

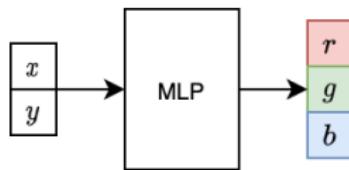
Meta Generation

Ivan Skorokhodov

July 13, 2020

Implicit Neural Representations

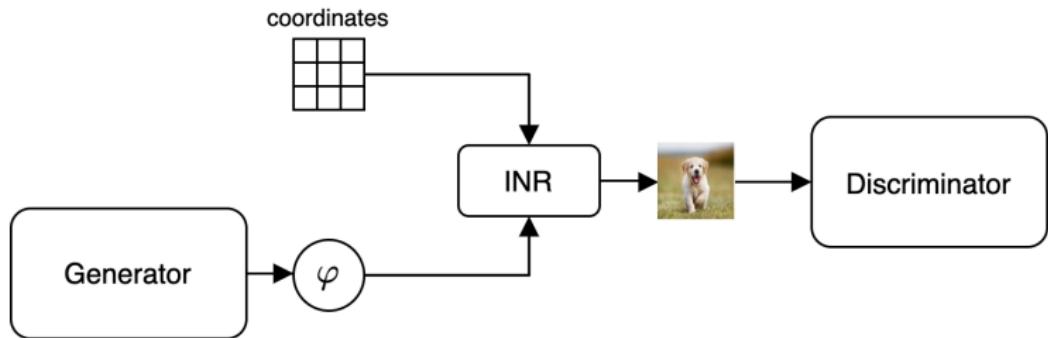
- ▶ INR is a neural network $f_\varphi(x, y)$ which takes coordinates (x, y) and produces a pixel value:



- ▶ We can generate the whole image by computing the value of $f_\varphi(x, y)$ at each coordinate
- ▶ I.e. we have $1 \text{ image} = 1 \text{ INR}$

INR-GAN

- ▶ Main idea: train a generator to produce INRs



- ▶ Discriminator can operate:
 - ▶ on top of images
 - ▶ on top of INRs (but this would require converting the whole dataset into INRs)

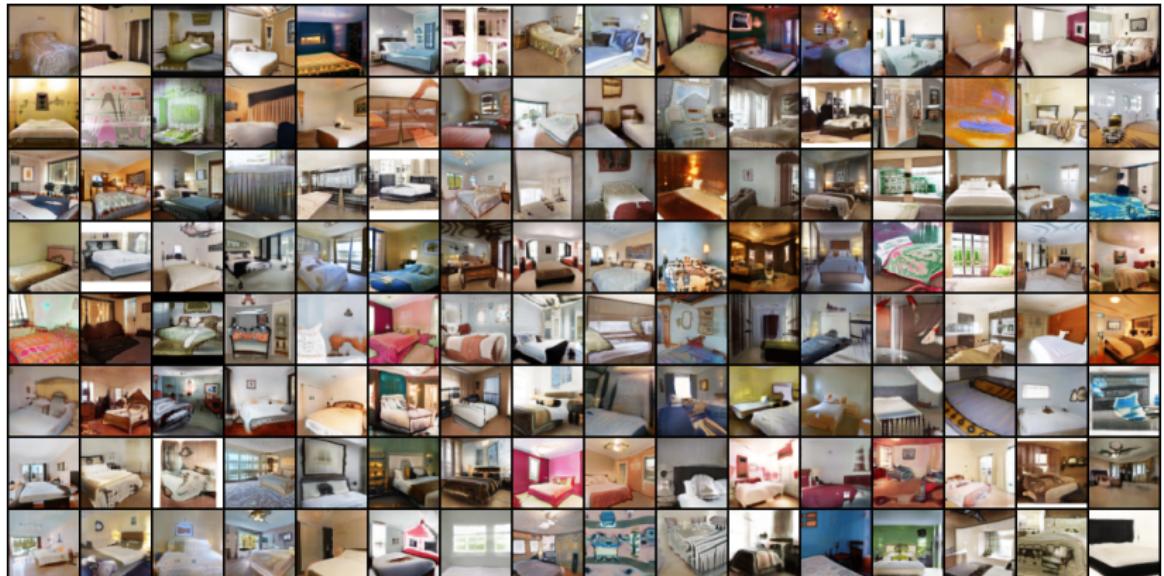
Advantages compared to traditional GANs

- ▶ A unified generator architecture for images/audio/video/3D-shapes
- ▶ Better biological plausibility
- ▶ We can generate different resolutions for different parts of an image
 - ▶ Imagine that we generate an image of a human on a sky background
 - ▶ We can use dense set of coordinates for a human and a sparse set of coordinates for the sky
 - ▶ It is good since it is impossible to do so for a normal GAN
- ▶ We can generate spherical images (and images on any surface)
- ▶ We can use “random resolutions” loss to train a high-resolution GAN
- ▶ We can train our generator on images of extreme resolution by using a “super-resolution” loss:
 - ▶ Train an INR-GAN normally on, for example, 256×256 images
 - ▶ Additionally, compute random dense patches and make the discriminator to distinguish between real/fake dense patches
 - ▶ (Provide global context if needed)
- ▶ It should be cheaper to use for some domains (audio, 3D, video, extreme-resolution images)
- ▶ Progressive growing is easier to incorporate

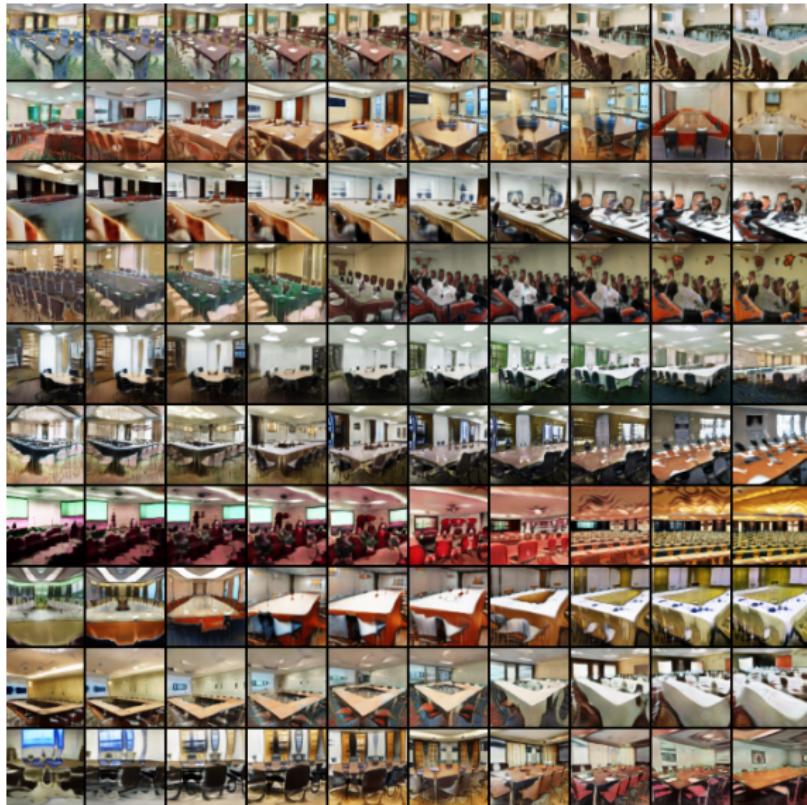
Question: what other benefits can an INR-GAN have?

What other benefits can an INR-GAN have compared to a traditional GAN?

Current samples on LSUN bedroom 64x64



Interpolations on LSUN conference room 64x64



Interpolations on LSUN bedroom 64x64



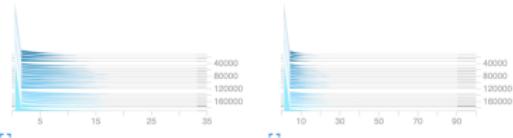
Question 1: how to show “geomtry understanding”

- ▶ CoordConv GAN paper claims that their interpolations are “geometry-aware”: they are done through rotations, zooms, translations, etc.
- ▶ In our experiments, we have found the same property
- ▶ We have not checked if it is true for StyleGAN (it is likely to be true)
- ▶ If it is true for StyleGAN, it does not mean that it was *more difficult* for it to learn this geometry information (i.e. it could spend a lot of capacity on it).
- ▶ How can we demonstrate that our model captures the geometry more easily?

Question 2: stability

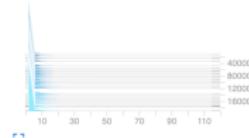
PREVIOUS PAGE

activations/ReLU-1
bedroom-9a0d6ba6b3-11215884/logs

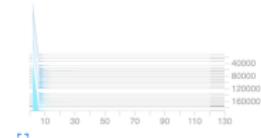


NEXT PAGE

activations/ReLU-5
bedroom-9a0d6ba6b3-11215884/logs



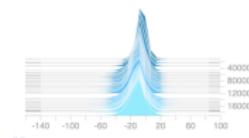
activations/ReLU-7
bedroom-9a0d6ba6b3-11215884/logs



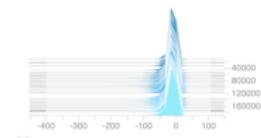
activations/ReLU-9
bedroom-9a0d6ba6b3-11215884/logs



activations/ScaledLinear-0
bedroom-9a0d6ba6b3-11215884/logs



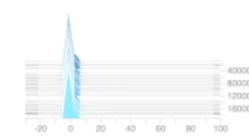
activations/ScaledLinear-4
bedroom-9a0d6ba6b3-11215884/logs



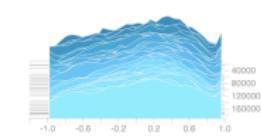
activations/ScaledLinear-6
bedroom-9a0d6ba6b3-11215884/logs



activations/ScaledLinear-10
bedroom-9a0d6ba6b3-11215884/logs



activations/Tanh-11
bedroom-9a0d6ba6b3-11215884/logs



- ▶ Do we have a stability problem? How to alleviate it?
- ▶ It's important to solve it since it's important to learn with a high LR

Question: abrupt interpolations problem

Sometimes during interpolations we have a very abrupt change



What can be the cause of it and how to alleviate it?