

GAN Variations

Researchers continue to find improved GAN techniques and new uses for GANs. Here's a sampling of GAN variations to give you a sense of the possibilities.

Progressive GANs

In a progressive GAN, the generator's first layers produce very low resolution images, and subsequent layers add details. This technique allows the GAN to train more quickly than comparable non-progressive GANs, and produces higher resolution images.

For more information see [Karras et al, 2017](https://arxiv.org/abs/1710.10196) (<https://arxiv.org/abs/1710.10196>).

Conditional GANs

Conditional GANs train on a labeled data set and let you specify the label for each generated instance. For example, an unconditional MNIST GAN would produce random digits, while a conditional MNIST GAN would let you specify which digit the GAN should generate.

Instead of modeling the joint probability $P(X, Y)$, conditional GANs model the conditional probability $P(X | Y)$.

For more information about conditional GANs, see [Mirza et al, 2014](https://arxiv.org/abs/1411.1784) (<https://arxiv.org/abs/1411.1784>).

Image-to-Image Translation

Image-to-Image translation GANs take an image as input and map it to a generated output image with different properties. For example, we can take a mask image with blob of color in the shape of a car, and the GAN can fill in the shape with photorealistic car details.

Similarly, you can train an image-to-image GAN to take sketches of handbags and turn them into photorealistic images of handbags.



In these cases, the loss is a weighted combination of the usual discriminator-based loss and a pixel-wise loss that penalizes the generator for departing from the source image.

For more information, see [Isola et al, 2016](https://arxiv.org/abs/1611.07004) (<https://arxiv.org/abs/1611.07004>).

CycleGAN

CycleGANs learn to transform images from one set into images that could plausibly belong to another set. For example, a CycleGAN produced the righthand image below when given the lefthand image as input. It took an image of a horse and turned it into an image of a zebra.



The training data for the CycleGAN is simply two sets of images (in this case, a set of horse images and a set of zebra images). The system requires no labels or pairwise correspondences between images.

For more information see [Zhu et al, 2017](#)

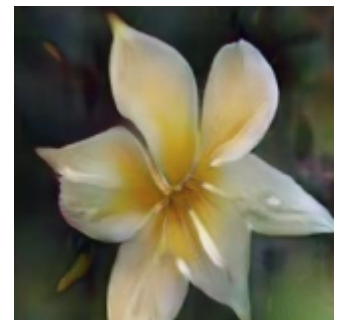
(http://openaccess.thecvf.com/content_ICCV_2017/papers/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.pdf)

, which illustrates the use of CycleGAN to perform image-to-image translation without paired data.

Text-to-Image Synthesis

Text-to-image GANs take text as input and produce images that are plausible and described by the text. For example, the flower image below was produced by feeding a text description to a GAN.

"This flower has petals that are yellow with shades of orange."



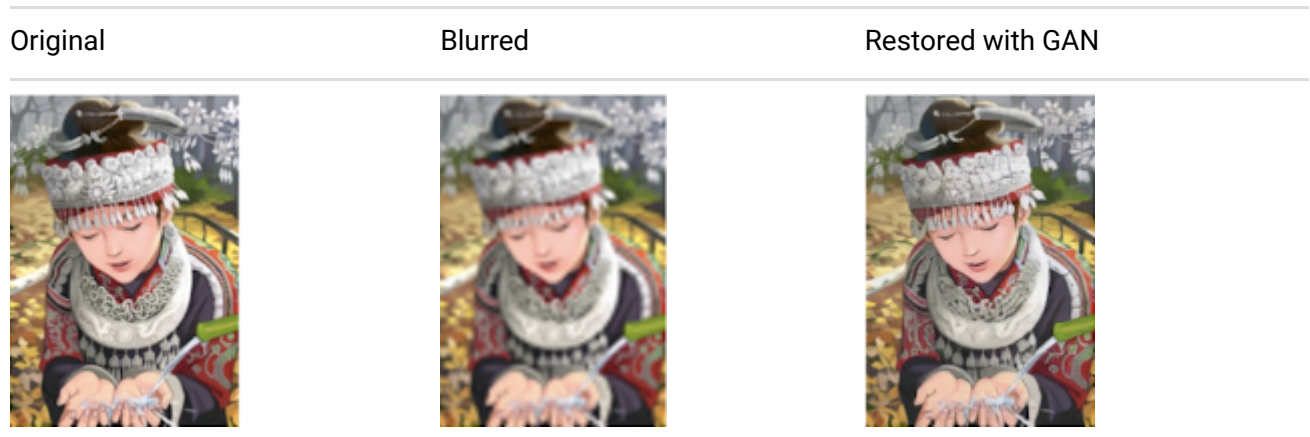
Note that in this system the GAN can only produce images from a small set of classes.

For more information, see [Zhang et al, 2016](#) (<https://arxiv.org/abs/1612.03242>).

Super-resolution

Super-resolution GANs increase the resolution of images, adding detail where necessary to fill in blurry areas. For example, the blurry middle image below is a downsampled version of

the original image on the left. Given the blurry image, a GAN produced the sharper image on the right:



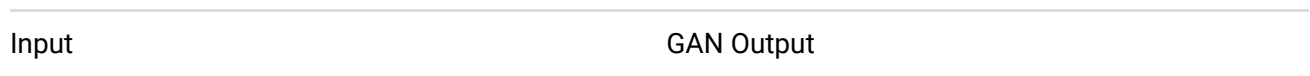
The GAN-generated image looks very similar to the original image, but if you look closely at the headband you'll see that the GAN didn't reproduce the starburst pattern from the original. Instead, it made up its own plausible pattern to replace the pattern erased by the down-sampling.

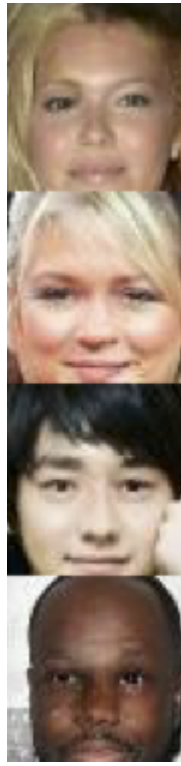
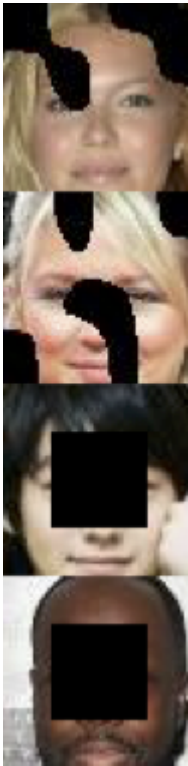
For more information, see [Ledig et al, 2017](https://arxiv.org/pdf/1609.04802.pdf) (<https://arxiv.org/pdf/1609.04802.pdf>).

Face Inpainting

GANs have been used for the *semantic image inpainting* task. In the inpainting task, chunks of an image are blacked out, and the system tries to fill in the missing chunks.

[Yeh et al, 2017](https://arxiv.org/pdf/1607.07539.pdf) (<https://arxiv.org/pdf/1607.07539.pdf>) used a GAN to outperform other techniques for inpainting images of faces:





Text-to-Speech

Not all GANs produce images. For example, researchers have also used GANs to produce synthesized speech from text input. For more information see [Yang et al, 2017](https://arxiv.org/pdf/1707.01670.pdf) (<https://arxiv.org/pdf/1707.01670.pdf>).

[Previous](#)

← [Common Problems](#) (/machine-learning/gan/problems)

[Next](#)

[TFGAN Colab Tutorials](#) (/machine-learning/gan/programming-exercise)



Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2019-04-28 UTC.