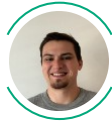


Must-Read Papers on GANs



Connor Shorten [Follow](#)

Mar 4 · 7 min read

Generative Adversarial Networks are one of the most interesting and popular applications of Deep Learning. This article will list 10 papers on GANs that will give you a great introduction to GAN as well as a foundation for understanding the state-of-the-art. Let's get to the list!

Here is the quick list if you are not interested in reading the descriptions of each paper:

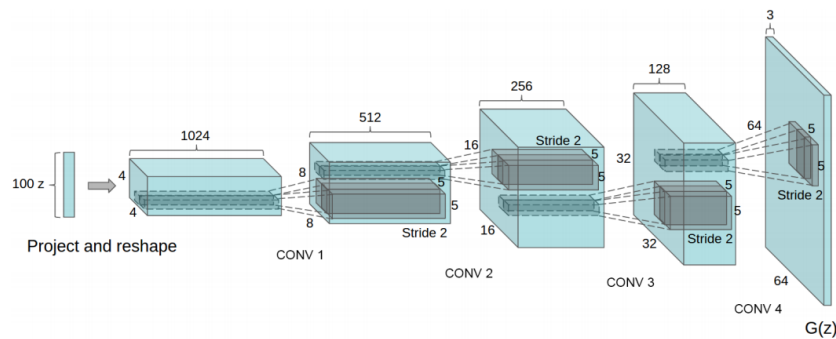
- 1—DCGANs
- 2—Improved Techniques for Training GANs
- 3— Conditional GANs
- 4—Progressively Growing GANs
- 5—BigGAN
- 6—StyleGAN
- 7—CycleGAN
- 8—Pix2Pix
- 9—StackGAN
- 10—Generative Adversarial Networks

. . .

DCGANs —Radford et al. (2015)

I would recommend starting your GAN journey with the DCGAN paper. This paper shows how convolutional layers can be used with

GANs and provides a series of additional architectural guidelines for doing this. The paper also discusses topics such as Visualizing GAN features, Latent space interpolation, using discriminator features to train classifiers, and evaluating results. All of these additional topics are bound to come up in your GAN research. In summation, the DCGAN paper is a must-read GAN paper because it defines the architecture in such a clear way that it is easy to get started with some code and begin developing an intuition for GANs.



DCGAN model—generator architecture with upsampling convolutional layers

Unsupervised Representation Learning with Deep Convolutional Generative...

In recent years, supervised learning with convolutional networks (CNNs) has seen...

[arxiv.org](#)

Improved Techniques for Training GANs—Salimans et al. (2016)

This paper, (whose authors include Ian Goodfellow), provide a series of recommendations for building on the architectural guidelines laid out in the DCGAN paper above. This paper will help you understand the best hypotheses for GAN instability. Additionally, this paper provides many additional techniques designed to stabilize the training of DCGANs. These include feature matching, minibatch

discrimination, historical averaging, one-sided label smoothing, and virtual batch normalization. Building onto a naive implementation of DCGANs with these suggestions is a great exercise for learning more about GANs.

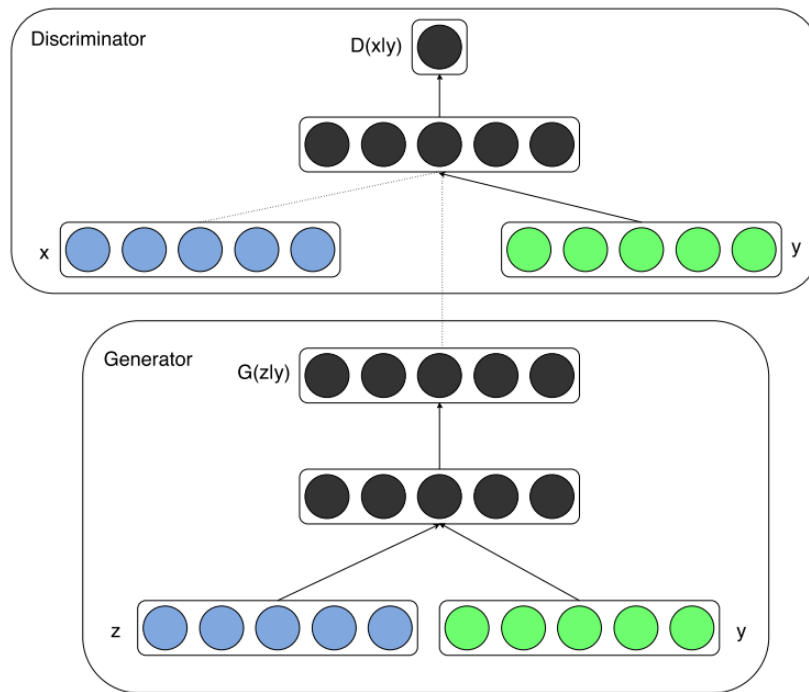
Improved Techniques for Training GANs

We present a variety of new architectural features and training procedures that we...

[arxiv.org](https://arxiv.org/abs/1606.03476)

Conditional GANs — Mirza and Osindero (2014)

This is a great paper which is a quick read. Conditional GANs are a central theme to state-of-the-art GANs. This paper shows how integrating the class labels of data results in more stable GAN training. This concept of conditioning GANs with prior information is a reoccurring theme in future works in GAN research and especially important for papers focusing on image-to-image or text-to-image.



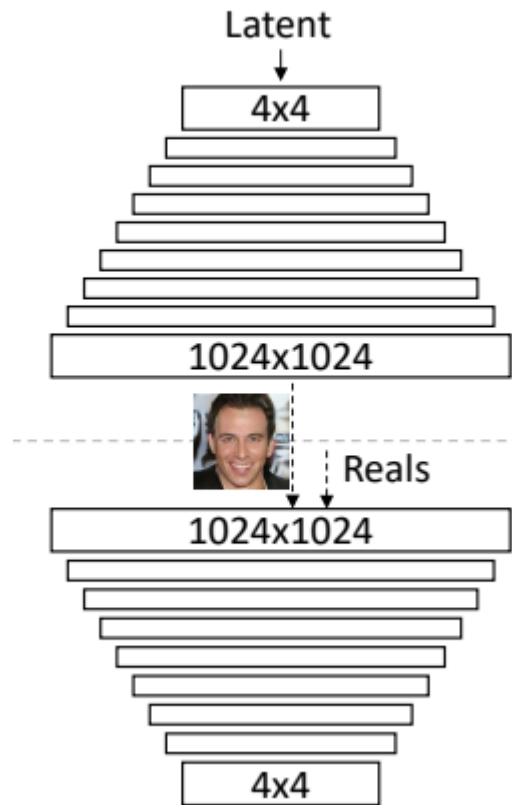
Architecture of the Conditional GANs, in addition to the random noise vector z , the class label y is concatenated together as input to the network

Conditional Generative Adversarial Nets

Generative Adversarial Nets [8] were recently introduced as a novel way to train...
arxiv.org

Progressively Growing of GANs for Improved Quality, Stability, and Variation —Karras et al. (2017)

The Progressively Growing GAN architecture is a must-read due to its impressive results and creative approach to the GAN problem. This paper uses a multi-scale architecture where the GAN builds up from 4^2 to 8^2 and up to 1024^2 resolution. GAN instability is largely increased with respect to the target image resolution size, and this paper shows a workaround to this problem.



This image depicts the multi-scale architecture of the Progressively Growing GAN architecture. The model goes from 4^2 progressively up to 1024^2

Progressive Growing of GANs for Improved Quality, Stability, and Variation

We describe a new training methodology for generative adversarial networks. The...

[arxiv.org](https://arxiv.org/abs/1809.02928)

BigGAN—Brock et al. (2019)

The BigGAN model is the current state-of-the-art on ImageNet generation. This model is difficult to implement on a local machine and there are many components to the architecture such as Self-Attention, Spectral Normalization, and cGAN with projection discriminators, that are each better explained in their own papers.

However, this paper provides a great overview to the ideas of the foundational papers that compose the current state-of-the-art.



Unbelievable samples from the BigGAN state-of-the-art model

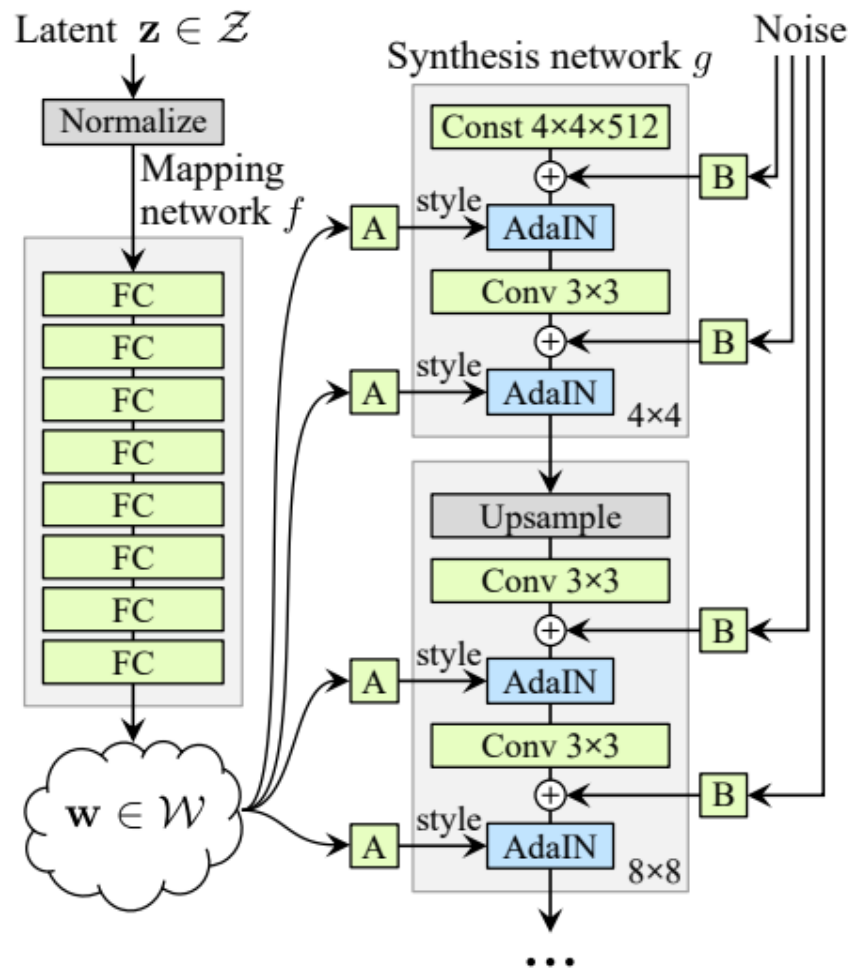
Large Scale GAN Training for High Fidelity Natural Image Synthesis

Despite recent progress in generative image modeling, successfully generating...

[arxiv.org](https://arxiv.org/abs/1809.02926)

StyleGAN—Karras et al. (2019)

The StyleGAN model is arguably the state-of-the-art in its way, especially in Latent Space control. This model borrows a mechanism from Neural Style Transfer known as Adaptive Instance Normalization, (AdaIN), to control the latent space vector z unlike anything before it. The combination of the mapping network and the distribution of AdaIN conditioning throughout the generator model makes this pretty difficult to implement yourself, but it is still a great read and contains many interesting ideas.



StyleGAN architecture that allows for state-of-the-art latent space control

A Style-Based Generator Architecture for Generative Adversarial Networks

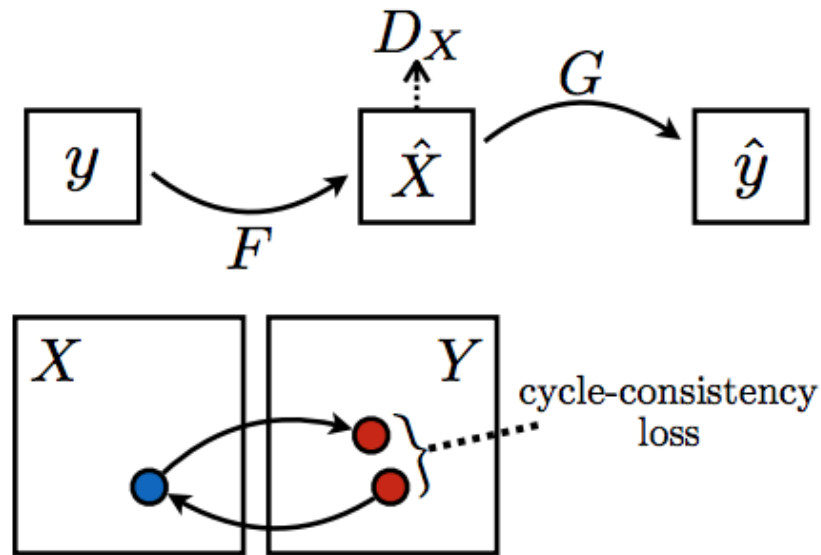
We propose an alternative generator architecture for generative adversarial...

[arxiv.org](https://arxiv.org/abs/1809.04875)

CycleGAN—Zhu et al. (2017)

The CycleGAN paper is different from the previous 6 papers mentioned because it discusses the problem of image-to-image translation rather than image synthesis from a random vector. CycleGAN more specifically deals with the case of image-to-image translation where you do not have paired training samples. However,

it is a great paper to read simply due to the elegance of the Cycle-Consistency loss formulation and the intuition on how this stabilizes GAN training. There are many cool applications CycleGAN can be used for as well such as super-resolution, style transfer, and horse to zebra.



Central Idea behind the Cycle Consistency Loss, a sentence translated from French to English and back to French should be the same sentence

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial...

Image-to-image translation is a class of vision and graphics problems where the...

arxiv.org

Pix2Pix—Isola et al. (2016)

Pix2Pix is another image-to-image translation GAN model. This framework uses paired training samples and investigates with many different configurations within the GAN model. One of the most interesting things to me when reading this paper was the discussion of the PatchGAN. The PatchGAN looks at 70×70 regions of the

image to determine if they are real or fake versus looking at the whole image. This model also shows an interesting U-Net style generator architecture as well as using ResNet-style skip connections in the generator model. This has many cool applications such as edge-maps to photo-realistic images.

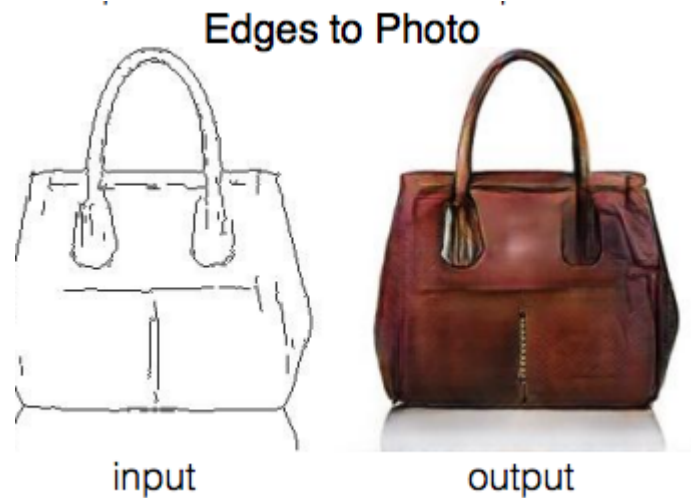


Image-to-Image translation with paired training samples

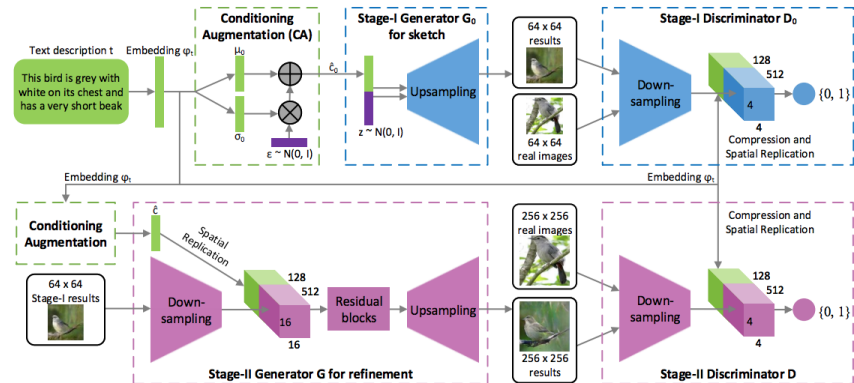
Image-to-Image Translation with Conditional Adversarial Networks

We investigate conditional adversarial networks as a general-purpose solution to...
[arxiv.org](https://arxiv.org/abs/1611.07017)

StackGAN—Zhang et al. (2017)

The StackGAN paper is very unique to the previous papers in this list. It is most similar to Conditional GANs and Progressively Growing GANs. The StackGAN model works similar to Progressively Growing GANs in the sense that it works on multiple scales. The StackGAN first outputs an image of resolution 64^2 and then takes this as prior information to generate an image of resolution 256^2 . The StackGAN is very unique to the other papers because it goes from natural language text to image. This is done by altering a text embedding such that it captures visual characteristics. This is a very interesting

paper to read and it would be amazing to see the latent space control exhibited in StyleGAN combined with the natural language interface defined in StackGAN.



Idea behind the StackGAN multi-scale architecture conditioned on a text embedding

StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative...

Synthesizing high-quality images from text descriptions is a challenging problem in...

arxiv.org

Generative Adversarial Networks —Goodfellow et al. (2014)

The original paper from Ian Goodfellow is a must-read for anyone studying GANs. This paper defines the GAN framework and discusses the ‘non-saturating’ loss function. This paper also gives the derivation for the optimal discriminator, a proof which frequently comes up in the more recent GAN papers. The paper also demonstrates the effectiveness of GAN empirically on the MNIST, TFD, and CIFAR-10 image datasets.

Generative Adversarial Networks

We propose a new framework for

estimating generative models via an...
arxiv.org

This embedded content is from a site that does not comply with the Do Not Track (DNT) setting now enabled on your browser.

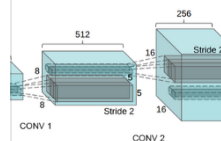
Please note, if you click through and view it anyway, you may be tracked by the website hosting the embed.

[Learn More about Medium's DNT policy](#)

Thanks for reading! I have provided blog post summaries of many of these papers published on TowardsDataScience listed below:

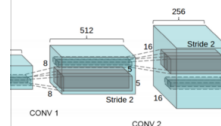
DCGANs (Deep Convolutional Generative Adversarial Networks)

One of the most interesting parts of Generative Adversarial Networks is the...
towardsdatascience.com



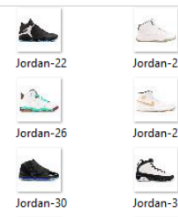
Deeper into DCGANs

Further exploring topics such as GAN overfitting, GAN feature visualization, and...
towardsdatascience.com



Generating Basketball Shoes with DCGANs

Code, results, and analysis on implementing DCGANs in Keras to generate 45x45...



medium.com



Jordan-34

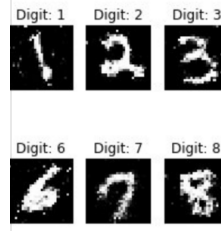


Jordan-35

Conditional GANs

Conditional GANs [1] are a very interesting extension to the GAN framework. This...

medium.com



Pix2Pix

This article will explain the fundamental mechanisms of a popular paper on Image...

towardsdatascience.com



Progressively-Growing GANs

The Progressively-Growing GAN architecture released from NVIDIA and...

towardsdatascience.com

