Automated Ad Creative Generation

Vishakha Kadam[1]**, Yiping Jin**[2,*]**, Bao-Dai Nguyen-Hoang**[1]
[1]Knorex, #04-01 21 Merchant Road, Singapore
[2]Department of Mathematics & Computer Science, Chulalongkorn University, Thailand
`{vishakha.kadam, jinyiping, dai.nguyen}@knorex.com`

Abstract

Ad creatives are ads served to users on a web-page, app, or other digital environments. The demand for compelling ad creatives surges drastically with the ever-increasing popularity of digital marketing. The two most essential elements of (display) ad creatives are the advertising message, such as headlines and description texts, and the visual component, such as images and videos. Traditionally, ad creatives are composed by professional copywriters and creative designers. The process requires significant human effort, limiting the scalability and efficiency of digital ad campaigns. This work introduces AUTOCREATIVE, a novel system to automatically generate ad creatives relying on natural language generation and computer vision techniques. The system generates multiple ad copies (ad headlines/description texts) using a sequence-to-sequence model and selects images most suitable to the generated ad copies based on heuristic-based visual appeal metrics and a text-image retrieval pipeline.

1 Introduction

Visually appealing ads with a compelling message will promote the brand image and lead to a better click-through rate. However, the ad composition process is time and labor-intensive, severely limiting the number of unique ads for each campaign. The ads' effectiveness deteriorates as the users are repeatedly exposed to the same ads, referred to as ad fatigue (Abrams and Vee, 2007). It underlines the importance of generating ad creatives automatically at scale.

This demo paper presents AUTOCREATIVE, a novel system for ad creative design which combines ad copy generation using a sequence-to-sequence Transformer model and ad image selection using object/scene detection and aesthetic appeal scoring. Figure 1 presents example ad creatives generated by our system.
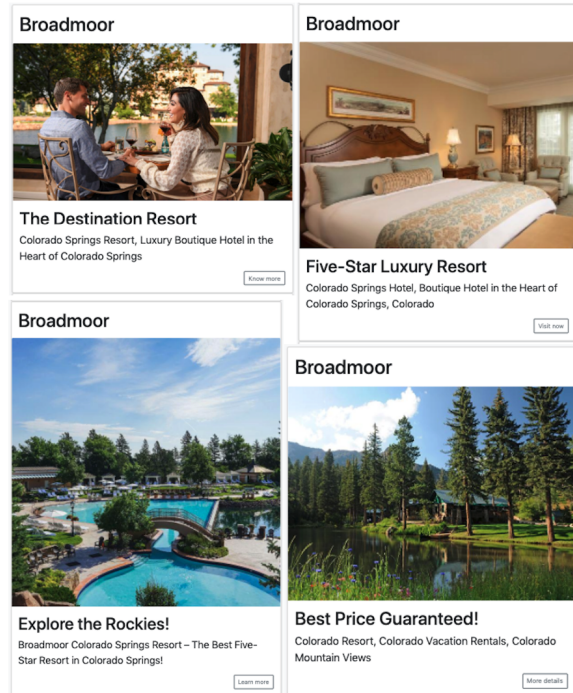
---

*Work done while at Knorex.

Figure 1: Examples ad creatives generated for the URL - https://www.broadmoor.com/.

2 System Overview

AUTOCREATIVE takes only the advertiser's URL as input. It first crawls the textual content and images from the advertiser's website. The ad copy generation module generates ad headlines and description texts conditioned on the content of the advertiser's website. The candidate images first go through a visual quality filter, then match against the generated ad copy using a text-image retrieval pipeline. Figure 2 overviews the framework.

2.1 Ad Copy Generation

We generate ad headlines and description texts conditioned on the textual content crawled from the advertiser's URL using a BART encoder-decoder model (Lewis et al., 2020). We use a relatively
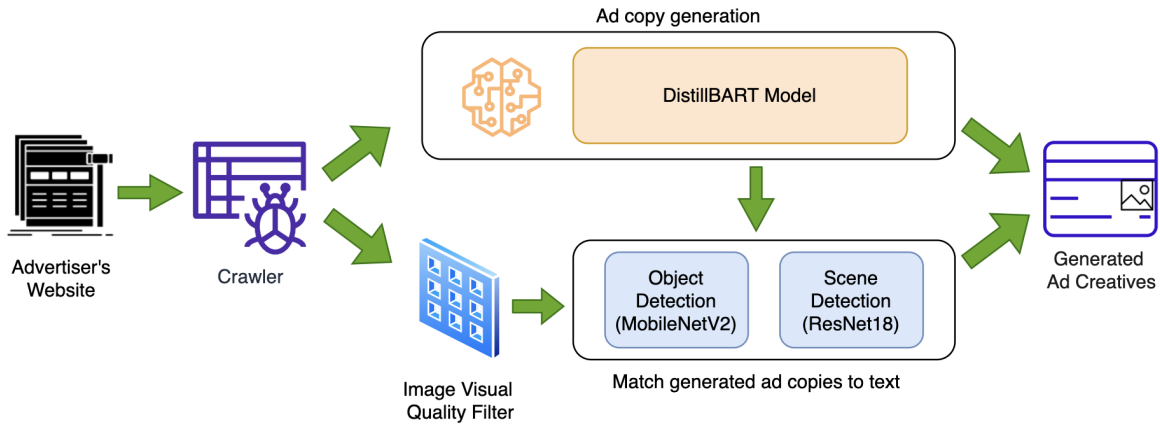
Figure 2: Overview of AUTOCREATIVE, an automated creative generation framework.

small DistilBART model [1] and fine-tuned it on a proprietary dataset of 300k (company description, advertising message) pairs (Jin et al., In press) .

Different message types (headline/description) and ad channels have different requirements for the ad message length. Therefore, we bucket the training data based on the ad message length and achieve fine-grained length control using conditional training. Similarly, we extract the POS tag of the first word in the ad message and use it as an additional control code to generate syntactically-diverse messages. The final generation probability is $P(message|description, len, pos)$, where len specifies the message length and pos specifies the POS tag of the first generated word. During inference, we choose the length based on the constraints and randomly sample POS tags to input to the model.

2.2 Image Selection

We first filter out noisy images like social media icons, background, or footer images, then measure the images' visual appeal based on factors like image colorfulness, contrast ratio, and lighting. Images that meet a pre-determined visual appeal threshold are input to a text-image matching module, where we utilize pre-trained object and scene detection models to match images with the generated headlines.

We use a MobileNet V2 (Sandler et al., 2018) model trained on Open Images Dataset V4 [2] for object detection and a ResNet-18 model trained

on Places-365 Dataset [3] for scene detection (Zhou et al., 2017). We represent each image with the average *word2vec* embedding for the detected objects and scenes, and we calculate the average word2vec embeddings of the ad headlines. Finally, we return the images with the highest cosine similarity with the ad headline embeddings, as illustrated in Figure 3.

Figure 3: Matching image to a given ad headline.

3 Feedback from Account Team and Clients

We deployed AUTOCREATIVE to production in March 2022 and integrated it to Knorex XPO [4], a self-serve cloud marketing automation platform. The system has enabled over 30 advertisers to generate ad creatives across various campaigns. The feedback we received from our internal account team and various clients was overwhelmingly pos-

itive. The technology was especially appreciated by small advertisers, who cannot afford an internal creative design team. Before the system was deployed, our creative designer team used to take more than a day to produce a set of ad creatives (including the communication overhead and the overhead of working on multiple creative design tasks simultaneously). Our system drastically reduced the turn-around time to a few minutes. Advertisers can now generate a set of appealing and diverse ad creatives from the UI without any creative design or copywriting knowledge. Critical feedback mostly relates to the cases where crawling is blocked on the advertiser landing page or no suitable image is available.

4   Conclusion

We introduced AUTOCREATIVE, a novel framework to automatically generate ad creatives. It has been deployed to production and used by clients of a global digital advertising company.

Acknowledgements

References

Zoë Abrams and Erik Vee. 2007. Personalized ad delivery when ads fatigue: An approximation algorithm. In *Proceedings of the International Workshop on Web and Internet Economics*, pages 535–540, Bangalore, India. Springer.

Yiping Jin, Akshay Bhatia, Dittaya Wanvarie, and Phu T. V. Le. In press. Toward improving coherence and diversity of slogan generation. *Natural Language Engineering*, pages 1–33. Cambridge University Press.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. 2020. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, page 7871–7880.

Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4510–4520, Salt Lake City, Utah, USA. IEEE.

Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. 2017. Places: A 10 million image database for scene recognition. *IEEE transactions on pattern analysis and machine intelligence*, 40(6):1452–1464.