Seismic resolution enhancement with self-supervised learning

*Shijun Cheng * and Tariq Alkhalifah, King Abdullah University of Science and Technology*
Haoran Zhang, China University of Petroleum-Beijing

SUMMARY

The concept of neural network (NN)-based seismic resolution enhancement has gained a lot of traction recently. Yet, the majority of work rely on training NNs on synthetic data via a supervised learning strategy, often encountering generalization issues on real data. To address this problem, we develop a self-supervised learning method for seismic resolution enhancement. Specifically, we reinterpret seismic resolution enhancement as a frequency extension task, particularly focusing on the reconstruction of high-frequency components. Initially, we warm up the NN using the original bandwidth-limited data as pseudo labels, with input data derived by filtering out high-frequency elements from the original data. Subsequently, the network undergoes iterative data refinement, where pseudo labels are predicted from the NN trained in the previous epoch on the original data, and input data are obtained by filtering high-frequency components from these predictions. The efficacy of our method is demonstrated through tests on both synthetic and field data.

INTRODUCTION

Seismic high-resolution images are crucial for accurately mapping subsurface structures and characterizing complex geological targets. However, due to the filtering effects of geological layers, noise, and limitations of seismic acquisition, the recorded seismic data are often bandwidth-limited, resulting in lower resolution. Therefore, effectively enhancing seismic resolution represents a significant challenge to address.

In the realm of conventional methods, numerous techniques have been developed to enhance seismic resolution. The least squares deconvolution (LSD) stands out as one of the most classic approaches, working to recover high-frequency content by minimizing the discrepancy between real seismic data and model predictions (Berkhout, 1977). While it can stably estimate high-frequency information, its capacity to enhance resolution is limited and it remains sensitive to noise. Alternatively, some studies have presented sparse spike inversion (SSI) techniques to extend the high-frequency in seismic data (Sacchi, 1997; Gholami and Sacchi, 2012). Typically, SSI assumes that the subsurface reflectivity consists of a sequence of discrete reflection spikes, and the objective is to use the minimal number of wavelet-convolved reflection spikes to simulate a seismic trace. It was initially applied on a trace-by-trace basis, resulting in some limitations in spatial continuity of the processing results. Consequently, further research resulted in multi-trace methods, thus offering a more stable high-resolution reconstruction product. Although SSI can improve seismic resolution, it shares a common limitation with LSD: the need to accurately estimating the seismic wavelet. However, the estimation of seismic wavelets for real data is often hard.

With the rapid development of deep learning algorithms in the field of seismic processing (Cheng et al., 2023b), some researchers have introduced techniques for enhancing seismic resolution based on neural networks (NNs) (Li et al., 2021; Gao et al., 2022). The key concept of this technology is to train an NN to approximate the nonlinear relationship between low-resolution and high-resolution data. Most efforts involve training a neural network on synthetic data using a supervised learning paradigm, which is then applied to enhance the resolution of field data. However, due to the difference in feature distribution between synthetic and field data, NNs trained on synthetic data often face generalization issues. To bridge the gap between synthetic and field data, Zhang et al. (2022) introduced a domain adaptation (DA) algorithm, inspired by MLReal (Alkhalifah et al., 2021), to transform synthetic and field data to have similar features (distributions). However, this DA algorithm might eliminate some characteristics of seismic data, such as phase, during the transformation process.

Therefore, a better alternative is to directly train training on field data, and thus, enable the NN to learn frequency characteristics directly from seismic data, thereby contributing to the resolution enhancement in field data. To the best of our knowledge, we have found only two contributions utilizing self-supervised learning (SSL) methods for seismic resolution enhancement. Chai et al. (2023) incorporated the convolutional model into the loss function in a self-supervised manner to constrain the network's predicted outcomes to match observational data. Similarly, Wang et al. (2023) integrated the Robinson convolutional model into the loss function to provide physical constraints, while also incorporating structural and sparsity constraints to train the NN in an SSL manner. Although the performances of both methods were validated on field data, they rely on estimating a relatively accurate wavelet, as they both incorporate convolutional models into the loss function.

In this abstract, we propose a novel SSL seismic resolution enhancement method that does not rely on wavelet estimation. As previously mentioned, a critical component of resolution enhancement is the extension of the frequency band information in seismic data. Thus, in our approach, we reinterpret seismic resolution enhancement as a problem of high-frequency extension. We rely on an iterative refinement mechanism to learn the data as we elevate the frequency. Furthermore, we present a multi-loss constraint to stabilize the network training and enhance its performance. We will demonstrate the effectiveness of our method on both synthetic and field data.

METHOD

Our framework principally consists of two stages: a warm-up and an iterative data refinement (IDR) phases. The whole workflow is illustrated in Figure 1. In the following, we will elucidate the key components therein.

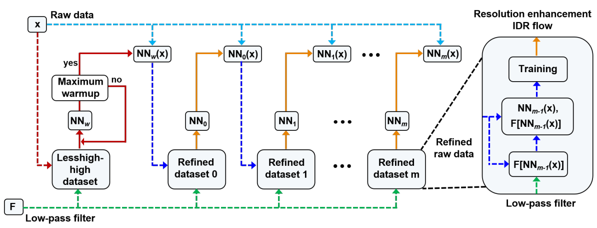

Figure 1: An illustration of the self-supervised seismic resolution enhancement workflow.

Firstly, the NN undergoes a warm-up phase. We first apply a low-pass filter to the original seismic data $\{x_i\}_{i=1}^{N}$, which is assumed already to be of low-resolution with a limited frequency band, thereby creating a training dataset $\{\mathtt{F}[x_i], x_i\}_{i=1}^{N}$ (we call lesshigh-high (LH2H)). In which, the original data become pseudo-labels, and the filtered results are input data. Subsequently, the NN is subjected to multiple epochs of optimization on this dataset. The objective of this training is to enable rapid stabilization of the NN, allowing it to preliminarily capture the characteristics of seismic data. Furthermore, this pre-trained network, denoted as $\mathtt{NN}_w$, has a certain degree of high-frequency extension ability with respect to the original seismic data. This capability forms the groundwork for the iterative refinement of the training set in subsequent stages.

Subsequently, the NN enters the IDR phase. In this stage, we first leverage the pre-trained network $\mathtt{NN}_w$ to predict the original seismic data. The predictions serve as the initial pseudo-labels for the IDR phase, with corresponding inputs derived from applying a low-pass filter to these predictions. This procedure facilitates the creation of a new LH2H dataset, employed for the first epoch of training during the IDR stage:

$$\mathtt{NN}_0 \leftarrow \{(\mathtt{F}[\mathtt{NN}_w(x_i)],\ \mathtt{NN}_w(x_i))\}_{i=1}^{N}, \qquad (1)$$

where the network $\mathtt{NN}_0$ is directly initialized from the pertrained network $\mathtt{NN}_w$.

We anticipate that after training for one epoch on the new training set, the high-frequency extension capability of $\mathtt{NN}_0$ will slightly surpass that of $\mathtt{NN}_w$. This improvement is attributed to the fact that, compared to the LH2H dataset used during the warm-up phase, the new LH2H dataset exhibits higher frequency representation in the labels. During the subsequent training, we iteratively perform this process to progressively diminish the frequency information bias between the predicted pseudo label and the broad band ground-truth data. Specifically, for each training epoch, we first employ the network trained in the previous epoch (e.g., $\mathtt{NN}_{j-1}$) to predict the original data $\{x_i\}_{i=1}^{N}$, and thus, obtain the frequency-extended pseudo-labels $\{\mathtt{NN}_{j-1}(x_i)\}_{i=1}^{N}$. Then, we apply a low-pass filter to these pseudo-labels to generate the corresponding input data $\{\mathtt{F}[\mathtt{NN}_{j-1}(x_i)]\}_{i=1}^{N}$. The resulting LH2H training set $\{(\mathtt{F}[\mathtt{NN}_{j-1}(x_i)],\ \mathtt{NN}_{j-1}(x_i))\}_{i=1}^{N}$ will optimize the network $\mathtt{NN}_j$ for one epoch, for example,

$$\mathtt{NN}_j \leftarrow \{(\mathtt{F}[\mathtt{NN}_{j-1}(x_i)],\ \mathtt{NN}_{j-1}(x_i))\}_{i=1}^{N}. \qquad (2)$$

After conducting multiple epochs of training in the IDR phase, our network incrementally enhances its high-frequency extension capabilities.

During the training process, we use a hybrid loss function to co-optimize the network. This hybrid loss function consists of a data loss $\mathscr{L}_d$, a focal frequency loss $\mathscr{L}_f$, and a sparsity-promotion loss $\mathscr{L}_s$. The data loss measures the difference between the NN's outputs O_i, $i = 1, \cdots, N$ and the corresponding pseudo-labels L_i, $i = 1, \cdots, N$, using the mean absolute error (MAE) as follows:

$$\mathscr{L}_d(L, O) = \frac{1}{N}\sum_{i=1}^{N} |L_i - O_i|. \qquad (3)$$

The focal frequency loss $\mathscr{L}_f$, which is developed by Jiang et al. (2021), allows an NN to adaptively focus on frequency components that are hard to represent by down-weighting the easy ones. It can adaptively emphasize the loss weight of the high-frequency components, and thus, mitigate the network's low-frequency bias characteristic. The loss $\mathscr{L}_f$ has the form:

$$\mathscr{L}_f(L, O) = \frac{1}{N}\sum_{i=1}^{N} \left|\tilde{L}_i - \tilde{O}_i\right| \cdot \left|\tilde{L}_i - \tilde{O}_i\right|^2, \qquad (4)$$

where the $\tilde{L}_i$ and $\tilde{O}_i$ represent the the spatial frequency spectrum for the pseudo-labels and network outputs, respectively. The sparsity-promotion loss can be expressed as follows:

$$\mathscr{L}_d(O) = \frac{1}{N}\sum_{i=1}^{N} |O_i|. \qquad (5)$$

The total loss function is defined as

$$\mathscr{L}(L, O) = \mathscr{L}_d(L, O) + \varepsilon_1 \cdot \mathscr{L}_f(L, O) + \varepsilon_2 \cdot \mathscr{L}_s(L, O), \qquad (6)$$

where ε_1 and ε_2 are hyperparameters, which are used to regulate the proportion of the focal frequency and sparsity-promotion losses within the total loss.

NUMERICAL EXAMPLES

To verify the efficacy of the proposed algorithm, we first use synthetic data generated with the Marmousi model to perform the tests. The modified model has 2650 traces and 934 sampling points with a time interval of 1.0 ms. We generate synthetic seismic data by convolving the reflectivity of the Marmousi model with a Ricker wavelet of 20 Hz peak frequency, which we consider as the original bandwidth-limited low-resolution data and is shown in Figure 2a. From the generated low-resolution data, we extract a total of 2236 data patches, each sized 128×128. During the warm-up phase, we randomly filter out frequency components higher than 20 $\sim$ 40 Hz from the original data to generate the input data. As previously described, the pseudo-label data at this time are the original observed data. In the IDR phase, the low-pass filter cutoff frequency range is still set between $20 \sim 40$ Hz. Within this frequency range, we randomly filter out the high frequencies from the pseudo-label data predicted by the network trained in the previous epoch, which serve as the input data. The network is trained for a total of 120 epochs, with the warm-up phase accounting for 20 epochs. We chose a learning rate of 2e-4 at the start, and then decrease by a factor of 0.8 at the 15, 30, and 60 epochs.

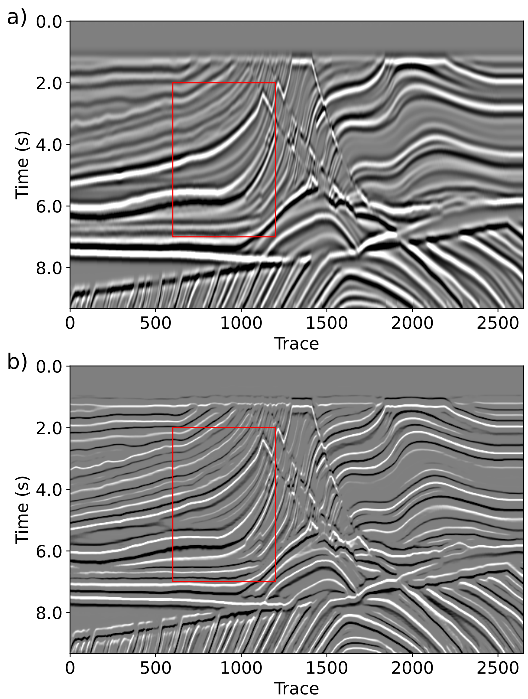

Figure 2: (a) The original low-resolution synthetic data. (b) The high-resolution prediction on the original synthetic data using our framework.

The prediction result of the raw data, derived from the network trained using our framework, is illustrated in Figure 2b. We can see that our method can significantly enhance the resolution of raw seismic data. As a result, it can offer a more refined characterization of subsurface structures. To present a more detailed comparison, a zoomed-in section of interest marked by the red boxes is shown in Figure 3. It is evident that our method delivers a high-resolution predictive output closely resembling the reflectivity model. To clearly verify the performance of the proposed algorithm in reconstructing the high-frequency components, we perform a spectral analysis between the original and predicted data, which is displayed in Figure 4. We emphasize that the amplitude spectrum represents the average value obtained after summing the amplitude spectrum of all traces. We can see that the amplitude spectrum of predicted data exhibits a good frequency extension after applying the proposed algorithm. In the low-frequency range, such as $0 \sim 40$ Hz, the amplitude spectrum of the prediction result and the original data generally align, indicating that we did not compromise the low-frequency information while expanding the high-frequency components. Therefore, the results demonstrate the accuracy, stability and effectiveness of the proposed SSL seismic resolution enhancement algorithm.

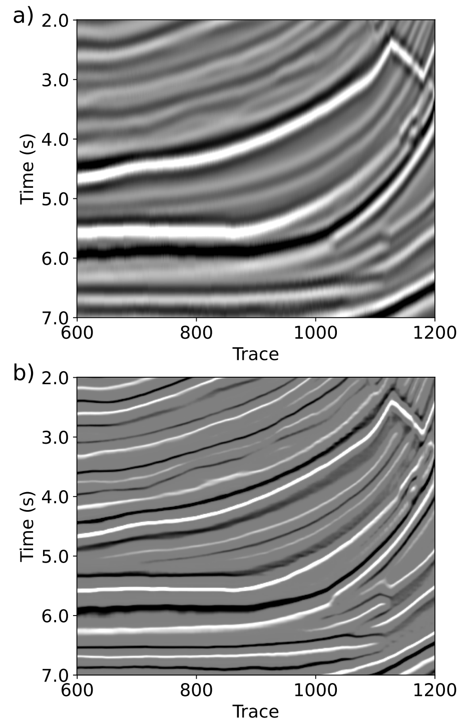

Figure 3: Zoomed-in view corresponding to the red box areas in Figure 2: (a) The original low-resolution synthetic data. (b) The high-resolution prediction on the original synthetic data.

We then test our method on a post-stack time-migrated image (see Figure 5a) from a China land field dataset. For training, we extract 528 patches, each sized 128×128, from the original image. Here, we observe that, in comparison to synthetic data, field data is contaminated with noise, posing a challenge for the resolution enhancement task. To better handle noisy field data, we incorporate a denoiser within our framework. A recent SSL paradigm utilizing IDR has been demonstrated to address various type of seismic noise effectively (Cheng et al., 2023a). Hence, by combining this SSL denoising framework with our resolution enhancement approach, we can achieve simultaneous seismic denoising and resolution enhancement. Specifically, during the warm-up and IDR phases, after filtering out the high-frequency components from the corresponding pseudo-label data, noise is added to form the input data. Given our prior knowledge that the field data is contaminated with noise, we employ the following formula $n_i = 0.01\varepsilon \cdot std(y_i) \cdot rand(0,1)$ to generate random noise, where ε denotes the noise level, $std(y_i)$ represents the standard deviation of extracted data patches y_i, and $rand(0,1)$ is the standard normal distribution. During both warm-up and IDR phases, the input data are generated by randomly filtering out frequencies higher than $20 \sim 60$ Hz from the pseudo-label data, and also, introducing the noise with a level between $20 \sim 100$. To achieve faster convergence, we directly employ models trained on synthetic data as initialization. The network is trained for a total of 50 epochs, where the network is pre-trained for 20 epochs in the warm-up phase. The initial learning rate is 1e-4, which is reduced by a factor of 0.6 at 10 and 25 epochs.

Figure 5b displays the processing results on the original field data by our method. We can see that our method significantly enhances seismic resolution, particularly for shallow reflectors, while effectively removing noise with the aid of the embedded denoiser. Figure 6 shows a zoomed-in view of the area marked by a red box in Figure 5, which clearly demonstrates the substantial improvement in the quality of the original imaging data achieved by our method. Meanwhile, Figure 7 presents the amplitude spectrum of the original data and our processing product. It is evident that our approach notably extends the high-frequency content of the original image, while preserving the low-frequency features of the original data.

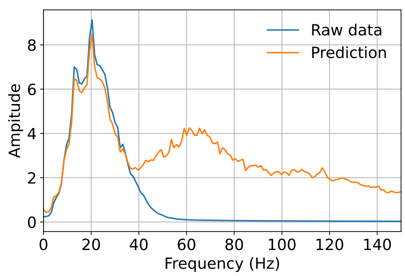

Figure 4: Amplitude spectrum comparison between the raw synthetic data and the high-resolution prediction.

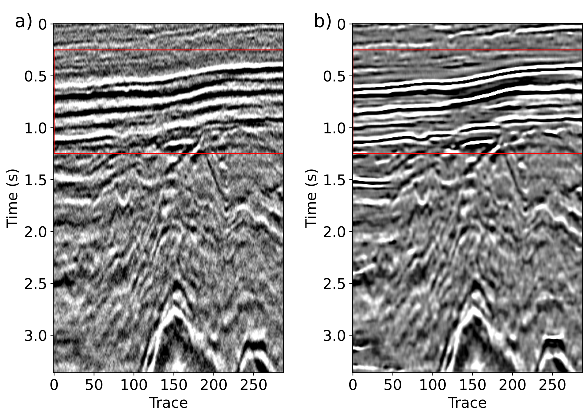

Figure 5: (a) The original post-stack time-migrated image, which comes from a China land field dataset. (b) The high-resolution prediction on the original image using our framework.

CONCLUSIONS

We developed a novel neural network (NN)-based seismic resolution enhancement method trained in a self-supervised learning (SSL) fashion. Under our framework, the NN sequentially undergoes two stages: warm-up and iterative data refinement (IDR). In the warm-up stage, we construct a lesshigh-high (LH2H) dataset, using the original observed data, which has a limited frequency band, as pseudo-labels. The input data

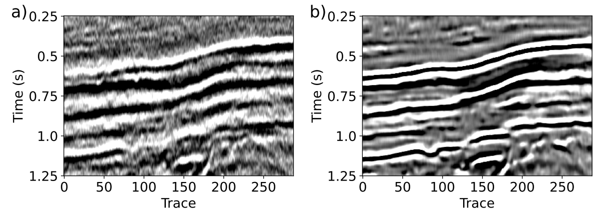

Figure 6: Zoomed-in view corresponding to the red box areas in Figure 5: (a) The original low-resolution field data. (b) The high-resolution prediction on the original field data.

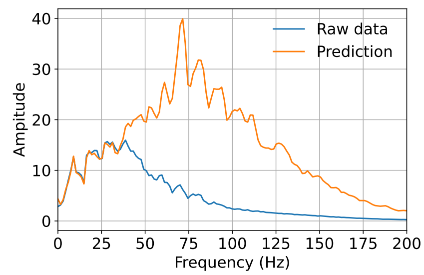

Figure 7: Amplitude spectrum comparison between the raw field data and the high-resolution prediction.

are obtained by applying a low-pass filter to these pseudo-labels, resulting in a further loss of high-frequency content. The NN rapidly warms up on this constructed dataset, initially extracting the original data's frequency characteristics and providing a degree of high-frequency extension capability. During the IDR stage, we update the LH2H dataset in each training epoch, where the pseudo-labels are derived from the network's predictions of the original data from the previous epoch, and the input data are created by applying a low-pass filter to these predicted pseudo-labels. Continually updating the training dataset allows us to gradually reduce the frequency information bias between the network's predictions and the ideal ground truth, and thus, steadily enhancing the network's high-frequency extension performance, thereby improving the seismic resolution. The test results demonstrated that our method effectively enhance seismic resolution.

ACKNOWLEDGMENTS

The authors thank KAUST and the DeepWave Consortium sponsors for supporting this research. We also thank SWAG members for the helpful discussions.