

Xing Zhao¹, Ping Lu², Yanyan Zhang², Jianxiong Chen², and Xiaoyang Li³

<https://doi.org/10.1190/tle38120934.1>

Abstract

Noise attenuation for ordinary images using machine learning technology has achieved great success in the computer vision field. However, directly applying these models to seismic data would not be effective since the evaluation criteria from the geophysical domain require a high-quality visualized image and the ability to maintain original seismic signals from the contaminated wavelets. This paper introduces an approach equipped with a specially designed deep learning model that can effectively attenuate swell noise with different intensities and characteristics from shot gathers with a relatively simple workflow applicable to marine seismic data sets. Three significant benefits are introduced from the proposed deep learning model. First, our deep learning model doesn't need to consume a pure swell-noise model. Instead, a contaminated swell-noise model derived from field data sets (which may contain other noises or primary signals) can be used for training. Second, inspired by the conventional algorithm for coherent noise attenuation, our neural network model is designed to learn and detect the swell noise rather than inferring the attenuated seismic data. Third, several comparisons (signal-to-noise ratio, mean squared error, and intensities of residual swell noises) indicate that the deep learning approach has the capability to remove swell noise without harming the primary signals. The proposed deep learning-based approach can be considered as an alternative approach that combines and takes advantage of both the conventional and data-driven method to better serve swell-noise attenuation. The comparable results also indicate that the deep learning method has strong potential to solve other coherent noise-attenuation tasks for seismic data.

Introduction

Swell noise, caused by long-period changes in the ocean surface, is a type of incoherent noise characterized by low frequencies and high-amplitude features. Swell-noise attenuation is usually the first step in processing marine seismic data. It is crucial to have clean results, as the high-amplitude features of swell noise can mask the signal of interest. The successful separation of true reflection signals and unwanted noise is a long-standing problem in seismic data processing. It greatly affects the fidelity of subsequent seismic imaging (Claerbout, 1985; Dai et al., 2012) and geophysical inversion, such as amplitude-variation-with-offset inversion (Buland and Omre, 2003; Li and Mallick, 2014), full-waveform inversion (Pratt, 1999; Chen et al., 2016, 2018), and geologic interpretation (Brown, 2011). Seismic data are inevitably affected by different types of noise. The existence of noise in prestack seismic data affects the amplitude information, which causes unreliable inversion results. For poststack seismic data, the existence of noise affects the ability of interpretation, which directly links the modeling of subsurface reservoirs.

Seismic noise attenuation has gone through a long history of development. Most conventional methods utilize signal features, such as wavenumber and frequency and domain transformation, to attenuate seismic noise. Noise attenuation for images using machine learning technology has also achieved success in the computer vision field. Deep learning in the denoising task for images developed in the past decades. Convolutional neural networks (CNNs) train the network through learning lower dimensional representations of the image features. By taking advantage of the CNN, many advanced noise-attenuation models have been proposed (He et al., 2016; Lehtinen et al., 2018; Zhang et al., 2018; Guo et al., 2019).

However, directly applying these methods to seismic data may not be practical since the geophysical domain requires the visual quality of the seismic image and the recovery quality of seismic signals. For example, for training purposes, the CNN would decrease the loss value (e.g., L1 loss or L2 loss) and make the predicted value converge to a certain level. As a result, some noise with significant amplitude changes, such as swell noise, will not be effectively detected and removed by these methods. This is due to the algorithm converging to a local optimum, and instead of keeping the phase information, such objective function aims to average the signal variations. Geophysical domains do not accept such results since the priority of noise attenuation is to keep the information from original signals.

In this paper, we would like to bring state-of-the-art techniques of noise attenuation from the computer vision field to the geosciences and make a variant version of the technique to fit the requirements in the geophysical domain, specifically for swell-noise attenuation. We will introduce related works about conventional solutions for seismic noise attenuation and state-of-the-art deep learning-based solutions for image noise attenuation. Then, we will introduce our model frame and apply our model to two cases. This will be followed by a results analysis. In summary, our main contributions are: (1) the proposal of a deep learning-based approach for swell-noise attenuation with consumption of a contaminated swell-noise model; (2) better denoising results in terms of noise detection and primary protection; and (3) the avoidance of additional steps.

Related works

Conventional approaches for seismic noise attenuation.

Seismic noise attenuation has gone through a long history of development. Most conventional methods utilize signal features, (e.g., wavenumber and frequency and domain transformation) to attenuate seismic noise. Specifically focusing on swell-noise attenuation, a straightforward approach is to apply FX filtering to the frequency range where the noise is present and replace noisy traces with their FX-filtered versions for these frequencies.

¹Texas A&M University, College Station, Texas, USA. E-mail: xingzhao@tamu.edu.

²Anadarko Petroleum Corp., The Woodlands, Texas, USA. E-mail: ping.lu@anadarko.com; yanyan.zhang@anadarko.com; jianxiong.chen@anadarko.com.

³University of Houston, Houston, Texas, USA. E-mail: xiaoyang.rebecca.li@gmail.com.

However, since noise amplitude can be high compared with the signal, residuals are often still unacceptable. Schonewille et al. (2008) introduce an improvement by applying FX filtering iteratively. Bekara and van der Baan (2010) develop an automatic method to exploit differences in statistical properties between swell noise and signal amplitudes to construct a detection criterion. Sternfels et al. (2015) model the coherent signal via its low-rank trajectory matrix and the erratic noise as a sparse component of the input data. Their method can effectively separate the signal and swell noise in the corrupted input. The method is the benchmark solution in this paper.

Deep learning for noise attenuation on ordinary images. Noise attenuation for images using machine learning technology has achieved great success in the computer vision field. Deep learning in the denoising task for images has been developed in the past decades (Rabie, 2005; Jain and Seung, 2009; Xie et al., 2012), and many research works indicate that Burger et al. (2012) made a giant leap in this field. CNNs train the network through learning lower dimensional representations of the image features. By utilizing CNNs, deep residual networks (ResNet) (He et al., 2016), and batch normalization (Ioffe and Szegedy, 2015), Zhang et al. (2017) propose the denoising CNN model, and it outperforms the traditional nonCNN-based methods. CNN-based techniques on noise attenuation have been widely and continually developed into many variants. Recently, the Noise2Noise (Lehtinen et al., 2018) model was introduced for the noise-attenuation task without providing ground-truth information. CBDNet (Guo et al., 2019) consists of two subnetworks (noise estimation and nonblind denoising). It achieves state-of-the-art results in terms of quantitative metrics and visual quality. Similarly, FFDNet (Zhang et al., 2018), RED30 (Mao et al., 2016), BM3D-Net (Yang and Sun, 2017), and CS-DIP (Van Veen et al., 2018) also achieved prominent performances.

Methodology

In this section, we propose our deep learning-based solution, Only2Noise, for the seismic swell-noise-attenuation problem. We introduce the learning mechanism and main component (ResNet) of our model. Then, we show the model structure and loss function.

Model design: Learning mechanism. The existence of swell noise is due to long-period changes in the ocean surface. Therefore, the phenomenon of swell noise is not universal, and it has extremely unstable amplitudes. Directly applying an existing deep learning-based denoising model to seismic swell-noise attenuation may obtain results without physical meaning. For example, the denoised seismic image using these methods can change the phase and spectrum of the signals and eventually hurt the primary signals due to high amplitudes from the swell noise. As a result, the swell noise may not be effectively detected and removed by using the traditional deep learning-based models while there exists significant residual swell noise or strong distortion on the primary signals.

Due to the coherent features of swell noise, the rationale is to design a model with residual networks to learn and predict the swell noise rather than the clean image. In other words, our residual network model is trained to extract the swell noise rather than to attenuate it. Specifically, for any training image

$s + n$, we design our model to map $F_s : (s + n) \rightarrow n$ instead of $F : (s + n) \rightarrow s$, where s is the clean signal and n is swell noise.

For random noise attenuation on ordinary images, learning the pattern of noise is difficult since random noise may not have strongly recognizable patterns. Some of the hybrid methods tend to derive denoised images by separating noise from raw data (e.g., Li et al., 2019). However, such methods not based on deep learning require additional steps to recover the primary signals due to excessive attenuations of primary signals. Unlike conventional methods, the proposed approach leverages the deep residual network, learns how to directly detect the swell noise, and then obtains attenuated data by removing swell noise from the corrupted raw data set. Therefore, one of the contributions of this paper is to directly capture the pattern of the swell noise using the deep learning-based residual network. Such learning processes can be beneficial to process different types and conditions of primary signals and to reduce the probability of overfitting issues. As far as we know, such deep learning residual network models, which only capture the swell-noise pattern on seismic data, have not been used in previous studies. Since only the patterns of swell noise are learned, we name our method Only2Noise.

Model design: ResNet components. In the machine learning society, CNNs have been widely used in image processing, such as image classification (Krizhevsky et al., 2012) and face recognition (Lawrence et al., 1997). ResNet (He et al., 2016), a deeper version of CNN, solved the problem that a deeper network may cause higher training/testing loss. ResNet splits the original mapping $x \Rightarrow H(x)$ into two parts:

$$x \Rightarrow f(x) \text{ (residual mapping)} \quad (1)$$

and

$$x + f(x) \Rightarrow H(x), \quad (2)$$

where x denotes the original identity, $f(x)$ denotes the residual mapping, and $H(x)$ denotes the final mapping.

In this way, the problem of vanishing or exploding gradient and degradation of traditional stacked deep CNN would be eliminated. Many variants of ResNet have been developed in recent years, such as the superresolution residual network (SRResNet) (Ledig et al., 2017) and EDSR+ (Lim et al., 2017). In this paper, we design our deep learning model based on the ResNet structure.

Only2Noise is designed to include residual blocks. Figure 1 shows the inside components of each residual unit. We use the shortcut connection to link the input and

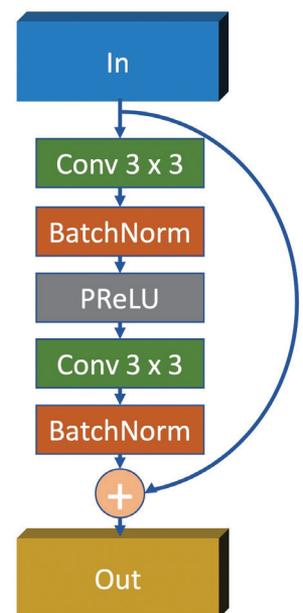


Figure 1. A residual unit with five internal layers.

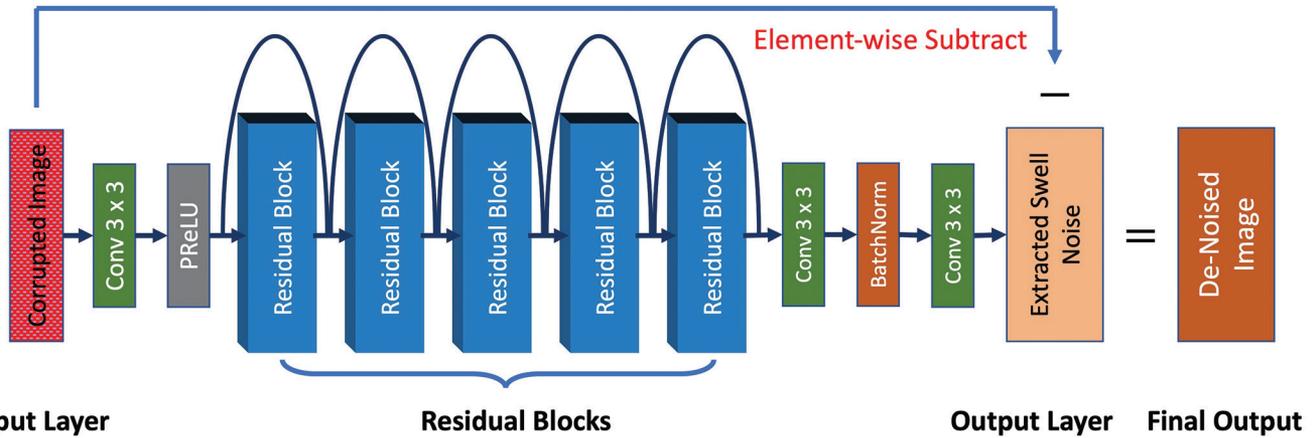


Figure 2. Residual neural network architecture. The model takes the corrupted image as input and returns extracted swell noise. The final output is the denoised image subtracting the extracted swell noise from corrupted input.

output of each residual block. In each residual unit, we use the same 3×3 convolutional layer followed by a batch normalization layer to expedite the convergence and to avoid overfitting. For the activation function, SRResNet, we employ Parametric ReLU (He et al., 2016) instead of ReLU used in the traditional ResNet.

Model design: Model structure. The model structure is shown in Figure 2. Our swell-noise-attenuation model, Only2Noise, generates the swell noise $F_s(S|\theta)$ when given the corrupted seismic data S as input and trained model parameter set θ . The final predicted seismic images S' are achieved by subtracting the predicted noise from the corrupted input data S defined as:

$$S' = S - S_f(S|\theta). \quad (3)$$

Similar to the traditional deep learning-based denoising model, we use the L2 norm to measure the difference. However, since our predicted value in this model is the swell noise (rather than the clean image), we change our loss function as follows. Given a corrupted signal S as input, the predicted swell noise $F_s(S|\theta)^{(M \times N)}$, and the target ground-truth image matrix $\hat{S}^{(M \times N)}$, we calculate the loss as follows:

$$loss = \sum_{i=1}^M \sum_{j=1}^N \left\| F_s(S|\theta)_{i,j} - (S_{i,j} - \hat{S}_{i,j}) \right\|^2. \quad (4)$$

For parameter tuning, we use the pair predicted swell noise $F_s(S|\theta)$ and the ground-truth signal \hat{S} to tune the parameters θ to minimize the pixel-wise loss:

$$\arg \min_{\theta} \sum_{i=1}^M \sum_{j=1}^N \left\| F_s(S|\theta)_{i,j} - (S_{i,j} - \hat{S}_{i,j}) \right\|^2. \quad (5)$$

Evaluation metrics. For evaluation purposes, we use the following measurements. The signal-to-noise-ratio (S/N) is defined as the ratio between the variance of the original gather and the

noise, where noise is the difference between the corrupted signal and the clean signal. Given corrupted seismic data S (or denoised seismic S') and its clean sample \hat{S} , S/N is defined as:

$$SNR = 10 \log_{10} \frac{(\hat{S})^2}{(S - \hat{S})^2}. \quad (6)$$

Mean squared error (MSE) is defined as the average of the element-wise squared difference between the predicted signal and true signal, calculated as:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N \left\| (S_{i,j} - \hat{S}_{i,j}) \right\|^2. \quad (7)$$

Swell-noise attenuation on synthetic data sets

As mentioned earlier, the swell noise, caused by long-period changes in the ocean surface, is neither universal nor zero-mean. It always has recognizable patterns and consistencies in different situations. Therefore, in this section, we test the performance of our Only2Noise model on synthetic seismic data to validate the performance of the deep learning model.

Data set preparation. First, a synthetic seismic data set representing geologic information S_{syn} is generated from real streamer data acquired from a Gulf of Mexico asset where there are 123 shot gathers associated with 556 traces for each gather. It is important to note that obtaining a pure swell-noise model is a difficult task. However, the proposed approach requires a learning target to proceed with the training process. In order to obtain a feasible solution, the idea of leveraging a swell-noise model derived from a conventional approach based on field data is introduced. A deswell workflow is applied to a small subset of the real data. Then, the swell-noise model n is generated from the subtraction of raw and deswell data. Eventually, the swell-noise model n is added back to the synthetic data set S_{syn} . The raw synthetic data with swell noise are defined as:

$$S_{raw} = S_{syn} + n. \quad (8)$$

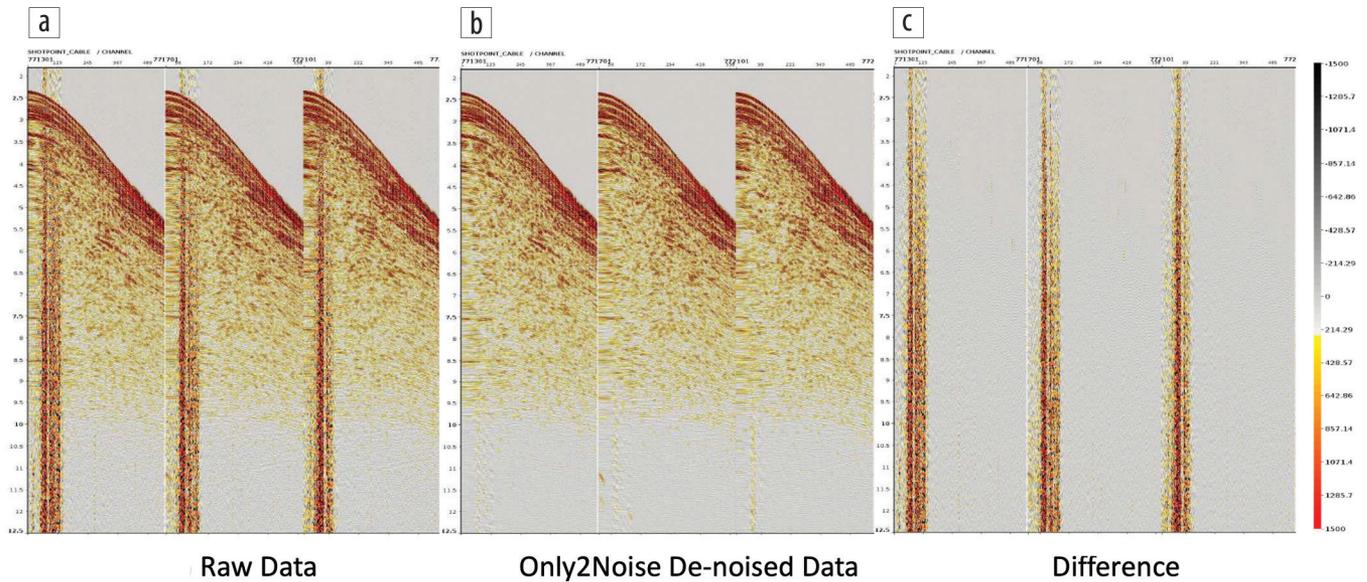


Figure 3. Synthetic data. Denoised results of selected gathers as examples. (a) Synthetic data with the swell. (b) Denoised results using the Only2Noise model. (c) Difference between (a) and (b), which is the detected swell.

Table 1. Optimal hyperparameters for Only2Noise.

Hyperparameters	Optimal value
Learning rate	0.05
Feature dimension	32
# of residual units	16
Optimizer	Adam
Steps per epoch	2000

Based on the data set mentioned earlier, we now have 123 shot gathers with 556 traces and 1800 timestamps. We randomly select 30% of the data as the training set for model training and parameter learning, 10% as the validation set for hyperparameter tuning, and the other 60% as the test set for final evaluation.

Experiment settings. We applied the loss function in equation 4. During the training process, MSE (equation 7) and S/N (equation 6) are used as evaluation metrics.

The training process is automatically terminated once the monitored performance converges. All parameter weights will be saved to a file. For hyperparameter tuning, we applied the validation data to adjust the hyperparameters (i.e., learning rate, feature dimension, number of residual units, and steps per epoch). We adopt the adaptive moment estimation (Kingma and Ba, 2014) as the optimizer for training, since it yields faster convergence compared to stochastic gradient descent. The optimal hyperparameters (Table 1) will be used in the final model training.

Results and analysis. In this section, we show the denoised results using the Only2Noise model on the testing set. Moreover, we present comparisons of the common channel and stacking results in the frequency panel between the conventional and Only2Noise approach. Rather than focusing only on the ability of noise attenuation, we also pay attention to the distortion of the primary signal. It is always a trade-off between attenuating

Table 2. Denoised results in respect to MSE and S/N.

	Raw	Denoised	Improved
MSE	975.848	9.740	99.00%
S/N	28.1 dB	48.1 dB	71.29%

noise and the ability to maintain primary signals in seismic data processing. In other words, an algorithm with stronger denoising ability is more likely to hurt the primary signals. For these conventional methods, such as frequency-domain transformation, it is apparent that remaining residual noise stays in high possibilities to prevent hurting the primary signals.

Figure 3 shows the denoised results on selected gathers. Figure 3a shows the synthetic data with the swell noise, Figure 3b shows the denoised results using the Only2Noise model, and Figure 3c shows the difference between Figures 3a and 3b, which is also the detected swell noise. Figure 3c shows synthetic noise-free data. On one hand, by visual comparison, most of the swell noise has been attenuated; however, there is still some mild residual swell noise in the denoised results. On the other hand, in Figure 3c, we observe that the primary signals have not been hurt. Focusing on the modeled swell noise, the swell noise has been successfully learned and detected from the input raw data.

MSE and S/N. Table 2 shows the denoised results in respect to MSE and S/N on the testing set. As we can see, when compared with the corrupted raw data, the denoised data decreased the MSE from 975.848 to 9.740 with 99.00% improvement, and denoised data improved the S/N from 28.1 to 48.1 dB with 71.29% improvement.

Frequency panel, common channel, and stacking results. Figures 4a and 4b show the comparison of testing raw data and denoised data on the frequency panel, respectively. By checking each range of frequency of the raw data, we can observe

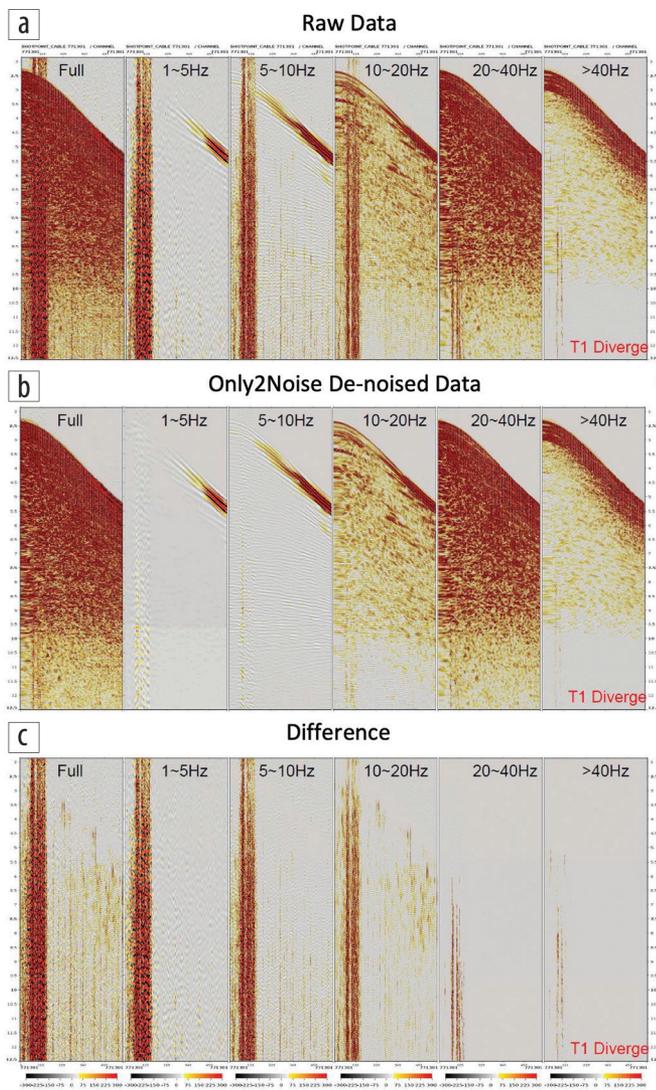


Figure 4. Synthetic data. Comparison of (a) corrupted raw data, (b) denoised data, and (c) detected swell (difference between [a] and [b]) on the frequency panel.

that the swell noise is more concentrated on the low-frequency range (i.e., 0 to 20 Hz), and a few exist on the high-frequency range (i.e., 20 Hz and more). From the perspective of geoscience, low-frequency data are essential to full-waveform inversion processes, and high-frequency data contribute to imaging resolution. Therefore, denoising for both the low- and high-frequency range are important. Figure 4c shows the detected swell noise on each frequency range.

Compared with the raw data, the Only2Noise model effectively removes most of the swell noise in each frequency model and performs better in the high-frequency part. In respect to keeping primary signals, as we can see in Figure 4c, most of the swell noise has been detected without strong primary hurts. There is only a mild hurt in the low-frequency range (0 to 10 Hz).

In respect to common channel (Figure 5) and stacking results (Figure 6), we can observe that the Only2Noise method has a strong ability of swell-noise attenuation for seismic data with mild hurts of primary signals. For detailed comparison, Figure 6c shows the stacking results of detected swell noise. As we can see,

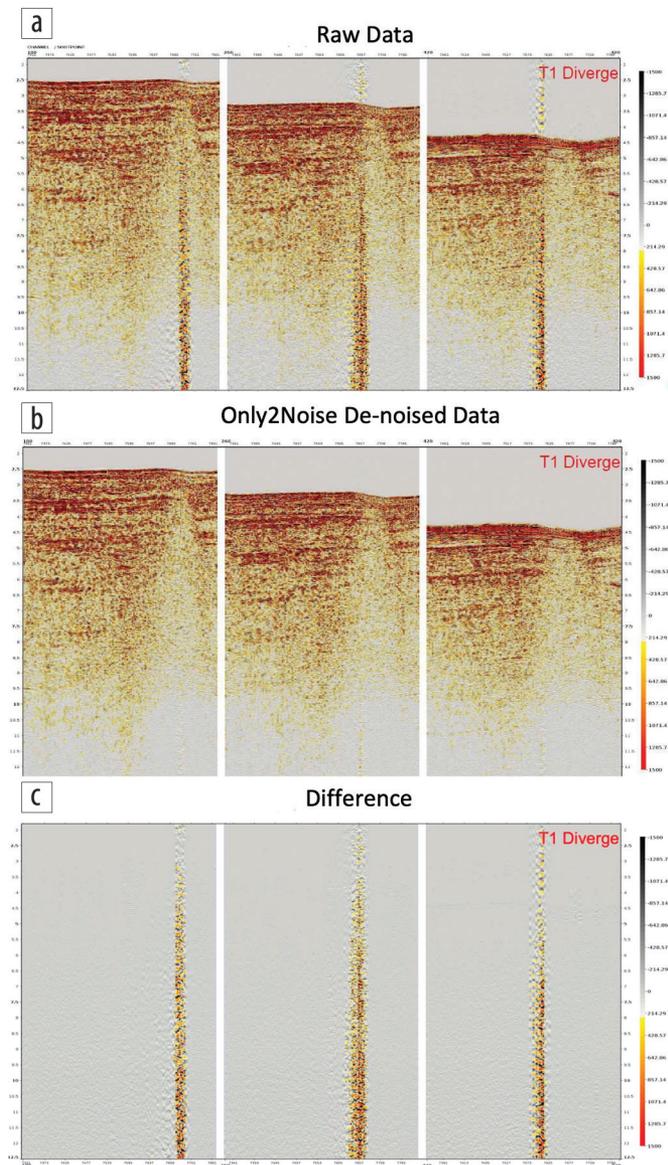


Figure 5. Synthetic data. Common channel comparison of (a) corrupted raw data, (b) Only2Noise denoised data, and (c) removed swell (difference between [a] and [b]).

primary hurts are slight and do not have a strong impact on the follow-up analysis.

Swell-noise attenuation on field data sets

In the previous section, we show the performance of our proposed Only2Noise model on the swell-noise-attenuation task on a synthetic data set. In respect to MSE and S/N, Only2Noise can effectively improve data quality. By detailed review of the frequency panel, common channel, and stacking results, our model can effectively detect the swell noise and return a remarkably better deswell result with mild primary hurts. In this section, we apply our learned model to the noise-attenuation task to field data sets and compare the results with conventional methods (Sternfels et al., 2015).

Seismic field data with swell noise in the Gulf of Mexico data set are used as raw input for noise attenuation. To fit the dimension

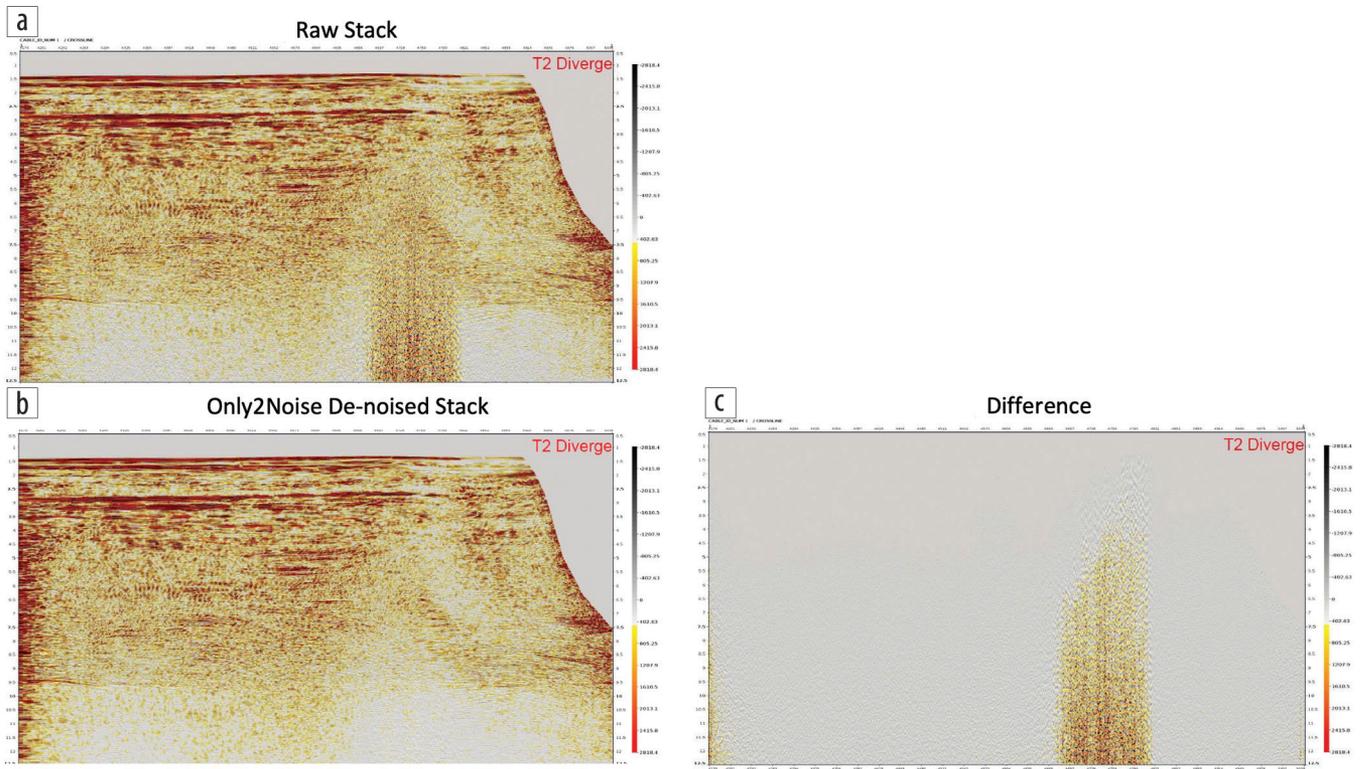


Figure 6. Synthetic data. Stack results comparison of (a) corrupted raw data, (b) Only2Noise denoised data, and (c) removed swell (difference between [a] and [b]).

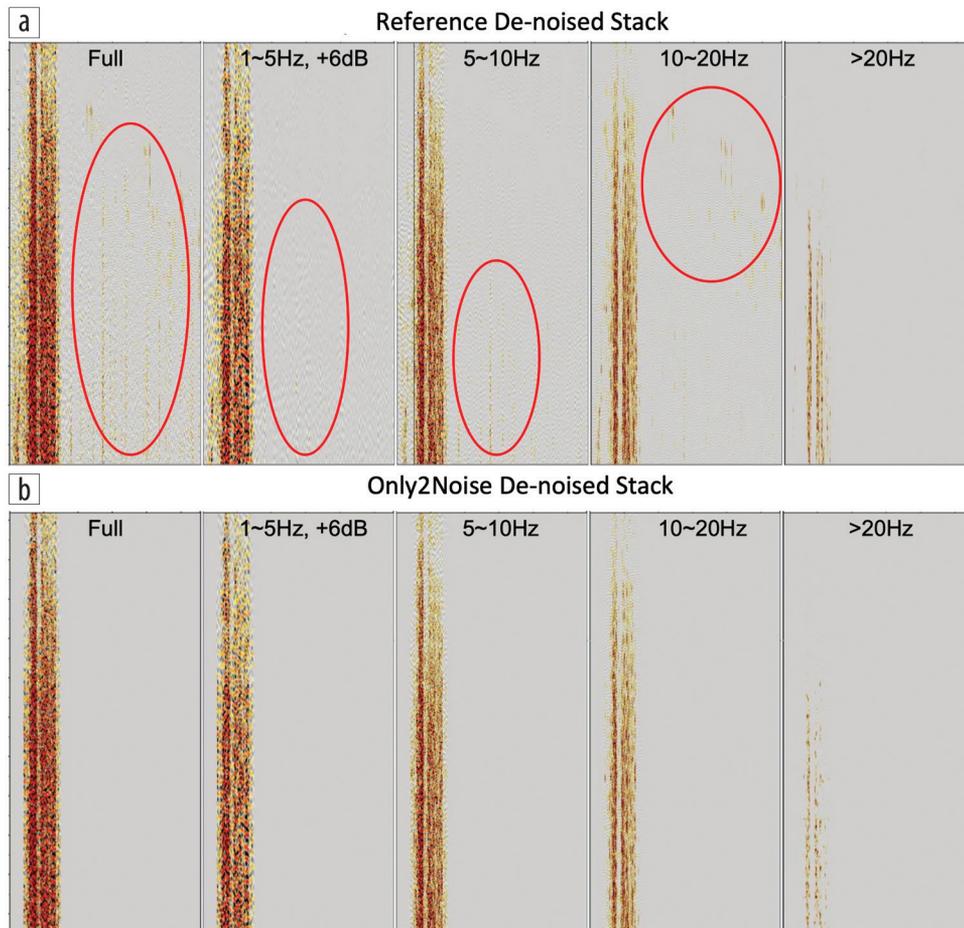


Figure 7. Field data. Comparison of (a) detected swell by the reference method and (b) detected swell by Only2Noise on the frequency panel.

of the synthetic data set, we chose 123 shot gathers with 556 traces and 1800 timestamps in our entire field data set. We directly applied our pretrained Only2Noise model to the field data set and show the deswell results in the following sections.

Results and analysis. Unlike experiments on the synthetic data set, in this experiment, we do not have the ground-truth data to compare with. Rather than checking the MSE and S/N, we visually compare our deswell results with the results of conventional methods (the reference method).

Figure 7 shows the frequency panel of the detected swell noise in a field data set by the reference method (Figure 7a) and Only2Noise method (Figure 7b). Compared with results from conventional methods, our results can detect similar swell-noise patterns albeit with fewer primary hurts in each frequency range. These improvements can be more prominent when we compare the common channels and stacking results.

Although a contaminated swell-noise model is provided for training the deep learning model, the proposed method still introduces apparent uplifts in terms of generating a stronger swell-noise model by comparing it with the original swell-noise model used for training. In respect to common channel (Figure 8) and stacking results (Figure 9), we observe that the Only2Noise method has a strong ability of swell-noise attenuation for seismic field data. Specifically, when compared with denoised common channel results by the reference method (Figures 8b and 8c), the Only2Noise model presents a trend of attenuating more swell noise in certain locations (red color). When compared with the stacking results (Figure 9) between denoised results by the reference method and by the Only2Noise method, we can see that our deep learning-based method has a strong ability to remove the swell noise with a sharper imaging quality and a better chance of interpretation of the geologic information.

Only2Noise on other marine data sets. The generalization of the deep learning model remains an active research topic in the academic community. Due to the complexity and variations of swell-noise characterizations, we do not expect to use one single model for swell-noise attenuation in all marine data sets. As long as the workflow introduced in this paper is followed, satisfactory results can be obtained for any marine data set. The corresponding workflow is summarized as:

- 1) Preparing a small synthetic primary model using geologic information from target marine data (similar to conventional methods)
- 2) Modeling a small portion of swell noise from target marine data
- 3) Training the deep learning model by leveraging the synthetic primary model and swell noise.

Conclusion

In this paper, we proposed a deep learning model with CNN-based residual neural networks for seismic swell-noise-attenuation tasks. Our model, Only2Noise, is applied to a synthetic data set and a field data set. It obtains strongly comparable results when compared with conventional approaches. With clear uplifts from two metrics of MSE and S/N, Only2Noise effectively and significantly enhances the quality of raw data.

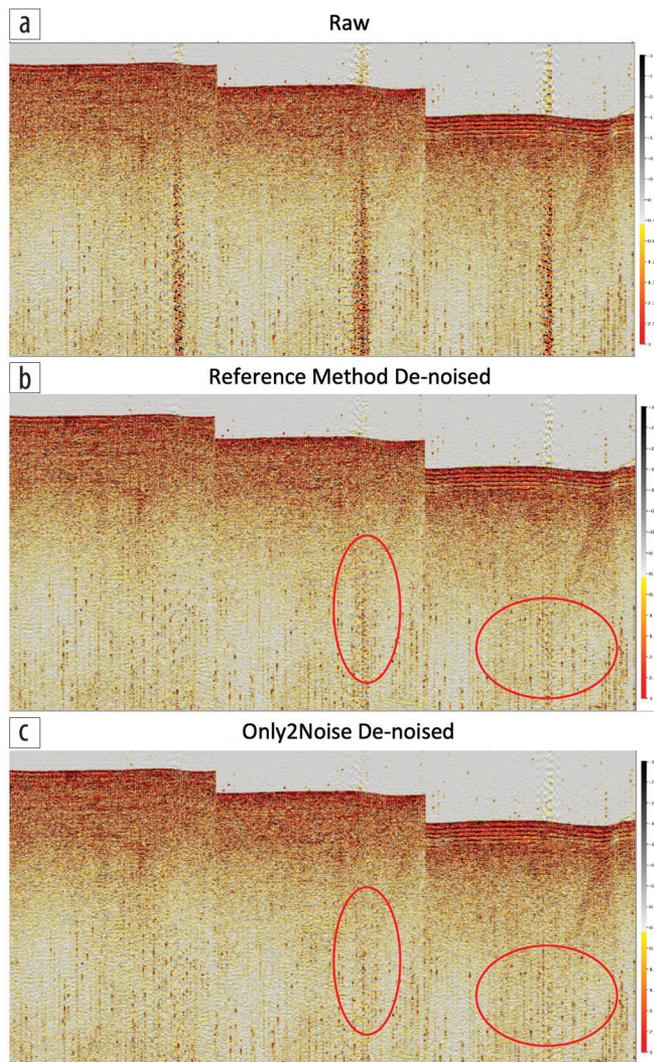


Figure 8. Field data. Common channel comparison of (a) corrupted raw data, (b) reference denoised data, and (c) Only2Noise denoised data.

By detailed and careful review of the frequency panels, common channels, and stacking results, our model indicates that the swell noise has been effectively detected and removed with mild primary hurts. The results show that the proposed technique can return a premium imaging quality with limited influence on the signal of interest.

The main contribution of this paper is to share a deep learning-based solution for swell-noise attenuation by directly training a swell-noise model. Such a swell-noise model can be derived from any marine data by using the conventional algorithm. Effectively and efficiently training on even a contaminated swell-noise model, guarantees an improved performance of attenuating swell noise from raw data. Such a method, in combination with conventional and data-driven algorithms, introduces a new means with strong potential to solve attenuation tasks relevant to swell noise as well as other coherent noise. **II**

Acknowledgments

The authors thank Anadarko Petroleum Corp. for permission to publish this work.

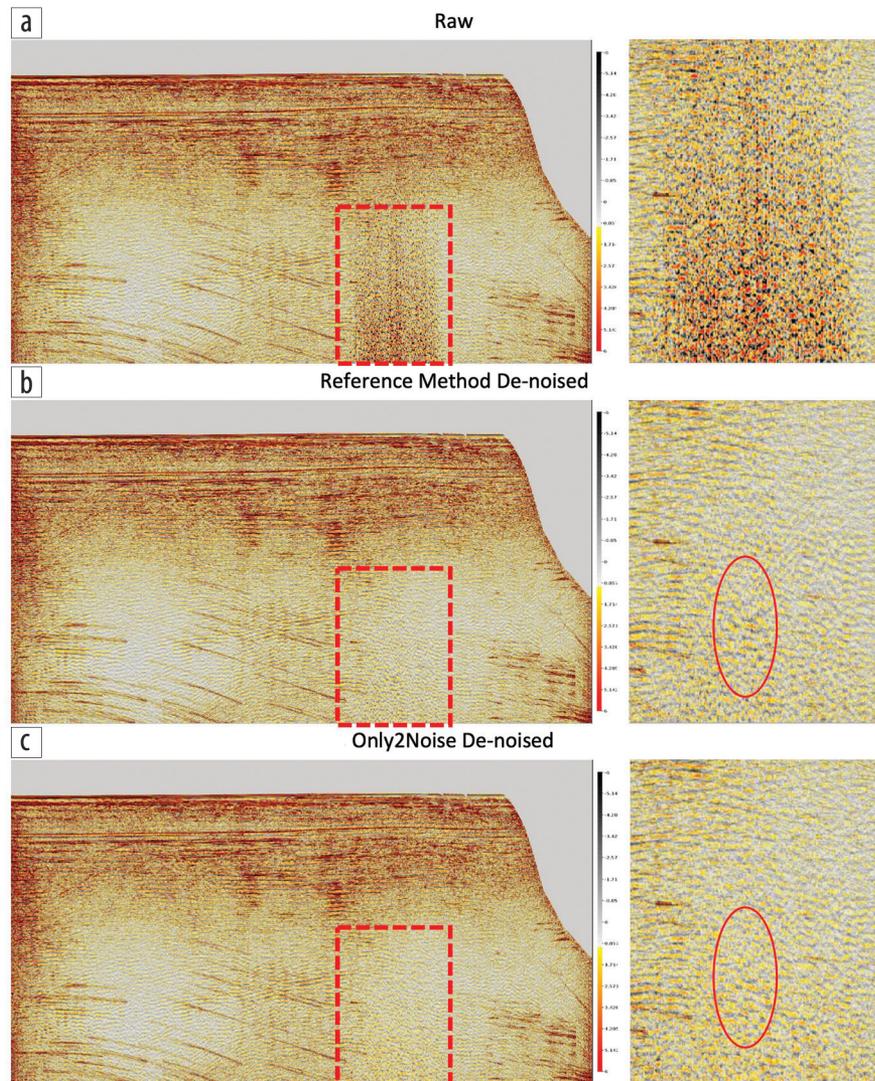


Figure 9. Field data. Stack results comparison of (a) corrupted raw data, (b) reference denoised data, and (c) Only2Noise denoised data.

Data and materials availability

Data associated with this research are available and can be obtained by contacting the corresponding author.

Corresponding author: xingzhao@tamu.edu

References

- Bekara, M., and M. van der Baan, 2010, High-amplitude noise detection by the expectation-maximization algorithm with application to swell-noise attenuation: *Geophysics*, **75**, no. 3, V39–V49, <https://doi.org/10.1190/1.3428749>.
- Brown, A. R., 2011, Interpretation of three-dimensional seismic data: AAPG and SEG.
- Buland, A., and H. Omre, 2003, Bayesian linearized AVO inversion: *Geophysics*, **68**, no. 1, 185–198, <https://doi.org/10.1190/1.1543206>.
- Burger, H. C., C. J. Schuler, and S. Harmeling, 2012, Image denoising: Can plain neural networks compete with BM3D?: *Conference on Computer Vision and Pattern Recognition*, IEEE, Extended Abstracts, 2392–2399, <https://doi.org/10.1109/CVPR.2012.6247952>.
- Chen, J., C. A. Zelt, and P. Jaiswal, 2016, Detecting a known near-surface target through application of frequency-dependent traveltimes tomography and full-waveform inversion to P- and SH-wave seismic refraction data: *Geophysics*, **82**, no. 1, R1–R17, <https://doi.org/10.1190/geo2016-0085.1>.
- Chen, J., D. Sixta, G. Raney, V. Mount, E. Riddle, A. Nicholson, H. Ma, H. Ji, and C. Peng, 2018, Improved subsalt imaging from reflection full-waveform inversion-guided salt scenario interpretation: A case history from deepwater Gulf of Mexico: 88th Annual International Meeting, SEG, Expanded Abstracts, 3773–3777, <https://doi.org/10.1190/segam2018-2996573.1>.
- Claerbout, J. F., 1985, *Fundamentals of geophysical data processing — With applications to petroleum prospecting*: Blackwell Scientific Publications.
- Dai, W., P. Fowler, and G. T. Schuster, 2012, Multi-source least-squares reverse time migration: *Geophysical Prospecting*, **60**, no. 4, 681–695, <https://doi.org/10.1111/j.1365-2478.2012.01092.x>.
- Guo, S., Z. Yan, K. Zhang, W. Zuo, and L. Zhang, 2019, Toward convolutional blind denoising of real photographs: *Conference on Computer Vision and Pattern Recognition*, IEEE, Extended Abstracts, 1712–1722.
- He, K., X. Zhang, S. Ren, and J. Sun, 2016, Deep residual learning for image recognition: *Conference on Computer Vision and Pattern Recognition*, IEEE, Extended Abstracts, 770–778, <https://doi.org/10.1109/CVPR.2016.90>.

- Ioffe, S., and C. Szegedy, 2015, Batch normalization: Accelerating deep network training by reducing internal covariate shift: arXiv:1502.03167.
- Jain, V., and S. Seung, 2009, Natural image denoising with convolutional networks: Presented at International Conference on Neural Information Processing Systems.
- Kingma, D. P., and J. Ba, 2014, Adam: A method for stochastic optimization: arXiv:1412.6980.
- Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012, ImageNet classification with deep convolutional neural networks: Presented at International Conference on Neural Information Processing Systems.
- Lawrence, S., C. L. Giles, A. C. Tsoi, and A. D. Back, 1997, Face recognition: A convolutional neural-network approach: *IEEE Transactions on Neural Networks*, **8**, no. 1, 98–113, <https://doi.org/10.1109/72.554195>.
- Ledig, C., L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, et al., 2017, Photo-realistic single image super-resolution using a generative adversarial network: Conference on Computer Vision and Pattern Recognition, IEEE, Extended Abstracts, 4681–4690, <https://doi.org/10.1109/CVPR.2017.19>.
- Lehtinen, J., J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, 2018, Noise2Noise: Learning image restoration without clean data: arXiv:1803.04189.
- Li, C., Y. Zhang, and C. Mosher, 2019, A hybrid learning-based framework for seismic denoising: *The Leading Edge*, **38**, no. 7, 542–549, <https://doi.org/10.1190/tle38070542.1>.
- Li, T., and S. Mallick, 2014, Multicomponent, multi-azimuth pre-stack seismic waveform inversion for azimuthally anisotropic media using a parallel and computationally efficient non-dominated sorting genetic algorithm: *Geophysical Journal International*, **200**, no. 2, 1136–1154, <https://doi.org/10.1093/gji/ggu445>.
- Lim, B., S. Son, H. Kim, S. Nah, and K. Mu Lee, 2017, Enhanced deep residual networks for single image super-resolution: Conference on Computer Vision and Pattern Recognition, IEEE, Extended Abstracts, 136–144, <https://doi.org/10.1109/CVPRW.2017.151>.
- Mao, X., C. Shen, and Y.-B. Yang, 2016, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections: Presented at International Conference on Neural Information Processing Systems.
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: *Geophysics*, **64**, no. 3, 888–901, <https://doi.org/10.1190/1.1444597>.
- Rabie, T., 2005, Robust estimation approach for blind denoising: *IEEE Transactions on Image Processing*, **14**, no. 11, 1755–1765, <https://doi.org/10.1109/TIP.2005.857276>.
- Schonewille, M., A. Vigner, and A. Ryder, 2008, Swell-noise attenuation using an iterative FX prediction filtering approach: 78th Annual International Meeting, SEG, Expanded Abstracts, 2647–2651, <https://doi.org/10.1190/1.3063892>.
- Sternfels, R., G. Viguier, R. Gondoin, and D. Le Meur, 2015, Joint low-rank and sparse inversion for multidimensional simultaneous random/erratic noise attenuation and interpolation: 77th Conference and Exhibition, EAGE, Extended Abstracts, <https://doi.org/10.3997/2214-4609.201412979>.
- Van Veen, D., A. Jalal, M. Soltanolkotabi, E. Price, S. Vishwanath, and A. G. Dimakis, 2018, Compressed sensing with deep image prior and learned regularization: arXiv:1806.06438.
- Xie, J., L. Xu, and E. Chen, 2012, Image denoising and inpainting with deep neural networks: Presented at International Conference on Neural Information Processing Systems.
- Yang, D., and J. Sun, 2017, BM3D-Net: A convolutional neural network for transform-domain collaborative filtering: *IEEE Signal Processing Letters*, **25**, no. 1, 55–59, <https://doi.org/10.1109/LSP.2017.2768660>.
- Zhang, K., W. Zuo, Y. Chen, D. Meng, and L. Zhang, 2017, Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising: *IEEE Transactions on Image Processing*, **26**, no. 7, 3142–3155, <https://doi.org/10.1109/TIP.2017.2662206>.
- Zhang, K., W. Zuo, and L. Zhang, 2018, FFDNet: Toward a fast and flexible solution for CNN-based image denoising: *IEEE Transactions on Image Processing*, **27**, no. 9, 4608–4622, <https://doi.org/10.1109/TIP.2018.2839891>.