

Defects Recognition of High-Density Polyethylene Pipe Butt-Fusion Joint via STSVD Filtering Total Focusing Method and Machine Learning

HAOWEN ZHANG, QIANG WANG* and FANGHUI FAN

ABSTRACT

The butt-fusion joints of high-density polyethylene (HDPE) pipes represent a critical vulnerability, being prone to various defects that potentially lead to structural failures. Therefore, it is essential to conduct nondestructive testing (NDT) on HDPE pipe butt-fusion joints for defect detection. This paper proposes a spatiotemporal singular value decomposition filtering total focusing method (STSVD-TFM), combined with machine learning (ML) to utilize A-scan signals for defect detection. Initially, the signal data is filtered using the STSVD method. Subsequently, feature parameters, including time domain features and spectral features, are extracted from the filtered data, and significant features are selected based on the Relief-F algorithm. Finally, the filtered signal data is employed for TFM imaging, with defect types determined by the training results of the ML models applied to the feature parameters. Detection experiments are conducted on HDPE pipe butt-fusion joint specimens, which included through-hole and square groove defects. The results demonstrate that the proposed method effectively reduces the amplitude of static clutter in the near-field areas, enhances the signal-to-noise ratio (SNR) of the detection images, and achieves high accuracy autonomous recognition of defect types.

Key Words: high-density polyethylene, butt-fusion joints, phased array ultrasonic detection, total focusing method, machine learning

INTRODUCTION

High-density polyethylene (HDPE) pipes are increasingly applied in urban natural gas pipeline due to their exceptional corrosion and oxidation resistance [1-2]. Butt-fusion welding technology, eminent for its cost-effectiveness and high reliability, which is one of the primary methods employed for welding HDPE pipes. Nevertheless, HDPE pipe joints are susceptible to manufacturing defects arising from operational errors during the welding process [3-4]. Such defects potentially lead to structural failures and may cause unintentional natural gas leaks. Therefore, to ensure the integrity of HDPE pipe butt-fusion joints, it is imperative to conduct nondestructive testing (NDT).

Phased array ultrasonic testing (PAUT), as an advanced NDT method, is widely used for defect detection in industrial pipelines [5-6]. The total focusing method (TFM) based on full matrix capture (FMC) technology has become the golden rule of post-processing technology due to its superior image quality and signal-to-noise ratio (SNR)

Haowen Zhang and Qiang Wang. College of Energy Environment and Safety Engineering, China Jiliang University, 258 Xueyuan Street, Qiantang District, Hangzhou, Zhejiang 310018, China

Fanghui Fan. Taian Quality Technical Inspection and Testing Institute (Taian Special Equipment Inspection and Research Institute), No. 339, Daizong Street, Tai'an, Shandong, 271000, China

*Corresponding author: qiangwang@cjlu.edu.cn

[7-8]. Due to the viscoelastic properties and low density of HDPE materials, ultrasonic waves propagating within the material experience significant attenuation and absorption, reducing the defect echo amplitude.

Singular value decomposition (SVD) has been justified to be effective for noise reduction of ultrasonic signals in the NDT field [9-10]. SVD is processed on a 2D Casorati matrix and filtered out the structural noise according to the singular vector distribution. Additionally, an advanced spatiotemporal singular value decomposition (STSVD) method has been applied to the PAUT of HDPE materials. Rao et al [11] used the STSVD method to filter the FMC ultrasonic 1/Q demodulated data of HDPE pipe materials, which effectively suppresses the structural noise and improves the SNR. Zhang et al [12] proposed an improved TFM algorithm based on STSVD, and the average SNR could be improved by up to 5.48 dB with imaging detection experiments of through-hole defects in HDPE test blocks.

The accuracy of PAUT for HDPE pipe butt-fusion joints often relies on manual defect recognition, which is a time-consuming and error-prone process. Recent advancements in artificial intelligence (AI) have enabled intelligent defect recognition through analyzing extensive datasets [13-14]. Machine learning (ML), a significant subset of AI, is integrated with PAUT to facilitate intelligent defect recognition by developing a model structure [15-16]. ML encompasses shallow learning (SL) and deep learning (DL). SL is characterized by its relatively simple model structure that can be trained on ultrasonic signal feature data. In contrast, DL models, known for their complex architectures, typically utilize large neural networks that are capable of processing features in imaging tasks. However, acquiring defect imaging data is challenging, and when using small datasets, image enhancement techniques are frequently employed to augment the dataset [17-18]. It is important to note that augmented defect images may cause overfitting and lead to significant errors from overlearning specific defects.

In summary, to ensure the welding quality of HDPE pipe butt-fusion joints, this paper proposes a spatiotemporal singular value decomposition filtering total focusing method (STSVD-TFM). Combined with the ML models, the intelligent recognition of defects is realized by conducting phased array ultrasonic detection experiments on HDPE pipe butt-fusion joints specimens, and training the ultrasonic A-scan data of FMC after STSVD filtering.

METHODS

SPATIOTEMPORAL SINGULAR VALUE DECOMPOSITION

The FMC matrix is converted into a 2D ($N_i \times N_j$, N_s) Casorati matrix $\mathbf{S}$, where N_i and N_j represent the number of excitation and reception elements, respectively, and N_s represent the number of sampling points. The Casorati matrix $\mathbf{S}$ using STSVD is as follows [11]:

$$\mathbf{S} = \mathbf{U} \mathbf{\Delta} \mathbf{V}^* \quad (1)$$

Where, $\mathbf{U}$ represents the left singular vector of the matrix $\mathbf{S}$. $\mathbf{\Delta}$ represents a nonsquare matrix whose diagonal elements are arranged in descending order as $[\sigma_1, \sigma_2, \dots, \sigma_r]$,

and r is the rank of the matrix $\mathbf{S}$; $\mathbf{V}$ represents the right singular vector of the matrix $\mathbf{S}$. Setting the low-order and high-order cutoff to reconstructed singular value matrices:

$$\Delta^{ST} = \Delta \times \mathbf{D}^{ST} \quad (2)$$

where, $\mathbf{D}^{ST}$ represents a diagonal matrix consisting only of elements 0 and 1. The STSVD inverse operation is then performed to obtain the filtered Casorati matrix $\mathbf{S}^{ST}$:

$$\mathbf{S}^{ST} = \mathbf{U} \Delta^{ST} \mathbf{V}^* \quad (3)$$

Finally, the 2D Casorati matrix $\mathbf{S}^{ST}$ is converted to a 3D matrix to obtain the STSVD spatiotemporal filtered FMC data matrix.

TOTAL FOCUSING METHOD

The pixel amplitude $I_{TFM}(x, z)$ of any imaging point $P(x, z)$ in a specified region of interest (ROI) is [7]:

$$I_{TFM}(x, z) = \left| \sum_{i=1}^N \sum_{j=1}^N H(A(x_i, x_j, \tau_{ij})) \right| \quad (4)$$

where, H represents the Hilbert transform of the FMC ultrasonic signal A . x_i and x_j represent the positions of the excitation and reception array elements, respectively. τ_{ij} represents the time delay between the array element and the pixel point. When the STSVD filtered FMC data A^{ST} is utilized for TFM imaging, the $I_{STSVD-TFM}(x, z)$ amplitude of the $P(x, z)$ in the ROI grid is:

$$I_{TFM}(x, z) = \left| \sum_{i=1}^N \sum_{j=1}^N H(A^{ST}(x_i, x_j, \tau_{ij})) \right| \quad (5)$$

DEFECT RECOGNITION METHOD BASED ON MACHINE LEARNING

The A-scan signals acquired by FMC are gated in the time domain to filter out both front-wall and back-wall echoes, thereby preserving the integrity of the defect echo signals. Subsequently, the feature parameters of the defective echo signals are extracted from the time and frequency domains. The Relief-F algorithm, leveraging multiple nearest neighbors for supervised classification. It evaluates feature differences between similar and dissimilar neighbors for each sample and computes the weight W for each feature parameter [14]:

$$W(\lambda) = W(\lambda) - \sum_{k=1}^K \frac{diff(\lambda, \eta, \psi_k)}{K} + \sum_{C \neq class(\eta)} \frac{(\frac{p(C)}{1 - p(class(\eta))} \sum_{k=1}^K diff(\lambda, \eta, M_k(C)))}{K} \quad (6)$$

where, λ represents the feature data, η represents the sample data. ψ_k represents the k th similar nearest neighbor of sample η , $diff$ represents the difference, $M_k(C)$ represents the k th dissimilar nearest neighbor of category C , and p represents the proportion.

This paper studies the recognition performance of three ML models: BP neural network, random forest (RF), and support vector machine (SVM). The significant feature parameters selected by the Relief-F algorithm are utilized as input data for training data. A five-fold cross-validation strategy is employed to enhance the utilization

and reliability of the data. Additionally, the Bayesian optimizer is implemented to search the optimal hyperparameters for each ML model, thereby reducing the likelihood of overfitting and underfitting during model training.

EXPERIMENTS

In this study, defect detection experiments are conducted using HDPE pipe butt-fusion joints specimens with an outer diameter with size of 315 mm and a thickness of 28.6 mm, as shown in Figure 1. The specimens contained through-hole defects of $\varnothing$ 2mm, $\varnothing$ 3mm and $\varnothing$ 4mm, and square groove defects of $3\text{mm} \times 7.5\text{mm}$ and $4.5\text{mm} \times 9\text{mm}$, respectively. Each type of defects included three samples, with defect depths of 10 mm, 16 mm, and 22 mm, respectively. The detection equipment utilizing 64 array elements, the spacing and width of the array elements are 0.75 mm and 0.6 mm, respectively. The center frequency of the transducer is 2.25 MHz, and the ultrasonic speed is 2300 m/s.

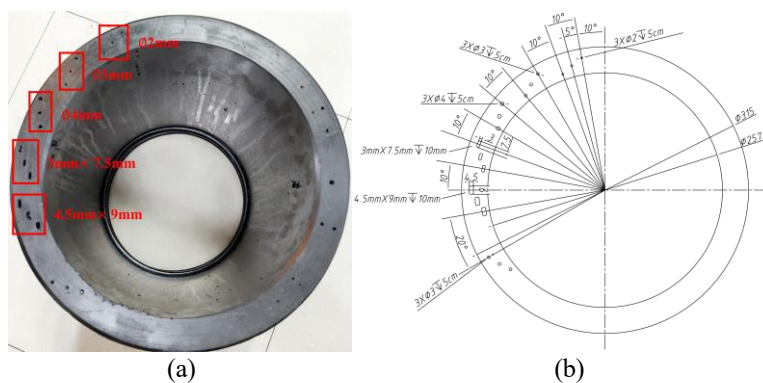


Figure 1. HDPE hot melt joint test block: (a) Physical drawing, (b) Structural drawing.

RESULTS AND DISCUSSION

DEFECTS IMAGING DETECTION RESULTS

The imaging detection results are shown in Figure 2 and Figure 3, respectively. In the TFM imaging results, as illustrated in Figure 2(a) and 2(b), the 1# defects measuring $\varnothing$ 2mm and $\varnothing$ 3mm are significantly affected by clutters in the near-field area. In contrast, the 2# defects, being farther from the near-field area, exhibit elevated amplitude. However, with increasing defect depth, the amplitude of the 3# defects are reduced by the viscoelastic attenuation. In Figure 2(d) and (e), the expansion of the reflection area of the square groove defects leads to the elevation of the noise amplitude near the defects, and even false positive indications such as 15mm-18mm. In the STSVD-TFM imaging results, the static clutter in the near-field area is effectively suppressed and the amplitude is reduced by approximately from 10 dB to 12 dB. In the imaging results of the through-hole defects, as shown in Figures 3(a), (b) and (c), the noise amplitude near the defects is lower than that of the TFM images. In the imaging results of the square groove defects (Figures 3(d), (e)), most of the noise near the defects

is suppressed, and the amplitude of the false-positive indications beneath the 1# defects is reduced compared to the TFM images, although it is not completely eliminated.

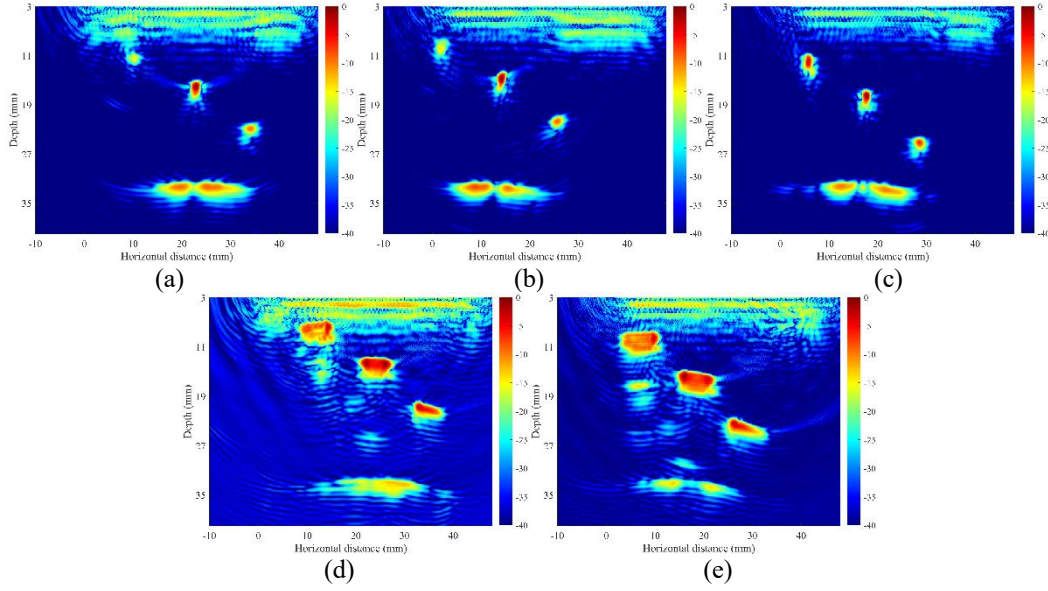


Figure 2. TFM imaging results: (a) $\varnothing$ 2mm, (b) $\varnothing$ 3mm, (c) $\varnothing$ 4mm, (d) 3mm $\times$ 7.5mm, (e) 3mm $\times$ 7.5mm.

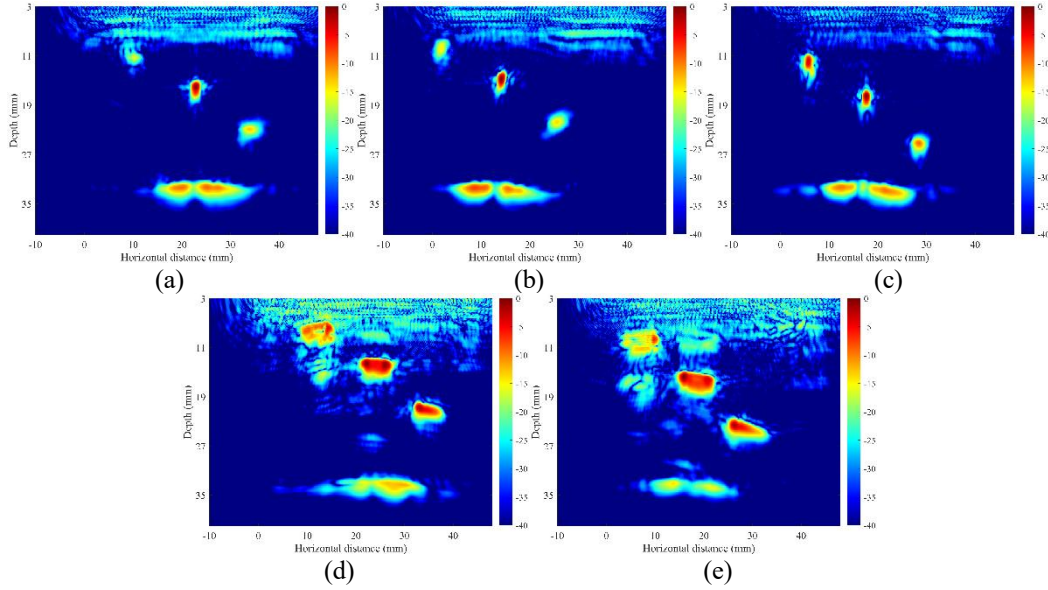


Figure 3. STSVD-TFM imaging results: (a) $\varnothing$ 2mm, (b) $\varnothing$ 3mm, (c) $\varnothing$ 4mm, (d) 3mm $\times$ 7.5mm, (e) 3mm $\times$ 7.5mm.

The SNR metric is utilized to quantitatively assess the imaging performance, and the SNR is calculated as follows [10]:

$$SNR = 20 \log_{10} \left| \frac{I_{\max}}{I_{\text{average}}} \right| \quad (7)$$

where, I_{max} represents the maximum amplitude in the defective areas, and $I_{average}$ represents the average amplitude in the non-defective areas. The calculated average SNR results are shown in Figure 4. It is evident that the average SNR of the STSVD-TFM imaging results consistently exceeds that of the TFM imaging, with improvements of up to 3.4 dB. This indicates the superior performance of STSVD-TFM imaging.

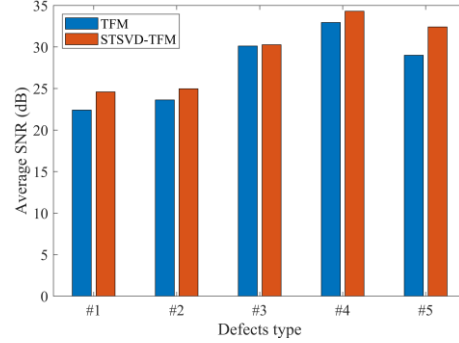


Figure 4. Results of average SNR: #1 to #3 represent $\varnothing 2\text{mm}$, $\varnothing 3\text{mm}$ and $\varnothing 4\text{mm}$ through-hole defects, #4 to #5 represent $3\text{mm} \times 7.5\text{mm}$ and $4.5\text{mm} \times 9\text{mm}$ square groove defects, respectively.

ML-BASED DEFECTS RECOGNITION

By manually setting the time gate on the FMC ultrasonic signals, the echoes from both the front and back walls are effectively filtered out. The A-scan time-domain waveforms for all defect types are presented in Figure 5. As the size of the defect increases, the ultrasonic waves reflect over a larger range, resulting in higher signal amplitude peaks and an increase in time sampling.

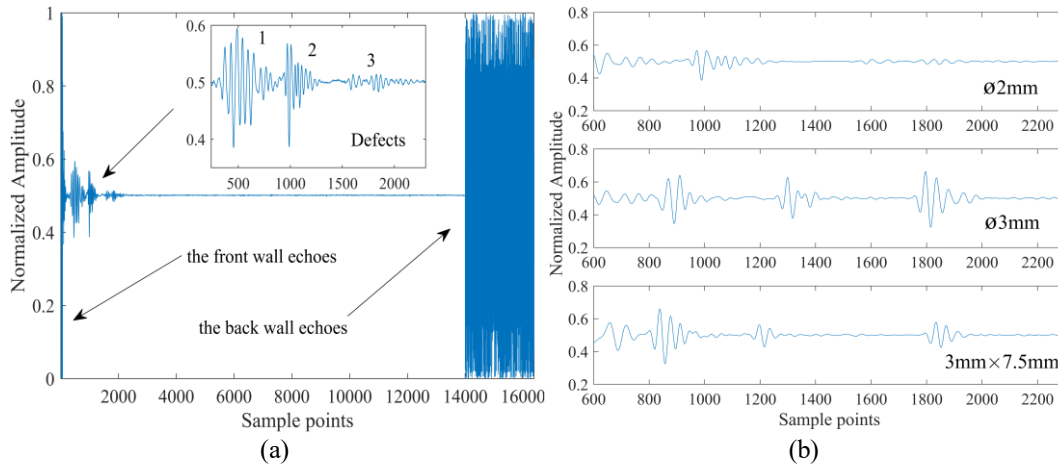


Figure 5. The time domain A-scan echo signals: (a) filter out the front and back wall echoes, (b) A-scan echoes of $\varnothing 2\text{mm}$ and $\varnothing 3\text{mm}$ through-hole defects, as well as $3\text{mm} \times 7.5\text{mm}$ square groove defects.

The Relief-F algorithm is utilized to select the 12 significant features to be used as the input data for the ML models. A-scan signals devoid of defects are excluded,

resulting in the creation of a feature data matrix with dimensions 13313×12 . The data were normalized, with 80% of them are allocated for training and the remaining 20% are designated as the test dataset. The ML models are trained using FMC and STSVD-FMC data, respectively, and the performance reliability of each ML model is quantitatively assessed using precision (P), recall (R), and F1 score, with the calculated performance evaluation metrics are shown in Table I.

TABLE I. INDICATOR VALUES FOR PERFORMANCE EVALUATION OF EACH ML MODEL.

Models	Dataset	Category	P	R	F1	Macro P	Weighted P
BP neural network	FMC	1	0.9234	0.8462	0.9139	0.9020	0.9020
		2	0.8806	0.9015	0.8920		
	STSVD-FMC	1	0.9400	0.9063	0.9508	0.9381	0.9437
		2	0.9362	0.9130	0.9304		
RF	FMC	1	0.9633	0.9684	0.9658	0.9577	0.9588
		2	0.9520	0.9445	0.9482		
	STSVD-FMC	1	0.9456	0.9669	0.9561	0.9470	0.9467
		2	0.9484	0.9162	0.9302		
SVM	FMC	1	0.9533	0.9724	0.9628	0.9554	0.9505
		2	0.9574	0.9286	0.9428		
	STSVD-FMC	1	0.9723	0.9844	0.9783	0.9741	0.9737
		2	0.9759	0.9575	0.9666		

Combining the performance metrics of the ML models for category 1 (through-hole defects) and category 2 (square groove defects), it can be seen that the recognition accuracy of through-hole defects is higher than that of square groove defects. Notably, the accuracy is higher than the FMC data training results when using STSVD-FMC data. Furthermore, the SVM model performs better compared to the other two methods, and the micro-accuracy and weighted accuracy reached 0.9741 and 0.9737, respectively, when trained on STSVD-FMC data.

CONCLUSION

In order to ensure the welding quality of HDPE pipe butt-fusion joints, this paper proposes a TFM imaging algorithm based on STSVD filter processing and combines the ML models to train the ultrasonic A-scan data for the intelligent recognition of defects. Through imaging detection experiments on HDPE pipe butt-fusion joints, the STSVD method effectively suppresses static clutter in the near-field area and noise near the defects, resulting in an improvement of the SNR by up to 3.4 dB compared with the TFM image. For intelligent recognition of defects, BP neural network, RF, and SVM models are established to train FMC and STSVD-FMC feature data. The model performance metrics indicate that the SVM model utilizing STSVD-FMC data has the highest defect recognition rate, and the Macro P and Weighted P can reach 0.9741 and 0.9737, respectively.

Funding: This research was supported by the National Key Research and Development Program of China (No. 2023YFF0611600).

Disclosure statement: The authors declare no conflicts of interest.

REFERENCES

1. Gong, Y., Wang, S. H., Zhang, Z., Yang, X., Yang, Z., Yang, H. 2021. "Degradation of Sunlight Exposure on the High-Density Polyethylene (HDPE) Pipes for Transportation of Natural Gases," *Polymer Degradation and Stability*, 194, 109752.
2. Guo, H., Rinaldi, R. G., Tayakout, S., Broudin, M., Lame, O. 2021. "The Correlation between the Mixed-Mode Oligo-Cyclic Loading Induced Mechanical and Microstructure Changes in HDPE," *Polymer*, 224, 123706.
3. Lai, H., Tun, N., Yoon, K., Kil, S. 2016. "Effects of Defects on Failure of Butt Fusion Welded Polyethylene Pipe," *International Journal of Pressure Vessels and Piping*, 139 – 140, 117 – 122.
4. Zheng, J., Zhang, Y., Hou, D., Qin, Y., Guo, W., Zhang, C., and Shi, J. 2018. "A Review of Nondestructive Examination Technology for Polyethylene Pipe in Nuclear Power Plant," *Frontiers of Mechanical Engineering*, 13, 535-545.
5. Egerton, J., Lowe, M., Huthwaite, P., Halai, H. 2017. "Ultrasonic Attenuation and Phase Velocity of High-Density Poly-ethylene Pipe Material," *The Journal of the Acoustical Society of America*, 141 (3), 1535 – 1545.
6. Li, W., Zhou, Z., and Li, Y. 2019. "Inspection of Butt Welds for Complex Surface Parts Using Ultrasonic Phased Array," *Ultrasonics*, 96, 75-82.
7. Holmes, C., Drinkwater, B., Wilcox, P. 2005. "Post-Processing of the Full Matrix of Ultrasonic Transmit – Receive Array Data for Non-Destructive Evaluation," *NDT & E International*, 38 (8), 701 – 711.
8. Wu, J., Hu, H., Song, Y., Ni, P., and Li, X. 2024. "Ultrasonic Phased Array Imaging of Multi-Layered Structures Using Full-Matrix Migration and Normalized Cross-Correlation Matching Technique," *NDT & E International*, 145, 103130.
9. Rao, J., Zeng, L., Liu, M., Fu, H. 2023. "Ultrasonic Defect Detection of High-Density Polyethylene Pipe Materials Using FIR Filtering and Block-Wise Singular Value Decomposition," *Ultrasonics*, 107088.
10. Zhang, H., Wang, Q., Song, Y., Zhou, H., Wu, L., Jin, Z., Wang, D. and Hu, D. 2024. "LNG Storage Tanks 9% Ni Steel Weld Inspection via Phased-Coherent TFM and SVD Enhancement," *Nondestructive Testing and Evaluation*, 1-17.
11. Rao, J., Qiu, H., Teng, G., Al Mukaddim, R., Xue, J. and He, J. 2022. "Ultrasonic Array Imaging of Highly Attenuative Materials with Spatio-Temporal Singular Value Decomposition," *Ultrasonics*, 124, 106764.
12. Zhang, H., Wang, Q., Zhou, J., Wu, L., Xu, W., and Wang, H. 2024. "High-Density Polyethylene Pipe Butt-Fusion Joint Detection via Total Focusing Method and Spatiotemporal Singular Value Decomposition," *Processes*, 12(6), 1267.
13. Shafiei, M., Scott, R., Seviaryn, F., and Maev, R. 2021. "Using Machine Learning to Automate Ultrasound-based Classification of Butt-Fused Joints in Medium-Density Polyethylene Gas Pipes," *The Journal of the Acoustical Society of America*, 150(1), 561-572.
14. Hawwat, S., Shah, J., and Wang, H. 2025. "Machine Learning Supported Ultrasonic Testing for Characterization of Cracks in Polyethylene Pipes," *Measurement*, 240, 115609.
15. Tunukovic, V., McKnight, S., Mohseni, E., Pierce, S., Pyle, R., Duernberger, E., Loukas, C., Vithanage, R.K.W., Lines, D., Dobie, G., Macleod, C.N., Cochran, S., and O'Hare, T. 2024. "A Study of Machine Learning Object Detection Performance for Phased Array Ultrasonic Testing of Carbon Fibre Reinforced Plastics," *NDT & E International*, 144, 103094.
16. Sorger, G., Virkkunen, I., and Söderholm, C. 2025. "Flaw Sizing with Plane Wave Imaging (PWI)–Total Focusing Method (TFM) and Deep Learning for Reactor Pressure Vessel," *NDT & E International*, 103332.
17. Medak, D., Posilović, L., Subašić, M., Budimir, M., and Lončarić, S. 2022. "DefectDet: A Deep Learning Architecture for Detection of Defects with Extreme Aspect Ratios in Ultrasonic Images," *Neurocomputing*, 473, 107-115.
18. Chen, H., and Tao, J. 2024. "Utilizing Improved YOLOv8 based on SPD-BRSA-AFPN for Ultrasonic Phased Array Non-Destructive Testing," *Ultrasonics*, 142, 107382.