On the Frequency Bias of Generative Models

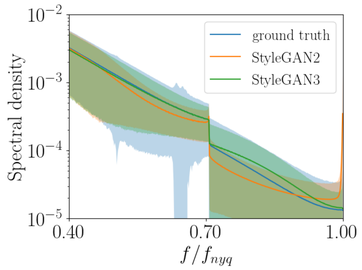
Katja Schwarz Yiyi Liao Andreas Geiger
University of Tübingen and MPI for Intelligent Systems

Max Planck Institute for Intelligent Systems
Autonomous Vision Group

Eberhard Karls Universität Tübingen

Motivation

- **2D GANs** achieve photorealistic image synthesis
- But generated images have **artifacts in their spectrum**



Spectrum of StyleGAN on FFHQ

- Simple classifier can **almost perfectly distinguish** real and fake images
- **Key objective** of GAN training not fulfilled
- Systematic issue **across GAN-architectures**

➡ Is there a systematic **frequency bias** in the generator and/or the discriminator?

Existing Observations

Generator

Most works attribute high-frequency artifacts to **upsampling operations** in generator
Bed-of-nails upsampling: peak at high frequencies
Bilinear/nearest neighbor upsampling: few high frequencies

- Does the upsampling operation introduce a frequency bias over the course of training?
- Can the learnable filters compensate for artifacts from upsampling?

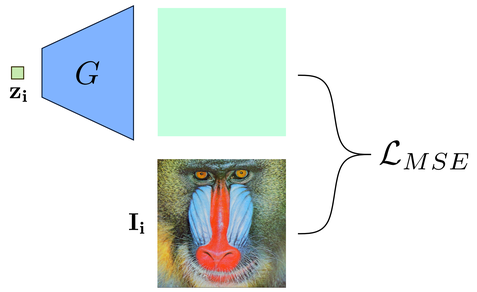
Discriminator

Some works propose additional training objectives but the discriminator is **not investigated in detail**.

- Can the discriminator detect high frequencies and provide the necessary supervision?
- Is aliasing due to the downsampling operations problematic?

Generator Testbed

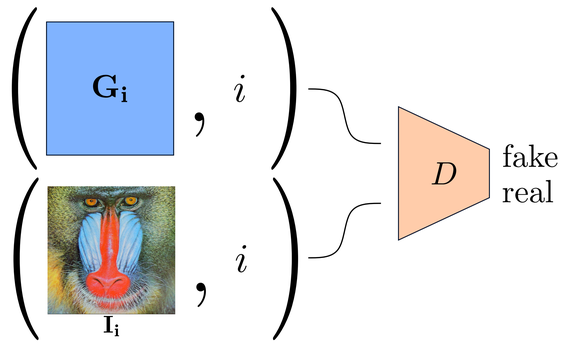
Idea: Analyze generator in **isolated testbed**.



- conditional reconstruction
- pair each image $\mathbf{I_i}$ with a fixed latent code $\mathbf{z_i}$

Discriminator Testbed

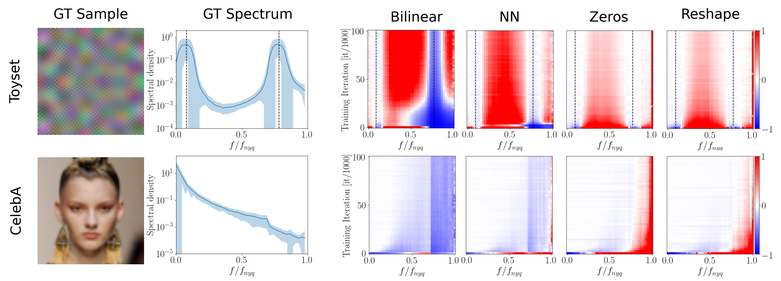
Idea: Analyze discriminator in **isolated testbed**.



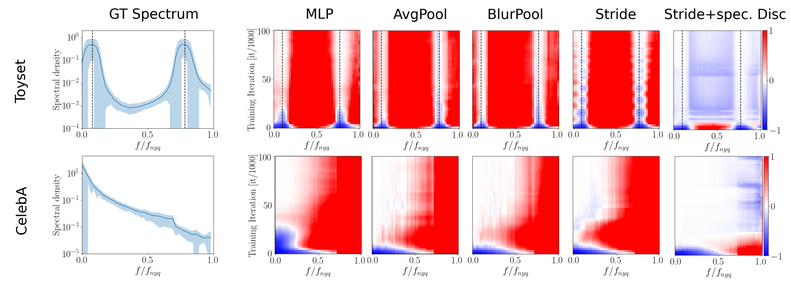
- class-conditional GAN with single sample per class
- pair each image $\mathbf{I_i}$ with a label i and a learnable tensor $\mathbf{G_i}$

Experiments

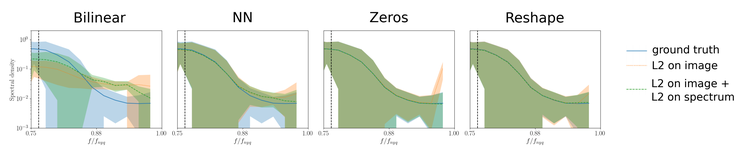
Is there a frequency bias over the course of training?



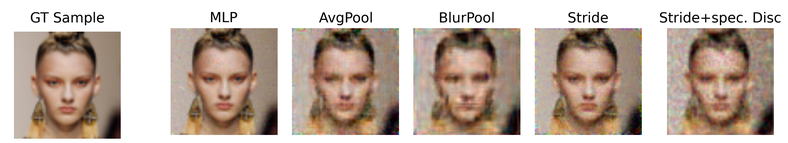
Can the discriminator detect high frequencies?



Can the generator compensate for upsampling artifacts?



Is downsampling problematic?



Main Findings

- Different **upsampling operations** bias the generator towards different spectral properties
- Checkerboard artifacts introduced by upsampling **cannot explain the spectral discrepancies alone** as the generator is able to compensate for these artifacts.
- The discriminator does not struggle with detecting high frequencies per se but rather **struggles with frequencies of low magnitude**.
- The **downsampling operations** in the discrimiantor can impair the quality of the training signal it provides.

katja.schwarz@uni-tuebingen.de