AI-Powered Banner Generation: Integrating GANs, LayoutDETR, and RetrieveAdapter for Contextual Marketing Designs

Abstract

This paper presents an AI-powered banner generation system designed to automate the creation of visually compelling and contextually relevant promotional banners. Our approach integrates Generative Adversarial Networks (GANs) with advanced layout analysis models—specifically, LayoutDETR and RetrieveAdapter—to intelligently arrange text and design elements based on user inputs and specific thematic requirements. The system features a responsive NEXT.js frontend coupled with a Flask backend, ensuring a smooth and interactive user experience. In addressing challenges such as model integration, data scarcity, and performance optimization, we employ techniques including transfer learning and zero-shot learning. Experimental results indicate that the proposed system not only automates the creative design process but also adapts effectively to diverse marketing themes, setting the stage for future extensions into video and 3D banner generation.

1. Introduction

Digital marketing increasingly demands dynamic and aesthetically appealing promotional content that can capture consumer attention quickly. Traditionally, banner design has been a labor-intensive process, relying on the expertise of designers to create content that is both visually attractive and contextually appropriate. As marketing campaigns grow in scale and complexity, there is a pressing need for automated systems that can generate high-quality designs efficiently.

In this work, we propose an AI-driven banner generation system that leverages recent advances in deep learning to address this need. Our system is built on the foundation of GANs, which are well known for their ability to produce high-fidelity images. We augment these capabilities by integrating LayoutDETR and RetrieveAdapter—models originally developed for layout detection and adaptive content retrieval—to ensure that the generated banners not only meet aesthetic standards but are also aligned with specific thematic and promotional contexts.

The architecture of our solution is deliberately modular. A NEXT.js frontend provides users with an intuitive interface to specify design parameters, select themes, and preview generated banners. Meanwhile, a Flask backend efficiently manages API communications and orchestrates the integration of the underlying AI models. This combination of technologies creates a flexible framework that can readily adapt to the evolving demands of digital marketing.

2. Previous Work and Existing Solutions

2.1 Generative Adversarial Networks (GANs) in Image Synthesis

Generative Adversarial Networks have revolutionized the field of image synthesis since their introduction. Initially applied to tasks such as image-to-image translation and artistic style transfer, GANs have evolved to generate increasingly realistic images under diverse conditions. Their ability to learn complex data distributions has made them indispensable in tasks where creative output is paramount. Prior research has demonstrated that conditional GANs can be tailored to generate images based on specific inputs, which is a critical capability for the banner generation task.

2.2 Layout Analysis and Content Adaptation Models

Recent progress in computer vision has given rise to models that specialize in understanding and predicting spatial layouts. LayoutDETR is one such model that applies transformer-based techniques to predict the positions of objects within a given space. Its effectiveness in document layout analysis has motivated its adaptation for dynamic design tasks, where the placement of text, images, and other decorative elements must be carefully managed.

Similarly, RetrieveAdapter has been utilized in contexts where content must be dynamically fetched and adapted to a given scenario. When combined with layout analysis models, it helps ensure that the selected design elements match the contextual requirements of the promotional content. These advancements have laid a strong foundation for developing systems that can automate complex design tasks.

2.3 Automated Marketing Design Tools

Existing solutions in automated advertisement generation typically focus on either static content creation or limited aspects of design automation. While several commercial tools offer template-based designs, they often lack the adaptability and creativity provided by AI-driven methods. Our work differentiates itself by integrating advanced generative models and layout prediction algorithms, thereby offering a more holistic and flexible approach to banner design that caters to the nuances of modern digital marketing.

3. Methodology

Our approach to automated banner generation is grounded in a modular architecture that seamlessly integrates various AI models with robust web technologies. This section details the components and techniques employed in our system.

3.1 System Architecture

The system is organized into two primary layers: the user interface and the backend processing. The frontend, implemented in NEXT.js, offers an interactive platform where users can input text, select images, and choose thematic elements (e.g., seasonal events such as Diwali or Dussehra). This real-time interface is designed to provide immediate visual feedback, thereby enhancing user engagement.

The Flask backend acts as the mediator between the user interface and the AI models. It exposes a set of API endpoints that receive user inputs, manage the coordination of model inferences, and return the generated banner in both HTML and PNG formats. This separation of concerns ensures that the system is both scalable and maintainable.

3.2 AI Model Integration

3.2.1 Generative Adversarial Networks

At the core of our banner generation process is a GAN architecture that synthesizes images conditioned on user inputs. This conditional generation enables the model to tailor its output according to specific themes and design parameters. The GAN is trained on a dataset augmented with synthetic examples to capture a wide range of design styles and contexts. Its objective is to produce high-quality visual content that meets the aesthetic and functional requirements of promotional banners.

3.2.2 LayoutDETR for Spatial Arrangement

LayoutDETR is integrated to ensure that the generated banners exhibit balanced and coherent layouts. By predicting optimal positions for various elements, the model facilitates a design that is both visually harmonious and functionally effective. The spatial predictions made by LayoutDETR are critical for maintaining clarity and ensuring that promotional messages are prominently displayed.

3.2.3 RetrieveAdapter for Content Adaptation

Complementing the spatial analysis, RetrieveAdapter dynamically selects and adapts design elements to enhance contextual relevance. This model retrieves the most appropriate graphical components from a curated library based on the banner's theme and the current layout. The synergy between RetrieveAdapter and LayoutDETR allows for a design process that is both adaptive and precise, ensuring that every element contributes effectively to the overall promotional message.

3.3 Training and Optimization Strategies

A significant challenge encountered during development was the limited availability of large, high-quality datasets specific to banner design. To address this, we employed two key strategies:

- **Zero-Shot Learning:** By leveraging pre-trained models from related domains, our system can generalize to the banner design task even in the absence of extensive labeled data.
- **Transfer Learning:** We fine-tuned pre-trained models on smaller, domain-specific datasets. This approach allowed us to adapt existing knowledge to the nuances of banner creation, resulting in improved performance and contextual accuracy.

In addition to these learning strategies, performance optimizations were implemented to reduce rendering times and enhance overall system responsiveness. Optimized image processing pipelines and resource management techniques ensured that the system could handle complex designs with minimal latency.

4. Discussion and Future Work

4.1 Discussion

The integration of GANs with layout and content adaptation models in our system has demonstrated the potential for automated, high-quality banner generation. The combined use of LayoutDETR and RetrieveAdapter has proven effective in addressing the dual challenges of spatial arrangement and contextual adaptation. Our experiments have shown that even with limited domain-specific data, advanced techniques such as transfer learning and zero-shot learning can significantly enhance the model's performance.

User feedback during initial testing highlighted the system's ability to produce aesthetically pleasing and contextually appropriate banners, thereby validating our approach. Moreover, the modular design of the system suggests that it can be extended to incorporate additional features without significant restructuring.

4.2 Future Work

The current implementation focuses on 2D banner generation. Future research directions include:

- **Video Banner Generation:** Extending the system to create dynamic video banners that can further engage modern audiences.
- **3D Banner Capabilities:** Exploring techniques for generating immersive 3D banners that leverage augmented and virtual reality environments.
- **Enhanced Personalization:** Integrating user data and behavioral analytics to produce highly personalized designs tailored to individual marketing strategies.
- **Social Media Integration:** Developing seamless sharing functionalities to facilitate direct publishing on various social media platforms, thereby expanding the system's utility in digital marketing.

These avenues represent promising directions for extending the capabilities of our system and addressing the evolving needs of digital content creation.

5. Conclusion

This paper has presented an integrated approach to AI-powered banner generation that harnesses the creative potential of GANs and the precision of layout and content adaptation models such as LayoutDETR and RetrieveAdapter. Our system not only automates the design process but also ensures that the generated banners are both visually appealing and contextually relevant. The use of advanced training techniques, coupled with a robust system architecture, addresses critical challenges in data scarcity and performance optimization.

The encouraging results and positive user feedback suggest that our approach can significantly streamline digital marketing workflows. With planned enhancements, including video and 3D banner generation, our work promises to further transform the landscape of automated creative design in marketing.

References:

[1] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," arXiv preprint, vol. 2005.11401, 2020. [Online]. Available: https://arxiv.org/abs/2005.11401

[2] P. Lewis, E. Denoyer, and S. Riedel, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," arXiv preprint, vol. 2005.12872, 2020. [Online]. Available: https://arxiv.org/abs/2005.12872

[3] A. Bala, "How Artificial Intelligence Will Change the Future of Marketing," ResearchGate, 2019. [Online]. Available: https://www.researchgate.net/publication/336430543_How_artificial_intelligence_will_change_the_future_of_marketing

[4] V. Xian, Z. Luo, J. Li, L. Wang, X. Xiong, and D. Tao, "Zero-Shot Learning—A Comprehensive Evaluation of the Good, the Bad and the Ugly," arXiv preprint, vol. 1707.00600, 2017. [Online]. Available: https://arxiv.org/abs/1707.00600

[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," Advances in Neural Information Processing Systems (NeurIPS), vol. 27, 2014.

[6] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," International Conference on Learning Representations (ICLR), 2015. [Online]. Available: https://arxiv.org/abs/1409.1556

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[8] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," arXiv preprint, vol. 1301.3781, 2013.