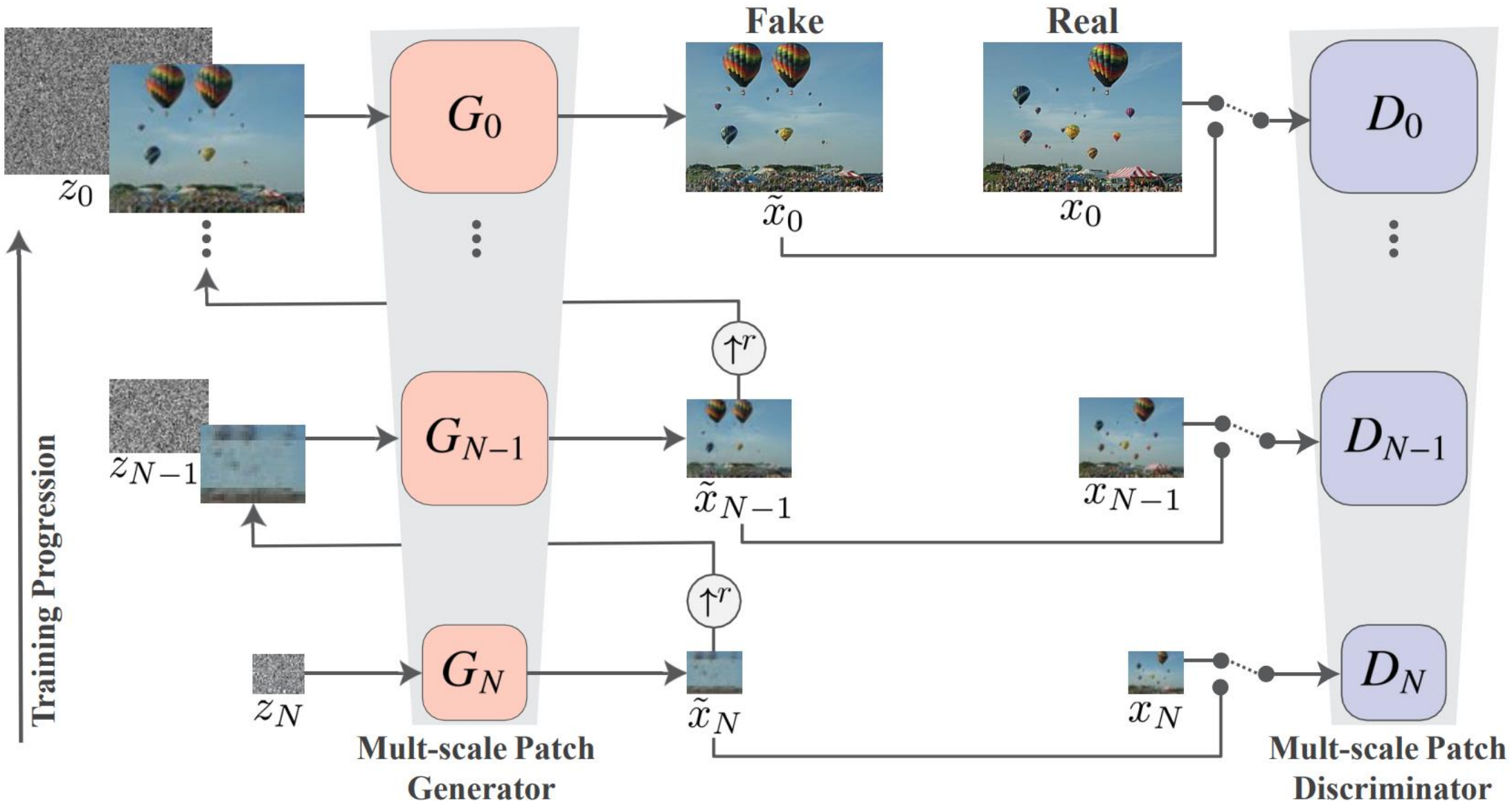
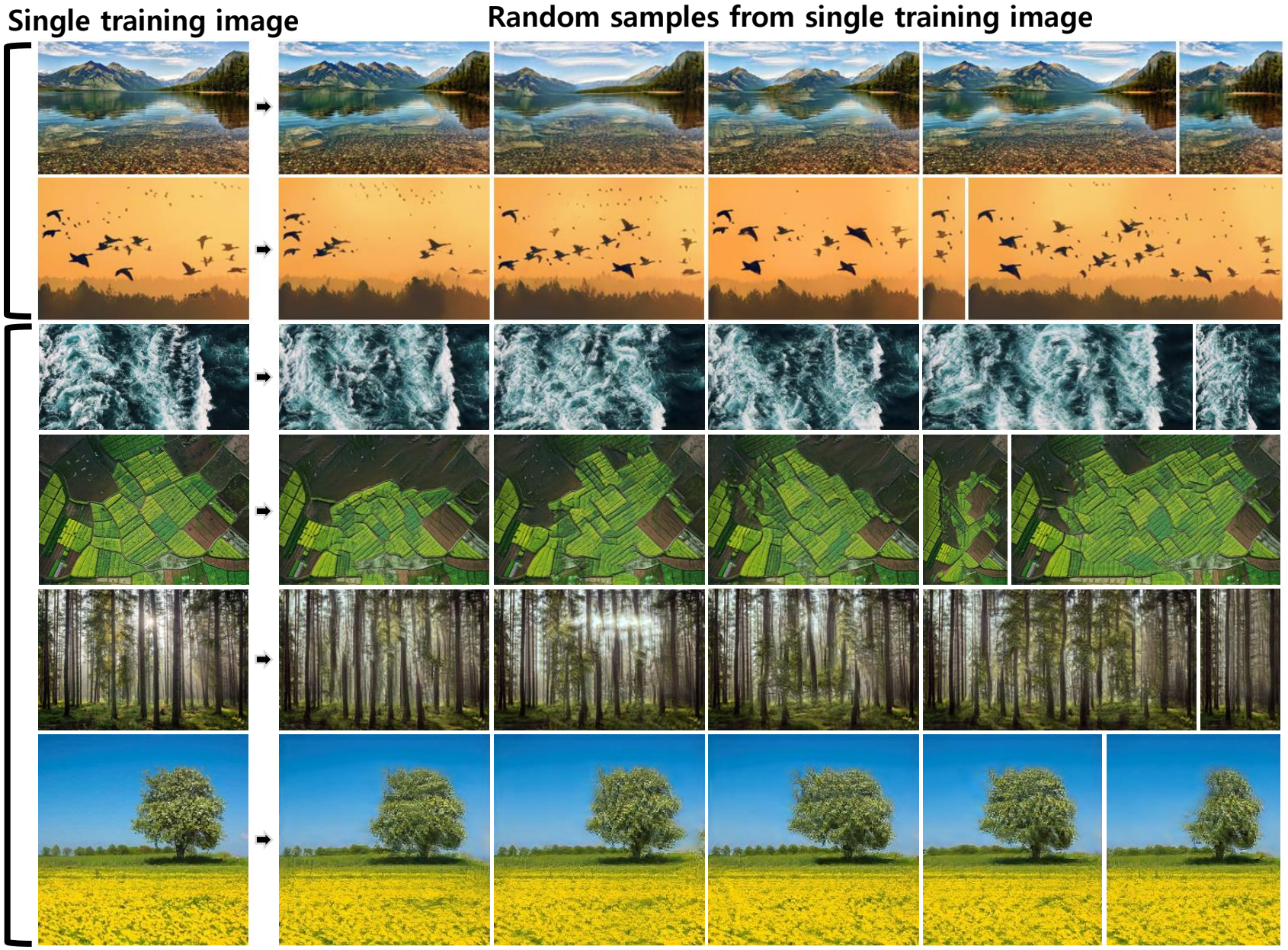
SinGAN: Learning a Generative Model from a Single Natural Image (CVPR 2019)

Sungwoo Son
Rehman Abdur
Rene Solzbacher

Overview of SinGAN



Reproduced Work (Random Samples)



Reproduced Work (Paint to Image)

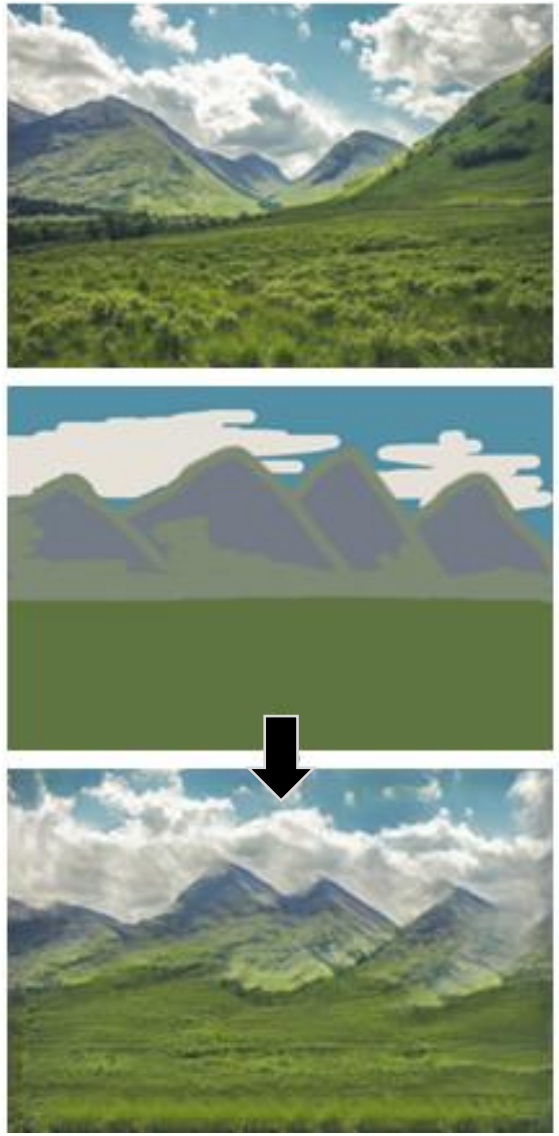
SinGAN Example

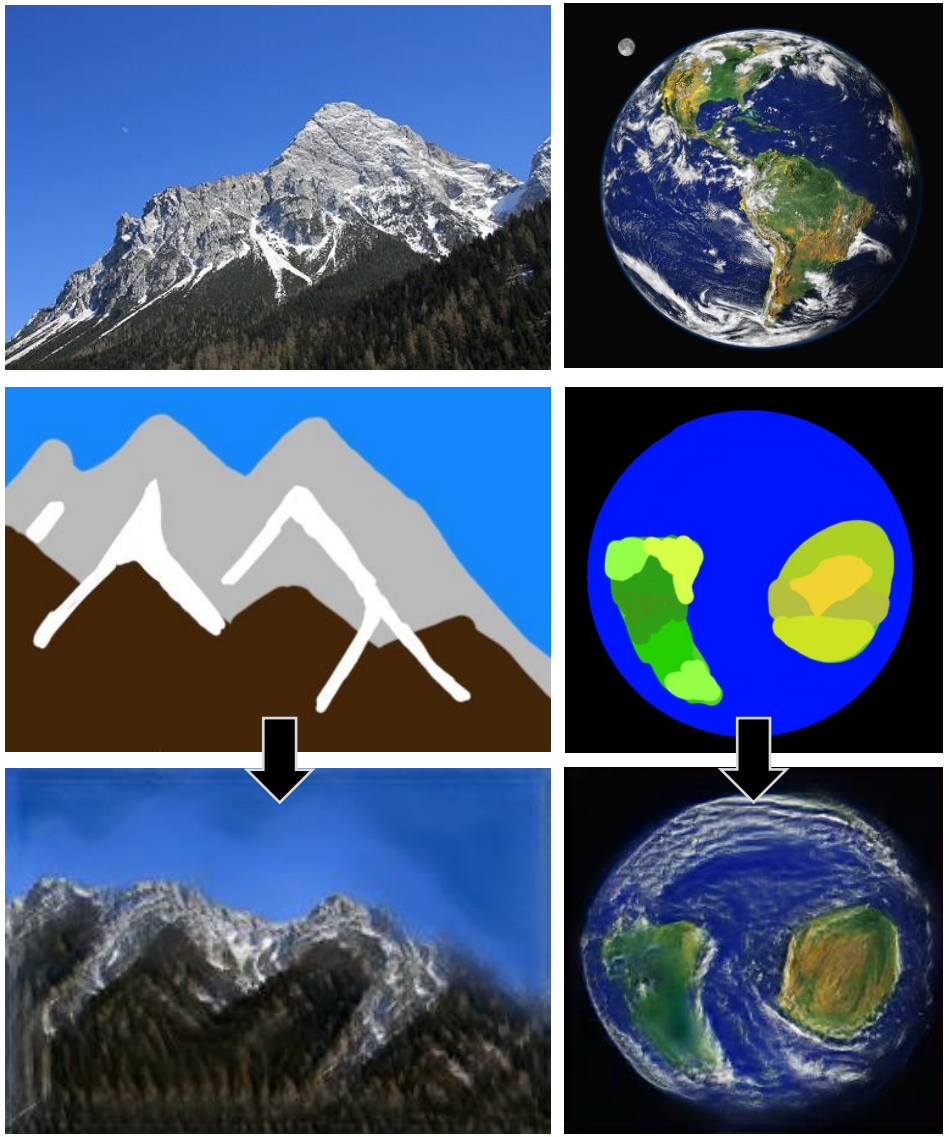
Reproduced Results

Training Image

Input

Output





Reproduced Work (Editing)

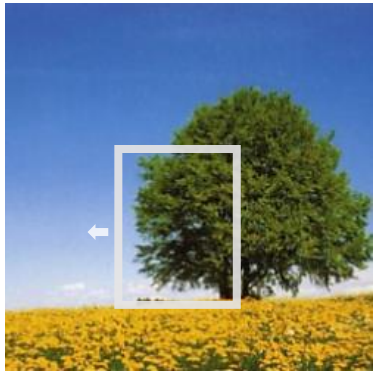
SinGAN Example

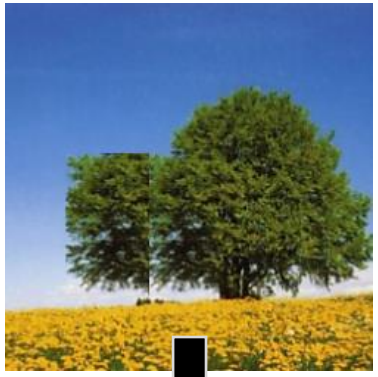
Reproduced Results

Training Image

Input

Output





Reproduced Work (Super Resolution)

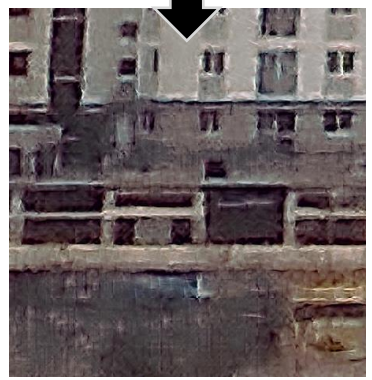
SinGAN Example

Training Image

Input

Output

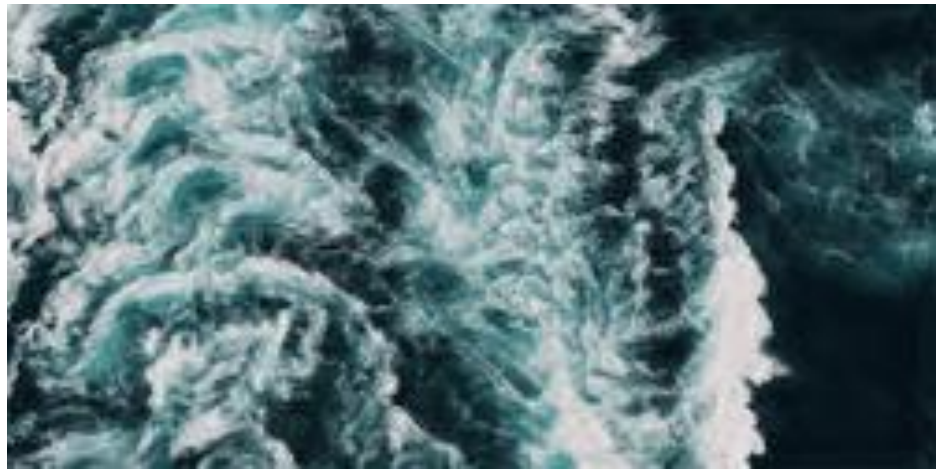
Reproduced Results



Reproduced Work (Animation)

SinGAN Example

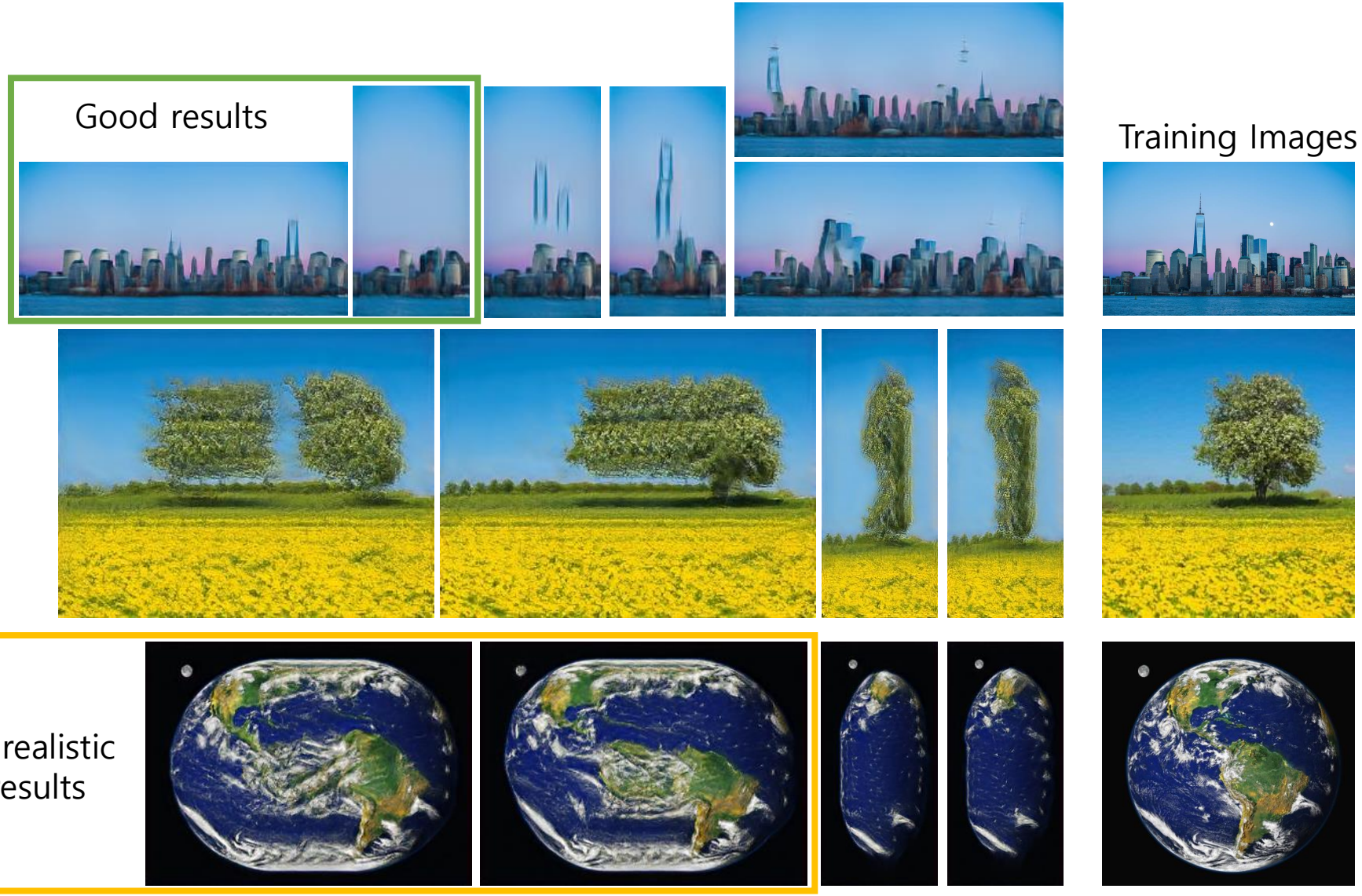
Reproduced Results



- Quality depends on subject
- Not all subjects are suitable

Limitations (Aspect Ratio)

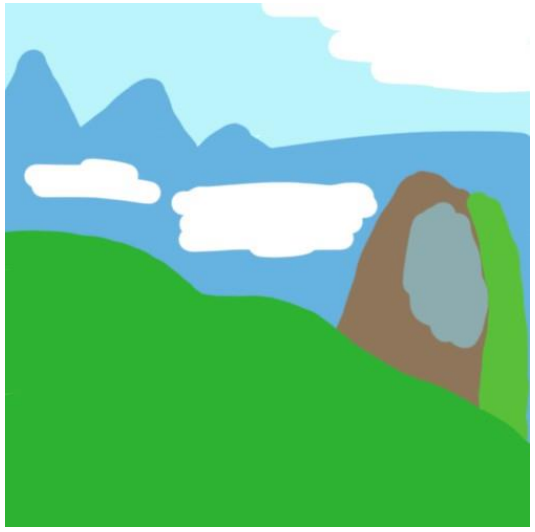
- **Only some** aspect ratio changes work
- Quality is dependent on subject and image composition
- Some changes may look good but are illogical



Limitations (Paint to Image)

Training Image

Input



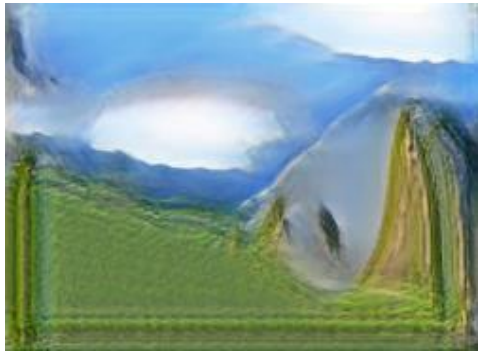
Output

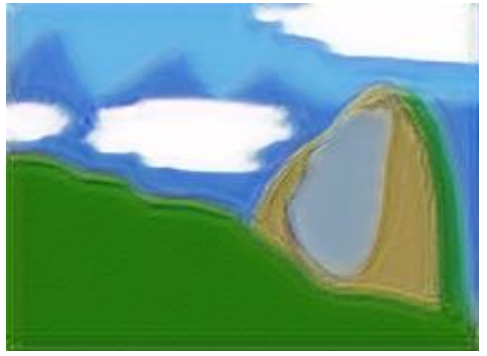
Scale 1

Scale 2

Scale 3

Output





Training Image

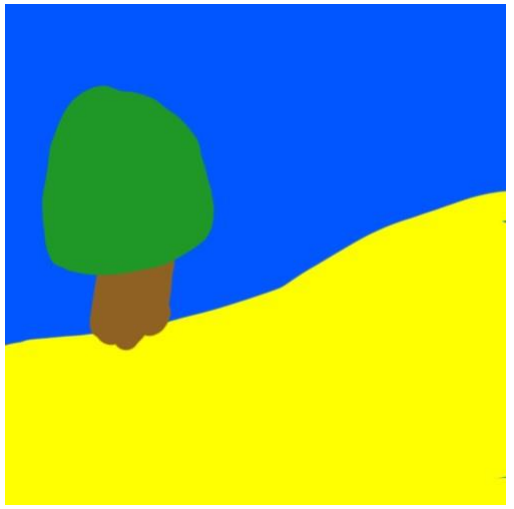
Input

Paint to Image Improvement

First Generation

Training Image

Input Image



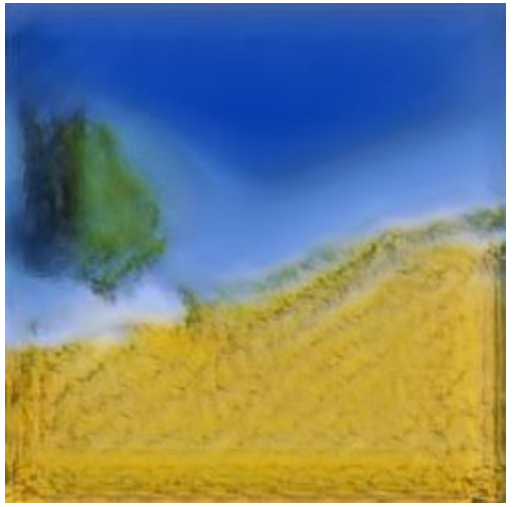
Output at scale 1

Second Generation
(Using output as input image)

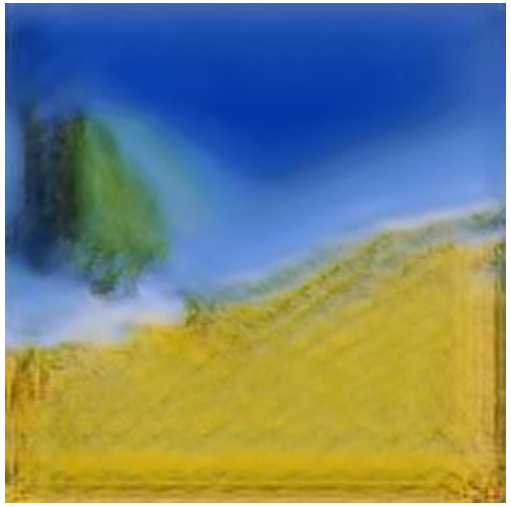
Training Image

Input Image

Output at scale 1



Output at scale 2



Output at scale 3

Proposed Work

x Limited to 250 pixels (longest side)
x Small images are less useful

Generate high resolution output images with SinGAN

- Increase image dimensions
- Keep comparable image quality
- Keep acceptable training times (currently 40-60 min)

- Change scale distribution
- Try image splicing
- Change hyperparameter
- Modify inner structure of generator

SinGAN Image Size

- SinGAN resizes training image to 250 pixels (longest side)
- Scales are based on resized image dimensions

```
0 = {Tensor: (1, 3, 25, 40)}
1 = {Tensor: (1, 3, 32, 51)}
2 = {Tensor: (1, 3, 40, 64)}
3 = {Tensor: (1, 3, 50, 80)}
4 = {Tensor: (1, 3, 63, 100)}
5 = {Tensor: (1, 3, 79, 126)}
6 = {Tensor: (1, 3, 100, 158)}
7 = {Tensor: (1, 3, 125, 199)}
8 = {Tensor: (1, 3, 157, 250)}
```

Input Image (585 x 366)

Resized Image (250 x 157)

Scale Distribution

- Lower scales (0,1) produce **major structural changes**
- Higher scales produce **fine detail**
- Scales are linearly implemented
- At each scale, the image is upscaled the same amount
- Anything beyond scale 2/3 produces too little change

Scale 1 Scale 2 Scale 3

Investigate different rescaling strategy that focuses on coarse scales

Author's High-Resolution Proposal

1. Train model to generate random samples
2. Train super resolution model to up-sample selected image

✓ SinGAN super resolution restores much detail

x Super resolution becomes problematic for large upscaling

Original crop

Down sampled (250px)

SR from 250px crop

High-Resolution (Image Splicing)

Structures largely remain the same towards the image edges

1. Split image into sections
2. Train model for each section
3. Generate random sample for each section
4. Train high resolution model
5. Splice sections together

Single generation

≠

Individual generation

?

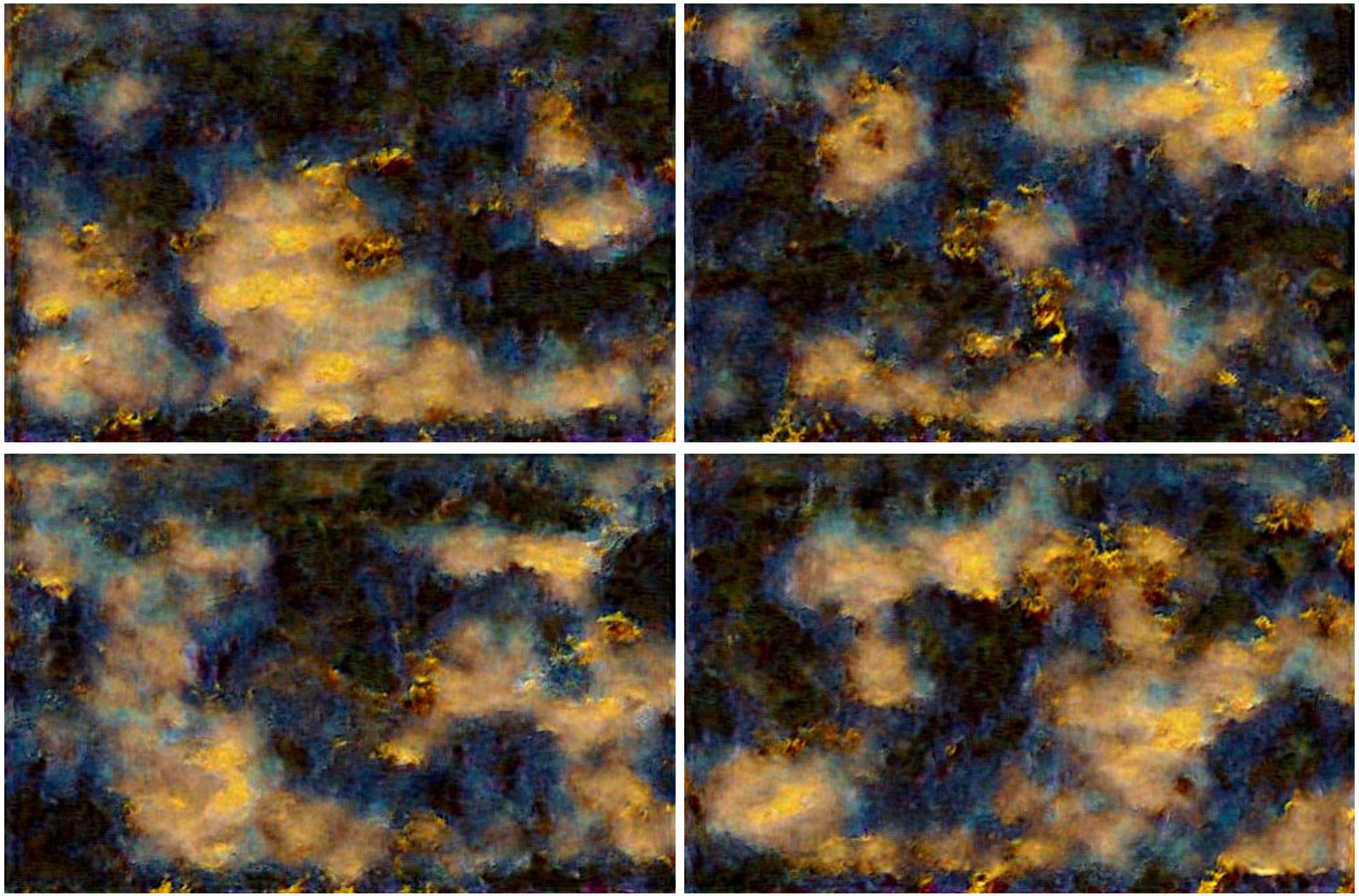
High-Resolution (Hyperparameter Changes)

- Values to fine-tune SinGAN
  - Receptive field size — default: 11x11px
  - Rescaling resolution — default: 250px
  - Conv kernels — default: 32
  - Learning rate — default: 0.0005
  - Epochs per scale — default: 2000

- Option to change the behaviour and performance of SinGAN
- SinGAN is optimized for images of **250px size**
- We want to optimize it for images of **500px size**

High-Resolution (Hyperparameter Changes)

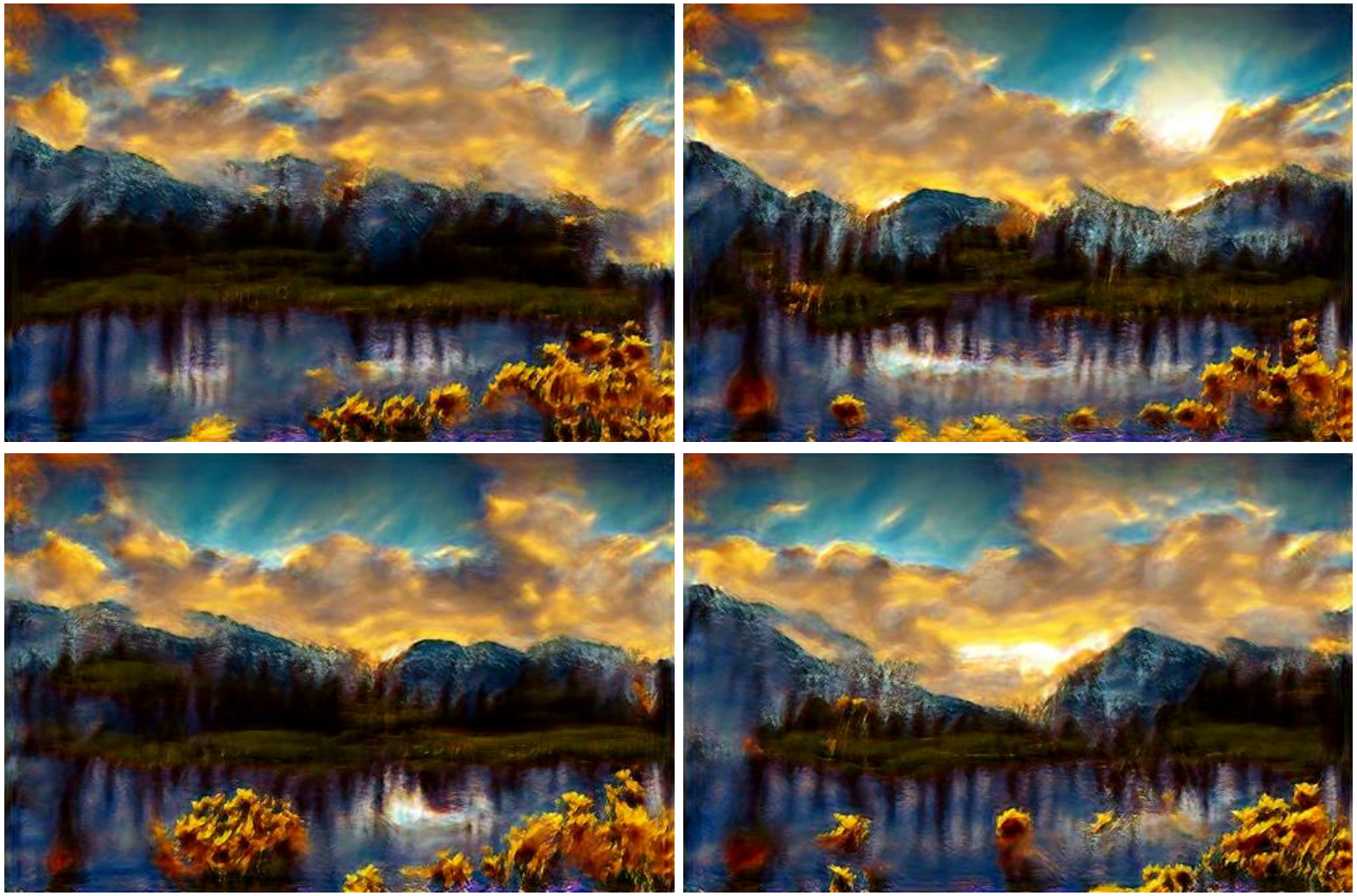
741 x 486



400 x 263

High-Resolution (Hyperparameter Changes)

741 x 486



400 x 263

High-Resolution (Hyperparameter Changes)

741 x 486

400 x 263

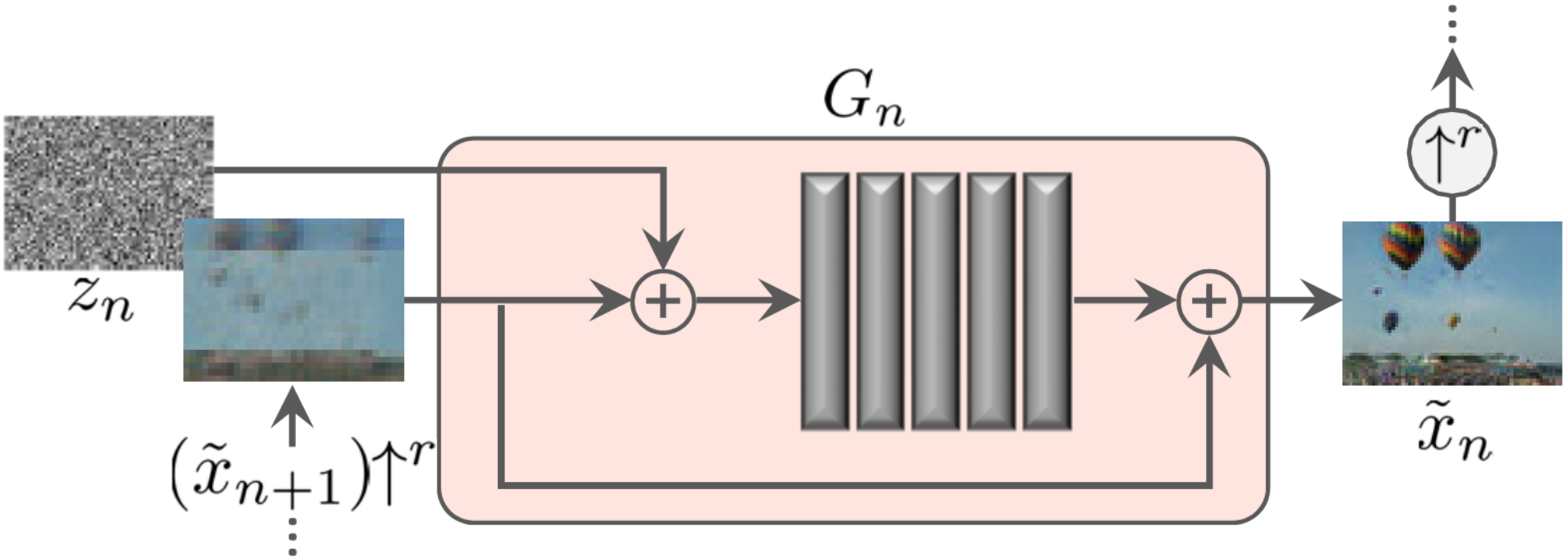
High-Resolution (Hyperparameter Changes)

- Hyperparameters that can increase quality:
  - Padding size
  - Scaling factor for up/down sampling images
  - Learning rate
  - Reconstruction loss weight (alpha)

Future Work

- Automate hyperparameters to adjust to image size automatically
  - SinGAN is only optimized for 250px images
  - Only basic starting parameters need to be set by user, the rest is automatically scaled for the desired image resolution

Modify Generator Structure



- Up-sampling in the structure and outside the structure
- Use Resnet for conformity to the original training image

Thank you