

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

1. Goodfellow, Ian; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). *Generative Adversarial Networks* (<https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>) (PDF). *Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014)*. pp. 2672–2680.
2. Salimans, Tim; Goodfellow, Ian; Zaremba, Wojciech; Cheung, Vicki; Radford, Alec; Chen, Xi (2016). "Improved Techniques for Training GANs". *arXiv:1606.03498* (<https://arxiv.org/abs/1606.03498>) [cs.LG (<https://arxiv.org/archive/cs.LG>)].
3. Isola, Phillip; Zhu, Jun-Yan; Zhou, Tinghui; Efros, Alexei (2017). "Image-to-Image Translation with Conditional Adversarial Nets" (<https://phillipi.github.io/pix2pix/>). *Computer Vision and Pattern Recog-*

nition.

4. Ho, Jonathon; Ermon, Stefano. "Generative Adversarial Imitation Learning" (<http://papers.nips.cc/paper/6391-generative-adversarial-imitation-learning>). Advances in Neural Information Processing Systems.
5. LeCun, Yann. "RL Seminar: The Next Frontier in AI: Unsupervised Learning" (<https://www.youtube.com/watch?v=IbjF5VjniVE>).
6. Luc, Pauline; Couprie, Camille; Chintala, Soumith; Verbeek, Jakob (2016-11-25). "Semantic Segmentation using Adversarial Networks". NIPS Workshop on Adversarial Training, Dec, Barcelona, Spain. 2016. arXiv:1611.08408 (<https://arxiv.org/abs/1611.08408>). Bibcode:2016arXiv161108408L (<http://adsabs.harvard.edu/abs/2016arXiv161108408L>).
7. Andrej Karpathy, Pieter Abbeel, Greg Brockman, Peter Chen, Vicki Cheung, Rocky Duan, Ian Goodfellow, Durk Kingma, Jonathan Ho, Rein Houthoofd, Tim Salimans, John Schulman, Ilya Sutskever, and Wojciech Zaremba, Generative Models (<https://openai.com/blog/generative-models/>), OpenAI, retrieved April 7, 2016
8. Samuel, Arthur (July 1959). "Some Studies in Machine Learning Using the Game of Checkers" (<https://ieeexplore.ieee.org/document/5392560>). IBM Journal of Research and Development.
9. Klok, Christie (February 21, 2018). "The GANfather: The man who's given machines the gift of imagination" (<https://www.technologyreview.com/s/610253/the-ganfather-the-man-whos-given-machines-the-gift-of-imagination/>). MIT Technology Review. Retrieved 14 May 2019.
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.
11. Cresswell, Antonia; White, Tom; Dumoulin, Vincent; Arulkumaran, Kai; Sengupta, Biswa; Bharath, Anil (Jan 1, 2018). "Generative Adversarial Networks: An Overview" (<https://ieeexplore.ieee.org/document/8253599/authors#authors>). IEEE Signal Processing.
12. Merity, Stephen. " "Schmidhuber interrupted the GAN tutorial" " (<https://twitter.com/Smerity/status/8057859999598>)
12. Merity, Stephen. " "Schmidhuber interrupted the GAN tutorial" " (<https://twitter.com/Smerity/status/805785999959879680>). Twitter.
13. Goodfellow, Ian. "Generative Adversarial Networks" (<http://www.iangoodfellow.com/slides/2016-12-04-NIPS.pdf>) (PDF).
14. Gutmann, Michael; Hyvärinen, Aapo. "Noise-Contrastive Estimation" (<http://proceedings.mlr.press/v9/gutmann10a/gutmann10a.pdf>) (PDF). International Conference on AI and Statistics.
15. Niemitalo, Olli (February 24, 2010). "A method for training artificial neural networks to generate missing data within a variable context" (<https://web.archive.org/web/20120312111546/http://yehar.com:80/blog/?p=167>). Internet Archive (Wayback Machine). Retrieved February 22, 2019.
16. Li, Wei; Gauci, Melvin; Gross, Roderich (July 6, 2013). "A Coevolutionary Approach to Learn Animal Behavior Through Controlled Interaction". Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation (GECCO 2013). Amsterdam, The Netherlands: ACM. pp. 223–230. doi:10.1145/2463372.2465801 (<https://doi.org/10.1145%2F2463372.2465801>).
17. Sajjadi, Mehdi S. M.; Schölkopf, Bernhard; Hirsch, Michael (2016-12-23). "EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis". arXiv:1612.07919 (<https://arxiv.org/abs/1612.07919>) [cs.CV (<https://arxiv.org/archive/cs.CV>)].
18. "This Person Does Not Exist: Neither Will Anything Eventually with AI" (<https://medium.com/@alagraphy/this-person-does-not-exist-neither-will-anything-if-artificial-intelligence-keeps-learning-1a9fcba728f>). March 20, 2019.
19. "ARTificial Intelligence enters the History of Art" (<https://www.issuewire.com/artificial-intelligence-enters-the-history-of-art-1620667772563815>). December 28, 2018.
20. Tom Février. "Le scandale de l'intelligence ARTificielle" (<https://link.medium.com/MYqBrGHIKV>).
21. "StyleGAN: Official TensorFlow Implementation" (<https://github.com/NVlabs/stylegan>). March 2, 2019 – via GitHub.

22. Paez, Danny (2019-02-13). "This Person Does Not Exist Is the Best One-Off Website of 2019" (<https://www.inverse.com/article/53280-this-person-does-not-exist-gans-website=website=Inverse>). Retrieved 2019-02-16.
23. BESCHIZZA, ROB (2019-02-15). "This Person Does Not Exist" (<https://boingboing.net/2019/02/15/this-person-does-not-exist.html>). Boing-Boing. Retrieved 2019-02-16.
24. Horev, Rani (2018-12-26). "Style-based GANs – Generating and Tuning Realistic Artificial Faces" (<https://www.lyrn.ai/2018/12/26/a-style-based-generator-architecture-for-generative-adversarial-networks/>). Lyrn.AI. Retrieved 2019-02-16.
25. Elgammal, Ahmed; Liu, Bingchen; Elhoseiny, Mohamed; Mazzone, Marian (2017). "CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms". arXiv:1706.07068 (<https://arxiv.org/abs/1706.07068>) [cs.AI (<https://arxiv.org/archive/cs/AI>)].
26. Cohn, Gabe (2018-10-25). "AI Art at Christie's Sells for \$432,500" (<https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html>). The New York Times.
27. Mazzone, Marian and Ahmed Elgammal (21 February 2019). "Art, Creativity, and the Potential of Artificial Intelligence" (<https://www.mdpi.com/2076-0752/8/1/26>). Arts. Retrieved 3 April 2019.
28. Caesar, Holger (2019-03-01), A list of papers on Generative Adversarial (Neural) Networks: nightrome/reallyawesome-gan (<https://github.com/nightrome/really-awesome-gan>), retrieved 2019-03-02
29. Greenemeier, Larry (June 20, 2016). "When Will Computers Have Common Sense? Ask Facebook" (<https://www.scientificamerican.com/article/when-will-computers-have-common-sense-ask-facebook/>). Scientific American. Retrieved July 31, 2016.
30. Schawinski, Kevin; Zhang, Ce; Zhang, Hantian; Fowler, Lucas; Santhanam, Gokula Krishnan (2017-02-01). "Generative Adversarial Networks recover features in astrophysical images of galaxies beyond the deconvolution limit". Monthly Notices of the Royal Astronomical Society: Letters. 467 (1): L110–L114. arXiv:1702.00403 (<https://arxiv.org/abs/1702.00403>). Bibcode:2017MNRAS.467L.110S (<http://adsabs.harvard.edu/abs/2017MNRAS.467L.110S>). doi:10.1093/mnrasl/slx008 (<https://doi.org/10.1093%2Fmnrasl%2Fslx008>).
31. Kincade, Kathy. "Researchers Train a Neural Network to Study Dark Matter" (<https://www.rdmag.com/news/2019/05/researchers-train-neural-network-study-dark-matter>). R&D Magazine.
32. Kincade, Kathy (May 16, 2019). "CosmoGAN: Training a neural network to study dark matter" (<https://phys.org/news/2019-05-cosmogon-neural-network-dark.html>). Phys.org.ws/2019-05-cosmogon-neural-network-dark.html). Phys.org.
33. "Training a neural network to study dark matter" (<https://www.sciencedaily.com/releases/2019/05/190516145206.htm>). Science Daily. May 16, 2019.
34. "3D Generative Adversarial Network" (<http://3dgan.csail.mit.edu/>). 3dgan.csail.mit.edu.
35. Vondrick, Carl; Pirsiavash, Hamed; Torralba, Antonio (2016). "Generating Videos with Scene Dynamics" (<http://www.cs.columbia.edu/~vondrick/tinyvideo/>). carlvondrick.com. arXiv:1609.02612 (<https://arxiv.org/abs/1609.02612>). Bibcode: 2016arXiv160902612V (<http://adsabs.harvard.edu/abs/2016arXiv160902612V>).
36. Antipov, Grigory; Baccouche, Moez; Dugelay, Jean-Luc (2017). "Face Aging With Conditional Generative Adversarial Networks". arXiv:1702.01983 (<https://arxiv.org/abs/1702.01983>) [cs.CV (<https://arxiv.org/archive/cs/CV>)].
37. Kang, Yuhao; Gao, Song; Roth, Rob (2019). "Transferring Multiscale Map Styles Using Generative Adversarial Networks" (<https://geods.geography.wisc.edu/archives/1192>). International Journal of Cartography.
38. Tang, Xiaoou; Qiao, Yu; Loy, Chen Change; Dong, Chao; Liu, Yihao; Gu, Jinjin; Wu, Shixiang; Yu, Ke; Wang, Xintao (2018-09-01). "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks" (<https://arxiv.org/abs/1809.00219v2>).
39. Roboticist, Javad Amirian (February 14, 2019). "#StyleGan, is ranked 6th among trending python projects on #Github. #DeepLearning #GenerativeNetworks" (<https://twitter.com/AmiryanJavad/status/1096126840434053120>).

40. Padhi, Radhakant; Unnikrishnan, Nishant. "A single network adaptive critic (SNAC) architecture for optimal control synthesis for a class of nonlinear systems" (<https://doi.org/10.1016/j.neunet.2006.08.010>). www.sciencedirect.com. doi: 10.1016/j.neunet.2006.08.010 (<https://doi.org/10.1016%2Fj.neunet.2006.08.010>). Retrieved 2019-05-06.
41. "AI can show us the ravages of climate change" (<https://www.technologyreview.com/f/613547/ai-can-show-us-the-ravages-of-climate-change/>). MIT Technology Review. May 16, 2019.
42. msmash (2019-02-14). " 'This Person Does Not Exist' Website Uses AI To Create Realistic Yet Horrifying Faces" (<https://tech.slashdot.org/story/19/02/14/199200/this-person-does-not-exist-website-uses-ai-to-create-realistic-yethorrifying-faces>). Slashdot. Retrieved 2019-02-16.
43. Doyle, Michael (May 16, 2019). "John Beasley lives on Saddlehorse Drive in Evansville. Or does he?" (<https://www.courierpress.com/story/news/crime/2019/05/16/john-beasley-lives-saddlehorse-drive-evansville-does-he/3700111002/>). Courier and Press.
44. Targett, Ed (May 16, 2019). "California moves closer to making deepfake pornography illegal". Computer Business Review.

External Links

- Knight, Will. "5 Big Predictions for Artificial Intelligence in 2017" (<https://www.technologyreview.com/s/603216/5-big-predictions-for-artificial-intelligence-in-2017/>). MIT Technology Review. Retrieved 2017-01-05.
- A Style-Based Generator Architecture for Generative Adversarial Networks (<https://arxiv.org/abs/1812.04948>)
- This Person Does Not Exist (<https://www.thispersondoesnotexist.com/>), Photorealistic images of people who do not exist, generated by StyleGAN Retrieved from "https://en.wikipedia.org/w/index.php?title=Generative_adversarial_network&oldid=897570264"