

# AI-Powered Multimodal Marketing Content Generator with Customizable Product Visualization

Rohith Sai Charan Eanugula | Vinay Joshva | Hrishikesh Komaragiri | Sunil Reddy Janapala

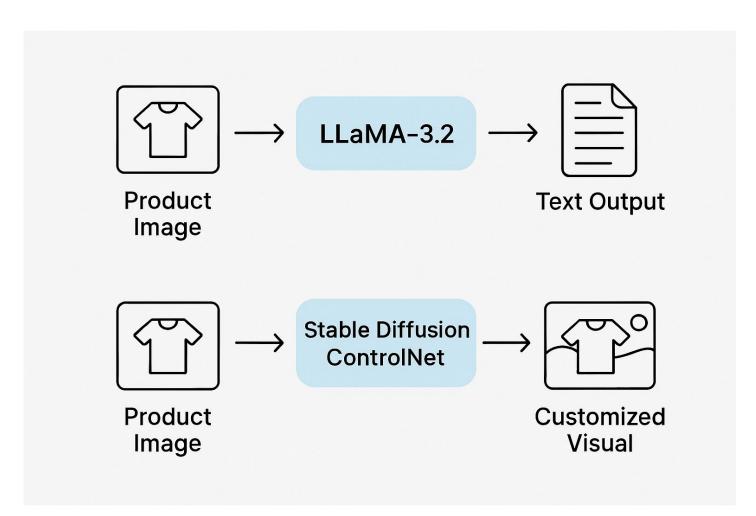
Master of Science in Computer Science

University of Missouri Kansas City, United States

## Introduction

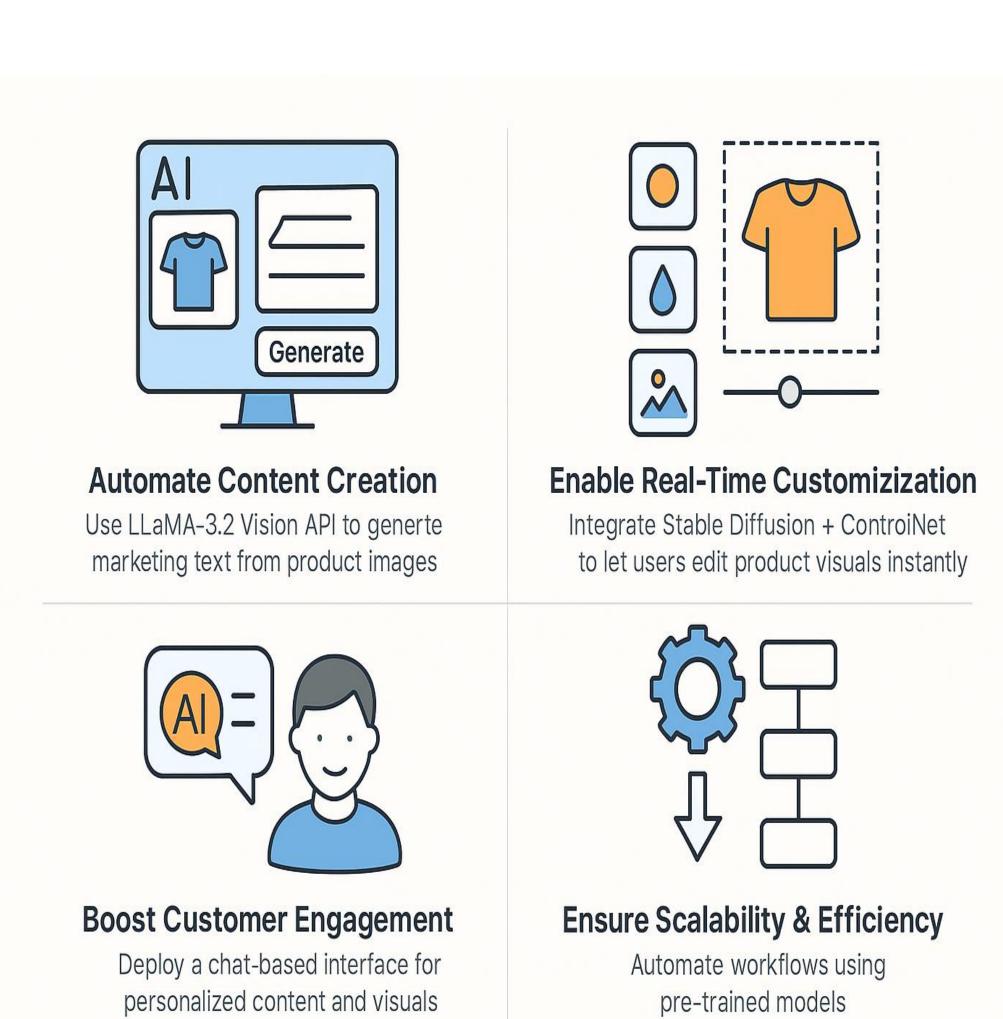
Creating personalized marketing content for large product catalogs is a major challenge in e-commerce manual photography and copywriting are timeconsuming, expensive, and slow.

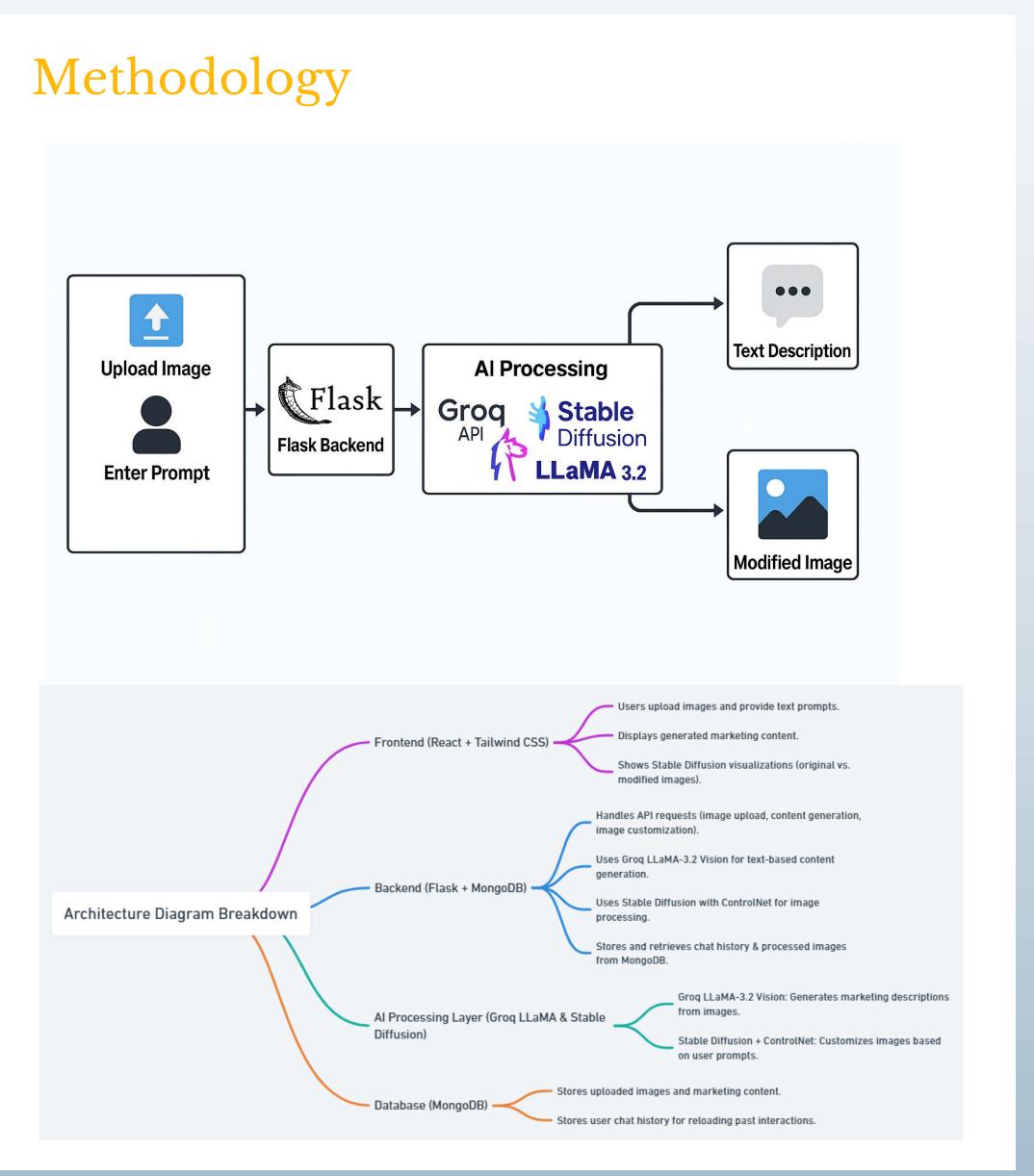
To solve this, we developed an Al-Powered Multimodal Marketing Content Generator with customizable Product Visualization that automates both text and visual content creation.

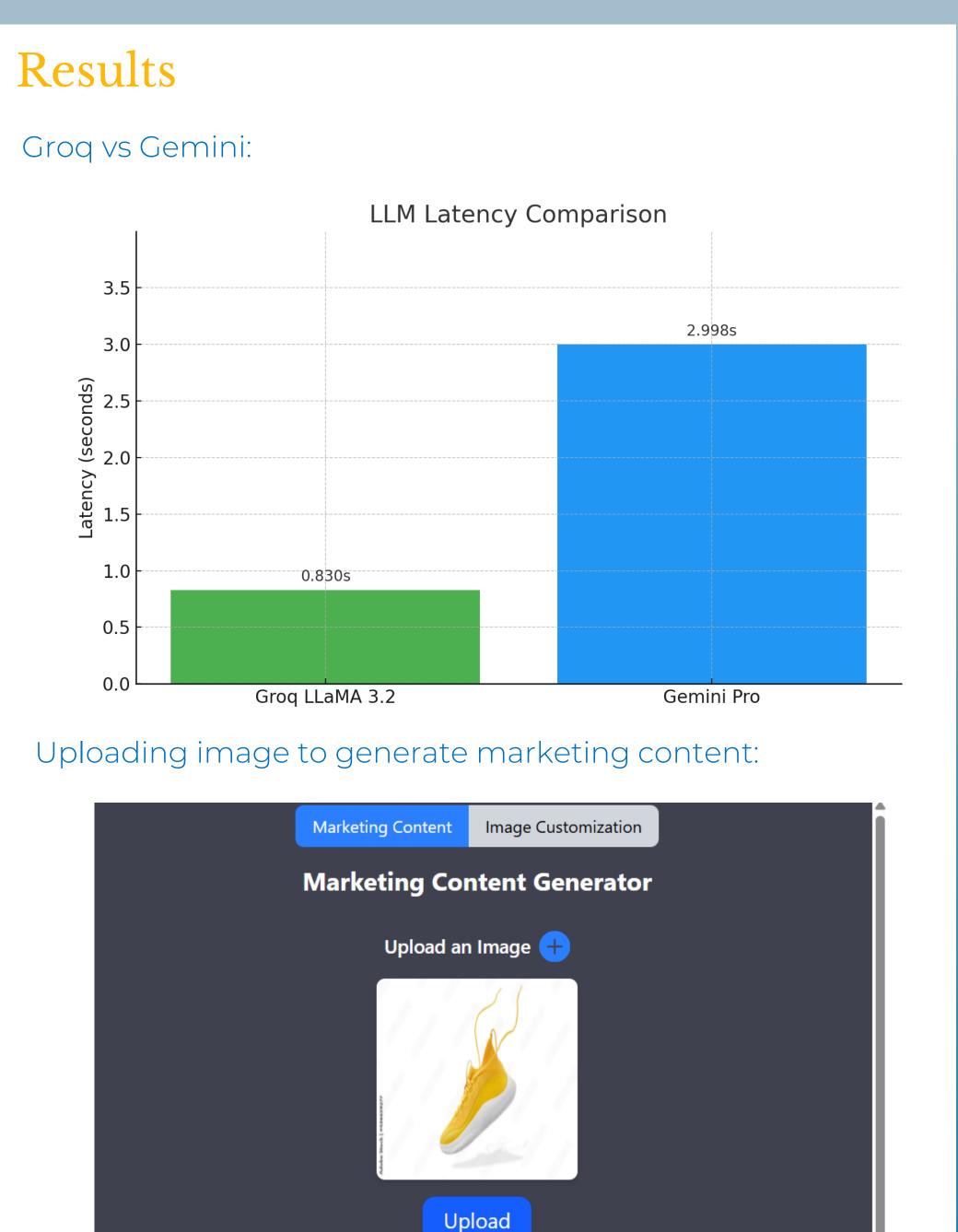


Our solution cuts costs, speeds up product rollouts, and enhances the customer experience with scalable, Aldriven marketing content.

## Objectives





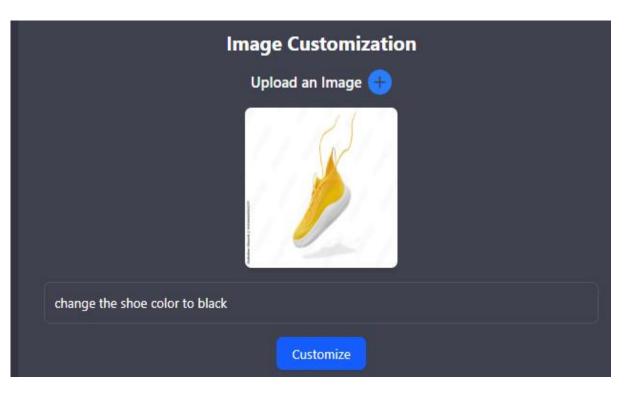


## Generated content and Chatting Interface: "Adobe Stock | #436659277". \*\*Product Description\*\* The image showcases a vibrant yellow sneaker with white soles and matching element to the overall visualization. \*\*Marketing Content\*\* This image is ideal for marketing materials related to fitness, sports, and overall atmosphere of the image is fun and carefree, suggesting a This image can be used on social media, websites, and ads that promote running, training, or any activity that implies moving and You: give me a tagline for the social media pos Al: Here are a few tagline options for the social media post: 1. "Kickstart Your Day" - This tagline captures the energetic vibe of the image and encourages users to start their day on a positive note. 2 referencing both the shoe's sole and the idea of finding like-minded emphasizes the fashionable aspect of the shoe and encourages users sense of freedom and joy, matching the dynamic movement of the shoe in the image. 5. "Walk on Sunshine" - This tagline captures the bright and cheerful atmosphere of the image and encourages users to walk into their day with positivity. Choose the one that resonates

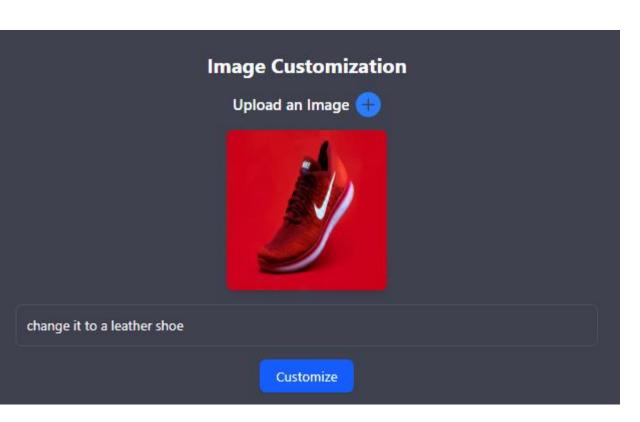
#### Customization of shoe color:

Enter prompt for Al..

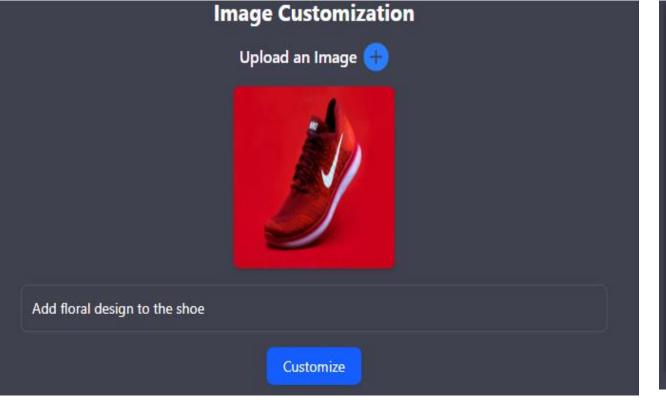
with your brand's tone and message.

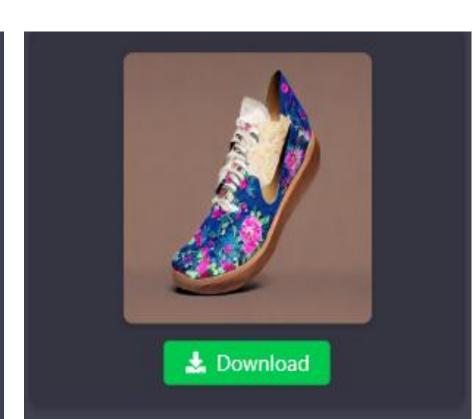






Customization of shoe design to floral:





▲ Download

Send

▲ Download

## Conclusion

This project showcases how multimodal Al can transform e-commerce by automating marketing content creation. Using Groq's LLaMA-3.2 Vision for text generation and Stable Diffusion + ControlNet for real-time image customization, our system generates product descriptions from images and enables interactive visual edits.

Key benefits include:

- Reduced content creation time and cost
- Enhanced customer engagement
- Scalable, personalized marketing workflows

#### Innovations:

Integration of vision-language and image synthesis models for end-to-end content generation.

#### Limitations:

- No fine-tuning due to GPU constraints
- No quantitative evaluation metrics used
- May require domain-specific customization

Despite these, the architecture is future-ready supporting LoRA fine-tuning, GANs, and RAG-based personalization, offering a scalable solution for Aldriven digital marketing.

### Future Works

- Fine-Tuning Stable Diffusion on Fashion Datasets
- 2. Integrate LoRA for Efficient Model Adaptation
- 3. Add Retrieval-Augmented Generation (RAG)
- 4. Expand Explainability & Metrics

## References:

High-Resolution Image Synthesis with Latent Diffusion Models Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B.

This paper introduces latent diffusion models, which form the backbone of Stable Diffusion, offering efficient high-resolution image synthesis.

https://arxiv.org/abs/2112.10752

Diffusion Models Beat GANs on Image Synthesis Prafulla Dhariwal and Alex Nichol (2021).

This paper demonstrates that diffusion models can outperform GANs in generating high-quality images, highlighting the advantages of diffusion-based approaches in image synthesis. https://arxiv.org/abs/2105.05233

Denoising Diffusion Implicit Models Jiaming Song, Chenlin Meng, and Stefano Ermon (2020). This work introduces an implicit formulation of diffusion models that improves sampling efficiency, providing key insights into the mechanics of denoising diffusion processes. https://arxiv.org/abs/2010.02502