Report Writing

Introduction:- In this report, I will be discussing a paper by Ian J. Goodfellow about a deep learning algorithm called Generative Adversarial Network(GAN). Firstly, I will give an introduction to GAN and state pros and cons of GAN. Secondly, I will be discussing main problems resolved and improvements since publication of the original work by introduction of new techniques such as SeqGAN and adversarial perturbation. Also, the main trends on the topic since the publication of the paper such as emergence of conditional models and semi-supervised learning using GAN and Chekhov GAN. Thirdly, I will be talking about the problems in GAN that need better solutions and further research work that is to be conducted.

Generative Adversarial Network(GAN) *estimates generative models via an adversarial process, training two models simultaneously. A generative model G is trained that captures the data distribution. It can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection. However, discriminative model D is trained that distinguishes the probability that a sample came from the training data rather than G. The discriminative model is analogous to the police, trying to detect counterfeit currency. The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game until nash equilibrium is achieved. Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable from the genuine instances. In the space of arbitrary functions G and D, a unique solution exists, with G recovering the training data distribution and D equal to 1/2 everywhere. The generative model generates samples by passing random noise through a multilayer perceptron, and the discriminative model is also a multilayer perceptron.*

GAN comes with ***pros and cons*** *relative to previous modeling frameworks. The advantages are that Markov chains are never needed, only backprop is used to obtain gradients, no inference is needed during learning, and a wide variety of functions can be incorporated into the model.The disadvantages are primarily that there is no explicit representation of pg(x), and that D must be synchronized well with G during training. Secondly, the model parameters oscillate, destabilize and never converge leading to non-convergence.Thirdly,the problem of diminished gradients when the discriminator gets too successful that the generator gradient vanishes and learns nothing. Lastly,Unbalance between the generator and discriminator causing overfitting.*

Additionally, ***GAN has limitations*** *when the goal is for generating sequences of discrete tokens. A major reason is that the discrete outputs from the generative model make it difficult to pass the gradient update from the discriminative model to the generative model.* ***SeqGAN*** *[4] is an approach to solve the problem. Modeling the data generator as Stochastic Policy In reinforcement learning, SeqGAN bypasses the generator differentiation problem by directly performing gradient policy update. The RL reward signal comes from the GAN discriminator judged on a complete sequence, and is passed back to the intermediate state-action steps* *using Monte Carlo search. Extensive experiments on synthetic data and real-world tasks demonstrate significant improvements over strong baselines.*

Another approach built is ***adversarial perturbations [5] with GAN****. It is alternately training both classiifer and generator networks. The generator network generates an adversarial perturbation that can easily fool the classiifer network by using a gradient of each image. Simultaneously, the classiifer network is trained to classify both original and adversarial images correctly generated by the generator. These procedures help the classiifer network to become more robust to adversarial perturbations. Furthermore, adversarial training framework efficiently reduces overfitting and outperforms other regularization methods such as Dropout.*

Furthermore, generative adversarial nets can be extended to a ***conditional model*** *[6] if both the generator and discriminator are conditioned on some extra information y. Y could be any kind of auxiliary information, such as class labels or data from other modalities. We can perform the conditioning by feeding y into both the discriminator and generator as additional input layers.*

An extension of GAN is using a novel training method named ***Chekhov GAN*** *[8].The method provably converges to an equilibrium for semi-shallow GAN architectures, i.e. architectures where the discriminator is a one layer network and the generator is arbitrary. On the practical side, developing an efficient heuristic guided by theoretical results, are applied to commonly used deep GAN architectures. On several real world tasks, Chekhov GAN exhibits improved stability and performance compared to standard GAN training.*

Generative Adversarial Networks (GANs) is further extended to the ***semi-supervised learning*** *[9] by forcing the discriminator network to output class labels.A generative model G and a discriminator D is trained on a dataset with inputs belonging to one of N classes. At training time, D is made to predict which N+1 classes the input belongs to,where an extra class is added to correspond to the outputs of G. This method can be used to create a more data-efficient classifier and that allows for generating higher quality samples than a regular GAN.*

However, ***the problem that needs yet to be resolved is mode collapse*** *in which a generator collapses and produces limited varieties of samples. This is a problem for which* ***DRAGAN*** *[7] could be used, resulting in improvements in modeling performance across a variety of settings. Nevertheless, this is where more ideas could be explored in depth and more research could be conducted to improve the efficiency of experimenting different hyperparameters.*

References

1. *Salakhutdinov, R. and Hinton, G. E. (2009). Deep Boltzmann machines. In AISTATS'2009, pages 448– 455.*

2. *Hinton, G. E., Osindero, S., and Teh, Y. (2006). A fast learning algorithm for deep belief nets. Neural Computation,18, 1527–1554.*

3. *Goodfellow, I. J., Mirza, M., Courville, A., and Bengio, Y. (2013b). Multi-prediction deep Boltzmann machines. In NIPS'2013.*

4. *Lantao Yu, Weinan Zhang, Jun Wang, Yong Yu (2017)*
   SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient

5. *HyeungillLee, SungyeobHan, JungwooLee .GenerativeAdversarialTrainer: Defenseto AdversarialPerturbationswithGAN*

6. *Mehdi Mirza. ConditionalGenerativeAdversarialNets*

7. *NaveenKodali,JacobAbernethy,JamesHays&ZsoltKira. ON CONVERGENCE AND STABILITY OF GANs*

8. *Paulina Grnarova, Kfir Y. Levy, Aurelien Lucchi, Thomas Hofmann, Andreas Krause: An Online Learning Approach to Generative Adversarial Networks*

9. *AugustusOdena: Semi-SupervisedLearningwithGenerativeAdversarialNetworks*