ISSN No:-2456-2165

A Survey on Generative Adversarial Networks (GANs)

Ashwini.R
B.E(CSE)
Rajalakshmi Engineering College
Chennai, India

Aro Immaculate.J
B.E(CSE)
Rajalakshmi Engineering College
Chennai, India

Abinaya.E
B.E(CSE)
Rajalakshmi Engineering College
Chennai, India

Abstract:- Generative adversarial networks are one of the recent research areas in deep learning. It is used in various applications in image/text/video generations etc. GANs are widely known for the adversarial process it follows and the two models in its architecture – the generator and the discriminator. Since it gives better results than other generative models it is preferred more. In this paper initially we discuss the basic architecture of GANs, the mathematical concept involved in it and how it outshines other generative models. The various applications of GANs to name a few, implementation of GANs in Radiology, beauty GANs, GANs deployed in cyber security and attention prediction have been investigated and mentioned with a brief description about the same. Further in the end we have discussed the challenges faced by it and its future scope in the fourth coming years.

Keywords:- *Generative Adversarial Network, Generator, Discriminator, Adversarial Process, Deep Learning.*

I. INTRODUCTION

Deep learning can be visualised as a subfield of machine learning based on artificial neural networks. Machine learning algorithms are generally linear whereas the deep learning algorithms are comprised as a net of complexity. It has various levels of complexity starting from single level perceptron and multi-layer perceptron to recurrent neural networks and convolutional neural networks. What makes a major difference in a neural network is that we don't need to tell the computer how to solve, instead, it will learn on its own from the data we provided, and deduce solution to the problem at hand. Today, many important problems in computer vision, autonomous vehicles, voice/speech recognition, and natural language processing are addressed by deep neural networks. Many large-scale companies such as Google, Microsoft, and Facebook deploy deep learning models to achieve better performance. Generative adversarial networks are one of those deep learning models which are relatively new.

Generative adversarial networks (GANs) are deep learning generative models introduced by Ian Goodfellow and other researchers in one of their papers in the year 2014. Though the idea has been in existence since 1990, the name GAN was coined out by Ian Goodfellow in 2014, and since then it has got familiarised. GANs have been in hype since the last few years for its versatile applications and results. GANs can essentially be called as a computer vision technique which follows adversarial process. GANs have been in use for applications like image/text/video generation, image editing, text to image generation, 3D object generation, security etc. GANs belong to the category of generative models and they have produced the most prominent results among others. The other models include real NVP, Variational Autoencoders (VAEs) and pixel CNN/pixel RNN.

II. POTENTIAL OF GAN

GANs' has a wide potential, because they can mimic any distribution of data. That is, GANs can be tutored to create worlds spookily similar to our own in any domain: images, music, speech, prose. They are in a sense can be considered as robot artists, and their output is poignant even. GANs have gained huge popularity even in the medical field but the lack of data lays it back.

III. ARCHITECTURE OF GANS

A. Overview of GAN Architecture

The Generative adversarial networks are basically comprised of two models- the discriminator and the generator. The generator produces new data and the discriminator checks for its authenticity. The two models are trained simultaneously, the generator is trained so that it will create new data similar to the real data and the discriminator is trained so that it will correctly classify the data whether its real or fake i.e. whether it belongs to the actual dataset or not. The discriminator can be assumed to be a standard convolutional network which can categorise the data fed to it and the generator can be assumed to be an inverse convolutional network.[1]

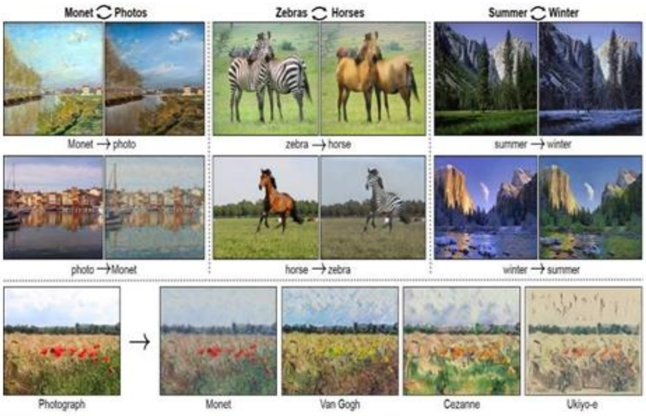

Fig 1

For example, take the case of an email, given there are two labels- spam and not spam, a discriminator will be able to classify an email as spam or not spam based on the words it contains. Here, the words are the features and the discriminator is mapping the features to its respective labels. When expressed mathematically, the problem is to find P(Y/X), where Y denotes the labels and X denotes the features, i.e. probability of Y given X which means probability that a mail is spam/not spam given the features. Now, coming to the generator, it does the opposite. The question a generative algorithm tries to answer is: Assuming this email is spam, how likely are these features? The problem is to find P(X/Y) i.e. the probability of X given Y. It gives the probability of the features given a label.

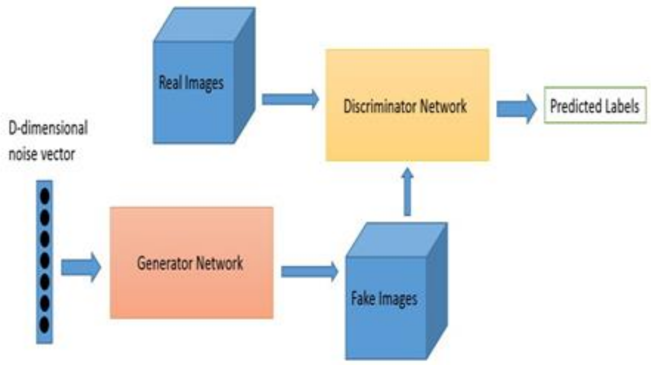

Fig 2

The generator network takes in random numbers as input and produces an image as output. This output image is a fake image and this is given as one of the inputs for the discriminator network along with the real images from the original dataset. Now the discriminator will classify whether the inputted images as real or fake. It will output a number between 0 and 1, with 1 representing a real image and 0 representing a fake image. This predicted output of the discriminator is given as feedback to the generator and discriminator is also given feedback based on the ground truth of the images which we know. This feedback help in the training of both the models and acts as its knowledge to grow.

B. *Mathematically Modelling A GAN*

A neural network *G(z, θ_1)* is used to model the Generator. It's role is to map the input noise variables **z** to the desired data space *x* (say images). Conversely, a second neural network *D(x, θ_2)* models the discriminator and outputs the probability that the data came from the real dataset, in the range (0,1). In both cases, θ_i represents the weights or parameters that define each neural network.

The discriminators weights are updated so as to maximize the probability that any real input image *x* is classified as belonging to the real dataset, while minimizing the probability that any fake image is classified as belonging to the real dataset. In more technical terms, the loss/error function used maximizes the function *D(x)*, and it also minimizes *D(G(z))*.

The Generator is trained to fool the Discriminator by generating data more similar to the real data, which means that the Generator's weight's are updated so as to maximize the probability that any fake image is classified as belonging to the real dataset. Formally this means that the loss/error function used for this network maximizes D(G(z)).

In practice, the logarithm of the probability is used in the loss functions instead of the raw probabilities. Since during training both the Discriminator and Generator are trying to optimize opposite loss functions, they can be thought of two agents playing a minimax game with value function V(G,D).

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

C. *GANs VS Other Generative Model*

As mentioned earlier, GANs are not the only type of generative models there are other algorithms as well. But GANs are superior to others and produce better results. The drawback with GANs are that they are difficult to work with and require a lot of data and tuning.

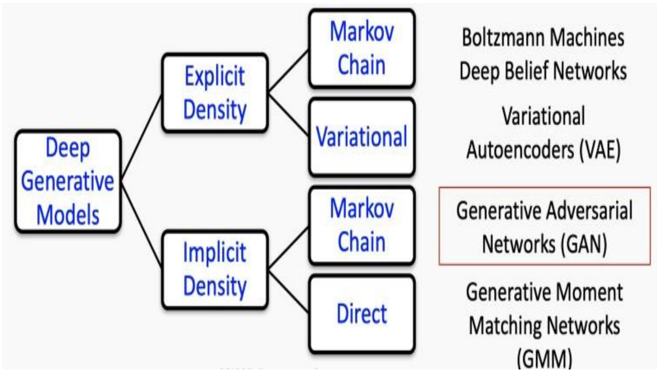

Fig 3

An autoencoder is pretty simple, it compresses its input down to a vector - with much fewer dimensions than its input data, and then transforms it back with the same shape as its input over several neural net layers. This compression means that only the most salient features of the image can stay—everything else is unnecessary. Autoencoders are more suitable for compressing data or generating semantic vectors from it whereas GANs are more suitable for generating data.

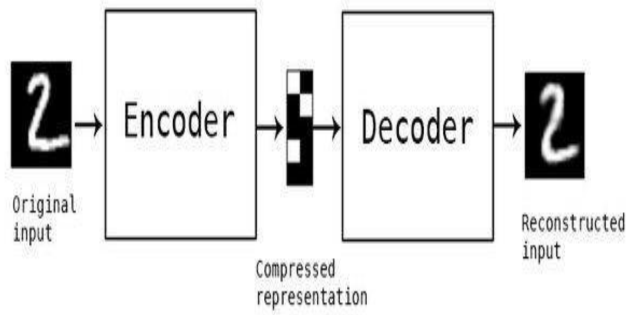

Fig 4

ISSN No:-2456-2165

The problem with variational autoencoders is that, the hidden representations are normalized. Variational autoencoders are also capable of both compressing and synthesizing data. However, GANs generate data in fine, granular detail whereas images generated by VAEs tend to be more blurred. hidden representations are normalized.

IV. APPLICATION OF GANS

A. *GANS IN RADIOLOGY*

Generative Adversarial Networks for Synthesizing Radiological Images of the Spine to be Used in In Silico Trials.

GAN can be used in silico trials (individualised computer simulation used in the development or regulatory evaluation of medical product, device, or intervention) to cut the costs related to development as well as the gradual time taken to reach the potential market. In these trials, virtual patients are recruited from a large database and their response to the treatment, such as the implantation of a medical device, is simulated by means of numerical models. In this work, we advocate the use of generative adversarial networks to produce synthetic radiological images to be used in silico trials. The generative models produced convincing synthetic sagittal X-rays of the lumbar spine based on a simple sketch, and were able to generate sagittal radiological images of the trunk using coronal projections as inputs, and vice versa. Although numerous fallacies in the anatomical details may still allow distinguishing synthetic and real images in the majority of cases, the present work showed that generative models are a feasible solution for fabricating synthetic imaging data to be used in in silico trials of novel medical devices.

The major drawbacks of this is:

1. The predominant and a relatively unexplored drawback is that the colossal databases of patient data, such as medical images, to be used in future in silico trials either do not exist at all or are at least not publicly available and ready for use.
2. Although early examples of in silico trials have been presented, the use of this form of investigation is not rampant yet, due to the major technical challenges involved. First, methods to simulate numerically the implantation of a medical device in the human body and to predict its outcome in an accurate and valid manner need to be available. The challenge has been widely confronted in the last decades and refined methods to create numerical models based on patient data, typically medical imaging such as X- rays, CTs and MRIs, have been developed and are currently widely employed.

B. *GAN IN IMAGE EDITING*

Most image editing software currently in the market don't provide much flexibility with creating new pictures with existing pictures.[2] With GAN its quite easy to alter images and make changes to existing images to create new ones. In the below example, the image of a female has been reconstructed drastically to represent a male.

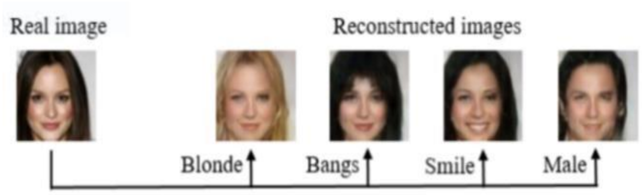

Fig 5

Another application of GAN in image editing is the image de-raining.

Fig 6

Another application of GAN in image editing is the image de-raining.

C. *USING GANs FOR CYBER SECURITY*

The upsurge and development of Artificial Intelligence has been wonderful for most industries. But there's a real perturb that has shadowed the entire AI revolution – cyber threats. Even deep neural networks are susceptible to being hacked. GANs are proving to be of immense help here, directly addressing the concern of "adversarial attacks".[5]

These adversarial attacks use a number of diverse techniques to deceive deep learning architectures. GANs are used to make existing deep learning models more robust to these techniques. How? By creating more such counterfeit examples and training the model to identify them. A technique called SSGAN is used to do steganalysis of images and detect deleterious encodings which shouldn't have been there.

D. *GENERATING DATA WITH GAN*

To build a deep learning model, huge amounts of data are required. The availability of data is a necessary criterion in several domains, especially where training data is needed to model training models.[4] The industry like healthcare comes to mind here. Using GANs, synthetic data for supervision could also be generated. For instance, the creation of synthetic data for training deep learning algorithms by creating realistic eye images with the help of GANs is given below.

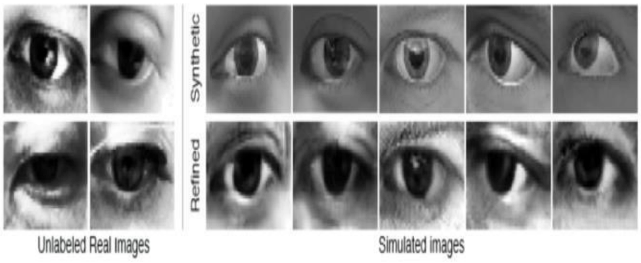

Fig 7

E. ATTENTION PREDICTION USING GANs

Attention is a cognitive process in which we focus on a particular aspect of information while ignoring other information. Generally, this is the way in which every human will look at an image and it is an important part of the human trait. In attention prediction we come to know where a person is looking at beforehand. Presently GANs are deployed for predicting the attention which can have many practical uses.

The GANs architecture used for attention prediction is known as SalGan (Saliency GAN). SalGAN consists of two networks: one predicts saliency maps from raw pixels of an input image; the other one takes the output of the first one to discriminate whether a saliency map is a predicted one or ground truth. SalGAN is expected to generate saliency maps that resembles the ground truth.

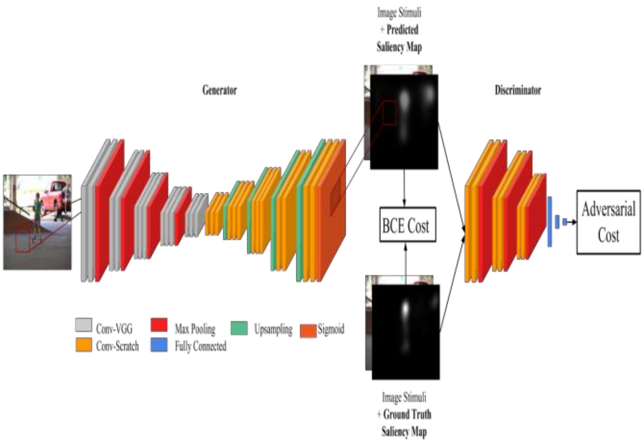

Fig 8:-Architecture of the saliency system

Generator: The generator network adopts a convolutional encoder-decoder architecture. Encoder part includes max pooling layers which decrease the size of the feature maps, while the decoder part uses up sampling layers followed by convolutional filters to construct an output that is the same resolution as the input.

Discriminator: The discriminator predicts whether the generated saliency map is correct or not.[3]

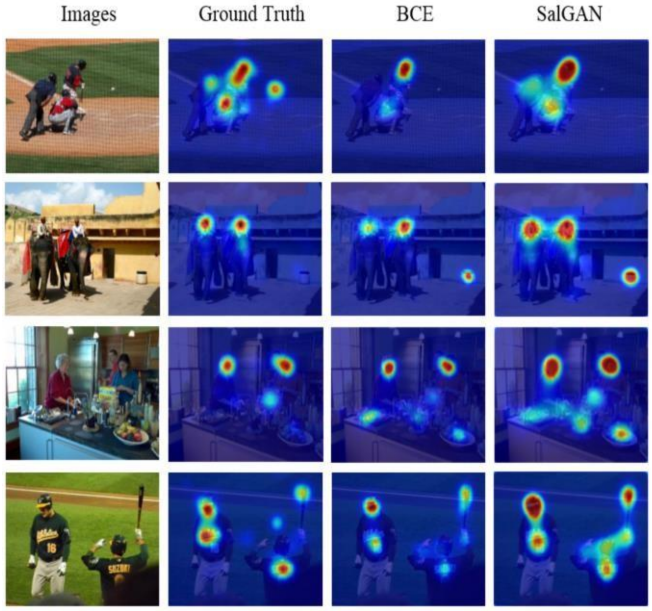

Fig 9

Qualitative results of SalGAN on the SALICON validation set, SalGAN predicts well those high salient regions which are missed by BCE model. Saliency maps of BCE model are very localized in a few salient regions, they tend to fail when the number of salient regions increases.

F. BEAUTY GAN

Makeup is the way in which someone or something is composed. The makeup you put on your face can be used to disguise your real self a bit. GANs can also be used for applying makeup on faces. When a face is given as an input GANS can be used to apply makeup while preserving the face identity.

Fig 10

V. CHALLENGES FACED BY GANS

GANs are by far the most prominent among the generative models and produce sharp results. But they are relatively more complex and computationally expensive. It needs a lot of images and high GPU capacity. The functions these networks are trying to optimize are loss functions and they do not have a closed form. Thus, optimizing this function is very hard and requires a lot of trial and error methods. One important factor to keep in mind while modelling GAN is that, the two models – the generator and discriminator will grow by outwitting each other and eventually learning by each other's output as feedback. But a major problem also comes from the same fact that one of the discriminator or generator can overpower the other. If the discriminator is superior, it will return values so close to 0 or 1 that the generator will struggle to read the gradient. If the generator is superior, it will persistently exploit weaknesses in the discriminator.

There are mainly three problems associated with GANs, MODE COLLAPSE: It is the situation in which the generator collapses by producing limited varieties of the samples or by producing the same data (complete collapse). Sometimes, GANs fail to model a multimodal probability distribution of data and suffer from mode collapse. VANISHING GRADIENT: the discriminator gets too successful that the generator gradient vanishes and learns nothing.

INTERNAL COVARIATE SHIFT: An internal covariate shift occurs when there is a change in the input distribution to our network. When the input distribution changes, hidden layers try to learn to adapt to the new distribution. This slows down the training process.

The problems faced by GANs can be overcome by optimizing the network design, cost function and through better implementation techniques.[6]

VI. CONCLUSION

The paper discuses on topics such as the potential of GAN , GAN architecture, GAN'S mathematical model, the various applications of GAN such as Radiology , image editing, cyber security , data generation, attention prediction , GAN in cosmetology. The survey also covers the challenges faced by GANs in various domains.

FUTURE SCOPE

GANs are relatively latest technologies in the of AI and are the next step in the deep learning evolution. No doubt that they are also growing day by day and evolving continuously. Research has been going on in various areas to improve the performance of GANs and produce much better results. Various optimization has been made in the network design in order to solve the problems that GANs face. Some of the areas where GANs are being introduced and are going to be a breakthrough in the near future are,

1. GANs in autonomous vehicles: GANs are being largely introduced in the concept of autonomous vehicles. The GAN architectures behind autonomous vehicles are called as Driving – GANs (DGANs).
2. GANs in Drug discovery: Researchers from Insilico Medicine proposed an approach of artificially intelligent drug discovery using GANs in which the generator when trained with existing drugs can produce new drugs for incurable diseases. Discriminator will check whether the new drug can actually cure the disease.
3. GANs in Fashion: The task is to model the input image given by a person with different outfits by keeping the wearer and his/her pose same. The concept of Stacking GANs are used where multiple GANs are used consecutively.
4. GANs in Medicine: MI-GANs are proposed for generating synthetic medical images and their segmented masks, which can then be used for the application of supervised analysis of medical images.

We have listed a few areas where GANs are being incorporated. There are various other areas and trends to these. There is no doubt that GANs are going to be life changing.

REFERENCES

[1]. Ian J. Goodfellow∗ , Jean Pouget-Abadie† , Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair‡ , Aaron Courville, Yoshua Bengio(2014). Generative Adversarial Nets
[2]. Guim Perarnau, Joost van de Weijer, Bogdan Raducanu, Jose M. Álvarez(2016). Invertible Conditional GANs for image editing
[3]. Junting Pana , Cristian Canton-Ferrerb , Kevin McGuinnessc , Noel E. O'Connorc , Jordi Torresd , Elisa Sayrola , Xavier Giro-i-Nieto(2018). SalGAN: visual saliency prediction with adversarial networks
[4]. Ashish Shrivastava, Tomas Pfister, Oncel Tuzel, Josh Susskind, Wenda Wang, Russ Webb Apple Inc(2017). Learning from Simulated and Unsupervised Images through Adversarial Training
[5]. Haichao Shi, Jing Dong, Wei Wang, Yinlong Qian, Xiaoyu Zhang(2017).SSGAN: Secure Steganography Based on Generative Adversarial Networks
[6]. Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen(2016). Improved Techniques for Training GANs.