Generative adversarial network
from Wikipedia

A generative adversarial network (GAN) is a class of machine learning systems invented by Ian Goodfellow [1].

Two neural networks contest with each other in a game (in the sense of game theory, often but not always in the form of a zero-sum game). Given a training set, this technique learns to generate new data with the same statistics as the training set. For example, a GAN trained on photographs can generate new photographs that look at least superficially authentic to human observers, having many realistic characteristics.

Though originally proposed as a form of generative model for unsupervised learning, GANs have also proven useful for semi-supervised learning[2], fully supervised learning[3], and reinforcement learning [4]. In a 2016 seminar, Yann LeCun described GANs as "the coolest idea in machine learning in the last twenty years" [5].

Method

The generative network generates candidates while the discriminative network evaluates them. [1] The contest operates in terms of data distributions. Typically, the generative network learns to map from a latent space to a data distribution of interest, while the discriminative network distinguishes candidates produced by the generator from the true data distribution. The generative network's training objective is to increase the error rate of the discriminative network (i.e., "fool" the discriminator network by producing novel candidates that the discriminator thinks are not synthesized (are part of the true data distribution). [1][6]

A known dataset serves as the initial training data for the discriminator. Training it involves presenting it with samples from the training dataset, until it achieves acceptable accuracy. The generator trains based on whether it succeeds in fooling the discriminator. Typically the generator is seeded with randomized input that is sampled from a predefined latent space (e.g. a multivariate normal distribution). Thereafter, candidates synthesized by the generator are evaluated by the discriminator. Backpropagation is applied in both networks so that the generator produces better images, while the discriminator becomes more skilled at flagging synthetic images. [7] The generator is typically a deconvolutional neural network, and the discriminator is a convolutional neural network.

History

Adversarial machine learning has other uses besides generative modeling and can be applied to models other than neural networks. The general idea of learning via competition between players dates back to at least 1959 with the influential work of Arthur Samuel, demonstrating that algorithms could learn to play checkers via adversarial self-play [8].

Ian Goodfellow is generally recognized as having invented GANs [9][10] in 2014.[1] This paper included the first working implementation of a generative model based on adversarial networks, as well as game theoretic analysis establishing that the method is sound.

Review papers about GANs from academic sources, such as [11] generally assign credit for the idea to Ian Goodfellow and make no mention of Juergen Schmidhuber. However, Schmidhuber interrupted [12] a presentation by Goodfellow in 2016 demanding more credit be given to his earlier work on predictability minimization. Goodfellow himself states [13] that the most direct inspiration for GANs was noise-contrastive estimation [14], which uses the same loss function as GANs and which Goodfellow studied during his PhD in 2010-2014. Following the publication and success of GANs, it has come to light that other people had similar ideas but did not develop them similarly. An idea involving adversarial networks was published in a 2010 blog post by Olli Niemitalo. [15] This idea was never implemented and did not involve stochasticity in the generator and thus was not a generative model. An idea similar to GANs was used to model animal behavior by Li, Gauci and Gross in 2013. [16]

In the late 2010s, GANs began to produce disturbingly accurate syntheses. In 2017 a GAN was used for image enhancement focusing on realistic textures rather than pixel-accuracy, producing a higher image quality at high magnification. [17] In 2017, the first faces were generated. [18] These were exhibited in
February 2018 at the Grand Palais. [19] [20] Faces generated by StyleGAN [21] in 2019, drew comparisons with deepfakes [22] and poker tells. [23] [24] Beginning in 2017, GAN technology began to make its presence felt in the fine arts arena with the appearance of a newly developed implementation which was said to have crossed the threshold of being able to generate unique and appealing abstract paintings, and thus dubbed a "CAN", for "creative adversarial network". [25] A GAN system was used to create the 2018 painting Edmond de Belamy, which sold for $432,500. [26] An early 2019 article by members of the original CAN team discussed further progress with that system, and gave consideration as well to the overall prospects for an AI-enabled art. [27]

**Applications**

GAN applications have increased rapidly. [28] GANs that produce photorealistic images can be used to visualize interior design, industrial design, shoes, bags, and clothing items or items for computer games' scenes. Such networks were reported to be used by Facebook. [29] GANs can improve astronomical images[30], simulate gravitational lensing for dark matter research [31][32][33], reconstruct 3D models of objects from images,[34] and model patterns of motion in video, [35] GANs can be used to age face photographs to show how an individual's appearance might change with age. [36] GANs can also be used to transfer map styles in Cartography. [37]

In 2018, GANs reached the video game modding community, as a method of up-scaling low-resolution 2D textures in old video games by recreating them in 4k or higher resolutions via image training, and then down-sampling them to fit the game's native resolution (with results resembling the supersampling method of anti-aliasing). [38] With proper training, GANs provide a clearer and sharper 2D texture image magnitudes higher in quality than the original, while fully retaining the original's level of details, colors, etc. Known examples of extensive GAN usage include Final Fantasy VIII, Final Fantasy IX, Resident Evil REMake HD Remaster, and Max Payne. GAN 2D texture modding can be applied only to PC game releases.

StyleGAN itself is currently the sixth most trending Python project on GitHub.[39]

A variation of the GANs are used in training a network to generate optimal control inputs to nonlinear dynamical systems. Where the discriminatory network is known as a critic that checks the optimality of the solution and the generative network is known as an Adaptive network that generates the optimal control. The critic and adaptive network train each other to approximate a nonlinear optimal control. [40] GANs have been used to visualize the effect that climate change will have on specific houses[41].

**Concerns**

Concerns have been raised about the potential use of GAN-based human image synthesis for sinister purposes, e.g., to produce fake and/or incriminating photographs and videos.[42] GANs can be used to generate unique, realistic profile photos of people who do not exist, in order to automate creation of fake social media profiles[43]. As of May 2019, the state of California is considering a bill that would ban the use of GANs and related technologies to make fake pornography without the consent of the people depicted [44].

**References**


10. "Invented a way for neural networks to get better by working together" (https://www.innovatorsunder35.com/the-list/ian-goodfellow/). Innovators under 35. MIT Technology Review.


External Links

- A Style-Based Generator Architecture for Generative Adversarial Networks (https://arxiv.org/abs/1812.04948)