PULSE:

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

Sachit Menon, *Columbia University/Duke University*

November 2020

Organization

1. Motivation/Problem Statement
2. Latent Space Exploration
3. Results
4. Bias

Motivation

- High-resolution image data often difficult to obtain
- Consumer demand for high-resolution image capture higher than ever before
- Many downstream tasks require high-resolution images

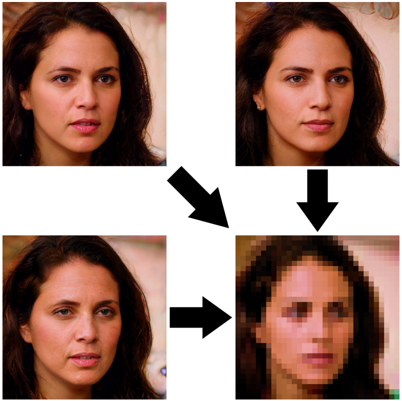
image super-resolution

Construct high-resolution (HR) images from corresponding low-resolution (LR) input

Challenges

- Problem of recovering true HR image depicted by an LR input inherently ill-posed
  - Number of such images grows exponentially with scale factor (Baker and Kanade 2000)

many HR images can correspond to same LR image



Prior Work

traditional methods

Supervised convolutional neural networks (CNNs) using (LR, HR) pairs (e.g., FSRNet)

goal

minimize $L := \| I_{HR} - I_{SR} \|_p^p$

I_{HR}: ground-truth HR image | I_{SR}: generated super-resolved (SR) image | $\| \cdot \|_p$: Some l^p norm

result

Blurring effect

- Some research extends this to optimize on metrics intended to encourage realism (e.g., FSRGAN)
  - "Drags" solution closer to natural image manifold $\mathcal{M} \subset \mathbb{R}^{M \times N}$

New Paradigm

goal

Generate realistic images within the set of feasible solutions

i.e., find points in $\mathcal{M} \cap \mathcal{R}$: points on $\mathcal{M}$ that also downscale correctly

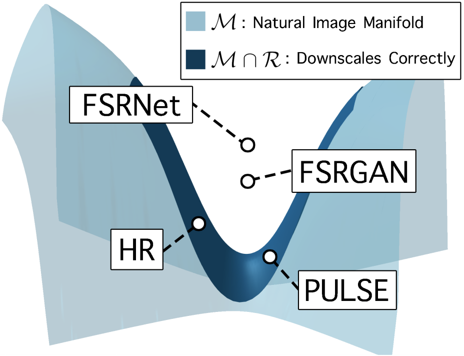
$\mathcal{R}$: set of images that downscale correctly

Given I_{LR} & $\epsilon > 0$, want $I_{SR} \in \mathcal{M}$

s.t.

$$\|DS(I_{SR}) - I_{LR}\|_p \leq \epsilon$$

$DS(\cdot)$: Downscaling operator



Latent Space Exploration

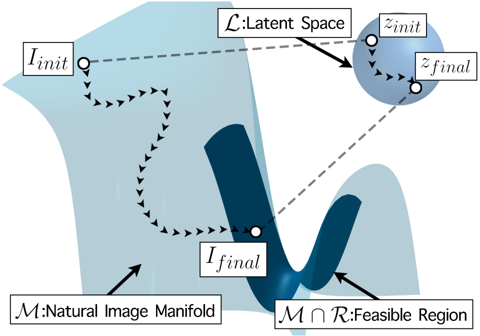
idea

Use generative models!

Given generator G with latent space $\mathcal{L}$ (i.e., $G : \mathcal{L} \to \mathbb{R}^{M \times N}$), find $z \in \mathcal{L}$ with

$$\|DS(G(z)) - I_{LR}\|_p^p \leq \epsilon$$

Simply requiring $z \in \mathcal{L}$ does **not** guarantee $G(z) \in \mathcal{M}$!



Latent Space Exploration

- For $G(z) \in \mathcal{M}$, must be in region of $\mathcal{L}$ seen during training
- **intuitive idea:** Add a loss term for the negative log-likelihood of the prior
  - ▷ Doesn't work!
- **soap-bubble effect:** Most of the mass of a high-dimensional Gaussian is located near the surface of a sphere of radius $\sqrt{d}$

Can replace Gaussian prior on $\mathbb{R}^d$ with a uniform prior on $\sqrt{d}\mathbb{S}^{d-1}$!

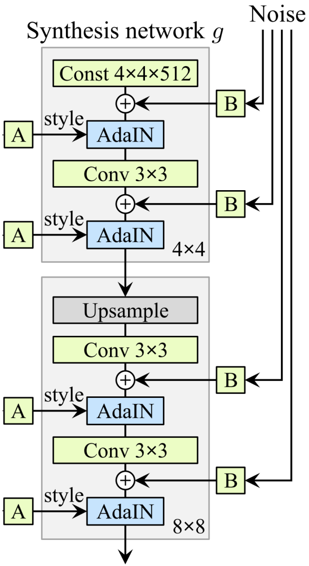
$\Downarrow$

Reduces problem from gradient descent in the entire latent space to projected gradient descent on a sphere

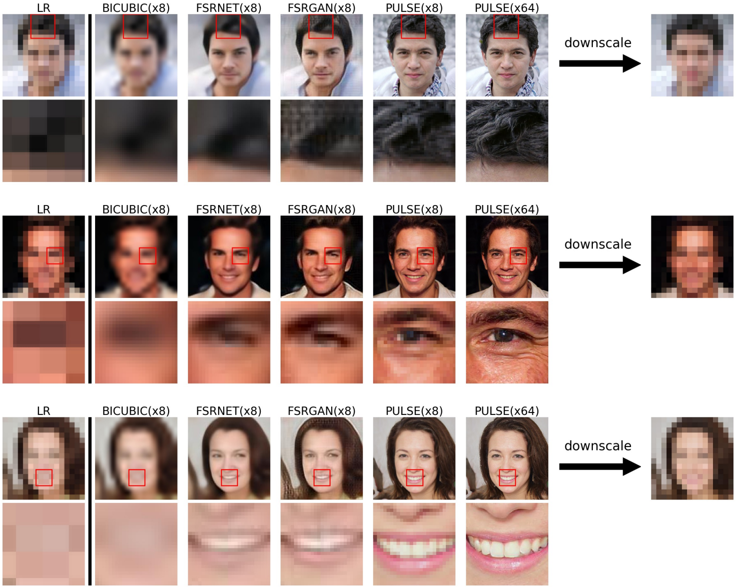
Implementation

- Used StyleGAN (Karras et al. 2019) as generative model due to capacity to generate sharp images
- Used 100 steps of spherical gradient descent with a learning rate of 0.4 starting with random initialization
- Can generate samples with global variation using new initializations and with fine-level variation using inherent stochasticity of StyleGAN

StyleGAN


Synthesis network g
Noise
Const 4×4×512
B
style
A
AdaIN
Conv 3×3
B
style
A
AdaIN
4×4
Upsample
Conv 3×3
B
style
A
AdaIN
Conv 3×3
B
style
A
AdaIN
8×8

Results



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

Results

Contributions

- New problem formulation for super-resolution
- Novel algorithm for addressing the super-resolution problem
  - Creates high-quality images at higher resolutions (1024×1024 vs 128×128) and higher scale factors ($32\times$ vs $8\times$) than state-of-the-art
  - Keeps up with advances in generative modeling without fundamental changes

Bias

- Immediate sources:
  - Pretrained generative model
  - Original training data
  - Evaluation data
- Fundamental source: **systemic bias**

What can we do? Promote inclusivity, create more balanced datasets, and make algorithms robust to imbalanced data

End

Thank you!