IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Text To Image Genernation Using AI And Deep Learning

[1]Akanksha Chinake, [2]Shreya Pawar , [3]Aishwarya Gaikwad, [4]Prajakta Kshirsagar, [5]Prof. S. R. Bhujbal

[1]Student, [2]Student, [3]Student, [4]Student, [5] Head of DepartmentBE (Department Of Computer Engineering)

P K Technical Campus, Pune, India

Abstract : Text to image synthesis refers to the method of generating images from the input text automatically. Deciphering data between picture and text is a major issue in artificial intelligence. Automatic image synthesis is highly beneficial in many ways. Generation of the image is one of the applications of conditional generative models. For generating images, GAN(Generative Adversarial Models) are used. Recent progress has been made using Generative Adversarial Networks (GAN). The conversion of the text to image is an extremely appropriate example of deep learning.

Index Terms **-** convolutional neural network; recurrent neural network; deep learning; generative adversarial networks; imagegeneration

- **INTRODUCTION**

Once the innovation is ready for business applications, there are a lot of future possible usesfor the creation of graphics from normal language. The layout of generative models has a place for generative adversarial networks. It suggests that they are able to produce new chemicals. Pixels are usedto transform text into images. Using the example: "Flower with pink petals." GAN is made up of a combination of two competing neural network models that track,capture, and replicate the variations present in a dataset. The main goal of text to picture synthesis is to turn suitable word descriptions into images. GAN models are frequently utilized today for betteroutcomes. Additionally, one issue with deeplearning is that there are numerous potential configurations for single text descriptions.

- **Motivation of project**

Converting text into a shape turns it into a vector layer, so you can edit it and cus- tomize it just like any other vector shape by adjusting points, direction handles, or combining it with other shapes.

- **Methodology**

This section describes the training details of deep learning-based generative models. Conditional GANs were used with recurrent neural networks (RNNs) and convolutional neural networks (CNNs) for generating meaningful images from a textual description. The dataset used consisted of images of flowers and their relevant textual descriptions. For generating plausible images from text using a GAN, preprocessing of textual data and image resizing was performed. We took textual descriptions from the dataset, preprocessed these caption sentences, and created a list of their vocabulary. Then, these captions were stored with their respectiveids in the list. The images were loaded and resized to a fixed dimension. These data were then given as input to our

proposed model. RNN was used for capturing the contextual information of text sequences by defining the relationship between words at altered time stamps. Text-to-image mapping was performed using an RNN and a CNN. The CNN recognized useful characteristics from the images without the need for human intervention

- **Data Flow Diagram 0**

In Data Flow Diagram,we Show that flow of data in our system in DFD0 we show that base DFD in which rectangle present input as well as output and circle show our system,In DFD1 we show actual input and actual output of system input of our system is text or image and output is rumor detected like wise in DFD 2 we present operation of user as well as admin.

Figure 4.2: Data Flow(0) diagram

- **Data Flow Diagram 1**

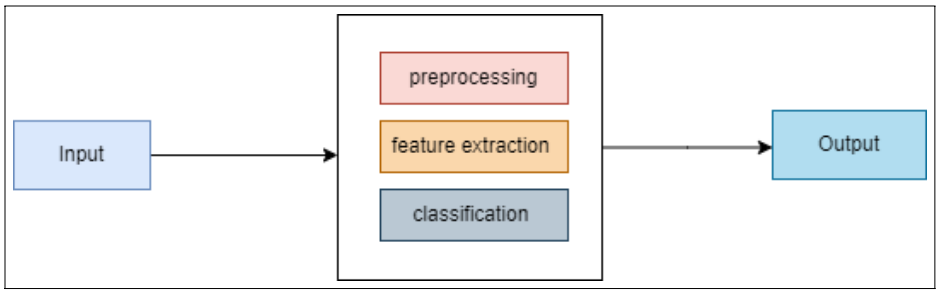
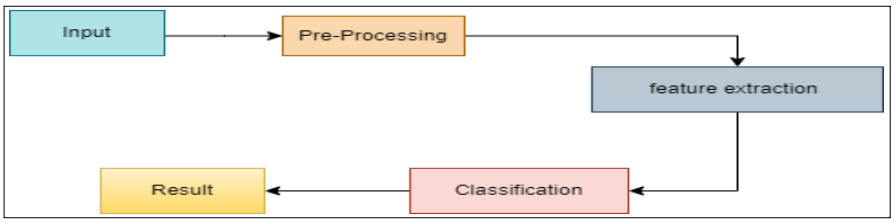
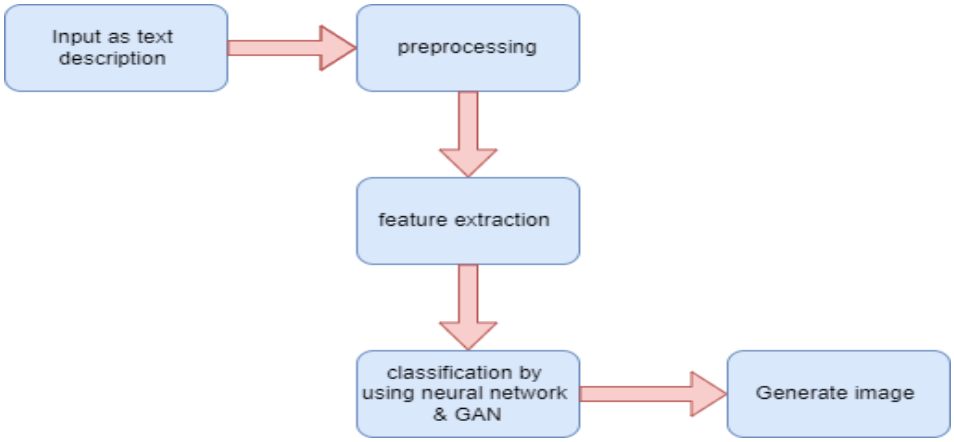


Figure 4.3: Data Flow(1) diagram



- **System Architecture**



- **Literature Survey**

1. **Paper Name:** LEARNING TEXT TO IMAGE SYNTHESIS WITH TEXTUAL DATAAUGMENTATION
   Author: Hao Dong, Jingqing Zhang, Douglas McIlwraith, Yike Guo
   Abstract ::- Translating information between text and image is a fundamental prob- lem in artificial intelligence that connects natural language processing and com- puter vision. In the past few years, performance in image caption generation has seen significant improvement through the adoption of recurrent neural networks (RNN). Meanwhile, text-to-image generation begun to generate plausible images using datasets of specific categories like birds and flowers. We've even seen image generation from multi-category datasets such as the Microsoft Common Objects in Context (MSCOCO) through the use of generative adversarial networks (GANs). Synthesizing objects with a complex shape, however, is still challenging. For ex- ample, animals and humans have many degrees of freedom, which means that they can take on many complex shapes..
2. **Paper Name:** A Hybrid Deep Neural Network for Urdu Text Recognition in Nat- ural Images
   Author: Asghar Ali,Mark Pickering
   Abstract : —In this work, we present a benchmark and a hybrid deep neural net- work for Urdu Text Recognition in natural scene images. Recognizing text in natural scene images is a challenging task, which has attracted the attention of computer vi- sion and pattern recognition communities. In recent years, scene text recognition has widely been studied where; state-ofthe-art results are achieved by using deep neu- ral network models. However, most of the research works are performed for English text and a less concentration is given to other languages. In this paper, we investigate the problem of Urdu text recognition in natural scene images. Urdu is a type of cur- sive text written from right to left direction where, two or more characters are joined to form a word. Recognizing cursive text in natural images is considered an open problem due to variations in its representation.

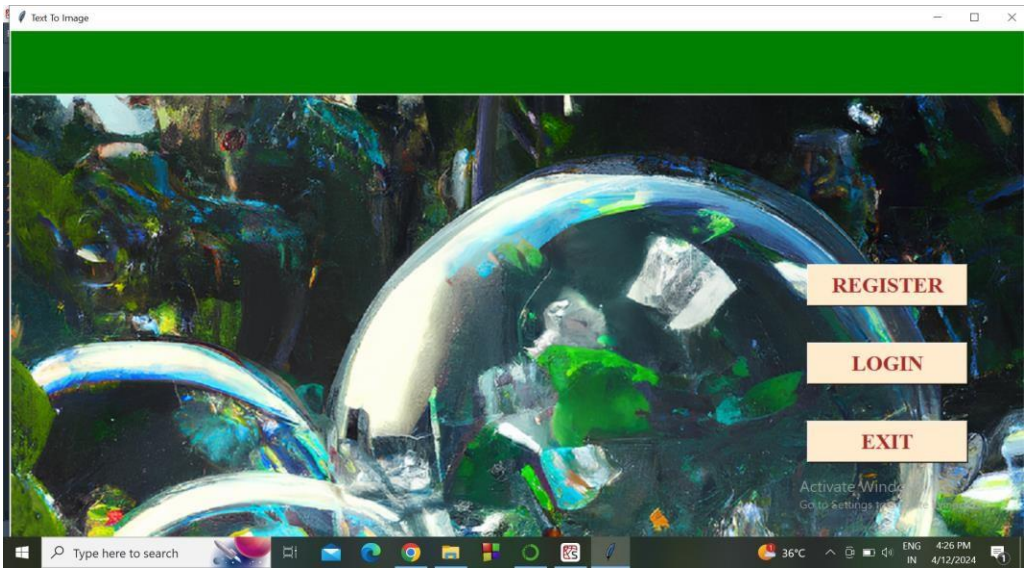
3. **Paper Name:**Google Image Search Refinement: Finding Text in Images Using LocalFeatures
   Author name: Yi Zhou, Kai Chen, Xiaokang Yang
   abstract : In this paper, we implement a Google image search refinement that uti- lizes local features to find Chinese characters in search results, including following stages: (1) Chinese characters images and their SIFT features (SDB) are generated offline, (2) A text-based image search results are retrieved from the Google, (3) SIFT features of results are matched to query-text SDB using MPLSH, (4) A geometric verification algorithm is used to find the query-text and rerank results. Experiment results show that our approach is simple and effective in recognition of text in natural images, and is helpful to refine the web image search.
4. **Paper Name:**Graph-based text segmentation using a selected channel image
   Author: Chao Zeng , Wenjing Jia
   abstract : This paper proposes a graph-based method for segmentation of a text im- age using a selected colour-channel image. The text colour information usually presents a twopolarity trend. According to the observationthat the histogram distri- butions of the respective colour channel images are usually different from each other, we select the colour channel image with the histogram having the biggest distance between the two main peaks, which represents the main foreground colour strength and background colour strength respectively. The peak distance is estimated by the mean-shift procedure performed on each individual channel image.

- **LITERATURE SURVEY TABLE:**

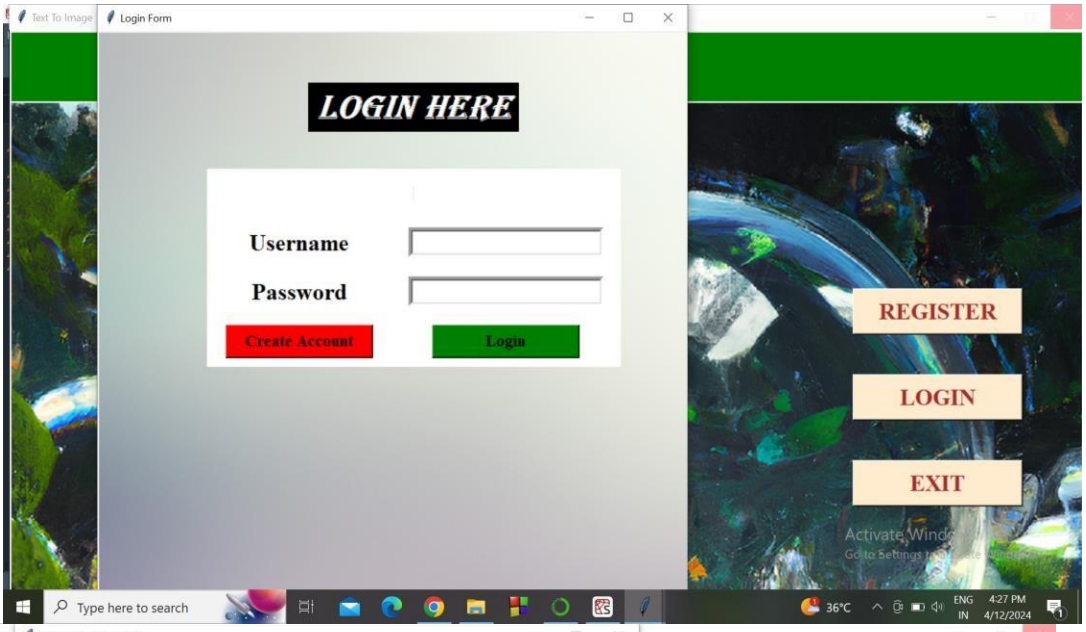
Sr.no	Title	Author	Abstract
1	Joint Embedding based Text- to-Image Synthesis	Menglan Wang, Yue Yu*	Learning joint embedding between image and text is significant for text-to-image synthesis as it bridges the semantic gap between image and text.

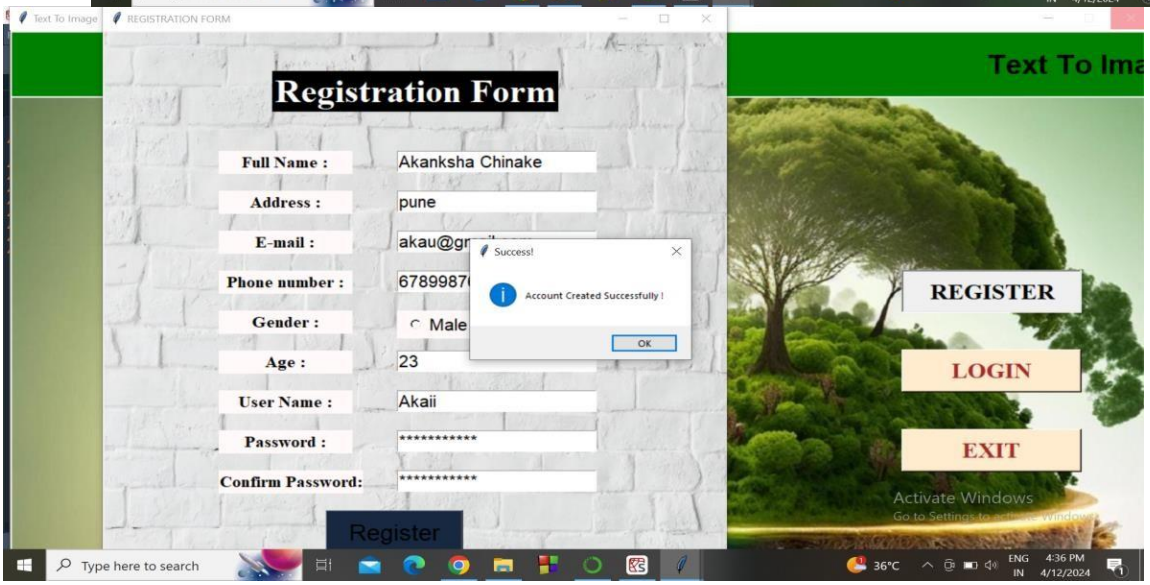
2	Graph-based text segmentation using a selected channel image	Chao Zeng, Wenjing Jia	This paper proposes a graph-based method for segmentation of a text image using a selectedcolour-channel image.
3	Google Image Search Refinement: Finding Text in Images Using Local Features	Yi Zhou, Kai Chen, Xiaokang Yan	In this paper, we implementa Google image search refinement that utilizes local features to find Chinese characters in search results
4	learning textto image synthesis with textual data augmentation	Hao Dong, Jingqing Zhang	Translating information betweentext and image is a fundamental problem in artificial intelligence that connects natural language processing and computer vision.

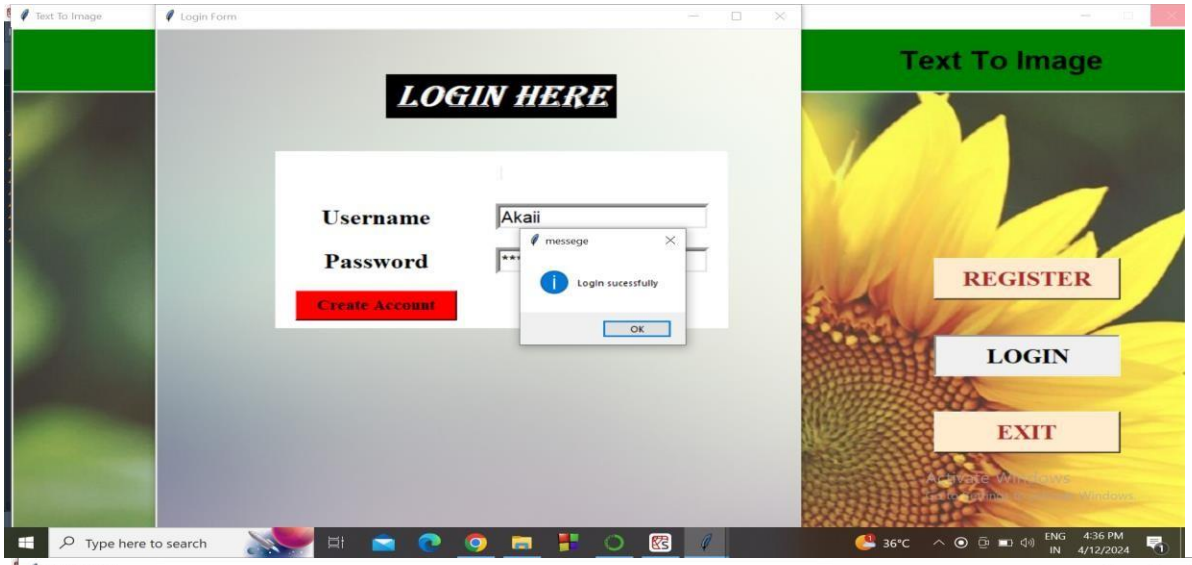
- **Result**

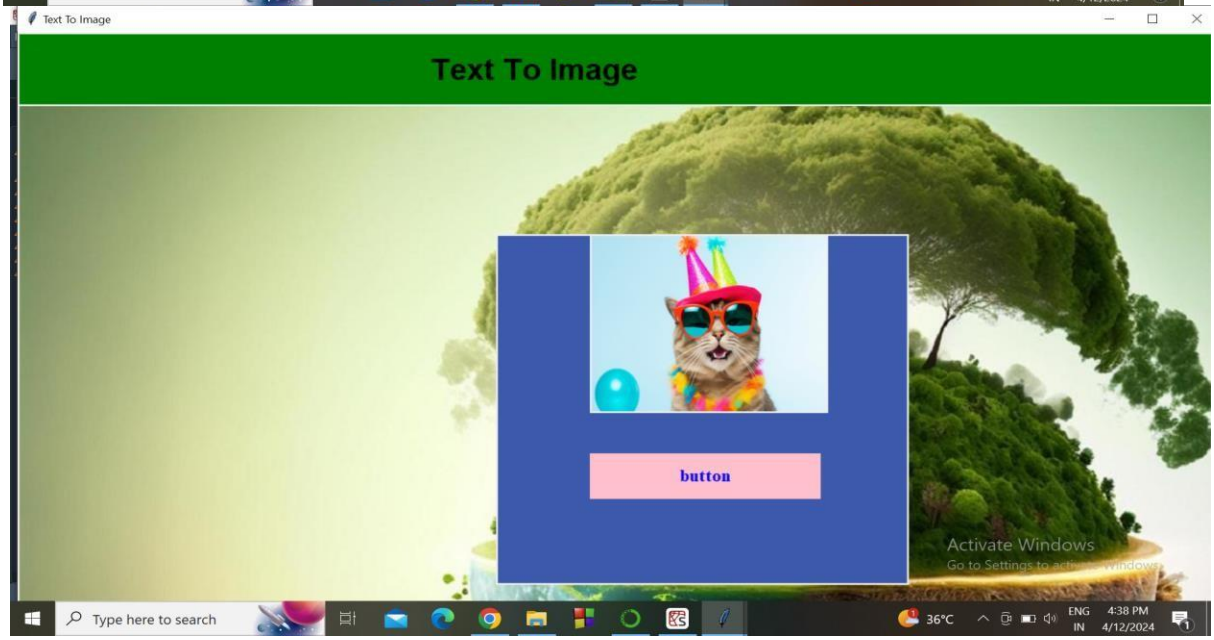


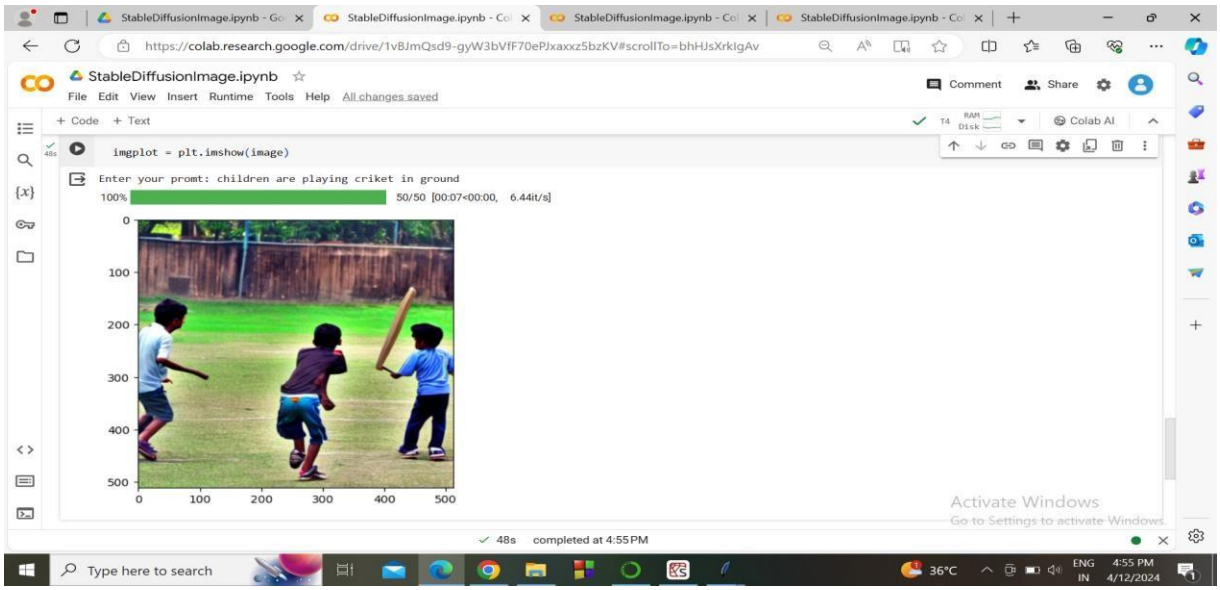

REGISTRATION FORM
Registration Form
Full Name :
Address :
E-mail :
Phone number :
0
Gender :
Male
Female
Age :
0
User Name :
Password :
Confirm Password:
Register
REGISTER
LOGIN
EXIT
Type here to search

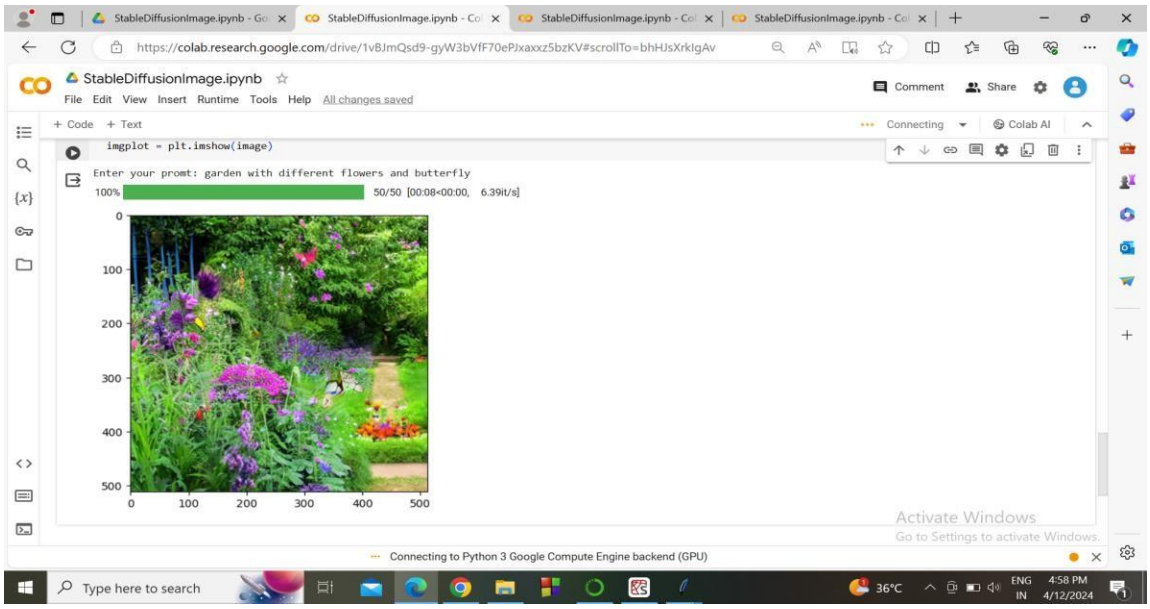

Login Form
LOGIN HERE
Username
Password
Create Account
Login
REGISTER
LOGIN
EXIT
Type here to search


REGISTRATION FORM
Registration Form
Full Name :
Akanksha Chinake
Address :
pune
E-mail :
Phone number :
Gender :
Male
Age :
23
User Name :
Akaii
Password :
Confirm Password:
Register
Success!
Account Created Successfully !
OK
REGISTER
LOGIN
EXIT
Type here to search


Text To Image
Login Form
LOGIN HERE
Username
Akaii
Password
Create Account
messege
Login sucessfully
OK
Text To Image
REGISTER
LOGIN
EXIT
Type here to search
36°C
ENG
IN
4:36 PM
4/12/2024


Text To Image
Text To Image
button
Activate Windows
Type here to search
36°C
ENG
IN
4:38 PM
4/12/2024


StableDiffusionImage.ipynb
https://colab.research.google.com/drive/1vBJmQsd9-gyW3bVfF70ePJxaxxz5bzKV#scrollTo=bhHJsXrkIgAv
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
Colab AI
imgplot = plt.imshow(image)
Enter your promt: children are playing criket in ground
100%
50/50 [00:07<00:00, 6.44it/s]
Activate Windows
48s completed at 4:55PM
Type here to search
36°C
ENG
IN
4:55 PM
4/12/2024



- **Conclusion**

After conducting a combined study of the papers and planning the project implemen- tation, we developed an easy and efficient model for image generation. In future, we would like to improve themodel so as to get pictures having a high resolution and will use this model on other dataset as well.

- **Acknowledge**

We would like to take this opportunity to thank our internal guide Prof. S. R. Bhujbal for giving us all the help and guidance we needed. We are really grateful to them for their kind support. Their valuable suggestions were very helpful.

We are also grateful to Prof. S. R. Bhujbal, Head of Computer Engineering Department, for his indispensable support,suggestions

- **Reference:**

- 1. Ankit Yadav1, Dinesh Kumar Vishwakarma2, Recent Developments in Generative Adversarial Networks: A Review (Workshop Paper),2020.
- 2. Gregor, K., Danihelka, I., Graves, A., Rezende, D., and Wierstra, D. Draw: A recurrent neural network for image generation. In ICML, 2015.
- 3. Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, “StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Net- works” in RutgersUniversity and Lehigh University August 2017.
- 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. Stian Bodnar, Jon Shapiro, “Text to Image Synthesis Using Generative Adversarial Networks” in The University of Manchester May 2018.
- 6. Zhang, Zhe Gan, Xiaolei Huang, and Xiaodong He. Attngan: Finegrained text to image generation with attentional generative adversarial networks. CoRR, abs/1711.10485, 2017.
- 7. Mehdi Mirza, Simon Osindero, Conditional Generative Adversarial Nets, 2014.