NYU

Text to Image Synthesizer using Deep Convolutional GANs

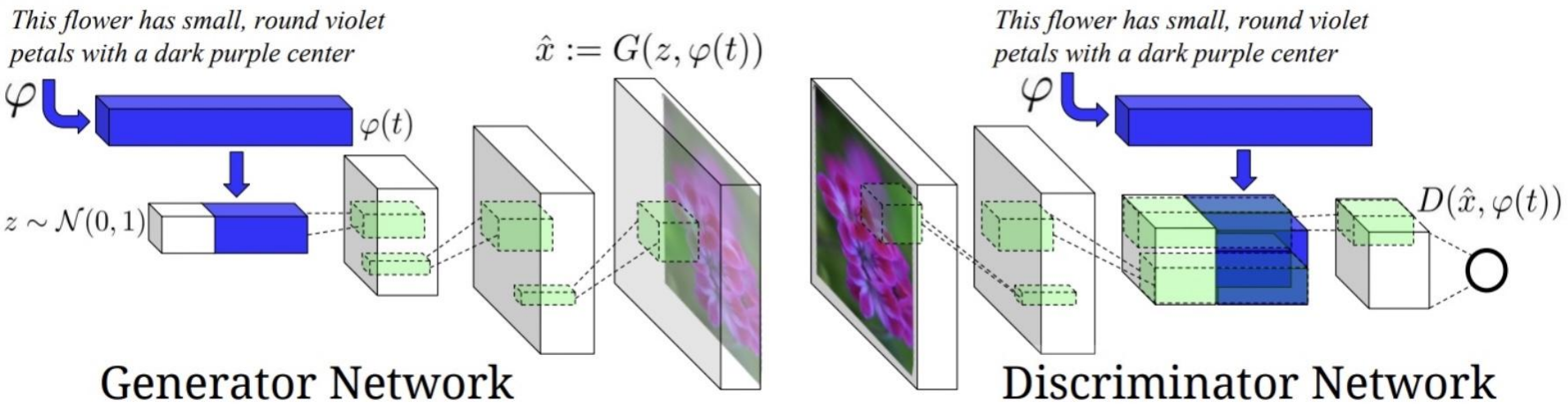
CV2271 Project Presentation

PRESENTED BY:
AYUSH JAIN - aj3152
HARSH DUBEY - hd2225
12/16/2021

Project Outline

- The objective of this project is to convert text to photo-realistic fake images which are indistinguishable from real images.
- We are utilising Deep Convolutional Generative Adversarial Networks (DCGANs) to create fake images using given textual descriptions.
- Why DCGANs? DCGANs are better than traditional machine learning algorithms as traditional models were unable to capture the necessary detail and randomness in the synthesized images.
- We trained our model for 200 epochs on Oxford-102 Flowers dataset.

Architecture



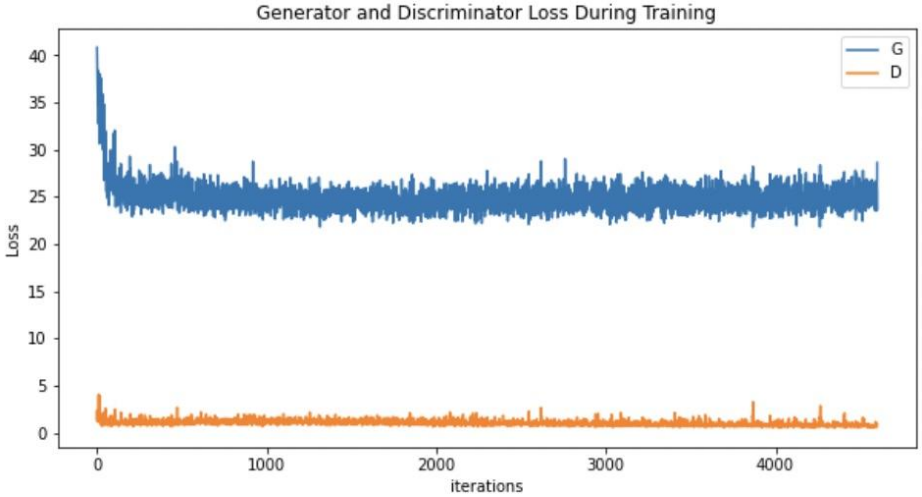
- Generator synthesizes fake images with the motive to fool the discriminator.
- Discriminator evaluates the fake images and provides feedback back to generator to generate better images from text descriptions.

Implementation

- The generator is fed with noise samples and text embeddings.
- Generator generates fake images from given inputs and feeds it to discriminator for evaluation.
- Discriminator is fed with random sample of fake and real training images to evaluate generator synthesized fake images against real images.
- Discriminator backpropagates the feedback to both discriminator and generator to optimise both of their jobs.
- It boils down to min-max optimization game involving 2 players trying to outsmart each other.

Results and Conclusions

- Nash equilibrium is established eventually when both generator and discriminator losses converge at an arbitrary threshold to synthesize most photo-realistic fake images.



Sample of fake images synthesized by generator

- Future scope: StackGANs and COCO dataset

NYU


Thank You!
NYU
6