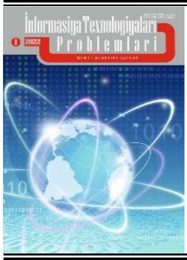

Available online at www.jpit.az

13 (1)
2022

Analysis of generative adversarial networks

Yadigar N. Imamverdiyev [a], Firangiz I. Musayeva [b]

[a]Azerbaijan Technical University, Institute of Artificial Intelligence, H. Javid, ave. 25, AZ1148 Baku, Azerbaijan
[b]Institute of Information Technology, Azerbaijan National Academy of Sciences, B. Vahabzade str., 9A, AZ1141 Baku, Azerbaijan

[a]yadigar@iit.science.az; [b]sadiyeva.firengiz@gmail.com

[a]0000-0002-3710-1046

ARTICLE INFO

http://doi.org/10.25045/jpit.v13.i1.03

Article history:
Received 2 September 2021
Received in revised form 18 November 2021
Accepted 21 January 2022

Keywords:
Neural networks
Generative models
Generative adversarial networks
Auto encoders
Generator
Discriminator

ABSTRACT

Recently, a lot of research has been done on the use of generative models in the field of computer vision and image classification. At the same time, effective work has been done with the help of an environment called generative adversarial networks, such as video generation, music generation, image synthesis, text-to-image conversion. Generative adversarial networks are artificial intelligence algorithms designed to solve the problems of generative models. The purpose of the generative model is to study the set of training patterns and their probable distribution. The article discusses generative adversarial networks, their types, problems, and advantages, as well as classification and regression, segmentation of medical images, music generation, best description capabilities, text image conversion, video generation, etc. general information is given. In addition, comparisons were made between the generative adversarial network algorithms analyzed on some criteria.

Introduction

Many available approaches to the development of artificial intelligence are based on machine learning. Supervised learning has been a widely used and successful type of machine learning so far. The Generative Adversarial Network (GAN), which has become very popular in recent years in the field of artificial intelligence, refers to the class of deep neural networks. These networks were proposed by Ian Goodfellow and colleagues in 2014 (Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, ... & Bengio, 2020), and interesting results were obtained in the fields of application such as video, photo, music, and synthetic data generation, etc. GANs are used in various fields such as computer vision, natural language processing, synthetic time series, and semantic segmentation. GANs refer to the family of generative models in machine learning. Compared to other generative models, such as variational autoencoders, GANs offer numerous advantages, such as precisely predictable density function control and efficient generation of any pattern (Warde-Farley, Bengio, 2017).

GANs are characterized by the training of two competing networks. There are some metaphors about these networks: one is thought to be a thief of art and the other one is the expert who defines him. In GANs, a painting thief known as a generator (G) generates false patterns in order to create real images, while an expert known as a discriminator (D) takes both false patterns and real (original) images and tries to separate them. Both are taught simultaneously and in competition with each other. Most importantly, the generator does not have a direct access to real images. The discriminator has access to both false patterns and patterns taken from a collection of real images. Via the discriminator, the same error signal is used to allow the generator to generate better quality and more realistic false patterns (Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, ... & Bengio, 2020).

This paper highlights the theoretical foundations and structure of generative competing networks, and comparatively analyzes their problems, advantages, learning algorithms, and studies their applications in many fields.

1. GAN theory

Since the first application of GANs, several interesting studies have been conducted for pattern generation and for its application to various fields. In general, generative models are grouped into two categories: Restricted Boltzmann Machine (RBM), Naive Bayes Model (NBM), and machine learning-based traditional generative models as the Hidden Markov Model (HMM) (Donahue, Kr¨ ahenb¨uhl and Darrell, 2016). Another is a training model consisting of Variational AutoEncoder (VAE), GANs and derivative models. GAN is a generative model generating target data using hidden variables. Significantly, game training is conducted between the generator and the discriminator in the model, and the target variables, which consist of the distribution of real data, are generated by random variables. Compared to traditional machine learning algorithms, the model includes more functional and more application scenario. They often perform better results in big data sets, such as ImageNet CIFAR. The development process of GANs can be explained in three phases. The first phase is the period from the creation of GAN models in 2014-2015 to the development of the Deep Convolutional Generative Adversarial Networks (DCGAN) model, the second phase covers the period from the application of the DCGAN model to the development of the Wasserstein GAN (WGAN) model in 2015-2017. Finally, the third phase is considered to be the period after the development of the WGAN model from 2017 to the present. The DCGAN and WGAN models constitute the basis of each stage. Compared to previous GANs, DCGAN can be applied more easily and eliminates the mode violation problem. The emergence of WGAN has become important for GAN models. This model is capable to generate the patterns with the highest quality. Various GAN models appeared according to different requirements of each of the different patterns and scenarios. Moreover, better results have been achieved in fields as medicine and security. BigGAN generated the images with high accuracy and low quality for the first time. StyleGAN developed completely differently in the field of GAN and achieved better performance in the human face image generation. In addition, this model also provides the generation of other high-quality images. The usage of algorithms such as AgecGAN is more appropriate for face recognition (Dumoulin, Belghazi, Poole, Mastropietro, Lamb, Arjovsky, & Courville, 2016, Creswell, White, Dumoulin, Arulkumaran, Sengupta, & Bharath, 2018).

2.1. The structure of GAN

The idea of generative adversarial networks is very simple: two neural networks (generator and discriminator) are taken and in the process of training they “compete” with each other, i.e., the generator builds a statistical model of data and accordingly creates realistic objects, while the discriminator determines whether they are real or false (Figure 1).

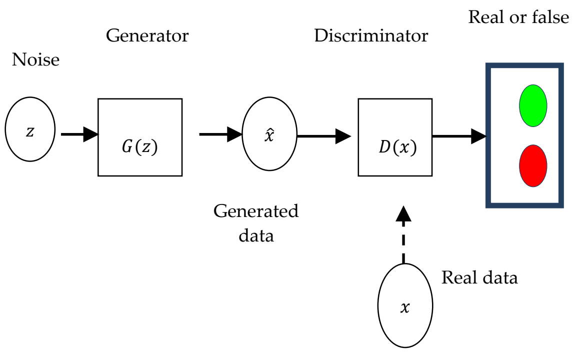

Figure 1. The scheme of Generative Adversarial

The generator generates the data by distributing real data using random noise. The discriminator distinguishes between real data and false data. Based on this, a two-player game is implemented for the GAN model. Game training is used to optimize the weight parameters between the two networks by developing the model's generalization ability. Consequently, the distribution of false data generated by the generator is adapted to the distribution of real data. The balanced state of the model is the situation in which the discriminator cannot distinguish false data from real ones. False patterns are generated by the generator $G(z)$ in the GAN architecture presented above (Figure 1.). The discriminator accepts two types of input data (real and false data) and determines whether they are real or not. GAN loss function is based on a min-max game of two players in a network competing with each other in a zero-sum game environment. The discriminator has to separate the data from the generator from the real data and optimize the weights of the network model with a rollback algorithm. Input parameters of the discriminator are x and $\theta(D)$, and the loss function is as follows (Yang, Yan, Zhang, Yu, Shi, Mou, & Wang, 2018).

$$V\left(D,\theta^{(D)}\right) = -E_{x\sim p_r(x)}[logD(x)] - E_{z\sim p_g(z)}\left[log\left(1 - D\left(G(z)\right)\right)\right]. \quad (1)$$

where, $p_r(x)$ is a distribution of real data, $p_g(z)$ is a distribution of input noise (mainly, a Gaussian random distribution is considered), $D(x)$ is a discriminator function $G(z)$ is a generator function. The loss function of the generator is demonstrated in formula (2).

$$V\left(G,\theta^{(G)}\right) = -E_{z\sim p_g}\left[-log\,D\left(G(z)\right)\right]. \quad (2)$$

The word "adversary" symbolizes the competition between these two networks because the goals of these networks are contradictory, and their relationship can be described as an antagonistic game. The generator and the discriminator "compete" with each other in the following min-max game:

$$\min_G \max_D V(G,D) = \mathbb{E}_{x\sim p_r(x)}\,[\log D(x)] + \mathbb{E}_{z\sim p_g(z)}[\log(1 - D(G(z)))]. \quad (3)$$

The discriminator tries to bring the probability of $D(G(z))$ closer to 0, that is, this tries to maximize $(1 - D(G(z)))$. Correspondingly, the Generator tries to bring $D(G(z))$ closer to 1 so that the discriminator makes a mistake and accepts the generated data as a real (formula (3)). That is, the generator tries to minimize $(1 - D(G(z)))$. In a competitive collaborative learning process, the system may reach Nash equilibrium, in which the differentiation of network-generated data from natural data is possible. The min-max game has only solution way in the space of arbitrary G and D functions. $p_r(x) = p_g(z)$, ∀x restores the distribution of training data and $D = 1/2$ for any x.

During GAN training, the generator and the discriminator are taught in turn (first the D discriminator, then the G generator). When one of the networks is trained, another remains constant. Theoretically, in order to obtain the most appropriate solution in $V(G,D)$, the discriminator must first be trained k times and then the generator must be trained once. However, in practice $k = 1$ is taken. In order to get the maximum value of $V(G,D)$ the following $D^*(x)$ is obtained after its derivative.

$$D^*(x) = \frac{p_r(x)}{p_r(x)+p_g(x)}. \quad (4)$$

Kullback–Leibler divergence between probability distribution $p_r(x)$ and $p_g(z)$ can be achieved by applying the above objective function. This can superiorly explain the training of the model (Lucic, Kurach, Michalski, Gelly, & Bousquet, 2017, Alguliyev, Abdullayeva, & Ojagverdiyeva, 2020).

2.2. Problems of GAN

GAN have several shortcomings. Mainly, training these networks is problematic and unstable process, as it is necessary to train both of them and to balance their recognition accuracy.

1) Mode collapse (generator gives the same object at any input values) is the lack of generally accepted approaches to assessing the effectiveness of GAN. When a generator generates similar data, it is called a partial collapse, and in the worst case, if only one data is generated, it is called a complete collapse. For balancing the distribution of real and false mass-generated data by the discriminator, the variety of the generator can be increased by using practical tricks or several GAN involving different probability distribution modes (Bang, & Shim, 2018, July, Che, Li, Jacob, Bengio, & Li, 2016, Metz, Poole, Pfau, & Sohl-Dickstein, 2016).

2) Training saddle points. The Hessian loss function in GRS is endless. Therefore, the optimal

solution is to find a saddle point, rather than a minimum point. In deep learning, most optimizers depend only on the primary derivation of the loss function. The approach to the saddle point is accepted as a good commencement for GAN (Arjovsky, & Bottou, 2017).

3) Evaluation of generative models. "How can the reliability of data generated by generative models be measured?", "Should probability assessment be used?", and other questions are not only for GAN, but for probability models in general. In (Theis, Oord, & Bethge, 2015), The GAN is evaluated by different metrics. The ways to get different conclusions about the quality of generated data are presented.

Offering various practical and theoretical solutions to solve these problems is significant. One of the main problems of GAN is to measure and minimize the distance between the probability distributions p_r and p_g. When the generator remains constant, the cross-entropy of discriminator training is minimized. In the structure of the above-mentioned GAN, the following formula can be obtained from formulas (4) and (3).

$$V\left(D, \theta^{(G)}\right) = KL(p_r || \frac{p_r+p_g}{2}) + KL(p_g || \frac{p_r+p_g}{2}) - 2log2. \quad (5)$$

Thus, the main problem of GAN today is the selection of the distance function.

The generative model is built to approximate data distribution. Assume that $x \sim p_{data}(x)$ and the finite dataset from this distribution $X = \{x | x \sim p_r(x)\}, |X| = n$ is given. We should create such a model so that $p_{model}(x, \theta) \sim p_r(x)$. The generative model can be used to generate new data similar to existing ones. For instance, a generative model can be used to construct a common distribution of objects and classes, and then by the use of Bayesian formula, we can find the probability that an object refers to a class.

2. Models of Generative Adversarial Network

Researchers distinguish different forms according to the type of neural network used in the GAN architecture:

A. Vanilla GAN is the first proposed GAN architecture. Both the discriminator and the generator are multilayer propagation networks (e.g., multilayer perceptron). This architecture is used for MNIST (Modified National Institute of Standards and Technology database), CIFAR (Canadian Institute for Advanced Research), and Toronto dataset. In practice, a missing gradient problem may arise in the optimization of the generator, and for the training of G the maximum value of the expression $D(G(z))$ should be found.

B. Deep Convolutional GAN - multilayer convolution networks are used and applied to image synthesis. This environment consists of two networks: a convolutional neural network (CNN) called a network generator, and a de-CNN called a discriminator. The shortcoming is the training process of the model being very long. Restrictions of CNN architecture include (Salimans, Goodfellow, Zaremba, Cheung, Radford, & Chen, 2016).

- Deletes all levels of pooling layers with step convolutions.
- Bolsman machines are used for G and D.
- ReLU (Rectified Linear Unit) and leaky-ReLU are used in generator and discriminator networks, respectively.

C. Conditional GAN (CGAN) – this generator receives hidden noise vector c (class attribute, text, or image) and z as input. Thus, $G(z/c)$ is the generation of real-like patterns. The generator and the discriminator can generate the distribution of a dataset by specifying a specific class. The loss function of conventional GAN is shown below.

$$L_{CGAN} = -E_{x \sim p_r(x)}\left[logD\left(\frac{x}{c}\right)\right] - E_{z \sim p_z(z)}\left[log\left(1 - D\big(G(z/c)\big)\right)\right]. \quad (6)$$

where, $p_z(z)$ is a noise and c is the conventional data generator, real x and the conditional c data are the inputs of the discriminator (Mirza, & Osindero, 2014, Isola, Zhu, Zhou, & Efros, 2017, Odena, Olah, & Shlens, 2017, July).

D. CycleGAN – is the most advanced GAN for image generation. These networks do not require binary datasets for passage between domains, as it is very difficult to obtain such data. However, periodic networks should be trained with data from two different X and Y domains (for instance, X - horses, Y - zebras). The "sequential loss of time" mechanism is used to restrict the transition from one domain to another (Yeh, Chen, Yian Lim, Schwing, Hasegawa-Johnson, & Do, 2017).

E. **Wasserstein GAN** loss function is modified by including Wasserstein distance. Consequently, the loss function of the Wasserstein network is related to image quality. In addition, learning sustainability improves and it does not depend on architecture (Arjovsky, Chintala, & Bottou, 2017, July).

F. **Progressive GAN (ProGAN) –** created on the basis of a developed Wasserstein network, and new layers are gradually added during training. Each of these layers increases the resolution of the images for both the discriminator and the generator. The generation of high qualified images is a big challenge. The larger the image, the easier it is for the network to make mistakes, because it should learn to create more complex and sensible details. In the ProGAN network, first small-scale layers are studied, then the model focuses on cleaning large-scale structures (Karras, Aila, Laine, & Lehtinen, 2017).

G. **StyleGAN** is used as a style-based generative model to solve the distortion problem of specific features of ProGAN during image generation. StyleGAN reconstruct the architecture of the generator network, allows properly to manage image synthesis by making small-scale changes to styles that compromise quality. StyleGAN is a GAN able to generate very high-resolution images (for instance, size 1024 * 1024). The idea is to build a stack of layers. The initial layers generate low-dimensional images (starting at 2 * 2) and the subsequent layers gradually increase in size (Karras, Laine, & Aila, 2019).

Various architectures of GAN have been offered for video (MocoGAN, Pose-GAN, VGAN), description (CycleGAN, DiscoGAN, PAN, Pix2pix), music generation (C-RNN-GAN, ORGAN, SeqGAN), text-to-image (StackGAN, TAC-GAN), and so on (Chen, Duan, Houthooft, Schulman, Sutskever, & Abbeel, 2016, December, Wang, Z., She, & Ward, 2021, Hitawala, 2018).

Numerous different generative models are available, and in fact, GAN is not the first generative model. The main alternatives to GAN are variation autoencoder, autoregressive models, flow models and hybrid models. Each of them has similarities and differences, advantages and disadvantages compared to GAN (Nowozin, Cseke, & Tomioka, 2016, December).

3. Application of GAN training algorithms

GAN model is a very effective generative model for creating real patterns after learning some data. These advantages provide the application of GANs in various fields of Computer Vision (CV) and Artificial Intelligence (AI). Here, we discuss various GAN applications in different fields such as image, sound, and vide.

3.1. Classification and regression

In recent years, training with convolutional neural networks (CNN) has been widely used in computer vision applications, and less attention has been paid to unsupervised learning with CNN. In (Radford, Metz, & Chintala, 2015), CNN class called DCGAN with certain architectural limitations is presented and considered to be a strong candidate for unsupervised learning. The hierarchy of images from small parts of the object to large scenes is taught in both the generator and the discriminator of a pair of deeply convolutional adversarial networks trained on a different descriptive dataset. In addition, learned attributes are used for new challenges.

3.2. Segmentation of medical images

In (Xue, Xu, Zhang, Long, & Huang, 2018), an innovative adversarial neural network called SegAN for segmentation of medical images derived from classical GAN is proposed. Since dense, pixel-level classification is required for the segmentation of an image, a single real or false output, the classic GRS discriminator may be ineffective in providing stable and sufficient feedback to networks. Alternatively, a fully convolutional neural network is used as a generator to generate segmentation class attribute dependencies, and a new competitor network with a multi-scale loss function L1 is proposed to force the discriminator and generator to learn both global and local features. In the SegAN environment, generator and discriminator networks are trained alternately in the min-max game: The discriminator takes a pair of images as input (original image- predicted class mark dependence, original image - basic real class mark dependence) and then trains by maximally multiplying the multivariate loss function; The segmenter is prepared only with gradients passed by the discriminator to minimize the multidimensional function. Such a SegAN environment is more efficient and sustainable for segmentation, which provides better performance than the most modern U-net segmentation method.

3.3. Music generation

Existing contemporary approaches to motion recognition are based more on Recurrent Neural Networks (RNN). In (Mao, Li, Xie, Lau, Wang, & Paul Smolley, 2017), CNN-based environment for classifying and determining movement is proposed. The validation set of NTU RGB + D database with a simple 7-layer network performs 89.3% accuracy. A network is developed to produce classified temporal segment patterns within the same network to identify motion in uncut videos. Consequently, the 93.7% map is obtained exceeding the base level in the PKU-MMD database by a large margin.

3.4. Super Resolution

(Ledig, Theis, Huszár, Caballero, Cunningham, Acosta, ... & Shi, 2017) proposes a generative adversarial network SRGAN for Super-Resolution (SR). To achieve this, a perception loss function consisting of a mutual loss and a loss of content is proposed.

Generative adversarial networks are strong generative models, but they have instability problem in training. The proposed Wasserstein GAN moves towards a stable training of GANs, but it can still generate only weak patterns. The proposed method in (Gulrajani, Ahmed, Arjovsky, Dumoulin, & Courville, 2017) outweighs the standard WGAN and enables numerous GAN architectures to be trained sequentially without any hyperparameters; including the 101-layer ResNets model.

3.5. Transformation of text into image

Synthesis of high-quality images from text images is a difficult problem for computer vision and requires many applications. Patterns created by existing text-image approaches may represent the meaning of the given descriptions approximately, but do not contain the necessary details and parts of a living object. (Zhang, Xu, Li, Zhang, Wang, Huang, & Metaxas, 2018) proposes GAN (StackGAN) collected to create real images of 256x256 pixels based on text-image. Stage-I GAN draws the primitive shape and colors of an object based on a given text image, and Stage-I produces low-quality images. Stage-I can eliminate defects in the results and add different details. The Conditioning Augmentation method is presented to develop the diversity of synthesized images and to stabilize conditional GAN training.

3.6. Video generation

The video signals can be groupped into two parts according to their content and action. While the content indicates the objects existing in the video, the action determines their dynamics. Based on this, GAN (MoCoGAN) environment consisting of motion and content parts is proposed for video generation (Tulyakov, Liu, Yang, & Kautz, 2018). The proposed environment generates a new video by adapting random vectors to the video environment. Each random vector consists of a content and a motion part. While the content remains constant, the motion part performs as a stochastic process. Experimental results on several data sets by qualitative and quantitative comparisons with the latest approaches evaluate the effectiveness of the proposed environment. Moreover, we observe that MoCoGAN can generate videos with the same content but different motion, as well as videos with different content and the same motion.

3.7. Transformation of face patterns

Synthetic aging of the human face was performed using the Age-cGAN method. Contrary to the GANs in the transformation of facial features, the features of a person's original aged face are preserved here. So, a new approach is proposed to optimize the identification of hidden vectors of GAN (Identity-Preserving). Objective evaluation of old and young faces by the most modern facial recognition and age assessment methods demonstrates the high potential of the proposed method (Antipov, Baccouche, & Dugelay, 2017, September, Tran, Yin, & Liu, 2017, Fedus, Rosca, Lakshminarayanan, Dai, Mohamed, & Goodfellow, 2017).

4. Conclusion

Generative adversarial networks have emerged in recent years and are growing rapidly. GANs are a type of generative model based on game theory. GANs have been practically successful in generating real data, especially images. As a strong class of generative models, GANs cannot accurately predict the distribution of data patterns, but they train to generate new patterns that match the same distribution as real data. Correspondingly, these networks have been applied to effective fields such as classification of various images, prediction, facial recognition. In this paper, we briefly analyzed and compared various important types of generative adversarial networks. We studied GAN for classification and regression, segmentation of medical images, music generation, best viewing ability, text conversion into the image, and video generation. The importance was that GANs tend to produce higher-quality images over long periods of training.

Reference

Alguliyev, R. M., Abdullayeva, F. J., & Ojagverdiyeva, S. S. (2020). Protecting children on the internet using deep generative adversarial networks. International Journal of Computational Systems Engineering, 6(2), 84-90. https://ieeexplore.ieee.org/document/771073

Antipov, G., Baccouche, M., & Dugelay, J. L. (2017, September). Face aging with conditional generative adversarial networks. In 2017 IEEE international conference on image processing (ICIP) (pp. 2089-2093). IEEE.

Arjovsky, M., & Bottou, L. (2017). Towards principled methods for training generative adversarial networks. https://openreview.net/pdf?id=Hk4_qw5xe Accessed 11 December 2021.

Arjovsky, M., Chintala, S., & Bottou, L. (2017, July). Wasserstein generative adversarial networks. In International conference on machine learning (pp. 214-223). PMLR.

Bang, D., & Shim, H. (2018, July). Improved training of generative adversarial networks using representative features. In International conference on machine learning (pp. 433-442). PMLR.

Che, T., Li, Y., Jacob, A. P., Bengio, Y., & Li, W. (2016). Mode regularized generative adversarial networks. https://arxiv.org/abs/1612.02136 Accessed 11 December 2021.

Chen, X., Duan, Y., Houthooft, R., Schulman, J., Sutskever, I., & Abbeel, P. (2016, December). Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In Proceedings of the 30th International Conference on Neural Information Processing Systems (pp. 2180-2188).

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: an overview. IEEE Signal Process Mag 35 (1): 53–65. https://arxiv.org/abs/1710.07035 Accessed 11 December 2021.

Donahue, J., Krähenbühl, P., & Darrell, T. (2016). Adversarial feature learning. https://arxiv.org/abs/1605.09782 Accessed 11 December 2021.

Dumoulin, V., Belghazi, I., Poole, B., Mastropietro, O., Lamb, A., Arjovsky, M., & Courville, A. (2016). Adversarially learned inference. https://arxiv.org/abs/1606.00704 Accessed 11 December 2021.

Fedus, W., Rosca, M., Lakshminarayanan, B., Dai, A. M., Mohamed, S., & Goodfellow, I. (2017). Many paths to equilibrium: GANs do not need to decrease a divergence at every step. https://arxiv.org/abs/1710.08446 Accessed 11 December 2021.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139144. https://doi.org/10.1145/3422622

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of wasserstein gans. https://arxiv.org/abs/1704.00028 Accessed 11 December 2021.

Hitawala, S. (2018). Comparative study on generative adversarial networks. https://arxiv.org/abs/1801.04271 Accessed 11 December 2021.

Xue, Y., Xu, T., Zhang, H., Long, L. R., & Huang, X. (2018). Segan: Adversarial network with multi-scale l 1 loss for medical image segmentation. Neuroinformatics, 16(3), 383-392. https://link.springer.com/article/10.1007%2Fs12021-018-9377-x

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of gans for improved quality, stability, and variation. https://arxiv.org/abs/1710.10196 Accessed 11 December 2021.

Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4401-4410).

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4681-4690).

Lucic, M., Kurach, K., Michalski, M., Gelly, S., & Bousquet, O. (2017). Are gans created equal? a large-scale study. https://arxiv.org/abs/1711.10337v4 Accessed 11 December 2021.

Mao, X., Li, Q., Xie, H., Lau, R. Y., Wang, Z., & Paul Smolley, S. (2017). Least squares generative adversarial networks. In Proceedings of the IEEE international conference on computer vision (pp. 2794-2802).

Metz, L., Poole, B., Pfau, D., & Sohl-Dickstein, J. (2016). Unrolled generative adversarial networks. https://arxiv.org/abs/1611.02163 Accessed 11 December 2021.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. https://arxiv.org/abs/1411.1784 Accessed 11 December 2021.

Nowozin, S., Cseke, B., & Tomioka, R. (2016, December). f-gan: Training generative neural samplers using variational divergence minimization. In Proceedings -of the 30th International Conference on Neural Information Processing Systems (pp. 271-279).

Odena, A., Olah, C., & Shlens, J. (2017, July). Conditional image synthesis with auxiliary classifier gans. In International conference on machine learning (pp. 2642-2651). PMLR.

Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. https://arxiv.org/pdf/1511.06434.pdf%5D Accessed 11 December 2021.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training gans. Advances in neural information processing systems, 29, 2234-2242. https://proceedings.mlr.press/v80/bang18a/bang18a.pdf

Theis, L., Oord, A. V. D., & Bethge, M. (2015). A note on the evaluation of generative models. https://arxiv.org/pdf/1511.01844.pdf Accessed 11 December 2021.

Tran, L., Yin, X., & Liu, X. (2017). Disentangled representation learning gan for pose-invariant face recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1415-1424).

Tulyakov, S., Liu, M. Y., Yang, X., & Kautz, J. (2018). Mocogan: Decomposing motion and content for video generation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1526-1535).

Wang, Z., She, Q., & Ward, T. E. (2021). Generative adversarial networks in computer vision: A survey and taxonomy. ACM Computing Surveys (CSUR), 54(2), 1-38. https://arxiv.org/pdf/1906.01529.pdf

Warde-Farley, D., & Bengio, Y. (2016). Improving generative adversarial networks with denoising feature matching. https://openreview.net/forum?id=S1X7nhsxl

Yang, Q., Yan, P., Zhang, Y., Yu, H., Shi, Y., Mou, X., ... & Wang, G. (2018). Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. IEEE transactions on medical imaging, 37(6), 1348-1357. https://arxiv.org/pdf/1708.00961.pdf

Yeh, R. A., Chen, C., Yian Lim, T., Schwing, A. G., Hasegawa-Johnson, M., & Do, M. N. (2017). Semantic image inpainting with deep generative models. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5485-5493).

Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., & Metaxas, D. N. (2018). Stackgan++: Realistic image synthesis with stacked generative adversarial networks. IEEE transactions on pattern analysis and machine intelligence, 41(8), 1947-1962. https://arxiv.org/abs/1710.10916