Generative Fusion: AI-Powered Image Synthesis Through Gan and NLP Integration

P KARTHIK
Karthikmissile2002@gmail.com
Students Computer Science and
Engineering Dr.Mgr Educational
Research Institute University,
Chennai,India

P SOHAIL KHAN
khansohail8809@gmail.com
Students Computer Science and
Engineering Dr.Mgr Educational
Research Institute University ,
Chennai,India

N DINIESH CHAITANYA
dineshchaitanyanagabathula@gmail.com
Students Computer Science and
Engineering Dr.Mgr Educational
Research Institute University ,
Chennai,India

COORDINATOR
DR.REKHA.K.K
DR.A.VINOTH KUMAR
ASSOCIATE PROFESSORS
DR.MGR EDUCATONAL AND
RESEARCH INSTITUTE UNIVERSITY

C.SUBBULAKSHMI
ASSOCIATE PROFESSOR
DR.MGR EDUCATIONAL AND
RESEARCH INSTITUTE UNIVERSITY

ABSTRACT
ABSTRACT
The real world which is working with artificial intelligence generates different kinds of algorithms which are helping us in solving the realistic problems and generating images using natural language processing(NLP) which helps in building a communication with computer and human. Now in the present generation generating images with verbal descriptions has gained much popularity. The generative adversarial network (GAN) which belongs to deep learning family model helps us in generating the image by the conjunction with AIML(Artificial Markup Language)which helps the users in generating better graphics directly with the user input instructions.

In general there are two main parts required for text to image synthesis with GANs:
1.　　Discriminator
2.　　Generator

The working of generator is it should take the input provided by the user and generate the realistic image. It takes the text as input and gives image as output. The job of discriminator is like a gatekeep who always stops the fake one to not let enter as output.

OBJECTIVE
The main objective of this text to image synthesis is that to create an AI model that uses the textual data to create visuals.

Inorder to create those visuals deep learning models were used specifically transformers and Genarative Adversarial Networks(GANs).

INTRODUCTION
Text to image generators are the result of the advanced picture generating by the combination of AIML (Artificial Markup Language),NLP(Natural Language Processing),GANs(Generative advrsarial N\etwork) which plays the major role in generating the images by taking the input and generating the outputs.

Originally it uses pytorch framework which is a deep learning library and the python as programming language and from NLP library it uses sentence transformers. When an user texts a input to the generate the image It doesn't directly take the text which is converted into embeddings. Inorder to generate the image as output by the generator firstly the generator model needs to be trained with .jpg files which is an ultimate task.

The GAN requires a multifaceted strategy to create Text to image by combining with AIML. So first the generator gets trained with the sizable dataset inorder to generate the output for the given user input.

LITERATURE SURVEY

Akansha Singh, Sonam Anekar, Rithika Shenoy, Prof. Sainath Patil in 2022 they developed an easy and efficient model for low resolution image but in future they will like to improve the quality of the image of generating in high resolution.It is limited control over style and details of generated images.

Sadia Ramzan,Muhammad Munwar Iqbal,Tehmina klasum also in 2022 they developed a GAN model which is a deep learning model with NLP that helps in generating the images of the user provided input. But there is limited image quality .

Scott Reed, Zeynep Akat,Xinchen Yen in 2016 they developed a transformer which is the function of NLP that helps in taking the taking the text as input and get convereted in to embeddings.the drawback is it is highly intensive and challenging to scale.

BowenLI,Xiaojaun QI, Thomas Lukasiewicz,P.philip in 2019 introduced control GAN that helps in generating the perfect images of the textual provided.

EXISTING SYSTEM

Generative adversarial network(GAN): is deep learning model which is an unsupervised learning technique that generates new images . Inorder to create the new data instances Neural networks were used by the Generative adversarial networks. The images can be generated through user input texts with those networks.

Generator: Generator is the part of Generative adversarial networks which helps in taking the inputs from the user for generating the images as output. Inorder to generate new images firstly the GAN needs to get trained by the sizable data set inorder to learn patterns and know about the image when the user inputs a image .

Discriminator : Discriminator is also a part of Generative adversarial networks which acts like gatekeeper for not allowing the image which doesn't looks real. So the discriminator also gets trained for identifying the images wether the images generated by the generated by generator are real or fake and lets us know by using sigmoid function which generates 0 or 1 for false or true statement .

The process here is the generator takes the user input after getting trained with a particular and sizable data set it starts to generate images that were real or fake and the images generated by that generator will be sended as input to the discriminator will identify that real or fake if that image looks like fake then it generates 0 as the value and doesn't allow the generator to display the image and this process

continuous till the generator generates the real images.

In specific terms generator G and a discriminator D engage in subsequent game on V(D,G):

Min min G max D V (D,G)=Ex~pdata(x) [log] D(x)]+(1)Ex~pz(z)[log(1-D(z)).
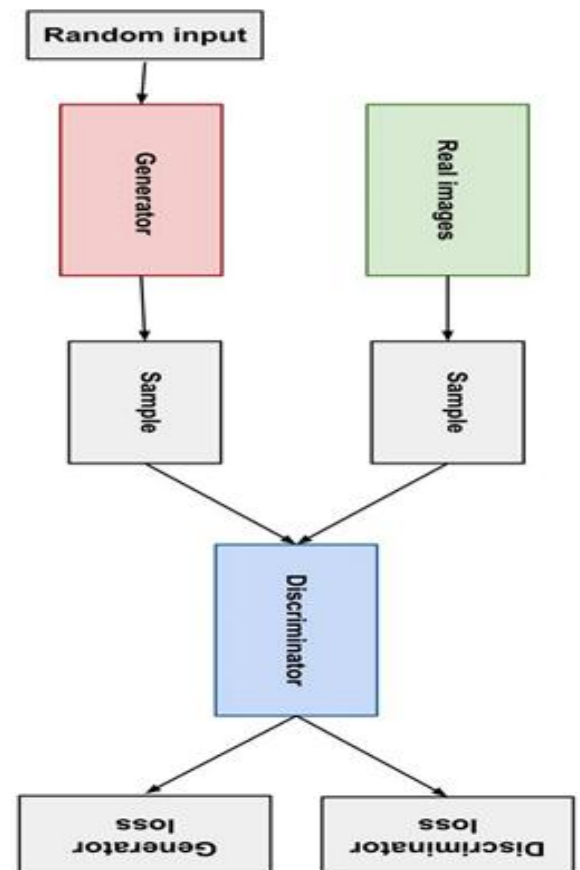

PROPOSED SYSTEM

By using the deep learning models which is an unsupervised technique helps in generating the images perfectly in the best possible way.

How the generative deep learning model works :

Generative deep learning models are unsupervised learning technique which helps in generating the images perfectly with different resolutions based on your model.

There are various applications in generating images but using GAN with NLP integration it help in generating the images according to the provided user

input also GAN is a popular deep learning model in generating images there are other architectures also like Variational Autoencoders (VAEs) and diffusion models.
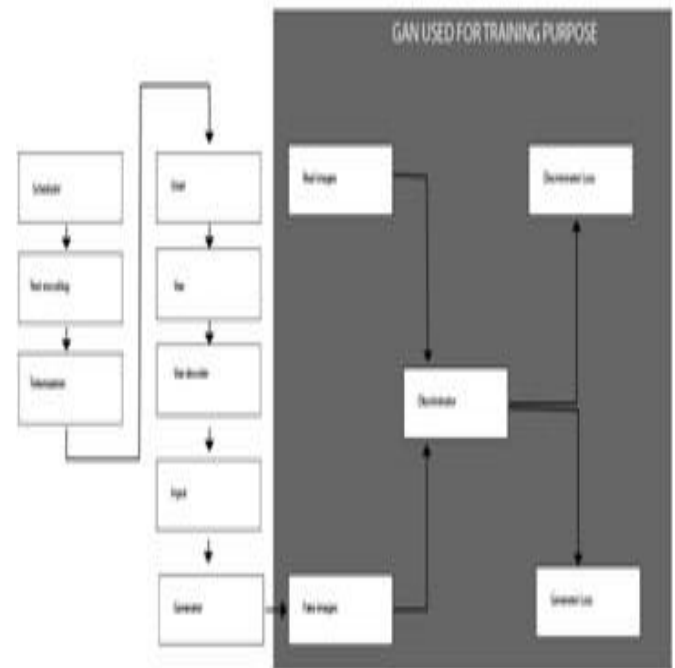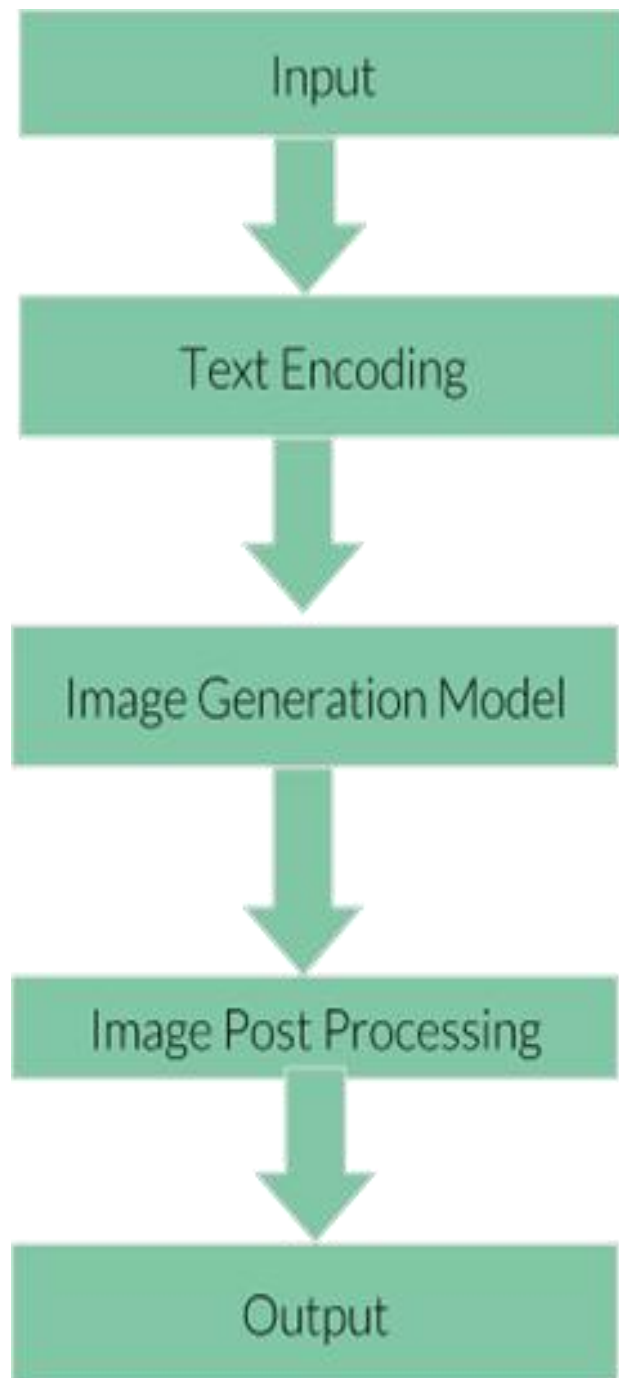

PROBLEM STATEMENT

Since it is very difficult to visualize the text when reading it this may be a problem.Certain mords can also have wrongly readen in specific times.

When it comes to image the image will be very lookable if you have used a perfect GAN model with NLP integration . But the main problem statement is Generator should be trained with a large dataset because training generator with a small quantity images helps generator in just detecting the colour patterns shapes so inorder to detect the exact image of the text provided it needed to train by a large dataset which may takes more than the time require
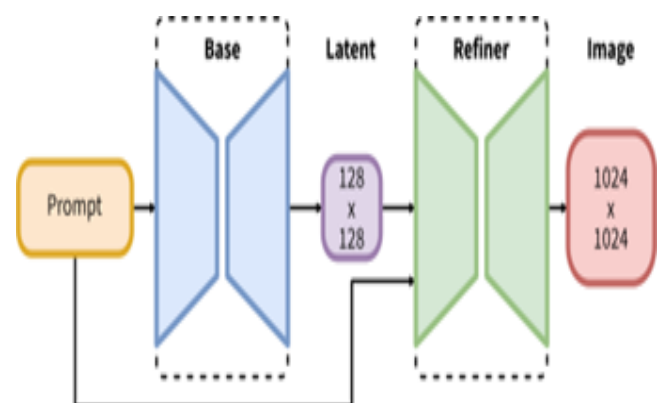
GAN   ARCHITECTURE
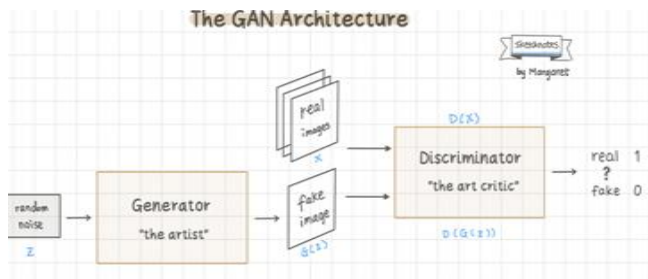
RESULTS:

The very first we need to train the train the deep learning model that is GAN with a large number of dataset that might contain cartoon pictures after training the GAN model we need to give an input that contains in the pictures we trained then its gives the output as shown in the results.

CONCLUSION

An important development in AI image synthesis is the use of Generative adversarial networks in text to picture generation.So how better the Model you train the high resolution images you get. GAN which is deep learning model is a an excellent model and is also a unsupervised learning technique which helps in generating images by the conjunction with AIML(Artificial Markup Language) also with NLP which helps in taking the texts and convert it in to embeddings which is understandable by the computer how better the GAN will be the perfect images you get and also there will be limited editing in styles as drawbacks in each and every case.

REFERENCES

1.   https://www.researchgate.net/publication/359441889_Text_to_Image_using_Deep_Learning

2.   B. Goertzel and C. Pennachin, Artificial general intelligence. Springer, 2007, vol. 2. 1

3.   Ankit Yadav1, Dinesh Kumar Vishwakarma2, Recent Developments in Generative Adversarial Networks: A Review (Workshop Paper),2020.

4.   Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee , "Generative Adversarial Text to Image Synthesis" in University of Michigan and Max Planc

Institute for Informatics June 2016.

5.   Zhang, Zhe Gan, Xiaolei Huang, and Xiaodong He. Attngan: Fine-grained text to image generation with attentional generative adversarial networks. CoRR, abs/1711.10485, 2017.

6.   Mehdi Mirza, Simon Osindero, Conditional Generative Adversarial Nets,2014.

7.   Stian Bodnar, Jon Shapiro, "Text to Image Synthesis Using Generative Adversarial Networks" in The University of Manchester May 2018.

8.   V. C. Muller and N. Bostrom, "Future progress in artificial in- ¨telligence: A survey of expert opinion," in Fundamental issues of artificial intelligence. Springer, 2016, pp. 555–572.

9.   https://proceedings.neurips.cc/paper_files/paper/2019/file/1d72310edc006dadf2190caad5802983-Paper.pdf

10.   Gregor, K., Danihelka, I., Graves, A., Rezende, D., and Wierstra, D. Draw: A recurrent neural network for image generation. In ICML, 2015.

11.   J. Clune, "Ai-gas: Ai-generating algorithms, an alternate paradigm for producing general artificial intelligence," arXiv preprint arXiv:1905.10985, 2019.

12.   R. Fjelland, "Why general artificial intelligence will not be realized," Humanities and Social Sciences

13.   Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, 2015

14.   https://www.mdpi.com/2673-4591/20/1/16

15.   Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, "StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks" in Rutgers University and Lehigh University August 2017.