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

Sungwoo Son
Rehman Abdur
Solzbacher Rene Marcel

Introduction

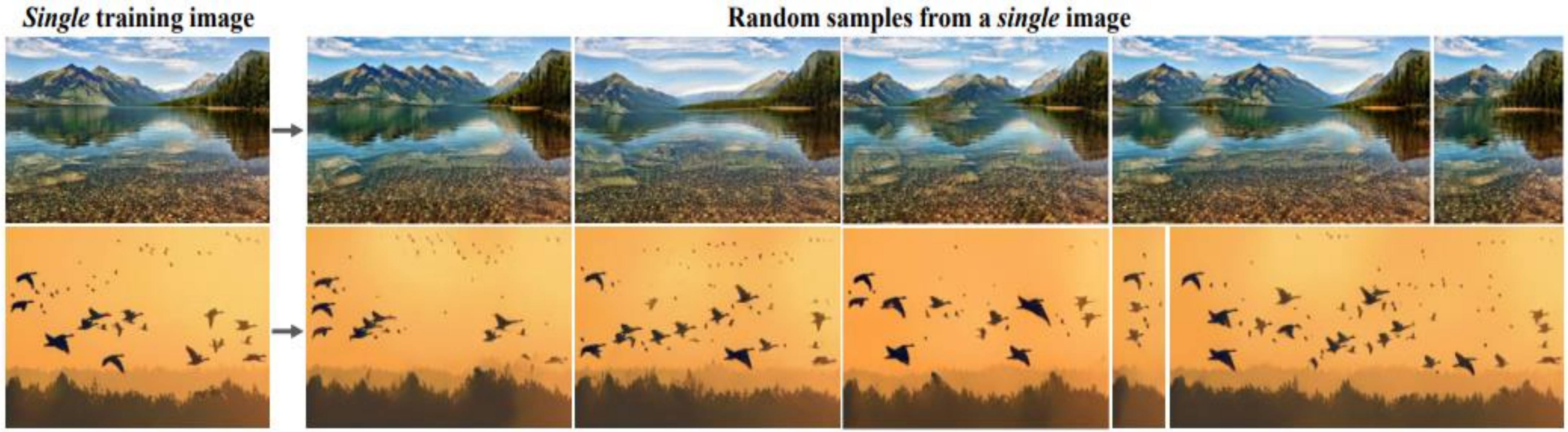
- Unconditional GANs have shown remarkable progress since their inception
- Learning to generate highly complex scenes requires a large amount of data

- What if we have a single image?

- A generative model that can learn form single image

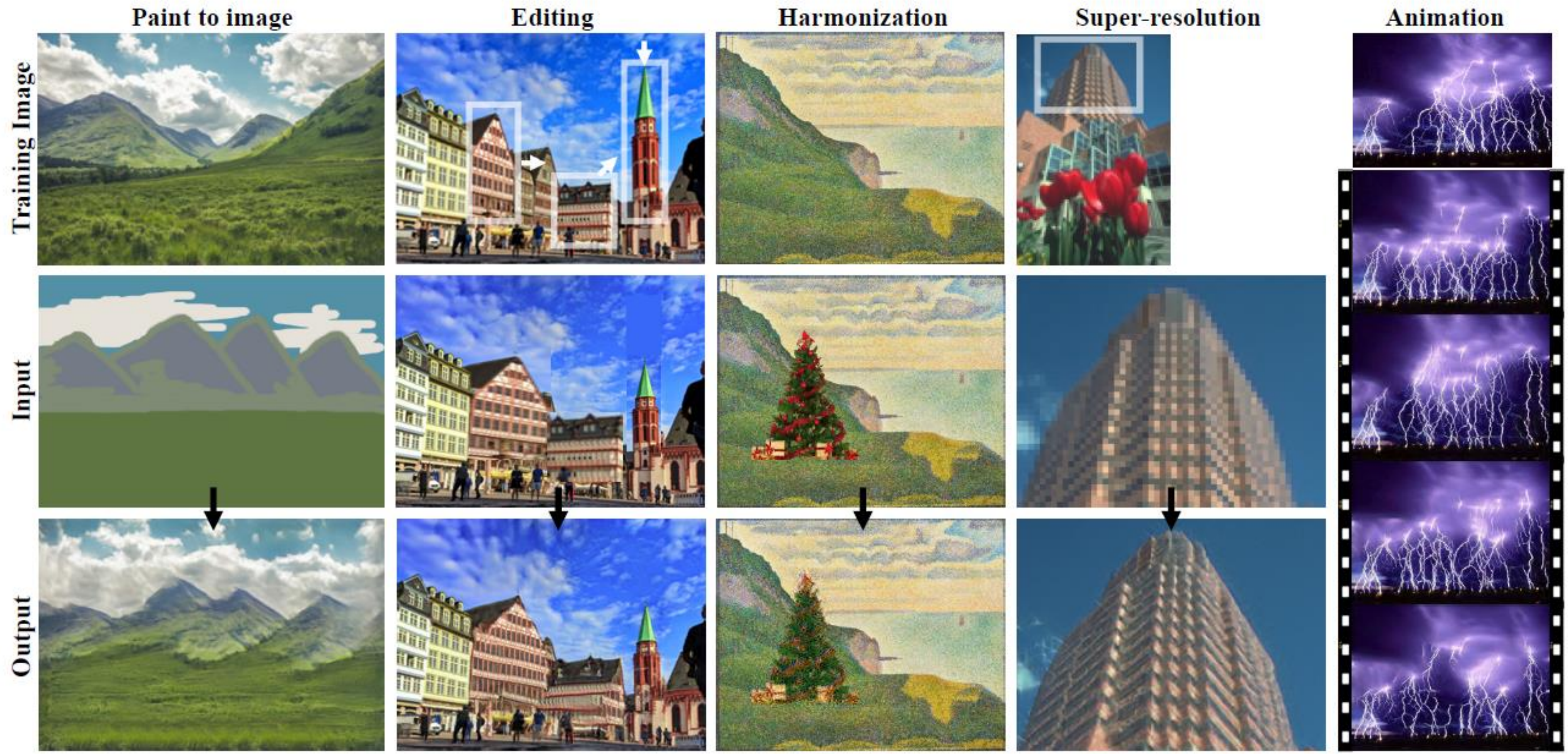
Introduction

- Generates high quality, diverse samples that contain same visual content as input image



- Learns about patch distribution at different scales of the image

Introduction



Related work

- Unconditional GANs have been explored only in the context of texture generation
- [Gatys et al.' 2015] **Texture generation**
- [Zhou et al.' 2018] **Texture expansion**

- These models do not produce meaning full results on non texture images

SinGAN Advantages

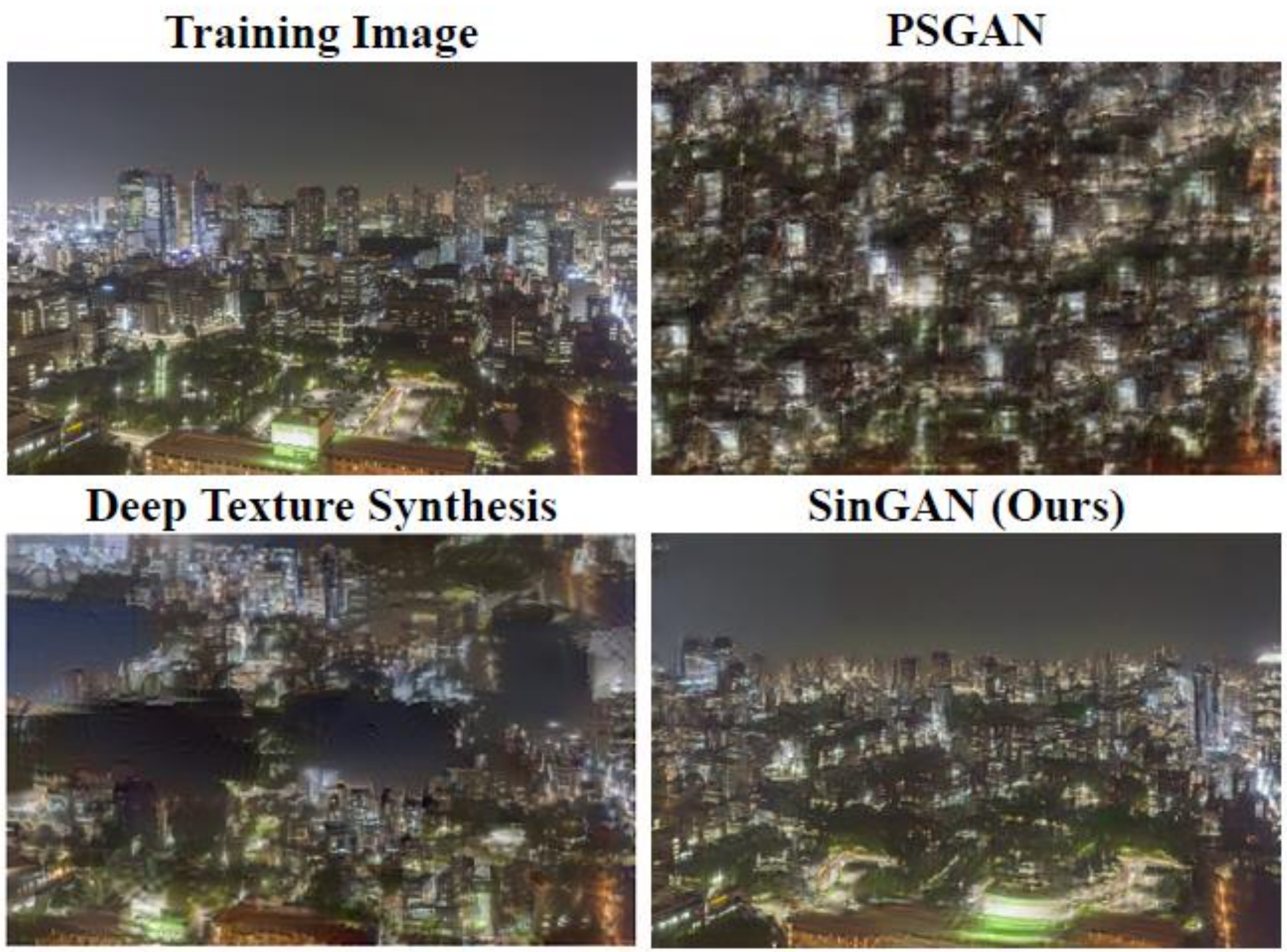
Classical GANs

- Requirement of large image dataset
- Dataset specific results
- Application specific (Super Resolution (SR), Texture synthesis)

SinGAN

- Not dependent on anything
- Given a random image, it can generate random samples via absorbing the patch distribution properties
- Results are not application or dataset specific
- Hence, useful in a wide range of applications

SinGAN Advantages



PSGAN, Deep Texture Synthesis

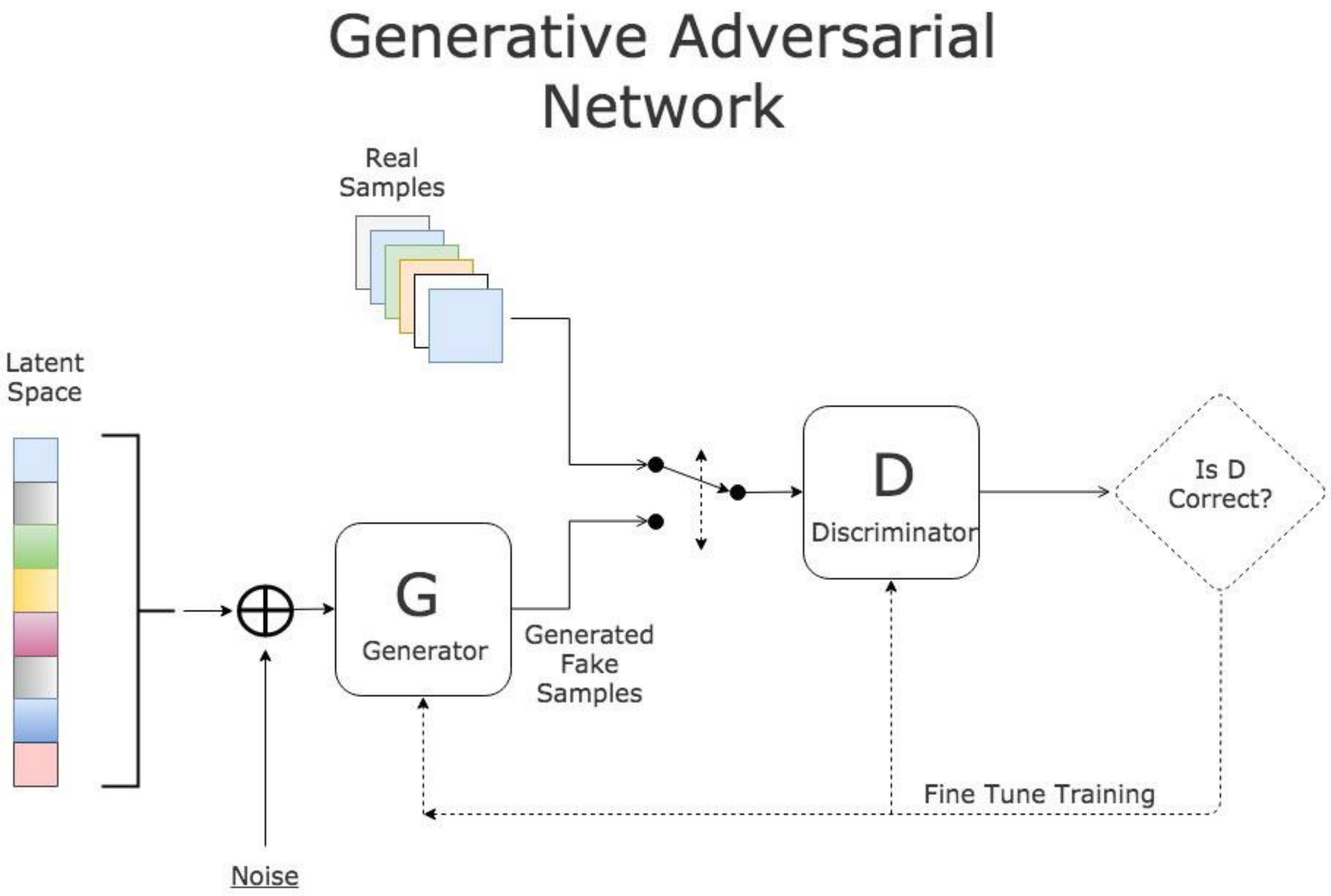
- Designed for specific task *(e.g., super resolution, texture expansion)*
- Limited to texture images

SinGAN

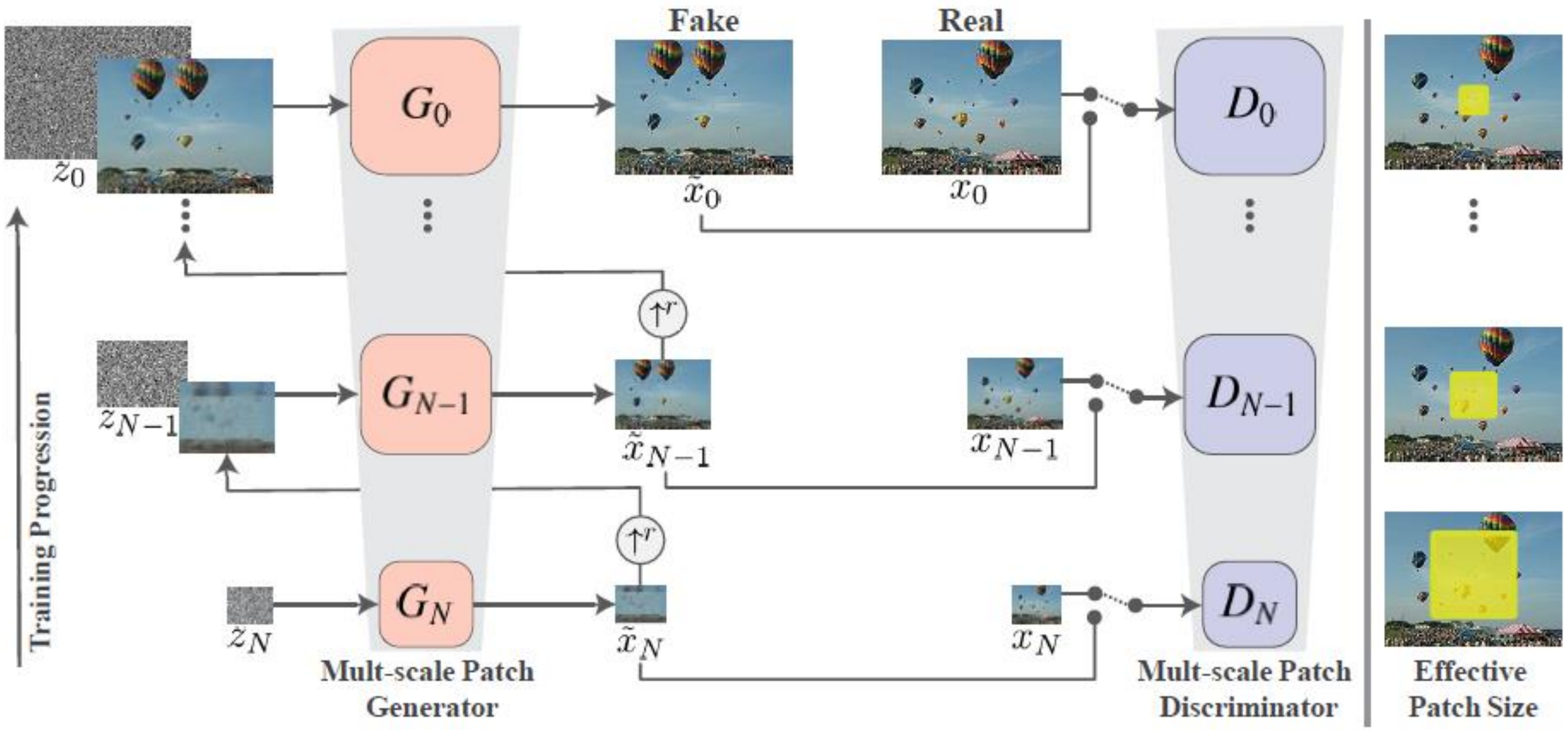
- Maps noise to image samples
  - Suits different image manipulation tasks
- Applicable to natural images (not only textures)

Model

DGIST 대구경북과학기술원 Daegu Gyeongbuk Institute of Science & Technology

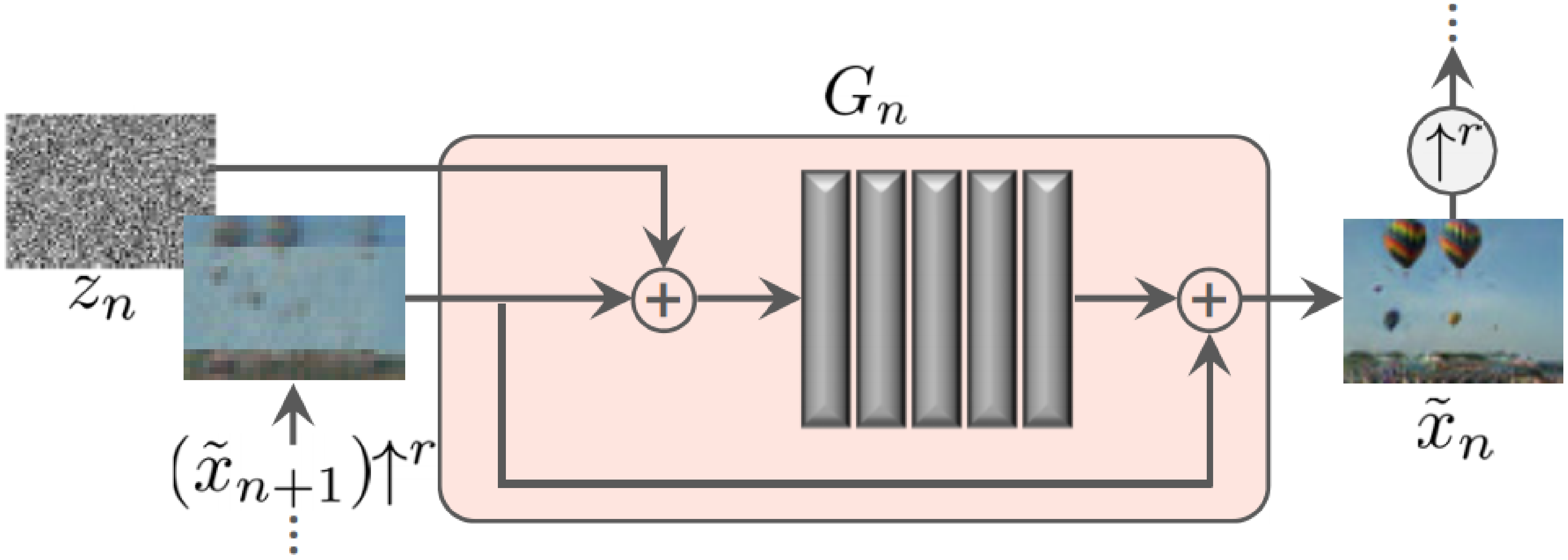


Model



Fake
Real
G_0
$\tilde{x}_0$
x_0
D_0
z_0
z_{N-1}
G_{N-1}
$\tilde{x}_{N-1}$
x_{N-1}
D_{N-1}
$\uparrow^r$
z_N
G_N
$\tilde{x}_N$
x_N
D_N
Training Progression
Mult-scale Patch Generator
Mult-scale Patch Discriminator
Effective Patch Size

Model(single scale generation)



Model

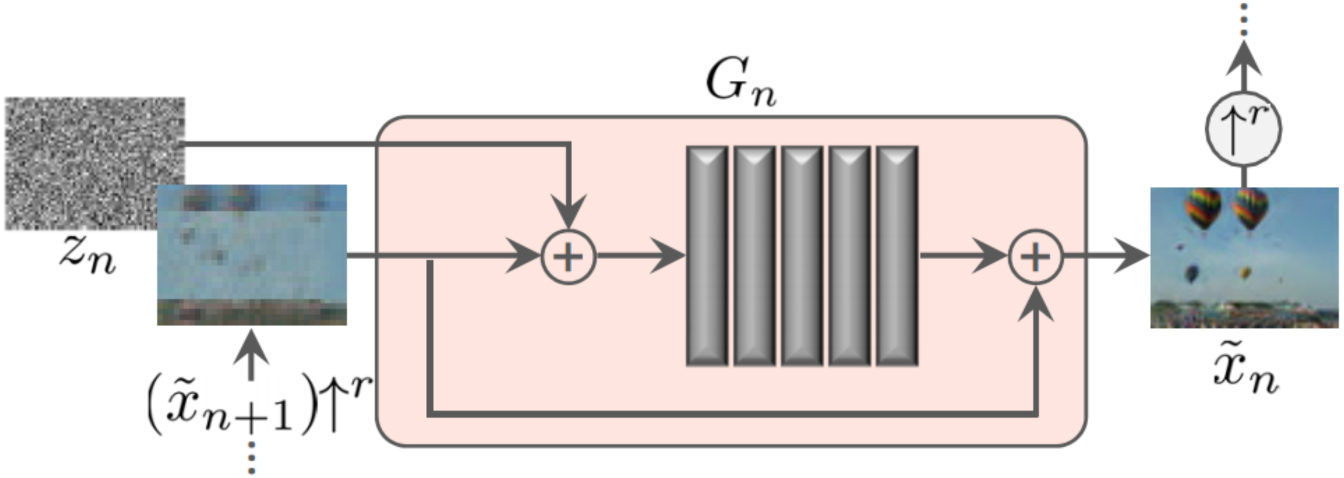
- Additional reconstructive loss added to generative loss

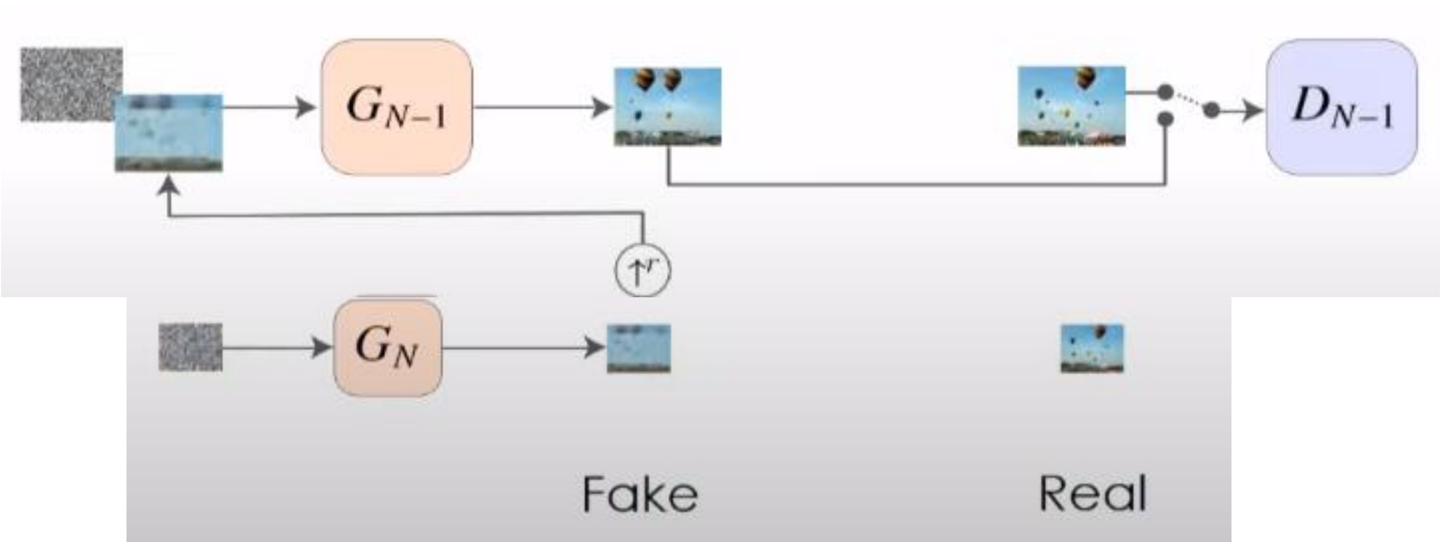
$$\min_{G_N} \max_{D_N} \mathcal{L}_{\mathrm{adv}}(G_N, D_N) + \alpha\mathcal{L}_{\mathrm{rec}}(G_N)$$

- Reconstruction loss ensures that at a particular input noise, model will generate original image
- This way we can include original image in our latent space

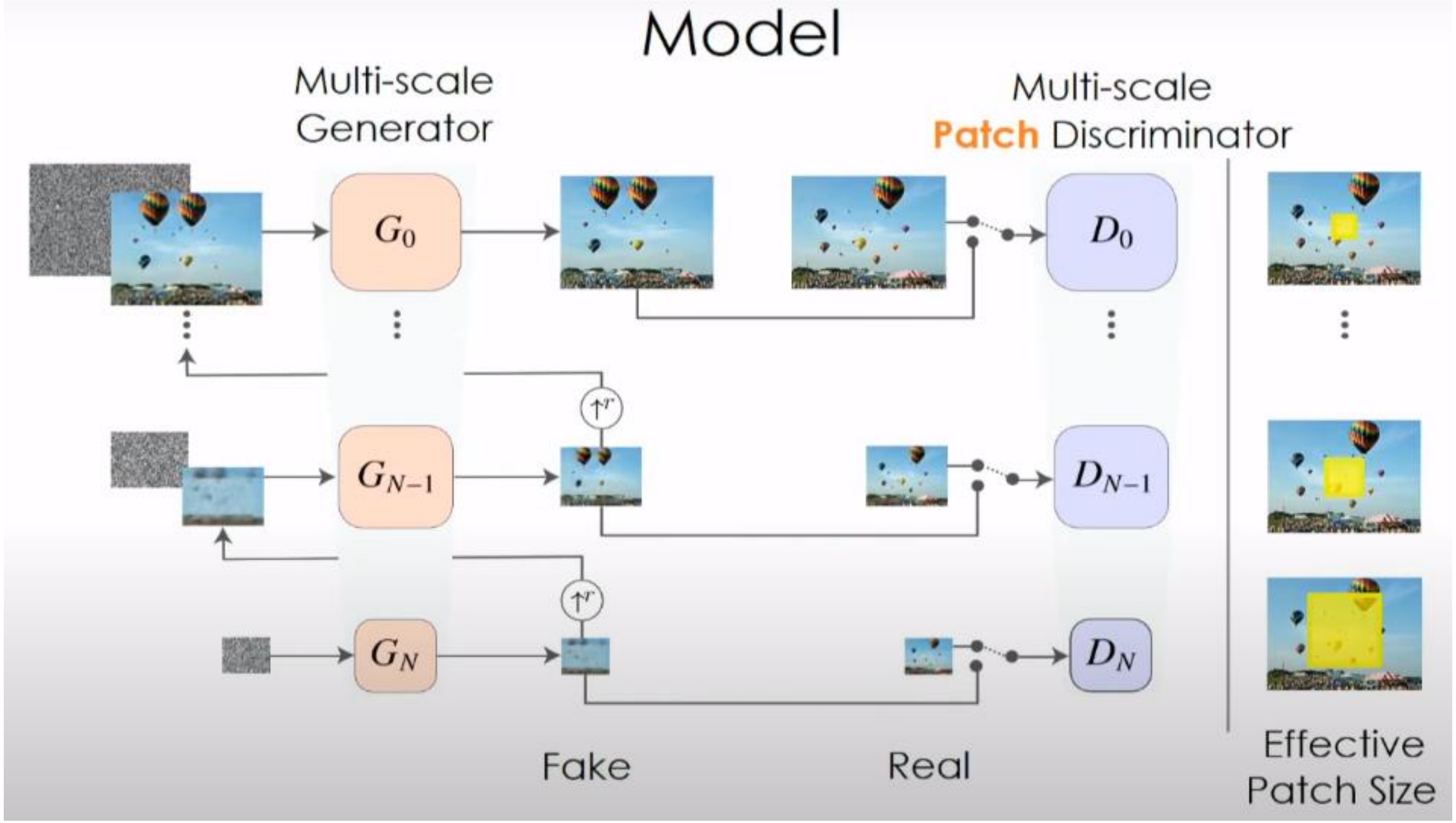
Model

Single Scale Architecture





Model

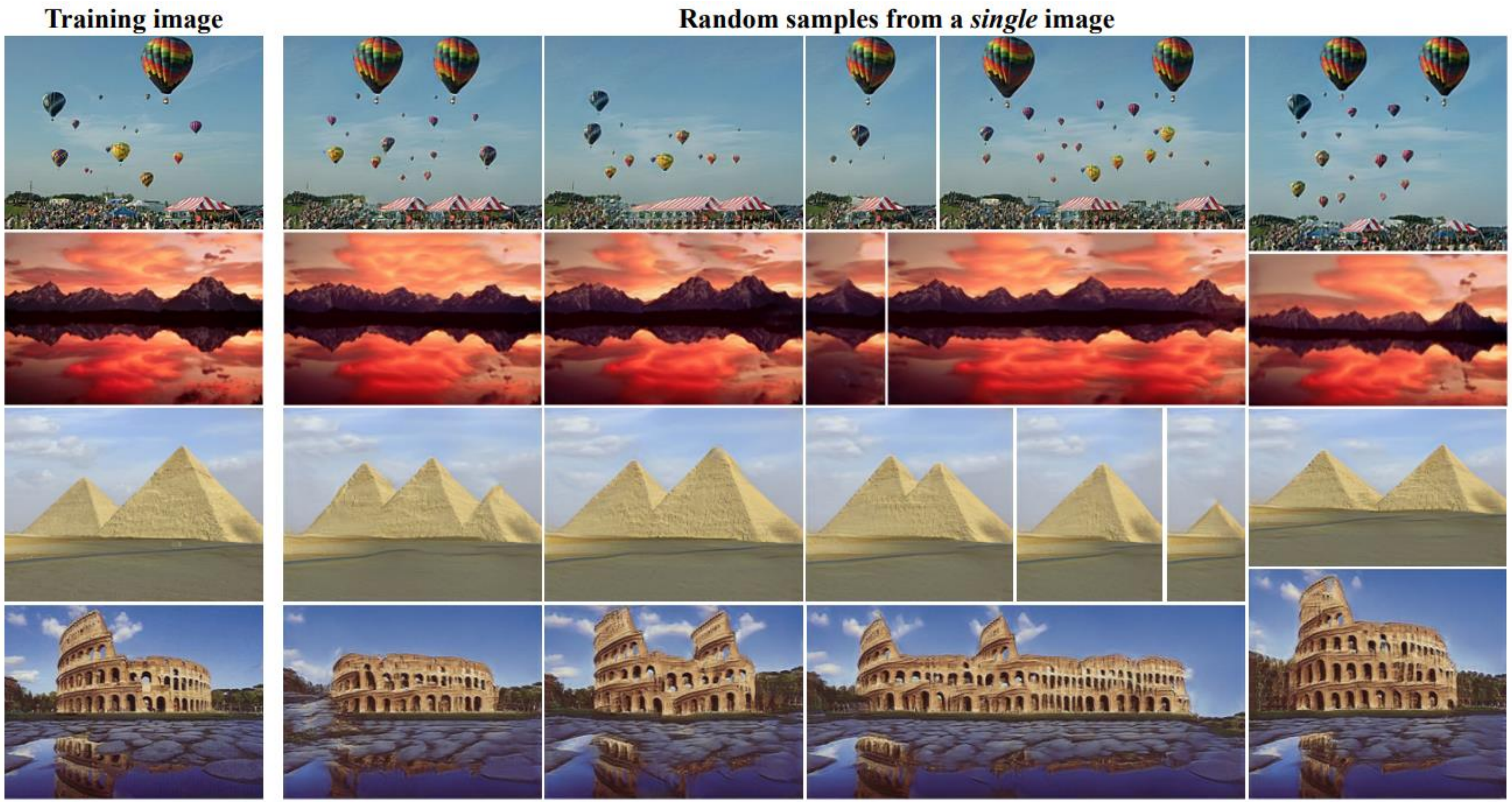


Model
Multi-scale
Generator
Multi-scale
Patch Discriminator
G_0
G_{N-1}
G_N
D_0
D_{N-1}
D_N
$\uparrow^r$
$\uparrow^r$
Fake
Real
Effective
Patch Size

Model

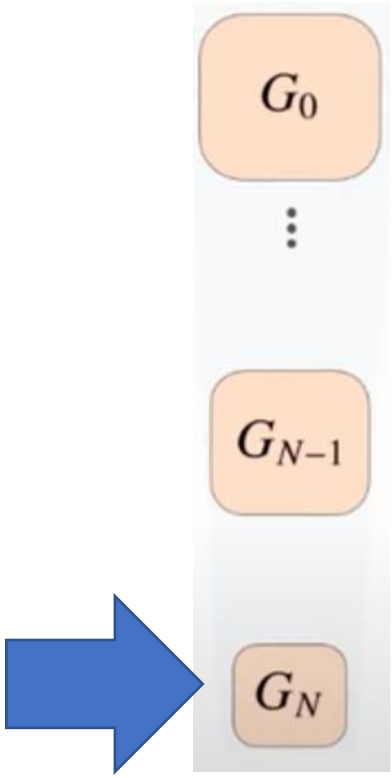
DGIST 대구경북과학기술원 Daegu Gyeongbuk Institute of Science & Technology

- SinGAN architecture is Fully Convolutional so we can inject noise of varying aspect ratio to obtain desired image.



Model

Effect of Scale at Test Time

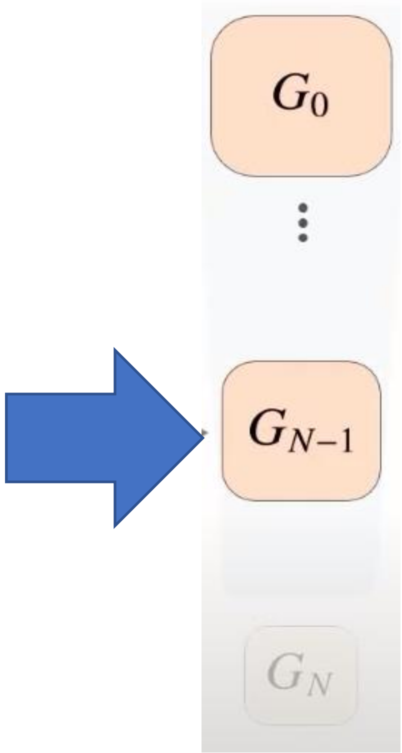


Multi-scale
Generator

Single training image

Generated images

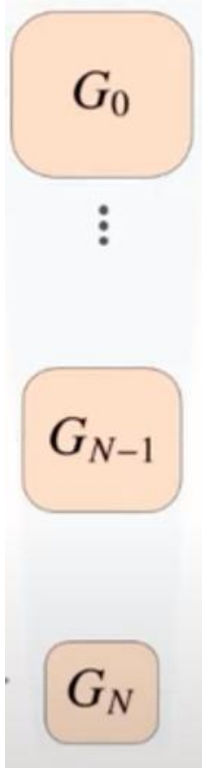
Model



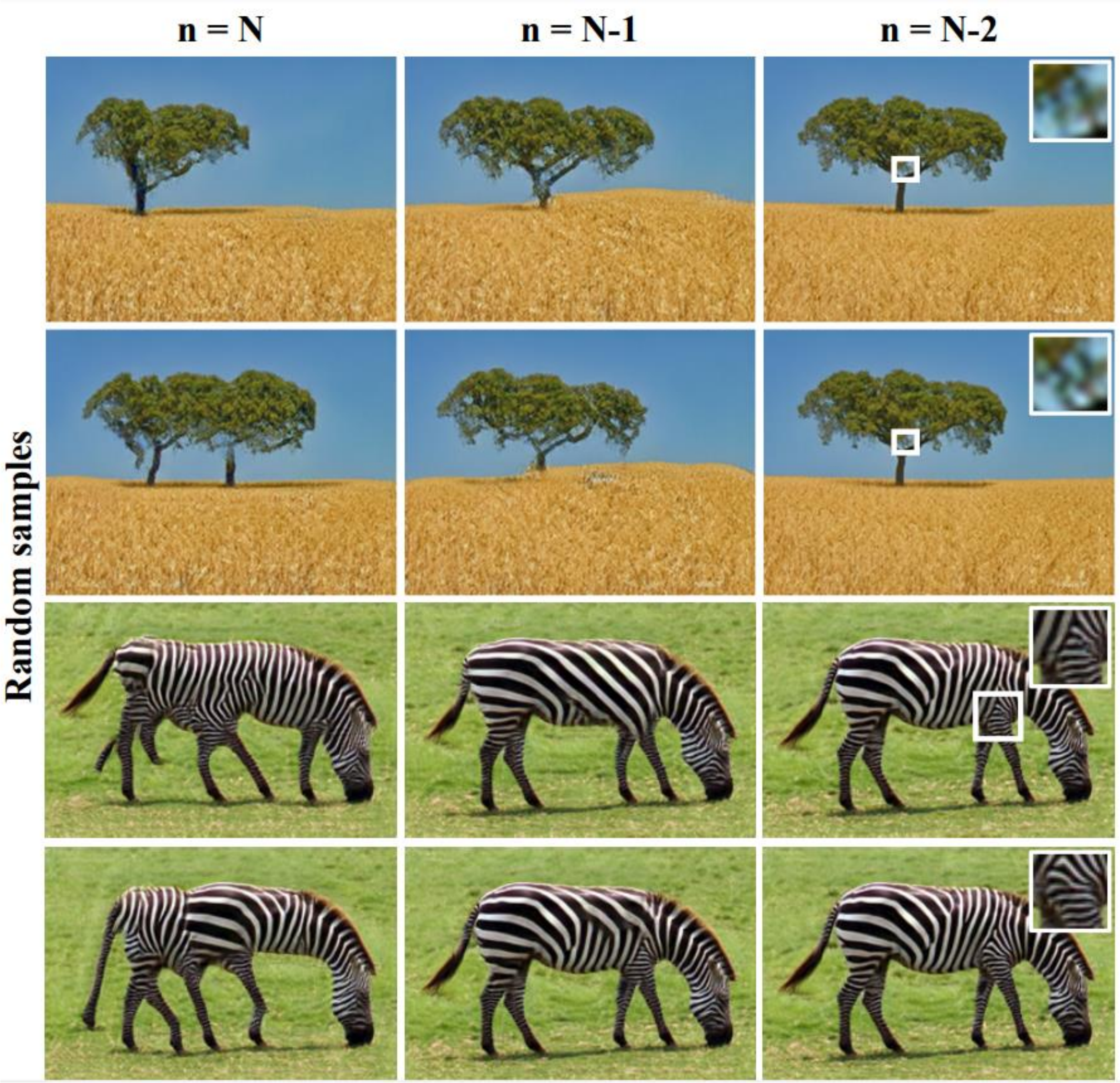
Single training image

Generated images

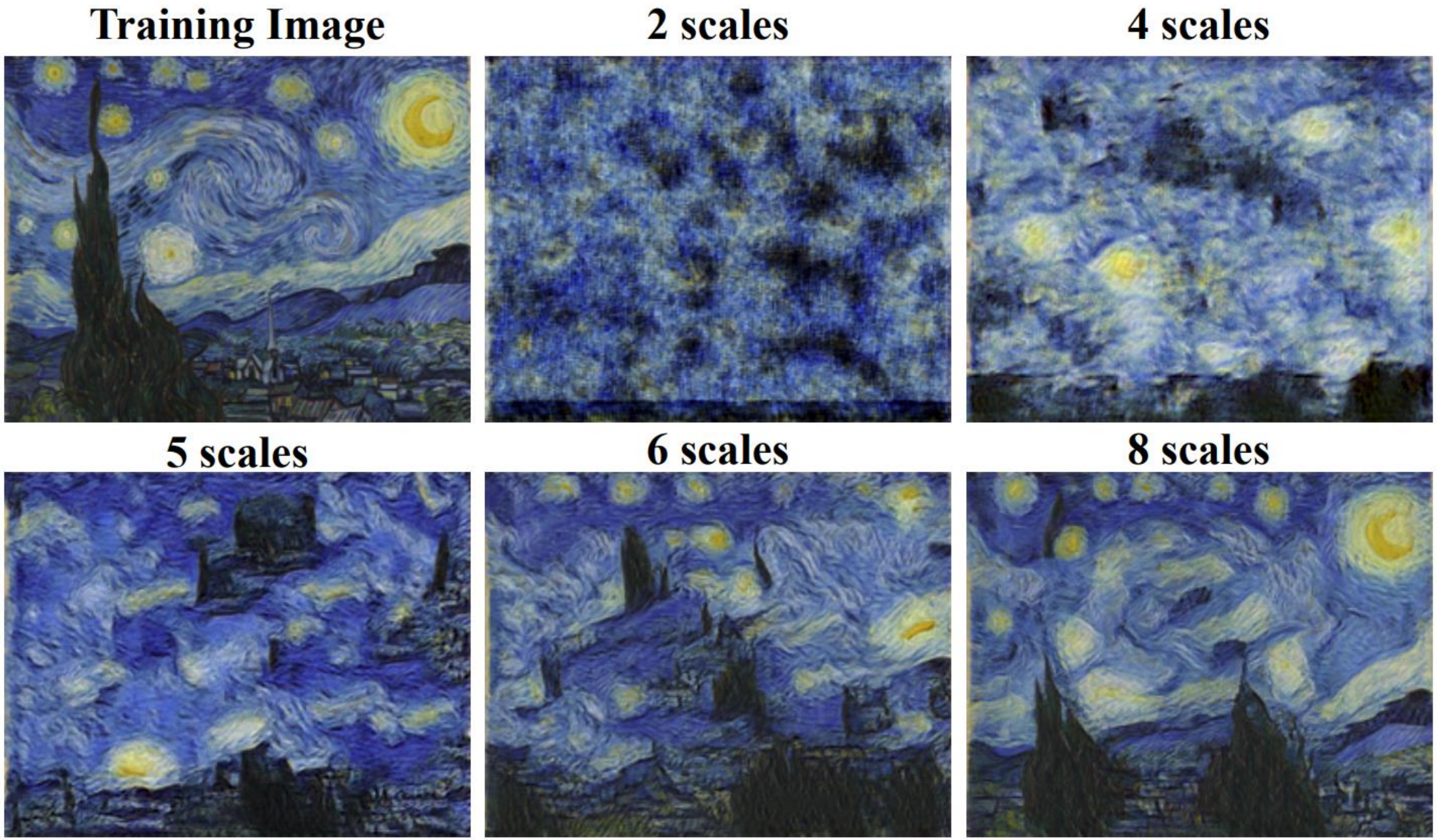
Model



Multi-scale
Generator

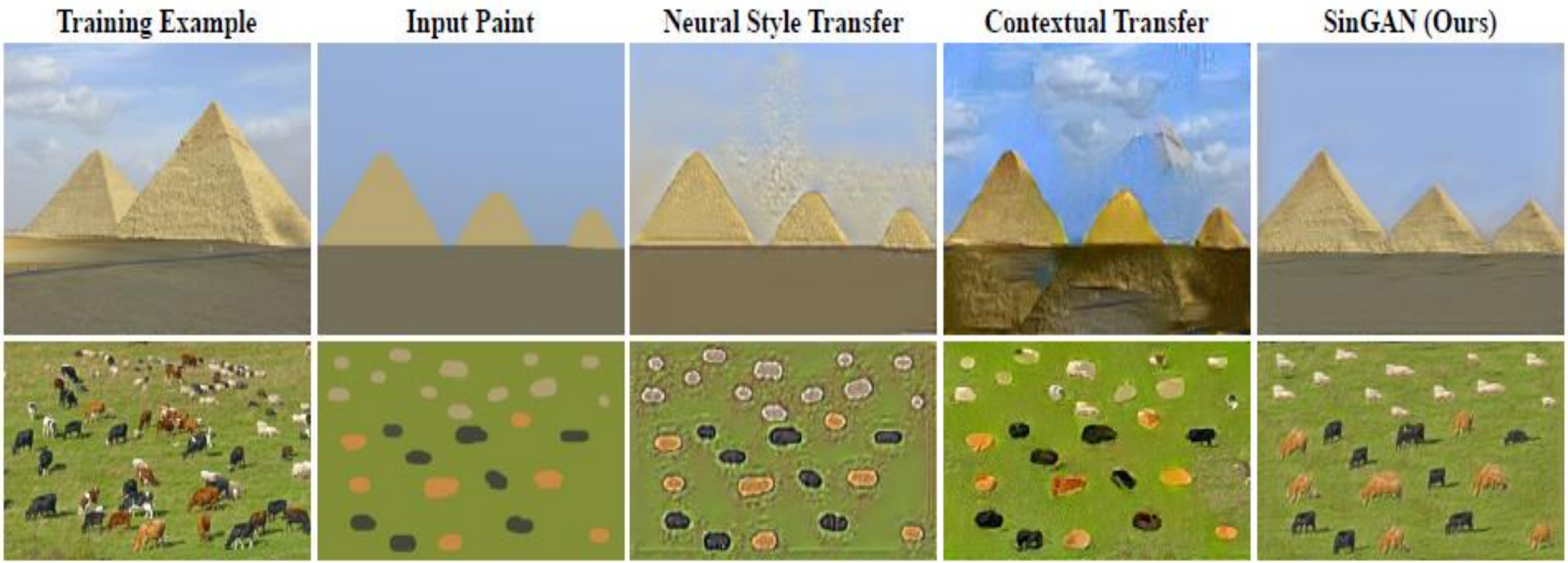


Effect of Scales during Training



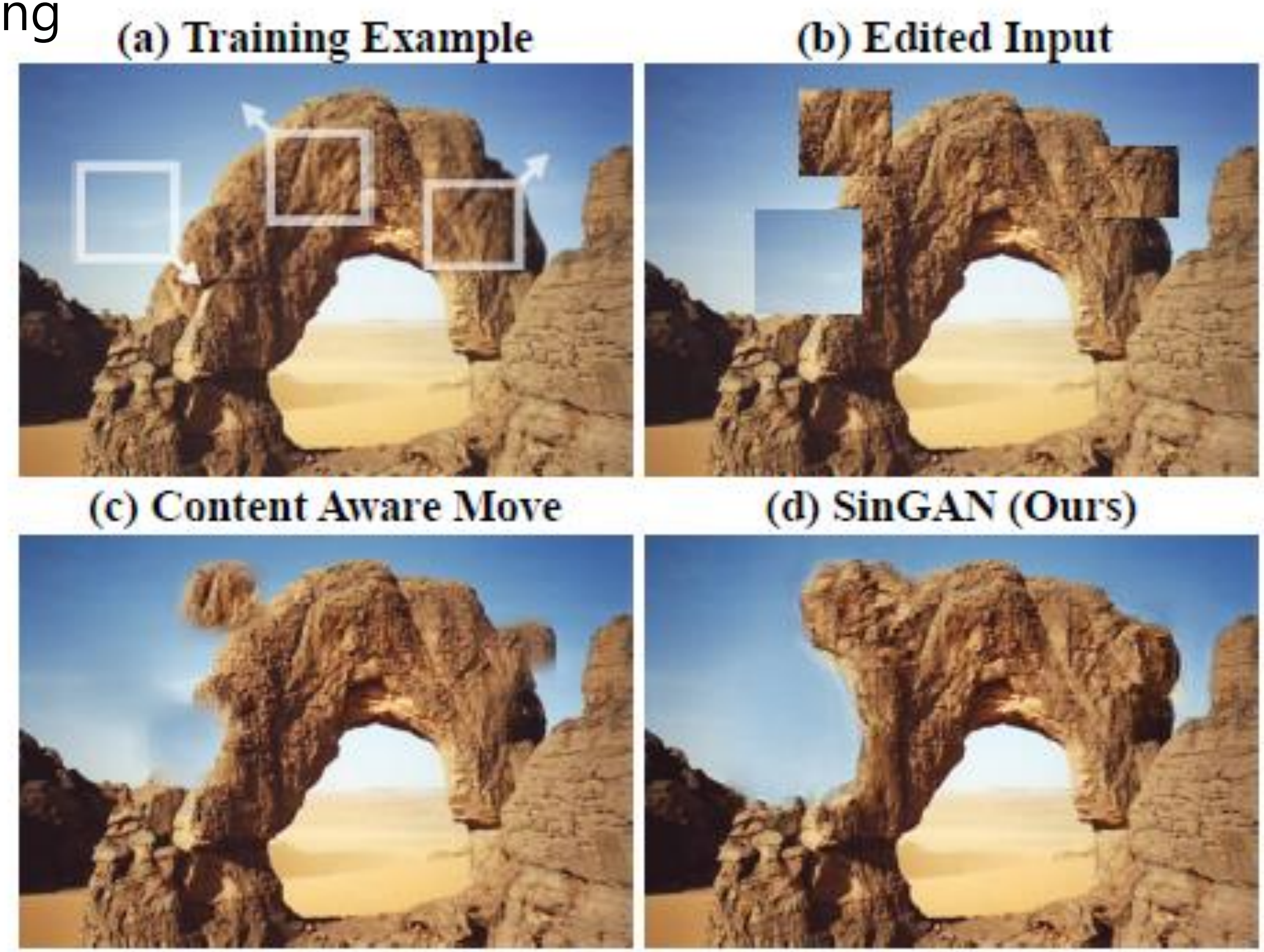
SinGAN Application

▶ Paint to Image



SinGAN Application

▶ Image Editing



SinGAN Application

▶ Harmonization

SinGAN Application

▶ Animation   [SinGAN for single image animation - YouTube]

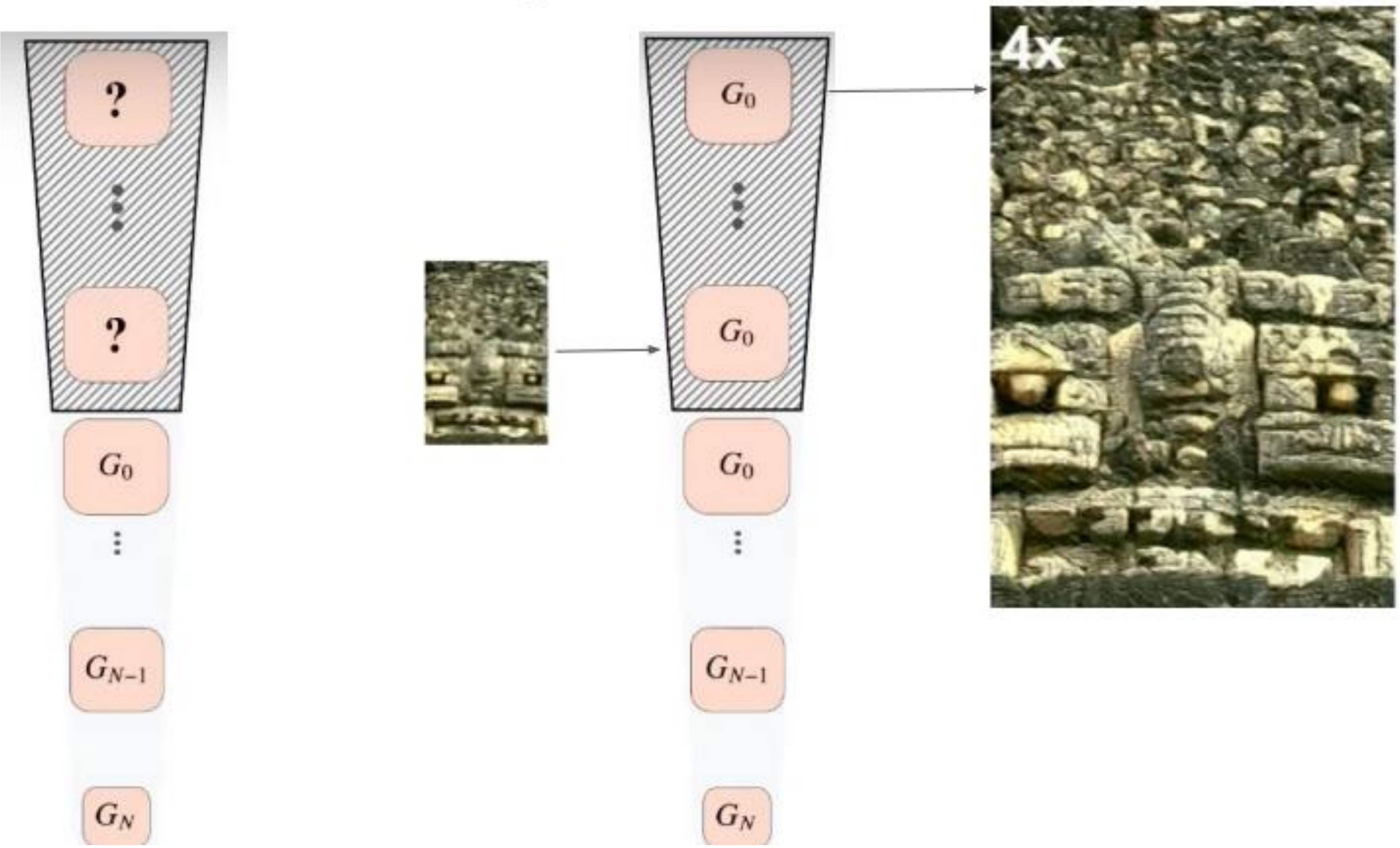
SinGAN Application

▶ Super Resolution

	External methods		Internal methods		
	SRGAN	EDSR	DIP	ZSSR	SinGAN
RMSE	16.34	12.29	13.82	13.08	16.22
NIQE	3.41	6.50	6.35	7.13	3.71

- distortion quality RMSE↓
- perceptual quality NIQE↓

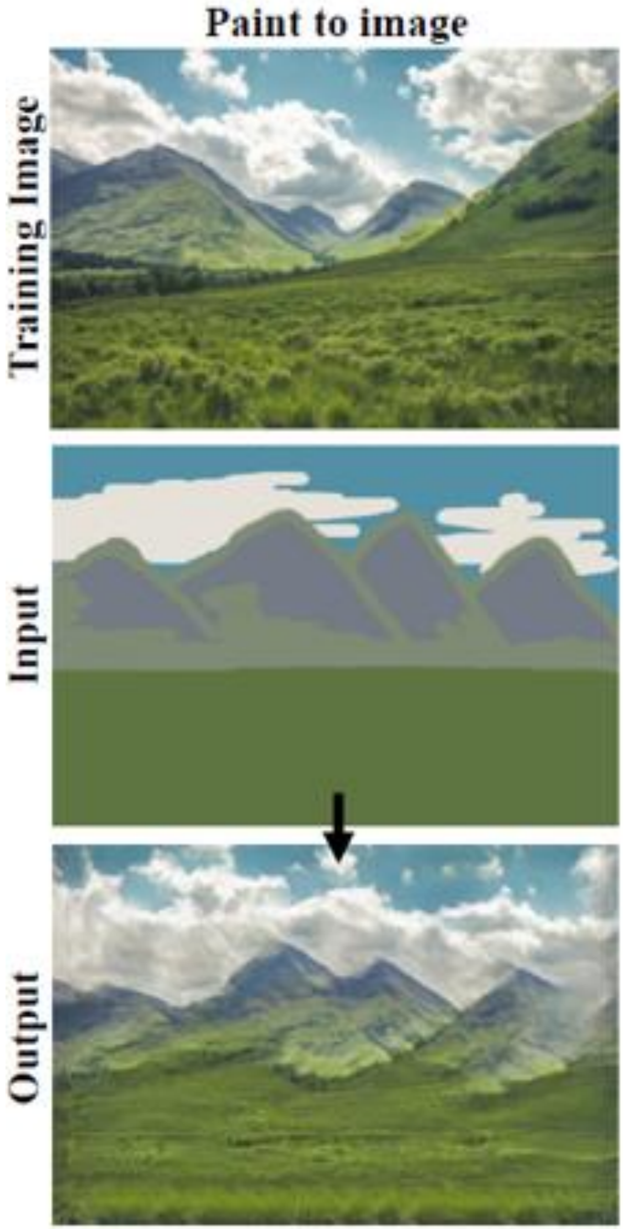
SinGAN Application

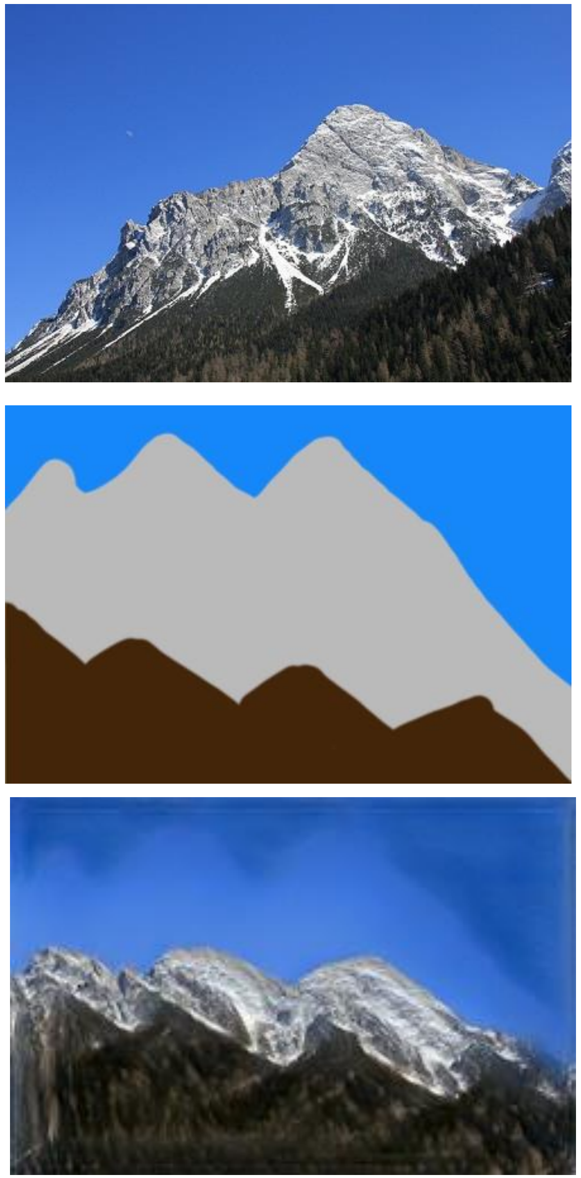


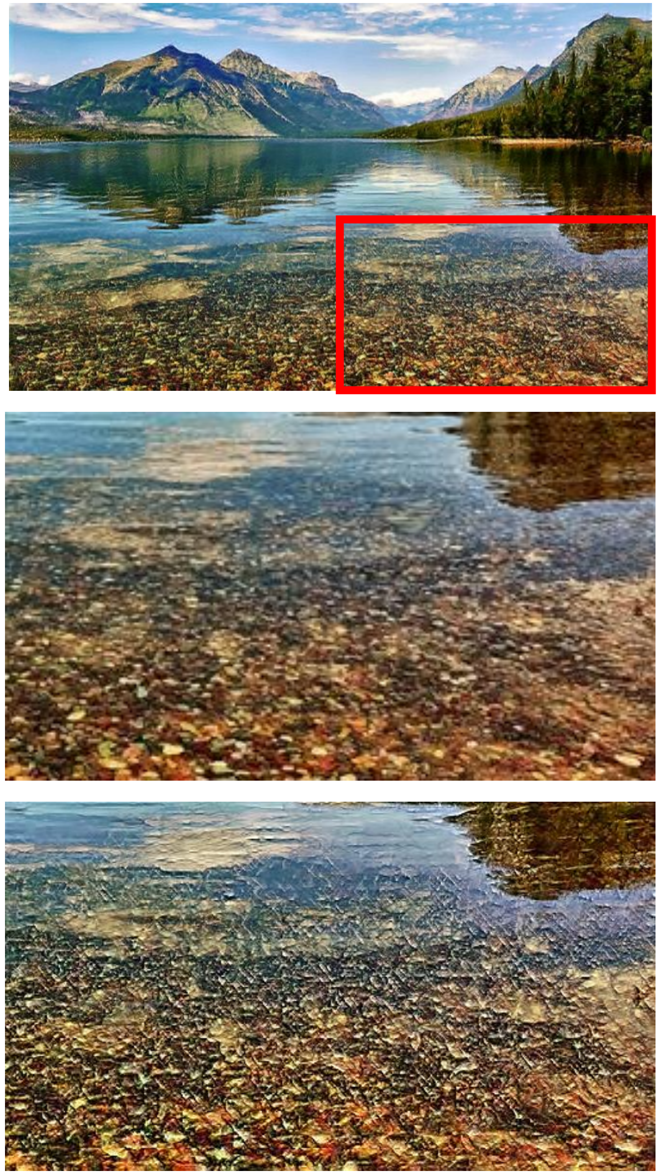
SinGAN modified for super resolution

- Pyramid extension leads to super resolution effect

Singan Reproduced







Thank you