ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Transposed arrangement strategy-based deep learning for seismic data crossline interpolation

Jiyun Yu[1†], Yonghwan Joo[2†], and Daeung Yoon[1*]

[1]Department of Energy and Resources Engineering, Chonnam National University, Gwangju, Republic of Korea

[2]Resources Exploration and Development Research Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), Daejeon, Republic of Korea

Abstract

Towed-streamer marine seismic acquisition systems generally have a dense receiver spacing in the inline receiver direction within common-shot gathers (along-streamer), while the streamer spacing is relatively sparse in the crossline receiver direction within common-shot gathers (cross-streamer). This disparity can lead to spatial aliasing issues in the crossline receiver direction within common-shot gathers and result in resolution degradation during the processing of 3D seismic data. To address this issue and enhance resolution, data interpolation in the crossline receiver direction within common-shot gathers is essential. Various supervised learning-based interpolation methods have been developed to this end. However, the absence of true data in the crossline receiver direction within common-shot gathers poses challenges for training supervised learning models with actual field data. To overcome this, we have developed a novel approach called the "transposed arrangement strategy" for a deep learning-based reconstruction model for crossline interpolation. This method involves training the model with 3D input and labels patched from existing field data, and then applying the trained model with transposed 3D input to reconstruct data in the crossline receiver direction within common-shot gathers. During this process, the 3D U-Net and U-Net+ models were utilized, demonstrating their superiority through comparisons with traditional interpolation methods.

Keywords: Data processing; Signal processing; Deep learning; Crossline interpolation; Transposed arrangement strategy; 3D U-Net+

[†]These authors contributed equally to this work.

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

Citation: Yu J, Joo Y, Yoon D. Transposed arrangement strategy-based deep learning for seismic data crossline interpolation. *J Seismic Explor*. doi: 10.36922/JSE025330056

Received: August 13, 2025

1st revised: September 22, 2025

2nd revised: October 24, 2025

Accepted: November 3, 2025

Published online: November 25, 2025

1. Introduction

Seismic survey data can be subject to loss due to environmental, economic constraints, or mechanical defects. Notably, the towed-streamer marine seismic systems are typically configured with dense inline and wide crossline spacing to maximize exploration coverage.[1] This setup can lead to spatial aliasing issues in the crossline receiver direction within common-shot gathers and resolution degradation during the processing of 3D seismic survey data. Consequently, data interpolation in the crossline receiver direction within common-shot gathers is necessary to address these problems and improve data resolution.

Before the introduction of deep learning, a wide range of conventional approaches were developed for seismic data interpolation. Traditional methods include bicubic

interpolation,[2] wave-equation–based methods,[3,4] and f–x interpolation methods.[5] Prediction error filtering[6] has also been widely applied for reconstructing missing traces by exploiting predictable structures in seismic data. In addition, projection onto convex sets (POCS)[7] methods and their extensions, such as curvelet-domain POCS,[8,9] have been shown to effectively exploit sparsity in transform domains. Rank reduction methods[10-12] and sparse transform-based methods, including Fourier, curvelet, and seislet transforms,[13-15] have also been extensively studied. These conventional methods perform well for linear or near-linear seismic events and can mitigate aliasing, but they often involve high computational cost, require careful parameter tuning (*e.g.*, window size, rank selection), and may struggle with complex or curved structures. With the increasing size and complexity of modern 3D seismic surveys, these limitations have motivated the exploration of more efficient and adaptive approaches, including deep learning–based methods.

Recent advancements in deep learning technology have been actively incorporated into seismic data interpolation research to provide high-accuracy interpolation results and more efficient data processing.[16-20] Research in deep learning-based seismic data interpolation can be categorized into techniques for 2D and 3D seismic data based on the dimensionality of the data.

2D seismic data interpolation techniques are primarily based on convolutional neural networks (CNNs), generative adversarial networks (GANs) and diffusion models. Yu and Wu[21] proposed a CNN that utilizes a hybrid loss function combining structural similarity index (SSIM) and L1 norm, along with an attention mechanism that explicitly leverages global information. Li *et al.*[22] proposed a method of integrating a coordinate attention block into U-Net for 2D successive missing traces interpolation. Park *et al.*[20] proposed a strategy using the Coarse-Refine U-Net (CFunet), which consists of two U-Nets and an upsampling process between them, along with the Fourier loss. Lou *et al.*[23] proposed a wavelet-based convolutional block attention deep learning (W-CBADL) network for the reconstruction of irregularly sampled seismic data. Dodda *et al.*[24] proposed the use of an attention-based wavelet convolutional neural network (AWUN) for simultaneous noise reduction and reconstruction of incomplete seismic data. Tian *et al.*[25] proposed feature restoration-based U-Net (FR–U-Net), a feature restoration-based interpolation method built upon the U-Net architecture to reconstruct consecutively missing seismic traces. Kaur *et al.*[26] proposed a model using the cycle GAN structure that comprises two generators and one discriminator. Chang *et al.*[27] proposed a method for interpolating seismic data using the conditional generative adversarial network in the time and frequency domains (TF-CGAN). Deng *et al.*[28] proposed conditional constraint diffusion model, a diffusion model with conditional constraints. However, these proposed methods are limited to interpolating 2D seismic data and have not been applied to multidimensional seismic data.

Consequently, 3D seismic data interpolation techniques have been developed, which predominantly involve models based on CNNs, GANs, diffusion models and additionally, autoencoder-based models. Kong *et al.*[29] proposed a multi-resolution U-Net model that utilizes the correlations in 3D data. Jin *et al.*[30] proposed a method to reduce the operational cost of U-Net by introducing depthwise separable convolution instead of standard convolution. Chang *et al.*[31] proposed a dual-domain conditional generative adversarial network that uses seismic data sets and discrete Fourier-transformed data sets in the frequency domain as input vectors. Dou *et al.*[32] proposed a multi-dimensional adversarial GAN that uses three discriminators. Yu and Yoon[33] applied the conditional Wasserstein GAN (cWGAN) model to 3D seismic data interpolation. Ding *et al.*[34] proposed Self-Attention Generative Adversarial Network, a deep learning-based model that integrates the self-attention mechanism with GAN. Wang *et al.*[35] proposed SeisFusion, a diffusion model combined with conditional constraints. Qian *et al.*[36] introduced the deep tensor autoencoder model, which is capable of learning data-driven, non-linear, and high-dimensional mappings.

Despite these technological advancements, these models often fall short of generating new data to reduce the spacing between traces compared to the original dataset. Typically, they demonstrate strong performance when interpolating from decimated data back to its full original format. This is because a supervised learning-based interpolation model requires a label, which is selected from the original dataset. For instance, if the original data has a spacing of 20 m, a model might be effectively trained to fill in missing traces corresponding to the label, thus decreasing the spacing from 40 m to 20 m by reconstructing the decimated data. However, this training does not guarantee a reduction in spacing from 20 m to 10 m, which is the ultimate goal of seismic data interpolation.

This limitation, however, can be mitigated in cross-streamer wavefield reconstruction due to the denser inline sensor spacing compared to the crossline receiver direction within common-shot gathers. In this scenario, an interpolation model can be trained along the inline receiver direction within common shot gathers and subsequently applied to reconstruct the coarser data in the crossline receiver direction within common-shot gathers. Larsen Greiner *et al.*[37] introduced a cross-streamer

wavefield reconstruction model using a wavelet domain approach, training the model along the inline receiver direction within common-shot gathers to reconstruct coarse crossline receiver direction within common-shot gathers. Yeeh *et al.*[38] proposed a trace-to-trace approach for crossline interpolation, which learns the relationship between consecutive traces to predict the intermediate trace along the dense inline receiver direction within common-shot gathers, ultimately bridging the gap in the coarse crossline receiver direction within common-shot gathers.

Recently, increasing attention has been given to self-supervised learning approaches, which mitigate the dependency on labeled crossline receiver direction within common-shot gathers. Chen *et al.*[39] proposed an interpolation method that combines POCS-Net, a CNN-based architecture built upon the POCS algorithm, with a self-supervised transfer learning framework. In addition, Wang *et al.*[40] presented a dip neural network (DINN) that leverages self-supervised learning for crossline interpolation. More recently, Goyes-Peñafiel *et al.*[41] introduced constrained diffusion-driven deep image prior (CDDIP), an unsupervised framework that combines diffusion processes with deep image prior (DIP) to reconstruct missing seismic traces without requiring labeled data.

In this study, we extend this strategy[39-41] for cross-wavefield reconstruction by utilizing not only 2D data from the inline receiver direction within common-shot gathers but also 3D volume data encompassing both inline and crossline receiver direction within common-shot gathers for training, effectively enhancing data interpolation by utilizing information across multiple spatial dimensions. The proposed approach, called the transposed arrangement strategy, alters the input array during the training and inference stages of the crossline interpolation model. This strategy allows for the reconstruction of the crossline receiver direction within common-shot gathers from original field data without the need for specific crossline labels. It involves training the model using a 3D input dataset from which 50% of the traces are regularly removed along the inline receiver direction within common-shot gathers. After training, this original input is transposed and fed into the model to facilitate data generation in the crossline receiver direction within common-shot gathers. In addition, to align event patterns between the inline and crossline receiver direction within common-shot gathers, linear moveout (LMO) correction is applied. Compared with existing self-supervised approaches, which rely on contrastive objectives, masked autoencoding, or physics-based regularization to avoid the need for labels, our method achieves self-supervision in a different way: It directly leverages dense inline traces as surrogate labels during training and then transposes the model for crossline inference. This provides a simple yet effective mechanism for eliminating the need for explicit crossline labels while remaining complementary to prior self-supervised paradigms. In addition, unlike other frameworks, such as DINN and POCS-Net, the proposed transposed strategy requires no auxiliary convex constraints, dip picking, or secondary network—the self-supervision emerges purely from a simple tensor transpose operation, making the method both conceptually elegant and computationally efficient. For this purpose, we employed two interpolation models: 3D U-Net and 3D U-Net+, whose effectiveness was measured against the conventional bicubic interpolation method,[2] the f–x interpolation method,[5] POCS[6] and CDDIP[41] through comparative performance evaluations. Among these, f–x interpolation was included as a widely used conventional method, POCS was selected as a representative compressive sensing–based method due to its simplicity, robustness, and computational efficiency, and CDDIP was considered to represent recent self-supervised deep learning approaches.

2. Methodology

2.1. Transposed arrangement strategy

Our proposed transposed arrangement strategy offers a label-substituting approach for interpolating actual field data. This strategy is universally applicable across deep learning models and is particularly optimized for scenarios where the receiver-to-streamer spacing ratio is 1–2. For example, in the dataset used in this study, the inline receiver spacing was 20 m and the crossline streamer spacing was 40 m, resulting in a ratio of 1:2. The proposed strategy unfolds as follows:

Initially, the original data D is acquired from the actual field and consists of N_i traces in the inline receiver direction within common-shot gathers, N_x traces in the crossline receiver direction within common-shot gathers, and N_t time samples. This can be represented as a matrix D with dimensions $N_x \times N_i \times N_t$, and the trace T can be illustrated along the time sample axis as follows:

$$\mathrm{D} = \begin{bmatrix} \mathrm{T}_{1,1} & \cdots & \mathrm{T}_{1,N_i} \\ \vdots & \ddots & \vdots \\ \mathrm{T}_{N_x,1} & \cdots & \mathrm{T}_{N_x,N_i} \end{bmatrix} \tag{I}$$

Where $T_{x,i}$ is a trace vector corresponding to the crossline index x and inline index i, each with a dimension of $N_t \times 1$. Here, N_x denotes the number of traces in the

crossline direction, N_i the number of traces in the inline direction, and N_t the number of time samples per trace. Thus, each element $T_{x,i}$ represents a seismic trace consisting of N_t amplitude values along the time axis.

From the original data D, a volume patch P of size $\tilde{N}_x \times \tilde{N}_i \times \tilde{N}_t$ is extracted. If the starting position of the patch is denoted as (i, j), then the patch P can be represented as:

$$P = \begin{bmatrix} T_{i,j} & \cdots & T_{i,j+\tilde{N}_i-1} \\ \vdots & \ddots & \vdots \\ T_{i+\tilde{N}_x-1,j} & \cdots & T_{i+\tilde{N}_x-1,j+\tilde{N}_i-1} \end{bmatrix} \quad \text{(II)}$$

Where $\tilde{N}_x$ and $\tilde{N}_i$ serve as the indices navigating the crossline and inline receiver direction within common-shot gathers, respectively, within the patch, and $\tilde{N}_i = 2\tilde{N}_x - 1$.

During the training phase, the complete patch P is used as a label, while an input P_{in} is created by regularly removing 50% of P in the inline receiver direction within common-shot gathers, represented as follows:

$$P_{in} = \begin{bmatrix} T_{i,j} & 0 & T_{i,j+2} & 0 & \cdots & T_{i,j+\tilde{N}_i-1} \\ T_{i+1,j} & 0 & T_{i+1,j+2} & 0 & \cdots & T_{i+1,j+\tilde{N}_i-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ T_{i+\tilde{N}_x-1,j} & 0 & T_{i+m-1,j+2} & 0 & \cdots & T_{i+\tilde{N}_x-1,j+\tilde{N}_i-1} \end{bmatrix} \quad \text{(III)}$$

The training process for the reconstruction model can be represented by the following equation:

$$\hat{P} = Net\left(P_{in}\right)$$

$$\min imize\, L(\hat{P}, P) \quad \text{(IV)}$$

Where $Net(.)$ denotes the reconstruction deep learning model, $\hat{P}$ is the prediction of the model, $L(.)$is the loss function used to train the model. This approach trains the reconstruction model to fill in the missing traces represented by the zero values in **Equation III.**

In the inference phase, a new input patch is created by: (i) Extracting a patch of size $\tilde{N}_x \times \tilde{N}_x$ with $\tilde{N}_t$ time sampling from the original data D, (ii) regularly adding zero values into the crossline receiver direction within common-shot gathers expanding the dimensions to $(2\tilde{N}_x - 1) \times \tilde{N}_x \times \tilde{N}_t$, and (iii) transposing the first and second dimensions to result in a final dimension of $\tilde{N}_x \times (2\tilde{N}_x - 1) \times \tilde{N}_t$. Then, the new input patch for reconstructing crossline receiver direction within common-shot gathers can be represented as follows:

$$\overset{\vee}{P}_{in} = \begin{bmatrix} T_{i,j} & T_{i,j+2} & \cdots & T_{i,j+2(\tilde{N}_x-1)} \\ 0 & 0 & \cdots & 0 \\ T_{i+1,j} & T_{i+1,j+2} & \cdots & T_{i+1,j+2(\tilde{N}_x-1)} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ T_{i+\tilde{N}_x-1,j} & T_{i+\tilde{N}_x-1,j+2} & \cdots & T_{i+\tilde{N}_x-1,j+2(\tilde{N}_x-1)} \end{bmatrix}^T \quad \text{(V)}$$

Subsequently, by inputting new input patches $\overset{\vee}{P}_{in}$ from the original data D into the trained model, Net(.), the model is able to generate the crossline receiver direction within common-shot gathers between streamers. Figure 1 visually illustrates examples for **Equations III** and **V**.

2.2. LMO correction

In the towed streamer system, the difference in offsets between the source and receivers generates varying moveout patterns in the inline and crossline receiver direction within common-shot gathers. Specifically, the moveouts in the inline receiver direction within common-shot gathers are steeper compared to those in the crossline. To harmonize these pattern discrepancies between the inline and crossline receiver direction within common-shot gathers, LMO correction is employed.[38]

LMO correction is a data processing technique used to adjust the arrival time differences of seismic waves captured by multiple receivers. This method is primarily used to improve the initial alignment of multi-channel data and to reduce the time differences between records at close ranges. Notably, this correction can be applied using just the streamer and receiver intervals, without the need for actual subsurface velocity information.[42] The formula for the LMO correction used in this study is as follows:

$$\tau = t\text{-}px \quad \text{(VI)}$$

Where t represents the original time of the seismic event at offset x, p is the ray parameter defined as dt/dx, and τ is the time-shift for the corrected events. In this study, LMO correction is applied to the original seismic data before fetching the volume patches, ensuring that event patterns are similar in both inline and crossline receiver direction within common-shot gathers, with a typical dip range reduced from approximately −8 ms/trace before correction to within ±2 ms/trace after correction. The ray parameter was estimated from the first arrivals and applied uniformly across the line for processing simplicity. In practice, this step could be replaced or refined using spatially varying normal moveout correction or pre-stack time migration velocity fields to achieve higher physical realism.

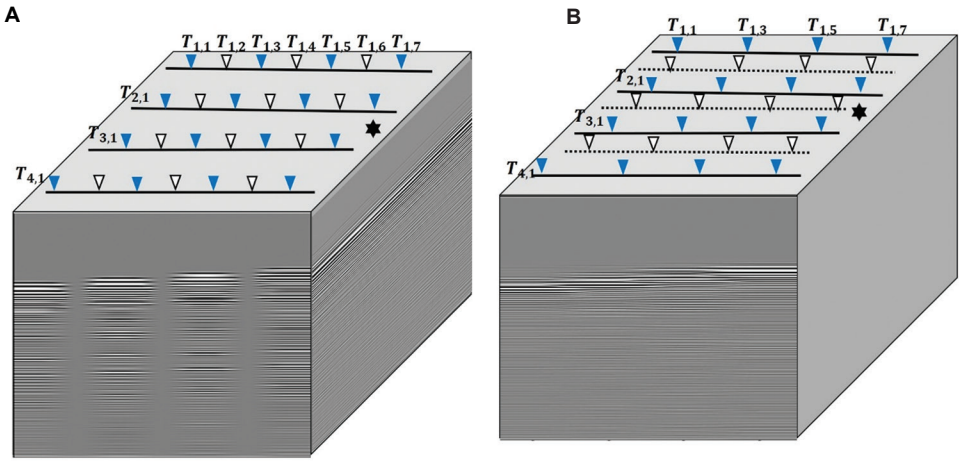

Figure 1. Input arrays for the transposed arrangement strategy during training (A) and inference stages (B). Blue triangles indicate active receivers, and white triangles denote absent receivers, represented as zeros in the data array. Black solid lines represent streamers, while dotted lines show the positions of absent streamers.

2.3. 3D U-Net and 3D U-Net+

Given that our proposed strategy can be applied to any deep learning model, we employed the widely used 3D U-Net and its extension, 3D U-Net+, as our interpolation models.[43,44] Both models were implemented using 3×3×3 convolution layers and 1×1×2 upsampling layers. The 3D U-Net utilizes rectified linear unit as its activation function, whereas 3D U-Net+ employs exponential linear unit. During the model training phase, we adopted the mean absolute error as the loss function and the Adam optimizer for optimization. The learning rate was set to 0.0001. Figure 2 illustrates the structures of 3D U-Net and 3D U-Net+. In this study, the "+" in 3D U-Net+ denotes an extension of the basic 3D U-Net by adding skip connections, which strengthens feature propagation and better preserves structural information during interpolation.

3. Data example

We applied our transposed arrangement strategy to the publicly available SEG/EAGE Salt Model, specifically to the narrow azimuth data. The acquisition parameters for this dataset are detailed in Table 1.[45] The receiver and streamer spacings are 20 m and 40 m, respectively, maintaining a 1:2 ratio, which is well-suited for the application of our strategy.

To evaluate the performance of our proposed strategy, we initially removed 50% of the data regularly in both inline and crossline receiver direction within common-shot gathers from the original dataset. Consequently, the number of streamers was adjusted from 8 to 4, the receiver spacing from 20 m to 40 m, and the streamer spacing from 40 m to 80 m. The removed crossline receiver direction within common-shot gathers were then utilized as the ground truth in our final model evaluation, and to distinguish this dataset from the test set used to train the model below, we call it the crossline test dataset.

Using the adjusted dataset, we initially applied LMO correction based on the first arrival to mitigate discrepancies of moveout patterns arising from inline and crossline receiver direction within common shot gathers, as illustrated in Figure 3. Then, we split the dataset into training, validation, and test sets in a 7:2:1 ratio, respectively, according to shot locations.

During the training phase, we extracted volume patches of size 4×7×624 from the training and validation datasets. To generate a sufficient number of patches, each patch was configured to include seven receivers. This setup ensured that there were overlapping sections between patches. Each patch consists of 4×7 traces, aligned along the crossline and inline receiver direction within common-shot gathers, respectively, with each trace containing 624 time samples. These complete patches are used for labels. Then, we regularly removed 50% of each volume patch in the inline receiver direction within common-shot gathers and replaced the removed values with zeros, to serve as input for our deep learning models as shown in Figure 1A. Thus, the optimal patch size (4×7) for our strategy is determined based on the number of streamers and the number of receivers per streamer.

We trained models using both 3D U-Net and 3D U-Net+ architectures. After training, we evaluated their performance on the test dataset. As illustrated in Figure 4, both models demonstrated similar residual levels and provided interpolation results that closely matched the labels. This performance assessment confirms that our

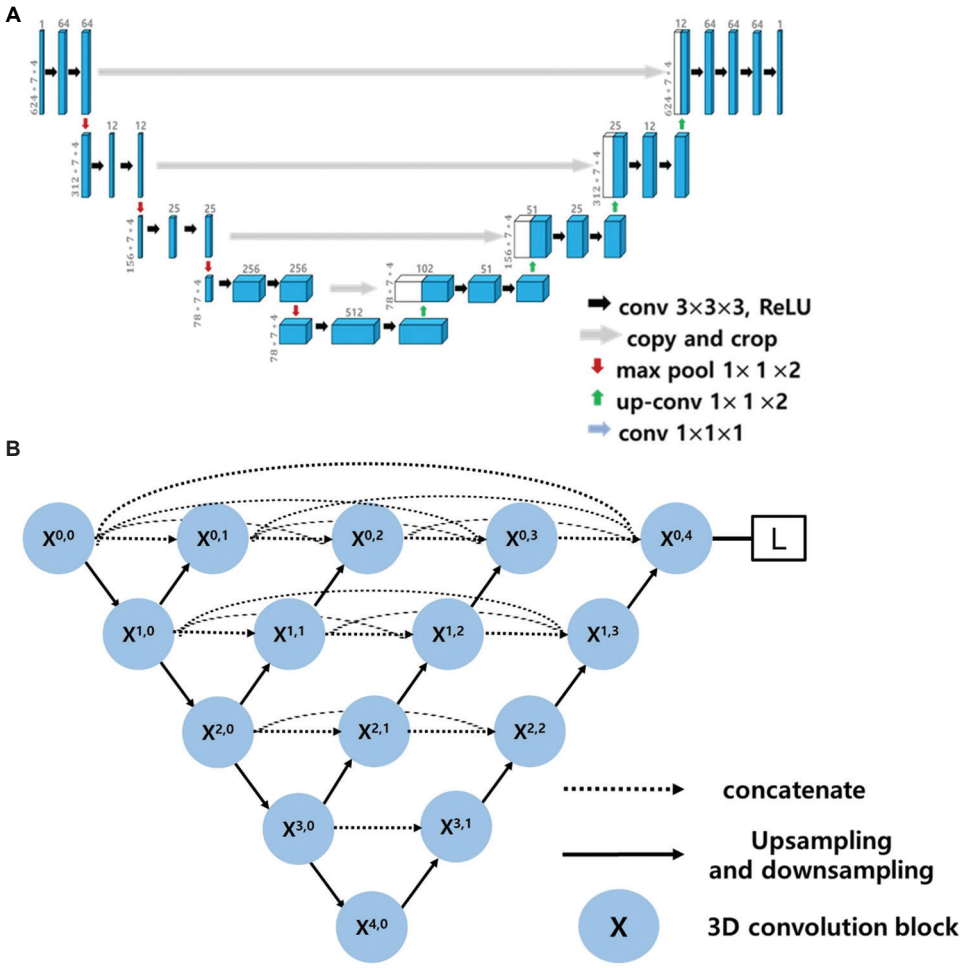

Figure 2. The architectures of 3D U-Net (A) and 3D U-Net+ (B)

Table 1. Data acquisition parameter

Acquisition geometry	Value
Number of sail lines	50
Spacing of sail lines (m)	160
Shots per line	96
Spacing of shot (m)	80
Number of streamers	8
Spacing of receiver (m)	20
Receivers per streamer	68
Max offset (m)	1,340
Number of samples	625
Spacing of samples (s)	0.008
Spacing of streamer (m)	40

models are effectively capable of interpolating data, with 50% of dataset removed regularly in the inline receiver direction within common-shot gathers.

Finally, to reconstruct the crossline receiver direction within common-shot gathers, we implemented the transposed arrangement strategy by feeding transposed input arrays into our trained models, and subsequently compared the outputs with the ground truth from the crossline test dataset. As depicted in Figure 5, the 3D U-Net+ model exhibited smaller residuals compared to the 3D U-Net, POCS, CDDIP, the f–x interpolation method, and the bicubic interpolation method, with the latter showing significantly larger discrepancies relative to the deep learning models. In addition, Figure 5H and 5O represent the results of training and testing the 3D U-Net+ on original data without applying LMO. These results demonstrate that applying LMO significantly improved the performance of the model.

To further highlight differences in fine-scale textures, Figure 6 presents enlarged views of the low amplitude regions marked by red boxes in Figure 5. The results show that the proposed 3D U-Net+ achieves more accurate

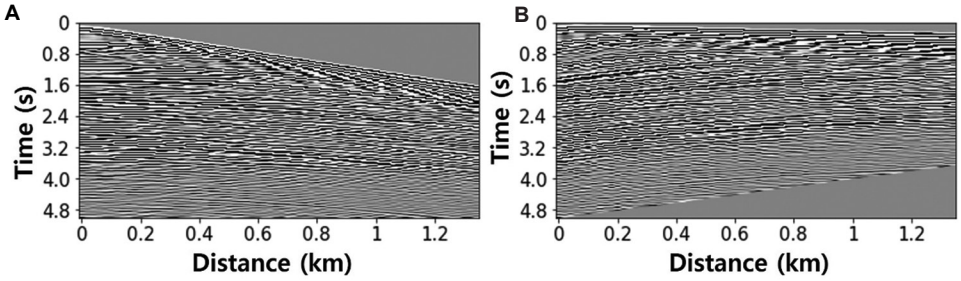

Figure 3. Inline receiver direction within common-shot gathers with before (A) and after applying linear moveout correction (B)

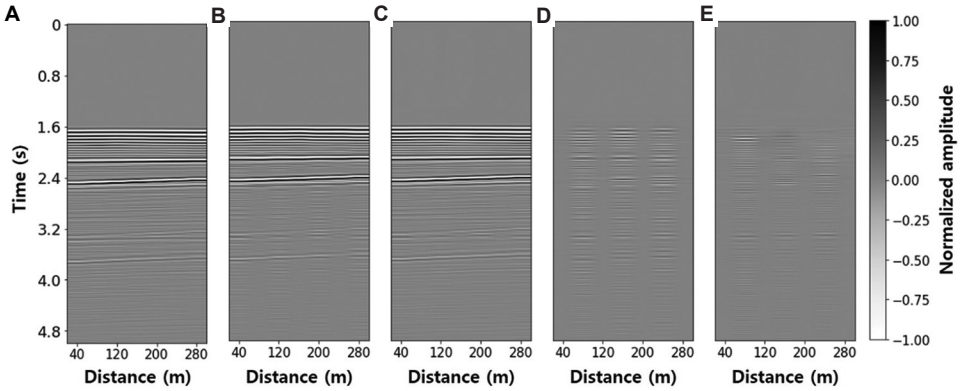

Figure 4. Interpolation results for the inline receiver direction within common shot gathers. (A) Ground truth. (B) 3D U-Net. (C) 3D U-Net+. (D) Difference between (A) and (B). (E) Difference between (A) and (C).

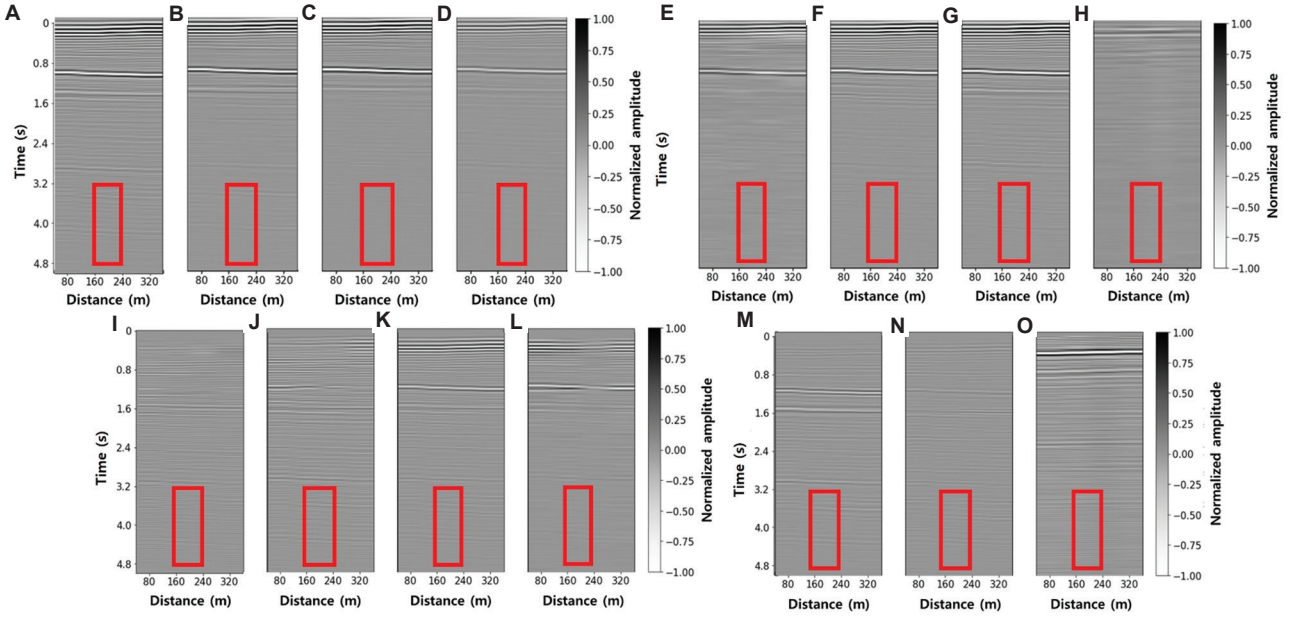

Figure 5. Linear moveout (LMO)–corrected interpolation results for crossline receiver direction within common-shot gathers. (A) Ground truth. (B) Bicubic. (C) f–x interpolation. (D) Projection onto convex sets. (E) Constrained diffusion-driven deep image prior. (F) 3D U-Net. (G) 3D U-Net+. (H) 3D U-Net+ (LMO not applied). (I) Difference between (A) and B). (J) Difference between (A) and (C). (K) Difference between (A) and (D). (L) Difference between (A) and (E). (M) Difference between (A) and (F). (N) Difference between (A) and (G). (O) Difference between (F) and ground truth without LMO applied.

recovery of subtle structures and produces smaller residuals compared to conventional methods, even in low amplitude zones.

Further analysis of trace comparisons, as illustrated in Figure 7, indicates that both 3D U-Net and 3D U-Net+ accurately predict large amplitude events with flat slopes

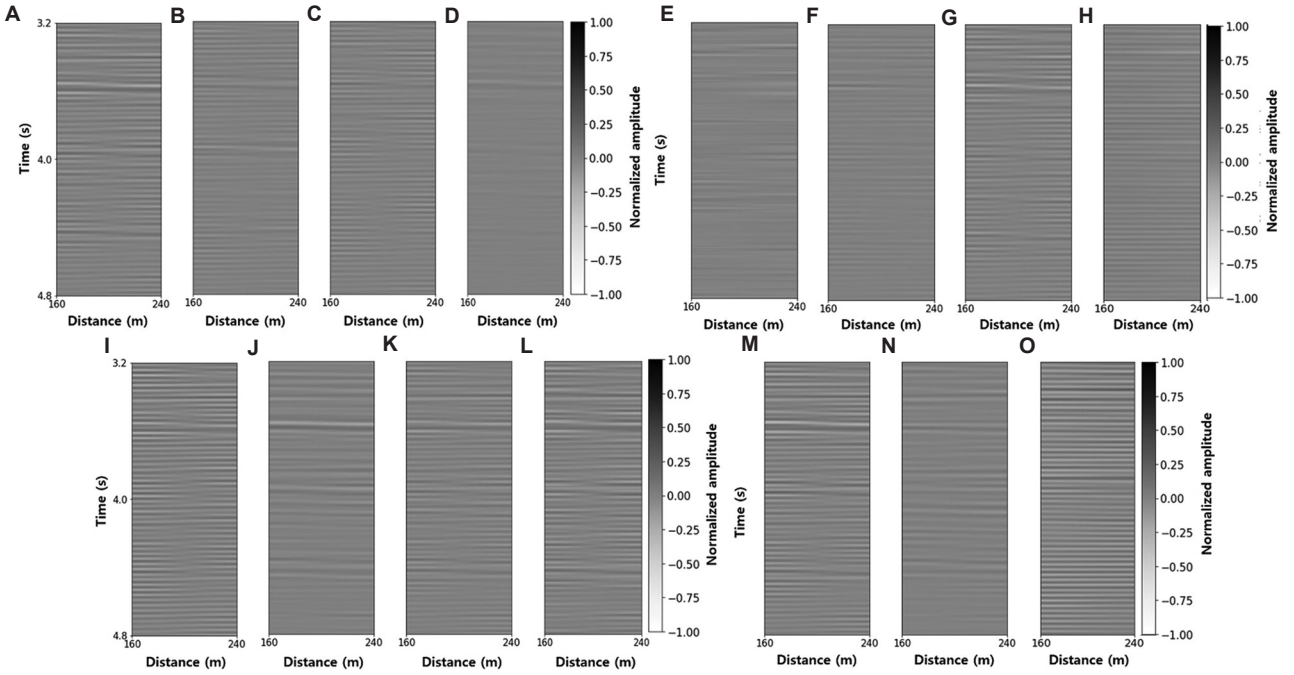

Figure 6. Enlarged views of the regions indicated by red boxes in Figure 5, focusing on the weak-amplitude zones. (A) Ground truth. (B) Bicubic. (C) f–x interpolation. (D) Projection onto convex sets. (E) Constrained diffusion-driven deep image prior. (F) 3D U-Net. (G) 3D U-Net+. (H) 3D U-Net+ (linear moveout [LMO] not applied). (I) Difference between (A) and (B). (J) Difference between (A) and (C). (K) Difference between (A) and (D). (L) Difference between (A) and (E). (M) Difference between (A) and (F). (N) Difference between (A) and (G). (O) Difference between (F) and ground truth without LMO applied.

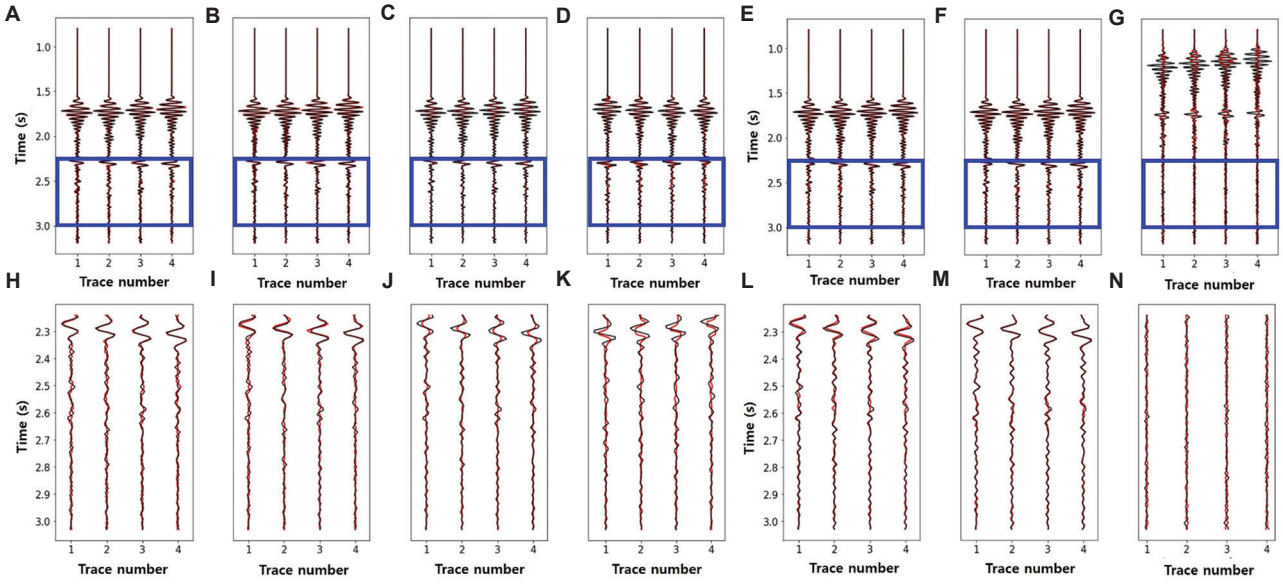

Figure 7. Comparison of traces for Figure 5. The black line represents the trace for the ground truth, and the red line represents the trace for the prediction result. (A) Bicubic. (B) f–x interpolation. (C) Projection onto convex sets. (D) Constrained diffusion-driven deep image prior. (E) 3D U-Net. (F) 3D U-Net+. (G) 3D U-Net+ (linear moveout [LMO] not applied). (H–N) Enlarged views of the blue box in (A), (B), (C), (D), (E), (F), and (G), respectively.

and also achieve high precision in reconstructing small amplitude events with steep slopes. In contrast, both the bicubic method, f–x interpolation method and POCS method struggle with the interpolation of small amplitude events with steep slopes. CDDIP is a self-supervised interpolation method whose performance is highly sensitive to hyperparameter choices, which can substantially affect the quality of reconstructed results. In our experiments, we adopted the publicly available code with default hyperparameters. As shown in Figure 7, while CDDIP successfully predicted the overall seismic signals, its reconstruction accuracy in fine-scale details was relatively low compared to our proposed method.

Table 2. PSNR and SSIM results

Method	PSNR (standard deviation)	SSIM (standard deviation)
bicubic	30.35 (3.93)	0.60 (0.07)
f–x interpolation	30.79 (3.22)	0.61 (0.11)
POCS	31.21 (3.56)	0.62 (0.14)
CDDIP	31.59 (4.08)	0.68 (0.22)
3D U-Net	32.23 (1.11)	0.71 (0.21)
3D U-Net+	32.86 (3.65)	0.74 (0.11)
3D U-Net+ (without LMO)	20.36 (1.71)	0.44 (0.01)

Abbreviations: CDDIP: Constrained diffusion-driven deep image prior; LMO: Linear moveout; POCS: Projection onto convex sets; PSNR: Peak signal-to-noise ratio; SSIM: Structural similarity index.

For quantitative performance evaluation, we used peak signal-to-noise ratio (PSNR) as performance metrics, which are calculated using following equations:

$$PSNR = 10log_{10}(peakval^2)/MSE \quad \text{(VII)}$$

Where peakval (peak value) is the maximum value in the data. A larger PSNR indicates better quality of the data.

Table 2 presents the PSNR results for the test dataset, with the 3D U-Net+ model demonstrating higher value of 32.86 dB, calculated using data normalized to a unit peak amplitude (max = 1). Consequently, it has been demonstrated that, even in the absence of label data for actual field data, high-accuracy interpolation can be

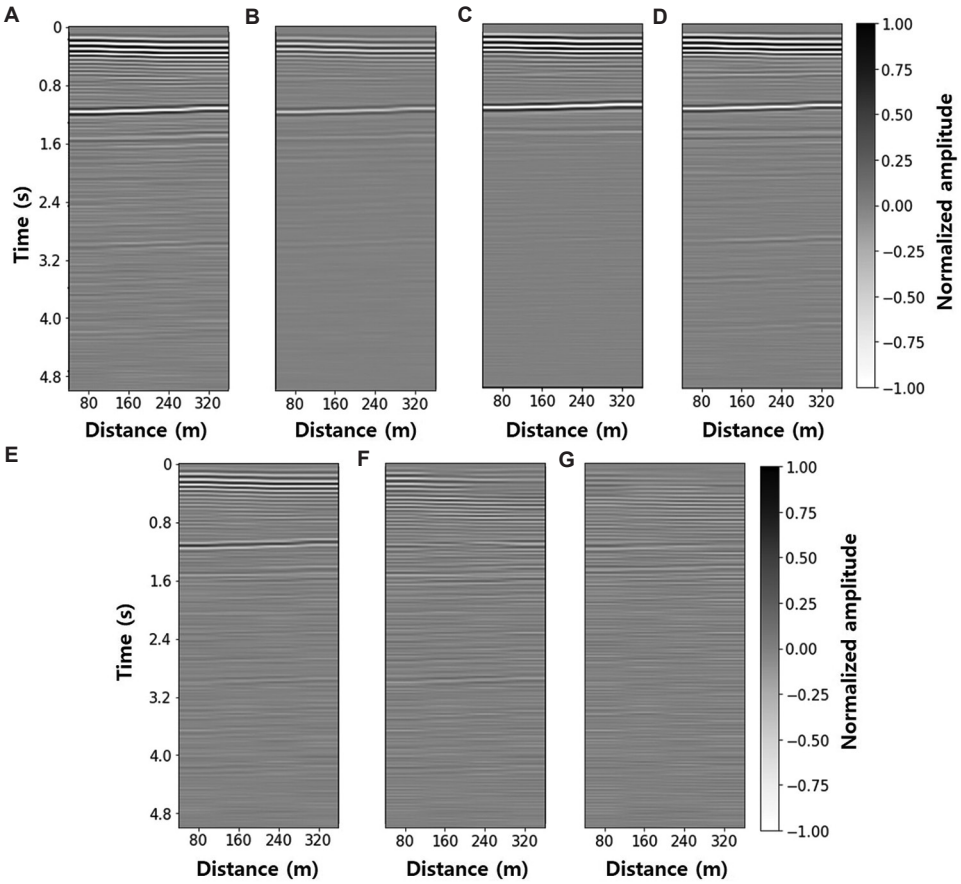

Figure 8. Interpolation results for crossline receiver direction within common-shot gathers with 5% Gaussian noise. (A) Ground truth. (B) Constrained diffusion-driven deep image prior (C) 3D U-Net. (D) 3D U-Net+. (E) Difference between (A) and (B). (F) Difference between (A) and (C). (G) Difference between (A) and (D).

performed through the transposed arrangement strategy-based deep learning interpolation method.

Figure 8 presents additional experiments conducted by adding 5% Gaussian noise to the original test data and evaluating the interpolation performance of CDDIP, 3D U-Net, and 3D U-Net+, which previously demonstrated relatively higher accuracy. The results show that all three methods were able to effectively handle a certain level of noise, but the proposed 3D U-Net+ produced smaller residuals compared to both CDDIP and 3D U-Net, indicating superior robustness in noisy conditions. Nevertheless, to obtain even better results, additional denoising would be necessary.

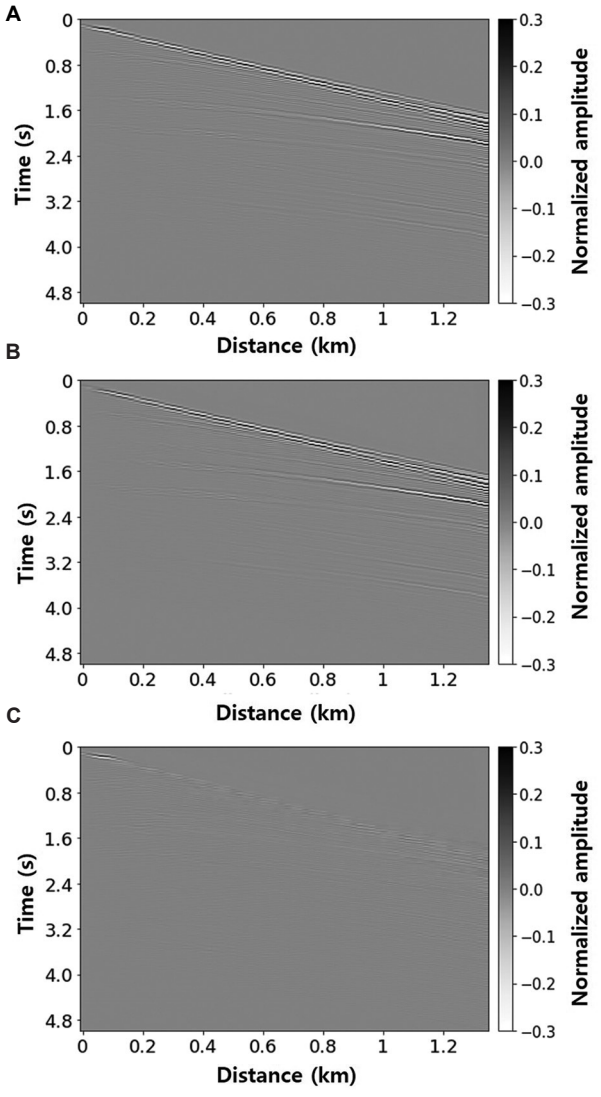

Figure 9. Inline receiver direction within common-shot gathers with the interpolated crossline location of 3D U-Net+. (A) Ground-truth. (B) Interpolation result. (C) Difference between (A) and (B).

Table 3 presents a comparison of the computational costs of the proposed 3D U-Net and 3D U-Net+ models against conventional interpolation methods, including bicubic, f–x interpolation, POCS, and CDDIP. The experiments were conducted on a server equipped with four NVIDIA TITAN RTX GPUs (each with 24 GB memory), running CUDA version 12.4 and driver version 550.76. The reported memory usage was measured as the peak GPU/CPU memory consumption during training and inference. The table summarizes training time (GPU hours), inference time, and GPU/CPU memory usage for each method.

Figure 9 shows the inline receiver direction within common-shot gathers with the reconstructed crossline position, combined with the output patches from the 3D U-Net+ model, which is identified as our optimal model. The results demonstrate that the proposed model is able to accurately predict the overall seismic signals, effectively capturing both strong reflections and subtle structures. Figure 10 presents the f–x spectrum corresponding to Figure 9, illustrating the frequency–wavenumber characteristics of the reconstructed seismic section. Although certain high-frequency components are not fully recovered, the overall spectrum closely resembles that of the label data, indicating that the proposed method preserves the dominant frequency content and structural consistency. Figure 11 presents a further trace-based analysis of the results in Figure 9 to assess the accuracy of the interpolation at the individual trace level.

4. Discussion

Our proposed method is designed for cross-wavefield reconstruction in marine towed-streamer systems with regularly sampled data and a 1:2 spacing ratio between receivers and streamers. The experiment utilizes complete

Table 3. Computational cost comparison

Method	Training time (GPU hours)	Inference time (min)	GPU/CPU memory usage
Bicubic	-	~0.1	<1 GB
f–x interpolation		~0.5	<1 GB
POCS		~10	<2 GB
CDDIP	~200	~30	~6 GB (training)/<2 GB (inference)
3D U-Net	~120	~2	~23 GB (training)/<1 GB (inference)
3D U-Net+	~150	~2	~24 GB (training)/<1 GB (inference)

Abbreviations: CDDIP: Constrained diffusion-driven deep image prior; POCS: Projection onto convex sets.

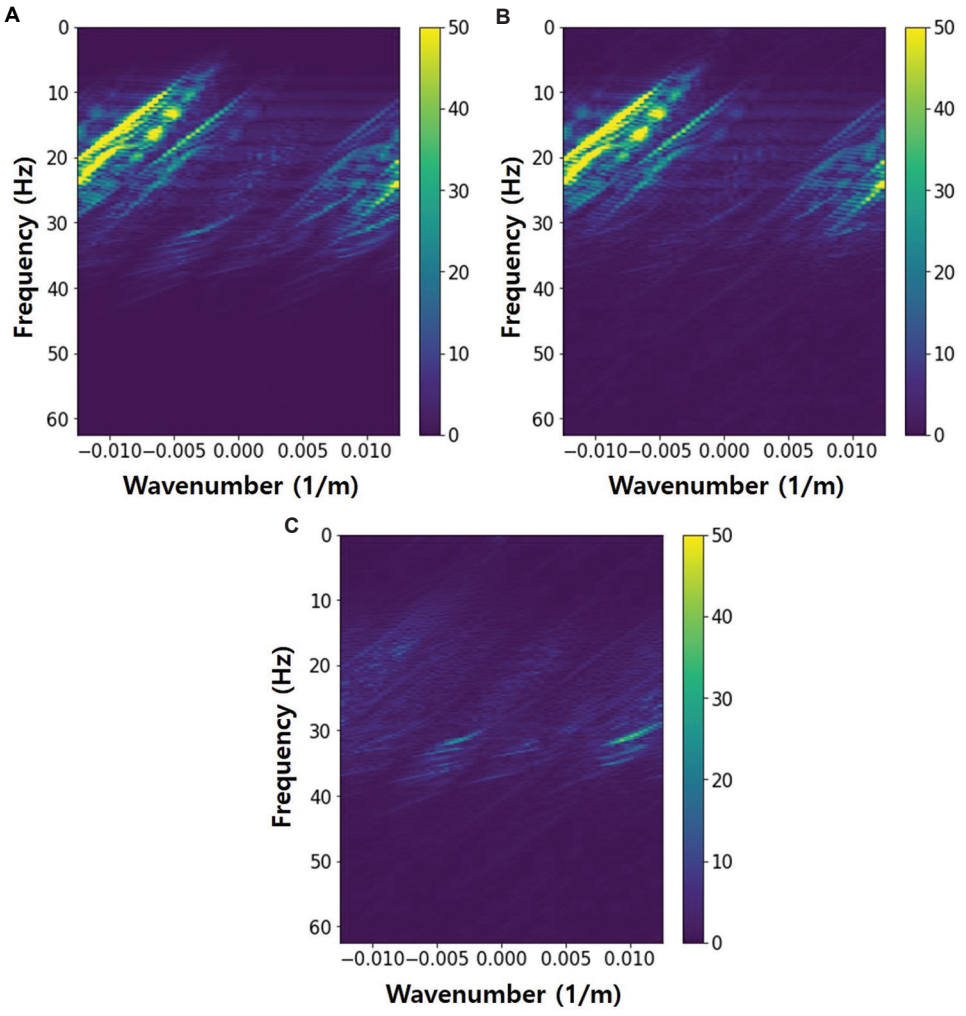

Figure 10. f–x spectrum corresponding to Figure 9. (A) Ground-truth. (B) Interpolation result. (C) Difference between (A) and (B).

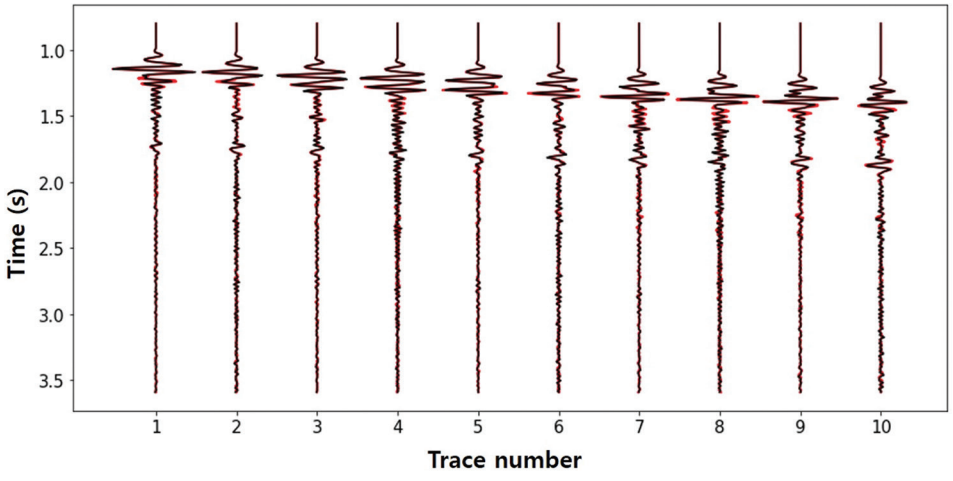

Figure 11. Comparison of traces 40–49 (out of 624) extracted from the gathers shown in Figure 9. The black line represents the trace for the ground truth, and the red line represents the trace for the prediction result.

common-shot gathers from towed-streamer data; however, regularization may be necessary if missing data is present. In addition, our method can be extended to CMP gathers or seismic image volumes, provided that the data maintain a regular 1:2 spacing ratio. If the spacing ratio is 2^n, where n >1, our method can be applied recursively to interpolate the data to a 1:1 ratio. If the spacing ratio is not in the form of $1:2^n$, the transposed arrangement strategy may be difficult to apply, and additional data processing may be necessary to adjust to a $1:2^n$ ratio. However, it should be noted that recursive applications may accumulate errors, potentially impacting accuracy.

In addition, the reconstruction performance of the proposed strategy is highly dependent on the quality of inline receiver direction within common-shot gathers, since dense inline traces are employed as surrogate labels during training. If the inline receiver direction within common-shot gathers are noisy or contain structural discontinuities, the errors may propagate into the crossline reconstruction. The method may also be sensitive to strong random or coherent noise, which could reduce accuracy in weak-amplitude regions. Future improvements could incorporate denoising pre-processing or noise-aware loss functions to enhance robustness.

The experiments were conducted on a server equipped with four NVIDIA TITAN RTX GPUs (each with 24 GB memory), running CUDA version 12.4 and driver version 550.76. During training, each model instance occupied approximately 23 GB of GPU memory per device, nearly exhausting the available capacity of a single TITAN RTX card. This high memory demand arises mainly from the use of 3D convolutional kernels and large input patches (*e.g.*, 4 × 7 × 624 traces). Therefore, multi-GPU environments with sufficient memory capacity (≥24 GB per GPU) are strongly recommended for training 3D U-Net and U-Net+ architectures with the proposed transposed arrangement strategy. In this study, the batch size during training was set to 128, and we note that reducing the batch size would alleviate the memory issue. In contrast, the inference stage required <1 GB of GPU memory, making the trained models practical for deployment in real-world seismic data processing tasks.

The seismic dataset used in this study consists of 50 sail lines, each containing 96 shots, 8 streamers, and 68 receivers per streamer. Each receiver records 625 samples at an 8 ms sampling interval, resulting in a total data volume of approximately 6.1 GB for a single-component (float32) dataset.

The proposed 3D U-Net+ model required about 24 GB of GPU memory per device during training and was trained using four TITAN RTX GPUs for approximately 150 GPU h. In contrast, the inference phase required <1 GB of GPU memory, demonstrating that the proposed framework can process seismic data efficiently even at moderate survey scales. Thus, although model training requires high-performance GPUs (≥24 GB each), the trained network can be efficiently applied to large-scale seismic data in practical production environments. This method was developed for streamer data on regular grids and is not directly applicable to irregular land geometries.

5. Conclusion

This study introduced the transposed arrangement strategy, a novel approach designed to enhance the performance of crossline interpolation. By adjusting input array configurations during the training and inference phases, our method successfully reconstructed crossline receiver direction within common-shot gathers without the need for labeled data in the crossline receiver direction within common-shot gathers. To address feature differences between inline and crossline receiver direction within common-shot gathers, we applied LMO correction, enabling effective crossline interpolation using a model trained with inline receiver direction within common-shot gathers.

The interpolation performance of our method was rigorously evaluated by comparing it with traditional bicubic interpolation methods, f–x interpolation method, and POCS using deep learning models, 3D U-Net and 3D U-Net+. The results demonstrated that 3D U-Net+ provided the clearest and most accurate interpolations, closely resembling the original data. Numerically, 3D U-Net+ also exhibited superior PSNR values, confirming its efficacy.

This research confirms that applying the proposed method in deep learning-based models significantly enhances the accuracy of crossline interpolation, suggesting a promising direction for future advancements in seismic data processing.

Acknowledgments

None.

Funding

This work was supported by the Korea Institute of Marine Science and Technology Promotion grant funded by the Ministry of Oceans and Fisheries (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: Yonghwan Joo, Daeung Yoon
Formal analysis: Jiyun Yu
Investigation: Jiyun Yu
Methodology: Yonghwan Joo, Daeung Yoon
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).

Further disclosure

Part of the findings of this study were presented at the 2023 Fall Joint Conference and the 9th International Symposium on Mine Reclamation, held at Pyeongchang on November 03, 2023.

References

1. Ikelle LT, Amundsen L. *Introduction to Petroleum Seismology*. Society of Exploration Geophysicists; 2005.

   doi: 10.1190/1.9781560803447.ch1

2. Keys R. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust.* 1981;29(6):1153-1160.

   doi: 10.1109/TASSP.1981.1163711

3. Ronen J. Wave-equation trace interpolation. *Geophysics*. 1987;52(7):973-984.

   doi: 10.1190/1.1442366

4. Fomel S. Seismic reflection data interpolation with differential offset and shot continuation. *Geophysics*. 2003;68(2):733-744.

   doi: 10.1190/1.1567243

5. Spitz S. Seismic trace interpolation in the F-X domain. *Geophysics*. 1991;56(6):785-794.

   doi: 10.1190/1.1443096

6. Spitz S. Pattern recognition, spatial predictability, and subtraction of multiple events. *Geophysics*. 1999;64(2): 431-436.

   doi: 10.1190/1.1438154

7. Bregman L. The method of successive projection for finding a common point of convex sets (theorems for determining common point of convex sets by method of successive projection). *Sov Math*. 1965;6:688-692.

8. 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

9. Naghizadeh M, Sacchi MD. Multistep autoregressive reconstruction of seismic records. *Geophysics*. 2007;72(6):V111-V118.

   doi: 10.1190/1.2771685

10. Trickett S, Burroughs L, Milton A, Walton L, Dack R. Rank-Reduction-Based Trace Interpolation. In: *Conference: SEG Technical Program Expanded Abstracts 2010*; 2010. p. 3829-3833.

    doi: 10.1190/1.3513645

11. Gao J, Sacchi MD, Chen X. A fast reduced rank interpolation method for prestack seismic volumes that depend on four spatial dimensions. *Geophysics*. 2013;78(1):V21-V30.

    doi: 10.1190/geo2012-0038.1

12. Ma J. Three-dimensional irregular seismic data reconstruction via low-rank matrix completion. *Geophys*. 2013;78(5):V181-V192.

    doi: 10.1190/geo2012-0465.1

13. Abma R, Kabir N. 3D interpolation of irregular data with a POCS algorithm. *Geophysics*. 2006;71(6):E91-E97.

    doi: 10.1190/1.2356088

14. Gan S, Wang S, Chen Y, Zhang Y, Jin Z. Dealiased seismic data interpolation using seislet transform with low-frequency constraint. *IEEE Geosci Remote Sens Lett*. 2015;12(10): 2150-2154.

    doi: 10.1109/LGRS.2015.2453119

15. Liu W, Cao S, Gan S, Chen Y, Zu S, Jin Z. One-step slope estimation for dealiased seismic data reconstruction via iterative seislet thresholding. *IEEE Geosci Remote Sens Lett*. 2016;13(10):1462-1466.

    doi: 10.1109/LGRS.2016.2591939

16. Jia Y, Ma J. What can machine learning do for seismic data processing? An interpolation application. *Geophysics*. 2017;82:163-177.

    doi: 10.1190/geo2016-0300.1

17. Oliveira DAB, Ferreira RS, Silva R, Brazil EV. Interpolating seismic data with conditional generative adversarial networks. *IEEE Geosci Remote Sens Lett*. 2018;15:1952-1956.

    doi: 10.1109/LGRS.2018.2866199

18. Wang B, Zhang N, Lu W, Geng J, Huang X. Intelligent missing shots' reconstruction using the spatial reciprocity of green's function based on deep learning. *IEEE Trans Geosci Remote Sens*. 2020;58:1587-1597.

    doi: 10.1109/TGRS.2019.2947085

19. Park J, Choi J, Seol SJ, Byun J, Kim Y. A method for adequate selection of training data sets to reconstruct seismic data using a convolutional U-Net. *Geophysics*. 2021;86:375-388.

doi: 10.1190/geo2019-0708.1

20. Park J, Yeeh Z, Seol SJ, *et al.* Seismic data interpolation using attention-based deep learning. *Eur Assoc Geosci Eng*. 2022;2022:1-5.

doi: 10.3997/2214-4609.202210245

21. Yu J, Wu B. Attention and hybrid loss guided deep learning for consecutively missing seismic data reconstruction. *IEEE Trans Geosci Remote Sens*. 2021;60:1-8.

doi: 10.1109/TGRS.2021.3068279

22. Li X, Wu B, Zhu X, Yang H. Consecutively missing seismic data interpolation based on coordinate attention unet. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/LGRS.2021.3128511

23. Lou Y, Wu L, Liu L, *et al.* Irregularly sampled seismic data interpolation via wavelet-based convolutional block attention deep learning. *Artif Intellig 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. 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

26. Kaur H, Pham N, Fomel S. Seismic data interpolation using CycleGAN. *SEG Tech Prog Expand Abstracts*. 2019;38:5407.

doi: 10.1190/segam2019-3207424.1

27. Chang DK, Yang WY, Yong XS, Li S. Seismic data interpolation with conditional generative adversarial network in time and frequency domain. *SEG Tech Prog Expand Abstracts*. 2019;38:2589-2593.

doi: 10.1190/segam2019-3210118.1

28. Deng F, Wang S, Wang X, Fang P. Seismic data reconstruction based on conditional constraint diffusion model. *IEEE Geosci Remote Sens Lett*. 2024;21:1-5.

doi: 10.1109/LGRS.2024.3371675

29. 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;19:1-5.

doi: 10.1109/LGRS.2020.3044455

30. Jin Z, Li X, Yang H, Wu B, Zhu X. Depthwise separable convolution Unet for 3D seismic data interpolation. *Front Earth Sci*. 2023;10:1-18.

doi: 10.3389/feart.2022.1005505

31. Chang D, Yang W, Young X, *et al.* Seismic data interpolation using dual-domain conditional generative adversarial networks. *IEEE Geosci Remote Sens Lett*. 2020;18:1856-1860.

doi: 10.1109/LGRS.2020.3008478

32. Dou Y, Li K, Duan H, *et al.* MDA GAN: Adversarial-learning-based 3-D seismic data interpolation and reconstruction for complex missing. *IEEE Trans Geosci Remote Sens*. 2023; 61:1-14.

doi: 10.48550/arXiv.2204.03197

33. Yu J, Yoon D. Crossline reconstruction of 3D seismic data using 3D cWGAN: A comparative study on sleipner seismic survey data. *Appl Sci*. 2023;13:5999.

doi: 10.3390/app13105999

34. Ding M, Zhou Y, Chi Y. Self-attention generative adversarial network interpolating and denoising seismic signals simultaneously. *Remote Sens.* 2024;16:305.

doi: 10.3390/rs16020305

35. Wang S, Deng F, Jiang P, Gong Z, Wei X, Wang Y. SeisFusion: Constrained diffusion model with input guidance for 3-D seismic data interpolation and reconstruction. *IEEE Trans Geosci Remote Sens*. 2024;62:1-15.

doi: 10.48550/arXiv.2403.11482

36. Qian F, Liu Z, Wang Y, Liao S, Pan S, Hu G. DTAE: Deep tensor autoencoder for 3-D seismic data interpolation. *IEEE Trans Geosci Remote Sens*. 2021;60:1-19.

doi: 10.1109/TGRS.2021.3075968

37. Larsen Greiner TA, Hlebnikov V, Lie JE, *et al.* Cross-streamer wavefield reconstruction through wavelet domain learning. *Geophysics*. 2020;85:457-471.

doi: 10.1190/geo2019-0771.1

38. Yeeh Z, Yoon D, Byun J. Cross-streamer wavefield reconstruction of a towed streamer system using bidirectional LSTM networks with a traces-to-trace approach. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

doi: 10.1109/LGRS.2023.3312712

39. Chen Y, Yu S, Lin R. Self-supervised transfer learning POCS-net for seismic data interpolation. *IEEE Trans Geosci Remote Sens*. 2024;62:1-18.

doi: 10.1109/TGRS.2024.3494723

40. Wang S, Wu X, Chen J. Dip-informed neural network for self-supervised anti-aliasing seismic data interpolation. *IEEE Trans Geosci Remote Sens*. 2024;62:1-4.

doi: 10.1109/TGRS.2024.3359247

41. Goyes-Peñafiel P, Kamilov US, Arguello H. CDDIP:

Constrain diffusion-driven deep image prior for seismic data reconstruction. *IEEE Geosci Remote Sens Lett.* 2025;22:1-5.

doi: 10.48550/arXiv.2407.17402

42. Souza WE, Cerqueira AG, Porsani MJ. First-break prediction in 3-D land seismic data using the dynamic time warping algorithm. *Geophys J Int*. 2024;237:402-418.

doi: 10.1093/gji/ggae048

43. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Cham: Springer; 2015. p. 234-241.

44. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A nested U-Net architecture for medical image segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support*. 2018;11045:3-11.

doi: 10.48550/arXiv.1807.10165

45. Ji J, Choi YG. 3D seismic data processing methodology using public domain software system. *Geophys Geophys Explor.* 2010;13:159-168.