

ARTICLE

CT-UNet: Curvelet transform-based component-wise deep learning framework for seismic trace interpolation

Jiyun Yu¹ , Young Kim², and Daeung Yoon^{1*} ¹Department of Energy and Resources Engineering, Chonnam National University, Gwangju, Republic of Korea²YK Geophysics, Seoul, Republic of Korea

Abstract

Seismic data may suffer from missing traces due to economic and environmental constraints as well as equipment malfunctions, where regular gaps caused by receiver or streamer spacing and irregular gaps resulting from topographic limitations occur simultaneously. Such missing data degrade the reliability of subsequent processing and interpretation, making accurate interpolation essential. High-resolution reconstruction requires balanced recovery of broadband components spanning low to high frequencies; however, existing deep learning-based interpolation methods exhibit degraded performance in certain spectral regions. In this study, we propose the curvelet transform (CT)-UNet, a deep learning interpolation method that leverages CT-based frequency- and directional-component separation to improve reconstruction balance across frequency bands, particularly under high-rate, regular undersampling conditions. The proposed method applies the uniform discrete CT (UDCT) to decompose input data into low-frequency and high-frequency components with multi-resolution and directional characteristics, trains specialized U-Net models for each component, and reconstructs the final result through inverse transformation. Performance was evaluated on synthetic data (Society of Exploration Geophysicists/European Association of Geoscientists and Engineers Salt Model) and field data (Mobil Amplitude Versus Offset Viking Graben Line 12) by comparing to f - x interpolation, Constrained Diffusion-Driven Deep Image Prior, a standard U-Net, a parameter-matched U-Net-Wide, and the wavelet-based wavelet transform-UNet. An exploratory experiment on irregular missing and a structural compatibility analysis of the UDCT level configuration were also conducted. Results show that CT-UNet improves reconstruction quality across the tested missing conditions, with the clearest gains observed under high-rate regular undersampling and the 70% irregular missing scenario. Shot-level analysis further supports its advantage in challenging cases, particularly for synthetic 80%, field 75%, field 80%, and 70% irregular missing conditions. The advantage is especially pronounced in recovering curvature-varying, conflicting-dip events, while structural similarity is generally maintained at a level comparable to or higher than that of the baseline methods.

Keywords: Seismic data interpolation; Curvelet transform; Deep learning; U-Net; Uniform discrete curvelet transform; Missing trace reconstruction

***Corresponding author:**
Daeung Yoon
(duyoon@jnu.ac.kr)

Citation: Yu J, Kim Y, Yoon D. CT-UNet: Curvelet transform-based component-wise deep learning framework for seismic trace interpolation. *J Seismic Explor.* 2026;35(3):026120048. doi: 10.36922/JSE026120048

Received: March 16, 2026

Revised: May 27, 2026

Accepted: May 27, 2026

Published online: June 15, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Seismic data may suffer from missing traces due to economic and environmental constraints, as well as equipment and mechanical malfunctions, where regular gaps caused by receiver or streamer spacing and irregular gaps resulting from topographic limitations or instrument failures occur simultaneously. Such missing data not only degrade the reliability of subsequent processing (e.g., noise suppression, multiple removal, velocity analysis, and migration) and interpretation, but their recovery through additional surveys demands considerable time and cost, making accurate interpolation at the data processing stage essential.

Conventional seismic data interpolation methods include f - x interpolation and Projection Onto Convex Sets (POCS). f - x interpolation estimates missing traces by exploiting the predictability of signals in the frequency-space domain and offers the advantages of relative simplicity and computational efficiency.¹ However, in zones where complex dip components coexist or non-stationarity is strong, the assumptions underlying the prediction model are weakened, potentially leading to residual and artifact problems. POCS iteratively updates the solution through projections based on sparsity or constraint sets, providing stable reconstruction even under irregular missing conditions,² but it suffers from increased computational cost and difficulty in guaranteeing uniform recovery across the full spectral bandwidth.

In recent years, deep learning-based interpolation methods capable of effectively modeling complex geological structures and non-stationarity through data-driven learning have been actively investigated. Convolutional neural network (CNN) and U-Net³ architectures have become representative approaches for seismic interpolation, demonstrating strength in recovering missing traces through multi-scale feature extraction and skip connections.⁴⁻¹⁰ More recently, Mukherjee *et al.*¹¹ demonstrated the practical applicability of a three-dimensional (3D) U-Net framework for predicting missing traces in structurally complex field settings. Generative adversarial network (GAN)-based approaches preserve high-frequency fine structures through adversarial training with a discriminator,^{12,13} and transformer-based models have shown promising results in interpolation by leveraging global contextual information.¹⁴ Research on applying diffusion models¹⁵ to seismic interpolation has also been actively progressing.¹⁶⁻¹⁹ Diffusion models learn the distribution of training data through a progressive denoising process and offer the advantage of flexible reconstruction for diverse missing patterns. Although these methods achieve higher reconstruction accuracy

than conventional techniques, the difficulty of a single network simultaneously satisfying all frequency and dip components results in spectral bias or residuals in certain regions.

To mitigate these limitations, various strategies for directly controlling spectral characteristics have been introduced. These include approaches that add Fourier loss or frequency-domain constraints to prevent excessive smoothing of high-frequency components (HFs),²⁰ dual-domain architectures that simultaneously incorporate time-frequency information,²¹ and spectral constraint-based training strategies that directly control spectral energy distribution and phase coherence.^{21,22} However, these approaches also tend to concentrate performance on bands that are relatively easier to learn, rather than fully guaranteeing balance across the entire bandwidth.

As an alternative approach, methods have been proposed that decompose signals into multi-resolution or subband representations and reconstruct each component with individual networks. The wavelet transform (WT), with its strength in time-frequency localization, has been widely used to separate and reconstruct scale-specific features of seismic signals.²³⁻²⁶ However, wavelets have limited directional representation, and when dip or curvilinear reflection events are complex, residuals may persist in certain directional components.²⁷ This necessitates a transform-based design capable of simultaneously and precisely representing both directionality and scale.

The curvelet transform²⁸ has already demonstrated its effectiveness in seismic data recovery. Herrmann and Hennenfent²⁹ showed that seismic missing-trace recovery is achievable through curvelet-based sparsity-promoting optimization, providing empirical evidence that the sparse representation capability of curvelets is well-suited for seismic interpolation. However, such optimization-based approaches incur high computational costs due to iterative operations and are sensitive to hyperparameters, with limited generalization to complex geological structures.³⁰ In this study, we combine the representational power of curvelets with data-driven deep learning rather than iterative optimization, simultaneously pursuing inference efficiency and reconstruction quality.

More recently, coordinate-based implicit neural representations (INRs) have emerged as a continuous-representation alternative for seismic data reconstruction. Sitzmann *et al.*³¹ introduced sinusoidal representation networks (SIRENs), which use periodic activation functions to overcome the spectral bias of conventional multilayer perceptrons (MLPs) and represent complex signals at fine detail. Building on this foundation, Gao *et al.*³² proposed Neural Implicit Representations for Five-Dimensional

(5D) Seismic Data Interpolation, a coordinate-based MLP for self-supervised 5D seismic interpolation that infers missing trace amplitudes directly from spatio-temporal coordinates. SIREN-based approaches have also been applied to unsupervised single-gather seismic interpolation, demonstrating robust reconstruction under regular missing rates of up to 75%.³³ INR-based methods offer flexibility across diverse missing patterns and do not require large labeled training datasets; however, they typically rely on per-instance optimization at inference time and face challenges in generalization to large-scale datasets.

In this study, we propose the CT-UNet framework, which integrates the curvelet transform into deep learning-based interpolation. The curvelet transform retains the multi-resolution characteristics of wavelets while explicitly providing directionality and anisotropy, enabling efficient representation of reflection events with curvilinear and dipping structures.²⁸ Although the contourlet,³⁴ a similar directional multi-resolution transform, has been proposed, the UDCT provides a uniform sampling lattice that is more suitable for integration with CNNs. CT-UNet decomposes input seismic data into multi-scale and directional components using the UDCT, interpolates each component with dedicated U-Net models, and reconstructs the final result across the full bandwidth through inverse transformation.

Unlike generic multi-network approaches, the design of CT-UNet is structurally motivated by the 2-lattice decimation property of the UDCT, which naturally partitions the high-frequency subbands into groups with distinct spatial sampling structures, enabling component-wise learning aligned with the transform's inherent decomposition. The main contributions of this paper are as follows: (i) We propose the CT-UNet pipeline, which decomposes seismic data into multi-scale and directional components via the UDCT, reconstructs each component with dedicated U-Net models, and synthesizes the final result through inverse transformation. (ii) We demonstrate that the proposed method is particularly effective in recovering curvature-varying (conflicting-dip) events, where conventional single-network approaches tend to fail, through reduction of residuals in specific frequency and dip components, improving reconstruction quality compared to conventional deep learning interpolation using a single network without component separation. (iii) Through comparison with WT-based methods, we verify the advantages that the directional and anisotropic representation of the curvelet transform provides for seismic event reconstruction. (iv) We exploratorily examine the characteristics of CT-UNet under irregular

missing scenarios and confirm the validity of the selected configuration through a structural compatibility analysis of the UDCT level settings.

2. Methods

2.1. Curvelet transform-UNet

The CT-UNet is a framework that takes two-dimensional seismic data containing missing traces as input, performs multi-resolution and multi-directional decomposition based on the UDCT, interpolates each decomposed component with dedicated U-Net models, and reconstructs a coherent section in the original domain via inverse transformation (Figure 1).

Applying the UDCT forward transform to the input data decomposes it into a low-frequency component (LF) and high-frequency directional components (HF_V , HF_H), where each component is fed into an independent U-Net interpolation model for component-wise prediction. The predicted components are then synthesized through the UDCT inverse transform to produce the final interpolation result (Figure 1A). Figure 1B illustrates the schematic of the UDCT forward and inverse transform processes.

In conventional original-domain interpolation, a single model must simultaneously recover both the low-frequency background and the high-frequency detailed reflection events, leading to imbalanced reconstruction quality across frequency bands and directions. In contrast, CT-UNet reformulates the reconstruction problem component-by-component using the scale and directional separation provided by the UDCT, and learns tailored to the statistical characteristics of each component, thereby aiming for stable interpolation across the full bandwidth.

2.2. Uniform discrete curvelet transform forward transform

Given seismic data X , CT-UNet first applies the UDCT forward transform operator $C(\cdot)$ to compute a set of subband coefficients decomposed into multiple scales and directions (Equation 1).

$$\{C_{j,k}\} = C(X) \quad (1)$$

where $\{C_{j,k}\}$ is the set of UDCT coefficients, j is the scale (resolution) index, and k is the directional index.

The UDCT forward transform consists of two stages. First, in the multi-resolution filter bank stage, analysis filters $F_0(z)$ through $F_N(z)$ are applied to the input signal to decompose the frequency bands by scale, and downsampling generates data at each scale (Figure 2A).

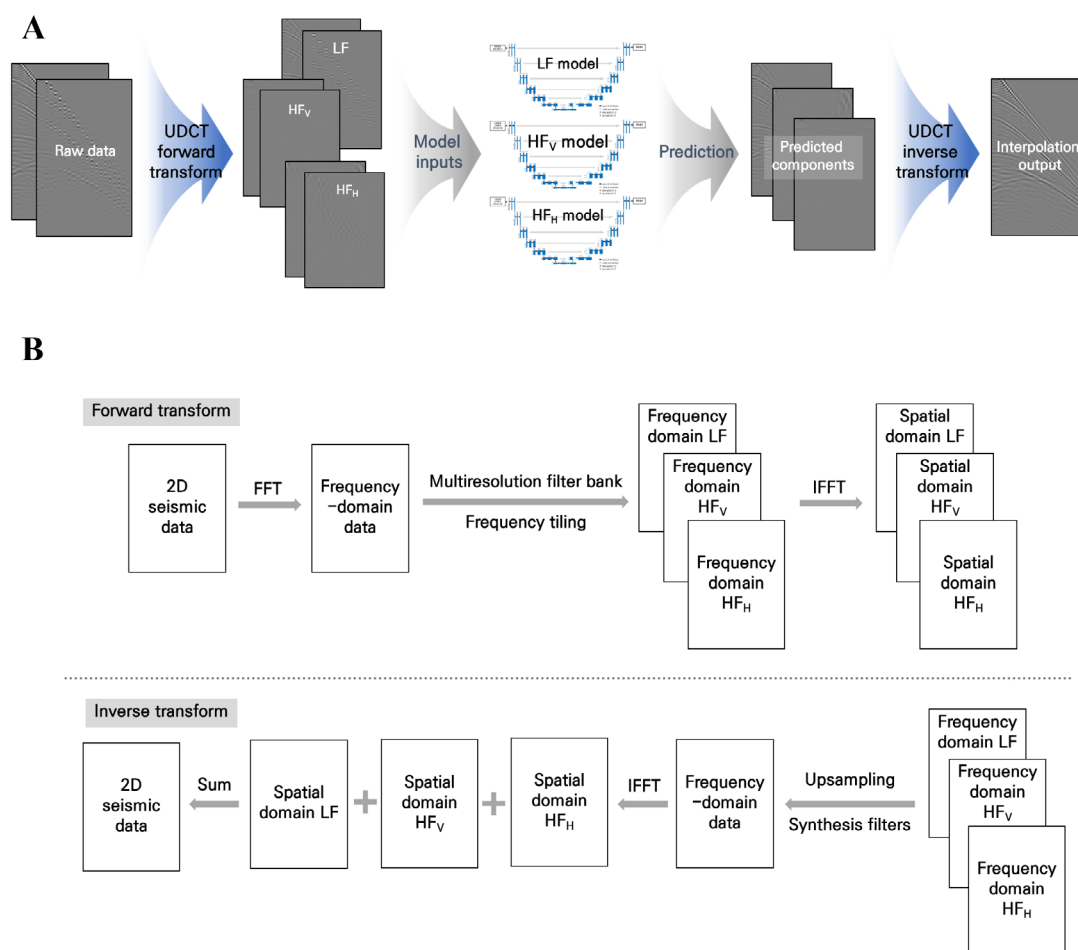


Figure 1. Overall workflow of CT-UNet, where (A) illustrates the process of component decomposition, component-wise interpolation, and inverse transform synthesis, and (B) shows the schematic of the UDCT forward and inverse transforms

Abbreviations: 2D: Two-dimensional; CT: Curvelet transform; FFT: Fast Fourier Transform; HF: High-frequency component; IFFT: Inverse Fast Fourier Transform; LF: Low-frequency component; UDCT: Uniform discrete curvelet transform.

Second, in the frequency tiling stage, the frequency plane of each scale-decomposed band component is subdivided into multiple directional tiles, so that components corresponding to directional events (dip) are mapped to distinct directional tiles.

In this study, the level configuration was set to [8,8], where the first parameter (8) specifies the number of directional decompositions per scale (8 directions) and the second parameter (8) specifies the resolution/scale depth parameter governing the multi-resolution filter bank. Reflecting the downsampling results, the tiled components are extracted at sizes smaller than the original input size (64, 256). For example, the LF is represented as (64, 256, 1), while the HFs take forms such as (32, 64, 8) and (16, 128, 8) with 8 directional channels (Figure 2B).

The selection of the level configuration [8,8] is based

on the following rationale. First, with eight directions, the spatial sizes of the decomposed HFs are extracted as powers of two (e.g., 32×64 , 16×128), maintaining an integer-multiple relationship with the 2×2 MaxPooling operations in the U-Net. Increasing the number of directions to 16 would excessively reduce the spatial size of the components, making meaningful feature extraction infeasible at the bottleneck stage of the four-level encoder. Second, according to Candès and Donoho,²⁷ eight directions in a single-level decomposition are consistent with the theoretical structure, whereas 4 directions are insufficient to capture the diversity of dips. Third, under the 64×256 patch-size condition, an eight-direction decomposition represents a practical balance that simultaneously provides sufficient spatial information and directional separation. The results of the structural compatibility analysis for alternative configurations ([4,4], [16,16]) are discussed in

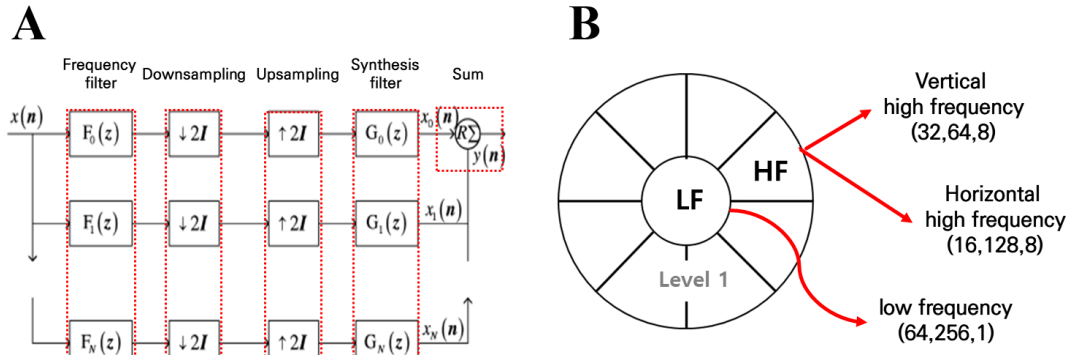


Figure 2. Examples of the UDCT multi-resolution filter bank and frequency tiling, where (A) shows the filter bank structure in which analysis filters $F_0(z)$ through $F_N(z)$ and downsampling constitute the forward transform while upsampling followed by synthesis filters $G_0(z)$ through $G_N(z)$ and band summation are used in the inverse transform, and (B) illustrates the frequency tiling and component configuration for the level setting [8,8]. Abbreviations: HF: High-frequency component; LF: Low-frequency component.

Section 4.

2.3. Component-wise interpolation and inverse transform

According to the theoretical structure of the UDCT, multi-directional HFs are naturally classified into two groups based on the characteristics of the decimation matrices in the frequency plane. The decimation matrices applied to the high-frequency directional subbands are

$$\text{defined in two forms, } D_{J,0} = \text{diag}\left\{2, \frac{2N_j}{3}\right\} \cdot 2^{J-j} \text{ and } D_{J,1} = \text{diag}\left\{\frac{2N_j}{3}, 2\right\} \cdot 2^{J-j}, \text{ indicating that the mostly}$$

horizontal group and the mostly vertical group are arranged on different lattice structures.³⁵ In this study, we directly utilize this 2-lattice structure, defining the LF, the vertically dominant HF as HF_V , and the horizontally dominant HF as HF_H .

Denoting the UDCT components of each group as C_{LF} , C_{HF_V} , and C_{HF_H} , and defining the corresponding reconstruction networks as $f_{LF}(\cdot)$, $f_{HF_V}(\cdot)$, and $f_{HF_H}(\cdot)$, respectively, three independent U-Net networks predict each group component in parallel, and the final interpolation result is generated through inverse transform synthesis (Equation 2).

$$\hat{C}_{LF} = f_{LF}(C_{LF}), \hat{C}_{HF_H} = f_{HF_H}(C_{HF_H}) \quad (2)$$

$$\hat{C}_{HF_V} = f_{HF_V}(C_{HF_V})$$

where, C_{LF} , C_{HF_V} , and C_{HF_H} are the group-wise UDCT components (coefficient maps) obtained from the frequency tiling, and $\hat{C}_{LF}$, $\hat{C}_{HF_V}$, and $\hat{C}_{HF_H}$ denote the interpolation result components predicted by each

network.

The component-wise reconstructed results are synthesized into the original domain through the inverse UDCT operator $C^{-1}(\cdot)$ to generate the final reconstructed section. This synthesis process can be summarized as Equation 3:

$$\hat{X} = C^{-1}(\hat{C}_{LF}, C_{HF_V}, C_{HF_H}) \quad (3)$$

where, $C^{-1}(\cdot)$ is the inverse UDCT operator, which aggregates the predicted components from the three networks according to the original scale-direction decomposition structure to produce the final interpolation result $\hat{X}$. The inverse transform stage combines the independently reconstructed components in accordance with the original scale-direction decomposition structure, ensuring that the information in the frequency-direction region covered by each group is preserved without loss.

2.4. Network architecture

The interpolation model is constructed as a two-dimensional U-Net-based convolutional neural network (Figure 3). The encoder consists of a four-stage downsampling path, where each stage performs 3×3 convolution followed by 2×2 max pooling, with the number of channels increasing as $32 \rightarrow 64 \rightarrow 128 \rightarrow 256$. The bottleneck has 512 channels, and the decoder adopts a symmetric structure to the encoder with skip connections. The activation function is the rectified linear unit, and the output layer uses 1×1 convolution with linear activation. The loss function employed is the mean squared error (MSE).

The same U-Net architecture is applied to all three components (LF, HF_V , HF_H), with only the spatial size

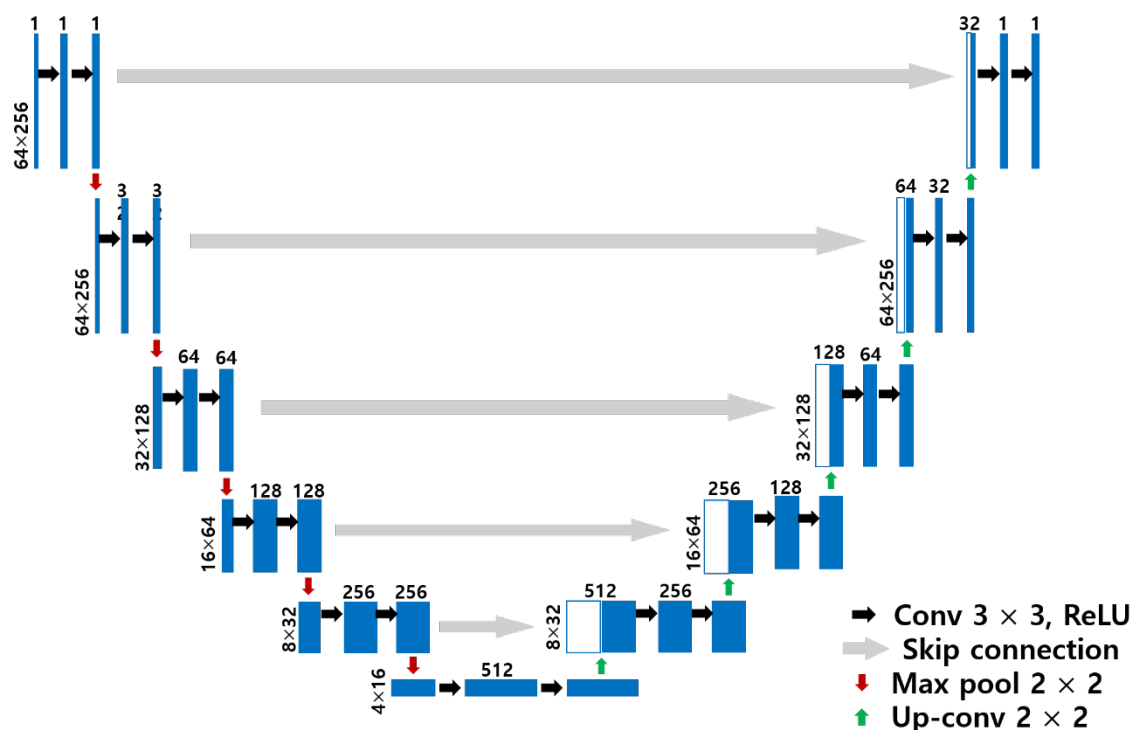


Figure 3. Architecture of the two-dimensional U-Net used in this study
Abbreviation: ReLU: Rectified linear unit.

of the input tensor varying according to the UDCT decomposition results. Training was configured with the Adam optimizer³⁶ (initial learning rate of 1×10^{-3}), a batch size of 128, and a maximum of 500 epochs, with early stopping applied when the validation loss did not improve for 10 consecutive epochs. Weight initialization used He initialization. The experimental environment consisted of an Intel Xeon Silver 4210R CPU (2.40 GHz, 2×10 cores), 125 GB RAM, and an NVIDIA TITAN RTX graphics processing unit (GPU) (24 GB VRAM).

The comparative experiments were designed to verify the contribution of curvelet-based component separation to interpolation performance. (i) F-x interpolation is a conventional frequency-domain interpolation method.¹ (ii) Constrained Diffusion-Driven Deep Image Prior (CDDIP) is a diffusion model-based self-supervised generative interpolation method.¹⁶ (iii) U-Net serves as the baseline model that directly interpolates in the original domain without any transform. (iv) WT-UNet is a hybrid structure consisting of one original-domain U-Net and one wavelet-domain U-Net, where the wavelet-domain model takes the four subbands (low-low, low-high, high-low, high-high) of the two-dimensional discrete wavelet transform concatenated along the channel dimension as input. U-Net-Wide is a parameter-matched baseline that uses a

single U-Net with increased channel widths to match the total parameter count of CT-UNet, without any transform-based decomposition. This baseline is included to verify whether the performance improvement of CT-UNet is attributable to curvelet-based decomposition or simply to increased model capacity. The comparative methods are selected to address distinct questions: F-x interpolation and CDDIP represent conventional and generative paradigms, respectively, situating the proposed method within the broader landscape of seismic interpolation. The standard U-Net serves as the primary ablation baseline, isolating the effect of curvelet-based decomposition from the underlying network architecture. WT-UNet evaluates the effect of transform selection (wavelet vs. curvelet) under a similar component-separation philosophy. U-Net-Wide controls for model capacity to confirm that the observed gains are attributable to the curvelet representation rather than to increased parameter count.

3. Experiments

3.1. Data generation and model training process

For synthetic data, the Narrow-azimuth data from the Society of Exploration Geophysicists (SEG)/European Association of Geoscientists and Engineers (EAGE) Salt Model (receiver spacing 20 m, streamer spacing 40 m)³⁷

was used. For field data, Mobil Amplitude Versus Offset (AVO) Viking Graben Line 12 (1001 shot records, 120 channels, recording time 6 s, sampling interval 4 ms, shot/receiver spacing 25 m)³⁸ was used.

Regular missing patterns were applied to the original data to generate input data with missing rates of 67%, 75%, and 80%, with the complete data prior to missing trace removal serving as ground truth. Specifically, 67% missing corresponds to retaining one trace and removing the following two traces (every three-trace cycle), 75% missing corresponds to retaining one trace and removing the following three traces (every four-trace cycle), and 80% missing corresponds to retaining one trace and removing the following four traces (every five-trace cycle). A patch-based training approach was adopted, with synthetic data configured as 64×256 patches (250 patches in the test set) and field data as $64 \times 1,024$ patches (505 patches in the test set). The difference in patch size reflects the spatial extent and event scale of each dataset, while the UDCT decomposition ratio and component structure remain identical under the level configuration [8,8]. Dataset splitting was performed by first randomly assigning all shot locations to training (70%), validation (20%), and test (10%) sets, and then extracting patches from the shots assigned to each set. This prevented data leakage where patches extracted from the same shot could be included in both the training and test sets simultaneously, and ensured the reliability of generalization performance evaluation by distributing shots from different spatial locations across each set.

A separate CT-UNet system (three component-wise U-Nets) was trained independently for each combination of dataset and missing rate. Missing traces were 0-filled prior to UDCT decomposition, and no binary sampling mask was provided as an auxiliary input to the networks. For the [8,8] level configuration, the real and imaginary parts of the complex-valued high-frequency coefficients were concatenated along the channel dimension, resulting in 16 input channels (eight directions $\times$ 2) for both the HF_v and HF_h networks, while the LF network received a single-channel real-valued input.

Irregular missing scenarios were generated by randomly removing 50% and 70% of all traces from the synthetic data, while maintaining all other conditions identical to those in the regular missing experiments. The UDCT forward transform was applied to decompose each dataset into three groups: LF, HF_v , and HF_h (Figure 4). The HFs are expressed in complex form, and in this study, the real and imaginary parts were separated and concatenated along the channel dimension to construct the U-Net input tensor. This approach maintains continuous real-valued inputs

suitable for CNN training while avoiding the 2π periodic discontinuity (phase wrapping) problem that arises in the alternative magnitude/phase separation approach. The validity of the real/imaginary separation method has been verified in image processing and MRI reconstruction.^{39,40} However, a limitation is that the phase coherence between the two channels is only implicitly learned, and the introduction of complex-valued convolution or a phase consistency loss is left for future work.

The three decomposed components were trained and predicted independently using U-Net models, and during prediction, the same UDCT decomposition followed by component-wise prediction and inverse transform synthesis was applied.

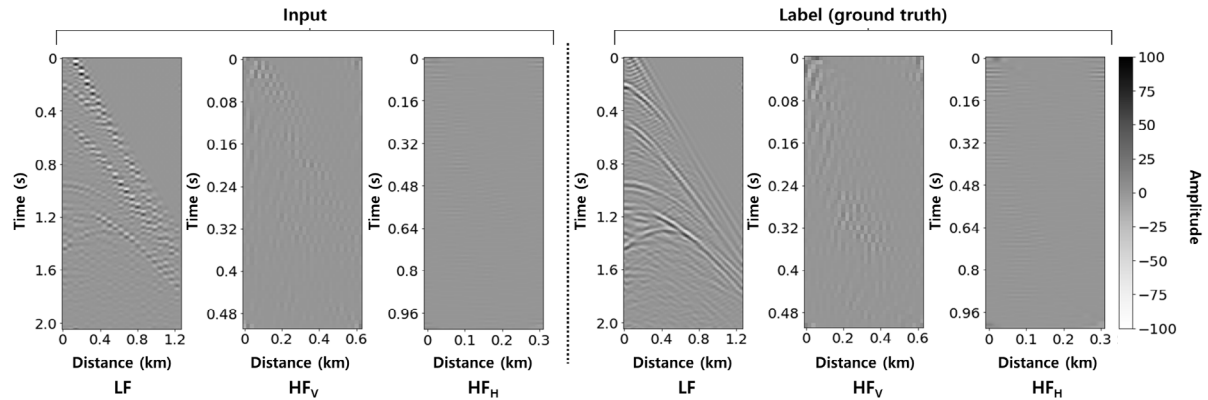
3.2. Results of synthetic data

Figure 5 shows the reconstruction results for missing rates of 67%, 75%, and 80%; Figure 6 presents the enlarged regions of interest (ROI); and Figure 7 compares the f-k spectra.

The f-x interpolation exhibits ringing and ripple residuals near the missing traces under varying dip conditions. CDDIP shows weak fidelity to the observed data, producing results that approximate plausible structure generation rather than precisely reproducing the phase and amplitude of the ground truth. U-Net performs stable interpolation overall; however, as the missing rate increases, event continuity in the ROI weakens, and a smoothing tendency is observed. U-Net-Wide, despite having a comparable total parameter count to CT-UNet, exhibited reconstruction quality similar to or slightly lower than U-Net across all missing rates, indicating that simply increasing model capacity without transform-based decomposition does not yield the improvement observed in CT-UNet. WT-UNet tends to over-amplify weak components in non-dominant directions. CT-UNet most naturally reproduces the continuity and curvature variations of curvilinear events in the ROI. In particular, in regions containing conflicting-dip events (i.e., reflections with dip opposite to the dominant event direction), U-Net and WT-UNet exhibit noticeable reconstruction degradation, while CT-UNet preserves these events with high fidelity — a pattern that is quantitatively supported by the ROI-based analysis below. In the f-k spectrum, CT-UNet maintains an energy distribution closest to the ground truth while effectively suppressing aliasing components.

Quantitative evaluation (Tables 1 and 2) shows that CT-UNet achieved the highest peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) across all missing rates. Paired *t*-test results indicate

A



B

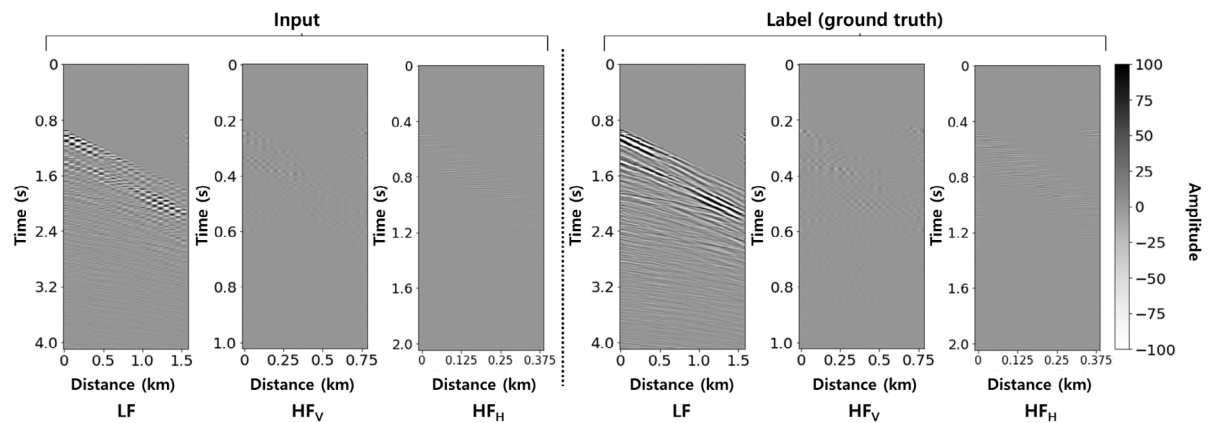


Figure 4. Example of UDCT decomposition for synthetic data (A) and field data (B). For each row, the left three panels show the decomposed components (LF, HF_V, HF_H) of the input with missing traces, and the right three panels show those of the ground truth. The dashed line separates the input and ground truth groups.

Abbreviations: HF: High-frequency component; LF: Low-frequency component.

that the PSNR difference between CT-UNet and U-Net was not statistically significant at 67% ($p = 0.139$) and 75% ($p = 0.878$), but was significant at 80% ($p = 0.039$). In terms of SSIM, consistent improvement was observed across all conditions. CT-UNet showed highly significant improvement over f-x interpolation, CDDIP, and WT-UNet across all conditions (all $p < 0.001$).

To further examine the reconstruction quality in the region containing curvature-varying events, the normalized correlation coefficient (CC) and SSIM were additionally computed over the ROI (blue boxes in Figure 5). Tables 3 and 4 present the ROI-based CC and SSIM, respectively. CT-UNet achieved the highest values across all missing rates, with a larger margin over U-Net than in the global metrics. This indicates that the advantage of curvelet-based decomposition is particularly evident in regions where complex dip variations are concentrated.

3.3. Results of field data

Figure 8 shows the reconstruction results of field data, and Figure 9 presents the corresponding f-k spectra. Field data contain noise, non-stationarity, and complex reflection patterns, requiring simultaneous event continuity and spectral consistency. The overall performance trend is similar to that of the synthetic data, with CT-UNet reproducing the continuity of major reflection events most naturally and maintaining the energy distribution closest to the ground truth in the f-k spectra.

Quantitative evaluation (Tables 1 and 2) shows that CT-UNet achieved the highest PSNR across all missing rates. At 67%, the difference from U-Net was only 0.01 dB, indicating equivalent performance ($p = 0.918$), while the gap widened to 0.60 dB at 75% ($p = 5.63 \times 10^{-8}$) and 1.61 dB at 80% ($p = 1.48 \times 10^{-66}$). U-Net-Wide achieved PSNR values of 35.29 dB (67%), 34.08 dB (75%), and 31.22

Table 1. Peak signal-to-noise results

Data	Missing ratio (%)	F-x interpolation	CDDIP	U-Net	U-Net-Wide	WT-UNet	CT-UNet
Synthetic data	67	23.64 ($\pm$ 0.84)	22.63 ($\pm$ 1.21)	31.96 ($\pm$ 1.43)	31.52 ($\pm$ 1.37)	28.82 ($\pm$ 1.35)	32.14 ($\pm$ 1.28) ^{n.s.}
	75	22.91 ($\pm$ 0.91)	22.23 ($\pm$ 1.18)	30.18 ($\pm$ 1.51)	30.01 ($\pm$ 1.45)	29.42 ($\pm$ 1.47)	30.20 ($\pm$ 1.39) ^{n.s.}
	80	22.06 ($\pm$ 1.02)	22.04 ($\pm$ 1.09)	29.42 ($\pm$ 1.58)	29.08 ($\pm$ 1.49)	28.73 ($\pm$ 1.52)	29.70 ($\pm$ 1.44)*
	Irregular missing traces 50%	24.76 ($\pm$ 0.73)	27.24 ($\pm$ 1.12)	32.61 ($\pm$ 1.45)	32.33 ($\pm$ 1.29)	31.92 ($\pm$ 1.47)	32.76 ($\pm$ 1.11) ^{n.s.}
	Irregular missing traces 70%	23.48 ($\pm$ 0.89)	24.65 ($\pm$ 1.15)	30.32 ($\pm$ 1.50)	30.46 ($\pm$ 1.44)	30.56 ($\pm$ 1.41)	30.98 ($\pm$ 1.33)*
Field data	67	27.17 ($\pm$ 1.12)	26.31 ($\pm$ 1.38)	36.29 ($\pm$ 1.61)	35.29 ($\pm$ 1.54)	34.84 ($\pm$ 1.55)	36.30 ($\pm$ 1.48) ^{n.s.}
	75	26.67 ($\pm$ 1.08)	23.93 ($\pm$ 1.72)	33.86 ($\pm$ 1.83)	34.08 ($\pm$ 1.69)	34.00 ($\pm$ 1.71)	34.46 ($\pm$ 1.65)***
	80	25.05 ($\pm$ 0.97)	23.85 ($\pm$ 1.31)	29.88 ($\pm$ 1.42)	31.22 ($\pm$ 1.36)	31.18 ($\pm$ 1.38)	31.49 ($\pm$ 1.33)***

Notes: For paired *t*-test results (CT-UNet vs. U-Net), * $p < 0.05$; *** $p < 0.001$; Values represent the mean ($\pm$ standard deviation) computed over all test patches (250 patches for synthetic data, 505 patches for field data); All metrics were computed on the original amplitude scale after denormalization; CDDIP is a self-supervised method that does not use training data and is included to provide insight into the performance characteristics across different reconstruction paradigms rather than for direct comparison with supervised methods.

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; n.s.: Not significant; WT: Wavelet transform.

Table 2. Structural similarity index measure results

Data	Missing ratio (%)	F-x interpolation	CDDIP	U-Net	U-Net-Wide	WT-UNet	CT-UNet
Synthetic data	67	0.71 ($\pm$ 0.03)	0.58 ($\pm$ 0.15)	0.89 ($\pm$ 0.03)	0.89 ($\pm$ 0.03)	0.82 ($\pm$ 0.04)	0.92 ($\pm$ 0.02)
	75	0.62 ($\pm$ 0.03)	0.54 ($\pm$ 0.05)	0.86 ($\pm$ 0.03)	0.83 ($\pm$ 0.04)	0.81 ($\pm$ 0.04)	0.87 ($\pm$ 0.03)
	80	0.56 ($\pm$ 0.04)	0.51 ($\pm$ 0.05)	0.84 ($\pm$ 0.04)	0.83 ($\pm$ 0.04)	0.82 ($\pm$ 0.04)	0.85 ($\pm$ 0.03)
	Irregular missing traces 50%	0.74 ($\pm$ 0.06)	0.76 ($\pm$ 0.11)	0.90 ($\pm$ 0.06)	0.90 ($\pm$ 0.05)	0.90 ($\pm$ 0.06)	0.91 ($\pm$ 0.05)
	Irregular missing traces 70%	0.61 ($\pm$ 0.06)	0.62 ($\pm$ 0.10)	0.83 ($\pm$ 0.05)	0.85 ($\pm$ 0.05)	0.86 ($\pm$ 0.05)	0.87 ($\pm$ 0.04)
Field data	67	0.87 ($\pm$ 0.02)	0.82 ($\pm$ 0.05)	0.98 ($\pm$ 0.01)	0.97 ($\pm$ 0.01)	0.98 ($\pm$ 0.01)	0.98 ($\pm$ 0.01)
	75	0.82 ($\pm$ 0.02)	0.78 ($\pm$ 0.02)	0.97 ($\pm$ 0.01)	0.97 ($\pm$ 0.01)	0.97 ($\pm$ 0.01)	0.98 ($\pm$ 0.01)
	80	0.79 ($\pm$ 0.02)	0.75 ($\pm$ 0.02)	0.96 ($\pm$ 0.01)	0.96 ($\pm$ 0.01)	0.96 ($\pm$ 0.01)	0.97 ($\pm$ 0.01)

Note: Values represent the mean ($\pm$ standard deviation) over 250 patches (synthetic data) and 505 patches (field data).

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; WT: Wavelet transform.

Table 3. Region of interest-based normalized correlation coefficient for synthetic data

Missing ratio (%)	F-x interpolation	CDDIP	U-Net	U-Net-Wide	WT-UNet	CT-UNet
67	0.42 ($\pm$ 0.20)	0.25 ($\pm$ 0.14)	0.88 ($\pm$ 0.06)	0.89 ($\pm$ 0.06)	0.82 ($\pm$ 0.09)	0.90 ($\pm$ 0.05)
75	0.28 ($\pm$ 0.16)	0.22 ($\pm$ 0.13)	0.80 ($\pm$ 0.10)	0.81 ($\pm$ 0.10)	0.75 ($\pm$ 0.11)	0.82 ($\pm$ 0.09)
80	0.21 ($\pm$ 0.13)	0.15 ($\pm$ 0.11)	0.77 ($\pm$ 0.11)	0.78 ($\pm$ 0.10)	0.73 ($\pm$ 0.12)	0.79 ($\pm$ 0.10)

Note: Values were computed over the region of interest (blue boxes in Figure 5) and represent the mean ($\pm$ standard deviation) over 250 test patches.

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; WT: Wavelet transform.

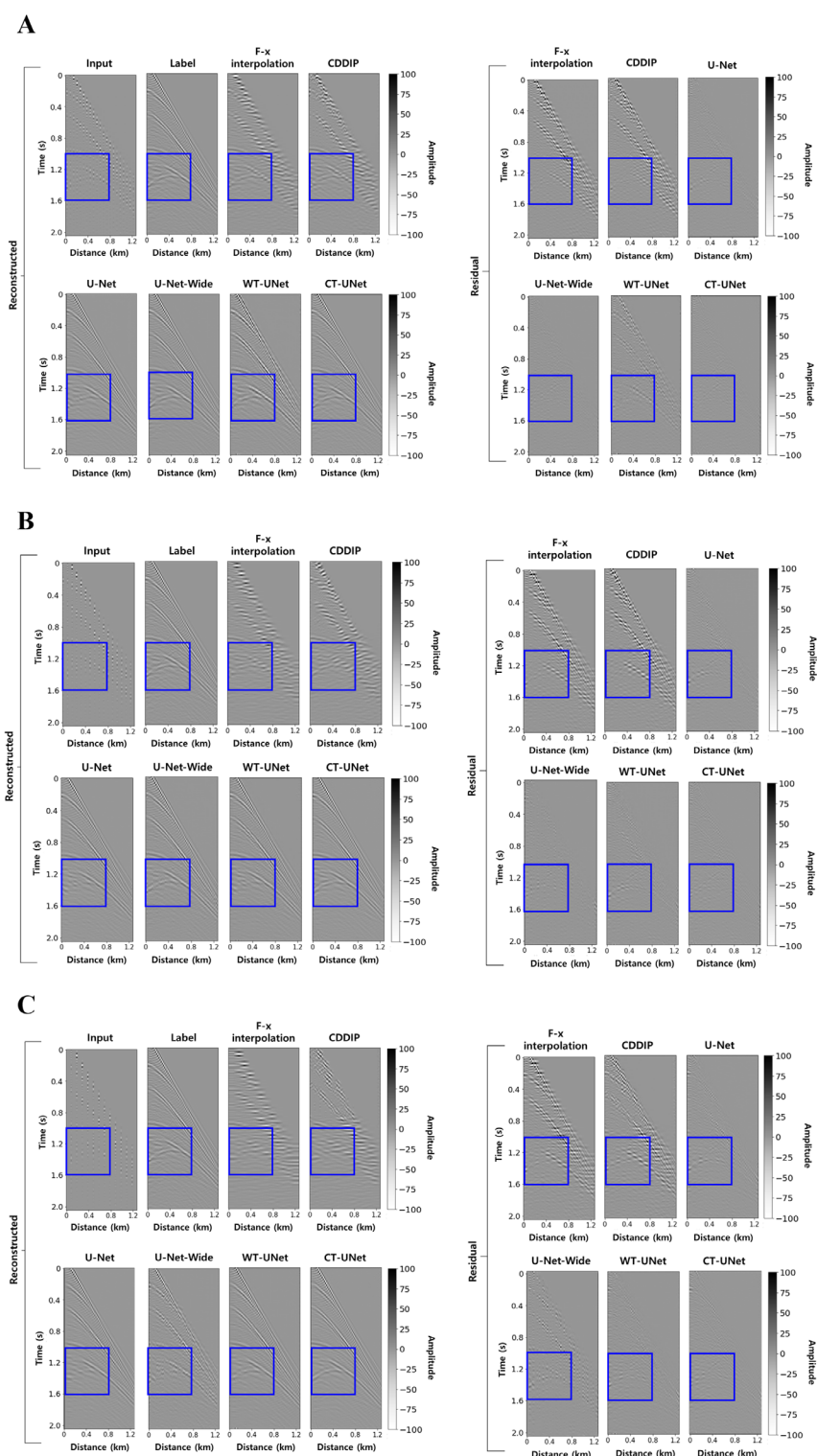


Figure 5. Reconstruction results and residuals of synthetic data under regular missing rates of (A) 67%, (B) 75%, and (C) 80%. The top row of each panel shows the reconstructed results, and the bottom row shows the residuals. From left to right, the reconstructed results are the input, ground truth, f-x interpolation, Constrained Diffusion-Driven Image Prior (CDDIP), U-Net, U-Net-Wide, wavelet transform (WT)-UNet, and curvelet transform (CT)-UNet. The residual panels, from left to right, correspond to f-x interpolation, CDDIP, U-Net, U-Net-Wide, WT-UNet, and CT-UNet.

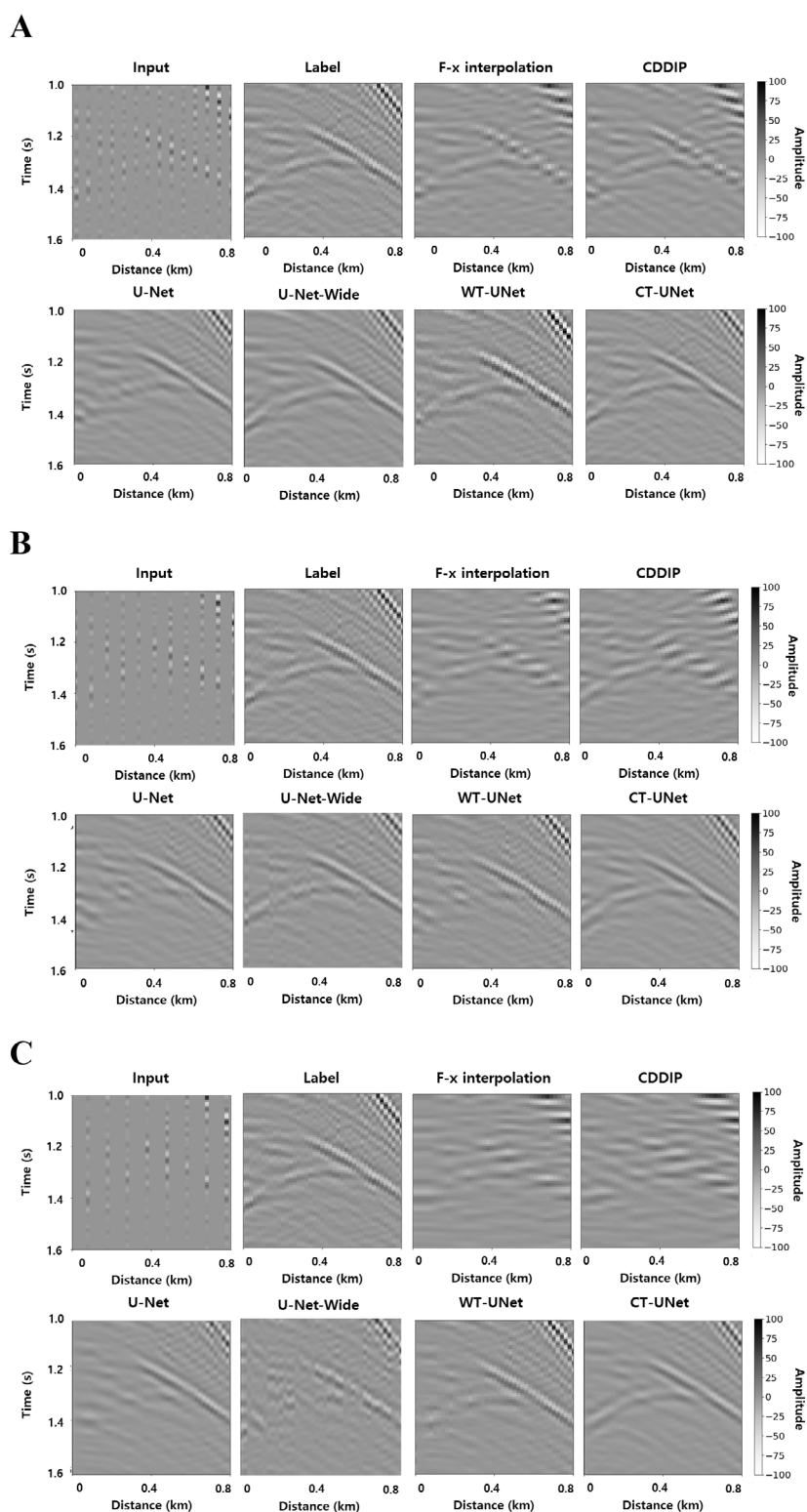


Figure 6. Enlarged views of the region of interest (blue boxes in [Figure 5](#)) for synthetic data under regular missing rates of (A) 67%, (B) 75%, and (C) 80%. From left to right, the top row shows the input, ground truth, f-x interpolation, and Constrained Diffusion-Driven Deep Image Prior results, while the bottom row shows the U-Net, U-Net-Wide, wavelet transform-UNet, and curvelet transform-UNet results.

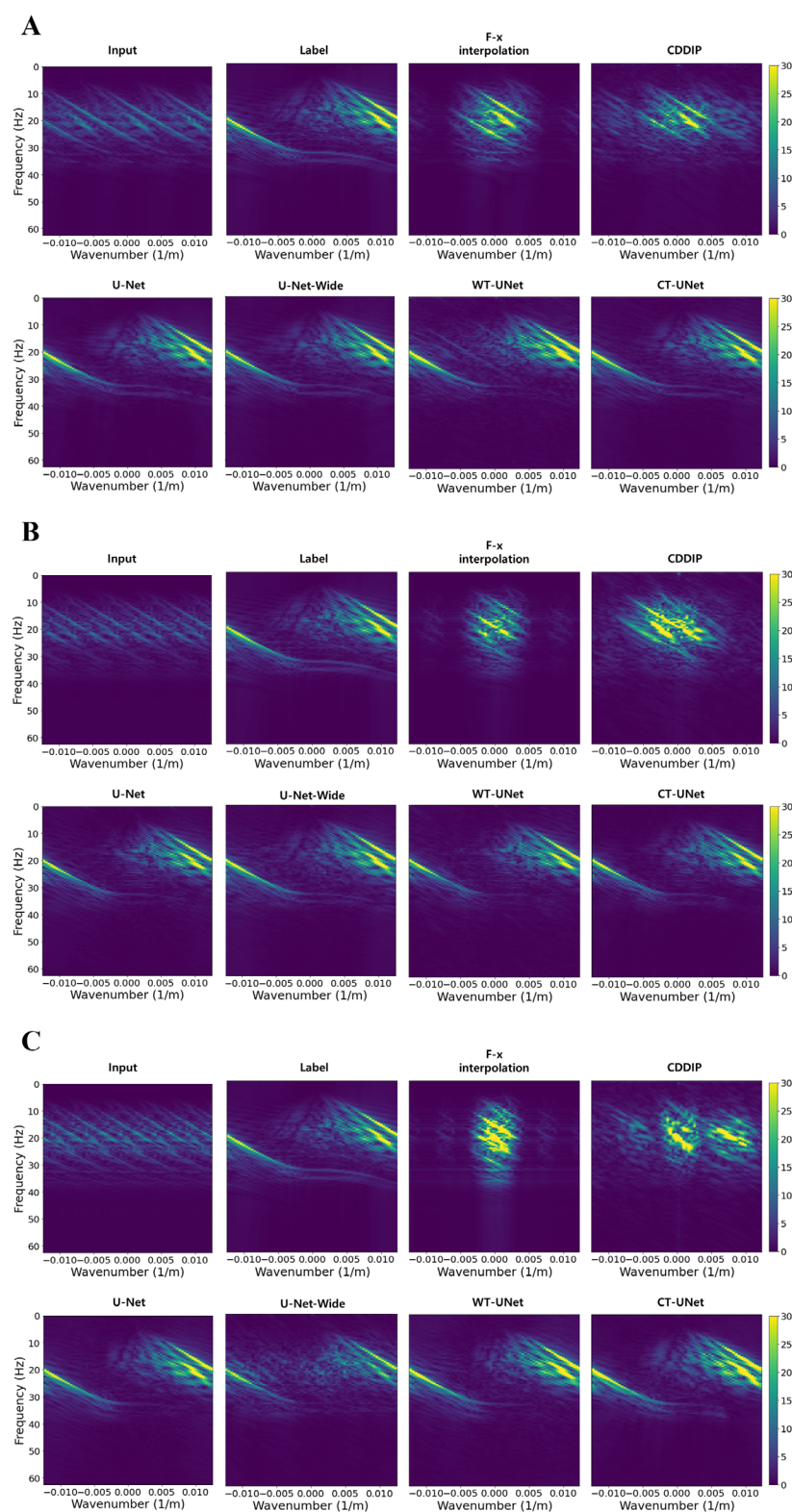


Figure 7. The f - k spectra corresponding to the reconstructed results in Figure 5 under regular missing rates of (A) 67%, (B) 75%, and (C) 80%. From left to right, the top row shows the input, ground truth, f - x interpolation, and Constrained Diffusion-Driven Deep Image Prior results, while the bottom row shows the U-Net, U-Net-Wide, wavelet transform-UNet, and curvelet transform-UNet results.

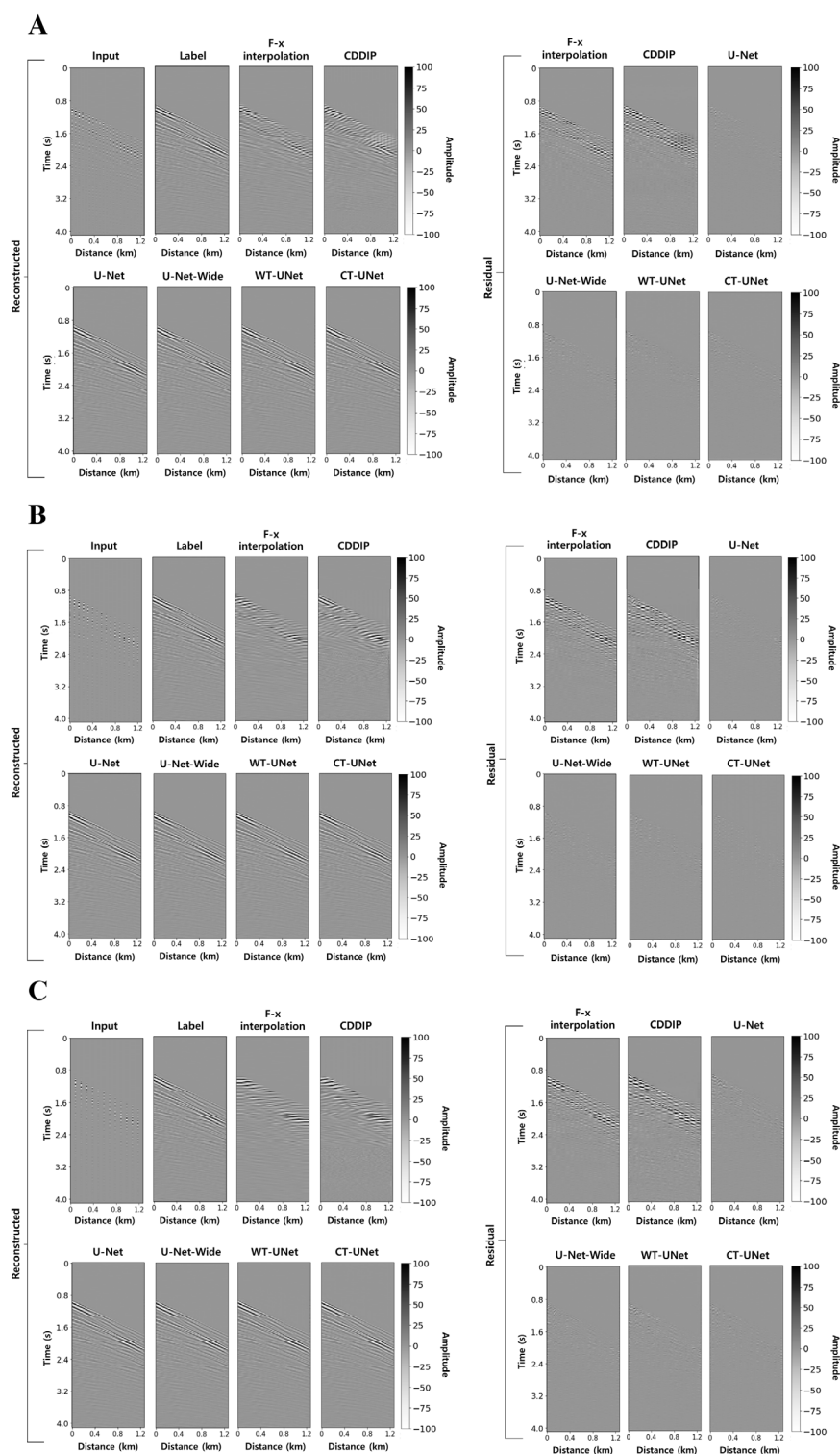


Figure 8. Reconstruction results and residuals of field data under regular missing rates of (A) 67%, (B) 75%, and (C) 80%. For each missing-rate case, the reconstructed results are shown in the first two rows, and the corresponding residuals are shown in the last two rows. From left to right, the reconstructed results are the input, ground truth, f -x interpolation, Constrained Diffusion-Driven Image Prior (CDDIP), U-Net, U-Net-Wide, wavelet transform (WT)-UNet, and curvelet transform (CT)-UNet. The residual panels, from left to right, correspond to f -x interpolation, CDDIP, U-Net, U-Net-Wide, WT-UNet, and CT-UNet.

dB (80%). At 75% and 80%, U-Net-Wide outperformed U-Net but remained below CT-UNet, confirming that the curvelet-based decomposition provides an additional benefit beyond what increased model capacity alone can achieve. CT-UNet showed highly significant improvement over WT-UNet, f-x interpolation, and CDDIP across all conditions (all $p < 0.001$).

To examine the frequency-dependent reconstruction characteristics, band-limited PSNR and SSIM were computed for three frequency bands (Tables 5 and 6). In both synthetic and field data, CT-UNet achieved the highest PSNR and SSIM across all three bands. The advantage over U-Net was most pronounced in the mid-frequency band, where CT-UNet achieved 31.0 dB on synthetic data and 33.5 dB on field data, compared with 30.6 dB and 32.4 dB, respectively, for U-Net. U-Net-Wide showed intermediate performance between U-Net and CT-UNet in most bands, confirming that increased model capacity alone does not fully account for CT-UNet's improvement.

To provide a quantitative comparison of the frequency-dependent reconstruction accuracy, the spectral residual energy profiles under the 80% regular missing condition are presented in Figure 10. For both synthetic and field data, CT-UNet exhibits the lowest residual energy across the entire frequency range, indicating the most balanced reconstruction among all compared methods. Notably, U-Net-Wide, despite having a comparable total parameter count to CT-UNet, shows higher residual energy than CT-UNet, particularly in the mid-frequency range (normalized frequency 0.1–0.3), confirming that the improvement is attributable to curvelet-based decomposition rather than increased model capacity.

3.4. Results of irregular missing data

To examine the behavior of CT-UNet beyond regular undersampling, two irregular (random) missing scenarios at 50% and 70% missing rates were tested on synthetic data (Figures 11 and 12). Unlike regular missing, no periodic aliasing pattern forms in the f-k domain, and the damaged components spread incoherently across the spectrum.

In the spatial domain, f-x interpolation and CDDIP exhibit relatively large residuals along event structures across both missing rates, whereas U-Net, U-Net-Wide, WT-UNet, and CT-UNet stably reconstruct the major reflection events. The visual differences among the learning-based methods are smaller compared to the regular missing conditions. In the f-k spectra, U-Net, U-Net-Wide, and CT-UNet show nearly identical levels of energy distribution recovery at both missing rates.

Quantitative evaluation (Tables 1 and 2) shows that at 50% irregular missing, CT-UNet achieved the highest

PSNR of 32.76 ± 1.11 dB and SSIM of 0.91 ± 0.05 , with U-Net at 32.61 ± 1.45 dB and U-Net-Wide at 32.33 ± 1.29 dB. At the more challenging 70% irregular missing rate, CT-UNet maintained the highest PSNR of 30.98 ± 1.33 dB ($p < 0.05$) and SSIM of 0.87 ± 0.04 , outperforming U-Net (30.32 ± 1.50 dB; 0.83 ± 0.05), U-Net-Wide (30.46 ± 1.44 dB; 0.85 ± 0.05), and WT-UNet (30.56 ± 1.41 dB; 0.86 ± 0.05). These results indicate that CT-UNet remains effective under a more challenging random, irregular, and missing condition.

Across both tested random irregular missing rates, CT-UNet maintained the highest PSNR and SSIM among the compared methods. These results suggest that curvelet-based component separation can remain effective beyond periodic regular undersampling, although further evaluation is required for more realistic acquisition-gap scenarios, such as blocky, mixed regular-irregular, consecutive, and spatially varying missing patterns.

To control for potential spatial autocorrelation between patches, paired *t*-tests were additionally conducted on shot-aggregated PSNR and SSIM values ($n = 50$ for synthetic and $n = 101$ for field data). Results for CT-UNet versus U-Net are summarized in Table 7. CT-UNet shows consistently non-negative differences across all tested conditions, with statistically significant PSNR advantages at synthetic 80% (+0.28 dB, $p < 0.05$), synthetic 70% irregular missing (+0.66 dB, $p < 0.05$), field 75% (+0.60 dB, $p < 0.001$), and field 80% (+1.61 dB, $p < 0.001$). For SSIM, CT-UNet maintains consistent advantages across all conditions. This pattern indicates that CT-UNet's improvement over U-Net is most pronounced under high regular missing rates and challenging irregular missing scenarios, consistent with the patch-level results in Tables 1 and 2.

4. Discussion

This section discusses the experimental results in terms of (i) the contribution of curvelet-based component separation, (ii) performance characteristics according to missing type and level configuration, (iii) design differences among comparison methods, (iv) computational cost, and (v) limitations and future work.

4.1. Contribution of curvelet-based component separation

The core of CT-UNet lies in separating components in the curvelet domain and performing learning tailored to the statistical characteristics of each component. The curvelet transform provides a near-optimal sparse representation for curvilinear discontinuities, and according to Candès and Donoho²⁷ and Candès *et al.*,²⁸ the approximation error decay rate for C^2 discontinuities is $O(k^{-2}(\log k)^3)$, which

Table 4. ROI-based structural similarity index measure for synthetic data

Missing ratio (%)	F-x interpolation	CDDIP	U-Net	U-Net-Wide	WT-UNet	CT-UNet
67	0.45 ($\pm$ 0.11)	0.20 ($\pm$ 0.13)	0.83 ($\pm$ 0.10)	0.84 ($\pm$ 0.09)	0.77 ($\pm$ 0.15)	0.86 ($\pm$ 0.09)
75	0.31 ($\pm$ 0.09)	0.20 ($\pm$ 0.10)	0.70 ($\pm$ 0.14)	0.69 ($\pm$ 0.14)	0.62 ($\pm$ 0.18)	0.73 ($\pm$ 0.13)
80	0.23 ($\pm$ 0.07)	0.15 ($\pm$ 0.08)	0.65 ($\pm$ 0.15)	0.63 ($\pm$ 0.16)	0.59 ($\pm$ 0.17)	0.68 ($\pm$ 0.14)

Notes: Values were computed over the ROI (blue boxes in Figure 5); mean ($\pm$ standard deviation) over 250 test patches.

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; ROI: Region of interest; WT: Wavelet transform.

Table 5. Band-limited PSNR (dB) and SSIM for synthetic data under 80% regular missing

Metric	F-x interpolation	CDDIP	U-Net	U-Net-Wide	WT-UNet	CT-UNet
Low PSNR (dB)/SSIM	27.6/0.78	27.1/0.73	34.6/0.93	34.8/0.92	33.3/0.90	35.0/0.94
Mid PSNR (dB)/SSIM	24.8/0.65	24.1/0.60	30.6/0.87	30.8/0.86	29.7/0.84	31.0/0.88
High PSNR (dB)/SSIM	22.9/0.54	22.2/0.49	28.8/0.80	29.0/0.79	27.9/0.77	29.2/0.81

Note: Metrics were computed separately for low (0–12.5 Hz), mid (12.5–37.5 Hz), and high (37.5–62.5 Hz) frequency bands using Butterworth bandpass filtering (dt = 0.008 s, Nyquist frequency = 62.5 Hz).

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index measure; WT: Wavelet transform.

Table 6. Band-limited PSNR (dB) and SSIM for field data under 80% regular missing

Metric	F-x interpolation	CDDIP	U-Net	U-Net-Wide	WT-UNet	CT-UNet
Low PSNR (dB)/SSIM	29.5/0.86	28.8/0.82	34.9/0.98	35.9/0.98	35.7/0.98	36.1/0.98
Mid PSNR(dB)/SSIM	27.6/0.82	26.9/0.78	32.4/0.96	33.2/0.97	33.0/0.97	33.5/0.97
High PSNR (dB)/SSIM	24.2/0.78	23.4/0.74	29.7/0.95	30.6/0.96	30.4/0.96	30.8/0.96

Note: Band definitions are identical to Table 5.

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index measure; WT: Wavelet transform.

Table 7. Shot-level comparison between CT-UNet and U-Net

Data	Condition	n_{shots}	Δ PSNR (dB)	Δ SSIM	Significance
Synthetic	Regular 67%	50	+0.18	+0.03	n.s.
Synthetic	Regular 75%	50	+0.02	+0.01	n.s.
Synthetic	Regular 80%	50	+0.28	+0.01	*
Synthetic	Irregular 50%	50	+0.15	+0.01	n.s.
Synthetic	Irregular 70%	50	+0.66	+0.04	*
Field	Regular 67%	101	+0.01	+0.00	n.s.
Field	Regular 75%	101	+0.60	+0.01	***
Field	Regular 80%	101	+1.61	+0.01	***

Notes: PSNR and SSIM were aggregated at the shot level ($n = 50$ for synthetic, $n = 101$ for field) to reduce spatial dependence among neighboring patches; $\Delta = \text{mean}(\text{CT-UNet}) - \text{mean}(\text{U-Net})$; positive values indicate CT-UNet advantage; * $p < 0.05$; *** $p < 0.001$.

Abbreviations: CT: Curvelet transform; n.s.: Not significant; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index measure.

challenges for conventional interpolation methods because of their opposing dip orientations relative to the dominant reflections. The ROI-based analysis (Tables 3 and 4) quantitatively confirms this, with CT-UNet exhibiting a more pronounced margin over U-Net in regions containing such events compared to the global metrics.

The band-limited analysis (Tables 5 and 6) quantitatively confirms that CT-UNet reduces reconstruction error more uniformly across frequency bands compared to all other methods, with consistent superiority in the mid-frequency band where complex dip components are concentrated.

The spectral residual energy analysis (Figure 10) provides further evidence for this conclusion. In the synthetic 80% condition, CT-UNet maintains the lowest residual energy across all frequencies, whereas U-Net and U-Net-Wide concentrate residuals in the mid-to-high frequency range. This pattern is consistent with the theoretical prediction that curvelet-based decomposition reduces the complexity of the reconstruction problem for each component-wise network, enabling more balanced recovery across the full bandwidth. The same trend is reproduced in the field data, where CT-UNet achieves the lowest residual energy despite noise and non-stationarity, reinforcing the generalizability of the observed benefit.

4.2. Performance characteristics according to missing type

In the regular case, aliasing components exhibit a periodic structure in the f - k domain, allowing the directional decomposition of curvelets to effectively separate signal and aliasing components. Under irregular missing, the damaged components spread incoherently across the f - k domain rather than forming periodic aliasing patterns. Despite this loss of structured aliasing, CT-UNet maintained the highest PSNR and SSIM among the compared methods at both 50% and 70% random irregular missing rates. These results suggest that curvelet-based separation remains beneficial even when the missing pattern is not periodic. However, because only random irregular missing was examined, further evaluation is required for more realistic acquisition-gap scenarios, including blocky, mixed regular-irregular, consecutive, and spatially varying missing patterns.

4.3. Influence of uniform discrete curvelet transform level configuration

With the [4,4] configuration, the HFs produced only two types of spatial sizes, making 3-group classification impossible. With the [16,16] configuration, the spatial size of the HF_H component was extracted as 7×128 , which was incompatible with the 2×2 MaxPooling operations of the

U-Net. These findings confirm the theoretical rationale presented in Section 2.2 and demonstrate that [8,8] is the only practical configuration that can fully constitute the CT-UNet framework under the combination of patch size (64×256) and a four-stage U-Net. The effects of asymmetric configurations (e.g., [8,4], [4,8]) and increased patch sizes are set as future tasks.

4.4. Design differences among comparison methods

The structural difference between WT-UNet and CT-UNet stems from the inherent decomposition characteristics of each transform. For wavelets, it is natural to concatenate four subbands along the channel dimension and feed them into a single network, whereas the UDCT produces components with different lattice structures for each direction, making group-wise independent networks consistent with the transform structure. Forcing an identical framework would distort the inherent advantages of each transform; therefore, this comparison is interpreted as evaluating the performance when each transform is applied in its optimal structure.

Given that CT-UNet employs three independent U-Nets and thus has approximately three times the total parameters of the baseline U-Net, the question may arise whether the performance improvement is attributable to curvelet decomposition or simply to increased model capacity. To directly address this concern, U-Net-Wide was included as a capacity-matched control: a single U-Net with channel widths scaled to match CT-UNet's total parameter count ($\sim 23M$) without applying any transform. U-Net-Wide achieved results comparable to, and in some conditions slightly better than, the standard U-Net, but it remained below CT-UNet in the key high-missing-rate cases. This indicates that increasing model capacity alone is insufficient to reproduce the performance of CT-UNet. The spectral residual energy profiles further support this interpretation, showing that U-Net-Wide does not match CT-UNet despite having a comparable parameter count.

Curvelet-based sparsity-promoting reconstruction methods (e.g., curvelet-POCS, L1-regularized inversion) utilize the same transform domain but follow an iterative-optimization paradigm rather than data-driven learning. The comparative experiments in this study are deliberately scoped to evaluate the effect of transform selection within the learning-based interpolation paradigm, where all compared methods (U-Net, U-Net-Wide, WT-UNet, CT-UNet) share a supervised learning framework. This design provides a controlled comparison for assessing the contribution of curvelet-domain decomposition within the learning-based interpolation framework. Comparison with optimization-based curvelet methods addresses a

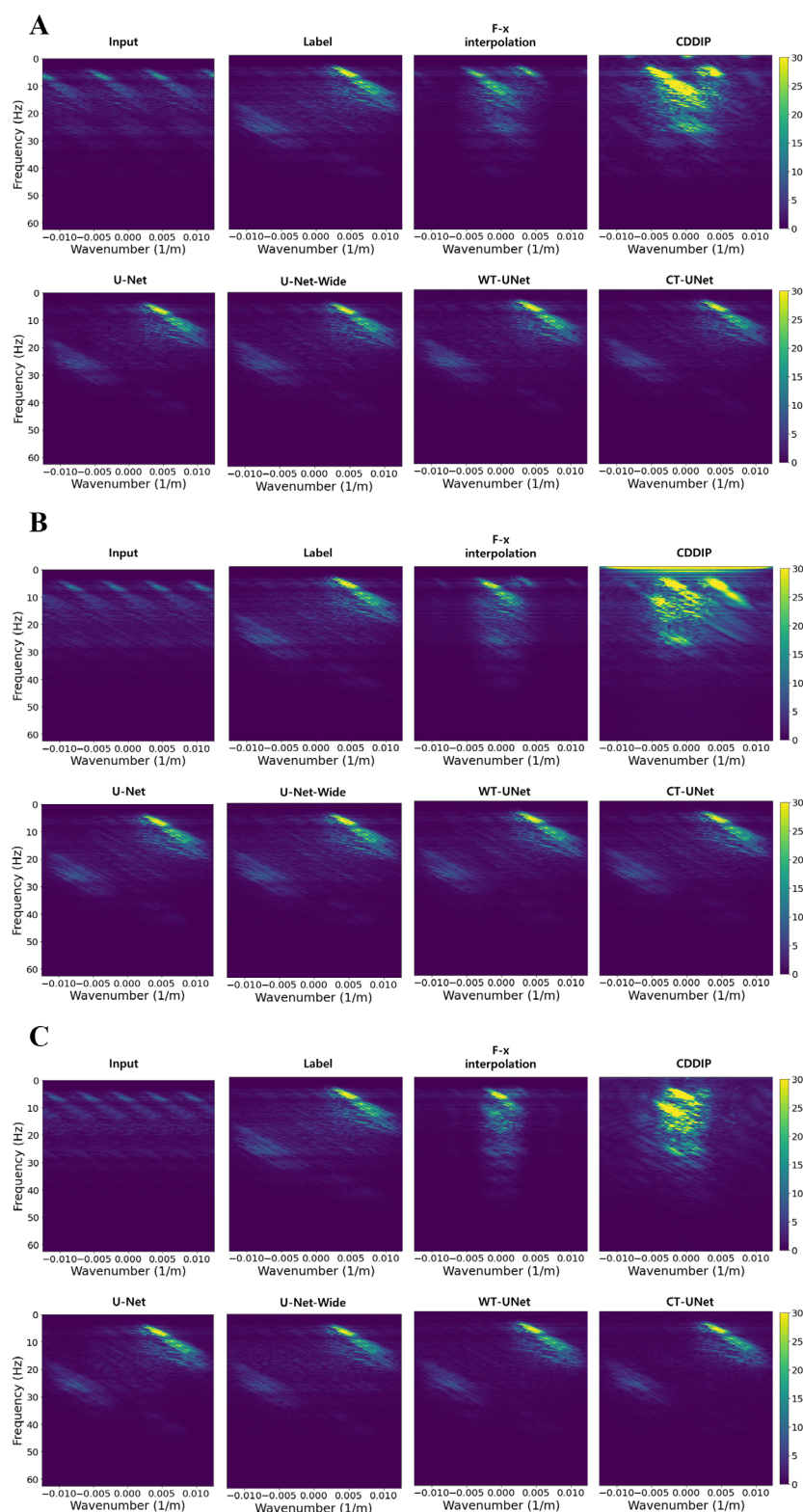


Figure 9. The f-k spectra corresponding to the reconstructed results in Figure 8 under regular missing rates of (A) 67%, (B) 75%, and (C) 80%. From left to right, the top row shows the input, ground truth, f-x interpolation, and Constrained Diffusion-Driven Deep Image Prior results, while the bottom row shows the U-Net, U-Net-Wide, wavelet transform-UNet, and curvelet transform-UNet results.

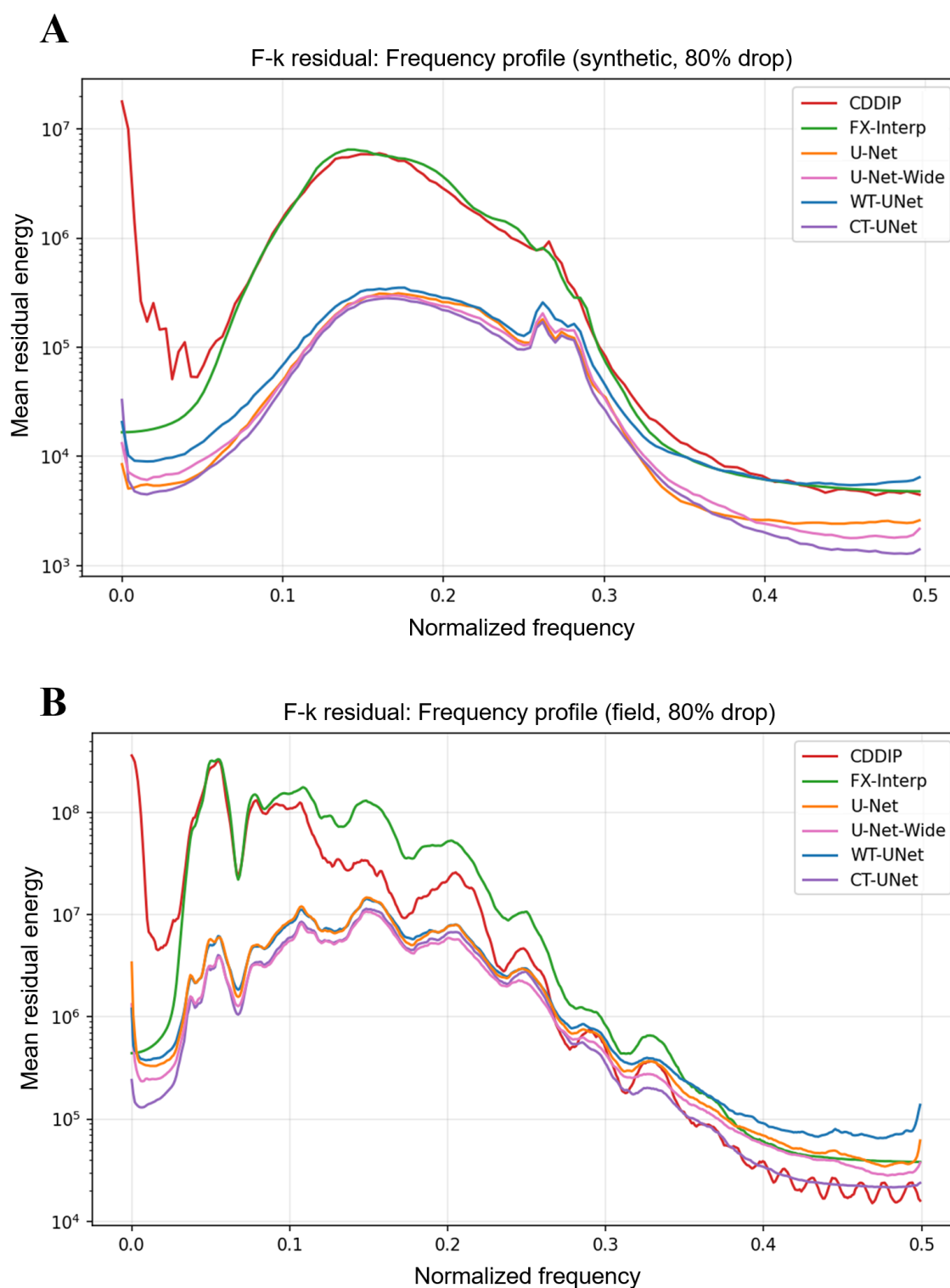


Figure 10. Spectral residual energy profiles under 80% regular missing for (A) synthetic data and (B) field data. Each curve represents the mean residual energy in the f-k domain, computed by averaging the squared magnitude of the residual spectrum along the wavenumber axis at each frequency. Lower values indicate a more accurate reconstruction.

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; WT: Wavelet transform.

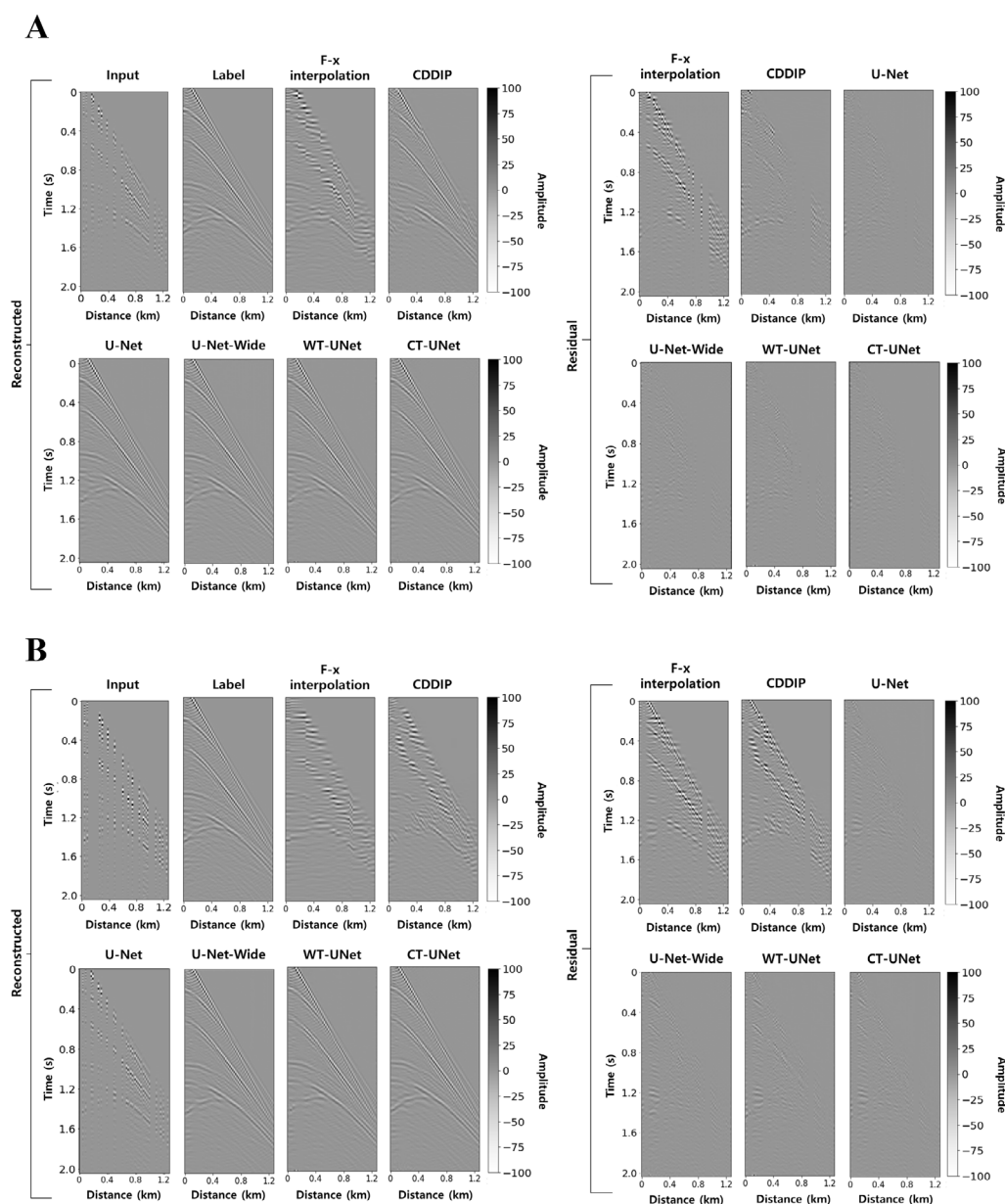


Figure 11. Reconstruction results and residuals of synthetic data under irregular (random) missing: (A) 50% missing rate and (B) 70% missing rate. In each panel, the top row shows the reconstructed results and the bottom row shows the residuals. From left to right, the reconstructed results are the input, ground truth, f-x interpolation, Constrained Diffusion-Driven Deep Image Prior (CDDIP), U-Net, U-Net-Wide, wavelet transform (WT)-UNet, and curvelet transform (CT)-UNet. The residual panels, from left to right, correspond to f-x interpolation, CDDIP, U-Net, U-Net-Wide, WT-UNet, and CT-UNet.

exhibits fundamentally faster convergence compared to $O(k^{-1})$ for wavelets. This reduces the complexity of the residuals that each component-wise U-Net must learn, and mitigates the spectral bias of a single network through component-wise learning.⁴¹

The experimental results are consistent with this theoretical prediction. Under high missing-rate conditions

(80%), CT-UNet achieved PSNR improvements of 0.28 dB on synthetic data and 1.61 dB on field data compared to U-Net, with the difference statistically significant. In contrast, no significant difference was observed under low missing-rate conditions. This theoretical property also translates into a practical advantage for reconstructing conflicting-dip events, which are known to pose significant

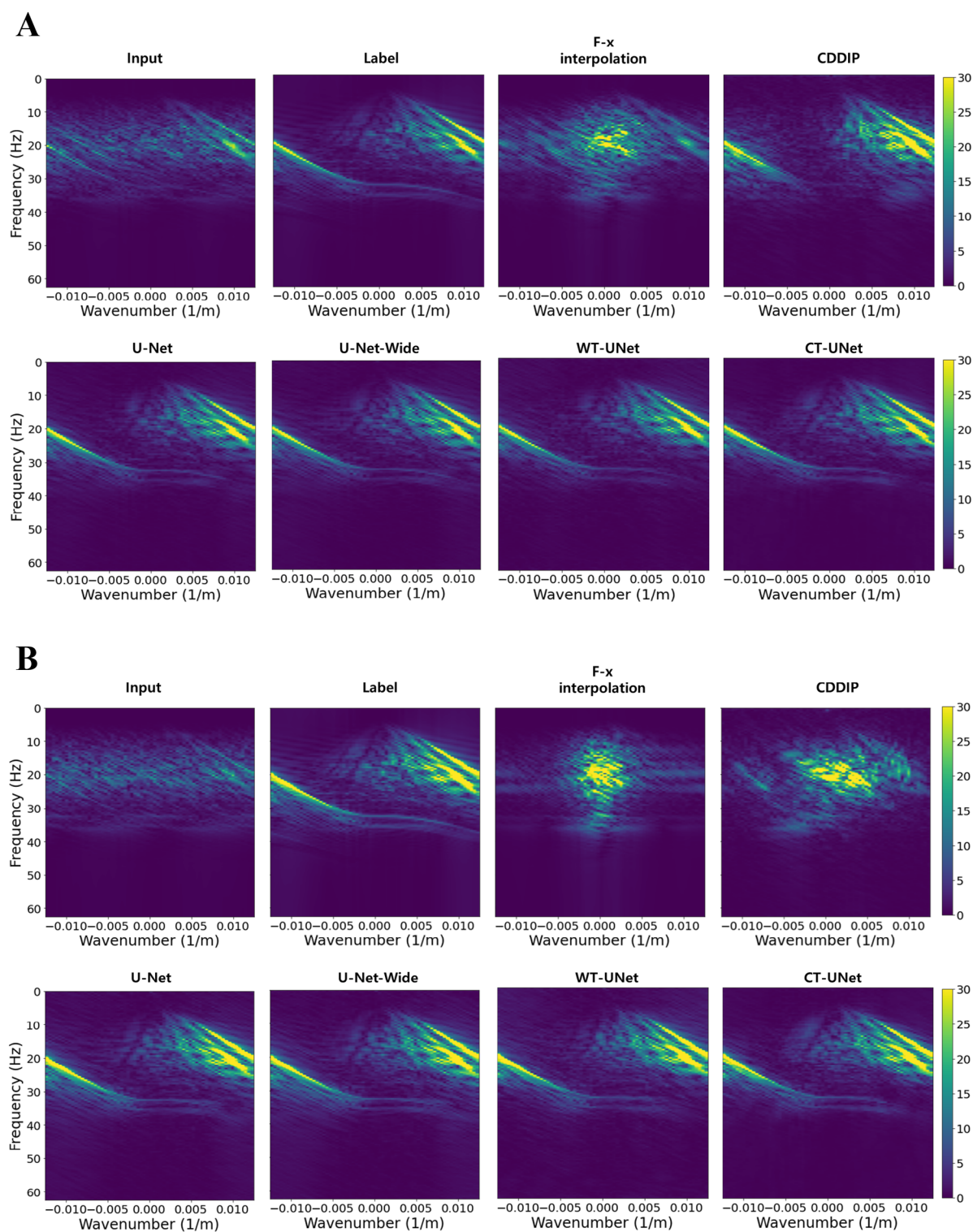


Figure 12. The f-k spectra corresponding to the reconstructed results in Figure 11 under irregular (random) missing: (A) 50% missing rate and (B) 70% missing rate. In each panel, from left to right, the top row shows the input, ground truth, f-x interpolation, and Constrained Diffusion-Driven Deep Image Prior results, while the bottom row shows the U-Net, U-Net-Wide, wavelet transform-UNet, and curvelet transform-UNet results.

fundamentally different research question—learning-based versus iterative-optimization-based reconstruction—and requires dedicated investigation of optimization hyperparameters (iteration count, thresholding schedule, regularization weights). Such a comparison is therefore outside the present scope and is left for independent investigation.

In comparison with CDDIP, it should be noted that CDDIP is fundamentally different, performing iterative optimization without training data, and the purpose of this comparison is to examine performance characteristics across different reconstruction paradigms.

4.5. Computational cost and recommended use cases

The CT-UNet incurs approximately a threefold increase in training time (~30 GPU-hours) and inference time (~30 s) compared to a single U-Net (Table 8); however, this can be reduced in a parallel GPU environment, and the peak memory remains identical at ~6 GB (training)/<0.25 GB (inference). The practical value of this additional cost is supported by the consistent quantitative gains: shot-level paired t-tests confirm statistically significant PSNR improvements over U-Net at synthetic 80% (+0.28 dB), synthetic 70% irregular missing (+0.66 dB), field 75% (+0.60 dB), and field 80% (+1.61 dB). CT-UNet also maintains SSIM values comparable to or higher than those of the baseline methods across the tested conditions, suggesting that its structural reconstruction quality is at least preserved and often improved beyond the high-missing regime. In conflicting-dip regions where conventional methods fail (ROI analysis, Tables 3 and 4), CT-UNet's advantage is particularly pronounced. Furthermore, the component-wise architecture enables selective deployment: in scenarios where only specific frequency bands require enhanced reconstruction, individual component networks

can be applied independently - an operational flexibility unavailable in single-network architectures. CT-UNet is therefore most recommended for applications where high missing rates, structural continuity, or conflicting-dip event recovery are critical. For lower-missing-rate cases where the performance gap relative to a standard U-Net is small, the simpler U-Net may remain a more cost-effective choice.

4.6. Validation methodology for field data

The field data validation in this study employed Mobil AVO Viking Graben Line 12,³⁸ a fully sampled, publicly available benchmark dataset widely used in seismic interpolation research. Controlled missing patterns were applied to the original data, which served as ground truth, enabling rigorous quantitative evaluation through PSNR and SSIM. This semi-synthetic validation protocol is a standard approach in the field, providing quantifiable benchmarks for objective comparison across methods, reproducibility through publicly available data, and systematic control over missing rate and pattern. Application to datasets with naturally occurring acquisition gaps constitutes a natural application study downstream of the methodological validation presented here.

4.7. Limitations and future work

This study has the following limitations: First, the real/imaginary separation of complex coefficients means that phase coherence is only implicitly learned, necessitating the introduction of complex-valued convolution or phase consistency loss. Second, only random irregular missing was examined in this study. More realistic acquisition-gap scenarios, including blocky missing traces, mixed regular-irregular missing, consecutive missing traces, and spatially varying missing patterns, require systematic evaluation with dedicated mask-generation protocols and

Table 8. Computational cost comparison

Method	Training time (GPU-hours)	Inference time (min)	GPU memory (train/infer)
F-x interpolation	–	~0.5	<1 GB
CDDIP	~16.7	~2.5	~1.5 GB/<0.5 GB
U-Net	~10	~0.17	~6 GB/<0.25 GB
U-Net-Wide	~12	~0.20	~7 GB/<0.3 GB
WT-UNet	~10	~0.17	~6 GB/<0.25 GB
CT-UNet	~30	~0.5	~6 GB/<0.25 GB

Abbreviations: CDDIP: Constrained Diffusion-Driven Deep Image Prior; CT: Curvelet transform; GPU: Graphics processing unit; WT: Wavelet transform.

mask-aware learning strategies. These cases are therefore left as future work. Third, only MSE loss was applied, but component-specific customized losses (e.g., MSE for LF and frequency-domain loss for HF) represent an extension direction naturally permitted by the modular framework of CT-UNet. Fourth, the standard U-Net can be replaced with attention-based models, transformers, or diffusion models. Fifth, the synthetic data were generated under relatively clean conditions, and robustness evaluation under controlled noise environments (signal-to-noise ratio 10/20 dB) remains a future task.

5. Conclusion

In this study, we proposed CT-UNet, a seismic data interpolation framework based on the curvelet transform. The proposed method decomposes seismic data into multi-scale and multi-directional components using the UDCT, reconstructs each component with independent U-Net models, and synthesizes them via inverse transformation, aiming to simultaneously enhance event continuity and preserve frequency components even under high missing-rate conditions.

The performance of CT-UNet was evaluated on synthetic data (SEG/EAGE Salt Model) and field data (Mobil AVO Viking Graben Line 12) at regular missing rates of 67%, 75%, and 80%, as well as on synthetic data at 50% and 70% random irregular missing rates. Combining spatial-domain reconstruction results, f-k spectral analysis, and PSNR/SSIM quantitative evaluation, CT-UNet showed performance comparable to U-Net under lower missing-rate conditions, while its advantages became more evident under challenging conditions, including synthetic 80%, synthetic 70% irregular missing, field 75%, and field 80%. The shot-level analysis further supports the conclusion that these gains are most robust under high missing-rate and challenging irregular-missing conditions. In the tested irregular missing scenarios, CT-UNet maintained the highest PSNR and SSIM among the compared methods, suggesting that curvelet-based component decomposition can remain effective beyond periodic regular undersampling. However, more realistic missing patterns, such as blocky, mixed regular-irregular, consecutive, and spatially varying gaps, require further investigation.

In summary, the results indicate that integrating curvelet-based multi-scale and multi-directional representation into deep learning interpolation is an effective strategy for improving event continuity and spectral consistency, particularly under high-rate regular undersampling and in regions containing curvature-varying, conflicting-dip events. The proposed CT-UNet is therefore expected to be

most useful in acquisition scenarios where high missing rates, structural continuity, and recovery from conflicting dip events are critical.

Acknowledgments

None.

Funding

This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (No. RS-2025-25435564), and by the Korea Institute of Marine Science and Technology Promotion grant funded by the Ministry of Oceans and Fisheries (No. 20220254, Development of technology for seabed classification based on machine learning).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Daeung Yoon

Formal analysis: Jiyun Yu

Investigation: Jiyun Yu

Methodology: Daeung Yoon, Young Kim

Writing-original draft: Jiyun Yu

Writing-review & editing: Jiyun Yu, Daeung Yoon

Availability of data

The data underlying this article are available in the SEG Wiki (https://wiki.seg.org/wiki/SEG_C3_NA) and the Mobil AVO Viking Graben Line 12 (https://wiki.seg.org/wiki/Mobil_AVO_viking_graben_line_12).

Further disclosure

The methodology (Section 2) and the synthetic-data results (Sections 3.1 and 3.2) of this study were presented at the 2025 Spring Conference of the Korean Society of Earth and Exploration Geophysicists (KSEG), held at the Korea Institute of Geoscience and Mineral Resources (KIGAM), Daejeon, Republic of Korea, on April 3, 2025.

References

1. Spitz S. Seismic trace interpolation in the F-X domain. *Geophysics*. 1991;56(6):785-794.
doi: 10.1190/1.1443096
2. Abma R, Kabir N. 3D interpolation of irregular data with a POCS algorithm. *Geophysics*. 2006;71(6):E91-E97.
doi: 10.1190/1.2356088

3. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: *Proceedings of MICCAI 2015*. LNCS 9351. 2015:234-241.
doi: 10.1007/978-3-319-24574-4_28
4. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27–30, 2016; Las Vegas, NV, USA. 2016:770-778.
doi: 10.1109/CVPR.2016.90
5. Wang B, Zhang N, Lu W, Wang J. Deep-learning-based seismic data interpolation: A preliminary result. *Geophysics*. 2019;84(1):V11-V20.
doi: 10.1190/geo2017-0495.1
6. Yu J, Wu B. Attention and hybrid loss guided deep learning for consecutively missing seismic data reconstruction. *IEEE Trans Geosci Remote Sens*. 2022;60:1-8.
doi: 10.1109/TGRS.2021.3068279
7. Li X, Wu B, Zhu X, Yang H. Consecutively missing seismic data interpolation based on coordinate attention Unet. *IEEE Geosci Remote Sens Lett*. 2022;19:3005005.
doi: 10.1109/LGRS.2021.3128511
8. He T, Wu B, Zhu X. Seismic data consecutively missing trace interpolation based on multistage neural network training process. *IEEE Geosci Remote Sens Lett*. 2022;19:7504105.
doi: 10.1109/LGRS.2021.3089585
9. Tian Y, Fu L, Fang W, Li T. FR-UNet: A feature restoration-based UNet for seismic data consecutively missing trace interpolation. *IEEE Trans Geosci Remote Sens*. 2025;63:1-10.
doi: 10.1109/TGRS.2025.3531934
10. Yu J, Joo Y, Yoon D. Transposed arrangement strategy-based deep learning for seismic data crossline interpolation. *J Seism Explor*. 2025;34(5):66-80.
doi: 10.36922/JSE025330056
11. Mukherjee B, Kar S, Banerjee R, Sain K. Seismic data recovery with the aid of 3D U-Net deep learning approach: A case study from Amguri prospect, Upper Assam Shelf, India. *J Appl Geophys*. 2026;250:106214.
doi: 10.1016/j.jappgeo.2026.106214
12. Yu J, Yoon D. Crossline reconstruction of 3D seismic data using 3D cWGAN: A comparative study on Sleipner seismic survey data. *Appl Sci*. 2023;13(10):5999.
doi: 10.3390/app13105999
13. Oliveira DAB, Ferreira RS, Silva R, Brazil EV. Interpolating seismic data with conditional generative adversarial networks. *IEEE Geosci Remote Sens Lett*. 2018;15(12):1952-1956.
doi: 10.1109/LGRS.2018.2866199
14. Guo Y, Fu L, Li H. Seismic data interpolation based on multi-scale Transformer. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.
doi: 10.1109/LGRS.2023.3298101
15. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2020;33:6840-6851. Accessed January 20, 2026. <https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html>
16. Goyes-Peñafiel P, Kamilov US, Arguello H. CDDIP: Constrained diffusion-driven deep image prior for seismic data reconstruction. *IEEE Geosci Remote Sens Lett*. 2025;22:7504105.
doi: 10.1109/LGRS.2025.3545860
17. Wei X, Zhang C, Wang H, Tan C, Xiang D, Jiang B. Seismic data interpolation via denoising diffusion implicit models with coherence-corrected resampling. *IEEE Trans Geosci Remote Sens*. 2024;62:1-14.
doi: 10.1109/TGRS.2024.3485573
18. Durall R, Ghanim A, Fernandez MR, Ettrich N, Keuper J. Deep diffusion models for seismic processing. *Comput Geosci*. 2023;177:105377.
doi: 10.1016/j.cageo.2023.105377
19. Liu Q, Ma J. Generative interpolation via a diffusion probabilistic model. *Geophysics*. 2024;89(1):65-85.
doi: 10.1190/geo2023-0182.1
20. Fang W, Fu L, Zhang M, Li Z. Seismic data interpolation based on U-net with texture loss. *Geophysics*. 2021;86(1):V41-V54.
doi: 10.1190/geo2019-0615.1
21. Chang D, Yang W, Yong X, et al. Seismic Data Interpolation Using Dual-Domain Conditional Generative Adversarial Networks. *IEEE Geosci Remote Sens Lett*. 2021;18(10):1856-1860.
doi: 10.1109/LGRS.2020.3008478
22. Chang DK, Yang WY, Yong XS, Li HS. Seismic data interpolation with conditional generative adversarial network in time and frequency domain. In: *SEG Technical Program Expanded Abstracts*. Houston, TX: Society of Exploration Geophysicists; 2019:2589-2593.
doi: 10.1190/segam2019-3210118.1
23. Lou Y, Wu L, Liu L, et al. Irregularly sampled seismic data interpolation via wavelet-based convolutional block attention deep learning. *Artif Intell Geosci*. 2022;3:192-202.
doi: 10.1016/j.aiig.2022.12.001
24. Dodda VC, Kuruguntla L, Mandpura AK, Elumalai K. Simultaneous seismic data denoising and reconstruction with attention-based wavelet-convolutional neural network. *IEEE Trans Geosci Remote Sens*. 2023;61:1-14.

- doi: 10.1109/TGRS.2023.3267037
25. Liu N, Wu L, Wang J, Wu H, Gao J, Wang D. Seismic data reconstruction via wavelet-based residual deep learning. *IEEE Trans Geosci Remote Sens.* 2022;60:4508213.
doi: 10.1109/TGRS.2022.3152984
 26. Kong F, Picetti F, Lipari V, Bestagini P, Tang X, Tubaro S. Deep prior-based unsupervised reconstruction of irregularly sampled seismic data. *IEEE Geosci Remote Sens Lett.* 2022;19:7501305.
doi: 10.1109/LGRS.2020.3044455
 27. Candès EJ, Donoho DL. New tight frames of curvelets and optimal representations of objects with C2 singularities. *Commun Pure Appl Math.* 2003;57(2):219-266.
doi: 10.1002/cpa.10116
 28. Candès EJ, Demanet L, Donoho DL, Ying L. Fast discrete curvelet transforms. *Multiscale Model Simul.* 2006;5(3):861-899.
doi: 10.1137/05064182X
 29. Herrmann FJ, Hennenfent G. Non-parametric seismic data recovery with curvelet frames. *Geophys J Int.* 2008;173(1):233-248.
doi: 10.1111/j.1365-246X.2007.03698.x
 30. Sun B, Ma J, Chauris H, Yang H. Solving wave equation in the curvelet domain: A multi-scale and multi-directional approach. *J Seism Explor.* 2009;18(4):385-399.
doi: 10.36922/JSE402
 31. Sitzmann V, Martel JNP, Bergman AW, Lindell DB, Wetzstein G. Implicit neural representations with periodic activation functions. In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2020;33:7462-7473. Accessed February 2, 2026. <https://proceedings.neurips.cc/paper/2020/hash/53c04118df112c13a8c34b38343b9c10-Abstract.html>
 32. Gao W, Liu D, Chen W, Sacchi MD, Wang X. NeRSI: Neural implicit representations for 5D seismic data interpolation. *Geophysics.* 2025;90(1):V29-V42.
doi: 10.1190/geo2023-0767.1
 33. Lee G, Cheong S, Choi Y. Implicit neural representations for unsupervised seismic data interpolation from single gather. *Geophys Prospect.* 2025;73(9):e70110.
doi: 10.1111/1365-2478.70110
 34. Do MN, Vetterli M. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Trans Image Process.* 2005;14(12):2091-2106.
doi: 10.1109/TIP.2005.859376
 35. Nguyen TT, Chauris H. Uniform discrete curvelet transform. *IEEE Trans Signal Process.* 2010;58(7):3618-3634.
doi: 10.1109/TSP.2010.2047666
 36. Kingma DP, Ba J. Adam: A method for stochastic optimization. In: *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. 2015. Accessed February 20, 2026. <https://arxiv.org/abs/1412.6980>
 37. Aminzadeh F, Brac J, Kunz T. *SEG/EAGE 3-D Salt and Overthrust Models*. SEG/EAGE 3-D Modeling Series, No. 1. Tulsa, OK: Society of Exploration Geophysicists; 1997.
 38. Keys RG, Foster DJ. A data set for evaluating and comparing seismic inversion methods. In: Keys RG, Foster DJ, eds. *Comparison of Seismic Inversion Methods on a Single Real Data Set*. Tulsa, OK: Society of Exploration Geophysicists; 1998:1-12.
doi: 10.1190/1.9781560802082.ch1
 39. Wang S, Su Z, Ying L, et al. Accelerating magnetic resonance imaging via deep learning. In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2016:514-517.
doi: 10.1109/ISBI.2016.7493320
 40. Cole EK, Ong F, Vasanaawala SS, Pauly JM. Fast unsupervised MRI reconstruction without fully-sampled ground truth data using generative adversarial networks. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. 2021:3988-3997.
doi: 10.1109/ICCVW54120.2021.00444
 41. Rahaman N, Baratin A, Arpit D, et al. On the spectral bias of neural networks. In: *Proceedings of the 36th International Conference on Machine Learning (ICML)*. 2019;97:5301-5310. Accessed February 20, 2026. <https://proceedings.mlr.press/v97/rahaman19a.html>