Critical Analysis: Generative Adversarial Nets

Bhanu Pratap Singh (210760006)

Generative Adversarial Nets

Generative Adversarial Networks or GANs are a deep learning based generative model. More generally GANs are a model architecture involves two sub models: a generator model for generating new examples and a discriminator model for classifying whether generated examples are real, from the domain, or fake, generated by the generator model.

- *Generator:* Model that is used to generate new plausible examples from the problem domain.
- *Discriminator:* Model that is used to classify the examples as real (from the domain) or fake (generated).

A generative adversarial network is a class of machine learning frameworks designed by Ian Goodfellow and his colleagues in June 2014. Two neural networks contest with each other in a game. Given a training set, this technique learns to generate new data with the same statistics as the training set.

What are the main trends on the topic since the publications :

The prospect of GANs has led to explosive development in the computer vision areas, including image translation, image super-resolution, image synthesis and video generation, etc. The details of these applications are introduced as follows:

- *Image super resolution-* To improve the resolution of image, a Super-Resolution Generative Adversarial Networks (SRGAN) was proposed by Ledig et al. [44].
- *Face synthesis***-** BigGAN is proposed to better image synthesis. This is advance version SAGAN
- *Image translation:* To convert the image content from one domain to another, an image-to-image translation approach they introduced CGANs.
- It is also used in text image or image to image translation.
- In medicine an AnoGAN is used for anomaly detection of medical images, and learned the characteristics of lesions by learning the characteristics of health data sets

At present, GANs also has some achievements in the field of language and speech processing. The SeqGAN can outperform traditional methods in terms of speech, poetry and music generation and RankGAN is used to generate sentences.

What are the key ideas of related published works since the original publication :

Various derived GANs models have emerged since 2014, a large number of relevant research have been made as diverse GANs variant with higher and stable training performance, evaluation metrics with the aim of reflecting the quality of the model reasonably. These are categorised into two category architectures optimization based GANs and objective functions optimization based GANs.

- Architecture optimization based GANs

1. *Convolution based GANs*: In this paper have improved the network structure of generator and the discriminator by adopting the Multi-Layer Perceptron (MLP) for its working.

However, CNN performs better than MLP in extracting image features, this is reason why deep convolutional generative adversarial networks (DCGANs) is proposed.

2. *Condition based GANS*: In this both D and G are conditioned on some extra information which can be label, text or other data for affecting the data generation process, conditional GANs is proposed. In additional proposed InfoGANs which decomposed the input vector into two parts, but the latent information of latter method is unknown.
3. *Autoencoder based GANs*: There are many attempts to combine the idea of adversarial networks with the GANs

- Objective functions optimization based GANs

1. *Least Square GANs (LSGANs)*: In this paper opt for the least squares rather opting for the cross-entropy loss in the original work for both the discriminator and the generator. It is proposed to overcome the problem of vanishing gradient. This method generates the better images as compared to regular GANs.
2. *Energy-based GANs:* In this it gives high energy to the fake samples from the generative and lower energy to the real samples.

What are the main problem solved or improvements over the original work :

GANs faces the issue of the non-convergence. It was very hard to achieve the Nash equilibrium. Probably the most common form of harmful non-convergence encountered in the GAN game is mode collapse. During the training, the generator may collapse to a setting where it always produces same output, this is called Mode Collapse, this is arising in multi-modes in which we are expecting the generator model to learn from both modes but it is learning from the single mode only. By using the minibatch features in which it allows the discriminator to compared an example to minibatch of generated samples to minibatch of real samples. We solve the problem of mode collapse to such extent that we can also resolve the problems of counting, perspective (it fails to adapt 3D object while generating 2D representation in background) and global structures (cannot understand the holistic structure) by designing the better model architectures.

What are the remaining problems from the published works so far:

The main problem which is remain unsolved is that it is still hard to select an appropriate evaluation, which should distinguish generated samples from the real one, verify the mode collapse, mode drop and detect the overfitting. The instability in the training process is also a challenge to the researcher.

There is no function for evaluating the metric of the GANs.

What is the unsolved problem on the topic most interesting to you to solve and why :

I would like to solve the problem of finding the better evaluation metric because in term of improving the improving the GANs performance there has been a majority of solutions, but in term of the evaluation metric of GANs there is no well behaved and unified metric approved by most of researcher, though there have been some tries for finding the better evaluation metric. By which we could quantitatively evaluate the model and thus helps in the future research of GANs.

References:

[1] Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I. J., Bergeron, A., Bouchard, N., and Bengio, Y. (2012). Theano: new features and speed improvements. Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop.

[2] Bengio, Y. (2009). Learning deep architectures for AI. Now Publishers.

[3] Bengio, Y., Mesnil, G., Dauphin, Y., and Rifai, S. (2013a). Better mixing via deep representations. In ICML'13.

[4] Bengio, Y., Yao, L., Alain, G., and Vincent, P. (2013b). Generalized denoising auto-encoders as generative models. In NIPS26. Nips Foundation.

[5] Bengio, Y., Thibodeau-Laufer, E., and Yosinski, J. (2014a). Deep generative stochastic networks trainable by backprop. In ICML'14.

[6] Bengio, Y., Thibodeau-Laufer, E., Alain, G., and Yosinski, J. (2014b). Deep generative stochastic networks trainable by backprop. In Proceedings of the 30th International Conference on Machine Learning (ICML'14).

[7] Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Turian, J., Warde-Farley, D., and Bengio, Y. (2010). Theano: a CPU and GPU math expression compiler. In Proceedings of the Python for Scientific Computing Conference (SciPy). Oral Presentation.

[8] Breuleux, O., Bengio, Y., and Vincent, P. (2011). Quickly generating representative samples from an RBM-derived process. Neural Computation, 23(8), 2053–2073.