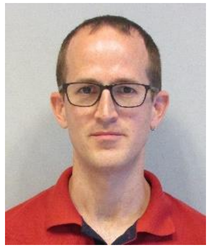
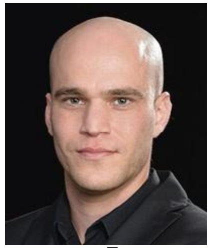
Deep Internal Learning

Assaf Shocher [1]

Collaborators:


Yossi Gandelsman [1]


Shai Bagon [1]


Nadav Cohen [2]


Phillip Isola [3]


Michal Irani [1]

1 מכון ויצמן למדע WEIZMANN INSTITUTE OF SCIENCE

2 INSTITUTE FOR ADVANCED STUDY

3 Massachusetts Institute of Technology

Witl

* Imag

[Efros ...

[Wexl

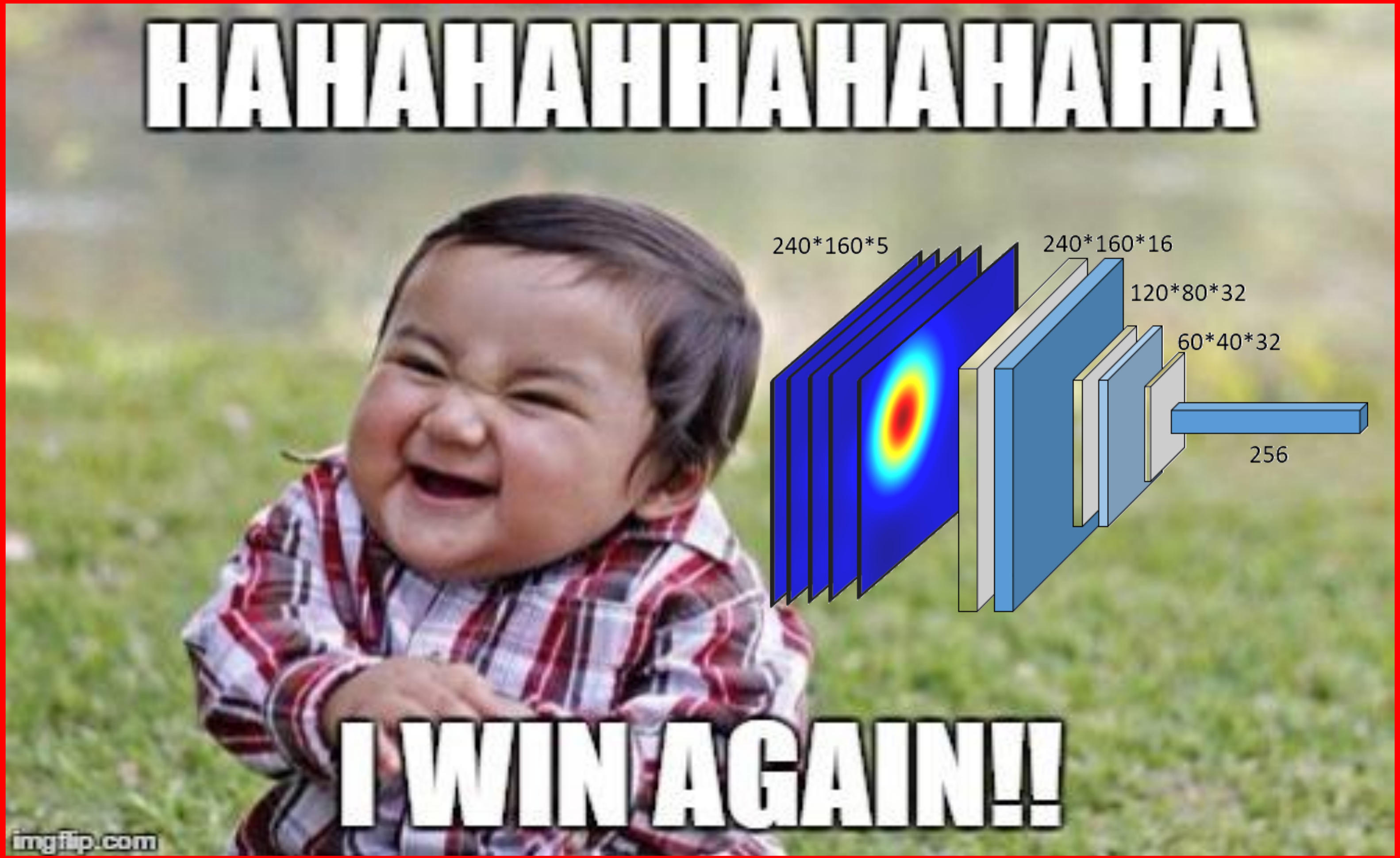
Irani '0 09],

* Imag ...

NLM, l

14]

* Visu



Outperformed by Deep-Learning

Main idea:

Deep learning + **Internal statistics**

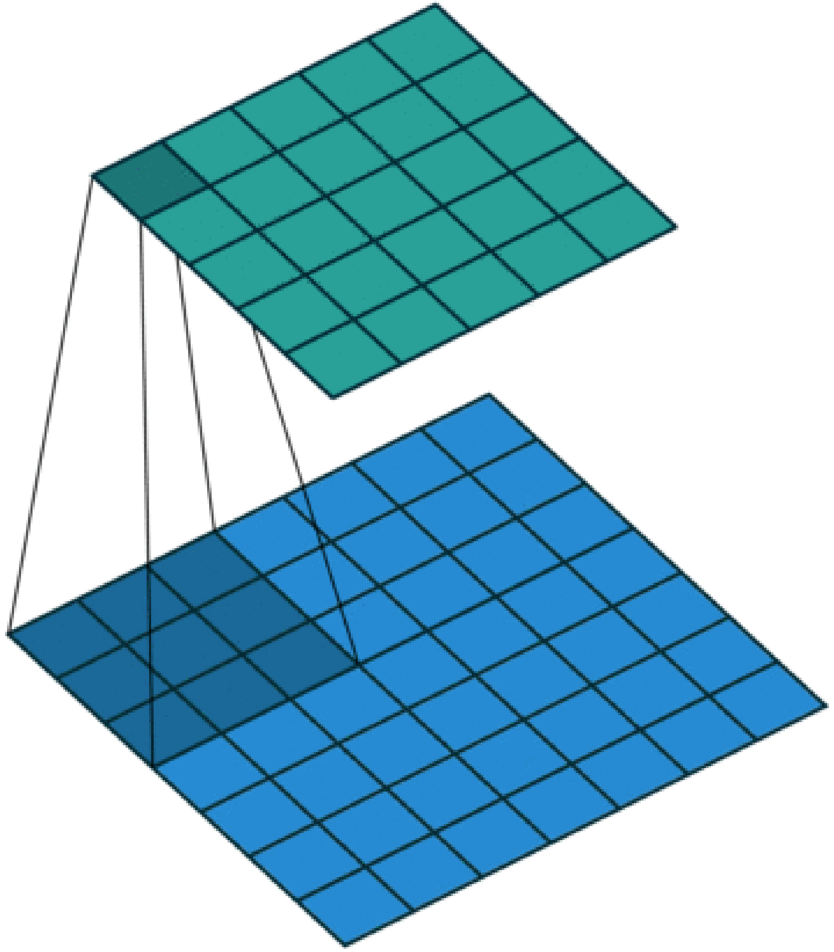
Deep Internal Learning

We train an Image-Specific CNN

At test time

On the test image only

A single image contains tons of data!



- A fully-conv net is actually learning from a <u>“bag of patches”</u>
- # of patches ≈ # of pixels of an image
- So, an image is a huge <u>batch</u> of many patch example pairs

How can a network be trained at test time?

Inside | **Outside**

[Zontak&Irani;2011]

Inside	Outside
Low entropy	High entropy
↓	↓
Light network	Heavy network
↓	↓
Fast training	Slow training

For SR:
train+inference time =
9 sec [V-100],
54 sec [K-80]

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

“Zero-Shot” Super-Resolution using Deep Internal Learning

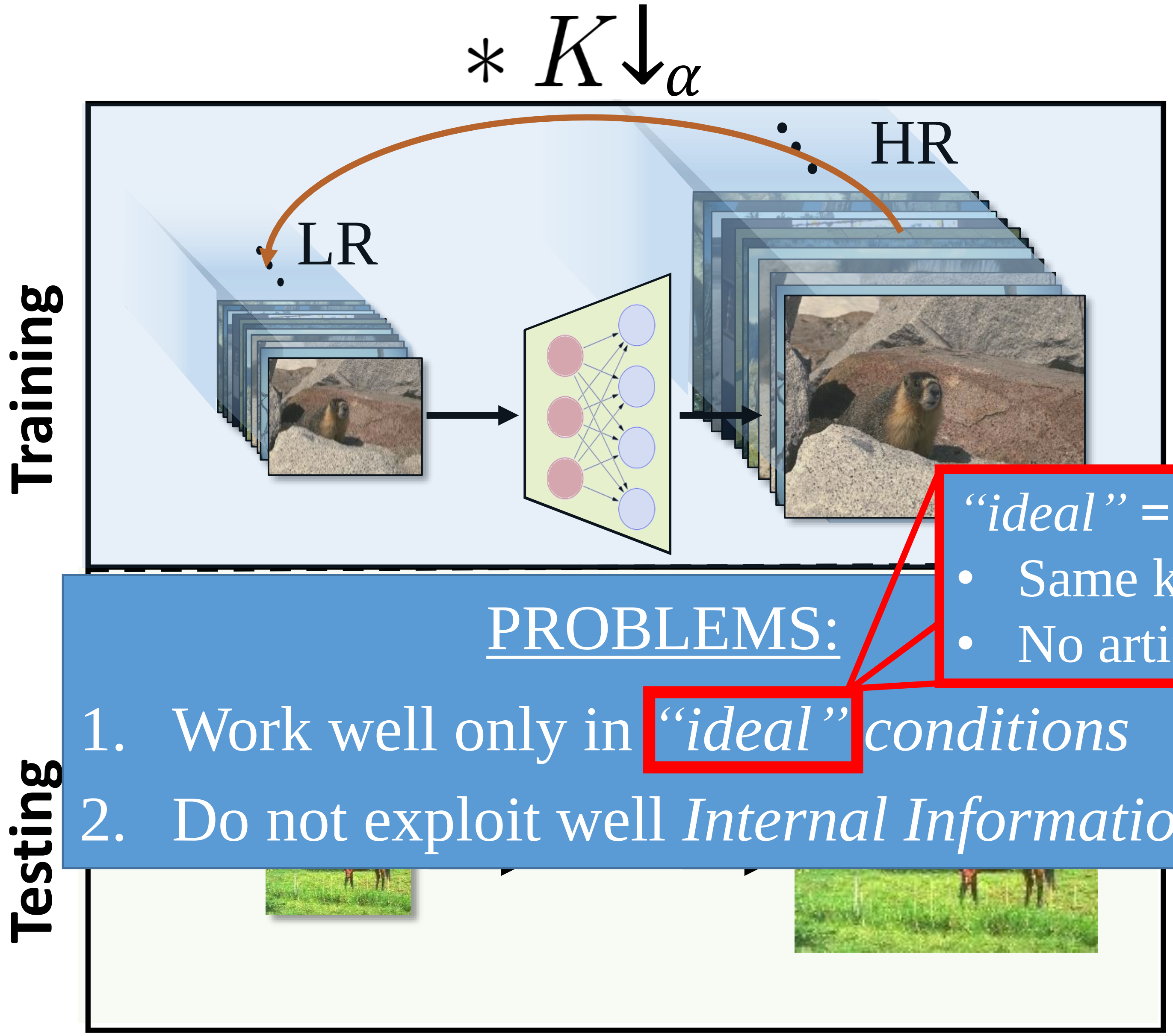
Assaf Shocher
Weizmann

Nadav Cohen
IAS

Michal Irani
Weizmann

CVPR’18

Prior Work: External CNN-based SR



SRCNN; Dong et al, 2014

VDSR; Kim et al, 2016

EDSR+; Lim et al, 2017
+2 dB improvement

PROBLEMS:

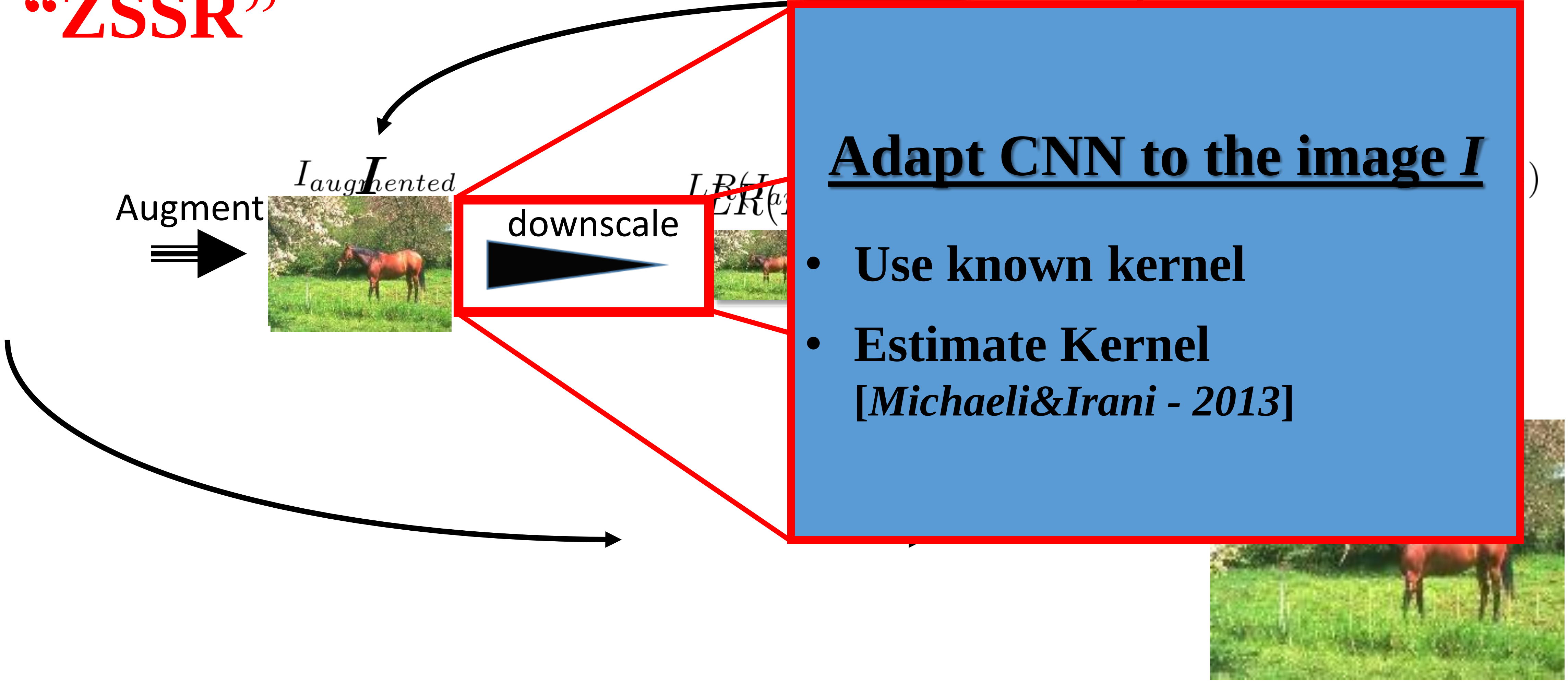
1. Work well only in *"ideal" conditions*
2. Do not exploit well *Internal Information*

"ideal" =
- Same kernel
- No artifacts

“Zero-Shot” SR using Deep Internal Learning

$$\text{Loss} = \left\| SR\left(LR(I)\right) - I \right\|_1$$

“ZSSR”



Adapt CNN to the image *I*

- **Use known kernel**
- **Estimate Kernel**
  [*Michaeli&Irani - 2013*]

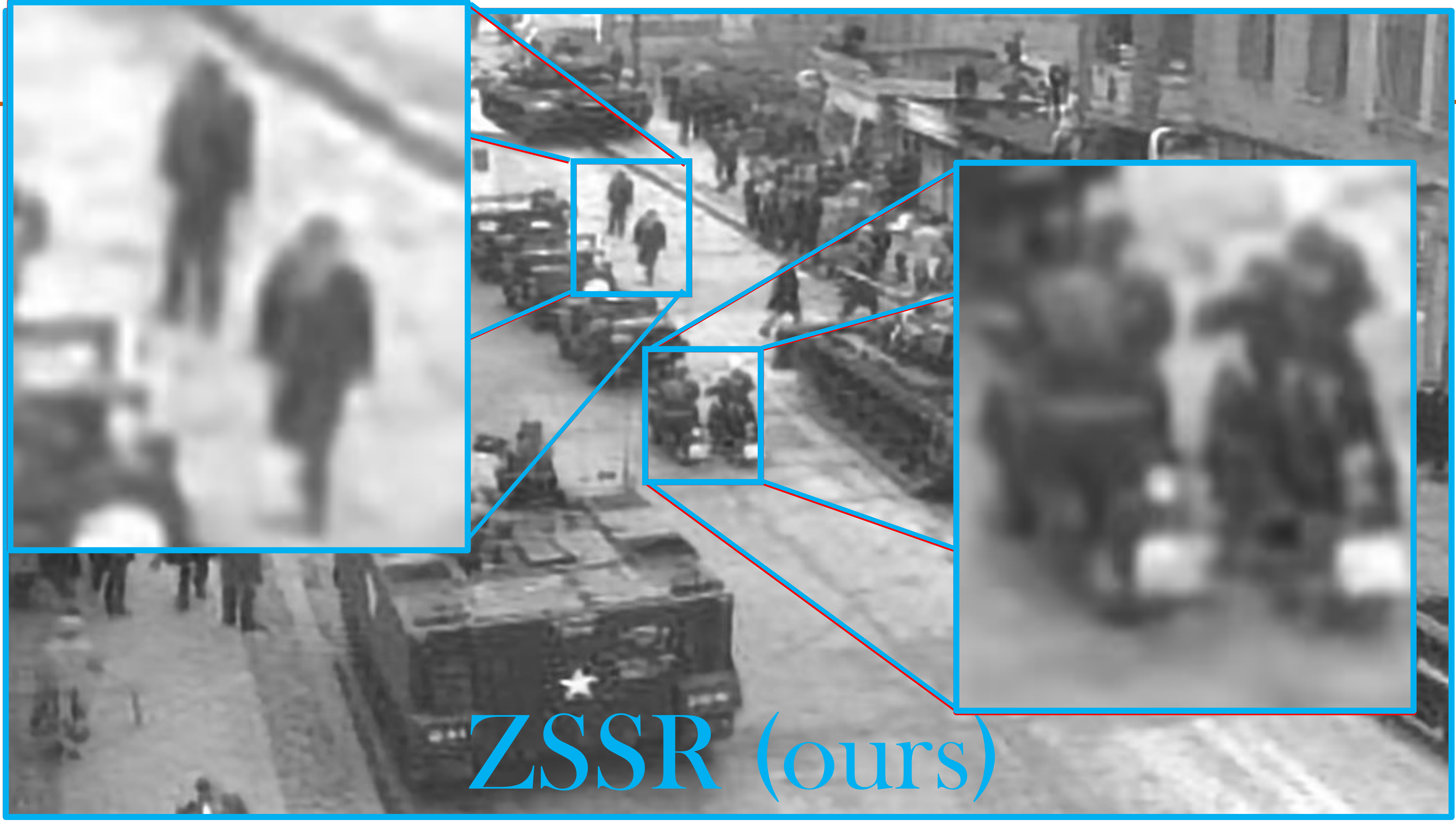
Int… …le

✓ **Any dat…**

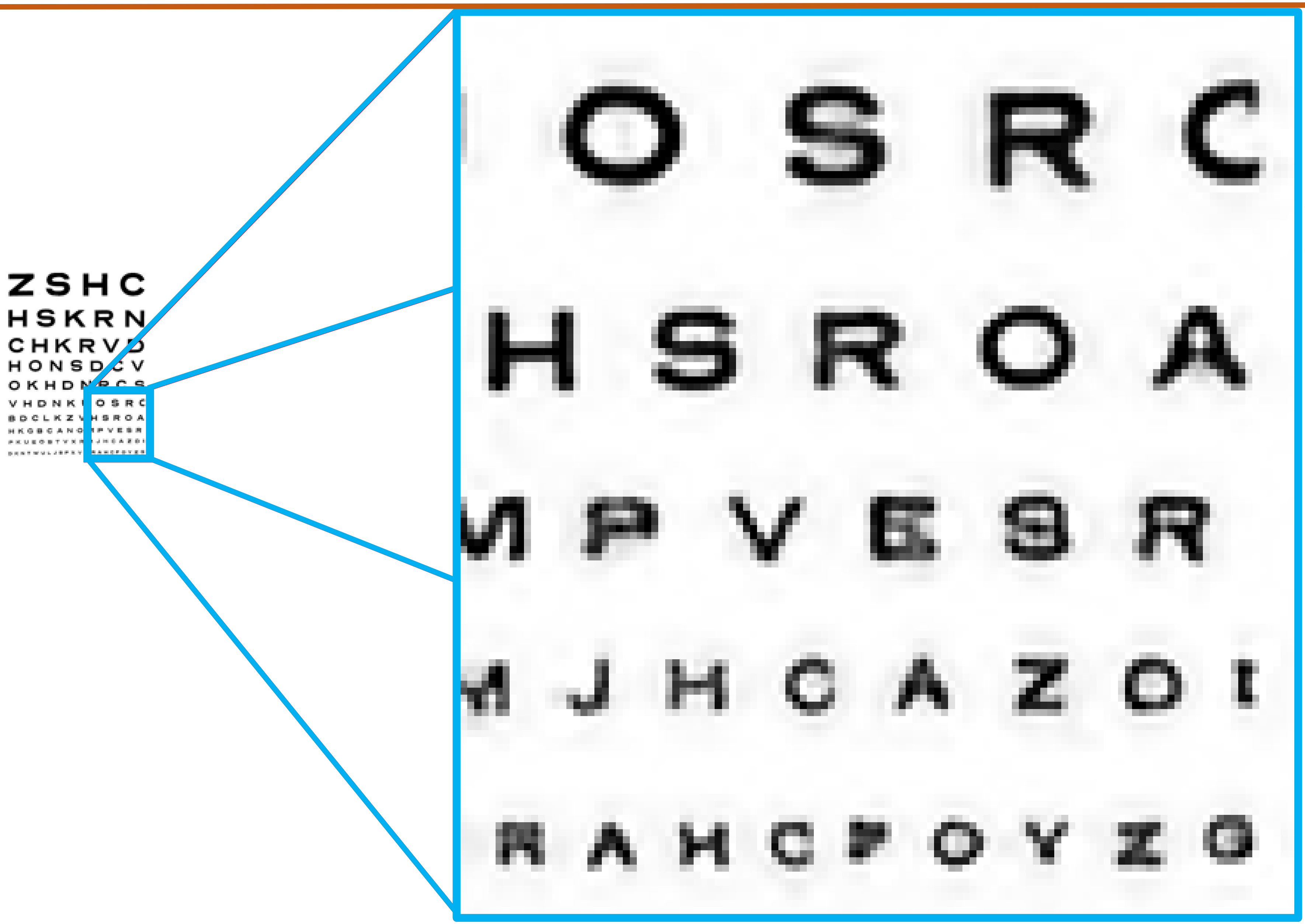
✓ **Any dov…**

✓ **Unknov…**

✓ **To any s…**


ZSSR (ours)

“Ideal” case



EDSR+

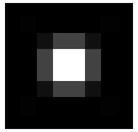
ZSSR (ours)

PSNR

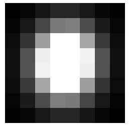
EDSR+: 25.29 dB

ZSSR: 25.68 dB

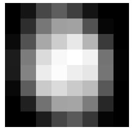
Non-ideal downscaling kernels



"ideal" (bicubic) kernel



The true (unknown) kernel



Estimated Kernel (via Blind-SR)

PSNR

EDSR+: 26.42 dB
Blind-SR: 27.29 dB
ZSSR: 29.42 dB

Non-ideal downscaling kernels

Poor-quality LR images

PSNR

EDSR+: 24.91 dB
ZSSR: 26.25 dB

Poor-quality LR images

PSNR

EDSR+: 35.88 dB
ZSSR: 37.80 dB

Empirical evaluations

- On “ideal” images → reasonably well
- On ***non-ideal*** images → 1dB - 2dB better than SotA

“ InGAN”: Capturing and Replicating the “DNA” of a Natural Image

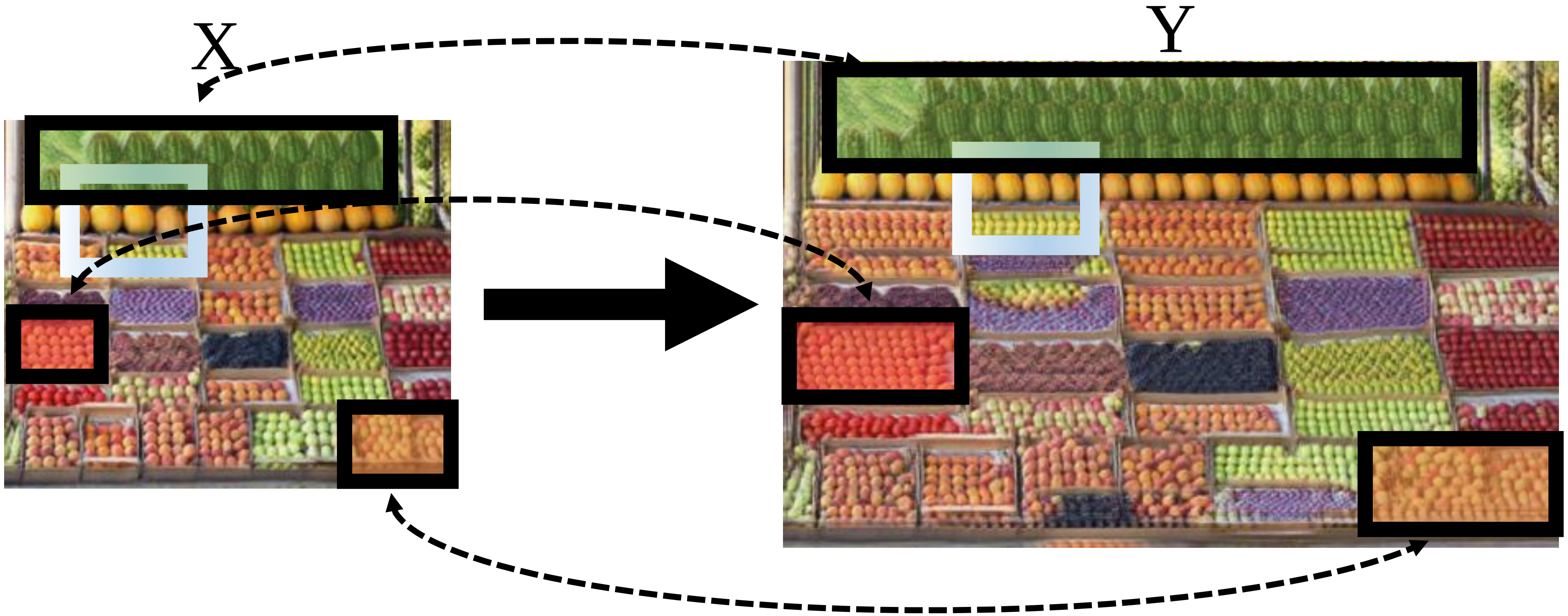
Assaf Shocher,[1] Shai Bagon,[1] Phillip Isola,[2] Michal Irani [1]

[1]Weizmann Institute of Science [2]Massachusetts Institute of Technology

Goal

Preserve various element sizes and relative locations

Our approach:



- Bidirectional similarity (Simakov Et al.)
- Distribution matching: $P(X) \approx P(Y)$ $\boxed{P(X|loc) \approx P(Y|loc)}$
- In all scales
- Preserve relative locations

Local

Of patches in all scales

How can we match distributions?

GANs

(Goodfellow et al. 2014)

Deep Internal Learning

(Shocher et al. 2018)

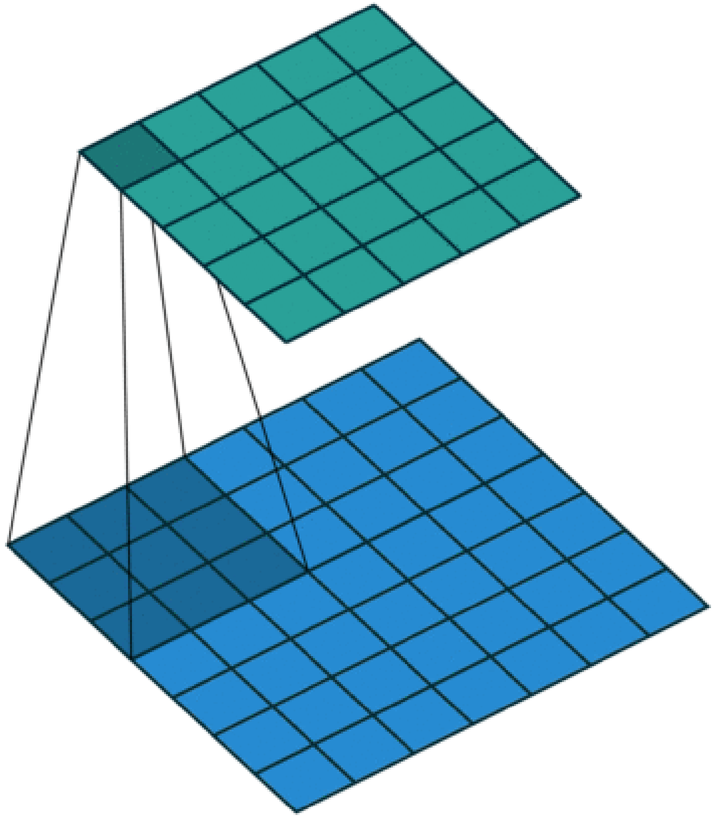
Deep Internal Learning:

On image patches?

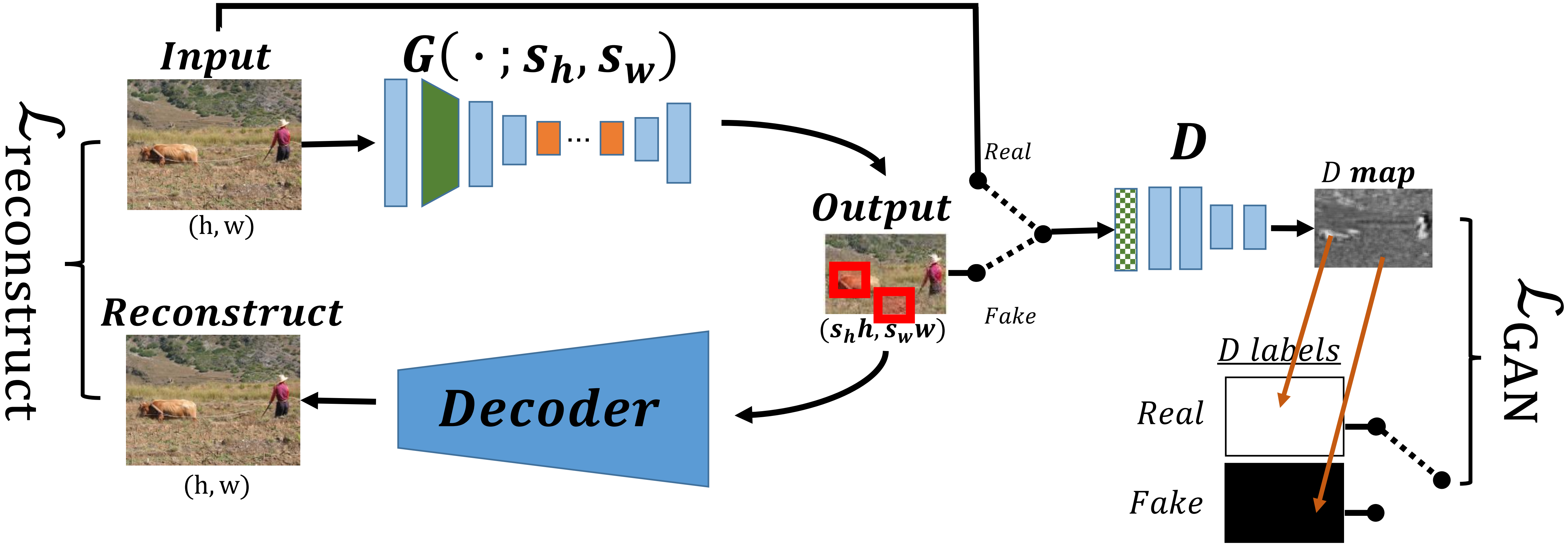
Tons of data in one image

Internal GA

We train a GAN on a singl

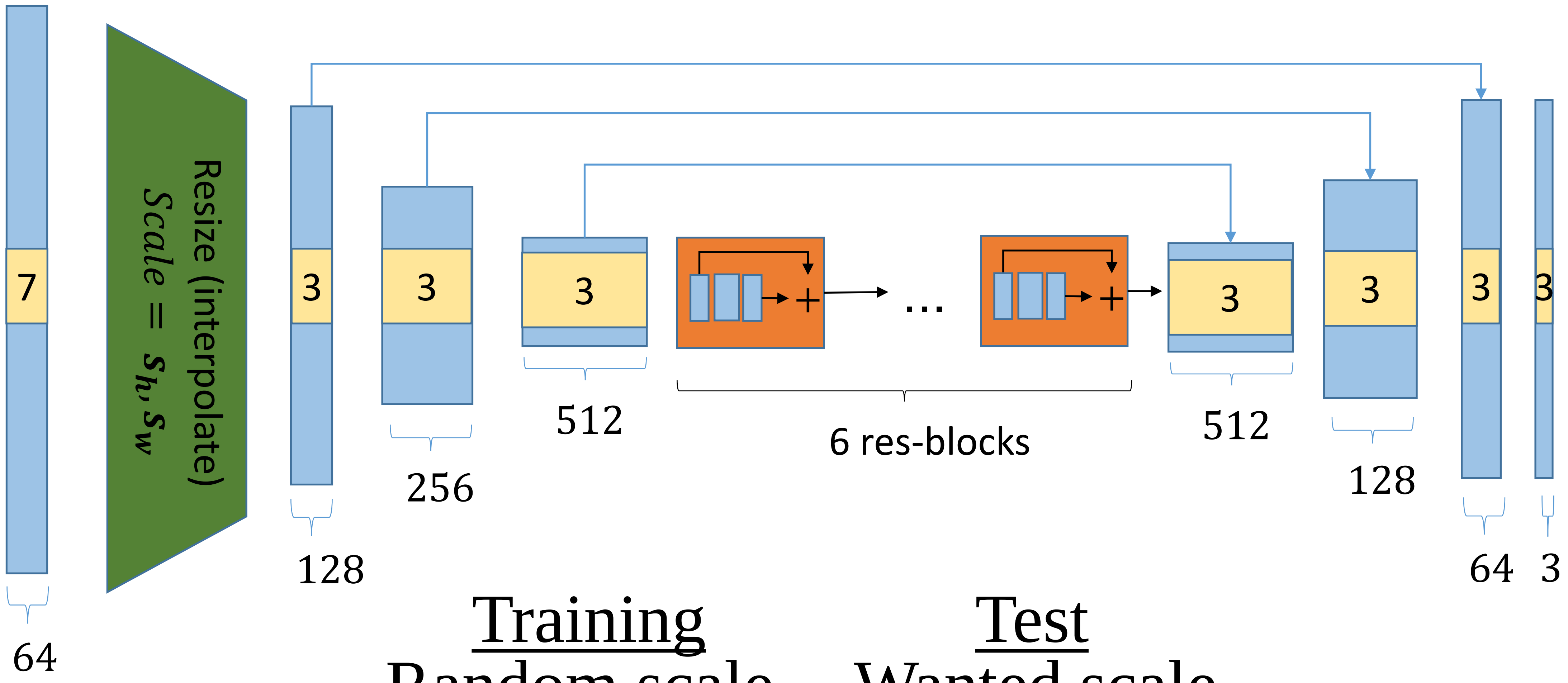


InGAN:



$$\mathcal{L} = \mathcal{L}_{\text{GAN}} + \lambda \mathcal{L}_{\text{reconstruct}}$$

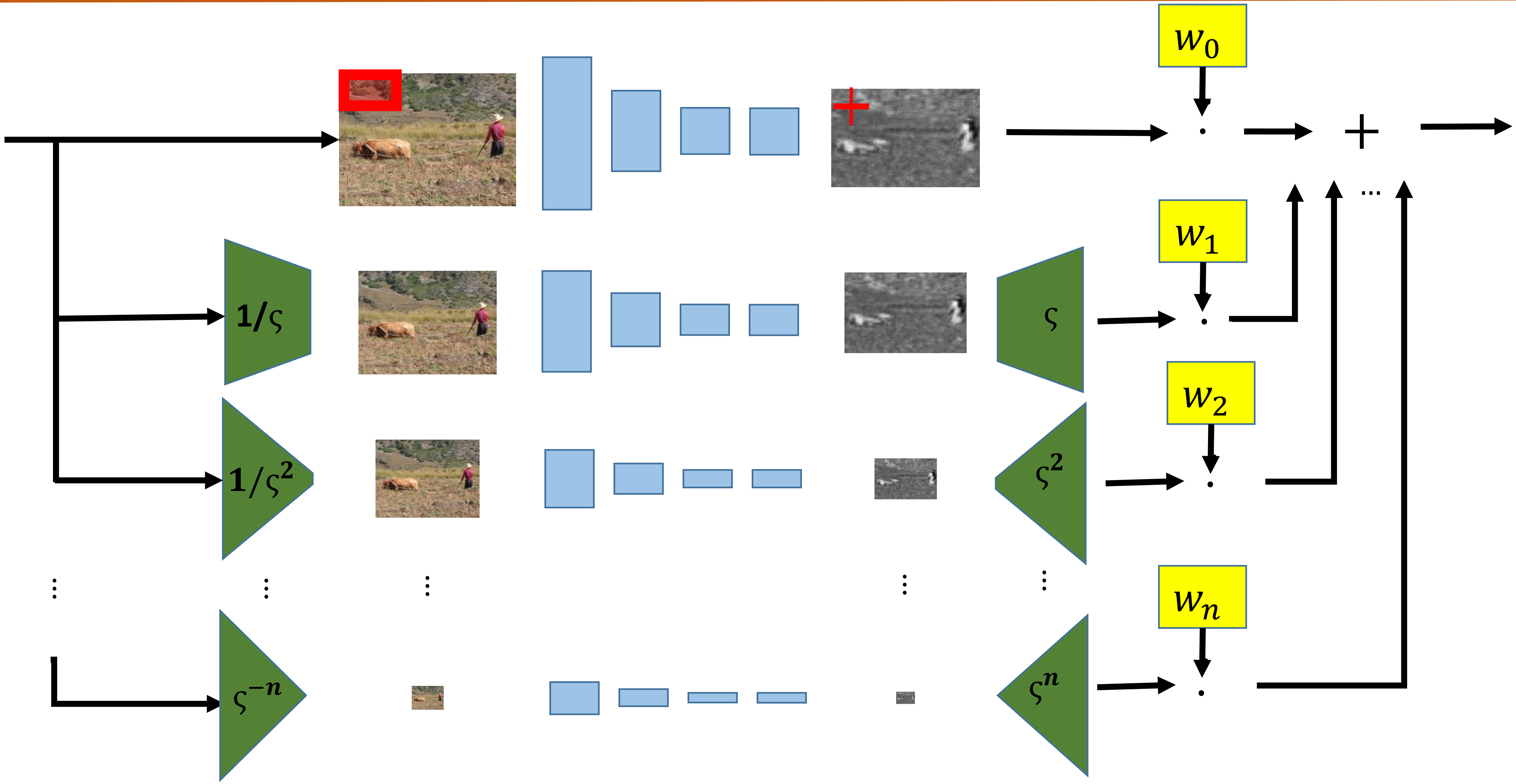
Generator Architecture:



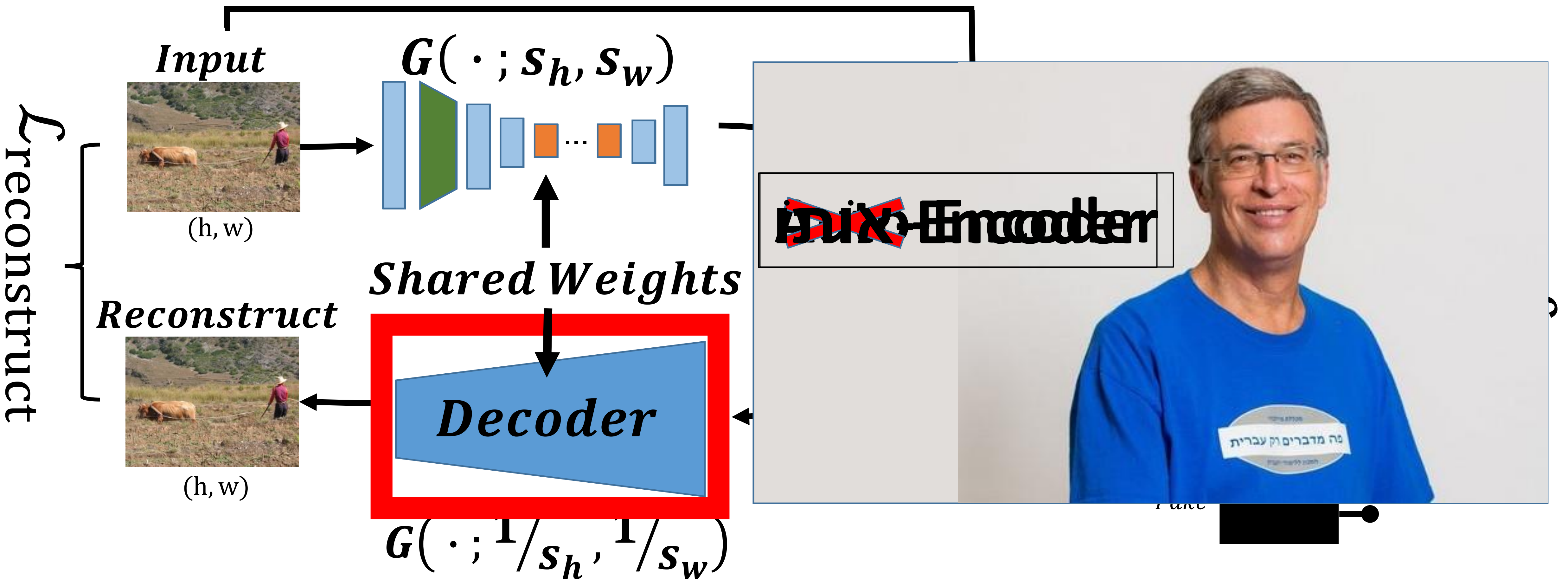
<u>Training</u>
Random scale

<u>Test</u>
Wanted scale

Adaptive Multiscale Patch Discriminator:



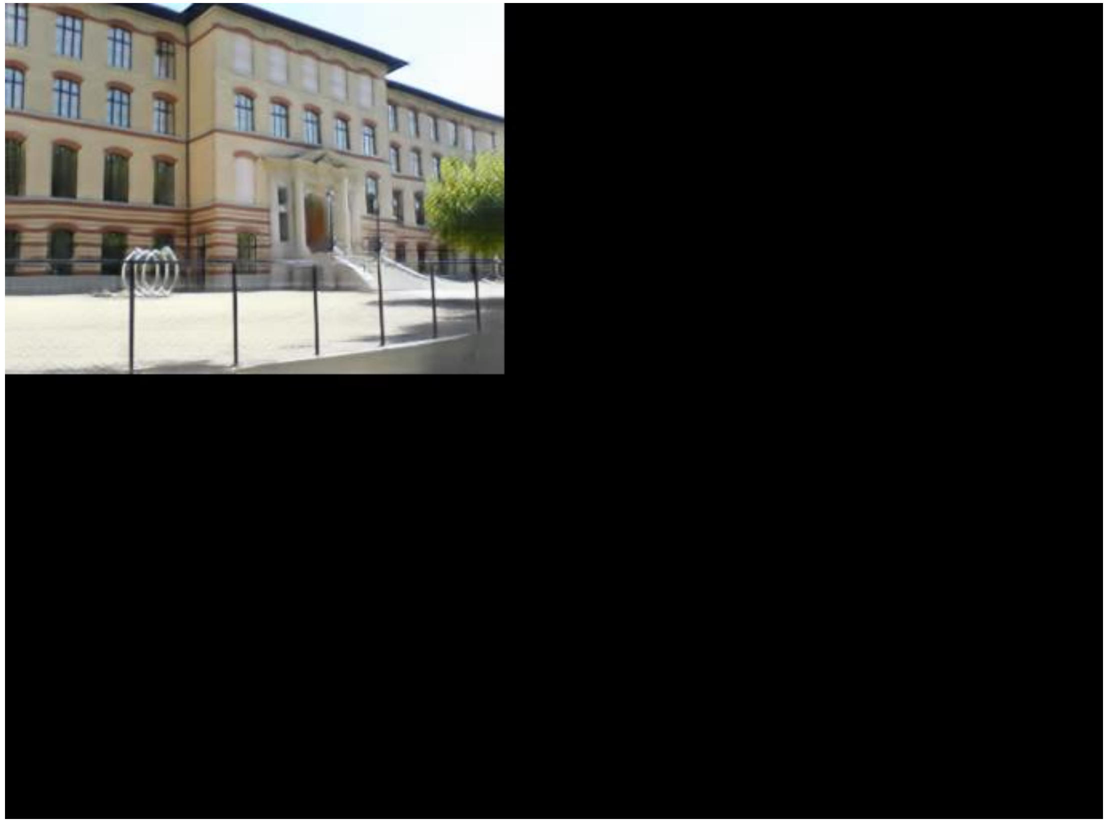
Invertible Generator:



Results:

Input

Results:



Results:

×2 Comparison:

InGAN

(ours)

Bidirectonal Similarity

(Simakov Et al.)

Seam Carving

(Avidan&Shamir)

Non-Stationary Texture-Synthesis

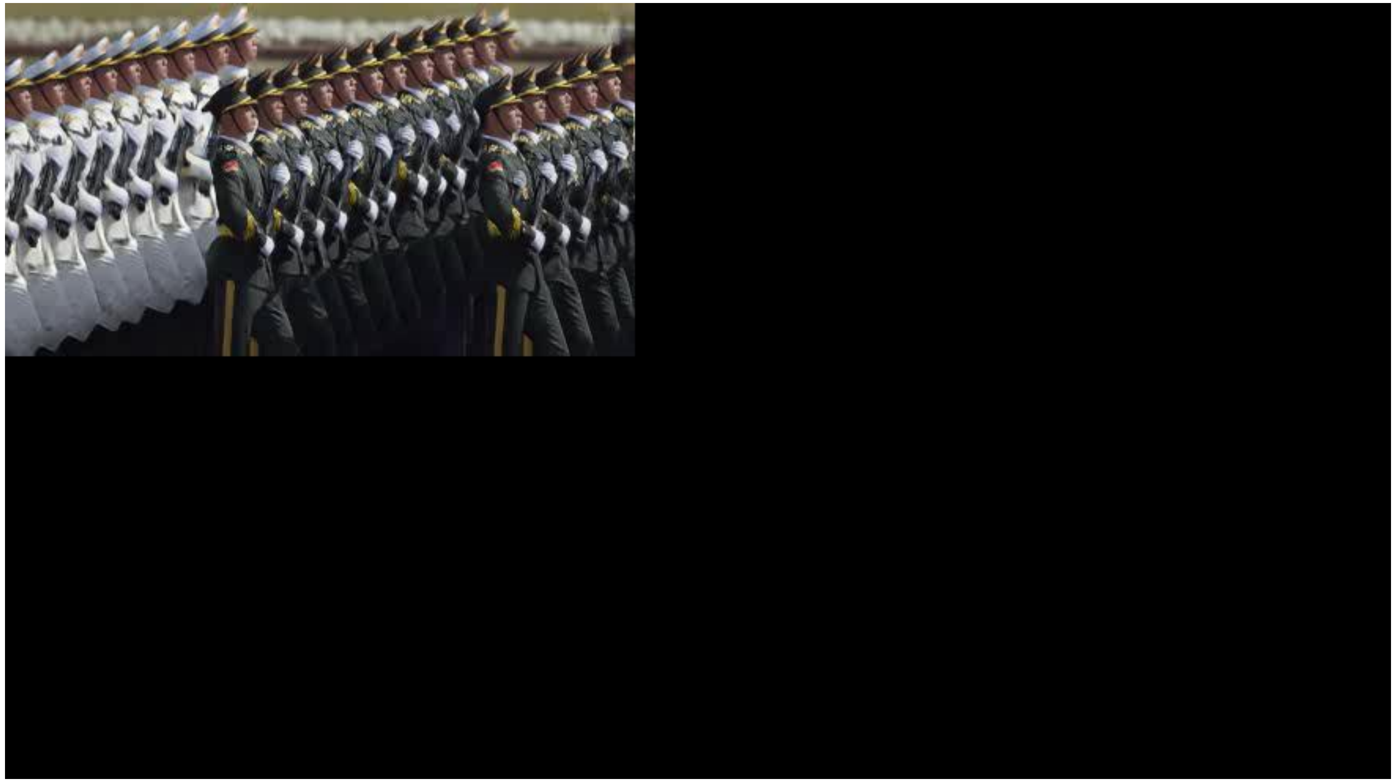
(Yang Et al.)

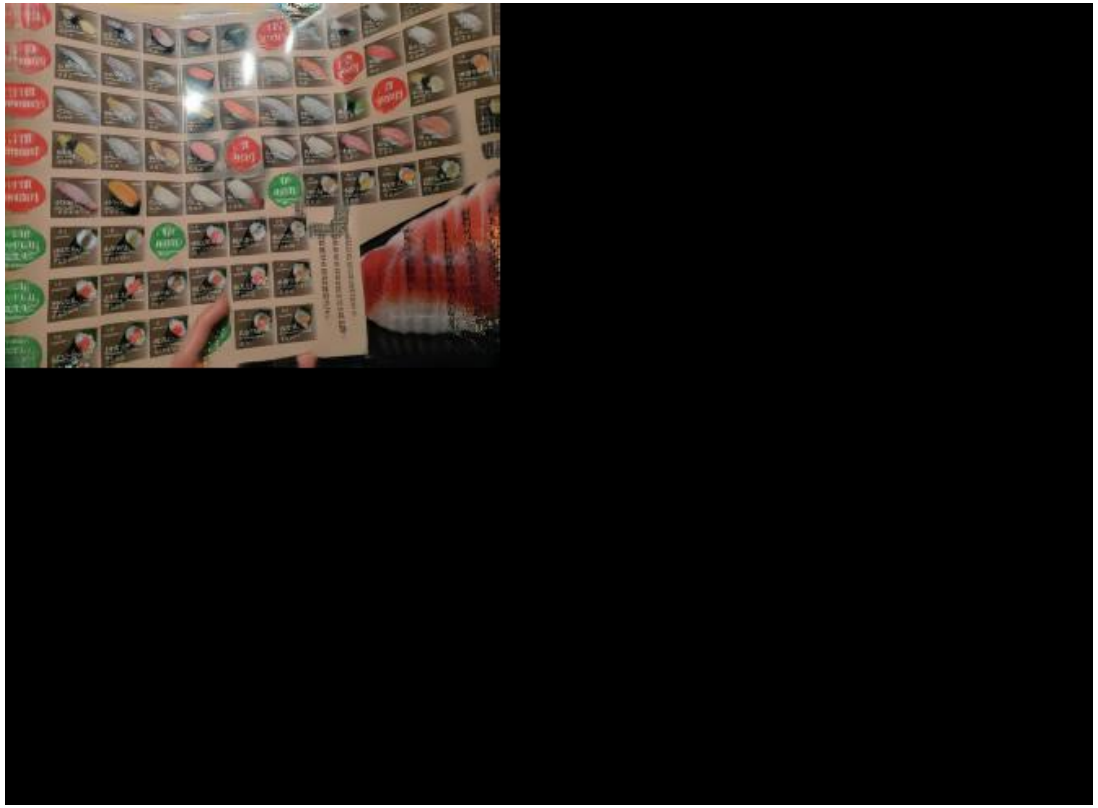
Comparison:

BiDir *SC* ***InGAN** (ours)*

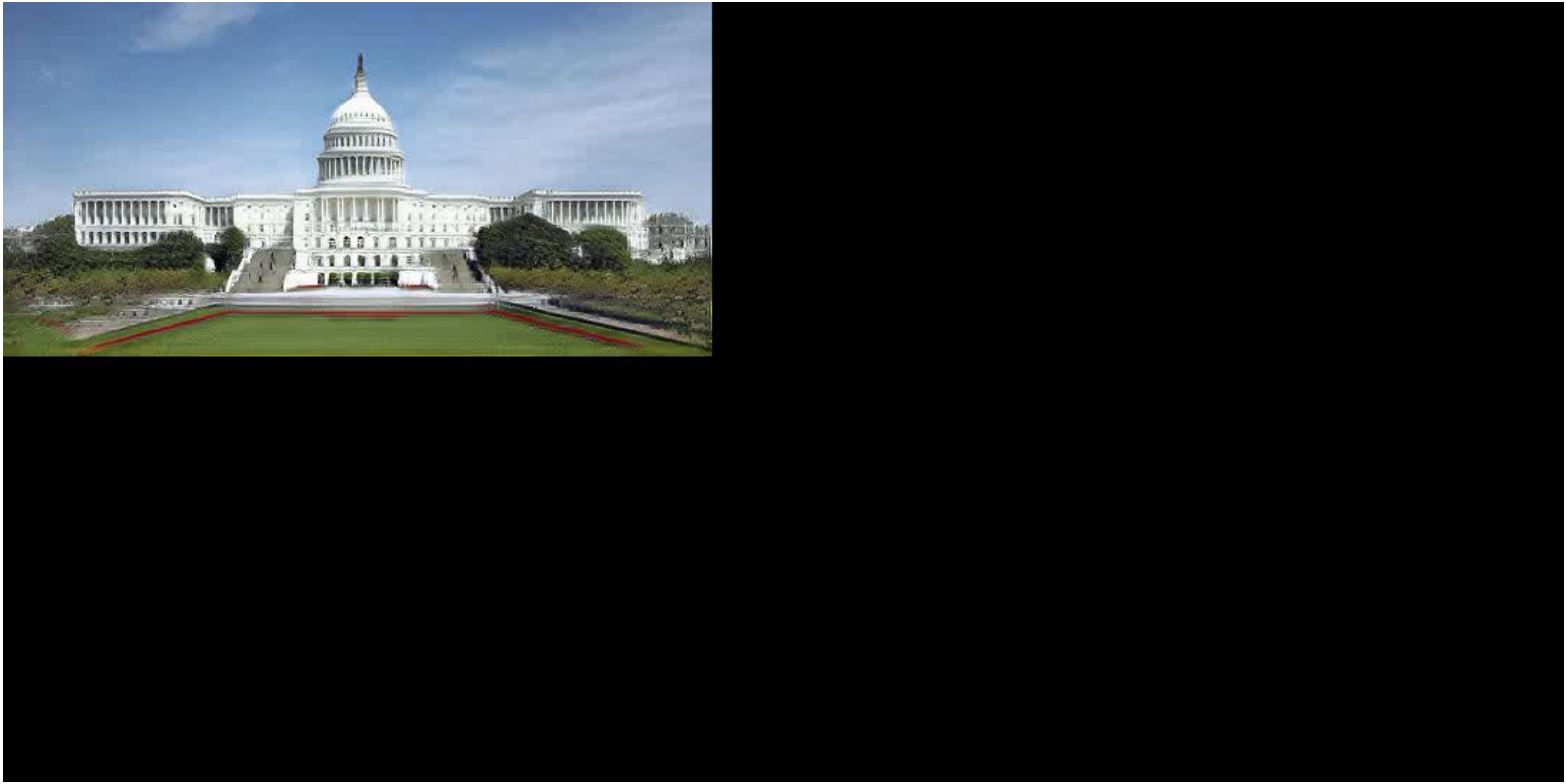
input BiDir SC **InGAN**

Results:

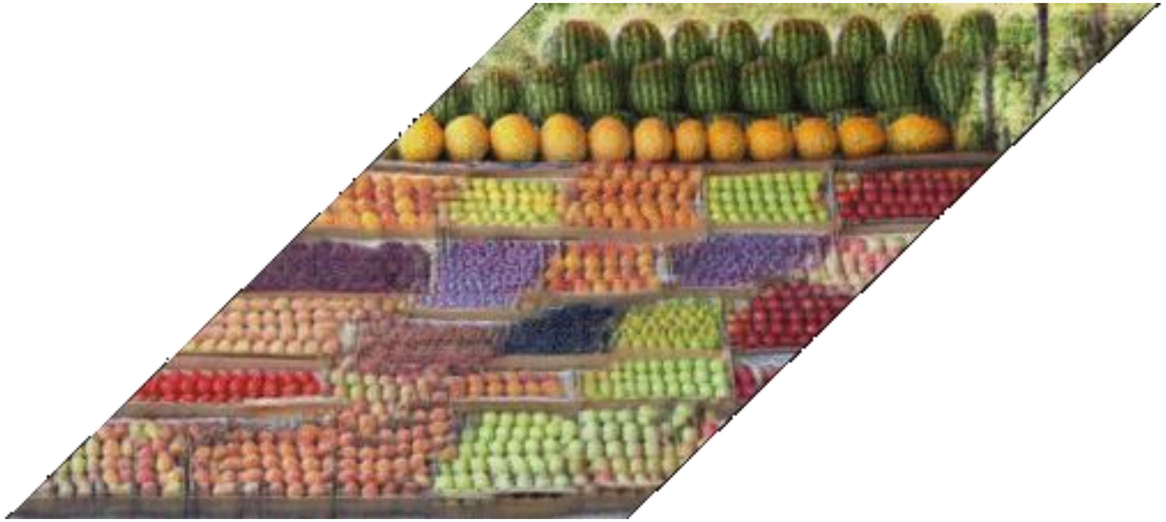




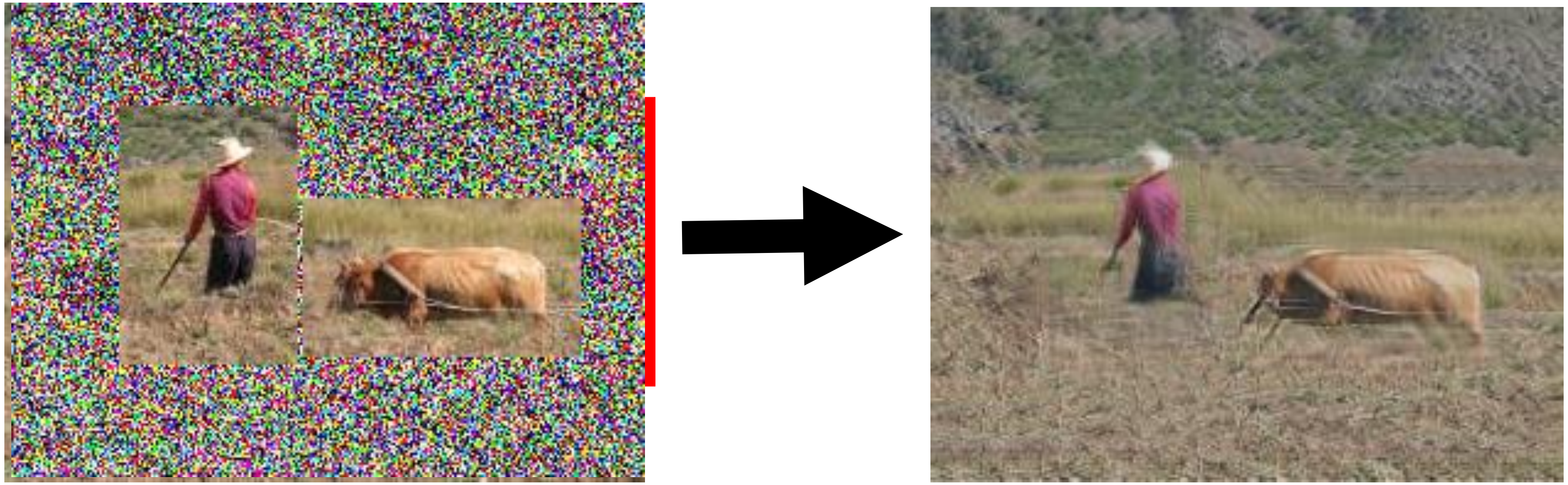
Results:



Not only resizing



Not only resizing



Double-DIP: Unsupervised Image Decomposition via Coupled Deep-Image-Priors

Yossi Gandelsman Assaf Shocher Michal Irani

Weizmann Institute of Science

Accepted CVPR’19 (oral)

Deep Image Prior [Ulyanov Et al.]

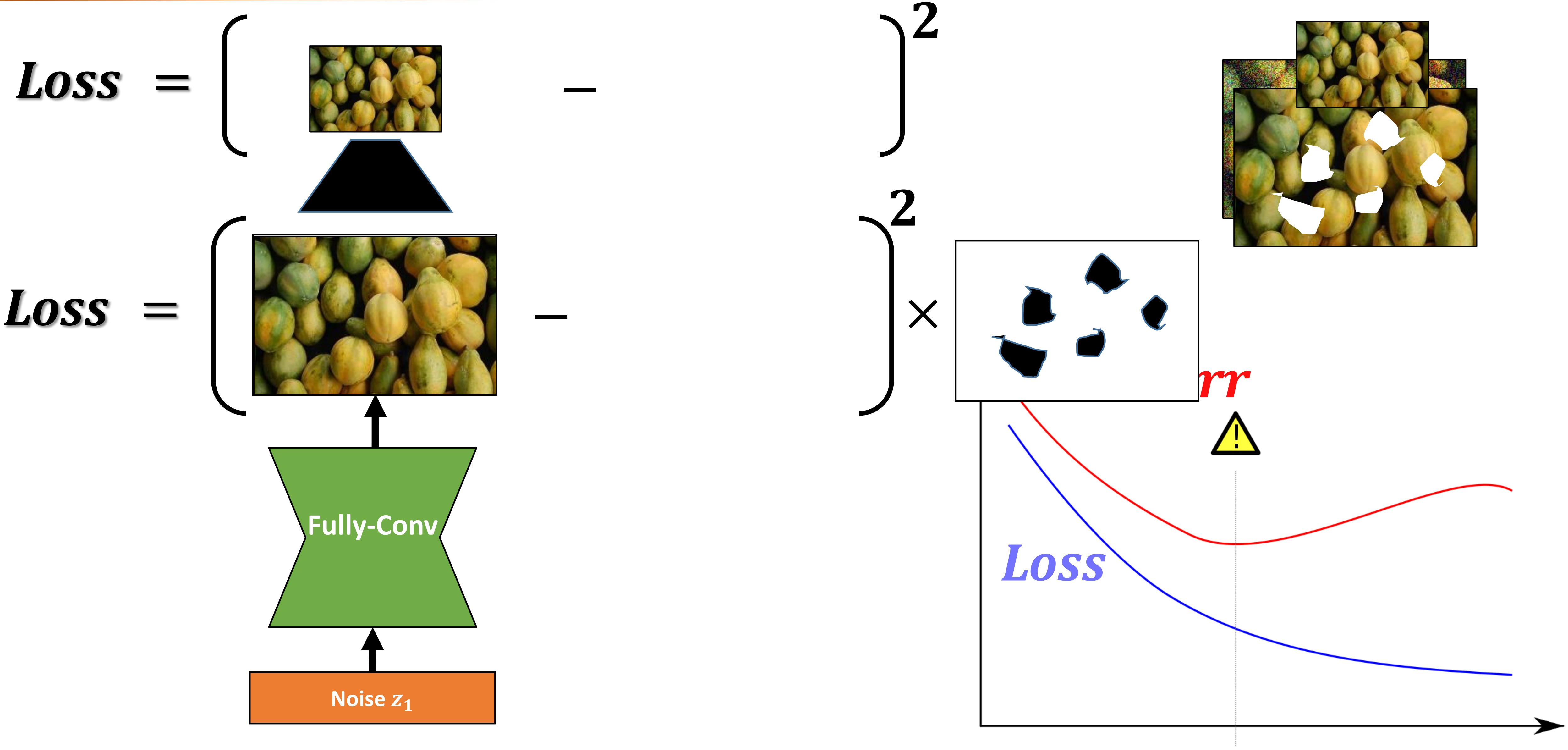
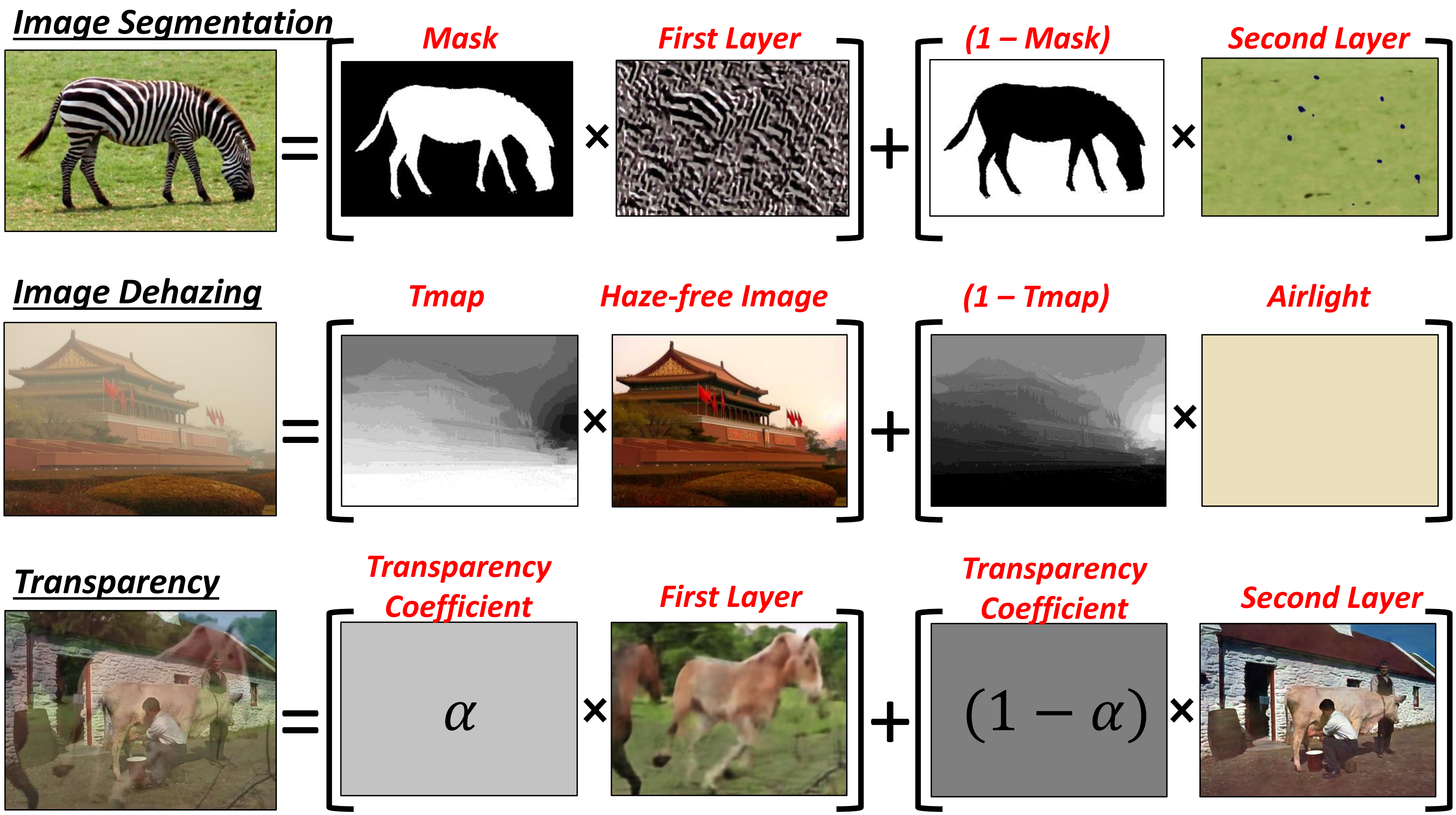
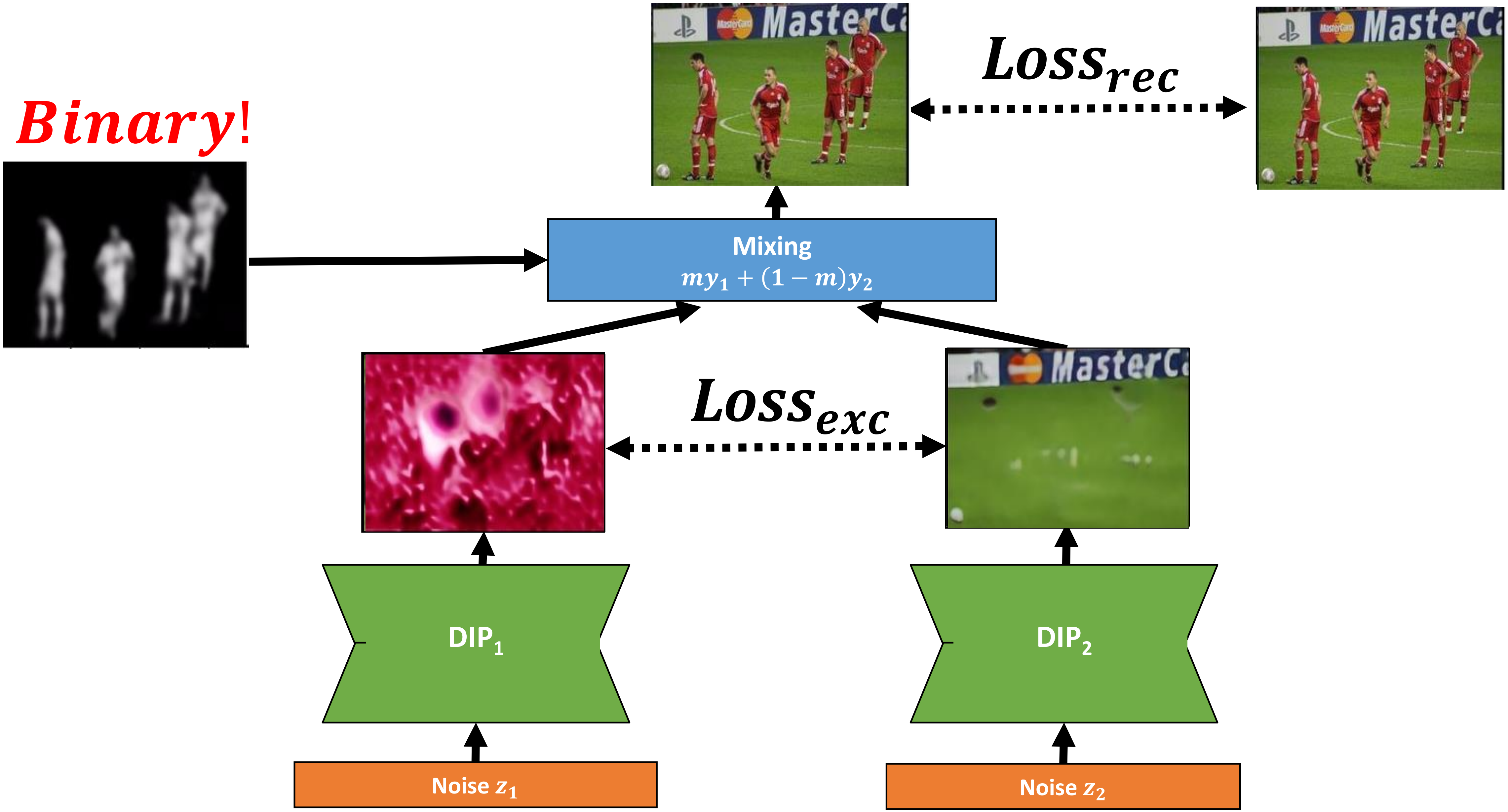


Image Decomposition

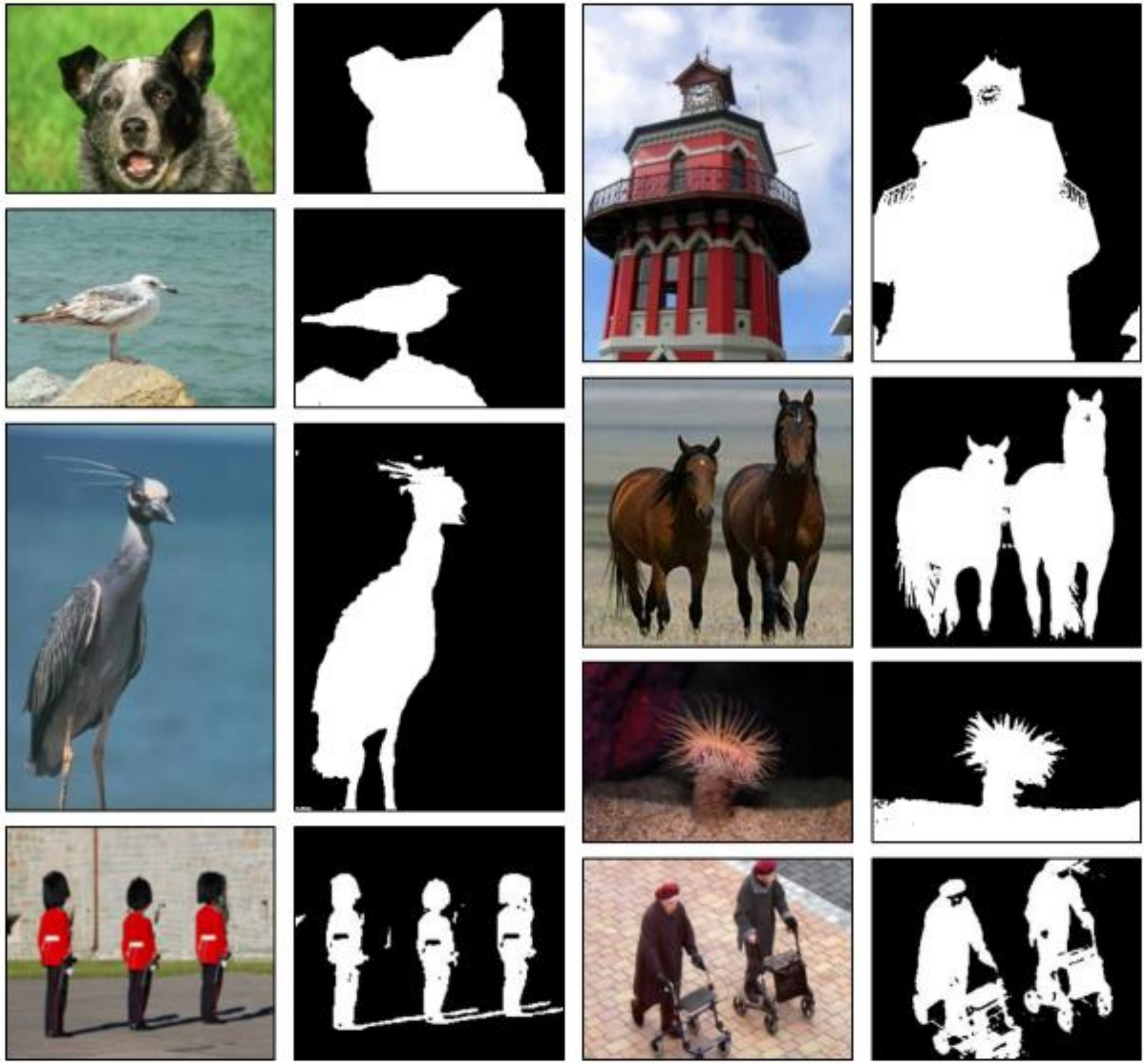


DoubleDIP

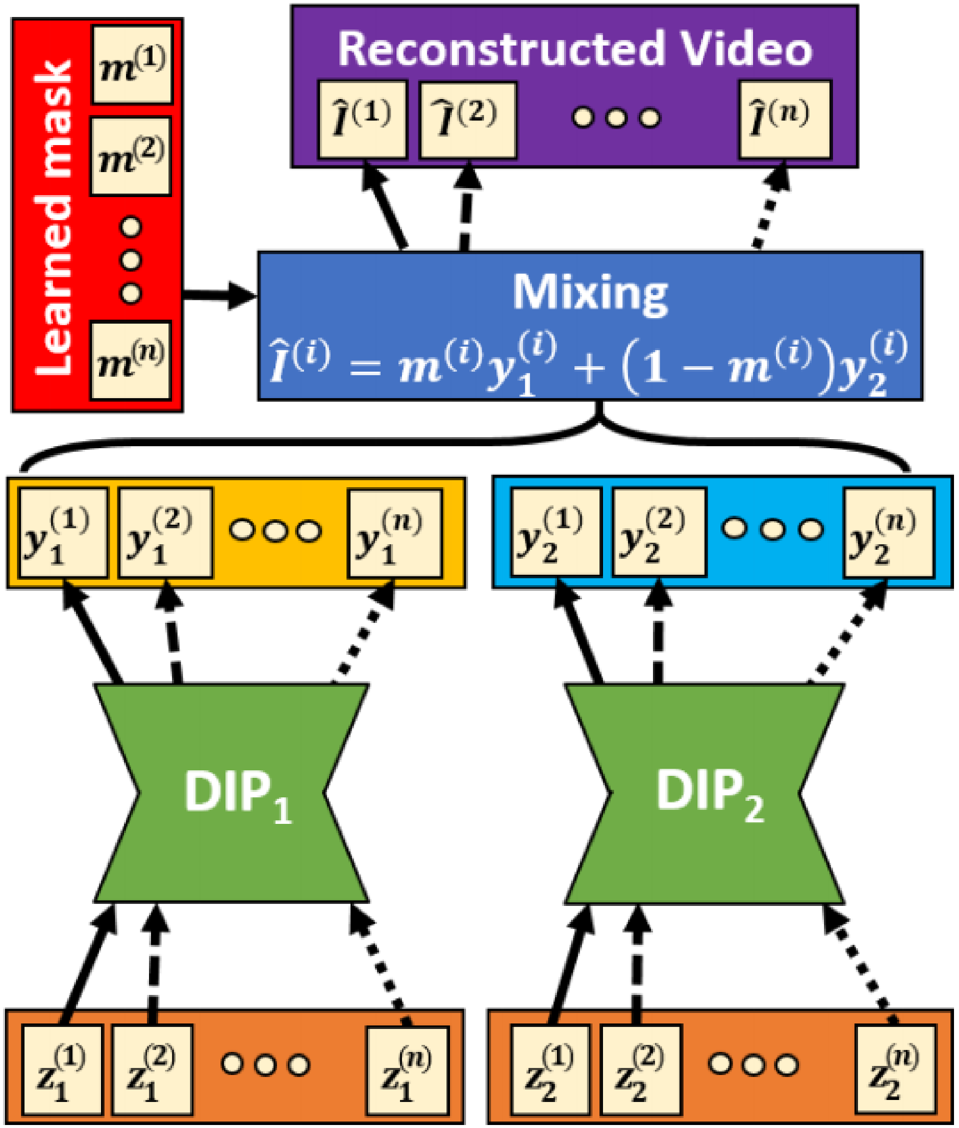


Binary!
$Loss_{rec}$
Mixing
$my_1 + (1-m)y_2$
$Loss_{exc}$
DIP_1
DIP_2
Noise z_1
Noise z_2

Segmentation



DoubleDIP for Video

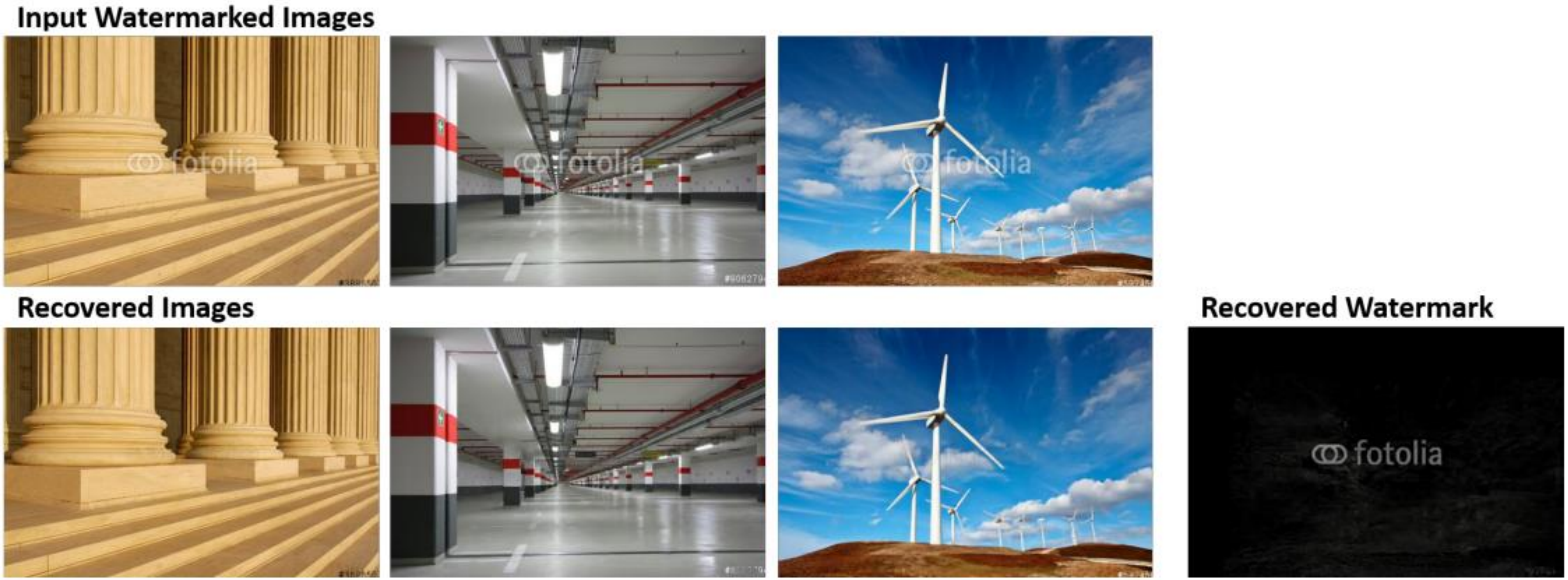


Watermark removal

A single input image (with watermark)

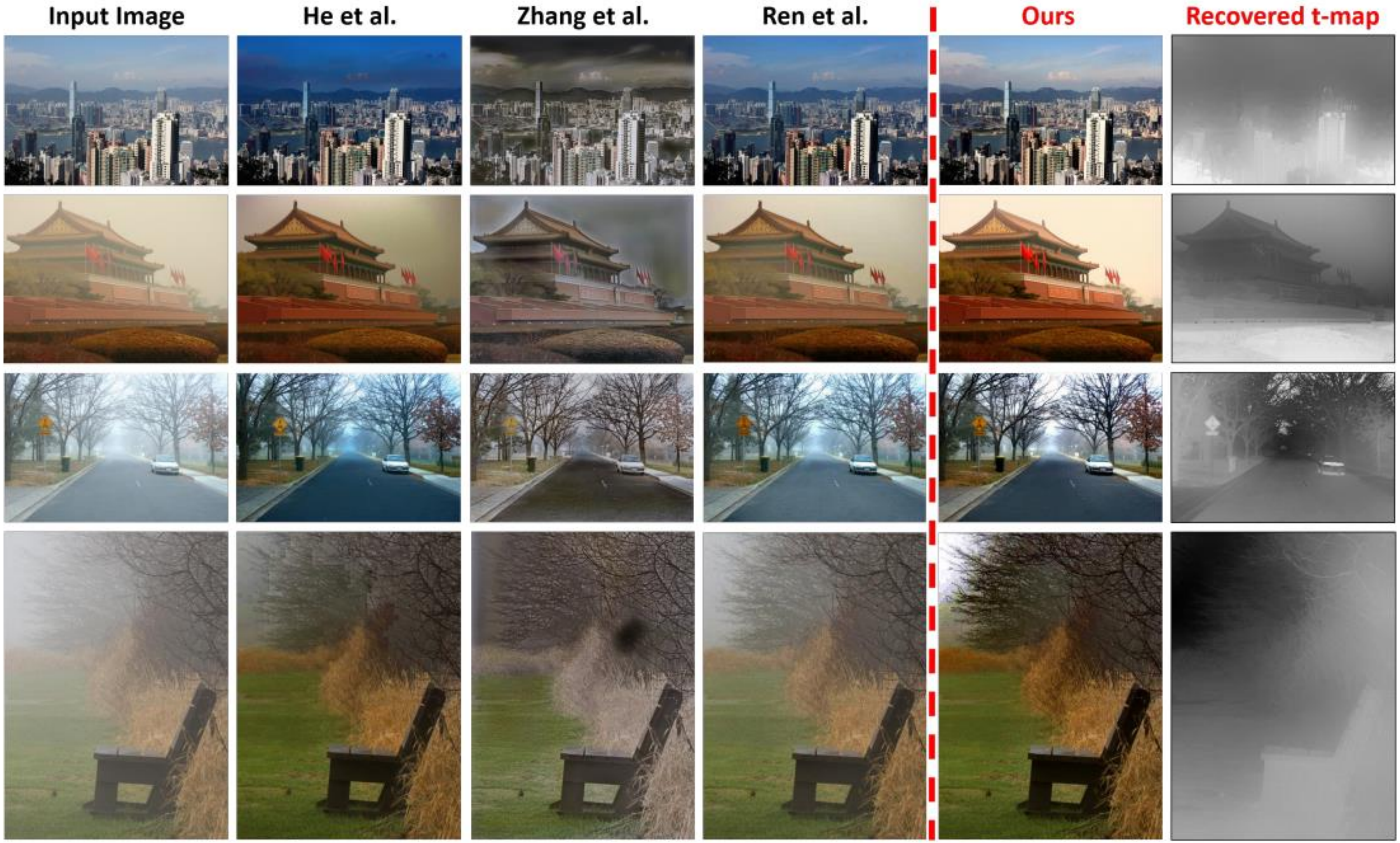
The recovered output image (watermark removed)

Watermark removal

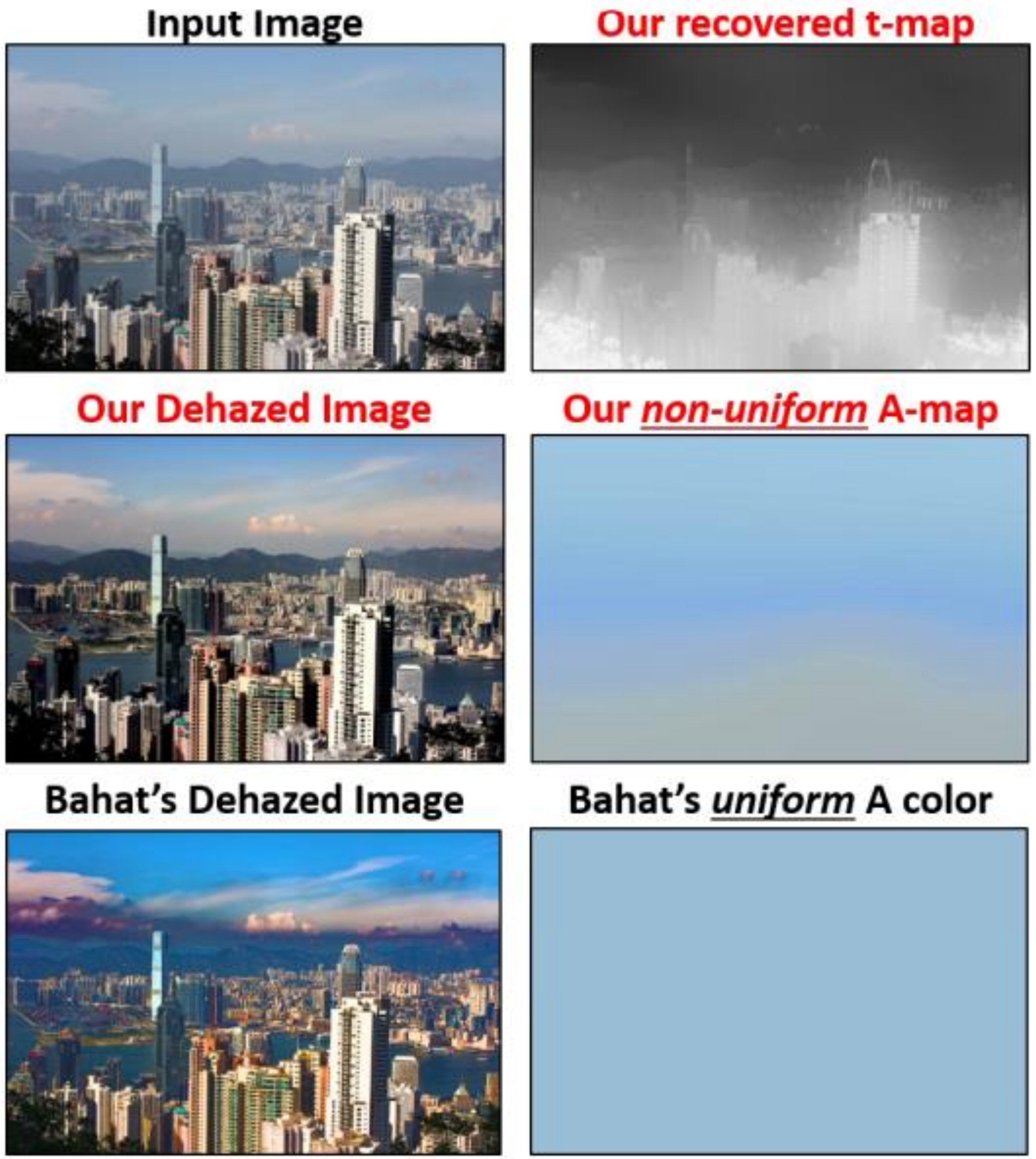


Dehazing

$$I(x) = t(x)J(x) + (1 - t(x))A(x)$$


Input Image
He et al.
Zhang et al.
Ren et al.
Ours
Recovered t-map

Dehazing: Non-uniform airlight



Take Home Message

- Tons of data in a single instance
- Local patch distribution is the DNA of an image
- Deep Internal Learning needs nothing but the input