Proceedings of the IASS 2024 Symposium
Redefining the Art of Structural Design
August 26-30, 2024, Zurich Switzerland
Philippe Block, Giulia Boller, Catherine DeWolf,
Jacqueline Pauli, Walter Kaufmann (eds.)

A Deep Learning-based Compressive Sensing Method for Vibration Monitoring of Spatial Structures

Guan-Sen Dong*, Hua-ping Wan, Yaozhi Luo

*College of Civil Engineering and Architecture, Zhejiang University
Hangzhou 310058, China
guansendong@zju.edu.cn

Abstract

Vibration monitoring is pivotal in Structural Health Monitoring (SHM) of spatial structures. However, the acquisition of high-frequency vibration data encounters challenges due to hardware limitations. Compressive Sensing (CS) is an innovative data collection and compression technique, which typically relies on the sparsity of signals. Nevertheless, the complex mechanical characteristics of spatial structures often lead to vibration signals that are not sparse, which degrades CS performance. To address this issue, we propose a novel deep learning-based CS method using Deep Convolutional Generative Adversarial Networks (DCGAN). This model learns an end-to-end mapping between compressed and original signals, thus eliminating the dependence on signal sparsity. The DCGAN architecture includes a generator that reconstructs compressed data and a discriminator that facilitates the training of the generator. The effectiveness of the proposed method is validated using vibration data from a large test steel grandstand. The results indicate that the proposed DCGAN model have high recovery accuracy, highlighting its potential for vibration monitoring of spatial structures.

Keywords: Structural health monitoring, vibration monitoring, compressive sensing, deep learning.

1. Introduction

Compressive sensing (CS) is a novel framework for data acquisition, compression, and recovery [1]. It merges data compression directly into the collection process, requiring only a few random samples to represent the compressed signal. It then reconstructs the complete signal by leveraging the inherent sparsity of the signal [2]. CS breaks the traditional Shannon-Nyquist sampling theorem, reducing the sampling frequency without losing important information. Its applications span various fields [3], including computer vision, signal processing, and medical image processing.

In structural engineering, CS has been effectively applied to vibration monitoring, since the structural vibration signals are often sparse in the frequency domain due to the modal characteristics. It offers potential solutions to challenges such as limited transmission bandwidth, storage capacity, and sensor battery life in structural health monitoring (SHM) systems [4, 5].

However, as for the spatial structures, the application of CS encounters a challenge. Unlike relatively simpler structural forms like bridges or tall buildings, spatial structures feature more components, complex joint connections, and intricate mechanical behaviors, leading to more complex vibration signal spectra. This complexity often renders the foundational assumption of signal sparsity in CS inappropriate, resulting in high recovery errors in practical applications.

Recent advancements [6, 7] in deep learning-based CS methods within the computer vision field provide a novel solution to these issues. These deep models learn direct mappings from extensive sets of low- and high-resolution exemplar pairs, bypassing the need for inherent signal sparsity and random sampling.

Inspired by these successes, this study introduces a novel CS approach for spatial structure vibration data using Deep Convolutional Generative Adversarial Networks (DCGAN). Unlike traditional sparsity-based CS methods, the proposed DCGAN approach does not rely on signal sparsity, thereby accommodating the complex frequency characteristics of spatial structure vibrations. The effectiveness of our DCGAN model is demonstrated through its application to vibration data from a test steel grandstand, showing significant improvements in accuracy in signal recovery.

2. GAN for CS of vibration data from spatial structures

2.1. Overview of CS theory

Considering a original vibration signal $\mathbf{x}$, a limited number of measurements can be collected as the compressed signal $\mathbf{y}$, expressed by:

$$\mathbf{y} = \mathbf{\Phi x} \tag{1}$$

where $\mathbf{\Phi}$ is a sampling matrix with far fewer rows than columns representing the down-sampling. The traditioanl sparisty-based CS methods requires that $\mathbf{\Phi}$ is a random matirx. However, the deep learning-based method eliminates the need for the random nature. Therefore, the compressed signal $\mathbf{y}$ can be obtained by decreasing the sampling frequency strightforward.

The general CS problem aims to recover the complete singal from its compressed form, formulated as:

$$\mathbf{x} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi x}\|_2^2 + \mathcal{R}(\mathbf{x}) \tag{2}$$

where $\mathcal{R}(\mathbf{x})$ is a structural prior to enhance the reconstruction. For tratidional sparsity-based CS methods, $\mathcal{R}(\mathbf{x})$ often involves minimizing the number of non-zero coefficients of $\mathbf{x}$ in some transformation domain, leveraging the inherent sparsity of the signal to make the underdetermined system of equations solvable.

2.2. Network architecture

Before establishing the DCGAN network, the raw vibration data, which typically comprises extensive time series, undergoes a series of preprocessing steps to ensure optimal input quality for the model. The preprocessing includes data normalization, slicing, padding, and masking. Normalization is used to ensure uniformity across the dataset, achieved by adjusting the data to have a consistent scale. Then, the raw data is sliced into shorter, more manageable sequences of 512 data points, which is easier to handle. Padding involves adding zeros to the sequences to reach the required length, ensuring uniformity across all input data. Finally, a Boolean mask is applied to the dataset, indicating which parts of the compressed signal $\mathbf{y}$ have been padded.

After the data preprocessing, the dataset would be fed to the DCGAN. The proposed architecture is composed of two deep convolutional networks [8], one is the generator *G*, and the other is the discriminator *D*. *G* learns the mapping from the compressed signal $\mathbf{y}$ to the real signal $\mathbf{x}$, and its primary objective is to generate a realistic reconstruction of $\mathbf{x}$ that mimics the true data as closely as possible. *D* undertakes a classification task to distinguish the recovered signal from original ones. The network architecture and scheme of training are illustrated in Figure 1.

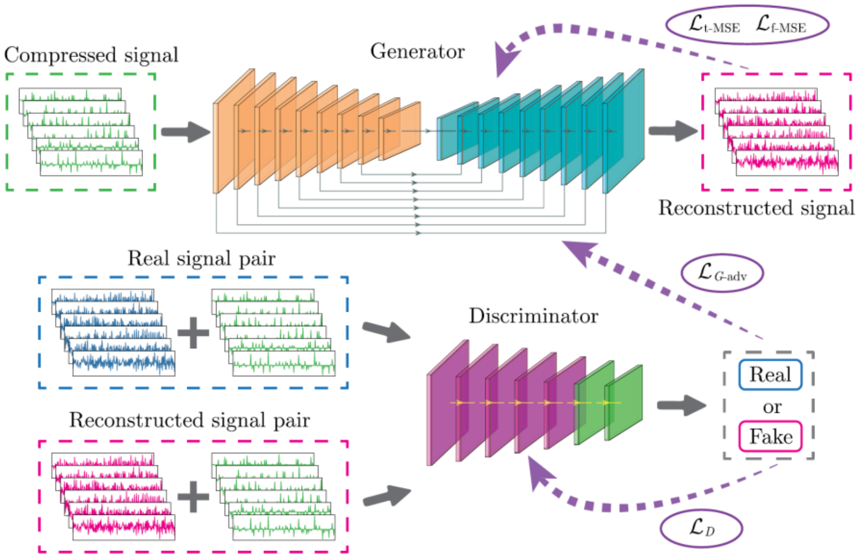

Figure 1: The network architecture and training scheme.

The generator G is a modified 1D U-net [9] with symmetric shortcuts, as shown in Figure 1. The first half of G is an encoder that comprises nine stacked convolutional layers, designed to incrementally extract more abstract features from the compressed input signal. Mirroring the encoder, the decoder consists of nine transposed convolutional layers. The transposed layers aim to reconstruct the target signal from the abstracted features, gradually recovering the detailed aspects of the original signal. Each convolutional and transposed convolutional layer is followed by a batch normalization layer and an activation function. Symmetric shortcut connections bridge each convolutional layer to its corresponding transposed layer in the decoder. These shortcuts serve dual purposes: First, they ensure that feature maps from each level of the encoder are directly fed into the corresponding decoder layers, preserving critical details that might otherwise be lost in deeper layers of the network. Second, they facilitate the back-propagation of gradients during training, making deep networks much easier to train.

The discriminator D is a typical convolutional classifier. It receives both the compressed signal $\mathbf{y}$ and the corresponding complete signal $\mathbf{x}$ as input, and attempts to classify whether they are the original or generated by G. D consists of 5 convolutional layers with a stride of 2, and each convolutional layer is followed by batch normalization and leaky ReLU layers. The output layer of D is a dense layer with the sigmoid activation function for binary (0 or 1) classification.

2.3. Loss function and network training

The discriminator D is a binary conditional classifier that differentiates between the reconstructed signal pair (labeled 0) and the real exemplar pair (labeled 1). The loss function for the discriminator utilizes the cross-entropy between the real and predicted labels, formulated as:

$$\max_{D} \mathcal{L}_D = \mathbb{E}\left[\log D(\mathbf{y},\mathbf{x})\right] + \mathbb{E}\left[\log(1 - D(\mathbf{y}, G(\mathbf{y})))\right] \tag{3}$$

On the other hand, the loss function of generator G consisits three components, expressed as:

$$\min_{G} \mathcal{L}_G = \lambda_1 \mathcal{L}_{G\text{-adv}} + \lambda_2 \mathcal{L}_{\text{t-MSE}} + \lambda_3 \mathcal{L}_{\text{f-MSE}} \tag{4}$$

$\mathcal{L}_{G\text{-adv}}$ is the adversarial loss the generator's success in deceiving the discriminator, defined as:

$$\min_{G} \mathcal{L}_{G\text{-adv}} = \mathbb{E}\left[\log(1 - D(\mathbf{y}, G(\mathbf{y})))\right] \tag{5}$$

To improve the recovery accuracy, additional mean squared error (MSE) terms $\mathcal{L}_{\text{t-MSE}}$ and $\mathcal{L}_{\text{f-MSE}}$ are included to account for the discrepancies in the time and frequency domains. The weights λ_1, λ_2, and

λ_3 balance the components, with recommanded values of $\lambda_1 = 1$, $\lambda_2 = 100$, and $\lambda_3 = 10$, respectively, derived from extensive testing.

The DCGAN model is optimized by alternately minimizing $\mathcal{L}_G$ and maximizing $\mathcal{L}_D$ using the Adam algorithm. The proposed network is implemented and trained using TensorFlow, and the hyperparameters are set as follows: the learning rate is 1×10^{-6}; the exponential decay rates for the 1st and 2nd moment estimates in Adam are 0.9 and 0.999, respectively; and the batch size is 128. In addition, the label smoothing technique is used to encourage the convergence of GANs training.

3. Demonstration using the experimental data from a steel grandstand

3.1. Description of the experimental model

The performance of the proposed DCGAN for spatial structure vibration data is demonstrated using a large-scale test structure. The vibration dataset is obtained from the study conducted by Abdeljaber et al [10]. The experiment was conducted on a grandstand simulator (Figure 2), which consists of 8 4.6 m-long main girders, 25 filler beams, and 4 columns. A modal shaker was used to excite the grandstand with white noise on the 8th joint for 256 seconds, and the acceleration data was acquired by accelerometers installed at the joints with a sampling frequency of 1024 Hz. The data acquired from the 22nd joint is chosen to demonstrate the data recovery performance. A total of about 80,000 slices of acceleration signals are used for network training, which are obtained through the sliding-window strategy.

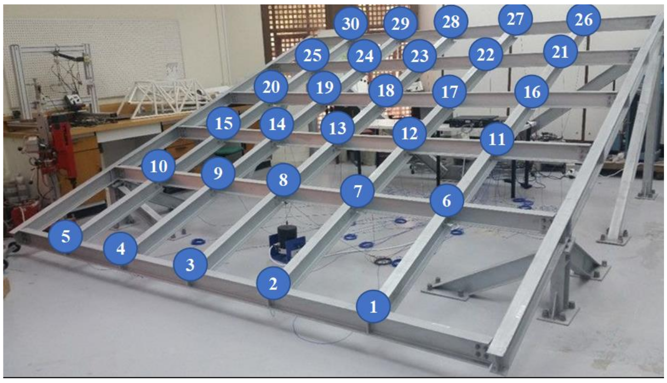

Figure 2: The tested steel grandstand.

3.2. Vibration signal recovery result

The proposed DCGAN is trained to recover the signal from a uniformly down-sampled sequence with a compression ratio of 4. The recovery results are shown in Figure 3 and 4. The recovered signal closely matches the original one, even only 25% samples are used. The major peaks and most minor peaks in the Fourier spectrum are accurately restored. However, it is observed that the spectrum of the reconstructed signal is smoother compared to the original, with minor peaks around 350 Hz showing some shifts or attenuation.

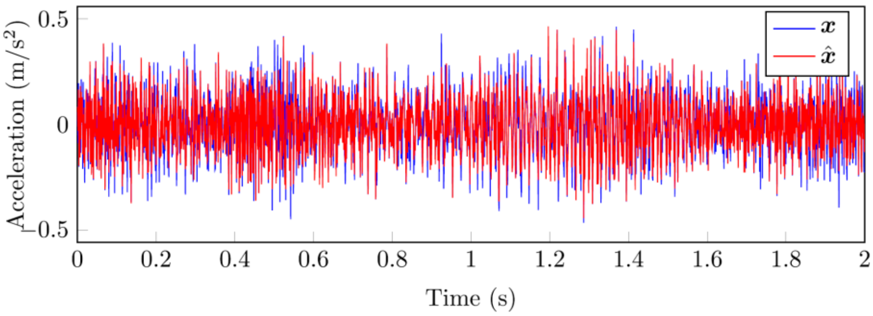

Figure 3: The recovered vibration signal in the time domain with a compression ratio of 4.

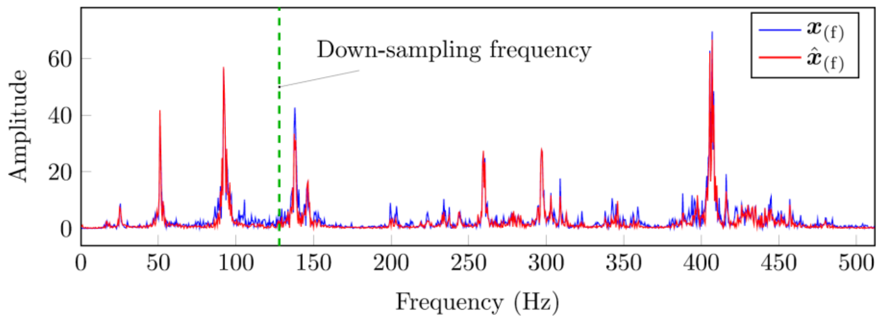

Figure 4: The recovered vibration signal in the frequency domain with a compression ratio of 4.

To further demonstrate the reliability of the proposed DCGAN, modal identification was conducted using both the real and recovered vibration signals. The acceleration data from all thirty nodes of the grandstand was compressed and recovered with a compression ratio of 4. Modal identification using the SSI-COV method revealed eleven distinct natural frequencies. The corresponding mode shapes are displayed in Figure 5, with Modal Assurance Criterion (MAC) values provided. The results highlight high accuracies as most MAC values exceeding 0.97, although the MAC values for high-frequency mode shapes were slightly lower than those for low-frequency mode shapes.

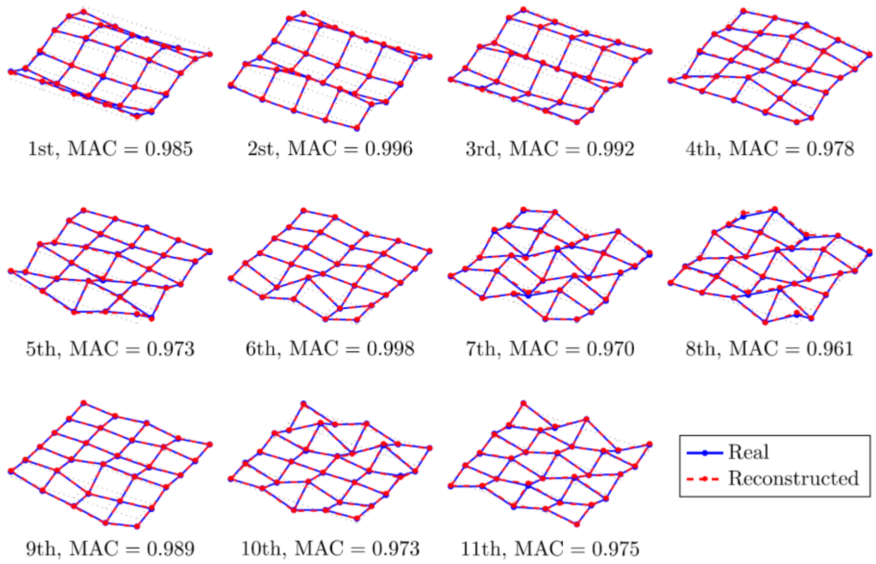

Figure 5: Mode shapes extracted from recovered vibration signals.

4. Conclusion

In this work, we propose a novel DCGAN designed for the challenges of using CS on vibration data from spatial structures. The DCGAN consists of a modified 1D convolutional U-net generator with shortcuts, and a conventional 1D convolutional classifier as the discriminator. A composite adversarial loss function is proposed to improve the recovery performance. The effectiveness of the proposed DCGAN is demonstrated through the experimental data collected from a test steel grandstand. The results show that despite the low sparsity of these signals, the proposed DCGAN can accurately recover the compressed vibration signal. Additionally, the results from modal identification further validate the reliability of the recovered signals. In summary, the DCGAN offers a promising solution for the challenges of vibration monitoring in spatial structures.

References

[1] CANDES E J, WAKIN M B, “An introduction to compressive sampling”, *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, 2008.

[2] CANDES E, ROMBERG J, "Sparsity and incoherence in compressive sampling", *Inverse Problems*, vol. 23, no. 3, pp. 969-985, 2007.

[3] RANI M, DHOK S B, DESHMUKH R B., "A systematic review of compressive sensing: Concepts, implementations and applications", *IEEE Access*, vol. 6, pp. 4875-4894, 2018.

[4] HUANG Y, BECK J L, WU S, et al., "Robust Bayesian compressive sensing for signals in structural health monitoring", *Computer-Aided Civil and Infrastructure Engineering*, vol. 29, no. 3, pp. 160-179, 2014.

[5] BAO Y Q, TANG Z Y, LI H, "Compressive-sensing data reconstruction for structural health monitoring: a machine-learning approach", *Structural Health Monitoring*, vol. 19, no. 1, pp. 293-304, 2020.

[6] DONG C, LOY C C, HE K, et al., "Learning a deep convolutional network for image super-resolution", *European Conference on Computer Vision*, Springer, 184-199, 2014.

[7] KIM J, LEE J K, LEE K M, "Accurate image super-resolution using very deep convolutional networks", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[8] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al., "Generative adversarial nets", *Advances in Neural Information Processing Systems*, vol. 27, 2014.

[9] RONNEBERGER O, FISCHER P, BROX T, "U-net: Convolutional networks for biomedical image Segmentation", *International Conference on Medical Image Computing and Computer-assisted Intervention*, pp. 234-241, 2015.

[10] ABDELJABER O, AVCI O, KIRANYAZ S, et al., "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks", *Journal of Sound and Vibration*, vol. 388, pp.154-170, 2017.