Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 46 No. 2 (2025)

AI Image Generator using Generative AI for Applications

[1] Suman S, [2]Dr. N. Shunmuga Karpagam, [3]Kanchana

[1]Student, Department of CSE, Er.Perumal Manimekalai college of Engineering ,Hosur

[2]Associate Professor, Department of CSE, Er.Perumal Manimekalai college of Engineering, Hosur

[3]Assistant Professor,Dept of CSE, Er.Perumal Manimekalai college of Engineering, Hosur

Abstract

The creative industry has experienced a transformation through Generative AI in image generation because machines generate realistic or abstract visuals from written prompts. The main architectural elements for image generation systems include GANs along with Transformer-based models such as DALL-E and Stable Diffusion which successfully produce artwork that resembles authentic human creativity. The main method entails training the model through massive datasets that contain both images and text descriptions so it can detect detailed visual patterns in such data. A typical AI image generator trains deep neural networks using extensive image databases to develop knowledge which lets it understand how objects relate through textures and spatial positions. After completion of training the generator system becomes able to create novel images by converting description texts and control settings into new visual outputs.

Keywords**:** *Generative Adversarial Networks (GANs), DALL-E system, web app, Express, AI, machine learning, natural language, image generation.*

1. Introduction

How AI Image Generators Work Generative Adversarial Networks (GANs) The main architectural foundation for numerous AI image generation systems consists of two essential parts called GAN components. The generating mechanism inside the model serves to produce images through its operations. The generator system ingests input data to produce pictures that it perceives as natural. The Discriminator unit analyses images to detect real content from the dataset or artificial creations from the generator (generated by the generator). A recurrent competitive process exists between the discriminator and the generator because the generator attempts to deceive the discriminator. The generator enhances its output production while attempting to fool the discriminator and the discriminator simultaneously strengthens its skills for recognizing genuine from artificial data. The generator reaches exceptional proficiency in its ability to generate images which closely match actual visual content during its training process.

2. Applications for AI Image Generators

These artificial intelligence image generators serve multiple sectors of both industrial and creative work. Art and Creative Industries Artists and designers implement AI to produce new digital artworks and illustrations which exist as complete pieces or initiate development for further human enhancement. The combination of AI systems with Fashion and Product Design produces quick creations of clothing designs in addition to product prototypes and interior designs which draw inspiration from contemporary trends and designated themes. AI uses its capabilities to develop special marketing visual content along with logos and advertisements that businesses create promptly and economically through their systems. Entertainment and Media Artificial Intelligence systems generate virtual landscapes and characters as well as game assets for the development of video games thus improving both development speeds and reducing production expenses. Artificial Intelligence enables the production of scenes with the help of visual effects for animations that derive from scripts and creative concepts.

3. Objective:

A generative AI-based AI image generator functions as an automated system which produces visually sophisticated results from artificial intelligence algorithms that respond to different inputs. The inputs the AI model receives include textual descriptions as well as visual sketches which it uses to create images meeting precise creative artistic or functional requirements. An AI image generator based on generative AI has three main goals for creating visual content which includes diverse high-quality imagery. AI image generation targets the production of appealing and realistic or abstract visuals according to the design purpose. This involves: The AI system creates images which retain the same identifying features as original photographs together with traditional artworks including complete details and textural elements and picture composition characteristics. The training of AI models enables them to create diverse image outputs which support users to find new artistic concepts and explore different artistic expressions. Generative AI functions to both decrease production time and decrease labor needs during image creation activities. This includes AI automation delivers rapid creation of complex visual content since the image generation procedure requires less time than other traditional creative approaches. AI functions as an assistance tool for Creative professionals who need it to prototype their work and experiment with design options and create draft content for time-efficient creative development.

4. Problem statement:

The fast development of Generative Artificial Intelligence (AI) achieved notable improvements in making image generation models generate high-quality realistic and creative pictures through text descriptions and sketches. The generators can process texts and illustrations or various alternative input formats. Several unresolved problems persist with AI image generators despite their powerful capabilities and these problems exist both in technological limitations and moral issues. Generative AI-based image generators tackle automation and optimization as their main problem because they aim to streamline processes of image production which formerly needed human artistic talent. The creation of systems generating beautiful visual content demands development through minimization of user requirements combined with maximal flexibility and overcoming multiple obstacles in technical, creative and ethical aspects.

5. Literature Survey

The advancement of Generative Artificial Intelligence (AI) for image generation has brought machines to produce advanced visual content that shows high quality as well as creativity and diversity. Machines utilize generative AI to create pictures that exhibit high quality and show diverse variations along with creative elements. The use of AI in image generation primarily hinges on two key technologies: Generative Adversarial Networks (GANs) and Transformer-based models. Current research will analyze fundamental contributions in AI image development through generative AI technologies. This document explores fundamental models alongside latest developments and their practical applications and system-level obstacles and moral aspects. Generative Adversarial Networks (GANs) The GANs framework which Ian Goodfellow introduced in 2014 stands today as the primary framework for much of the research in generative image modeling. GANs include two neural networks which are known as the Generator and the Discriminator. The two components which comprise GANs are referred to as Generator and Discriminator. These operate together as competitors. The synthetic images emerge from the generator while its counterpart maintains identification skills between genuine images and artificial outputs. Noxious images put the discriminators through the test for detecting real images from false ones. The training system operates on both networks at once. The training objective aims at enhancing the generator's capacity to produce images which resemble actual images without distinction.

6. Key Developments:

The studies conducted by Karras et al from 2017 to 2020 demonstrated significant progression through their multiple papers including Progressive GANs and StyleGAN. The research investigated new methods that resulted in improved stability along with realistic performance for GANs. StyleGAN (2019), was groundbreaking for generating The system produces high-quality photorealistic faces based on its ability to learn different image representations which enable precise control of output. over generated output. Image GPT (2020) implemented the Image GPT model through a transformer structure for creating images according to OpenAI. It is based on the

The neural network operates through transformers which process sequential data like GPT-3 but it produces images rather than text. DALL·E serves as an OpenAI product which employs transformer architecture to transform written texts into visual outputs. textual descriptions. DALL·E generates visuals that keep semantic elements intact with the original textual description. text. CLIP (2021) received training as a model to interpret images together with text through its Contrastive Language-Image Pre-training method. unified space. DALL·E received its foundation from the Clip platform to transform written descriptions into picture outputs. prompts. Impact: GANs enable various applications which include face generation along with art generation and super-resolution and image-to-image translation. image-to-image translation. GANs encounter two main training difficulties which include instability and mode collapse phenomena where the generator system generates limited diversity. The model can struggle with both simple and complex or multi-object visual scenes and exhibit a restricted ability to display diverse outputs.

The Attention Mechanism (Vaswani et al., 2017) established self-attention mechanisms that completely transformed NLP and provided the basis for transformer. The development of transformer18 received basis from NLP through foundational work done by Vaswani et al. DALL-E (2021) generates a wide range of images through its transformer-based architecture that operates on textual descriptions. from textual descriptions. The system merges methods from transformers and GANs to produce remarkable results. The system produces new artistic images from sophisticated and difficult prompts such as "an armchair in the shape of an avocado." avocado"). The OpenAI platform released CLIP (2021) which performs Contrastive Language-Image Pre-training through extensive training on text-description pairs. Vision and language form common understanding through the combination of image files and written descriptions. A feature of CLIP is its ability to both process text descriptions through images which results in pattern matching. The text descriptions and guidance function for additional image models represent two capabilities of this system. A groundbreaking development in AI art creation occurred when Stable Diffusion emerged as a denoising diffusion probabilistic model (DDPM) that utilizes CLIP pre-trained models (2022). A probabilistic model known as DDPM joined with CLIP software equipped through training leads to the production of high-quality images from text descriptions. prompts. The software received widespread adoption because it comes with open-source licensing and provides users with powerful accessibility and practicality features. Impact: The emergence of Transformer-based models DALL-E and Stable Diffusion created fresh opportunities in creative image generation that let users guide image output through text commands. Users gain more control and flexibility through text-prompts which enable them to generate images using this new system. control compared to GANs. AI-generated artworks cause both academic and public debates about machine creativity while determining the impact of artificial intelligence on creative field ownership. concerning ownership and authorship of AI-generated works

7. Applications for AI Image Generation

AI image generation models have found extensive use throughout different fields of application. Some of the most prominent areas of application include: The art and design sector benefits from AI platforms DeepArt Artbreeder and RunwayML which let artists unite with AI for designing novel artistic works and creative visual arrangements. The technology enables artists to team up with AI systems to develop fresh artistic styles together with artistic visual elements and creative design work. The fashion industry makes use of artificial intelligence-generated designs for two areas: fashion computer-aided design and trend prediction and customized fashion development. GAN-based clothing design tools let users create customized fashion designs. AI models utilize advertising and marketing resources by producing targeted images which expand the range of advertising visuals at low costs. quickly and at a reduced cost. AI solutions produce banner visuals and individualized images that industry businesses regularly use. becoming common in digital marketing. AI image generation models within game development scenarios produce virtual environments and textures together with game characters. The use of video games shortens development times while cutting down production expenses. The Healthcare sector utilizes AI models of Generative Adversarial Networks for medical imaging purposes to produce synthetic medical data. AI models use medical images to train systems when combined with artificial simulations of uncommon medical conditions for research analysis.

8. System analysis existing system

AI image generation through generative AI technology has experienced fast development in the recent time span. Various highly successful systems with GANs and VAEs and Transformer components develop realistic images that retain user-provided content while creating original visual results. The following section summarizes several vital systems that exist for AI image generation. Generative Adversarial Networks (GANs) GANs represent the fundamental basis for creating AI image generation systems. Models based on GANs consist of two components: a generator alongside a discriminator so the models learn how to produce images that cannot be distinguished from genuine photographs. AI image generation benefits from essential systems that use GAN-based technology. StyleGAN and StyleGAN2 1.Developer: NVIDIA 2.Technology: GAN(Generative Adversarial Network) Key Features: StyleGAN (2018) and its subsequent version StyleGAN2 (2019) are recognized among the leading GAN architectural frameworks that specialize in high-quality face and animal image production along with other objects. StyleGAN includes a style-based generator that offers precise control for various feature elements including texture along with pose and lighting thus generating photorealistic outputs. The successor StyleGAN2 solved the image quality and training instability as well as consistency flaws which existed in its original form. Applications: 1. The system demonstrated extensive usage in producing real-looking human face images. 2. Artists together with designers have employed StyleGAN to produce art while creating landscapes in addition to building surreal imagery. 3. Virtual Character Creation serves game developers for making realistic character designs through this system.

9. Proposed system

A new AI Image Generator system built with Generative AI technologies employs GANs and Transformer-based models and Diffusion Models to develop a modern generating engine for users to produce professional-quality images from directives. The system addresses the drawbacks of current image generators by developing solutions for output control issues together with training stability, bias reduction and computational efficiency problems. The proposed system includes three crucial features that enable high-quality image generation from user input. 1. The system transforms written descriptions into images which enables users to conduct creative work like art production and design work and content development. 2. Users will gain precise management of generated output by letting them modify particular components including drawing style and basic color selection patterns and positioning objects into scenes. 3. Users will experience real-time image generation followed by refinement capabilities which enable them to modify their pictures during production until their artistic vision is achieved. 4. The system offers different image design options which range from photorealistic through abstract art to cartoon-style and artistic visual creation based on specified user preferences. 5. The proposed system features built-in mechanisms to decrease bias while upholding ethical standards which resolves issues stemming from dangerous stereotypes together with objectionable images and intellectual property violations. 6. The system employs model distillation along with quantization as optimization techniques to operate efficiently for wider accessibility to tool users.

10. Project modules

The AI Image Generator builds its operation through various key modules which handle distinct image generation responsibilities. Multiple modules work together to develop images of high quality which match user specifications and meet ethical standards. The proposed system consists of these main functional modules which follow. PURPOSE: The AI Image Generator using Generative AI generates automated images of high quality along with diversity through user-provided inputs including text descriptions or sketches and reference pictures. The combination of Generative Adversarial Networks (GANs) and Diffusion Models together with transformer-based models creates a system that transforms operations in art and design alongside the entertainment sector as well as marketing and education. This AI-powered system contains the following essential objectives which will be discussed.

11. Scope of the project:

An AI Image Generator project implemented with Generative AI includes specific boundaries for system areas of impact together with functional capabilities and user groups while relying on particular technological systems. This advanced system employs combination technology of Generative Adversarial Networks (GANs) and Diffusion Models together with Transformer-based Models as well as additional generative methods to generate images from textual prompts and sketch inputs or illustration sources. The following document presents a complete breakdown of the project requirements for the AI Image Generator system. Functional Scope Image Generation from Text (Text-to-Image): Users can enter text into the system for which the solution automatically produces corresponding visual output. The system accepts descriptions that detail particular scenes together with objects and actions as well as artistic styles. A user can input "A futuristic city skyline at sunset with flying cars" to receive an image generated according to the description. The text-to-image generation process will be powered by transformer-based models including DALL-E, Stable Diffusion, or CLIP.

12. System architecture feasibility analysis

Generative AI models built from deep learning technology have found enormous popularity for generating images in their field. The models generate authentic artistic content from input information which helps multiple domains such as artistic production, commercial promotion, entertainment industries and medical picture analysis. The study evaluates the viability of building an AI image generator with generative AI as the base technology through examination of technical execution alongside monetary assessment and operational matters.

Operational Feasibility

a. Development Team:

The realization of this project demands a team composed of professionals who excel in deep learning as well as computer vision and natural language processing disciplines. For the successful product development team members will consist of data scientists along with AI researchers in addition to software engineers together with domain experts who specialize in art or design. AI models need regular maintenance together with model updates to preserve their performance value and topicality. The system requires periodic retraining of models through updated data input to empower improved performance alongside user need adjustments.

b. User Interface:

A user-friendly design is mandatory for the interface to ensure easy accessibility. Users must find it easy to enter definitions through natural language or add reference photos to the platform. The user experience will improve by letting people change different aspects of generated images like style selection and object positioning along with color options. The AI image generator operates seamlessly inside Photoshop Illustrator and several web-based applications to create a smooth operating environment for creators. c. Ethical Considerations: The model needs training to prevent the propagation of discriminatory biases which could create stereotypes mainly during human image generation. The system needs a process of rigorous testing as well as fairness audits to function properly. The application should present complete information about AI algorithms and highlight their capacities along with their performance boundaries particularly involving image attributes. Multiple legal challenges related to AI content usage exist in specific regions because their regulatory frameworks create difficulties with advertising needs and media and intellectual property rights.

13. Analysis

Generative AI technology enables the AI Image Generator to use machine learning models which transform different varieties of input materials including text prompts and existing images as well as random noise into new visual content. Both the technical architecture with its generative model type selection and training frameworks alongside operational features for deployment methods and user interfaces and ethical requirements define the design process. Requirements Analysis When developing a technical design for an AI image generator system the fundamental requirements must be clarified first. The necessary specifications establish two primary sections which include functional requirements and non-functional requirements. Functional Requirements The system delivers high-quality pictures which accept distinct input data including text explanations and graphic conversion

as well as picture conversion and style modification processes. Users must obtain the ability to modify style elements while adjusting color options and choosing image resolution settings with other image-related variables. The system enables users to provide text or image input as well as define output limitations between abstract and realistic results. User Authentication functions should enable premium features by permitting user registration together with implementation of login systems and management of accounts. The system produces images with resolutions from 512x512 px to 1024x1024 px together with low artifact levels. The system needs to integrate with established image editing tools such as Photoshop and Canva for users to export content and continue editing in the same environment. The system needs to provide a batch generation process to create multiple images when producing projects or commercial products.

14. Non-Functional Requirements

User interaction needs the system to produce images which run either in real-time or close to it at production speeds of about 10-30 seconds. The system design requires capabilities to grow efficiently when processing numerous users at a time. The efficient management of growing scale requires using infrastructure solutions in the cloud. User data protection along with secure transactions form part of the security requirements which especially apply to premium users. The system demands a simple interface which users can operate without difficulty. All system users who are not technology experts must be able to use it through simple methods. The system needs to comply with moral standards to stop producing dangerous or prejudiced material. The system requires the ability to update and retrain models through simple procedures because technologies and user needs change over time.

15. Data flow diagram:

A DFD for an AI Image Generator based on Generative AI requires a system analysis of data flow between components and identification of major processing operations. AI Image Generator operations use three key steps that start with inputting specified data elements such as text or images before processing this content with AI models like Generative Adversarial Networks or GANs to produce outputs that can include images or videos. The following section provides an explained overview of essential elements together with DFD graphical elements and step-by-step processes leading to a graphical representation of the DFD: Components: 1. An end-user supplies text descriptions along with image inputs to the system for generating new images through the processing stage. 2. The preprocessing module contains two essential functions including data normalization through text tokenization or image resizing and optional data augmentation to increase input variety. 3. Generative Model (AI Model): o Generative Adversarial Network (GAN) or other models like VQ-VAE, StyleGAN, or Diffusion Models: The generator function of the model produces new images from the input data. The discriminator module performs quality assessment of images which the generator has produced. 4. The last stage includes the Image Refining function to reduce noise and increase resolution. The output format of images can be adjusted for user demands through specifying file types and image resolution. 5. The output includes a generated image which represents the final user result while feedback also exists as an optional feature that helps enhance model performance through user feedback or reinforcement learning.

16. Data Flow Diagram (Level 1)

• Process 1: Receive User Input The system stores user-provided text and image data in its Data Store through user interface data entry. In Process 2 the system processes raw input data which includes both text and images. The data processing yields pre-processed information of normalized text alongside tokenized text and resized and augmented images for use by the AI model input system. • Process 3: Generative Model Pre-processed data (text, image) serves as input while the generated image appears as output. Both generator weights and discriminator architecture store their static parameters as model data. • Process 4: Postprocessing The raw image produced by AI goes into the system for the AI model to generate refined output images. The storage system functions by final image preservation for both retrieval or presentation operations. Process 5 Obtains Output from Users by Showing Them Finalized Images Which Are Delivered to the User Interface.

17. Conclusion

Modern AI systems that use both GANs and VAEs convert textual information into real-looking pictures containing unique imagery. The technological progress has opened creative possibilities for professionals in design and art production as well as businesses to produce unique images at speed. OpenAI's DALL·E 3 alongside Midjourney present tools which enable users to generate unique images through user-friendly interfaces.

Technology advancement in AI-generated images creates immediate difficulties in the present era. People actively discuss ethical data usage along with biases within AI-generated materials and false associations between AI-made creations and human-made art. The use of generative technologies requires proper solutions to maintain ethical and fair practices.

References

[1] C. H. Mai, R. Nakatsu, N. Tosa, T. Kusumi and K. Koyamada, "Learning of Art Style Using AI and Its Evaluation Based on Psychological Experiments", *Entertain. Comput. – ICEC 2020*, pp. 308-316, 2020.

[2] M. Mazzone and A. Elgammal, "Art Creativity and the Potential of Artificial Intelligence", *Arts*, vol. 8, no. 1, pp. 26, 2019.

[3] S. Audry and J. Ippolito, "Can Artificial Intelligence Make Art without Artists? Ask the Viewer", *Arts*, vol. 8, no. 1, pp. 35, 2019.

[4] P. Riccio, K. Bergaust, B. Christensen-Scheel, J.-C. De Martin, M. A. Zuluaga and S. Nichele, "AI-based artistic representation of emotions from EEG signals: a discussion on fairness inclusion and aesthetics", *Proc. Polit. Mach. - Rogue Res. 2021*, pp. 1-8, 2021.

[5] V. Dignum, "Responsible Artificial Intelligence: Designing AI for Human Values", *ITU J.*, no. 1, pp. 1-8, 2017, [online] Available: http://europa.eu/rapid/press-release_SPEECH-14-421_en.htm%0A.

[6] D. Moher et al., "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement", *PLoS Med.*, vol. 6, no. 7, 2009.

[7] J.-W. Hong and N. M. Curran, "Artificial Intelligence Artists and Art: Attitudes Toward Artwork Produced by Humans vs. Artificial Intelligence", *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 15, no. 2s, pp. 1-16, 2019.

[8] C. Oh, J. Song, J. Choi, S. Kim, S. Lee and B. Suh, "I lead you help but only with enough details: Understanding the user experience of co-creation with artificial intelligence", *Conf. Hum. Factors Comput. Syst. - Proc*, pp. 1-13, 2018.

[9] S. M. C. Loureiro, J. Guerreiro and I. Tussyadiah, "Artificial intelligence in business: State of the art and future research agenda", *J. Bus. Res.*, vol. 129, no. 2020, pp. 911-926, 2021.

[10] A. Van Den, N. Kalchbrenner, O. Vinyals, L. Espeholt, A. Graves and K. Kavukcuoglu, "Conditional Image Generation with PixelCNN Decoders", *30th Conf. Neural Inf. Process. Syst*, pp. 4790-4798, 2016, [online] Available: https://papers.nips.cc/paper/6527-conditional-image-generation-with-pixelcnn-decoders.pdf.