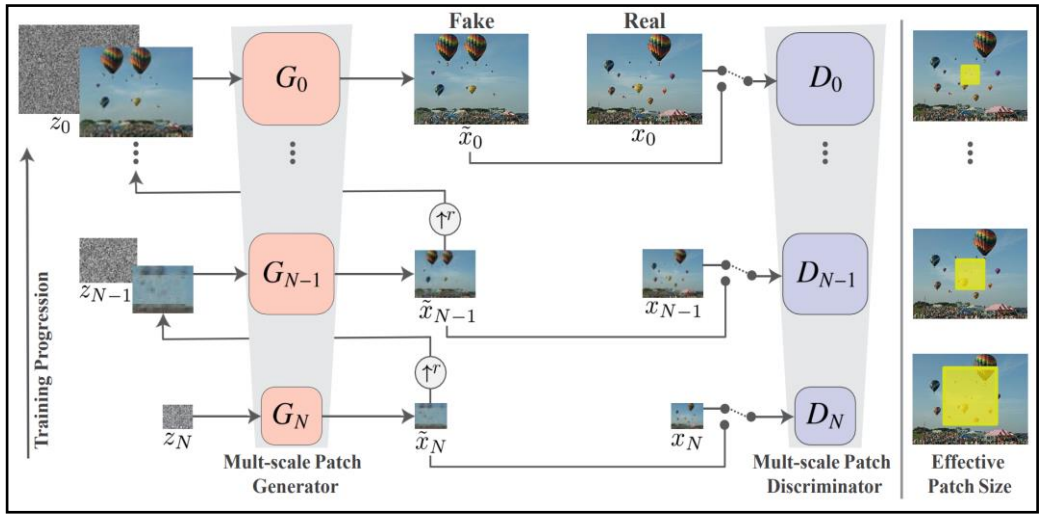
High-Res SinGAN:
A High-Resolution Capable SinGAN Model

Sungwoo Son
Abdur Rehman
Rene Solzbacher

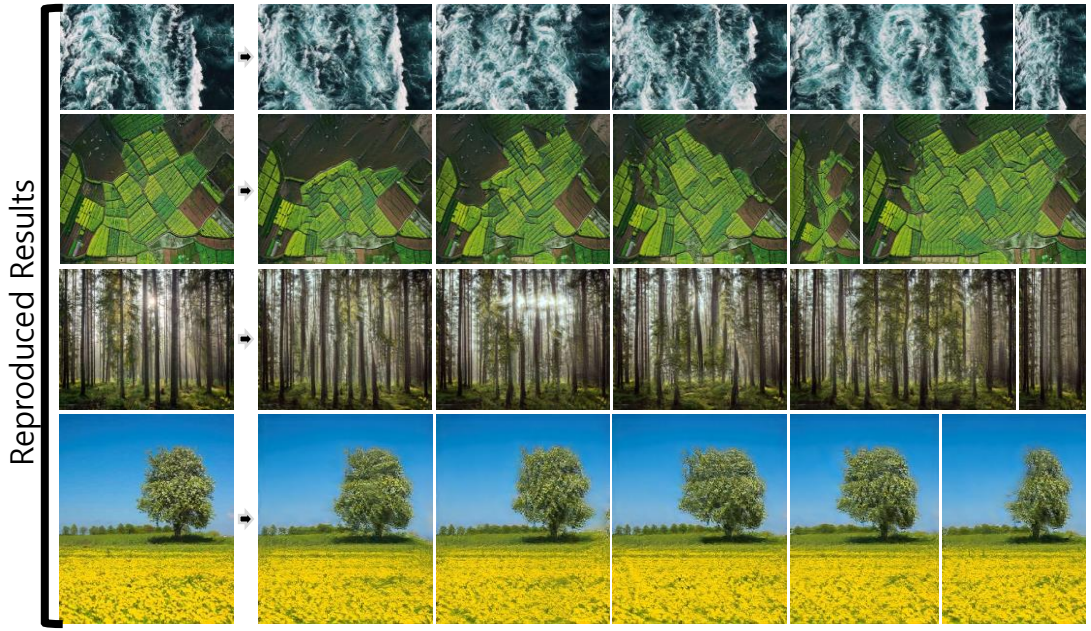
SinGAN

- **Random image creation from single natural image**
- Based on stacked GANs
- Maximum image size **250px**
- Training time **40-60 min**
- Additional functions available
  - Editing
  - Harmonization
  - Animation
  - Paint to image
  - Super resolution



Single training image

Random samples from single training image



High-Res SinGAN

- **Amplified** version of SinGAN
- Can produce output images **up to 500px**
- Introduction of two different methods
  1. Change SinGAN structure
  2. Splice images

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

Method 1 – SinGAN Optimization

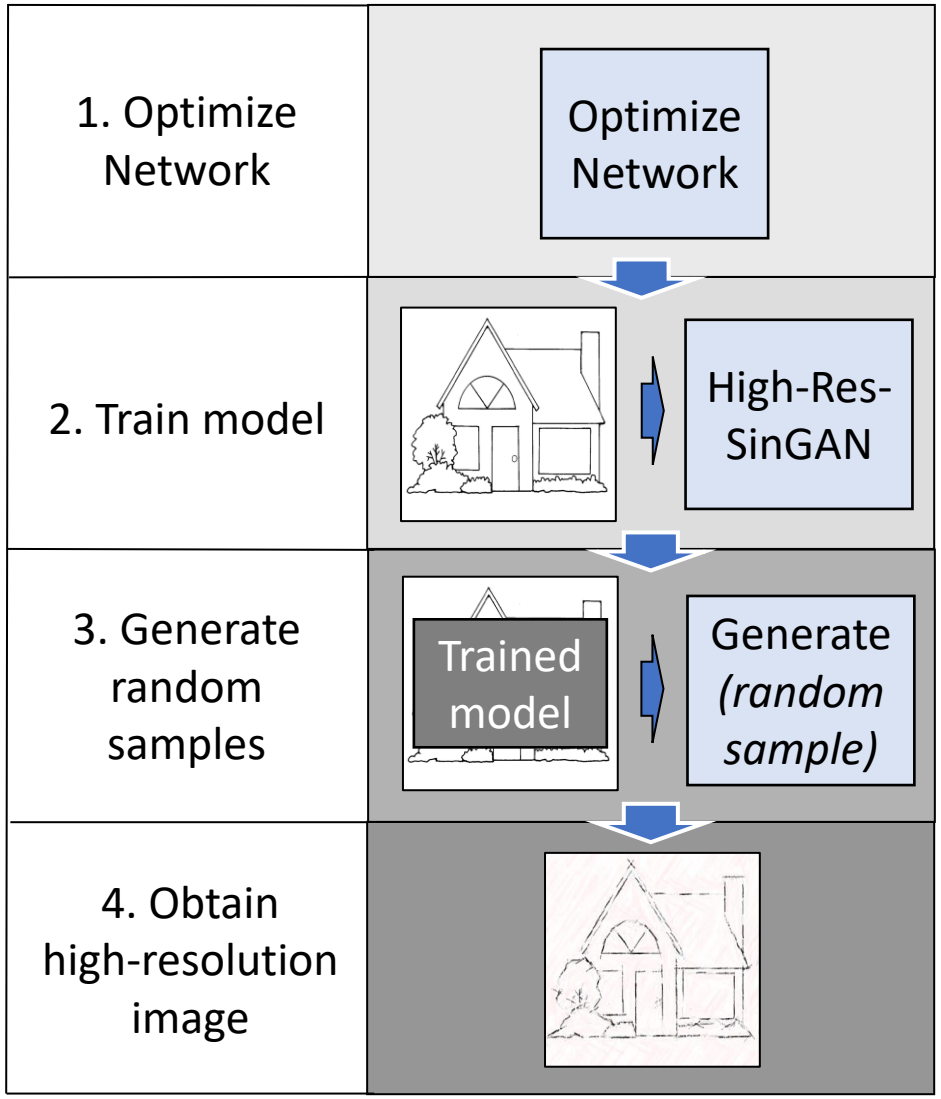
- **Automated** optimization of network based on the input image

Optimization parameters

- **Receptive Field**
  - Affects global structure
- **Number of Kernels**
  - Adds fine texture detail
- **Number of scales**
  - controls transition from global structure to texture detail

- Possible with **any size input** image

```
python HR.py --input_name image.png --size 500
```



Method 1 – Results

Input image (900 x 600)

Random output samples (500 x 334)

Method 1 – Results

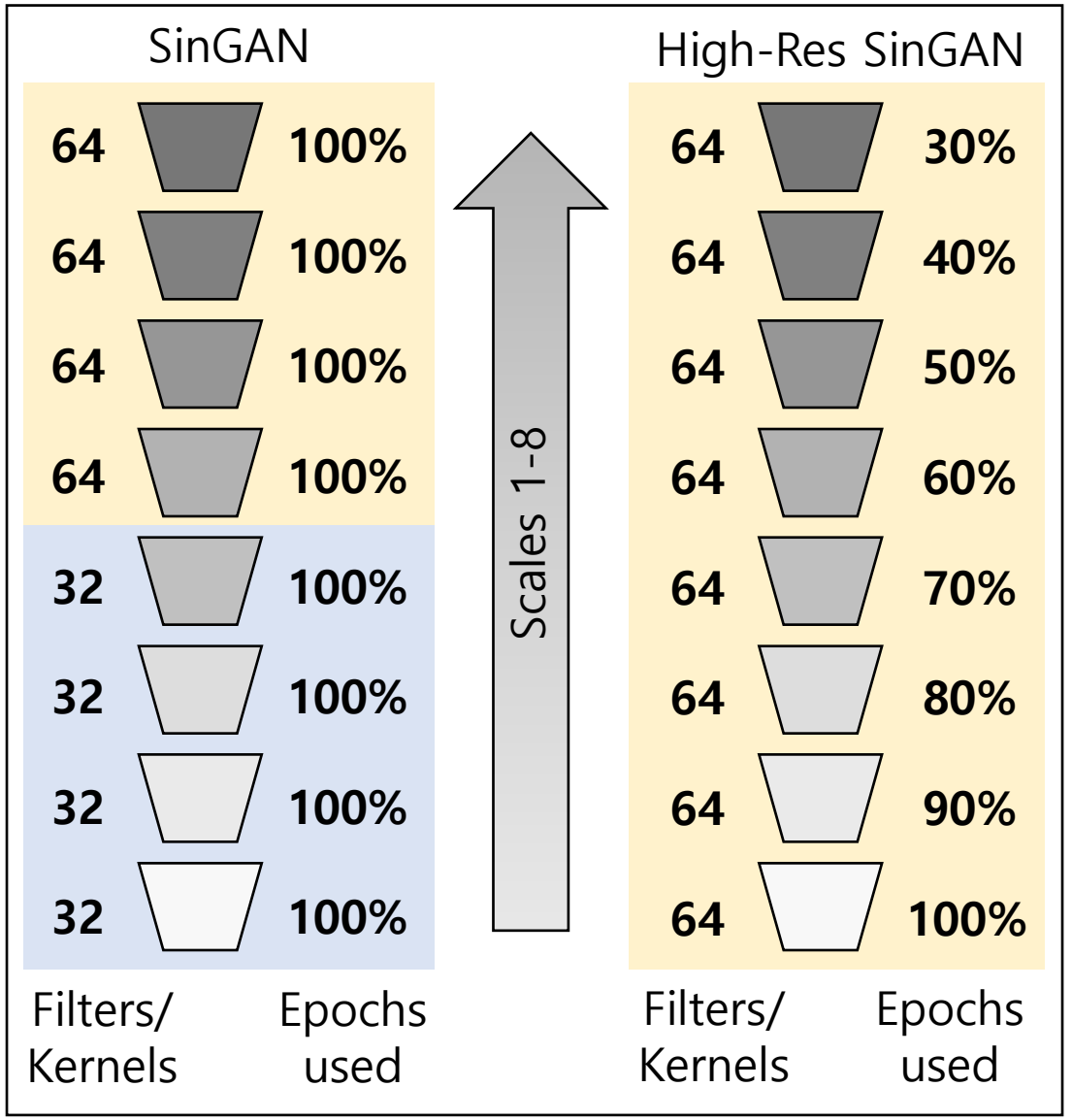
Input image (1028 x 482)

Random output samples (400 x 170)

Method 1 – Training Time Reduction

- Architecture change after scale 4
  - x All weights are reinitialized
- ✓ **Use same weights for all scales**
- Higher scales only add minor detail
- ✓ **Introduce epoch decay**
  - No. of epochs reduced by 10% for each scale

Image size (pixels)	**Normal training time** (min)	**Optimized training time** (min)	**Time reduction** (%)
250	70	40	42.9
400	170	100	41.2
500	330	210	36.4



Hardware for tests - GPU: Nvidia RTX 3060 (12GB), CPU: Intel Core i5 10400F

Method – TTR Results

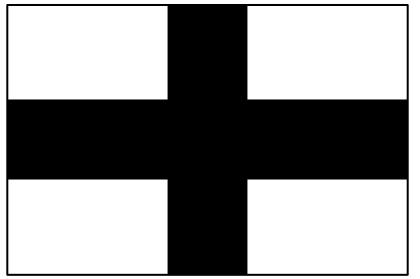
Input image (900 x 520)

Random output samples (400 x 232)

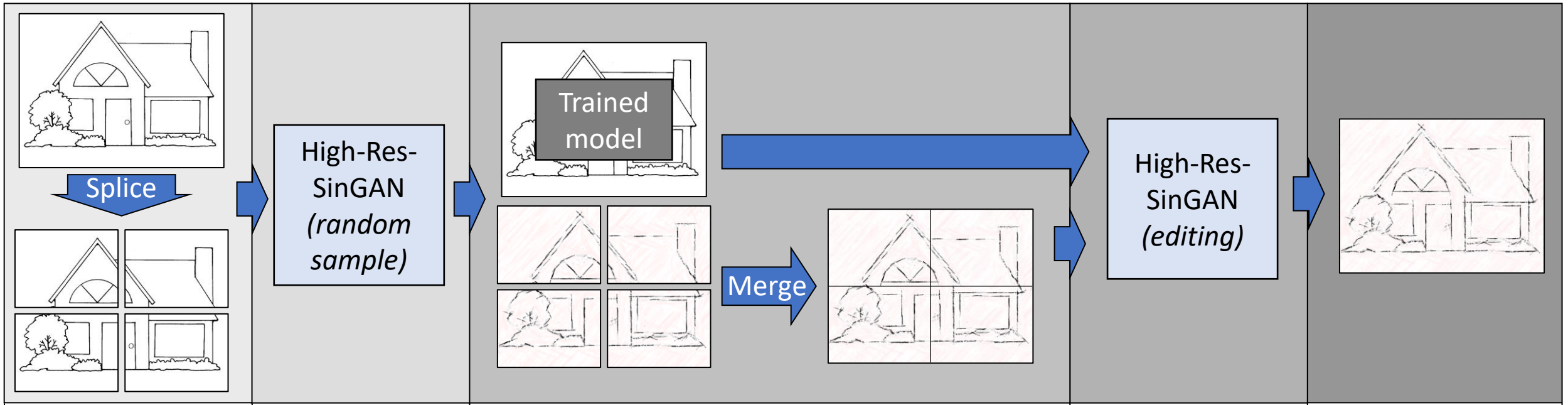
Method 2 – Image Splicing

- **Fully automated** training process
- Image splicing and editing with **pillow**
- Possible with images **up to 500px**

Image editing mask



```
python splicing.py --input_name image.png
```



1. Splice image	2. Train 5 models (4 splices & original full resolution image)	3. Merge 4 random sample slices into one image	4. Edit image (using trained original model)	5. Obtain high-resolution image

Method 2 – Results

Input image

Raw Spliced Image

High-Resolution Output Image

400 x 320

250 x 168

Method Comparison

Method 1

Structure modification

- ✓ Realistic high-res random samples
- ✓ Acceptable training time with training time optimization
- ~ Higher memory requirement
- Non-linear scaling might further reduce memory and training time

Method 2

Image slicing

- ~ Some splicing artifacts remain
- ~ Changes not always realistic
- ~ Higher memory requirement
- x Training 5 networks takes long time
- May produce better results with different splicing algorithm

Structure modification output

Input image

Image slicing output

Design Goals and Discussion

- **Increase image dimension**
  - ✓ Image dimensions of up to 500 pixels possible
- **Keep comparable image quality**
  - ✓ Method 1 produced good results
  - x Method 2 produced less than acceptable results
- **Keep acceptable training times** (SinGAN 40-60 min)
  - Method 1:
    - ~ Training times longer (1.5-3.5 hours)
    - ✓ Still acceptable for neural network training
  - Method 2:
    - x Very long training times (6-7 hours)

The source code and full instructions for both HR methods are available on GitHub
*https://github.com/**Cust0s/High-Res-SinGAN**.git*

Thank you