remote sensing

MDPI

Article

GPR-Based Leakage Reconstruction of Shallow-Buried Water Supply Pipelines Using an Improved UNet++ Network

Qingqi Xu, Qinghua Liu * and Shan Ouyang

School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China; xuqingqi@mails.guet.edu.cn (Q.X.); hmoysh@guet.edu.cn (S.O.)
* Correspondence: qhliu@guet.edu.cn

Abstract

Ground-penetrating radar (GPR) plays a critical role in detecting underground targets, particularly locating and characterizing leaks in buried pipelines. However, the complex nature of GPR images related to pipeline leaks, combined with the limitations of existing neural network-based inversion methods, such as insufficient feature extraction and low inversion accuracy, poses significant challenges for effective leakage reconstruction. To address these challenges, this paper proposes an enhanced UNet++-based model: the Multi-Scale Directional Network PlusPlus (MSDNet++). The network employs an encoder–decoder architecture, in which the encoder incorporates multi-scale directional convolutions with coordinate attention to extract and compress features across different scales effectively. The decoder fuses multi-level features through dense skip connections and further enhances the representation of critical information via coordinate attention, enabling the accurate inversion of dielectric constant images. Experimental results on both simulated and real-world data demonstrate that MSDNet++ can accurately invert the location and extent of buried pipeline leaks from GPR B-scan images.

Keywords: ground penetrating radar; pipeline leakage reconstruction; multi-scale directional convolution; coordinate attention; dense skip connections

check for updates

Academic Editors: Xuan Feng, Haoqiu Zhou and Zejun Dong

Received: 21 April 2025
Revised: 22 June 2025
Accepted: 23 June 2025
Published: 25 June 2025

Citation: Xu, Q.; Liu, Q.; Ouyang, S. GPR-Based Leakage Reconstruction of Shallow-Buried Water Supply Pipelines Using an Improved UNet++ Network. *Remote Sens.* **2025**, *17*, 2174. https://doi.org/10.3390/rs17132174

Copyright: © 2025 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

In recent years, the rapid pace of urbanization has exacerbated the aging of water supply and drainage pipelines in cities, making it an increasingly severe issue. Many cities began constructing their water supply pipelines in the 1970s and 1980s. After decades of use, these pipelines have entered a phase of aging, leading to frequent ruptures and leaks. Leakage in drainage pipelines not only leads to significant water loss but also compromises the integrity of surrounding foundations, creates underground voids, and poses serious public safety risks [1].

Ground-penetrating radar has emerged as a rapidly developing, high-resolution, and efficient non-destructive testing technology. Its key characteristics—non-destructive nature, portability, continuity, high efficiency, broad applicability, and ease of use—have led to its widespread application across various fields, including road detection [2], tunnel defect identification [3], and military target detection [4]. Among these applications, the two-dimensional reconstruction of buried pipeline leaks using GPR is particularly valuable because it enables the early detection and localization of underground water losses, which helps prevent soil erosion, structural damage, and urban road collapses. In

urban infrastructure maintenance, the timely identification of leaks can significantly reduce repair costs and water loss, thus enhancing public safety and resource efficiency.

GPR inversion, in the broader context of geophysics, refers to the quantitative process of estimating the subsurface physical properties from the observed electromagnetic wave data [5]. Specifically for GPR, this involves recovering parameters such as the dielectric permittivity, electrical conductivity, and magnetic permeability of the subsurface materials, which govern the propagation and reflection of the radar waves. While magnetic permeability is often close to the free-space value for most common non-magnetic subsurface materials, permittivity and conductivity are the primary targets for inversion. By analyzing the inversion results, researchers can characterize subsurface structures or detect underground defects within the surveyed area. GPR inversion algorithms are generally categorized into traditional inversion methods and deep learning-based methods. The commonly used conventional inversion method is the Full Waveform Inversion (FWI) [6] method, which aims to estimate subsurface parameters by minimizing the difference between the observed and simulated waveforms. However, these conventional inversion methods are highly dependent on the initial model and tend to fall into local minima during the optimization process, which not only affects the convergence of the algorithm but also reduces the accuracy of the inversion.

In recent years, deep neural networks (DNNs) have demonstrated exceptional performance in various applications, including image classification [7], object detection [8], semantic segmentation [9], and image synthesis [10]. To enhance inversion efficiency and accuracy, many researchers have leveraged DNNs to explore the relationship between GPR B-scans and dielectric constant maps. In 2018, Alvarez et al. [11] employed three different network architectures—encoder–decoder, U-Net, and generative adversarial network (GAN)—for reconstructing the dielectric constant from ground-penetrating radar data, demonstrating the feasibility of using deep learning for GPR inversion. Xie et al. [12] introduced data compression and output units based on the U-Net framework and improved the network with instance normalization to invert both target locations and dielectric constants. Liu et al. [13] proposed a DNN-based GPR inversion network, GPRInvNet, which integrates a path-to-path encoder and a dielectric constant decoder. By incorporating neighborhood information-enhanced GPR A-scan features and a fully connected layer-based feature encoding method in the path-to-path encoder, the network addressed the spatial correspondence issue between B-scan data and dielectric constant maps, thereby facilitating tunnel lining reconstruction. Ji et al. [14] introduced an improved dielectric constant inversion network, PINet, which first compresses B-scan dimensions and extracts features using a dimension compressor. It then optimizes and aligns these features using a global feature encoder and finally upsamples them in the dielectric constant decoder to complete the reconstruction. Although PINet improved imaging results compared to GPRInvNet, its use of 128 fully connected layers as global feature encoders significantly increased model complexity and computational demands. Wu et al. [15] proposed a two-stage inversion network, MHInvNet, which uses multi-scale dilated convolution (MSDC) and hybrid attention gates (HAG). The network first denoises B-scans using MHInvNet1 and then jointly inputs the original data into MHInvNet2 for dielectric constant reconstruction. By combining the multi-scale feature extraction capability of MSDC with the key feature enhancement mechanism of HAG, the network significantly improved target localization accuracy and parameter inversion in complex scenarios. However, existing methods still face challenges in specific scenarios, especially in reconstructing buried pipeline leaks, which require inversion techniques capable of accurately capturing the size and precise location of the target, and there is little relevant research in pipeline leak reconstruction.

As a result, the reconstruction of pipeline leak locations and characteristics using deep learning has become a key research focus.

In response to these challenges, this paper proposes a network called MSDNet++, which is based on an enhanced version of UNet++, for the reconstruction of buried pipeline leaks. The network employs an encoder–decoder architecture to achieve dielectric constant image inversion through feature compression and reconstruction. Specifically, the network extracts multi-scale features in the encoder by designing multi-scale directional convolutional blocks (MSD), thereby enhancing its ability to perceive features across multiple scales and directions. The decoder then effectively fuses multi-level feature information through dense skip connections, enabling the network to learn both low-level detailed features and high-level abstract features simultaneously. This fusion helps mitigate information loss and enhances the network's ability to restore image details during upsampling. Following the fusion process, the network incorporates the coordinate attention (CA) mechanism and MSD modules to emphasize important spatial regions and channels, further optimizing the fused features and improving the representation of key features. Residual connections are incorporated to alleviate the vanishing gradient problem, accelerate training, and, in combination with multi-depth supervision outputs, strengthen multi-level feature learning. This approach significantly improves both the inversion accuracy and stability. The effectiveness and applicability of the proposed network are validated through data obtained from numerical simulations and real-world measurements.

2. Inversion Method Based on Improved UNet++

UNet++ [16] is an enhanced, fully convolutional neural network specifically designed to improve segmentation performance. Building upon the traditional U-Net architecture, UNet++ introduces dense skip connections through nested and redesigned decoder modules, progressively bridging the semantic gap between features at different levels. Specifically, UNet++ not only includes encoder and decoder components but also establishes dense connections between them via multiple intermediate nodes, thereby facilitating hierarchical feature fusion. In the encoder, the network progressively downsamples the input image to extract multi-scale features, while in the decoder, it refines and reconstructs these features through upsampling and skip connections. This nested structure enables a more effective fusion of multi-level features, alleviating the vanishing gradient problem and significantly improving the accuracy and robustness of both image segmentation and regression tasks.

Additionally, UNet++ incorporates a deep supervision mechanism that enhances both training efficiency and prediction performance. In the context of GPR image inversion, the primary task is to establish a non-linear mapping from GPR B-scan images to the corresponding dielectric constant images. Leveraging deep supervision in deep learning, this non-linear mapping is learned during training by minimizing the loss function via gradient descent and backpropagation, effectively overcoming many challenges associated with direct solution methods. Unlike the traditional U-Net structure, where the loss is calculated solely from the final decoder output, UNet++ adds supervision signals at multiple intermediate decoder outputs, allowing features from various network depths to contribute to the training process. This deep supervision mechanism accelerates convergence, enhances feature representation, and improves the inversion and segmentation performance for targets of different scales. Moreover, deep supervision mitigates the vanishing gradient problem, facilitates optimization, and ultimately improves accuracy and robustness for both regression and segmentation tasks.

2.1. Overall Framework and Improvements

MSDNet++, proposed in this paper, is an enhanced version of the UNet++ architecture, specifically optimized to address key challenges in GPR image inversion. Although traditional UNet++ enhances feature fusion via dense skip connections, it still has limitations in directional perception and localization accuracy when processing GPR images. To address these limitations, MSDNet++ integrates three key components: the multi-scale directional convolutional Block (MSD), the coordinate attention (CA) mechanism, and dense skip connections.

The MSD module enhances the model's capacity to capture anisotropic features by employing parallel horizontal and vertical convolutional kernels combined with a multi-scale feature extraction strategy. The CA mechanism applies 1D directional pooling and spatial weight generation to enhance the network's spatial sensitivity while maintaining computational efficiency. Dense skip connections optimize gradient propagation paths and enhance detail reconstruction by facilitating cross-level feature fusion and incorporating residual connections.

The overall architecture of MSDNet++ is illustrated in Figure 1. The cooperation of these three components enables MSDNet++ to significantly improve the localization and shape inversion accuracy of pipeline leak targets, providing a more powerful and robust solution for GPR image inversion.

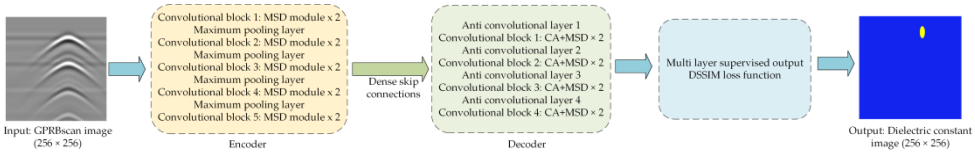


Figure 1. Network framework diagram.

2.2. Coordinate Attention Mechanism

The Coordinate attention mechanism decomposes channel attention into two one-dimensional encoding processes: one along the horizontal axis and the other along the vertical. This design facilitates the capture of long-range dependencies while preserving essential positional information. In contrast to the traditional 2D global pooling method used in channel attention, CA utilizes two parallel 1D global pooling operations to generate direction-specific feature maps, which are then encoded into spatial attention maps. This approach effectively enhances the perception of wide-area features in mobile neural networks while maintaining low computational overhead.

Figure 2 illustrates the structure of the CA mechanism. Given an input feature map $X \in R^{C\times H\times W}$, it first passes through a residual connection to preserve the original information. Then, global average pooling is applied along the height and width directions to obtain the horizontal feature $X^w_{avg} \in R^{C\times 1\times W}$ and vertical feature $X^h_{avg} \in R^{C\times H\times 1}$, respectively. The two directional features are concatenated and fused using a 1×1 convolution, followed by batch normalization and a non-linear activation function resulting in a fused representation $F \in R^{\frac{C}{r}\times(W+H)}$. A scaling factor r is introduced at this stage to reduce computational cost. F is then passed through two separate 1×1 convolutional layers with a Sigmoid activation function to generate the attention weights $A_w \in R^{C\times 1\times W}$ and $A_h \in R^{C\times H\times 1}$ for the height and width directions, respectively. Finally, these attention weights are applied to the corresponding spatial dimensions of the input feature map to produce the final weighted output:

$$Y = X \cdot A_h + X \cdot A_w \tag{1}$$

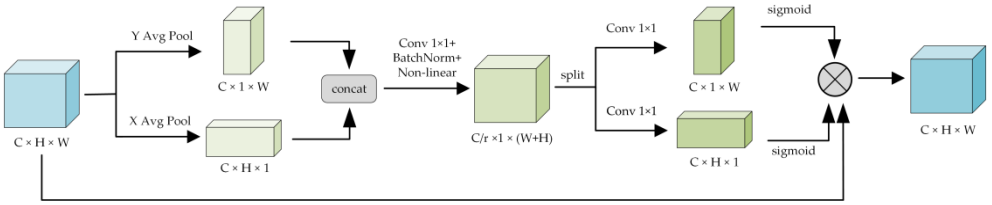


Figure 2. Coordinate attention mechanism.

This mechanism effectively captures long-range dependencies while retaining positional sensitivity, making it well suited for tasks requiring precise localization. Moreover, it incurs low computational overhead and can be seamlessly integrated into various network architectures.

2.3. Multi-Scale Directional Convolution

During the propagation of GPR signals, electromagnetic waves are scattered at the interfaces of differing dielectric properties in the subsurface medium, resulting in correlations among echoes across different traces. To enhance inversion accuracy, incorporating inter-trace correlation information is essential. Multi-scale convolution is a technique that uses parallel convolutional kernels of varying sizes to extract and fuse features across different scales. Traditional single-scale convolution can only capture features within a fixed receptive field and is limited in its ability to comprehensively represent multi-scale information. In contrast, multi-scale convolution combines global features from larger receptive fields with local features from smaller receptive fields, thereby enhancing the overall representational power and depth of feature extraction.

To address the complex feature extraction needs of GPR B-scan images, this paper proposes the multi-scale directional convolution module, illustrated in Figure 3. This module not only extracts features from multi-scale receptive fields but also integrates directional information, significantly enhancing the model's ability to identify anisotropic and complex structures. Specifically, the multi-scale directional convolution module includes the following key components:

(1) Multi-directional Convolution: The module incorporates horizontal and vertical convolutions to capture feature information from multiple directions within the image. These directional convolution layers use kernels of various scales, including four directional convolution layers with kernel sizes of 1×3, 3×1, 1×5, 5×1, ensuring the comprehensive extraction of features in each direction.

(2) Multi-scale Feature Extraction: The module combines small-scale kernels to capture fine-grained details and large-scale kernels to extract broader, global information, thereby enhancing the model's ability to capture a wide range of feature diversity.

(3) Deep Fusion: Each directional convolution is followed by batch normalization and an activation function (e.g., ReLU or LeakyReLU) to enhance feature expressiveness. The outputs from different directions are fused via channel concatenation, resulting in a high-dimensional representation that integrates both multi-scale and multi-directional information.

The concatenated feature map is then compressed through an additional convolution to generate a unified multi-scale, multi-directional representation. This design effectively captures information across diverse spatial resolutions and directions, thereby enriching the representational capacity of subsequent network layers.

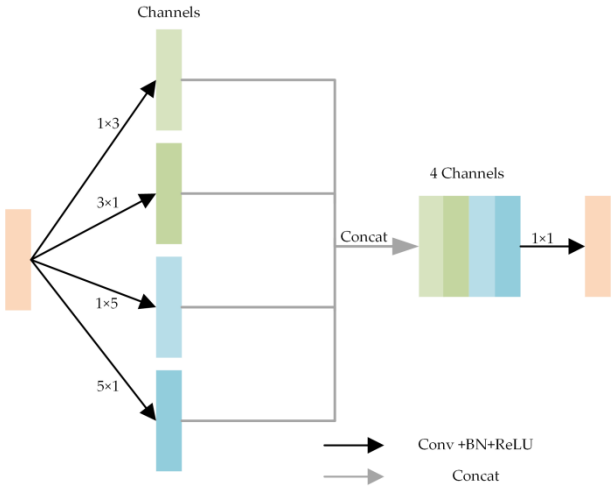


Figure 3. Multi-scale directional convolution module.

2.4. MSDNet++ Network Design

To establish a non-linear mapping between ground-penetrating radar (GPR) data and the corresponding dielectric constant images, this paper presents an enhanced GPR inversion network based on UNet++, termed MSDNet++. The overall network architecture is illustrated in Figure 4. MSDNet++ takes GPR B-scan images as input and uses the corresponding dielectric constant images as ground truth, performing inversion via feature compression and reconstruction.

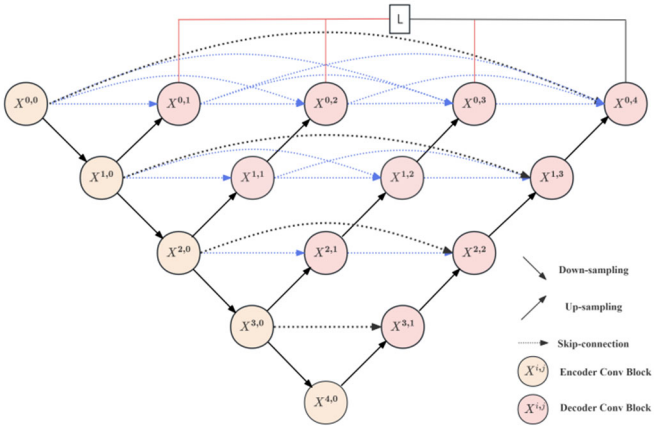


Figure 4. Network architecture of MSDNet++.

In Figure 4, the "L" marker represents the deep supervision mechanism, which is implemented by adding multiple output layers throughout the MSDNet++ network. The red solid lines indicate the connections between these output layers, and each output layer is supervised using a DSSIM loss function. These losses are combined into a total loss using a weighted average approach. This design enables the network to learn effectively at various levels by providing early-stage feedback during training. By supervising the network at different depths, the model learns features at multiple scales, thereby enhancing its ability to detect objects of varying sizes and shapes. This structure further enhances the fusion of low-level and high-level features, improving inversion accuracy in the presence of complex backgrounds and fine structural details.

The blue dashed lines in Figure 4 represent the dense skip connections, which are a feature unique to UNet++ compared to the standard UNet architecture which only has black solid skip connections. These dense skip connections allow each decoder layer to receive feature maps not only from its corresponding encoder layer but also from other

encoder stages. This design enables a more effective fusion of multi-level features by learning both low-level details and high-level semantic representations. As a result, it helps reduce information loss and improves detail recovery during upsampling, thereby enhancing inversion accuracy.

Specifically, the feature extraction network comprises five encoder convolutional blocks and a max-pooling layer with a stride of 2, as illustrated in Figure 5. Each encoder convolutional block includes two MSD modules and a residual connection. The dielectric constant decoder consists of four decoder convolutional blocks and an upsampling layer, which together reconstruct the output image. Each decoder convolutional block contains a CA attention mechanism and two MSD modules, as shown in Figure 6. Residual connections are incorporated in both the encoder and decoder to mitigate the vanishing gradient problem and accelerate training. Additionally, the number of convolutional channels in the encoder and decoder layers is set to [32, 64, 128, 256, 512] and [256, 128, 64, 32], respectively.

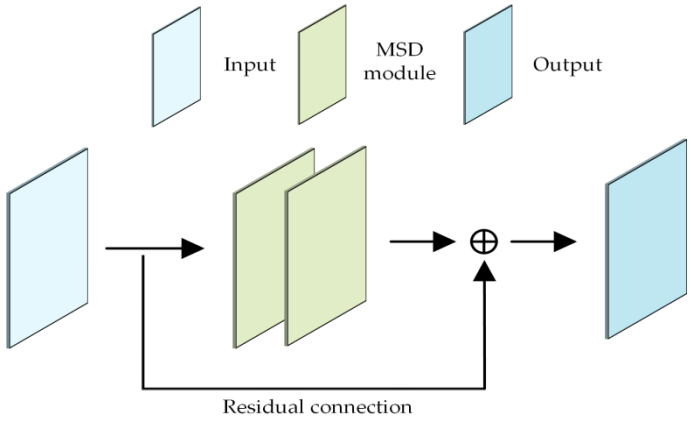


Figure 5. Encoder conv block.

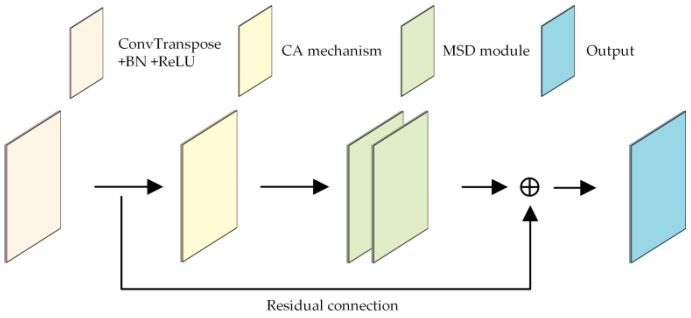


Figure 6. Decoder conv block.

2.5. Loss Function

In neural network models, the loss function plays a crucial role by quantifying the discrepancy between the model's predictions and the actual labels. Mean Squared Error (MSE) is one of the most commonly used loss functions. Minimizing the loss function via the backpropagation algorithm enables the effective adjustment of the network weights, thereby improving prediction accuracy. In this study, a variant of the Structural Similarity Index Measure (SSIM) [17] was adopted as the loss function for GPR image inversion, as defined below:

$$\mathrm{DSSIM}(x, y) = 1 - \mathrm{SSIM}(x, y) \tag{2}$$

SSIM (Structural Similarity Index Measure) evaluates the similarity between two images based on three key factors: luminance, contrast, and structure. Its specific formulation is as follows:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{3}$$

where μ_x and μ_y are the means of images x and y, respectively; σ_x^2 and σ_y^2 are the variances of images x and y, respectively; σ_{xy} is the covariance of images x and y; c_1 and c_2 are stabilizing constants.

3. Experimental Validation

3.1. Dataset and Evaluation Criteria

To evaluate the effectiveness of MSDNet++ for GPR dielectric constant inversion, this study constructed a dedicated inversion dataset for validation. GprMax, an open-source electromagnetic wave propagation simulator, was utilized to perform the forward modeling of GPR detection using the finite-difference time-domain (FDTD) method. Based on the characteristics of water supply pipeline leaks and their spatial distribution patterns, Shen Yupeng [18] used GeoStudio seepage field simulation software and found that water from a leak typically forms an approximately circular distribution around the pipeline, with the leak point at the center. Therefore, this study simulated buried pipeline leak scenarios with varying leakage radii and constructed a GPR inversion dataset using GprMax.

The simulation model had dimensions of 2 m × 1 m × 0.0025 m, with a uniform cell size of 0.0025 m in all directions and a perfectly matched layer (PML) thickness of 10 cells. A 25 ns time window was used with a Ricker wavelet centered at 400 MHz. The antenna transmit–receive distance was set to 0.1 m. In the common-offset mode, the GPR device moved horizontally from left to right with a measurement step of 0.02 m, resulting in 93 traces, each containing 4241 sampling points in the B-scan data. In the homogeneous half-space model, the background medium was dry sand, and the pipeline was made of PVC. The burial depth of the pipeline was randomly selected between 0.45 m and 0.75 m, with the leak point located directly beneath the pipeline. The radius of the leak area was randomly set between 0.05 m and 0.15 m, and the water content was set to either 30% or 40%. The electrical parameters of the predefined target media are listed in Table 1. In addition, in most natural media (e.g., air, soil, rock, etc.), for non-magnetic materials, the relative permeability is close to 1, i.e., these materials respond to the magnetic field similarly to a vacuum, and do not significantly enhance or diminish the magnetic field. Therefore, the magnetic permeability is set to 1 for all the materials simulated in this simulation.

Table 1. Dielectric constant and conductivity of the medium.

Media	Dielectric Constant	Conductivity (S/m)
Dry Sand	4	0.0068
PVC	3	0
Water	81	0
30% moisture content	17	0.0537
40% moisture content	25	0.1041

Based on the above settings, a total of 4000 data samples were generated, including 1000 non-leaking and 3000 leaking pipeline scenarios. All images were resized to 256 × 256 pixels and split into training and validation sets in a ratio of 8:2.

To further evaluate the performance of MSDNet++ under different time windows, an additional dataset with a 60 ns time window was generated while keeping all other

parameters constant concerning the 25 ns dataset. This dataset also included 4000 samples, with the same distribution of 1000 non-leaking and 3000 leaking scenarios. A comparison of the inversion results from the two datasets was conducted to more comprehensively evaluate the adaptability and robustness of MSDNet++ under varying detection conditions.

To evaluate performance, standard image quality metrics were employed, including the Structural Similarity Index Measure (SSIM), Mean Absolute Error (MAE), and Mean Squared Error (MSE). SSIM quantifies the structural similarity between predicted and ground-truth images, where higher values indicate better perceptual similarity. MAE and MSE measure pixel-wise errors, with lower values indicating more accurate predictions.

3.2. Experimental Environment and Parameter Settings

The hardware and software environment during the experiment is shown in Table 2.

Table 2. Experimental software and hardware environment parameters.

Name	Particulars
System environment	Windows10 64bit
Graphics chip	NVIDIA TITAN RTX
Memory capacity	24 GB
Deep learning frameworks	TensorFlow

The MSDNet++ architecture was implemented in TensorFlow, with the DSSIM loss function employed for training. The initial learning rate was set to 0.0005, and the Adam optimizer was adopted with a batch size of four. The network was trained for 150 epochs to ensure convergence and achieve accurate GPR image inversion.

Regarding the model validation and preservation strategy, this study adopted an optimal weight-saving mechanism based on validation loss. Specifically, the validation loss was continuously monitored in real time through a callback function, which automatically saved model weights corresponding to the minimal validation loss. This approach ensures that the final model selected for evaluation represents the optimal weights demonstrating peak performance on the validation set throughout the entire training process, a design principle established to guarantee the reliability and generalization capability of model performance.

3.3. Simulation Data Experiments

3.3.1. The 25 ns Simulation Data Experiment

We compared our model with other deep learning-based inversion networks, such as UNet [19], MRF-UNet [20], and UNet++. Among them, MRF-UNet adopts a U-shaped architecture containing an encoder and decoder, which preserves spatial information by hopping connections. The Multi-Receptive Field (MRF) module in the network is capable of extracting features at different spatial scales, capturing multi-scale information in the image, and enhancing the network's ability to perceive image details and overall structure. Table 3 shows the comparison of the computational metrics MSE, MAE, and SSIM or the inversion of the different networks, as well as the number of parameters for each network and the average time required to invert each B-scan image. It can be observed from the table that MSDNet++ shows superior inversion accuracy, but there was a sacrifice in the number of parameters, and it takes longer to process each image.

Table 3. Comparison of inversion performance at 25 ns.

Method	MSE	MAE	SSIM	Parameter/M	Time/ms
UNet	0.00225	0.00506	0.97628	29.64	344.81
MRF-UNet	0.00037	0.00206	0.99297	59.75	476.33
UNet++	0.00285	0.01499	0.96491	34.55	519.51
MSDNet++	0.00018	0.00288	0.99667	75.13	1043.01

Figure 7a–c show the variation curves of MSE, MAE, and SSIM during the dataset training of different networks. It can be observed that MSDNet++ shows superior performance on the whole. Figure 8 illustrates the inversion results produced by the different networks. The color bar quantitatively represents the dielectric permittivity constant range of 0–81 for network inversion. Figure 8a–c are B-scan images of ground-penetrating radar. After the leakage of PVC water pipes, the electromagnetic wave will be reflected when it comes into contact with the wall of the pipe, the internal body of water, and the boundary of the leakage area, so that a complex oscillating hyperbolic signal is presented in the B-scan image. These oscillating waves show a specific shape and distribution pattern in the B-scan image. When the pipeline is not leaking, as shown in Figure 8a, the electromagnetic wave reflects at the interface of the pipeline wall and the homogeneous medium such as the water inside the pipeline, forming a relatively regular and symmetrical hyperbolic signal. After a leakage occurs in the pipeline, as illustrated in Figure 8b,c, the waveforms become more complex and irregular, and in addition to the reflected signals from the pipeline itself, the leakage region triggers additional scattering and reflections, forming complex oscillatory hyperbolic signals. In order to make full use of these oscillatory wave features, the MSDNet++ network is designed, which utilizes the MSD module to extract oscillatory wave features at different scales and directions. The small-scale convolution kernel captures the fine structure of the oscillatory waveform, while the large-scale convolution kernel helps to identify the overall waveform pattern and trend. Parallel horizontal and vertical convolution operations enhance the model's ability to perceive oscillatory waves in different directions. On top of this, dense skip connections incorporate multiple layers of feature information, enabling the decoder to learn both low-level detailed features and high-level abstract features. This helps to reduce information loss during the upsampling process and improves the recovery of image details. The CA mechanism is introduced to further optimize the fused features by emphasizing key spatial regions and channels. This mechanism highlights the most representative parts of the oscillating waveforms, resulting in a more accurate reconstruction of the leakage features.

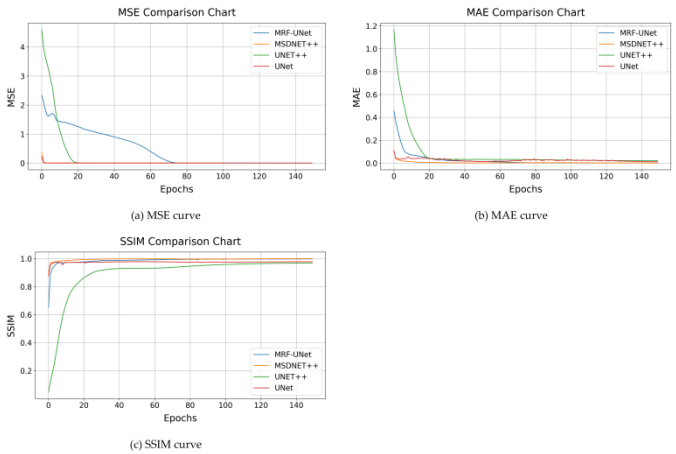


Figure 7. Comparison results of different network variation curves at 25 ns.

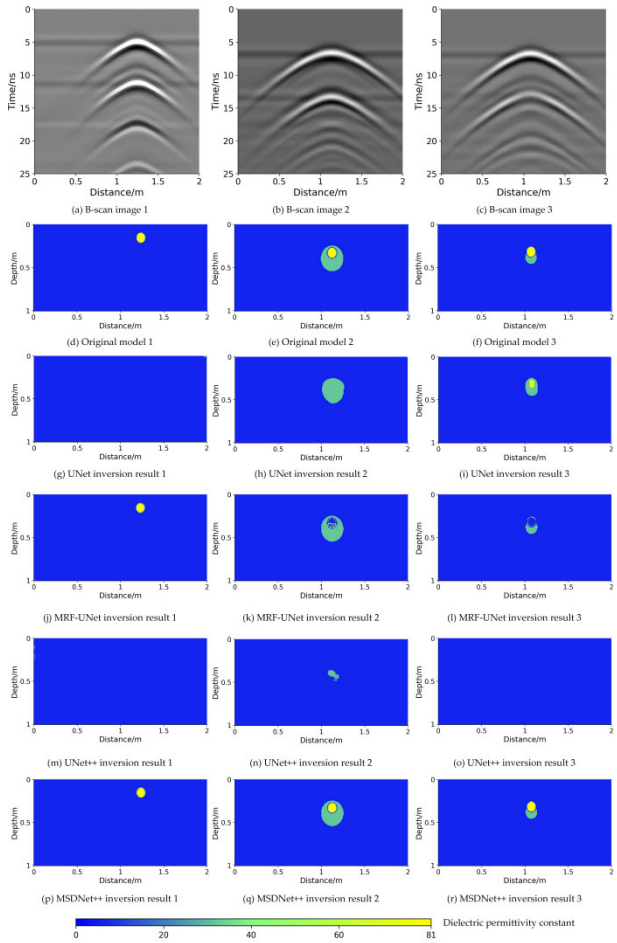


Figure 8. The inversion results of different networks at 25 ns.

Figure 8d–f are real dielectric constant images. As shown in Figure 8g–i, the pipeline contours generated by UNet are incomplete, and the leak areas are irregularly shaped, significantly deviating from the actual underground conditions. Figure 8j–l show that MRF-UNet, which incorporates a multi-receiver field module for multi-scale feature extraction, improves the inversion performance. However, the pipeline still exhibits unclear boundaries between the pipe wall and the internal water. Figure 8m–o indicate that the inversion results of UNet++ are blurry and suffer from ghosting artifacts. In contrast, Figure 8p–r show that MSDNet++, benefiting from the integration of multi-scale directional convolution, coordinate attention mechanisms, and dense skip connections, produces inversion results that closely match the ground truth model and outperform the other three networks.

3.3.2. The 25 ns Generalization Ability Test

In the previous section, we verified the superiority of MSDNet++ using simulation data, but in the actual detection process of ground-penetrating radar, the ground-penetrating radar data obtained is usually more complicated than the simulation data because of various noises and interferences. In order to verify the noise resistance of the model, we added Gaussian white noise with mean value 0 and standard deviation 1 to the ground-penetrating radar B-scan to simulate the background noise interference in the real detection environment, and at the same time, we used the peak signal-to-noise ratio, which is widely used in the field of denoising, as one of the indexes for evaluating the anti-noise performance of the network. The reconstruction test with background noise is obtained by fine-tuning 10 epochs on the trained model.

The results of the reconstruction test with background noise are shown in Table 4. It can be seen that compared with the other three networks, MSDNet++ still achieves better reconstruction performance.

Table 4. Comparison of noise data inversion index at 25 ns.

Method	MSE	MAE	PSNR (db)	Parameter/M	Time/ms
UNet	0.00435	0.02900	24.6693	29.64	399.24
MRF-UNet	0.00192	0.00897	27.5979	59.75	558.11
UNet++	0.05023	0.22237	23.8661	34.55	542.95
MSDNet++	0.00148	0.01106	29.6295	75.13	1082.50

Figure 9 illustrates the inversion results of the different networks for noisy data. Figure 9a–c show the B-scan images of the ground-penetrating radar with added Gaussian white noise, and Figure 9d–f are real dielectric constant images. As shown in Figure 9g–i, the UNet inversion of the target is not clear and is significantly inconsistent with the actual subsurface conditions. Figure 9j–l show that MRF-UNet is able to invert the pipeline target, but the leakage contour area is blurred. Figure 9m–o show that the UNet++ inversion results have artifacts and the inversion does not produce results. In contrast, Figure 9p–r show that the MSDNet++ inversion results are more consistent with the real model and superior to the other three networks, although the contours are somewhat blurred.

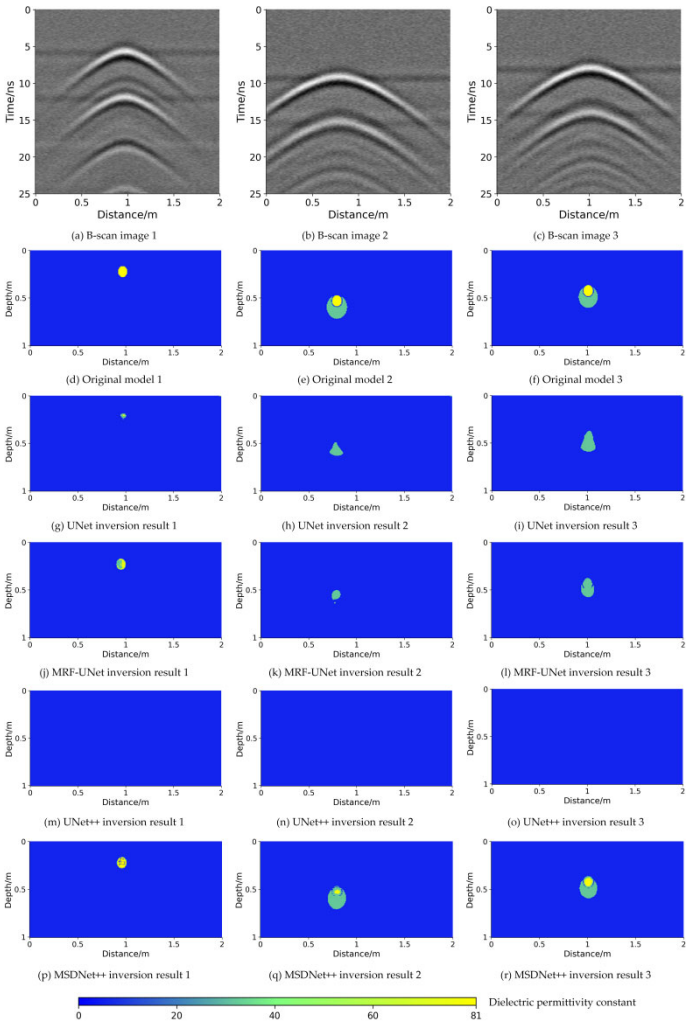


Figure 9. Noisy data inversion results for different networks at 25 ns.

3.3.3. The 60 ns Simulation Data Experiment

Similarly, a comparison was conducted with other inversion networks, including UNet, MRF-UNet, and UNet++, as shown in Table 5. The results indicate that MSDNet++ outperforms these networks in terms of overall performance.

Table 5. Comparison of inversion performance at 60 ns.

Method	MSE	MAE	SSIM	Parameter/M	Time/ms
UNet	0.00619	0.04033	0.96179	29.64	344.81
MRF-UNet	0.00148	0.02403	0.98175	59.75	476.33
UNet++	0.00241	0.00601	0.97471	34.55	519.51
MSDNet++	0.00058	0.00515	0.98838	75.13	1043.01

Figure 10a–c show the variation curves of MSE, MAE, and SSIM during the dataset training of different networks. It can be observed that MSDNet++ shows superior performance on the whole. Figure 11 displays the inversion results produced by each network. Figure 10a–c are B-scan images of ground-penetrating radar, and Figure 11d–f are real dielectric constant images. As shown in Figure 11g–i, UNet failed to generate effective inversion results. Figure 11j–l demonstrate that while MRF-UNet improved the inversion quality, the pipeline target remained blurry and the water–pipe interface was indistinct. Figure 11m–o indicate that UNet++ also produced blurry results, with poorly defined boundaries between the pipe wall and the internal water. In contrast, Figure 11p–r show that although the inversion results from MSDNet++ exhibited some ghosting artifacts, they were significantly more accurate and superior to those of the other three networks.

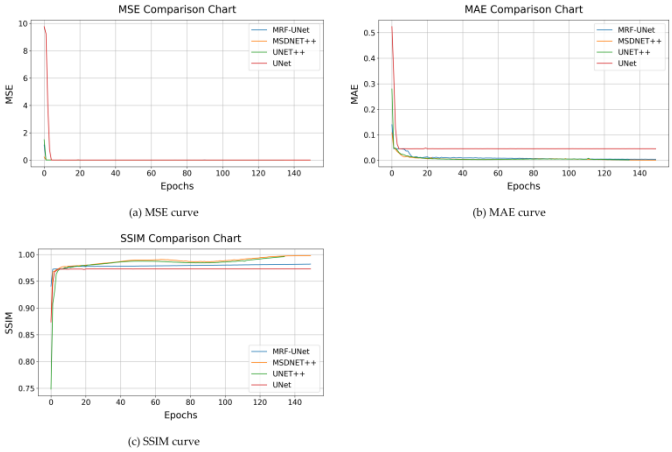


Figure 10. Comparison results of different network variation curves at 60 ns.

3.3.4. The 60 ns Generalization Ability Test

Similarly, to verify the model noise immunity at 60 ns, we also added Gaussian white noise with mean 0 and standard deviation 1 to the ground-penetrating radar B-scans. And the peak-to-noise ratio was used as a metric to assess the network's noise immunity performance. On the trained model micro-tuning for 10 epochs, we obtained the reconstruction test results with background noise. The results of the reconstruction test with background noise are shown in Table 6. It can be seen that compared with the other three networks, MSDNet++ still achieves better reconstruction performance.

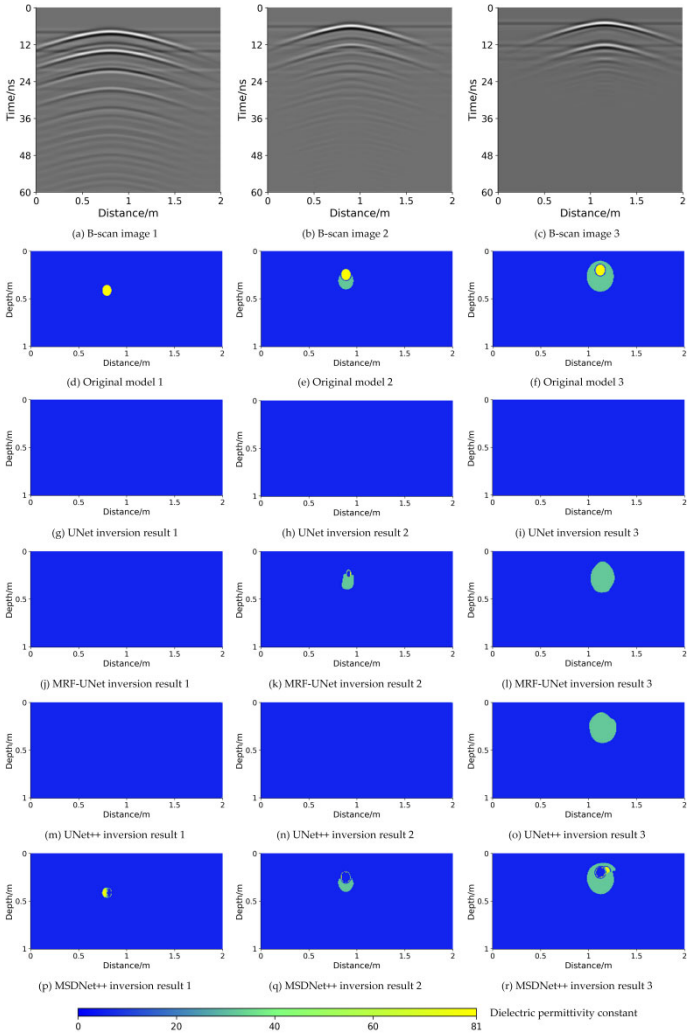


Figure 11. The inversion results of different networks at 60 ns.

Table 6. Comparison of noise data inversion index at 60 ns.

Method	MSE	MAE	PSNR (db)	Parameter/M	Time/ms
UNet	0.04086	0.20001	13.8894	29.64	399.24
MRF-UNet	0.02632	0.15819	15.9096	59.75	558.11
UNet++	0.02499	0.15415	16.0504	34.55	542.95
MSDNet++	0.00659	0.07954	21.8130	75.13	1082.50

Figure 12 illustrates the inversion results of different networks for noisy data. Figure 12a–c show the B-scan images of the ground-penetrating radar with added Gaussian white noise, and Figure 12d–f are real dielectric constant images. As shown in Figure 12g–i, the UNet inversion does not produce a target. Figure 12j–l show that the MRF-UNet inversion of the target has blurred contours, which are different from the real model. Figure 12m–o show that UNet++ is able to invert the pipeline target, but also suffers from blurred contours. In contrast, Figure 12p–r show that MSDNet++ inverts the inner water body of the pipeline indistinctly, but is more consistent with the real model, outperforming the other three networks.

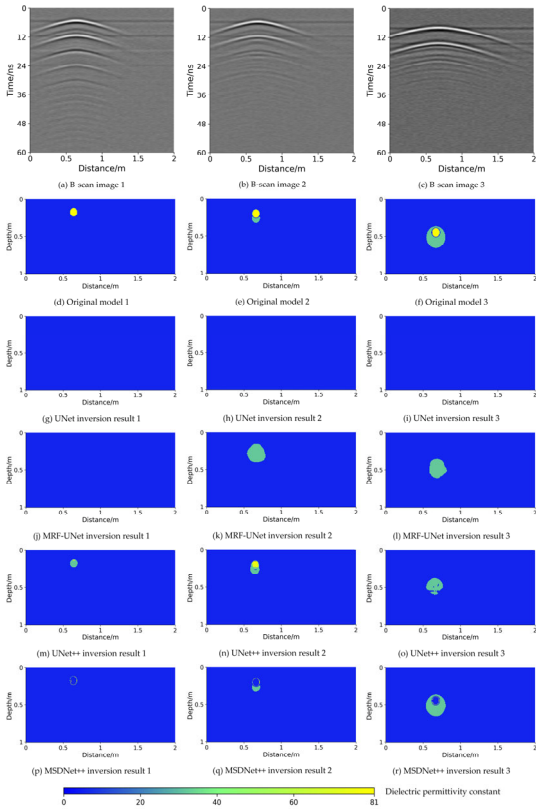


Figure 12. Noisy data inversion results for different networks at 60 ns.

3.4. Ablation Study

3.4.1. The 25 ns Ablation Experiment

To evaluate the effectiveness of the proposed MSD convolution module, the introduced CA mechanism, and the dense skip connections, ablation experiments were conducted. The inversion results on the test set were quantitatively analyzed using SSIM, MSE, and MAE as evaluation metrics, as presented in Table 7. The number of parameters per network and the average time required to invert each B-scan image can also be observed in Table 7.

Table 7. Test results under different modules at 25 ns.

Network Structure	MSE	MAE	SSIM	Parameter/M	Time/ms
① UNet++ only adds the CA mechanism	0.00112	0.00347	0.98408	36.60	647.27
② UNet++ only adds the MSD module	0.00122	0.00376	0.98538	74.74	913.06
③ UNet adds CA and MSD	0.00114	0.00455	0.97638	65.65	703.32
④ MSDNet++	0.00105	0.00341	0.99334	75.13	1043.01

The ablation strategy was based on the UNet++ architecture. First, only the CA attention mechanism was added, resulting in Network ①. Then, only the MSD module was added to form Network ②. These configurations were used to evaluate the individual contributions of each module. Subsequently, based on the UNet++ framework, both the CA mechanism and the MSD module were included, while the dense skip connections were removed to construct Network ③, aiming to assess the impact of dense skip connections. The comparative results of the ablation experiments are illustrated in Figure 13. Figure 13a–c are B-scan images of ground-penetrating radar, and Figure 13d–f are real dielectric constant images. As shown in the figure, Networks ① and ②—each incorporating only one of the modules—produced inversion results with blurred boundaries between the pipe wall and the internal water region. Similarly, Network ③, which lacked dense

skip connections, exhibited comparable deficiencies. In contrast, the complete MSDNet++ model, which integrates all three components, achieved the most accurate inversion results, with well-defined contours and precise boundaries for the inverted targets.

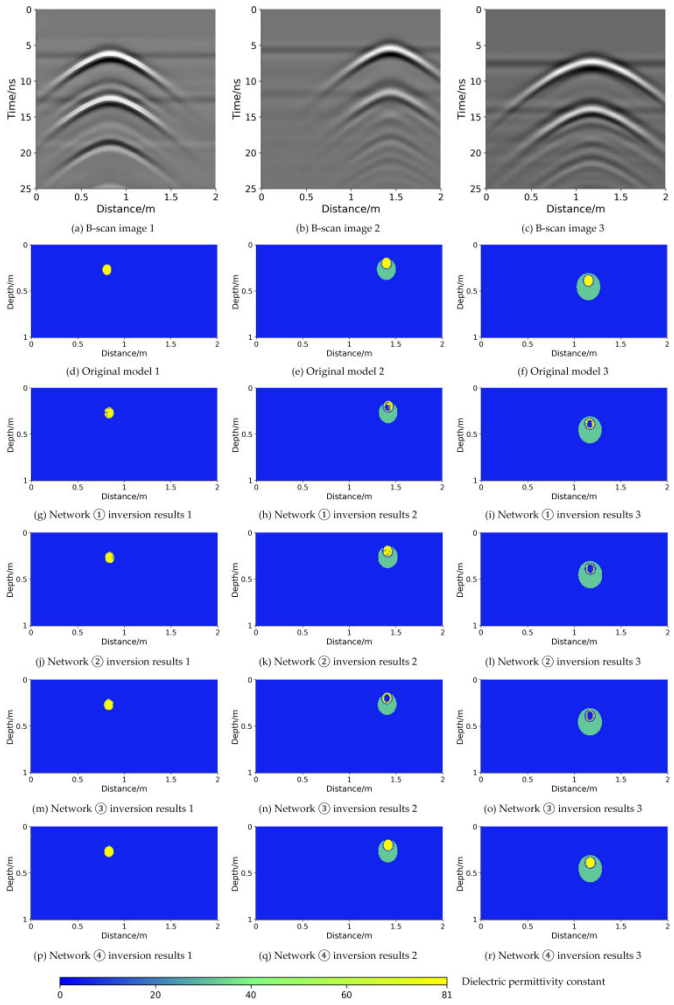


Figure 13. Comparison of ablation results at 25 ns.

3.4.2. The 60 ns Ablation Experiment

The ablation experiments were conducted using the same settings as the 25 ns time window experiments. SSIM, MSE, and MAE were again employed as evaluation metrics to quantitatively assess the inversion results on the test set, as presented in Table 8.

Table 8. Test results under different modules at 60 ns.

Network Structure	MSE	MAE	SSIM	Parameter/M	Time/ms
① UNet++ only adds the CA mechanism	0.00121	0.00802	0.98210	36.60	647.27
② UNet++ only adds the MSD module	0.00141	0.01647	0.98516	74.74	913.06
③ UNet adds CA and MSD	0.01854	0.13457	0.96331	65.65	703.32
④ MSDNet++	0.00111	0.00596	0.99121	75.13	1043.01

The comparative results of the ablation experiments are illustrated in Figure 14. Figure 14a–c are B-scan images of ground-penetrating radar, and Figure 14d–f are real dielectric constant images. As shown in the figure, Networks ① and ②, which incorporated only the CA attention mechanism or the MSD module, respectively, exhibited ghosting artifacts in the inversion results, with unclear boundaries between the pipe wall

and the internal water region. Similarly, Network ③, in which the dense skip connections were removed, demonstrated comparable deficiencies. Although the MSDNet++ model—which integrates all three components—also exhibited some blurring at the boundaries between the pipe wall and the internal water region, it still achieved the best overall performance, producing inversion results with more apparent target contours. In conclusion, MSDNet++, which incorporates the CA mechanism, the MSD module, and dense skip connections, demonstrated superior performance compared to the other variants.

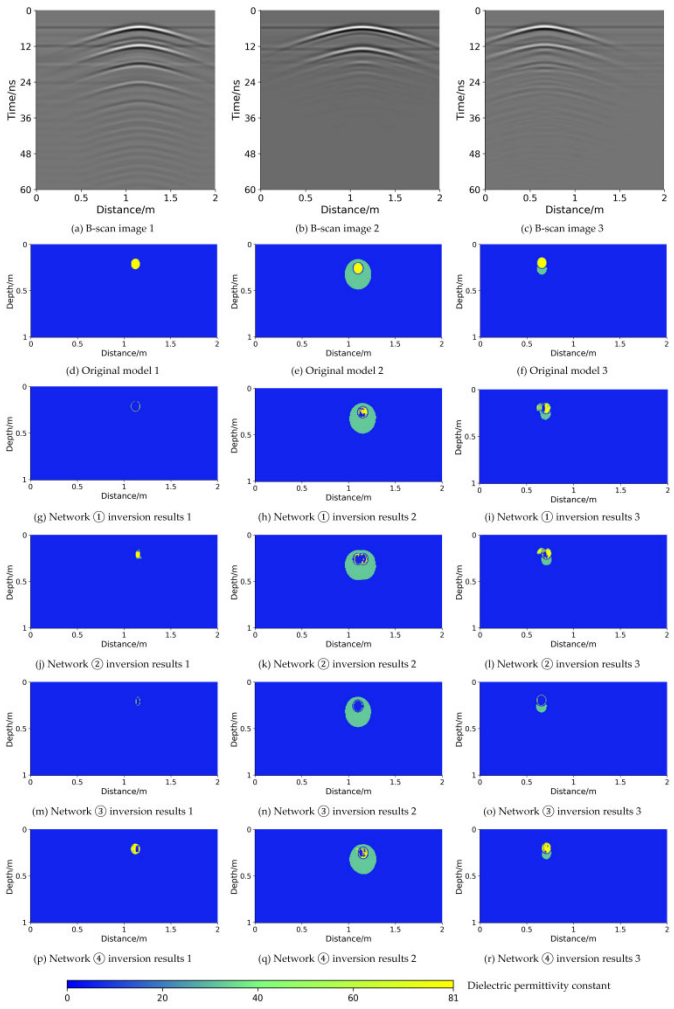


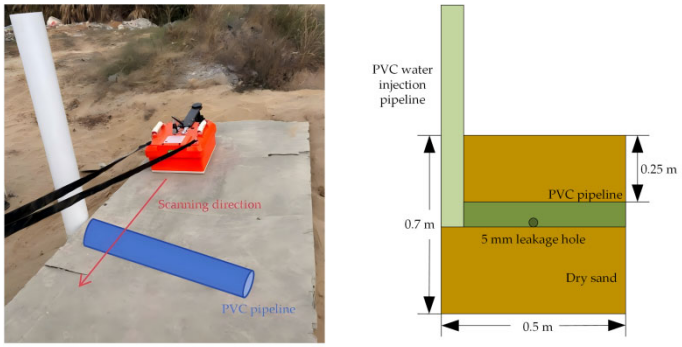
Figure 14. Comparison of ablation results at 60 ns.

By comparing the simulation experiments with 25 ns and 60 ns time windows, it was observed that the inversion results obtained from the 25 ns data were generally more accurate. This is primarily because the shorter 25 ns time window more effectively captures reflection signals from shallow layers, thereby highlighting the characteristics of pipeline leaks more clearly. And the 60 ns time window signals show multi-layer oscillations or multiple noise artifacts, which are more likely to occur in longer time windows, significantly affecting the clarity of the shallow reflection signals. Furthermore, the shorter time window reduces signal attenuation during propagation, preserving signal strength and clarity, which facilitates the extraction of useful features by the inversion network.

3.5. MSDNet++ Testing with Real Measurement Data

To validate the inversion results of MSDNet++ on real measurement data, an experimental setup was established, as shown in Figure 15. The overall experimental model consisted of a sandpit with dimensions of 1.8 m (length) × 0.5 m (width) × 0.7 m (height),

as shown in Figure 15a. Inside the model, a PVC pipe with a diameter of 110 mm, wall thickness of 4.2 mm, and length of 0.5 m was buried at a depth of 0.25 m and positioned horizontally at 0.9 m. A leakage hole with a diameter of 5 mm was drilled at the midpoint of the pipe's bottom. One end of the pipe was connected to a one m long water pipe, as shown in the schematic diagram in Figure 15b. The small leakage holes were first plugged and the pipe was filled with water, and measurements were taken to obtain the results when the pipe was not leaking. Subsequently, the leakage hole was opened and a leakage experiment was conducted at a water pressure of 0.01 MPa, with a continuous filling of the pipe, which leaked 24 L of water after 45 min, followed by the excavation of sand for leakage assessment. The leaked water formed a circular pattern with a radius of approximately 10 cm, centered around the point of leakage. The experimental setting parameters are summarized in Table 9.



(a) Site diagram of test layout (b) Schematic diagram of the model

Figure 15. Experimental scenario.

Table 9. Summary of experimental setup parameters.

Name	Particulars
Dimensions	1.8 m × 0.5 m × 0.7 m
Pipeline diameter	110 mm
Depth of burial	0.25 m
Diameter of the leakage hole	5 mm
Capacity of leakage water	24 L
Leak time	45 min
Leakage radius	10 cm

In this experiment, a YL-GPR ground-penetrating radar from Y-Line Company, with a center frequency of 400 MHz, was used. During detection, the radar time window was set to 60 ns, the working mode was distance mode, and the sampling interval was 0.02 m. The collected data is preprocessed to include direct wave removal, filtering, gain, and finally cropping so that the processed data matches the size of the dataset data. Because the time window is long and the scanning distance is short, the hyperbolic curvature feature is not obvious when cropped to 256 × 256 size. Finally, the data was input into MSDNet++ for inversion testing, and the inversion results are shown in Figure 16.

Figure 16a,b are B-scan images of real scans from ground-penetrating radar, and Figure 16c,d show the original dielectric constant model images. As seen in Figure 16e,f, MSDNet++ successfully performed the inversion of real measurement data for both shallowly buried pipelines without leakage and those with leakage in the sand. However, deviations in the shape and size of the inverted targets were observed. This may be due to the presence of noise in the actual environment, which caused the B-scan image data

to be less refined and significantly different from the simulated data, thereby affecting the network's prediction accuracy.

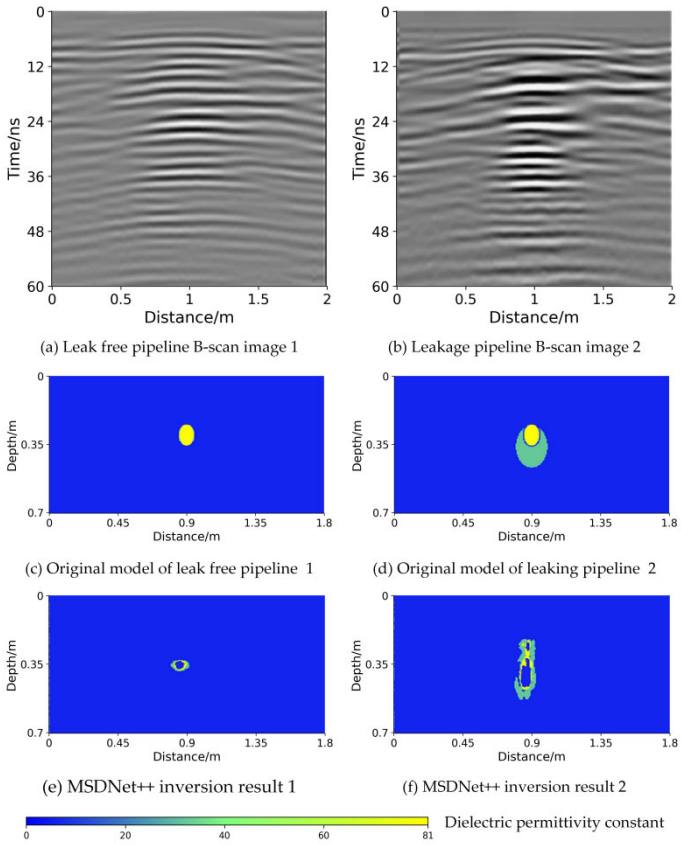


Figure 16. Results of inversion of measured data.

4. Discussion

This study proposes an improved UNet++-based neural network, MSDNet++, for reconstructing shallowly buried water supply pipeline leaks using ground-penetrating radar (GPR) data. By integrating multi-scale directional convolution (MSD), coordinate attention (CA) mechanisms, and dense skip connections, the network effectively addresses the challenges of feature extraction and spatial accuracy in GPR image inversion. Experimental validations were conducted on both simulated and real-world data. The key findings and implications are discussed as follows:

(1) The MSD Module: The MSD module combines multi-scale and directional convolutional kernels to capture anisotropic features in GPR images. Small-scale kernels extract fine-grained details of pipeline leaks, while large-scale kernels capture global structural information. Parallel horizontal and vertical convolutions enhance the model's ability to recognize features along different directions, which is crucial for accurately reconstructing irregular leak geometries. This design significantly improves the network's performance compared to traditional single-scale convolutions, as demonstrated by the clear leak boundaries in the inversion results (Figure 8p–r).

(2) The CA Mechanism: The CA mechanism decomposes channel attention into 1D feature encoding processes along the horizontal and vertical directions, preserving positional sensitivity while capturing long-range dependencies. This approach addresses the blurring and ghosting issues observed in UNet++ (Figure 8m–o) by emphasizing key spatial regions in the feature maps. The lightweight design of the CA mechanism

ensures computational efficiency, making it suitable for integration into complex network architectures.

(3) Dense Skip Connections: Building upon the UNet++ framework, MSDNet++ employs dense skip connections to fuse multi-level feature information. This design enables the decoder to simultaneously learn low-level, detailed features and high-level, abstract features, thereby minimizing information loss during upsampling. The residual connections further alleviate the vanishing gradient problem and accelerate the training process. The ablation experiments (Tables 7 and 8) confirm that the combination of these three components achieves the best inversion performance.

(4) Experimental Validation: The proposed network outperforms existing methods (UNet, MRF-UNet, and UNet++) in both 25 ns and 60 ns simulation experiments, achieving higher SSIM (0.99667) and lower MSE (0.00018) values (Tables 3 and 5). For real-world data, MSDNet++ successfully reconstructs pipeline leaks, though deviations in shape and size occur due to environmental noise (Figure 16). This highlights the need for noise-resistant training strategies or preprocessing techniques to bridge the gap between simulated and measured data performance.

(5) Limitations and Future Work: While MSDNet++ demonstrates promising results, its performance in harsh underground environments with strong interference signals requires further improvement. Future research should focus on the following:

① Enhancing noise robustness by incorporating advanced denoising modules or adversarial training.

② Expanding the training dataset with diverse leakage scenarios and soil conditions to improve generalization.

③ Optimizing the network for real-time deployment on portable GPR devices.

In conclusion, MSDNet++ provides an effective solution for GPR-based pipeline leak inversion by integrating multi-scale directional perception, attention mechanisms, and hierarchical feature fusion. Future efforts should enhance its adaptability to complex real-world conditions and extend its applications to other subsurface imaging tasks.

5. Conclusions

This paper presents an improved U-Net++ neural network architecture, MSDNet++, designed to efficiently convert ground-penetrating radar (GPR) B-scan images into dielectric constant distribution images of underground media. By integrating multi-scale directional convolution and coordinate attention mechanisms, the network enhances the extraction and representation of complex features in GPR images, enabling the more accurate inversion of the location and size of shallowly buried pipeline leaks. MSDNet++ combines multi-scale directional convolutional (MSD) blocks and coordinate attention mechanisms, performing feature extraction, compression, and reconstruction through an encoder–decoder architecture. The encoder utilizes MSD modules to extract multi-scale features, thereby enhancing the recognition of targets at various scales and directions. The decoder fuses multi-level feature information through dense skip connections, reducing information loss, and further optimizes feature representation with the coordinate attention mechanism to highlight key areas. Training and testing were conducted on a simulation dataset generated by GprMax, and the results demonstrate that MSDNet++ achieves higher inversion accuracy and robustness compared to traditional networks such as UNet, MRF-UNet, and UNet++. Testing with real measurement data further confirms the effectiveness of MSDNet++ in real-world GPR scenarios. However, due to environmental noise interference in real measurement data, discrepancies were observed between the inversion results and actual conditions. Future work will focus on developing more advanced GPR physical models, expanding the network's training dataset, and further optimizing the

MSDNet++ deep learning inversion method. By enhancing feature extraction and boosting the network's generalization capability, MSDNet++ will become even more applicable to real-world GPR signal inversion tasks, providing more accurate and reliable solutions for urban underground pipeline leak detection. However, the presence of ambient noise interference in the actual measurement data led to discrepancies between the inversion results and the actual situation. Future work will focus on developing more advanced physical models for GPR, expanding the size of the network training dataset, dealing with deeper buried pipeline leaks, and further investigating the inversion of underground pipeline targets and the inversion of specific values of dielectric constant images at the time of leakage for a more comprehensive and reliable 2D reconstruction of pipeline leaks. Meanwhile, the MSDNet++ deep learning inversion method will be further optimized to make MSDNet++ more suitable for practical GPR signal inversion tasks by enhancing feature extraction and improving the generalization ability of the network, so as to provide a more accurate and reliable solution for the reconstruction of urban underground pipeline leakage.

Author Contributions: Conceptualization, Q.X. and Q.L.; methodology, Q.X. and Q.L.; validation, Q.X. and Q.L.; writing—original draft preparation, Q.X.; writing—review and editing, Q.L. and S.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under Grant 62361015 and the Guangxi Natural Science Foundation under Grant 2025GXNSFAA069204.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, S.C. Analysis and Study on Subsidence Mechanisms of Road Caused by Leakage of Urban Underground Pipeline. Master's Thesis, Zhengzhou University, Zhengzhou, China, 2017. (In Chinese)
2. Zhang, J.; Hu, Q.W.; Zhou, Y.M.; Zhao, P.; Duan, X. A multi-level robust positioning method for three-dimensional ground penetrating radar (3D GPR) road underground imaging in dense urban areas. *Remote Sens.* **2024**, *16*, 1559. [CrossRef]
3. Qin, H.; Xie, X.Y.; Tang, Y.; Wang, Z. Experimental study on GPR detection of voids inside and behind tunnel linings. *J. Environ. Eng. Geophys.* **2020**, *25*, 65–74. [CrossRef]
4. Daniels, D.J. A review of GPR for landmine detection. *Sens. Imaging* **2006**, *7*, 90–123. [CrossRef]
5. Abubakar, A.; Habashy, T.M.; Druskin, V.; Alumbaugh, D.; Knizhnerman, L. 2.5D forward and inversion modeling for interpreting low-frequency electromagnetic measurements. *Geophysics* **2008**, *73*, F165–F177. [CrossRef]
6. Lei, L.L.; Liu, S.X.; Fu, L.; Wu, J.J. Reverse time migration applied to GPR data based on full waveform inversion. *Chin. J. Geophys.* **2015**, *58*, 3346–3355. (In Chinese)
7. LeCun, Y.; Bengio, Y.; Hinton, G.E. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
8. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef] [PubMed]
9. Miao, X.; Wang, J.; Wang, Z.; Sui, Q.; Gao, Y.; Jiang, P. Automatic recognition of highway tunnel defects based on an improved U-Net model. *IEEE Sens. J.* **2019**, *19*, 11413–11423. [CrossRef]
10. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, A. Generative adversarial nets. *Proc. Adv. Neural Inf. Process. Syst.* **2014**, *3*, 2672–2680.
11. Alvarez, J.K.; Kodagoda, S. Application of deep learning image-to-image transformation networks to GPR radargrams for sub-surface imaging in infrastructure monitoring. In Proceedings of the 13th IEEE Conference on Industrial Electronics and Applications, Wuhan, China, 31 May–2 June 2018; pp. 611–616.
12. Xie, L.; Zhao, Q.; Ma, C.; Liao, B.; Huo, J. Ü-net: Deep-learning schemes for ground penetrating radar data inversion. *J. Environ. Eng. Geophys.* **2020**, *25*, 287–292. [CrossRef]
13. Liu, B.; Ren, Y.; Liu, H.; Xu, H.; Wang, Z.; Cohn, A.G.; Jiang, P. GPRInvNet: Deep learning-based ground-penetrating radar data inversion for tunnel linings. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 8305–8325. [CrossRef]
14. Ji, Y.; Zhang, F.; Wang, J.; Wang, Z.; Jiang, P.; Liu, H.; Sui, Q. Deep neural network-based permittivity inversions for ground penetrating radar data. *IEEE Sens. J.* **2021**, *21*, 8172–8183. [CrossRef]

15. Wu, M.; Liu, Q.; Ouyang, S. Two-stage GPR image inversion method based on multi-scale dilated convolution and hybrid attention gate. *Remote Sens.* **2025**, *17*, 322. [CrossRef]
16. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imaging* **2020**, *39*, 1856–1867. [CrossRef] [PubMed]
17. Bakurov, I.; Buzzelli, M.; Schettini, R.; Castelli, M.; Vanneschi, L. Structural similarity index (SSIM) revisited: A data-driven approach. *Expert Syst. Appl.* **2022**, *189*, 116087. [CrossRef]
18. Shen, Y.P.; Dong, S.H.; Wang, Q.; Huang, L.Y. Seepage model analysis of urban water pipeline leakage degree and forward modeling of GPR signals. In Proceedings of the 2016 National Conference on Engineering Geology, Chengdu, China, 15–17 October 2016. (In Chinese)
19. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
20. Dai, Q.; Lee, Y.H.; Sun, H.H.; Ow, G.; Yusof, M.L.M.; Yucel, A.C. DMRF-UNet: A two-stage deep learning scheme for GPR data inversion under heterogeneous soil conditions. *IEEE Trans. Antennas Propag.* **2022**, *70*, 6313–6328. [CrossRef]

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.