A Style-Based Generator Architecture for Generative Adversarial Networks

Authors:
Terro Karras
Samuli Laine
Timo Aila

Presented By:
Mayank Gupta
Swatik
Sanyam

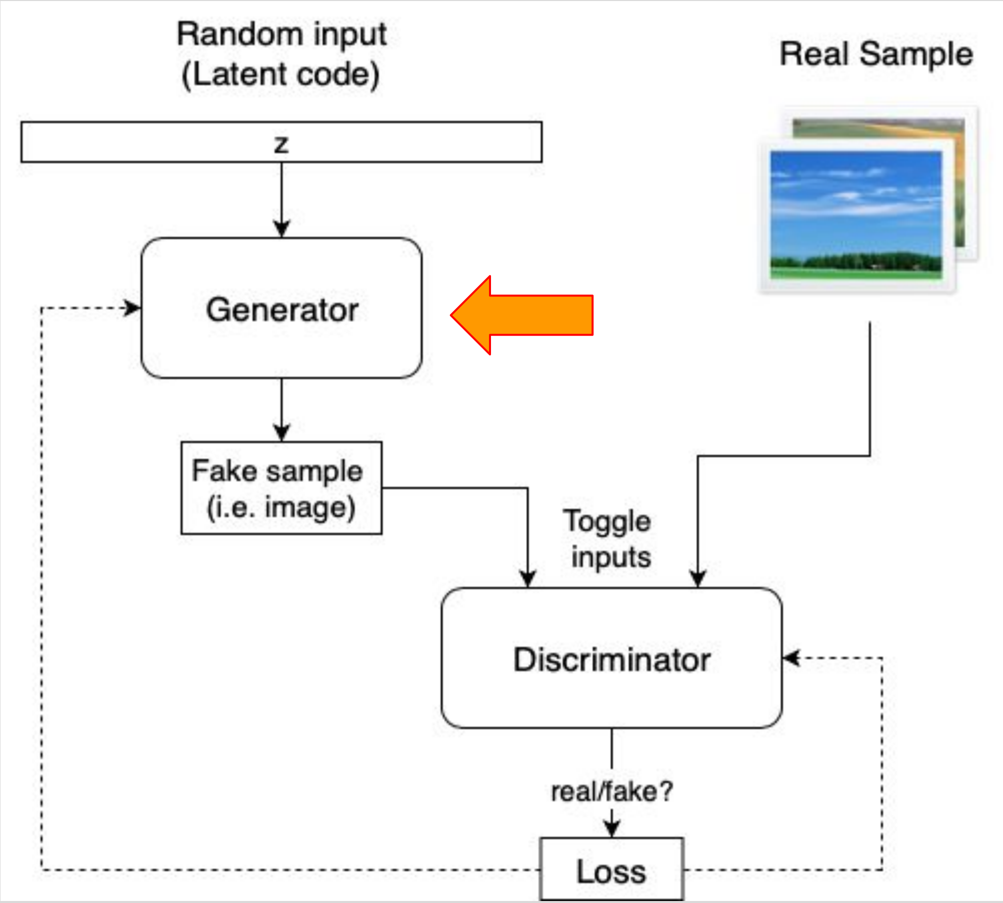
Outline

1. Motivation
2. Style transfer network
3. ProGAN + StyleGAN architecture
4. Generator architecture
5.
6.
7.
8. Style Mixing
9. Separation of global effects from stochasticity
10. Disentanglement Studies
11. Limitations and Future Scopes

Motivation

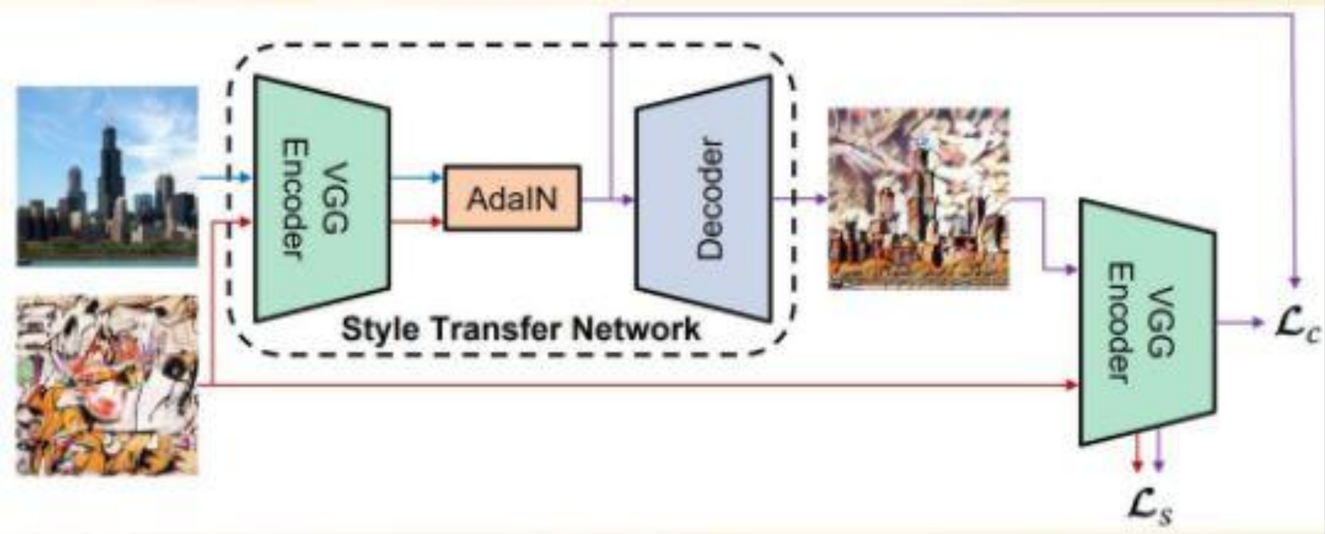
1. GAN images became realistic overtime but still generator works as a black box.

2. Understanding of image synthesis is poor.



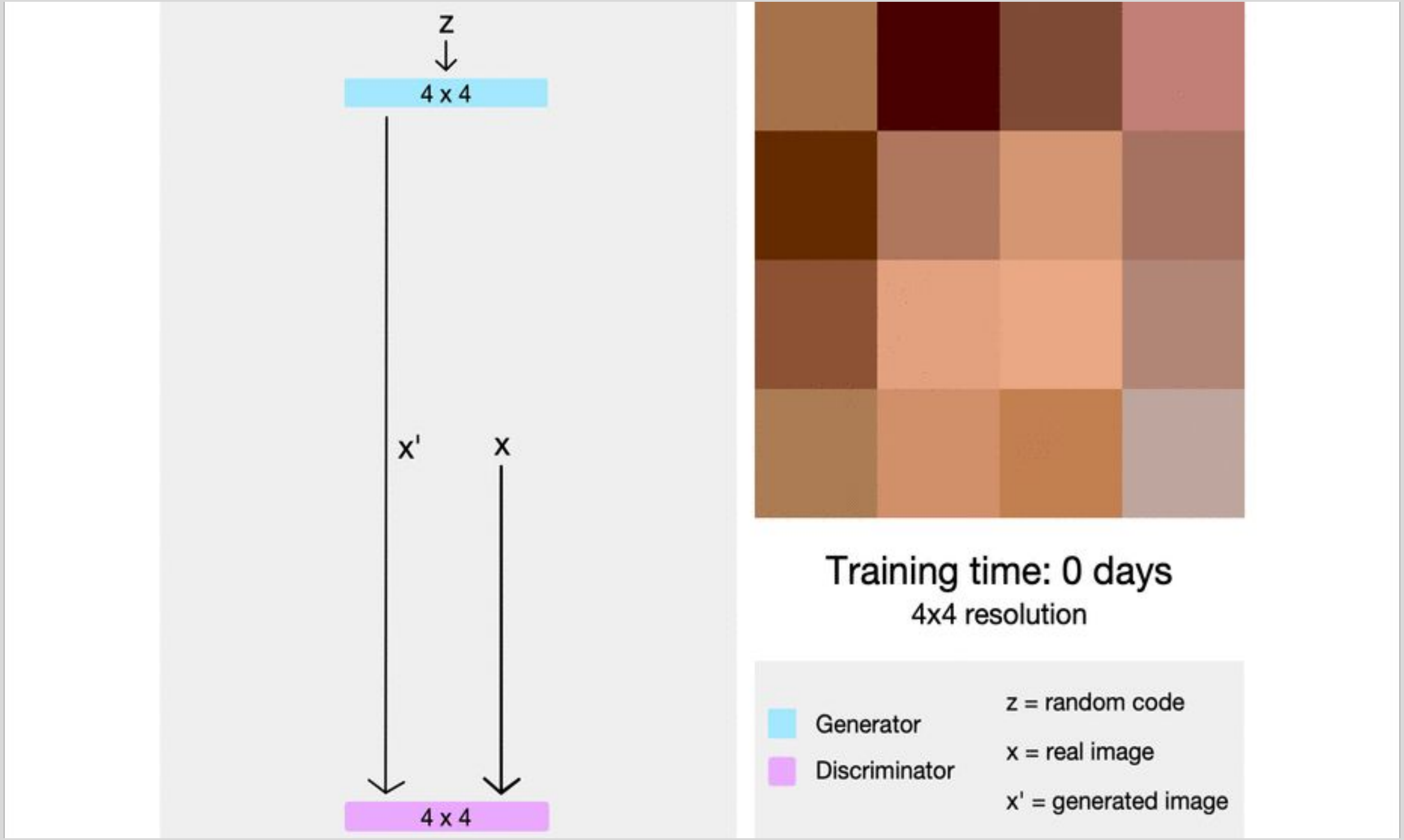
Source: towardsdatascience-StyleGAN and tuning

This work proposes a model for the generator that is inspired by the **style transfer network**



Source: towards datascience- Analysing how StyleGAN works: style incorporation in high quality images

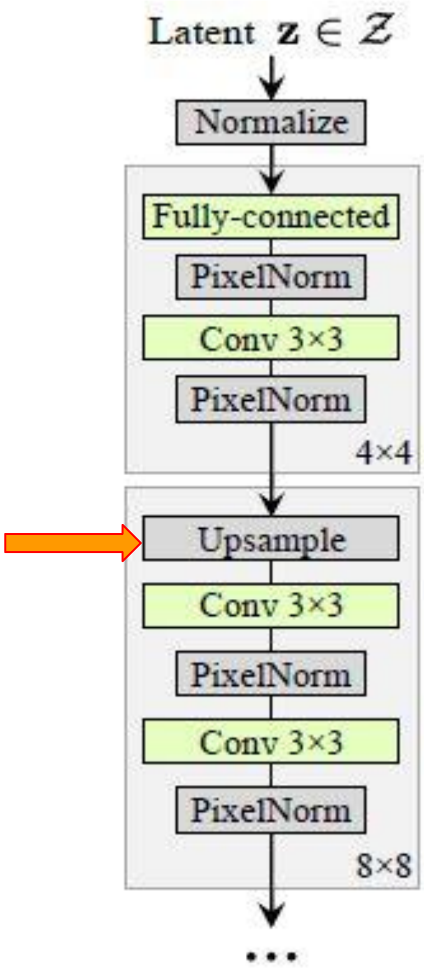
ProGAN + StyleGAN architecture

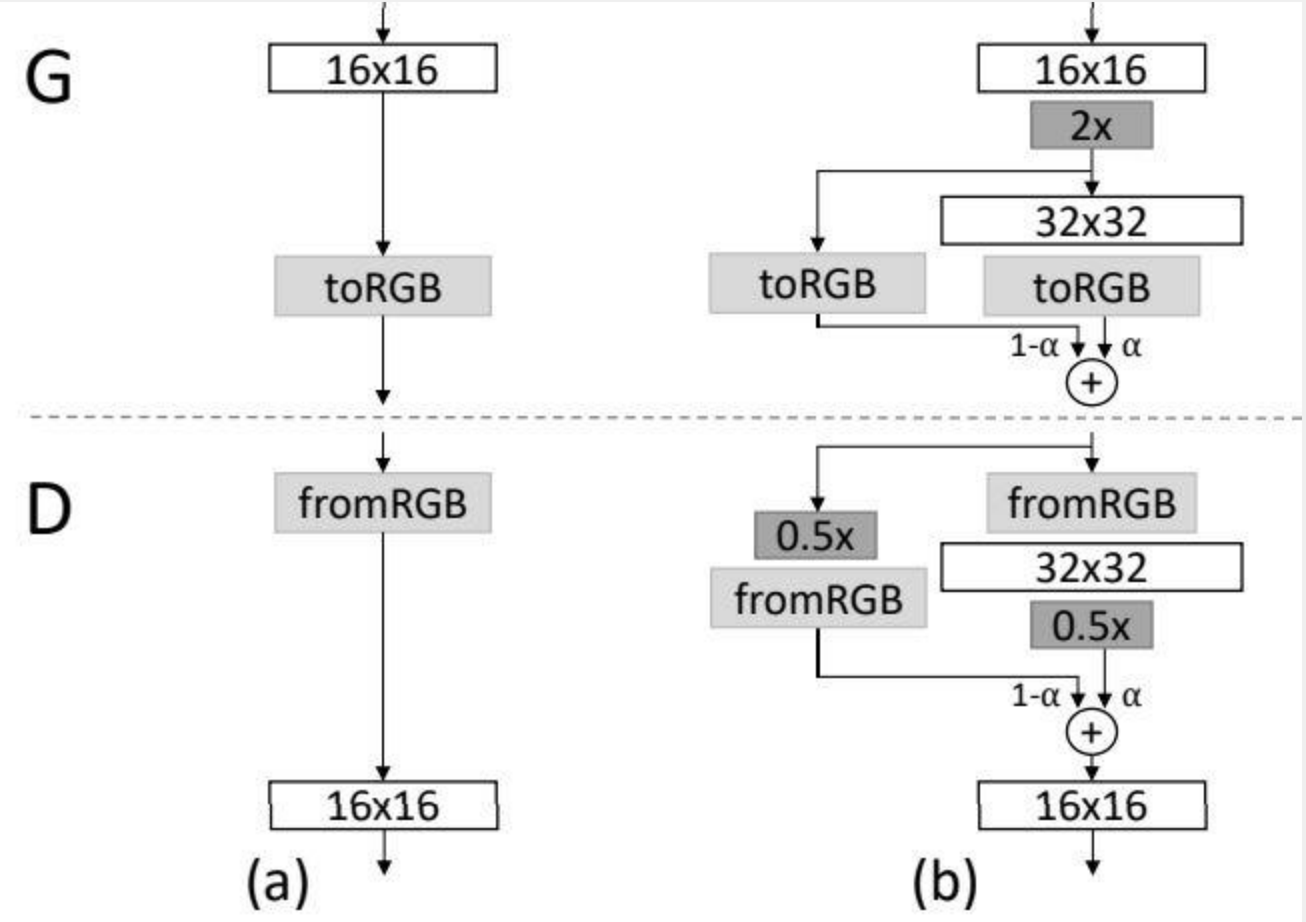


Source: towards datascience- Analysing how StyleGAN works: style incorporation in high quality images

- Replace nearest neighbor with bilinear upsampling

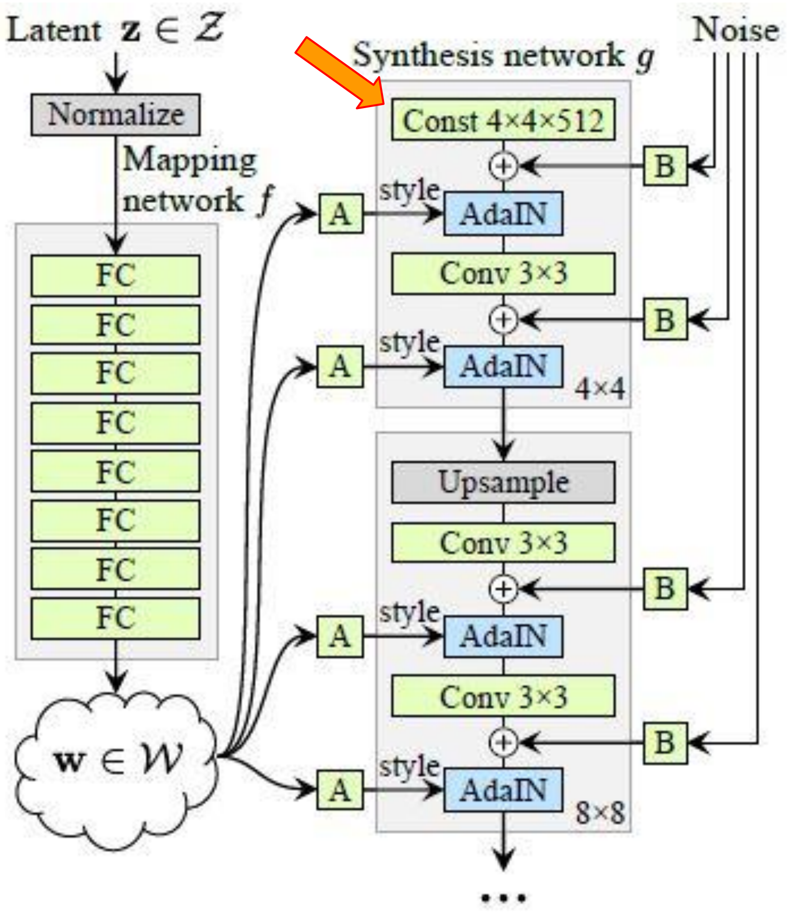
- Replace pooling with bilinear downsampling(in the discriminator)



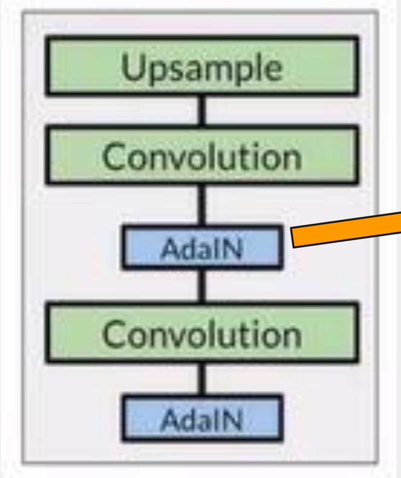


Source: https://arxiv.org/pdf/1710.10196.pdf

- Add mapping network and styles. Helped in preventing entanglement.
- Styles are generated in W and used in AdaIN operations.
- Remove traditional input.



Adaptive Instance Normalization simply scales the normalized input with style spatial statistics. This has profound implications.



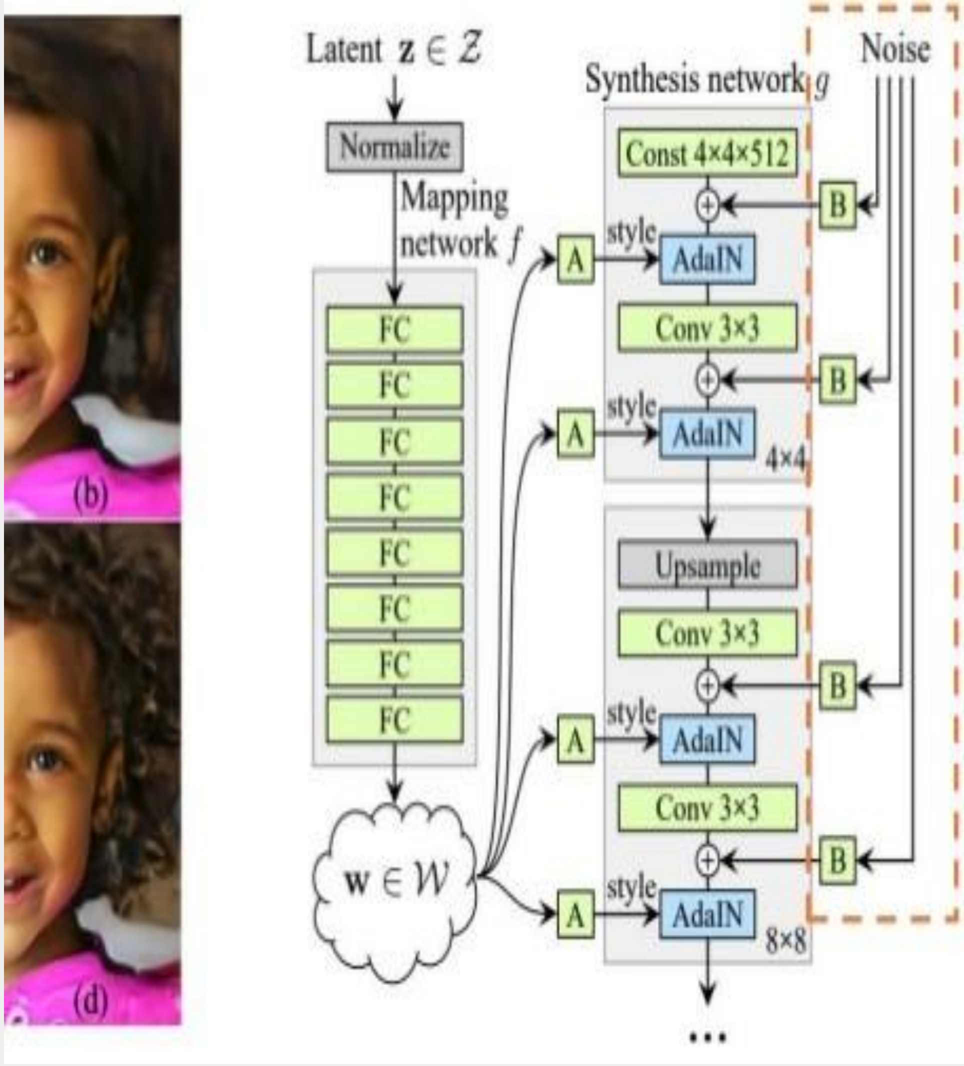
$$\mathrm{AdaIN}(\overset{\text{input}}{x}, \overset{\text{style}}{y}) = \underbrace{\sigma(y)}_{\text{Scaling}} \left(\frac{x - \mu(x)}{\sigma(x)} \right) + \underbrace{\mu(y)}_{\text{Bias(shifting)}}$$

Source: towards datascience- Analysing how StyleGAN works: style incorporation in high quality images

Stochastic Variation

There are many aspects in people's faces that are small and can be seen as stochastic, such as freckles, exact placement of hair, wrinkles, features which make the image more realistic and increase the variety of outputs.

The common method to insert these small features into GAN images is adding random **noise** to the input vector.



Tricks to avoid generating poor image

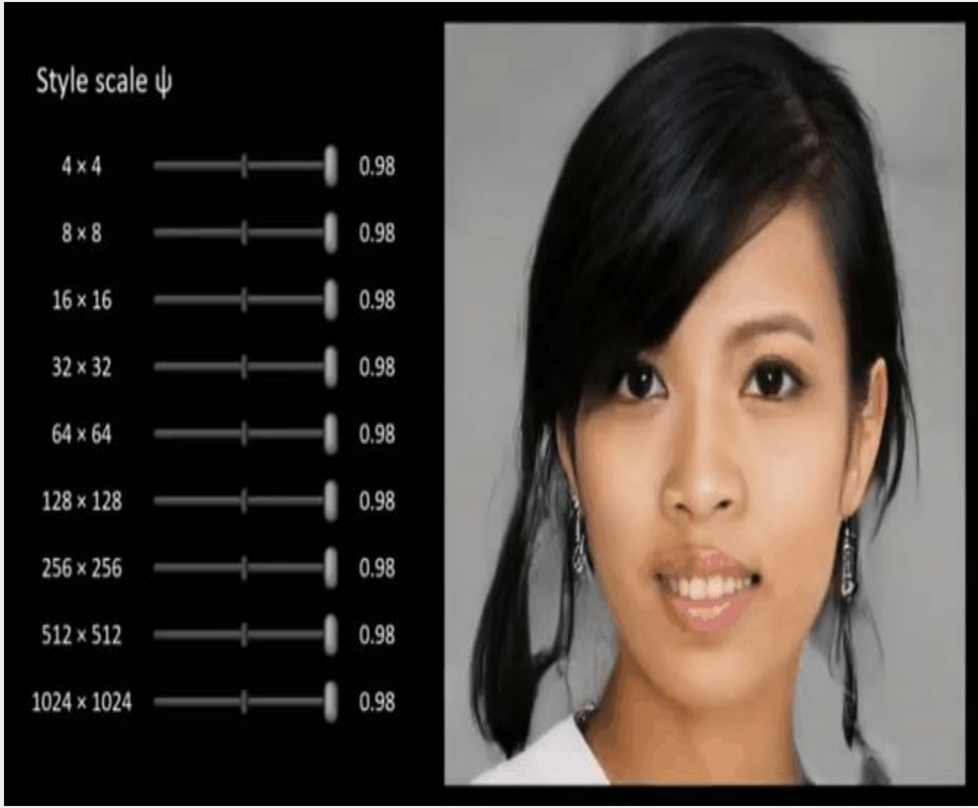
One of the challenges in generative models is dealing with areas that are poorly represented in the training data.

The generator isn’t able to learn them and create images that resemble them.

Hence creates bad-looking images.

$$W_{new} = W_{avg} + \boldsymbol{\psi} (W - W_{avg})$$

We can choose the strength at which each style is applied with respect to an “average face”



Mixing Regularization

How image quality is improved progressively:

Further Improvements in StyleGAN:

1. **Style Mixing**
2. **Stochastic Variation**

This created amazing results where a single image is released from two types of latent codes w1 and w2 from z1 and z2.

Adding per pixel noise will help in better understanding of the minute textures in the images

Prevents the network to memorise the features that are correlated

https://youtu.be/kSLJriaOumA

Style Mixing:

Generator thinks of an image as a collection of "styles" where each style controls the effects at a particular scale :

- Coarse styles —› pose, hair, face shape

- Middle styles —› facial features, eyes

- Fine styles —› color scheme

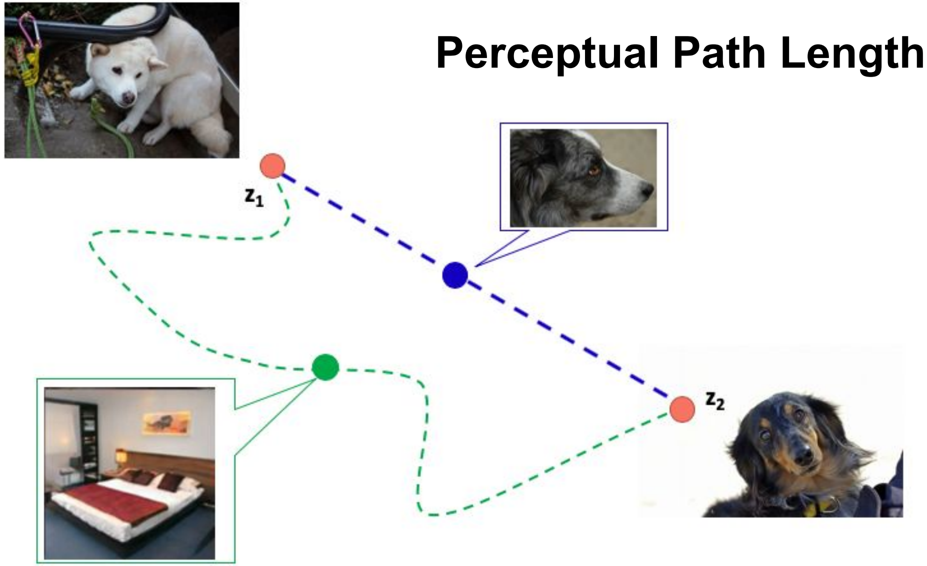
Disentanglement Studies:

it is true that if we had latent codes of each of the face features we would be able to control the features in the image and come up with a completely different representation

1. **Perceptual Path Length**
2. **Linear Separability (Known already)**

This will find the distance between feature maps visually. If separating the images far off and we observe a drastic change, this means the feature maps are entangled. Suppose removing the hair feature removes ears also, then we need to take care of it as removing hair shall not remove ears.

Perceptual Path Length



Conclusion and Misc. Takeouts:

1. StyleGAN Allows a control over the features in the image.
2. Major changes from previous tools is AdaIN and Mapping Network.
3. Advantages over ProGAN : *Now you can easily manipulate entangled features. High quality and realistic images are generated.*
4. **Limitation**: Check for the blue line in the image. Smile is not changing wrt the Face Movement

Small Comparisons:

Conditional GAN (Exploiting the given labels and preserving class identities in early stages)	These GANs use extra label information and result in better quality images and are able to control how generated images will look. cGANs learn to produce better images by exploiting the information fed to the model.
Cycle GAN (For unpiared data)	CycleGAN is an extension to the GAN for image-to-image translation without paired image data. That means that examples of the target image are not required as is the case with conditional GANs. This we have briefly covered in class
PROGAN (Improving images progressively and increasing depth of GANS)	PROGan is a change to the architecture and training of GAN models that involves progressively increasing the model depth during the training process. *starting from a low resolution, we add new layers that model increasingly fine details as training progresses.*
StyleGAN	(Idea is to grab features from latent code and reducing the load on generator, which now focuses on constant latent code)

References

StyleGAN2. This article explores changes made in… | by Connor Shorten | Towards Data Science
https://towardsdatascience.com/stylegan2-ace6d3da405d

Which Face is Real? Applying StyleGAN to Create Fake People - Exxact
https://blog.exxactcorp.com/which-face-is-real-applying-stylegan-to-create-fake-people/

From GAN basic to StyleGAN2. This post describes GAN basic… | by Akihiro FUJII | Analytics Vidhya | Medium
https://medium.com/analytics-vidhya/from-gan-basic-to-stylegan2-680add7abe82

Can anyone inform me please about the advantages and limitations of Generative Adversarial Networks (GANs) ?
https://www.researchgate.net/post/Can-anyone-inform-me-please-about-the-advantages-and-limitations-of-Generative-Adversarial-Networks-GANs

StyleGAN: Use machine learning to generate and customize realistic images | by Jamshed Khan | Heartbeat
https://heartbeat.fritz.ai/stylegans-use-machine-learning-to-generate-and-customize-realistic-images-c943388dc672

Which Face is Real?
https://www.kdnuggets.com/2019/04/which-face-real-stylegan.html#:~:text=Generative%20models%20have%20a%20limitation,the%20differences%20in%20the%20photographs.

Introduction to Generative Adversarial Networks (GANs): Types, and Applications, and Implementation | by Derrick Mwiti | Heartbeat
https://heartbeat.fritz.ai/introduction-to-generative-adversarial-networks-gans-35ef44f21193

6 GAN Architectures You Really Should Know - neptune.ai
https://neptune.ai/blog/6-gan-architectures

3 different types of generative adversarial networks (GANs) and how they work | Packt Hub
https://hub.packtpub.com/3-different-types-of-generative-adversarial-networks-gans-and-how-they-work/

GANs vs. Autoencoders: Comparison of Deep Generative Models | by Matthew Stewart, PhD Researcher | Towards Data Science
https://towardsdatascience.com/gans-vs-autoencoders-comparison-of-deep-generative-models-985cf15936ea

[StyleGAN] A Style-Based Generator Architecture for GANs, part 1 (algorithm review) | TDLS - YouTube
https://www.youtube.com/watch?v=SPI5uGCnxlc&ab_channel=MLExplained-AggregateIntellect-AI.SCIENCE

A Gentle Introduction to StyleGAN the Style Generative Adversarial Network
https://machinelearningmastery.com/introduction-to-style-generative-adversarial-network-stylegan/

No ownership for any image