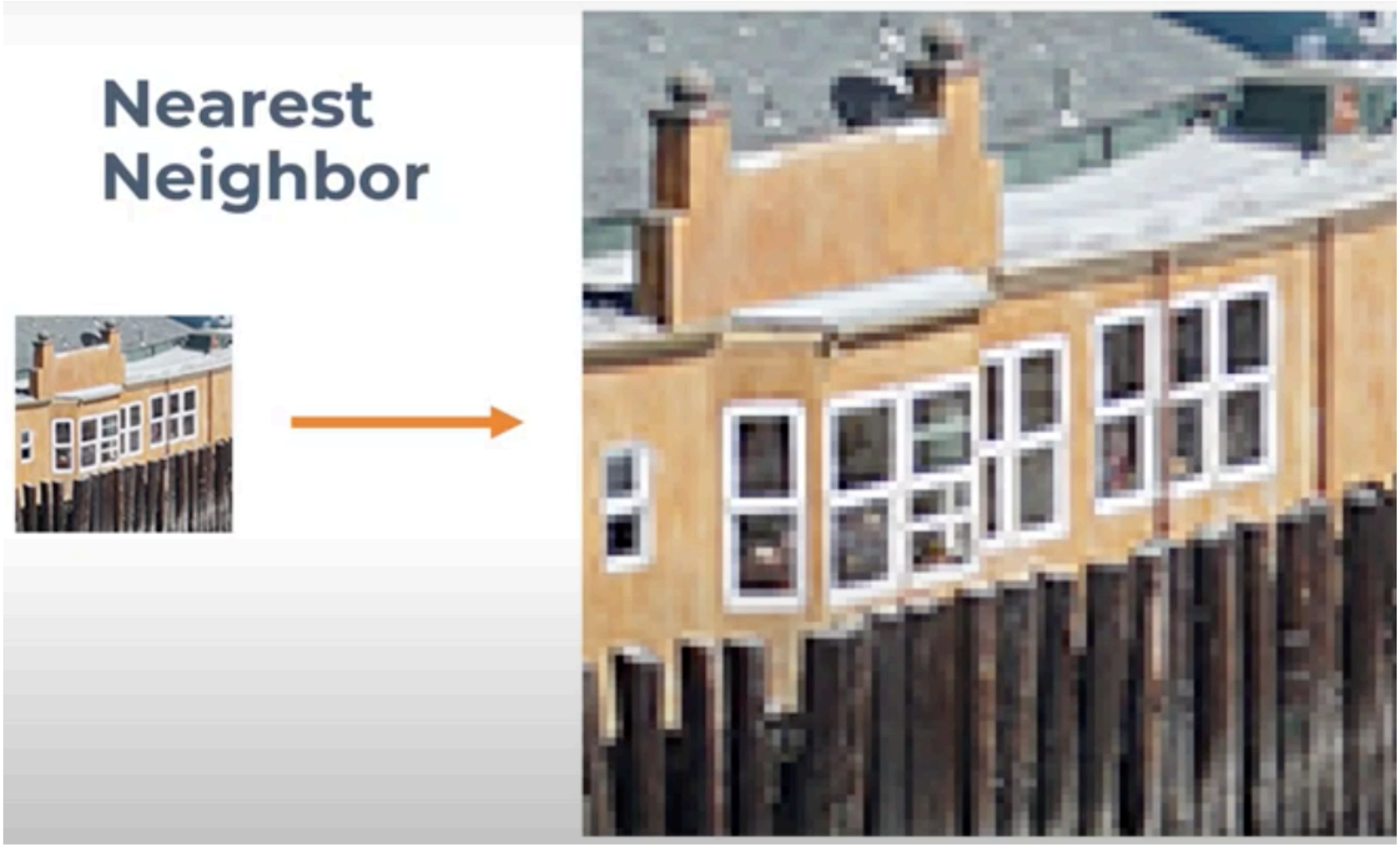
Super-Resolution

- Super resolution is the process of upscaling and or improving the details within an image.
- The details in the high resolution output are filled in where the details are essentially unknown.

Techniques

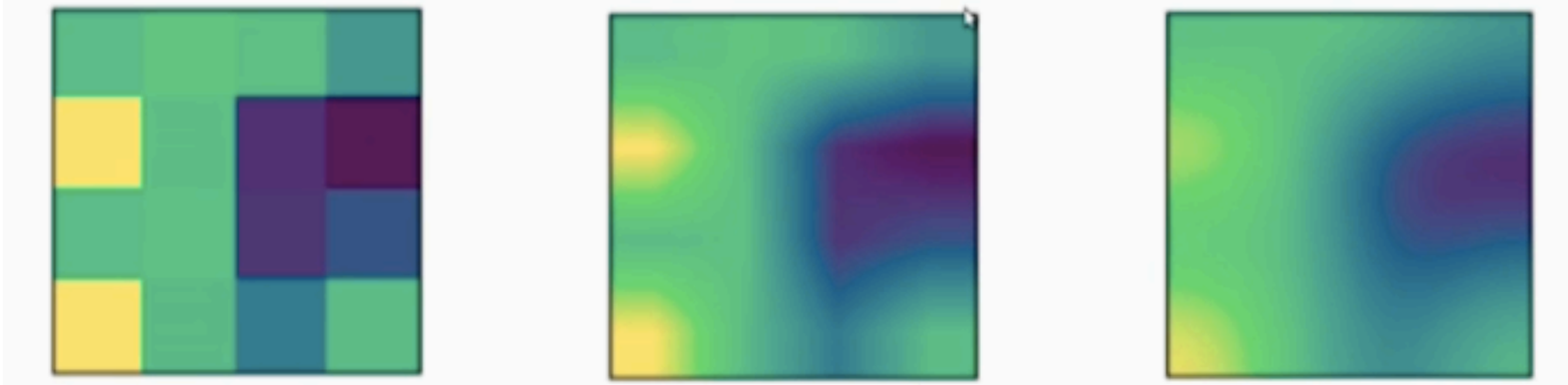
- Nearest neighbor interpolation
- Bilinear interpolation
- Bicubic interpolation
- Neural networks



Checker box effects in nearest neighbor interpolation

Bilinear Interpolation

Blurred effect in bilinear interpolation



nearest neighbour bilinear bicubic

Why are these results not perfect?

Data Processing Inequality

- Concept of information theory stating that no clever transformation of the received code Y can give more information about the sent code X than Y itself.

$$X \to Y \to Z$$

$$I(X,Y) \geq I(X,Z)$$

Neural Network

Neural Network

SRCNN

(2015)

LR image

Bicubic interpolation

Conv 9x9

Conv 1x1

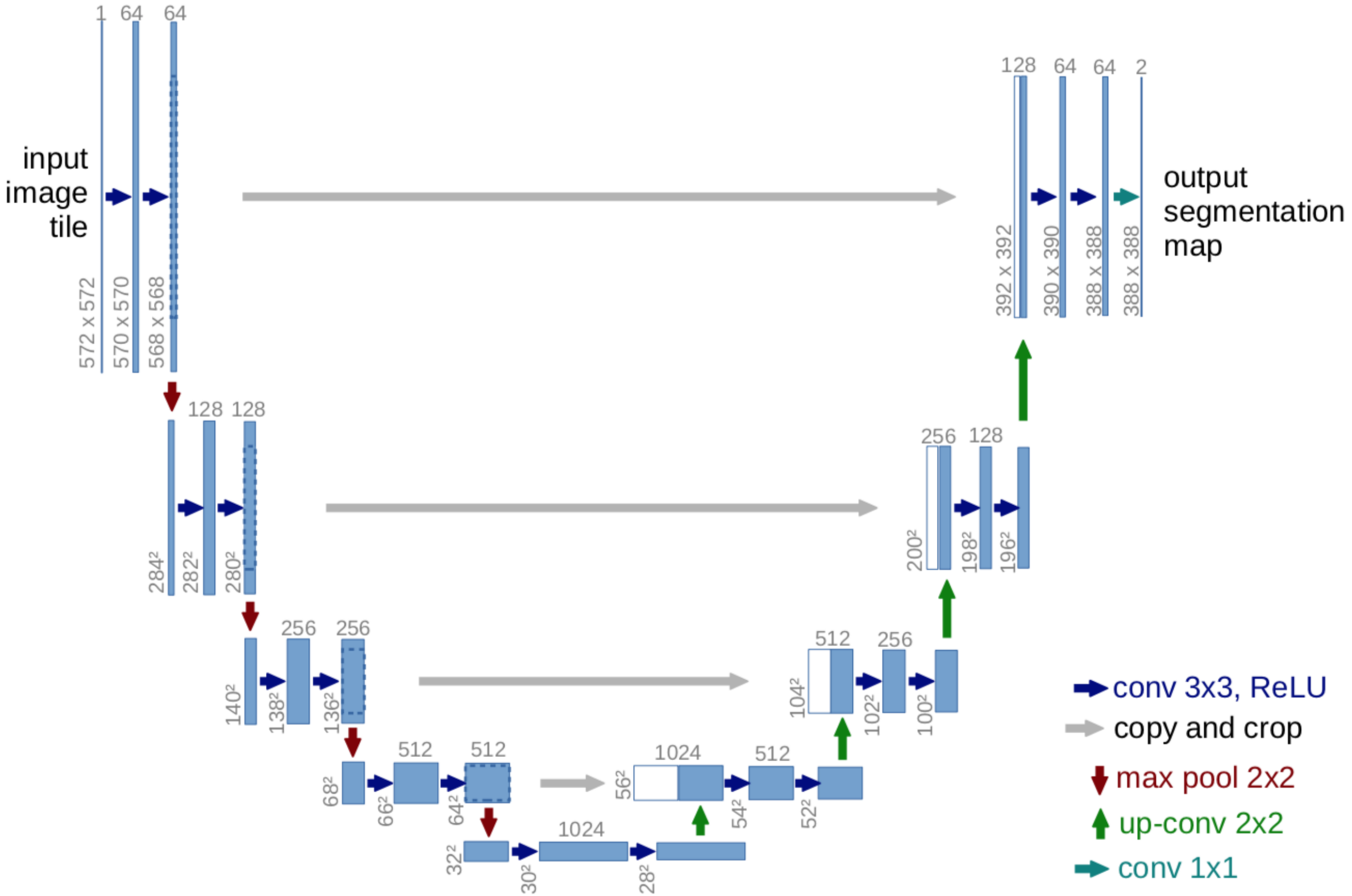
Conv 5x5

HR image

Model output

Original

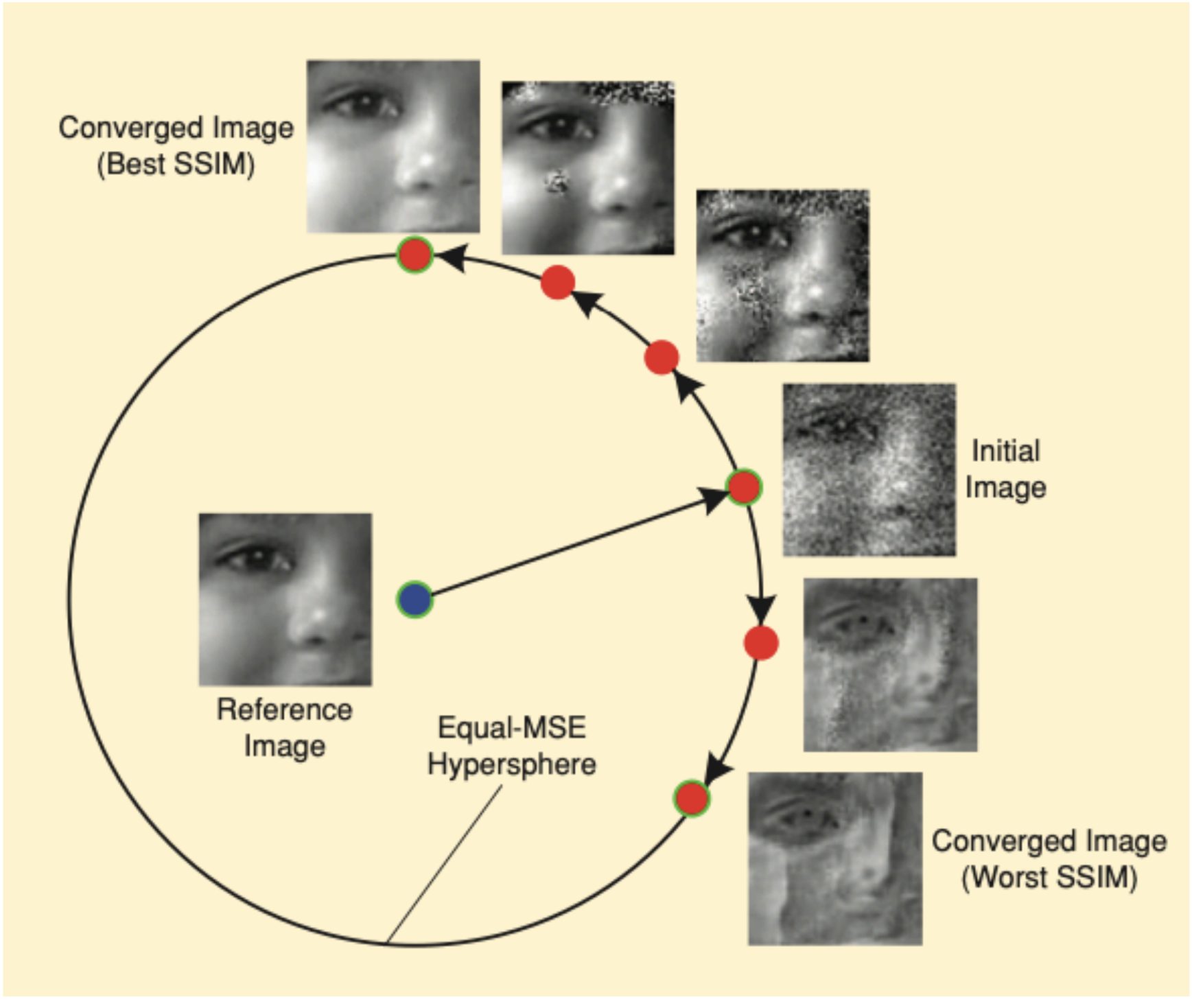
U-Net


input image tile
1
64
64
572 x 572
570 x 570
568 x 568
128
128
284²
282²
280²
256
256
140²
138²
136²
512
512
68²
66²
64²
1024
32²
30²
1024
512
56²
54²
28²
512
256
104²
102²
100²
256
128
200²
198²
196²
128
64
64
2
392 x 392
390 x 390
388 x 388
388 x 388
output segmentation map
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1

Loss Functions

Mean squared error
Structural similarity index
Perceptual loss
Pixel loss
Gram loss
Content loss
Texture loss
.
.

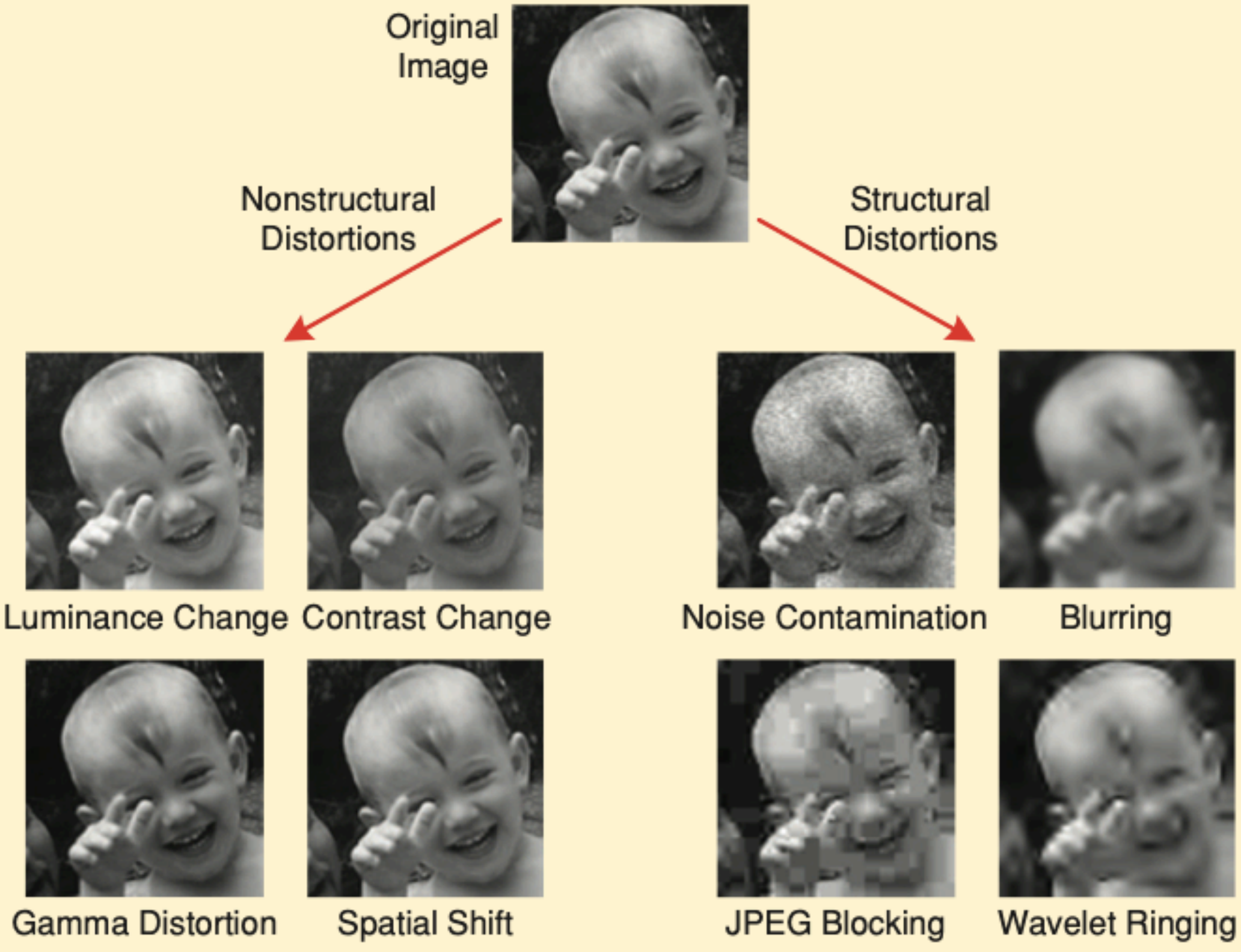
Mean Squared Error



Structural Similarity Index (SSIM)

- Natural image signals are highly structured i.e. strong neighbor dependencies.
- Measures similarity on the basis of
  - Luminance
  - Contrast
  - Structure

$$S(\mathbf{x}, \mathbf{y}) = l(\mathbf{x}, \mathbf{y}) \cdot c(\mathbf{x}, \mathbf{y}) \cdot s(\mathbf{x}, \mathbf{y}) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \cdot \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \cdot \left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \right),$$


Original Image
Nonstructural Distortions
Structural Distortions
Luminance Change
Contrast Change
Noise Contamination
Blurring
Gamma Distortion
Spatial Shift
JPEG Blocking
Wavelet Ringing

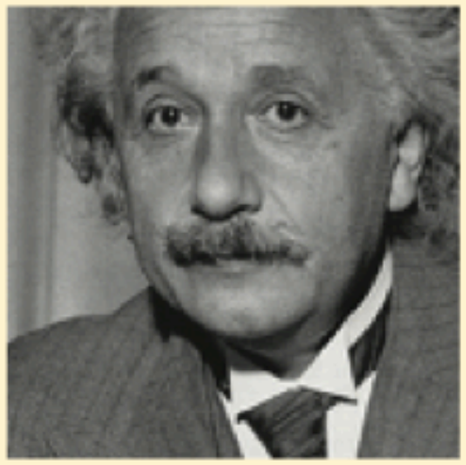
- The SSIM index is symmetric: $S(\mathbf{x}, \mathbf{y}) = S(\mathbf{y}, \mathbf{x})$

- It is also bounded: $-1 < S(\mathbf{x}, \mathbf{y}) \leq 1$, achieving maximum value $S(\mathbf{x}, \mathbf{y}) = 1$ if and only if $\mathbf{x} = \mathbf{y}$.

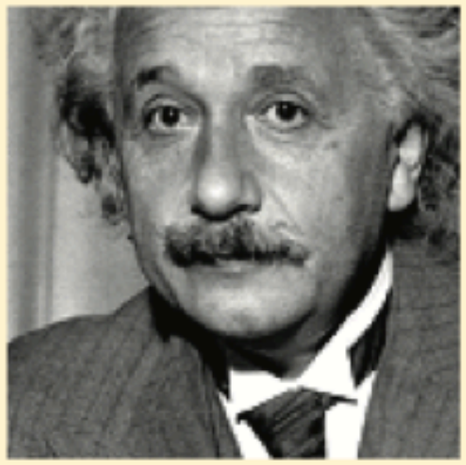
- A drawback of the SSIM index is its sensitivity to relative translations, scalings and rotations of images.

- To handle such situations, a wavelet-domain version of SSIM, complex wavelet SSIM (CW-SSIM) index was developed.

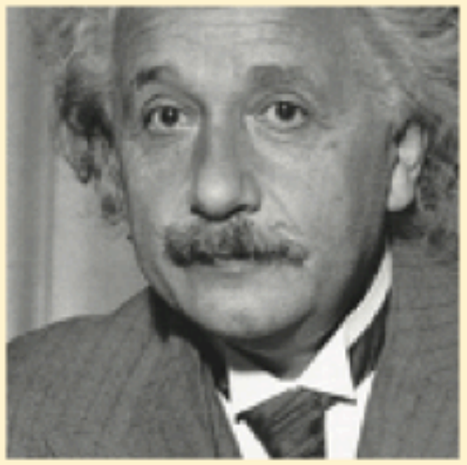
- But they usually work well for grayscale and not colored images.



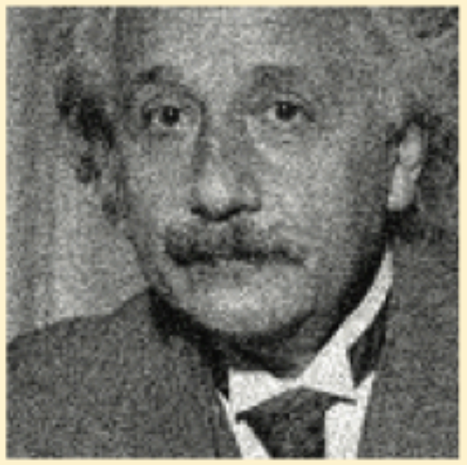
MSE=0, SSIM=1
CW-SSIM=1
(a)



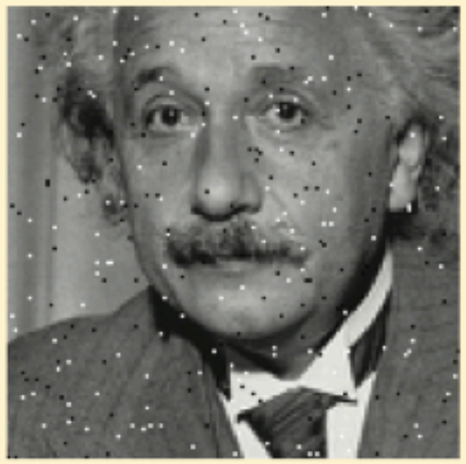
MSE=306, SSIM=0.928
CW-SSIM=0.938
(b)



MSE=309, SSIM=0.987
CW-SSIM=1.000
(c)



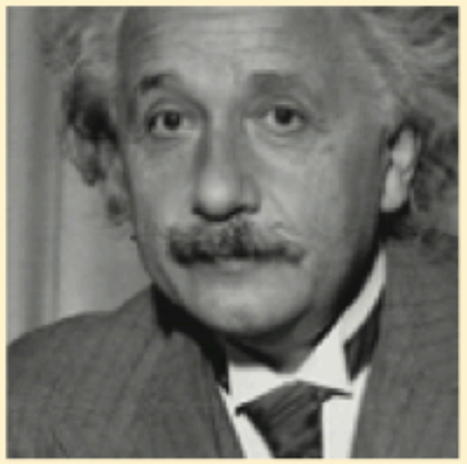
MSE=309, SSIM=0.576
CW-SSIM=0.814
(d)



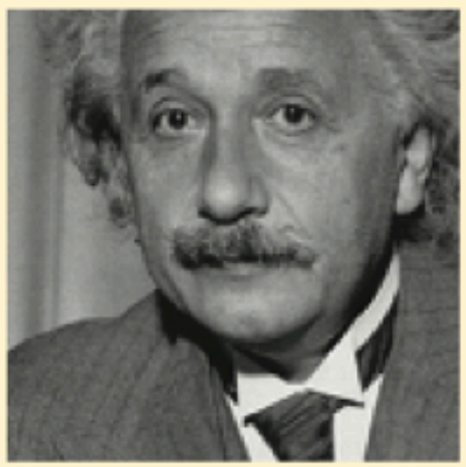
MSE=313, SSIM=0.730
CW-SSIM=0.811
(e)

MSE=309, SSIM=0.580
CW-SSIM=0.633
(f)

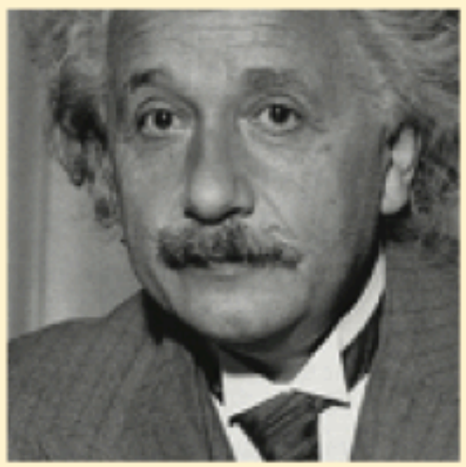
MSE=308, SSIM=0.641
CW-SSIM=0.603
(g)



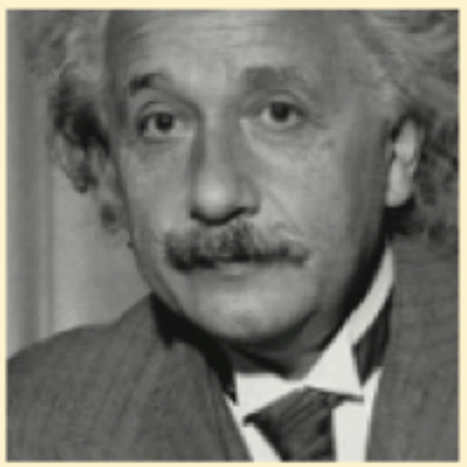
MSE=694, SSIM=0.505
CW-SSIM=0.925
(h)



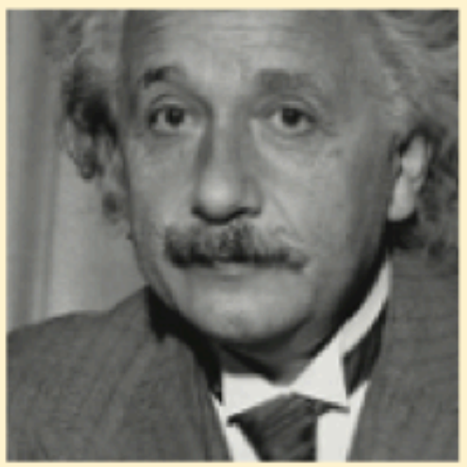
MSE=871, SSIM=0.404
CW-SSIM=0.933
(i)



MSE=873, SSIM=0.399
CW-SSIM=0.933
(j)

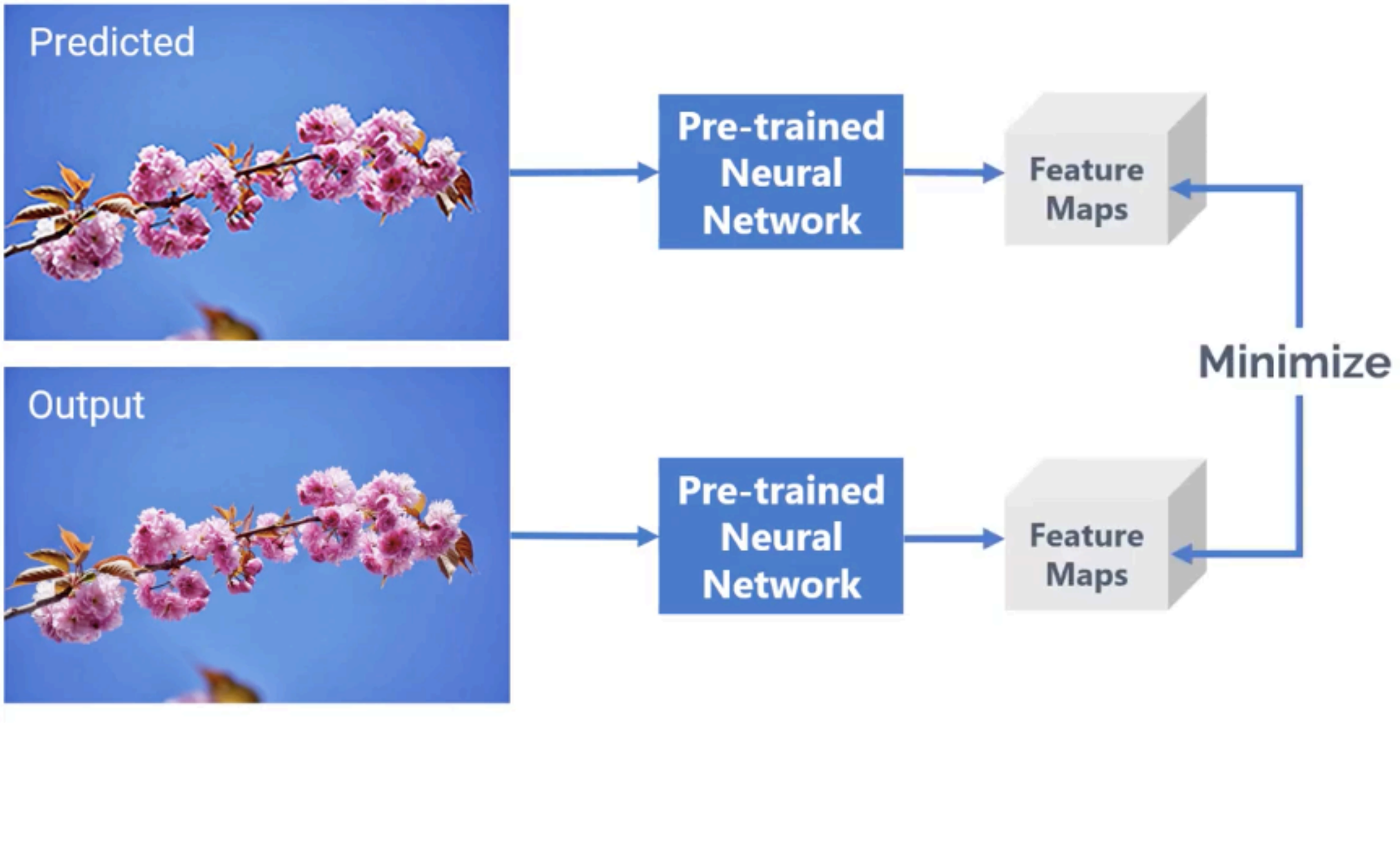


MSE=590, SSIM=0.549
CW-SSIM=0.917
(k)

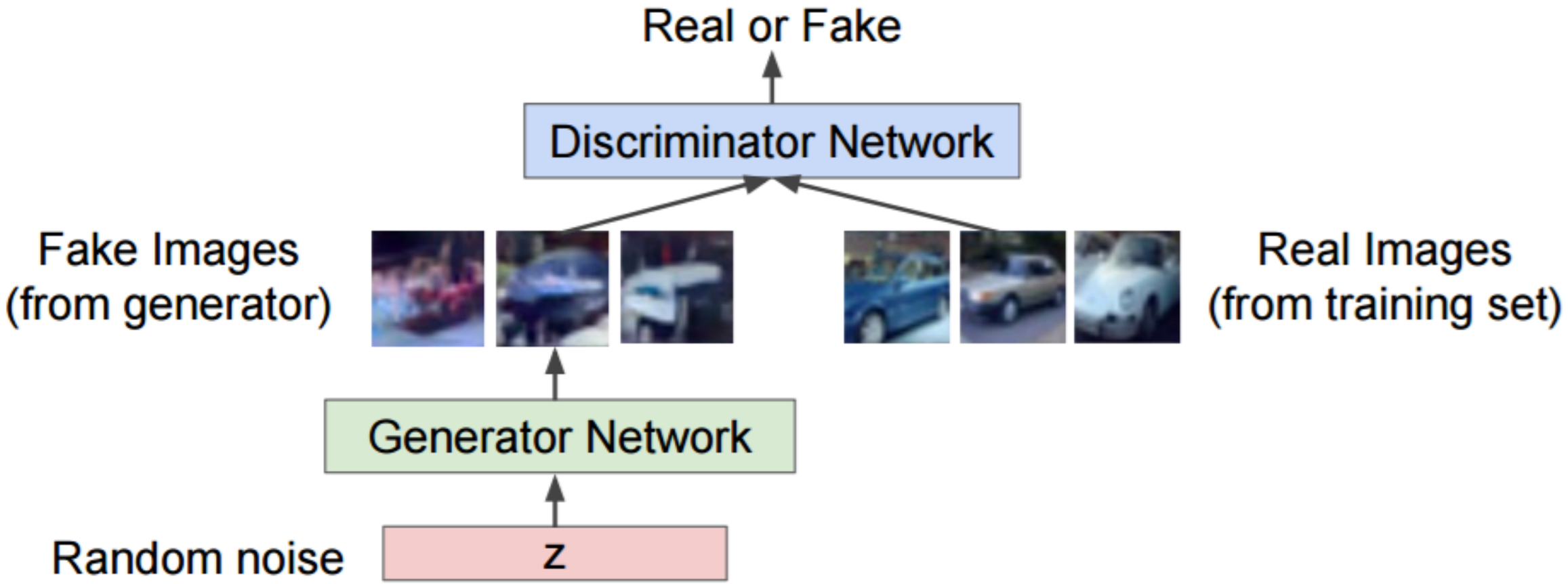


MSE=577, SSIM=0.551
CW-SSIM=0.916
(l)

Perceptual loss



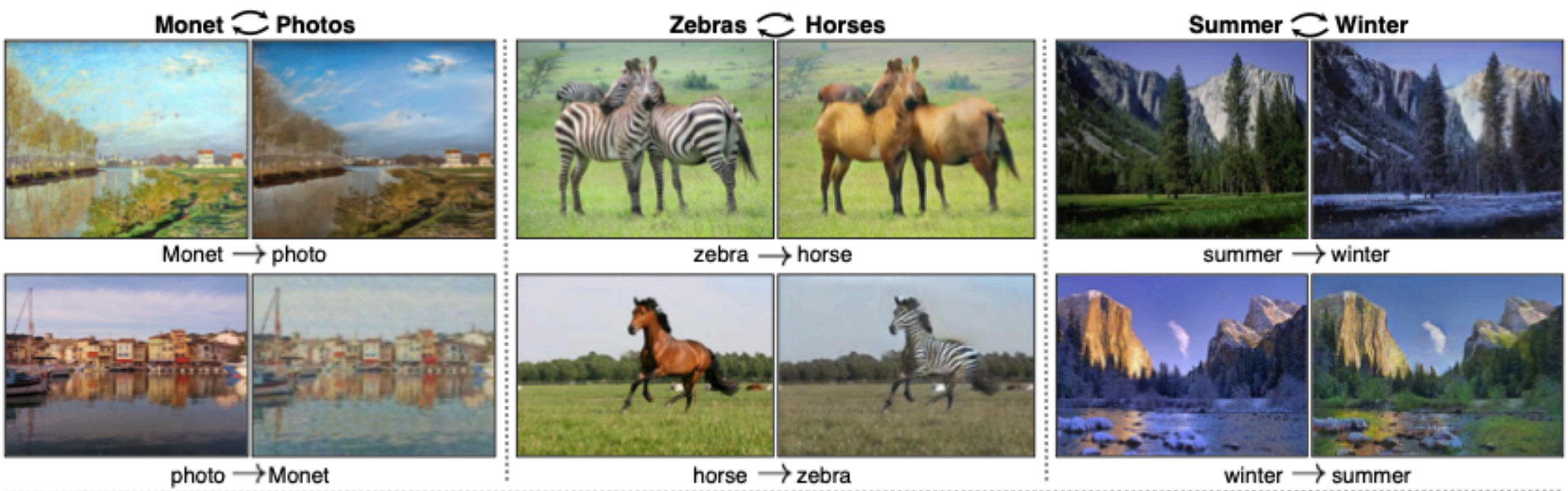
Adversarial loss



CycleGAN

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Jun-Yan Zhu* Taesung Park* Phillip Isola Alexei A. Efros
Berkeley AI Research (BAIR) laboratory, UC Berkeley



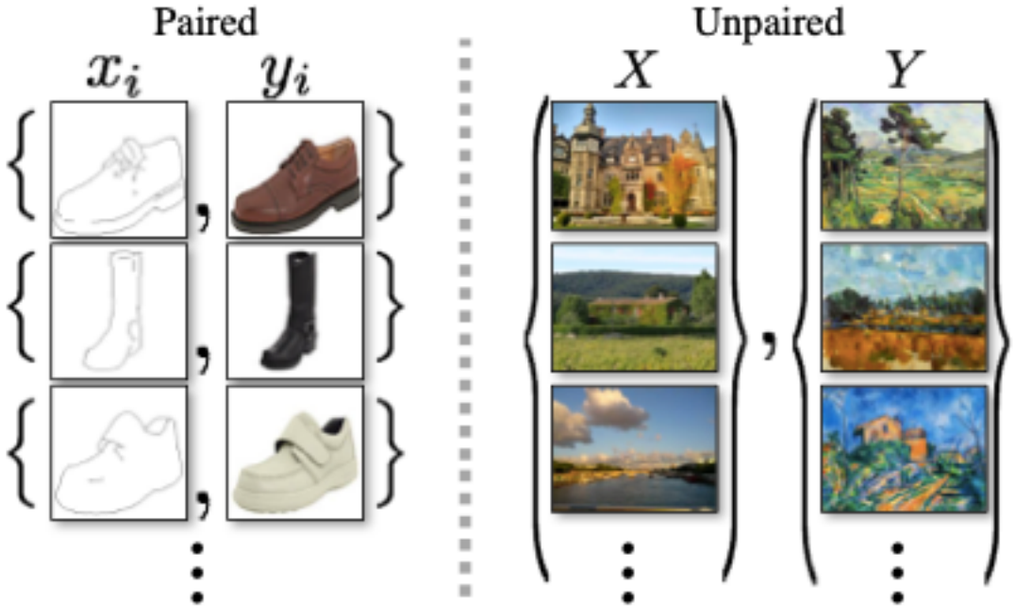
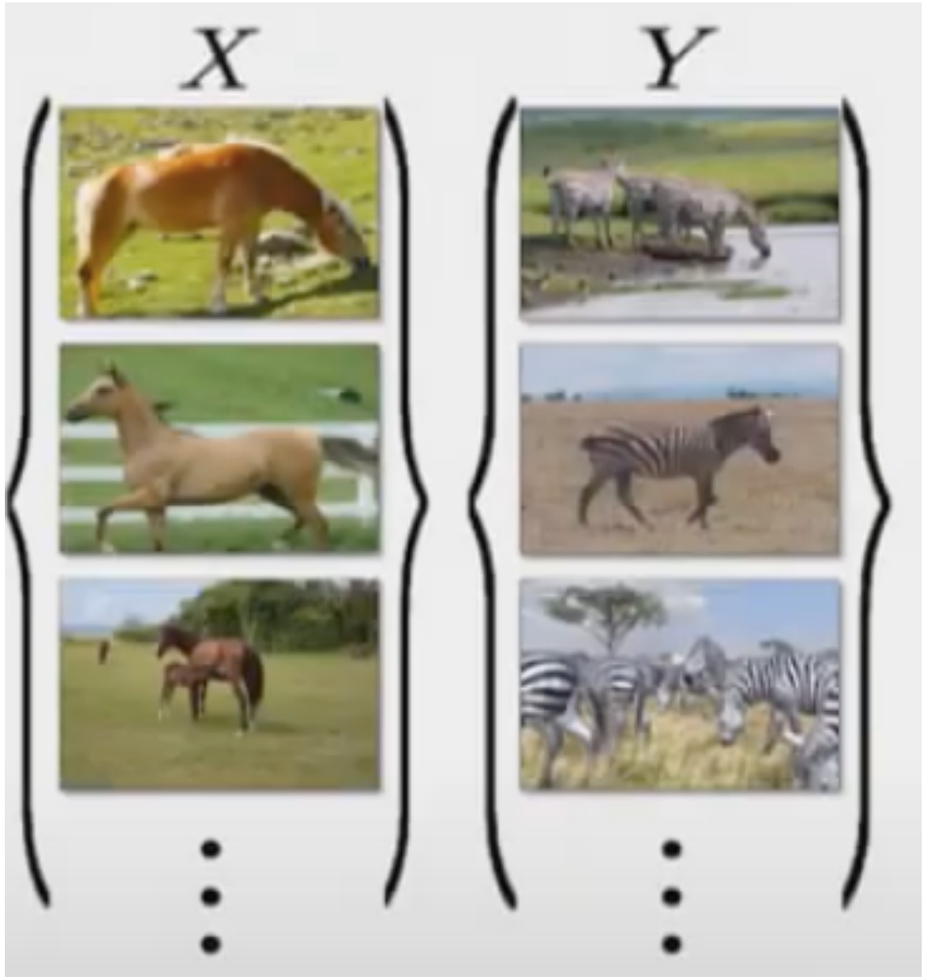
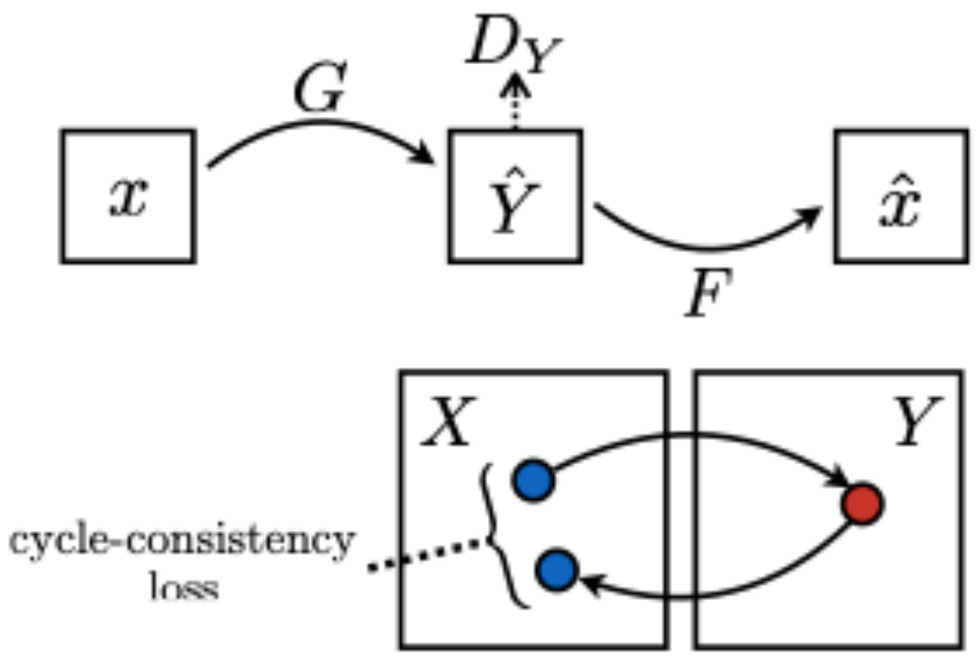
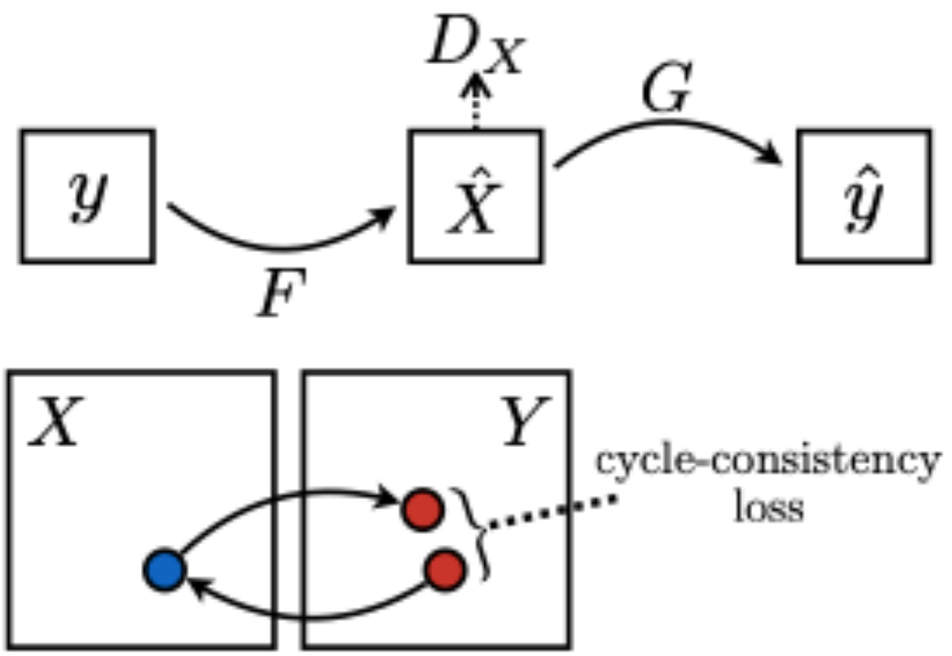


Figure 2: *Paired* training data (left) consists of training examples $\{x_i, y_i\}_{i=1}^N$, where the correspondence between x_i and y_i exists [22]. We instead consider *unpaired* training data (right), consisting of a source set $\{x_i\}_{i=1}^N$ ($x_i \in X$) and a target set $\{y_j\}_{j=1}$ ($y_j \in Y$), with no information provided as to which x_i matches which y_j.







Problem with GAN

- Hard to train
  - Slow
  - Unstable

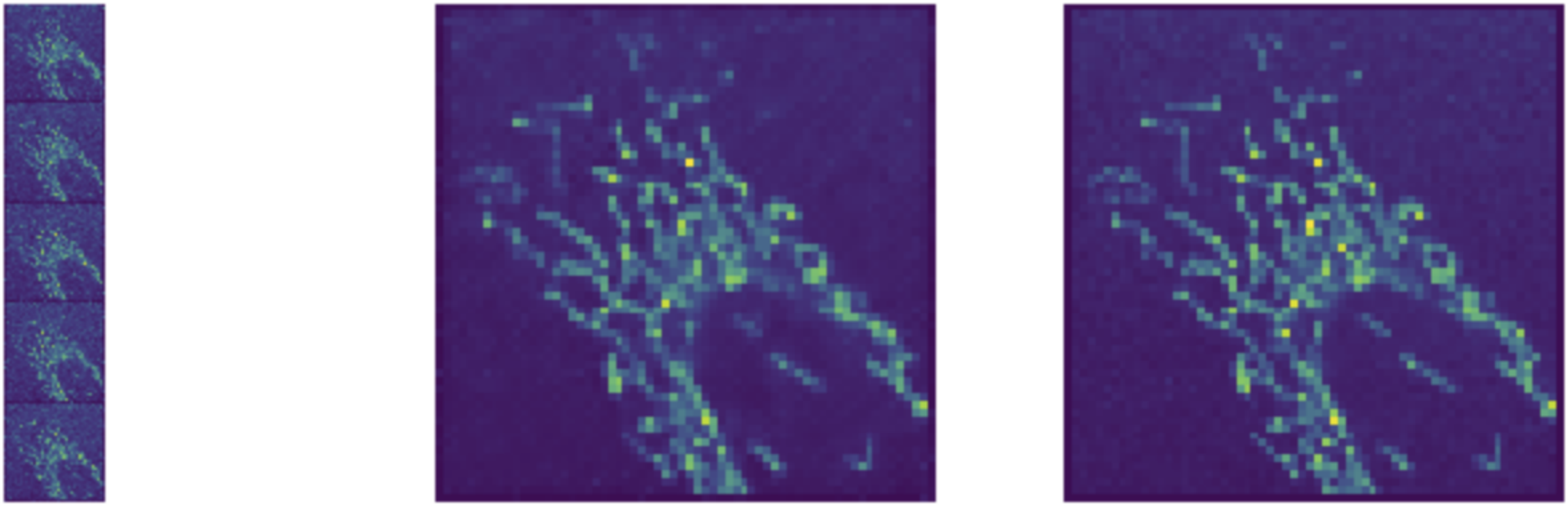
- May lead to unrealistic results (Hallucination)

Solution


Pre-Training
GAN Training
Hours Training
0
5
10
15
20
25
30
35
40
45
50

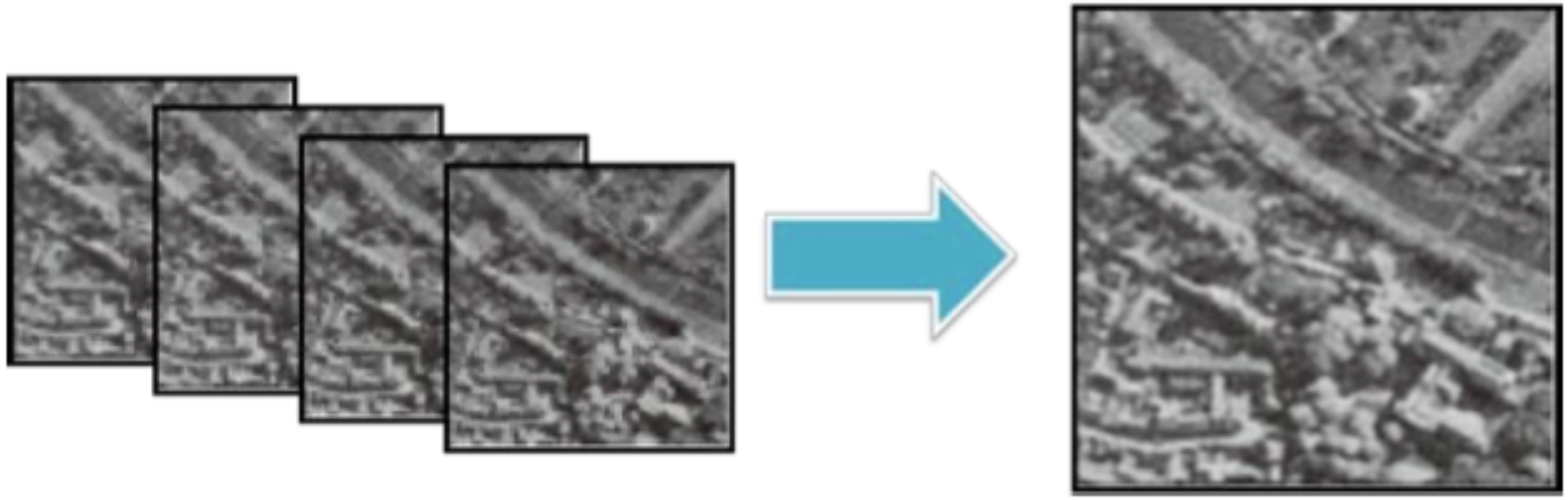
Applications

- Medical imaging



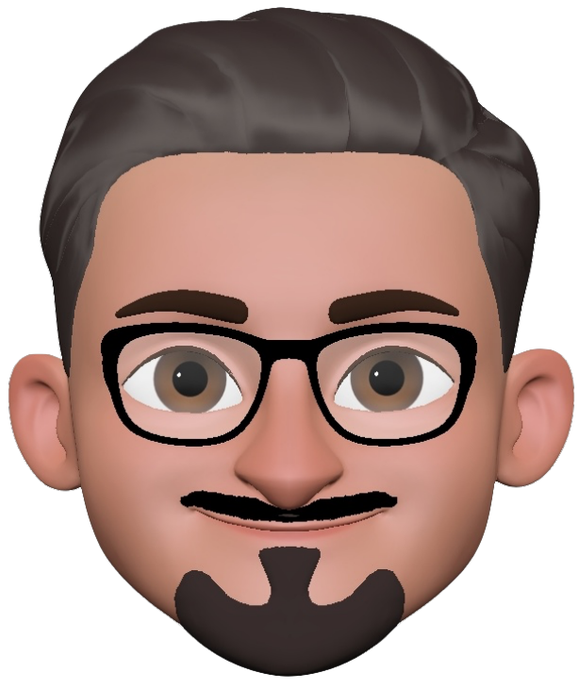
Applications

- Satellite imaging



Applications

- DeOldification/Colorization



AMAN AGARWAL

Software Developer @ HSBC | Intel Software Innovator of AI

+91 75740 14476 | amanag.11@gmail.com | amanbasu.github.io/portfolio/

Aman Agarwal amanbasu @TheAbecedarian_

References

- Structural similarity index
- How super resolution works
- Deep learning based super resolution
- Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks