Visual Content Generation Guided by Product Branding

Fong Bi Qi, Dr. John See

INTRODUCTION

Brand visual identity builds customer recognizability, trust, and emotional connection. Notable advancements have been made with the advent of state-of-the-art models like DALL-E, MidJourney, and Stable Diffusion. Leveraging the power of generative AI, designers can streamline the design process, expand their creative horizons, and deliver visually appealing designs that strengthen brand identity.

OBJECTIVES

- to introduce a user-oriented framework that can generate marketing graphics that incorporate unique and consistent brand visual identity
- to develop an application that implement the framework

METHODOLOGY

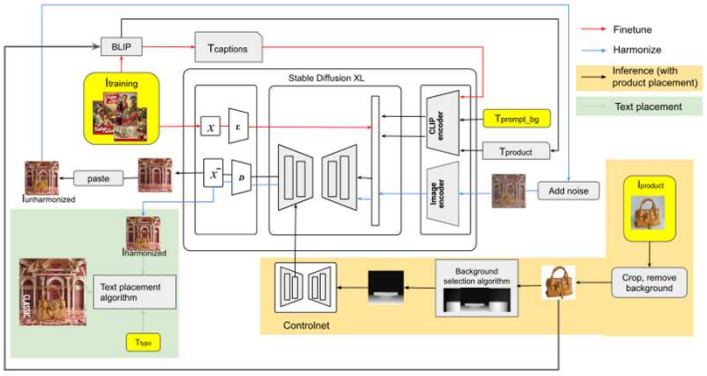

Figure 1. Proposed framework

- A pretrained SDXL is the backbone of the framework.
- **Finetuning:** BLIP generates captions for the 20-40 user-uploaded training images. The image-caption pairs finetune the U-Net of SDXL using the DreamBooth-LoRA method.
- **Product Placement and Background Generation:** ControlNet is connected to the U-Net decoder (Lvmin Zhang, 2023), which takes the best-fitting depth map and user prompt to generate the background for product placement.
- **Image harmonization with SDEdit method**: The product image is resized and pasted on the generated background. Composite image is noised then undergo a backward diffusion process for harmonization (Chenlin Meng, 2021).
- **Text Placement**: Performs saliency detection on generated images and places text on the largest non-salient region.

RESULTS

We conducted a quantitative evaluation using the CLIP score (prompt to image relevancy) and LAION aesthetic score (aesthetic quality), and a qualitative evaluation with Marketing professionals (group A) and the general public (group B).

	VintageRed	PinkVector	Baseline
CLIP score (/100)	20.29	21.57	19.83 *
LAION score (/10)	6.39	5.67	5.4 **

* obtained from images generated by SDXL vanilla.
** obtained from 55 Canva posters without typography.

Table 1. Quantitative Evaluation

	Group A	Group B
General Image Quality	3.71	4.03
Brand Style Preservation	4.36	3.92
Product Placement and Background Generation	3.67	3.4
Typography	3.43	3.69

Rating is in a scale of 1 (worst)-5 (best)

Table 2. Human Evaluation

Figure 2. Generated image samples

CONCLUSION

DreamBooth-LoRA finetuning method on SDXL can effectively reflect brand style in generated images. Product placement and background generation tasks can be accomplished by pipelining ControlNet depth, BLIP and SDEdit method along with template match algorithm.

Acknowledgment

Deep appreciation to supervisor Dr John See, the HuggingFace community, my dearest friend and family.

References (40 points, bold)

Lvmin Zhang, A. R. and Agrawala, M. (2023). Adding conditional control to text-to-image diffusion models. *arXiv:2302.05543v2*

Chenlin Meng, Yutong He. (2021). Sdedit: Guided image synthesis and editing with stochastic differential equations