

# Selected Applications of Generative Adversarial Networks: Mini Review

Gokce Iymen, Gizem Tanriver, and Onur Ergen\*

Graduate School of Sciences and Engineering, Turkey

ISSN: 2832-4463



\*Corresponding author: Onur Ergen,  
Graduate School of Sciences and  
Engineering, Turkey

Submission: 📅 July 08, 2020

Published: 📅 August 06, 2020

Volume 1 - Issue 2

**How to cite this article:** Gokce Iymen,  
Gizem Tanriver, Onur Ergen. Selected  
Applications of Generative Adversarial  
Networks: Mini Review. COJ Rob Artificial  
Intel. 1(2).COJRA. 000506. 2020.  
DOI: [10.31031/COJRA.2020.01.000506](https://doi.org/10.31031/COJRA.2020.01.000506)

**Copyright@** Onur Ergen, This article is  
distributed under the terms of the Creative  
Commons Attribution 4.0 International  
License, which permits unrestricted use  
and redistribution provided that the  
original author and source are credited.

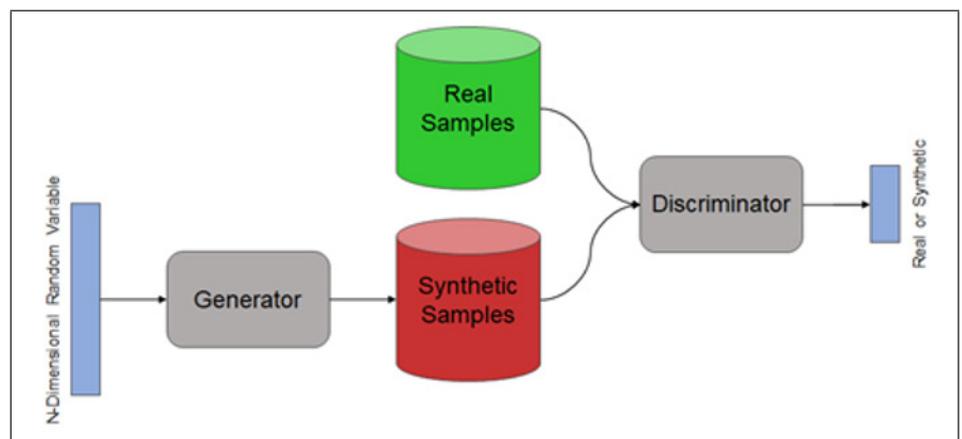
## Abstract

Generative adversarial networks have become increasingly popular since they were first introduced in 2014. Many variants of GANs have been developed over the years and employed in a range of applications from computer vision to audio generation and medical imaging. As its applications in computer vision have been widely explored by the artificial intelligence community, here, we focus on more specific applications of GANs, namely audio generation and medical image synthesis. In the age of big data, these two fields still struggle with the scarcity of labelled data, hence they benefit greatly from the capabilities of GANs.

**Keywords:** Audio Generation; Generative Adversarial Networks; Generative Models; Medical Image Synthesis

## Introduction

Generative models producing synthetic but real-like data are one of the most exciting research topics in the field of artificial intelligence. Generative adversarial networks (GANs) which were introduced in 2014 by Goodfellow [1], are a type of generative models using adversarial training for two neural network models, namely generator and discriminator [1]. The main difference of GANs from other generative models is its simplicity. Figure 1 illustrates the structure of a typical GAN, where the generator is trained so that the discriminator cannot distinguish synthetic data from real ones.



**Figure 1:** Structure of GAN consisting of a generator and a discriminator [1].

While generating new instances, models aim to generate a random variable from the probability distribution of a pre-existing dataset. The task can be extremely challenging since the parameters and even the existence of this probability distribution is not fully known. Furthermore, the probability distribution for high-dimensional data is generally very complex over a high-dimensional space; therefore, neural networks are commonly used to learn and mimic the unknown probability distribution from which the original data are sampled [1-3].

During training, GANs do not try to explicitly make approximations on the parametric features of a probability distribution, which requires complex computations as in [4-6].

Instead, they attempt to produce data samples from the probability distribution on target while forcing these samples to be as similar as possible to the ones from the original probability distribution.

Currently, many different versions of GANs such as Conditional GANs [7], CycleGAN [8], DCGANs [9], DiscoGAN [10], LSGAN [11], and MelGAN [12] can be found in the literature, each of which is proposed for different application areas. In this review, we will focus on selected applications of GANs which have attracted great attention in recent years, namely audio generation, and medical imaging.

## Selected Applications of GANs

### Audio generation

Although GANs are best known for their use in image generation, they have also been used successfully for generating sequential data as in [13-15]. Sequential data such as audio, natural language, and time-series can be generated by GANs with high performance in terms of both speed of generation and goodness of output, compared to the use of other generative counterparts. Audio generation can be applicable in specific domains such as speech synthesis and music generation. For speech synthesis, concatenative and parametric approaches were previously used before the advent of generative models. Generative models using autoregressive models such as WaveNet [16], Fast WaveNet [17] and SampleRNN [18] have since been developed, yet they work extremely slow due to their sample-level nature (i.e. they produce one sample at a time). On the other hand, GAN-based models [19-21] for speech synthesis have been shown to work much more efficiently. Owing to their rapid sampling characteristics, GANs hold great potential for data augmentation in speech recognition models. Moreover, the methods used in speech synthesis can be generalized to any form of audio. Music generation is another application area in which GANs are used for producing new music as in [13,20].

Different representations of sound can be more desirable in some applications of audio generation. In [20], WaveGAN and SpecGAN, which use raw waveform of audio and spectrogram respectively, are presented. When comparing these two representations, they obtained promising results in both approaches. However, in [13], it was shown that use of spectrograms instead of waveforms yields more coherent output. Although spectrograms are inherently non-invertible, which may make them disadvantageous in certain conditions, it is possible to approximate them back to their waveforms. Since human perception is sensitive to coherence in speech or music, it is important to successfully convert generated spectrograms into waveforms in order not to lose fidelity.

### Medical image synthesis

Deep learning algorithms are routinely used in medical imaging tasks such as classification and segmentation, whose performance relies heavily on availability of large amounts of labelled data [22]. Nonetheless, the medical field hugely suffers from the scarcity of labelled data more than any other, primarily due to the labor-intensive annotation process for medical images [22]. According

to Hou [23], manual segmentation of nucleus in a small dataset of 50 tissue image patches (each 600 × 600 pixel) takes about 225 hours of a pathologist's time [23]. Considering the data-hungry characteristic of deep learning algorithms, annotating sufficient amounts of medical data requires unrealistic time and effort of medical experts. Data augmentation techniques are often utilized to increase the size of a training set; yet they generate augmented images that are too similar to the original ones, providing a very limited performance improvement. Other challenges also exist in medical imaging such as high-class imbalance (i.e. underrepresentation of diagnostically less common conditions in a dataset) and continuous spectrum of features (i.e. classes are not inherently distinct due to progressive nature of diseases) [24]. Various GAN architectures have been proposed to address these challenges in a range of medical applications [24-29], which provided promising results in generating realistic looking but synthetic medical images while improving model performance. One study comparing traditional data augmentation with synthetic data augmentation utilizing DCGAN for a liver lesion classification task demonstrated that the use of synthetic samples significantly improves classification performance even on a small dataset consisting of computed tomography images of 182 liver lesions [29]. While unconditional GAN architectures such as DCGAN address the instability problem of GANs, typically, they do not work well at relatively low resolutions [22]. Baur [30] exploited progressive growing of GANs (PGGAN) to synthesize skin lesion images at high resolution, which produced highly realistic synthetic images that expert dermatologists had difficulty distinguishing them from real images [30]. As the data scarcity remains a major obstacle for medical imaging, GANs are likely to become a standard practice to fill this gap.

## Conclusion

GANs are used in a plethora of applications and their success has excited the deep learning community greatly. Although the trustability of generated data and the lack of established evaluation metrics for GAN-based methods remain as major limitations to their wider adoption, GANs have proven to be powerful even in specific domains such as audio generation and medical image synthesis. We hope that this mini review gives readers a sense of how GANs open up new possibilities in these two domains.

## References

1. Goodfellow J, Abadie JP, Mirza M, Xu B, Farley DW (2014) Generative adversarial nets. NIPS'14: Proceedings of the 27<sup>th</sup> International Conference on Neural Information Processing Systems 2: 2672-2680.
2. Santos CND, Mroueh Y, Padhi I, Dognin PL (2019) Learning implicit generative models by matching perceptual features. Computer Vision and Pattern Recognition.
3. Li Y, Swersky K, Zemel RS (2015) Generative moment matching networks.
4. Salakhutdinov R, Hinton G (2009) "Deep boltzmann machines," in proceedings of the twelfth international conference on artificial intelligence and statistics. Proceedings of Machine Learning Research, Florida, USA, 5: 448-455.
5. Kingma DP, Welling M (2013) Auto-encoding variational bayes.

6. Rezende DJ, Mohamed S, Wierstra D (2014) Stochastic backpropagation and approximate inference in deep generative models. Proceedings of the 31<sup>st</sup> International Conference on Machine Learning, Beijing, China.
7. Mirza M, Osindero S (2014) Conditional generative adversarial nets.
8. Zhu J, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp. 260-267.
9. Metz RL, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. 9<sup>th</sup> Conference on Artificial Intelligence and Robotics and 2<sup>nd</sup> Asia-Pacific International Symposium, Kish Island, Iran.
10. Kim T, Cha M, Kim H, Lee JK, Kim J (2017) Learning to discover cross-domain relations with generative adversarial networks. Proceedings of Machine Learning Research.
11. Mao X, Li Q, Xie H, Lau RYK, Wang Z (2016) Multi-class generative adversarial networks with the L2 loss function. Computer Vision and Pattern Recognition.
12. Kumar K, Kumar R, Boissiere TD, Geste L, Teoh WZ, et al. (2019) MelGAN: Generative adversarial networks for conditional waveform synthesis.
13. Engel JH, Agrawal KK, Chen S, Gulrajani I, Donahue C, et al. (2019) GANsynth: Adversarial neural audio synthesis.
14. Yu L, Zhang W, Wang J, Yu Y (2016) SeqGAN: Sequence generative adversarial nets with policy gradient. Proceedings of the 31<sup>st</sup> AAAI Conference on Artificial Intelligence, California, USA.
15. Li D, Chen D, Shi L, Jin B, Goh J, et al. (2019) MAD-GAN: multivariate anomaly detection for time series data with generative adversarial networks. International Conference on Artificial Neural Networks.
16. Oord V, Dieleman S, Zen H, Simonyan K, Vinyals O, et al. (2016) Wavenet: A generative model for raw audio.
17. Paine TL, Khorrami P, Chang S, Zhang Y, Ramachandran P, et al. (2016)
18. Mehri S, Kumar K, Gulrajani I, Kumar R, Jain S, et al. (2016) Samplernn: An unconditional end-to-end neural audio generation model.
19. Kaneko T, Kameoka H, Hojo N, Ijima Y, Hiramatsu K, et al. (2017) Generative adversarial network-based postfilter for statistical parametric speech synthesis. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Louisiana, USA, pp. 4910-4914.
20. Donahue C, McAuley JJ, Puckette MS (2018) Adversarial audio synthesis.
21. Binkowski M, Donahue J, Dieleman S, Clark A, Elsen E, et al. (2019) High fidelity speech synthesis with adversarial networks.
22. Kazemina S, Baur C, Kuijper A, Ginneken BV, Navab N, et al. (2018) GANs for medical image analysis.
23. Hou L, Agarwal A, Samarasinghe D, Kurc TM, Gupta RR, et al. (2017) Unsupervised histopathology image synthesis.
24. Wei J, Suriawinata A, Liu X, Ren B, Moin MN, et al. (2020) Difficulty translation in histopathology images.
25. Wei J, Suriawinata A, Vaickus L, Ren B, Liu X, et al. (2019) Generative image translation for data augmentation in colorectal histopathology images.
26. Guibas JT, Virdi TS, Li PS (2017) Synthetic medical images from dual generative adversarial networks.
27. Costa P, Galdran A, Meyer MI, Abramoff MD, Niemeijer M, et al. (2017) Towards adversarial retinal image synthesis.
28. Ghorbani A, Natarajan V, Coz D, Liu Y (2019) DermGAN: Synthetic generation of clinical skin images with pathology.
29. Adar MF, Diamant I, Klang E, Amitai M, Goldberger J, et al. (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification.
30. Baur C, Albarqouni S, Navab N (2018) Generating highly realistic images of skin lesions with gans. OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis, pp. 260-267.

For possible submissions Click below:

Submit Article