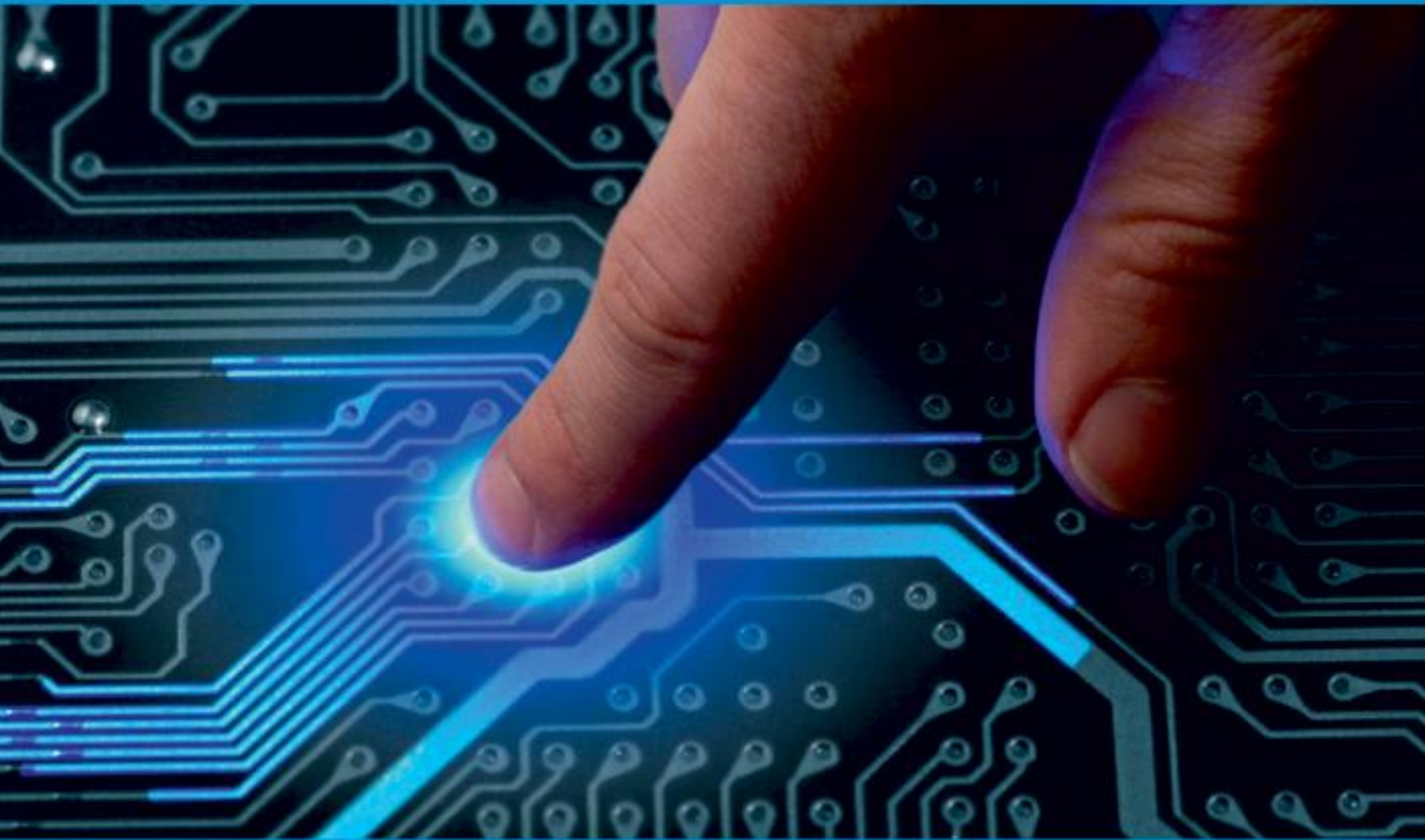
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204298 |

Fashion Classification Using Open AI & Tensorflow

Dr. Nirmala C R[1] , Chandrashekhar M V[2] , Akshatha C M[3], Prasad N Patil[4] , Rajesh K L[5], Renu R U[6]

Head of the Department, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[1]

Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[2]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[3 4 5 6]

ABSTRACT: The aim of the project is to explores the integration of OpenAI and TensorFlow for fashion classification, aiming to enhance accuracy and efficiency in image recognition tasks. Leveraging OpenAI's language capabilities, the model gains a understanding of fashion-related context. TensorFlow, a powerful machine learning framework, is employed for image processing and classification tasks. The synergy between OpenAI's natural language processing and TensorFlow's image recognition contributes to a robust fashion classification system. The research showcases the potential of combining these technologies to create a versatile and high-performing model, bridging the gap between textual and visual information in the realm of fashion analysis. Results demonstrate improved accuracy and effectiveness in classifying diverse fashion items, laying the foundation for advanced applications in the fashion industry.

KEYWORDS: Chatbot, Keyword-based processing, Artificial intelligence.

I. INTRODUCTION

The fashion industry is undergoing a significant transformation in the age of technology, much like every other sector. With the boundaries between tech and fashion becoming increasingly hazy, there is a growing need for advanced tools to handle this changing environment. The 'Fashion Classification Using TensorFlow and OpenAI' project is at the center of this convergence. It is a groundbreaking initiative that has the potential to completely transform how we view, classify, and engage with fashion in the digital sphere.

This project harnesses the power of TensorFlow, an all-in-one open-source machine learning platform that enables us to build a model for classifying fashion items. TensorFlow's flexibility and extensive range of tools make it an ideal choice for our endeavor. To enhance the richness of metadata and descriptions associated with each fashion item, we also capitalize on the expertise of OpenAI's cutting-edge language models. By doing so, we ensure that our classifications are not only visually accurate but also semantically informative.

In addition to TensorFlow, the project integrates the cutting-edge language models developed by OpenAI. This inclusion aims to enhance the richness of metadata and descriptions associated with each fashion item. By combining visual accuracy with semantic informativeness, the classification process transcends traditional boundaries, providing a holistic understanding of fashion items in the digital space. Transitioning from the written word to the visual realm, TensorFlow emerges as a powerful tool for image classification. As we embark on a deep dive into the architecture of neural networks, this chapter elucidates how TensorFlow can be harnessed to develop models capable of recognizing and categorizing diverse fashion images. From basic garment identification to intricate style patterns, the visual palette becomes the canvas upon which the AI artist paints its understanding of fashion

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204298 |

II. LITERATURE REVIEW

Hamam Mokayed, Rajkumar Saini, Hum Yan Chai. proposed a work on "**Study of AI-Driven Fashion Recommender Systems. "** in **2023.** In this Study, paper Amidst the ever-churning tide of fashion trends and exploding product landscapes, choosing what to wear can feel like an intricate game of dress-up. Traditional recommender systems, while adept at suggesting similar movies or books, falter in this domain. Fashion demands a nuanced understanding of compatibility, personal style, and cultural context, a need traditional approaches fail to grasp. Enter AI, stepping onto the virtual runway to revolutionize the shopping experience with next-generation fashion recommender systems. **[1]**.

Chenshuang Zhang. has demonstrated a work on **"AI Text-to-Image Generation in the Era of Large Model"** in **2023.** In this paper, Authors claimed that Step aside, Photoshop, there's a new creative force in town - and it's powered by words. Text-to-image generation (TTI) is rapidly transforming how they create and access images, ushering in an era where imagination takes center stage. Gone are the days of painstakingly manipulating pixels or scouring stock photo libraries; with TTI, all you need is a vivid description, and boom! - a photorealistic masterpiece materializes on your screen. Imagine conjuring up the perfect image to illustrate your novel chapter, designing a dream outfit based on your wildest fashion fantasy, or bringing historical events to life with stunningly accurate visuals.. **[2]**.

Guendalina Caldarini , Sardar Jaf and Kenneth McGarr. proposed a work on **"Recent Advances in Chatbots,** in **2022.** This paper provides an overview of the evolutionand applications Chatbots are intelligent conversational computer systems designed to mimic human conversationto enable automated online guidance and support. The increased benefits of chatbots led to their wide adoption by many industries in order to provide virtual assistance to customers. Chatbots utilise methods and algorithms from two Artificial Intelligence domains: Natural Language Processing and Machine Learning. However, there are many challenges and limitations in their application. In this survey they review recent advances on chatbots, where Artificial Intelligence and Natural Language processing are used. They highlight the main challenges and limitations of current work and make recommendations for future research investigation **[3]**.

Jason Wei, Yi Tay, Rishi Bommasani. demonstrated a work on **"Emergent Abilities of Large Language Models"** in **2022.** In this research they are Scaling up language models has been shown to predictably improve performance and sample efficiency on a wide range of downstream tasks. This paper instead discusses an unpredictable phenomenon that they refer to as emergent abilities of large language models. We consider an ability to be emergent if it is not present in smaller models but is present in larger models. Thus, emergent abilities cannot be predicted simply by extrapolating the performance of smaller models. The existence of such emergence raises the question of whether additional scaling could potentially further expand the range of capabilities of language models **[4]**.

Elad Richardson, Yuval Alaluf, Or Patashnik, Yotam Nitzan, Yaniv Azar, Stav Shapiro, Daniel Cohen-Or. demonstrated a work on **"A StyleGAN Encoder for Image-to-Image Translation"** in **2021.** In this work, they propose a novel encoder architecture that can be used to directly map a real image into the W+ latent space with no optimization required. There, styles are extracted in a hierarchical fashion and fed into the corresponding inputs of a fixed StyleGAN generator. Combining our encoder with a StyleGAN decoder, we present a generic framework for solving various image-to-image translation tasks, all using the same architecture. Notably, in contrast to the "invert first, edit later" approach of previous StyleGAN encoders, they show pSp can be used to directly encode these translation tasks into StyleGAN, thereby supporting input images that do not reside in the StyleGAN domain Additionally, differing from previous works that typically rely on dedicated architectures for solving a single translation task, we show pSp to be capable of solving a wide variety of problems, requiring only minimal changes to the training losses and methodology. We hope that the ease-of-use of our approach will encourage further research into utilizing StyleGAN for real image-to-image translation tasks **[5]**.

Mikhail Komarov , and Felix Mödritscher. proposed work on **"Image Classification for the Automatic Feature Extraction in Human Worn Fashion Data"** in **2021.** In this work, For humans, it does not take too much effort to tell apart trousers from a sweater or to recognize the outfit of a person. However, assigning features in an image to a certain category is still a hard problem to solve for computers.Images are captured everywhere. On Facebook alone, about 350 million images are uploaded every day , and many of them contain fashion objects or apparel. With the continuously increasing amount of data, it is crucial to automatically extract information out of image data. Over the last decade, the progress to address these deep learning problems has been enormous. The latest common method to understand features in images is a model called a convolutional neural network (CNN), a subtype of neural networks **[6]**.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204298 |

Mingchen Zhuge1,Dehong Gao1,Deng-Ping Fan. presented a work on **"Kaleido-BERT: Vision-Language Pre-training on Fashion Domain"** in **2021.** In this work, they propose a novel framework for the fashion-based tasks. The core idea is to focus on fine-grained representation learning and to bridge the semantic gaps between text and image. To achieve this goal, we first introduce an efficient "kaleido" strategy, which extracts a series of multi-grained image patches for the image modality. As a result, our model is named as Kaleido-BERT. This strategy is scalable and largely alleviates the aforementioned coarse presentation issue by introducing the patch-variant pre-training scheme. Furthermore, to bridge the semantic gap between different modalities, attention mechanism is employed to build pre-alignments between kaleido patches and text tokens. **[7]**.

Tingting Qiao, Jing Zhang , Duanqing Xu1 , and Dacheng Tao. proposed a work on **"Learning Text-to-image Generation by Redescription"** in **2018.** In this paperthey address the challenging T2I generation problem by proposing a novel global-local attentive and semantic-preserving text-to-image-to-text framework called MirrorGAN. MirrorGAN successfully exploits the idea of learning text-to-image generation by redescription. STEM generates word- and sentence-level embeddings. GLAM has a cascaded architecture for generating target images from coarse to fine scales, leveraging both local word attention and global sentence attention to progressively enhance the diversity and semantic consistency of the generated images. STREAM further supervises the generators by regenerating the text description from the generated image, which semantically aligns with the given text description. We show that MirrorGAN achieves new stateof-the-art performance on two benchmark datasets **[8]**.

YUPENG CHANG and XU WANG. demonstrated a work on **"A Survey on Evaluation of Large Language Models**" in **2018**. In this work they Evaluation carries profound significance, becoming imperative in the advancement of AI models, especially within the context of large language models. This paper presents the first survey to give a comprehensive overview of the evaluation on LLMs from three aspects: what to evaluate, how to evaluate, and where to evaluate. By encapsulating evaluation tasks, protocols, and benchmarks, our aim is to augment understanding of the current status of LLMs, elucidate their strengths and limitations, and furnish insights for future LLMs progression. Our survey reveals that current LLMs exhibit certain limitations in numerous tasks, notably reasoning and robustness tasks. Concurrently, the need for contemporary evaluation systems to adapt and evolve remains evident, ensuring the accurate assessment of LLMs' inherent capabilities and limitations. We identify several grand challenges that future research should address, with the aspiration that LLMs can progressively enhance their service to humanity**[9]**.

2.1 Literature Review Summary

The literature review aimed to explore Significant advances in artificial intelligence (AI) have occurred in recent years, particularly in the domains of fashion recommender systems, text-to-image generation, chatbots, large language models (LLMs), and image classification. Traditional recommender systems have struggled to meet fashion's nuanced demands, but AI-powered solutions are poised to transform the shopping experience by leveraging a deep understanding of compatibility, personal style, and cultural contexts. Text-to-image generation techniques allow for the creation of photorealistic images from detailed descriptions, transforming how images are created and consumed. Chatbots, powered by AI and natural language processing, are providing virtual assistance in a variety of industries, improving customer support and guidance. Large language models demonstrate emergent abilities, which broaden their capabilities beyond traditional predictions. Also, image classification techniques Using convolutional neural networks (CNNs), features are extracted from fashion images automatically, allowing for deeper insights into human-worn fashion data. The evaluation of large language models is also gaining popularity, with a focus on determining their strengths, limitations, and potential research challenges. These advancements collectively represent a significant step forward in AI's intersection with various domains, with the potential to reshape how we interact with technology and information.

III. EXISTING SYSTEM

In today's e-commerce and fashion design landscape, product availability filters are essential tools for users looking for specific items. These filters aim to simplify the search process by allowing users to refine their choices based on size, colour, style, and other criteria. While these filters are effective in theory, their practical application can be difficult for users with varying levels of technical knowledge. Navigating through an array of filtering options takes time and patience, which can lead to frustration for those unfamiliar with the interface. Furthermore, the sheer number of available filters can overwhelm users, making it difficult to determine which ones are most relevant to their needs. As a result, some users may be able to use these filters to find the products they want. while others may struggle to navigate the filtering options to locate their desired products.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 4, April 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1204298 |

IV. PROBLEM STATEMENT

The challenge is to create a system that can generate textual descriptions and metadata for fashion items while also accurately categorising them based on visual characteristics. This requires combining a visual processing mechanism, such as TensorFlow, with a semantic generation and comprehension system. The system must recognise the visual characteristics of fashion items and convert them into coherent textual descriptions. Furthermore, it should include the nuanced characteristics and contextual information required for accurate classification and metadata generation. This requires a sophisticated combination of image recognition and natural language understanding technologies. To provide comprehensive and meaningful descriptions of a wide range of fashion items, the system must navigate both the visual and textual domains. Addressing this challenge can significantly improve efficiency and the user experience of fashion analysis and recommendation platforms.

V. PROPOSED SYSTEM

The proposed system incorporates a variety of features, including cutting-edge image generation and conversational AI and image generation capabilities. Its implementation uses cutting-edge technologies such as OpenAI and TensorFlow. Users can easily provide descriptions, and the system, powered by these technologies, expertly processes the data to produce the desired result. This seamless integration of cutting-edge technologies improves the system's overall functionality and user experience. By combining the strengths of textual and visual processing, the system produces comprehensive results that address the diverse needs of users. Whether producing textual descriptions or visual representations, the system ensures accuracy and efficiency. It achieves remarkable performance in content understanding and generation by leveraging OpenAI and TensorFlow capabilities. and get personalized recommendations. This seamless integration goes beyond traditional text-based browsing, providing an immersive, intuitive, and tailored fashion exploration experience unrivalled in the digital world.

VI. OBJECTIVES

- **To develop robust and efficient chatbot system :** Implement a chat bot with natural language processing tailored for fashion related queries.
- **To design the state of art text to image conversion model** : Image conversion model with the focus on text encoder and image decoder the objective is to generate usually coherent and meaningful images.
- **To develop a comprehensive tool for fashion designer :** To create a comprehensive tool for fashion designer integrating features for trend analytics, fabric selection and collaboration.

VII. METHODOLOGY

This section provides a summary of the datasets, the suggested method, the structural design of the system, and the algorithms utilized for categorizing liver disease.

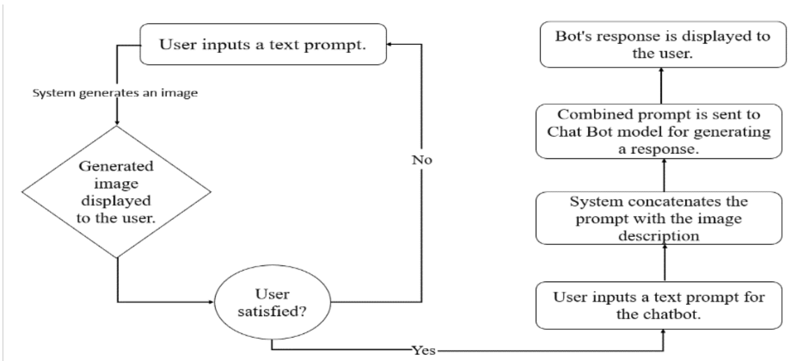


Figure 7.1: Design Frame work

The generative image model begins with receiving the text prompt, the model goes through a complex process to understand the semantics and context embedded in the input. This uses natural language processing techniques to parse the text and extract key concepts and attributes. Next, the model uses its knowledge of language and image

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 4, April 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1204298 |

relationships to convert the textual description into visual representations. This mapping process usually includes several steps:

1. Concept Analysis: The model divides the text prompt into constituent concepts and identifies nouns, adjectives, and other relevant linguistic elements.
2. Image Search: Using its knowledge base, the model searches through a large database of images for visual elements that correspond to the concepts extracted from the text prompt.
3. Image Synthesis: Using the retrieved visual elements, the model uses generative techniques like neural networks or probabilistic graphical models to create a new image that matches the textual description. This synthesis may include combining multiple image components, adjusting colours and textures, and fine-tuning details to match the input prompt.
4. Evaluation and Refinement: The generated image is evaluated to ensure consistency with the text prompt. This could include assessing the relevance of image features to the input text is iteratively refined, as is the image generation process.

The generative image model is a newly generated image that corresponds to the user's text prompt. This image encapsulates the visual interpretation of the textual description provided by the user, offering a creative and personalized representation of the input concept.

Upon receiving the inputs, the chatbot model engages in a multifaceted analysis to understand both the textual and visual components of the user's query. This analysis involves:

1. Text Understanding: The model analyses the user's original text prompt to determine its semantics, intent, and underlying context. This step may include syntactic analysis, sentiment analysis, and semantic comprehension in order to extract relevant information from the text.
2. Image Analysis: The model examines the generated image to determine its visual content, features, and potential associations with the text prompt. This analysis may include object recognition, scene comprehension, and aesthetic evaluation in order to extract insights from
image.
3. Context Integration: The chatbot creates a coherent response by combining information from both the text prompt and the generated image. . This entails creating a textual response that addresses the user's query while drawing on the visual context provided by the image. The response may include descriptions, opinions, recommendations, or other relevant information based on the user's input.

The chatbot model is a text-based response that encompasses insights from both the user's original text prompt and the generated image. This response is designed to engage with the user's query in a meaningful and contextually relevant manner, leveraging the combined textual and visual information to provide a personalized and enriching interaction experience.

VIII. RESULTS

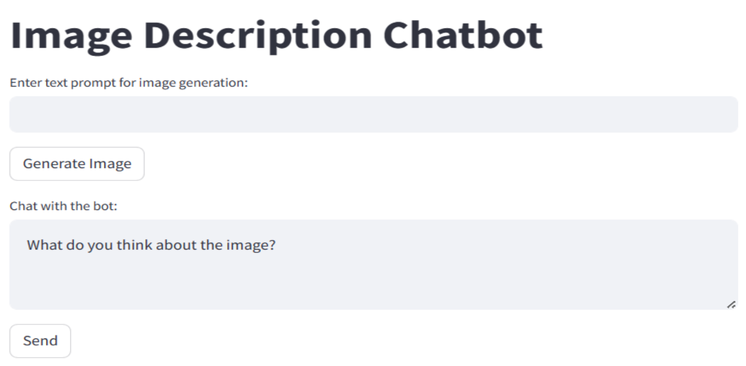


Fig. 8.1: home page of model where user can give text prompt to generate images.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204298 |

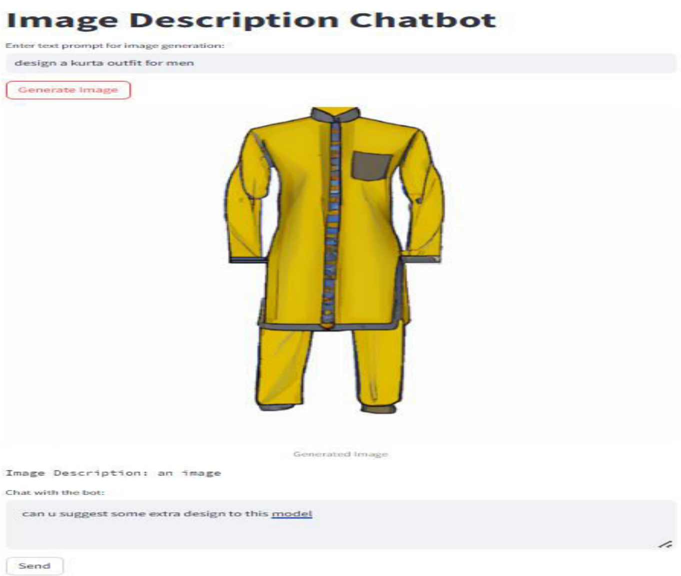


Fig. 8.2: System generates an image based on the prompt.

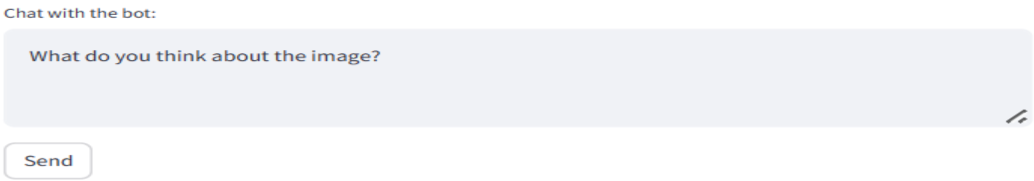


Fig. 8.3: User inputs a text prompt for the chatbot.

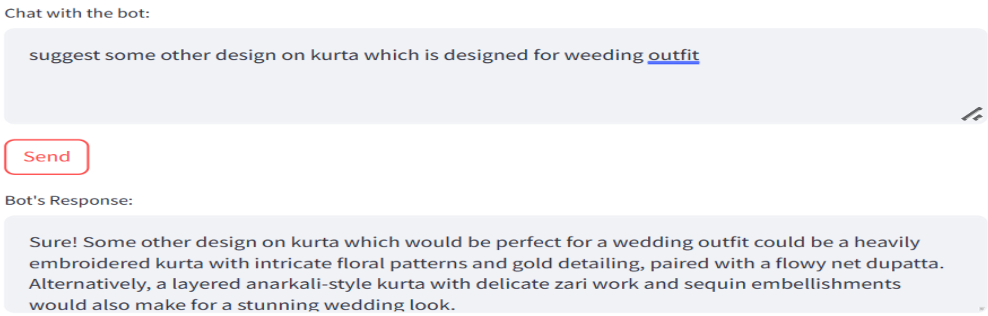


Fig. 8.4: The bot's response is displayed to the user.

REFERENCES

1. Hamam Mokayed, Rajkumar Saini, Hum Yan Chai18 May 2023, “Study of AI-Driven Fashion Recommender Systems ” International journal of creative research thoughts (IJCRT).
2. Chenshuang Zhang 14 March 2023. “The Survey into AI Text-to-Image Generation in the Era of Large Model”. Journal of Applied Artificial Intelligence.
3. Guendalina Caldarini , Sardar Jaf and Kenneth McGarr 15 Jan 2022, “Recent Advances in Chatbots”. International journal of creative research thoughts (IJCRT).
4. Jason Wei, Yi Tay, Rishi Bommasani Aug 2022, “Emergent Abilities of Large Language Models” arXiv preprint.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204298 |

5. Elad Richardson, Yuval Alaluf, Or Patashnik 2021, Yotam Nitzan, Yaniv Azar, Stav Shapiro, Daniel Cohen-Or, “A StyleGAN Encoder for Image-to-Image Translation”.
6. Mikhail Komarov , and Felix Mödritscher 16 March 2021, “Image Classification for the Automatic Feature Extraction in Human Worn Fashion Data”.
7. Paula E. Mayol, ModelJohn Paul A March 2019, “Madulid Clothing Classification using the Convolutional Neural Network Inception”.
8. Mingchen Zhuge1,Dehong Gao1,Deng-Ping Fan 2017, “Kaleido-BERT: Vision-Language Pre-training on Fashion Domain”.
9. Umar Jameell, Hashim Khan, Aqid Anwar. 31-Dec-2021. “Doctor Recommendation Chatbot: A Research Study”. Journal of Applied Artificial Intelligence.
10. Zhi-Song Liu, Li-Wen Wang, Wan-Chi Siu, and Vicky Kalogeiton. Name your style: An arbitrary artist-aware image style transfer. arXiv preprint arXiv:2202.13562, 2022.
11. Tao, M., Tang, H., Wu, S., Sebe, N., Wu, F., and Jing,X.-Y. Df-gan: Deep fusion generative adversarial networks for text-to-image synthesis. arXiv preprintarXiv:2008.05865, 2020.
12. Sun, C., Shrivastava, A., Singh, S., and Gupta, A. Revisiting unreasonable effectiveness of data in deep learning era. In Proceedings of the IEEE international conference on computer vision, pp. 843–852, 2019.
13. Razavi, A., Oord, A. v. d., and Vinyals, O. Generating diverse high-fidelity images with vq-vae-2. arXiv preprint arXiv:1906.00446, 2019.
14. Rameen Abdal, Yipeng Qin, and Peter Wonka. Image2stylegan: How to embed images into the stylegan latent space? In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 4432–4441,2019.
15. Sennrich, R., Haddow, B., and Birch, A. Neural machine translation of rare words with subword units. arXiv preprint arXiv:1508.07909, 2018.
16. Köster, U., Webb, T. J., Wang, X., Nassar, M., Bansal, A. K., Constable, W. H., Elibol, O. H., Gray, S., Hall, S., Hornof, L., et al. Flexpoint: An adaptive numerical format for efficient training of deep neural networks. arXiv preprint arXiv:1711.02213, 2018.
17. Sennrich, R., Haddow, B., and Birch, A. Neural machine translation of rare words with subword units. arXiv preprint arXiv:1508.07909, 2018.
18. Cer, D., Yang, Y., Kong, S.-y., Hua, N., Limtiaco, N., John, R. S., Constant, N., Guajardo-Cespedes, M., Yuan, S., Tar, C., et al. Universal sentence encoder. arXiv preprint arXiv:1803.11175 (2018).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com