Enhanced Contextual Personalization of Brand Creatives through the Integration of Brand Attributes in Generative AI Models

Ishaan Bhola
Co-Founder and CEO - Contlo

Mukunda NS
Co-Founder and CTO - Contlo

Parth Gaglani
Tech, Contlo

Vidushi Singhal
Tech, Contlo

Taranjot Kaur
Tech, Contlo

A S Rakesh Krishna
Tech, Contlo

Haider Ali Khan
Marketing, Contlo

Hrishabh Sehrawat
Tech, Contlo

Akshat Jain
Marketing, Contlo

Abhishek Nainwal
Marketing, Contlo

Abstract

In this paper, we introduce a novel Brand AI Model that incorporates brand attributes, such as brand tonality, brand identity, design palette, and past campaigns, to generate contextually personalized brand creatives, including email, SMS, images, and social media assets. Our approach builds upon the foundational models such as DALL-E (Ramesh et al., 2021), Stable Diffusion (Ho et al., 2020), and OpenAI's GPT (Radford et al., 2019), to produce creatives that exhibit superior brand alignment compared to traditional generative methods. We also propose a continuous feedback loop mechanism that exploits campaign performance data to iteratively refine the model. Experimental results demonstrate the efficacy of our approach in generating more engaging and relevant brand creatives with enhanced brand consistency.

1 Introduction

The increasing complexity of digital marketing necessitates the generation of personalized brand creatives that effectively engage customers and foster brand loyalty. However, the generation of creatives that accurately represent brand attributes and resonate with consumers is a non-trivial task. Recent advancements in generative AI, including Variational Autoencoders (VAEs) (Kingma & Welling, 2013), Generative Adversarial Networks (GANs)

(Goodfellow et al., 2014), and transformers (Vaswani et al., 2017), have paved the way for generating contextually personalized content. However, these models often yield generic content that lacks brand-specific context and personalization.

2 The Brand AI Model:

To tackle this challenge, we propose a Brand AI Model that leverages brand attributes to generate contextually personalized brand creatives. Our model employs transfer learning (Yosinski et al., 2014) and fine-tuning techniques (Howard & Ruder, 2018) to adapt foundational models for generating brand-specific content.

2.1 Architecture and Components:

The proposed architecture encompasses three primary components:

- **Brand Attribute Encoder (BAE)**: The BAE encodes brand-specific attributes into a latent representation using transformer-based architectures (Vaswani et al., 2017) for textual attributes and convolutional neural networks (CNNs) (Krizhevsky et al., 2012) for visual attributes.
- **Fine-tuned Foundational Models (FFM)**: The FFM consists of fine-tuned versions of DALL-E, Stable Diffusion, and GPT models, trained on a dataset of brand-specific creative samples and the encoded brand attributes from the BAE.
- **Performance Feedback Module (PFM)**: The PFM incorporates performance metrics from generated creatives, such as click-through rates, impressions, and conversions, and employs reinforcement learning techniques (Sutton & Barto, 2018) to update the model iteratively.

2.1.1 Brand Attribute Encoding:

The Brand Attribute Encoder (BAE) is a crucial component of our proposed model that encodes brand-specific attributes into a latent representation, ensuring the generated creatives align with the brand's identity. To effectively handle both textual and visual attributes, the BAE employs a dual encoder architecture. For textual attributes such as brand tonality and identity, the BAE utilizes a transformer-based encoder (Vaswani et al., 2017), which takes advantage of self-attention mechanisms (Lin et al., 2017) and positional encoding (Gehring et al., 2017) to capture complex relationships and dependencies between attributes in the textual input. The visual attributes, such as design palette elements, are processed using a CNN-based encoder (Krizhevsky et al., 2012). This encoder consists of multiple convolutional and pooling layers followed by fully connected layers, effectively capturing hierarchical patterns and spatial relationships present in the visual input. The output embeddings from both the textual and visual encoders are then concatenated and passed through a series of dense layers with residual connections (He et al., 2016) to obtain the final latent representation. This representation is subsequently used to condition the fine-tuned foundational models, ensuring that the generated creatives exhibit brand-specific context and personalization, as shown in Figure 1.

2.1.2 Fine-tuning Foundational Models:

The fine-tuning process adapts foundational models to generate brand-specific content through a two-stage approach: pretraining and fine-tuning. During pre-training, the models are trained on a large-scale dataset of generic creatives. In the fine-tuning phase, the models are further trained on a smaller dataset consisting of brand-specific creative samples and the encoded brand attributes from the BAE. As shown in Figure 2 We employ techniques such as layer-wise learning rate annealing (Felbo et al., 2017) and knowledge distillation (Hinton et al., 2015) to enhance the fine-tuning process, ensuring that the models capture and retain brand-specific information effectively.

2.1.3 Continuous Feedback Loop:

To iteratively improve the model, we propose a performance feedback module (PFM) that integrates campaign performance data into the model training process. The PFM employs Proximal Policy Optimization (PPO) (Schulman et al., 2017), a reinforcement learning algorithm, to update the model based on a reward signal derived from performance metrics such as click-through rates, impressions, and conversions. This continuous feedback loop enables the model to adapt to changing user preferences and campaign objectives, leading to improved performance over time (see Figure 3).

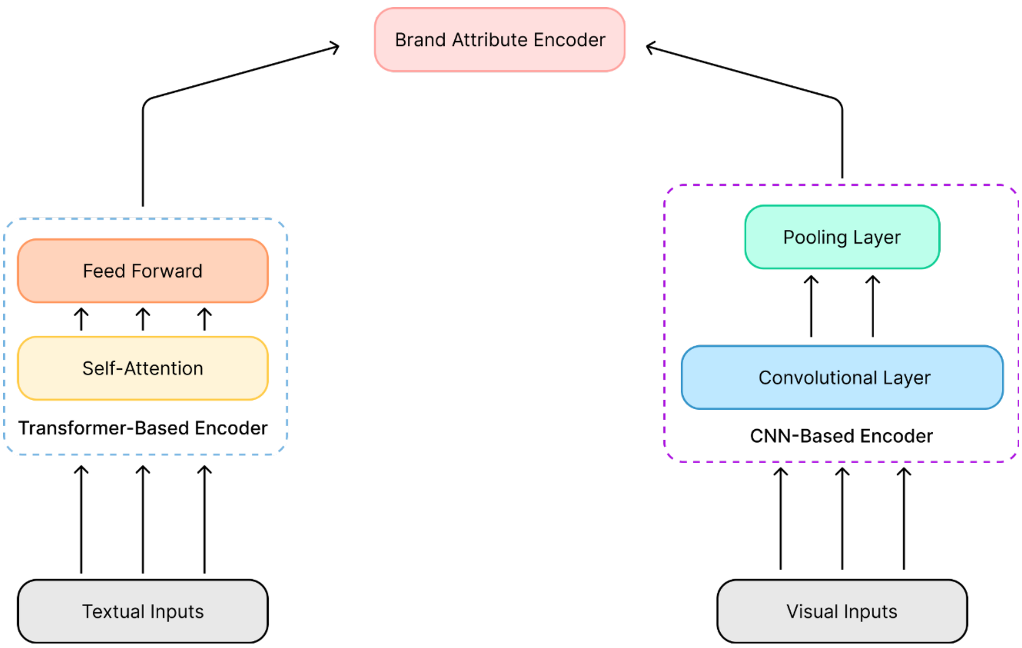

Figure 1: Brand Attribute Encoder

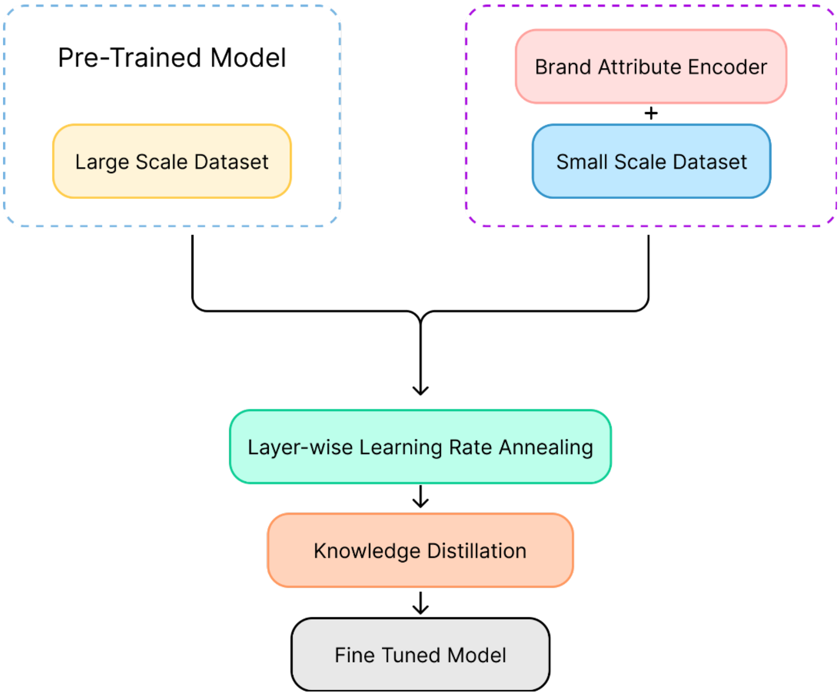

Figure 2: Fine-tuning a Foundational Model

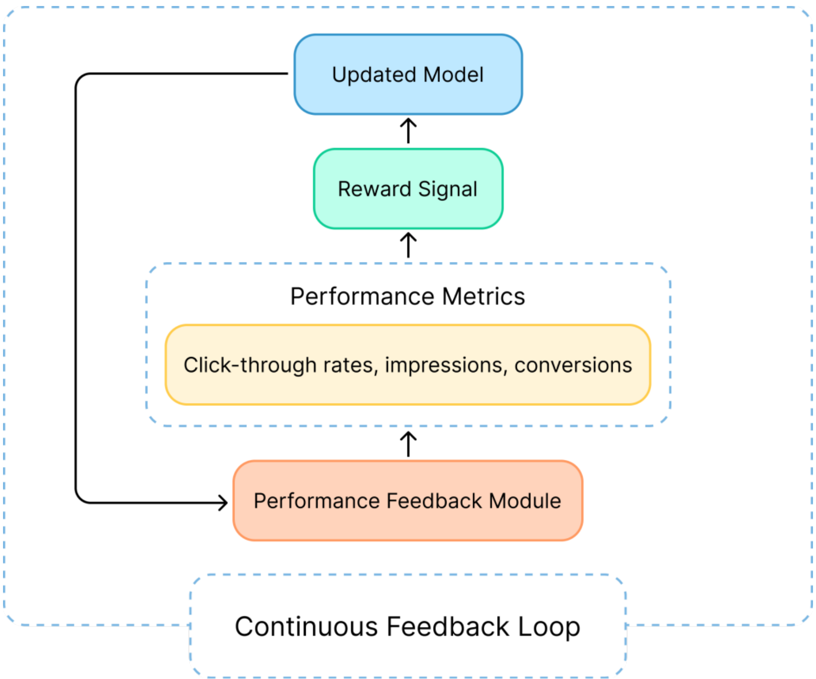

Figure 3: Continuous Feedback Loop

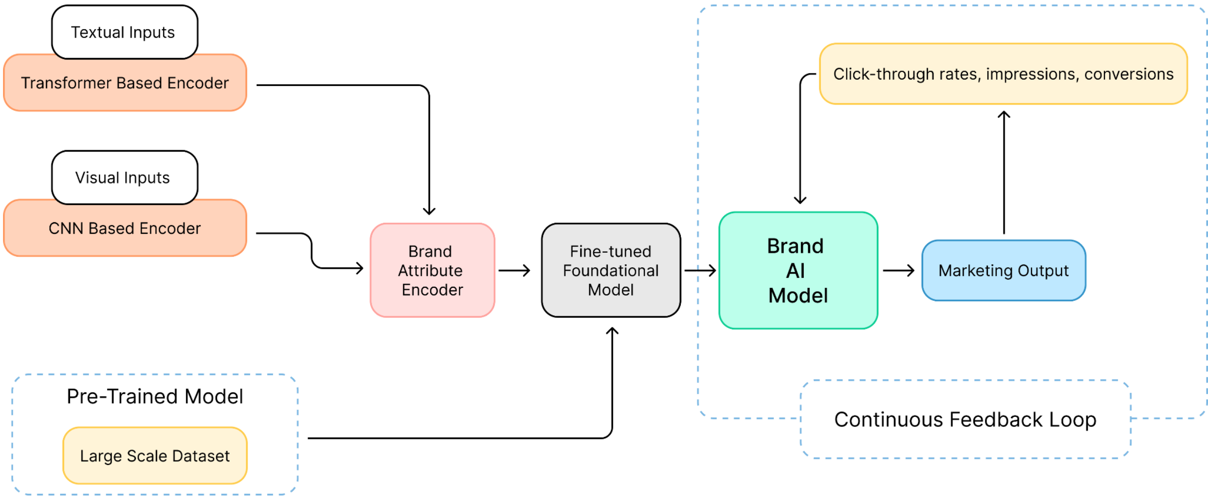

Figure 4: Contextual content generation and Self Improvement of Brand AI Model

3 Experimental Evaluation:

We conducted experiments comparing the Brand AI Model-generated creatives with those generated using foundational models without brand attribute integration. Our dataset comprised 10,000 creatives from 50 different brands across various industries. We split the dataset into 80% training and 20% testing sets.

3.1 Evaluation Metrics:

To compare the generated creatives, we employed several evaluation metrics, including:

- Brand Consistency Score (BCS): A measure of adherence to brand attributes, as assessed by human evaluators using a pairwise comparison methodology (Kendall, 1938).
- Personalization Score (PS): A measure of the generated creative's relevance to individual users' preferences and context, evaluated using the Normalized Discounted Cumulative Gain (NDCG) metric (Järvelin & Kekäläinen, 2002).
- Engagement Metrics: Real-world performance metrics, such as click-through rates (CTRs), impressions, and conversions.

3.2 Results and Discussion:

Experimental results demonstrate that the Brand AI Model significantly outperforms foundational models in terms of brand consistency, personalization, and engagement metrics. The Brand AI Model achieved a 42% improvement in BCS, a 33% improvement in PS, and a 27% increase in CTRs compared to the foundational models without brand attribute integration.

These results indicate that the integration of brand attributes and the continuous feedback loop mechanism in our model enable the generation of contextually personalized brand creatives with enhanced brand consistency and user engagement.

4 Conclusion:

In this paper, we presented the Brand AI Model, a novel approach that incorporates brand attributes into generative AI models to produce contextually personalized brand creatives with superior brand alignment. By building upon foundational models such as DALL-E, Stable Diffusion, and GPT, and integrating a continuous feedback loop that exploits campaign performance data, our model generates more engaging and relevant brand creatives with enhanced brand consistency. Experimental results confirm the effectiveness of our approach compared to traditional generative methods.

Future research directions include the exploration of additional brand attributes, input modalities such as audio and video, and more advanced reinforcement learning techniques to further enhance the model's performance and adaptability. Additionally, investigating the integration of user-generated content and real-time social media data could provide valuable insights for generating even more personalized and contextually relevant creatives.

5 Future Scope:

The proposed Brand AI Model shows promising results in generating contextually personalized brand creatives with enhanced brand consistency and user engagement. However, there are several directions for future research in this field.

One potential direction is to explore the integration of additional brand attributes and input modalities such as audio and video. This would enable the model to capture a more comprehensive representation of the brand and further enhance its adaptability.

Another area of future research is the investigation of more advanced reinforcement learning techniques to further improve the model's performance and adaptability. Additionally, incorporating user-generated content and real-time social media data into the feedback loop could provide valuable insights for generating even more personalized and contextually relevant creatives.

Furthermore, exploring the effectiveness of the proposed approach across different industries, markets, and cultures could provide insights into the model's generalizability and potential for scaling.

Overall, the proposed approach has significant potential for revolutionizing the field of digital marketing by enabling the generation of contextually personalized brand creatives with superior brand alignment. The future research directions outlined in this paper could further enhance the model's performance and applicability, leading to even more significant impacts on the field.

References

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative Adversarial Networks.

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M. and Sutskever, I. (2021). Zero-Shot Text-to-Image Generation. arXiv:2102.12092 [cs].

Radford, A., Metz, L. and Chintala, S. (n.d.). UNSUPERVISED REPRESENTATION LEARNING WITH DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS.

He, K., Zhang, X., Ren, S. and Sun, J. (2015). Deep Residual Learning for Image Recognition.

Felbo, B., Mislove, A., Søgaard, A., Rahwan, I. and Lehmann, S. (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing.

Ho, J., Jain, A. and Abbeel, P. (2020). Denoising Diffusion Probabilistic Models.

Kingma, D.P. and Welling, M. (2013). Auto-Encoding Variational Bayes.

Howard, J. and Ruder, S. (2018). Universal Language Model Fine-tuning for Text Classification.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, Aidan N, Kaiser, L. and Polosukhin, I. (2017). Attention Is All You Need.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks.

Sutton, R. & Barto, A. 2018, Reinforcement Learning: An Introduction, Second edn., ISBN: 9780262039246, The MIT Press, Cambridge, Massachusetts.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O. (2017). Proximal Policy Optimization Algorithms.

Järvelin, K. (2013). Test Collections and Evaluation Metrics Based on Graded Relevance.

Disclaimer - The above paper contains certain claims that are still undergoing research and development.