Image Restoration without Clean Data using CNN

Zhongling Wang

University of Waterloo

Department of Electrical & Computer Engineering

July 11, 2019

Outline

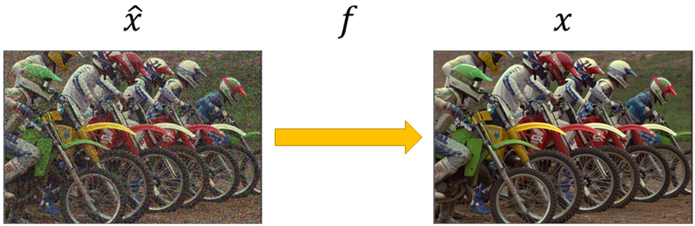
1. Introduction
   - Supervised Method
   - Unsupervised Method
2. Unsupervised CNN-based Method
   - Deep Image Prior [Lempitsky, 2018]
   - Noise2Noise [Lehtinen, 2018]
   - Net-SURE [Shakarim, 2018]
3. Summary

Outline

Learning-based Image Restoration

- Pristine Image: x
- Corrupted Image: $\hat{x}$
- **Objective**: Find the restoration function $f(\cdot)$, such that

$$x = f(\hat{x})$$



Outline

Supervised Image Restoration

- Objective: find optimal f **with** access to the clean image x.

$$\min_{f} \|f(\hat{x}) - x\|_2^2$$

- $f(\cdot)$ could be:
  - Autoencoder
  - CNN
- Drawbacks:
  - Clean image can be hard to obtain in some scenarios (CT scans).
  - Generalizability may be limited.

Outline

Unsupervised Image Restoration

- Objective: find optimal f **without** access to the clean image x.

$$\min_f \|f(\hat{x}) - \hat{x}\|_2^2 + g(f(\hat{x}))$$

- $\|f(\hat{x}) - \hat{x}\|_2^2$: content similarity
- $g(y)$: regularizer
  - Image **smoothness** measure:

$$\sum_i |y_{i+1} - y_i| \quad \text{(Total Variation)}$$

  - Image **naturalness** measure:

$$-log\{p(y)\} \quad \text{(Image Prior)}$$

Outline

Outline

Deep Image Prior [Lempitsky, 2018]

- Assumption:
  - CNN itself can capture low-level image statistics **before any learning**.
- Objective:

$$\begin{aligned} &\min_f \|f(\hat{x}) - \hat{x}\|_2^2 + g(f(\hat{x})) \\ =&\min_f \|f(\hat{x}) - \hat{x}\|_2^2 \end{aligned} \tag{1}$$

  - $f(\cdot)$ is a CNN.
  - $g(\cdot) \equiv 0$.
- Remarks:
  - Need to train a model for each image.
  - Need early stopping in training.

How to improve ?

- Original objective: $\min_f \|f(\hat{x}) - \hat{x}\|_2^2 + g(f(\hat{x}))$
- CNN is already a good prior.
- How can we better estimate $\|f(\hat{x}) - x\|_2^2$ from the corrupted data $\hat{x}$?

Outline

Noise2Noise [Lehtinen, 2018]

- Assumptions:
  - Noise n has **zero mean**: $\mathbb{E}[n] = 0$.
- Estimating the optimal point:

$$f_1^* = \underset{x}{\operatorname{argmin}} \underset{a}{\mathbb{E}}[\|x - a\|_2^2]$$

$$f_2^* = \underset{x}{\operatorname{argmin}} \underset{b}{\mathbb{E}}[\|x - b\|_2^2]$$

$$\Rightarrow f_1^* = f_2^* \text{ as long as } \mathbb{E}[a] = \mathbb{E}[b]$$

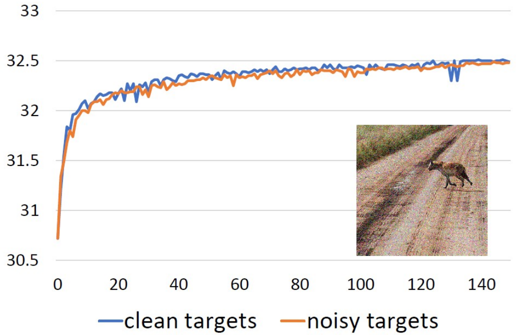
- Objective:

$$\begin{aligned} &\min_f \underset{x \in X}{\mathbb{E}}[\|f(\hat{x}) - x\|_2^2] \\ = &\min_f \underset{x \in X}{\mathbb{E}}[\|f(\hat{x}) - \hat{x}\|_2^2] \end{aligned} \tag{2}$$

  - $\mathbb{E}[\hat{x}] = \mathbb{E}[x]$ since noise has zero mean.
  - $f(\cdot)$ is a CNN.

Noise2Noise

- Convergence speed:



(a) **White Gaussian,** $\sigma = 25$

Noise2Noise

- PSNR Comparison:

Table 1. PSNR results from three test datasets KODAK, BSD300, and SET14 for Gaussian, Poisson, and Bernoulli noise. The comparison methods are BM3D, Inverse Anscombe transform (ANSC), and deep image prior (DIP).

	Gaussian (σ=25)			**Poisson (λ=30)**			**Bernoulli (p=0.5)**		
	clean	noisy	BM3D	clean	noisy	ANSC	clean	noisy	DIP
Kodak	32.50	32.48	31.82	31.52	31.50	29.15	33.01	33.17	30.78
BSD300	31.07	31.06	30.34	30.18	30.16	27.56	31.04	31.16	28.97
Set14	31.31	31.28	30.50	30.07	30.06	28.36	31.51	31.72	30.67
Average	**31.63**	**31.61**	**30.89**	**30.59**	**30.57**	**28.36**	**31.85**	**32.02**	**30.14**

Noise2Noise

- Remarks:
  - A single model can be applied to all images with the same type of noise.
  - Training dataset does not need to be really large in order to get zero mean.
  - Real world noise may not has zero mean.

Outline

Net-SURE [Shakarim, 2018]

- Assumptions:
  - Noise n is **Additive Gaussian**: $\hat{x} = x + n$ and $n \backsim N(\mu, \sigma^2)$, $\hat{x}, x, n \in \mathbb{R}^K$.
- **Estimate** the MSE without target
  - Stein's Unbiased Risk Estimator (SURE):

$$\mathop{\mathbb{E}}_{n}[\|f(\hat{x}) - x\|_2^2] = K \cdot \mathop{\mathbb{E}}_{n}[SURE(f(\hat{x}))]$$

  - $SURE(f(\hat{x}))$:

$$SURE(f(\hat{x})) = \sigma^2 + \frac{\|h(\hat{x})\|_2^2}{K} + \frac{2\sigma^2}{K}\sum_{i=1}^{K}\frac{\partial h_i(\hat{x})}{\partial \hat{x}_i}$$

$$f(\hat{x}) = \hat{x} + h(\hat{x})$$

Net-SURE

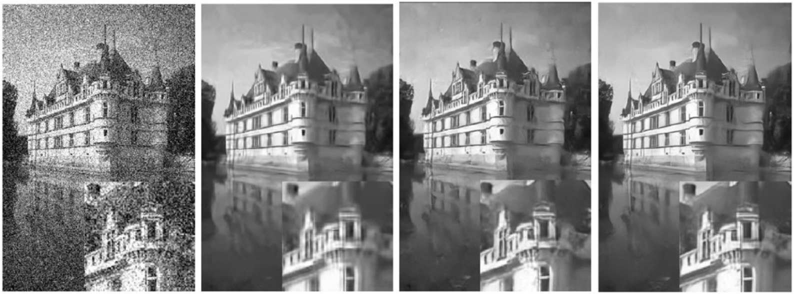
- Objective:

$$\begin{aligned} &\min_f \mathbb{E}[\|f(\hat{x}) - x\|_2^2] \\ =& \min_f \mathbb{E}[SURE(f(\hat{x}))] \\ =& \min_f \mathbb{E}[\|f(\hat{x}) - \hat{x}\|_2^2] + \mathbb{E}[\cdots] \end{aligned} \tag{3}$$

  Basically, it's Noise2Noise with **regularization**.
- Remarks:
  - $f(\cdot)$ should be differentiable.
  - SURE fit in CNN based $f(\cdot)$ perfectly.
  - Can deal with non-zero-mean noise, but it has to be Gaussian.

Net-SURE

- Performance comparison



Noisy image / 14.76dB | BM3D / 26.14dB (baseline) | SURE / 26.46dB (unsupervised) | MSE / 26.85dB (supervised)

Outline

Summary

- Comparison of Denoising Ability

	Noise Type	Application	Approximation
Deep Image Prior	Anything	1 for 1	N.A.
Noise2Noise	Zero Mean	1 for many	Optimal Point
Net-SURE	Gaussian	1 for many	MSE Loss

- Remarks
  - Unsupervised methods are still inferior than there supervised counterpart.
  - Hard to generalize on real world noisy images.

Thank you

References

V. Lempitsky, A. Vedaldi, D. Ulyanov. (2018)
Deep Image Prior.
In *IEEE conference on Computer Vision and Pattern Recognition (CVPR).*

J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila. (2018)
Noise2Noise: Learning Image Restoration without Clean Data.
In *International Conference on Machine Learning (ICML).*

S. Shakarim, and S. Y. Chun. (2018)
Training Deep Learning Based Denoisers without Ground Truth Data.
In *Advances in Neural Information Processing Systems (NIPS).*