International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

AI IMAGE GENERATOR THROUGH TEXT PROMT

[1]Sarthak Shivaji Bagul, [2] Atharv Vasant Ahire, [3] Abhijit Rajendra Chavhan,[4] Rugved Samadhan Dumbare, [5] Mr. S.K. Mahajan

[1] sarthakbagul84@gmail.com, Student of Dept. of Computer Technology K.K.Wagh Poly, Nashik, India
[2] atharvahire206@gmail.com, Student of Dept. of Computer Technology K.K.Wagh Poly, Nashik, India
[3]abhijitchavhan1724@gmail.com, Student of Dept. of Computer Technology K.K.Wagh Poly, Nashik, India
[4] rugveddumbre@gmail.com, Student of Dept. of Computer Technology K.K.Wagh Poly, Nashik, India
[5]Internal Guide Dept. of Computer *Technology K.K.Wagh Poly, Nashik, India*

Abstract -

In the era of digital content creation, the Text-to-Image Generator has emerged as a powerful and innovative tool that transforms textual descriptions into captivating visual representations. This website leverages advanced deep learning techniques to bridge the gap between imagination and reality, enabling users to effortlessly generate stunning images from their written words.

Our Text-to-Image Generator website provides a user-friendly interface where individuals can input textual prompts, whether they are vivid descriptions, imaginative stories, or abstract concepts. Behind the scenes, state-of-the-art generative models decode the text and generate high-quality, contextually relevant images that reflect the essence of the input. In the digital landscape, the fusion of text and images has become a quintessential form of communication. Text-to-image generators serve as pivotal tools in this synergy, bridging the conceptual realm of language with the visual spectrum. This abstract delves into the conceptualization of a cutting-edge Text Prompt to Image Generator website, epitomizing the amalgamation of natural language processing and computer vision technologies.

This innovative platform harnesses the power of advanced deep learning algorithms, empowering users to input textual descriptions and witness them materialize into vivid, high- resolution images. The generator employs state-of-the-art techniques such as Generative Adversarial Networks (GANs) and Transformer architectures to interpret nuanced textual cues, capturing abstract ideas and intricate details. Users can explore a diverse array of categories, from nature and architecture to fantastical realms, ensuring a versatile and captivating user experience.

Key Words: *"Text to Image Generator" , "AI Text to Image", "Text-Driven Image Creation" , "Generate Images from Text" , "Image Generator from Words".*

1.INTRODUCTION

The purpose of an AI Image Generator Through Text Prompt project is to create a system or application that can generate images based on textual descriptions or prompts using artificial intelligence techniques, particularly deep learning and generative models. This technology serves several key purposes: Content Creation : It enables the automatic creation of images or visual content from text, which can be valuable for various applications such as digital art, graphic design, marketing, and more. Creative Assistance : Artists and content creators can use AI image generators to aid their creative process. They can describe their ideas in text, and the AI can provide visual representations or drafts for further refinement.Accessibility : It can assist individuals who may have difficulty creating images using traditional graphic design tools. They can describe what they want, and the AI can generate the visual content for them.Prototyping and Mockups : It's a useful tool for quickly prototyping or creating mockup images for websites, applications, or products based on textual descriptions.

Augmented Reality and Virtual Reality : In AR and VR applications, AI image generators can help create realistic environments, objects, or characters based on textual input.

Storytelling : Authors and writers can use AI image generators to visualize scenes, characters, or settings in their stories based on text descriptions. Data Augmentation : In machine learning, AI image generators can be used to create additional training data for computer vision tasks, which is especially valuable when labeled datasets are limited.

1. User Interface (UI): Create a user-friendly web interface using HTML, CSS, and JavaScript. Include a text input field for users to enter their prompts. Implement a "Generate" button to trigger image generation. Display the generated image on the webpage.

2. Backend: Use a server-side language (e.g., Python with Flask or Node.js with Express) to handle user requests. Set up routes to receive text prompts from the UI.

3. Natural Language Processing (NLP): Integrate an NLP library or API (e.g., spaCy, NLTK, or Hugging Face Transformers) to process and understand user input. Extract keywords and context from the text prompt to provide more context for image generation.

4. Image Generation: Utilize a pre-trained image generation model (e.g., GANs, VQ-VAE-2, CLIP) to create images based on the input text. Pass the processed text to the model, which generates an image representation.

5. Image Display: Return the generated image to the frontend for display. Use JavaScript to dynamically update the webpage with the generated image.

6. Error Handling: Implement error handling to address issues such as invalid input, server errors, or model failures. Provide meaningful error messages to users.

7. Security: Secure user data and server endpoints to prevent unauthorized access. Implement rate limiting and CAPTCHA to protect against abuse.

8. Scalability: Consider server scalability to handle a large number of user requests. Use cloud hosting solutions like AWS, Azure, or Google Cloud if needed.

9. Optimization: Optimize image generation for speed and resource efficiency. Implement caching to store and reuse generated images when possible.

10. User Experience: Add features like image zoom, download, and sharing options. Provide feedback to users during image generation to indicate progress.

2. LITERATURE SURVEY

1. Show and Tell: A Neural Image Caption Generator

Automatically describing the content of an image is a fundamental problem in artificial intelligence that connects computer vision and natural language processing. In this paper, we present a generative model based on a deep recurrent architecture that combines recent advances in computer vision and machine translation and that can be used to generate natural sentences describing an image. The model is trained to maximize the likelihood of the target description sentence given the training image. Experiments on several datasets show the accuracy of the model and the fluency of the language it learns solely from image descriptions. Our model is often quite accurate, which we verify both qualitatively and quantitatively. For instance, while the current state-of-the-art BLEU-1 score (the higher the better) on the Pascal dataset is 25, our approach yields 59, to be compared to human performance around 69. We also show BLEU-1 score improvements on Flickr30k, from 56 to 66, and on SBU, from 19 to 28. Lastly, on the newly released COCO dataset, we achieve a BLEU-4 of 27.7, which is the current state-of-the-art.

2. Analysis of Appeal for Realistic AI-Generated Photos

AI-generated images have gained in popularity in recent years due to improvements and developments in the field of artificial intelligence. This has led to several new AI generators, which may produce realistic, funny, and impressive images using a simple text prompt. DALL-E-2, Midjourney, and Craiyon are a few examples of the mentioned approaches. In general, it can be seen that the quality, realism, and appeal of the images vary depending on the used approach. Therefore, in this paper, we analyze to what extent such AI-generated images are realistic or of high appeal from a more photographic point of view and how users perceive them. To evaluate the appeal of several state-of-the-art AI generators, we develop a dataset consisting of 27 different text prompts, with some of them being based on the DrawBench prompts.

3. Zero-Shot Text-to-Image Generation

Text-to-image generation has traditionally focused on finding better modeling assumptions for training on a fixed dataset. These assumptions might involve complex architectures, auxiliary losses, or side information such as object part labels or segmentation masks supplied during training. We describe a simple approach for this task based on a transformer that auto regressively models the text and image tokens as a single stream of data. With sufficient data and scale, our approach is competitive with previous domain-specific models when evaluated in a zero-shot fashion.

4. StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks

Synthesizing high-quality images from text descriptions is a challenging problem in computer vision and has many practical applications. Samples generated by existing textto- image approaches can roughly reflect the meaning of the given descriptions, but they fail to contain necessary details and vivid object parts. In this paper, we propose Stacked Generative Adversarial Networks (StackGAN) to generate photo-realistic images conditioned on text descriptions.We decompose the hard problem into more manageable sub-problems through a sketch-refinement process. The Stage-I GAN sketches the primitive shape and colors of the object based on the given text description, yielding Stage-I low-resolution images. The Stage-II GAN takes Stage-I results and text descriptions as inputs, and generates high-resolution images with photo-realistic details. It is able to rectify defects in Stage-I results and add compelling details with the refinement process.

3. MODULE DESCRIPTION

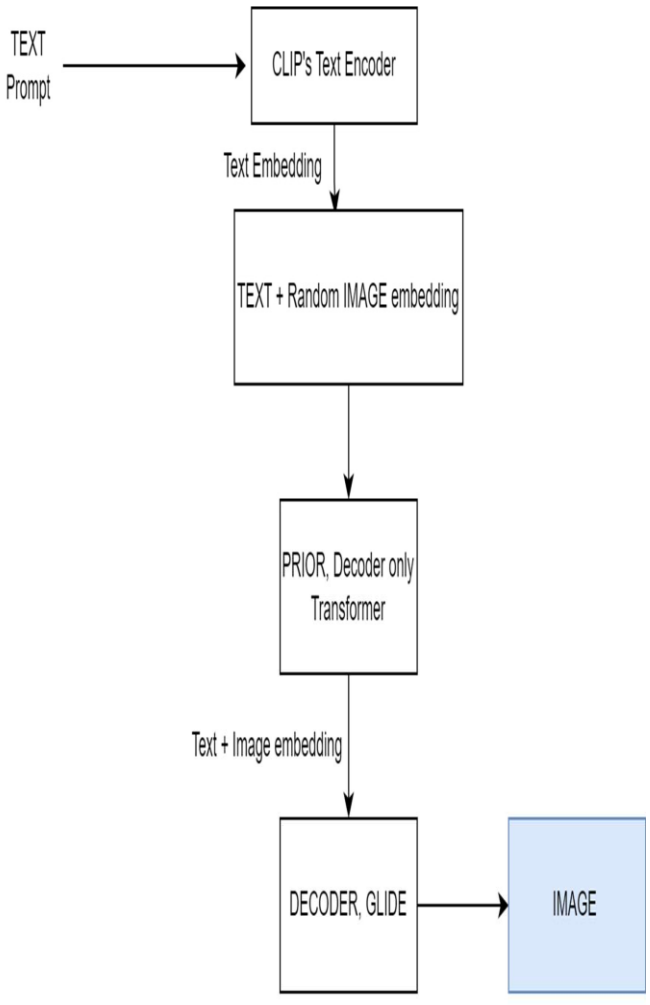

Fig3.1 Design Concept

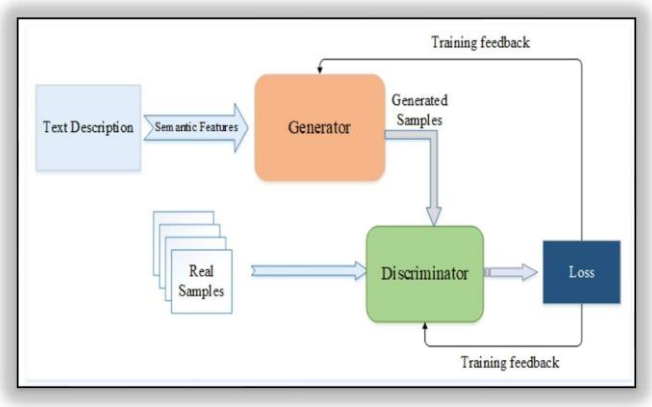

Fig3.1 System Architecture

6. CONCLUSIONS

The text prompt to image generator website offers a revolutionary way to create visual content. With its advanced AI algorithms, it can transform written descriptions into stunning images. This technology has significant implications for content creators, marketers, and designers. It streamlines the creative process, saving time and effort. The website's versatility allows for various artistic styles and themes, making it adaptable to diverse needs. However, potential concerns include ethical considerations, such as the misuse of generated images. Despite this, the website represents a promising leap in AI-generated visuals, opening up exciting possibilities for innovation and creativity in the digital realm. Its future evolution will undoubtedly be closely watched by both professionals and the broader online community.

7. REFERENCES

1. PMBOK Guide, A guide to the project management body of know, Third Edition, 2004, p.
2. C.L.Hwang & K.P.Yoon , " Multiple Attribute Decision Making and Introduction " , London , Sage Publication , 1995,PP2.
3. Ghodsypour.H,"Multiple Objective Decision Making", Amir Kabir University Publication, 2003.
4. R. P. Mohanty, project selection by a multi-criteria decision making method: an example from a developing country, International Journal of Project Management, 10, (1992), 31-38.
5. S. Mahmoodzadeh, J. Shahrabi, M. Pariazar and M. S. Zaeri, project selection by using fuzzy AHP and TOPSIS technique, International Journal of Human and Social Sciences, 1, (2007).
6. Eddie W. L. Cheng, Heng Li, exploring quantitative methods for project location selection, Building and Environment, 39, (2004), 1467-76.
7. R. P. Mohanty, R. Agarval, A. K. Choudhury and M. K. Tiwari, International Journal of Production Research, International Journal of Production Research, 43(23), (2005), 5199-5216.
8. S. Ghornani, M. Rabbani, a new multi objective algorithm for a project selection problem, Advance in Engineering Software, accepted 14 march 2008-12-05.
9. Ringuest JL. , Graves SB., the linear multi objective R&D project selection problem, Eng Manage, IEEE Trans, (1989), 36-54.