

# Image Denoising

... Not What You Think

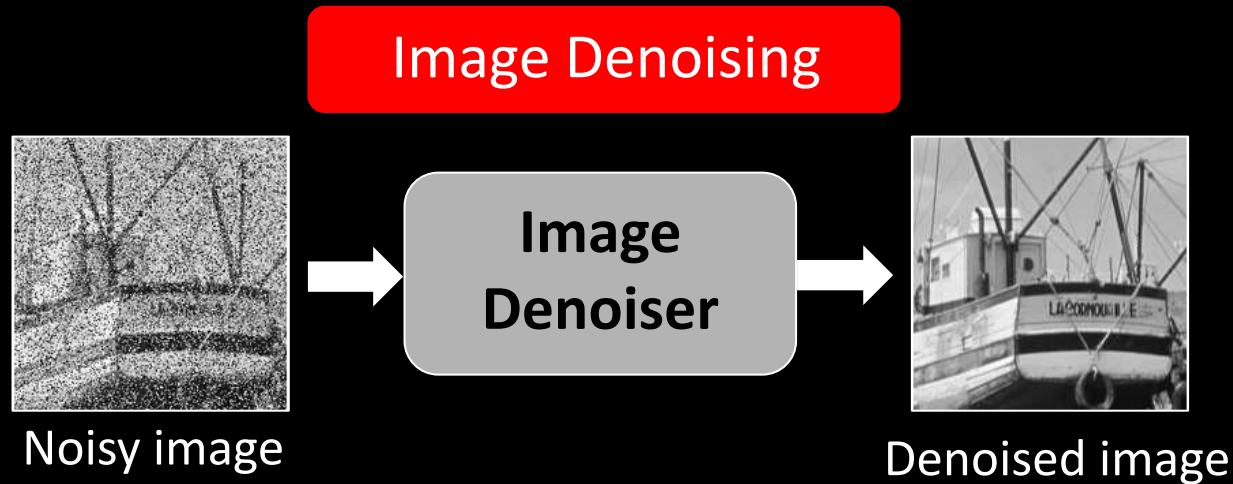
Michael Elad

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

September 10<sup>th</sup> 2021



# This Lecture is About ...



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



# Our Agenda

1. Brief Introduction & History
2. Image Denoising: The Classic Era
3. The Deep Learning Revolution
4. Synergy: Classic + Deep Learning
5. Our Focus Today: Denoising for ...
  - Solving general inverse problems
  - Image Synthesis
  - High perceptual quality recovery
6. Summary



# Introduction & History

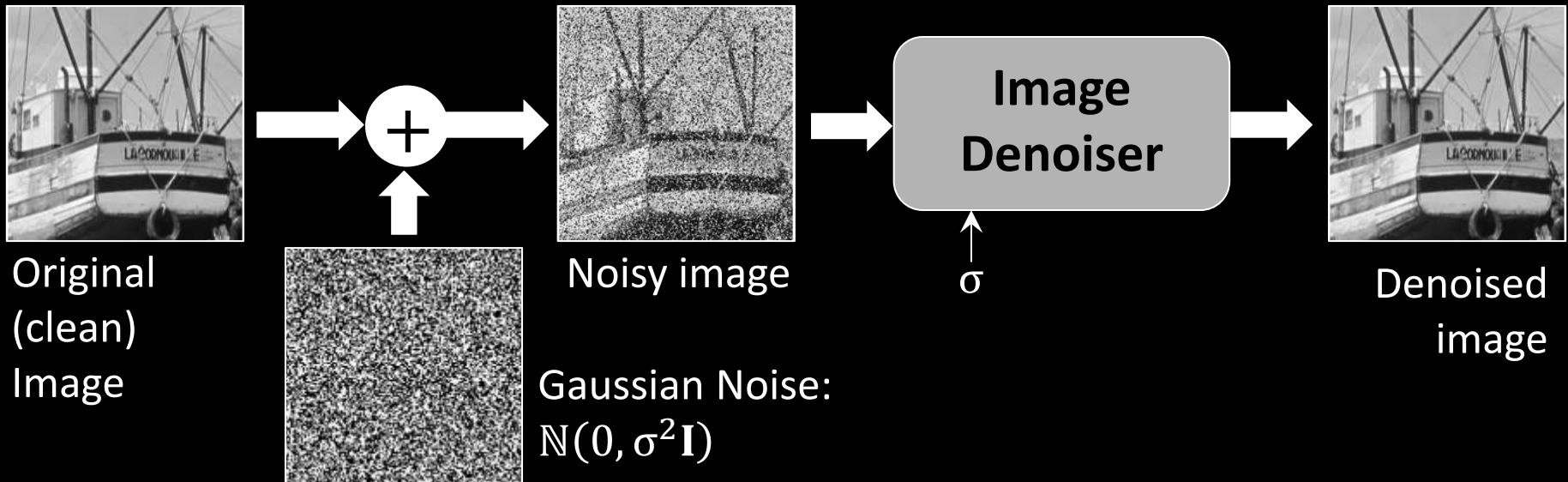


# So, Let's Talk About ...

## Image Denoising

or more accurately

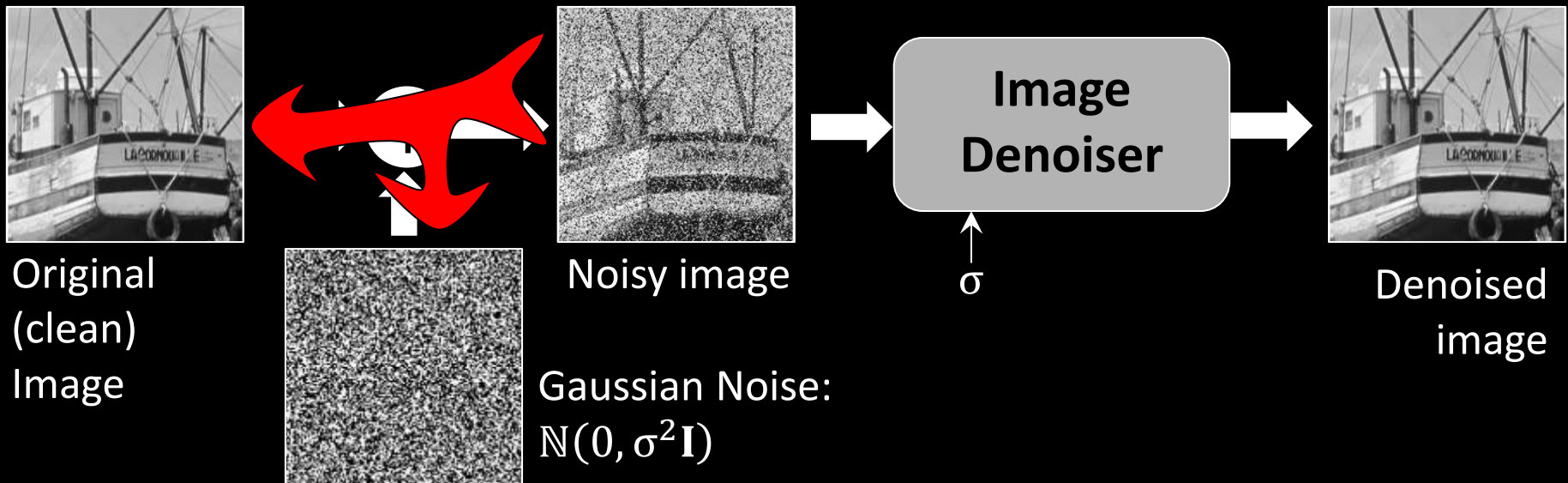
Removal of **White Additive Gaussian Noise** from an Image



# Image Denoising is Challenging

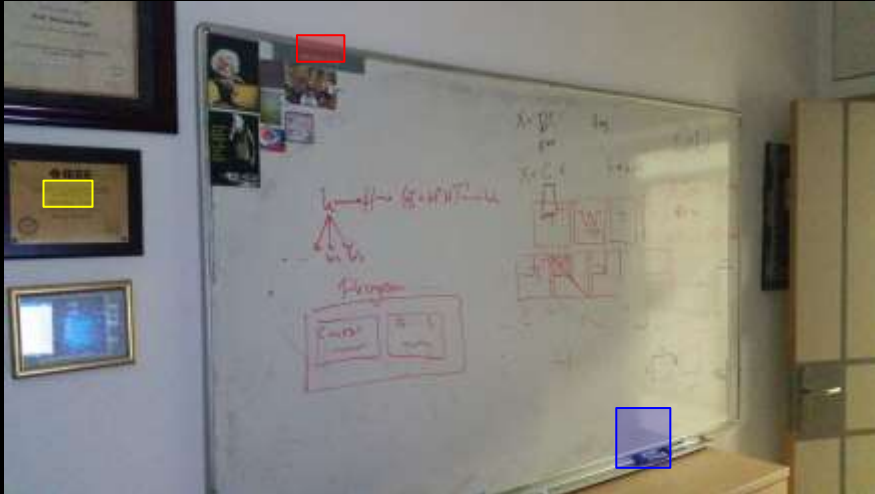
Image denoising is far from trivial task! **Why?**

- ❑ Because our goal is to remove noise as much as possible while **preserving** the details in the image
- ❑ Denoising is essentially a highly ill-posed **separation** task



# Why Work on Image Denoising?

1. **Practical:** It is a real-world problem, arising in all cameras,



Noisy image

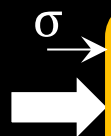


Image  
Denoiser

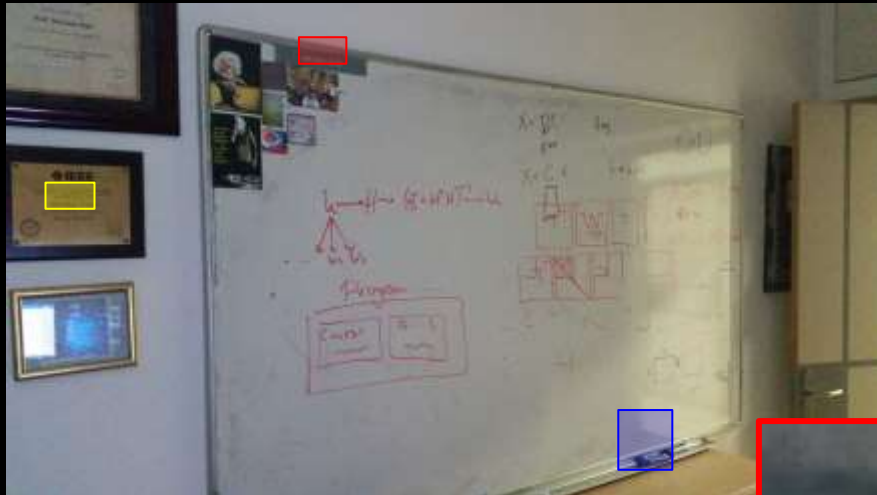


Denoised image

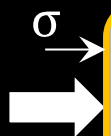


# Why Work on Image Denoising?

1. **Practical:** It is a real-world problem, arising in all cameras,

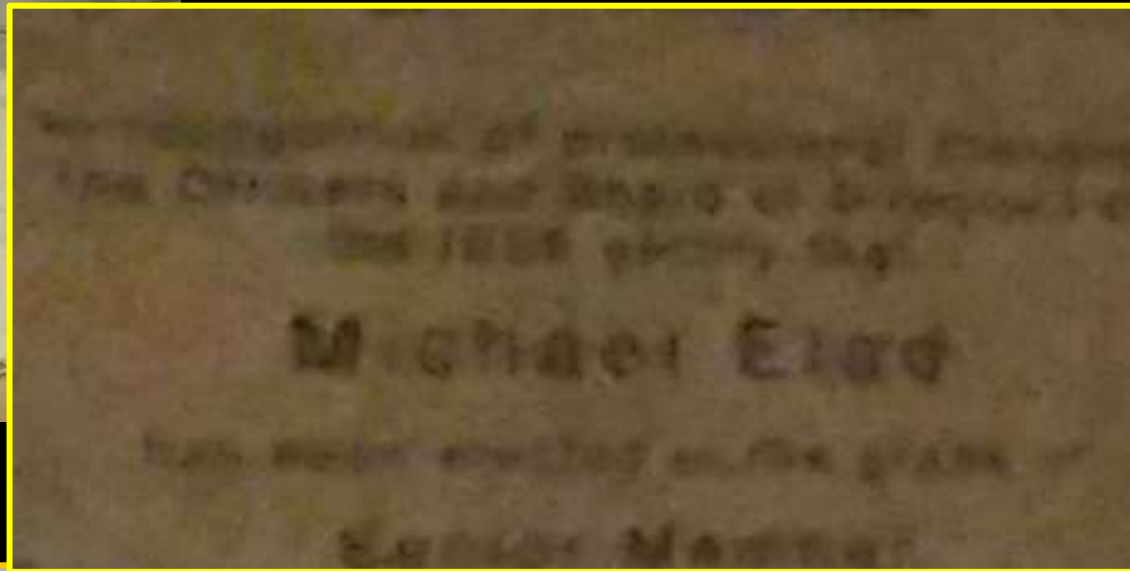
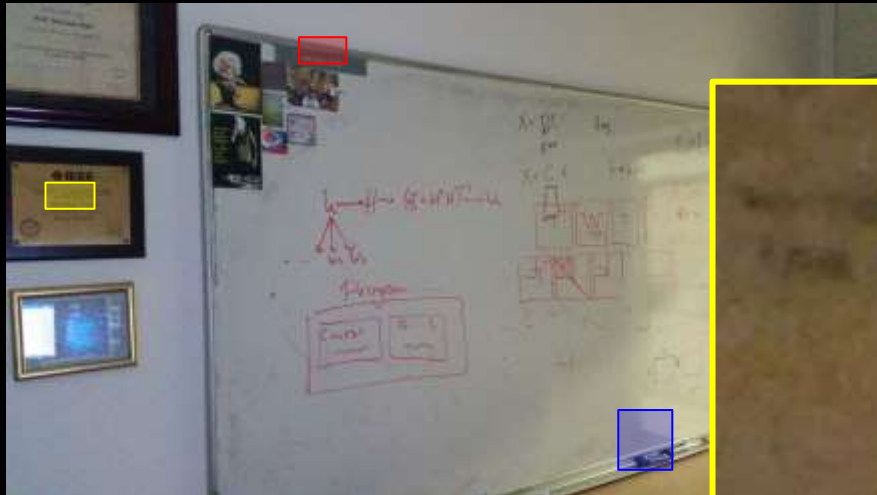


Noisy image

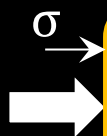


# Why Work on Image Denoising?

1. **Practical:** It is a real-world problem, arising in all cameras,

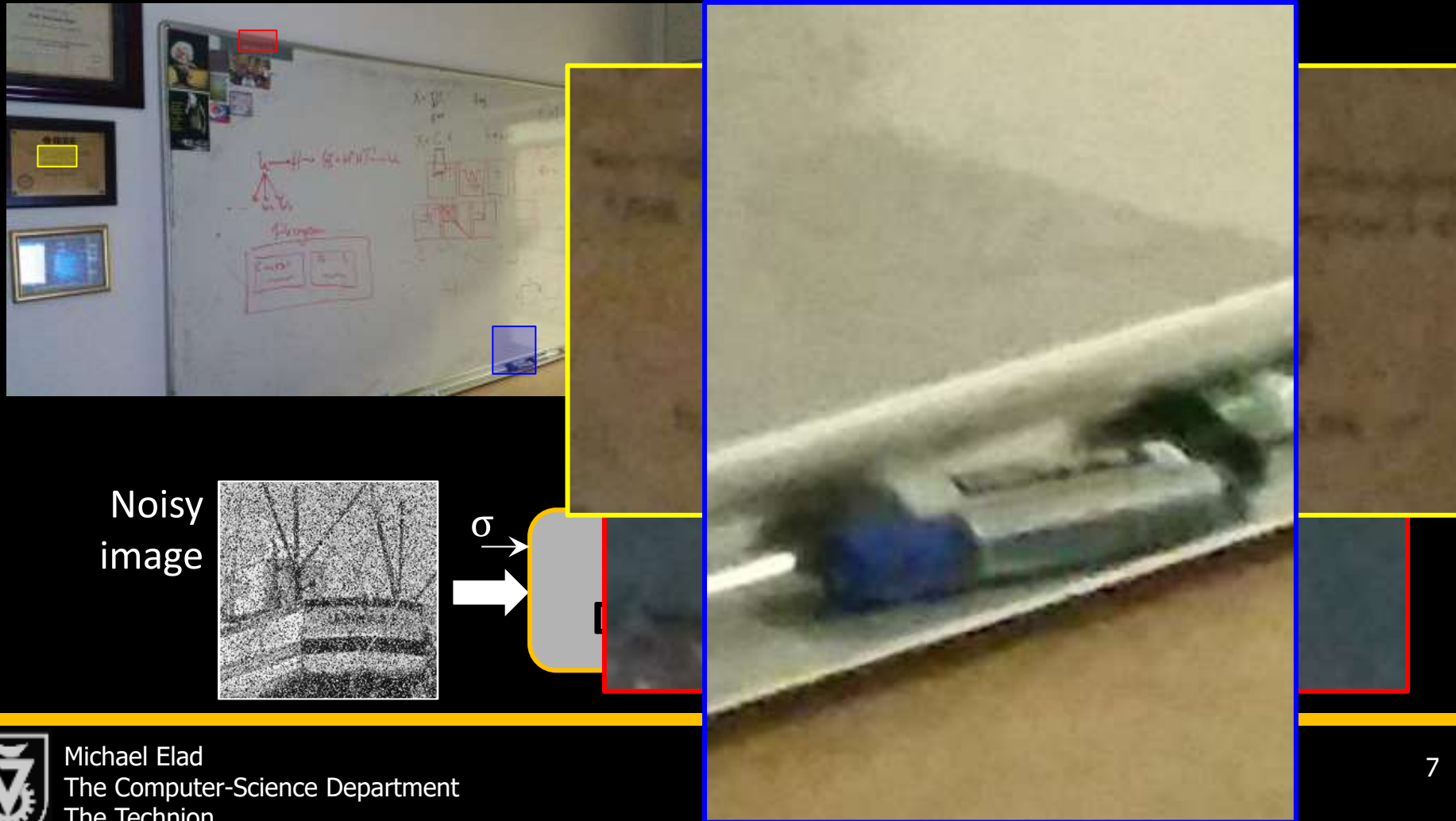


Noisy  
image



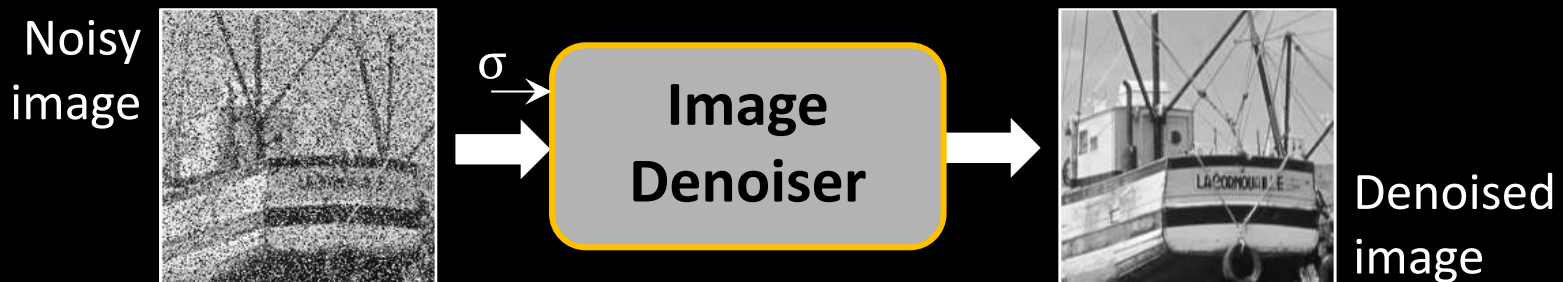
# Why Work on Image Denoising?

1. **Practical:** It is a real-world problem, arising in all cameras,



# Why Work on Image Denoising?

1. **Practical:** It is a real-world problem, arising in all cameras,
2. **Front-Gate to Image Processing:** Being the simplest inverse problem, it is a platform for assessing new ideas in our field, &
3. **Other Uses for the Denoiser Engine:** Recent work has shown that given a denoiser, there are other fascinating uses for it that go far beyond noise removal

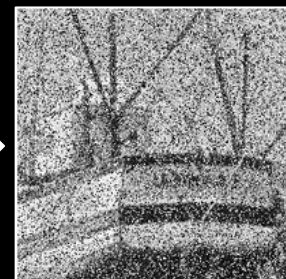
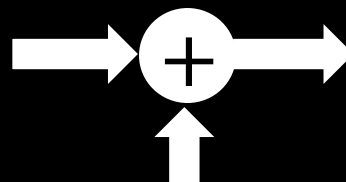


# Why Assume Gaussian Noise?

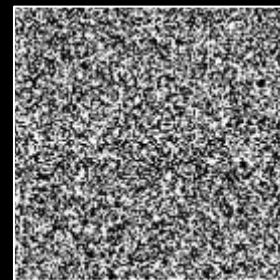
- ❑ The Gaussian case is more common and much more important
- ❑ When considering a Poisson noise,
  - High count of photons – The distribution gets closer and closer to the Gaussian case
  - Low-count Poisson-distributed image can be converted to a Gaussian-noisy one by **Anscomb** - Variance Stabilizing 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)



Original  
(clean)  
Image



Noisy image

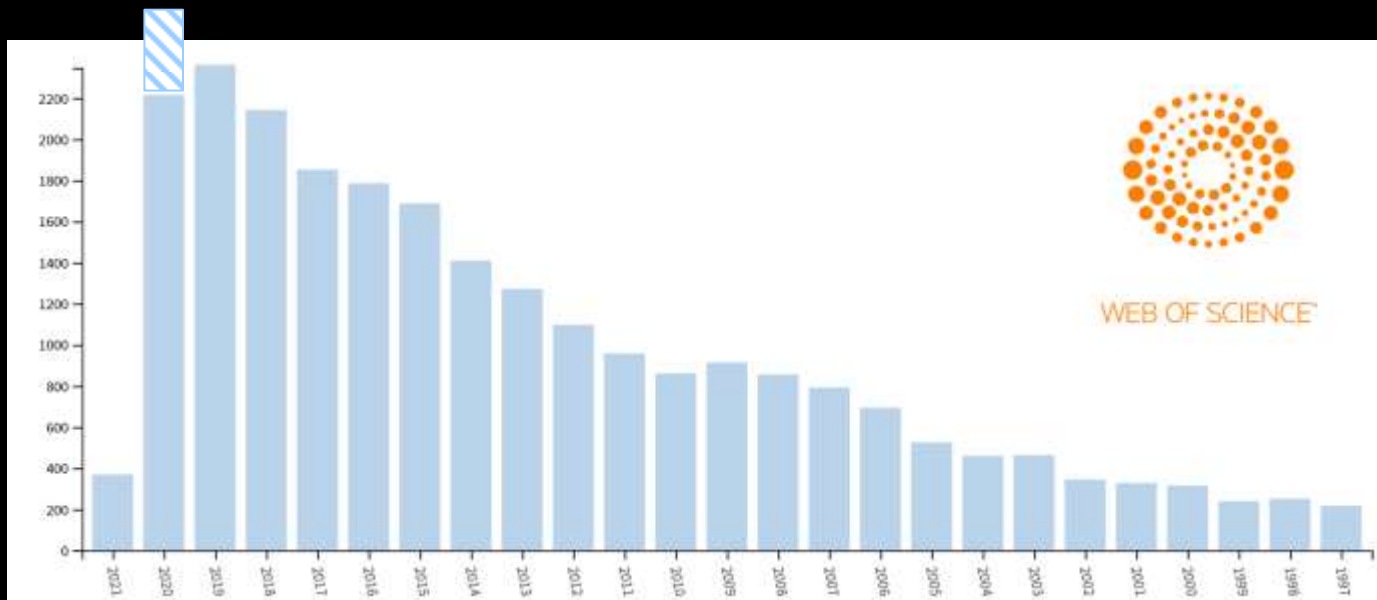


Gaussian  
Noise:  
 $\mathcal{N}(0, \sigma^2 \mathbf{I})$

# Image Denoising: Little bit of History

Roughly speaking, there are ~25,000 papers\* on this subject, offering algorithms, theoretical analysis and so much more

My speculation



\* Search done on June 1<sup>st</sup> in WoS, topic: ((image or video ) and (denoising or (noise and remov) or clean))



# Image Denoising: Little bit of History

## Citing Articles:

USA:	140524
China:	45284
Germany:	29272
France:	35585
England:	24090
Canada:	18325
Spain:	17880
Israel:	13988
Australia:	13358
Switz.:	12504
Japan:	12389
Italy:	11754
Netherland:	10455
India:	8830
Finland:	7842
Korea:	7558
Belgium:	5027
Singapore:	4964
Brazil:	4849
Taiwan:	4134
Iran:	3112
Russia:	2595

This research comes from all over the globe



... and it is heavily cited



# The Classic Era




# Design of Image Denoising Algorithms

## How can we design a denoiser?

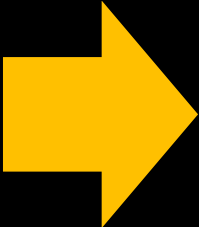
The classic Bayesian approach (1960-2014):

- Model image content with a **prior** expression (e.g., forcing smoothness, sparsity, low-rank, self-similarity, ... ), and
- Formulate the denoising task as an optimization problem


$$\hat{x} = \min_x \underbrace{\|x - y\|^2}_{\text{Likelihood}} + \underbrace{\rho(x)}_{\text{Prior}}$$

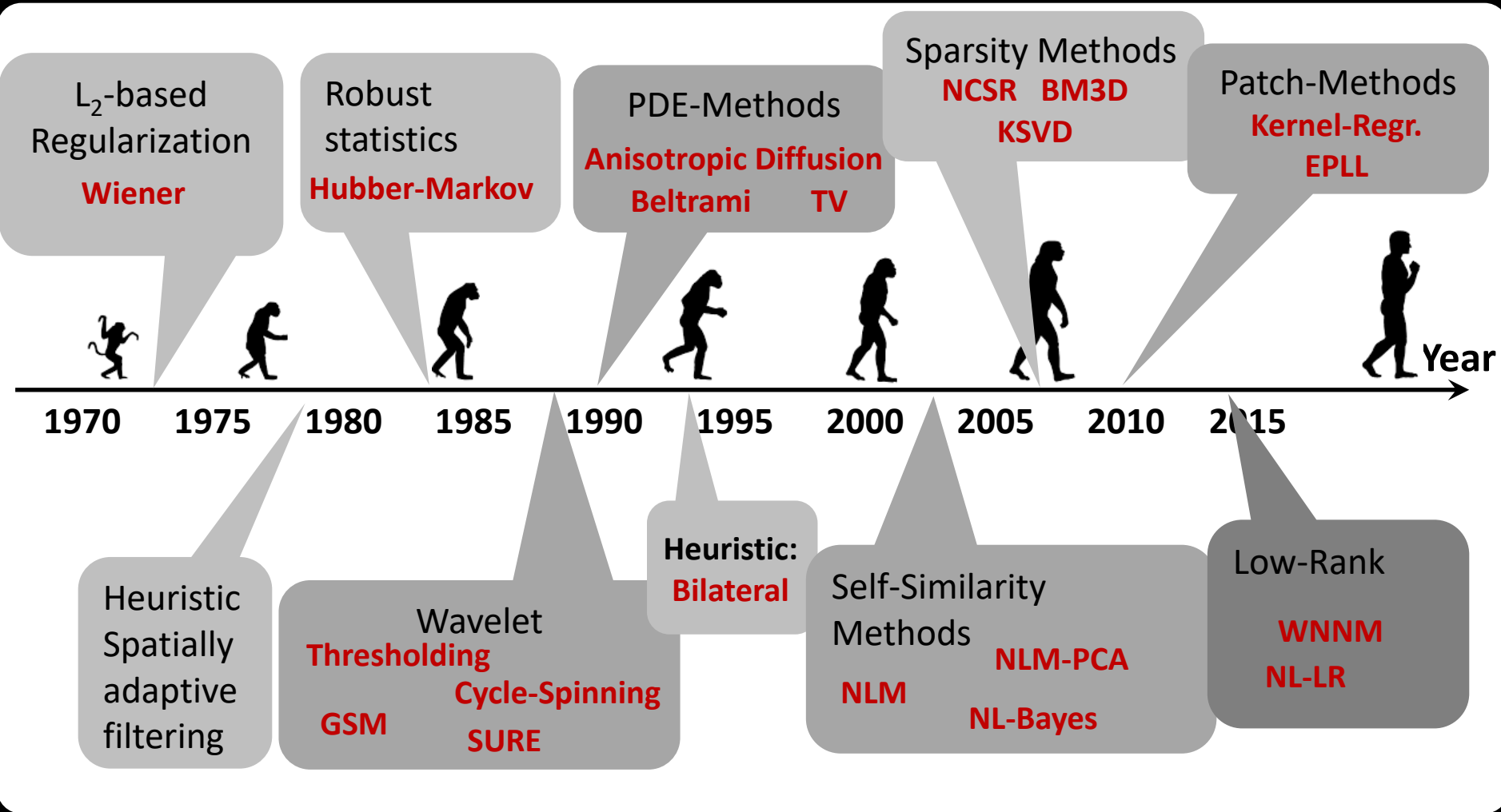
$y$  : Given noisy image

$\hat{x}$  : Denoised result



Design an iterative  
or a direct algorithm  
for getting  $\hat{x}$  from  $y$

# Image Denoising: Evolution



# End of an Era?

This evolution of algorithms and the tendency of different methods to perform very similarly has led to a feeling that “Denoising is Dead”

IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 19, NO. 4, APRIL 2010

895

## Is Denoising Dead?

Priyam Chatterjee, *Student Member, IEEE*, and Peyman Milanfar, *Fellow, IEEE*

**Abstract**—Image denoising has been a well studied problem in the field of image processing. Yet researchers continue to focus attention on it to better the current state-of-the-art. Recently proposed methods take different approaches to the problem and yet their denoising performances are comparable. A pertinent question then to ask is whether there is a theoretical limit to denoising performance and, more importantly, are we there yet? As camera manufacturers continue to pack increasing numbers of pixels per unit area, an increase in noise sensitivity manifests itself in the form of a noisier image. We study the performance bounds for the image denoising problem. Our work in this paper estimates a lower bound on the mean squared error of the denoised result and compares the performance of current state-of-the-art denoising methods with this bound. We show that despite the phenomenal recent progress in the quality of denoising algorithms, some room for improvement still remains for a wide class of general images, and at certain signal-to-noise levels. Therefore, image denoising is not dead—yet.

**Index Terms**—Bayesian Cramér–Rao lower bound (CRLB), bias, bootstrapping, image denoising, mean squared error.

### I. INTRODUCTION

**I**MAGE denoising has been a well-studied problem in the image processing community and continues to attract

erature on such performance limits exists for some of the more complex image processing problems such as image registration [7], [8] and super-resolution [9]–[12]. Performance limits to object or feature recovery in images in the presence of point-wise degradation has been studied by Treibitz *et al.* [13]. In their work, the authors study the effects of noise among other degradations and formulate expressions for the optimal filtering parameters that define the resolution limits to recovering any given feature in the image. While their study is practical, it does not define statistical performance limits to denoising general images. In [14], Voloshynovskiy *et al.* briefly analyze the performance of MAP estimators for the denoising problem. However, our bounds are developed in a much more general setting and, to the best of our knowledge, no comparable study currently exists for the problem of denoising. The present study will enable us to understand how well the state-of-the-art denoising algorithms perform as compared to these limits. From a practical perspective, it will also lead to understanding the fundamental limits of increasing the number of sensors in the imaging system with acceptable image quality being made possible by noise suppression algorithms.

Before we analyze image denoising statistically, we first de-



# End of an Era?

This evolution of algorithms and the tendency of different methods to perform very similarly has led to a feeling that “Denoising is Dead”

IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 19, NO. 4, APRIL 2010

895

## Is Denoising Dead?

Priyam Chatterjee, Student Member

**Abstract**—Image denoising has been a well studied field in the field of image processing. Yet researchers continue to put attention on it to better the current state-of-the-art. Reported methods take different approaches to the problem and their denoising performances are comparable. A pertinent question then to ask is whether there is a theoretical limit to denoising performance and, more importantly, are we there yet? As manufacturers continue to pack increasing numbers of pixels per unit area, an increase in noise sensitivity manifests itself as a noisier image. We study the performance bounds for image denoising problem. Our work in this paper estimates a lower bound on the mean squared error of the denoised result and compares the performance of current state-of-the-art denoising methods to this bound. We show that despite the phenomenal recent improvements in the quality of denoising algorithms, some room for improvement still remains for a wide class of general images, at least at signal-to-noise levels. Therefore, image denoising is not dead.

**Index Terms**—Bayesian Cramér–Rao lower bound, bootstrapping, image denoising, mean squared error

### I. INTRODUCTION

IMAGE denoising has been a well-studied problem in the image processing community and continues to be so.

## Natural Image Denoising: Optimality and Inherent Bounds

Anat Levin and Boaz Nadler

Department of Computer Science and Applied Math  
The Weizmann Institute of Science

### Abstract

*The goal of natural image denoising is to estimate a clean version of a given noisy image, utilizing prior knowledge on the statistics of natural images. The problem has been studied intensively with considerable progress made in recent years. However, it seems that image denoising algorithms are starting to converge and recent algorithms improve over previous ones by only fractional dB values. It is thus important to understand how much more can we still improve natural image denoising algorithms and what are the inherent limits imposed by the actual statistics of the data. The challenge in evaluating such limits is that constructing proper models of natural image statistics is a long standing and yet unsolved problem.*

*To overcome the absence of accurate image priors, this paper takes a non parametric approach and represents the distribution of natural images using a huge set of  $10^{10}$  patches. We then derive a simple statistical measure which provides a lower bound on the optimal Bayesian minimum mean square error (MMSE). This imposes a limit on the best possible results of denoising algorithms which utilize a*

ever, it seems that the performance of denoising algorithms is starting to converge. Recent techniques typically improve over previous ones by only fractional dB values. In some cases the difference between the results of competing algorithms is so small and inconclusive, that one actually has to successively toggle between images on a monitor to visually compare their denoising quality. This raises the question of whether the error rates of current denoising algorithms can be reduced much further, or whether there are inherent limitations imposed by the statistical structure of natural images? The goal of this paper is to derive a lower bound on the best possible denoising error under a well defined statistical framework. Such a bound can help us understand if there is hope to significantly improve the current state-of-the-art image denoising with even better algorithms, or whether we have nearly approached the fundamental limits.

Understanding the limits of natural image denoising is also important as an instance of a more fundamental computer and human vision challenge: modeling the statistics of natural images and understanding the inherent limits of their statistical power. Several works attempted to estimate the entropy of natural images [15, 4]. However, there is

CVPR 2011



Michael Elad  
The Computer-Science Dep  
The Technion

# End of an Era?

And so, somewhere around 2010-2012, the general feeling in our community was that ...

We are currently touching the ceiling in denoising performance and chances of improving them are very slim



There is no point in devising new denoising methods



Work in this field has diminishing returns

**Well, We Were Wrong !**



# End of an Era?

## Wrong ? How?

The past decade has taught us that image denoising is still

**very much alive and kicking**

due to several branches of novel activity on:

- Obtaining better performing denoisers with **deep learning**
- New frontiers in denoising:
  - Better **adaptation** to image content
  - Denoising strategies that go **beyond PSNR**
  - Identifying **alternative** methods for designing/training denoisers
  - Extending the denoising task to **realistic noise**, and
- Discovering new ways for **leveraging** denoisers for other needs



# The Deep Learning Revolution



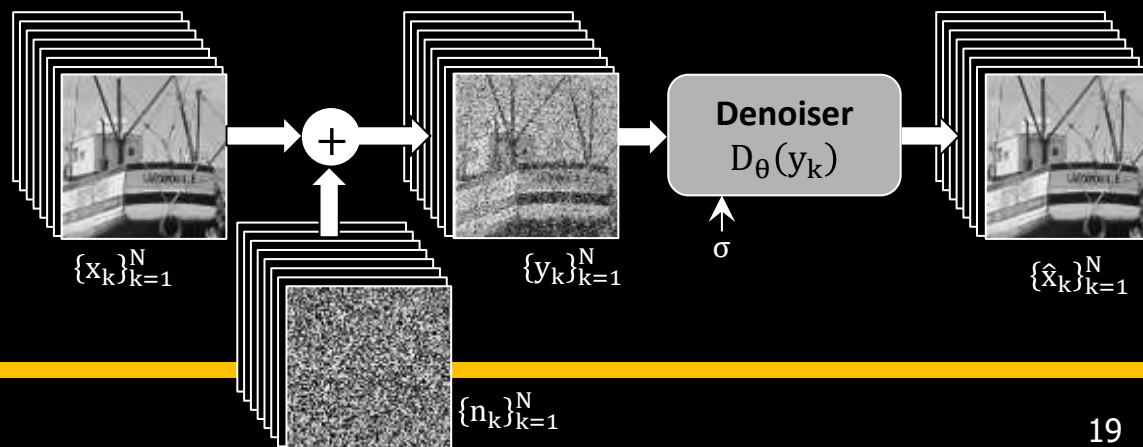
# Design of Algorithms: Take 2

How can we **ALTERNATIVELY** design a denoiser?

The machine learning approach (2012-Now):

- Gather a LARGE dataset of clean images  $\{x_k\}_{k=1}^N$
- Add AWGN these images:  $\{y_k = x_k + n_k\}_{k=1}^N$
- Define a parametric denoising machine  $D_\theta(y)$
- Train  $D_\theta(\blacksquare)$  by setting its parameters  $\theta$ :

$$\min_{\theta} \sum_{k=1}^N \|x_k - \underbrace{D_\theta(y_k)}_{\hat{x}_k}\|^2$$



# 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

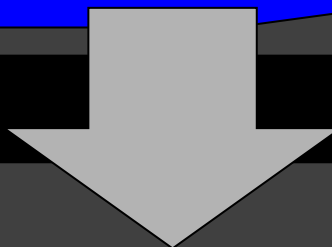
Low Rank

Piecewise Smoothness

Non-Local Self-Similarity

GMM

Low dimensionality



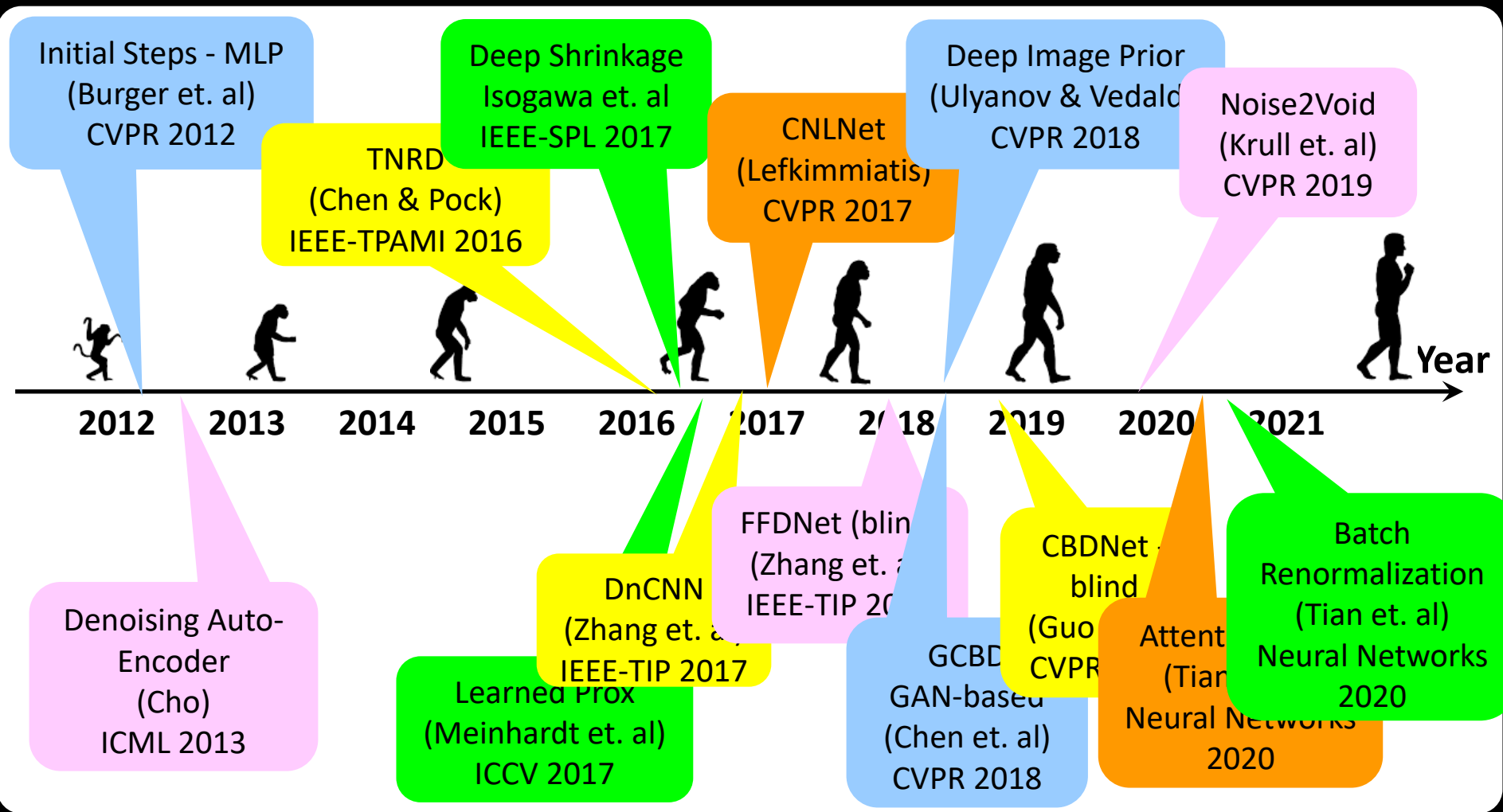
Learning

**Supervised Training**

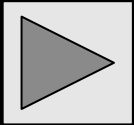
Observe that with this trend, all the knowledge and knowhow accumulated carefully over decades in image processing became **TOTALLY OBSOLETE**



# Image Denoising: Recent Evolution

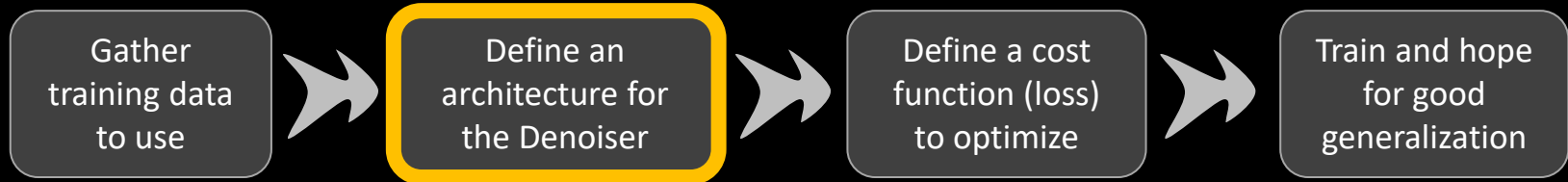
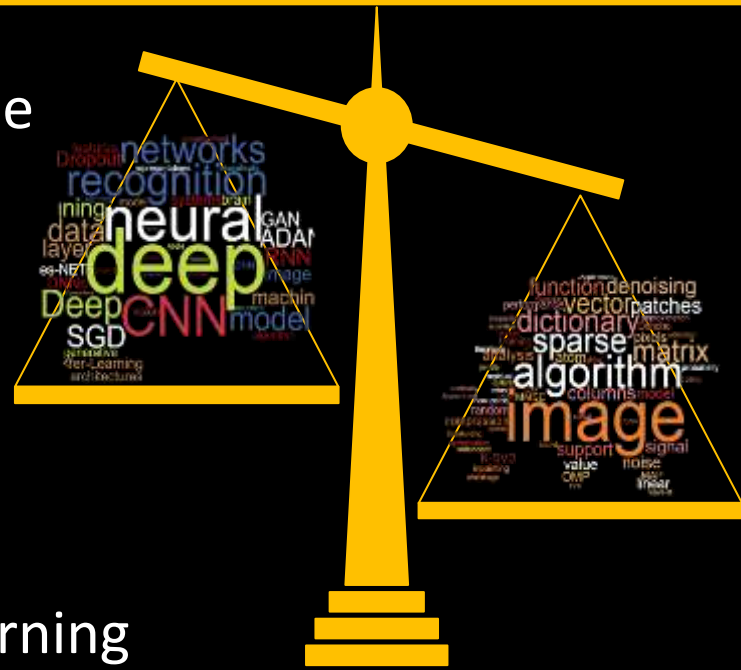


# Synergy: Classics + Deep Learning



# Image Denoising: Return of the Classics

- ❑ In recent years deep learning is ruling the image denoising domain, pushing aside all the classical methods, along with their great achievements
- ❑ Recently, however, we do see a synergy between the two paradigms
- ❑ Recall: In building a **supervised** deep learning denoiser solution, we operate along the following lines:



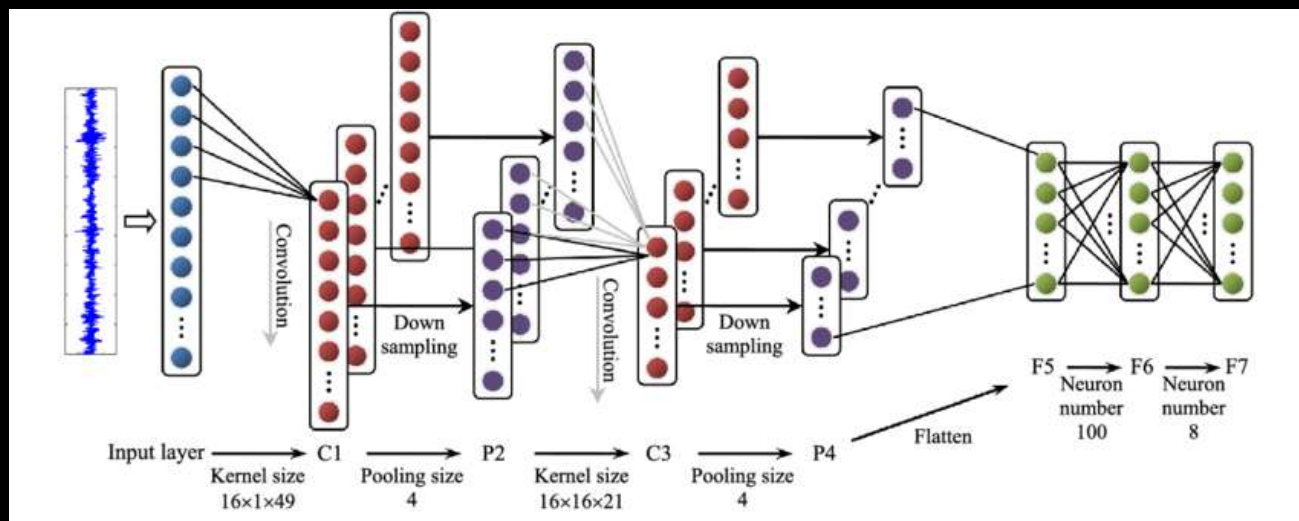
# Image Denoising Architectures

So, how do we choose an architecture for a given task?

**Option 1** - Copy an existing network that has shown good results in earlier work (VGG, U-Net, ...), and slightly modify it

**Option 2** – Pile and Guess a series of steps that mix known pieces such as convolutions, fully connected layer, batch-norm, ReLU, pooling, stride, skips, upscale/downscale, connections, ... and add new “tricks”

**Option 3** – Neural Architecture Search



# Image Denoising Architectures

Here are several paper examples from CVPR/NIPS 2019 that illustrate these architectures

Meta-SR: A Magni

Noise2Void - Learning Denoising from Single Noisy Images

Alexander Krull<sup>1,2</sup>, Tim-Oliver Buchholz<sup>2</sup>, Florian Jug

<sup>1</sup> krull@mpi-cbg.de

<sup>2</sup> Authors contributed equally

MPI-CBG/PKS (CSBD), Dresden, Germany

CVPR 2019: U-Net-based with 3.8e6 params

CVPR 2019: DnCNN-based with 5.5e5 params

Samuli Laine  
NVIDIA\*

CVPR 2019: U-Net-based with 5.3e6 params

NIPS 2019: U-Net-based with 1.1e6 params

\*The Chinese University of Hong Kong

CVPR 2019: DnCNN-based with 1.2e6 params



# Alternative Architecture Design

- ❑ Message: Do far better in choosing architectures by relying on unfolding algorithms from the classics of image processing
- ❑ The benefits in such architectures:
  - They are far **more concise** yet just as effective as leading methods
  - They are **easier to train** because they are lighter
  - They have the potential to break current performance **barriers**
  - They may bring better understanding and **explainability**
  - They enable better **adaptation** to out of distribution images
- ❑ Here are few representative examples:
  - Rethinking the CSC Model [[Simon & Elad, NIPS '19](#)]
  - Non-Local & Multi-Scale Denoising [[Vaksman, Milanfar & Elad, CVPR \(NTIRE\) '20](#)]
  - Deep KSVD Denoising [[Scetbon, Milanfar & Elad, IEEE-TIP '21](#)]
  - PatchCraft: Non-Local Video Denoising [[Vaksman, Elad & Milanfar, Arxiv '21](#)]



# Our Focus Today: Recent Discoveries

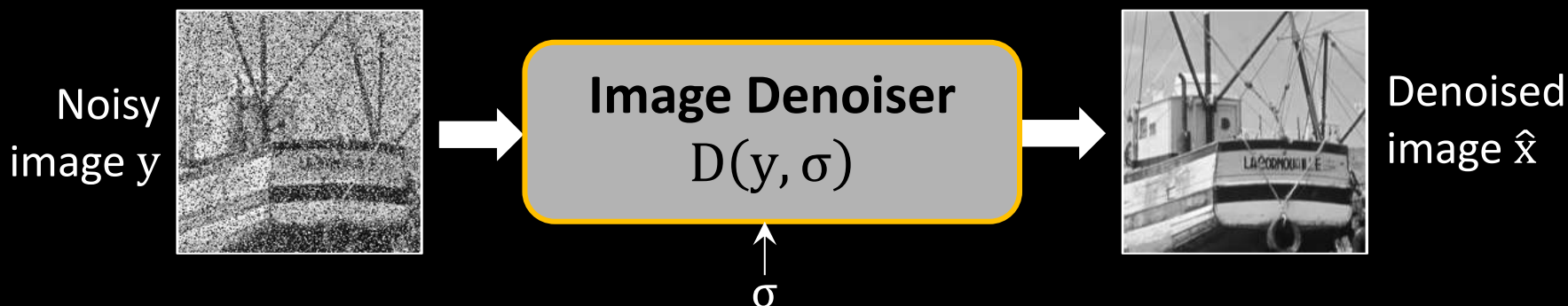


# Our Focus Today

Recent findings on using denoisers for other tasks:

- ❑ **Discovery 1:** Solving general inverse problems [2013-]
- ❑ **Discovery 2:** Image Synthesis [2019-]
- ❑ **Discovery 3:** High perceptual quality recovery [2021-]

We turn to describe these results



# Discovery 1: Solving Inverse Problems

## Inverse Problems: Recovery of images from corrupted measurements

- ❑ De-Noising
- ❑ De-Blurring
- ❑ In-Painting
- ❑ De-Mosaicing
- ❑ Tomography
- ❑ Image Scale-Up & super-resolution

... and more ...




# Discovery 1: Solving Inverse Problems

## How can we solve inverse problems?


We can return to the classic Bayesian approach:

- Model image content with a **prior** expression (e.g., forcing smoothness, sparsity, low-rank, self-similarity, ... ), and
- Formulate the inversion task as an optimization problem


$$\hat{x} = \min_x \underbrace{\|Hx - y\|^2}_{\text{Likelihood}} + \underbrace{\rho(x)}_{\text{Prior}}$$

$y$  : Given measurements

$\hat{x}$  : Denoised result

- 
- This is known as MAP estimation
  - It is an extension of the classic path for denoising, tailoring methods for inverse problems
  - This approach leads to iterative algorithm for getting  $\hat{x}$  from  $y$
  - Is there a supervised learning alternative? **Definitely!**



# Discovery 1: Solving Inverse Problems

**Question:** Given a denoiser  $D(y, \sigma)$   
how can one solve inverse problems with it?

Plug-and-play priors for model based reconstruction

382 2013

SV Venkatakrisnan, CA Bouman, B Wohlberg

2013 IEEE Global Conference on Signal and Information Processing, 945-948

The little engine that could: Regularization by denoising (RED)

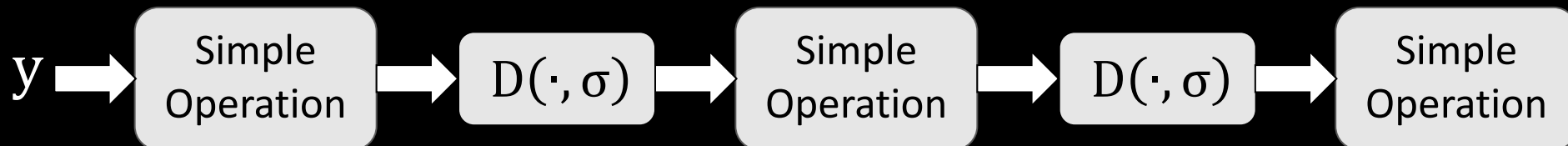
261 2017

Y Romano, M Elad, P Milanfar

SIAM Journal on Imaging Sciences 10 (4), 1804-1844

**Answer:** Use  $D(y, \sigma)$  as a *regularizer*

**Practical Implication:** Iterated use of  $D(\cdot, \sigma)$



# Discovery 1: Solving Inverse Problems

Here is (roughly) the PnP Perspective in a nutshell:

- Recall: Inverse problems can be formulated as optimization tasks:

$$\hat{x} = \min_x \frac{1}{2} \|Hx - y\|^2 + \rho(x)$$

- Let's do something "stupid" and split the unknown:

$$\hat{x} = \min_{x,v} \frac{1}{2} \|Hx - y\|^2 + \rho(v) \quad \text{s.t. } x = v$$

- Now, turn the constraint into a penalty\*

$$\hat{x} = \min_{x,v} \frac{1}{2} \|Hx - y\|^2 + \rho(v) + \beta \|x - v\|^2$$

- And solve by alternating between  $x$  and  $v$

- Least-Squares:  $\hat{x} = \min_x \frac{1}{2} \|Hx - y\|^2 + \beta \|x - v\|^2$

- A denoiser:  $\hat{v} = \min_v \rho(v) + \beta \|x - v\|^2$

... and this way we  
got an iterated  
algorithm that  
keeps calling to a  
denoiser,  
for solving the  
inverse problem

\* The PnP uses the Augmented  
Lagrange which is more accurate and  
less sensitive to the choice of  $\beta$



# Discovery 1: Solving Inverse Problems

Here is the RED Perspective in a nutshell:

Let's start again with the formulated optimization task,  
and suggest a very specific regularization term:

$$\hat{x} = \min_x \frac{1}{2} \|Hx - y\|^2 + \rho(x) = \min_x \frac{1}{2} \|Hx - y\|^2 + \underbrace{\lambda x^T [x - D(x, \sigma)]}_{\substack{\uparrow \\ x^T(I - S)x}}$$

Let's use the  
Steepest Descent

Under mild conditions\* the  
gradient of this is  $[x - D(x, \sigma)]$

$$\hat{x}_{k+1} = \hat{x}_k - \mu [H^T(H\hat{x}_k - y) + \lambda[\hat{x}_k - D(\hat{x}_k, \sigma)]]$$

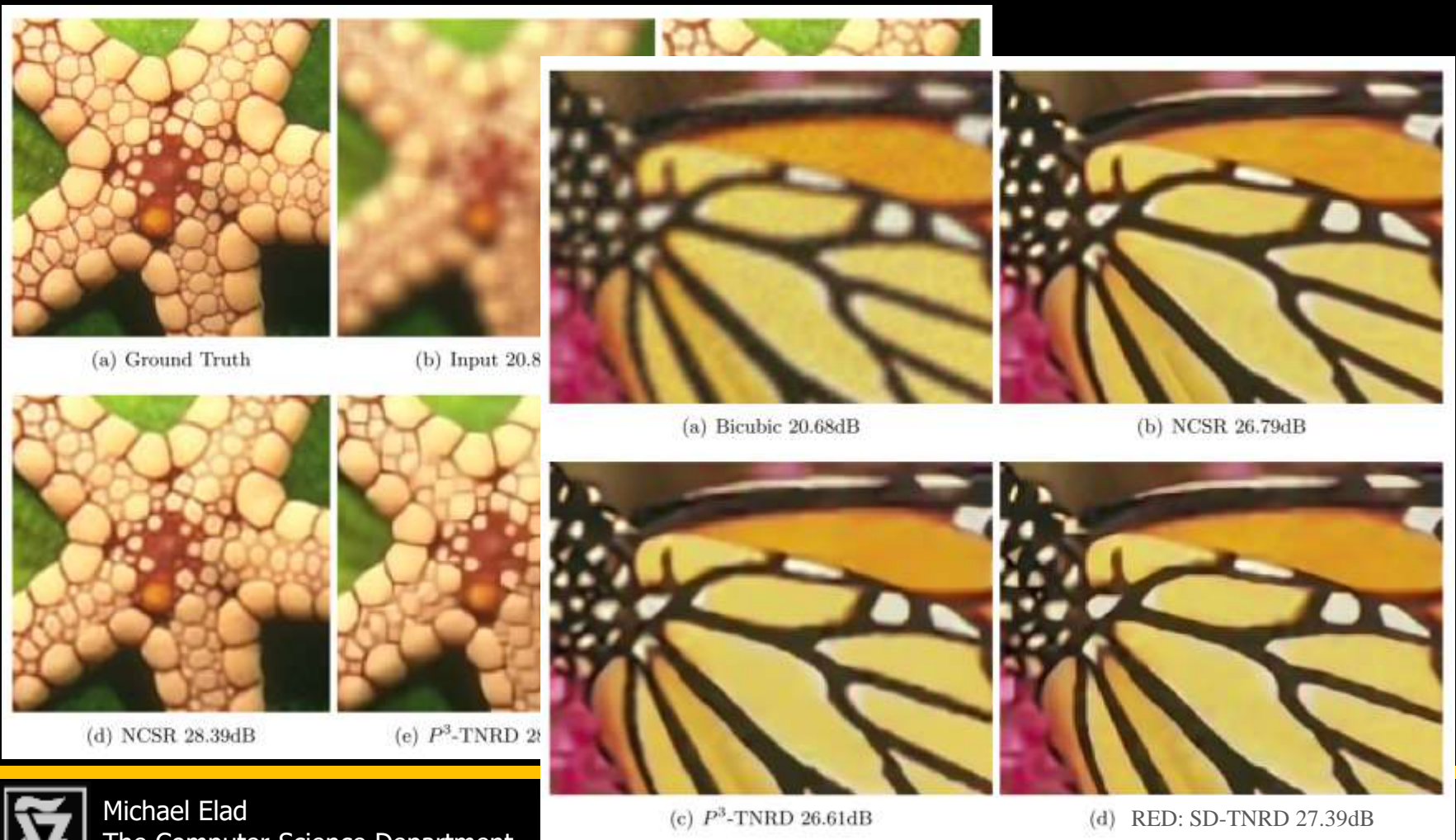
... and this way we got an iterated algorithm that keeps calling to a denoiser,  
and is guaranteed to achieve the minimum

\* Differentiability, local homogeneity, passivity and symmetric Jacobian (MMSE)



# Discovery 1: Solving Inverse Problems

Here are some results for Deblurring and Super-Resolution



# Discovery 1: Solving Inverse Problems

❑ PnP and RED are heavily cited and extensively studied, owing to their generality and elegance

Plug-and-play priors for model based reconstruction

382 2013

SV Venkatakrishnan, CA Bouman, B Wohlberg

2013 IEEE Global Conference on Signal and Information Processing, 945-948

The little engine that could: Regularization by denoising (RED)

261 2017

Y Romano, M Elad, P Milanfar

SIAM Journal on Imaging Sciences 10 (4), 1804-1844

❑ Follow-up work focuses on

- Proving convergence to the desired solution and tying these to properties of the permissible denoisers (e.g. MMSE ...)
- Deployment in various applications
- Creation of new variants of these two methods ... and ...

❑ PnP/RED can be used to define **well-motivated architectures** for solving general inverse problems, built around a core learned denoising engine



# Discovery 2: Image Synthesis

- ❑ In recent years, and with the deep-learning revolution, there is a growing interest in 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?



# Discovery 2: Image Synthesis

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

Generative modeling by estimating gradients of the data distribution

104

2019

Y Song, S Ermon

arXiv preprint arXiv:1907.05600

Improved techniques for training score-based generative models

20

2020

Y Song, S Ermon

arXiv preprint arXiv:2006.09011

Solving linear inverse problems using the prior implicit in a denoiser

4

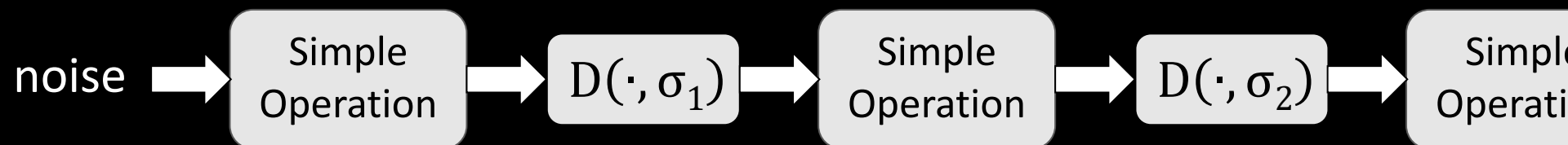
2020

Z Kadkhodaei, EP Simoncelli

arXiv preprint arXiv:2007.13640

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

**Practical Implication:** Iterated use of  $D(\cdot, \sigma)$  with varying  $\sigma$



# Discovery 2: Image Synthesis

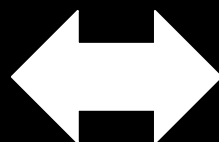
Here is the core idea in a nutshell:

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

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

$$\hat{x}_{k+1} = \hat{x}_k + a \cdot \underbrace{\nabla \log P(\hat{x}_k)}_{\text{Score Function}} + b \cdot z_k \quad (\text{Langevin Dynamics})$$

This is known as the **Score Function** and it is approximately proportional to  $[\hat{x}_k - D(\hat{x}_k, \sigma)]$  for a small value of  $\sigma$



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

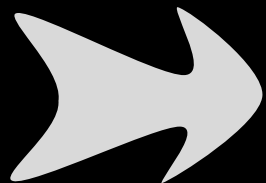
... and this way we got an iterated algorithm that keeps calling to a denoiser, and is guaranteed to obtain a sample from  $P(x)$



# Discovery 2: Image Synthesis

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

$$\begin{aligned} P(\mathbf{x} + \mathbf{v}) \quad \text{for } \mathbf{v} \sim \mathcal{N}(0, \sigma_k^2 \mathbf{I}) \\ = P(\mathbf{x}) \otimes c \cdot \exp \left\{ -\frac{1}{2\sigma^2} \|\mathbf{x}\|^2 \right\} \\ \text{with } \sigma_0 > \sigma_1 > \sigma_2 \cdots > \sigma_N > 0 \end{aligned}$$



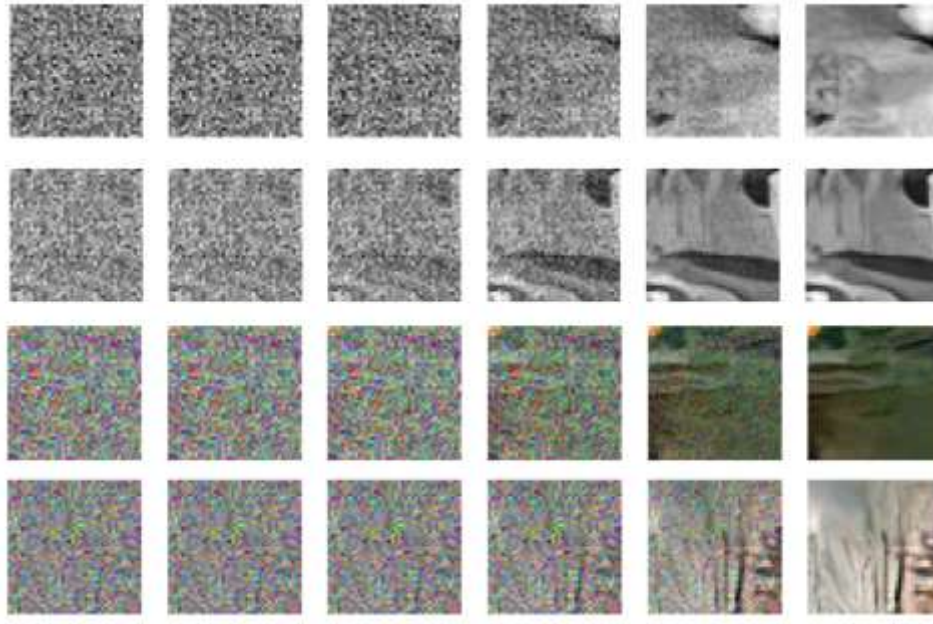
The core idea: start by drawing from a wider distribution and gradually narrow it, leading to a faster sampling performance

Blurred Image  
Manifold

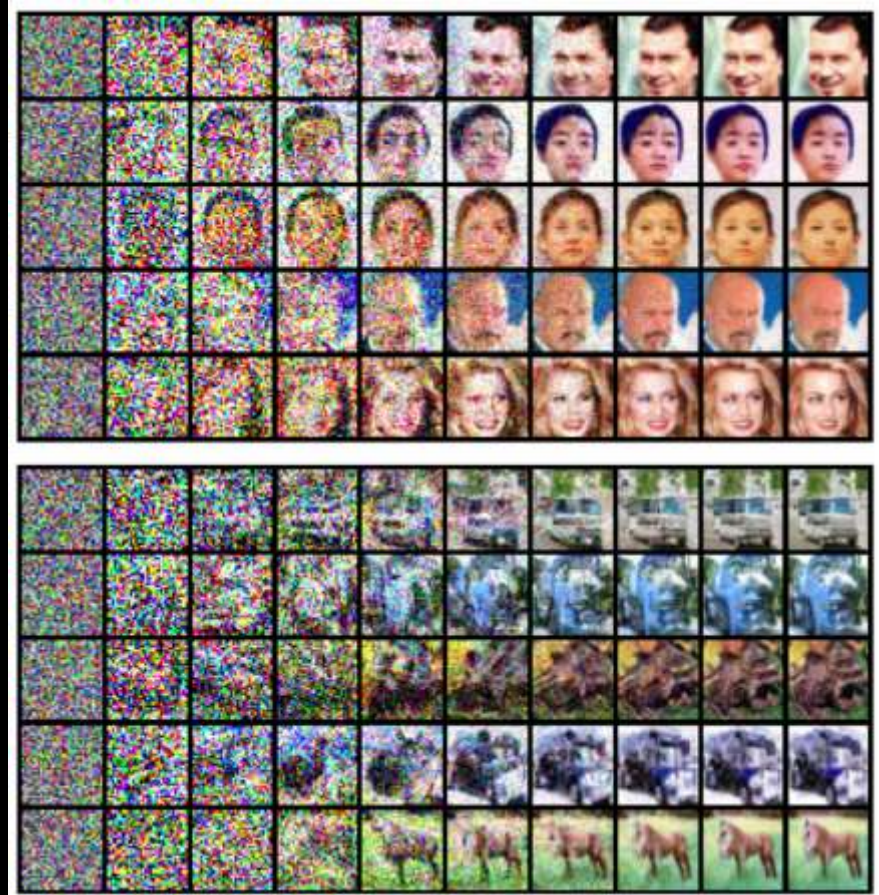


# Discovery 2: Image Synthesis

Does it work? Here are some results



Kadkhodaie & Simoncelli



Song &  
Ermon



# Discovery 2: Image Synthesis

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

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



BigGAN (FID 6.95)

Diffusion (FID 4.59)

## 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  $128 \times 128$ , 4.59 on ImageNet  $256 \times 256$ , and 7.72 on ImageNet  $512 \times 512$ , 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  $256 \times 256$  and 3.85 on ImageNet  $512 \times 512$ . We release our code at <https://github.com/openai/guided-diffusion>.

### 1 Introduction

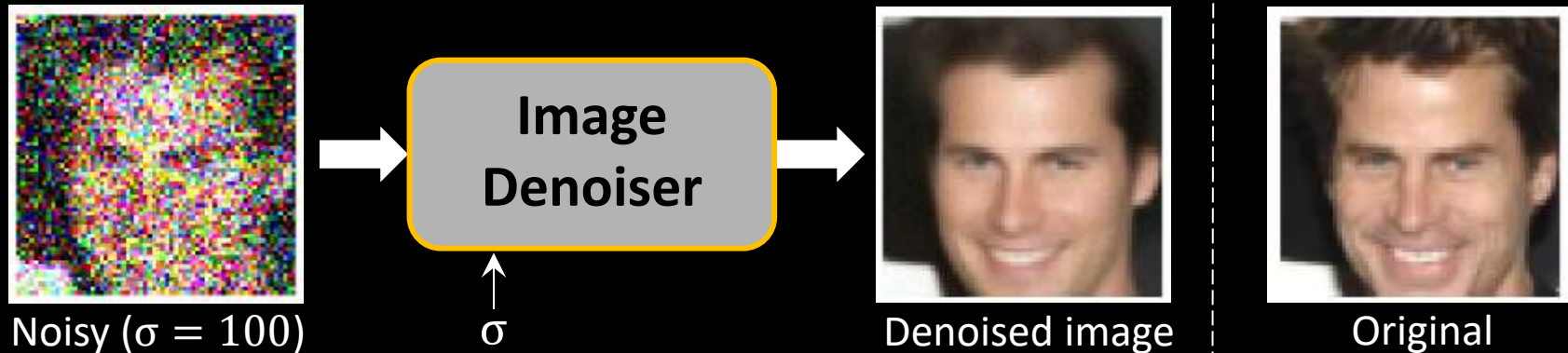


Figure 1: Selected samples from our best ImageNet  $512 \times 512$  model (FID 3.85)

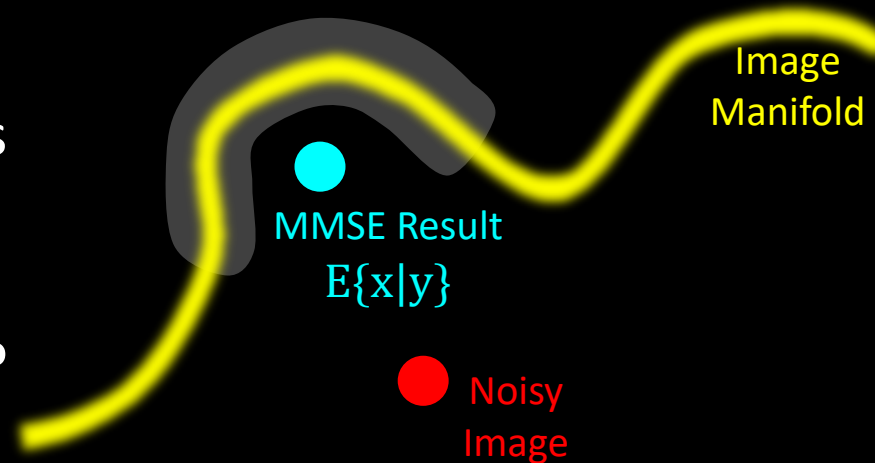


# Discovery 3: Targeting Perceptual Quality

Suppose that we need to denoise the following image:



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

# Discovery 3: 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

2021

G Ohayon, T Adrai, G Vaksman, M Elad, P Milanfar  
arXiv preprint arXiv:2103.04192

Stochastic Image Denoising by Sampling from the Posterior Distribution

2021

B Kawar, G Vaksman, M Elad  
arXiv preprint arXiv:2101.09552

**Answer:** Denoise by sampling from the posterior  $P(x|y)$

**Practical Implication:** We consider 2 methods

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

These methods  
produce a multitude  
of possible solutions



# Discovery 3: Targeting Perceptual Quality

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

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

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

$$\hat{x}_{k+1} = \hat{x}_k + a \cdot \underbrace{\nabla \log P(\hat{x}_k | y)}_{\text{Langevin with a conditional Score}} + b \cdot z_k$$

Bayes rule

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

$$= \underbrace{\hat{x}_k - D(\hat{x}_k, \sigma)}_{\text{Approx. Score}} + \underbrace{\nabla \log P(y | \hat{x}_k)}_{\text{A Gaussian Distribution}}$$

Approx. Score

A Gaussian Distribution



# Discovery 3: 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:  $n \sim \mathcal{N}(0, \sigma_0^2 \mathbf{I})$
- Synthetic annealing noise:  $v \sim \mathcal{N}(0, \sigma_k^2 \mathbf{I})$  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  $n$

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



# Discovery 3: 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



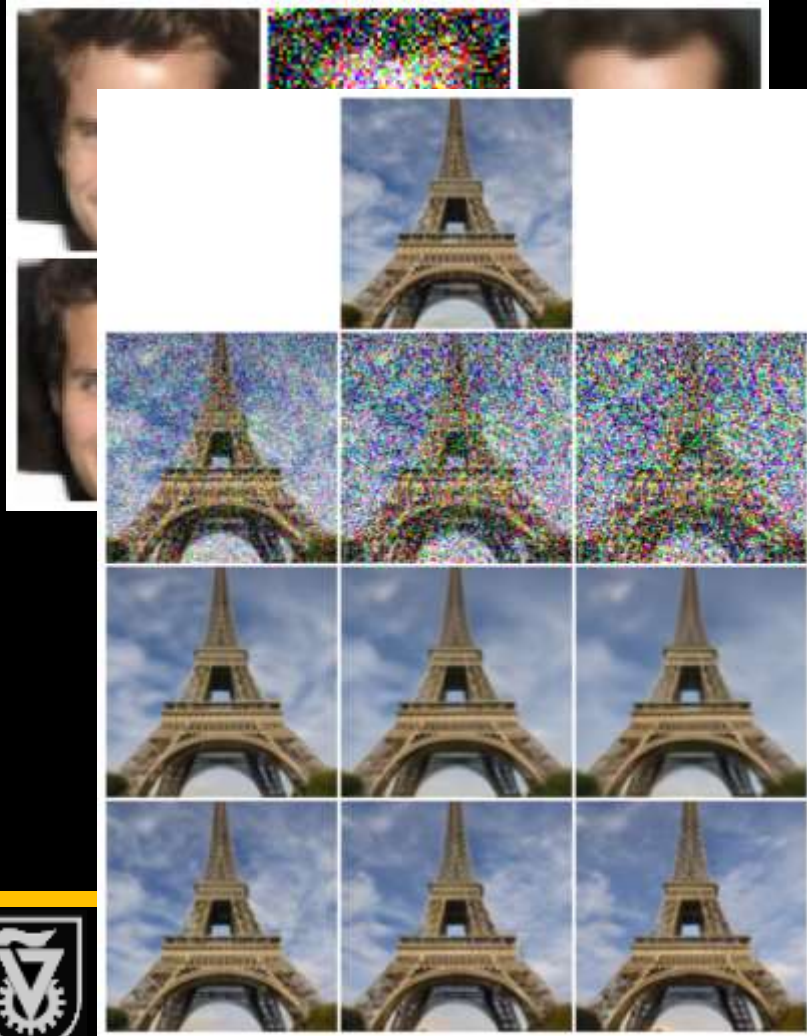
# Discovery 3: Targeting Perceptual Quality

## Stochastic Image Denoiser:



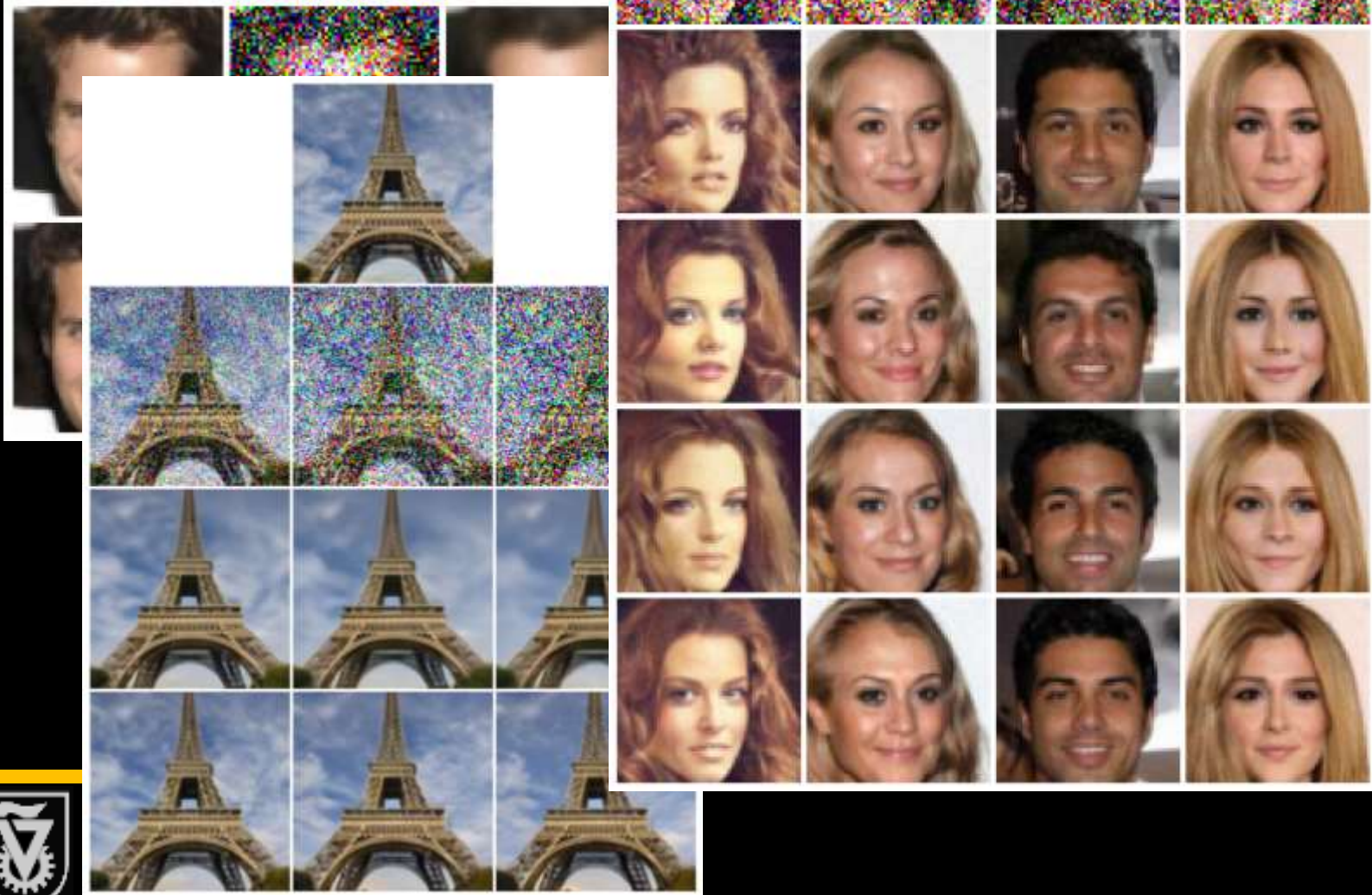
# Discovery 3: Targeting Perceptual Quality

## Stochastic Image Denoiser:



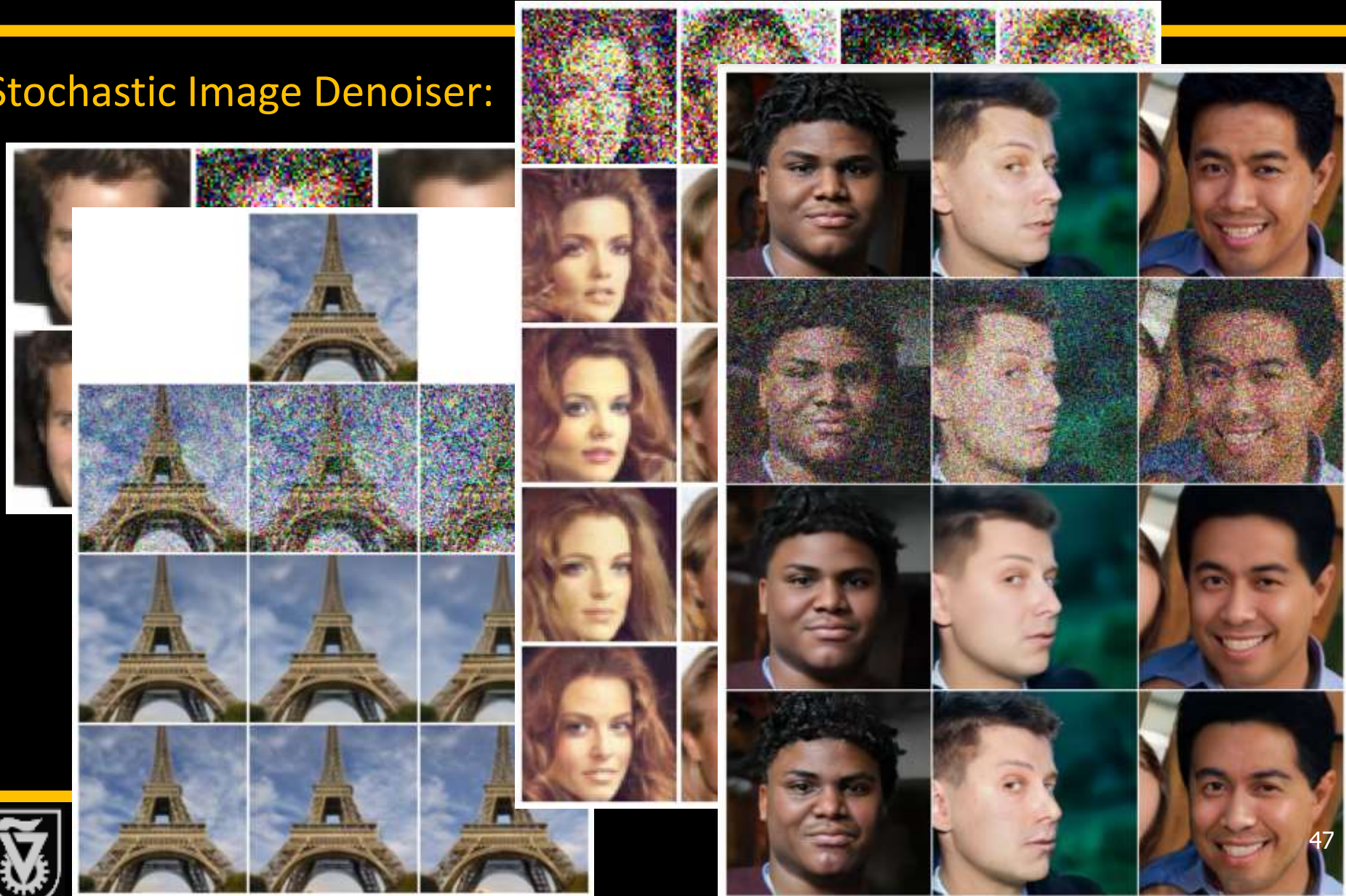
# Discovery 3: Targeting Perceptual Quality

Stochastic Image Denoiser:



# Discovery 3: Targeting Perceptual Quality

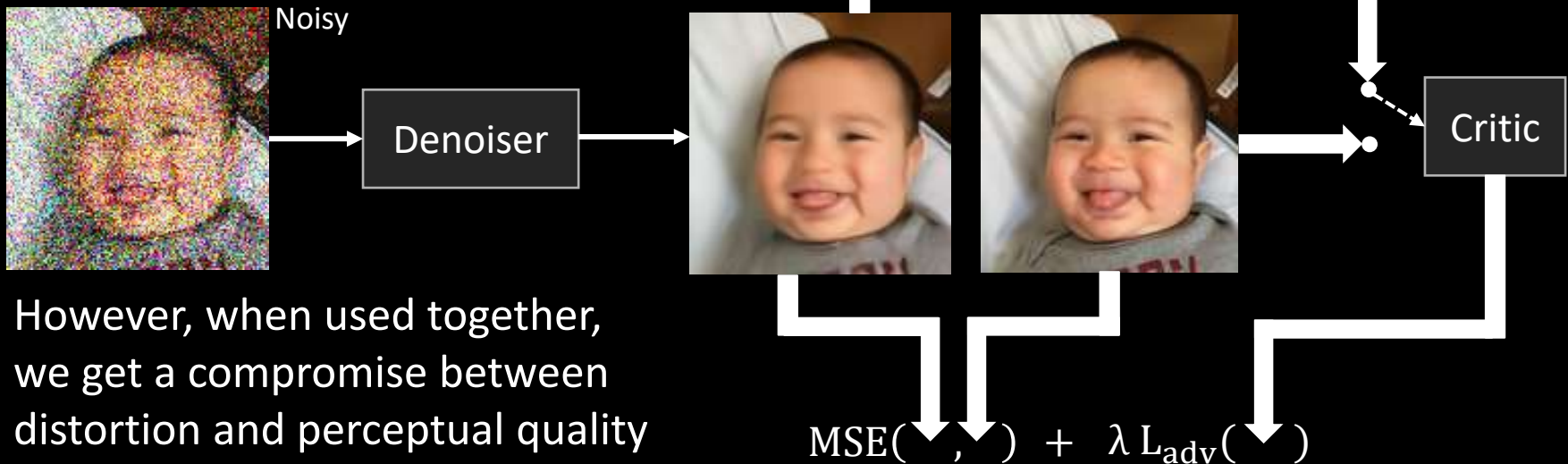
## Stochastic Image Denoiser:



# Discovery 3: Targeting Perceptual Quality

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

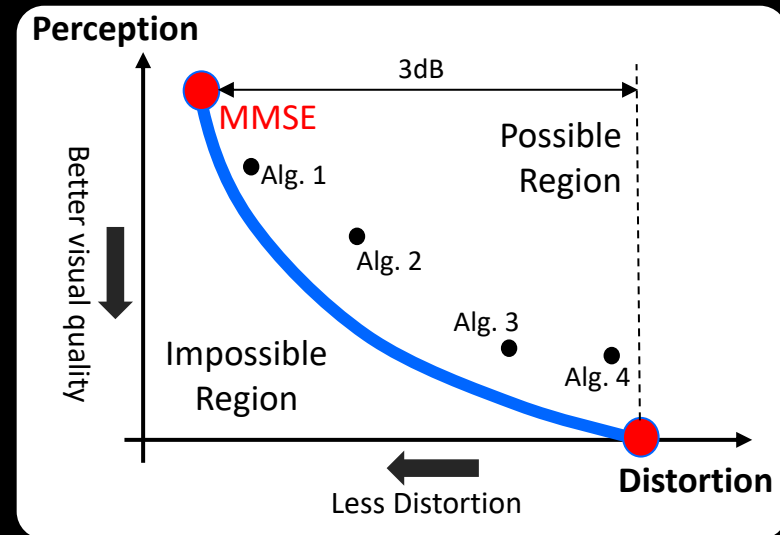
- ❑ Typical design approach: Optimize a distortion measure (e.g. MSE) between the denoised and the ideal images
- ❑ Adversarial loss could be added to improve the perceptual quality



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

# Discovery 3: 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**)



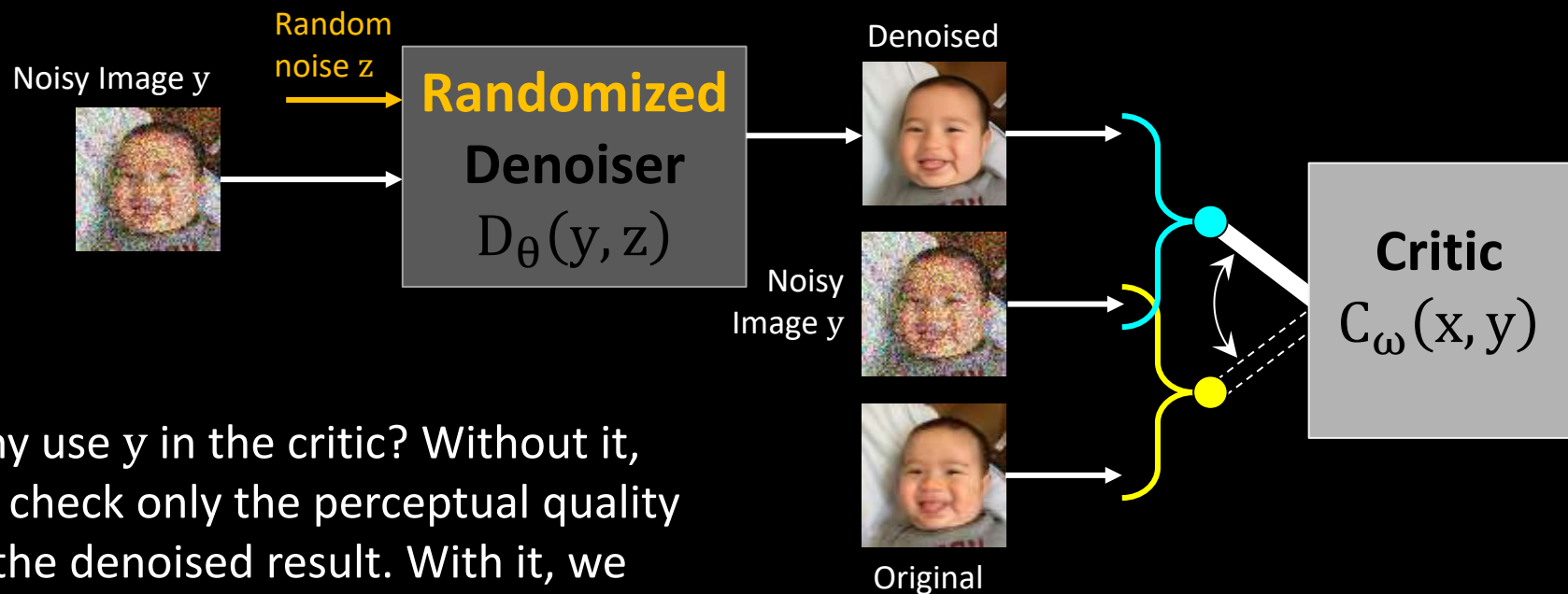
$$x \sim P_X \quad y \sim P_{Y|X=x}$$

Samples from  $P_{X|Y=y}$



# Discovery 3: 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 3: Targeting Perceptual Quality

## What about the Loss?

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

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

- ❑ 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

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

- ❑ Our remedy: adding an MMSE oriented penalty term:

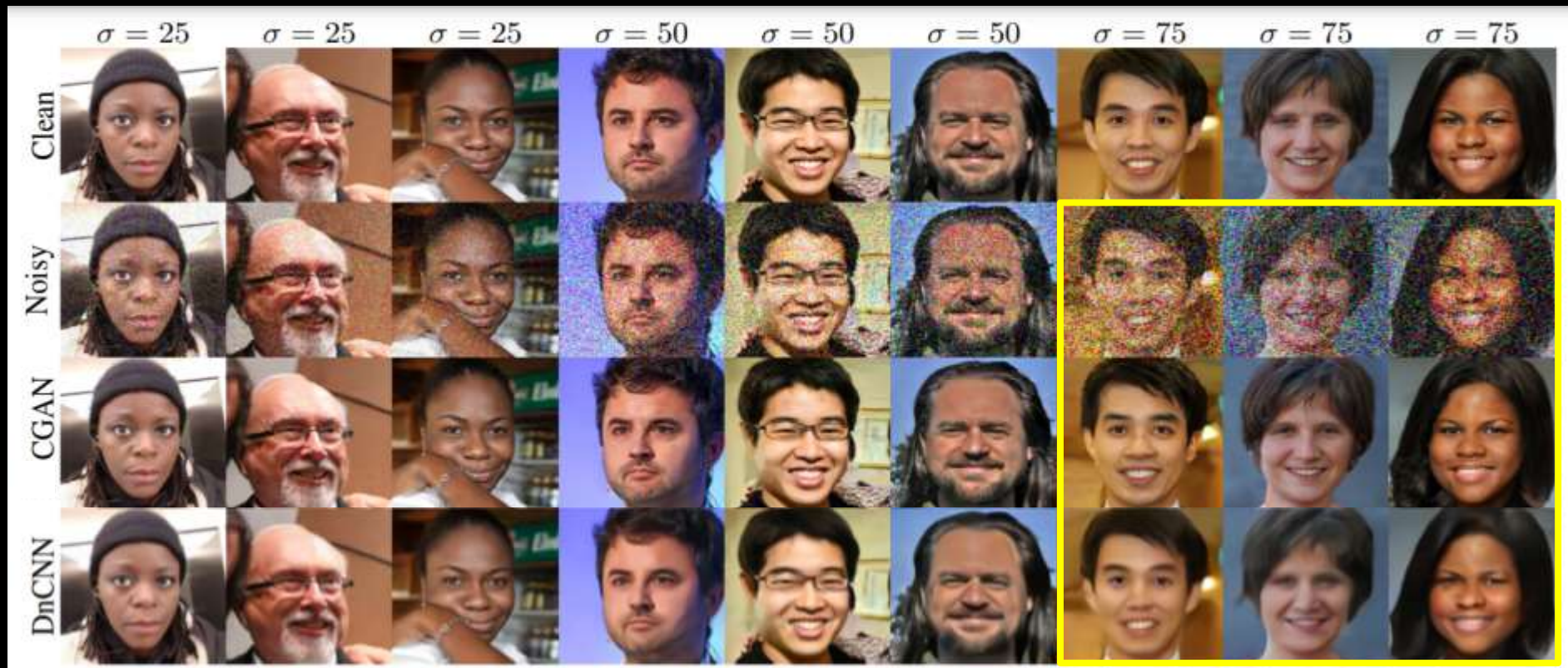
$$\mathbb{E}_{X,Y} [\|x - \mathbb{E}_Z [D_{\theta}|y]\|_2^2]$$

- ❑ This gives the MMSE result “for free” (averaging many instances)



# Discovery 3: Targeting Perceptual Quality

## CGAN:

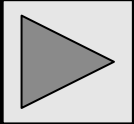


# Discovery 3: Targeting Perceptual Quality

## CGAN:



# Oh ... and One Last Thing



# Back to 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

B Kawar, G Vaksman, M Elad  
arXiv preprint arXiv:2105.14951

2021



# Back to 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

$$\nabla \log P(\mathbf{x}_i | \mathbf{y}) = \nabla \log P(\mathbf{x}_i) + \nabla \log P(\mathbf{y} | \mathbf{x}_i)$$

- ❑ We know that  $\mathbf{y} = H\mathbf{x} + \mathbf{n}$  and thus:

$$\nabla \log P(\mathbf{y} | \mathbf{x}_i) = \nabla \log P(\mathbf{y} - H\mathbf{x}_i | \mathbf{x}_i) =$$

$$\nabla \log P(H\mathbf{x} + \mathbf{n} - H\mathbf{x} - H\mathbf{v}_i | \mathbf{x}_i) = \nabla \log P(\mathbf{n} - H\mathbf{v}_i | \mathbf{x}_i)$$

- ❑ However, ... while  $\mathbf{n} - H\mathbf{v}_i$  is a simple Gaussian, its dependency on  $\mathbf{x}_i$  is non-trivial, so how do we proceed from here?



# Back to Inverse Problems

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

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



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

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

- Thus, we can apply the Langevin dynamics algorithm on  $\tilde{x}_T = V^T x_i$  given  $y_T = U^T y$  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



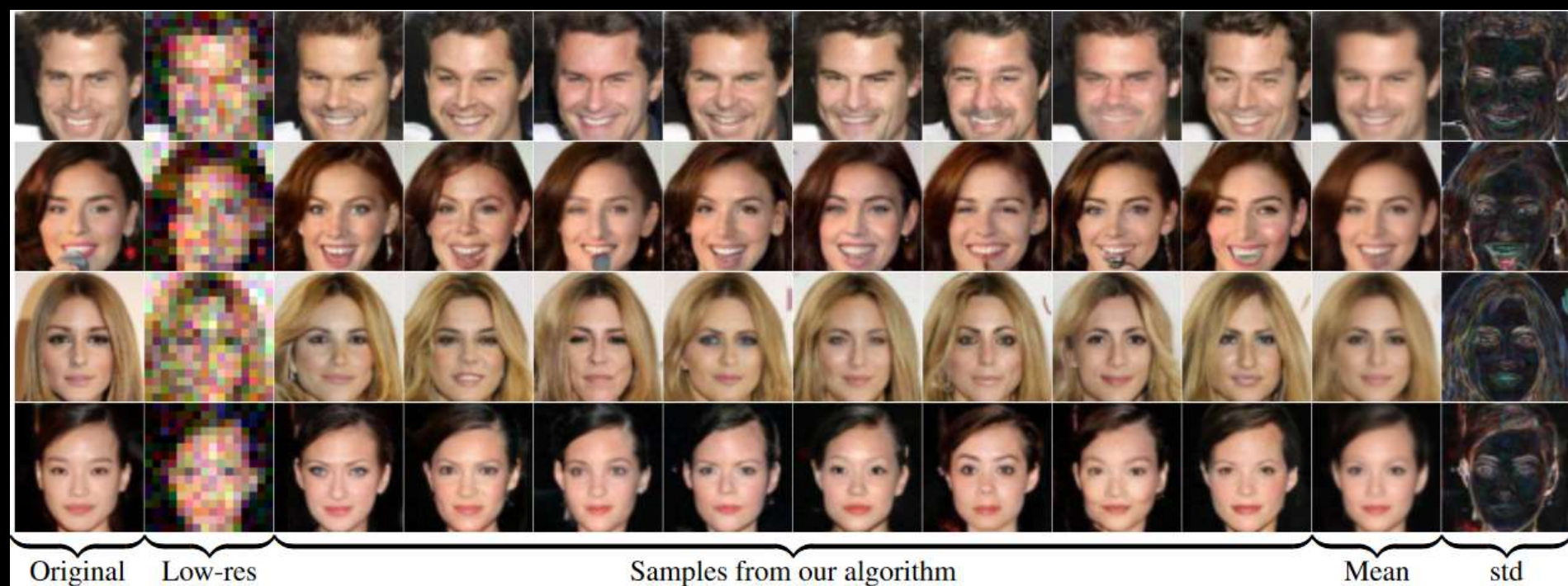
# Back to Inverse Problems

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



# Back to Inverse Problems

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



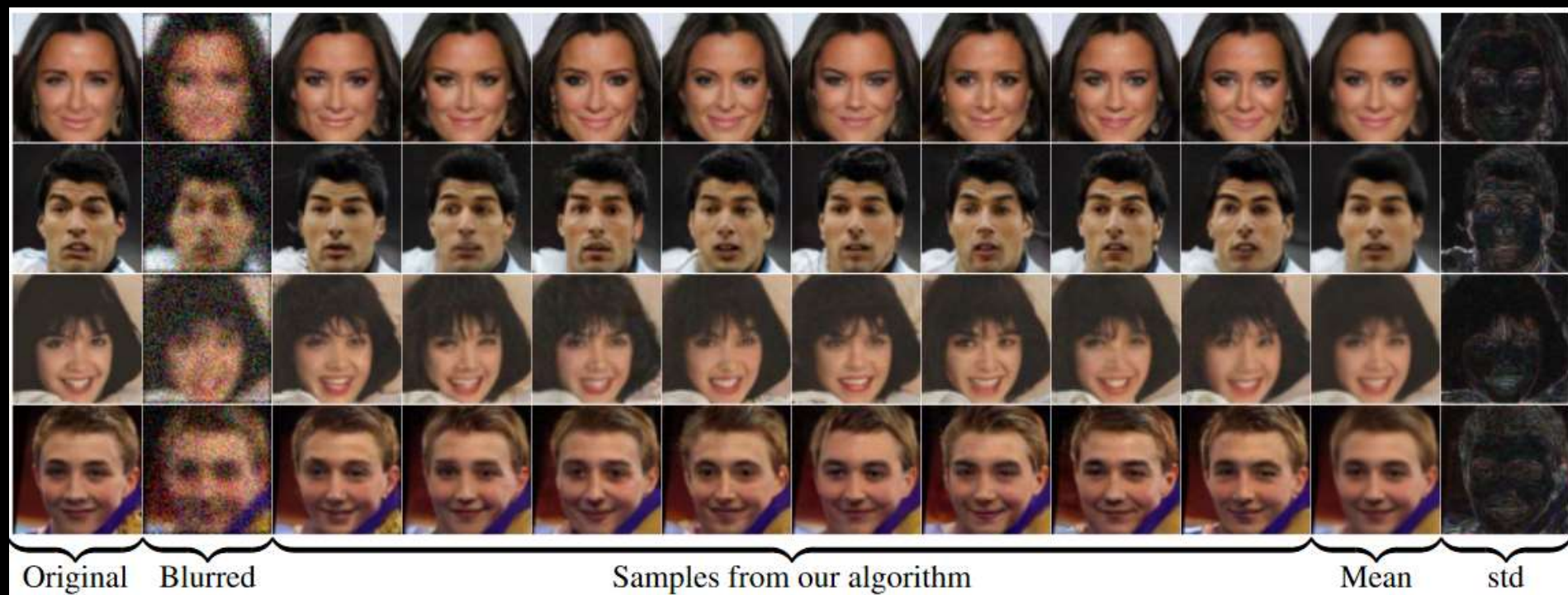
# Back to Inverse Problems

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



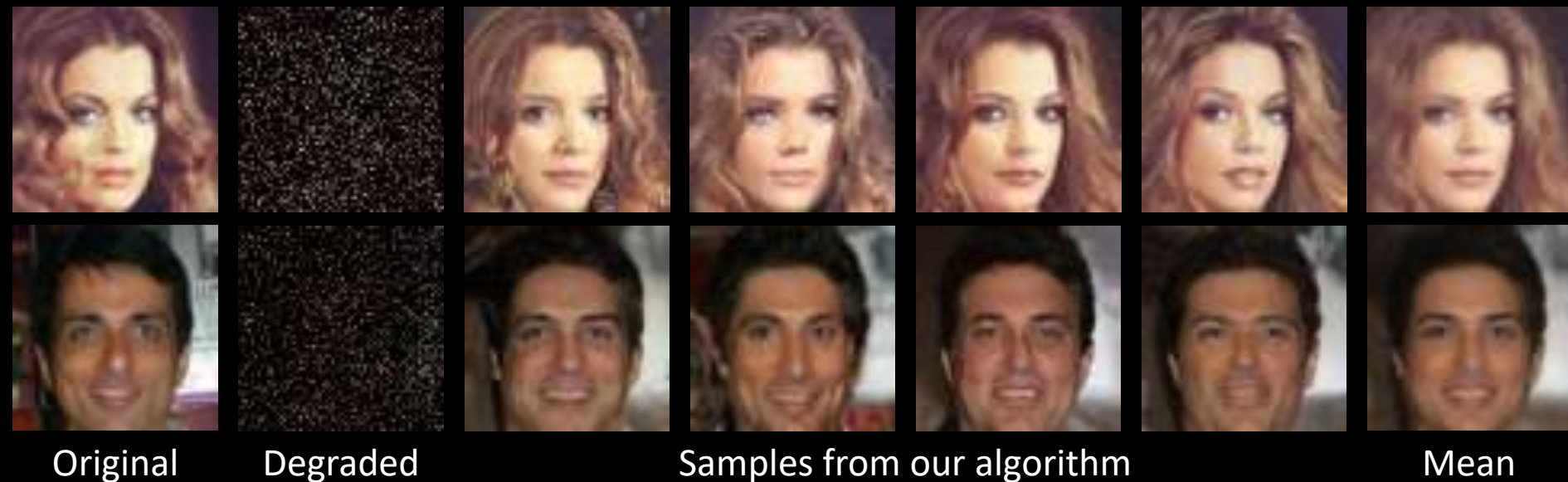
# Back to Inverse Problems

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



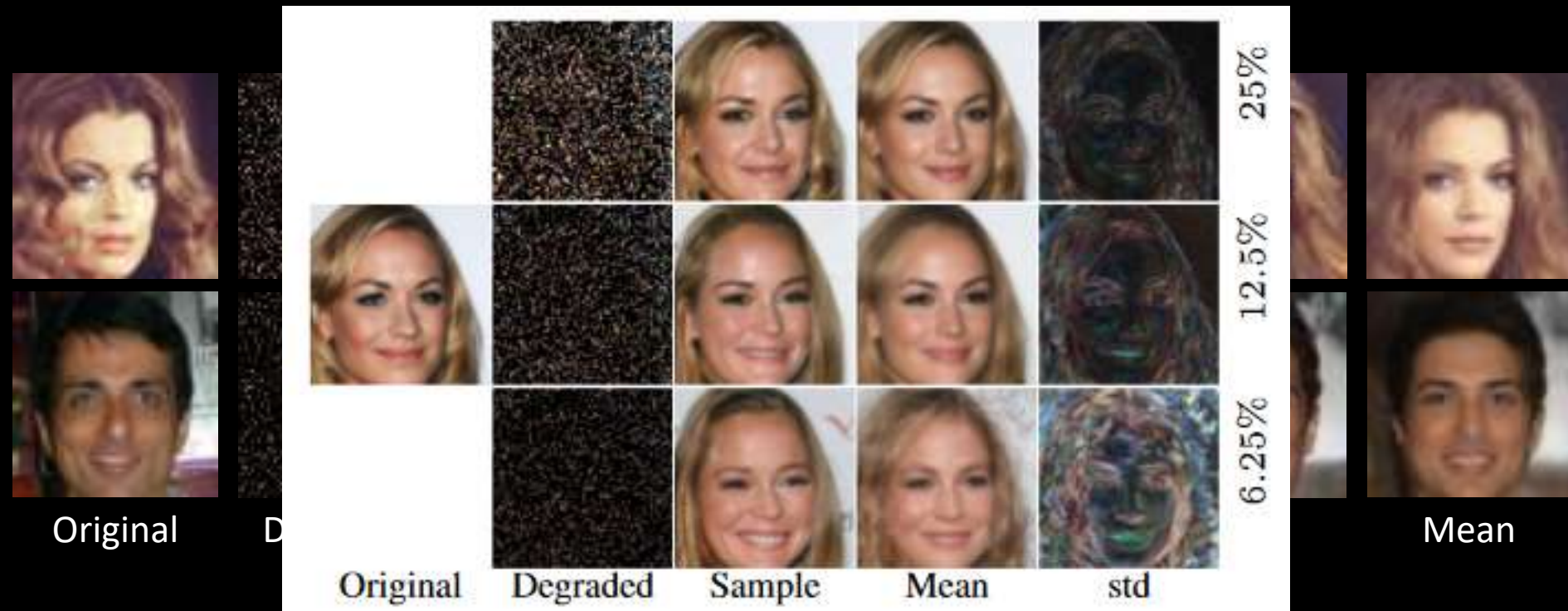
# Back to Inverse Problems

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



# Back to Inverse Problems

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



# 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



# Time to Summarize



# 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

