Recent Advances in Machine Learning - Zero-Shot Noise2Noise

University of Siegen

Aakash Deshpande
aakash.deshpande@student.uni-siegen.de
1701953

Ahmad Alhaj Jnaid
ahmad.alhajjnaid@student.uni-siegen.de
1595707

Gopi Ravi Teja Gudupu
gopi.gudupu@student.uni-siegen.de
1710623

Abstract

In this report, we focused on the denoising of microscopic images. Initially, we added noise to clean images and then applied various filters, such as Median, wiener, BM3D to remove the noise. The effectiveness of these filters was measured using the Peak Signal-to-Noise Ratio (PSNR).

Subsequently, we utilized the code from the study "Zero-Shot Noise2Noise: Efficient Image Denoising without any Data" [6] to denoise the images. We compared all the PSNR values obtained from the different methods to evaluate their performance.

Our findings indicate that, depending on the type of noise and the image, BM3D sometimes outperformed deep learning (DL) methods. However, in general, DL methods achieved good results across most types of noise. This comprehensive comparison highlights the strengths and weaknesses of each denoising approach in preserving image quality and structural details essential for microscopic analysis.

1 Introduction

In recent years, the field of image denoising has witnessed significant advancements, mainly through the use of self-supervised neural networks. Traditional denoising methods typically rely on extensive datasets comprising pairs of clean and noisy images, which train neural networks to map noisy inputs to their clean counterparts. However, the collection and maintenance of such datasets are often expensive and time-consuming. Moreover, the performance of these dataset-dependent methods tends to degrade when tested on images from different distributions than those in the training set. These challenges have prompted interest in dataset-free denoising methods.

This paper introduces a novel approach to image denoising called the Zero-Shot Noise2Noise (ZS-N2N)[6] method. This method, which does not require any training data or prior knowledge of noise distribution, offers high-quality denoising. Inspired by the Noise2Noise [4] and Neighbor2Neighbor [3]techniques, this method utilizes a simple 2-layer network with minimal computational cost, making it suitable for scenarios with limited data and computing resources. Through experiments involving artificial, real-world camera, and microscope noise, ZS-N2N[6] consistently outperforms existing dataset-free methods, achieving superior denoising quality while significantly reducing computational demands. This approach leverages the concept of pixel-wise independent noise and operates effectively even when the noise statistics are unknown, thus providing a robust solution for diverse denoising applications.

2 Related Work

Image denoising has seen considerable advancements with the advent of machine learning, particularly deep learning techniques. Traditionally, supervised learning methods have been the cornerstone of state-of-the-art denoising performance. These networks learn to map noisy images to their clean counterparts through extensive training. While highly effective, these supervised methods are often limited by the availability of large datasets, the high cost of data acquisition, and the time-consuming nature of training.

Noise2Noise [4] marked a significant innovation by demonstrating that denoising networks could be trained without clean images. By using pairs of noisy images of the same scene, Noise2Noise [4] exploits the assumption that noise is zero-mean. This method maintains high denoising performance comparable to supervised methods, provided the noisy pairs are well-aligned, and the noise characteristics remain consistent. Following this, Neighbour2Neighbour (NB2NB)[3]extended the Noise2Noise [4] concept by generating pairs of noisy images from a single noisy image through sub-sampling.

Several approaches have emerged in zero-shot or dataset-free denoising, focusing on eliminating the need for training data. Noise2Fast[5] is one such method that builds on the principles of Noise2Noise [4] and Neighbour2Neighbour[3]. It applies these principles to grayscale images and uses a relatively large neural network, necessitating an early stopping criterion to prevent overfitting. Despite its innovative approach, Noise2Fast[5] is limited by its computational demands and restriction to grayscale images.

The Zero-Shot Noise2Noise (ZS-N2N)[6] method proposed in this paper advances the field of dataset-free denoising by employing a lightweight network with only 20k parameters and a novel regularization strategy. This approach reduces computational costs and generalizes well to various noise types and levels without requiring prior knowledge of the noise distribution. By building on the strengths of Noise2Noise [4] and Neighbour2Neighbour[3], ZS-N2N [6]achieves competitive denoising quality and often outperforms more complex, computationally intensive methods. This makes it particularly suitable for applications with limited data availability and computational resources, representing a significant step forward in practical image denoising.

3 Methodology

3.1 Non-DL Method: Filters & Working procedure

Prior to the widespread adoption of neural networks and deep learning techniques, images were pre-processed using conventional or non-deep learning methods known as filters. These filters were employed for a variety of purposes, including enhancing signal strength, removing noise or disturbance and more.

These days, engineers have a plethora of filters at their disposal to experiment with and apply the same to the image dataset in order to pre-process the images. Filters are broadly categorised into 2 types, Linear filters and Non-linear filters.While mean, gaussian, wiener are some of the filters classified as linear filters,there are numerous other filters that belong to the non-linear filters, including bilateral, median, and BM3D (block-matching and 3D filtering).

To understand the filtering better, additional noise can be added to captured image. These additional noises are added in the form of certain statistical disturbances with an intention that they mimic the real noises which may be introduced while the image is being captured, in this case by a microscope. The manifestation of these noises, including their intensity, depends on various factors such as the skill of the operator, capturing conditions, and the quality of the microscope.

To make sure the results are comparable with deep learning method, the following noises and filters were selected in line with the paper "Zero-Shot Noise2Noise: Efficient Image Denoising without any Data" [6].

Noises:

1.Gaussian noise: caused by inadequate lighting, electronic circuit noise or excessive temperatures.

2.Poisson noise: caused by inadequate photon detection, which is a particle that carries energy.

3.Uniform noise: caused by quantization errors and hence also called quantization noise.

Filters:

1.Median filter.

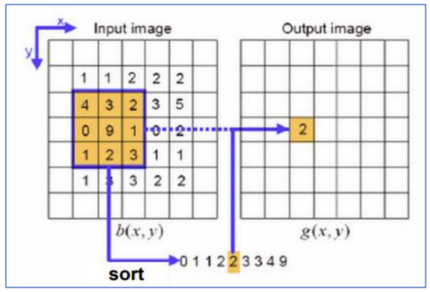

Figure 1: median filtering with kernel size 3 x 3[2]

2.Wiener filter. This filter's aim is to minimise the mean square error between the clean image and the denoised image. It modifies each pixel based on the local mean and variance within a defined window. The pixels are either smoothened or the details/edges are retained depending upon whether the local variance is low or high when compared to the noise variance.

3.BM3D filter(block-matching and 3D filtering).

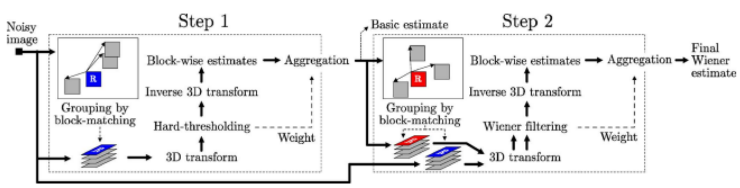

Figure 2: bm3d has 2 cascades: a hard-thresholding and a Wiener filter stage[1]

3.2 DL Method: Loss Functions & Working Model of Zero-Shot

ZS-N2N builds[6] on the N2N [4] and NB2NB [3] methods, while also utilizing a small, efficient neural network with 20k parameters.

Traditionally, a network f_θ is trained to map a noisy image $y_1 = x + e_1$ to a clean image x. Noise2Noise trains the network to map a noisy image to another noisy image of the same scene, i.e., y_1 to $y_2 = x + e_2$, where e_1 and e_2 are independent noise components. Formally, the network is trained by minimizing the mean squared error (MSE):

$$\theta_{N2N} = \arg\min_{\theta} \mathbb{E}\left[\|f_\theta(y_1) - y_2\|_2^2\right] \tag{1}$$

It can be shown that this is equivalent to minimizing the error between the noisy image and the clean image, assuming the noise has a zero mean:

$$\arg\min_{\theta} \mathbb{E}\left[\|f_\theta(y_1) - x\|_2^2\right] = \arg\min_{\theta} \mathbb{E}\left[\|f_\theta(y_1) - y_2\|_2^2\right] \tag{2}$$

NB2NB[3] extends N2N [4]by generating pairs of noisy images from a single noisy image through subsampling, making it more flexible. The pairs $D_1(y)$ and $D_2(y)$ are created by applying fixed filters:

$$D_1(y) = y * k_1 \quad \text{and} \quad D_2(y) = y * k_2 \tag{3}$$

where k_1 and k_2 are specific filters.

A lightweight network f_θ is trained with the following loss function:

$$L(\theta) = \|f_\theta(D_1(y)) - D_2(y)\|_2^2 \tag{4}$$

To enhance performance, residual learning and a symmetric loss are used:

$$L_{\text{res}}(\theta) = \frac{1}{2}\left(\|D_1(y) - f_\theta(D_1(y)) - D_2(y)\|_2^2 + \|D_2(y) - f_\theta(D_2(y)) - D_1(y)\|_2^2\right) \tag{5}$$

A consistency loss ensures consistent results:

$$L_{\text{cons}}(\theta) = \frac{1}{2}\left(\|D_1(y) - f_\theta(D_1(y)) - D_1(y - f_\theta(y))\|_2^2 + \|D_2(y) - f_\theta(D_2(y)) - D_2(y - f_\theta(y))\|_2^2\right) \tag{6}$$

The final loss function is:

$$L(\theta) = L_{\text{res}}(\theta) + L_{\text{cons}}(\theta) \tag{7}$$

The ZS-N2N[6] model works using the same mathematical method, in addition to the capability to operate with just one image and employing a very simple network with only 20K parameters and two layers. To avoid overfitting, an explicit regularization term $R(\theta)$ is added:

$$L_{\text{total}}(\theta) = \|f_\theta(D_1(y)) - D_2(y)\|_2^2 + \lambda R(\theta) \tag{8}$$

where λ is a regularization parameter.

After training, the network f_θ is applied to the original noisy image y to obtain the denoised image $\hat{x}$:

$$\hat{x} = f_\theta(y) \tag{9}$$

4 Experiments

Having a metric to measure the effectiveness of any experiment is very crucial as it helps us to make a fair comparison. In this particular experiment we have opted for PSNR(peak signal-to-noise ratio). Following are the results for an image with 3 different noises and their denoised versions along with their PSNR values.

4.1 Parameters used in Zero-shot DL Method

Hyperparameters: Zero-Shot Noise2Noise (ZS-N2N)[6] operates with a straightforward architecture, necessitating only a minimal set of hyperparameters. These key hyperparameters include:

- Learning rate: Determines the size of adjustments made to the model during training. Set to 0.001.
- Step size: Specifies how often the learning rate is reduced. Set to 1000 epochs.
- Number of epochs: Indicates the total training cycles. Set to 3000.
- Regularization: Helps prevent overfitting by penalizing complex models. Weight decay and dropout are used with a strength of 0.5.
- Optimizer: The algorithm used to minimize the model's error. Adam optimizer is employed.

4.2 Results Comparison

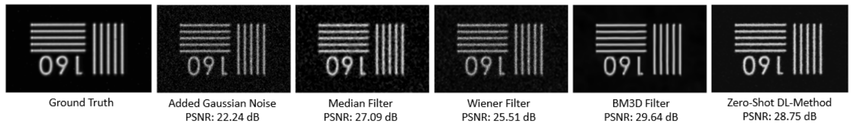

Figure 3: Gaussian noise added with mean 0 and std deviation 25 along with denoised images

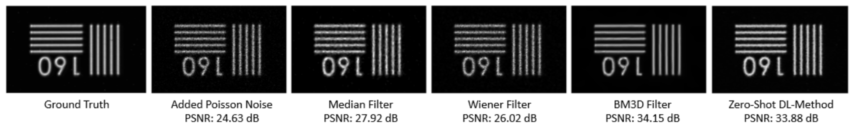

Figure 4: Poisson noise added with lambda 25 along with denoised images

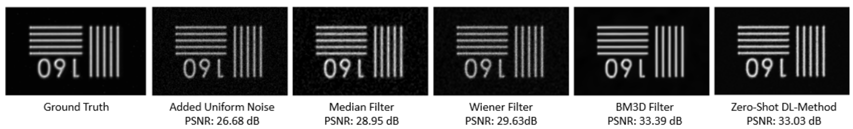

Figure 5: Uniform noise added with a noise parameter of 0.2 along with denoised images

Noise Type	**Noisy Image**	**Median Filter**	**Wiener Filter**	**BM3D Filter**	**ZS DL Method**
Gaussian	22.24	27.09	25.51	29.64	28.75
Poisson	24.63	27.92	26.02	34.15	33.88
Uniform	26.68	28.95	29.63	33.39	33.03

Table 1: PSNR (dB) values for different noise types and denoising methods.

5 Conclusion

This study explored various methods for denoising microscopic images, comparing traditional filters with deep learning approaches. Our results indicate that while filters like BM3D occasionally outperformed deep learning models for specific noise types, ZS-N2N demonstrates robust performance across various noise types, achieving not just competitive but superior Peak Signal-to-Noise Ratio (PSNR) values. The method's efficiency, requiring no prior data or extensive computational resources, makes it highly suitable for practical applications, particularly in scenarios with limited data. This study underscores the potential of zero-shot learning approaches in advancing the field of image denoising, offering a viable alternative to conventional methods. The results highlight the promise of ZS-N2N in maintaining image quality and structural integrity, making it a valuable tool for microscopic image analysis and other applications.

References

[1] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007. Student Member, IEEE and Senior Member, IEEE.

[2] Mahmoud Elmokadem. Discrete implementation for a mean and median filters, 2024. GitHub. Retrieved July 23, 2024.

[3] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu. Neighbor2neighbor: Self-supervised denoising from single noisy images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14781–14790, 2021.

[4] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. *arXiv preprint arXiv:1803.04189*, 2018.

[5] Jason Lequyer, Reuben Philip, Amit Sharma, Wen-Hsin Hsu, and Laurence Pelletier. A fast blind zero-shot denoiser. *Nature Machine Intelligence*, 4(11):953–963, October 2022.

[6] Youssef Mansour and Reinhard Heckel. Zero-shot noise2noise: Efficient image denoising without any data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14018–14027, 2023.