MIT | Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

# SCHOOL OF COMPUTER ENGINEERING

# Text To Face Generation Using DCGAN


**Guide:**

Ms. Sharmila B Kharat


A brown hair girl with bow cut

**Students Names :**

Abhishek Pawar(35)

Riya Hiwanj (244)

Asrar Shaikh (144)

Ashwini Jadhav (157)

# INDEX


- Introduction & Motivation
- Problem Statement & Objectives
- Literature Review
- Implementation
- Conclusion & References

# Introduction

*The Headmaster at Hogwarts. Dumbledore is an old wizard with a long white beard and half-moon-shaped glasses.*

- Imagination is a natural intelligence.
- Text-to-Face is sub-domain of Text-to-Image.
- How to generate a face ?
- Generating characters from novel & stories.
- Identifying persons for criminal investigations.

labs.openai.com/e/jSJnqEzF8d7V9VTQaJ6CFRdv

(2) WhatsApp | Dashboard | My Purchases | Cou... | Machine Learning C... | Data Scientist Learn... | Abhishek Pawar | Li... | Abhipawar02 (Abhi... | Home | Infosys Spri... | Triplebyte - Create... | Fellowship program

DALL·E My collection

Edit the detailed description

Surprise me | Upload

an old wizard with long white beard and half moon glasses

Generate

Recent

Report issue

# Problem Statement

To visualize textual description of a human face we develop Text to Face generative model using deep convolutional generative networks.

# Objectives

1. To survey the existing text to image and text to face generative models.
2. To develop a generative model to construct face images of a person from textual input.

# Objectives

3. To design architecture diagrams  that would be able to handle the variability and complexity of natural language descriptions and generate realistic images that are consistent with the input.

4. To Implement generative model that would be a functional system that can generate realistic facial images from textual descriptions.

# Objectives

5. To improve the performance of the developed system by comparing it with the existing models.

6.To draft and communicate the improved results in a research paper

# LITERATURE REVIEW

- Text to Face/Image techniques
- Generative Adversarial Models
- Transformers Models
- Image Manipulation Models
- Image to Image Translation
- Application

# TEXT TO FACE/IMAGE TECHNIQUES

- In research by A. Kumar et al [1] they were able to generate a Realistic Image of a Face From a Text Description, they have shown an approach of encoders and decoders with Long short-term memory, how to encode the text into semantic vectors, and the approach of decoding the semantic features of text into natural images.
- In the Model Proposed by Jichao Zhang et al [15], they have demonstrated that their model can produce high-quality editing with higher view consistency while retaining the non-target regions when compared to the baselines.
- With the aid of a straight-forward optimization technique, the model can alter geometry while maintaining look. With the aid of a straightforward optimization technique, the model can alter geometry while maintaining look.

# GENERATIVE ADVERSARIAL NETWORK

- GAN is a combination of a generator and a discriminator with a built-in noise detector. Han Zhang et al [9] Explain StackGan++, how a generative adversarial network (GAN) might have high-resolution images based on the given text description.
- the paper proposed by Tero Karras et al [12] was A Style-Based Generator Architecture for Generative Adversarial Networks, in which they conducted research to demonstrate how Traditional GANs were improved by style-based design, for which they trained and tested two different models. Muhammad Zeeshan Khan et al [8] used a high-quality, sizable dataset for training purposes, and more dynamic images were used to train the model.

# TRANSFORMERS MODELS

- The Model Proposed by Brian Lovell et al [13] suggests a pre-trained BERT multi-label model for natural language processing, The model generates 40-word sparse text embeddings. Then, utilizes training data, where the images contain paired labels. high-resolution images are created using the cutting-edge StyleGAN2 image generator, which maps the noise vectors into a feature-disentangled latent space. The Multi-Label Text Classification feature of the Text to Face HD model enables the output of numerous diverse images for a given set of texts. With too many feature axis adjustments is difficult, the text vector has been moved to a low-density region of the latent space distribution, which leads to the mode collapse problem.

# IMAGE MANIPULATION MODELS

- A method proposed by  Lee B et al [2] that works on text-guided picture manipulation problems with multi-turn image manipulation (LatteGAN). both a Text-Conditioned U-Net discriminator architecture, which distinguishes between the global and local representations of fake or real images, and a module, which extracts fine-grained text representations for the generator.
- In the next paper of image manipulation, Bidirectional Generative Adversarial Network was proposed by Zixu Wang et al [6]. To generate semantically consistent images, they have propose two semantics-enhanced modules and Textual-Visual Bidirectional Generative Adversarial Network (TVBi-GAN). Specifically, this paper proposes a semantic enhanced attention module and a semantics-enhanced batch normalization module .During the adversarial process, the encoder could guide generator to explore corresponding features behind descriptions.

# IMAGE TO IMAGE TRANSLATION

- Image-to-image translation is the process of mapping images from one domain to another.
- A new Method was Developed by Jian Zhao et al [5] about Photographic Face using GANs. He developed a neural network that uses a combination of GAN and Skip connection layers to extract high-level semantic information from sketch images and attribute vectors.
- S. Nikkath Bushra et al [10] Discovered, In Deep Convolutional Generative Adversarial Network ,It is used to transform a forensic sketch into a real photo as part of a facial recognition system employing advanced artificial intelligence techniques (DCGAN).
- In Research by Saira Jabeen et al [7] Face Generation using Text Description is a model based on the sketch-refinement process which is implemented using StackGAN and a two-staged architecture.

# APPLICATION

- Varsha Hole et al [11] explains information about Face Generation in Forensic , In this they have propose a deep learning implementation based on GAN that consists of generators which generates an image based on textual description and discriminator which validates the image generated by generator with the data set.
- The Forensics Sketch Artist’s Job is Automated by a Model although, there is one drawback of this strategy is that low resolution images are produced due to the dataset’s comparatively small quantity of low-resolution shots.

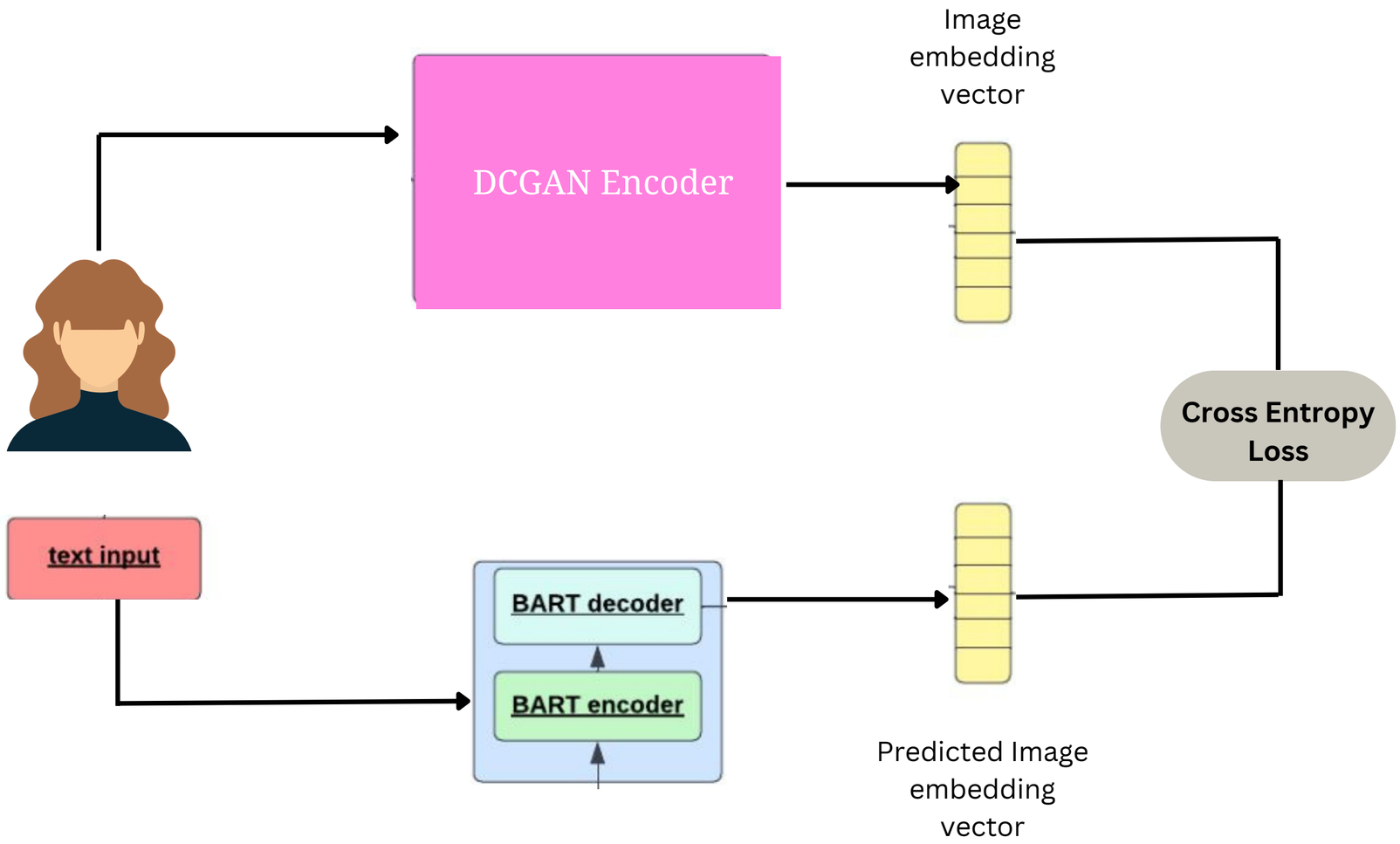
# Gap Identified

- Lack of diversity: Most models tend to generate faces that are very similar to each other and lack diversity in terms of facial features, age.
- Poor resolution: The generated faces often have low resolution and lack fine details, making them look unnatural.
- Limited control: The models often have limited control over the specific attributes of the generated face, making it difficult to generate faces that match a desired description.
- Bias and ethical considerations: Some models may perpetuate and amplify existing biases in the training data, leading to unequal representation and unfair outcomes.

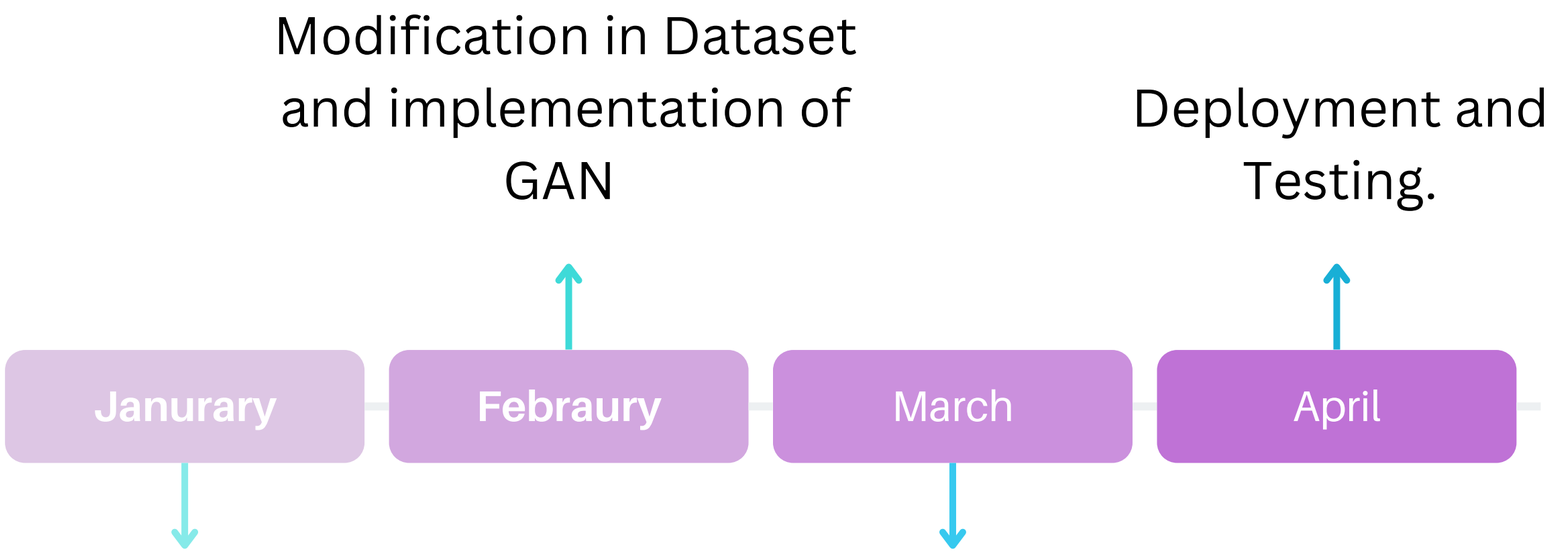
# Implementation

- Datasets - MNIST , CELEBA
- Discriminator :
  - sequential(conv2d, LeakyRelu, block, conv2d,sigmoid )
  - block( Conv2d, BatchNorm2d, LeakyReLu)
- Generator :
  - sequential(block, ConvTranspose2d, Tanh )
  - block(ConvTranspose2d, BatchNorm, Relu)
- Training

# Training Diagram



# Implementation Timeline



# Application

- AI generated Characters.
- Portrait a characters face from text/novel.
- Use in commercials
- Game Industries.
- Criminal Investigations.

# Conclusion

- Text to image generation techinques
- Text to face generation
- Architecture diagram and Methedology
- Implementation detials
- Aim and Objectives
- Future Scope

# Thank You