Visual Aids to the Rescue: Predicting Creativity in Multimodal Artwork

Serra Sinem Tekiroglu, Gözde Özbal, and Carlo Strapparava

FBK-irst, Trento, Italy
{serrasinem,gozbalde}@gmail.com, strappa@fbk.eu

Abstract. Creativity is the key factor in successful advertising where catchy and memorable media is produced to persuade the audience. Considering not only advertising slogans but also the visual design of the same advertisements would provide a perceptual grounding for the overall creativity, consequently the overall message of the advertisement. In this study, we propose the exploitation of visual modality in creativity assessment of naturally multimodal design. To the best of our knowledge, this is the first study focusing on the computational detection of multimodal creative work. To achieve our goal, we employ several linguistic creativity detection features in combination with bag of visual words model and observable artistic visual features. The results of the creativity detection experiment show that combining linguistic and visual features significantly improves the unimodal creativity detection performances.

Keywords: NLP for Creative Language · Images and Language · Multimodality.

1 Introduction

Making an advertisement catching and memorable is the core task of creative people behind any original and effective campaign. Especially analysing award-winning ads, it is possible to appreciate a range of approaches including ways of visualizing concepts, the use of rhetorical devices, such as exaggeration, paradox, metaphor and analogy, and taking advantage of shock tactics and humour [5]. In any case, visual and text modalities are carefully thought to have a complementary and coordinated effect. While a computational treatment of the most subtle techniques is still very challenging, we think it is worthwhile to start exploring this topic in order to have a better understanding of contributing factors that can be utilized in computational creativity.

Accordingly, automatically quantifying the creativity level of multimodal design might be beneficial in various purposes such as choosing the potential successful advertisements, creative language and image generation for educational material or even a computational assessment of artistic value of the multimodal artwork.

As a topic being on the rise in computational linguistics, multimodality is mostly exploited on top of the linguistic models to perceptually ground the

current tasks. For instance, semantic representations benefit from the reinforcement of linguistic modality with visual [2, 1, 3] and auditory [4] modalities. In the same manner, we propose devising visual modality in collaboration with linguistic modality in creativity detection task. To the best of our knowledge, this is the first study aiming to identify multimodal creativity in a computational manner. Moreover, the multimodality type of the dataset stands out amongst the others since the linguistic channel of an advertisement is complementary to the visual channel instead of being a scene description or an image label.

We used a set of naturally creative images; an advertising dataset which is composed of 500 images and corresponding slogans. As the counterpart of the creative data class, we investigated WordNet synset definitions and corresponding images from ImageNet. Figure **??** exemplifies the creative, non-creative tuples in the final dataset. Although these images seem to be very similar at the first glance with a child in the middle of the frame, the subtle and creative details in the left picture of Figure **??**, such as the male figure held by the child, immediately draw audience's attention. In this study, our focus will be to capture these creative properties both in visual and linguistic modalities to generate a multimodal creativity detection model.

We conjecture that in order to get meaningful information from a small dataset, we need to use the features that are as generalized as possible. To this end, we employ a Bag-of-Visual-Words (BoVW) model to determine if creativity in images display common visual characteristics and if these characteristics have a positive effect on overall creativity of a multimodal advertisement.

The rest of the paper is organized as follows. We first give a brief summary of the relevant multimodality and computational creativity studies in Section 2. In Section 3, we give the basis of our work in terms of the dimensions of advertising creativity. We present the creativity dataset that we collected in Section 4. Section 5 and 6 include the visual and linguistic creativity detection models, while Section 7 summarizes the experiments that we conducted.

References

1. Bruni, E., Tran, G.B., Baroni, M.: Distributional semantics from text and images. In: Proceedings of the GEMS 2011 workshop on geometrical models of natural language semantics. pp. 22–32. Association for Computational Linguistics (2011)
2. Bruni, E., Tran, N.K., Baroni, M.: Multimodal distributional semantics. J. Artif. Intell. Res.(JAIR) **49**(1-47) (2014)
3. Feng, Y., Lapata, M.: Visual information in semantic representation. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. pp. 91–99. Association for Computational Linguistics (2010)
4. Kiela, D., Clark, S.: Multi-and cross-modal semantics beyond vision: Grounding in auditory perception. In: Proceedings of EMNLP (2015)
5. Pricken, M.: Creative Advertising Ideas and Techniques from the World's Best Campaigns. Thames & Hudson, 2^{nd} edn. (2008)