iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.65289

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

AIGraphy: Image to Advetisement

Piyush Mallick[1], Mann Gupta[2], Manas Karadia[3], Priyanshu Jaj[4]
Computer Science, Acropolis Institute of Technology Research, Indore, India

Abstract: ***Since e-commerce is growing so quickly, excellent product photography is now crucial for High-quality product photography has become crucial for influencing consumer purchasing decisions as e-commerce has grown rapidly. Traditional photography techniques, however, are frequently expensive and time-consuming, which prevents small firms from using them and makes them difficult for larger corporations with dynamic product catalogues. This study investigates the use of generative artificial intelligence (AI) models to improve and automate product photography, providing a scalable and reasonably priced substitute. Using three essential models: ControlNet for regulated image alteration, Stable Diffusion with inpainting capabilities for backdrop generation, and Segment Anything Model (SAM) for accurate segmentation. According to our research, AI-driven product photography is a feasible answer to the demands of contemporary e-commerce since it can generate excellent images on par with conventional techniques. Our research shows that AI-driven product photography can create images of the same caliber as conventional techniques, which makes it a good option for today's e-commerce requirements. Establishing the groundwork for future research into automated, AI-based product photography and its potential to transform the production of visual content in online retail is the goal of this project.***

Keywords: Generative AI, Product Photography, E-commerce, ControlNet, Stable Diffusion, Segment Anything Model.

I. INTRODUCTION

Product photography has become a crucial element of consumer impression and intent as online buying grows in popularity. Research indicates that more than 76% of internet buyers are aware of how important high-quality photos are when making judgements about what to buy. Despite its significance, traditional product photography is frequently unaffordable and time-consuming; depending on complexity and specifications, each shot can cost anywhere from $35 to $400. These expenses pose serious obstacles for small firms, and even larger brands can find it difficult to use traditional photography techniques to keep up with the regular catalogue updates. Recent developments in Generative AI offer a creative answer to these problems in product photography, allowing companies to produce high-quality, editable photographs without requiring complex setups or a substantial financial outlay.

This strategy offers a versatile, effective, and economical substitute for conventional techniques by separating items, improving visual backgrounds, and tailoring outputs to match brand aesthetics. This AI-driven solution democratizes access to high-quality product photography and lays the groundwork for future research into AI-powered visual media in retail by enabling companies of all sizes to create visually appealing .

II. METHODOLOGY

The Hugging Face Hub serves as a robust platform providing access to a vast array of open-source models, datasets, and interactive demos, making it an ideal environment for exploring and experimenting with the latest advancements in powerful machine learning models. This project leverages the Hugging Face Hub to integrate and fine-tune three key models crucial to our AI-driven product photography pipeline: Segment Anything Model (SAM), Stable Diffusion with inpainting capabilities, and ControlNet. Together, these models form the core of a streamlined, automated workflow that enables the generation of high-quality, customizable product images.

A. SAM (Segment Anything Model)

As in Meta AI Research's Segment Anything Model (SAM) introduces an advanced solution for image segmentation through a prompt able approach. This model addresses segmentation by generating a precise mask based on prompts, such as points or bounding boxes, which specify the areas or objects in an image that need to be isolated. SAM's architecture consists of an image encoder, a prompt encoder, and a mask decoder, with a Vision Transformer (ViT) at its core, ensuring detailed image feature capture. SAM processes each image with these prompts, generating a segmentation mask with probabilities for each pixel

©IJRASET: All Rights are Reserved

B. ControlNet

ControlNet provides additional control over the image generation process in models like Stable Diffusion. By duplicating the U-Net encoder weights into trainable and locked versions, ControlNet enhances precision in generating images based on new prompts, allowing users to modify visual elements without impacting the original model's core knowledge. This makes it ideal for creating tailored visuals in product photography applications.

C. Stable Diffusion + Inpainting

The Stable Diffusion model enables the creation of high-quality visuals based on input descriptions, using a text encoder, U-Net, and autoencoder (VAE) to generate images from text prompts. The model begins with Gaussian noise and iteratively refines it to produce a clear image. Our project employs stable-diffusion-2-inpainting, which supports inpainting and out painting functionalities, effectively generating new image regions that integrate smoothly with the existing product image

III. APPLICATIONS

A. E-Commerce and Retail

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, SC, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Real Estate and Interior Design:

AI-driven image generation provides virtual staging and enhanced visualization options for real estate listings, reducing costs and making spaces more marketable. Interior designers can present rendered room settings tailored to client preferences, expediting decision-making

C. Advertising and Marketing Agencies

Marketing agencies can use AI to create tailored campaigns for different audience segments, adapting visuals for social media or specific product placements, thus improving campaign relevance and audience engagement.

D. E-Learning

AI-generated images provide accurate visuals for educational content in subjects like science or geography. E-learning platforms can use AI to create custom illustrations and engaging classroom backgrounds, enriching the learning experience.

E. Gaming and VR

AI-driven tools streamline the creation of immersive backgrounds and objects in games and VR, supporting real-time interaction and personalization. This improves user engagement by allowing for extensive customization of in-game elements

IV. ACKNOWLEDGMENT

The authors would like to thank the research teams at Meta AI, CompVis, Stability AI, and LAION for their contributions to open-source AI models and datasets, which played a crucial role in this study. Special appreciation goes to the Hugging Face community for providing access to resources, models, and tools that facilitated the fine-tuning and development of the AI-driven product photography pipeline. The authors also extend gratitude to colleagues and mentors for their valuable insights and encouragement throughout the research process.

REFERENCES

References given below:

[1] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, A., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W. Y., & Dollár, P. (2023). Segment Anything. Meta AI Research. from the Meta Research website.

[2] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. In CVPR 2022

[3] Zhang, Z., Sun, J., Gao, H., & Wang, Z. (2023). ControlNet: Adding Conditional Control to Text-to-Image Diffusion Models. Available at ControlNet GitHub.

[4] Hugging Face. (2023). Hugging Face Model Hub [Online]. Available: https://huggingface.co.

[5] J. Wang, H. Xu, and M. Xiao, "AI and e-commerce: The role of artificial intelligence in the transformation of retail," International Journal of Retail & Distribution Management

[6] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., & Brew, J. (2020). Transformers: State-of-the-Art Natural Language Processing.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

Copyright Transfer and Declaration form

The undersigned authors hereby transfer all rights, titles, and interests in the above-titled article (the "Article") to [Journal/Publisher Name]. This includes the exclusive rights to reproduce, publish, distribute, and sell the Article in any format or media, whether now known or hereafter developed.

This transfer of copyright applies only if the Article is accepted for publication. If the Article is not accepted, this agreement is null and void, and the copyright remains with the authors.

Author Warranties and Responsibilities

The undersigned authors warrant that:

The Article is an original work and has not been published elsewhere.

All necessary permissions have been obtained for any copyrighted material included in the Article.

The Article does not infringe on any third-party rights, including copyrights, trademarks, patents, or trade secrets.

The authors have the authority to enter into this agreement and that all co-authors have agreed to the terms.

The undersigned authors agree to indemnify and hold harmless [Journal/Publisher Name] from any claims or damages resulting from a breach of these warranties.

Author's Profile



Piyush Mallick is a Computer Science student at Acropolis Institute of Technology and Research in Indore, India. His academic interests lie in artificial intelligence and its applications in digital solutions. Piyush is the corresponding author for this research paper and can be reached at piyushmallick001@gmail.com.



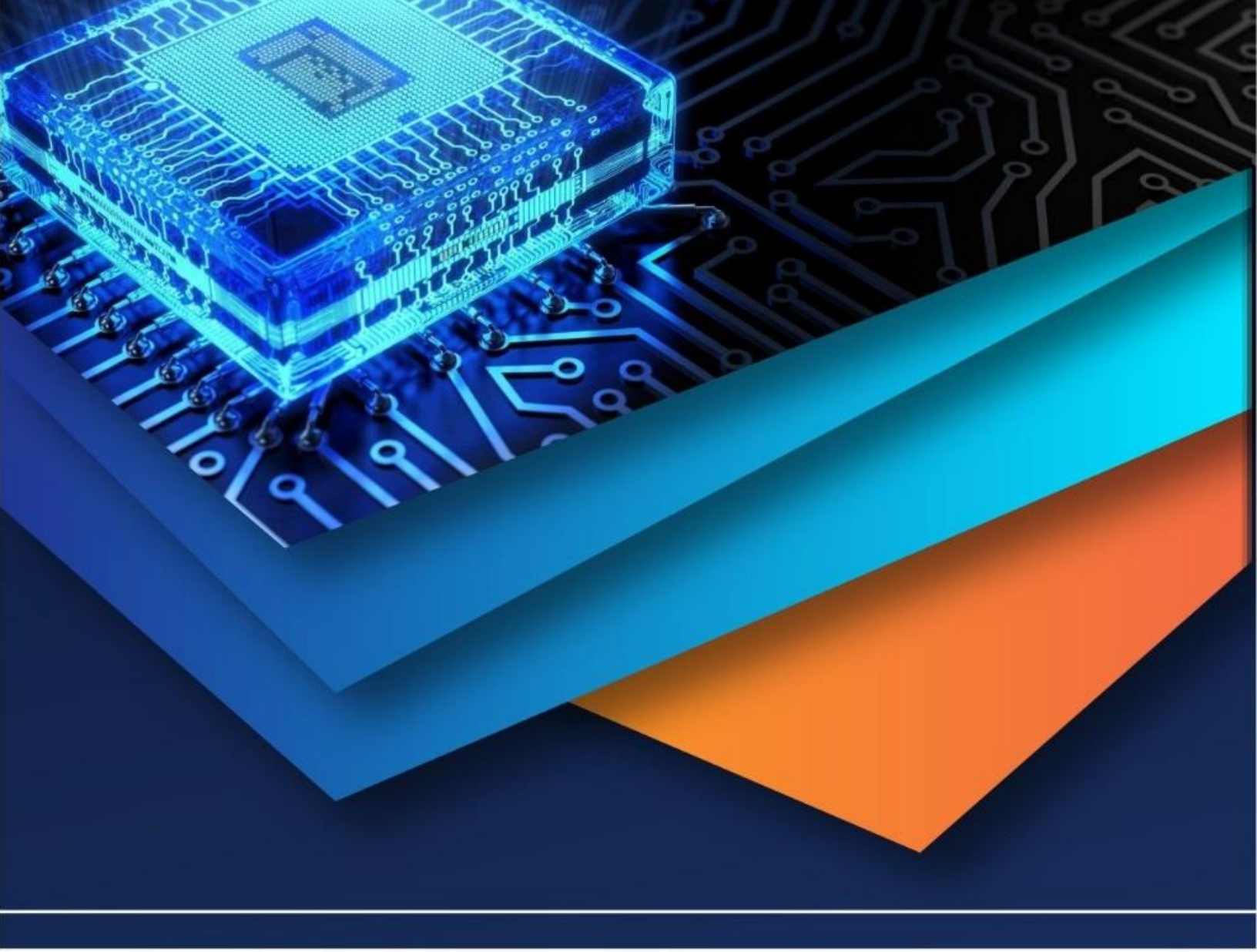
Mann Gupta is a Computer Science student at Acropolis Institute of Technology and Research, Indore, India. His areas of focus include AI and software development for enhanced user experience in the digital domain. Mann's email is **guptamann108@gmail.com.**



Priyanshu Jaj is pursuing a degree in Computer Science at Acropolis Institute of Technology and Research, Indore, India. He has a keen interest in e-commerce technologies and AI-driven systems for practical implementations. Priyanshu can be contacted at priyanshujaj11@gmail.com .



Manas Karadia is a Computer Science student at Acropolis Institute of Technology and Research, Indore, India. He is passionate about AI applications and technology innovation. Manas can be reached at manaskaradia@gmail.com.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)