International Journal of Science and Engineering Applications
Volume 15-Issue 03, 72 – 79, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1503.1015

Intelligent Identification of Pipeline Weld Defects by Fusing Transformer and Faster R-CNN

Chen Cai
Yangtze University
School of Electronics Information and Electrical Engineering
Jingzhou, Hubei, China

Peng Zheng
Yangtze University
School of Electronics Information and Electrical Engineering
Jingzhou, Hubei, China

JunBin Duan
Yangtze University
School of Electronics Information and Electrical Engineering
Jingzhou, Hubei, China

Abstract: To address the issues in ultrasonic testing of pipeline welds, including reliance on manual expertise, insufficient generalization capability of deep learning models, and scarcity of annotated samples, this paper proposes an intelligent identification method based on Gramian Angular Field (GAF) and an improved Transformer-Faster R-CNN. By converting one-dimensional ultrasonic time-series signals into two-dimensional feature images via GAF encoding, the method preserves the temporal dependencies and spatial patterns of defects. A hybrid backbone network integrating ResNet-101 and a Transformer encoder is introduced to combine local feature extraction with global contextual awareness. Additionally, a spatial attention module is embedded in the Region Proposal Network (RPN) to enhance the mod-el's focus on discriminative defect features. Experimental results demonstrate that, compared to models such as ResNet, VGG, Vision Transformer (Vit), and the original Faster R-CNN, the proposed method achieves significant improvements in accuracy, precision, recall, and F1-score, reaching up to 96.94%, 97.21%, 96.96%, and 96.90%, respectively. It exhibits superior identification accuracy and robust-ness, particularly for small-sized and complex-shaped defects. This approach provides an efficient solution for the automated and intelli-gent identification of pipeline weld defects.

Keywords: Ultrasonic Testing; Pipeline Welds; Defect Identification; Deep Learning; Transformer; Faster R-CNN; Gramian Angular Field (GAF); Feature Encoding

1. INTRODUCTION

Defect recognition is the ultimate objective of ultrasonic nondestructive testing for pipeline welds. Its core task is to automatically and accurately determine the presence of defects, classify their types (e.g., porosity, cracks, lack of fusion), evaluate their geometric dimensions (length, height), and precisely locate them from processed signals or images[1-3]. With the surge in data volume from modern inspection techniques and the growing demand for intelligent solutions, defect recognition technology has progressively transitioned from an experience-dependent model reliant on manual interpretation to a new stage of automation and intelligence. This evolution is of great significance for ensuring the safe operation of energy pipelines.

Traditional defect recognition methods heavily depend on the professional knowledge and practical experience of inspectors. Operators typically assess defect characteristics subjectively by observing features such as echo amplitude, temporal position, and waveform morphology in A-scan displays. The widespread application of Phased Array Ultrasonic Testing (PAUT) technology[4-5] has introduced two-dimensional sector scans and three-dimensional volumetric imaging, transforming traditional one-dimensional waveform signals into intuitive 2D/3D images. This advancement significantly reduces reliance on operator experience and enhances the intuitiveness and reliability of defect recognition[6-7]. In these images, different types of defects usually appear as bright spots or regional distributions with specific morphological characteristics, enabling operators to perform identification and quantitative analysis more directly. However, this approach fundamentally remains within the realm of "human intelligence," where recognition efficiency and consistency are difficult to guarantee, and it is susceptible to subjective factors from operators.

Currently, the forefront of defect recognition research has fully entered the era of deep learning, particularly target detection and image segmentation techniques based on Convolutional Neural Networks (CNN)[8-9]. Researchers extensively adopt established CNN-based object detection models (e.g., Faster R-CNN[10-11], YOLO[12-13] series) by training them on large, annotated ultrasonic image datasets. This allows the models to automatically learn deep visual features of defects and directly output bounding box locations and class confidence scores. Going a step further, semantic segmentation models (e.g., U-Net [14-15]) can classify every pixel in an image accurately, enabling fine delineation of defect contours. This offers unparalleled advantages for the quantitative assessment of irregularly shaped defects. For one-dimensional A-scan signals, researchers have also developed 1D CNNs or hybrid models combining Recurrent Neural Networks (RNNs), capable of end-to-end learning directly from raw waveform data to accomplish defect classification and localization tasks [16-17].

Despite the significant success of deep learning in defect recognition, its practical application in pipeline weld ultrasonic testing still faces specific challenges. Firstly, acquiring large quantities of high-quality, accurately annotated defect samples in industrial field settings is extremely difficult and costly, especially for rare but hazardous defect types. This necessitates techniques such as few-shot learning [18-19], transfer learning[20][21], and intelligent data augmentation[22][23]. Secondly, model generalization capability and robustness are critical, as factors like material variations, changing operating conditions, and noise interference in real-world inspection environments can lead to significant performance degradation [24]. Furthermore, the limited interpretability of deep learning models hinders their application in safety-critical scenarios.

International Journal of Science and Engineering Applications
Volume 15-Issue 03, 72 – 79, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1503.1015

Addressing the aforementioned challenges, this paper proposes an improved Transformer-Faster R-CNN architecture as the core recognition model. Building upon the traditional Faster R-CNN, this model incorporates a Transformer-based multi-scale feature enhancement module, effectively capturing global contextual information of defect regions through the self-attention mechanism. This architecture fully leverages the Transformer's strength in global feature modeling while maintaining the precise localization capability of Faster R-CNN, providing reliable technical support for the accurate identification of weld defects.

2. PRINCIPLES AND METHODS

2.1 Gram-Charlier expansion principle

Ultrasonic testing signals are inherently one-dimensional time-series data, where defect features are embedded in amplitude, frequency, and temporal correlations. Traditional waveform analysis methods struggle to extract deep feature patterns effectively. To address this, the study introduces the Gramian Angular Field (GAF) transformation technique to encode one-dimensional ultrasonic signals into two-dimensional feature images.

The core idea of GAF is to map time-series signals to a polar coordinate system, preserving the angular relationships between signal points to construct a Gramian matrix, which is then converted into a 2D image. The specific steps are:

Data normalization: The one-dimensional ultrasonic signal is normalized to the interval [-1, 1] to eliminate amplitude variations and enhance feature sensitivity. The normalized signal is then transformed into a polar coordinate system, where the angle of each data point is calculated using Equation (2):

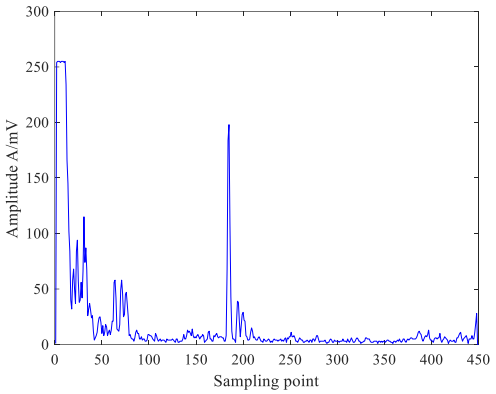
$$t_{i,norm} = 2 \times \frac{t_i - \min(T)}{\max(T) - \min(T)} - 1 \quad (1)$$

$$\phi_i = \arccos(t_i), \quad t_i \in [-1,1] \quad (2)$$

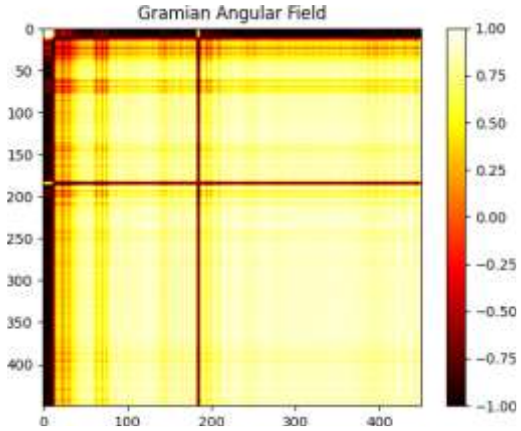
Construction of GASF matrix: The Gramian Angular Summation Fields (GASF) matrix is generated using the cosine function of summed angles to capture global correlation features of the spectrum. The calculation formula is as follows:

$$\mathrm{GASF} = \cos\left(\phi_i + \phi_j\right) = \begin{bmatrix} \cos(\phi_1+\phi_1) & \cos(\phi_1+\phi_2) & \dots & \cos(\phi_1+\phi_n) \\ \cos(\phi_2+\phi_1) & \cos(\phi_2+\phi_2) & \dots & \cos(\phi_2+\phi_n) \\ \vdots & \vdots & \vdots & \vdots \\ \cos(\phi_n+\phi_1) & \cos(\phi_n+\phi_2) & \dots & \cos(\phi_n+\phi_n) \end{bmatrix} \quad (3)$$

This transformation preserves the temporal dependencies and relative angular information of the original signal, mapping deep features such as amplitude variations and phase characteristics into textural patterns and spatial configurations in the image. Ultrasonic signals from pipeline welds collected at a reference temperature of 25°C were transformed using GAF, successfully constructing an image dataset encompassing five major defect types: lack of fusion, porosity, cracks, slag inclusion, and incomplete penetration. Each defect category contains over one thousand samples, providing a sufficient data foundation for subsequent deep learning model training. The results are shown in Figure 1.



(a) Measurement signal



(b) 2D ultrasonic image from GAF

Fig. 1 Clean signal and image conversion

2.2 Improved Transformer-Faster R-CNN Recognition Model

After obtaining the two-dimensional feature images transformed by GAF, this study employs an improved Transformer-Faster R-CNN model for automatic defect recognition and classification. Built upon the traditional Faster R-CNN, this model incorporates a Transformer module, forming a hybrid architecture that integrates local feature extraction with global context awareness.

First, the GAF images are input into a hybrid backbone network consisting of ResNet-101 and a Transformer encoder. ResNet-101 extracts local features from the images, while the Transformer encoder captures long-range dependencies among features through its self-attention mechanism. The multi-head attention mechanism can be expressed as:

$$MultiHead(Q,K,V) = Concat(head_1,...,head_h)W^O \quad (4)$$

where MultiHead denotes multi-head attention, Q, K, V represent the queries, keys, and values, Concat combines multiple heads, $head$ is a single attention head, and W^O is another learnable weight matrix. The calculation for each head is:

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (5)$$

$$Attention(Q,K,V) = softmax((QK^T)/sqrt(d_k))V \quad (6)$$

where d_k is the dimension of each head, and softmax represents the feed-forward process of multi-head attention.

International Journal of Science and Engineering Applications
Volume 15-Issue 03, 72 – 79, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1503.1015

The overall loss function of the model consists of two parts: RPN loss and Fast R-CNN loss: $L = L_{RPN} + L_{FastRCNN}$ (7)

$$L_{RPN} = L_{cls\,RPN} + L_{reg\,RPN} \quad (8)$$

$$L_{FastRCNN} = L_{cls\,RCNN} + L_{reg\,RCNN} \quad (9)$$

Among them, the classification loss(L_{cls}) uses Focal Loss to address the class imbalance problem, and the regression loss (L_{reg})uses Smooth L1 Loss.

The network structure of the Transformer-Faster R-CNN recognition model is shown in Table 1.

Table 1 Parameters of each layer of the model

Stage	Name	Input/Output/ Kernel/Stride/ Padding	Activation
CNN	Conv1t	3/64/7*7/2/3	ReLU
	MaxPool1	64/64/3*3/2/1	-
	Conv2_1	64/128/3*3/1/1	ReLU
	Conv2_2	128/256/3*3/1/1	ReLU
	AdaptiveAvgPool	256/256/-/-/-	-
Projection	Projection	256/512/1*1/1/0	-
Transformer	Trans Layer 1	512/512/-/-/-	ReLU
	MH Attention	512/512/-/-/-	-
	Feed Forward	512/2048→512/-/-/-	ReLU
	Trans Layer 2	512/512/-/-/-	ReLU
	MH Attention	512/512/-/-/-	-
	Feed Forward	512/2048→512/-/-/-	ReLU
FPN	FPN Level 0-2	512/256/-/-/-	-
RPN	RPN Conv	256/256/3*3/1/1	ReLU
	RPN Classifier	256/18/1*1/1/0	-
	RPN Regressor	256/36/1*1/1/0	-
ROI	ROI Align	-	-
	ROI Classifier	256×7²/C+1	-
	ROI Regressor	256×7²/4(C+1)	-

2.3 Evaluation metrics

To comprehensively evaluate model performance, four metrics—Accuracy, Recall, Precision, and F1-score—are employed for result analysis. Accuracy reflects the overall correctness of the model across all samples. Precision indicates the proportion of samples predicted as positive that are truly positive. Recall measures the proportion of actual positive samples correctly identified. The F1-score balances Precision and Recall to avoid biases from single metrics. The calculation formulas for each evaluation metric are as follows:

$$\text{Acurracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (10)$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (11)$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (12)$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Presicion} + \text{Recall}} \quad (13)$$

where: TP represents the number of positive samples correctly identified; TN represents the number of negative samples correctly identified; FN represents the number of positive samples misclassified; FP represents the number of negative samples misclassified.

3. EXPERIMENTAL PREPARATION

3.1 Experimental equipment

To verify the effectiveness of the Transformer-Faster R-CNN model for classifying ultrasonic signals from pipeline welds, signals were acquired using a self-developed ultrasonic flaw detector (Figure 2). The detector consists of a host unit and an ultrasonic probe. The host unit integrates a high-speed signal acquisition and processing module, signal amplification and filtering circuits, a data storage unit, and a control system, enabling multi-channel synchronous signal acquisition, real-time analysis, and intelligent defect recognition. The ultrasonic probe comprises a metal protective casing, piezoelectric film, backing damping block, acoustic impedance matching layer, wedge, and sound-absorbing block. The piezoelectric film is made of PZT-5H material, shaped as a 10×10 mm square. To simulate real defects, different types of artificial defects were prefabricated in the weld. Ultrasonic signals from pipeline welds were acquired in a test chamber (Figure 3), with 450 sampling points per signal.

(a) Main unit of the detector

(b)PMUT linear array ultrasound probe

Fig. 2 Ultrasonic flaw detection intelligent detector

(a) Environmental simulation chamber

(b) Pipe specimens with weld defects

Fig.3 Environmental simulation chamber and pipeline specimens with weld defects

3.2 Pipeline Weld Joint Classification

According to the "Classification and Description of Metal Fusion Weld Defects", weld defects can be divided into six major categories: cracks, porosity, slag inclusion, lack of fusion, incomplete penetration, and shape defects. Each main category can be further subdivided into multiple subclasses based on their location and morphological characteristics. For example, cracks can be classified as transverse cracks in the weld or transverse cracks in the heat-affected zone. However, due to the limitations of current detection technology, precise discrimination of each defect subclass is still difficult to achieve. Therefore, in engineering practice, pipeline weld defects are typically categorized into the above six main types, with shape defects assessed based on conventional welding conditions. In this experiment, based on the statistics from the factory inspection certificate of pipeline weld defects (as shown in Table 2), ten weld defect detection points were established, covering five typical defect types. Specifically, defect points 1–5 correspond to porosity, incomplete penetration, cracks, lack of fusion, and slag inclusion, respectively, while defect points 6–10 correspond to the same five defect types in the same order.

Table 2. Factory inspection certificate data of pipeline weld defects

No.	Defect Type	Defect No.	Position (mm)	Depth (mm)	Length (mm)
1	Porosity	1S	1S1=73	5	L1=25
2	Incomplete penetration	2S	2S1=307	10	L2=23
3	Crack	3S	3S1=458	Surface	L3=10
4	Lack of fusion	4S	4S1=661	8	L4=18
5	Slag inclusion	5S	5S1=848	7	L5=24
6	Porosity	6S	6S1=1041	5	L6=14
7	Incomplete penetration	7S	7S1=1231	10	L7=40
8	Crack	8S	8S1=1445	Surface	L8=30
9	Lack of fusion	9S	9S1=1592	8	L9=63
10	Slag inclusion	10S	10S1=1802	7	L10=47

Porosity refers to cavities formed when bubbles in the molten pool fail to escape before metal solidification. These bubbles primarily originate from gases generated by physical and chemical reactions in the welding material. In the Gramian Angular Field (GAF) image, porosity defects appear as short-duration, high-amplitude, and energy-concentrated features. The overall image distribution is relatively isolated, with no obvious connections between pores, consistent with their discrete nature. The corresponding image is shown in Figure 4(a).

Slag inclusion occurs when molten slag and other non-metallic impurities fail to fully float to the surface and become trapped in the weld during welding, often due to excessively low current or high travel speed. Inclusions mainly consist of non-metallic compounds, flux residues, and oxides. In the GAF image, slag inclusion exhibits no distinct central concentration, with drastic blackness variations, reflecting its complex composition and random distribution. The corresponding image is shown in Figure 4(b).

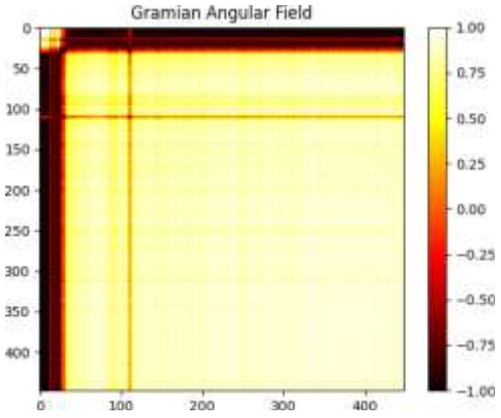
Cracks are fissures formed in the welded joint region due to the combined effects of multiple factors, representing one of the most hazardous common defects in welded structures. In the GAF image, cracks are characterized by directional reflections, concentrated energy, and short duration. The image exhibits linear, parallel, or intersecting textures, reflecting the propagative nature and severity of cracks. The corresponding image is shown in Figure 4(c).

Lack of fusion refers to interface-type defects where incomplete bonding occurs between weld metal and base material or between adjacent weld passes. In the GAF image, lack of fusion displays strong interface reflections and uneven energy distribution. The defect region shows clear contrast with the background, consistent with its distinct interface and fixed location characteristics. The corresponding image is shown in Figure 4(d).

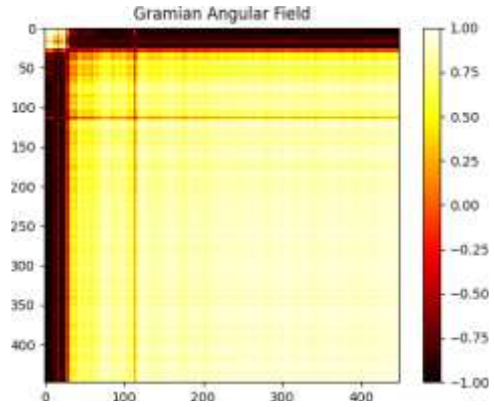
Incomplete penetration refers to the failure of complete fusion at the weld root, commonly observed in single-sided and double-sided welding. In the GAF image, incomplete

International Journal of Science and Engineering Applications
Volume 15-Issue 03, 72 – 79, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1503.1015

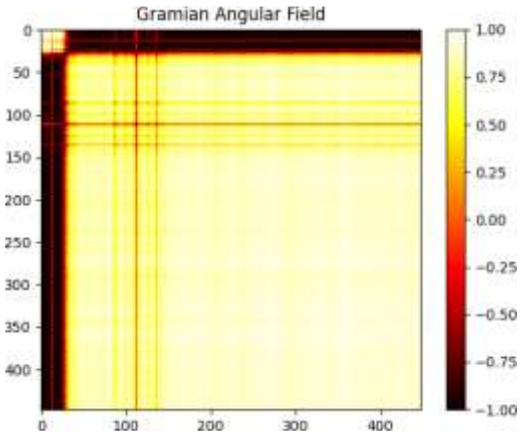
penetration typically appears as a continuous or intermittent dark line with relatively straight edges and a certain degree of continuity, reflecting its fixed position and regular morphology. The corresponding image is shown in Figure 4(e).



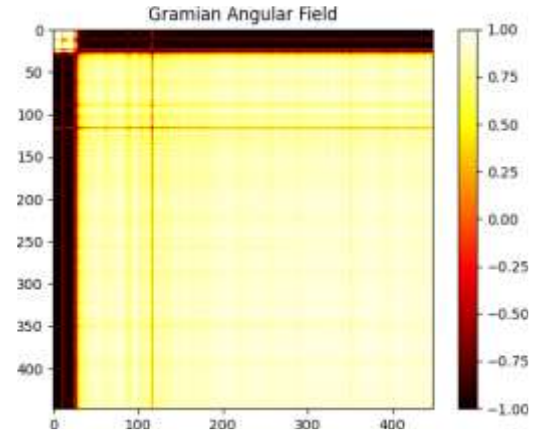
(a)Porosity



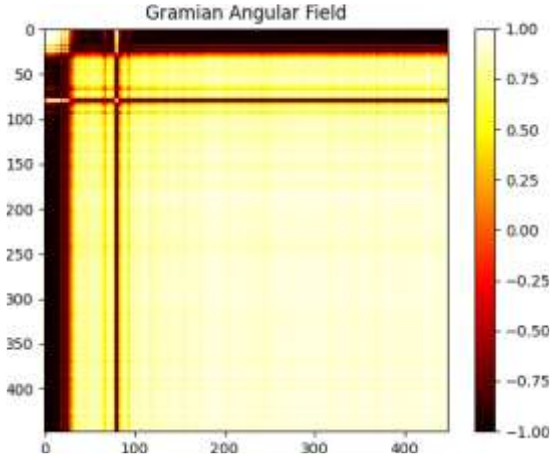
(b)Slag Inclusion



(c)Crack



(d)Lack of Fusion



(e)Incomplete Penetration

Fig. 4 GAF-based Images of Different Weld Defects

3.3 Dataset

As shown in Table 2, a total of 1,199 GAF images were obtained for porosity (599 from defect No. 1 and 600 from No. 6), 1,199 for incomplete penetration (599 from No. 2 and 600 from No. 7), 1,199 for cracks (599 from No. 3 and 600 from No. 8), 1,200 for lack of fusion (600 each from Nos. 4 and 9), and 1,199 for slag inclusion (600 from No. 5 and 599 from No. 10). From the acquired dataset, 70% of the images were randomly selected as the training set, and the remaining 30% as the test set.

4. TRAINING AND RESULTS ANALYSIS

Due to the complex structure and high computational demands of neural networks, training on CPUs is often time-consuming and insufficient for practical applications. Therefore, a GPU (GeForce RTX 4060, 8GB memory) was used for training in this study. The models were implemented in Python using the PyTorch framework on the Windows operating system.

4.1 Experimental Methods

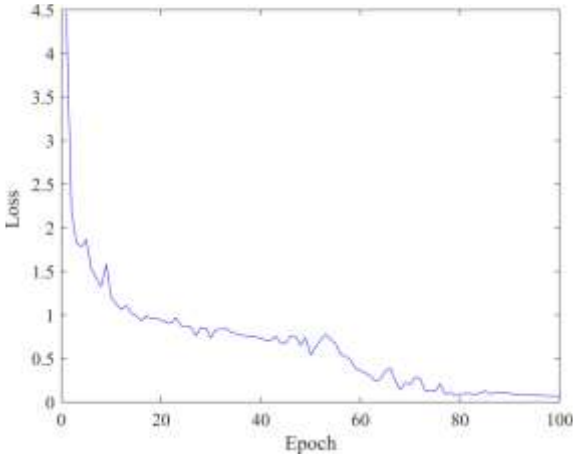
To validate the effectiveness of the Transformer-Faster R-CNN model for classifying ultrasonic signals from pipeline welds, the following evaluation was conducted. Ultrasonic signals were converted into GAF images, and the model's performance was compared with ResNet, VGG, ViT, and

International Journal of Science and Engineering Applications
Volume 15-Issue 03, 72 – 79, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1503.1015

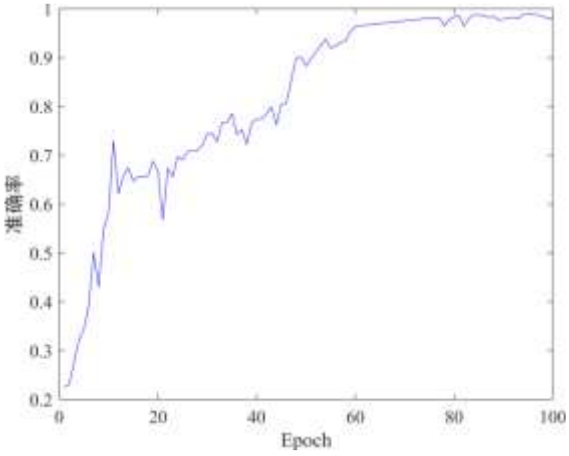
Faster R-CNN using Accuracy, Recall, Precision, and F1-score.

4.2 Training Results

During model training, the Transformer-Faster R-CNN was trained using the training set images, with accuracy (ACC1) and loss curves recorded for each epoch to evaluate its recognition performance. The trained model was then validated on the test set. As shown in Figure 5, with increasing iterations, ACC1 gradually rose and stabilized at 0.98, while the loss converged to near zero and remained stable.



(a) Loss variation curve during model training

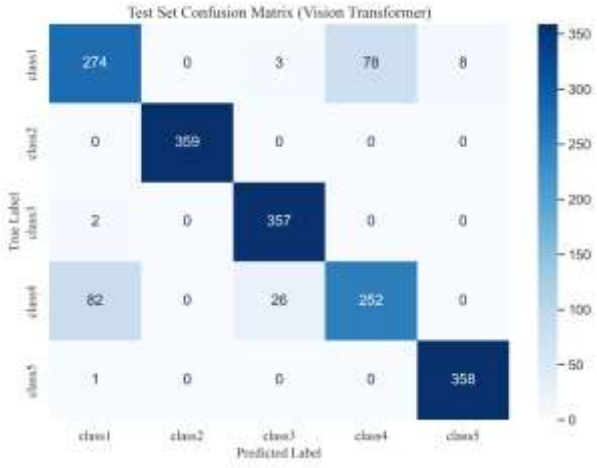


(b)ACC1 variation curve during model training

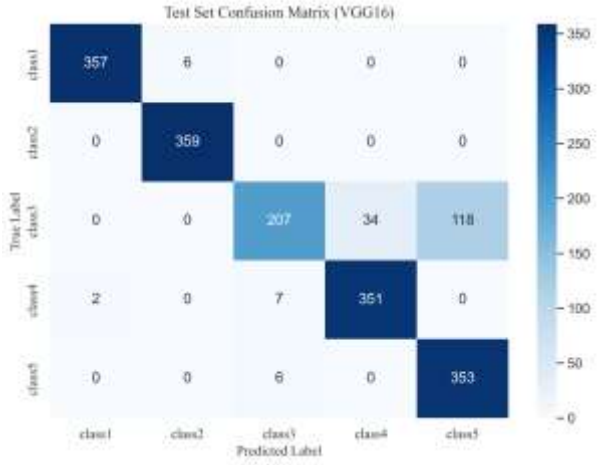
Fig. 5 Training results of the Transformer-Faster R-CNN model

4.3 Test Result Comparison Experiment

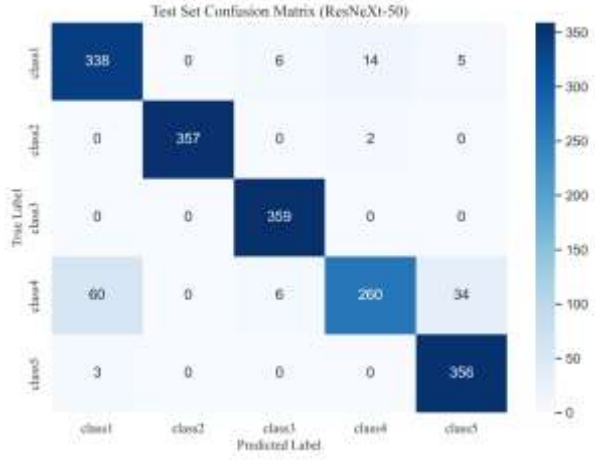
Figure 6 presents the confusion matrix for ultrasonic identification of pipeline weld defect types, where Classes 1-5 correspond to porosity, incomplete penetration, cracks, lack of fusion, and slag inclusion, respectively. The vertical axis represents the actual defect types, while the horizontal axis represents the defect types identified by the model. All correct identifications are recorded in the diagonal cells of the matrix. As shown in Figure 6, the Transformer-Faster R-CNN network model adopted in this study demonstrates excellent performance in identifying and classifying pipeline weld defect types when compared with ResNet, VGG, ViT, and Faster R-CNN. The model exhibits strong performance across all evaluation metrics, indicating high recognition accuracy and good generalization ability.



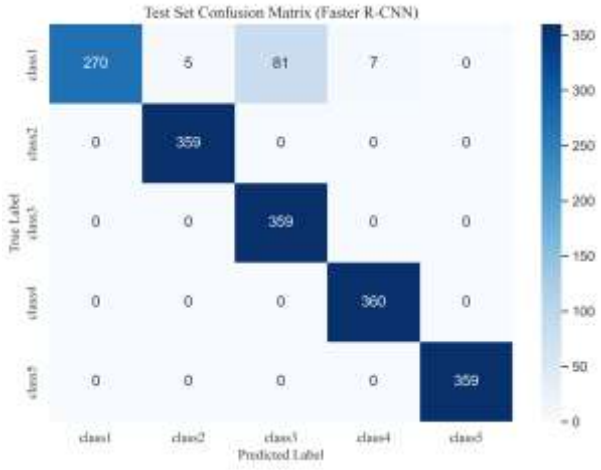
(a) Vit



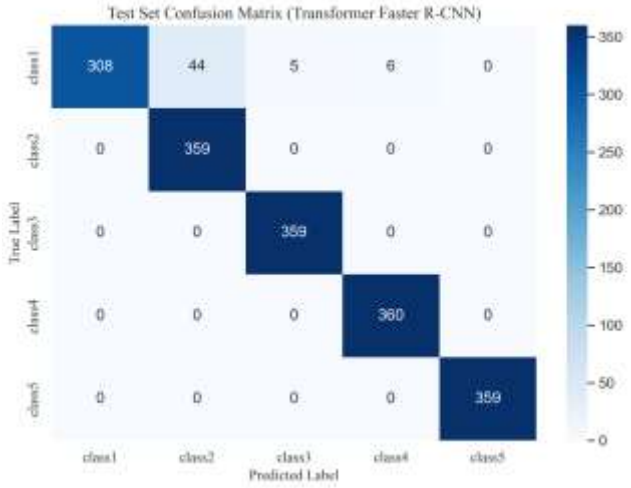
(b) VGG



(c)ResNet50



(d) Faster-RCNN



(e) Transformer-Faster R-CNN

Fig. 6 Comparison of confusion matrices for weld defect identification across different models

Table 3 presents the overall average metrics of different models tested in this study. Table 4 shows the overall evaluation results of the Transformer-Faster R-CNN model for ultrasonic identification of pipeline weld defects, including Precision, Recall, and F1-score validated on original images of five defect types. The F1-scores and average metrics for the five defect types are all close to 96.90%, indicating that the model achieves high recognition accuracy for weld defect classification.

Table 3 Overall average metrics for ultrasonic data across different models

Type	Accuracy	Precision	Recall	F1-score
Transformer-Faster R-CNN	0.9694	0.9721	0.9696	0.9690
Faster-RCNN	0.9483	0.9566	0.9487	0.9470
ResNet50	0.9277	0.9307	0.9278	0.9254
VGG	0.903889	0.9160	0.9036	0.8977
Vit	0.8888	0.8859	0.8892	0.8870

Table 4 Comprehensive classification evaluation results for ultrasonic data using the Transformer-Faster R-CNN model

Type	Precision	Recall	F1-score
Porosity	1	0.8484	0.9180
Incomplete penetration	0.8908	1	0.9422
Crack	0.9862	1	0.9930
Lack of fusion	0.9836	1	0.9917
Slag inclusion	1	1	1

5. CONCLUSION

This paper proposes an intelligent recognition method for pipeline weld defects based on Gramian Angular Field and an improved Transformer-Faster R-CNN. The conclusions are as follows:

(1) The GAF transformation technique maps one-dimensional ultrasonic time-series signals into two-dimensional feature images. By preserving amplitude, frequency, and temporal correlation characteristics through polar coordinate encoding, it significantly enhances the capability of CNNs and Transformer models to extract features of typical defects such as porosity, cracks, and lack of fusion, addressing the limitations of one-dimensional signal feature representation.

(2) The improved Transformer-Faster R-CNN model, through a hybrid architecture combining ResNet-101 and a Transformer encoder, integrates both local feature perception and global dependency modeling advantages. Combined with a spatial attention mechanism and multi-scale feature pyramid, it effectively improves recognition accuracy and robustness for small-sized defects and those with complex morphologies.

(3) Compared with traditional CNN models and the standard Faster R-CNN method, the proposed approach achieves significantly superior performance across all evaluation metrics in five typical defect recognition tasks. Its end-to-end processing pipeline and GPU acceleration substantially enhance recognition efficiency. The method maintains high generalization performance even with limited annotated samples, providing a reliable technical pathway for the intelligent advancement of pipeline weld nondestructive testing.

6. REFERENCES

[1] Chen L ,Li J ,Ren J , et al.End-to-end leakage localization in buried hydrogen-blended natural gas pipelines based on multiscale spatiotemporal feature fusion[J].International Journal of Hydrogen Energy,2026,222154250-154250.

[2] Ochoa M S ,Sadovnychiy S .Application of an adaptive observer for the detection and location of pipeline

leak[J].Petroleum Science and Technology,2026,44(9):1396-1409.

[3] Wu J ,Yang Y ,Lin Z , et al.Weak ultrasonic guided wave signal recognition based on one-dimensional convolutional neural network denoising autoencoder and its application to small defect detection in pipelines[J].Measurement,2025,242(PE):116234-116234.

[4] Hu F ,Gou Y H ,Yang Z H , et al.Automatic PAUT crack detection and depth identification framework based on inspection robot and deep learning method[J].Journal of Infrastructure Intelligence and Resilience,2025,4(1):100113-100113.

[5] Peloquin E .Phased Array Ultrasound (PAUT) Vs. Phase Coherence Imaging (PCI) For Pipeline Inspection[J].Quality,2023,62(11):29-29.

[6] Cao W ,Sun X ,Liu Z , et al.The detection of PAUT pseudo defects in ultra-thick stainless-steel welds with a multimodal deep learning model[J].Measurement,2025,241115662-115662.

[7] P.L. S ,Gantala T ,Balasubramaniam K .Multi modal data fusion of PAUT with thermography assisted by Automatic Defect Recognition System (M-ADR) for NDE Applications[J].NDT and E International,2024,143103062-.

[8] Wang N ,Zhang K ,Wang X , et al.Pipeline Leakage Identification Based on Acoustic Sensors and EPSO-1D-CNN-Bi-LSTM Model[J].Sensors,2025,25(23):7355-7355.

[9] Zhang Z ,Wei M ,Wang Z .An Ultrasonic Echo Defect Recognition Method for Oil and Gas Pipelines Combining CNN-LSTM and Multi-Head Self-Attention Mechanism[J].Russian Journal of Nondestructive Testing,2025,61(6):633-653.

[10] Wei Y ,Yancai X ,Haikuo S , et al.Generalized weld bead region of interest localization and improved faster R-CNN for weld defect recognition[J].Measurement,2023,222.

[11] Baizhan X ,Hao L ,Shiguang S .Improved Faster R-CNN Based Surface Defect Detection Algorithm for Plates.[J].Computational intelligence and neuroscience,2022,20223248722-3248722.

[12] Xie J ,Yang J ,Fu K , et al.Quantitative assessment of pipeline defects utilizing a dual-stage deep learning framework: Integration of pretrained YOLO network and multi-input parallel convolution architectures on magnetic flux leakage data[J].Journal of Pipeline Science and Engineering,2026,6(1):100282-100282.

[13] Wen X ,Weng M ,Dong J , et al.YOLO–DE: An enhanced YOLOv11 model for the study of pipeline flow monitoring technology[J].Engineering Research Express,2026,8(3):035224-035224.

[14] Yuxin X ,Cun Z ,Taiji D , et al.Research on Pipeline Leak Signal Reconstruction and Multi-Aperture Classification Method Based on Multi-Task 1D U-Net[J].Advances in Computer, Signals and Systems,2025,9(4):.

[15] Pan G ,Zheng Y ,Guo S , et al.Automatic sewer pipe defect semantic segmentation based on improved U-Net[J].Automation in Construction,2020,119.

[16] X. Y S ,J. Y Z ,H. Z W , et al.A classification and location of surface defects method in hot rolled steel strips based on YOLOV7[J].Metalurgija,2023,62(2):240-242.

[17] Prabhakaran S ,Uthra A R ,Preetharoselyn J .Deep Learning-Based Model for Defect Detection and Localization on Photovoltaic Panels[J].COMPUTER SYSTEMS SCIENCE AND ENGINEERING,2023,44(3):2683-2700.

[18] Bae M Y ,He Y ,He Z , et al.Sub-class discovery in semiconductor defect detection through clustering and few-shot learning[J].Expert Systems With Applications,2026,311131353-131353.

[19] Wang Y ,Yan J ,Zhang Z , et al.A novel multi-sensor fusion method for diagnosing insulation defects in gas-insulated substations guided by adaptive-attention and contrastive-based few-shot learning[J].Applied Soft Computing,2026,190114600-114600.

[20] Lin Z ,Xu Y ,Wang X , et al.Transfer learning based on multimodal feature fusion: A study on intelligent detection of robotic MAG welding defects under cross-condition and small-sample scenarios[J].Journal of Manufacturing Processes,2026,162269-284.

[21] Lin D .A Transfer Learning-Based Model for Predicting Internal Control Deficiencies[J].Academic Journal of Management and Social Sciences,2025,12(3):77-79.

[22] Liu B ,Ding L ,He L , et al.Residual generative adversarial network-driven Data enhancement for magnetic flux leakage-based defect recognition in oil and gas pipelines[J].Engineering Applications of Artificial Intelligence,2025,162(PD):112720-112720.

[23] Qianqian Z ,Zuxiang S ,Shuai T , et al.Comparative Effectiveness of Data Augmentation Using Traditional Approaches versus StyleGANs in Automated Sewer Defect Detection[J].Journal of Water Resources Planning and Management,2023,149(9):.

[24] Wang J ,Sun A ,Wang C , et al.Improved deep learning-based detection of drainage pipeline defects using an enhanced YOLOv8 framework[J].Journal of Water Process Engineering,2025,78108605-108605.