Zero-Shot Denoising of Distributed Acoustic Sensing Data using Deep Priors

Jannik Mänzer

Bachelor of Science

March 2025

Supervisors:
Prof. Dr. Stefan Harmeling
Sebastian Konietzny

Artificial Intelligence (VIII)
Department of Computer Science
TU Dortmund University

Abstract

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Contents

Chapter 1

Introduction

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Chapter 2

Background

This chapter provides the necessary background for the denoising methods explored in this work. We begin by defining the general denoising problem and discussing its inherent challenges. We then introduce distributed acoustic sensing (DAS) as a real-world application and highlight the unique difficulties it poses. Finally, we present key deep learning concepts and techniques that are used in the context of denoising.

2.1 Denoising

In general, denoising refers to the process of recovering a clean signal from a noisy observation. Formally, it can be described by the inverse problem

$$y = x + n \tag{2.1}$$

where y is the noisy observation, x is the underlying clean signal and n represents some form of noise, for example Gaussian noise $n \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$. Since both the noise and its distribution are often unknown, denoising is an inherently ill-posed problem. Therefore, additional assumptions about the solution are essential for solving the denoising problem. This process is known as regularization and typically involves imposing certain constraints on the solution space to favor more natural solutions [22]. The choice of regularizer depends on the specific problem setting and the type of data involved.

2.2 Distributed Acoustic Sensing

Distributed acoustic sensing (DAS), also known as distributed vibration sensing (DVS) [6], is an innovative technology for high-resolution vibration measurements over large distances by utilizing fiber optic cables as sensor arrays. When a short laser pulse is sent through the fiber by a DAS interrogator unit, a fraction of the light is scattered back due to small variations or imperfections in the fiber. This phenomenon is referred to as Rayleigh scattering. Vibrations along the cable caused by external influences, e.g. seismic events, strain the fiber, which in turn causes phase shifts in the backscattered light.

These shifts are detected by the interrogator and, since the travel time of the light is known, can be used to accurately locate the strain along the cable [12]. In order to extract meaningful measurements, strain is analyzed over sections of the fiber, rather than at individual points. The length of these sections is called the gauge length, while another parameter, the channel spacing, determines how much this section is moved for each measurement, or channel, along the cable [4]. In practice, each channel corresponds to a virtual sensor capturing the average strain within its gauge length. Typically, the gauge length is selected to be bigger than the channel spacing, meaning that the measurement sections of neighboring channels overlap, as visualized in Figure 2.1. This concept of virtual sensors leads to high cost-effectiveness and, paired with the high sample rates enabled by the optical approach, allows measurements with significantly higher spatial and temporal resolution compared to conventional seismographs.

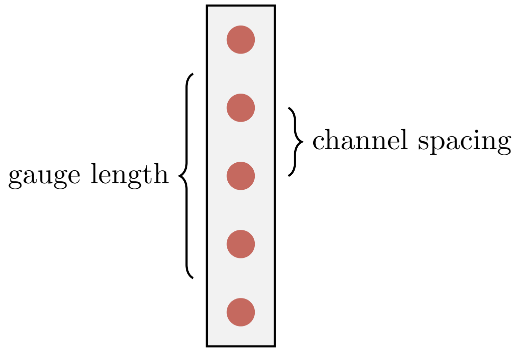

Figure 2.1: Gauge length and channel spacing. Red dots represent the individual channels along the fiber. Figure adapted from [4].

Despite these advantages, DAS systems often suffer from much lower signal quality than conventional seismographs, as they are more sensitive to various sources of noise. These can be divided into environmental noise and optical noise. Environmental noise includes natural phenomena such as winds or ocean waves, but also vibrations caused by vehicular and pedestrian traffic. Optical noise originates from various interactions between the light and the fiber. It includes high-amplitude erratic noise and common mode noise [3].

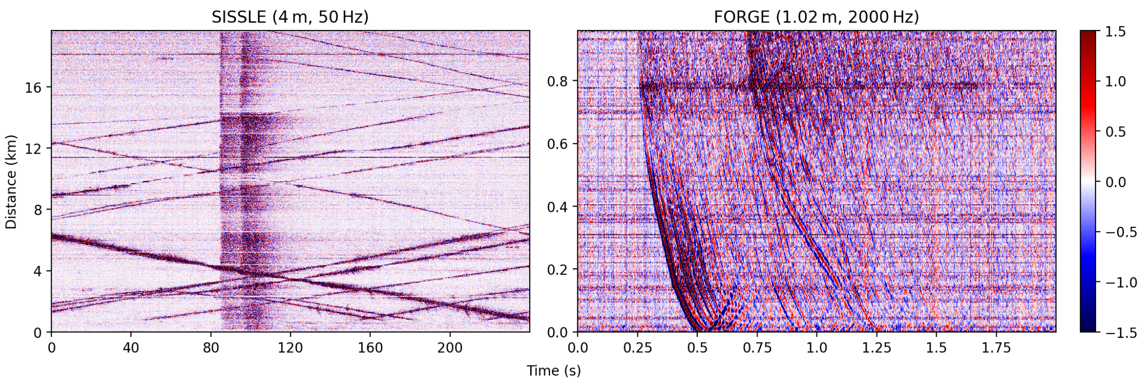

Figure 2.2: Types of noise in different DAS setups. Both measurements capture seismic activity; however, the SISSLE data mainly suffers from traffic noise (the diagonal lines). In contrast, erratic and common mode noise (the horizontal and vertical lines, respectively) are most prominent in the FORGE data.

The actual noise characteristics of DAS data not only depend on the environment, but also the measurement parameters such as channel spacing and sample rate, as visualized in Figure 2.2 for data from the SISSLE experiment near Haast, New Zealand [17] and the FORGE site in Utah [15].

2.3 Deep Learning

Deep learning is a subfield of machine learning that utilizes deep neural networks to learn complex patterns from data. Over the past decade, deep learning has established itself as the state-of-the-art approach for a wide range of problems across various different fields.

2.3.1 Deep Neural Networks

In its most basic form, a neural network consists of neurons organized in layers, where each neuron applies a linear transformation followed by a non-linear activation function. The output of a single neuron is given by

$$y = \varphi(\mathbf{w}^T \mathbf{x} + b) \tag{2.2}$$

for an input $\mathbf{x} \in \mathbb{R}^n$, a weight vector $\mathbf{w} \in \mathbb{R}^n$, a bias $b \in \mathbb{R}$ and an activation function $\varphi : \mathbb{R} \to \mathbb{R}$. The outputs of each layer are then passed as inputs to the next layer, which is why this architecture is known as a fully-connected neural network. The activation functions are needed in order to avoid the network collapsing into a single linear transformation. Therefore, a neural network can be described as a function $f_\theta : \mathcal{X} \to \mathcal{Y}$ parameterized by θ, where θ represents the weights and biases across all layers [5].

In order to optimize these parameters, a loss function $\mathcal{L} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ is defined, which measures the difference between the predicted output and the target value. Now, the gradient of the loss function with respect to the parameters, $\nabla_\theta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \theta}$, represents the direction of steepest ascent. Therefore, by moving the parameters in the opposite direction of the gradient, the loss function can be minimized. Typically, the gradient is not calculated for a single data point or for the whole dataset, but instead for a small subset of the dataset, which is why this approach is referred to as (mini-batch) gradient descent. Backpropagation [20] is used to efficiently compute the gradient by making use of the chain rule, enabling fast optimization.

While traditionally neural networks only consisted of a few layers and required handcrafted features to work effectively, advances in computing power allow modern architectures to automate feature extraction by using additional layers, hence the term *deep* neural network.

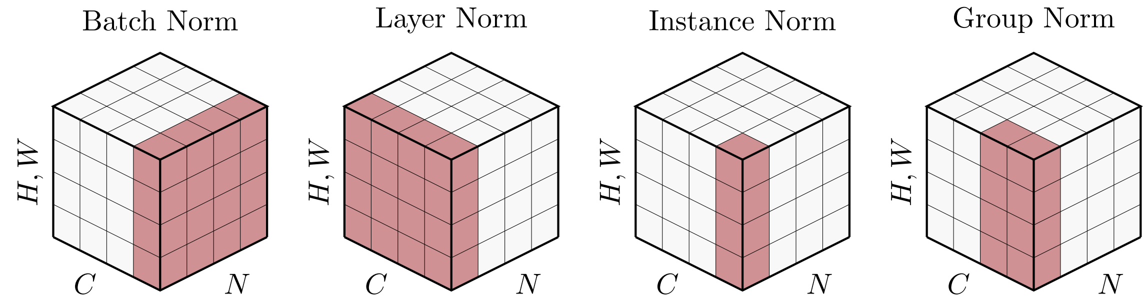

Figure 2.3: Different normalization techniques. N is the batch dimension, C is the channel dimension and H and W are the spatial dimensions of a 4D tensor. The input is normalized across the dimensions highlighted in red. Figure adapted from [26].

2.3.2 Convolutional Neural Networks

Convolutional neural networks (CNNs) [13] are a specific type of neural network that learns features using kernels. Prior to the rise of deep learning, such kernels were designed manually for various computer vision tasks, for example the Sobel kernel [21] used for edge detection. In CNNs, these kernels are automatically learned from data. In contrast to fully-connected layers, the output of a convolutional layer is obtained by convolution with one or multiple kernels, replacing the matrix multiplication. For a kernel $\mathbf{K} \in \mathbb{R}^{n \times m}$, the convolution is defined as

$$(\mathbf{X} * \mathbf{K})_{i,j} = \sum_{k=1}^{n} \sum_{l=1}^{m} X_{i+n,j+m} \cdot K_{k,l}. \tag{2.3}$$

The output of the convolution is then passed through a non-linear activation function, just like in fully-connected layers. CNNs provide two main advantages: First, since the weights are shared across the spatial dimensions, convolutional layers drastically reduce the number of parameters compared to fully-connected layers. Second, convolutions are translationally equivariant, meaning that local patterns in the input can be recognized regardless of their position, which makes CNNs very suitable for image data [5].

2.3.3 Normalization

During the training process, the inputs of each layer change with each iteration as the parameters are optimized. That slows down the training because now each layer has to adapt to the new distribution of its inputs. This process is often referred to as internal covariate shift. To counteract this issue, Ioffe et al. proposed Batch Normalization (BN) [9]. The idea behind BN is to normalize the inputs across the whole mini-batch and their spatial dimensions. The normalized input for a channel i is given by

$$\hat{x}^{(i)} = \frac{x^{(i)} - \mu^{(i)}}{\sigma^{(i)}}, \tag{2.4}$$

where $\mu^{(i)}$ and $\sigma^{(i)}$ are the per-channel mean and standard deviation of the mini-batch, respectively. In order to allow the model to learn the identity if that were the optimal transformation, two additional learnable parameters, γ and β, are introduced. The output of the BN layer is then defined as

$$y^{(i)} = \gamma^{(i)}\hat{x}^{(i)} + \beta^{(i)}. \tag{2.5}$$

Since there are no batch statistics available at inference time, BN keeps track of the running mean and variance during training and uses these values for normalization. While BN is widely used, there exist various similar normalization techniques [1, 24, 26] mainly differing in the dimensions across which they are applied. A selection of them is visualized in Figure 2.3.

2.3.4 Attention Mechanisms

In neural networks, some input features are typically more important than others. An attention mechanism helps the network to focus on (*attend to*) the most relevant parts of the input, rather than processing all inputs equally. This works by dynamically reweighting the features based on their importance [19]. While attention is often associated with natural language processing (NLP), especially since the introduction of the Transformer architecture [25] where it is the underlying key principle, it also has applications beyond NLP. In computer vision, for example, it can help CNNs by reweighting feature channels [7] or highlighting important spatial regions.

2.4 Fourier Transform

Chapter 3

Related Work

In the last few years, deep-learning-based methods have been successfully applied to various image denoising tasks [22] and achieve state-of-the-art results. While these networks were traditionally trained in a supervised fashion, requiring clean target images, recent methods remove the dependency on clean data by leveraging self-supervision [28]. In this chapter, we give an overview of existing supervised and self-supervised approaches, presenting key principles and discussing their limitations.

3.1 Supervised Methods

Traditional supervised methods typically employ a neural network f_θ to learn a mapping from a noisy image y to its clean counterpart x. Therefore, a dataset of paired clean and noisy images, denoted $\{(y^i, x^i)\}_i^n$, is essential for the training process. The corresponding optimization problem is given by

$$\underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{n} ||f_\theta(y^i) - x^i||_2^2. \tag{3.1}$$

Zhang et al. proposed the denoising convolutional neural network (DnCNN) [29] which improves denoising performance by making use of residual learning, i.e. instead of directly predicting the clean image, it is trained to predict the noise in the noisy image. The denoised image is then obtained as $\hat{x} = y - f_{\theta^*}(y)$ for trained parameters θ^*. However, depending on the problem setting, acquiring the needed clean data can be difficult or even impossible, for example in medical imaging or DAS.

To address this issue, Lehtinen et al. proposed Noise2Noise (N2N) [14], which does not require any clean data. Instead, it utilizes two independent noisy observations $y_1 = x + n_1$ and $y_2 = x + n_2$ of the same underlying clean signal x as input and target, respectively. This method relies on the assumption that the noise is zero-mean, i.e. $\mathbb{E}[n] = 0$, which, due to linearity of expectation, implies that $\mathbb{E}[y] = x$. Thus, by training a neural network as in (3.1), replacing the clean target with the second noisy observation, the network learns to predict x implicitly, as the MSE is a mean-seeking loss function. Given infinite data, the optimal solution is actually equivalent to the one obtained by training with clean targets. Although N2N is often impractical in practice because the required noisy-noisy pairs are difficult to obtain, it led to the development of other *self*-supervised approaches.

3.2 Self-Supervised Methods

Self-supervised methods are trained similarly to traditional supervised methods, but they do not rely on externally-provided target values. In the context of denoising, these approaches can be broadly categorized into two main strategies: Target-based methods generate their own supervisory signals using only the noisy inputs. Blind-spot-based methods, on the other hand, exploit spatial correlations in the image using different masking strategies, either in the input or in the network architecture itself.

3.2.1 Target-Based Methods

Noisier2Noise [18] builds upon N2N, but unlike N2N, it does not require a set of paired noisy-noisy images. Instead, it constructs these training pairs from individual noisy images only. Given a noisy input y, it generates an even noisier image $z = y + m = x + n + m$, with additional independent noise m following the same distribution as n. Once again, it is optimized through (3.1), using z as the input and y as the target. The authors argue that $\mathbb{E}[y|z] = x + \frac{n+m}{2}$, since $\mathbb{E}[n] = \mathbb{E}[m]$. Therefore, by the same reasoning as in N2N, given a sufficient amount of noisy images, the network should learn to predict the mean of x and z, which can then be used to obtain the denoised estimate as $\hat{x} = 2f_{\theta^*}(z) - z$. While this method removes the need for a paired dataset, it requires knowledge of the noise distribution in order to sample the additional noise.

Another approach based on N2N is Neighbor2Neighbor [8]. The key idea behind this method is to construct training pairs from the noisy input y by leveraging spatial redundancy through a sub-sampling strategy: y is divided into 2×2 cells from each of which two neighboring pixels are randomly selected — one pixel is assigned to the first sub-sampled image and the other to the second. These sub-sampled images then build the noisy training pair. As a result of the sub-sampling, unlike in N2N, the underlying clean signal x is not exactly identical in the two noisy images. To address this, the authors extend the training strategy given by (3.1) by using an additional regularization term that encourages minimizing differences between sub-sampled versions of the denoised estimate.

Zero-Shot Noise2Noise [16] takes this idea one step further by enabling training on just one single noisy image instead of a set of noisy images. The term *zero-shot* refers to a training setup where the model is supposed to make predictions for types of data it has never observed before without any training examples. This approach employs a similar sub-sampling strategy to obtain input and target values. In order to avoid overfitting to the noisy target, it makes use of residual learning, a symmetric loss and an additional regularization term enforcing consistency with respect to the order in which downsampling and inference are performed.

3.2.2 Blind-Spot-Based Methods

All blind-spot-based methods assume that noise is zero-mean and spatially independent, while the clean image signal exhibits spatial correlations. The underlying key principle for all of them is that a network should predict the value of a given pixel in the denoised image without directly observing its noisy counterpart, hence the term *blind-spot*. Therefore, the network can only learn from the neighboring pixels, which — under the assumption of independent noise — do not carry any information about the noise affecting the target pixel, thus preventing the network from predicting a noisy image.

Krull et al. first introduced this concept in their Noise2Void (N2V) paper [10]. The authors consider training a network to predict the center pixel of a single patch of the input image in a supervised fashion, using the actual pixel value as the target. To prevent the network from simply learning the identity, they propose restricting the output pixel's receptive field by masking the center pixel. The receptive field refers to the set of pixels in the input that influences a particular pixel in the output, as visualized in Figure 3.1. However, this process is not feasible in practice, as a whole patch has to be processed to obtain a single output pixel. In order to allow efficient training, they approximate this behavior by training on random patches, for each of which a fixed number of pixels are randomly replaced by local neighbors, using their respective original noisy values as targets.

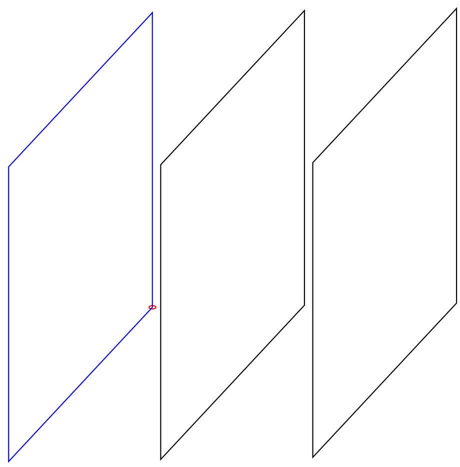

Figure 3.1: Receptive field in CNNs.

In Noise2Self (N2S) [2], Batson et al. generalize this concept to sets of variables, instead of single pixels only, by introducing the notion of $\mathcal{J}$-invariance.

Laine et al. [11] choose a different approach; instead of relying on masking strategies, they directly manipulate the receptive field by adapting the network architecture itself. In Noise2Same [27], the authors show that both N2V and N2S are not strictly $\mathcal{J}$-invariant and conclude that strict $\mathcal{J}$-invariance is thus not necessary for achieving good denoising performance. Therefore, they propose to do without explicit manipulation of the receptive field and instead add a regularization term that encourages the network to learn an approximately $\mathcal{J}$-invariant mapping by itself.

Chapter 4

Methods

While the self-supervised denoising methods discussed so far eliminate the need for clean data, most still require large datasets of noisy images or are limited to specific noise types. A notable exception is the Deep Image Prior (DIP) introduced by Ulyanov et al. [23]. DIP is a zero-shot method, i.e. it operates on a single noisy sample, and does not make any explicit assumptions about the noise distribution.

In this chapter, we present the DIP and various extensions. First, we describe the fundamental principles of DIP, followed by common regularization techniques such as early stopping and total variation. Finally, we explore other DIP-based approaches that build upon these foundations.

4.1 Deep Image Prior

4.1.1 Early Stopping

4.1.2 Total Variation

4.2 Deep Diffusion Image Prior

4.3 Self-Guided Deep Image Prior

Chapter 5

Experimental Setup

5.1 Architecture

5.2 Metrics

5.3 Datasets

Chapter 6

Results

Chapter 7

Discussion

Chapter 8

Conclusion

Bibliography

[1] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. *Layer Normalization*. 2016. arXiv: 1607.06450 [stat.ML]. URL: https://arxiv.org/abs/1607.06450.

[2] Joshua Batson and Loic Royer. "Noise2Self: Blind Denoising by Self-Supervision". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, Sept. 2019, pp. 524–533. URL: https://proceedings.mlr.press/v97/batson19a.html.

[3] Yangkang Chen et al. "Denoising of Distributed Acoustic Sensing Seismic Data Using an Integrated Framework". In: *Seismological Research Letters* 94.1 (Nov. 2022), pp. 457–472.

[4] Tim Dean, T. Cuny, and Arthur Hartog. "Determination of the Optimum Gauge Length for Borehole Seismic Surveys Acquired Using Distributed Vibration Sensing". In: June 2015. DOI: 10.3997/2214-4609.201412740.

[5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[6] Arthur H. Hartog. *An Introduction to Distributed Optical Fibre Sensors*. CRC Press, 2017. URL: http://dx.doi.org/10.1201/9781315119014.

[7] Jie Hu et al. *Squeeze-and-Excitation Networks*. 2019. arXiv: 1709.01507 [cs.CV]. URL: https://arxiv.org/abs/1709.01507.

[8] Tao Huang et al. "Neighbor2Neighbor: Self-Supervised Denoising From Single Noisy Images". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2021, pp. 14781–14790.

[9] Sergey Ioffe and Christian Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *CoRR* abs/1502.03167 (2015). arXiv: 1502.03167. URL: http://arxiv.org/abs/1502.03167.

[10] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. "Noise2Void - Learning Denoising From Single Noisy Images". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019, pp. 2124–2132. DOI: 10.1109/CVPR.2019.00223.

[11] Samuli Laine et al. "High-Quality Self-Supervised Deep Image Denoising". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019. URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/2119b8d43eafcf353e07d7cb5554170b-Paper.pdf.

[12] S Lapins et al. "DAS-N2N: machine learning distributed acoustic sensing (DAS) signal denoising without clean data". In: *Geophysical Journal International* 236.2 (Nov. 2023), pp. 1026–1041. ISSN: 1365-246X. DOI: 10.1093/gji/ggad460. URL: http://dx.doi.org/10.1093/gji/ggad460.

[13] Y. LeCun et al. "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4 (1989), pp. 541–551. DOI: 10.1162/neco.1989.1.4.541.

[14] Jaakko Lehtinen et al. "Noise2Noise: Learning Image Restoration without Clean Data". In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by Jennifer Dy and Andreas Krause. Vol. 80. Proceedings of Machine Learning Research. PMLR, Oct. 2018, pp. 2965–2974. URL: https://proceedings.mlr.press/v80/lehtinen18a.html.

[15] A. Lellouch et al. "Low-Magnitude Seismicity With a Downhole Distributed Acoustic Sensing Array—Examples From the FORGE Geothermal Experiment". In: *Journal of Geophysical Research: Solid Earth* 126.1 (Dec. 2020). ISSN: 2169-9356. DOI: 10.1029/2020jb020462. URL: http://dx.doi.org/10.1029/2020JB020462.

[16] Youssef Mansour and Reinhard Heckel. "Zero-Shot Noise2Noise: Efficient Image Denoising without any Data". In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2023, pp. 14018–14027. DOI: 10.1109/CVPR52729.2023.01347.

[17] Meghan S. Miller, John Townend, and Voon Hui Lai. "The South Island Seismology at the Speed of Light Experiment (SISSLE): Distributed Acoustic Sensing Across and Along the Alpine Fault, South Westland, New Zealand". In: *Seismological Research Letters* (Dec. 2024). ISSN: 0895-0695. DOI: 10.1785/0220240322. eprint: https://pubs.geoscienceworld.org/ssa/srl/article-pdf/doi/10.1785/0220240322/7068598/srl-2024322.1.pdf. URL: https://doi.org/10.1785/0220240322.

[18] Nick Moran et al. "Noisier2Noise: Learning to Denoise From Unpaired Noisy Data". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[19] Zhaoyang Niu, Guoqiang Zhong, and Hui Yu. "A review on the attention mechanism of deep learning". In: *Neurocomputing* 452 (2021), pp. 48–62. ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2021.03.091. URL: https://www.sciencedirect.com/science/article/pii/S092523122100477X.

[20] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors". In: *Nature* 323 (Oct. 1986), pp. 533–. URL: http://dx.doi.org/10.1038/323533a0.

[21] Irwin Sobel. "An Isotropic 3x3 Image Gradient Operator". In: *Presentation at Stanford AI Lab (SAIL)* (1968).

[22] "TODO". In: ().

[23] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. "Deep Image Prior". In: *International Journal of Computer Vision* 128.7 (Mar. 2020), pp. 1867–1888. ISSN: 1573-1405. DOI: 10.1007/s11263-020-01303-4. URL: http://dx.doi.org/10.1007/s11263-020-01303-4.

[24] Dmitry Ulyanov, Andrea Vedaldi, and Victor S. Lempitsky. "Instance Normalization: The Missing Ingredient for Fast Stylization". In: *CoRR* abs/1607.08022 (2016). arXiv: 1607.08022. URL: http://arxiv.org/abs/1607.08022.

[25] Ashish Vaswani et al. *Attention Is All You Need*. 2023. arXiv: 1706.03762 [cs.CL]. URL: https://arxiv.org/abs/1706.03762.

[26] Yuxin Wu and Kaiming He. "Group Normalization". In: *CoRR* abs/1803.08494 (2018). arXiv: 1803.08494. URL: http://arxiv.org/abs/1803.08494.

[27] Yaochen Xie, Zhengyang Wang, and Shuiwang Ji. "Noise2Same: Optimizing A Self-Supervised Bound for Image Denoising". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle et al. Vol. 33. Curran Associates, Inc., 2020, pp. 20320–20330. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/ea6b2efbdd4255a9f1b3bbc6399b58f4-Paper.pdf.

[28] Dan Zhang et al. *Unleashing the Power of Self-Supervised Image Denoising: A Comprehensive Review*. 2024. arXiv: 2308.00247 [eess.IV]. URL: https://arxiv.org/abs/2308.00247.

[29] Kai Zhang et al. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". In: *IEEE Transactions on Image Processing* 26.7 (July 2017), pp. 3142–3155. ISSN: 1941-0042. DOI: 10.1109/tip.2017.2662206. URL: http://dx.doi.org/10.1109/TIP.2017.2662206.

Appendix A

Appendix