Course Project

Review of An Unsupervised Deep Learning Approach for
Real-World Image Denoising paper

Performed by Dmitry Pogrebnoy
ITMO Master student of a J42332c group

Introduction

In today's world, there are a large variety of systems that use digital images for their work. Various recognition and tracking systems, photo and video enhancement systems, intelligent systems for working with medical or geological images. All these require high-quality images with a minimal noise, because the quality of the result of such systems directly depends on it.

One of the works devoted to removing noise from images is the work of researchers Dihan Zheng et al. entitled "[An Unsupervised Deep Learning Approach For Real-World Image Denoising](#)" [1]. This paper was published at the International Conference on Learning Representations 2021 (ICLR 2021).

The authors draw attention to the fact that a large number of existing image noise removal methods use the assumption that the image contains only additive white Gaussian noise (AWGN). That is, that the noise in different parts of the image does not depend on what is depicted in these parts. It is convenient to synthetically generate such noise having a clean image and most denoising methods work reasonably well with such noise. But in the real world, noise in photos and images is almost always not AWGN. This happens because of various technical features of cameras and optical systems and there is no way to get rid of it completely. Due to the fact that the assumption that noise is AWGN is not held, these methods show mediocre results on real photos.

In addition, the authors point out that despite the great breakthrough in the quality of deep learning-based methods, these methods still require large datasets of better quality with clean and noisy photos for training, which requires huge resources in science-intensive fields such as medicine or geology. And even then, the data is often insufficient to train the model in all possible noise patterns.

Trying to solve the described problems with the existing methods, the authors propose their single-image dataset based method based on a combination of deep learning with traditional methods for unsupervised image denoising. Thus, the authors propose a method that will improve the quality of existing unsupervised denoising methods on real images.

The main idea is to use a neural network to convert the noisy image into the same image but with AWGN and then apply any of the traditional unsupervised image denoising techniques such as BM3D [2], DnCNN [3] or NLM [4] to the resulting image.

Model

The model consists of two parts. The first part is responsible for transforming an image with arbitrary noise into a space with additive white Gaussian noise. For this preconversion a neural network of encoder-decoder architecture is used. The second part of the model is any traditional unsupervised denoising model, the authors measured performance with BM3D, DnCNN and NLM models.

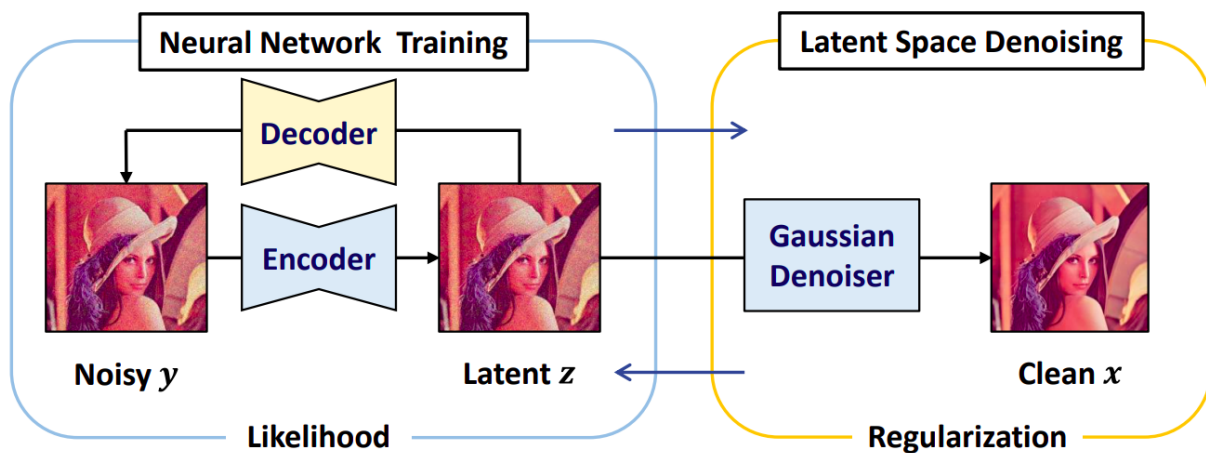A diagram with the model is shown below. The picture was taken from the reviewed article.

Figure 1. Diagram of the proposed model.

Denoising takes place in steps. At the first step the initial image is encoded, the change of the encoded image is applied and then decoding takes place. This operation is performed a certain number of times. The loss function takes into account the errors between the original image and the decoded image, as well as between the image in the latent space with noise AWGN and the image in the latent space after applying the traditional

denoising method to it. Then the traditional denoising method is applied to the obtained image in the latent space with AWGN noise. This completes the stage. In the next stage, everything is repeated, but using the results from the previous stage. The process continues for a predetermined number of steps. The detailed mathematical model and algorithms are described in the reviewed paper.

This approach has some advantages
- only one image is needed for denoising
- allows to correct any noise, not just AWGN
- allows to improve existing data processing pipelines without any additional costs

But the approach also has disadvantages:
- despite the lack of strong sensitivity to hyperparameters, it is necessary to adjust the parameters to achieve the best quality
- requires significant computing and time resources for quality results, respectively not suitable for real-time systems
- the results of our own experiments will show that the method makes the images a bit darker, probably it can be corrected by fine-tuning the parameters

Experimental result

For my experiments, I first took a photo from my laptop webcam and cut out the part with the noisy cabinet from this photo at 368x416 pixels in jpeg format. The image is shown below.
As you can see there is a lot of noise in the image and its intensity differs in different parts of the image. I then applied the described approach to this image. I used BM3D as a traditional denoiser.
The results of using NN+BM3D are shown below.
The NN+BM3D result after the first epoch.

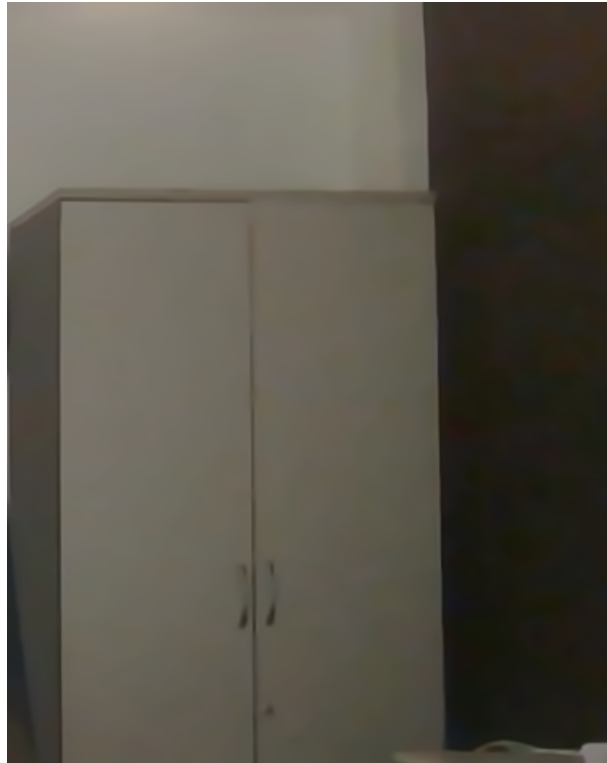Figure 2. Example of real noisy photo

Figure 3. Result of denoising with NN+BM3D after the first epoch.

The NN+BM3D after tenth epochs and the only BM3D results presented in Figure 4.
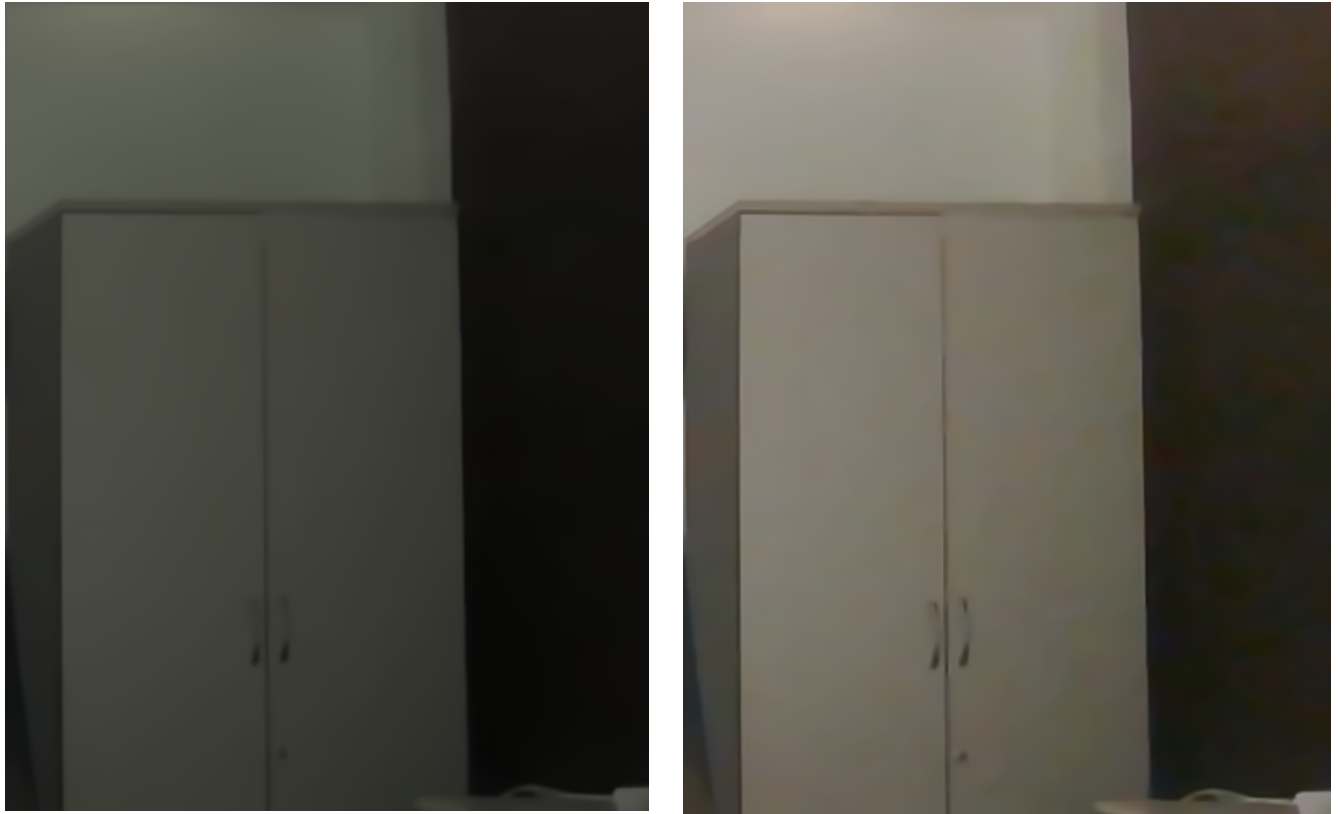

Figure 4. Left photo - NN+BM3D denoising result after tenth epochs. Right photo - BM3D denoising result.

It can be noted that even though the photo has become a little darker, the noise on it has been eliminated better than with the BM3D method. This can be seen even with the naked eye without any metrics.

But one picture is not enough for a good experiment. The authors of the article tested their approach on five of the most popular datasets. However, they left out the NIND dataset [5], which is the one I took for my experiments. This dataset contains a large set of high-resolution images. Each image has several variations with different noise content, from almost complete absence of noise, to noticeable noise.

I took 78 sets of images from this dataset. In each set, I selected the two images with the most and the least noise. I cut out the top left square of 250x250 pixels from those images and used that data for the experiment. All these operations are necessary to reduce the size of the task and to reduce the time of the experiment. Even with this volume experiment with denoising these images on Google Colab with Nvidia T4 graphics card took more than three hours.

As a result, the average PSRN [6] and SSIM [7] metrics for NN+BM3D and BM3D are presented in Table 1.

	AVG PSRN	AVG SSIM
BM3D	24.059	0.381
NN + BM3D	**28.857**	**0.651**

Table 1. Average PSRN and SSIM metrics for NN+BM3D and BM3D denoising methods results.

And the few results are shown in Figure 5 and Figure 6. The left are NN + BM3D, the right are BM3D.
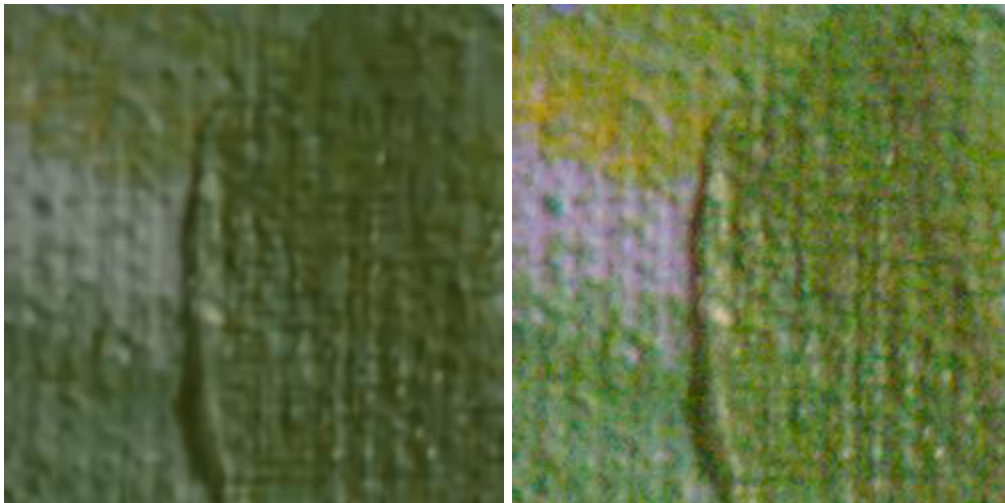

Figure 5. Left photo - denoising result by NN + BM3D method.
Right photo - denoising result by BM3D method.

Figure 5. Left photos - denoising results by NN + BM3D method.
Right photos - denoising results by BM3D method.

Conclusion

In this work, the article "An Unsupervised Deep Learning Approach For Real-World Image Denoising" was studied and reviewed. The aims of the work and the advantages and disadvantages of the proposed approach were described. The proposed model is briefly described. Experiments with our data as well as with a [public dataset of photos NIND](#) were carried out. The results of the experiments are described. As expected, the results of the proposed method were better than those of the traditional BM3D. Examples of images from the experiments are presented. All code can be found in the fork of the original repository at [this link](#).

References

[1] Zheng, Dihan, Sia Huat Tan, Xiaowen Zhang, Zuoqiang Shi, Kaisheng Ma and Chenglong Bao. "An Unsupervised Deep Learning Approach for Real-World Image Denoising." ICLR (2021).

[2] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans. Image Process., 16(8):2080– 2095, 2007b.

[3] Zhang, Kai & Zuo, Wangmeng & Chen, Yunjin & Meng, Deyu & Zhang, Lei. (2016). Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. IEEE Transactions on Image Processing.

[4] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In CVPR, volume 2, pp. 60–65. IEEE, 2005.

[5] Brummer, Benoit and De Vleeschouwer, Christophe. Natural Image Noise Dataset. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2019.

[6] Fardo, Fernando & Conforto, Victor & Oliveira, Francisco & Rodrigues, Paulo. (2016). A Formal Evaluation of PSNR as Quality Measurement Parameter for Image Segmentation Algorithms.

[7] Zhou Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, 2004.