THE ANDREW & ERNA VITERBI FACULTY OF ELECTRICAL ENGINEERING

SIPL
Signal and Image Processing Lab

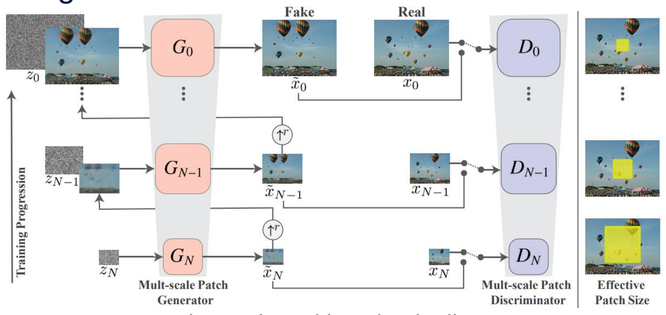
TECHNION Israel Institute of Technology

# SinGANxSR
# Super-Resolution Transfer with SinGAN

**Ofry Livney, Supervised by Tamar Rott Shaham**

## Introduction

- Learning based systems have advanced the performance of single image super-resolution immeasurably.
- These can be split into 2 main groups:
  - External learning systems that require big datasets of images and can generalize to unseen images.
  - Internal learning systems that learn a single low-resolution image and can then perform super-resolution on the training image.
- SinGAN is a multi-scale patch-GAN model that is part of the Internal learning class. It has been shown to perform perceptually sound super-resolution on a given low resolution training image.




SinGAN's multi-scale pipeline.

## Goals

- Exploring the SinGAN architecture's ability to perform super-resolution on unseen images that have a semantic link with the training image.
- Use case: Video Super-Resolution as a method to check generalizability over frames with semantic resemblance to the training image.
- Discriminator Super-Resolution Fitness (DSRF) metric: creating an a-priori task fitness metric that is built-in the model and can predict the fit for a model in performance of the given task.

## Challenges

- The prior knowledge of the model is only of a single image.
- Natural flickering effect created when using SISR systems for the VSR task.

## Super-Resolution Inference Methods

- SMI – Single Model Inference – performing SR on an unseen image using a single pretrained SinGAN model
- DSRFI – DSRF best Inference – a method that decides which pretrained model to use based on the most fitting model for the task based on the minimum absolute DSRF score out of the optional models.
- AMI – Alternating Models Inference – a method that utilizes the iterative nature of the SR process to add details in each iteration based on a different generator.
- AVGI – Averaging the VSR results of 15 different pretrained models – all trained on the same image, in hope to solve the flickering challenge in the VSR task.

## Model Comparison


| | ↓ LPIPSx10 | ↓ NIQE | ↑ VMAF | ↑ PSNR | ↓ RMSE | ↑ SSIM | ↓ tLPx100 | ↓ tOFx10 | DSRFx1000 |
|---|---|---|---|---|---|---|---|---|---|
| ZSSR | 4.68±0.05 | 7.80±0.24 | 66.57±2.23 | 29.48±0.03 | 8.56±0.03 | 0.65±0.00 | 1.49±0.51 | 3.62±0.69 | 0.000±0.000 |
| SinGAN | 2.70±0.05 | 3.52±0.13 | 64.83±1.93 | 28.68±0.02 | 9.39±0.02 | 0.47±0.00 | 6.44±1.35 | 4.01±0.62 | 0.000±0.000 |
| ESRGAN | 2.15±0.10 | 3.49±0.20 | 72.68±2.24 | **28.86**±0.03 | 9.19±0.03 | 0.52±0.01 | 6.16±1.06 | 4.66±0.81 | 0.000±0.000 |

SR comparison between our model and ZSSR and ESRGAN models. Each representing A different class of SR learning sytems.

- ZSSR – Zero Shot Super Resolution – is an image specific CNN that is part of the internal learning class. We show that it produces less high frequency details in comparison to SinGAN.
- ESRGAN – An impressive SR, GAN based model that belongs to the external learning class. Even though the model is trained on a big dataset, SinGAN equals its results perceptualy.

## DSRF Metric

- The goal is to explore the usefulness of the pretrained SinGAN's patch discriminator as a possible a-priori score of it's ability to perform super-resolution on a given unseen LR image.
- The formula for calculating the DSRF metric:

$$DSRF(I_{LR-input}, I_{LR-train}, D) = \left\langle \underline{D(I_{LR-input})} - \underline{D(I_{LR-train})} \right\rangle$$

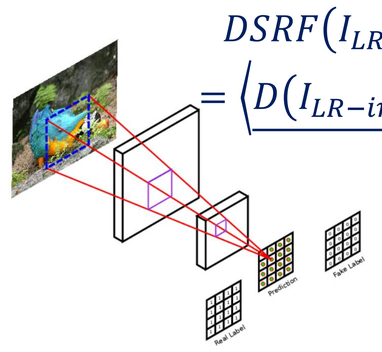
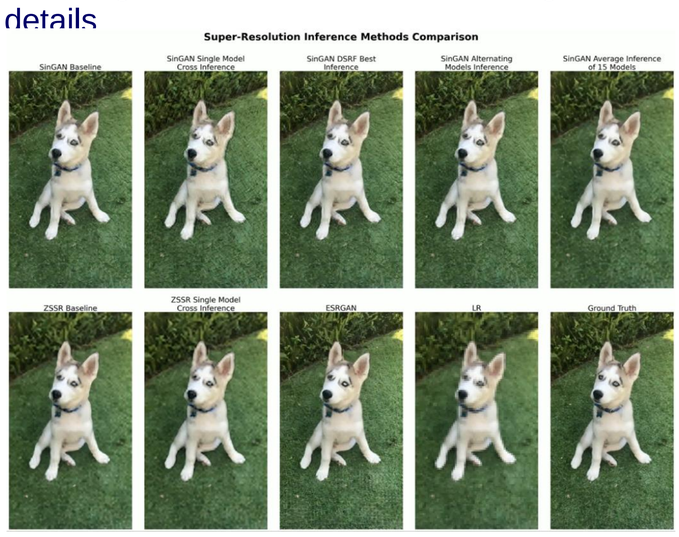



Illustration of a discrimination map of a patch-GAN

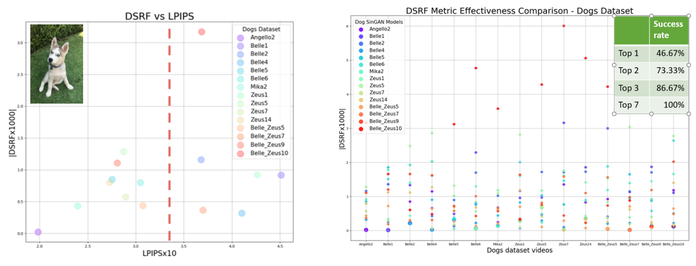
## Results

- In terms of inference methods, the baseline performs best as expected.
- The DSRFI method fails to choose the right model each time, over the tested dataset.
- The AMI method works well, and allows use of 2 pretrained models without a major harm to the SR results in case there is a strong semantic link between them.
- The AVGI method, successfully reduces the flickering, but comes at a cost of losing texture details




Inference methods comparison

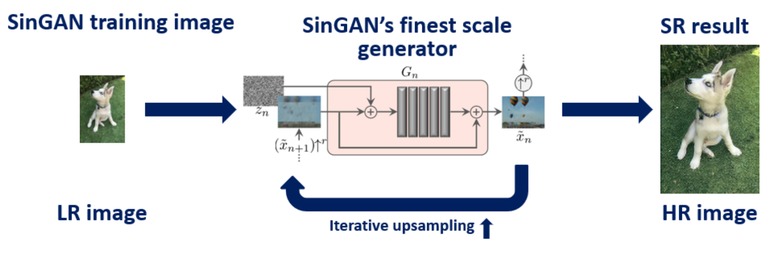
- The DSRF metric proved unreliable in terms of correlation with successful SR – as measured by the LPIPS metric.
- However, the metric has shown an ability to give a confidence interval, as can be seen in the top 2 and top 3 accuracies.




DSRF correlation with successful SR & top 1 accuracy

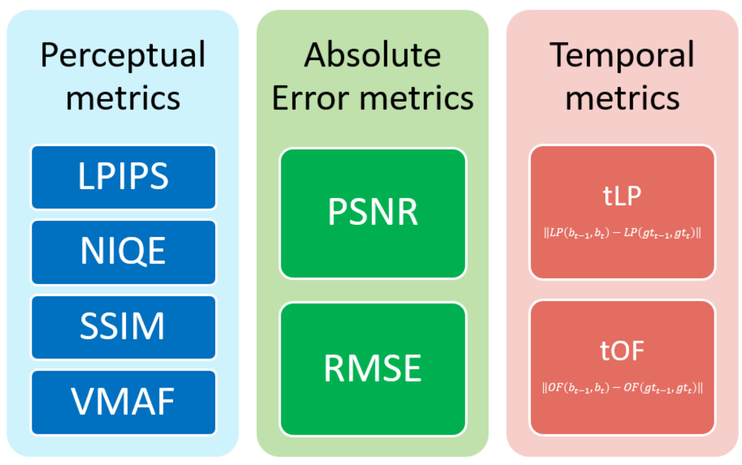
## SinGAN SR Process

- Single Scale SinGAN architecture.
- We Utilize the SinGAN's finest scale generator in an iterative process that upsamples the input image from the LR dimension into the HR result. At each iteration we forward pass the image through the pretrained generator, which adds fine texture details to create the SR result.
- The fully-convolutional patch-GAN architecture allows us to super resolve images of any given LR size.




SinGAN SR Process scheme

## Image Quality Metrics




- Advances in the image quality metrics research have allowed us to measure the results with high confidence in correlation to the human judgement, both spatially and temporally.
- Perceptual metrics such as LPIPS and NIQE are commonly used to evaluate modern SR research.

## Conclusions

- SinGAN has shown a generalization ability in terms of performing SR on unseen images with semantic resemblance.
- The DSRF metric has shown an ability to narrow a search for a fitting model – based on a confidence interval

April 2021