Beyond the Product: Discovering Image Posts for Brands in Social Media

Francesco Gelli
National University of Singapore
francesco.gelli@u.nus.edu

Tiberio Uricchio
Università degli Studi di Firenze
tiberio.uricchio@unifi.it

Xiangnan He*
National University of Singapore
xiangnanhe@gmail.com

Alberto Del Bimbo
Università degli Studi di Firenze
alberto.delbimbo@unifi.it

Tat-Seng Chua
National University of Singapore
chuats@comp.nus.edu.sg

ABSTRACT

Brands and organizations are using social networks such as Instagram to share image or video posts regularly, in order to engage and maximize their presence to the users. Differently from the traditional advertising paradigm, these posts feature not only specific products, but also the value and philosophy of the brand, known as *brand associations* in marketing literature. In fact, marketers are spending considerable resources to generate their content in-house, and increasingly often, to discover and repost the content generated by users. However, to choose the right posts for a brand in social media remains an open problem. Driven by this real-life application, we define the new task of *content discovery for brands*, which aims to discover posts that match the marketing value and brand associations of a target brand. We identify two main challenges in this new task: high inter-brand similarity and brand-post sparsity; and propose a tailored content-based learning-to-rank system to discover content for a target brand. Specifically, our method learns fine-grained brand representation via explicit modeling of brand associations, which can be interpreted as visual words shared among brands. We collected a new large-scale Instagram dataset, consisting of more than 1.1 million image and video posts from the history of 927 brands of fourteen verticals such as food and fashion. Extensive experiments indicate that our model can effectively learn fine-grained brand representations and outperform the closest state-of-the-art solutions.

KEYWORDS

Content Discovery, Image Ranking, Computational Marketing

ACM Reference Format:
Francesco Gelli, Tiberio Uricchio, Xiangnan He, Alberto Del Bimbo, and Tat-Seng Chua. 2018. Beyond the Product: Discovering Image Posts for Brands in Social Media. In *2018 ACM Multimedia Conference (MM '18), October 22–26, 2018, Seoul, Republic of Korea.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3240508.3240689

*Xiangnan He is the corresponding author.

MM '18, October 22–26, 2018, Seoul, Republic of Korea

ACM ISBN 978-1-4503-5665-7/18/10…$15.00
https://doi.org/10.1145/3240508.3240689

Figure 1: Three recent sample posts from Instagram accounts of the brands Airfrance and Gucci (first and second line respectively). Brand associations are aircrafts, natural and urban landscapes for the former brand; and extravagant models, colorful clothes and vintage art for the latter.

1 INTRODUCTION

With the advance of Social Network (SN) websites, we are experiencing an unprecedented growth in user generated media items such as images or videos. According to Hootsuite[1], in 2016 Instagram users alone were sharing 95 million photos and videos per day. This phenomenon is widely affecting media industry, advertising and politics, which are adapting in order to leverage social media and User Generated Content (UGC). Most brands and organizations widely use SN websites such as Instagram to regularly share posts and engage social network users. According to recent marketing trends [30], consumers are becoming increasingly resistant to messages directly asking them to buy a product, hence such direct Ads are losing their appeal. For this reason, recent posts from brands rarely feature a product directly, but are often centered on a set of *brand associations* which reflect the philosophy and lifestyle of a brand (see Figure 1) [4]. For instance, BMW posts aim to project the ideas of sophistication, fun driving and superior engineering to the mind of the consumers. Since many amateur photographers may share content that deeply associate with the marketing value of a brand, discovery of UGC has become a major trend in marketing

[1]https://blog.hootsuite.com/instagram-statistics/

[26]. Brands aim to promote, engage and often re-share this kind of UGC to reinforce their marketing efforts[8, 9, 14].

However, how to choose automatically the right posts for a brand is an open problem. The current process involves a manual selection made by a social network marketing expert who understands the brand ideas and has the daunting task to analyze and choose the right content from millions of possible posts. Fully automating this process is thus desirable. While several companies (e.g. Olapic, ShareIQ, Chute and Curalate) offer hashtag-based services to marketers, relying only on hashtags severely limits the amount of relevant UCG that can be discovered. Brand-related hashtags in fact requires users to be aware of a brand's campaign and deliberately take action to be discovered, while a lot of relevant UGC is still spontaneously posted and they may be a better match for the philosophy of a brand.

To our knowledge, no work in the literature address the problem of automatically discovering visual content that matches the unique posting style of a social network account for a brand or organization. We thus name this problem as **Content Discovery for Brands** and formulate it as a learning-to-rank task. Given a set of social network image posts, the goal is to generate a personalized ranking based on their likelihood to be adopted by a brand. The most significant challenge here is high *inter-brand similarity*: many brand associations are often shared across multiple brands, resulting in photographs posted by brands having subtle differences from that of their competitors. For example, visual elements of aircrafts as in Figure 1 often appear in posts by other airline companies as well, such as United or Emirates, where the differences may only be in the logo on the aircraft or the cities served. It is thus a fine-grained problem, where it is important to distinguish as much as possible the specific set of unique values and associations of each brand from the competitors.

The problem of discovering images for brands is close to the task addressed by image recommendation systems [21, 25, 29], which recommend images to social network users based on personal preferences. However, *recommending images* and *content discovery for brands* are indeed different problems. The first predicts *what a user would like* while the second *what a brand user should post* to project its ideas to the audience. One significant distinction is that in recommendations scenarios, for each image there are often multiple users interacting with it (e.g. likes, views). However, in our case it is extremely rare that different brands post the same exact images, either for reason of copyright infringement or marketing tendency to project brand uniqueness [3]. This leads to extremely sparse brand-image interaction, which needs to be addressed with advanced content-based methods.

We propose a learning-to-rank method named Personalized Content Discovery (PCD) to discover image posts for brands, learning from their posting history. PCD is designed to learn fine-grained brand representation by explicitly modeling high-level semantics brand associations, which correspond to recurring visual aspects of the posting history of brands. Being purely content based, PCD can be used to rank a set of unseen image for a brand by simply using cosine similarity measure. We jointly trained the model end-to-end using a large scale Instagram dataset of more than 1.1 million posts, which we collected by crawling the posting history of 927 brands from fourteen different verticals such as food, fashion and auto. We benchmarked the model against the closest works in the state-of-the-art works, using three different quantitative metrics, and performed ablation studies and qualitative analysis of case studies.

We summarize the contributions of this work as follows: 1) We cast the emerging problem of content discovery for brands as a content-based learning-to-rank problem, highlighting the main challenges. 2) We design and benchmark the Personalized Content Discovery (PCD) as a novel content discovery framework which explicitly models the brand's values and associations. The method is able to outperform the closest state-of-the-art works in content-based recommendations, demonstrating its ability to learn a more fine-grained brand representation. 3) We collected a large scale Instagram dataset containing more than 1.1 million image and video posts with associated metadata. The dataset is released to the research communities for further studies on content discovery, popularity predictions and computational marketing.

2 RELATED WORK

The task of content discovery for brands and our approach is mainly related to the literature on social media marketing for the background scenario and image recommendations which solve a similar problem.

2.1 Social Media Marketing and Computational Marketing

Social Media Marketing (SMM) is a real-life field that is strongly related to the industry world, particularly branding. Nonetheless, it have constantly solicited a vast amount of research across different disciplines, from marketing itself to psychology [2]. Gensler et al. [13] outlined a list of research questions related to SMM, such as investigating what kind of brand content will stimulate consumers and spread in social networks. Such research question inspired several works using computational approaches to explain marketing dynamics in social media, establishing the subfield of literature named Computational Marketing. One popular example is the study by De Vries et al [7], who employed regression models to investigate which features of social brand posts engage users. The authors outlined a set of post features which have a positive impact on indicators such as the number of likes and comments, including image vividness and post interactivity. In the same line, several other works investigated similar indicators with multiple social features on different networks like Facebook [20, 32], Twitter [1] and Instagram [28, 31].

A second group of works tackled the problem of identifying the brand associations from social media posts [22, 23]. These papers propose clustering-based methods to automatically detect visual brand associations from social networks post as image clusters. They treat each brand independently from the others. Different from those, we argue that the same or similar associations are shared among different brands. We thus propose a model that learns brand associations jointly among a large set of brands.

2.2 Image Recommendations

Discovering content for brands share some similarities to the task of recommending images to social users, which solutions are mostly based on recommender systems.

A first group of works is based on Factorization Machines (FM), which are designed to model the interaction between each combination of entities, including users, items and any additional numerical value information available [5, 17, 33]. Factorization Machines have been extended to content-based scenarios, adopting context information such as hashtags, text in the image [6], or integrating dense image features extracted with a pre-trained CNN [11]. Other frameworks make use of denoising auto-encoders [38, 41], attention models [5] or recurrent neural networks [19] to integrate content features in recommendation architectures. All these works rely on the hypothesis that rated items are shared by multiple users, which does not hold in the task of content discovery for brands.

A second line of recommendation works adopt the pairwise optimization framework Bayesian Personalized Ranking (BPR) by Rendle et al. [34]. Among these, VBPR enriched the item latent vectors with image features [15], while Yu et al. [39] applied BPR to the fashion domain. The more recent Neural Personalized Ranking (NPR) by Niu et al. [29] complements existing matrix factorization and VBPR with a new more flexible neural network architecture. The model supports image features as input, which can be used together with other context clues such as topic or location.

However, these models are not purely content-based, since they only integrate content-features in traditional collaborative filtering frameworks [18]. Other works propose instead solutions that rely solely on content[12, 21, 25, 27], which has the advantage of preventing the item cold-start problem [36]. For example, Kang et al. [21] proposed a method based on Bayes Personalized Ranking in the fashion domain. Differently, Lei et al. [25] proposed a Comparative Deep learning (CDL) triplet network, where a positive and a negative sample image are simultaneously forwarded through two convolutional subnetworks with shared weights, together with information related to a user who interacted with the positive, but not with the negative photograph. The user features are processed with a separate subnetwork, then compared with the two image latent features with element-wise subtraction.

Even if this last group of works is closer to our task, all recommendation models are not designed for the problem of content discovery, since they either rely on multiple user-item interactions or are not explicitly designed for learning fine-grained brand representations.

3 PERSONALIZED CONTENT DISCOVERY

This section is dedicated to present our Personalized Content Discovery (PCD) framework for the problem of content discovery for brands. We first present the notations and define the problem of content discovery for brands. We describe the two main components of PCD: brand and post representation learning. Finally, we describe the details of the optimization method that we used to train the model.

3.1 Notations and Problem Formulation

We adopt in this paper the widely used notation or representing matrices and vectors with bold capital letters and bold lowercase letters respectively. In the case of matrices, we use the subscript notation to denote a specific row. For example, $\mathbf{A_i}$ indicates the i-th row of matrix $\mathbf{A}$. To indicate sets, we use capital cursive letters and we use regular lower case letters for elements of the set, such as $a \in \mathcal{A}$. We indicate brands as $\mathcal{B} = \{b_1, ..., b_N\}$, where $b_i \in \mathcal{B}$ is a brand with an active social media account. For the image posts set we use the notation $\mathcal{P} = \{p_1, ..., p_M\}$, where the elements are all the images posted on a social network platform. Brands and images are linked with a single type of relation, which is the adoptation relation. We refer to the posting history of a brand b as $\mathcal{H}(b)$. An image p is posted by the author b with the notation $p \in \mathcal{H}(b)$, where $\mathcal{H}(b) \subseteq \mathcal{P}$. Because of the brand-post sparsity, posting histories have little or no overlap.

Given these notations for the input data, the goal of the problem of content discovery for brands is to learn a function $f : \mathcal{B}x\mathcal{P} \mapsto \mathbb{R}$ such that for each new posts p_x of any brand $b \in B$:

$$f(b, p_x) > f(b, p_y) \tag{1}$$

where p_y is a new post of any other brand $\hat{b} \neq b$.

In other words, given a set of brands and related posting histories, we aim to learn a relevance function f that can be used to rank a set of unseen image posts, such those relevant and likely to be adopted by the brand will be ranked higher. With this problem formulation, we need an effective method to browse a large set of new social media posts to discover photographs that match the marketing value and brand associations.

3.2 Proposed Method

The PCD model is inspired by the popular matrix factorization approach, which is based on the principle of mapping users and items in the same latent space [35]. Since these architectures can easily be extended to leverage the predictive power of deep neural networks, we chose to adopt a similar model to learn a common latent space for both brands and image posts. As a result, a first component of PCD is dedicated to learning the brand representation and a second component learns the post representation. When a post has a similar representation as the brand representation, it is considered a good match for the brand. Specifically, let $\mathbf{b} \in \mathbb{R}^k$ and $\mathbf{p} \in \mathbb{R}^k$ denote the latent representation of brand b and post p respectively; the function in equation 1 is the cosine similarity between the two:

$$f(b, p) = \frac{\mathbf{b}^T \mathbf{p}}{||\mathbf{b}|| \; ||\mathbf{p}||}$$

The whole PCD framework is represented in Figure 2, where the two components of brand and post representation learning are illustrated on the left and right part respectively.

3.2.1 Brand Representation Learning. The first component is designed to map a brand into the common latent space.

The most common approach is to learn user latent vectors directly from one-hot ID [21, 29]. However, because of high interbrand similarity, this approach may fail to learn discriminative brand representation capable of making fine-grained distinctions between competitor brands. For this reason, PCD learns the fine-grained brand representation by explicitly modeling the brand associations. In order to capture the high-level semantics of brand associations, our model automatically learns a fixed-sized latent vector for each of them. We define a matrix of association vectors $\mathbf{A} \in \mathbb{R}^{hk}$, where h is an arbitrary number of associations and k is the number of dimensions of the latent space. Since PCD learns

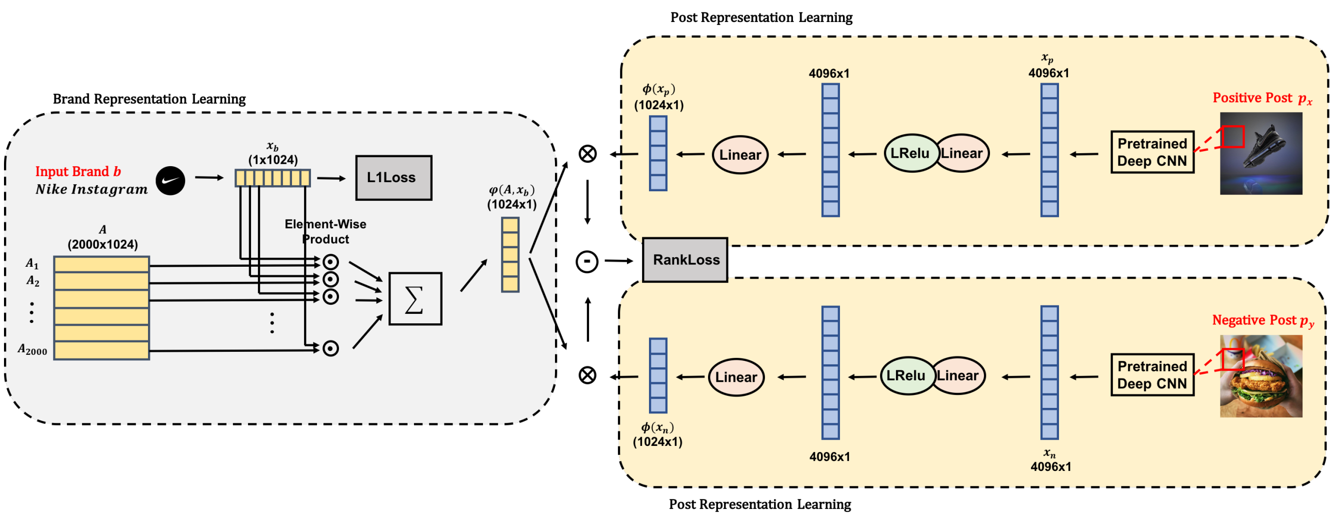

Figure 2: Model architecture. The left component performs the brand representation learning while the right learns the post representation. The two post representation networks have shared parameters.

brand representation from both the one-hot brand ID and association matrix, we define $\mathbf{b} = \varphi(\mathbf{A}, b)$. Different from other works, which pre-allocate aspects with fixed handcrafted features [16], PCD learns the importance of each single association in an end-to-end manner for each brand.

Our final brand representation vector is then computed as:

$$\mathbf{b} = \varphi(\mathbf{A}, \mathbf{x_b}) = \sum_{i=1}^{n} \mathbf{A_i} \circ \mathbf{x_b} \tag{2}$$

where $\mathbf{x_b} \in \mathbb{R}^h$ are the importance weights for brand b and $\circ$ indicates element-wise multiplication.

Since brands are free to assume any weighted combination of h high-semantics association vectors, this method allows a richer learning of fine-grained brand representations compared with learning directly from one-hot brand ID.

3.2.2 Post Representation Learning. The second component of PCD is responsible to learn the representation for a post p and project it to the same k-dimensional space of brand representations.

Because of the brand-post sparsity problem, the one-hot post ID as input feature does not provide any information. For this reason PCD uses uniquely image content to learn image post representation, similar to [25]. Secondly, this design prevents the item cold-start problem [36], since our model is intended to rank a set of new posts which were not used to train the model.

We designed a two-layer neural network ϕ for this task, whose solely input is image content. To achieve faster training, we adopted the same strategy as in [15] and utilize features extracted with pretrained CNN as input. We denote the extracted image feature of post p as $\mathbf{x_p}$.

As illustrated in Figure 2, PCD processes image posts using a linear network $\phi(\mathbf{x_p})$ with two linear layers and a Leaky ReLU activation:

$$\mathbf{p} = \phi(\mathbf{x_p}) = \mathbf{W}_2(\xi(\mathbf{W}_1\mathbf{x_p} + \boldsymbol{\gamma_1})) + \boldsymbol{\gamma_2} \tag{3}$$

where

$$\xi(x) = \begin{cases} x & \text{if } x > 0 \\ 0.01x & \text{otherwise} \end{cases} \tag{4}$$

Our preliminary experiments showed that using Leaky ReLU leads to better performance compared to standard ReLU.

3.3 Optimization

3.3.1 Loss Functions. Our training data T consist in a set of brand-post pairs (b, p) such that $p \in H(p)\,\forall(b, p) \in T$. Similar to [15], we adopt pairwise learning, which consists in sampling a certain number of negative posts for each pair. With pairwise learning, a single training data point has the following form: (b, p_x, p_y) where b is a brand, $p_x \in H(p)$ and $p_y \notin H(p)$. The basic intuition is that for a brand, the positive sample post should have a higher score than the negative one.

We train our CDP model $f : \mathcal{B}x\mathcal{P} \mapsto \mathbb{R}$ using the following ranking loss:

$$\mathcal{L}(b, p_x, p_y) = max(0, (f(b, p_y) - f(b, p_x)) + \eta) \tag{5}$$

where η is the minimum distance desired between the negative and the positive samples.

In addition to the ranking loss, we also add two regularization terms, in order to increase the interpretability of the aspects associated to a brand and reduce overfitting. For the first term, we recall that each brand has weights of each association latent vector $\mathbf{x_b}$ (Sect. 3.2.1). Our desired effect is that these weights operate as "selectors", positively or negatively affecting only a small set of associations for each brand. As a result, we adopt a L1 regularization on $\mathbf{x_b}$ to encourage a sparse representation [24]. For the second term, we adopt L2 regularization on every weight of the model.

The final equation of the loss is:

$$\mathcal{L}(b, p_x, p_y) + \alpha * \sum_i |\mathbf{x_b}| + \beta * ||\theta||_2 \tag{6}$$

where θ are the set of all the weights of the model and α and β control the importance of the regularization terms.

3.3.2 Learning Details. Parameters $\mathbf{A}$, $\mathbf{x_b}$ and ψ can be learned from the model using any optimizer algorithm such as SGD or Adam. We trained our model for twenty epochs using the Adadelta algorithm, which adapts the step interval dynamically over time without requiring to set any learning rate [40]. We trained using mini batches of 256 brand-post pairs, which we shuffled at the beginning of each epoch before batching. To improve generalization performance, we employed dropout with 50% dropping rate and a loss margin of 0.3.

One of the major difficulties of pairwise learning is negative sampling. Computational requirements prevent training using all the possible of brand-post pairs, hence negative data must be selected using a sampling strategy. For this work we adopted uniform negative sampling: given each brand-post pair (b, p), we randomly sampled ten negative sample posts $p_{y1}, p_{y2}, \ldots$ such that $p_{yi} \notin H(b) \, \forall \, i$.

4 EXPERIMENTS

In this section we conduct experiments to evaluate the performance of the proposed method. We first describe the process of dataset collection and the experimental setup. We then report several experiments to demonstrate the effectiveness of our method, including comparison with baselines, the performance of the brand representation, and the impact of modeling brand associations. Finally we report several qualitative experiments where we show the brand associations and perform case studies on some brands.

4.1 Dataset

To benchmark our model on real-world data, we need a large-scale image dataset with social media posting history, and the dataset should contain brands of different verticals. This is required in order to learn fine-grained brand representation from all brand associations.

4.1.1 Existing Datasets. Among public datasets of brand images posts in social network, Gao et al. released a popular large scale collection of over 1.2 million microblogs [10]. The dataset is not suitable for our task, since it only contains images with a brand logo, when one of the main points of our work is that a more complex set of brand associations exists beyond the product. Other works that study the popularity of brands [1, 31, 32] use datasets where the task of discovering content for brands may be applied, however they are private, made ad-hoc for the popularity task or publicly unavailable. We hence decided to build our own dataset.

4.1.2 A new dataset for Content Discovery. We chose Instagram as our source of data, since it has a more vibrant brand community due to its higher engagement rate [2]. We selected fourteen verticals on the marketing Website Iconosquare and collected the list of Instagram accounts from the Website. We filtered out brands with less than 100 posts to avoid sample insufficiency, retaining a set of 927 brands. For each of these, we crawled at most 2,000 recent posts from their history, for a total of 1,158,474 posts (approximately 1,250 posts per brand on average). For each post we collected the image or video together with all the metadata available such as the posting time, number of likes and comments. A possible reason for vertical *Energy* having only four brands (see Table 1) is that energy brands, such as OCTG oil & gas, target businesses rather than people and hence don't make a wide use social media. We release dataset and code to the public [3].

Table 1: Number or brands for each vertical

Alcohol 69	Airlines 57	Auto 83	Fashion 98	Food 85
Furnishing 49	Electron. 79	Nonprof. 71	Jewellery 71	Finance 37
Services 69	Entertain. 88	Energy 4	Beverages 67	**TOT 927**

Table 2: Metrics for comparison of PCD with baselines

Metric	Range	Description
AUC	[0-1]	Probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative one.
cAUC	[0-1]	Probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative sample from a competitor.
NDCG$_\mathbf{x}$	[0-1]	Measures the quality of a ranking list based on the post position in the sorted result list. Truncated at x. The higher the better.
MedR	[0-inf]	The median position of the first relevant document. The lower the better.

4.2 Experimental Setup

We split the dataset into training and testing sets, where the test set contains the 10 most recent posts for each brand and all the remaining data was used for training. This gives rise to a total of 1,149,204 posts for training and 9,270 for testing. We denote the training posts for a brand b as $H_{train}(b)$, and the testing posts as $H_{test}(b)$.

4.2.1 Metrics. We adopted the metrics described in Table 2. To measure the probability of choosing the most relevant examples like in [34], we adopt AUC as in [15]:

$$AUC = \frac{1}{|B|} \sum_{b} \frac{1}{|E(b)|} \sum_{(p_x, p_y) \in E(b)} \delta(f(b, p_x) > f(b, p_y)) \quad (7)$$

where δ is the indicator function and the evaluation pairs per brand b are:

$$E(b) = \{(p_x, p_y) | p_x \in H_{test}(b) \wedge p_y \in H_{test}(c), c \neq b\} \quad (8)$$

We introduced a novel metric called Competitors AUC (cAUC), which is computed exactly as the regular AUC, but restricting the evaluation pairs to only those involving competitor brands.

$$E(b) = \{(p_x, p_y) | p_x \in H_{test}(b) \wedge p_y \in H_{test}(c), V(c) = V(b)\} \quad (9)$$

[2]https://locowise.com/blog/instagram-engagement-rate-is-higher-than-facebook

[3]https://github.com/GelliFrancesco/pcd

where $V(b)$ is the vertical of brand b. This metric is evaluated to assess if our model is able to learn fine-grained brand representation for discriminating between subtle differences among competitors. We employ NDCG (Normalized Discounted Cumulative Gain) to evaluate the performance of post discovery by taking into account the ranking of relevant posts. We compute the relevance score of position i as follow:

$$r_b(i) = \begin{cases} 1 & \text{if } p_i \in H_{test}(b) \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where b is the target brand, p_i is the post ranked at position i and $H_{test}(b)$ are the testing set images posted by brand b. Intuitively, given a brand, a high performance discovery model will rank as high as possible the test images posted by that brand. Thus, in addition to NDCG, we also introduce the metric MedR, which is the median position of the first relevant post retrieved. A low MedR indicates that the first relevant post is ranked in the most relevant results most of the times.

4.2.2 Baselines. Since there are not existing methods specifically designed for content discovery for brands, we compare our method against a set of baselines inspired by pairwise models in image recommendations, which are the closest to PCD.

Random: we generate a random ranking for testing posts.

BrandAVG: we perform nearest neighbor retrieval with respect to brand representation which is the mean feature vector among the image features of all images appearing in the brand's posting history.

DVBPR [21]: Visually-aware Deep Bayesian Personalized Ranking is a pairwise model inspired by VBPR [15], which excludes non-visual latent factors. We adopt the variant with pre-trained image features as described in the paper.

CDL [25]: Comparative Deep Learning is a pure content-based pairwise architecture. We use pre-trained image features and one-hot brand ID as user information.

NPR [29]: Neural Personalized Ranking is one of the most recent pairwise content-based architecture. Since our formulation is a pure content-based scenario, we use image features as the sole item input, using pre-trained image features after PCA as described in the paper.

4.2.3 Implementation details. For all the experiments performed in this paper, we use pre-trained image features extracted with VGG16 network [37] from the first fully connected layer. We select the image preview as input for video posts and selected the first photograph for multiple-images posts. We first optimize the model on the training set, then use it to rank all the 9,270 testing posts for each brand separately.

4.3 Experiment 1: PCD vs Others

In this experiment we evaluate quantitatively the performance of PCD versus the baselines. We report in Table 3 the results of PCD against the baselines.

Our method has the best performance according to all metrics, outperforming all baselines. In particular, PCD obtains the best AUC followed by DVBPR and NPR. The cAUC values are consistently lower than AUC, confirming that discriminating between the subtle differences of competitors is harder than the general problem. The cAUC values are also consistent with AUC across the metrics, with PCD having the highest score. This confirms that our method learns finer-grained brand representation compared to content-based recommender systems.

Table 3: Comparison of PCD with the baselines. We used cut-offs of 10 and 50 for NDCG.

	AUC	cAUC	$NDCG_{10}$	$NDCG_{50}$	MedR
Random	0.503	0.503	0.001	0.003	568
BrandAVG	0.796	0.687	0.068	0.105	29
DVBPR	0.862	0.734	0.059	0.102	20
CDL	0.807	0.703	0.079	0.119	19
NPR	0.838	0.716	0.040	0.076	33
PCD	**0.880**	**0.785**	**0.151**	**0.213**	**5**

Considering the NDCG and Medr metrics, we observe that NPR has inferior performances compared to the other baselines. We believe the reason is that the architecture of NPR is designed for a less sparse collaborative filtering scenario, while all the other methods are natively proposed to rely solely on content.

Finally, PCD has the best NDCG values and the lower MedR, indicating that the learned brand and post embedding have higher capability of discovering a small number of relevant posts in the large test set. The value for MedR for PCD is almost four time smaller than in the case of CDL. This indicates that whenever a social network user will post a relevant content for a brand, our method is far more likely to discover it.

4.4 Experiment 2: Brand Representation Learning

In this experiment we evaluate that explicitly modeling brand associations yields better rankings than directly learning brand representation from one-hot brand ID. For this purpose we defined PCD1H, which is a variation of PCD without the brand representation learning component, learning instead a brand embedding from one-hot brand ID.

In Figure 3 we compare the NDCG of PCD and PCD1H for increasing cut-off values. We notice that PCD values are consistently higher than PCD1H, for both small and high cut-offs. This result confirms the effectiveness of our brand representation learning and the importance of explicitly modeling brand associations.

We observe that PCD has a more marked v-shape, particularly on the left side. For example, considering the cut-off of 1, PCD retrieves a relevant result in the top position for 202 out of 927 brands, while in the case of PCD1H this only happens in 127 cases. This indicates that our brand representation learning is particularly important in order to retrieve relevant posts at the top ten positions of the ranking. The reason for the curves to invert their trends is due to the discount effect of NDCG.

4.5 Experiment 3: Brand Verticals

Different from previous works on brand associations [23], PCD jointly processes all the brands of our dataset with an end-to-end framework. For this reason the evaluation metrics we used in the

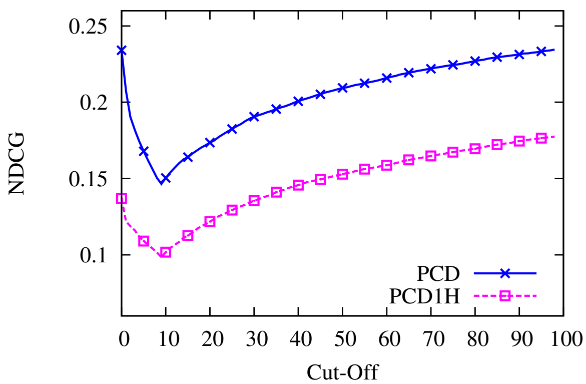

Figure 3: NDCG curves: the major performance of learning brand representation from associations is shown by a higher curve in the case of PCD for all the cut-off points.

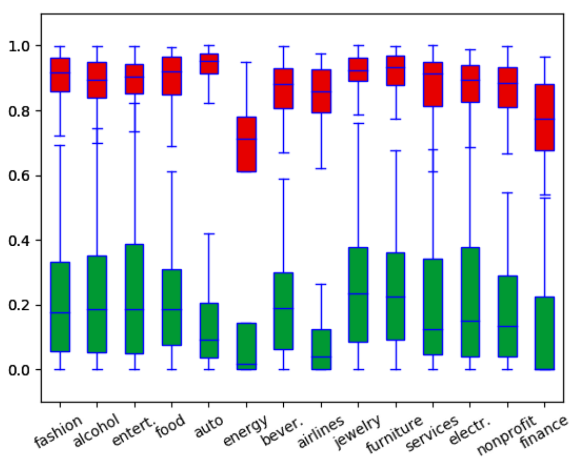

Figure 4: Box plot of performances for brand verticals. AUC and NDCG@50 are represented in red and green boxes respectively.

previous studies average results among the 927 brands. This motivates us in performing an additional analysis to investigate the performance of the model one step deeper in terms of brand verticals.

We computed AUC and NDCG@50 for each brand and plotted these results using respectively a red and green box plot, organized by verticals. Each box in Figure 4 represents the distribution of performance results for all the brands belonging to a certain vertical, such food or alcohol. On boxes it is indicated the median value, the upper and lower quartile and extremes. We omitted outliers for a clearer representation.

We can observe similar performance among brands in fashion, alcohol, entertainment, food, beverage, jewelry and furniture, with a median NDCG and AUC of approximately 0.19 and 0.86 respectively. In terms to NDCG, we notice that content for automobile, airlines, energy and finance brands is the hardest to discover. However, the former two verticals still achieve high AUC, while for the latter ones, these metric is also poor. One possible explanation is that energy and finance brands are the hardest cases, probably because they lack clear recurring visual element. On the other side, since posts from automobile and airline verticals commonly share extremely similar visual elements, we think that these are the hardest brands in terms of fine-grained distinctions between competitors.

4.6 Brand Associations Visualization

A selection of qualitative examples of brand associations is illustrated in Figure 5. In order to visualize associations, we used our trained model to project all training images into the latent space and selected the nearest neighbor posts to the association latent vector using cosine similarity. The example shows some high-level semantics aspects which are captured by brand associations latent vectors. For example dedicate brand associations are automatically learned to represent the visual aspects of coffee cups, sea water, cars, alcohol bottles, rings, cyclists, dogs, fashion items and classical buildings. By looking into a brand's weight vector $\mathbf{x_b}$, we can understand which aspects have the largest positive or negative contribution. For the associations in Figure 5 we retrieved the 50 brands who were positively affected the most. Among these, the top left association affects brands such as *Costa Coffee*, *Starbucks* and *Salt Spring Coffee*, while the one below is adopted by alcohol brands such as *Dom Pérignon* and *Moët & Chandon*. Finally, the top right association is adopted by most of the car manufacturer brands in our dataset, such as *Rolls-Royce*, *Tesla*, *Cadillac* and *Volvo*.

4.7 Case Studies

In order to achieve a deeper understanding of what image posts our model discovers for brands, we offer a qualitative analysis of eight case studies. We selected eight among the most popular brands in our dataset, such that most readers will be familiar with them. For each of them the posts of our testing dataset were ranked using our method. Same as in previous experiment settings, the goal of the model is to rank the ten positive images for each brand higher than the others. We aim to show for each of these brands what kind of posts were correctly discovered, what relevant posts were missed and what are the posts from other brands who erroneously obtained high positions in the ranking. Figure 6 tabulates results for the eight selected cases: beer brand *Carlsberg*, *Qatar Airways*, computer manufacturer *Lenovo*, *Ford* motor company, *Coca-Cola*, Italian fashion brand *Gucci* and video-game companies *Nintendo* and *Ubisoft*. For each brand the first column shows a relevant post that PCD ranked in one of the very first positions of the ranking (true positive), while the central example is another relevant posts for which the model failed to attribute a high ranking score (false negative). Finally, the third column shows an example of a post from another brand (either competitors or not) that erroneously achieved high positions in the ranking (false positives), together with the name of the brand who adopted it.

While the examples in the first column evidently match their brand expected style and content, the ones in the second column are much harder to discover. For example, in the case of *Carlsberg*, the method is able to easily retrieve an image of a beer glass on a green

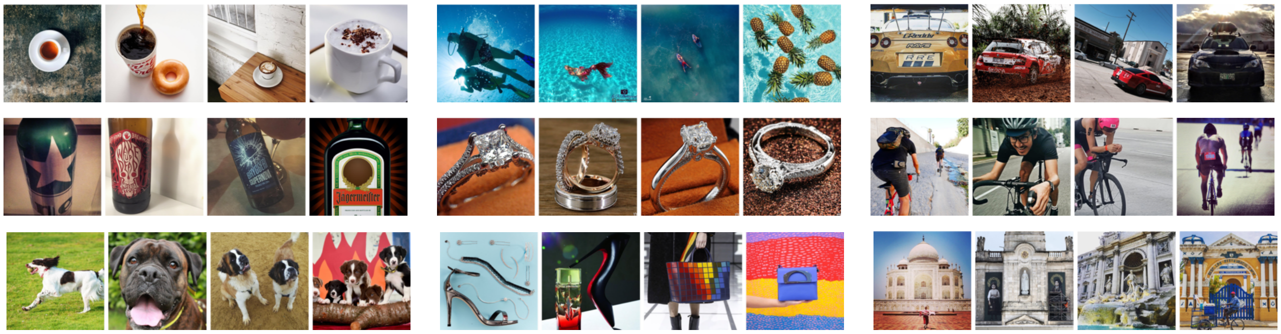

Figure 5: Nine examples of brand associations. For each association, four example images are selected for display from the training set using nearest neighbors.

Figure 6: Case study for eight brands. For each brand, the three columns show one example of true positive, false negative and false positive respectively. Please note that all examples were manually selected, hence this picture has no indication of performance.

background, but fails to retrieve another photograph posted by the brand, featuring two astronauts visiting the beer company on the first anniversary of their space mission. As a false positive example, we selected another picture of a beer glass on a grass background by German beer brand *Astra*. We can easily notice how all the other false negatives in this case studies are to a certain extent sporadic and occasional for the brands, which partially explains why they were missed.

Finally, we observe that not all the false positives are from competitor brands. For example, a post from the financial services company *Allianz* is retrieved for the brand *Ford*, featuring a truck in an outdoor environment. This confirms that in these cases PCD didn't learn high-level category level representation, but rather captured the fine-grained stylistic features.

5 CONCLUSIONS AND FUTURE WORKS

In this work we take inspiration from a real-life marketing problem and introduce the new problem of content discovery for brands in order to automatically filter which photographs in a large set of new posts are likely to be adopted by a brand. A content-based method called Personalized Content Discovery (PCD) is proposed to generate a personalized ranking by learning fine-grained brand representation from their previous posting history. We crawled and released a large-scaled Instagram dataset, which we used to train, evaluate and visualize our model with extensive experiments.

As future works, we are working on incorporating the temporal dimension, since we believe that brands' posting strategies largely depend on context. For example, what post a brand is looking for may depend on a specific marketing campaign featuring current events and trends. Secondly, we will extend our analysis to videos and incorporate signals from multiple modalities, such as text, engagement and social network links.

ACKNOWLEDGMENTS

NExT research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IRC@SG Funding Initiative.

REFERENCES

[1] Hassan Alboqami, Wafi Al-Karaghouli, Yasser Baeshen, Ismail Erkan, Chris Evans, and Ahmad Ghoneim. 2015. Electronic word of mouth in social media: the common characteristics of retweeted and favourited marketer-generated content posted on Twitter. *International Journal of Internet Marketing and Advertising* 9, 4 (2015), 338–358.

[2] Christy Ashley and Tracy Tuten. 2015. Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. *Psychology & Marketing* 32, 1 (2015), 15–27.

[3] Janelle Barlow and Paul Stewart. 2004. *Branded customer service: The new competitive edge.* Berrett-Koehler Publishers.

[4] Chia Yu Chang. 2014. *Visualizing brand personality and personal branding: case analysis on Starbucks and Nike's brand value co-creation on Instagram.* The University of Iowa.

[5] Jingyuan Chen, Hanwang Zhang, Xiangnan He, Liqiang Nie, Wei Liu, and Tat-Seng Chua. 2017. Attentive Collaborative Filtering: Multimedia Recommendation with Item- and Component-Level Attention. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17)*. 335–344.

[6] Tao Chen, Xiangnan He, and Min-Yen Kan. 2016. Context-aware Image Tweet Modelling and Recommendation. In *Proceedings of the 2016 ACM on Multimedia Conference (MM '16)*. 1018–1027.

[7] Lisette de Vries, Sonja Gensler, and Peter S.H. Leeflang. 2012. Popularity of Brand Posts on Brand Fan Pages: An Investigation of the Effects of Social Media Marketing. *Journal of Interactive Marketing* 26, 2 (2012), 83 – 91.

[8] Yi Ding, Chee Wei Phang, Xianghua Lu, Chuan-Hoo Tan, and Juliana Sutanto. 2014. The Role of Marketer- and User-Generated Content in Sustaining the Growth of a Social Media Brand Community. In *2014 47th Hawaii International Conference on System Sciences*. 1785–1792.

[9] Aleksandr Farseev, Kirill Lepikhin, Kenny Powar, Eu Khoon Ang, and Hendrik Schwartz. 2018. SoMin.ai: Social Multimedia Influencer Discovery Marketplace. In *Proceedings of the 26th ACM International Conference on Multimedia (MM '18)*. ACM.

[10] Yue Gao, Fanglin Wang, Huanbo Luan, and Tat-Seng Chua. 2014. Brand Data Gathering From Live Social Media Streams. In *Proceedings of International Conference on Multimedia Retrieval (ICMR '14)*. 169–176.

[11] Francesco Gelli, Xiangnan He, Tao Chen, and Tat-Seng Chua. 2017. How Personality Affects Our Likes: Towards a Better Understanding of Actionable Images. In *Proceedings of the 2017 ACM on Multimedia Conference (MM '17)*. 1828–1837.

[12] Xue Geng, Hanwang Zhang, Jingwen Bian, and Tat-Seng Chua. 2015. Learning Image and User Features for Recommendation in Social Networks. In *International Conference on Computer Vision (ICCV '15)*. 4274–4282.

[13] Sonja Gensler, Franziska Völckner, Yuping Liu-Thompkins, and Caroline Wiertz. 2013. Managing Brands in the Social Media Environment. *Journal of Interactive Marketing* 27, 4 (2013), 242 – 256.

[14] Khim-Yong Goh, Cheng-Suang Heng, and Zhijie Lin. 2013. Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content. *Information Systems Research* 24, 1 (2013), 88–107.

[15] Ruining He and Julian McAuley. 2016. VBPR: Visual Bayesian Personalized Ranking from Implicit Feedback. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*. 144–150.

[16] Xiangnan He, Tao Chen, Min-Yen Kan, and Xiao Chen. 2015. TriRank: Review-aware Explainable Recommendation by Modeling Aspects. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM '15)*. 1661–1670.

[17] Xiangnan He and Tat-Seng Chua. 2017. Neural Factorization Machines for Sparse Predictive Analytics. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17)*. 355–364.

[18] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. 173–182.

[19] Balázs Hidasi, Massimo Quadrana, Alexandros Karatzoglou, and Domonkos Tikk. 2016. Parallel Recurrent Neural Network Architectures for Feature-rich Session-based Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*. 241–248.

[20] Sudarsan Jayasingh and Venkatesh Rajagopalan. 2015. Customer Engagement Factors in Facebook Brand Pages. *Asian Social Science* 11 (08 2015).

[21] Wang-Cheng Kang, Chen Fang, Zhaowen Wang, and Julian McAuley. 2017. Visually-Aware Fashion Recommendation and Design with Generative Image Models. In *International Conference on Data Mining (ICDM '17)*. 207–216.

[22] Gunhee Kim and Eric P. Xing. 2013. Discovering Pictorial Brand Associations from Large-Scale Online Image Data. In *2013 IEEE International Conference on Computer Vision Workshops (ICDMW '15)*. 404–411.

[23] Gunhee Kim and Eric P. Xing. 2014. Visualizing Brand Associations from Web Community Photos. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining (WSDM '14)*. 623–632.

[24] Jingu Kim, Renato D. C. Monteiro, and Haesun Park. 2012. Group Sparsity in Nonnegative Matrix Factorization. In *Proceedings of the 2012 SIAM International Conference on Data Mining (SDM '12)*. 851–862.

[25] Chenyi Lei, Dong Liu, Weiping Li, Zheng-Jun Zha, and Houqiang Li. 2016. Comparative deep learning of hybrid representations for image recommendations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '16)*. 2545–2553.

[26] Tereza Litsa. 2016. How the rise of user-generated content changes marketing. (October 2016). https://www.clickz.com/how-the-rise-of-user-generated-content-changes-marketing

[27] Shuang Ma and Chang Wen Chen. 2018. D-Sempre: Learning Deep Semantic-Preserving Embeddings for User interests-Social Contents Modeling. *CoRR* abs/1802.06451 (2018).

[28] Masoud Mazloom, Robert Rietveld, Stevan Rudinac, Marcel Worring, and Willemijn van Dolen. 2016. Multimodal Popularity Prediction of Brand-related Social Media Posts. In *Proceedings of the 2016 ACM on Multimedia Conference (MM '16)*. 197–201.

[29] Wei Niu, James Caverlee, and Haokai Lu. 2018. Neural Personalized Ranking for Image Recommendation. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*. 423–431.

[30] Steve Olenski. 2016. Can Your Content Sell A Lifestyle And Not Just Products? (July 2016). https://www.forbes.com/sites/steveolenski/2016/07/02/can-your-content-sell-a-lifestyle-and-not-just-products

[31] Gijs Overgoor, Masoud Mazloom, Marcel Worring, Robert Rietveld, and Willemijn van Dolen. 2017. A Spatio-Temporal Category Representation for Brand Popularity Prediction. In *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval (ICMR '17)*. 233–241.

[32] Irena Pletikosa Cvijikj and Florian Michahelles. 2013. Online engagement factors on Facebook brand pages. *Social Network Analysis and Mining* 3, 4 (2013), 843–861.

[33] Steffen Rendle. 2010. Factorization Machines. In *Proceedings of the 2010 IEEE International Conference on Data Mining (ICDM '10)*. 995–1000.

[34] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI '09)*. 452–461.

[35] Ruslan Salakhutdinov and Andriy Mnih. 2007. Probabilistic Matrix Factorization. In *Proceedings of the 20th International Conference on Neural Information Processing Systems (NIPS'07)*. 1257–1264.

[36] Martin Saveski and Amin Mantrach. 2014. Item cold-start recommendations: learning local collective embeddings. In *Eighth ACM Conference on Recommender Systems (RecSys '14)*. 89–96.

[37] Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations (ICLR '14)*.

[38] Yao Wu, Christopher DuBois, Alice X. Zheng, and Martin Ester. 2016. Collaborative Denoising Auto-Encoders for Top-N Recommender Systems. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining (WSDM '16)*. 153–162.

[39] Wenhui Yu, Huidi Zhang, Xiangnan He, Xu Chen, Li Xiong, and Zheng Qin. 2018. Aesthetic-based Clothing Recommendation. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. 649–658.

[40] Matthew D. Zeiler. 2012. ADADELTA: An Adaptive Learning Rate Method. *CoRR* abs/1212.5701 (2012).

[41] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. 2016. Collaborative Knowledge Base Embedding for Recommender Systems. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. 353–362.