W. Wang, "Fully convolutional neural network with GRU for 3D braided composite material flaw detection," *IEEE Access*, vol. 7, pp. 151180–151188, 2019, doi: 10.1109/ACCESS.2019.2946447.

[16] J. C. Aldrin, N. D. Schehl, V. A. Kramb, D. Zainey, J. T. Welter, J. N. Wertz, S. M. Wallentine, E. A. Lindgren, and M. D. Uchi, "Investigations of pitch-catch angled-beam ultrasonic NDE for characterization of hidden regions of impact damage in composites," in *Proc. 45th Annu. Rev. Prog. Quant. Non-Destructive Eval.*, vol. 38, Jul. 2018, pp. 15–19, doi: 10.1063/1.5099762.

[17] F. Honarvar and A. Varvani-Farahani, "A review of ultrasonic testing applications in additive manufacturing: Defect evaluation, material characterization, and process control," *Ultrasonics*, vol. 108, Dec. 2020, Art. no. 106227, doi: 10.1016/j.ultras.2020.106227.

[18] A. E. Bouzenad, S. Yaacoubi, S. Montresor, and M. Bentahar, "A model-based approach for in-situ automatic defect detection in welds using ultrasonic phased array," *Expert Syst. Appl.*, vol. 206, Nov. 2022, Art. no. 117747, doi: 10.1016/j.eswa.2022.117747.

[19] S. Cantero-Chinchilla, P. D. Wilcox, and A. J. Croxford, "Deep learning in automated ultrasonic NDE—Developments, axioms and opportunities," *NDT & E Int.*, vol. 131, Oct. 2022, Art. no. 102703, doi: 10.1016/j.ndteint.2022.102703.

[20] G. Liu, "Surface defect detection methods based on deep learning: A brief review," in *Proc. 2nd Int. Conf. Inf. Technol. Comput. Appl. (ITCA)*, Guangzhou, China, Dec. 2020, pp. 200–203, doi: 10.1109/ITCA52113.2020.00049.

[21] X. Meng, C. Shang, J. Dong, X. Fu, and P. Lang, "Automatic modulation classification of noise-like radar intrapulse signals using cascade classifier," *ETRI J.*, vol. 43, no. 6, pp. 991–1003, Dec. 2021, doi: 10.4218/etrij.2020-0338.

[22] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," *Eng. Appl. Artif. Intell.*, vol. 115, Oct. 2022, Art. no. 105151, doi: 10.1016/j.engappai.2022.105151.

[23] X. Cao, M. Jahazi, J. P. Immarigeon, and W. Wallace, "A review of laser welding techniques for magnesium alloys," *J. Mater. Process. Technol.*, vol. 171, no. 2, pp. 188–204, Jan. 2006, doi: 10.1016/j.jmatprotec.2005.06.068.

[24] J. Wegrzyn, M. Mazur, A. Szymański, and B. Balcerowska, "Development of a filler for welding magnesium alloy GA8," *Weld. Int.*, vol. 1, no. 2, pp. 146–150, Jan. 1987, doi: 10.1080/09507118709452102.

[25] Y. Fang et al., "Detection of transverse crack in weld by dual probe cross-weld scanning imaging," in *Proc. 2nd Int. Conf. Frontiers Mater. Synth. Process.* China: FMSP, 2018, Art. no. 493, doi: 10.1088/1757-899X/493/1/012055.

[26] A. C. S. Douglass et al., "Segmentation of hidden delaminations with pitch–catch ultrasonic testing and agglomerative clustering," *J. Nondestruct. Eval.*, vol. 39, no. 1, Jan. 2020, Art. no. 8, doi: 10.1007/s10921-019-0649-7.

[27] M. Alkhatib, A. Hafiane, and P. Vieyres, "Merged 1D–2D deep convolutional neural networks for nerve detection in ultrasound images," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 4774–4780, doi: 10.1109/ICPR48806.2021.9412988.

[28] J. Zhao, X. Mao, and L. Chen, "Speech emotion recognition using deep 1D & 2D CNN LSTM networks," *Biomed. Signal Process. Control*, vol. 47, pp. 312–323, Jan. 2019, doi: 10.1016/j.bspc.2018.08.035.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

KYEEUN KIM received the M.S. degree from the Department of Advanced Materials Engineering, Chonnam National University, in 2016. Since 2016, she has been with the Electronics and Telecommunications Research Institute, Yuseong-gu, Daejeon, Republic of Korea, as a Researcher. Her research interests include deep learning and ultrasonic testing.

KEO SIK KIM received the Ph.D. degree in electronics engineering from Chonbuk National University, Jeonju, South Korea, in 2011. Since 2011, he has been with the Electronics and Telecommunications Research Institute (ETRI), Yuseong-gu, Daejeon, Republic of Korea, as a Principal Researcher. His research interests include deep learning, machine vision, optical imaging, and non-destructive testing.

HYOUNG-JUN PARK received the Ph.D. degree in electronics engineering from Chonbuk National University, in 2009. He is currently the Director of the Optical ICT Convergence Section, Electronics and Telecommunications Research Institute (ETRI), Yuseong-gu, Daejeon, Republic of Korea. His current research interests include optical information technology convergence, optical engine systems, and optical subassembly.