SUPER-RESOLUTION FOR ALIASED IMAGES

Marco Foco

NVIDIA

WHAT IS SUPERRESOLUTION

SUPERRESOLUTION

Build a high resolution version of a given low resolution image

3

ZOOM! ENHANCE!

Sure!

Can you enhance that?

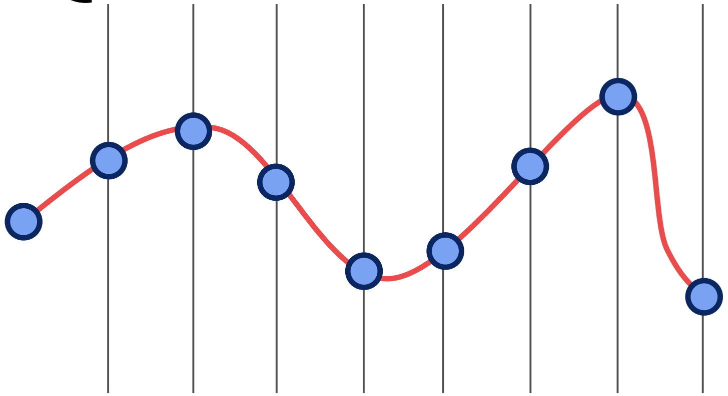
oom on the
cense plate

EVEN THE INTERNET KNOWS…

EXISTING TECHNIQUES

Interpolation (bilinear, bicubic, lanczos, etc.)

Interpolation + Sharpening (and other filtration)



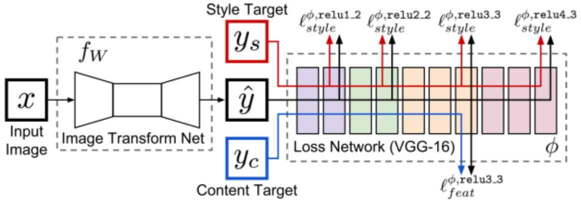
interpolation

filter-based sharpening

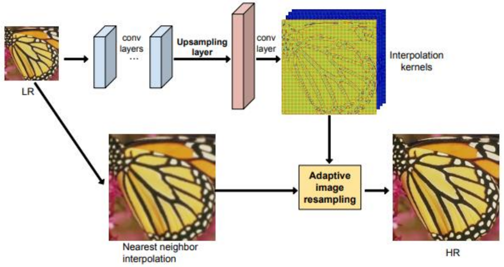
Such methods are data-independent

Very rough estimation of the data behavior

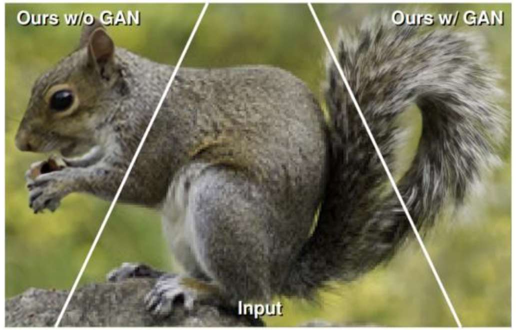
EXISTING TECHNIQUES (DEEP)



Perceptual Losses for Real-Time Style Transfer and Super-Res: 2016



Super-Res with Deep Adaptive Image Resampling: Jia et al. 2017



A Fully Progressive Approach to Single-Image Super-Res: Wang 2018

EnhaceNet: Mehdi et al. 2017

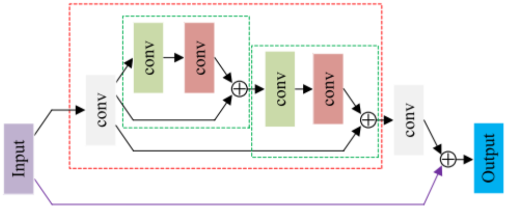


Image Super-Res via Deep Recursive ResNets: 2018

OUR SOLUTION

TRAINING PIPELINE

TRAINING PIPELINE

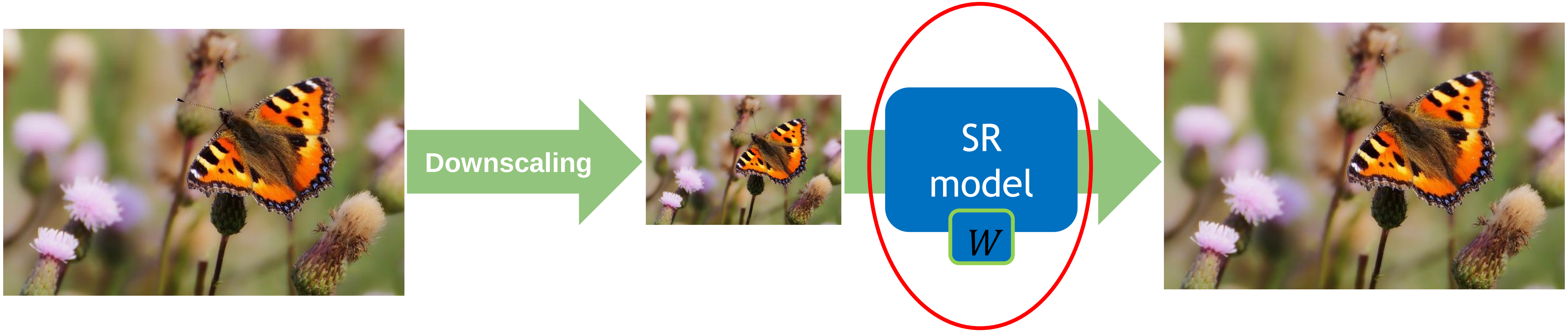
Model Input

Downscaling = **Filtering** + **Decimation**

cutoff frequency at
(or below) nyquist

TRAINING PIPELINE

Model optimization



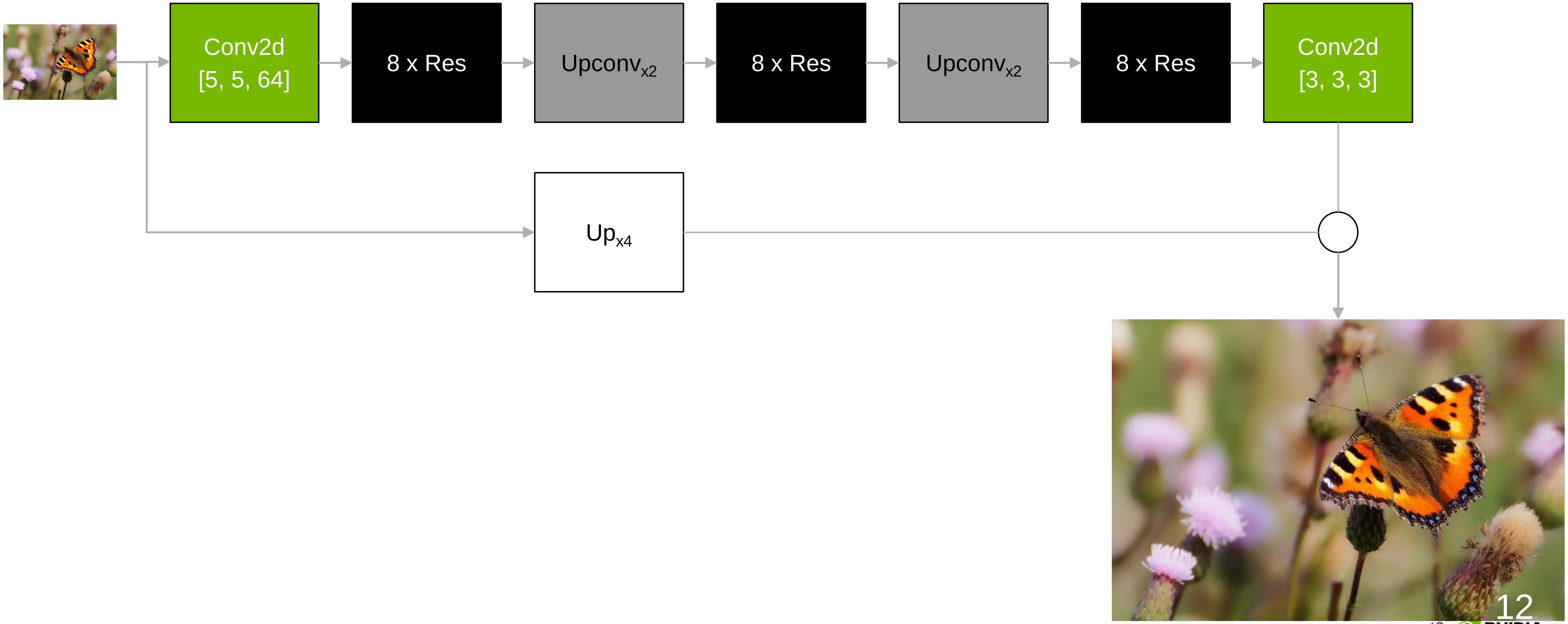
Solve the optimization problem:

$$W = argmin \sum_{i} Dist(x_i , F_W(D(x_i)))$$

$\{x_i\}$- training set

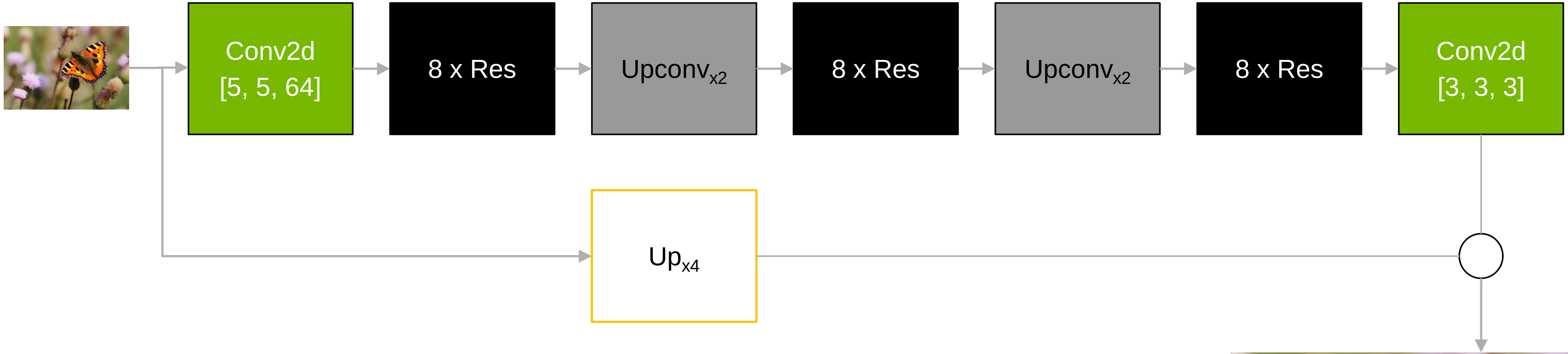
MODEL (GWMT)

4x upscaling model



MODEL (GWMT)

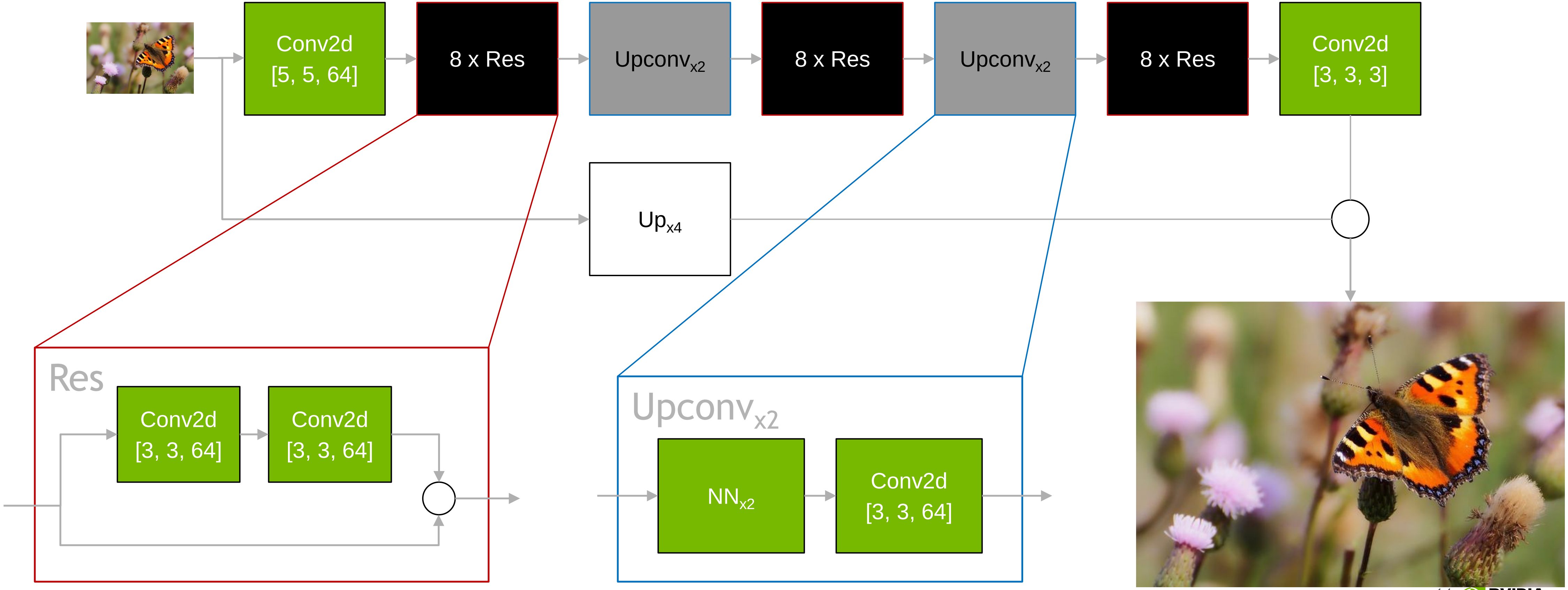
4x upscaling model



Low-pass: Bilinear up-scaling of the input image.

MODEL (GWMT)

4x upscaling model



DATASET

OpenImagesV4*

Training on fixed-size random crops

Input data issues

JPEG compression artifacts

Raw

JPEG (over-compressed)

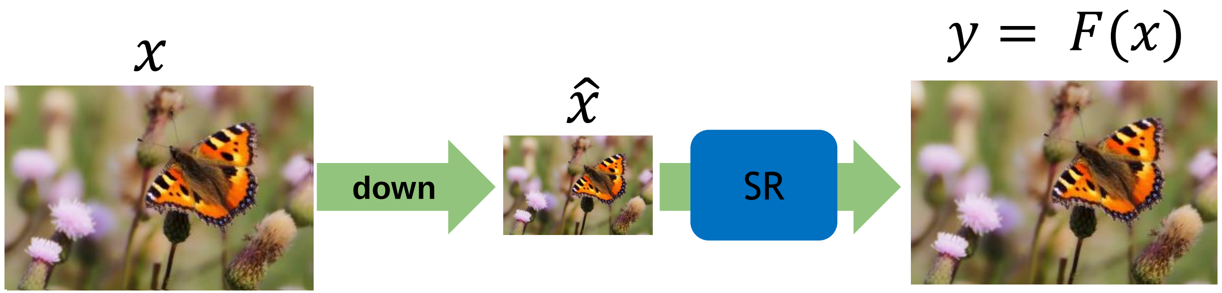
* https://storage.googleapis.com/openimages/web/index.html

LOSS FUNCTION

MSE | HFEN | VGG | TV | GAN

LOSS FUNCTION

MSE | HFEN | VGG | TV | GAN



PSNR
Peak Signal-to-Noise Ratio

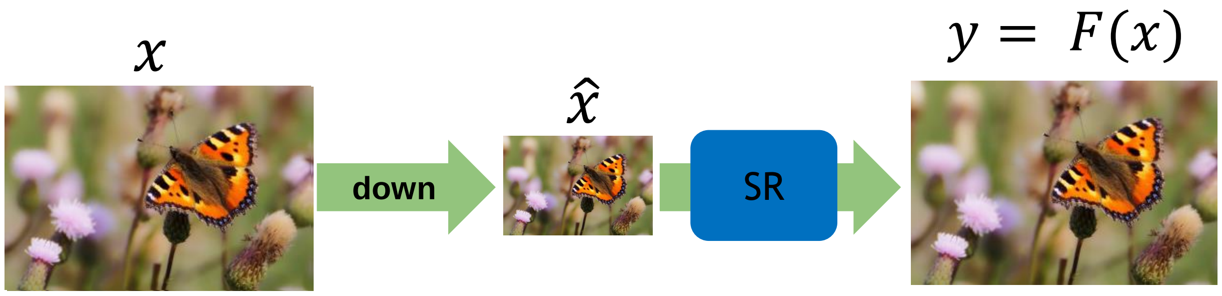
MSE loss: $L = \frac{1}{N}\|x - F(x)\|^2$

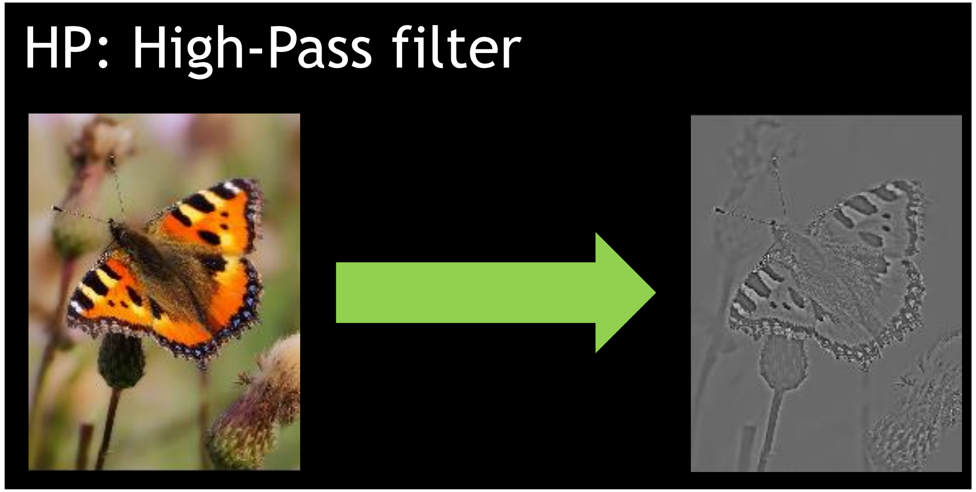
Related to

$$10 * log_{10}\left(\frac{MAX^2}{MSE}\right)$$

LOSS FUNCTION

MSE | HFEN | VGG | TV | GAN





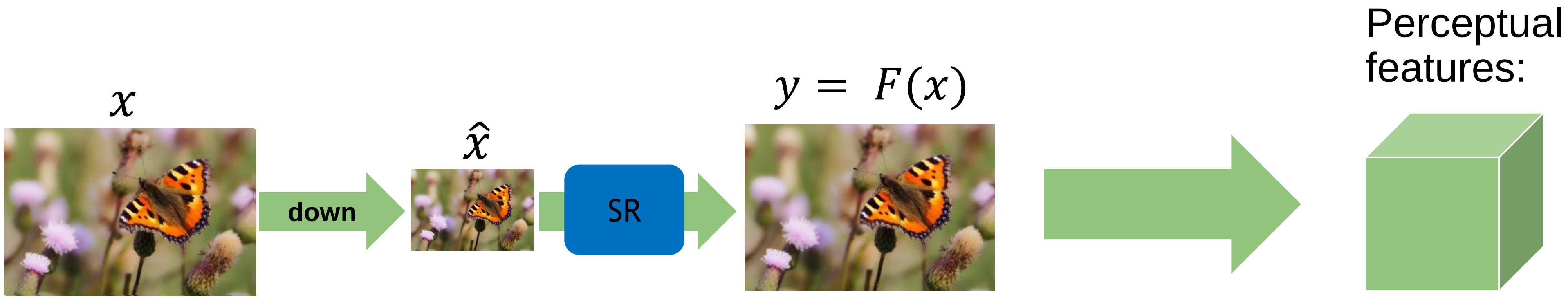
HFEN* loss: $\mathrm{L} = \alpha_1 ||HP(x - F(x))||^2$

- HFEN*: High Frequency Error Norm

http://ieeexplore.ieee.org/document/5617283/

LOSS FUNCTION

MSE | HFEN | VGG | TV | GAN



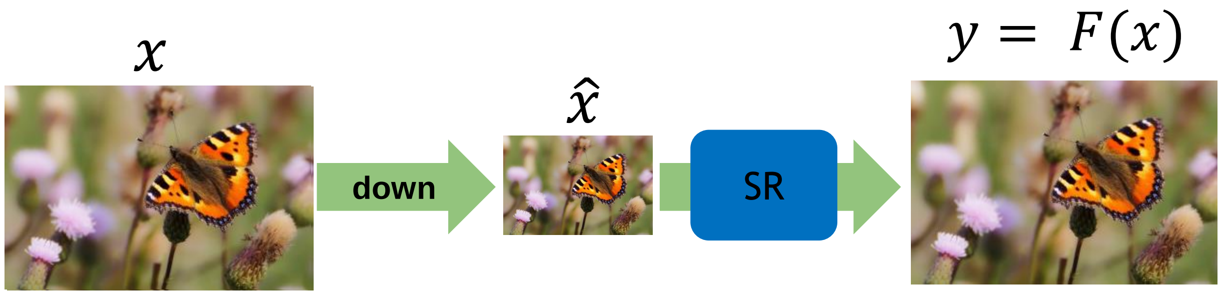
VGG* loss: $\mathrm{L} = \alpha_2 \| G(x) - G(F(x)) \|^2$

- VGG19 features taken after the 4^{th} convolutional layer (before 5^{th} max-pooling)

https://arxiv.org/abs/1409.1556

LOSS FUNCTION

MSE | HFEN | VGG | TV | GAN

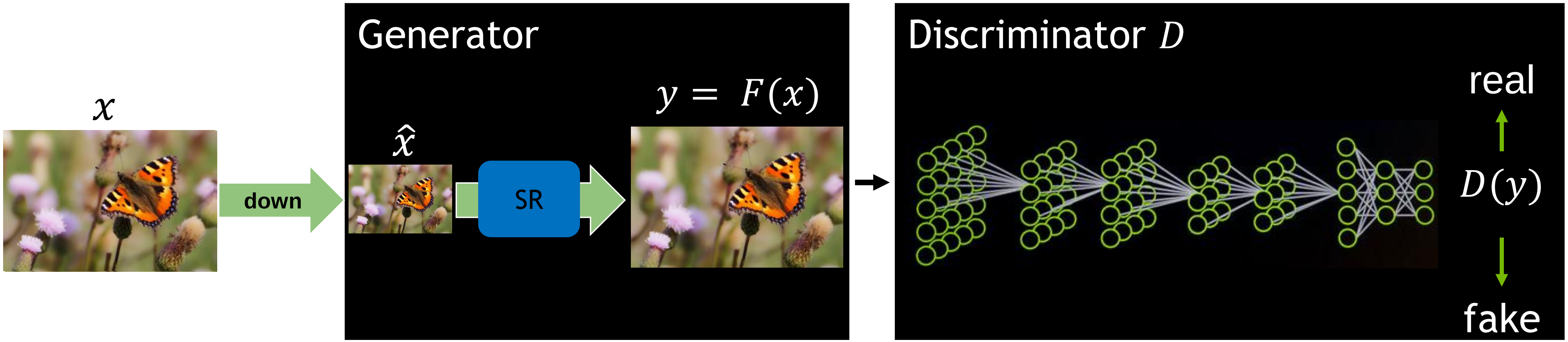


TV loss: $L = \alpha_3 \int_{\Omega} |\nabla F(x)|$

- Serves as a regularizer and has little influence on the optimization

LOSS FUNCTION

MSE | HFEN | VGG | TV | GAN



GAN loss $= -\alpha_4 lnD(F(x))$

https://arxiv.org/pdf/1609.04802.pdf

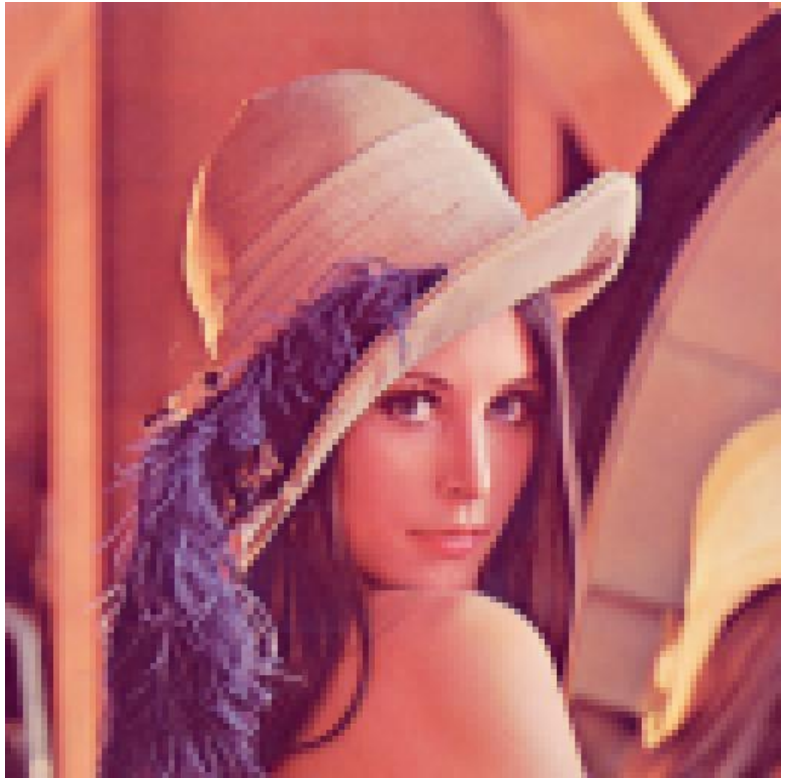
RESULTS


ONE DOES NOT SIMPLY
ENHANCE THE IMAGE


ONE DOES NOT SIMPLY
ENHANCE THE IMAGE

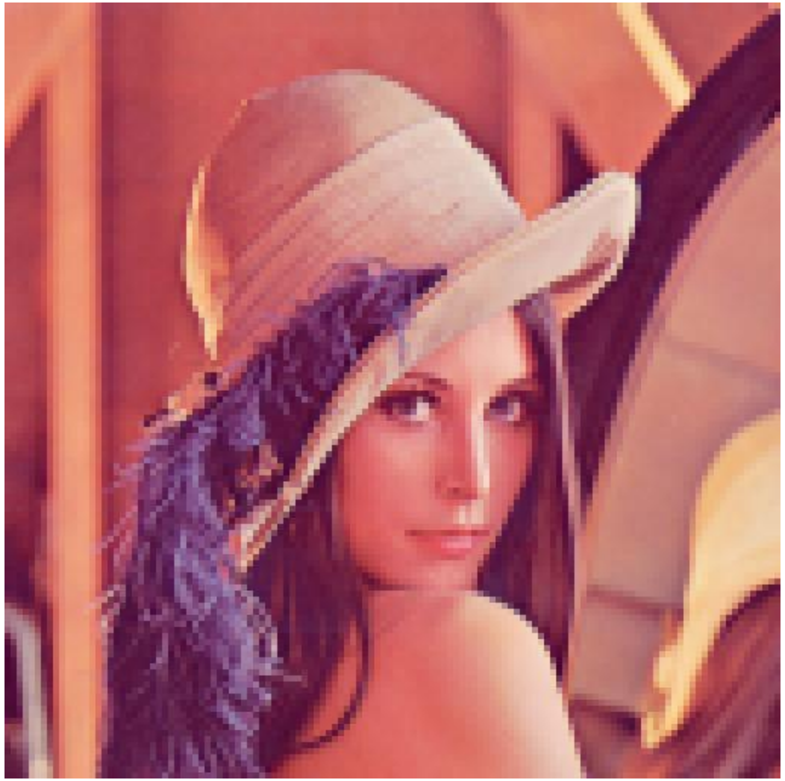
COMPARISON

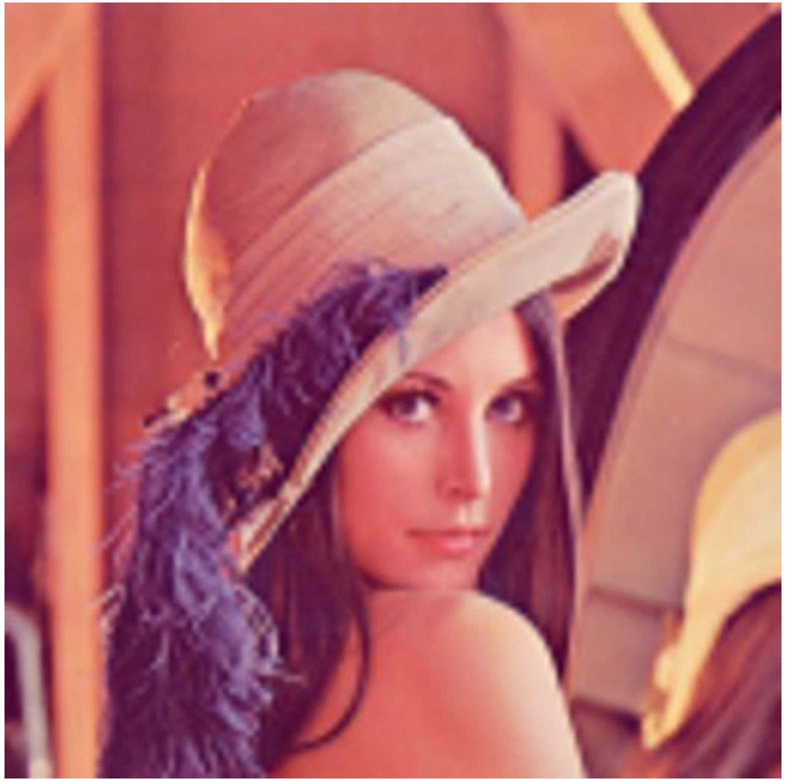
Original vs downscaled



COMPARISON

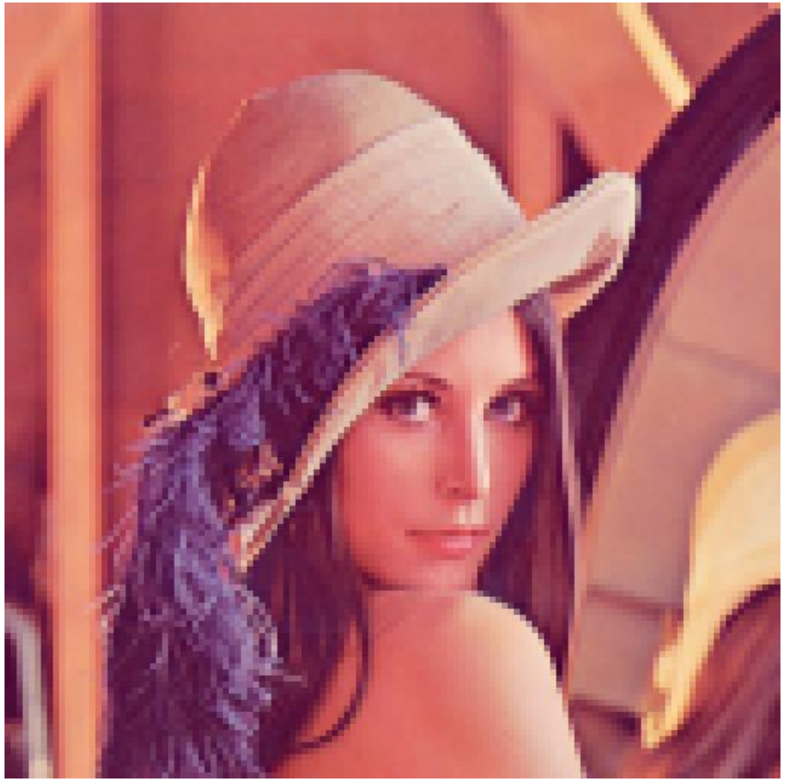
downscaled vs bicubic





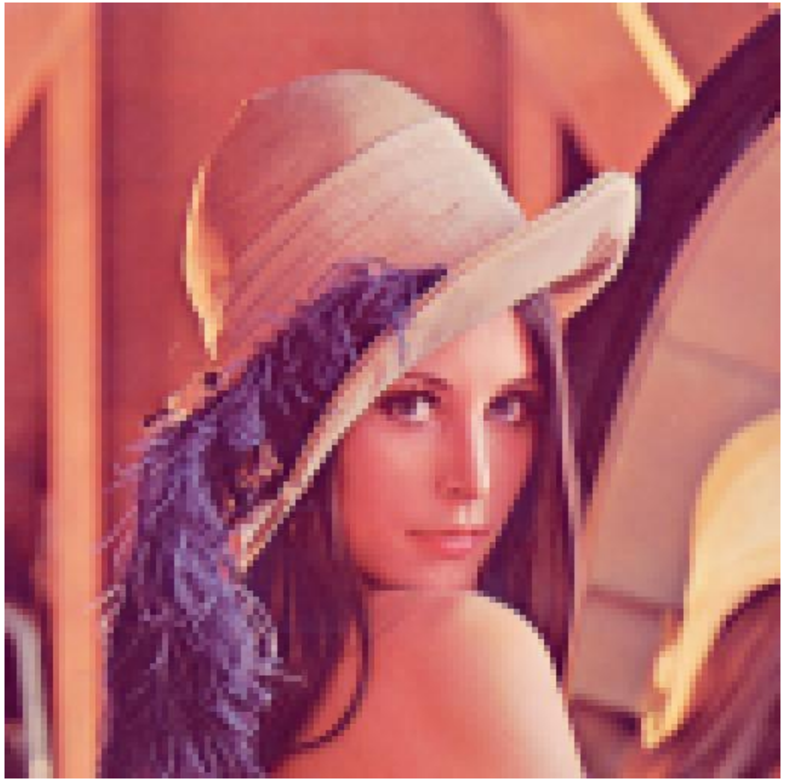
COMPARISON

downscaled vs perceptual



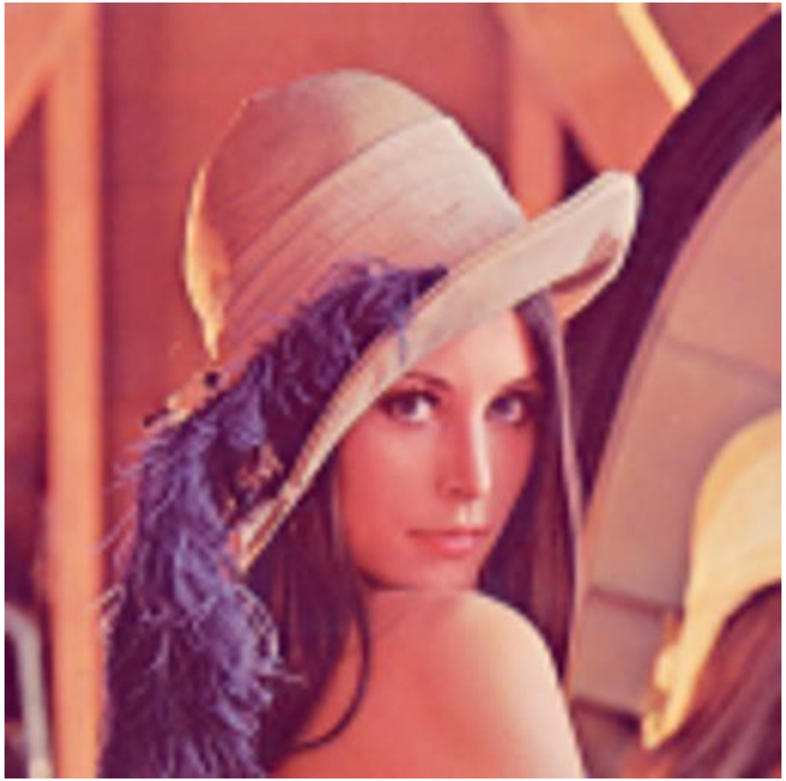
COMPARISON

downscaled vs perceptual+GAN



COMPARISON

original vs bicubic



COMPARISON

original vs perceptual

COMPARISON

original vs perceptual+GAN

COMPARISON

easy details (hat)

Original

Perceptual

Downscaled (input)

Bicubic

Perceptual + GAN

COMPARISON

details (eye)

Original

Perceptual

Downscaled (input)

Bicubic

Perceptual + GAN

COMPARISON

hard details (feathers plume)

Original

Perceptual

Downscaled (input)

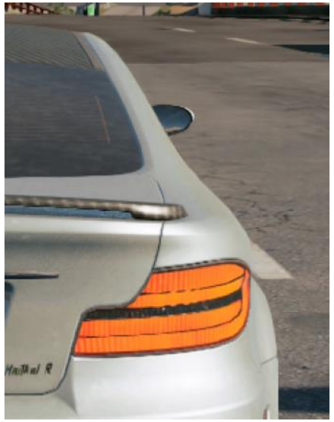
Bicubic

Perceptual + GAN

WHAT ABOUT SYNTHETIC IMAGES?

COMPARISON

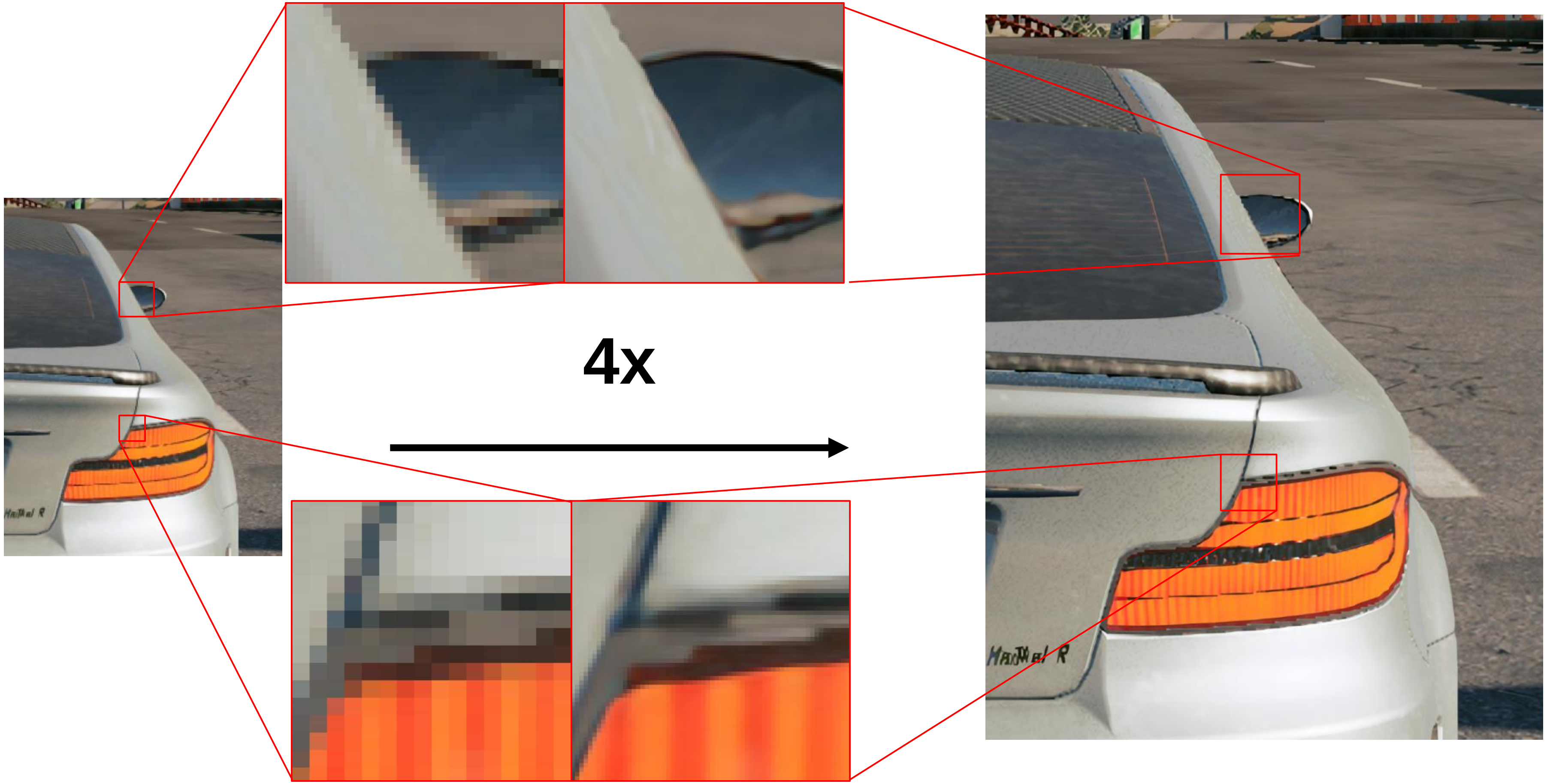
Synthetic Images



4x

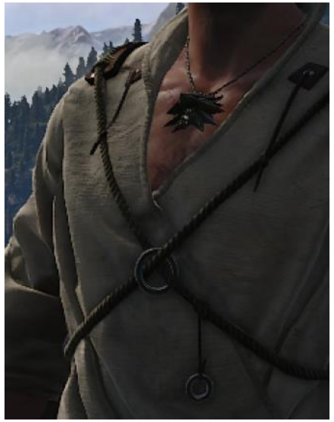
COMPARISON

Synthetic Images



COMPARISON

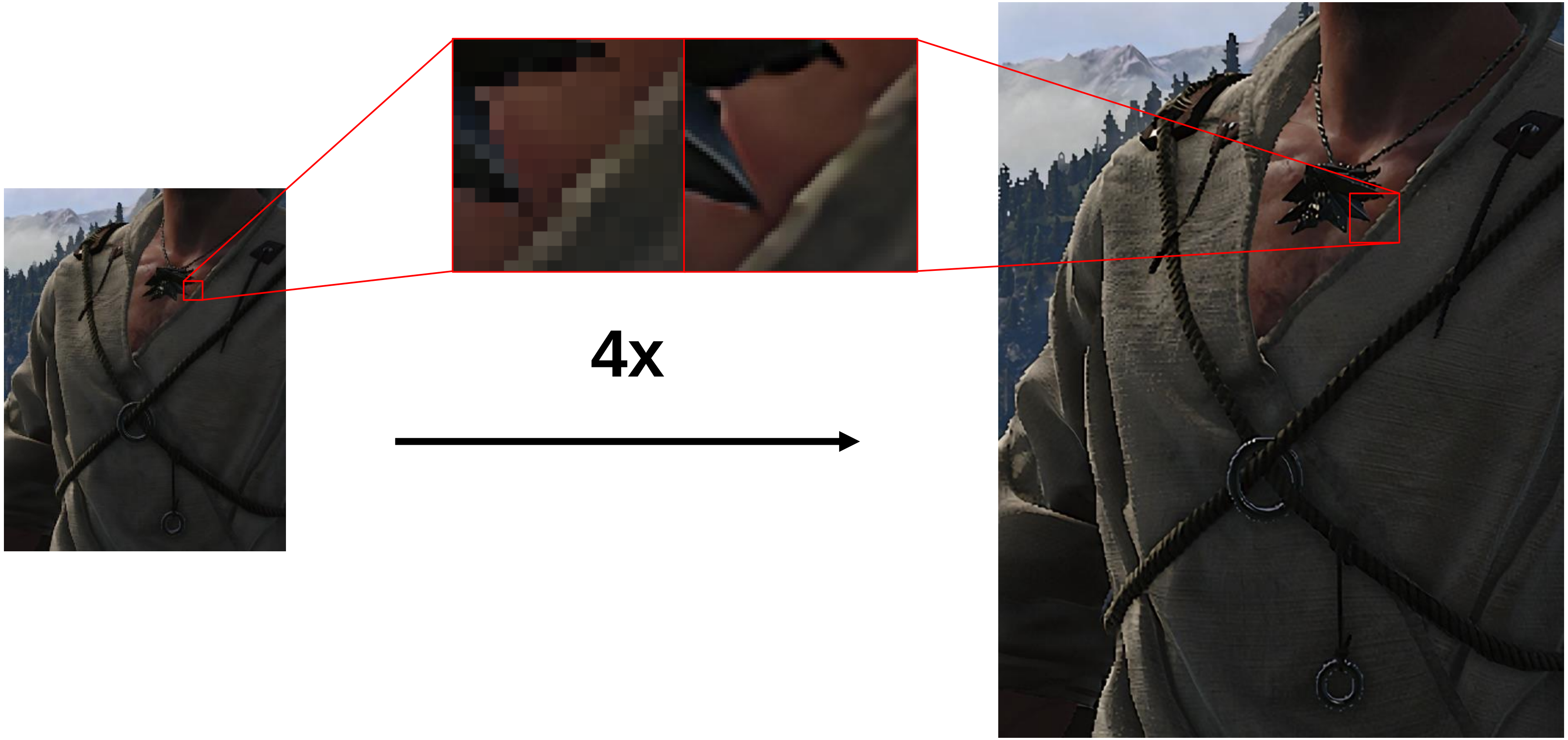
Synthetic Images



4x

COMPARISON

Synthetic Images



OBSERVATIONS

On synthetic image upscaling

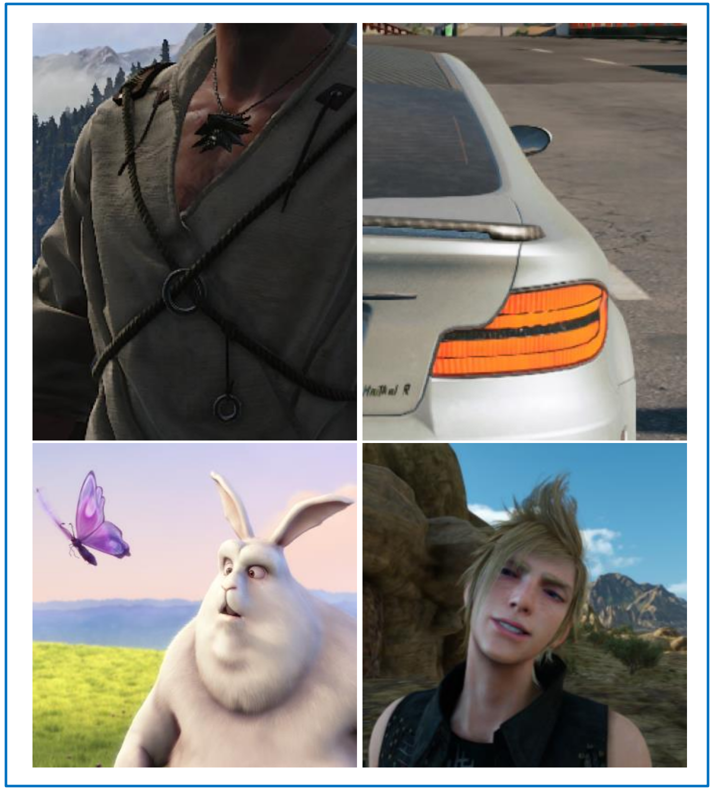
- Synthetic images have more high frequency details
- Synthetic images with dithering contains noise-like artifacts
- The Network has never seen synthetic images during trainings
- Presence of artifacts in training image is reflected into upscaling artifact
  - Especially with GANs

- We can probably improve these results

UPSCALING SYNTHETIC IMAGES

GOAL

Train Super Resolution for synthetic images



SOLUTION?

Train on game images!

SOLUTION?

Train on game images!

- Difficult to produce

SOLUTION?

Train on game images!

- Difficult to produce

- Extremely biased dataset

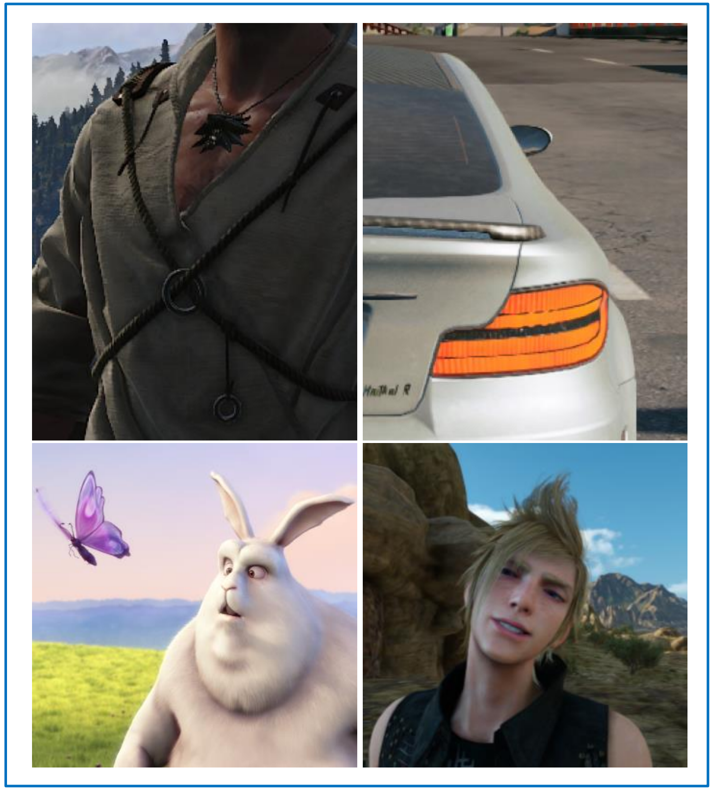
SOLUTION?

Train on game images!

- Difficult to produce
- Extremely biased dataset
- License issues?

NEW GOAL

Train SuperRes with natural images and apply to synthetic images

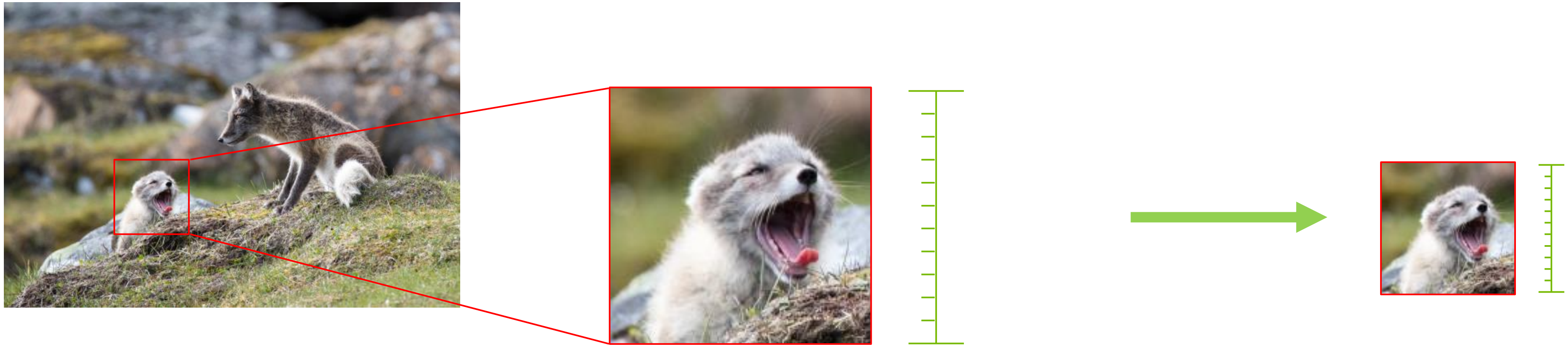


AUGMENTATION

SOLUTION

Augment photographic images

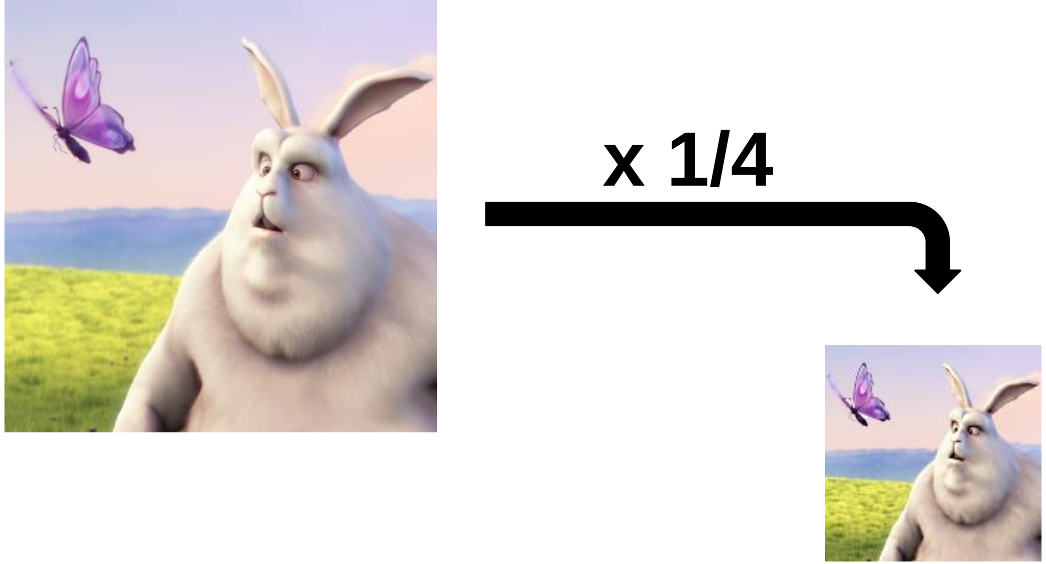
To reduce the compression artifacts, we will extract random crops and downscale them to our training crop size



SOLUTION

Downscale with aliasing

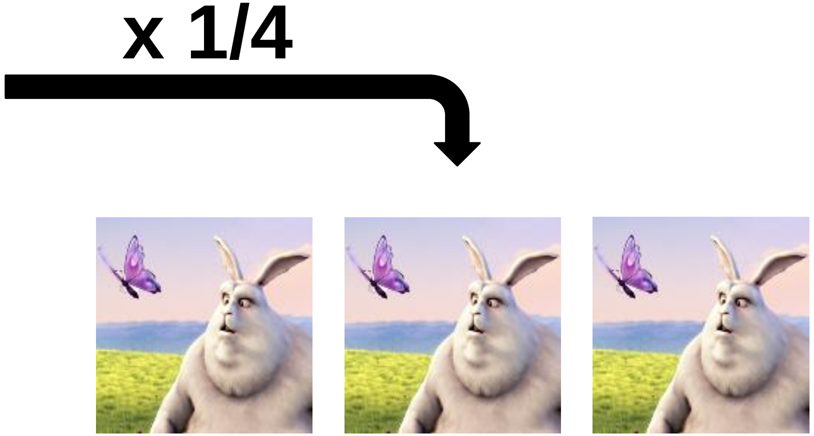
Filter the image with a cutoff above Nyquist limit



SOLUTION

Downscale with variable aliasing

Use different cutoff limits above Nyquist



Every downscaling now generate examples with different aliasing features.

SOLUTION

Stochastic decimation

After filtering, instead of sampling on a regular grid, jitter each sampling point

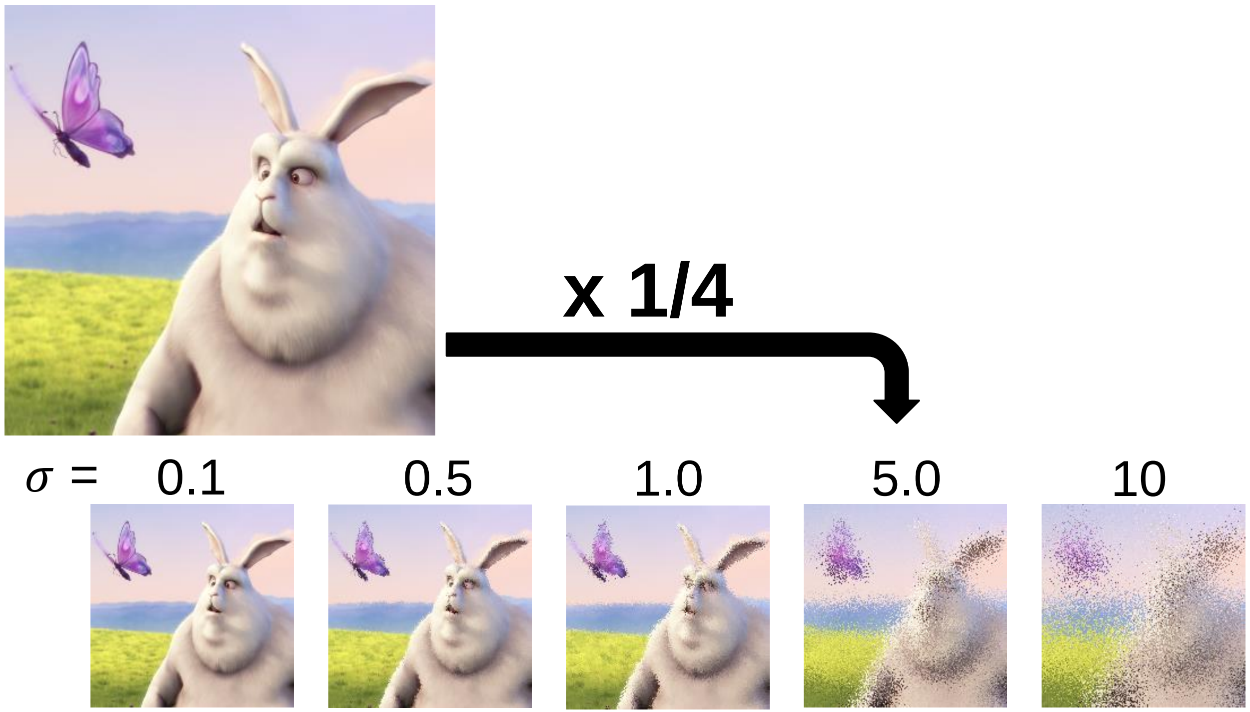
x 1/4

Every downscaling now generate very different examples

SOLUTION

Variable stochastic decimation

Full control over introduced noise/ aliasing effect



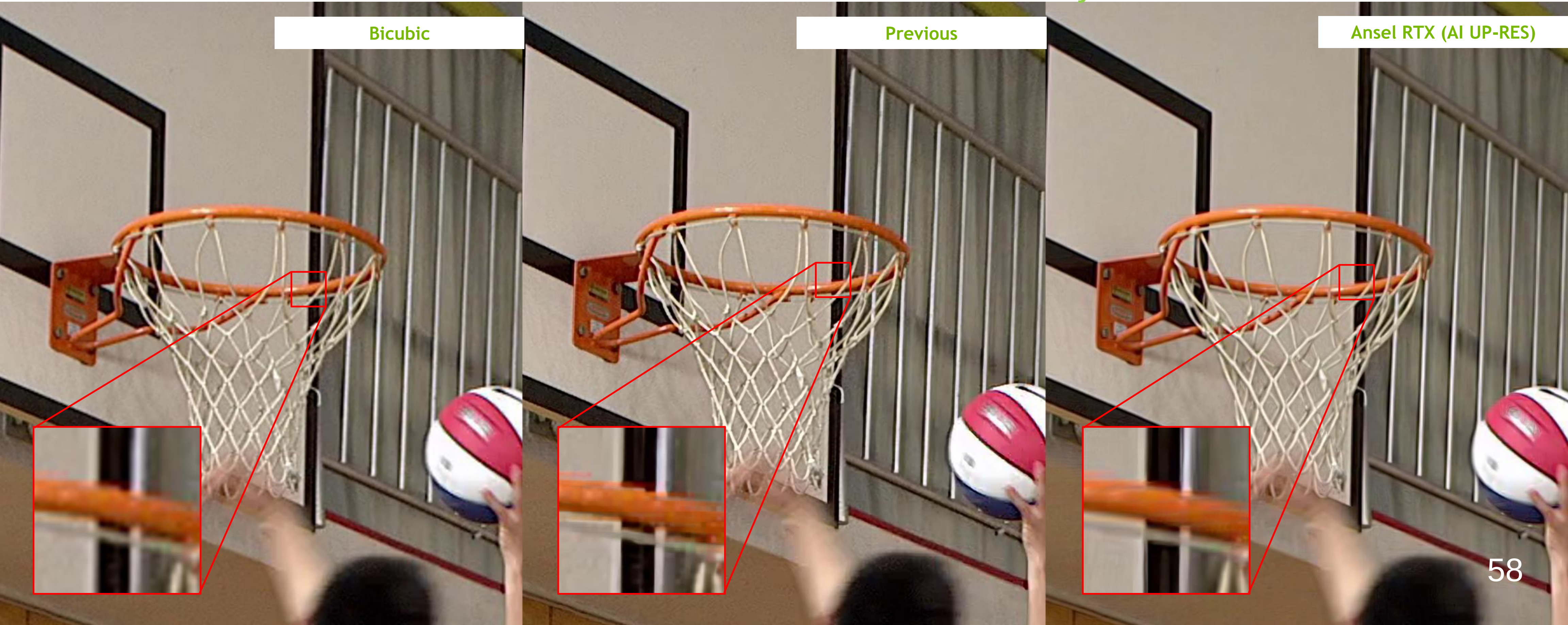
COMPARISON WITH PREVIOUS METHOD

EVALUATION

EVALUATION

EVALUATION

Note. Not only unaliased but also denoised!



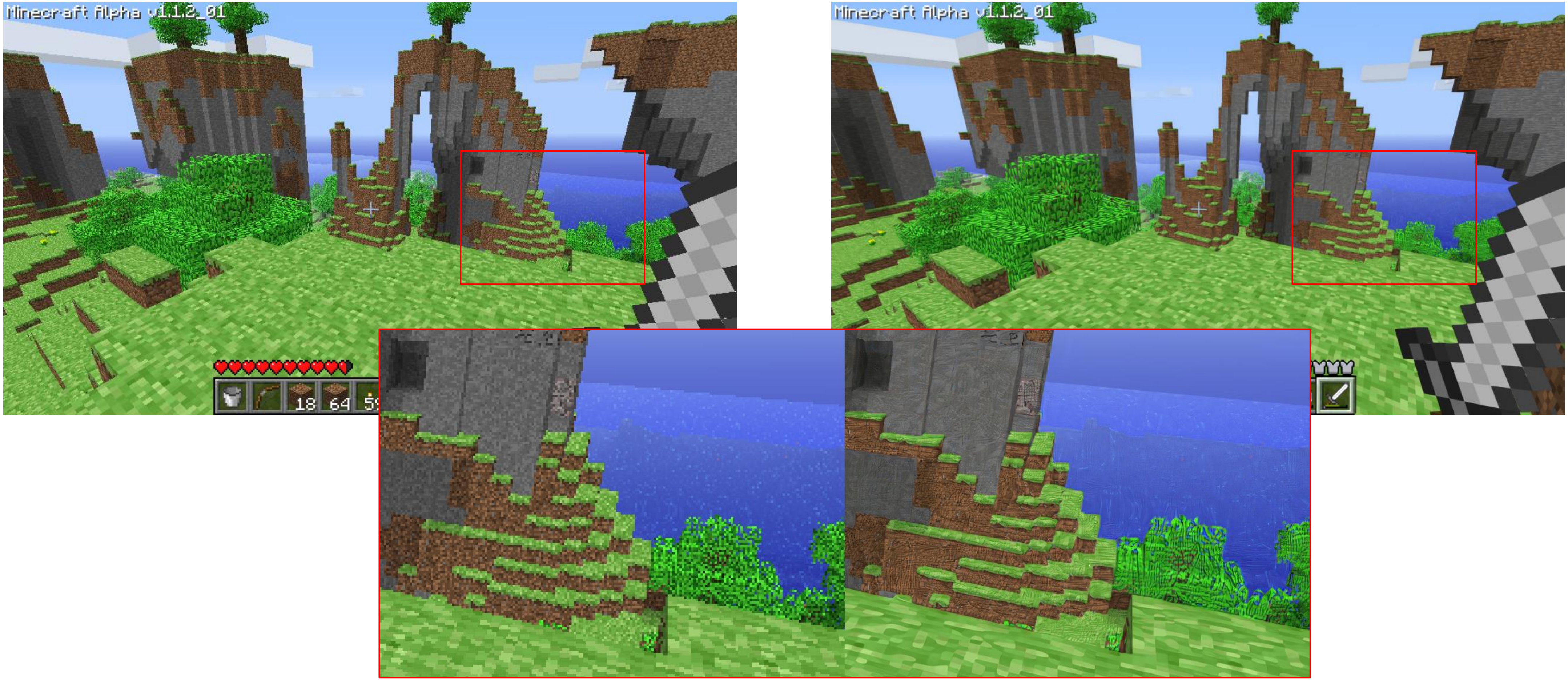
COMPARISON WITH INPUT IMAGES

INPUT VS OUTPUT

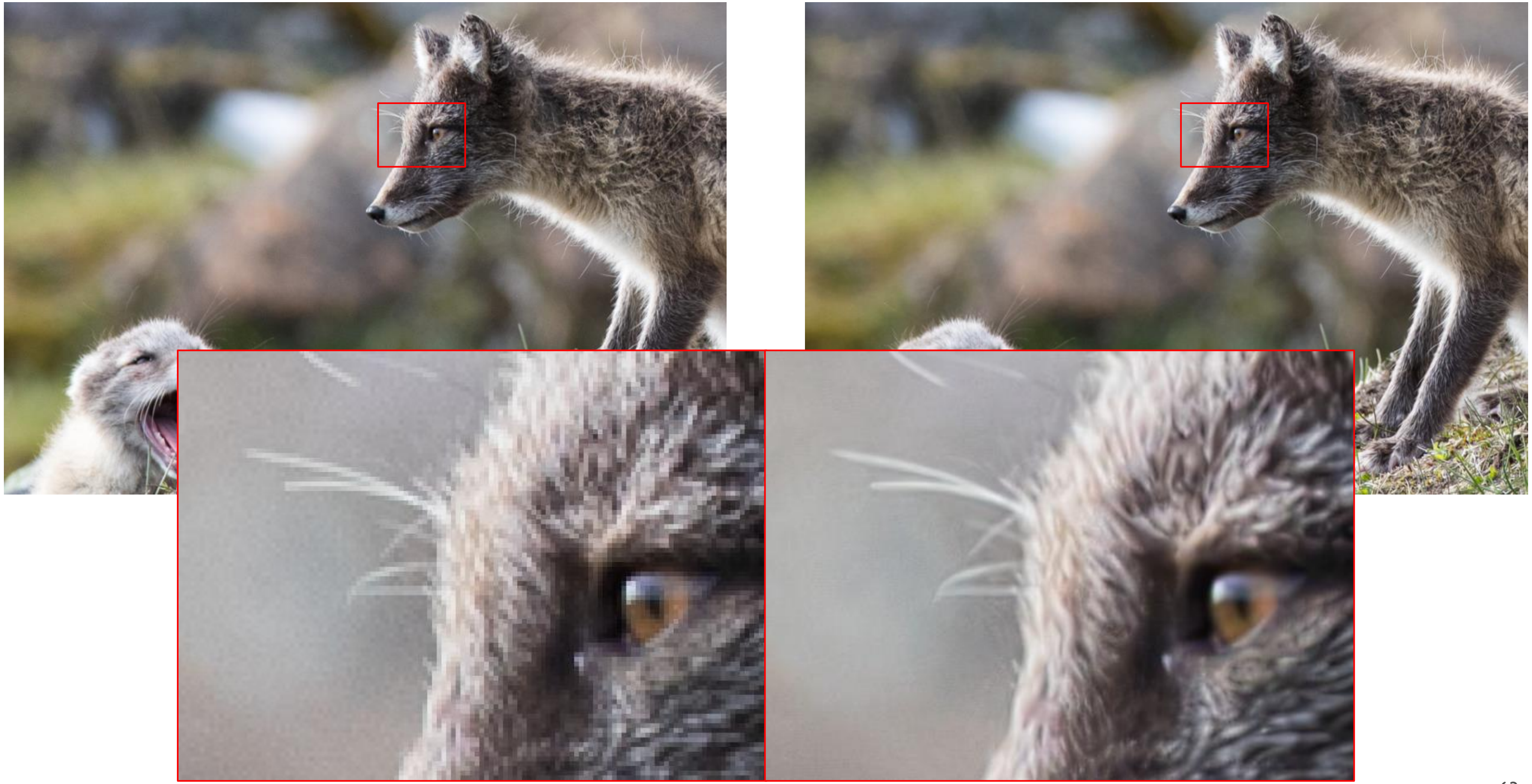
Note. The input images are interpolated by Nearest Neighbor algorithm to make it same size with upscaled image

INPUT VS OUTPUT

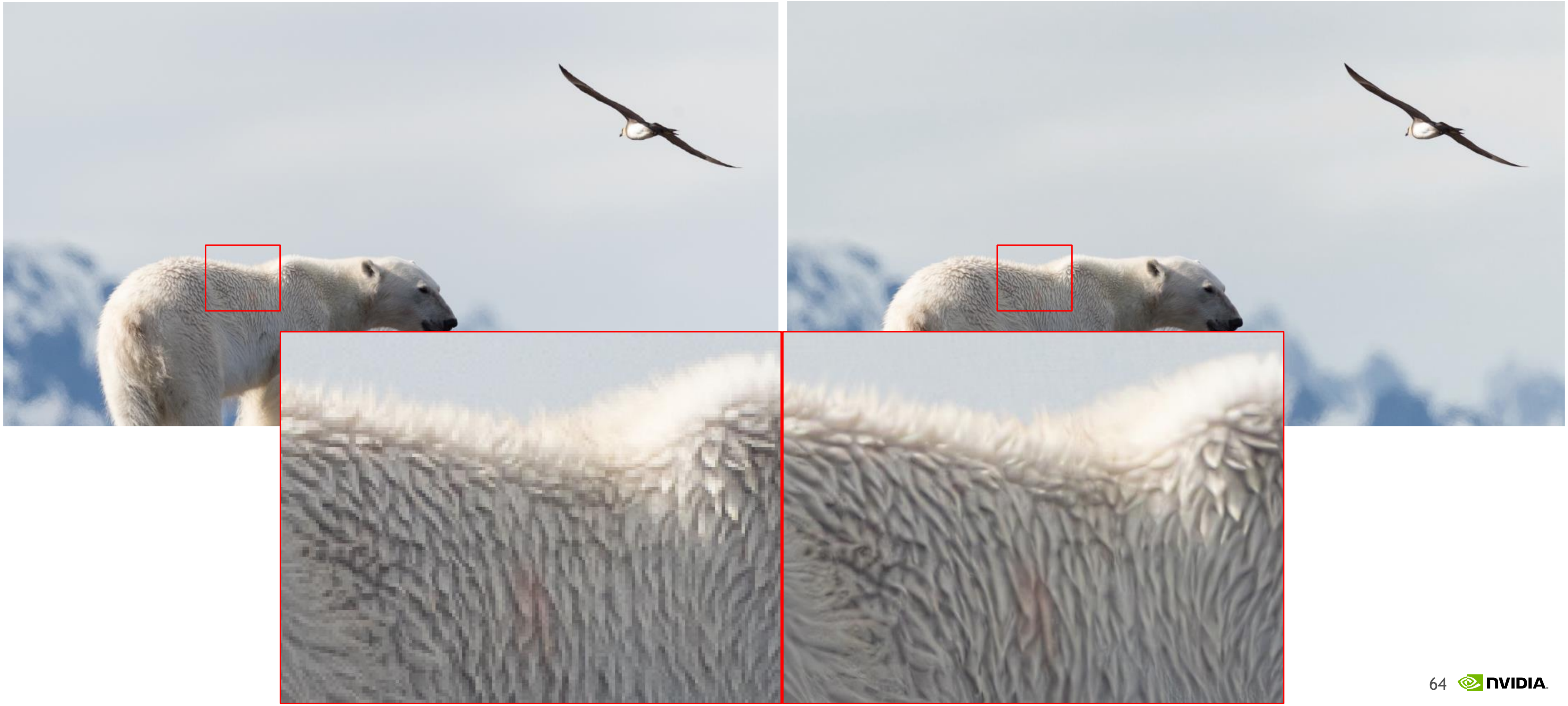
INPUT VS OUTPUT



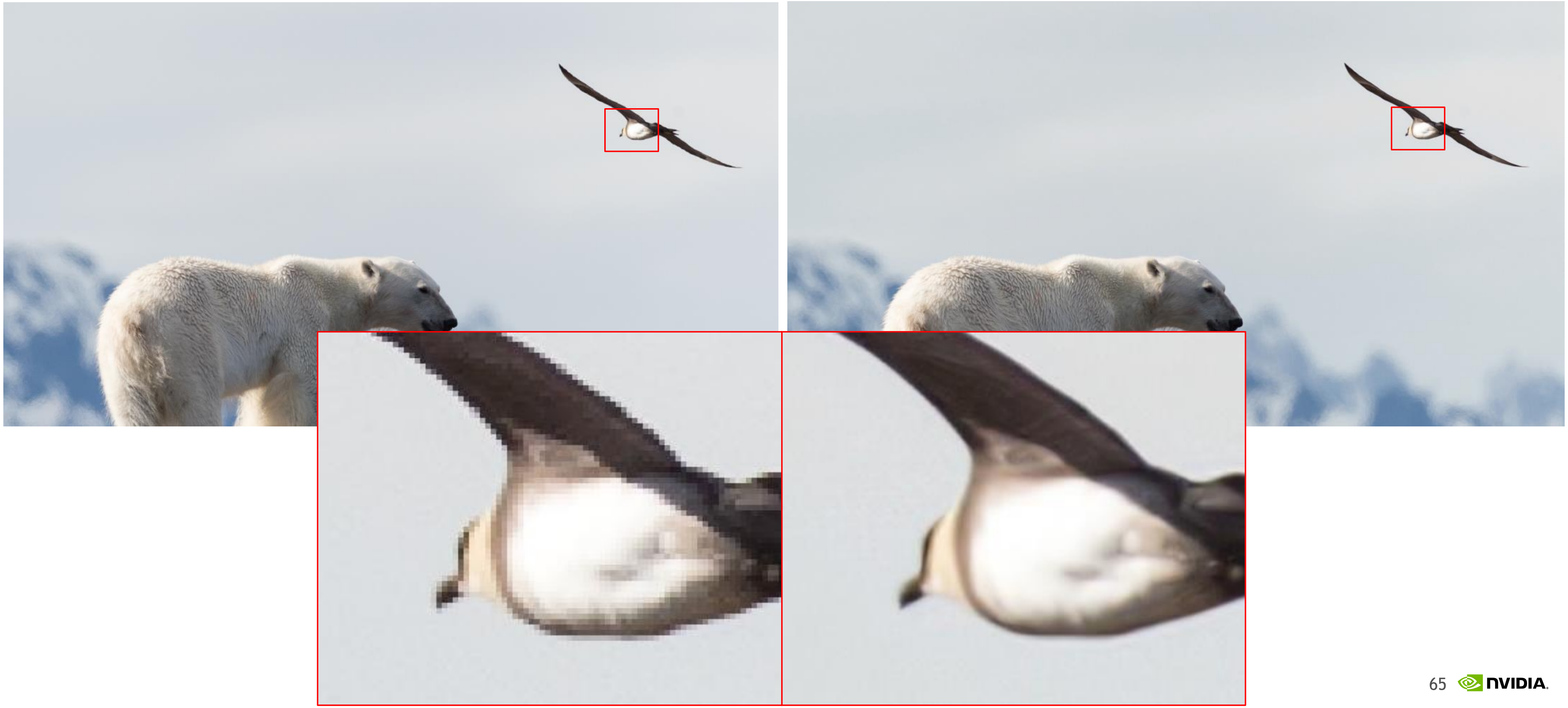
INPUT VS OUTPUT (REAL IMAGE)



INPUT VS OUTPUT (REAL IMAGE)



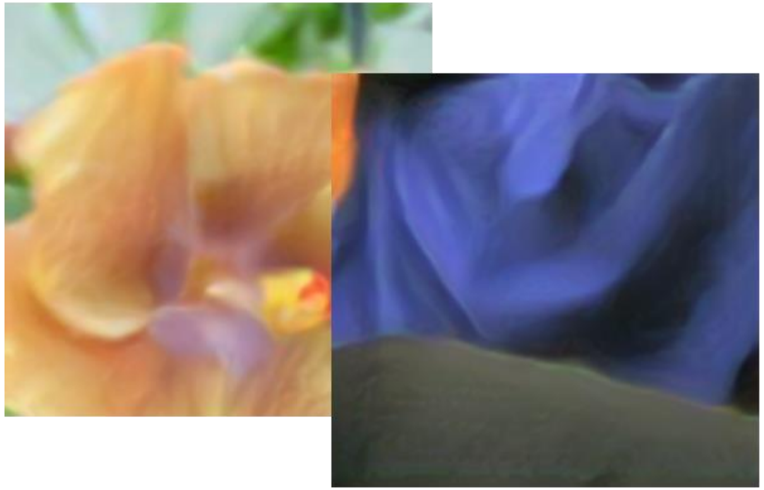
INPUT VS OUTPUT (REAL IMAGE)



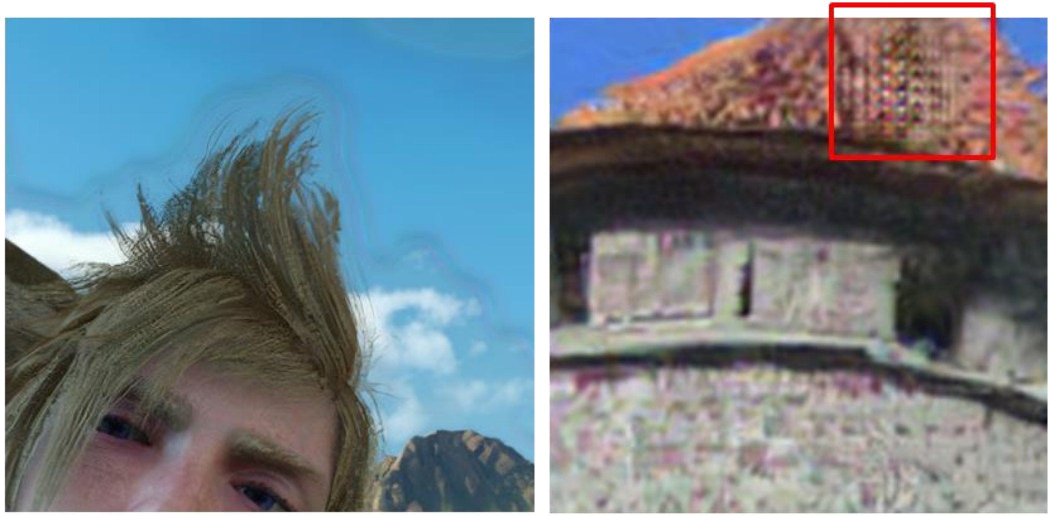
QUESTIONS?

THANK YOU!

DISCUSSION



VS



Low weight for GAN

High weight for GAN

Blurry image

GAN artifacts & color shift