International Journal of Engineering &Technology, *7 (3.12) (2018) 864 -870*

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Understanding Trending Variants of Generative Adversarial Networks

Tanvi Bhandarkar[1], A. Murugan[2]

Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur
**Corresponding Author Email: [1]tanvi_bhandarkar@srmuniv.edu.in, [2]murugan.a@ktr.srmuniv.ac.in*

Abstract

Generative Adversarial Networks (GAN) have its major contribution to the field of Artificial Intelligence. It is becoming so powerful by paving its way in numerous applications of intelligent systems. This is primarily due to its astute prospect of learning and solving complex and high-dimensional problems from the latent space. With the growing demands of GANs, it is necessary to seek its potential and impact in implementations. In short span of time, it has witnessed several variants and extensions in image translation, domain-adaptation and other academic fields. This paper provides an understanding of such imperative GANs mutants and surveys the existing adversarial models which are prominent in their applied field.

Index Terms: *Generative Adversarial Networks (GANs), generative models, adversarial learning.*

1. Introduction

Generative Adversarial Networks, proposed by Ian Good fellow et al. [1], have become popular these days due to its capacity to learn and generate complex and high dimensional data. This has made it vulnerable to tremendous accomplishments in short span of time. Be-ing comprehensible in nature, it has contributed towards semi-supervised learning[46],[52],[53], image synthesis and translation [47],[48],[56], object detection and transfiguration [49],[51], super-resolution[10], generating 3D images and models [50],[55] and many other academic fields. This unsupervised learning approach is based on a minimax game. There are two main components of GAN - Generator Network, (G) and Discriminator Network, (D) which compete with each other. The generator network learns to generate a sample of data and the discriminator network takes the real data as input and discriminates with the output of generator network, i.e. it tries to evaluate if the data is real or generated. This can be referred from Figure 1. Without going into further theoretical and mathematical details, which can be referred from [1], [59],[53],[54], and [57], this paper directly focuses on the popular variants of GANs. Though today there are several extensions of GANs available, this paper aims to seek the potential of these adversarial networks by familiarizing with selective variants who have outperformed others. The paper starts by presenting how the GANs have its flair in the field of Image enhancement, modeling, and processing. We extend the section by summarizing the impact of GANs in image quality, text-to-image modeling, and image-to-image translation. After that, we adduce the applications of GANs in other academic fields such as Medicine, Music, Video generation, and other miscellaneous areas.

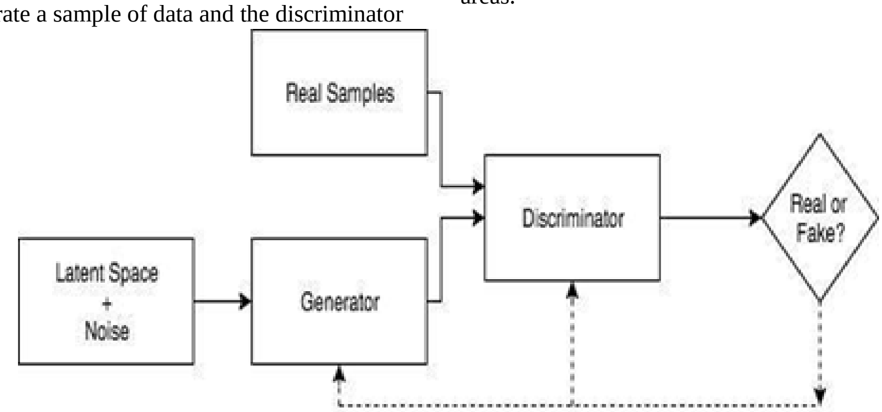

Fig. 1: Schematic diagram of GAN. The generator is fed with noise and latent space to produce sample images similar to the real dataset. These sampled and real images are sent to a discriminator which by learning through iterations distinguishes it as fake or real

2. Applications of GAN

GANs has the dominant architecture due to its ability to map from unlabeled the data and generate sample structures. By providing random input vector of visual signals or text or any form of data, it can produce astounding results.

A. Image Enhancement, Modeling and Processing

The GANs have varied applications in generating 3D images, enhancing resolution and quality of an image as well as image blending [2] and inpainting [3], [4] and many more. After the release of [1], the subsequent year followed the maximum of its implementation in this field. GANs are like the boon to the artificial intelligence in many ways but started to surge in generating high-quality sample data.

DCGANs. Radford et al. [5] proposed Deep Convolutional Generative Adversarial Networks which are highly effective when it comes to stability and usability of learning features from images in unsupervised learning. Through their constrained architecture, image stability is achieved in the training process: (1) Substituting pooling layers with Discriminator (strided-convolutions) and Generator (fractional-strided convolutions); (2) Introducing batch norm layer [6] in both generator and discriminator models; (3) removing fully connected hidden layers; (4) In Generator, all layers have ReLU activation function and the output layer uses Tanh function; (5) In Discriminator, all layers have LeakyReLU activation function. The generator in the DCGAN model can perform vector arithmetic operations to manipulate qualities and properties of generated samples. This can be witnessed from Figure 2. This is the basic constrained architecture of DCGANs which was then trained on three popular datasets: LSUN [7], Imagenet-1000, Faces dataset. The model is ought to give better representations of images for supervised learning and generative modeling as well. Though the framework bears a limitation of completely attaining stability when they are trained for a longer time. It finds its future implementation in video frame prediction and speech synthesis.

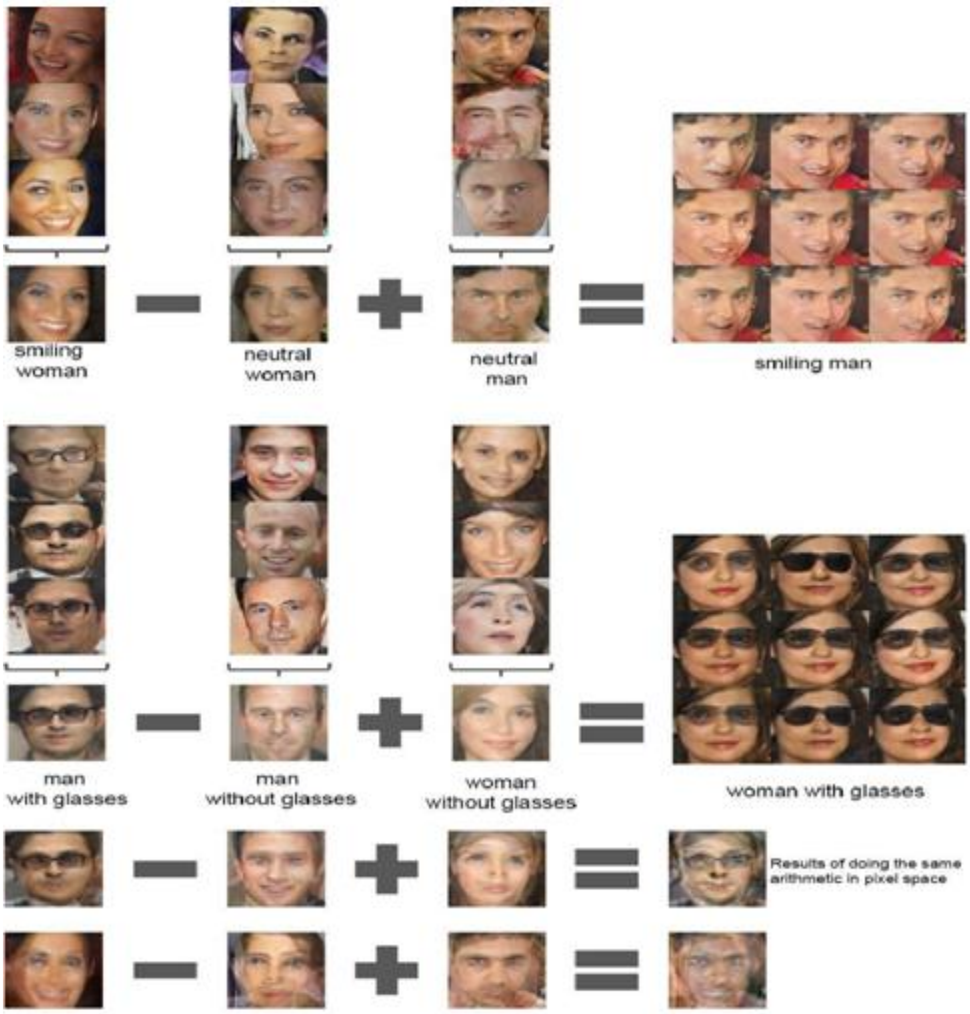

Fig. 2: Vector arithmetic of facial images using DCGAN. Photo by [5]

cGANs. Mirza et al. [9] proposed a conditional version of the GAN [1] called Conditional Generative Adversarial Network. In this extended version of GAN, both the adversarial models, generator and discriminator, are conditioned with some auxiliary information, y, as an additional input. This y could be any class labels or data. Experiments were performed on the MNIST digit dataset which was conditioned on class labels and the MIR Flickr 25,000 dataset [58] for multi-modal learning.

1) Game of Quality and Resolution: **SRGAN**.Ledig et al.

[10] presented a Super-Resolution using Generative Adversarial Network, Figure 3. This is the first framework which can upscale the down sampled image into photo-realistic image of 4x resolution using its perceptual loss function. This function consists of: (1) adversarial loss, which makes the solution to look more of the original kind by training through discriminator. This network is trained to differentiate between the original and resolved images;

(2) content loss, is the Euclidean distance between the feature maps of the resolved image and the original image. These feature representations are extracted from ReLU activation layers of pre-trained VGG19 network [11], [12]. Experiments performed on low-resolution images were tested by mean-opinion-score (MOS). The obtained MOS scores on many public datasets where significantly closer to the natural image of high resolution than those obtained by any state-of-the-art approaches.

LAPGAN.Denton et al. [13] introduced a combination of conditional GANs with Laplacian pyramid [14] framework to generate high quality images. Each layer of the Laplacian pyramid consists of separate convolutional-adversarial network model which takes the random input vector (only at first stage) and generates the residual image, conditioned on the image from the previous layer. This image now feeds this output as input to the next layer. Experiments performed on three datasets showed the following results: (1) CIFAR10 - 32 x 32 pixel; (2) STL - 96 x 96 pixel; (3) LSUN - 64 x 64 pixels.

BEGAN. Berthelot et al. [15] introduced a new equilibrium method called Boundary Equilibrium Generative Adversarial Network which cogently generates the images of faces at resolution 128 X 128. In this method, the discriminator D functions as an auto encoder, which is inspired by Energy-based Generative Adversarial Network(EBGAN)[16], and re-constructs the real images. The weights of the D are updated to minimize the reconstruction loss to generate the real image. This is the main goal of the method, to produce the coherent image at considerable resolution by balancing the loses. This stabilizes the training of generator and discriminator to get the approximate measure of convergence. This also adjusts the trade-off between image diversity and realism. It sees the potential application in dynamically weighing regularization or other heterogeneous objectives.

2) Text to Image Synthesis: The trend of generating plausible images from the text or captions using generative adversarial nets began from 2016. Reed et al. [17] proposed the deep convolutional adversarial model which successfully generated compelling visual results of birds and flowers from the human-written detailed text descriptions. The datasets used for the purpose were Caltech-UCSD birds database [19] and Oxford-102 Flowers database [20] along with five text descriptions per image. This model was also then tested on the MS COCO dataset [18]. Both the discriminator and generator network works as a feed-forward network, learning to map from the character and words to pixel-level with the conditions on textual descriptions, instead of a class label. The remarkable results about the model are in generating sharp samples, diversity in samples and delivering detailed visual outputs. Most of the generated scenes are not coherent and sometimes lack to handle complex multi-object scenes. These are few of its limitations but it incorporates to scale up in the game of resolution and more variety of text types in its future work.

PPGNs.Nguyen et al. [21] proposed a variant of GAN [1] called Plug & Play Generative nets which can produce the high-resolution images of 227 X 227 pixels for 1000 categories of the ImageNet dataset. This is achieved by improving both sample quality and diversity of the image at the training stage. Model is composed of: (1) a generator network G, which can draw a wide range of different categories of images and (2) condition C, which is conditioned on class and caption to direct generator to draw sample image. This text-to-image generative model can describe the image with words or captions by attaching the recurrent image-captioning network to the output of the generator. Thus generating the high resolution images by iterative sampling. It holds future application in producing images for videos and creating various artworks with multiple conditioned networks at same time.

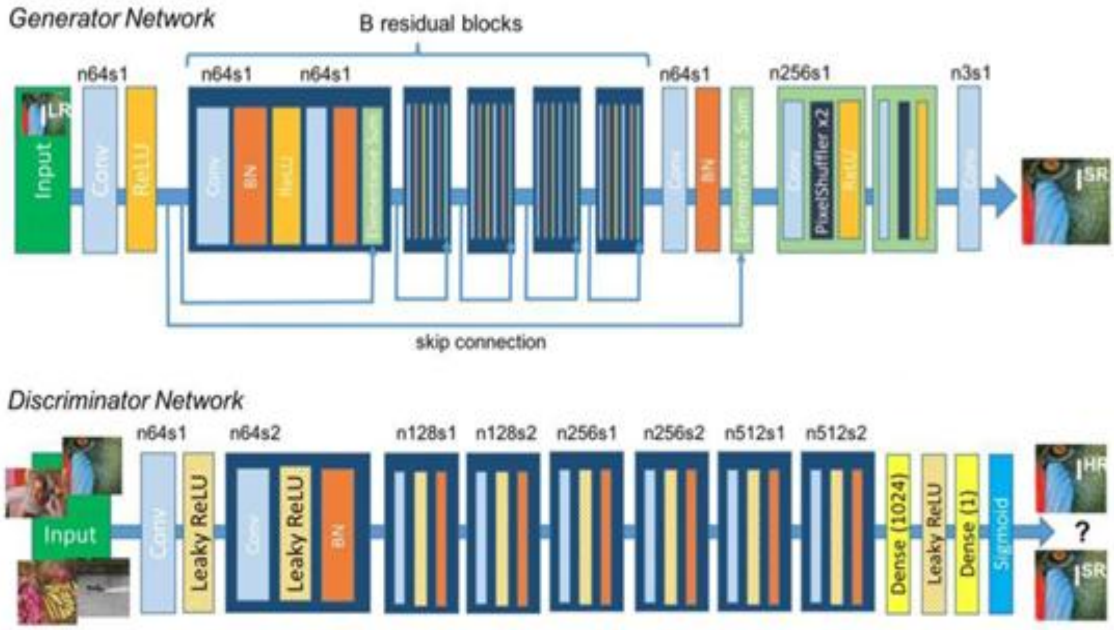

Fig. 3: Architecture of Generator and Discriminator Network with corresponding number of feature maps(n) & stride(s) in SRGAN. Photo by [10]

Stack GAN. Zhang et al. [22] introduced two-stage architecture of GAN [1] called Stacked GAN which produces the photo-realistic images of 256 X 256 resolution conditioned on only text descriptions as shown in Figure 4. This is highly stable and effective method compared to other state-of-the-art methods which could generate at most of 128 X 128 resolution. Thus it holds huge applications in photo editing, computer designing, etc. This two stages generates compelling results and achieves 28.7% and 20.30% improvements of inception scores on CUB[19] and Oxford-102[20] datasets, respectively. The Stage-I produces the basic contour of the image at low resolution with light background and nominal colours. Then this resulted image is feed as input to Stage-II along with text descriptions, which corrects the defects of the image obtained from Stage-I and adds more realistic details in the photo. Failure of the image is when the Stage-I messes up the image.

Fig. 4: Photo-realistic images generated from Text to image conversion by StackGAN performed on CUB[19] and Oxford-102[20] dataset. Photo by StackGAN [22]

3) Image to Image Translation: pix2pix.Isola et al. [23] proposed a conditional generative model which learns a loss function to map from the input image to output image. When learning from the dataset, the two standard loss functions Least absolute deviations (L1) and Least square errors (L2) decides which function should be minimized. StarGAN uses simple mask vector method to employ partially labeled datasets to perform image-to-image translation using a single model among multiple domains. This approach holds application in augmenting photos from label maps, adding colors, rebuilding the objects from edge maps, etc.

Cycle GAN.Zhu et al. [24] presented a method that uses the training set of aligned pairs and learns through the mapping of function G : X ! Y where the source domain X is translated to output image domain Y using adversarial loss, which is indistinguishable from G(X). Then the cycle consistency loss is introduced by coupling it with inverse mapping F : Y ! X, as it is highly under-constrained in nature. It finds its numerous applications in photo generation and enhancement, style transfer, object transfiguration and other such tasks.

StarGAN. Choi et al. [25] introduced the approach to perform image-to-image translation. This single model network takes input mask vectors and produces scalable facial attribute transfer and a facial expression synthesis tasks, regardless of the number of domains. The task involves changing of the facial attributes such as from normal input to smiling, fearful, angry, and similar other when experimented on CelebA and RaFD datasets. StarGAN stands in need of just single pair of generator and discriminator compared to other proposed methods [24],[26],[27].

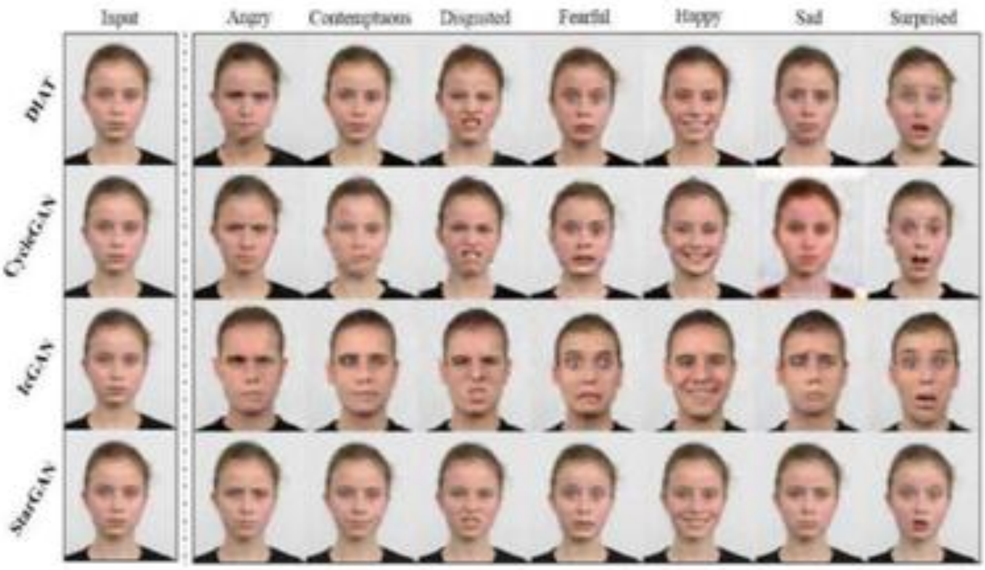

Fig. 5: Synthesis of Facial expressions comparison of different variants on RaFD dataset. Star GAN outperforms others. Photo by Star GAN[25]

B. Video Prediction and Translation

VGAN. Vondrick et al. [28] presented a model that leverages unlabelled data to learn the scene dynamics through the generative adversarial concept [1]. Background is considered stationary using 2D CNN and the separated foreground uses 3D CNN to predict the future frames. The network learns whether the object is moving or not and thus incorporates the frames of tiny videos. This two-streamed generative model holds potential applications in video representation learning, forecasting and simulating.

MoCo GAN. Tulyakov et al. [29] proposed a generative model which decomposes the visual signals into content and motion. These are then used as input to the framework which generates the sequence of video frames. The latent space is divided significantly by the content part, which is built with the Gaussian distribution and the motion part, which is built with the recurrent neural network(RNN). This has impacted many applications and has achieved edge over VGAN [28] and TGAN [30] on several challenging datasets.

C. Medicine

Nie et al. [31] applied the adversarial strategy to train the 3D fully convolutional network to generate Computed Tomography (CT) from the given magnetic resonance imaging (MRI) image. MRI being safer than CT, it is technically difficult to deduce the CT image. So this model maps the non-linear relationship between the CT and MRI images by using the loss function and auto-context algorithm [32] to train each stage using GAN framework and make it context-aware. This method promises to show the applications in the medical image analysis, denoising and super-resolution.

Se GAN. Xue et al. [33] proposed a segmentor-critic architecture of GAN for segmenting the medical images. This is an end-to-end adversarial model implies the segmentor to predict the segmented image and the critic to maximize the features differences between this segmented image and the actual one. This trains the segmentor to learn the features of actual segmentation adversarially. This model has outperformed other image segmentation algorithms such as DI2IN [34] and SCAN [35].

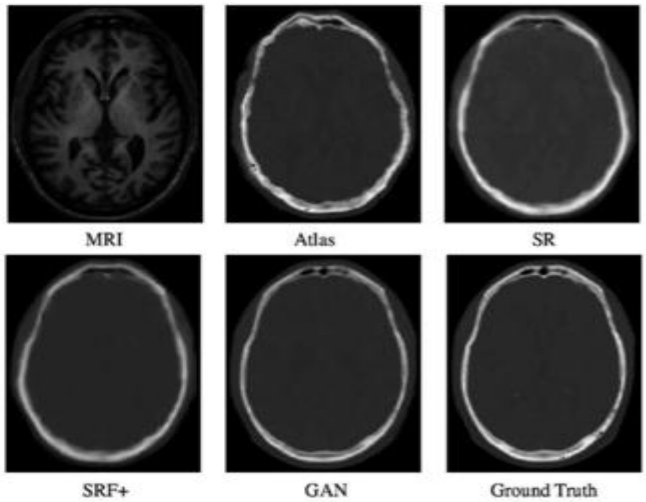

Fig. 6: Comparison of CT image obtained from MRI of various models. It can be seen that image produced by GAN is closed to ground truth. Photo by [31]

D. Music

C-RNN-GAN. Mogren et al. [36] used the adversarial concept to produce the sequential data using the RNN network called long-short term memory (LSTM) [37]. By training the collection of classical music it learns patterns from original dataset and generates various tones, the span of different intensities of played tones and so on. It generates the higher polyphony score by producing more than one tone per LSTM cell.

MidiNet.Yang et al. [38] presented a CNN-GAN based model to generate the multiple MIDI tracks by exploiting the known information from the scratch. It then produces musical note by the sequencing the chords or conditioning the melody bars. The Figure 7 explains the architecture of Generator and Discriminator models of MidiNet where the conditions and noise are modeled as per the functions of respective models. They showed that CNN based approach is more realistic than Google's MelodyRNN models [39]. It holds future work in generating multi-track music by training on the larger MIDI dataset.

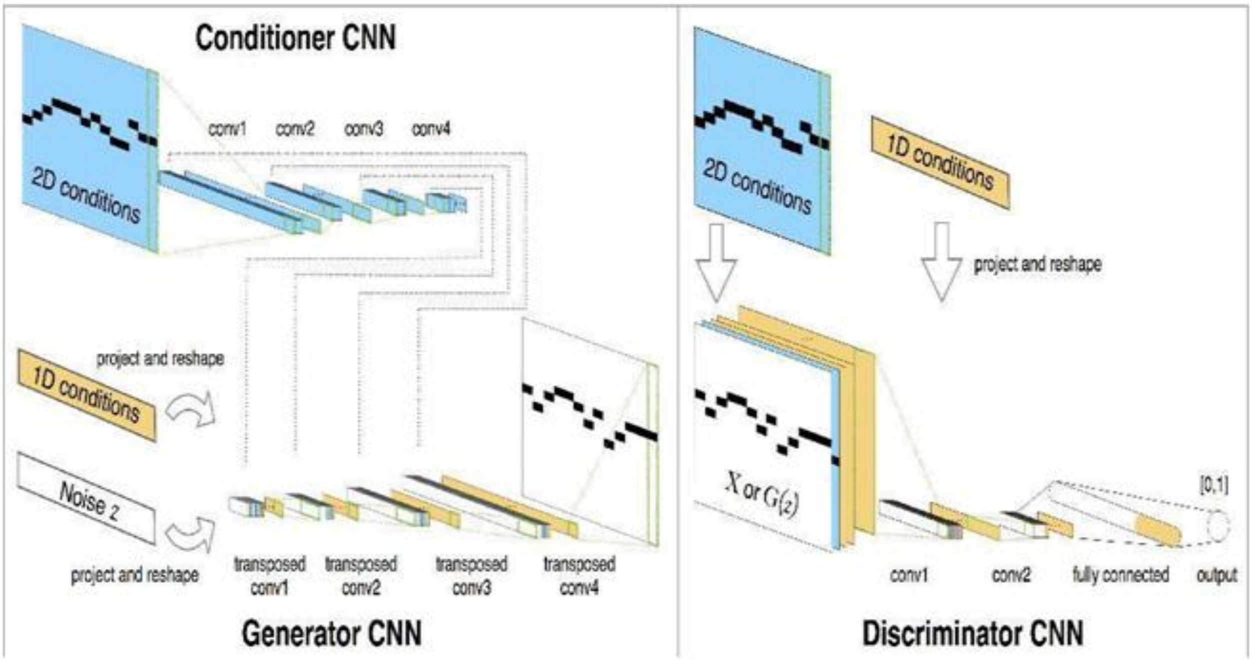

Fig. 7: Architecture of MidiNet to produce symbolic-domain music. Photo by [38]

E. Miscellaneous

GANs are popularly used in Physics lately. The proposed models such as CaloGAN [40] and LAGAN [41] represent patterns of energy distributions of particles by generating its images.

Adversarial models have shown many applications in language and speech analysis, generation and voice conversion by synthesis of conditional data - SEGAN[42], RankGAN[43], VAW-GAN [44], and so on.

Attempts are also made in incorporating GAN in steganography, SSGAN[45] by assigning one generator, S and two discriminators, D, and S. G generates the sample images, D labels the sampled images as real or fake and S, steganalyser checks for the secret message in the image.

Wu et al. [55] proposed 3D-GAN method to generate 3D models from the images. Advances in volumetric convolutional networks have led to the creation of such frameworks which can seamlessly generate 3D models and show impressive results on 3D object recognition. Being inspired from Radford et al. [5], this framework can also be extended to 3D-VAE-GAN which can produce 3D object from one or two dimensional image.

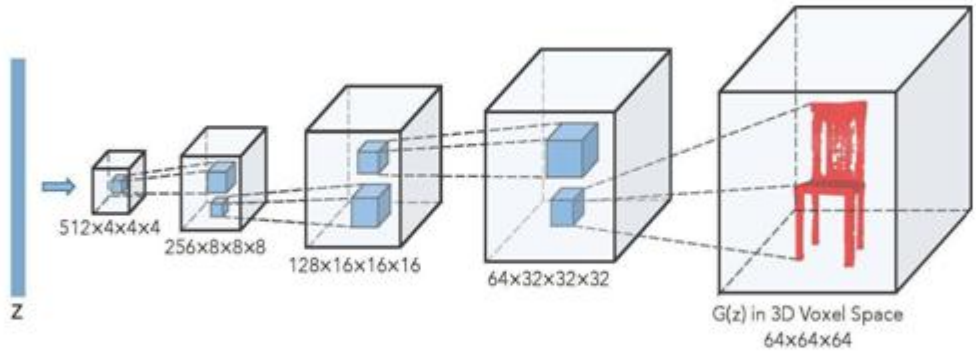

Fig. 8: The generator in 3D-GAN. From the random input vector z the generator captures the structure of the object and synthesis it to generate high quality 3D object. The discriminator learns features of these 3D models in unsupervised manner. Photo by [55]

3. Conclusion

We discussed how the generative adversarial networks have strong impact on the field of artificial intelligence and also it is not limited to it. It is applied to various practical and creative fields, besides just machine learning, such as medicine, music, art, video prediction, language and speech recognition, object transfiguration and many more. The reason for its endless accomplishments is due to its way of solving high-dimensional and complex data from the latent space. The two of its major components, generator G and discriminator D competes with each other to generate the data samples. D learns the features from the original dataset by distinguishing the real and fake data. We also saw that it effectively works on unlabeled and missing data as well.

4. Acknowledgment

The authors would like to thank Anurag Maravi for his advice, support, and feedback for this paperwork.

References

[1] Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial nets, in Advances in Neural Information Processing Systems 27, Montreal, Quebec, Canada, 2014, pp. 26722680.

[2] H. K. Wu, S. Zheng, J. G. Zhang, and K. Q. Huang, GP-GAN: Towards Realistic High-Resolution Image Blending, arXiv preprint arXiv: 1703.07195, 2017.

[3] R. A. Yeh, C. Chen, T. Y. Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do, Semantic image inpainting with deep generative models, in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR),Honolulu, HI, USA, 2017

[4] W. Y. Wang, Q. G. Huang, S. Y. You, C. Yang, and U. Neumann, Shape inpainting using 3D generative adversarial network and recurrent convolutional networks, in The IEEE Int. Conf. Computer Vision (ICCV), Venice, Italy, 2017, pp. 22982306.

[5] Radford, L. Metz, and S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks , in Int. Conf. Learning Representations (ICLR), San Juan, Puerto Rico, 2016.

[6] Sergey Ioffe, Christian Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in Proceedings of the 32nd International Conference on Machine Learning, PMLR 37:448-456, 2015.

[7] Fisher Yu, Ari Seff, Yinda Zhang, Shuran Song, Thomas Funkhouser, Jianxiong Xiao, LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop, in arXiv:1506.03365 [cs.CV], 2016.

[8] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, Generative Adversarial Text to Image Synthesis, arXiv preprint arXiv:1605.05396, 2016.

[9] Mirza, M. and Osindero, S., Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784, 2014.

[10] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, arXiv preprint arXiv:1609.04802, 2016.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, Going deeper with convolutions. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 19, 2015.

[12] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (ICLR), 2015.

[13] Denton, Emily, Chintala, Soumith, Szlam, Arthur, and Fergus, Rob, Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks, arXiv preprint arXiv:1506.05751, 2015.

[14] P. J. Burt, Edward, and E. H. Adelson, The laplacian pyramid as a compact image code. IEEE Transactions on Communications, 31:532540, 1983.

[15] D. Berthelot, T. Schumm, and L. Metz, Began: Boundary equilibrium generative adversarial networks, arXiv preprint arXiv:1703.10717, 2017.

[16] J. Zhao, M. Mathieu, and Y. LeCun. Energy-based generative adversarial network. In 5th International Conference on Learning Representations (ICLR), 2017.

[17] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, Generative adversarial text to image synthesis, arXiv preprint arXiv:1605.05396, 2016.

[18] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, Microsoft COCO: Common objects in context. In ECCV. 2014.

[19] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, The Caltech-UCSD Birds-200-2011 Dataset. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011.

[20] M.-E. Nilsback and A. Zisserman, Automated flower classification over a large number of classes. In Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing, Dec 2008.

[21] Nguyen, A., Yosinski, J., Bengio, Y., Dosovitskiy, A.,and Clune, J. (2016), Plug & play generative networks: Conditional iterative generation of images in latent space. arXiv preprint arXiv:1612.00005.

[22] H. Zhang, T. Xu, H. Li, S. Zhang, X. Huang, X. Wang, and D. Metaxas, Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. arXiv preprint arXiv:1612.03242, 2016.

[23] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, Image-to-image translation with conditional adversarial networks.In CVPR, 2017.

[24] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.

[25] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, StarGAN: Unified Generative Adversarial Networks for MultiDomain Image-to-Image Translation, ArXiv e-prints, Nov. 2017.
[26] M. Li, W. Zuo, and D. Zhang. Deep identity-aware transfer of facial attributes. arXiv preprint arXiv:1610.05586, 2016.
[27] G. Perarnau, J. van de Weijer, B. Raducanu, and J. M. A lvarez. Invert-ible conditional gans for image editing.arXiv preprint arXiv:1611.06355, 2016.
[28] C. Vondrick, H. Pirsiavash, and A. Torralba. Generating videos with scene dynamics. In Neural Information Processing Systems (NIPS). 2016.
[29] Tulyakov, S.; Liu, M.; Yang, X.; and Kautz, J. 2017. MoCoGAN: Decomposing motion and content for video generation. arXiv preprint arXiv:1707.04993.
[30] M. Saito, E. Matsumoto, and S. Saito. Temporal generative adversarial nets with singular value clipping. In IEEE International Conference on Computer Vision (ICCV), 2017.
[31] D. Nie, R. Trullo, C. Petitjean, S. Ruan, and D. Shen, Medical image synthesis with context-aware generative adversarial networks, arXiv preprint arXiv:1612.05362, 2016.
[32] Zhuowen Tu, Auto-context and Its Application to High-level Vision Tasks, Computer Vision and Pattern Recognition, CVPR 2008.
[33] Yuan Xue, Tao Xu, Han Zhang, Rodney Long, and Xiaolei Huang. Segan: Adversarial network with multi-scale L1 loss for medical image segmentation. arXiv preprint arXiv:1706.01805, 2017.
[34] Dong Yang, Tao Xiong, Daguang Xu, Qiangui Huang, David Liu, S Kevin Zhou, Zhoubing Xu, JinHyeong Park, Mingqing Chen, Trac D Tran, et al. Automatic vertebra labeling in large-scale 3d ct using deep image-to-image network with message passing and sparsity regularization, In International Conference on Information Processing in Medical Imaging, pages 633644. Springer, 2017.
[35] Wei Dai, Joseph Doyle, Xiaodan Liang, Hao Zhang, Nanqing Dong, Yuan Li, and Eric P Xing, Scan: Structure correcting adversarial network for chest x-rays organ segmentation, arXiv preprint arXiv:1703.08770, 2017.
[36] Olof Mogren, C-rnn-gan: Continuous recurrent neural networks with adversarial training, arXiv preprint arXiv:1611.09904, 2016.
[37] Sepp Hochreiter and Jrgen Schmidhuber. Long short-term memory. Neural computation, 9 (8):17351780, 1997.
[38] Li-Chia Yang, Szu-Yu Chou, and Yi-Hsuan Yang, MidiNet: A convolutional generative adversarial network for symbolic-domain music generation using 1d and 2d conditions, arXiv preprint arXiv:1703.10847, 2017.
[39] Elliot Waite, Douglas Eck, Adam Roberts, and Dan Abolafia. Project Magenta: Generating longterm structure in songs and stories, 2016. https://magenta.tensorflow.org/blog/2016/07/15/lookback-rnn-attentio n-rnn/.
[40] Michela Paganini, Luke de Oliveira, and Benjamin Nachman. Calogan: Simulating 3d high energy particle showers in multilayer electromagnetic calorimeters with generative adversarial networks. arXiv preprint arXiv:1705.02355, 2017.
[41] Luke de Oliveira, Michela Paganini, and Benjamin Nachman. Learning particle physics by example: Location-aware generative adversarial networks for physics synthesis. arXiv preprint arXiv:1701.05927, 2017.
[42] Santiago Pascual, Antonio Bonafonte, and Joan Serr. Segan: Speech enhancement genera- tive adversarial network. arXiv preprint arXiv:1703.09452, 2017.
[43] Kevin Lin, Dianqi Li, Xiaodong He, Zhengyou Zhang, and Ming-Ting Sun. Adversarial ranking for language generation. arXiv preprint arXiv:1705.11001, 2017
[44] Chin-Cheng Hsu, Hsin-Te Hwang, Yi-Chiao Wu, Yu Tsao, and Hsin-Min Wang. Voice conversion from unaligned corpora using variational autoencoding wasserstein generative adversarial networks.arXiv preprint arXiv:1704.00849, 2017.
[45] Haichao Shi, Jing Dong, Wei Wang, Yinlong Qian, and Xiaoyu Zhang, Ssgan: Secure steganography based on generative adversarial networks.arXiv preprint arXiv:1707.01613, 2017.
[46] Emily Denton, Sam Gross, and Rob Fergus. Semi-supervised learning with context- conditional generative adversarial networks. arXiv preprint arXiv:1611.06430, 2016.
[47] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou , and Alexei Efros. Image-to- image translation with conditional adversarial networks. arXiv preprint arXiv:1611.07004, 2016.
[48] Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jungkwon Lee, and Jiwon Kim. Learning to discover cross-domain relations with generative adversarial networks. arXiv preprint arXiv:1703.05192, 2017.
[49] Shuchang Zhou, Taihong Xiao, Yi Yang, Dieqiao Feng, Qinyao He, and Weiran He. Genegan: Learning object transfiguration and attribute subspace from unpaired data. arXiv preprint arXiv:1705.04932, 2017.
[50] Matheus Gadelha, Subhransu Maji, and Rui Wang, 3d shape induction from 2d views of multiple objects. arXiv preprint arXiv:1612.05872, 2016.
[51] X.L.Wang, A.Shrivastava, and A.Gupta, A-Fast-RCNN: hardpositive generation via adversary for object detection, arXiv: 1704.03414, 2017.
[52] Chongxuan Li, Kun Xu, Jun Zhu, and Bo Zhang, Triple generative adversarial nets.arXiv preprint arXiv:1703.02291, 2017.
[53] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen, Improved techniques for training gans. In Advances in Neural Information Processing Systems, pages 22342242, 2016.
[54] Martin Arjovsky and Leon Bottou, Towards principled methods for training generative adversarial networks. In International Conference on Learning Representations, 2017. Under review.
[55] Jiajun Wu, Chengkai Zhang, Tianfan Xue, Bill Freeman, and Josh Tenenbaum. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In Advances in Neural Information Processing Systems, pages 8290, 2016.
[56] Sagie Benaim and Lior Wolf, One-sided unsupervised domain mapping,arXiv preprint arXiv:1706.00826, 2017.
[57] White, T. Sampling Generative Networks, 2016; http://arxiv.org/abs/1609.04468.
[58] Huiskes,M.J.andLew,M.S.(2008).Themirflickrretrievalevaluation. In MIR 08: Proceedings of the 2008 ACM International Conference on Multimedia Information Retrieval, New York, NY, USA. ACM.
[59] Ian J. Goodfellow, Jonathon Shlens and Christian Szegedy, EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES, arXiv:1412.6572v3, 2015.