Fashionpedia-Ads
Do Your Favorite Advertisements Reveal Your Fashion Taste?
Shi, Mengyun; Cardie, Claire; Belongie, Serge

Publication date:
2023

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):

Download date: 31. mar., 2024
Fashionpedia-Ads: Do Your Favorite Advertisements Reveal Your Fashion Taste?

Mengyun Shi¹ Claire Cardie¹ Serge Belongie²
¹Cornell University  ²University of Copenhagen

Abstract

Consumers are exposed to advertisements across many different domains on the internet, such as fashion, beauty, car, food, and others. On the other hand, fashion represents second highest e-commerce shopping category. Does consumers’ digital record behavior on various fashion ad images reveal their fashion taste? Does ads from other domains infer their fashion taste as well? In this paper, we study the correlation between advertisements and fashion taste. Towards this goal, we introduce a new dataset, Fashionpedia-Ads, which asks subjects to provide their preferences on both ad (fashion, beauty, car, and dessert) and fashion product (social network and e-commerce style) images. Furthermore, we exhaustively collect and annotate the emotional, visual and textual information on the ad images from multi-perspectives (abstractive level, physical level, captions, and brands). We open-source Fashionpedia-Ads to enable future studies and encourage more approaches to interpretability research between advertisements and fashion taste. Fashionpedia-Ads can be found at: ¹

1. Introduction

It is understandable that there could be some correlation between ads and products for a same domain. For example, a user likes the style of a neckline in a fashion ads and might also like a fashion product that has similar style (Fig. 1). However, is there any correlation between ads and products from different domains? Specifically, can we interpret a consumer’s product preference from her website browsing?

¹Fashionpedia project page: fashionpedia.github.io/home/
The aim of this work is to enable future studies and encourage more exploration to interpretability research between advertisements and fashion taste. The contributions of this work are: 1) we introduce Fashionpedia-Ads, consisting of three datasets (Ads, Social network style and E-commerce style fashion products). We bridge the connection among them through the subjects’ preference (like or dislike) on these images and the annotation from multi-perspectives (e.g. abstract & physical attributes). 2) we formalize a new task that not only requires models to predict whether a subject like or dislike a fashion product image based on given ad images of various domains, but also provide a rationale explanation why it makes this prediction from multi-perspectives.

### 2. Dataset Creation Details

**3 sub-datasets** Fashionpedia-Ads dataset consists of 3 sub-datasets. 1) **Advertisement dataset**: this dataset consists of ad images from fashion, beauty, car, and dessert domains that consumers could see on the internet. We use the images from the ads dataset [11] for our study; 2) **Social network style fashion product dataset**: this dataset consists of street and runway style fashion product images (with human body), which simulates images that consumers could see on social network websites. We use the images from Fashionpedia dataset [13] for our study; 3) **E-commerce style fashion product dataset**: this dataset consists of online shopping style fashion product images (without human body), which simulates images that consumers could see on E-commerce websites. We collect the images and associated product information from an online shopping website called nuji.com.

**Build the relationship among the 3 sub-datasets through subject preference** We build the connection among the 3 sub-datasets by asking the same 100 subjects to annotate their preference (like or dislike) for all these 3 sub-datasets, as illustrated in Fig. 1.

### 3. Advertisement Annotation

**Visual & Abstractive perspective** We extract the sentiment and question-answer pairs (Action & Reason) annotation from the ads dataset [11] directly. These annotations are collected from MTurk workers and can help us understand subjects’ preference from sentiment and emotional level.

**Visual & Physical perspective** We follow the method proposed by Fashionpedia dataset [13] and construct taxonomies for the 4 ads domains. By using these taxonomies, we annotate localized objects, sub-objects and fine-grained attributes with associated masks for the ad images of these 4 domains. Why we annotate this? In contrast to the sentimental and emotional perspectives of the ads, a subject’s preference on ads can also be impacted by the visual effect of products demonstrated in the ads. With this annotation, we could analyze whether subjects’ preference is impacted by the physical perspective of the products shown on the ads.

**Textual perspective** We annotate the leading captions indicated in the ad images with the masks. Because the subjects’ preference and emotional feeling can be aroused by the textual information displayed on the ads.

**Brand perspective** We also annotate the brand name indicated in the ad image. The brands and their associated brand culture could also tie to subjects’ ads preference. For example, if a subject like fashion products from Dior, she could also like beauty products from Dior, as illustrated in Fig. 1.
4. Fashion Taste Annotation

To fully understand the contextual information from fashion product images, we annotate the social network style and E-commerce style fashion product images respectively, as mentioned in Sec. 2. 1) The social network style dataset: we directly use the annotation (task1/2/3/4) created in Fashionpedia-Taste, which asked the 100 subjects to explain their fashion taste in 3 perspectives: a) localized attributes; b) human attention; c) caption, as shown in Fig. 1; 2) The E-commerce style dataset: similar to Fashionpedia dataset [13], we annotate localized objects, sub-objects and fine-grained attributes with associated masks. Additionally, we annotate the dress length and introversion/extroversion for each image. Each image also contains detailed product information, as shown in Fig. 1.

5. Dataset Analysis

5.1. Visual-Abstractive Attribute Level

Sentiment Fig. 3 shows the frequency of the sentiment of the ads images for each ads domain. The results show the sentiment from fashion and beauty ads are more correlated compared to car and dessert ads (such as ‘fashionable’). Furthermore, fashion and beauty ads are more correlated to ‘feminine’ sentiment. However, car ads has more tendency of ‘manly’ sentiment. Dessert ads is invariant of both ‘feminine’ and ‘manly’ sentiment. All the 4 ads domains are correlated to ‘creative’ and ‘active’ sentiment.

Q&A word count statistics We use SGRank from Textacy [24] to calculate the frequency of words. Fig. 4 shows the most frequent 1, 2, 3 grams for QA. Similar to the sentiment analysis, fashion and beauty ads share similar emotional feeling. ‘Attractive’, ‘sexy’, ‘stylish’, and ‘beautiful’ are the most frequent adjectives annotated by MTurk workers. For car ads, the emotional feeling is more related to the functional aspect of cars. The high frequent words include ‘high quality’, ‘reliable’, ‘great performance’, and ‘powerful car’. For dessert ads, the emotional words are more related to the dessert flavors, such as ‘delicious’, ‘new flavor’, and ‘tasty’.

Q&A linguistic statistics We use part-of-speech (POS) tagging from Spacy [22] to tag nouns and adjectives in QA. Table 1 shows the number of most frequent unique nouns by POS. We find the most frequent common nouns are more associated with high level description of a product line indicated in ads images, such as clothes, jeans, perfume, lipstick, ice cream, and cookies. In contrast, the most frequent proper nouns is more related to the brands mentioned in the ads images, such as Gucci, Chanel, Audi, and Häagen-Dazs. This shows the linguistic diversity of our dataset.

5.2. Visual-Physical Attribute Level

Localized Attribute distribution Fig. 5 shows the distribution of attributes annotated in the 4 ads domains. For fashion ads, ‘symmetrical’ has highest frequency because most of the garments have balanced silhouette. For beauty ads, ‘glossy’ has high frequency because the high frequency of lipstick objects in the beauty ads. The similar observation is found for ‘cocoa chocolate’ in dessert ads. For car ads, ‘elegant & luxury’ style has high frequency, which could correlate to the fashion ads and products with ‘elegant & luxury’ feeling.

5.3. Textual information

Ads caption - Word count statistics For Fashion ads, Fig. 4 shows Kenneth cole, Calvin Klein, and Banana republic are most mentioned brands from emotional perspective. However, Chanel, Boss and Lacoste are most frequent brands according to the caption written on the ads images (Fig. 6). This indicates Kenneth cole, Calvin Klein, and Banana republic ads might contain some more provocative visual or textual information that arouses the MTurk workers’ sentiment. The similar observation is also found in beauty, car, and dessert ads.

Ads caption - Linguistic statistics Table 2 shows number of most frequent nouns from ads caption. Compared to QA (Table 1) annotated from emotional perspective, the caption written on the ad images is more concentrated to brand related keywords and less diverse. This is expected since the signal from QA contains more expression of the viewers’ emotional feeling and the caption written on the ads images is more focused on describing the brand itself or products that are advertised in the images.

5.4. Brand name

Fig. 7 shows the distribution of brand for all 4 ads domains. We found the brands with highest frequency mentioned in the ads images do not necessary always correspond to the highest brand keywords mentioned in QA (Fig. 4) and ads caption (Fig. 6). Fig. 4 shows Kenneth cole, Calvin Klein, and Banana republic as most mentioned brands from emotional perspective. On the other hand,
Chanel, Boss and Lacoste are the most frequent brands according to the caption written on the ads images (Fig. 6). However, the brand distribution ((Fig. 7)) shows Ralph Lauren has the highest number of images. This indicates some brands might contain some more provocative visual or textual information that arouses the viewers’ feeling even they have relatively lower number of images in the dataset. This needs to be further explored in this study. The similar observation is also found in beauty, car, and dessert ads.

6. Conclusion

In this work, we studied the problem of modeling human fashion taste in advertisements. We exhaustively collect and annotate the emotional, visual and textual information on the ad images from multi-perspectives (abstractive level, physical level, captions, and brands). We therefore hope that Fashionpedia-Ads will facilitate future research of interpretability between advertisements and fashion taste.

References


Figure 6: Caption word count statistics: Number of 1, 2, 3-grams.

Figure 7: Ads - Brand distribution.


