Cara Evangeline Chandra Mohan
190539421

GENERATIVE ADVERSARIAL NETWORKS

GAN and it's need:

- In machine learning generative modelling is an unsupervised task (although in GAN architecture the training of generative modelling is framed as supervised learning problem) that involves automatically discovering and learning the patterns of input data so that the model can be used to generate new examples that could have come from the original dataset.
- GAN's are used in various fields where we can retrieve images for historical archives, generate images from a given text, discover the drug pattern, super-resolution of images and in various other fields.

Main Trends in Generative Adversarial Networks:

1. In **GAN[3]**, Generator is pitted against an adversarial network called Discriminator. Generator's objective is to generate data that is very similar to the training data. Discriminator's objective is to identify if the data coming from generator and original dataset is real or fake.

- Avoids Markov chains which has high computational cost
- Generator function has lesser number of restrictions compared to boltzmann machine[4]

2. A variant of GAN is **Conditional GAN[1]*** proposed by Mirza, where generator and discriminator is conditioned on some extra information y.

- In 2016, an unsupervised adaptation of the Conditional GAN[1] **(ie.Information Maximising GAN, *InfoGAN[2]*)** was proposed where the discriminator estimates both if an image is real or fake and the class label of the image.
- Miyato et al.[5] proposed a projection based approach to condition information in the discriminator**(*cGAN*)**. cGAN takes the inner product between the embedded condition vector and the feature vector instead of concatenating the conditioning label with the training data.

3. **Deep Convolutional GAN (DCGAN)[6]*** was proposed by Radford et al. where discriminator and generator are deep CNNs.

- An approach for image-to-image translation called ***pix2pix*** was proposed by Isola et al.[7]. The model is able to reconstruct objects from edge maps, add colour to black and white images. This model implicitly uses deep CNN and conditional GAN.

4. Zhang et al.[8] proposed a two-stage GAN **(StackGAN)*** that generates a photo-realistic images from text descriptions.

Stage 1 : Sketches the general outline of what is described in the text and adds background and image colours.

Stage 2 : Takes as input both Stage 1 output and text description, making it more photo realistic.

5. In 2018, **Self-Attention GAN (SAGAN)[9]*** was proposed to model long-range dependency to generate images where objects and scenarios are related in consistent with realistic images.

Key ideas of related published works:

- GAN's [1] main idea is to train the generative network by making true and generated distributions go through a downstream task such that the optimisation process of the generative network with respect to the downstream task will make sure that the generated distribution is close to the true distribution.
- [1], [2], [5] conditions the model on additional information and makes it possible to direct the data generation process. Such conditioning can be based on class labels or on data from different modality as well.

***Variants of GAN**
- Conditional GAN
- Deep Convolutional GAN
- Stack GAN
- Self- Attention GAN

- Deep convolutional layers are used as generators and discriminators in [6] and [7].
- StackGAN[8] is based on conditioning augumentation for two stage adversarial tasks (StackGAN-v1) and conditional & unconditional image for multi-stage generative adversarial tasks(StackGAN-v2).
- [9] applies self-attention to both generator and discriminator to model both local and global dependencies in an image.

Problems solved:

[3] GAN

- Doesn't rely on Markov chains to train or generate samples which is computationally expensive
- They help to solve tasks such as image generation from textual descriptions, getting high resolution images from low resolution image, predicting which drug can treat a particular disease.

[2] InfoGAN

- Enforces high mutual information between the condition and the generator distribution.
- InfoGAN successfully disentangles writing styles from digit shapes on the MNIST dataset and was able to discover meaningful hidden representations of a number of image datasets.

[5] cGAN

- Because of Mode collapse, the discriminator stops causing the generator outputs to be dissimilar, resulting in the generator only outputting very similar looking images with no diversity and this problem is overcome by cGAN.
- Helps in producing highly discriminative super-resolution images.

[7]pix2pix

- Formulating a loss function has always been a challenge, but this model network learns loss function specific to a problem and maps input image to the output image.
- Generates photo-realistic images from label maps.

[8]StackGAN

- Produces high super-resolution photo-realistic images from textual description.

[9]SAGAN

- CNNs cannot solve long range dependencies but this model works well for long range dependencies.

Unsolved Problems:

- Some intiatives have been taken to train the GAN but yet there is no stable method to train them, eg. there is no proper loss function defined
- There is no one method to evaluate the overall performance of GAN which makes it difficult for comparisons between different variants.
- To improve efficiency we should devise methods to coordinate G and D or determine better distribution to sample z during training[3].

Interesting unresolved Problem:

Modelling an evaluation metric to compare the performance seems to be an interesting unresolved problem since we need a measure to know which model performs better.

References:

1. Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." *arXiv preprint arXiv:1411.1784* (2014).
2. Chen, Xi, et al. "Infogan: Interpretable representation learning by information maximizing generative adversarial nets." *Advances in neural information processing systems*. 2016.
3. Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems*. 2014.
4. Salakhutdinov, Ruslan, and Geoffrey Hinton. "Deep boltzmann machines." *Artificial intelligence and statistics*. 2009.
5. Miyato, Takeru, and Masanori Koyama. "cGANs with projection discriminator." *arXiv preprint arXiv:1802.05637* (2018)
6. Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." *arXiv preprint arXiv:1511.06434* (2015).
7. Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
8. Zhang, Han, et al. "Stackgan++: Realistic image synthesis with stacked generative adversarial networks." *IEEE transactions on pattern analysis and machine intelligence* 41.8 (2018): 1947-1962.
9. Zhang, Han, et al. "Self-attention generative adversarial networks." *arXiv preprint arXiv:1805.08318* (2018).