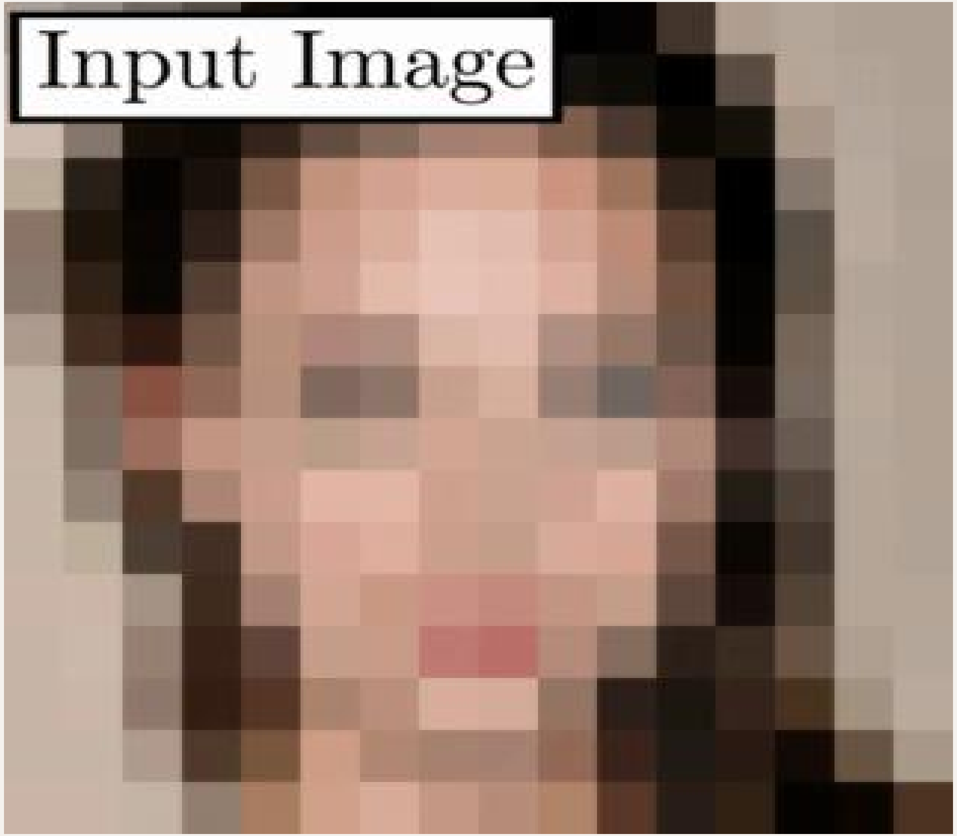
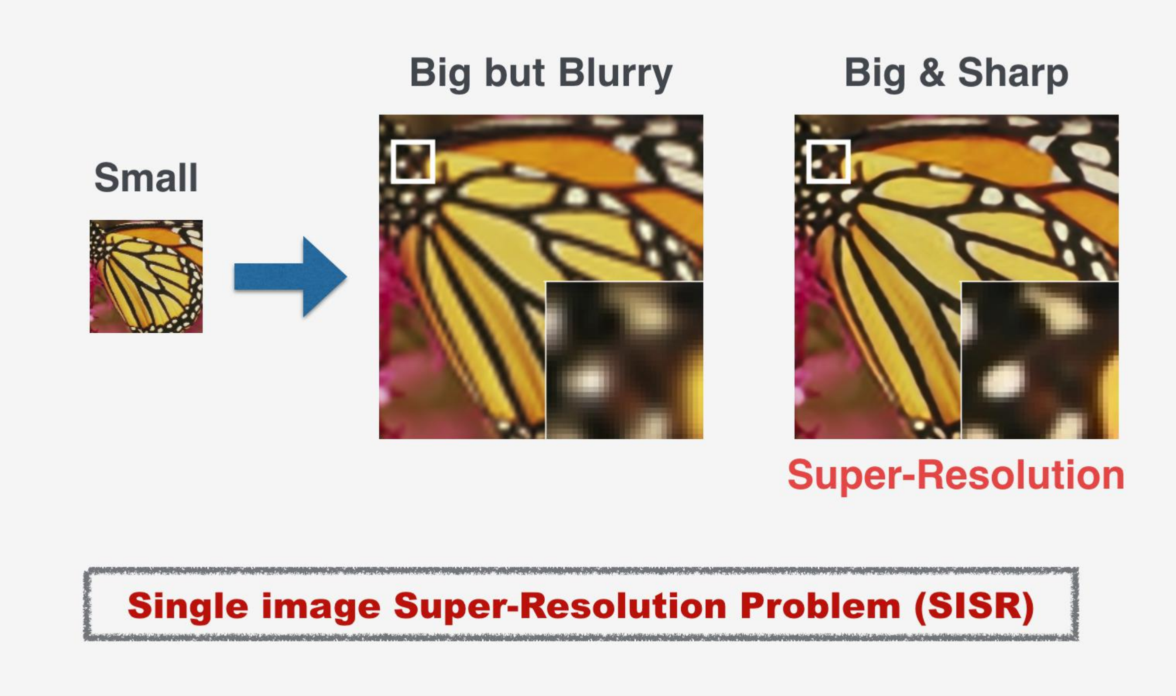
P U L S E

Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models


Input Image

Super-Resolution



超解析度是計算機視覺的一個經典應用。SR是指通過軟體或硬體的方法，從觀測到的低解析度影象重建出相應的高解析度影象，在監控裝置、衛星影象遙感、數字高清、顯微成像、視訊編碼通訊、視訊復原和醫學影像等領域都有重要的應用價值。

Well-posed problem

1. a solution exists.
2. the solution is unique.
3. the solution's behaviour changes continuously with the initial conditions.

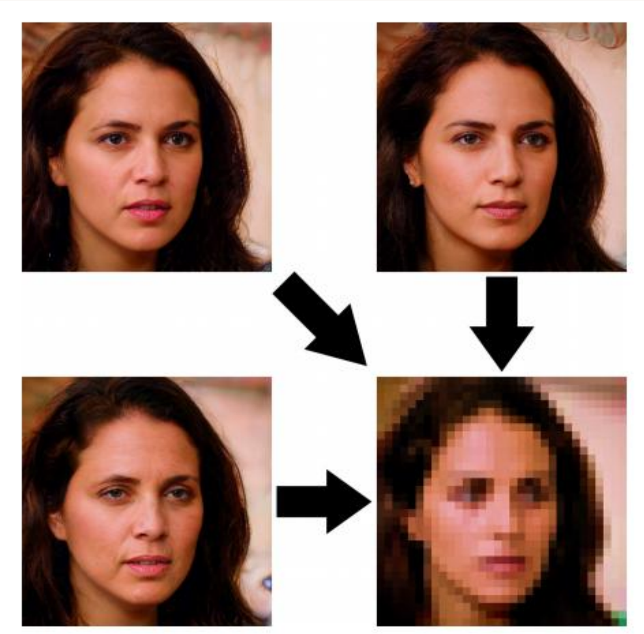
Ill-posed problem

Jaeyoung talked about CV's ill-posed problem in the paper of CVPR:

In most cases, there are several possible output images corresponding to a given input image and the problem can be seen as a task of selecting the most proper one from all the possible outputs.

Classical ill-posed problem on Image processing

- ◆ 影像去噪 Image De-nosing
- ◆ 影像恢復 Image Restorsion
- ◆ 影像放大 Image Zooming
- ◆ 影像修補 Image Inpainting
- ◆ 影像去馬賽克 Image Demosaicing
- ◆ **影像超解析 Image super-resolution**



Intro

Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi, Cynthia Rudin

Duke University

Durham, NC

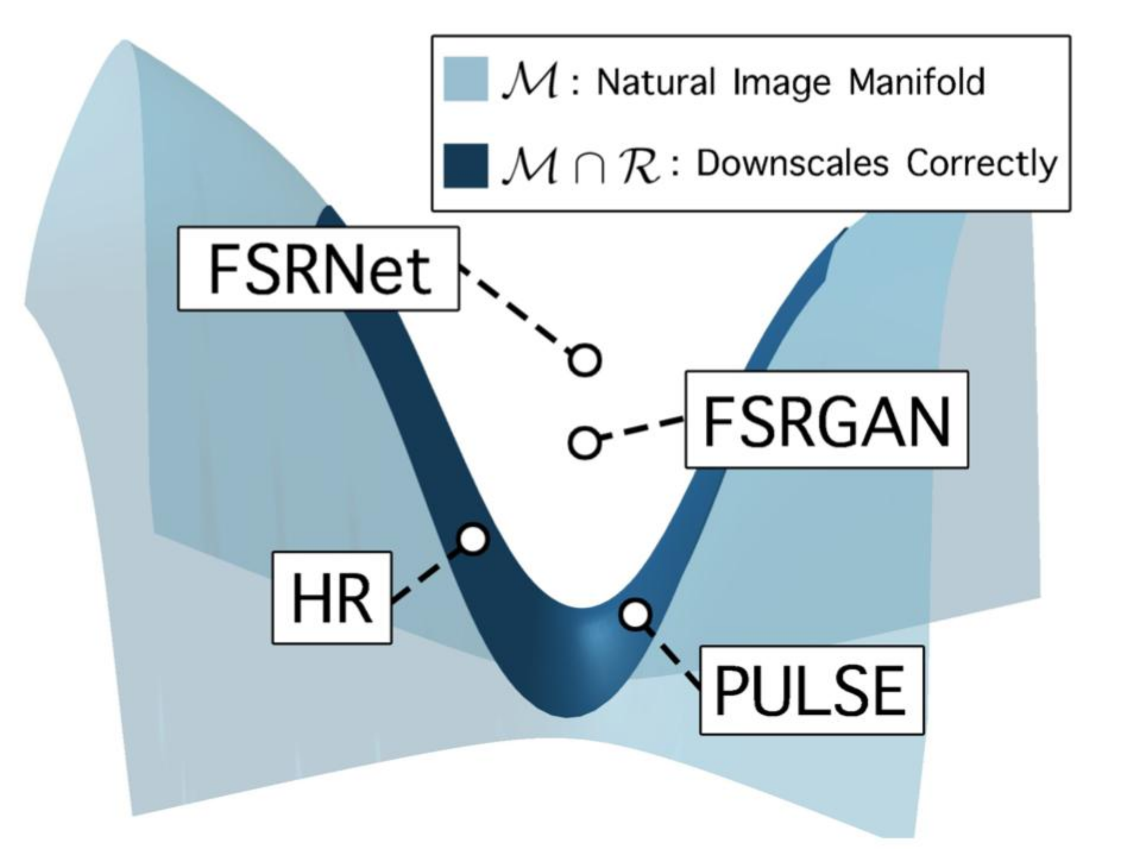
CVPR 2020

Abstract

1. It accomplishes this in an entirely self-supervised fashion.
2. It is not confined to a specific degradation operator used during training.
3. This is formalized through the “down- scaling loss,” which guides exploration through the latent space of a generative model. PULSE thereby generates super-resolved images that both are realistic and downscale correctly.

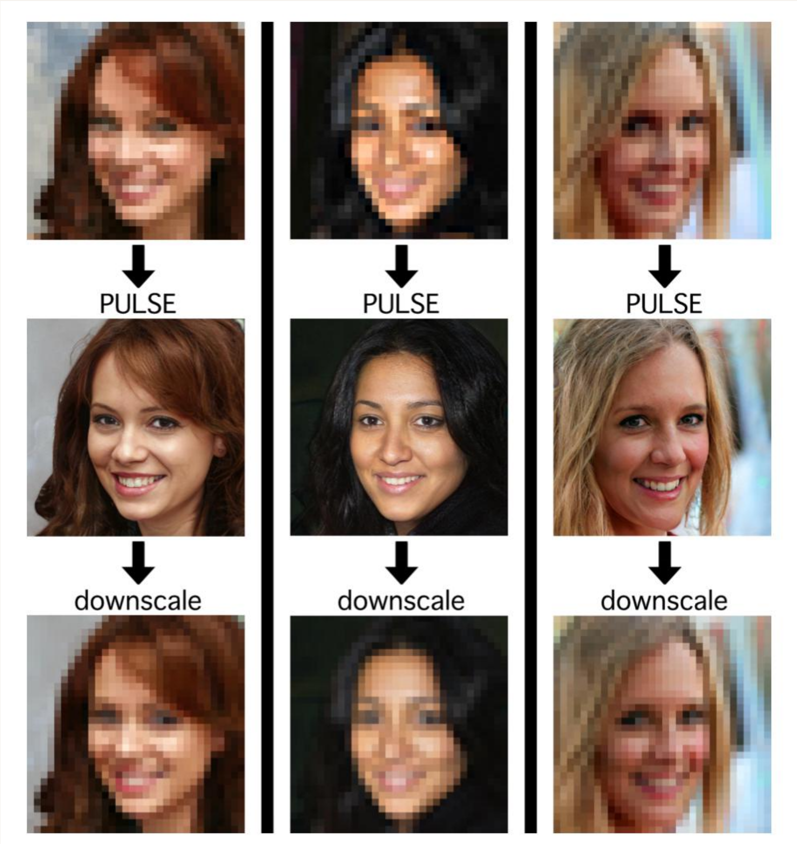
Traditional super-resolution

Traditional supervised super-resolution algorithms train a model (usually CNN) to minimize the pixel-wise mean-squared error (MSE) be- tween the generated super-resolved (SR) images and the corresponding ground-truth HR images [15] [8]. However, this approach has been noted to neglect perceptually rele- vant details critical to photorealism in HR images, such as texture.



FSRNet tends towards an average of the images that downscale properly. The discriminator loss in FSRGAN pulls it in the direction of the natural image manifold, whereas PULSE always moves along this manifold.

To avoid these issues, we propose a new paradigm for super-resolution. The goal should be to generate realistic images within the set of feasible solutions; that is, to find points which **actually lie on the natural image manifold and also downscale correctly.**



Contributions

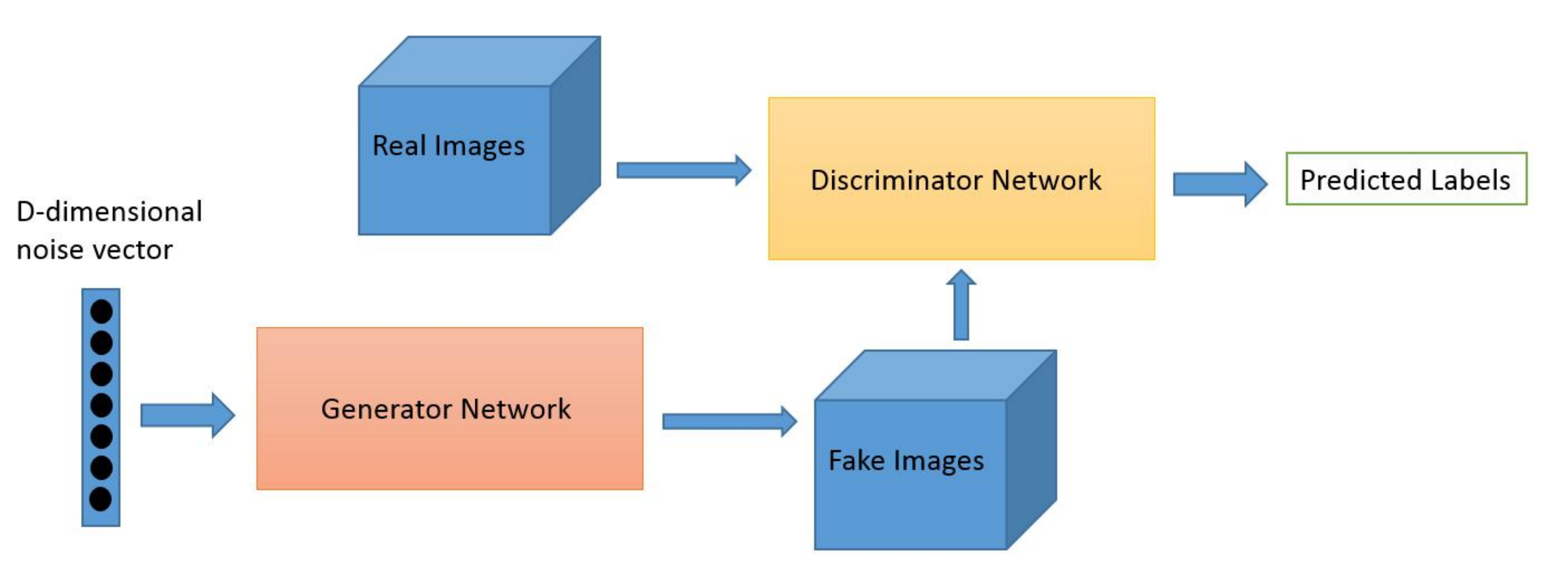
1. A new paradigm for image super-resolution
2. A novel method for solving the super-resolution task
3. An original method for latent space search under high-dimensional Gaussian priors

Generative Adversarial Network

GAN是非監督式學習的一種方法，由兩個網路構成，分別是鑑別網路（Discriminating Network）與生成網路（Generative Network），通過讓兩個神經網路相互博弈的方式進行學習。… 而生成網絡則要盡可能地欺騙判別網絡。 兩個網絡相互對抗、不斷調整參數，最終目的是使判別網絡無法判斷生成網絡的輸出結果是否真實。

生成對抗網路也可以與其他更多不同的網路結合，讓應用範圍加廣泛。具備其他神經網路沒有的「雙胞胎競爭」特性，也使它成為深度學習的一顆閃亮新星。

GAN Basic Flow Chart



Current trends

Currently, there exist two general trends: one, towards networks that **primarily better optimize pixel-wise average distance between SR and HR**, and two, **networks that focus on perceptual quality**.

Fail to enhance detail

$$\int_{\mathcal{M}\cap R} \|I_{HR} - I_{SR}\|_p^p \, dP(I_{HR}). \quad (1)$$

$$I_{SR} = \int_{\mathcal{M}\cap R} I_{HR} \, dP(I_{HR}), \quad (2)$$

In a traditional approach to super-resolution, one considers that the low-resolution image could represent the same information as a theoretical high-resolution image. In practice, even when trained correctly, these algorithms fail to enhance detail in high variance areas. Let M be the natural image manifold in $R^{M\times N}$, and let P be a probability distribution over M, and et R be the set of images that downscale correctly. Then in the limit as the size of our dataset tends to infinity, our expected loss(1) when the algorithm outputs a fixed image I_{SR} is minimized when I_{SR} is an l_p average of I_{HR} over M ∩ R. In fact, when p = 2, this is minimized, so the optimal I_{SR}(2) is a weighted pixelwise average of the set of high resolution images that downscale properly.

Generative networks

Our algorithm does not simply use GAN-style training; rather, it uses a truly unsupervised GAN. It searches the latent space of this generative model for latents that map to images that downscale correctly. The quality of cutting-edge generative models is therefore of interest to us.

As GANs have produced the highest-quality high-resolution images of deep generative models to date, we chose to focus on these for our implementation.Here we provide a brief review of relevant GAN methods with high- resolution outputs. **StyleGAN provides a very rich latent space for expressing different features, especially in relation to faces.**

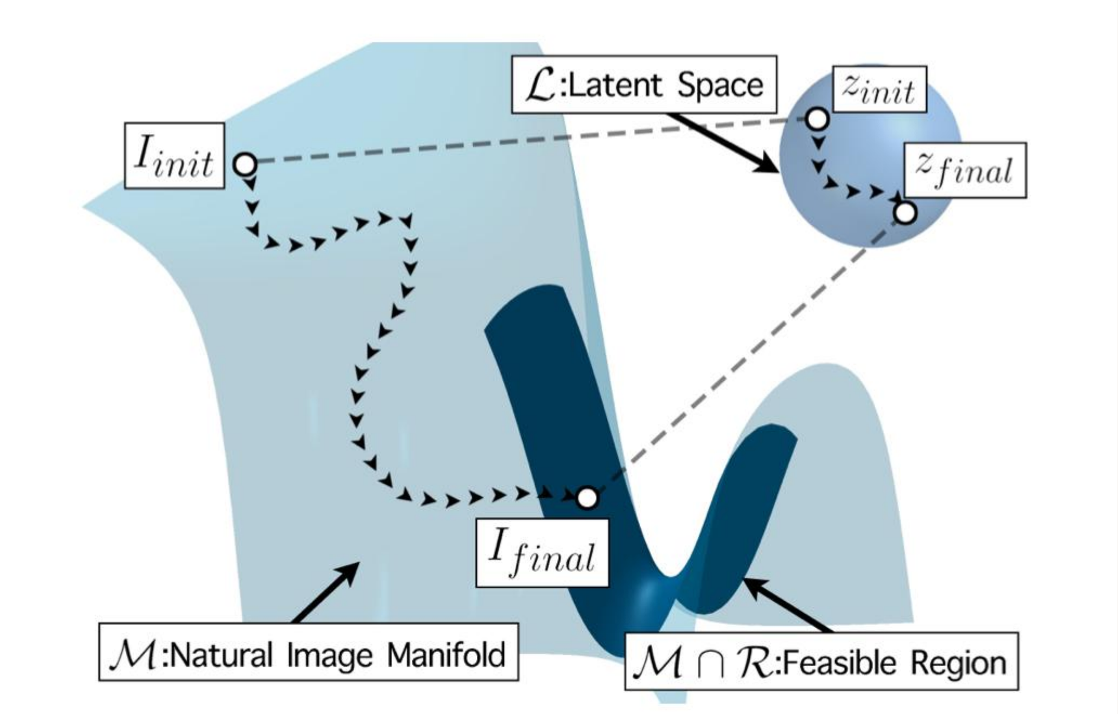
New framework

We therefore propose a new framework for single image super resolution. For a given I_{LR} image, $I_{LR} \in R^{m\times n}$ and $\varepsilon > 0$, our goal is to find an image $I_{SR} \in M$. In particular, we can let $R_\varepsilon \subset R^{N\times M}$ be the set of images that downscale properly. Then we are seeking an image $I_{SR} \in M \cap R_\varepsilon$. The set $M \cap R_\varepsilon$ is the set of feasible solutions, because a solution is not feasible if it did not downscale properly and look realistic.

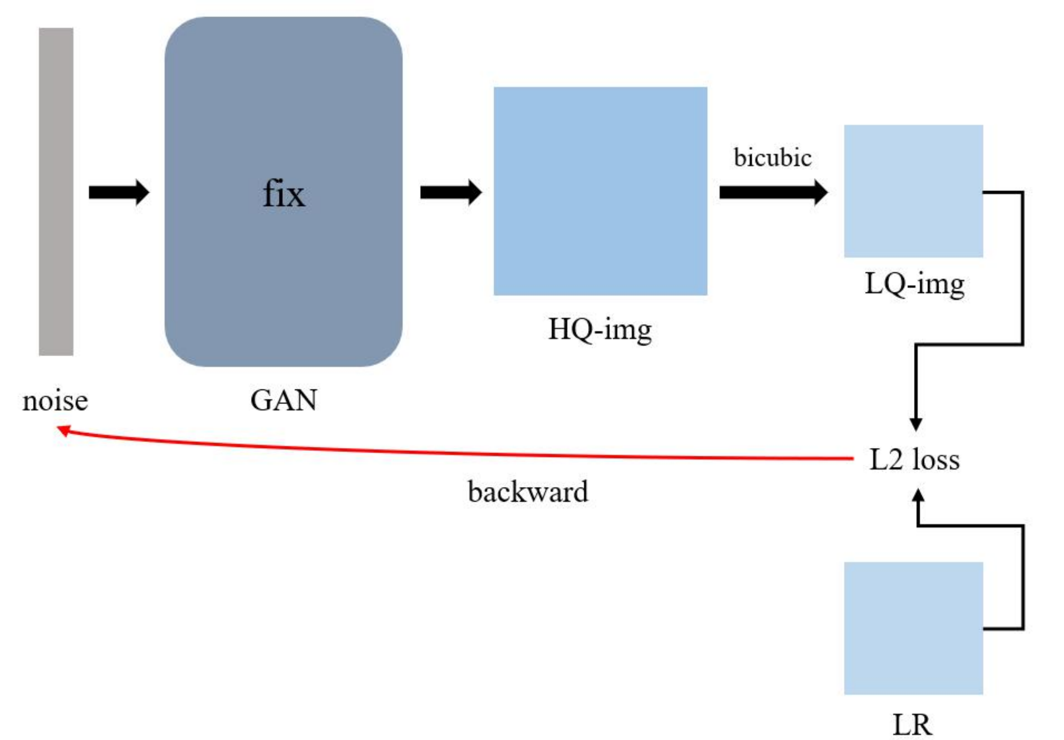
It is also interesting to note that the intersections $M \cap R_\varepsilon$ and in particular $M \cap R_0$ are guaranteed to be nonempty, because they must contain the original HR image.

Method

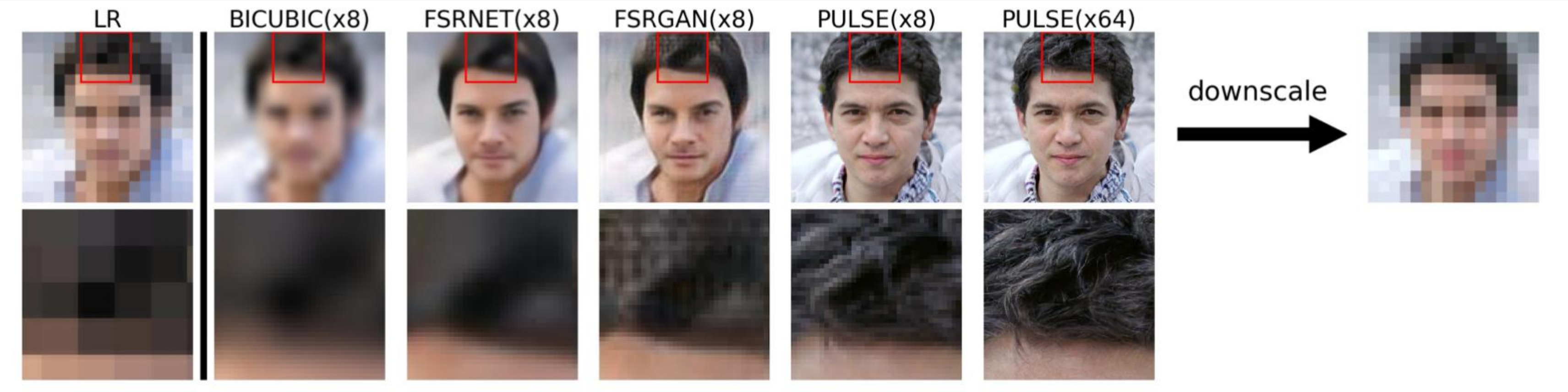
While traveling from z_{init} to z_{final} in the latent space L, we travel from $I_{init} \in M$ to $I_{final} \in M \cap R$.



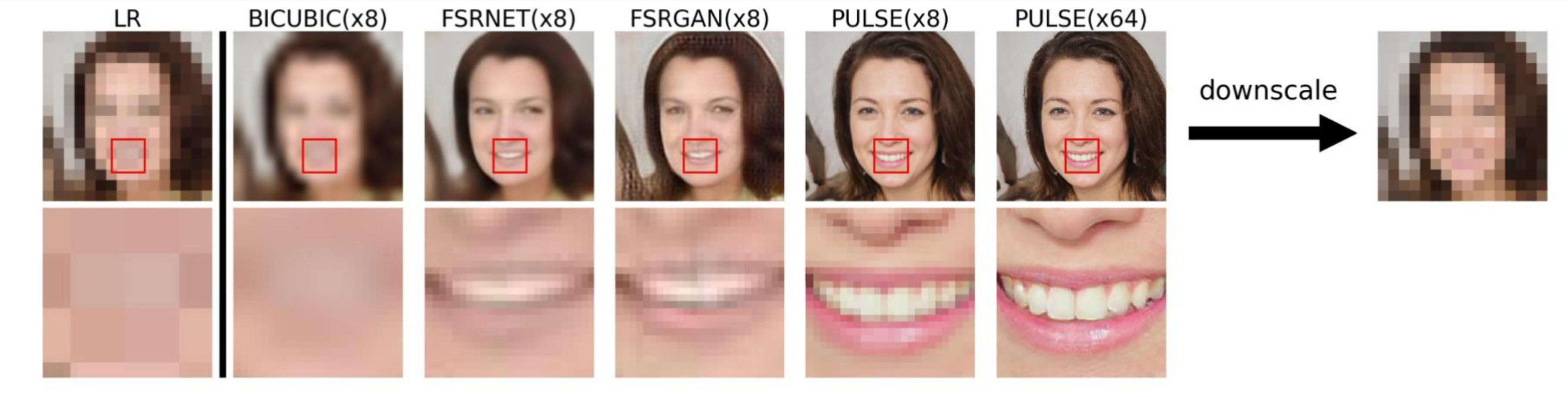
PULSE flow chart



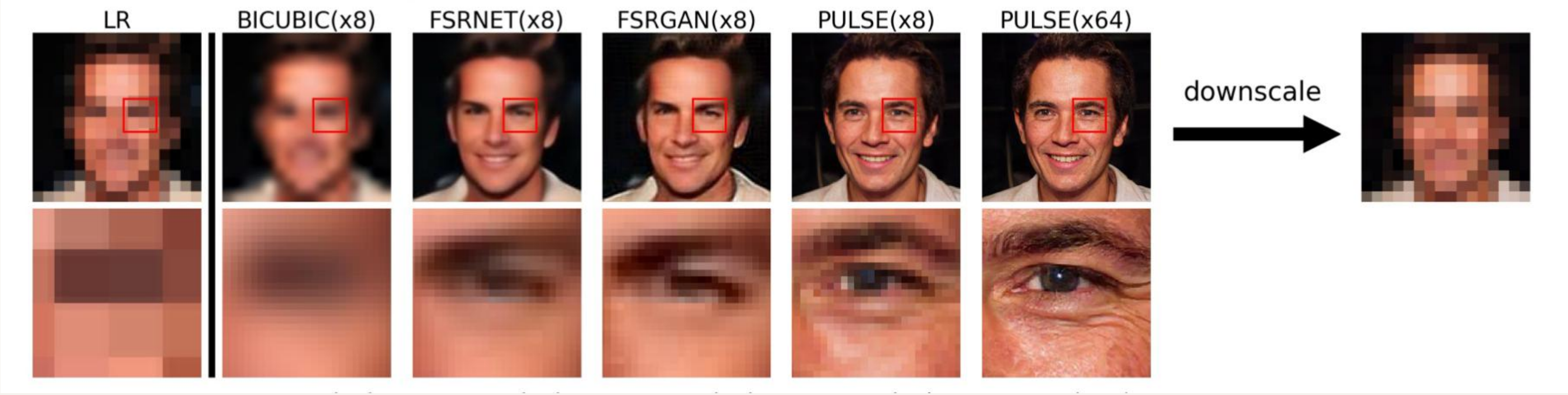
Result


LR
BICUBIC(x8)
FSRNET(x8)
FSRGAN(x8)
PULSE(x8)
PULSE(x64)
downscale

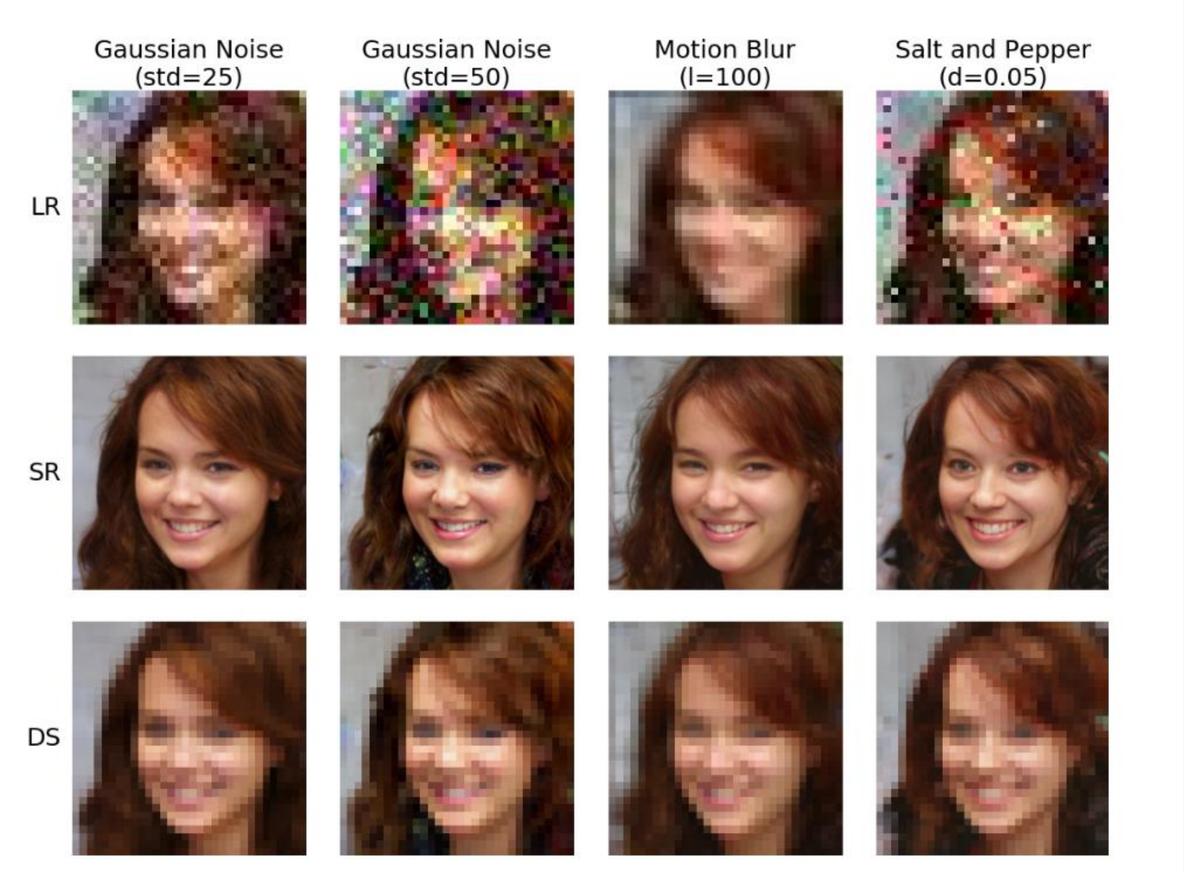
Result


LR
BICUBIC(x8)
FSRNET(x8)
FSRGAN(x8)
PULSE(x8)
PULSE(x64)
downscale

Result



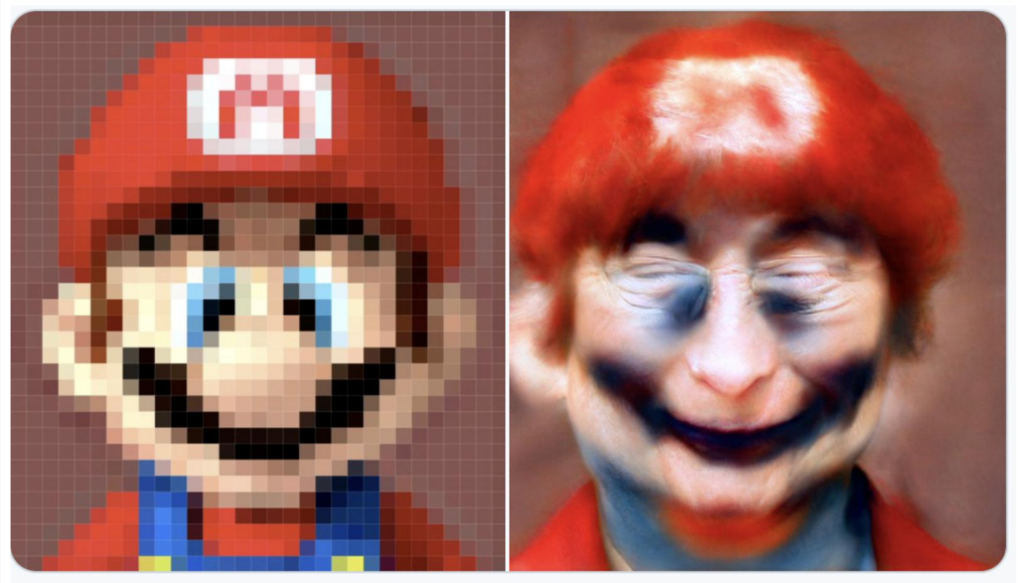
LR
BICUBIC(x8)
FSRNET(x8)
FSRGAN(x8)
PULSE(x8)
PULSE(x64)
downscale


Gaussian Noise (std=25)
Gaussian Noise (std=50)
Motion Blur (l=100)
Salt and Pepper (d=0.05)
LR
SR
DS

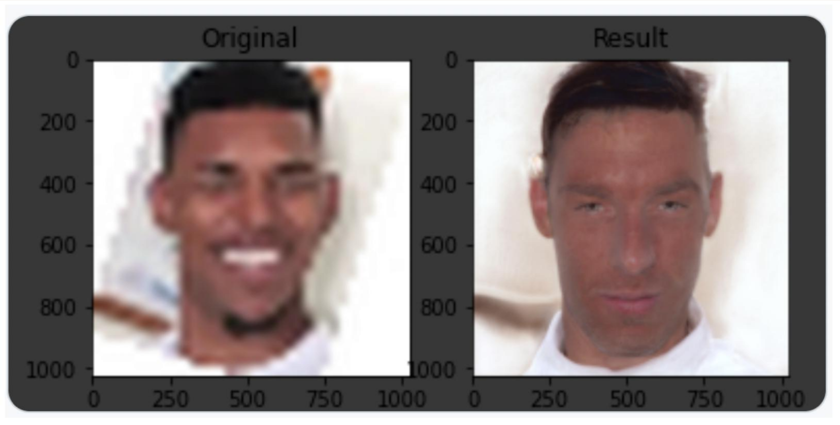
Bias

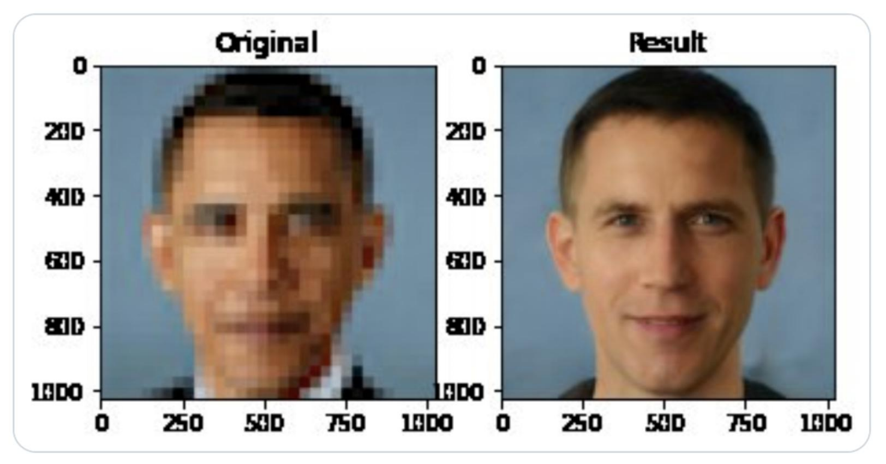
- 该模型对于超高分辨率的恢复任务比较有效，然而对于常规的超分倍数，如4倍、8倍，结果比较差
- 在本文中将其分开每一个单独进行优化，从而导致结果生成效果不如原始生成模型
- 在优化过程中，并没有对输入噪声进行正态分布的约束，因此会使输入跑到极端值，也就是正态分布几乎不可能采样得到的区域，从而导致生成结果不真实
- 以上问题大部分都可以归结为使用的生成模型训练数据与超分辨测试数据的分布不一致问题，如在原始的生成模型中，很少会出现比较奇怪角度的人脸，在超分模型的测试过程中如果输入为该类数据，就会得到比较差的结果

More



More


Original
Result
0
200
400
600
800
1000
0
250
500
750
1000


Original
Result
0
200
400
600
800
1000
0
250
500
750
1000

Resource

- PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models Paper
- PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models Code
- Face Depixelizer Colab

Reference

- 如何理解影象處理領域的不適定問題（ill-posed problem）
- PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models
- Epoch 23 - 超解析度 x VDSR
- [論文筆記]PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models
- 這個 AI 讓模糊的臉看起來銳利 60 倍
- [機器學習 ML NOTE]Generative Adversarial Network, GAN 生成對抗網路
- 生成對抗網路到底在GAN麻？

- 你知道機器學習(Machine Learning)，有幾種學習方式嗎?
- 高糊马赛克秒变高清，「脑补」面部细节，表情帝：这还是我吗？
- StyleGAN-基于样式的生成对抗网络

Thanks

Do you have any questions?