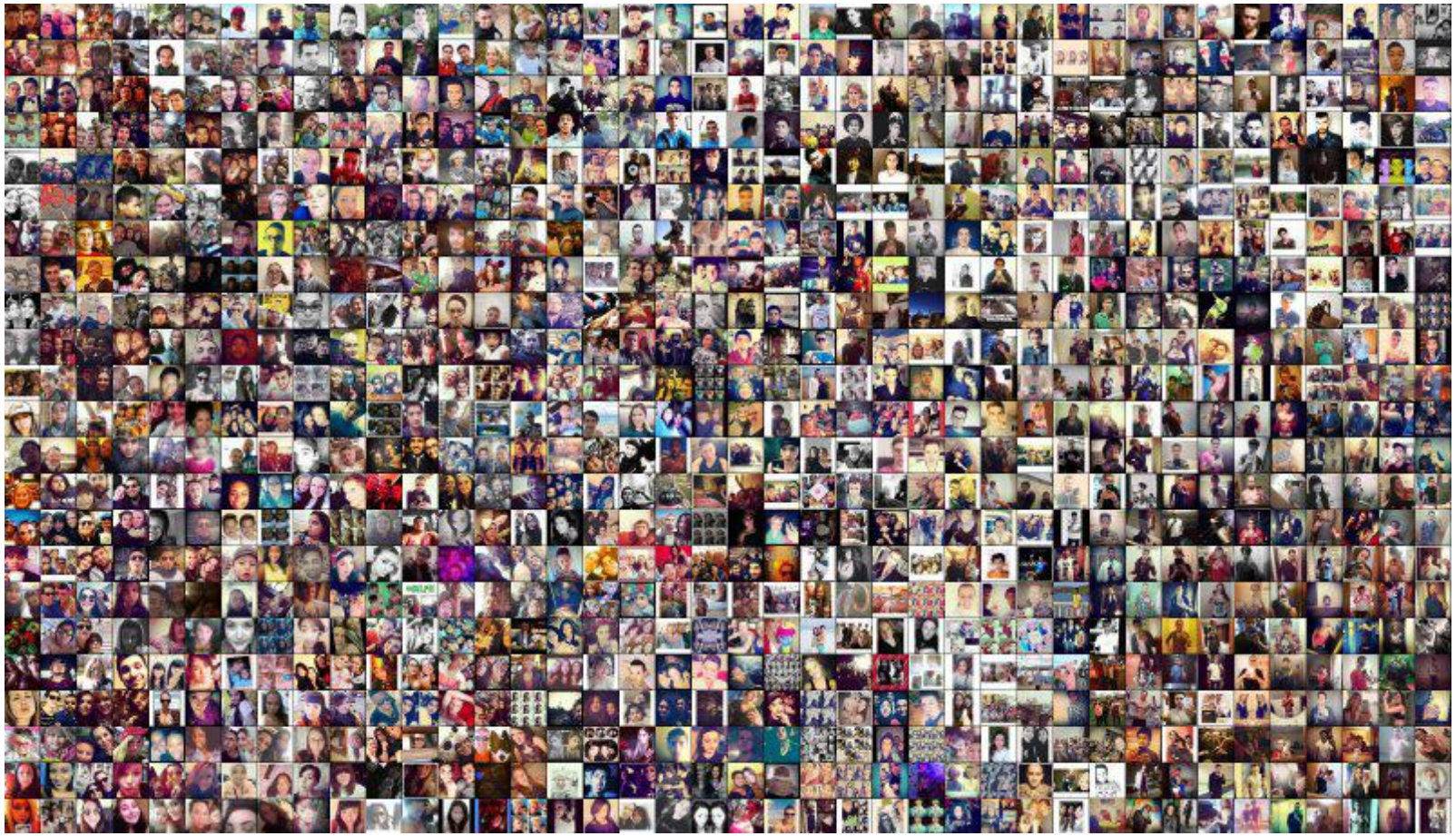
Deep Image Prior

Dmitry Ulyanov · Andrea Vedaldi · Victor Lempitsky

- Deep convolutional networks' excellent performance is due to their ability to learn realistic image priors from **a large number of example images**.



- The **structure** of a neural network is sufficient to capture a great deal of low-level image statistics **prior to any learning**.

- Instead of following the common paradigm of training a ConvNet on a large dataset of example images, they **fit a generator network to a single degraded image**.

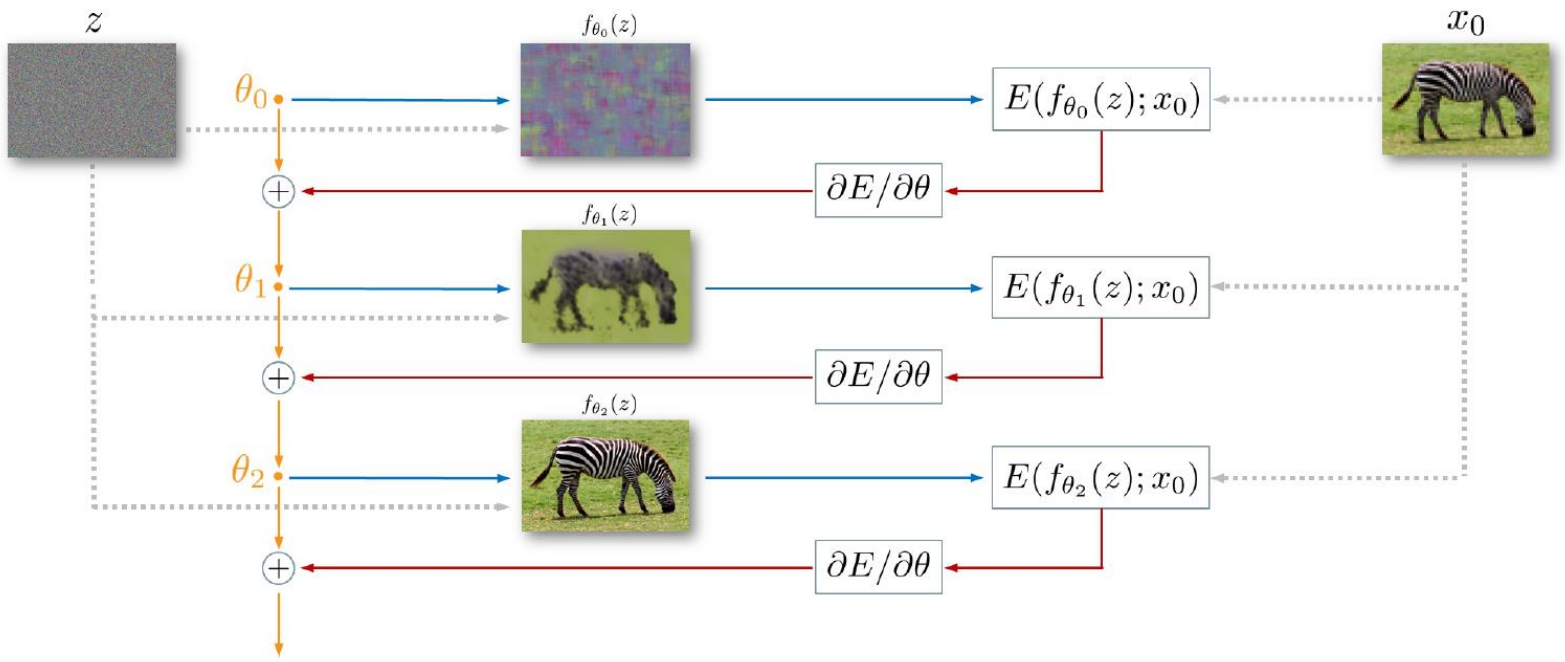
- The weights are **randomly initialized** and fitted to maximize their likelihood given a specific degraded image and a task-dependent observation model.

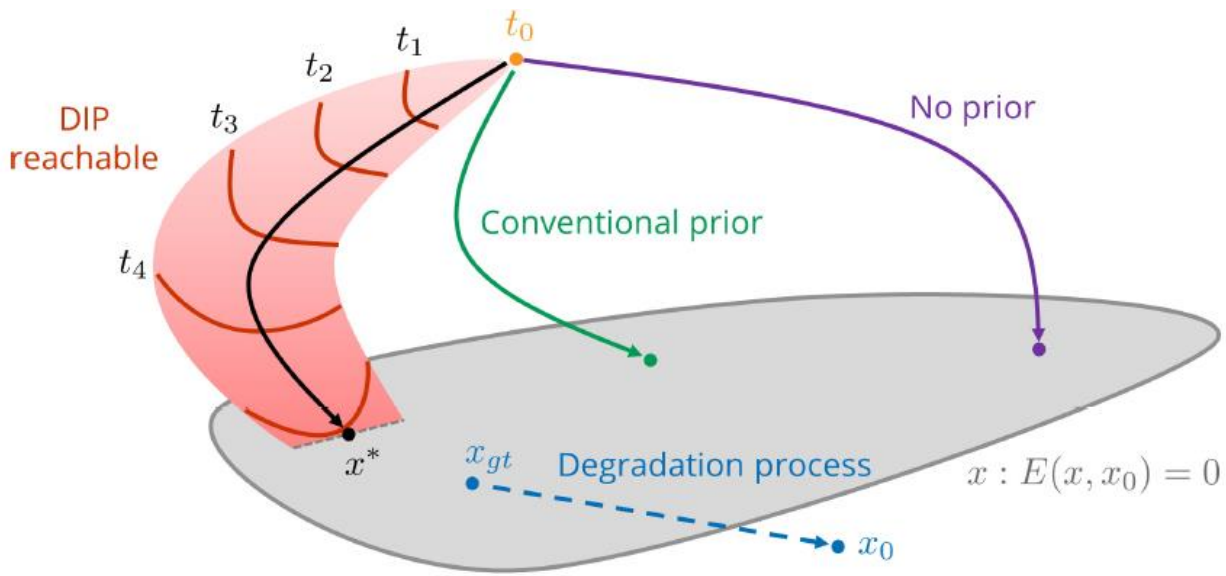
parametrization

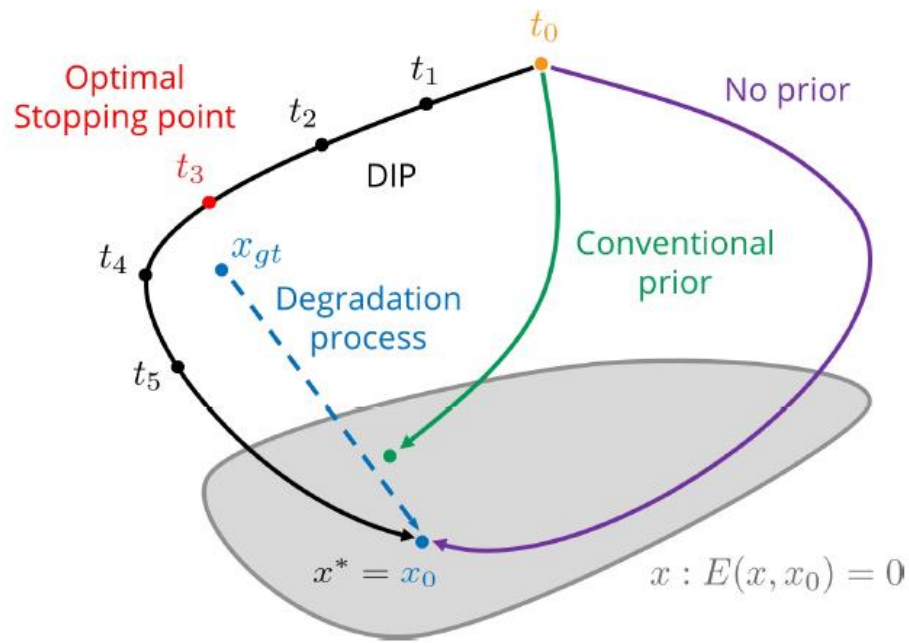
$$x^* = \min_x E(x; x_0) + R(x)$$

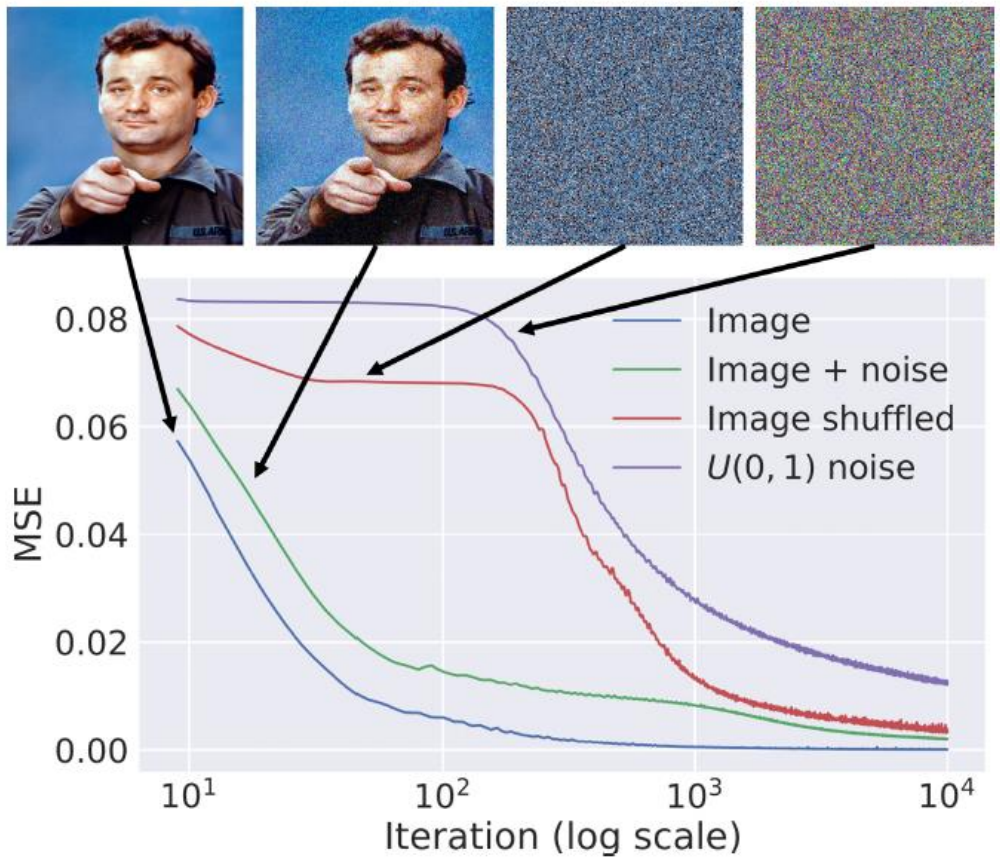
$$\theta^* = \operatorname*{argmin}_\theta E(f_\theta(z); x_0), \qquad x^* = f_{\theta^*}(z)$$

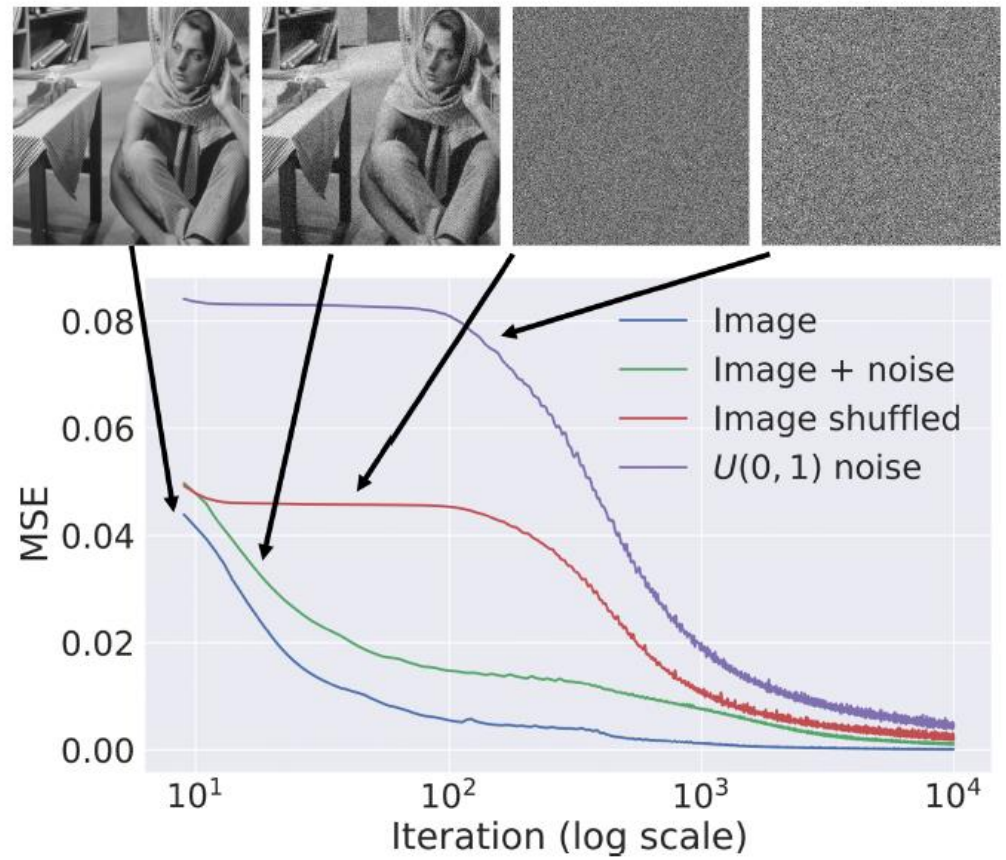
$$\min_\theta \|f_\theta(z) - x_0\|^2$$




t_0
t_1
t_2
t_3
t_4
DIP
reachable
No prior
Conventional prior
x^*
x_{gt}
Degradation process
x_0
$x : E(x, x_0) = 0$

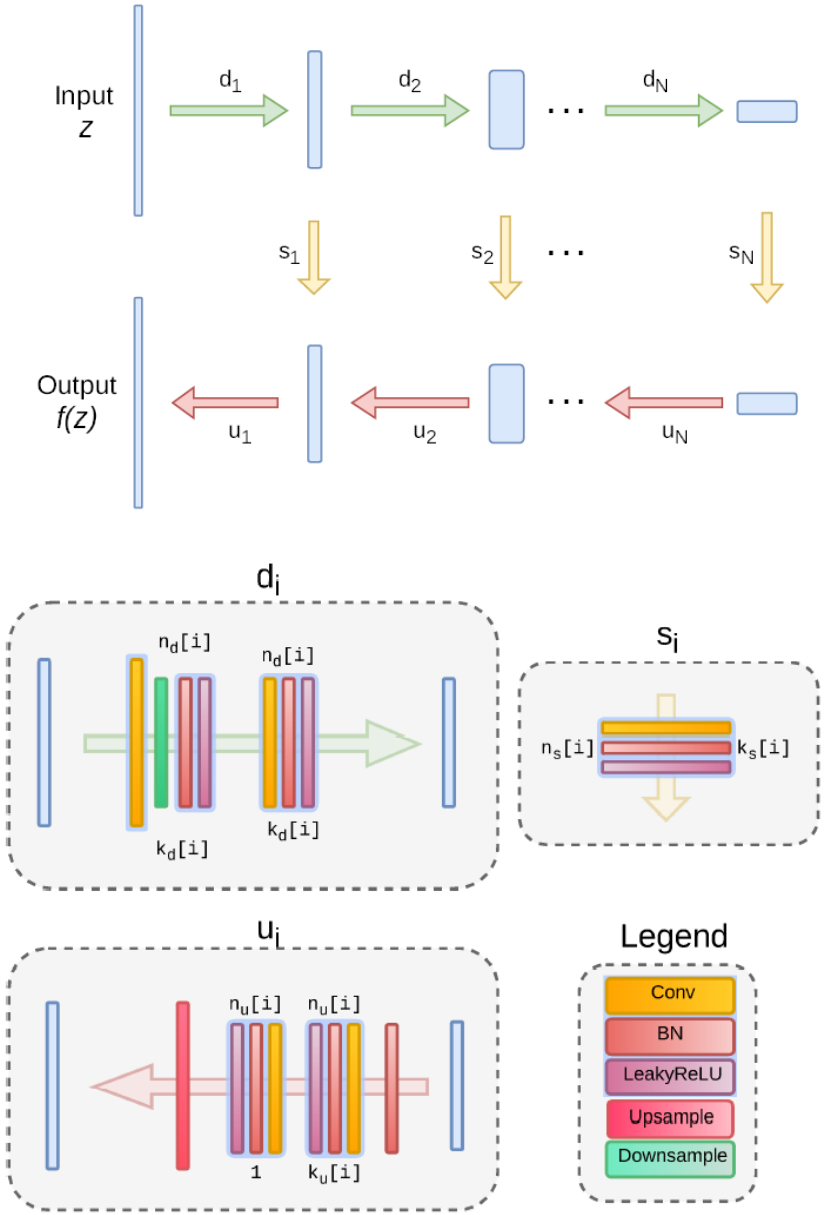

t_0
Optimal
Stopping point
t_1
t_2
t_3
DIP
No prior
t_4
x_{gt}
Degradation
process
Conventional
prior
t_5
$x^* = x_0$
$x : E(x, x_0) = 0$


0.08
0.06
0.04
0.02
0.00
MSE
Image
Image + noise
Image shuffled
$U(0,1)$ noise
10^1
10^2
10^3
10^4
Iteration (log scale)

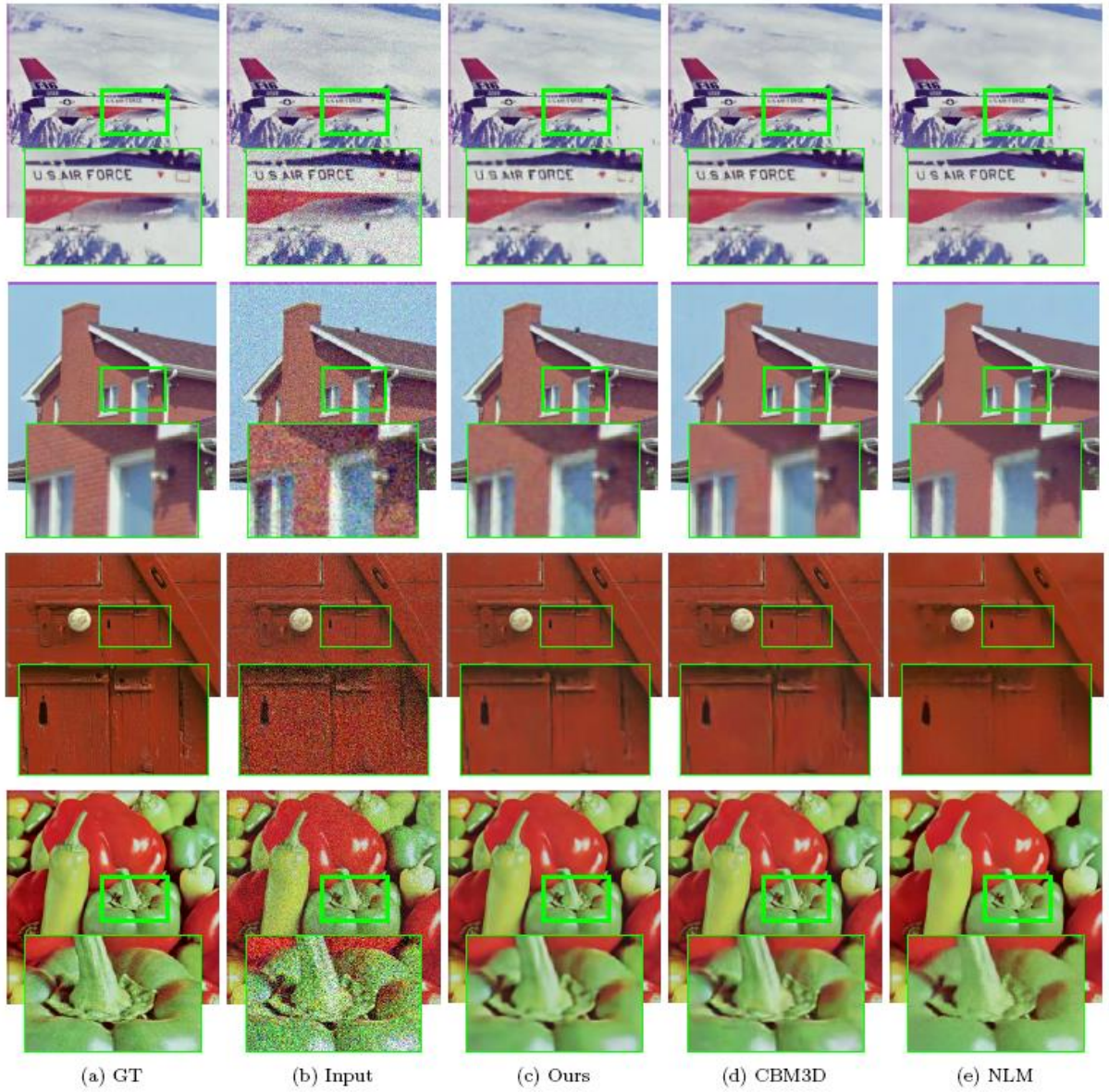

0.08
0.06
0.04
0.02
0.00
MSE
Image
Image + noise
Image shuffled
$U(0,1)$ noise
10^1
10^2
10^3
10^4
Iteration (log scale)

Applications

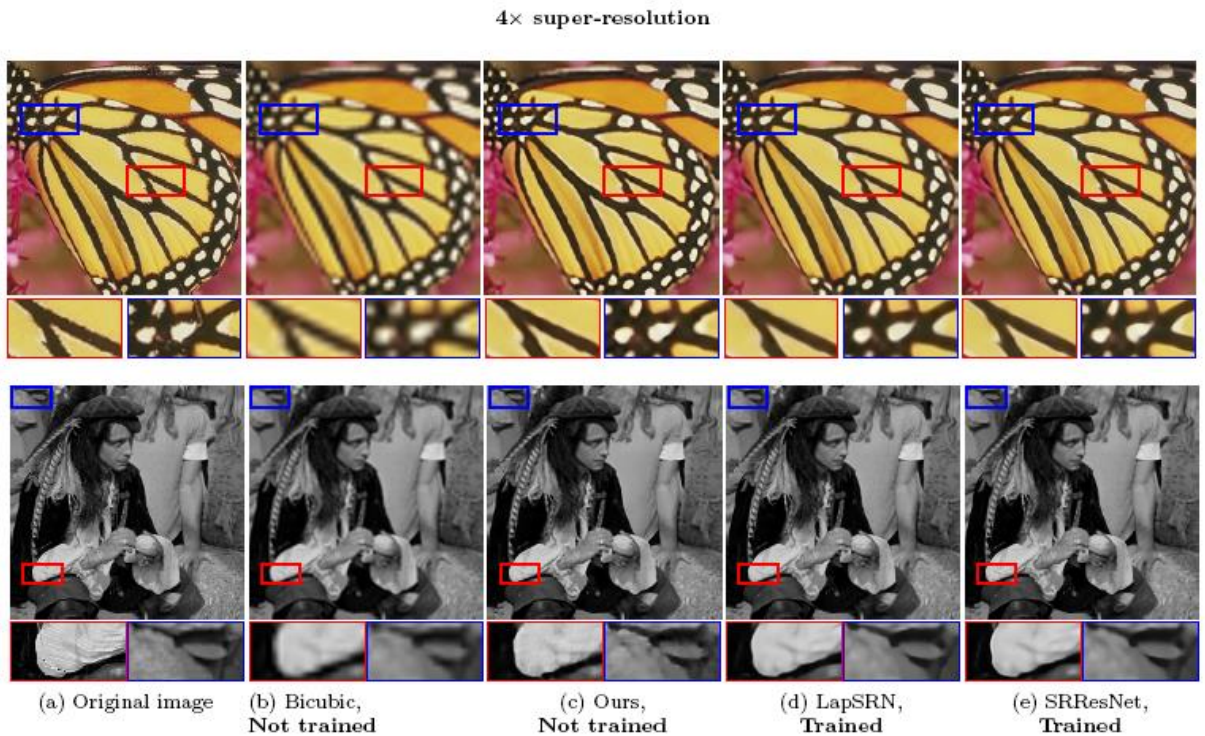
- Denoising
- Super-resolution
- Inpainting

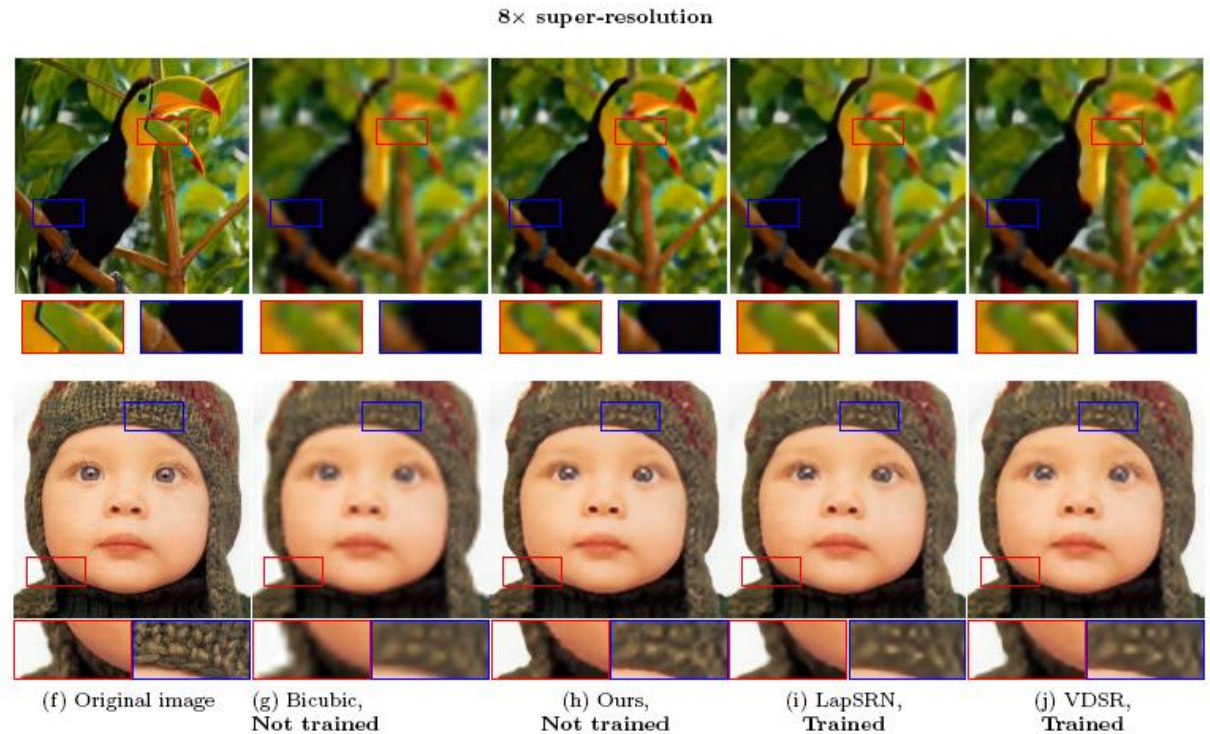


Denoising



SR

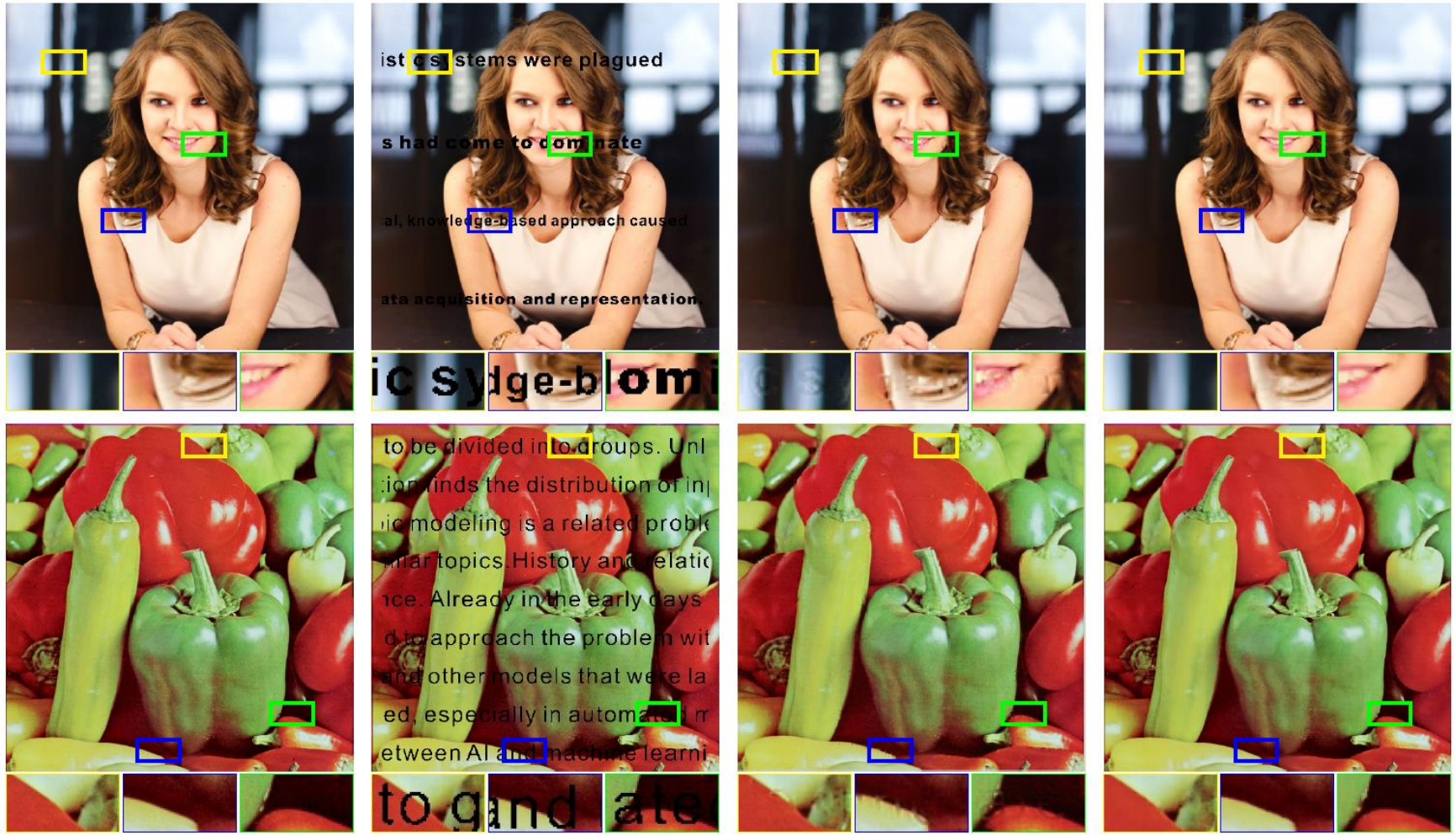

4× super-resolution
(a) Original image
(b) Bicubic, Not trained
(c) Ours, Not trained
(d) LapSRN, Trained
(e) SRResNet, Trained


8× super-resolution
(f) Original image
(g) Bicubic, Not trained
(h) Ours, Not trained
(i) LapSRN, Trained
(j) VDSR, Trained

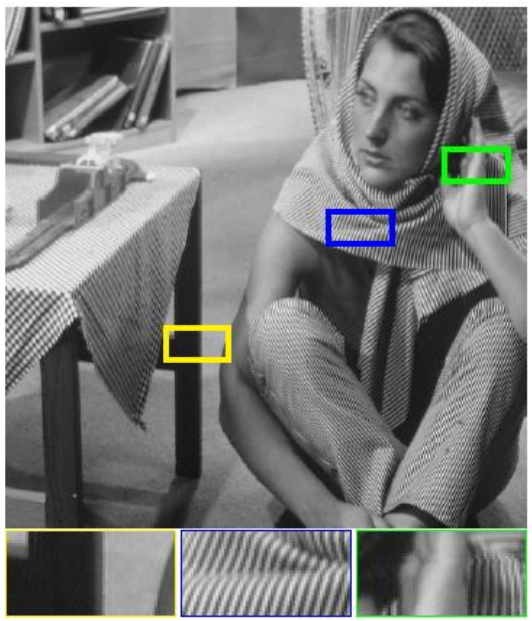
$$x^* = \min_x E(x; x_0) + R(x)$$

(a) HR image (b) Bicubic upsampling (c) No prior (d) TV prior (e) Deep image prior

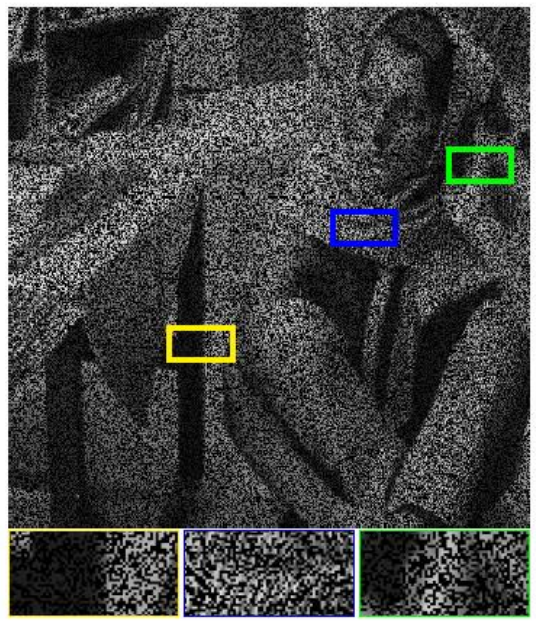
Inpainting



(a) Original image (b) Corrupted image (c) Shepard networks [44] (d) Deep Image Prior



(a) Original image



(b) Corrupted image

(c) CSC [42]

(d) Deep image prior

(a) Input (white=masked)

(b) Encoder-decoder, depth=6

(c) Encoder-decoder, depth=4

(d) Encoder-decoder, depth=2

(e) ResNet, depth=8

(f) U-net, depth=5

Conclusion

- Revealed the contribution of the prior imposed by the neural network architecture in processing images.
- Showed neural network tend to extract useful information contained in an input image during the initial training phase

Limitation

- Time costing