The New Era of Image Denoising

Michael Elad

Computer Science Department
The Technion - Israel Institute of Technology
Haifa 32000, Israel

Verily Research

April 10-21, 2023

A Multiscale Tour of Harmonic Analysis & Machine Learning
Celebrating Stéphane Mallat's 60 Birthday

Few Preliminary Words …

Thank You to Gabriel Peyre and Joan Bruna for inviting me to

T h i s   A m a z i n g   E v e n t ! ! !

Being here is both a Pleasure and an Honor

Few Preliminary Words …

Happy Birthday, Stephane
and
Thank you for your guidance over the years

Invariant Scattering Convolution Networks

Joan Bruna and Stéphane Mallat
CMAP, Ecole Polytechnique, Palaiseau, France

Abstract—A wavelet scattering network computes a translation invariant image representation, which is stable to deformations and preserves high frequency information for classification. It cascades wavelet transform convolutions with non-linear modulus and averaging operators. The first network layer outputs SIFT-type descriptors whereas the next layers provide complementary invariant information which improves classification. The mathematical analysis of wavelet scattering networks explain important properties of deep convolution networks for classification.

as wavelets. However, wavelet transforms are not invariant to translations. Building invariant representations from wavelet coefficients requires introducing non-linear operators, which leads to a convolution network architecture.

Deep convolution networks have the ability to build large-scale invariants which are stable to deformations

The New Era of Image Denoising

Michael Elad

Computer Science Department
The Technion - Israel Institute of Technology
Haifa 32000, Israel

Verily Research

April 10-21, 2023
A Multiscale Tour of Harmonic Analysis & Machine Learning
Celebrating Stéphane Mallat's 60 Birthday

This Lecture is About …

Image Denoising

Noisy image

Image Denoiser

Denoised image

Removal of noise from images is a heavily studied problem in image processing

In this talk we expand on recent discoveries and developments around this seemingly dead topic

Michael Elad
The Computer-Science Department
The Technion

Our Agenda

1. Brief Introduction & History
2. Our Focus Today: Denoising for …
   - Image Synthesis
   - High perceptual quality recovery
3. Summary

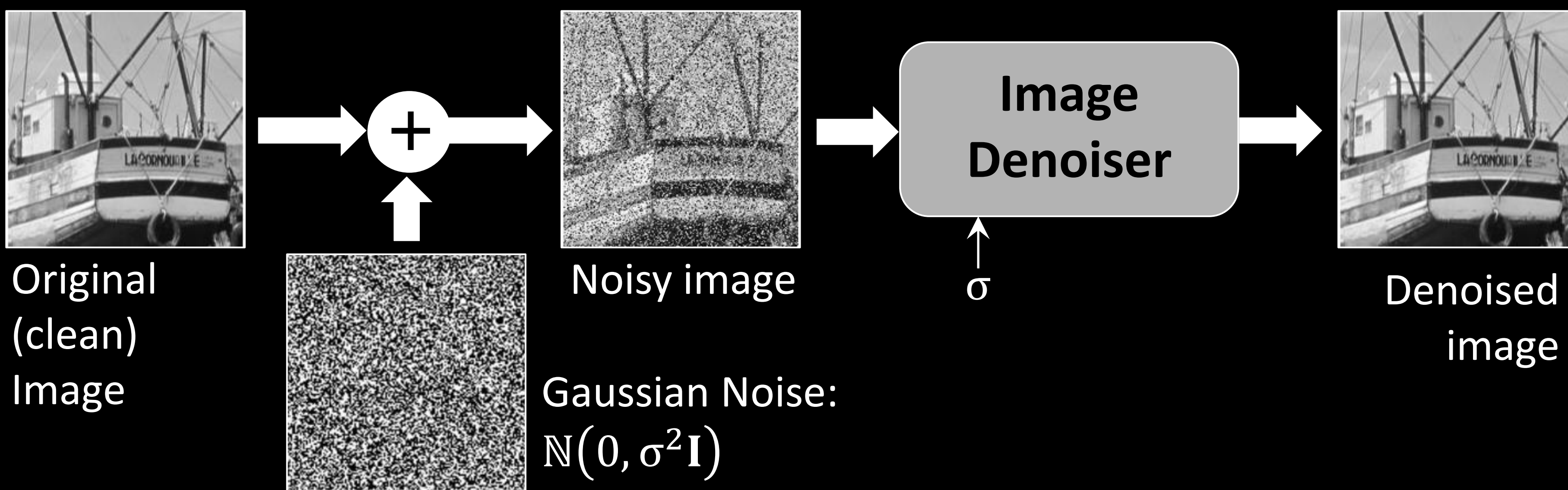
Introduction & History

So, Let's Talk About …

Image Denoising

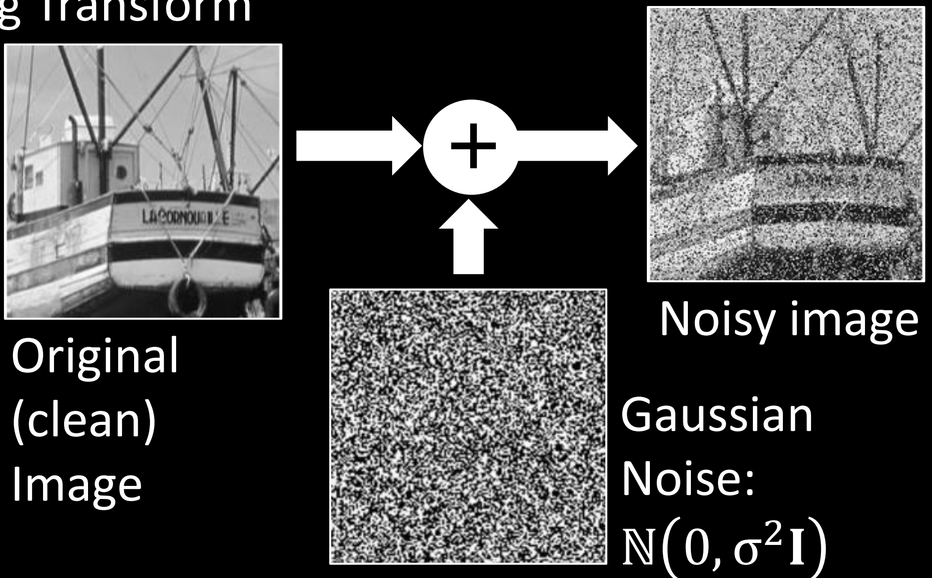
or more accurately

Removal of White Additive Gaussian Noise from an Image



Why Assume Gaussian Noise?

- The Gaussian case is more common and much more important
- When considering a Poisson noise,
  - High count of photons – The distribution gets closer and closer to the Gaussian case
  - Low-count Poisson-distributed image can be converted to a Gaussian-noisy one by Anscomb - Variance Stabilizing Transform
- Many of the developed ideas for the Gaussian case can be converted to other noise models
- MMSE denoisers for the Gaussian case are of extreme theoretical value (see later)



Michael Elad
The Computer-Science Department
The Technion

Image Denoising: Evolution

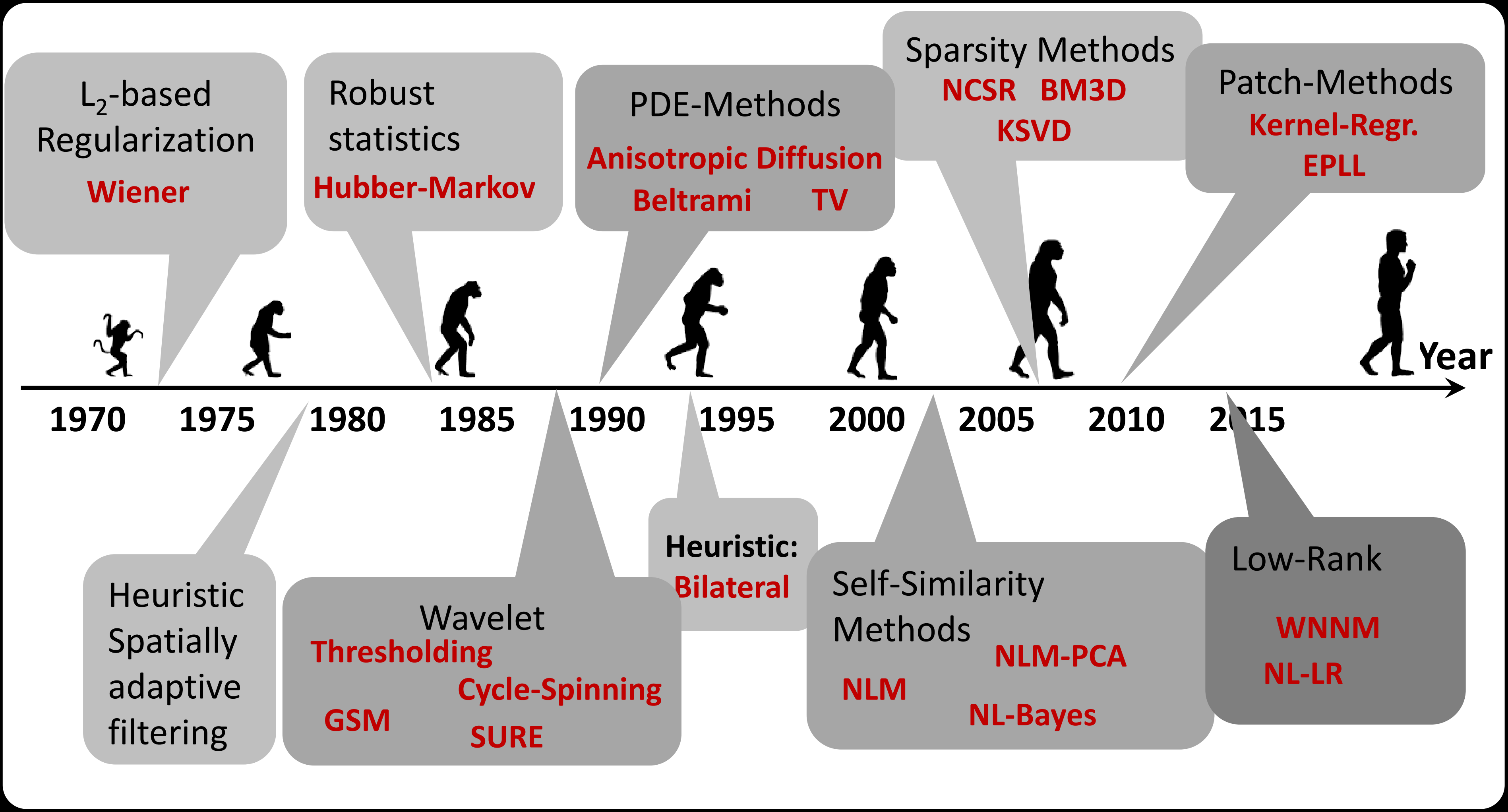


Image Denoising: A Paradigm Shift

How can we design a denoiser?

By **modeling** image content and leveraging it for noise filtering:

Classics

Sparse Representation

Scale Invariance

Low Rank

Piecewise Smoothness

Non-Local Self-Similarity

GMM

Low dimensionality

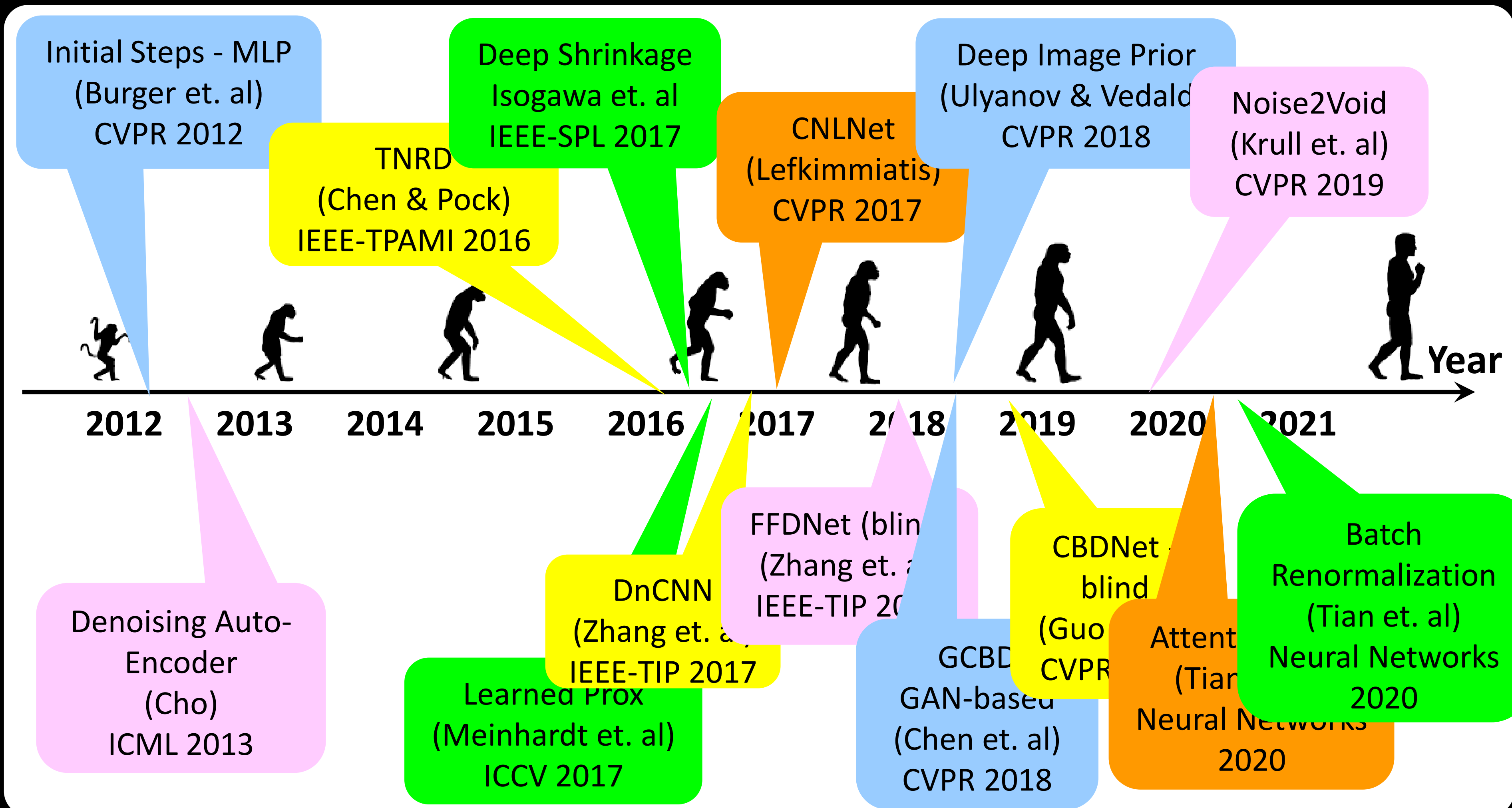
Learning

Supervised Training

Observe that with this trend, all the knowledge and knowhow accumulated carefully over decades in image processing became obsolete – is it true?

Michael Elad
The Computer-Science Department
The Technion

Image Denoising: Recent Evolution



Michael Elad
The Computer-Science Department
The Technion

Our Focus Today:
Recent Discoveries

Our Focus Today

Recent findings on using denoisers for other tasks:

- ❑ ~~Discovery 0: Solving general inverse problems [2013-]~~
- ❑ Discovery 1: Image Synthesis [2019-]
- ❑ Discovery 2: High perceptual quality recovery [2021-]

We turn to describe these results

Noisy image y

Image Denoiser

$D(y, \sigma)$

σ

Denoised image $\hat{x}$

Michael Elad
The Computer-Science Department
The Technion

Discovery 1: Image Synthesis

- In recent years, and with the deep-learning revolution, there is a growing interesting is synthesizing images “out of thin air”
- The popular tool of interest is called GAN – Generative Adversarial Network, built of two competing networks – a generator and a critique
- Why synthesize? Because
  - We can, and it is fascinating
  - It testifies that we have seized the distribution of images, and
  - It could be used for other needs
- Could we synthesize images differently?

Michael Elad
The Computer-Science Departm
The Technion

Discovery 1: Image Synthesis

Question: Given a denoiser $\mathrm{D}(\mathrm{y}, \sigma)$ how can one synthesize images with it?

Generative modeling by estimating gradients of the data distribution Y Song, S Ermon Advances in Neural Information Processing Systems 32	235	2019
Improved techniques for training score-based generative models Y Song, S Ermon Advances in neural information processing systems 33, 12438-12448	88	2020
Stochastic Solutions for Linear Inverse Problems using the Prior Implicit in a Denoiser Z Kadkhodaie, EP Simoncelli Advances in Neural Information Processing Systems 34	2	2021

Answer: Use $\mathrm{D}(\mathrm{y}, \sigma)$ as a ***Projector*** onto the image manifold

Practical Implication: Iterated use of $\mathrm{D}(\cdot, \sigma)$ with varying σ

noise → Simple Operation → $\mathrm{D}(\cdot, \sigma_1)$ → Simple Operation → $\mathrm{D}(\cdot, \sigma_2)$ → Simpl Operati

Michael Elad
The Computer-Science Department
The Technion

Discovery 1: Image Synthesis

Here is the core idea in a nutshell:

Our goal: draw a sample from the distribution of images $\mathrm{P}(\mathrm{x})$

- Start with a random noise image $\hat{\mathrm{x}}_0$
- Climb to a more probable image by the iterative equation:

$$\hat{\mathrm{x}}_{\mathrm{k}+1} = \hat{\mathrm{x}}_{\mathrm{k}} + \mathrm{a} \cdot \underbrace{\boxed{\nabla \log \mathrm{P}(\hat{\mathrm{x}}_{\mathrm{k}})}} + \mathrm{b} \cdot \mathrm{z}_{\mathrm{k}}$$ (Langevin Dynamics)

This is known (Miyasawa `61) as the Score Function and it is approximately proportional to $[\hat{\mathrm{x}}_{\mathrm{k}} - \mathrm{D}(\hat{\mathrm{x}}_{\mathrm{k}}, \sigma)]$ for a small value of σ

$\Leftrightarrow$

This suggests an implicit relation between MMSE denoisers and Priors: $\mathrm{D}(\mathrm{x}, \sigma) \leftrightarrow \mathrm{P}(\mathrm{x})$

… and this way we got an iterated algorithm that keeps calling to a denoiser, and is guaranteed to obtain a sample from $\mathrm{P}(\mathrm{x})$

Discovery 1: Image Synthesis

In practice, instead of the plain Langevin with a fixed (and small) value of σ we use the Annealed Langevin Algorithm that considers a sequence of blurred priors:

$$P(x+v) \quad \text{for} \quad v \sim \mathbb{N}\left(0, \sigma_k^2 \mathbf{I}\right)$$

$$= P(x) \otimes c \cdot \exp\left\{-\frac{1}{2\sigma^2}\|x\|^2\right\}$$

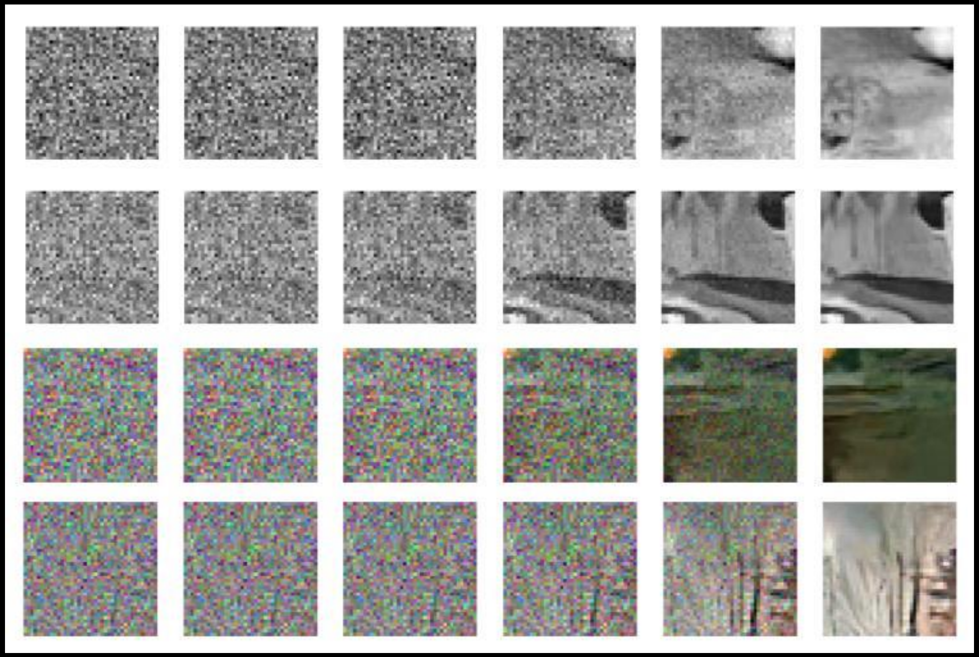
$$\text{with} \quad \sigma_0 > \sigma_1 > \sigma_2 \; \cdots \; > \sigma_N > 0$$

The core idea: start by drawing from a wider distribution and gradually narrow it, leading to a faster sampling performance

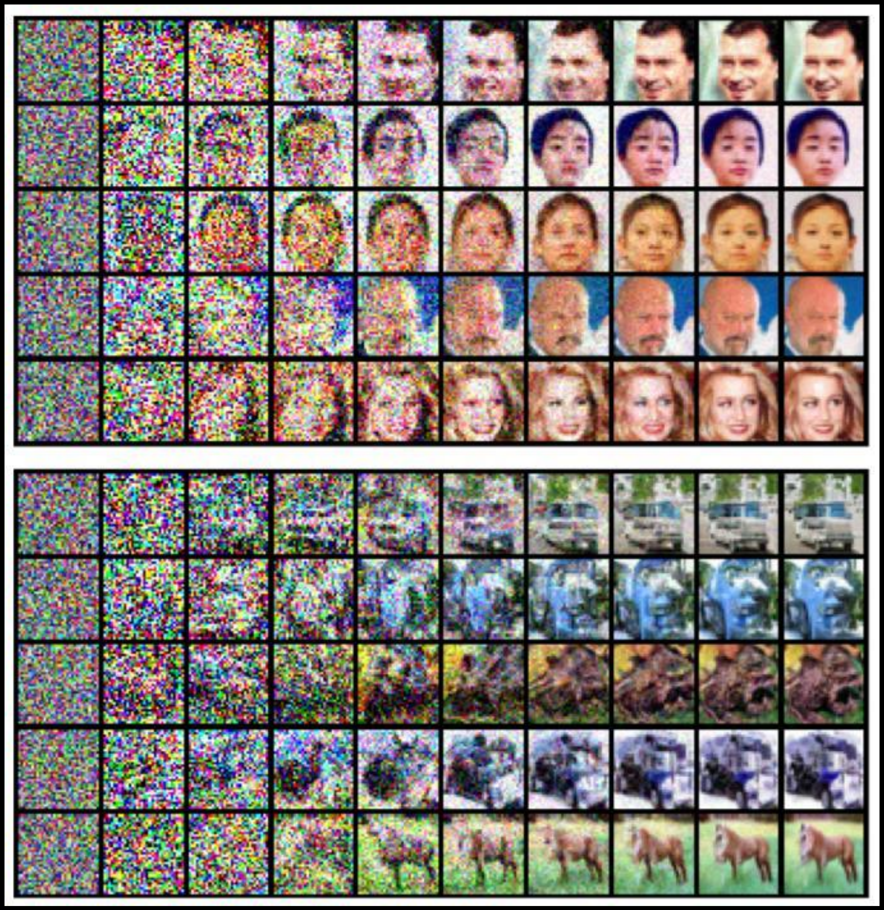
Michael Elad
The Computer-Science Department
The Technion

Discovery 1: Image Synthesis

Does it work? Here are some results



Kadkhodaie & Simoncelli

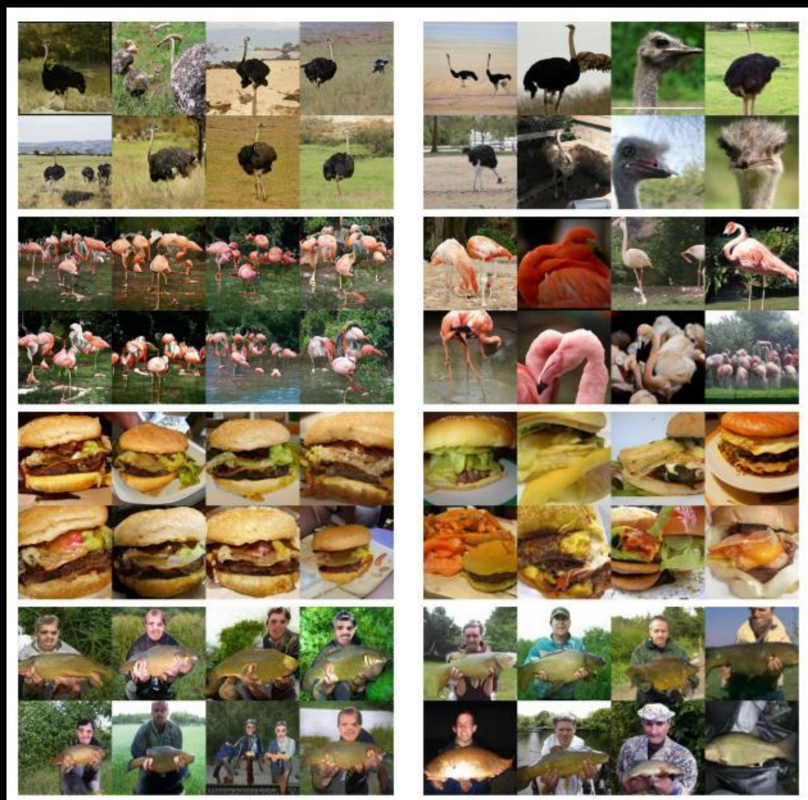


Song & Ermon

Michael Elad
The Computer-Science Department
The Technion

Discovery 1: Image Synthesis

Claim: diffusion-based methods are the best in image synthesis



BigGAN (FID 6.95) Diffusion (FID 4.59)

arXiv:2105.05233v4 [cs.LG] 1 Jun 2021

Diffusion Models Beat GANs on Image Synthesis

Prafulla Dhariwal* OpenAI prafulla@openai.com

Alex Nichol* OpenAI alex@openai.com

Abstract

We show that diffusion models can achieve image sample quality superior to the current state-of-the-art generative models. We achieve this on unconditional image synthesis by finding a better architecture through a series of ablations. For conditional image synthesis, we further improve sample quality with classifier guidance: a simple, compute-efficient method for trading off diversity for fidelity using gradients from a classifier. We achieve an FID of 2.97 on ImageNet 128×128, 4.59 on ImageNet 256×256, and 7.72 on ImageNet 512×512, and we match BigGAN-deep even with as few as 25 forward passes per sample, all while maintaining better coverage of the distribution. Finally, we find that classifier guidance combines well with upsampling diffusion models, further improving FID to 3.94 on ImageNet 256×256 and 3.85 on ImageNet 512×512. We release our code at https://github.com/openai/guided-diffusion.

1 Introduction

Figure 1: Selected samples from our best ImageNet 512×512 model (FID 3.85)

Discovery 1: Image Synthesis

Surely, you have heard of …

Imagen

unprecedented photorealism × deep level of language understanding

Google

DALL·E 2

OpenAI

Michael Elad
The Computer-Science Department
The Technion

Discovery 1: Image Synthesis

Surely, you have heard of …

Stable Diffusion

stability.ai

WAKEUP THE ARTIST IN YOU

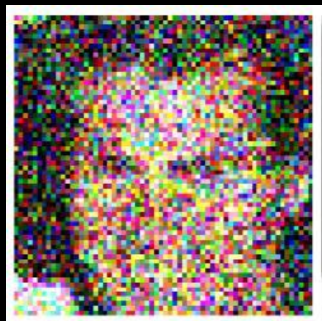
Midjourney

ARTIFICIAL INTELLIGENCE ART GENERATOR

Michael Elad
The Computer-Science Department
The Technion

Discovery 2: Targeting Perceptual Quality

Suppose that we need to denoise the following image:



Noisy ($\sigma = 100$)

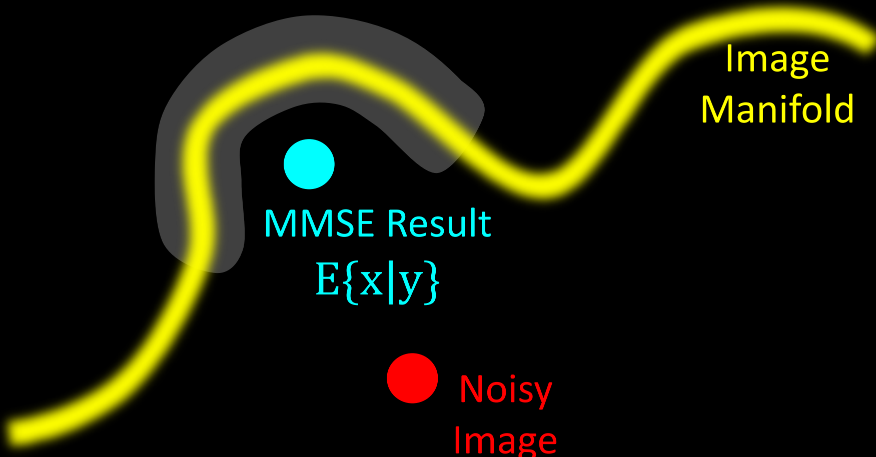
Image Denoiser

σ

Denoised image

Original

Should we be pleased with this result? It seems to be a bit … blurry, no? Why?



Minimum Mean-Squared-Error (MMSE) denoisers are great for MSE result, but their result falls outside the manifold

Michael Elad
The Computer-Science Department
The Technion

Discovery 2: Targeting Perceptual Quality

Question: How can we denoise an image while targeting “High Perceptual Quality”?

High perceptual quality image denoising with a posterior sampling cgan G Ohayon, T Adrai, G Vaksman, M Elad, P Milanfar Proceedings of the IEEE/CVF International Conference on Computer Vision …	3	2021
Stochastic image denoising by sampling from the posterior distribution B Kawar, G Vaksman, M Elad Proceedings of the IEEE/CVF International Conference on Computer Vision …	5	2021

Answer: Denoise by sampling from the posterior $\mathrm{P}(\mathrm{x}|\mathrm{y})$

Practical Implication: We consider 2 methods

- Training a deep denoiser via CGAN, or
- Iterated use of an MMSE denoiser $\mathrm{D}(\cdot, \sigma)$

These methods produce a multitude of possible solutions

Discovery 2: Targeting Perceptual Quality

Let's have a closer look at the Stochastic Image Denoiser:

Task: Draw a sample from $\mathrm{P}(\mathrm{x}|\mathrm{y})$ where $[\mathrm{y}=\mathrm{x}+\mathrm{n}, \mathrm{n}\sim\mathbb{N}\left(0,\sigma_0^2\mathbf{I}\right)]$

- Start with a random noise image $\hat{\mathrm{x}}_0$
- Climb to a more probable image by the iterative equation:

$$\hat{\mathrm{x}}_{\mathrm{k}+1}=\hat{\mathrm{x}}_{\mathrm{k}}+\mathrm{a}\cdot\underbrace{\nabla\mathrm{logP}(\hat{\mathrm{x}}_{\mathrm{k}}|\mathrm{y})}+\mathrm{b}\cdot\mathrm{z}_{\mathrm{k}}$$

Langevin with a conditional Score

$$=\nabla\mathrm{logP}(\hat{\mathrm{x}}_{\mathrm{k}})+\nabla\mathrm{logP}(\mathrm{y}|\hat{\mathrm{x}}_{\mathrm{k}})$$

Bayes rule

$$=\underbrace{\hat{\mathrm{x}}_{\mathrm{k}}-\mathrm{D}(\hat{\mathrm{x}}_{\mathrm{k}},\sigma)}_{\text{Approx. Score}}+\underbrace{\nabla\mathrm{logP}(\mathrm{y}|\hat{\mathrm{x}}_{\mathrm{k}})}_{\text{A Gaussian Distribution ?}}$$

Discovery 2: Targeting Perceptual Quality

Let's have a closer look at the Stochastic Image Denoiser:

$$\nabla \log P(\hat{x}_k|y) = \hat{x}_k - D(\hat{x}_k, \sigma) + \nabla \log P(y|\hat{x}_k)$$

- As we use the Annealed Langevin algorithm, there are two noise signals to consider:
  - Measurement's noise: $n \sim \mathbb{N}\left(0, \sigma_0^2 \mathbf{I}\right)$
  - Synthetic annealing noise: $v \sim \mathbb{N}\left(0, \sigma_k^2 \mathbf{I}\right)$ for $\sigma_0 > \sigma_1 > \sigma_2 \ \cdots \ > \sigma_N > 0$
- Implication: We recover a sequence of gradually less noisy images $\hat{x}_k$ where their noise v is assumed to be a portion of n

$$\nabla \log P(\hat{x}_k|y) = \hat{x}_k - D(\hat{x}_k, \sigma_k) + \frac{y - \hat{x}_k}{\sigma_0^2 - \sigma_k^2}$$

Michael Elad
The Computer-Science Department
The Technion

Discovery 2: Targeting Perceptual Quality

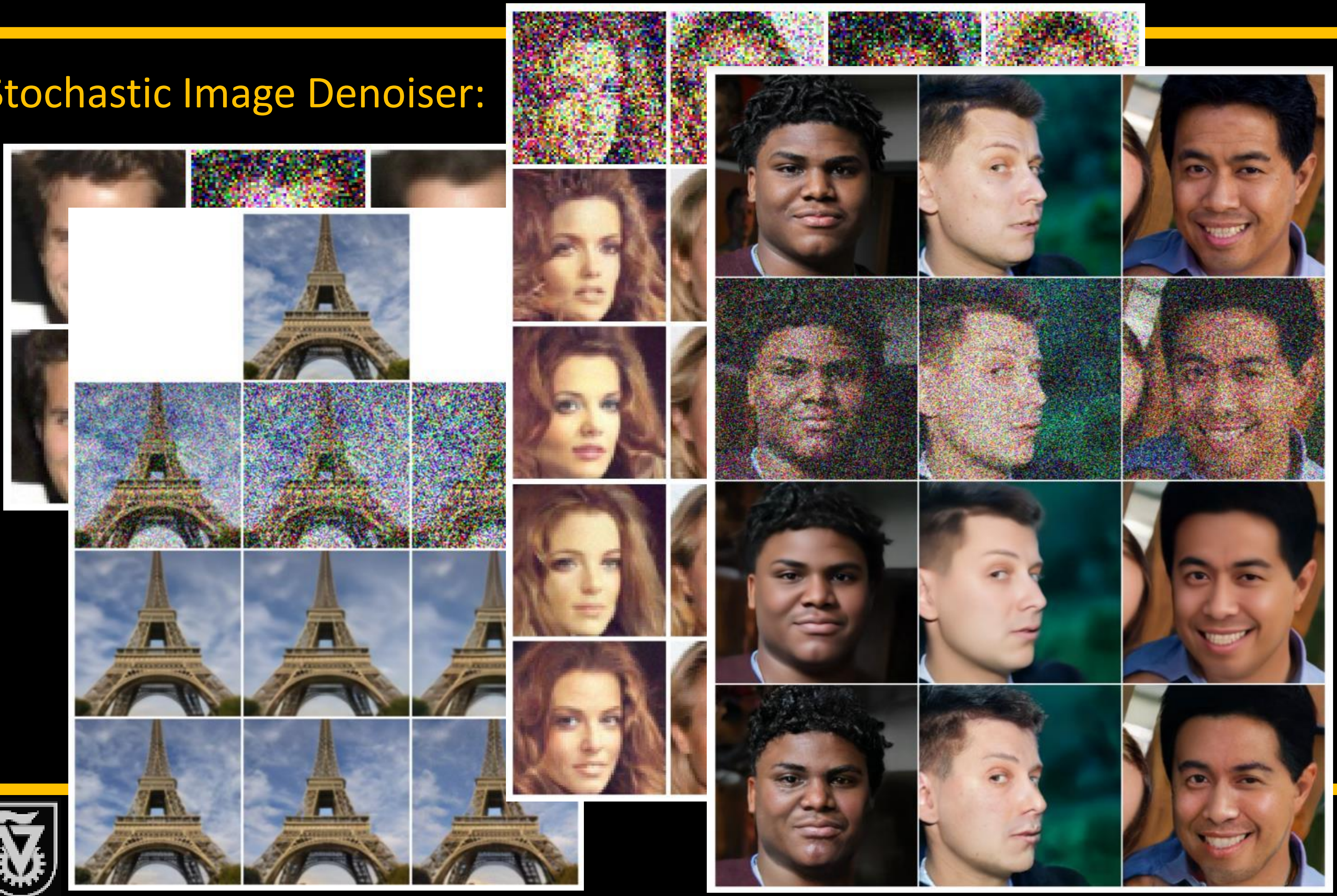
Stochastic Image Denoiser:

- We start from a noisy image ($\sigma \approx 150$ in this example)
- Then gradually denoise it using (conditional) annealed Langevin dynamics

Michael Elad
The Computer-Science Department
The Technion

Discovery 2: Targeting Perceptual Quality

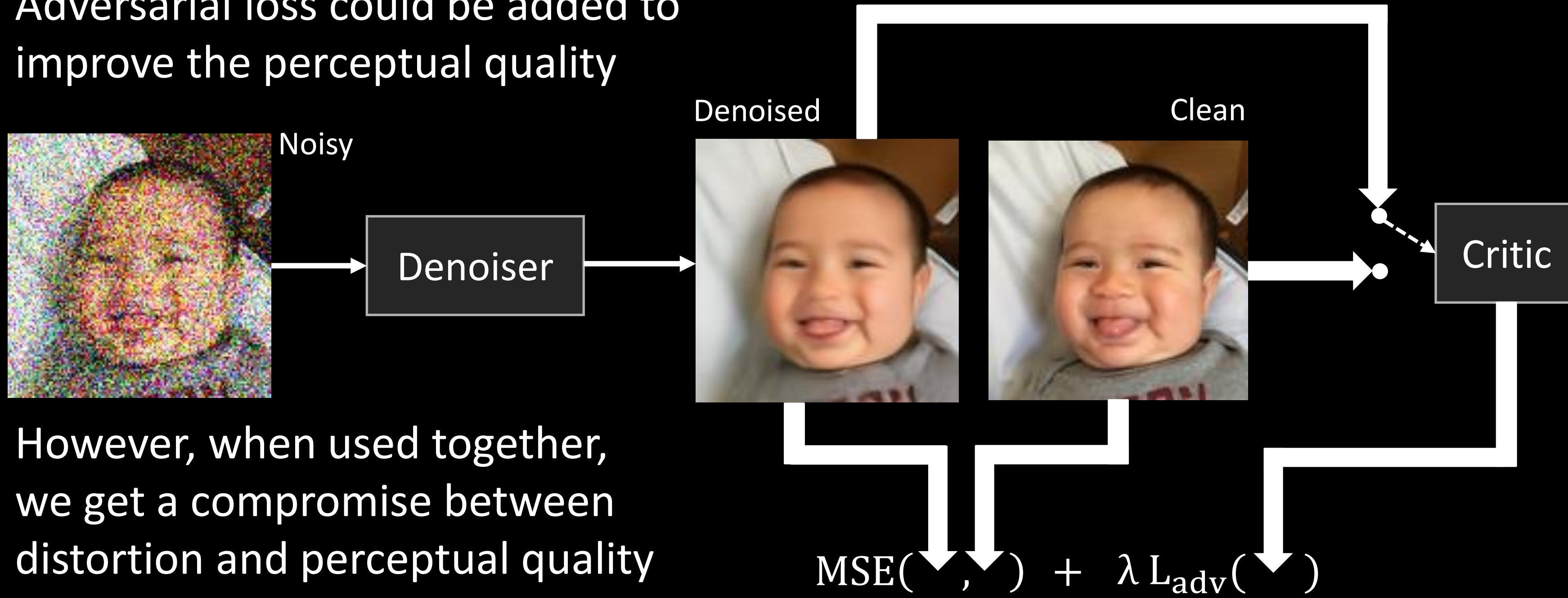
Stochastic Image Denoiser:



Discovery 2: Targeting Perceptual Quality

Let's have a closer look at the Conditional GAN Denoiser:

- Typical design approach: Optimize a distortion measure (e.g. MSE) between the denoised and the ideal images
- Adversarial loss could be added to improve the perceptual quality

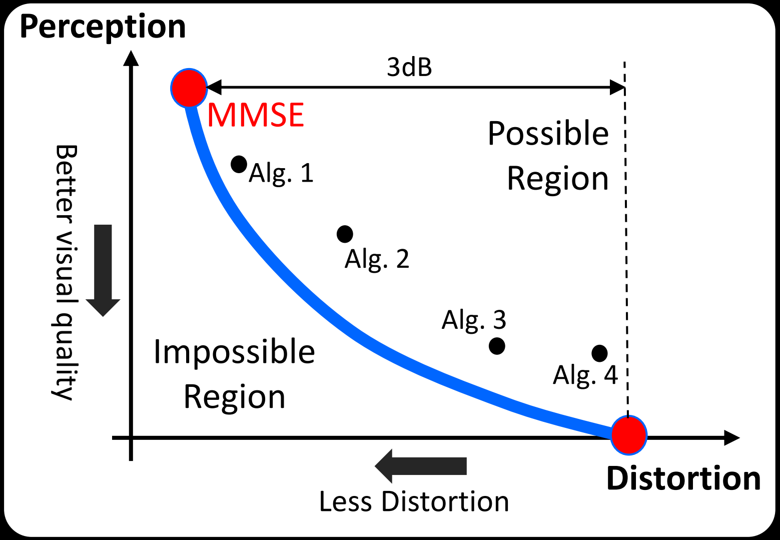


- However, when used together, we get a compromise between distortion and perceptual quality

Michael Elad
The Computer-Science Department
The Technion

Discovery 2: Targeting Perceptual Quality

- For ill-posed restoration tasks, perceptual quality performance comes at the expense of its distortion [Blau & Michaeli 2017]
- We aim for best perceptual quality
- The posterior distribution attains perfect perceptual quality, compromising 3dB compared to the MMSE [Blau & Michaeli 2017]
- We propose to sample from the posterior via a Conditional GAN mechanism (PSCGAN)

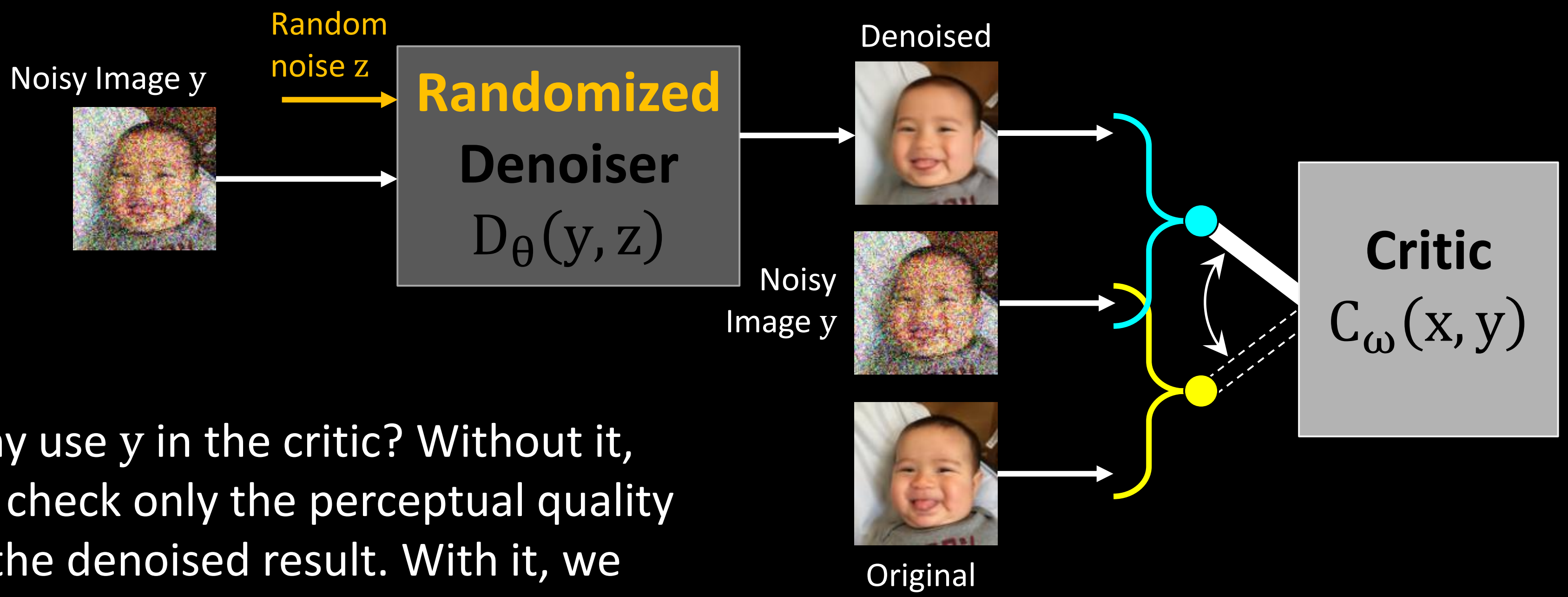


$x \sim P_X$ $y \sim P_{Y|X=x}$ Samples from $P_{X|Y=y}$

Michael Elad
The Computer-Science Department
The Technion

Discovery 2: Targeting Perceptual Quality

The PSCGAN Architecture:



Why use y in the critic? Without it, we check only the perceptual quality of the denoised result. With it, we also assess its denoising validity

Discovery 2: Targeting Perceptual Quality

What about the Loss?

- CGAN optimization leads to posterior sampling [Adler et al. 2018]:

$$\min_{\theta} \max_{\omega} \mathbb{E}_{X,Y}\left[C_{\omega}(x,y)\right] - \mathbb{E}_{D_{\theta},Y,Z}\left[C_{\omega}(D_{\theta},y)\right]$$

- However, this requires an unavailable balanced dataset to train on (many x's for each y and many y's for each x)
- On the other hand, we would like to avoid a penalty of the form

$$\mathbb{E}_{X,Y,Z}\left[\|x - D_{\theta}(y,z)\|_2^2\right]$$

- Our remedy: adding an MMSE oriented penalty term:

$$\mathbb{E}_{X,Y}\left[\|x - \mathbb{E}_{z}[D_{\theta}|y]\|_2^2\right]$$

- This gives the MMSE result "for free" (averaging many instances)

Discovery 2: Targeting Perceptual Quality

CGAN:

Oh … and One Last Thing

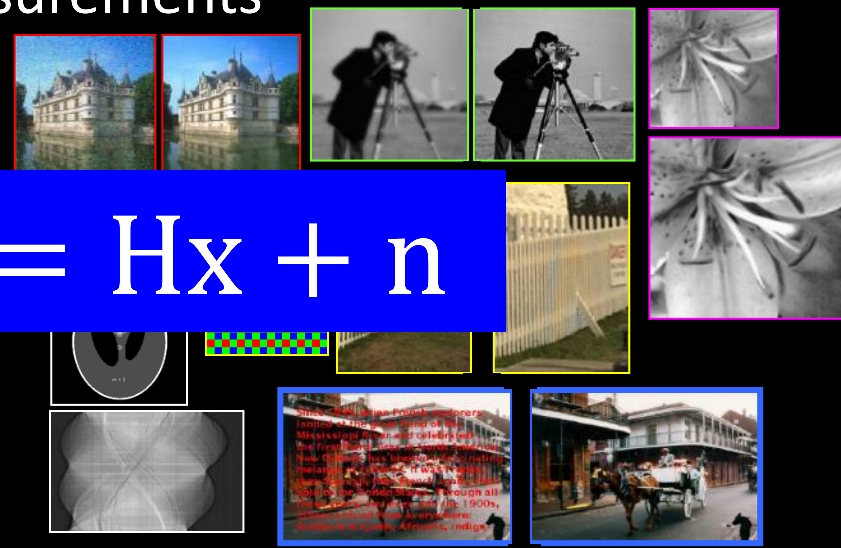
What about Inverse Problems?

- Goal: Recovery from corrupted measurements

  De-Noising · De-Blurring

  In-Painting · De-Mosaicing

  Tomography · Image Scale-Up & super-resolution



$$y = Hx + n$$

- Can we suggest a "sampler" from $P(x|y)$ for handling all these problems, in order to obtain "perfect looking" results?

- Answer: Yes! Use Langevin dynamics again, in an adapted form

Snips: Solving noisy inverse problems stochastically — 4 — 2021
B Kawar, G Vaksman, M Elad
Advances in Neural Information Processing Systems 34

Denoising Diffusion Restoration Models — 2022
B Kawar, M Elad, S Ermon, J Song
Advances in Neural Information Processing Systems (NeurIPS)

What about Inverse Problems?

- The idea is similar to our high-perceptual denoising, with necessary changes for considering the degradation operator H …
- Starting naively, using Bayes theorem, we need to calculate

$$\nabla \log \mathrm{P}(\mathrm{x_i}|\mathrm{y}) = \nabla \log \mathrm{P}(\mathrm{x_i}) + \nabla \log \mathrm{P}(\mathrm{y}|\mathrm{x_i})$$

- We know that $\mathrm{y} = \mathrm{Hx} + \mathrm{n}$ and $\mathrm{x_i} = \mathrm{x} + \mathrm{v_i}$ and thus:

$$\nabla \log \mathrm{P}(\mathrm{y}|\mathrm{x_i}) = \nabla \log \mathrm{P}(\mathrm{y} - \mathrm{Hx_i}|\mathrm{x_i}) =$$

$$\nabla \log \mathrm{P}(\mathrm{Hx} + \mathrm{n} - \mathrm{Hx} - \mathrm{Hv_i}|\mathrm{x_i}) = \nabla \log \mathrm{P}(\mathrm{n} - \mathrm{Hv_i}|\mathrm{x_i})$$

- However, … while $\mathrm{n} - \mathrm{Hv_i}$ is a simple Gaussian, it's dependency on $\mathrm{x_i}$ in non-trivial, so how do we proceed from here?

Michael Elad
The Computer-Science Department
The Technion

What about Inverse Problems?

- Step 1: Use SVD for decoupling the measurements $\mathrm{H} = \mathrm{U\Sigma V^T}$:

$$\mathrm{U^T y = U^T \underbrace{\left[U\Sigma V^T (x_i - v_i) + n\right]}_{y\,=\,Hx+n} = \Sigma V^T (x_i - v_i) + U^T n}$$

$$\boxed{\mathrm{y_T[k] = \sigma_k \tilde{x}_T[k] - \sigma_k \tilde{v}_T[k] + n_T[k]}}$$

Decouple $\mathrm{\tilde{x}_T[k]} \leftrightarrow \mathrm{\tilde{v}_T[k]}$ by choosing $\mathrm{\tilde{v}_T[k]}$ to be a portion of $\mathrm{n_T[k]}$

- Thus, we can apply the Langevin dynamics algorithm on $\mathrm{\tilde{x}_T = V^T x_i}$ given $\mathrm{y_T = U^T y}$ and sample from the conditional
- Bottom line: An MMSE denoiser is used for a novel solution of inverse problems, this time targeting best perceptual quality

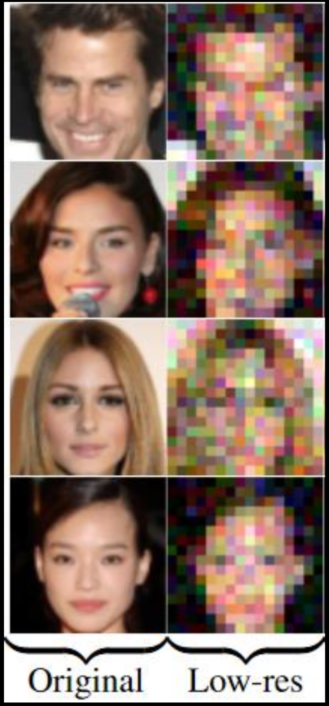
Michael Elad
The Computer-Science Department
The Technion

What about Inverse Problems?

Noisy Inpainting: A portion missing and noise with $\sigma_0 \approx 25$

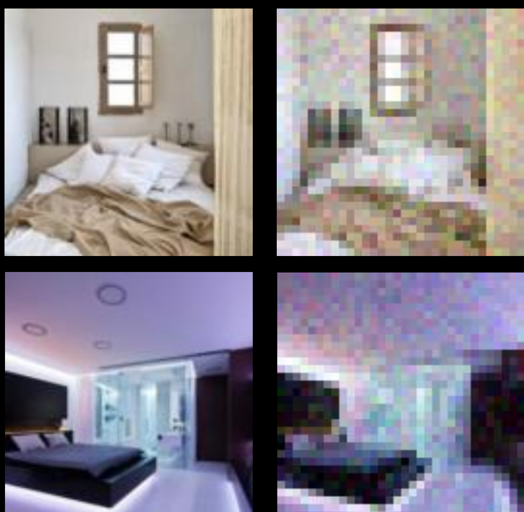
What about Inverse Problems?

Super resolution: downscaling by 4 with additive noise of $\sigma_0 \approx 25$



What about Inverse Problems?

Super resolution: downscaling by 4 with additive noise of $\sigma_0 \approx 12$



Original Degraded

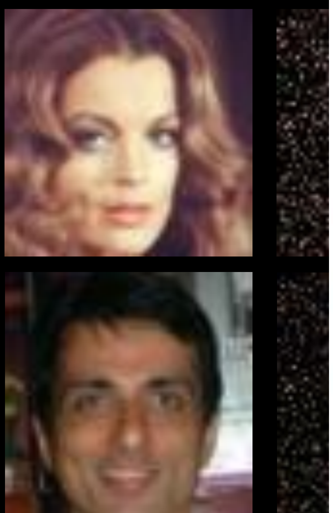
Michael Elad
The Computer-Science Department
The Technion

What about Inverse Problems?

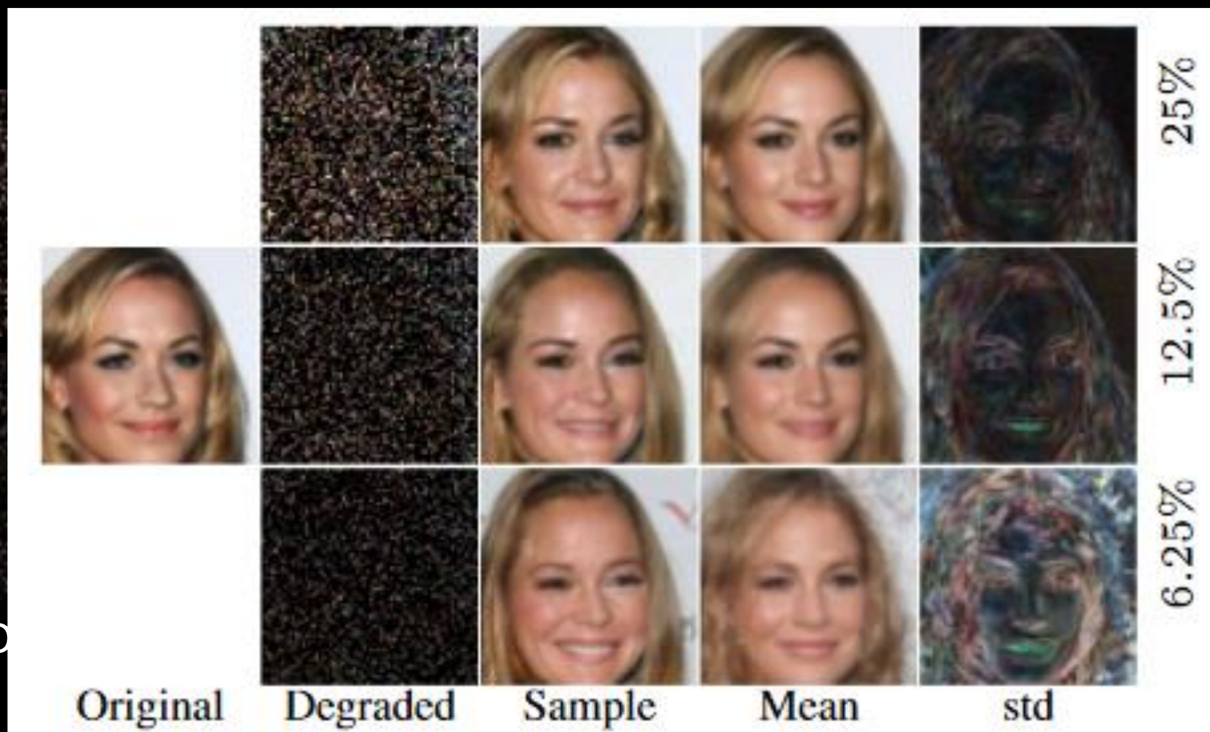
Deblurring: uniform 5×5 blur with additive noise of $\sigma_0 \approx 25$

What about Inverse Problems?

Compressive sensing (12.5%) with additive noise of $\sigma_0 \approx 25$



Original D



Back to Inverse Problems

And just to remind you …

The proposed diffusion-based sampling scheme, while quite appealing, suffers from several key shortcomings:

- ❑ It is rather S L O W (many denoising activations)
- ❑ It is limited (as of now) to specific families of images
- ❑ Relying on SVD is cumbersome

See our recent work that answers (some of) these challenges:

Denoising Diffusion Restoration Models B Kawar, M Elad, S Ermon, J Song Advances in Neural Information Processing Systems (NeurIPS)	53	2022

Time to Summarize

What Have we Seen Today?

Suppose that we are given an MMSE denoiser $\mathrm{D(y)}$

WE CAN USE $\mathrm{D(y)}$ FOR ....

- solving ANY inverse problem (PnP/RED)
- synthesizing natural-looking images
- denoising images while targeting high perceptual quality
- solving ANY inverse problem with high perceptual quality

All the above are achieved by simply applying $\mathrm{D(y)}$ iteratively

Summary

Image Denoising

… Not What You Think

1. There are so many opportunities and challenges in better understanding, designing, and proposing creative usage of image denoisers
2. Despite the fact that this has not been a talk about Deep-Learning, the presence of this field in the topics covered is prominent