Available online at www.sciencedirect.com

ScienceDirect

Underground Space 26 (2026) 458–478

www.keaipublishing.com/undsp

Research Paper

GPR-HIDiff: A diffusion-based model for horizontal interference suppression in urban underground detection radar profiles

Xiaosong Tang [a,*], Feng Yang [a], Xu Qiao [a], Jialin Liu [a], Haitao Zuo [a], Liang Gao [a], Jianshe Zhao [a], Suping Peng [b]

[a] School of Artificial Intelligence, China University of Mining and Technology, Beijing 100083, China
[b] State Key Laboratory of Coal Resources and Safe Mining, Beijing 100083, China

Received 9 July 2025; received in revised form 16 August 2025; accepted 25 August 2025
Available online 13 December 2025

Abstract

Automated subsurface utility detection systems in construction rely heavily on the quality of ground-penetrating radar (GPR) profiles, which are often degraded by high-amplitude horizontal interference. Existing low-rank decomposition methods lack the intelligence and flexibility required for multi-site data processing and involve labor-intensive parameter tuning, impeding their integration into intelligent construction workflows. To address these challenges, this paper proposes a horizontal interference suppression algorithm based on a diffusion model, termed GPR-HIDiff. The proposed model replaces conventional sequential convolutional operators with ResBlocks throughout the encoder, intermediate layer, and decoder of the UNet architecture, enhancing training stability. Lightweight agent attention modules are embedded between ResBlocks at each level to improve global information modeling capability. A spatial attention mechanism is deployed between the encoder and decoder to achieve adaptive spatial feature optimization. Furthermore, the forward diffusion phase adopts a cosθ schedule-based strategy to ensure a smooth temporal variation of noise variance. A standardized dataset comprising real-world measured samples and finite difference time domain simulation samples of urban road models has also been constructed. The effectiveness of the hybrid dataset, the introduced modules, the robustness analysis, and the cosθ schedule is validated through training with single/mixed datasets, ablation studies, evaluation of metric variations before and after the introduction of different noise levels, and comparative experiments with constant, linear, and cosθ schedules. Experimental results demonstrate that GPR-HIDiff significantly outperforms both traditional methods and state-of-the-art deep learning models on both simulated and real-world test samples. It effectively suppresses horizontal artifacts, preserves target hyperbolic contours, and avoids excessive reduction of target scattering, showcasing its exceptional performance. This method provides a powerful algorithmic foundation for high-resolution GPR imaging and target detection.

Keywords: Ground-penetrating radar; Horizontal interference; Diffusion model; Agent attention module; Spatial attention; Hybrid dataset

1 Introduction

In ground-penetrating radar (GPR) profiles, horizontal interference is an unavoidable type of noise present in all types of GPR data. It degrades the quality of the high-resolution images provided by the GPR system and mainly originates from the electronic design of the radar antenna, impedance mismatches between the antenna and the ground, or multiple reflections of electromagnetic waves between subsurface interfaces (Atef & Rashed, 2023). Given the growing applications of GPR in various fields and the continuous development of automatic GPR target detection techniques, it is necessary to propose an effective method to suppress such interference. Horizontal interference typically manifests as high-amplitude horizontal bands spanning the entire GPR profile, often displaying

* Corresponding author.
E-mail address: bqt2310405005@student.cumtb.edu.cn (X. Tang).
Peer review under the responsibility of Tongji University

https://doi.org/10.1016/j.undsp.2025.08.002

fluctuating characteristics that may obscure the true hyperbolic or inclined features of subsurface targets. Therefore, developing an intelligent and effective method for suppressing horizontal interference would provide important algorithmic support for emerging automatic GPR target detection techniques.

In recent years, conventional data processing methods have made significant progress in suppressing horizontal interference in GPR profiles. Researchers have proposed various approaches to effectively separate and remove the low-rank horizontal components in GPR profiles, such as principal component analysis (PCA) (Marukatat, 2023), robust principal component analysis (RPCA) (Li et al., 2019), nonnegative matrix factorization (NMF) (Wang & Zhang, 2013), and go decom position (GoDec) (Guo et al., 2018). However, these conventional methods do not always yield satisfactory results. On the one hand, they have limitations that hinder compatibility with all types of GPR data; on the other hand, these methods often rely on expert-driven subjective design and parameter tuning based on radar echo features, resulting in a low level of automation.

To further enhance the effectiveness of horizontal interference suppression and improve the intelligence of the method, this study focuses on deep learning techniques. Among them, generative adversarial networks (GANs) (Ledig et al., 2017) have been applied to restore visually realistic details. Such methods leverage adversarial optimization between a generator and a discriminator to encourage the generator to produce more realistic images. Generally, GANs require carefully designed auxiliary loss functions (e.g., perceptual loss (Wang et al., 2018) and gradient loss (Ma et al., 2022) to optimize distances in the feature domain. Although GANs can generate rich details, their training often suffers from instability and is prone to mode collapse, resulting in undesirable artifacts.

The task of horizontal interference suppression in GPR profiles using deep learning falls within the domain of "image restoration." Recently, diffusion models have demonstrated outstanding performance in various image generation tasks. Their fundamental principle is to model the diffusion process and learn its reverse process (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song & Ermon, 2019; Song et al., 2020; Rombach et al., 2022; Rissanen et al., 2023). Based on various existing diffusion model frameworks (Yang et al., 2023), this study adopts a diffusion model based on stochastic differential equations (SDEs). The method gradually diffuses the image to a pure noise distribution via SDEs and then generates samples by learning and simulating the corresponding reverse-time SDE (Anderson, 1982). The core lies in training a neural network to estimate the gradient field of the noisy data distribution, i.e., the score function (Song & Ermon, 2019). In current diffusion model training, most networks are based on UNet variants, which lack deeper architectural innovations in feature extraction and context modeling. Meanwhile, in deep learning models for horizontal interference suppression in GPR profiles, existing methods often suffer from insufficient noise resistance and excessive loss of target reflection energy.

This paper proposes the ground penetrating radar-horizontal interference diffusion model (GPR-HIDiff), an innovative diffusion-based approach designed to suppress horizontal low-rank components in GPR profiles. Leveraging the fundamental principle of diffusion models, which progressively add noise to transform the original data distribution into a Gaussian distribution and then perform stepwise denoising through a learned reverse diffusion process, GPR-HIDiff effectively removes horizontal interference while preserving key signal features. This results in enhanced GPR image quality and facilitates more accurate and reliable interpretation of subsurface structures. Specifically, the main contributions of this paper are as follows:

(1) GPR-HIDiff is a generative model designed for GPR horizontal interference suppression. Its core components include: replacing double convolution units with ResBlocks throughout UNet to enhance training stability; introducing agent attention modules (Han et al., 2024) with residual connections between ResBlocks to strengthen global modeling while reducing computational cost; integrating a spatial attention module in the intermediate layer to adaptively optimize features and improve interference suppression; and employing a cosine schedule to guide noise addition during the forward diffusion phase.
(2) A dataset targeting GPR horizontal interference has been constructed, consisting of both real-world and simulated data. For simulation, a numerical model of urban roadbeds based on real structures is first established, followed by forward modeling using finite difference time domain (FDTD) to expand the dataset.
(3) Whether on simulated or real samples, the proposed GPR-HIDiff outperforms advanced image restoration methods in deep learning as well as traditional horizontal interference suppression algorithms in conventional GPR data processing, offering a new perspective for exploring more efficient diffusion model frameworks.

Although this study primarily focuses on suppressing horizontal interference in GPR profiles, the GPR-HIDiff framework possesses no theoretical limitations on its general applicability. Its practical performance is determined by the interference characteristics of the target task, and it can effectively handle various typical interference patterns, including vertical striping noise, strong oscillatory interference induced by metallic media, as well as Gaussian, Poisson, and salt-and-pepper noise. The framework's design demonstrates promising transferability for broader GPR applications.

The remainder of this paper is organized as follows. Section 2 reviews conventional processing algorithms and recent advances in image restoration using GANs and diffusion models. Section 3 introduces some preliminaries of the diffusion process. Section 4 describes the implementation details of GPR-HIDiff. Section 5 presents experiments and analysis. Section 6 concludes the paper.

2 Related work

2.1 Conventional processing methods

Over the past few decades, various methods have been proposed and tested to suppress horizontal interference in GPR data, achieving varying degrees of success. Among them, the earliest and most traditional method is mean scan subtraction (Nobes, 1999). However, this conventional horizontal interference removal approach does not always yield satisfactory results. Consequently, over the past 20 years, many horizontal interference suppression techniques have been successively introduced and extensively tested. For example, Zhou and Tao (2011) proposed GoDec, which offers a more refined decomposition of the original signal matrix and replaces the continuous singular value decomposition operations with bilateral random projections, achieving higher efficiency compared to RPCA. Song et al. (2019) proposed a pre-screening method for GPR-based anti-personnel mine detection based on GoDec. This method first uses the GoDec algorithm to extract landmine images and then applies thresholding for target detection, effectively reducing the probability of false alarms. Kumlu and Erer (2018) studied the cost function of NMF and employed peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) to objectively evaluate the interference-suppressed images. The results showed that NMF with Kullback–Leibler (KL) divergence performs better.

The performance of the aforementioned methods largely depends on the selection of key parameters (e.g., rank constraints, sparsity thresholds, and regularization coefficients). These parameters are typically determined empirically, making it difficult to adapt optimally to the processing requirements of different GPR datasets. To address this limitation, numerous studies have integrated swarm intelligence optimization algorithms—such as genetic algorithm (GA) optimization, the grey wolf optimizer, and particle swarm optimization (He et al., 2021; Liu et al., 2024a; B. Liu et al., 2022a; B. Liu et al., 2022b; S. Liu et al., 2022; Xia et al., 2023)—with conventional signal processing techniques to adaptively optimize parameter configurations, thereby significantly improving signal processing accuracy. Nevertheless, it should be emphasized that conventional signal processing methods are inherently designed for specific types of interference, a characteristic that inevitably limits their generalization capability when confronted with complex and dynamically changing real-world detection scenarios.

2.2 GAN-based models

To improve automation and visual quality, GAN-based image restoration methods introduce refined auxiliary losses to guide the network in generating more realistic results. For instance, Ledig et al. (2017) first proposed the perceptual loss, which calculates the feature distance between the restored image and the real image in the visual geometry group feature space (Simonyan & Zisserman, 2014). Mao et al. (2017) proposed the least squares generative adversarial networks (LSGANs), which use a least squares loss function for the discriminator. Two comparative experiments between LSGANs and conventional GANs were also conducted to demonstrate the stability of LSGANs. To further improve training stability, Wang et al. (2018) introduced an enhanced super-resolution generative adversarial network, which removes batch normalization while adjusting the discriminator constraint to avoid artifact generation. Although GANs can significantly enhance visual quality, their optimization process is often challenging. Moreover, balancing among elaborately designed loss functions typically requires cumbersome techniques (Xiao et al., 2024). In contrast, the proposed GPR-HIDiff benefits from a clearly defined diffusion process, providing a stable training procedure.

2.3 Diffusion-based models

In recent years, numerous generative models based on diffusion mechanisms have been proposed, including the diffusion probabilistic model (DPM) (Sohl-Dickstein et al., 2015), conditional score-based diffusion model (Batzolis et al., 2021), and the denoising diffusion probabilistic model (DDPM) (Ho et al., 2020).

In 2015, Sohl-Dickstein et al. (2015) introduced DPM, which gradually corrupts the structure of the data distribution through a forward diffusion process and then reconstructs it by learning the reverse diffusion process, resulting in a highly flexible and tractable data generation model. Subsequently, Ho et al. (2020) proposed DDPM and demonstrated that diffusion models are capable of generating high-quality samples. The diffusion probabilistic model is a parameterized Markov chain that can be trained via variational inference. In particular, DDPM has shown strong performance in producing high-quality samples and has thus been widely applied in various tasks such as image colorization, super-resolution, inpainting, and semantic editing.

There have been only a few cases applying diffusion models to GPR data processing. For instance, Lan et al. (2024) used low-frequency images with inserted clutter as prior knowledge input to fit the Gaussian-distributed clutter in the forward process of the diffusion model. Huang et al. (2025) enhanced the robustness of GPR inversion by gradually adding and removing noise using the diffusion model. They also integrated

UNet++, vision transformer (ViT) structures, and a simple attention module during the reverse generation process to improve multi-scale feature extraction and contextual understanding, thereby achieving high-accuracy dielectric constant inversion. However, current research applying diffusion models to GPR data processing is extremely limited, particularly in suppressing horizontal interference. Moreover, most existing diffusion model architectures are relatively simple, limiting their capacity to extract features from complex GPR signals. As a result, they struggle to fully capture deep contextual information and multi-scale features, thus hindering the effective separation of target echoes and interference.

2.4 Summary

In summary, GANs and variational autoencoders (VAEs) (Rezende et al., 2014) represent two of the most prominent frameworks in generative modeling. GANs are capable of rapidly producing high-quality GPR samples; however, they suffer from limited mode coverage and insufficient sampling diversity. A common drawback of GANs lies in their unstable training dynamics, which can lead to issues such as mode collapse, vanishing gradients, and convergence issues (Wiatrak et al., 2019). In contrast, while VAEs offer superior mode coverage, they inherently tend to generate GPR samples of relatively low quality. VAEs optimize the log-likelihood of the data by maximizing the evidence lower bound. Despite notable progress, their performance remains unsatisfactory due to challenges such as the balancing problem (Davidson et al., 2018) and the phenomenon of variable collapse (Asperti, 2019). Diffusion models, on the other hand, are a powerful class of probabilistic generative models designed to learn complex data distributions. In the forward diffusion process, noise is progressively added to the input data, increasing its level until the data are transformed into pure Gaussian noise, thereby systematically disrupting the structure of the original distribution. The reverse diffusion process is then applied to recover the original structure from the perturbed distribution, effectively mitigating the degradation introduced during forward diffusion. As a result, diffusion models constitute a highly flexible and tractable generative framework capable of accurately modeling complex GPR data distributions from random noise (Kazerouni et al., 2023).

Considering the complex physical laws in GPR signal propagation and interference formation, future research could incorporate Maxwell's equations as physical constraints into model training, using architectures such as diffusion models to accelerate the solution of electromagnetic field partial differential equations (Thombre et al., 2025; Wang et al., 2024, 2025, 2022), thereby enabling precise characterization of underground electromagnetic wave propagation for efficient modeling of complex physical processes and enhanced suppression of horizontal interference components.

3 Preliminary

3.1 Forward diffusion process

The goal of the forward diffusion process is to gradually convert the initial data distribution $x(0)$ into a noisy image $x(T)$ at time step T. We define the high-resolution image I_{HR} as $x(0)$. Therefore, $x(T)$ can be approximately represented as a combination of a low-resolution image μ and pure Gaussian noise $\varepsilon \sim \mathcal{N}(0, \lambda^2)$, where λ^2 denotes the stationary variance. This paper adopts mean-reverting SDEs (Luo et al., 2023) to define the diffusion process, as this approach can improve sampling efficiency. Specifically, the forward diffusion process is expressed as follows:

$$\mathrm{d}x = \theta_t(\mu - x)\mathrm{d}t + \sigma_t \mathrm{d}w, \tag{1}$$

where θ_t, σ_t are two time-dependent parameters that control the speed of mean reversion and stochastic volatility. w refers to a standard Wiener process. To make Eq. (1) have a closed-form solution, we set$\sigma_t^2/\theta_t = 2\lambda^2$. As illustrated in Fig. 1, given the HR image $x(0)$ and $t \in [0, T]$, the corresponding state $x(t)$ at any intermediate time t can be strictly represented by the closed-form solution of Eq. (1).

$$x(t) = \mu + [x(0) - \mu]\mathrm{e}^{-\bar{\theta}_t} + \int_0^t \sigma_z \mathrm{e}^{-\bar{\theta}_t}\mathrm{d}w(z), \tag{2}$$

where $\bar{\theta}_t$ is equal to $\int_0^t \theta_z \mathrm{d}z$. In this case, $x(t)$ follows a Gaussian probability distribution $p_t(x)$, which is expressed as follows:

$$x(t) \sim p_t(x) = \mathcal{N}[x(t)\dot{m}_t(x), v_t], \tag{3}$$

where $m_t(x) = \mu + [x(0) - \mu]\mathrm{e}^{-\bar{\theta}_t}$ and $v_t = \lambda^2\left(1 - \mathrm{e}^{-2\bar{\theta}_t}\right)$ are the mean and variance of this Gaussian distribution, respectively. It can be readily observed that as the diffusion time $t \to \infty$, m_t and v_t will converge to μ and λ^2, respectively, i.e., the terminal state satisfies $x(T) \approx \mu + \varepsilon$, which aligns with the objective of the forward diffusion process.

In the forward diffusion process, a cosine-based noise variance schedule based on the cosine function is employed to achieve a smooth temporal transition, as detailed in the following equation (Nichol & Dhariwal, 2021):

$$\theta_t = 1 - \frac{f(t)}{f(0)}, \; f(t) = \cos\left(\frac{\frac{t}{T} + s}{1 + s} \cdot \frac{\pi}{2}\right)^2, \tag{4}$$

where $s = 0.008$, the same as that in (Nichol & Dhariwal, 2021). Once the θ function is determined, the corresponding diffusion coefficient σ_t can be computed from the following stationary condition: $\sigma_t^2/\theta_t = 2\ \lambda^2$.

3.2 Reverse diffusion process

The reverse diffusion process aims to recover the HR image from the terminal state $x(T)$. We can define the reverse diffusion process by simulating the reverse-time SDE as follows:

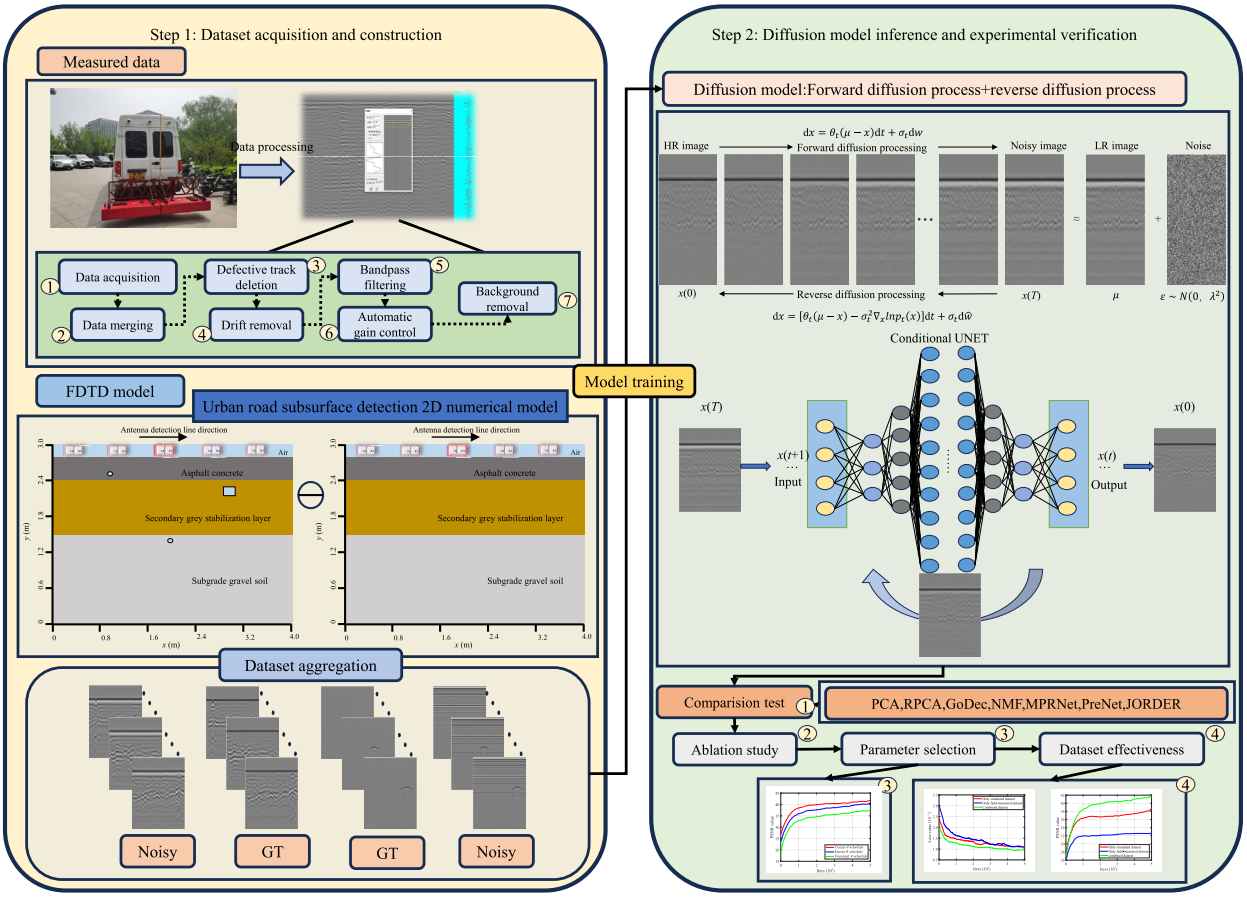

Fig. 1. Overarching framework for horizontal interference suppression.

$$\mathrm{d}x = \left[\theta_t(\mu - x) - \sigma_t^2 \nabla_x \ln p_t(x)\right]\mathrm{d}t + \sigma_t \mathrm{d}\hat{w}, \tag{5}$$

where $\hat{w}$ denotes the reverse-time Wiener process, and $\nabla_x \ln p_t(x)$ is the ground-truth score during the inference phase. It is important to note that during the training phase, the true image $x(0)$ is known, so more reliable conditional scores can be used to train the model. Specifically, it can be defined by

$$\nabla_x \ln p_t(x \mid x(0)) = -\frac{x(t) - m_t(x)}{v_t}. \tag{6}$$

In addition, we reparameterize $x(t)$ to $x(t) = m_t(x) + \sqrt{v_t}\epsilon_t$, where ϵ_t is a standard Gaussian noise with the distribution $\mathcal{N}(0, I)$. The "I" in $\mathcal{N}(0, I)$ represents the identity matrix. The ground-truth score can be represented as $-\frac{\epsilon_t}{\sqrt{v_t}}$. Since $m_t(x)$ and v_t are known, we only need to use a noise prediction network $\widetilde{\epsilon}_\phi$ to estimate the noise.

Similar to DPM, we compute the Euclidean distance between the predicted noise and the true noise ϵ_t using the following formula:

$$\mathscr{L}(\phi) = \sum_{t=0}^{T} \gamma_t \mathbb{E}\left[\left|\left|\widetilde{\epsilon}_\phi(x_t, \mu, t) - \epsilon_t\right|\right|\right], \tag{7}$$

where γ_t denotes the positive weight.

4 Proposed method

The overall research framework shown in Fig. 1 integrates multi-source datasets from both field GPR measurements and FDTD numerical simulations, employing a novel diffusion model to train the dataset for suppressing low-rank horizontal interference in GPR profiles, and demonstrating the framework's advancement through comparative algorithm analysis and experimental validation.

4.1 Overview of the GPR-HIDiff framework

Figure 2 illustrates the detailed flowchart of the proposed GPR-HIDiff. This study adopts a conditional UNet architecture, where the input consists of the noisy image $x(t)$ and μ. By computing $x(t)-\mu$ and concatenating it as the enriched condition I_t into the network input, the network's ability to perceive noise patterns is enhanced. Furthermore, the proposed network replaces the conventional sequential convolution units of UNet with Res-Blocks at a global scale (as shown in Fig. 3), which helps improve deep information transmission and alleviates the gradient vanishing problem. In addition, an agent attention

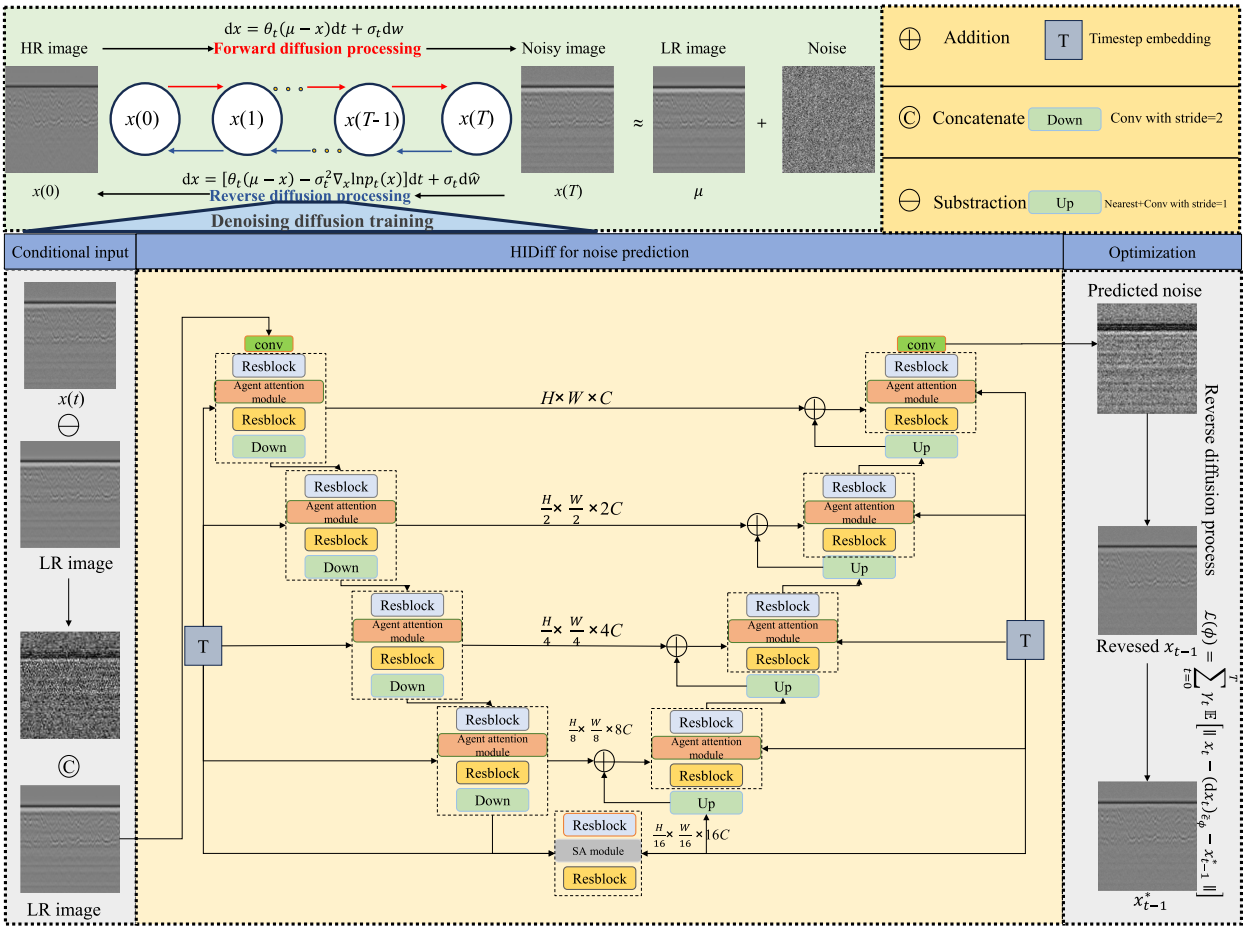

Fig. 2. Overall framework of GPR-HIDiff.

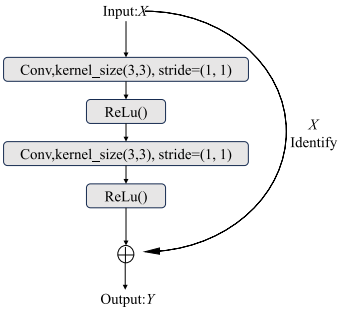

Fig. 3. Internal structure of the ResBlock module.

module is inserted between each ResBlock layer, and a spatial attention module is added between the encoder and decoder. Subsequently, the time-dependent network $\widetilde{\epsilon}_\phi$ takes the external condition I_t and time t as inputs, aiming to predict the output pure noise $\bar{\epsilon}_t$:

$$\bar{\epsilon}_t = \widetilde{\epsilon}_\phi(I_t, t). \tag{8}$$

Here, the study proposes a half instance normalization network (HINet) for noise prediction. Finally, $\widetilde{\epsilon}_\phi$ is optimized until convergence.

4.2 Linear attention module

Most previous methods typically construct the layers of UNet by simply stacking consecutive convolutional units. However, this approach may limit the ability to model global information, thereby affecting reconstruction quality. In this work, an agent attention module with residual connections is introduced between ResBlocks at each layer to enhance the capacity for global information capture. At the same time, this module has low computational complexity, making it more efficient for high-resolution GPR data processing.

Recent studies have found that using the Softmax function in self-attention computations forces pairwise calculations between all queries and keys, resulting in a computational complexity as high as $o(n^2)$. To address this issue, linear attention adopts simple activation functions (Katharopoulos et al., 2020; Shen et al., 2021) or specific mapping functions (Choromanski et al., 2020; Lu et al., 2021) to approximate the original Softmax mechanism. As shown in Fig. 4(a) and (b), linear attention adjusts the computation order from (query × key) × value to query × (key × value), thereby reducing the overall complexity to $o(n)$ and significantly improving computational efficiency.

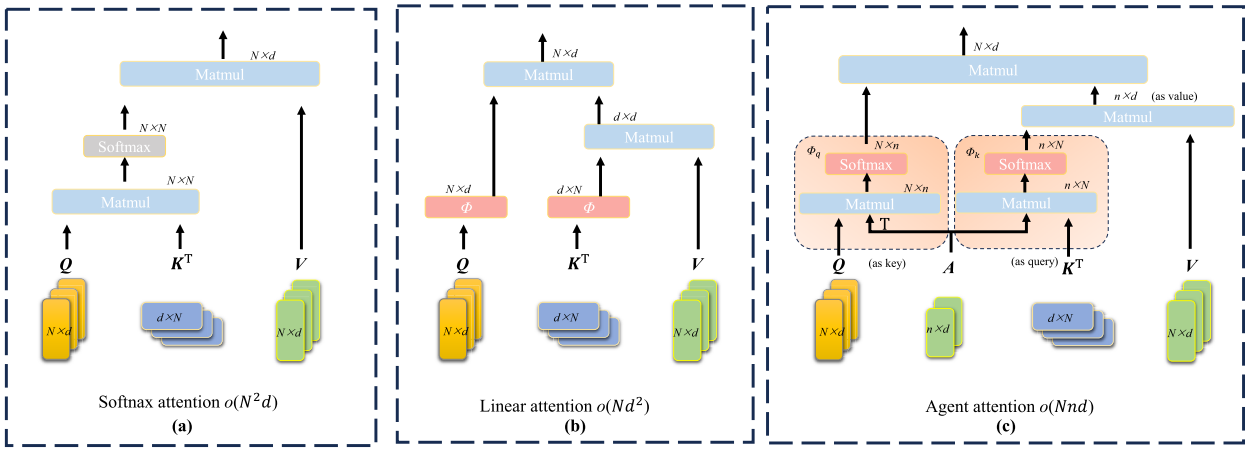

Fig. 4. Difference among Softmax attention, linear attention, and agent attention.

In this study, the agent attention module is introduced. This module innovatively incorporates an additional set of agent vectors $\boldsymbol{A}$ into the attention triplet ($\boldsymbol{Q}$, $\boldsymbol{K}$, $\boldsymbol{V}$), defining a quaternary agent attention paradigm ($\boldsymbol{Q}$, $\boldsymbol{A}$, $\boldsymbol{K}$, $\boldsymbol{V}$). As shown in Fig. 4(c), agent attention does not directly compute the pairwise similarity between $\boldsymbol{Q}$ and $\boldsymbol{K}$, but instead uses a small number of agent vectors $\boldsymbol{A}$ to gather information from $\boldsymbol{K}$ and $\boldsymbol{V}$, which is then delivered to $\boldsymbol{Q}$, thereby achieving global information modeling at a very low computational cost. Structurally, agent attention consists of two conventional Softmax attention operations and is equivalent to a generalized linear attention, naturally fusing the high performance of Softmax attention with the efficiency of linear attention. As a result, it combines the advantages of both—namely, low computational complexity and strong model expressiveness.

4.3 Spatial attention module

Although many UNet variants perform well, they inevitably make the network more complex and less interpretable (Guo et al., 2021). To address these issues, spatial attention modules have been introduced into UNet. This spatial attention helps the network focus on important features while suppressing unnecessary ones, thereby improving the representational capacity of the network.

The spatial attention (SA) module is introduced as a part of the convolutional block attention module (Woo et al., 2018). SA uses the spatial relationship between features to produce a spatial attention map. To calculate spatial attention, SA first applies max pooling and average pooling operations along the channel axis and concatenates them to produce an efficient feature descriptor, as shown in Fig. 5. Formally, the input feature $F \in R^{H\times W\times C}$ is forwarded through the channel-wise max pooling and average pooling to generate outputs $F^{s}_{mp} \in R^{H\times W\times 1}$ and $F^{s}_{ap} \in R^{H\times W\times 1}$, respectively. Then, a convolutional layer followed by the Sigmoid activation function on the concatenated feature descriptor is used to generate a spatial attention map $M^{s}(F) \in R^{H\times W\times 1}$. In short, the output feature $F^{s} \in R^{H\times W\times C}$ of spatial attention module is calculated as (Guo et al., 2021):

$$\begin{aligned} F^{s} &= F \cdot M^{s}(F) = \\ & F \cdot \sigma\left(f^{7\times 7}([\mathrm{MaxPool}(F); \mathrm{AvgPool}(F)])\right) = \\ & F \cdot \sigma\left(f^{7\times 7}\left(\left[F^{s}_{mp}; F^{s}_{ap}\right]\right)\right), \end{aligned} \tag{9}$$

where $f^{7\times 7}(\cdot)$ denotes a convolution operation with a kernel size of 7 and $\sigma(\cdot)$ represents the Sigmoid function.

4.4 Optimization and inference

Although Eq. (7) provides a direct solution for optimizing HINet, diffusion models often encounter instability during training. The key reason for this is that predicting instantaneous noise distribution is not an easy task. Therefore, the study draws on the maximum likelihood learning strategy from (Luo et al., 2023) and modifies the training objective. Specifically, to optimize HINet, we choose to minimize the Euclidean distance as follows:

$$\mathscr{L}(\phi) = \sum_{t=0}^{T} \gamma_t \mathbb{E}\left[\left|\left|x_t - (\mathrm{d}x_t)_{\widetilde{\epsilon}_\phi} - x^{*}_{t-1}\right|\right|\right], \tag{10}$$

where x^{*}_{t-1} is the ideal state reversed from x_t. The closed form of x^{*}_{t-1} can be determined by the following formula:

$$x^{*}_{t-1} = \frac{1-e^{-2\bar{\theta}_{t-1}}}{1-e^{-2\bar{\theta}_t}} e^{-\theta'_t}(x_t - \mu) + \frac{1-e^{-2\theta'_t}}{1-e^{-2\bar{\theta}_t}} e^{-\bar{\theta}_{t-1}}(x_0 - \mu) + \mu. \tag{11}$$

During the inference process, the study uses the pretrained $\widetilde{\epsilon}_\phi$ by sampling from the random state $x(T)$ and employs numerical methods (such as the Euler–Maruyama (EM) method (Kloeden & Platen, 1992)) to iteratively solve the SDE in order to predict high resolution (HR) images. To better understand the training and inference process of GPR-HIDiff, the study summarizes it in Algorithm 1 and Algorithm 2.

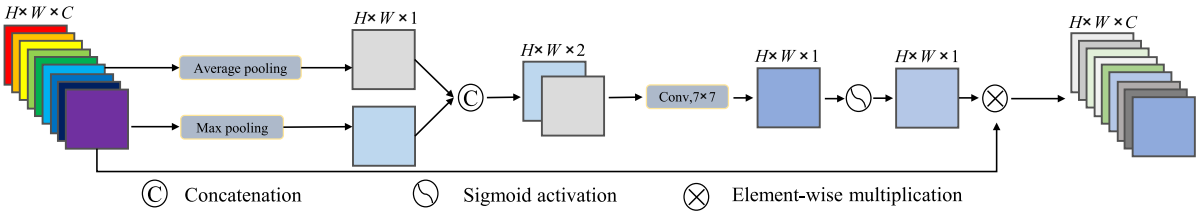

Fig. 5. Internal structure diagram of the SA module.

Algorithm 1 Training of GPR-HIDiff

Input: HR image $x(0) = I_{\mathrm{HR}}$, LR image $\mu{=}I_{\mathrm{LR}}$, total step T, noisy image $x(t)$.
1 Initialization: Random sample $\varepsilon_t \sim \mathcal{N}\left(0, \lambda^2\right), t \in [0, T], T = 100$.
2 repeat
3 I_t= Concate$(x(t) - \mu, \mu)$;
4 $\bar{\epsilon}_t = \widetilde{\epsilon}_\phi(I_t, t)$; // Predict noise
// Substitute score into Eq. (7)
5 $\mathrm{d}x_t = \left[\theta_t(\mu - x_t) + \sigma_t^2 \frac{\bar{\epsilon}_t}{\sqrt{v_t}}\right]\mathrm{d}t + \sigma_t \mathrm{d}\widehat{w}$;
6 $\mathscr{L}(\phi) = \gamma_t \mathbb{E}\left[\left|\left|x_t - (\mathrm{d}x_t)_{\widetilde{\epsilon}_\phi} - x_{t-1}^*\right|\right|\right]$; // Loss
7 $\nabla_\phi \mathscr{L}$; // Gradient descent
8 until converged

Algorithm 2 Inference of GPR-HIDiff

Input: LR image $\mu{=}I_{\mathrm{LR}}$, total step T.
Output: Image after horizontal interference suppression I_{HI}.
1 Initialization: Random sample $x(T)$ $\mathcal{N}\left(0, \lambda^2\right)$, $\widetilde{\epsilon}_\phi$ is the pre-trained HINet, EM$(\cdot)$ is Euler–Maruyama method, $T = 100$.
2 for $t = T$: 1 do
3 $\bar{\epsilon}_t = \widetilde{\epsilon}_\phi(x_t, \mu, t)$; // Predict noise
// Substitute score into Eq. (7)
4 $\mathrm{d}x_t = \left[\theta_t(\mu - x_t) + \sigma_t^2 \frac{\bar{\epsilon}_t}{\sqrt{v_t}}\right]\mathrm{d}t + \sigma_t \mathrm{d}\widehat{w}$;
5 $x_{t-1} = x_t - \mathrm{EM}(\mathrm{d}x_t)$; //Reverse SDE
6 end
7 $I_{\mathrm{HI}} = x_0$;

5 Experiment and discussion

5.1 Dataset

The dataset used in this study consists of two parts: measured samples and simulated samples. For the measured samples, as shown in Fig. 6, the study uses the MTGR-400 GPR from China University of Mining and Technology (Beijing) to detect urban roads. The antenna center frequency is 400 MHz, and the sampling interval is 0.023 m. Among them, the ground truth images for the measured samples are the results after mean scan subtraction. However, as the processing effect of this method is not always ideal, only the profiles with better results are selected. To expand the database and obtain more convincing horizontal interference suppression samples, the study uses FDTD forward modeling for data extension. As shown in Fig. 7, based on the urban road structure, the study randomly stacks small particles of different sizes and dielectric properties within the road layers and randomly generates the shapes of anomalous cavities to simulate the non-homogeneous underground environment. In addition, the simulation process involves 11 random control variables to enhance the algorithm's ability to suppress horizontal interference under different conditions. In the simulated samples, GPR profiles with defects are considered low-quality images, while profiles obtained by subtracting defect-free GPR profiles (to eliminate horizontal interference) are considered as ground truth images. The specific settings for the dielectric parameters of the model medium and signal parameters are detailed in Tables 1 and 2.

In total, the dataset contains 1189 pairs of samples, of which 755 pairs are measured samples and 434 pairs are simulated samples. Figure 8 lists some of the sample pairs, representing all categories in the sample library. Among them, (a) and (b) are 400 MHz GPR measured sample pairs; (c) and (d) are 100 MHz GPR measured sample pairs; and (e) is a 500 MHz GPR simulated sample pair. In addition, the dataset constructed in this study not only includes samples acquired under different frequency conditions, but also exhibits the following representative characteristics: first, the mixed samples display substantial energy variation between individual samples, encompassing horizontal interference with a wide dynamic range of amplitudes; second, particularly in the case of field-measured samples, the dataset covers a variety of typical and complex subsurface scenarios, including but not limited to: underground interfaces with pronounced undulating features, stratigraphic structures with significant differences in phase-aligned reflector continuity, and complex situations where non-geological structures such as pipelines intersect with geological strata. This broadly representative data distribution can effectively enhance the model's generalization capability under complex geological conditions.

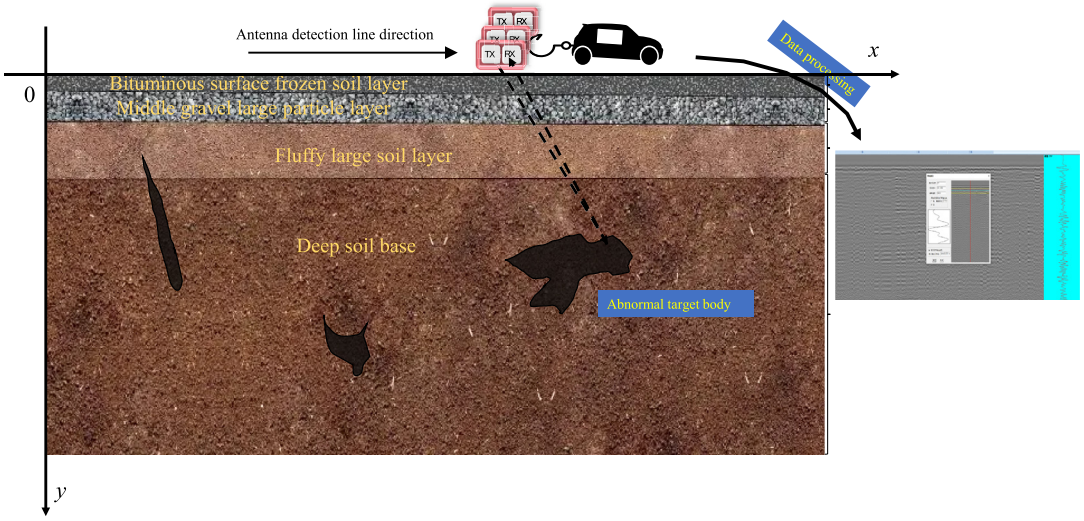

Fig. 6. Ground-penetrating radar detection principle diagram.

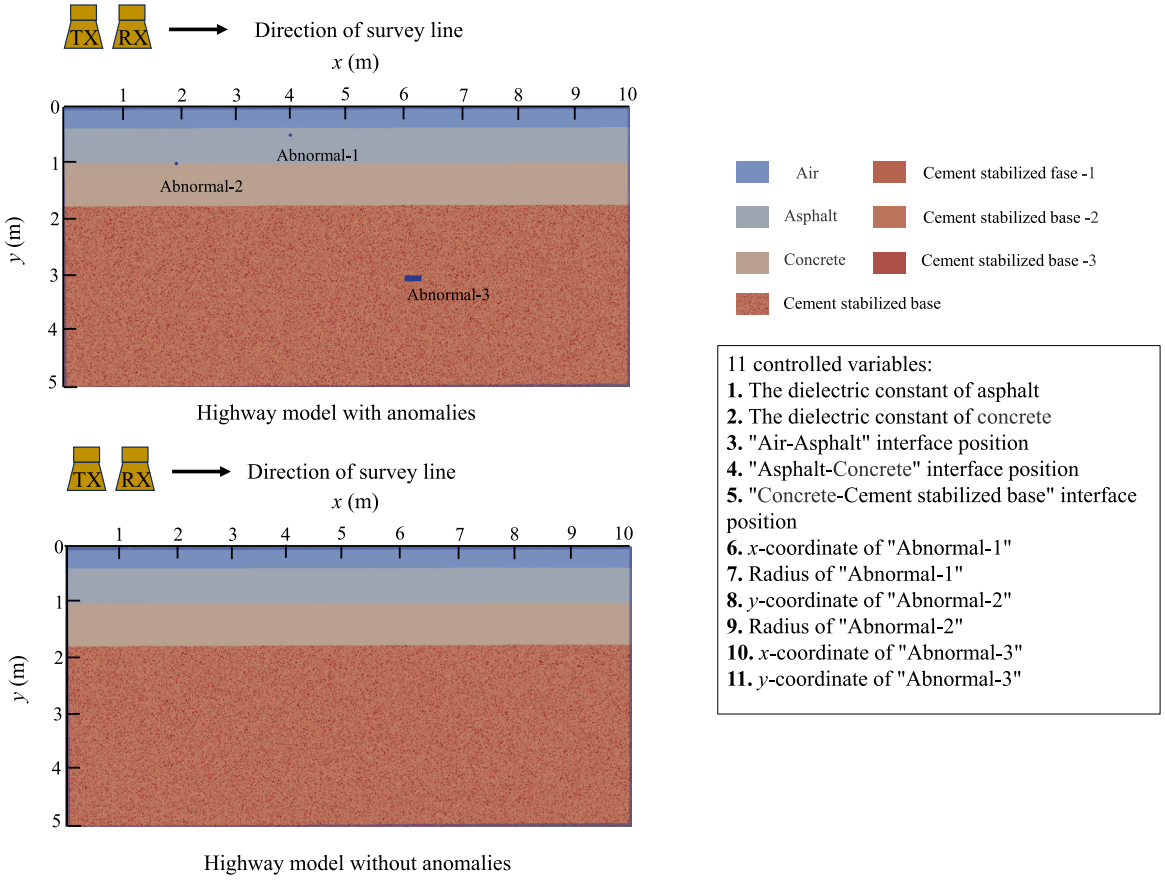

Fig. 7. Model for GPR-HIDiff-FDTD simulation training samples.

5.2 Experimental implementation details

For the proposed GPR-HIDiff network, the patch size is 128×128, and the batch size is 16. The Adam optimizer is employed with parameters $\beta_1 = 0.9$ and $\beta_2 = 0.99$. The total training steps are set to 5×10^5, with an initial learning rate of 10^{-4}, which decays by half every 2×10^5 iterations. The model requires approximately four days to train on a single RTX 4090 GPU (24 GB VRAM). Employing multi-GPU parallelization (e.g., a 4-GPU configuration) can reduce the training time to less than one day. During inference, the model demonstrates excellent

Table 1
Material parameters of GPRMAX.

Material	Relative dielectric constant (ε_r)	Conductivity (S/m)
Air	1.0	0
Asphalt	4.0/5.0	0.0045
Concrete	7.0/8.0	0.0082
Cement stabilized base-1	3.0	0.0070
Cement stabilized base-1	3.2	0.0070
Cement stabilized base-1	2.8	0.0070

Table 2
Signal channel setting parameters of GPRMAX.

Name of the parameter	Value
Domain (m)	10, 5, 0.005
d*x*_d*y*_d*z* (m)	0.005, 0.005, 0.005
Time window (ns)	100
Frequency of the antenna (MHz)	500
Transmitter start position (m)	0.1, 0.1, 0.005
Receiver start position (m)	0.2, 0.1, 0.005
Transmitter and receiver steps (m)	0.1, 0, 0
Number of sampling points	98

performance, processing a single 512 × 512 image in 4.984 s (well under 10 s), thereby meeting the requirements of real-time engineering applications. The final trained model produces a weight file (pth) of 521 MB. For on-site workflows such as construction, the model achieves a favorable balance among accuracy, speed, and computational resource consumption.

5.3 Metrics

The study uses six metrics to comprehensively assess the performance of the GPR-HIDiff model. These include three full-reference metrics: PSNR (Zhang et al., 2017), SSIM (Wang et al., 2004), and learned perceptual image patch similarity (LPIPS) (Yauri-Lozano et al., 2024). These metrics help to evaluate the distance between the generated image and the ground truth images. In addition, there are three no-reference metrics: natural image quality evaluator (NIQE) (Wu et al., 2022), perception based image quality evaluator (PIQE) (Pandey et al., 2020), and the maximum-to-average (M/A) amplitude ratio (Rashed, 2003). These metrics offer insights into the perceptual quality and the high-frequency details of the generated images. Specifically, M/A amplitude ratio is a measure of the strength of a signal relative to background noise on a seismic section, and the same principle can be applied to GPR sections. All amplitudes are squared, then square rooted before calculating the ratio to avoid cancellation of amplitudes having opposite signs. The ratio is calculated using the following formula:

$$(M/A)_j = \frac{\max\left(|A_{i,j}|\right)}{\frac{1}{N}\sum_{i=1}^{N}|A_{i,j}|}, \tag{12}$$

$$\overline{M/A} = \frac{1}{M}\sum_{j=1}^{M}(M/A)_j, \tag{13}$$

where M and N represent the number of survey lines and time sampling points in the GPR profile, respectively. $A_{i,j}$ is the amplitude value at the i-th time sampling point on the j-th survey line.

5.4 Comparisons

We compare GPR-HIDiff with mainstream conventional GPR processing methods, including PCA, RPCA, GoDec, and NMF, as well as advanced deep learning algorithms: Multi-stage progressive image restoration (MPRNet) (Zamir et al., 2021), progressive recurrent network (PReNet) (Ren et al., 2019), joint rain detection

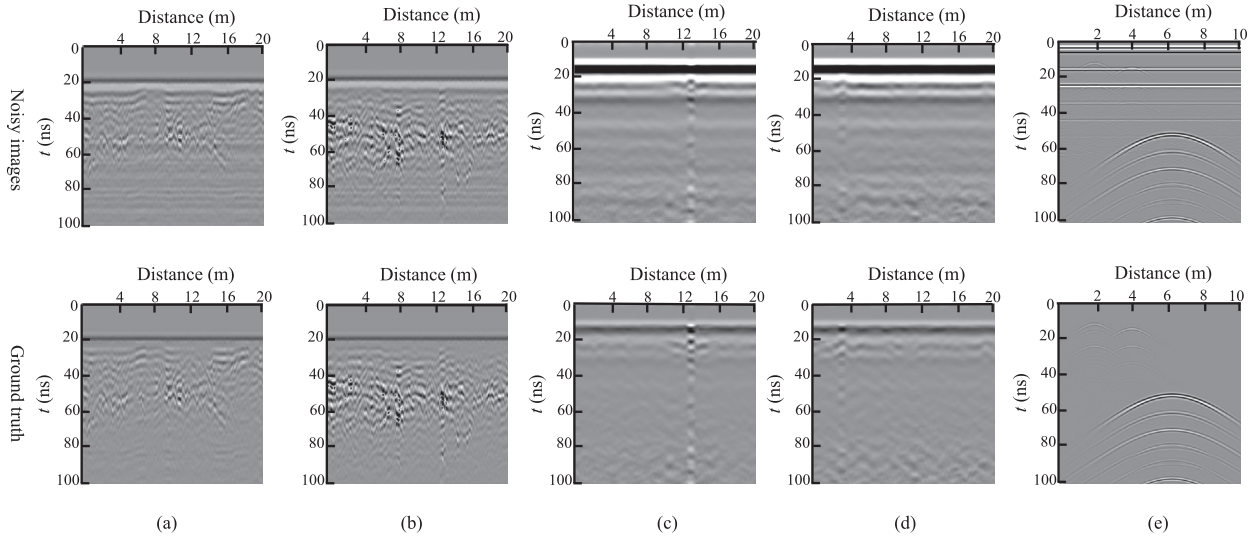

Fig. 8. Example of measured and simulated samples required for GPR-HIDiff training.

and removal (JORDER) (Yang et al., 2017), and image restoration using swin transformer (SwinIR) (Liang et al., 2021). These methods represent the mainstream in the fields of horizontal interference suppression and image restoration, ensuring the comprehensiveness of the evaluation. Among them, MPRNet is a multi-stage restoration network with adaptive attention, excelling in deraining, deblurring, and denoising. PReNet is a progressive recurrent network that unfolds shallow ResNet and recurrent layers to exploit feature dependencies, maintaining excellent performance while reducing parameters. The SwinIR combines the advantages of convolutional networks and transformers and can be divided into three components: shallow feature extraction, deep feature extraction, and high-quality image reconstruction modules. For these advanced deep learning methods, the experiments strictly follow the official implementations, and all comparative methods are trained from scratch on the dataset described in Section 5.1.

In summary, based on eight indicators (both full-reference and no-reference), the study compares the GPR-HIDiff model with eight advanced algorithms (including conventional GPR processing methods and deep learning methods) on a dataset composed of both real and simulated samples. Table 3 presents the quantitative results of different algorithms on the test set. In each row, we highlight the best performance for each metric in bold red. It can be observed that GPR-HIDiff achieves the best performance across the board, and the deep learning algorithms consistently outperform conventional GPR processing methods. Among the deep learning algorithms, MPRNet and the proposed GPR-HIDiff perform particularly well; among the conventional GPR methods, RPCA and GoDec are more prominent.

Additionally, we compare the average inference time and GPU memory usage. Conventional methods exhibit significantly faster inference than deep learning models because they typically only require direct matrix operations, without involving the multi-layer convolutional computations or iterative optimization processes inherent to deep learning. Although the inference time of deep learning models is generally longer, it usually remains under 10 s, which is sufficiently short to avoid noticeably affecting real-time performance in practical applications. In terms of GPU memory usage, MPRNet requires substantial memory due to its multi-stage design, whereas diffusion models demand considerable memory because their multi-step iterative reverse denoising process necessitates storing numerous intermediate features; both approaches lead to significantly increased memory consumption when processing high-resolution images such as 512×512. While conventional methods have advantages in speed and memory efficiency, they generally fall short of deep learning models in terms of accuracy.

Next, the study conducts a qualitative analysis of different algorithms based on four groups of sample pairs.

5.5 Experiment on simulated samples

The study conducts evaluations on simulated data. The suppression results of different algorithms are shown in Fig. 9. For conventional GPR processing methods, PCA retains all principal components from the 2nd to the 512th; RPCA is configured with a regularization multiplier of 0.1, a convergence tolerance of 1×10^{-7}e–7, and a maximum of 500 iterations; the rank for NMF decomposition is set to 1; for GoDec, the rank of the low-rank matrix is set to 1 and the number of iterations to 10. Deep learning algorithms are implemented strictly following their official configurations.

In the original B-scan noisy profile, numerous target hyperbolic reflections are obscured by horizontal interference and clutter, making their contours and spatial locations difficult to discern. Specifically, the term "hyperbolic contours" refers to the pattern formed when a radar antenna passes over a point-like subsurface target: the receiving antenna records the reflected signals, producing an initial trace. As the antenna approaches the target, the propagation path shortens, reaching its minimum directly above the target; as the antenna moves away, the path lengthens again. The variation in travel time with horizontal distance forms a characteristic hyperbolic shape on the radar profile, resulting from the combined effects of geometric path changes and the target's reflection properties. As indicated by the yellow boxes in Fig. 9, conventional GPR methods cannot suppress the random clutter

Table 3
Quantitative experimental results on the test set. The best PSNR, SSIM, LPIPS, NIQE, PIQE, and $\bar{M}/A$ performances are shown in bold.

Metrics/Method	PCA	RPCA	GoDec	NMF	MPRNet	PReNet	JORDER	SwinIR	GPR-HIDiff
PSNR (↑) (dB)	30.7247	32.3551	34.8543	29.9641	37.1111	36.5903	36.5558	36.7843	**38.9498**
SSIM (↑)	0.9575	0.9591	0.9653	0.9435	0.9748	0.9727	0.9730	0.9735	**0.9812**
LPIPS (↓)	0.0711	0.0781	0.0594	0.0877	0.0375	0.0470	0.0426	0.0403	**0.0225**
NIQE (↓)	13.7192	12.3571	11.8772	15.5520	10.5311	11.2934	11.3821	11.2251	**9.6214**
PIQE (↓)	62.9452	58.1890	55.9247	66.7447	47.3556	49.0184	52.1792	46.4612	**45.2318**
$\bar{M}/A$ ratio (↑)	1.4853	1.5391	1.5414	1.4520	1.7256	1.7128	1.6913	1.7124	**1.8452**
Average inference time (↓) (s)	**0.0220**	0.964	0.093	0.024	5.324	4.127	5.261	3.561	4.984
GPU memory usage (↓) (MB)	No	No	No	No	3827	**2216**	2352	3168	3304

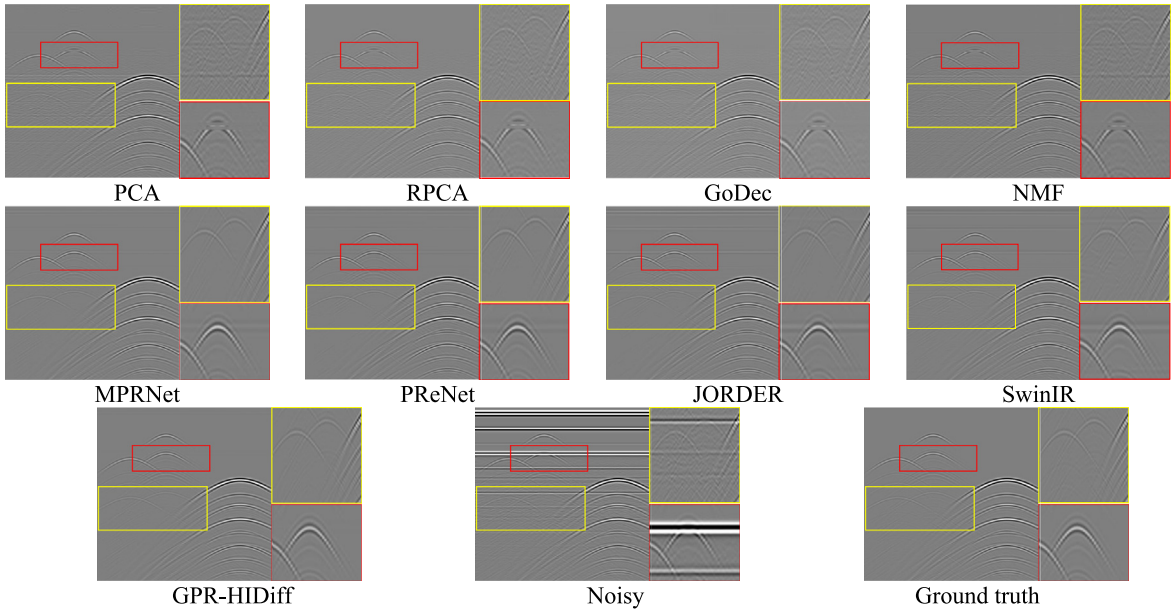

Fig. 9. Suppression effect of different algorithms on horizontal interference in simulated profiles.

caused by heterogeneous underground environments, leaving substantial residual noise. Some horizontal interference persists in the PCA and NMF results, while RPCA and GoDec perform relatively better. In contrast, deep learning methods show more significant suppression of both horizontal interference and clutter. Notably, the proposed method reveals more scattered contours of effective target hyperbolas compared to the other three advanced algorithms, and the hyperbolas extend further. In particular, the proposed method also unveils a hyperbolic contour formed by the overlapping of two upper hyperbolas, even though this composite hyperbolic signal has very weak energy, demonstrating the strong performance of the proposed algorithm.

As shown in the red boxes in Fig. 9, although conventional algorithms suppress the horizontal interference overlapping the target hyperbolas, they fail to adequately reveal the target contours. The segments obscured by interference are merely suppressed, without exposing the underlying hyperbolic features. In contrast, deep learning algorithms reveal these contours more naturally and realistically. However, visible horizontal interference with considerable amplitude still remains in the profiles processed by MPRNet, PReNet, JORDER, and SwinIR.

Residual maps between the profiles after suppression and the ground truth are also drawn, as shown in Fig. 10. The residual energy from deep learning methods is significantly lower than that of conventional methods, indicating better performance. Moreover, the residual profiles of the proposed method show minimal horizontal interference. For conventional methods, PCA and NMF leave substantial horizontal interference in the residual maps and are ineffective at suppressing it. RPCA, GoDec, and NMF retain stronger target hyperbolic responses in the residuals, indicating that these methods remove more target scattering contributions.

In summary, only the proposed algorithm effectively suppresses interference while preserving the morphological features of the majority of target hyperbolas. It enhances the target detection capability of GPR profiles more than the other algorithms and thus demonstrates an overwhelming advantage on simulated data.

5.6 Experiment on measured samples

Subsequently, the study evaluates two measured samples from test sets and one measured sample without ground truth.

In the first example, the suppression results of different algorithms are shown in Fig. 11. Compared with deep learning methods, conventional methods produce weaker overall in-phase reflection group amplitudes in the GPR profiles. However, conventional methods consistently suppress the direct wave, whereas deep learning methods cannot completely suppress the direct wave due to the characteristics of the measured dataset. For conventional methods, weak horizontal artifacts appear in the red and yellow boxes of the PCA-, RPCA-, GoDec-, and NMF-processed profiles. These artifacts blur the target wave groups and reduce their energy. In particular, the NMF-processed profiles exhibit vertical strip artifacts and color distortion in some regions, severely affecting the overall profile quality. RPCA and GoDec perform relatively better in this regard. For deep learning algorithms, such as MPRNet, PReNet, and JORDER, some residual horizontal interference remains in the red, yellow, and green boxes,

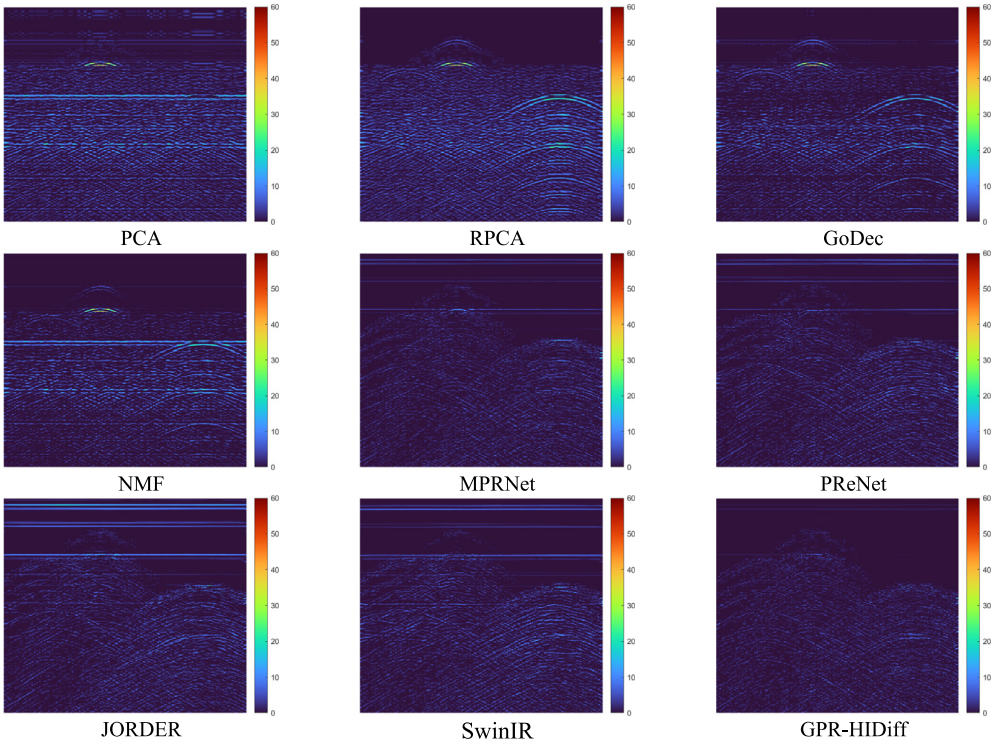

Fig. 10. Residual maps between profiles after suppression and ground truth profiles by different algorithms.

especially near the top or bottom of the profiles. In the red box, SwinIR exhibits residual weak-amplitude horizontal artifacts that have not been fully suppressed.

Notably, the in-phase reflection group energy in deep learning-processed profiles is generally stronger than that of conventional methods. They achieve strong suppression of high-amplitude horizontal interference at the bottom of the profile without significantly weakening the energy of the wave groups. As shown in Fig. 12, we extract the region with pixel coordinates x: 1–512, y: 179–215 (from the full area x: 1–512, y: 1–512). For this 37 × 512 region, we square the pixel values and sum them column-wise to obtain the energy value for each of the 512 columns. Each algorithm yields 512 such values, and there are 9 algorithms in total (i.e., 9 × 512 energy values). The study plots all 9 curves together. Furthermore, for quantitative analysis, a bar chart of the total absolute energy across all pixels in this region is created. To facilitate observation, a section of the curve plot is extracted. Combining the two figures, it can be seen that the energy values from deep learning algorithms are consistently higher than those of conventional methods.

Moreover, this example is a sample pair from the test set. Its ground truth is actually obtained by denoising the background of the corresponding noisy profile. However, as mentioned earlier, background denoising does not always yield satisfactory results. In particular, for this example, the suppression effect is poor due to the selection of an unsuitable background during the denoising process, resulting in residual weak-amplitude horizontal artifacts within the red box. Nevertheless, the proposed algorithm can still achieve a good effect. Additionally, there are some similar cases (with poor ground truth) in the training set established in the study. However, the proposed algorithm can still correct the erroneous sample pairs by learning from a large number of correct sample pairs. The excellent performance of the proposed algorithm in this example is attributed to the strong learning capability of the adopted model.

As shown in Fig. 13, similar to the first sample pair, apart from the proposed algorithm, the profiles processed by other methods more or less exhibit some weak-amplitude horizontal artifacts, strong-amplitude horizontal interference, striped artifacts in random regions, or color distortion. Furthermore, the profile processed by the proposed algorithm retains stronger wavelet energy overall, indicating its relatively optimal performance.

In addition, the study also tests the Noisy-3 profile without ground truth. The results are shown in Fig. 14. Weak-amplitude horizontal artifacts appear within the red boxes of PCA and GoDec; striped artifacts are present in the red boxes and weak-amplitude horizontal artifacts in the yellow boxes of RPCA and NMF. At the same location, MPRNet, PReNet, JORDER, and SwinIR produce strong-amplitude horizontal interference that does not exist in the original Noisy-3 profile.

From the above three groups of examples, conventional algorithms do show significant suppression of strong-amplitude horizontal interference. However, the resulting profiles often contain horizontal or striped artifacts,

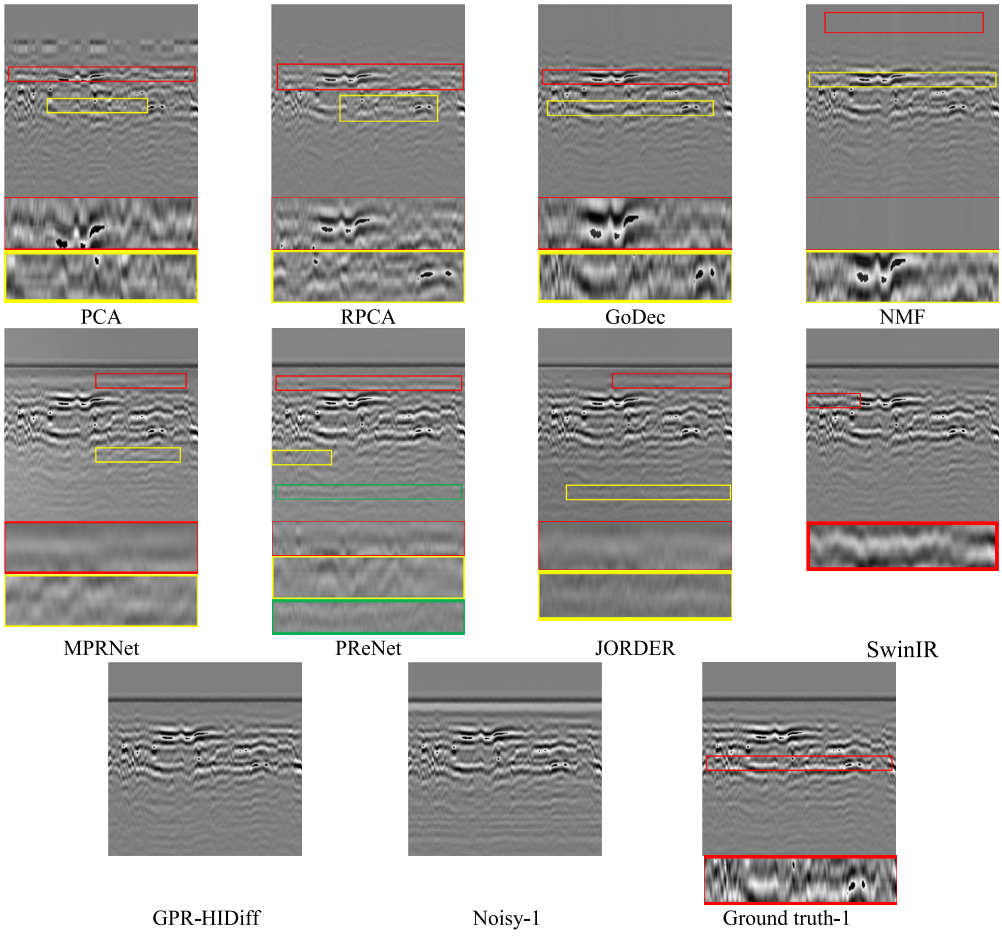

Fig. 11. Suppression effect of different algorithms on horizontal interference in measured profile Noisy-1.

particularly color distortion and excessive reflection group energy loss. Thus, their metrics in Table 3 are not advantageous. Although some deep learning algorithms are less effective in suppressing strong-amplitude horizontal interference compared to conventional methods, they result in minimal reflection group energy loss, giving them a relative advantage in the evaluation metrics.

5.7 Experimental analysis

5.7.1 Ablation study

To investigate the effectiveness of the agent attention module and the SA module in GPR-HIDiff, we conduct a controlled experiment by selectively discarding or retaining the two modules. By comparing the quantitative results on the test set for different module combinations in Table 4, we observe that adding the SA module to the baseline yields slightly better metrics than adding the agent attention module. Notably, the SA module is used only once in the model, while the agent attention module is utilized in most layers of the UNet. When both the agent attention module and the SA module are included, the model performs best across all metrics. Compared with the baseline, the proposed algorithm achieves a 9.0% improvement in PSNR, a 1.1% improvement in SSIM, a 56.8% improvement in LPIPS, a 16.5% improvement in NIQE, a 14.9% improvement in PIQE, and a 15.2% improvement in the M/A ratio. These results demonstrate that the introduced agent attention module and SA module enhance the performance of the diffusion model.

5.7.2 Robustness analysis

To evaluate the robustness of the model, we construct simulated models with varying numbers of layers using the method described in Section 5.1, producing low-quality images. High-quality images are obtained by subtracting the simulated images without targets from those containing targets. Examples of these simulated data pairs with different PSNR/SSIM values are shown in Fig. 15.

Based on 144 data pairs generated using the above approach, we compare the PSNR and SSIM between the clean images and the radar profiles both before and after suppression by the GPR-HIDiff model to evaluate the effectiveness of interference removal. Scatter plots are drawn, and linear fits are applied, as shown in Fig. 16. In the figure, R denotes the overall correlation coefficient

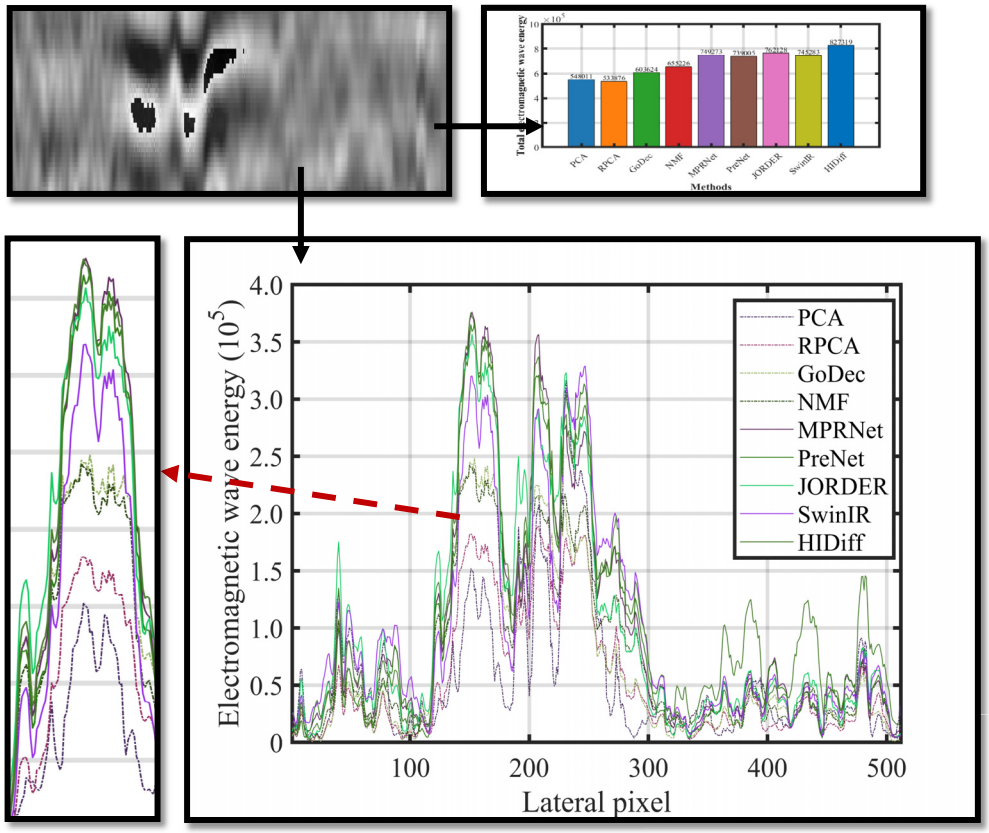

Fig. 12. Energy curve plots and bar charts of different algorithms for the sample pair of Example 1 within the region x: 1–512, y: 179–215.

between PSNR before and after suppression, as well as between SSIM before and after suppression. Overall, the results indicate that both PSNR and SSIM generally improve after suppression as their pre-suppression values increase. However, the correlation coefficient for PSNR before and after suppression is relatively stronger than that for SSIM.

5.7.3 Parameter selection

We can simply adjust θ to construct different noise schedules in GPR-HIDiff, as shown in Fig. 17. It should be noted that the loss and PSNR evaluation curves are smoothed using the Savitzky-Golay method (Sadeghi et al., 2020). We explore three different schedules for how to vary θ: constant, linear, and cosine. When θ is constant, the GPR-HIDiff simplifies to the Ornstein–Uhlenbeck (OU) process (Gillespie, 1996), which is widely used to solve mean-reverting problems. The linear/cosine schedules are widely used in existing diffusion probabilistic models. The constant schedule increases noise at a fixed rate throughout the diffusion process, lacking flexibility. The linear schedule decreases to zero much more rapidly, resulting in the loss of information faster than necessary. In contrast, the cosine schedule is designed to decay smoothly in the middle of the process, while changing very little near the extremes of $t = 0$ and $t = T$ to prevent abrupt changes in noise level (Nichol & Dhariwal, 2021). From Fig. 17, it is evident that while all three schedules perform well in the task of horizontal interference suppression in GPR, with the cos θ schedule significantly outperforming the others.

5.7.4 Dataset effectiveness

The dataset used in this study mainly consists of two types of samples: measured and simulated samples. To further verify the impact of dataset composition on model performance, we adopt a controlled variable approach and divide the training set into three types: (1) containing only measured samples; (2) containing only simulated samples; and (3) a mixture of measured and simulated samples. To ensure fairness in comparison, all three training sets use the same validation set, which is composed solely of simulated samples. We train the GPR-HIDiff model on these three types of training sets, respectively, and the training results are shown in Fig. 18. As can be seen from the results, the model trained with mixed samples performs significantly better in both loss and PSNR compared to models trained on a single type of sample. This indicates that measured and simulated samples have a certain degree of complementarity in feature distribution, and their combination can significantly enhance the model's performance and generalization ability.

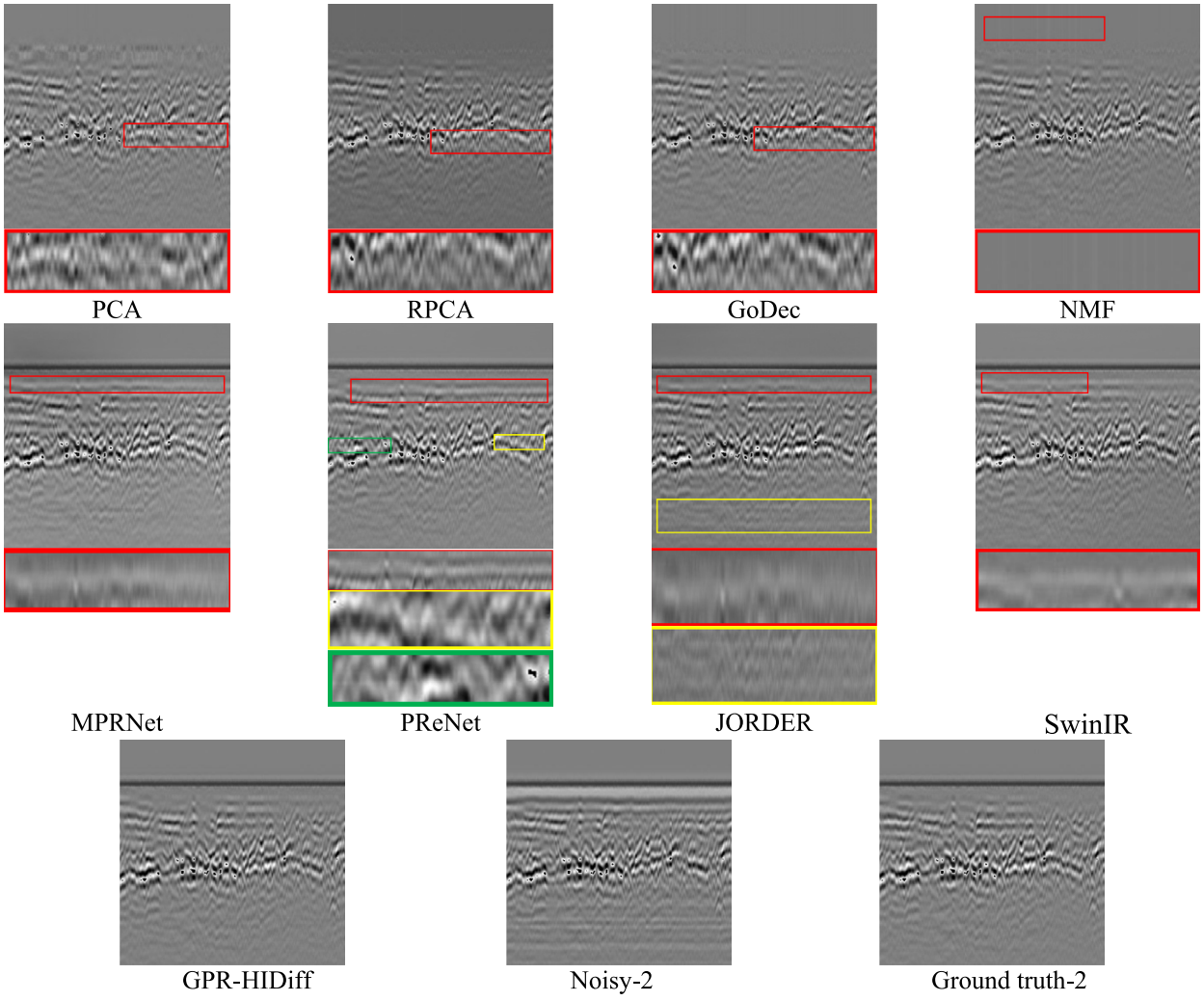

Fig. 13. Suppression results of different algorithms on horizontal interference in the measured profile Noisy-2.

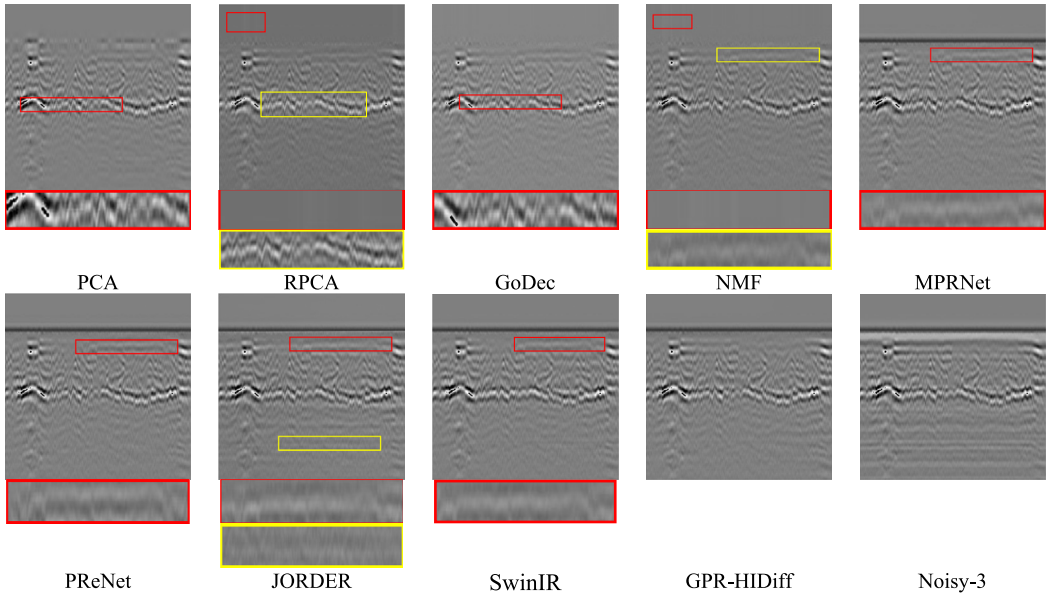

Fig. 14. Suppression results of different algorithms on horizontal interference in the measured profile Noisy-3.

Table 4
Ablation analysis of GPR-HIDiff with different components. The best PSNR, SSIM, LPIPS, NIQE, PIQE and $\bar{M}/A$ performance is shown in bold.

Agent attention module	SA module	PSNR (↑)	SSIM (↑)	LPIPS (↓)	NIQE (↓)	PIQE (↓)	$\bar{M}/A$ ratio (↑)
✗	✗	35.7329	0.9703	0.0521	11.5282	53.1398	1.6021
✓	✗	36.8920	0.9743	0.0389	10.8723	48.2425	1.7143
✗	✓	37.1294	0.9782	0.0337	10.2531	46.9532	1.7241
✓	✓	**38.9498**	**0.9812**	**0.0225**	**9.6214**	**45.2318**	**1.8452**

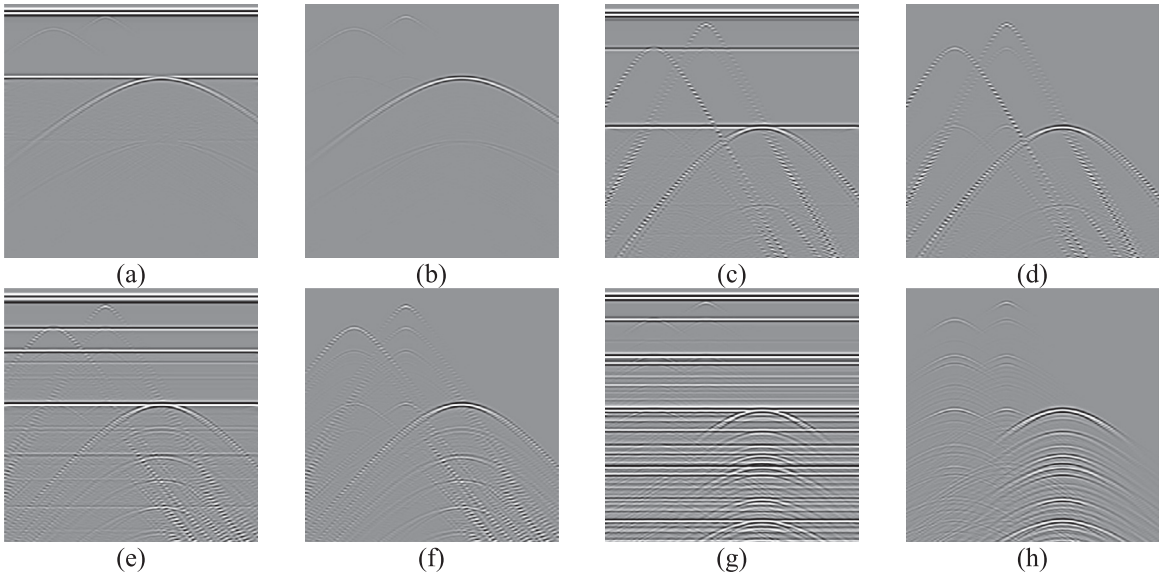

Fig. 15. Examples of simulated data pairs (low-quality image – high-quality image) with different PSNR/SSIM values. Specifically, (a) and (b) form one data pair with a PSNR of 19.91 dB and SSIM of 0.4551; (c) and (d) form a pair with a PSNR of 19.46 dB and SSIM of 0.5937; (e) and (f) form a pair with a PSNR of 17.51 dB and SSIM of 0.4669; (g) and (h) form a pair with a PSNR of 12.97 dB and SSIM of 0.1711.

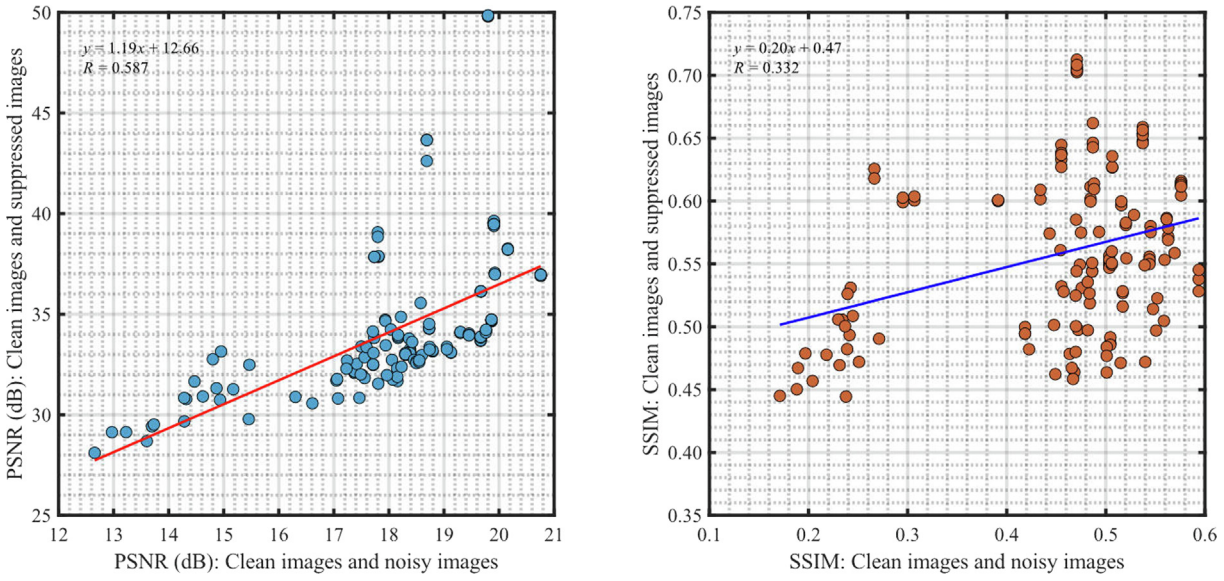

Fig. 16. Scatter plots of PSNR and SSIM between the clean images and the radar profiles before and after suppression by the GPR-HIDiff model (with the horizontal axis representing the x-axis and the vertical axis representing the y-axis).

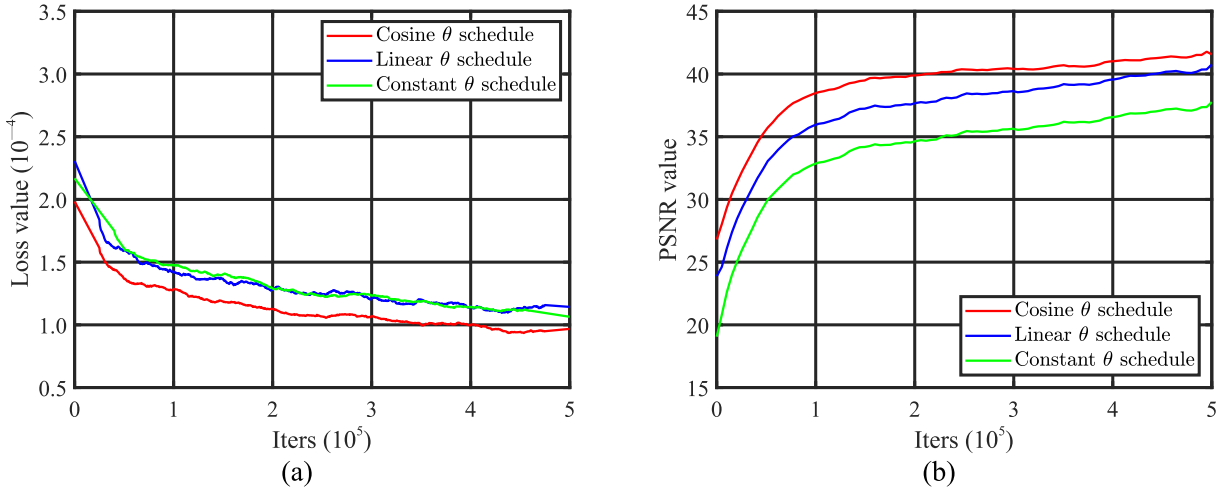

Fig. 17. Performance comparison of GPR-HIDiff under different θ schedules. (a) Loss curves of GPR-HIDiff under different θ schedules, and (b) PSNR evaluation curves of GPR-HIDiff under different θ schedules.

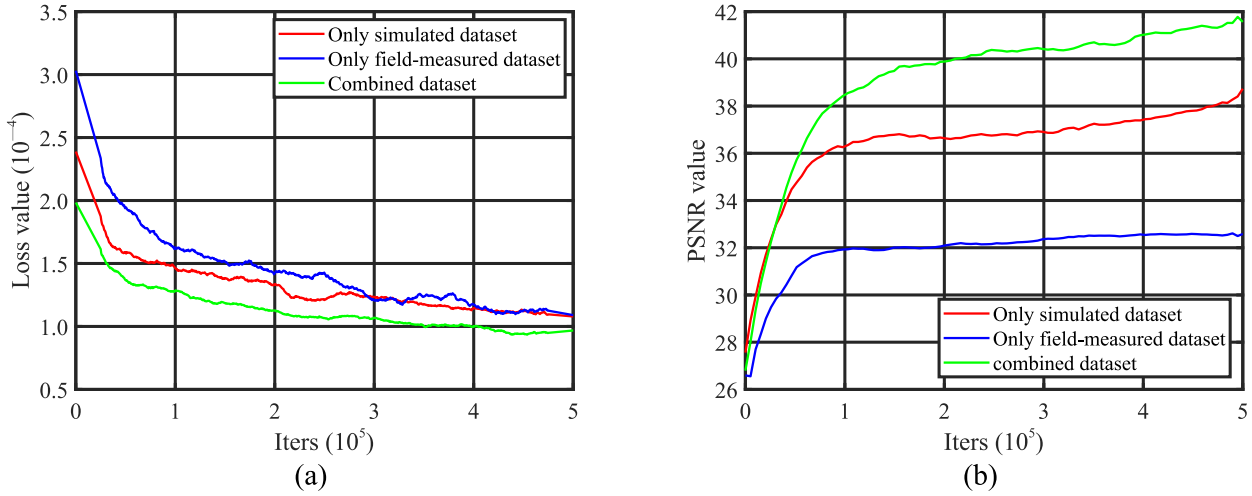

Fig. 18. Performance comparison of GPR-HIDiff under different training sets. (a) Loss curves of GPR-HIDiff under different training sets, and (b) PSNR evaluation curves of GPR-HIDiff under different training sets.

6 Conclusions

To address the issue of horizontal interference components in GPR data that severely degrade the accuracy of subsurface utility interpretation in construction environments, this study proposes a novel diffusion model framework for automated interference suppression. To the best of our knowledge, this represents one of the earliest attempts to apply deep learning to this specific challenge within automated underground utility detection systems.

(1) The proposed model, GPR-HIDiff, comprehensively replaces traditional convolutional units with ResBlocks in the encoder, intermediate layers, and decoder of the UNet architecture to enhance training stability. Lightweight agent attention modules are embedded between ResBlocks at each layer to strengthen global information modeling capability. A spatial attention mechanism is deployed between the encoder and decoder to achieve adaptive feature optimization in the spatial dimension. In the forward diffusion phase, a cosθ schedule is introduced to ensure a smooth temporal variation of noise variance.

(2) GPR-HIDiff outperforms conventional GPR processing algorithms and other advanced deep learning methods on the test set. For simulated test samples, GPR-HIDiff effectively suppresses horizontal interference while preserving the hyperbolic contours of valid targets, without excessively reducing the scattering contributions of the targets. For real test samples, the profiles processed by GPR-HIDiff exhibit no low-amplitude horizontal artifacts, strip-like artifacts, blurred reflection wave groups, or color distortions, while maintaining clear and high-energy reflection wave groups. Furthermore, even when some ground truth data in the training set are of suboptimal qual-

ity, GPR-HIDiff is capable of correcting them through its strong learning ability, demonstrating excellent robustness and generalization.

(3) Ablation experiments show that introducing either agent attention or spatial attention individually can improve performance, with the best results achieved when both are used in combination; removing both leads to a significant performance drop. The robustness analysis compares the PSNR and SSIM between the radar profiles before and after suppression by the GPR-HIDiff model and the clean images, based on 144 data pairs, and scatter plots are drawn with linear fits applied. Among the three θ schedule strategies—constant, linear, and cosine—the $\cos\theta$ schedule performs the best. A standardized GPR horizontal interference dataset is constructed using real measurements and FDTD simulations of urban road models. Performance comparisons between models trained on real samples only, simulated samples only, and a mix of both indicate a degree of complementarity in the feature space of the two sample sets, and the use of mixed samples enhances the model's performance.

In the future, we aim to develop more flexible approaches capable of effectively handling various low-rank noise types across complex geological scenarios and integrating this framework into automated on-site data acquisition systems. In addition, we plan to incorporate sensitivity analysis techniques, such as the Sobol method (Liu et al., 2025; Liu et al., 2020), to quantify the contributions of key network parameters—such as the number of ResBlocks, the placement of attention modules, and the number of diffusion steps—to the final suppression performance, as measured by metrics like SSIM and PSNR. Furthermore, given the critical role of physical laws in GPR signal propagation and interference formation, advanced techniques such as physics-informed neural networks (PINNs) (Cai et al., 2021; Liu et al., 2024b) could be employed in the future to implicitly approximate solutions of partial differential equations during model training by using the Maxwell equations, which describe subsurface electromagnetic wave propagation, as physical constraints combined with automatic differentiation. This innovative approach, which deeply integrates physical principles with data-driven features, is expected to significantly enhance both the interpretability and generalization capability of the model.

Data availability

The data that used in this study are available on https://github.com/txs129129/code–data–horizontal-iInterference.

CRediT authorship contribution statement

Xiaosong Tang: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Feng Yang:** Validation, Supervision. **Xu Qiao:** Visualization, Validation. **Jialin Liu:** Writing – original draft, Validation. **Haitao Zuo:** Visualization, Validation. **Liang Gao:** Writing – original draft. **Jianshe Zhao:** Visualization, Validation. **Suping Peng:** Validation, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant No. 52427901) and the Fundamental Research Funds for the Central Universities (Ph.D. Top Innovative Talents Fund of CUMTB) (Grant No. BBJ2025073).

References

Anderson, B. D. O. (1982). Reverse-time diffusion equation models. *Stochastic Processes and their Applications, 12*(3), 313–326.

Asperti, A. (2019). Variational autoencoders and the variable collapse phenomenon. *Sensors & Transducers, 234*(6), 1–8.

Atef, A. H., & Rashed, M. A. (2023). GPR ringing suppression using lateral outliers' swap filter. *Journal of Applied Geophysics, 208*, 104873.

Batzolis, G., Stanczuk, J., Schönlieb, C. B., & Etmann, C. (2021). Conditional image generation with score-based diffusion models. preprint. https://arxiv.org/abs/2111.13606.

Cai, S. Z., Mao, Z. P., Wang, Z. C., Yin, M. L., & Karniadakis, G. E. (2021). Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mechanica Sinica, 37*, 1727–1738.

Choromanski, K. M., Likhosherstov, V., Dohan, D., Song, X. Y., Gane, A., Sarlos, T., Hawkins, P., Davis, J. Q., Mohiuddin, A., Kaiser, L., Belanger, D. B., Colwell, L. J., & Weller, A. (2020). Rethinking attention with performers. *Advances in Neural Information Processing Systems*. preprint. https://arxiv.org/abs/2009.14794.

Davidson, T. R., Falorsi, L., Cao, N. D., Kipf, T., & Tomczak, J. M. (2018). Hyperspherical variational auto-encoders. preprint. https://arxiv.org/abs/1804.00891.

Gillespie, D. T. (1996). Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Physical Review E, 54*(2), 2084–2091.

Guo, C. L., Szemenyei, M., Yi, Y. G., Wang, W. L., Chen, B., & Fan, C. Q. (2021). SA-UNet: Spatial attention U-Net for retinal vessel segmentation. In *Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR)* (pp. 1236–1242).

Guo, K. L., Liu, L., Xu, X. M., Xu, D., & Tao, D. C. (2018). GoDec+: Fast and robust low-rank matrix decomposition based on maximum correntropy. *IEEE Transactions on Neural Networks and Learning Systems, 29*(6), 2323–2336.

Han, D. C., Ye, T. Z., Han, Y. Z., Xia, Z. F., Pan, S. Y., Wan, P. F., Song, S. J., & Huang, G. (2024). Agent attention: On the integration of softmax and linear attention. In *Proceedings of the 18th European Conference on Computer Vision* (pp. 124–140).

He, X. K., Wang, C., Zheng, R. Y., & Li, X. W. (2021). GPR image noise removal using grey wolf optimisation in the NSST domain. *Remote Sensing, 13*(21), 4416.

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems, 33*, 6840–6851.

Huang, M. J., Wang, Y. H., Wu, Y. Q., & Jia, Z. (2025). Ground Penetrating Radar Inversion Via Steady-State Diffusion Processes. *IEEE Transactions on Geoscience and Remote Sensing, 63*, 1–9.

Katharopoulos, A., Vyas, A., Pappas, N., & Fleuret, F. (2020). Transformers are rnns: Fast autoregressive transformers with linear atten-

tion. In *Proceedings of the 37th International Conference on Machine Learning* (pp. 5156–5165).
Kazerouni, A., Aghdam, E. K., Heidari, M., Azad, R., Fayyaz, M., Hacihaliloglu, I., & Merhof, D. (2023). Diffusion models in medical imaging: A comprehensive survey. *Medical Image Analysis, 88*, 102846.
Kloeden, P. E., & Platen, E. (1992). *Numerical solution of stochastic differential equations*. Springer.
Kumlu, D., & Erer, I. (2018). Performance evaluation of NMF methods with different divergence metrics for landmine detection in GPR. In *Proceedings of the International Society for Optical Engineering (SPIE)* (pp. 158–167).
Lan, T., Luo, X., Yang, X. P., Gong, J. B., Li, X. J., & Qu, X. D. (2024). A constrained diffusion model for deep GPR image enhancement. *IEEE Geoscience and Remote Sensing Letters, 21*, 1–5.
Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. H., & Shi, W. Z. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 105–114).
Li, Y., Liu, G. C., Liu, Q. S., Sun, Y. B., & Chen, S. Y. (2019). Moving object detection via segmentation and saliency constrained RPCA. *Neurocomputing, 323*, 352–362.
Liang, J. Y., Cao, J. Z., Sun, G. L., Zhang, K., Van Gool, L., & Timofte, R. (2021). SwinIR: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 1833–1844).
Liu, B. K., Liu, P. J., Lu, W. Z., & Olofsson, T. (2025). Explainable artificial intelligence (XAI) for material design and engineering applications: A quantitative computational framework. *International Journal of Mechanical System Dynamics, 5*(2), 236–265.
Liu, B. K., Lu, W. Z., Olofsson, T., Zhuang, X. Y., & Rabczuk, T. (2024a). Stochastic interpretable machine learning based multiscale modeling in thermal conductivity of Polymeric graphene-enhanced composites. *Composite Structures, 327*, 117601.
Liu, B. K., Vu-Bac, N., Zhuang, X. Y., Fu, X. L., & Rabczuk, T. (2022a). Stochastic full-range multiscale modeling of thermal conductivity of Polymeric carbon nanotubes composites: A machine learning approach. *Composite Structures, 289*, 115393.
Liu, B. K., Vu-Bac, N., Zhuang, X. Y., Fu, X. L., & Rabczuk, T. (2022b). Stochastic integrated machine learning based multiscale approach for the prediction of the thermal conductivity in carbon nanotube reinforced polymeric composites. *Composites Science and Technology, 224*, 109425.
Liu, B. K., Vu-Bac, N., Zhuang, X. Y., & Rabczuk, T. (2020). Stochastic multiscale modeling of heat conductivity of polymeric clay nanocomposites. *Mechanics of Materials, 142*, 103280.
Liu, B. K., Wang, Y. Z., Rabczuk, T., Olofsson, T., & Lu, W. Z. (2024b). Multi-scale modeling in thermal conductivity of polyurethane incorporated with phase change materials using physics-informed neural networks. *Renewable Energy, 220*, 119565.
Liu, S. X., Chen, Y. H., Luo, C. P., Jiang, H. J., Li, H., Li, H. Q., & Lu, Q. (2022). Particle swarm optimization-based variational mode decomposition for ground penetrating radar data denoising. *Remote Sensing, 14*(13), 2973.
Lu, J. C., Yao, J. H., Zhang, J. G., Zhu, X. T., Xu, H., Gao, W. G., Xu, C. J., Xiang, T., & Zhang, L. (2021). Soft: Softmax-free transformer with linear complexity. preprint. https://arxiv.org/abs/2110.11945.
Luo, Z. W., Gustafsson, F. K., Zhao, Z., Sjölund, J., & Schön, T. B. (2023). Image restoration with mean-reverting stochastic differential equations. In *Proceedings of the 40th International Conference on Machine Learning* (pp. 23045–23066).
Ma, C., Rao, Y. M., Lu, J. W., & Zhou, J. (2022). Structure-preserving image super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(11), 7898–7911.
Mao, X. D., Li, Q., Xie, H. R., Lau, R. Y., Wang, Z., & Smolley, S. P. (2017). Least squares generative adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 2813–2821).
Marukatat, S. (2023). Tutorial on PCA and approximate PCA and approximate kernel PCA. *Artificial Intelligence Review, 56*, 5445–5477.
Nichol, A. Q., & Dhariwal, P. (2021). Improved denoising diffusion probabilistic models. In *Proceedings of the 38th International Conference on Machine Learning* (pp. 8162–8171).
Nobes, D. C. (1999). Geophysical surveys of burial sites: A case study of the Oaro urupa. *Geophysics, 64*(2), 357–367.
Pandey, A., Yadav, D., Sharma, A., Sonker, D., Patel, C., Bal, C., & Kumar, R. (2020). Evaluation of perception based image quality evaluator (PIQE) no-reference image quality score for 99mTc-MDP bone scan images. *Journal of Nuclear Medicine, 61*(S1), 1415.
Rashed, M. A. (2003). *Optimum-offset weighted stacking: A novel Processing technique to enhance signal-to-noise ratio in seismic data acquired in urban areas and its application on Uemach Fault, Osaka, Japan*. [Doctoral dissertation, Osaka City University, Japan].
Ren, D. W., Zuo, W. M., Hu, Q. H., Zhu, P. F., & Meng, D. Y. (2019). Progressive image deraining networks: A better and simpler baseline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3932–3941).
Rezende, D. J., Mohamed, S., & Wierstra, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. In *Proceedings of the 31st International Conference on Machine Learning* (pp. 1278–1286).
Rissanen, S., Heinonen, M., & Solin, A. (2023). Generative modelling with inverse heat dissipation. In *Proceedings of the 11th International Conference on Learning Representations (ICLR)* (pp. 1–54).
Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 10684–10695).
Sadeghi, M., Behnia, F., & Amiri, R. (2020). Window selection of the Savitzky–Golay filters for signal recovery from noisy measurements. *IEEE Transactions on Instrumentation and Measurement, 69*(8), 5418–5427.
Shen, Z. R., Zhang, M. Y., Zhao, H. Y., Yi, S., & Li, H. S. (2021). Efficient attention: Attention with linear complexities. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 3530–3538).
Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representations* (ICLR). April 14–16, 2b014, Banff, Canada.
Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., & Ganguli, S. (2015). Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning* (pp. 2256–2265).
Song, J. M., Meng, C. L., & Ermon, S. (2020). Denoising diffusion implicit models. preprint. https://arxiv.org/abs/2010.02502.
Song, X. J., Xiang, D. L., Zhou, K., & Su, Y. (2019). Fast prescreening for GPR antipersonnel mine detection via go decomposition. *IEEE Geoscience and Remote Sensing Letters, 16*(1), 15–19.
Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representations* (ICLR). April 14–16, 2014, Banff, Canada.
Song, Y., & Ermon, S. (2019). Generative modeling by estimating gradients of the data distribution. In *Proceedings of the 34th International Conference on Neural Information Processing Systems* (pp. 6840–6851).
Thombre, M., Anitescu, C., Bharadwaja, B., Wang, Y. Z., Rabczuk, T., & Alankar, A. (2025). Energy-based methods for solving forward and inverse linear elasticity problems in 2D structures. *Computers & Structures, 316*, 107899.
Wang, X. T., Yu, K., Wu, S. X., Gu, J. J., Liu, Y. H., Dong, C., Qiao, Y., & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 63–79).
Wang, Y. X., & Zhang, Y. J. (2013). Nonnegative matrix factorization: A comprehensive review. *IEEE Transactions on Knowledge and Data Engineering, 25*(6), 1336–1353.
Wang, Y. Z., Bai, J. S., Lin, Z. Y., Wang, Q. M., Anitescu, C., Sun, J., Eshaghi, M. S., Gu, Y. T., Feng, X. Q., Zhuang, X. Y., Rabczuk, T., & Liu, Y. H. (2024). Artificial intelligence for partial differential equations in computational mechanics: A review. preprint. https://arxiv.org/abs/2410.19843.
Wang, Y. Z., Sun, J., Bai, J. S., Anitescu, C., Eshaghi, M. S., Zhuang, X. Y., Rabczuk, T., & Liu, Y. H. (2025). Kolmogorov–Arnold-Informed neural network: A physics-informed deep learning framework for solving forward and inverse problems based on Kolmogorov–Arnold Networks. *Computer Methods in Applied Mechanics and Engineering, 433*, 117518.

Wang, Y. Z., Sun, J., Li, W., Lu, Z. Y., & Liu, Y. H. (2022). CENN: Conservative energy method based on neural networks with subdomains for solving variational problems involving heterogeneous and complex geometries. *Computer Methods in Applied Mechanics and Engineering, 400*, 115491.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*(4), 600–612.

Wiatrak, M., Albrecht, S. V., & Nystrom, A. (2019). Stabilizing generative adversarial networks: A survey. preprint. https://arxiv.org/abs/1910.00927.

Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In *Proceedings of the 15th European Conference on Computer Vision (ECCV)* (pp. 3–19).

Wu, L. Y., Zhang, X. G., Chen, H., Zhou, Y. C., Wang, L. H., & Wang, D. J. (2022). An efficient unsupervised image quality metric with application for condition recognition in kiln. *Engineering Applications of Artificial Intelligence, 107*, 104547.

Xia, Y. Y., Zhang, C., Wang, C. X., Liu, H. J., Sang, X. X., Liu, R., Zhao, P., An, G. F., Fang, H. Y., Shi, M. S., Li, B., Yuan, Y. M., & Liu, B. K. (2023). Prediction of bending strength of glass fiber reinforced methacrylate-based pipeline UV-CIPP rehabilitation materials based on machine learning. *Tunnelling and Undeground Space Technology, 140*, 105319.

Xiao, Y., Yuan, Q. Q., Jiang, K., He, J., Jin, X. Y., & Zhang, L. P. (2024). EDiffSR: An efficient diffusion probabilistic model for remote sensing image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 5601514.

Yang, L., Zhang, Z. L., Song, Y., Hong, S. D., Xu, R. S., Zhao, Y., Zhang, W. T., Cui, B., & Yang, M. H. (2023). Diffusion models: A comprehensive survey of methods and applications. *ACM Computing Surveys, 56*(4), 1–39.

Yang, W. H., Tan, R. T., Feng, J. S., Liu, J. Y., Guo, Z. M., & Yan, S. C. (2017). Deep joint rain detection and removal from a single image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1685–1694).

Yauri-Lozano, E., Castillo-Cara, M., Orozco-Barbosa, L., & García-Castro, R. (2024). Generative adversarial networks for text-to-face synthesis & generation: A quantitative–qualitative analysis of natural language processing encoders for Spanish. *Information Processing & Management, 61*(3), 103667.

Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., Yang, M. H., & Shao, L. (2021). Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 14816–14826).

Zhang, K., Zuo, W. M., Gu, S. H., & Zhang, L. (2017). Learning deep CNN denoiser prior for image restoration. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2808–2817).

Zhou, T. Y., & Tao, D. C. (2011). GoDec: Randomized low-rank & sparse matrix decomposition in noisy case. In *Proceedings of the 28th International Conference on Machine Learning* (pp. 33–40).