**Review on: Generative Adversarial Text to Image Synthesis**

In this paper the authors have tried to come up with a GAN that will read from text and form an image accordingly. In this paper they have generated photos of birds and flowers only from detailed text description.

In this paper text is referred to a detailed description in a sentence form. GANs consist of a generator G and a discriminator D that competed in a two player minimax game such that the discriminator tries to generate real data and the generator tries to fool the discriminator (1). Minimax game has global optimum (1) (2). Both the discriminator D and generator G perform feed-forward inference conditioned on text feature by conditioned text features encoded by hybrid character-level CNN (1).

The method of work incorporated a network architecture, matching-aware discriminator (GAN-CLS), learning with manifold interpolation (GAN-INT), inverting the generator for style transfer.

Dataset and procedure:   
The dataset included CUB dataset for birds and Oxford-102 dataset of flower images with 11788 images of birds of 200 categories and 8189 images of flowers from 102 categories. These dataset were divided into test and training data. Using 5 captions per image.  
For text features 1024-dimentional GoogleNet image embedding were pre-trained in a convolutional recurrent text encoder (CNN-RNN) (3). Text embedding was used which generated images of the object with its attribute but not its style such as background etc.

Benchmark:   
The four methods have generated plausible flower images that match the description. The basic GAN was seen to have the most variety in flower morphology while other methods generated more class-consistent images. It was concluded that creating a flower was easier than a bird as birds have more structural regulations.

Conclusion:   
Besides birds and flowers, this model was able to generate photos of other things with detailed description too.

References:

1. Generative Adversarial Text to Image Synthesis, Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee, feaarXiv: 1605.05396v2 [cs.NE] 5 Jun 2016
2. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B, Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In NIPS, 2014.
3. Reed, S., Akata, Z., Lee, H., and Schiele, B. Learning deep representations for fine-grained visual descriptions. In CVPR, 2016.