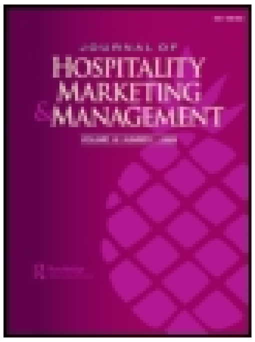


Journal of Hospitality Marketing & Management

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/whmm20

Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: implications for hospitality marketers

Meng Ren , Huy Quan Vu , Gang Li & Rob Law

JOURNAL OF HOSPITALITY MARKETING & MANAGEMENT
https://doi.org/10.1080/19368623.2020.1765226

Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: implications for hospitality marketers

Meng Ren[a], Huy Quan Vu [b], Gang Li [c,d] and Rob Law [e]

[a]School of Computer Science, Xi'an Shiyou University, Xi'an, China; [b]Department of Information Systems and Business Analytics, Deakin University, Melbourne, Australia; [c]School of Information Technology, Deakin University, Geelong, Australia; [d]Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Urumqi, China; [e]School of Hotel & Tourism Management, The Hong Kong Polytechnic University, Hong Kong SAR, China

ABSTRACT

With the prevalence of social media and Web 2.0, online visual contents such as photos or videos have quickly evolved into one popular information-disseminating channel for hotel managers and travelers. The current study aims to obtain a comprehensive understanding of the preconceptions as reflected in online photos posted by travelers. This paper presents a novel approach to online photo content analysis based on deep learning theory and computer vision framework, which can comprehensively analyze the content of large-scale photo datasets. We demonstrate and evaluate this approach through a case study, wherein we analyze over 53,000 photos collected from hotel review platform, TripAdvisor. We identified interesting differences in the contents of photos posted by hotel managers and travelers, including the differences in photo contents between low- and high-rating hotels. Our findings provide valuable implication for hotel marketing using visual assets.

随着社交媒体和网路 2.0的流行，在线的视觉内容，如照片或视频，已经迅速演变成一个流行的信息传播渠道，为酒店经理和旅客. 目前的研究旨在全面了解旅行者在网上发布的照片中所反映的先入为主的观念. 本文提出了一种基于深度学习理论和计算机视觉框架的在线照片内容分析新方法，可以全面分析大型照片数据集的内容. 我们通过一个案例研究来演示和评估这种方法，其中我们分析了从酒店点评平台TripAdvisor上收集的53,000多张照片. 我们发现了酒店经理和旅行者发布的照片内容的有趣差异，包括低等级酒店和高等级酒店的照片内容的差异. 我们的研究结果为利用视觉资产进行酒店营销提供了有价值的启示.

KEYWORDS
Hotel photo; content analysis; visual feature; deep learning; computer vision

Introduction

As one of the world's major industries, tourism has been playing a vital role in the metabolism of the global economy (Law et al., 2011). People are now traveling more than ever due to the development of transportation systems and the economic growth of many countries (Chou, 2013). As a result, the hotel industry has also witnessed continuous prosperity as a major component of the tourism sector (Yip, 2017). Hotel is an important component of travel

CONTACT Gang Li gang.li@deakin.edu.au School of Information Technology, Deakin University, Geelong, VIC 3216, Australia

experiences, which moderately impact travelers' satisfaction (Xiang et al., 2015). A good hotel experience will lead to a friendly relationship between hotel and travelers, whereas unfavorable hotel experiences could negatively influence the experience of travelers (Xia et al., 2018). Thus, selecting a suitable accommodation is important for travelers who are planning for future trips (Li et al., 2013).

The development of the Internet has changed the way travelers seek travel information (Lee et al., 2013). Travelers tend to rely on the information available on various travel-related websites as a ground for purchase decisions (Kim et al., 2017). Online visual contents, especially photos, strongly influence traveler's emotions and purchasing decisions (Chuang, 2007). Well-designed websites with appropriate visual content aid hotel managers in obtaining increased customer recognition and loyalty (Bai et al., 2008). In addition to online photos posted by hotel managers for marketing purpose, the purchasing decision of travelers is influenced by the online photos, which were captured and posted by other travelers (Leung & Bai, 2013). Travelers' photos capture genuinely the various physical representations of specific tourism experience attributes (Garrod, 2007). Understanding the contents captured from those photos provides hotel managers with insights into travelers' perception and preferences (Trpkovski et al., 2018), which are useful for developing effective marketing materials and online visual contents to attract future customers. However, comprehensive online photo content analysis is a challenge because online photos are usually available at large scales, and the captured contents are often diverse. Manual analysis is time consuming and limited to a few photos (Ma & Takagi, 2012). A method that can help tourism researchers analyzes the contents of online travel photos at a large scale has not been reported. Tourism managers are still facing challenges in obtaining comprehensive understanding of the travelers' preconceptions as reflected in their posted photos. In other words, important questions, such as the following have not been thoroughly answered, "*What are the differences between the projected image in the promotional photos of hotel managers and the perceived image reflected in the photos of travelers?,*" "*What are the major themes captured in travelers' photos?,*" or "*What features of promotional photos could be improved to increase the attractiveness of hotels?.*"

The theoretical development of computer science, especially in the computer vision field, has enabled the recent development of advanced photo processing and analysis techniques (Ma et al., 2018). The color, shape, and semantic concepts in photos could be automatically extracted and mathematically presented as visual features (Tian et al., 2012). Such visual features are then used as inputs into various machine-learning algorithms, such as multi-label image classification algorithm (Yang et al., 2018), support vector machine (Garcıa-Floriano et al., 2019), and neural network for automatic recognition of photo contents (Göring et al., 2018). Recently developed techniques, such as recurrent neural networks (Trabelsi et al., 2019), deep neural network (DNN) (Yao et al., 2012), convolutional neural network (CNN) (Kao et al., 2015), and deep belief network (DBN) (Al-antari et al., 2017), have allowed a deep understanding of visual content in photos with superior performance. However, prior attempts have scarcely employed this theory to analyze travel photos for tourism applications. Recent attempts by Zhang et al. (2019, 2020) focused on general travel photos in a tourism destination. Specific applications of deep learning on hotel photo analysis have not been demonstrated to address the above-mentioned research questions.

The current study aims to obtain a comprehensive understanding of the preconceptions as reflected in online hotel photos posted by travelers by adopting the recent theory of deep learning. We introduce and evaluate a framework that can automatically recognize the contents of online travel photos, especially those about hotels. We selected Macau as a target tourism destination for our study and extracted hotel photos as our dataset. Macau has a rich Chinese and Portuguese heritage, which includes many outstanding examples of Western and Oriental art and culture (Leong & Li, 2010). Recent developments in Macau, especially in hotels and gaming industries, have turned it into a popular tourism destination among domestic and international travelers. More than 32 million tourists arrived Macau in 2017 (Chen, 2019). We analyzed a large-scale photo dataset of more than 53,000 photos for 75 hotels. This dataset provides an overview of the popular contents captured in the hotel photos. Contrast analysis revealed interesting differences in projected image of hotel managers and perceived images of travelers, which are reflected in the photo contents. Exploratory analysis of visual contents and color themes in the photos provides valuable guidance for hotel managers to improve the attractiveness of their hotel websites.

This study distinguishes itself in adopting deep learning theory to examine the travelers' preconceptions as reflected in their posted photos. Moreover, it contributes to the literature of visual assets in the tourism and hospitality industry. The introduced techniques are beneficial to researchers studying large-scale online hotel photos. The findings are useful for hotel managers, especially those in Macau, in developing effective marketing materials using visual assets to attract and retain potential customers.

Literature review

This section first provides background on the relationships between photos and traveler perceptions. Representative works on online photo analysis in the hospitality and tourism literature are reviewed, with a critical analysis of their limitation. On the basis of the computer vision framework, the motivation for adopting deep learning theory into photo content analysis is then elaborated.

Photos and travelers' perception

Photography and tourism are widely regarded as intrinsically related because photos play a vital role in forming destination image and promoting tourism destination (Garrod, 2009). Dobni and Zinkhan (1990) referred to images as the sum of beliefs, ideas, or an overall impression of an individual on specific products or services. The image of a destination is formed through the process of collecting information about the destination via different information sources, such as promotion, marketing materials, online reviews, and travel photos (Tapachai & Waryszak, 2000).

For economy hotel brands and online travel agents, visual assets like photos are important marketing and promotion resources for hospitality and tourism products due to the low monetary cost and high benefits in advancing the emotional or hedonic experience with the brand. Kwortnik and Ross (2007) showed that positive emotions could be stimulated in response to the photo and provide a mental experience of a trip before the actual one takes place. Ert et al. (2016) found that travelers tend to judge the host's trustworthiness based on photos posted online. The visual content is robust and has an additive effect on trust building. Photos have cognitive and affective components (Stern et al., 2001), which influence travelers' perception. Numerous studies have been performed on the visual data in holiday brochures and postcards (Markwick, 2001) for the effective communication with target markets and promotion of resources to attract visitors. Moreover, customers have been recognized to spend excessive time reading those visual contents to assist in making their own decisions (Pritchard, 2001).

Tourism services are intangible products; thus, the images of a tourism destination or products projected in online spaces greatly influence travelers' perception (Gallarza et al., 2002). Hospitality and tourism managers tend to introduce promotional and marketing materials, which direct travelers toward a set of favorable experiences to create positive images about the destination (Govers et al., 2007). However, the promotion of destination image may result in negative evaluation when the actual settings or experiences encountered by travelers are significantly different from their expectation (Fairweather & Swaffield, 2002). Understanding travelers' perception is important for tourism managers in designing appropriate marketing materials that can attract and retain customers. Studying online travel photos posted by travelers is an effective approach to study travelers' perception because taking photos is an essential part of trip activities, and it closely reflects travel experiences (Donaire et al., 2014). Studies on destination photos have received much attention from hospitality and tourism researchers (Pan et al., 2014). However, they focus mainly on a few photos produced by destination management organizations or businesses, rather than those generated by travelers themselves. Few studies are devoted to examining the online travel photos to improve the understanding of the travelers' perceptions and travel experiences.

Online photo analysis

Content analysis is a popular approach in existing works attempting to analyze online photos for insights into traveler's perceptions about destinations. For instance, Pan et al. (2014) analyzed photo contents to uncover the connection among a destination's motivation, image dimensions, and affective qualities. Garrod (2009) combined content analysis and quantitative statistical techniques to discover the relationship between tourism destination imagery and tourist photography. Stepchenkova and Zhan (2013) compared the contents of photos posted by travelers and marketers in Peru to identify the differences between projected and perceived destination images. On the one hand, previous researchers found that photos taken by travelers are likely to contain daily activities, plants, domesticated animals, and food, rather than photos posted by marketers. On the other hand, marketers' photos tend to include traditional clothing, art objects, festivals, and ritual into their photos for marketing purposes. A similar approach to content analysis was adopted in a case study of Japan. Song and Kim (2016) found that the most dominant theme among the photos taken by marketers is modern architecture, whereas the most popular theme among the photos posted by travelers onto photo-sharing site Pinterest is nature and landscape. Hao et al. (2016) combined visual content analysis and photo interpretation for an outdoor tourism performance in China.

Realizing the influence of visual contents on travelers' emotion and purchasing intention, Negri and Vigolo (2015) regarded photos as an important asset in marketing and promotional strategies of hotel managers (Negri & Vigolo, 2015). Several attempts to analyze the effectiveness of advertising photos have been carried out to improve and reinforce hotel brand images. For instance, Jeong and Choi (2005) analyzed photo contents and measured the influence of picture presentations in a hotel website on customers' attitudes and intentions. They found that customers tend to have favorable attitudes toward a hotel website in the presence of various photos (Jeong & Choi, 2005). Bender Stringam and Gerdes (2010) focus on exploring the impression of users about hotel websites' visual contents, such as pictures, color, and other information. The presence of photographs on a hotel website was identified as the most significant factor affecting site appeal and booking decision (Ert et al., 2016). Content analysis of hotel photo is still not a well-studied topic. Existing works have mainly focused on photos available on hotel websites or posted by hotel marketers. Numerous hotel photos captured and posted by travelers on online social platforms have not been utilized in prior works. As such, no comprehensive view exists regarding the difference between projected image in the promotional photos and the perceived image reflected in travelers' photos. This shortcoming is probably due to the limitation of the manual content analysis, which cannot handle large-scale photo datasets.

Computer vision

Computer vision is a branch of computer science concerned with the theory and technology that enable computers to gain a high-level understanding of visual contents, such as digital photos or videos (Sonka et al., 2014). The aim of computer vision is on automatic extraction (Nixon et al., 2011), analysis, and understanding of useful visual information similar to a human visual system. The development of computer vision systems has provided increased flexibility and further automation options to manufacturers, thereby helping them find defects, sort products, and complete tasks faster and more efficiently than humans alone ever could (Tzimiropoulos & Pantic, 2017). Computer vision techniques have also been applied to diagnostic and surgical systems for biomedical image processing (Dholey et al., 2017). Other industrial applications of computer visions are postal automation and fruit quality evaluation for agriculture production (Cavallo et al., 2019), traffic monitoring (Jing et al., 2017), and security and surveillance systems (Hadjkacem et al., 2018).

Early applications of computer vision often utilized handcrafted visual features, which extract and represent visual information in a mathematical form for processing and analysis. Examples of these features are local binary pattern (Ojala et al., 2002), local ternary pattern (Tan & Triggs, 2007), and local phase quantization (Ojansivu & Heikkilä, 2008). Advanced techniques for visual feature extraction have been developed for complex visual analysis applications, such as scale-invariant feature transform (Lowe, 2004), speeded up robust features (Bay et al., 2008) and other shape and color features for photo quality assessment (Trpkovski et al., 2018). These visual features are often used as inputs into machine-learning algorithms for various automatic photo analysis tasks, such as object recognition (Napiorkowska et al., 2018) and classification (Hu et al., 2018). The aids from machine learning help recognizing and labeling photo contents much quicker than manual approaches, which allows large-scale photo data set to be processed and analyzed efficiently.

Despite the greater potential, computer visions have not been applied widely into tourism and hospitality areas. The barrier is probability due to the diversity of the contents captured in travel photos. The applications of computer vision in other domains often involve specific types of images, such as medical images (Dholey et al., 2017), license plate photos (Jing et al., 2017), and fruit photos (Cavallo et al., 2019). Such analyses are usually limited to small number of specific objects, which are predetermined by researchers. On the other hand, the analysis of travel photos contents can involve many possible objects, scenes, and themes, and researchers have no prior knowledge on what to expect from those photos. This challenge also presents great opportunities for tourism and hospitality researchers, as discover new and useful understanding about travelers' perception and experiences could be discovered if large-scale travel photo data set can be properly processed and analyzed.

Theoretical foundation of deep learning

Machine-learning provides the key methods for content recognition in most computer vision applications. In fact, machine-learning techniques have been utilized in various tourism and hospitality applications but mainly for textual data analysis, such as hotel reviews using topic model and decision tree (Calheiros et al., 2017), and restaurant reviews using random forest (Lee & Whaley, 2019; Lu & Stepchenkova, 2015). More advanced learning methods, such as support vector machine (Garcıa-Floriano et al., 2019) and neural network (Göring et al., 2018) were popularly used for photo analysis, due to their ease of integration and reasonable performance in most ordinary object recognition applications. However, one issue with these theories is their dependence on the feature engineering process, where visual features are extracted from raw photo data (Bay et al., 2008; Ojansivu & Heikkilä, 2008; Tan & Triggs, 2007; Trpkovski et al., 2018). This process often requires domain knowledge about a given problem and predetermination of features to be extracted, which is impractical for the case of travel photos analysis with diverse contents.

Recent advances in the theory of deep learning have enabled us to address above-mentioned challenges. Theoretically, deep learning is an artificial neural network model with multilayer architecture (Schmidhuber, 2015). The model construction is carried out with multiple neural networks stacked on top each other to progressively and automatically extract high-level features from raw input (L. Deng & Yu, 2014; Yao et al., 2012). Feature engineering with predetermined features is not required in deep learning theory. Due to the advanced modeling capability, deep learning has outperformed traditional machine learning and produced results comparable to and in some cases superior to human experts (Russell, 2017). In the past decade, deep learning theory has been successfully applied in various fields, such as natural language processing, audio signal processing, manufacturing, and particularly computer vision (Ciregan et al., 2012). Some applications of deep learning in photo processing and analysis are personalized recommendation of photography (Ji et al., 2019), location detection from the photos (N. Deng & Li, 2018), and object detection in satellite imagery (Napiorkowska et al., 2018).

Despite its great potential in the field of tourism, deep learning has only been adopted in very recent studies. In their seminal study, Law et al. (2019) developed a deep learning framework for tourism demand forecasting, which showed a significant increase in accuracy compared with that of other models. Recently, Y. Zhang et al. (2020) further developed the theory with group pooling strategy to address the challenge of overfitting. Similarly, Ma et al. (2018) introduced deep learning to evaluate online hotel reviews and photo data, but the work did not focus on the analysis of photo contents. Instead, visual features in photos were used to improve the performance of their prediction model on the helpfulness of online reviews. Some attempts were made recently to analyze photo contents using deep learning (Zhang et al., 2019), though their work focused on the general analysis of travel photos rather than hotel photos for hotel marketing and management. As such, this study proposes to adopt deep learning theory in online hotel photo analysis, which can further support practitioners in utilizing online photos to gain comprehensive insights into traveler perceptions and improve hotel marketing practices.

Methodology

This section presents our framework for hotel photo content analysis on the basis of deep learning theory. Our framework comprises three major stages: *Online Hotel Photo Extraction*, where a set of hotel photos is collected from online review platforms; *Deep Learning of Photo Content*, where the theory of deep learning is applied to automatically recognize the content captured in the photos; and *Exploratory Analysis*, where photo contents are explored based on descriptive statistics, contrast analysis, and graph visualization. The subsequent sections present additional details of each stage.

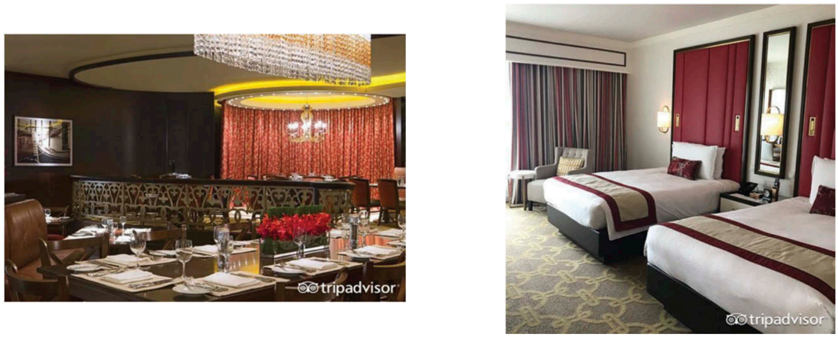
Online hotel photo extraction

This stage collected data for our study, which involves the extraction of hotel photos posted on online review platforms. Hotel photos are available on various platforms where travelers usually share their travel experiences. Examples of these platforms are TripAdvisor, Booking, Facebook, and Yelp. These platforms are popular resources for tourism researchers (Li et al., 2015; Liu et al., 2013). In this paper, we utilized TripAdvisor as our main data source for demonstration purpose. Photos on TripAdvisor reliably capture traveler experiences and perceptions (Negri & Vigolo, 2015). We collected photos posted by hotel officials and travelers. This method is convenient for comparative analysis of projected and perceived images between hotel marketers and travelers. Numerous photos are available on TripAdvisor. Hence, we developed a web crawler to automatically navigate through the photo gallery of each hotel and download photos. We downloaded the actual photos because their contents are the main source in our study.

Deep learning of photo content

We incorporated deep learning techniques into our framework for automatic analysis of photo content. Deep learning models require considerable data with label annotations for training purpose. Thus, building deep learning models from scratch is not practical in real application. Fortunately, varieties of deep learning tools with pre-trained models, such as Google Vision API (https://cloud.google.com/vision/), are available. This tool was developed based on an open-source library for machine learning developed by Google, named TensorFlow (www.tensorflow.org/). Deep learning techniques, such as CNN and DBN, were integrated into this library. Thousands of deep learning models exist in Google Vision API. These models were pre-trained on a large-scale dataset of approximately 920 million photos. The use of pre-trained models allows researchers to save time in constructing models, whereas the recognition performance is superior to most traditional machine-learning methods (Felzenszwalb et al., 2014).

The inputs into the toolbox were photos, whose contents were analyzed using pre-trained models to produce estimation scores corresponding to various entities, such as categories, scenes, objects, color themes, and human faces. Figure 1 shows two photos posted on an online review platform, whereas Table 1 shows the estimated score for the



(a) Hall (b) Bed Room

Figure 1. Online hotel photos (Source: TripAdvisor).

Table 1. Recognized entities in hotel photos.

Hall photo	Score	Bedroom photo	Score
Restaurant	0.917	Room	0.929
Interior design	0.778	Suite	0.876
Dining room	0.745	Interior design	0.777
Function hall	0.742	Hotel	0.769
Ceiling	0.550	Bed frame	0.728
Table	0.513	Ceiling	0.684
		Bedroom	0.567
		Window treatment	0.548
		Window	0.514
		Curtain	0.504

entities recognized in those photos. The entity scores ranged from 0 (least confidence) to 1 (most confidence). The first photo of a hotel hall had multiple related entities, describing the scene (e.g., restaurant, interior design, and dining room) and the objects (e.g., table). Similarly, the second photos of the bedroom contained various entities relating to a bedroom scene. The estimation score tends to be correlated to the occupation of the entity in the photo. For instance, the entity table was a component of a bigger entity dining room, as reflected in the photo scene. Therefore, the estimation score for the table was less than the dining room. Score entities varied according to individual photos as well as the scene and setting of the content. Treating a photo as having a particular entity is a common practice if its score is more than 0.5 (Tan et al., 2019).

Exploratory analysis

Once the entities were recognized from the photos for the entire data collection, we carried out numerous analyses to explore the photo contents. First, we computed a support value to describe how frequent an entity is in the photo collection. Suppose $P = \{p_1, p_2, \ldots, p_m\}$ is a photo collection having m photos. Each photo contained a set of entities $e_1, e_2, \ldots$ as identified by the deep learning algorithm, $e \in E$, where E is a list of all possible entities in the entire photo collection. The support of an entity e_i in the photo collection P was determined as follows:

$$Supp(e_i, P) = \frac{|e_i \in P|}{|P|}$$

where $|e_i \in P|$ is the count of photos that contain an entity e_i, and $|P| = m$ is the total number of photo in the collection. Here, $Supp(.)$ can be used to rank the entities based on their occurrence in the photos, so that the researcher can identify the objects frequently captured in the photos.

In the context of online hotel photo analysis, hotel managers would be interested in identifying the difference between the photos posted by travelers and the photos they posted online for marketing purpose so that adjustments can be made. We can apply the contrast analysis technique to identify such differences automatically (Li et al., 2015). Suppose G_x and G_y are two groups of photos to be compared. The difference in support for an entity e_i between two photo groups was measured by the following:

$$Diff_{Supp\left(e_i,\left(G_x,G_y\right)\right)} = \left|Supp(e_i, G_x) - Supp\left(e_i, G_y\right)\right|$$

$Diff_{Supp(e_i,(G_1,G_2))}$ value can be used to identify those entities whose occurrence is the most different between two groups of photos. Furthermore, we also computed another value called *GrowthRate*, which was determined by the following:

$$GrowthRate\left(e_i, G_x, G_y\right) = \begin{cases} 0, \ if \ supp(e_i, G_x) = 0 \ and \ supp\left(e_i, \ G_y\right) = 0 \\ \infty, \ if \ supp(e_i, G_x) = 0 \ and \ supp\left(e_i, G_y\right) \neq 0 \\ \frac{supp(e_i,G_x)}{supp\left(e_i,G_y\right)}, \qquad otherwise \end{cases}$$

GrowthRate measures how many times an entity are more popular in one group than the other. $GrowthRate\left(e_i, G_x, G_y\right) = \infty$ means that the entity e_i appears in one group but not in the other, which can reflect the interesting difference in photo contents between groups. Further statistical tests can be applied to verify significant differences.

In addition, we applied a visualization technique, based on graph theory (Otte & Rousseau, 2002) to visualize the relationship between entities in the photos. In the graph, entities are regarded as nodes, and the co-occurrence of entities in photos is regarded links. The applications of graph theory in hospitality and tourism are often in a form of network analysis (Tran et al., 2016). We adopted this technique mainly for visualization purpose of entity relationship in this paper.

Case study

This section presents a case study on online hotel photo analysis to demonstrate our introduced approach using deep learning. Details about our data collection are first provided, followed by the presentation of result analysis and discussions on practical implications.

Data collection

The data collection process involved the extraction of hotel photos from online review platforms using a web crawler, as outlined in Section 3.1. Our method could be applied to analyze hotel photos at any location. However, we focused on Macau as our target destination for demonstration purpose because the finding can be beneficial to the destination's growing hotel industry. The crawler navigated through the gallery page of each hotel page to extract photos and associated hotel information. In total, 53,813 photos were collected from 75 hotels in Macau. These data were significantly more than the datasets included in prior works on hotel photo analysis. It should be noted that a unique identification number is used to index each photo on the platform. Our data collection process ensures that each photo with unique photo identification number is only downloaded once. Some photos may have similar or identical content as user may take multiple photos at the same time. We treat them as individual photos, as travelers in fact took them. Considering the large scale of the collected data set, minor noise does not influence the overall results of discovered patterns. Table 2 shows the statistic of our photo dataset. Among the collected photos, hotel managers posted 2,397 photos for marketing purpose, whereas travelers posted 51,416 photos. Four- and five-star hotels have more photos than two- and three-star hotels, which was probably because high-rating hotels are usually at larger scales and thus could accommodate more travelers than their low-rating counterparts. We applied a deep learning technique using Google Vision API to identify photo contents and then carried out numerous exploratory analyses to demonstrate our method in providing insights into travelers' experiences and support marketing material development.

Visual content exploration

The entire photo dataset was input into Google Vision API for entity recognition. A photo was treated as containing certain concepts if its entity score was at least 0.5, which is a default threshold in most classification problem (Tan et al., 2019). A total of 2,703 different entities were recognized in the content of our photo dataset. Most photos produced multiple entities as a reflection of their contents. The support of each entity was then computed with respect to the number of photos in which it appeared. A support threshold of 0.05 was applied to remove the infrequent entity, while keeping the frequent ones for subsequent analysis. We adopted this support threshold as it was commonly used in previous works to identify frequent items in a data set (Li et al., 2015). We also removed entities with a general meaning, such as property, real estate, and home, as they were not useful in supporting the analysis of hotel features. We were left with 182 entities with specific meaning for subsequence analyses.

Figure 2 shows the top 50 entities and their corresponding supports from the most to the least frequent. The identified entities described specific hotel features. Examples of the

Table 2. Hotel photo dataset.

Photo owner	Star rating	No. of hotels	No. of photos
Hotel Manager	Two	8	105
	Three	19	222
	Four	26	849
	Five	22	1,221
	Total	75	2,397
Traveler	Two	8	517
	Three	19	3,627
	Four	26	15,249
	Five	22	32,023
	Total	75	51,416

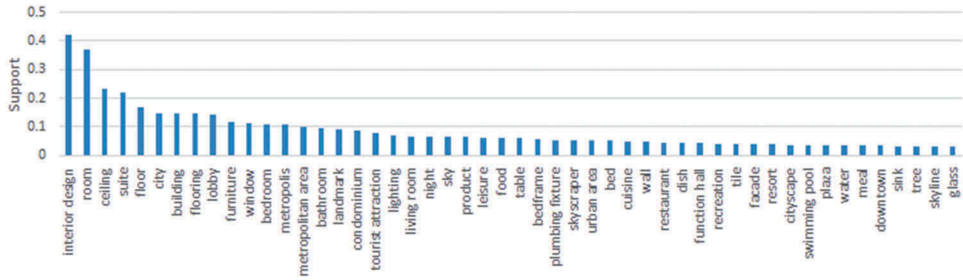
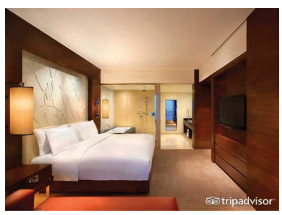


Figure 2. Popular entities captured from online hotel photos.

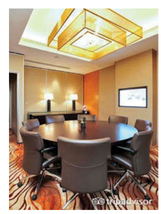
most frequent entities were interior design, room, ceiling, suite, and floor with high support values. These outcomes were consistent with the fact that such entities were common indoor hotel features, which were likely to be captured in the photos taken inside hotels. Notably, other entities describing outdoor scene of hotels, such as a landmark, night, sky, skyscraper, and urban area, were also identified.

We further viewed the actual photos with a high estimation score for several entities to further understand the scene and context where such entities appeared. Figure 3 shows sample photos with entity scores. Each photo might contain multiple entities. Hence, we only displayed the entity with the highest score in the caption. For example, the photo in Figure 3a had the highest score for entity interior design at 0.889, whereas the photo in Figure 3b had the highest

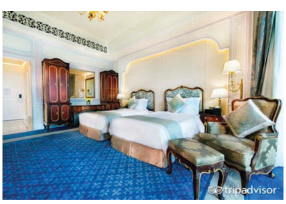
(a) interior design: 0.889



(b) room: 0.932



(c) ceiling: 0.876

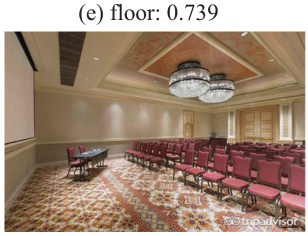


(d) suite: 0.805

(e) floor: 0.739

(f) city: 0.920

(g) building: 0.894



(h) flooring: 0.666

(i) lobby: 0.925

Figure 3. Sample photos with high entity score .

score for room entity at 0.932. The identified entity in the photos agreed with the common sense of human interpretation. Furthermore, deep learning could distinguish scene with similar contents such as an ordinary hotel room (Figure 3b) versus hotel suite (Figure 3d) or a plain floor (Figure 3e) versus flooring (Figure 3h). Such differences in the semantic meaning are high-level concepts, which could be easily overlooked by humans in manual analysis. Automatic analysis of photo content based on deep learning effectively identifies all possible semantic meaning in the photo contents.

One photo might contain multiple entities. Hence, we further analyzed the co-occurrence of entities to improve our understanding of the overall themes reflected in the hotel photo collections. Graph theory is adopted to visualize the relationship among entities in photos. A matrix of co-occurrence entity was first computed, where the value for each element represents the frequency of co-occurrence between two entities. The relationships between the nodes are then visualized using an undirected graph. The size of the node represented the popularity of the entity, and the thickness of link presented the frequency of concurrent. A big or thick size indicated high values. For easy interpretation, we set a frequency threshold of a link to display only the nodes with the most popular links. The frequency threshold was experimentally determined depending on the collected dataset. Lower values would include more nodes and links than higher values. The threshold corresponding to 5% of the maximum link value resulted in a graph showing the relatively clear structure of entity relationships, as in Figure 4. The graphs had two major clusters, one at the top right and another one on the left. Thus, the majority of photos capture indoor and outdoor scenes. In addition, the cluster at the bottom right indicated that numerous photos in our dataset were about food. Hence, dining is an important part of

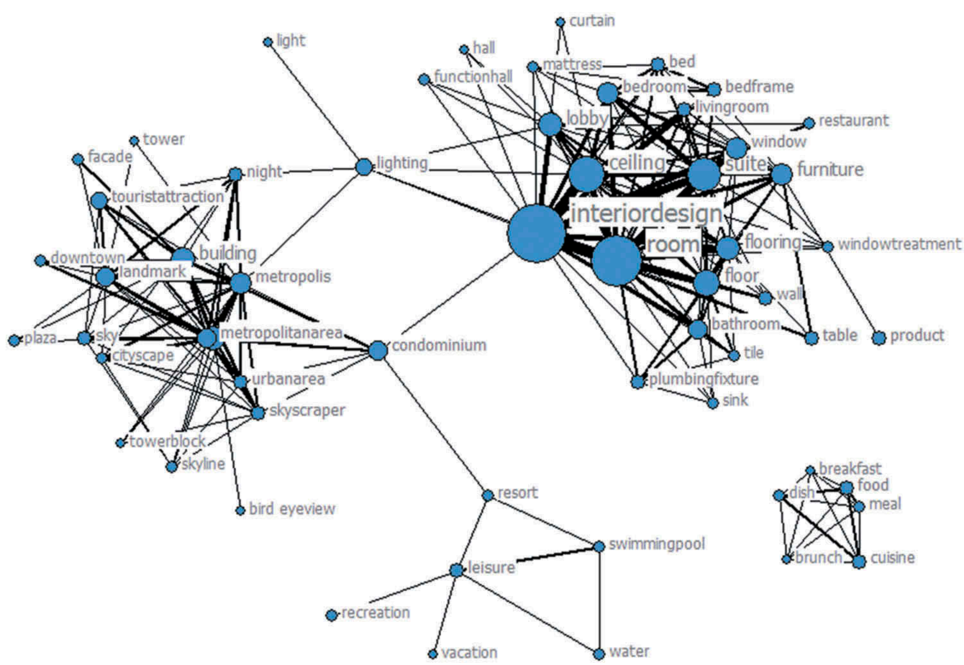


Figure 4. Entity-relationship visualization.

Table 3. Photo contents of hotel managers versus travelers.

Entity	Manager	Traveler	Difference	Growth Rate	Z-score	P-value*
Interior design	0.680	0.409	0.272	1.665	25.952	0.000
Ceiling	0.469	0.224	0.245	2.096	27.315	0.000
Lobby	0.290	0.136	0.154	2.130	20.723	0.000
Restaurant	0.174	0.037	0.137	4.673	31.917	0.000
Suite	0.332	0.216	0.116	1.538	13.075	0.000
Function hall	0.150	0.038	0.112	3.957	26.215	0.000
Living room	0.156	0.063	0.093	2.467	17.489	0.000
Room	0.417	0.366	0.051	1.140	4.664	0.000
City	0.054	0.153	0.099	2.836	−13.489	0.000
Building	0.074	0.151	0.077	2.043	−10.589	0.000
Metropolis	0.050	0.112	0.061	2.215	−9.568	0.000
Furniture	0.062	0.120	0.058	1.946	−8.842	0.000
Bathroom	0.047	0.097	0.050	2.067	−8.346	0.000

*Significant at p-value ≤ 0.05.

hotel experiences. Another cluster at the bottom of the graph indicated that a sizable set of photos capture the leisure experience of tourists, especially at resorts and swimming pools.

Contrast analysis of visual contents

This section examines the differences in the captured content between the photos of hotel managers and travelers. The support values for each entity were computed with respect to the number of photos in each group. The difference and growth rate were then computed following the method in Section 3.3. Table 3 shows the entity with the largest differences in support values (0.05 or more). The statistical significance of the differences was verified via z-test. A positive Z-score indicated high support for an entity in hotel manager photos, whereas a negative Z-score indicated otherwise. In comparison with the photos taken by travelers, the images from hotel managers were more likely to show hotel facilities and indoor features, such as interior design, ceiling, lobby, restaurant, suite, and function hall for marketing purpose. In particular, the restaurant and function hall were around four times more likely to be shown in photos of hotel managers than in photos of travelers. Hotel managers are keen to advertise these hotel facilities, which could be of little interests to travelers. Outdoor scene photos, such as city, building, and metropolis appeared more frequently in the photos taken by travelers than in those taken by managers. Furniture and bathroom attract travelers' attention more than that of managers.

We carried out another analysis between managers' photos of hotels with different ratings for an overview of their advertising focuses. Managers' photos of two- and three-star hotels are treated as one group, and those of four- and five-star hotels are treated as another group. Contrast analysis was carried out and the entities with significant differences are shown in Table 4. Some hotel features, including function hall, leisure, resort swimming pool, and ballroom, were more emphasized at high-rating hotels than at low-rating hotels, with a ratio of more than five. Low-rating hotels focused more on advertising about room features, such as bedroom, floor, furniture, bed, table, and bed frame with a ratio of more than two. This was probably because high-rating hotels are more likely to have other facilities for conference or leisure activities than low-rating hotels. Hence, they included them in the marketing photos. Low-rating hotels focused more on the room and basic hotel features due to the lack of such facilities.

Table 4. Photo contents of four- and five-star hotels versus two- and three-star hotels.

Entity	Four and Five stars	Two and three stars	Difference	Growth rate	Z-score	P-value*
Function hall	0.169	0.028	0.142	6.144	6.691	0.000
Lobby	0.304	0.199	0.105	1.529	3.939	0.000
Leisure	0.102	0.012	0.09	8.374	5.309	0.000
Living room	0.166	0.092	0.074	1.811	3.475	0.001
Resort	0.086	0.015	0.07	5.593	4.469	0.000
Restaurant	0.183	0.122	0.06	1.493	2.700	0.007
Swimming pool	0.065	0.006	0.059	10.657	4.290	0.000
Condominium	0.141	0.083	0.058	1.703	2.898	0.004
Ballroom	0.062	0.006	0.056	10.183	4.167	0.000
Room	0.387	0.602	0.215	1.555	−7.285	0.000
Bedroom	0.109	0.266	0.157	2.448	−7.827	0.000
Suite	0.313	0.45	0.137	1.436	−4.828	0.000
Floor	0.119	0.254	0.135	2.127	−6.528	0.000
Furniture	0.049	0.141	0.091	2.855	−6.356	0.000
Interior design	0.669	0.752	0.083	1.124	−2.925	0.003
Bed	0.007	0.083	0.075	11.389	−9.625	0.000
Table	0.066	0.138	0.071	2.079	−4.507	0.000
Flooring	0.111	0.171	0.06	1.541	−3.090	0.002
Bed frame	0.029	0.086	0.057	2.954	−5.042	0.000

*Significant at p-value ≤ 0.05.

Face and color analysis

The use of facial image and color influences viewer emotions (Negri & Vigolo, 2015). This section explores the usage of such elements in the marketing strategy of hotel managers through online photos. We applied a face detection function of Google Vision API to determine the presence of human face(s) in each photo. The support values were then computed with respect to different photo groups, as shown in Table 5. No significant difference was identified for the popularity of photo with a human face between photos of hotel managers and travelers, as the p-value was greater than 0.05. Although the popularities of photos with face are not different, the setting and context of such photos may vary. We further browsed several actual photos with the human face of photo groups for a detailed understanding. Most photos posted by hotel managers were usually from the third person's point of view, and the context was mainly about friends or partners getting together (Figure 5a and 5b). Photos posted by travelers were about selfies and family get together (Figure 5c and 5d). We blurred out the faces of people in the photos to preserve privacy. Among the photos posted by hotel managers, low-rating hotels did not post any photo with human faces, whereas numerous photos in high-rating hotel contained human faces. The differences between these photos are significant, as indicated by a p-value of less than 0.05.

We then applied the color theme estimation function of Google Vision API to determine the color theme of the photos. The support values were then computed with respect to each photo groups for contrast analysis. Table 6 lists the significantly different color

Table 5. Face appearance in photos.

Face appearance		Difference	Growth Rate	Z-score	P-value*
Manager 0.015	Traveler 0.021	0.005	1.358	−1.903	0.057
Four and five stars 0.017	Two and three stars 0	0.017	0	2.436	0.014

*Significant at p-value ≤ 0.05.

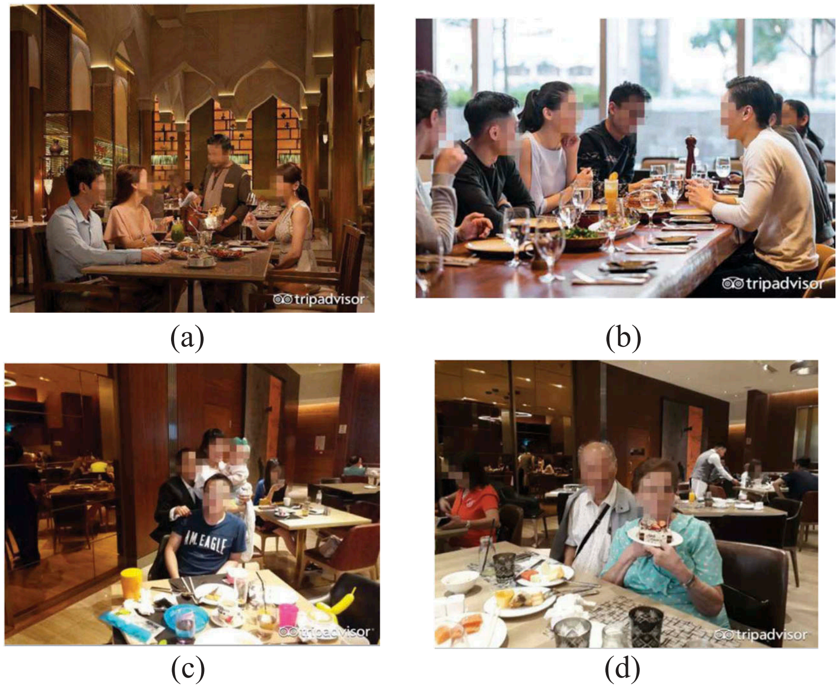


Figure 5. Sample photos with faces posted by managers (A, B) and travelers (C, D) .

Table 6. Color theme of hotel managers versus travelers.

Color Theme	Manager	Traveler	Difference	Growth Rate	Z-score	P-value*
Red	0.504	0.404	0.100	1.248	9.728	0.000
White	0.374	0.332	0.043	1.128	4.321	0.000
Blue	0.338	0.258	0.080	1.312	8.761	0.000
Black	0.405	0.493	0.089	1.219	−8.469	0.000

*Significant at p-value $\leq$ 0.05.

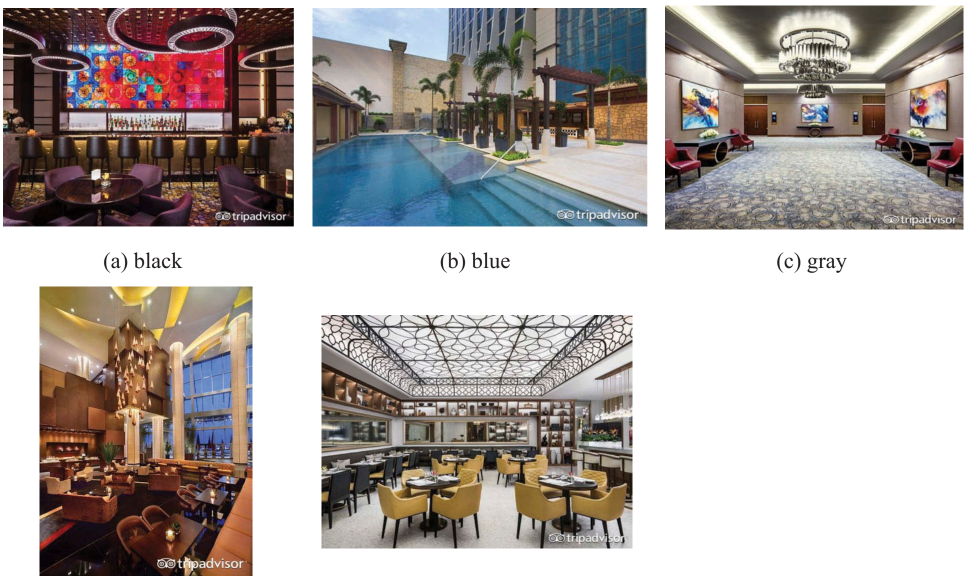
themes between the photos of hotel managers and travelers. More specifically, warm and calm color themes, such as red, white, and blue, were more popular among hotel managers' photos. By contrast, a black color theme was more popular among travelers' photos. A possible explanation for this pattern is that professional photographers took the photos uploaded by hotel managers. On the one hand, managers tend to be selective in choosing the scene and color setting to provide a positive and calm emotion to viewers. On the other hand, travelers are usually not professional photographers. Their photos are usually plain and often not adjusted with the light setting, which results in darker photos compared with those taken by hotel managers.

Table 7 lists significantly different colors between low- and high-rating hotels. The photos uploaded by hotel managers at high-rating hotels tend to be more colorful, with the red and blue theme. By contrast, photos of low-rating hotels tend to be less colorful with gray and white themes. Figure 6 shows sample photos of different color themes. Black and red themes often appeared in photos taken at a cafe or restaurant. The blue theme often contained

Table 7. Color themes of hotels with different ratings.

Color Theme	Four and Five Stars	Two and Three Stars	Difference	Growth Rate	Z-score	P-value*
Red	0.524	0.372	0.152	1.408	5.090	0.000
Blue	0.355	0.234	0.121	1.516	4.278	0.000
Gray	0.493	0.557	0.064	1.130	−2.151	0.032
White	0.351	0.520	0.169	1.481	−5.846	0.000

*Significant at p-value ≤ 0.05.



(a) black (b) blue (c) gray

(d) red (e) white

Figure 6. Sample photos with different color themes .

a swimming pool. Swimming pools are not a popular facility at low-rating hotels, which explains why the blue theme is more popular in high-than in low-rating hotels.

Discussion

Theoretical implications

This paper presented an approach to large-scale analysis of photo contents, which supports hotel managers in addressing important questions for effective marketing and thereby prompting their business. The approach is based on the theory of deep learning, with superior performance to traditional techniques in the literature (Calheiros et al., 2017; N. Deng & Li, 2018; Garcıa-Floriano et al., 2019; Göring et al., 2018). Notably, although deep learning is deemed advanced for photo content analysis, its primarily functions to recognize the contents of individual photos. The outputs of deep learning algorithm must be combined with other analysis techniques, such as contrast analysis (Li et al., 2015) to obtain a comprehensive view about the contents of an entire photo collection. Moreover, the entities contained in the photos are often scored independently by a set of pre-trained deep learning algorithms. The analysis of photo contents must consider the co-occurrence of entities by using graph theory (Otte & Rousseau, 2002) rather than independently analyzing the entity in prior research (Zhang et al., 20192019a), such that the understanding of general themes in the hotel photos can be obtained.

Given that most prior works often focused on photos produced by destination management organizations or businesses (Donaire et al., 2014; Govers et al., 2007; Pan et al., 2014), the current study extended the understanding about the travelers' perceptions and experiences through a large-scale content analysis of online hotel photos taken by travelers. Projected and perceived images of tourism destinations are important topics in hospitality and tourism research (Andreu et al., 2000; Marine-Roig & Ferrer-Rosell, 2018), as making an accurate impression is vital to business success in these competitive industries. Although photos have been known as a useful data source for such studies (Picazo & Moreno-Gil, 2019), their potentials in providing valuable insights have not been fully utilized probably due to barriers in traditional manual approaches to photo content analysis. Our case study demonstrated that the approach based on the theory of deep learning could facilitate the studies on the projected and perceived images from the photos with valuable implications.

Practical implications

The following several implications from a managerial perspective can be derived from the study. The exploration of visual content in Section 4.2 showed that the identified entities agreed with human interpretation. A comprehensive list of entities was constructed, which provides hotel managers with a comprehensive understanding of the common contents frequently appearing in online hotel photos. The visualization of entity relationships provides an overview of the structure of photo collections and helps identify common photo themes and contexts (Figure 4). Thus, in addition to basic indoor features, hotel managers should consider other aspects when advertising their hotels, such as outdoor/surrounding environment, leisure activities, and dining, as they play important roles of the hotel experiences of travelers.

Contrast analysis between the photos of hotel managers and travelers (Table 3) suggested that hotel managers might pay more attention to advertising indoor features, such as furniture and bathroom, because they receive much attention from travelers. More photos about the surrounding science of city and building can be included in marketing materials to attract travelers' attention. The differences in photo contents between four- and five-star hotels versus two- and three-star hotels are probably due to the difference in facility availability. Major facilities, such as function hall, swimming pool, and ballroom, are hard to improve. Hence, low-rating hotel can consider improving their existing facilities and market about them via photos of the lobby, living room, and restaurant to improve their attractiveness to potential customers.

The analysis in Section 4.4 suggested that low-rating hotels can consider including photos with human faces, such as staff greetings or get together scenes, to provide viewers with a warm welcome feeling, which may result in future purchases. High-rating hotels may include photos from a second person's point of view or family get together because such settings are commonly captured in travelers' photos with human faces. Color themes could also influence viewer emotion. Hence, managers of low-rating hotels can consider using photos with warm color themes to provide viewers with a positive feeling and impression about their hotels. The ability to determine color theme in photos is another advantage of deep learning because researchers with limited photography background could not do this task easily (Trpkovski et al., 2018).

Regarding the methodological aspect, deep learning theory depends on large-scale data set and a considerable amount of training time for performance. Therefore, constructing deep learning models from scratch for hotel photo analysis is not recommended due to the involved high cost. Employing pre-trained deep learning models are more feasible, such as those in Google Vision API, which has the state-of-the-art deep learning tool with proven performance in various applications (Felzenszwalb et al., 2014). Other deep learning tools, such as Clarifai (clarifai.com) and Amazon Rekognition (aws.amazon.com/rekognition), could also be utilized, depending on researchers' preferences.

Limitations

Despite the contribution in the content analysis of a large-scale hotel photos, our study is not without limitations. Hotels data were only collected for a single tourism destination. The identified entities only reflected the characteristics of hotels in Macau and the nature of this tourism destination, which might not be generalized well to other destinations. The case study was a demonstration of the proposed framework based on deep learning theory. Hence, the analysis mainly focused on the evaluation of the performance and exploration of the photo collection and did not take the profile of travelers into account. The profiles of Macau travelers may differ from those in other destinations. The identified content preferences in the photos may not be generalized well to travelers in other destinations. Nevertheless, the presented approach could process and analyze photo content automatically, which would take minimal effort for the researchers to extend their studies to multiple destinations for comprehensive understanding. Although this study focused on analyzing photo contents, such analysis could also be combined with user reviews (Liu et al., 2013; Ma et al., 2018; Xia et al., 2019) for other insights, such as identifying and improving hotel features that are receiving negative comments from travelers.

Conclusions

With the prevalence of social media and Web 2.0, an immense growth in the diversity of online content has occurred. Due to its visual representation, many travelers posted considerable travel photos online, which potentially capture various aspects of their experiences about tourism products and services (N. Deng & Li, 2018; Donaire et al., 2014). Such data sources greatly support researchers and hotel managers in gaining insights into travelers' perception and experiences to improve decision-making and marketing. However, such potential has not been effectively utilized due to the barrier in manual analysis approach in traditional studies (Fairweather & Swaffield, 2002; Garrod, 2007; Luo et al., 2017). In addition, content analysis based on manual visual inspection is a subjective process, which does not effectively recognize all possible semantic meanings in photos.

This study distinguishes itself in adopting deep learning theory to examine the travelers' preconceptions as reflected in their posted photos. It also contributes to the literature of visual assets in the tourism and hospitality industry. Based on deep learning theory and state-of-the-art computer vision framework, an approach to photo content analysis is introduced for practical applications in hotel marketing and management. We demonstrated the effectiveness of this approach through a case study of hotel photo content analysis in Macau. With nearly more than

53,000 hotel photos contributed by both hotel managers and travelers, this work is the first to analyze hotel photos at a scale larger than those used in previous tourism and hospitality studies. The content of those hotel photos was also analyzed comprehensively to recognize all possible captured contents. Deep learning could recognize contents in hotel photos in agreement with and, in some cases, beyond the capability of human viewers. Practical implications for hotel managers, especially those in Macau, have been provided.

It is worth mentioning that the aim of this paper is not to investigate the gaps between projected and perceived images about hotels in Macau, but to introduce an approach to facilitating such visual assets analysis using photos as data sources, which is an important but still understudied topic in hospitality and tourism research (Marine-Roig & Ferrer-Rosell, 2018; Picazo & Moreno-Gil, 2019).

This work can be further extended by incorporating traveler profiles into the content analysis of hotel photos for insights in their perceptions and preference differences. Contrast content analysis of hotel managers' photos could be carried out with respect to hotel of different branches, types, or scales for more detailed insights into projected images created by hotel management (Picazo & Moreno-Gil, 2019). Furthermore, the presented approach can be applied for analyzing travel photos in various domains for wider applications in hospitality and tourism marketing. With the approach presented in this work, the analysis of photo contents can be entirely automated, accordingly, future studies can also analyze large scales photos from other hotels, tourism destinations, and travelers for a comprehensive understanding.

Acknowledgments

We also would like to thank Xinjiang research fund with the Chinese Academy of Sciences to host Li's study leave in 2019.

Funding

The work described in this paper was partly supported by a research grant funded by the Hong Kong Polytechnic University and Shaanxi Province Key R&D Fund (2018KW-010). The work was completed when Gang Li is on ASL in the Chinese Academy of Sciences.

ORCID

Huy Quan Vu http://orcid.org/0000-0003-1947-2879
Gang Li http://orcid.org/0000-0003-1583-641X
Rob Law http://orcid.org/0000-0001-7199-3757

References

Al-antari, M. A., Al-masni, M., Park, S.-U., Park, J., Metwally, M., Kadah, Y., Han, S.-M., & Kim, T.-S. (2017). An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network. *Journal of Medical and Biological Engineering*, *38*(3), 443–456. https://doi.org/10.1007/s40846-017-0321-6

Andreu, L., Bigne, J. E., & Cooper, C. (2000). Projected and perceived image of spain as a tourist destination for british travellers. *Journal of Travel & Tourism Marketing*, *9*(4), 47–67. https://doi.org/10.1300/J073v09n04_03

Bai, B., Law, R., & Wen, I. (2008). The impact of website quality on customer satisfaction and purchase intentions: Evidence from Chinese online visitors. *International Journal of Hospitality Management*, *27*(3), 391–402. https://doi.org/10.1016/j.ijhm.2007.10.008

Bay, H., Ess, A., Tuytelaars, T., & Gool, L. V. (2008). Speeded-up robust features (SURF). *Journal Computer Vision and Image Understanding*, *100*(3), 346–359. https://doi.org/10.1016/j.cviu.2007.09.014

Bender Stringam, B., & Gerdes, J. J. (2010). Are pictures worth a thousand room nights? Success factors for hotel web site design. *Journal of Hospitality and Tourism Technology*, *1*(1), 30–49. https://doi.org/10.1108/17579881011023007

Calheiros, A. C., Moro, S., & Rita, P. (2017). Sentiment classification of consumer-generated online reviews using topic modeling. *Journal of Hospitality Marketing and Management*, *26*(7), 675–693. https://doi.org/10.1080/19368623.2017.1310075

Cavallo, D. P., Cefola, M., Pace, B., Logrieco, A. F., & Attolico, G. (2019). Non-destructive and contactless quality evaluation of table grapes by a computer vision system. *Computers and Electronics in Agriculture*, *156*, January, 558–564. https://doi.org/10.1016/j.compag.2018.12.019

Chen, Z. (2019). A qualitative pilot study exploring tourists' pre- and post-trip perceptions on the destination image of Macau. *Journal of Travel & Tourism Marketing*, *36*(3), 330–344. https://doi.org/10.1080/10548408.2018.1541777

Chou, M. C. (2013). Does tourism development promote economic growth in transition countries? A panel data analysis. *Economic Modelling*, *33*(2), 226–232. https://doi.org/10.1016/j.econmod.2013.04.024

Chuang, S.-C. (2007). The effects of emotions on the purchase of tour commodities. *Journal of Travel & Tourism Marketing*, *22*(1), 1–13. https://doi.org/10.1300/J073v22n01_01

Ciregan, D., Meier, U., & Schmidhuber, J. (2012). Multi-column deep neural networks for image classification. In *IEEE conference on computer vision and pattern recognition* (pp. 3642–3649). Providence, RI, USA.

Deng, L., & Yu, D. (2014). Deep learning - methods and applications. *Foundations and Trends in Signal Processing*, *7*(3–4), 197–387. https://doi.org/10.1561/2000000039

Deng, N., & Li, X. R. (2018). Feeling a destination through the 'right' photos: A machine learning model for DMOs photo selection. *Tourism Management*, *65*, April, 267–278. https://doi.org/10.1016/j.tourman.2017.09.010

Dholey, M., Sarkar, A., Maity, M., Giri, A., Sadhu, A., Chaudhury, K., Das, S., & Chatterjee, J. (2017). A computer vision approach for lung cancer classification using fnac-based cytological images. In *2nd international conference on computer vision & image processing* (pp. 181–195), Roorkee, India.

Dobni, D., & Zinkhan, G. M. (1990). In search of brand image: A foundation analysis. *Advances in Consumer Research*, *17*, 110–119.

Donaire, J. A., Camprub, R., & Gal, N. (2014). Tourist clusters from Flickr travel photography. *Tourism Management Perspectives*, *11*, July, 26–33. https://doi.org/10.1016/j.tmp.2014.02.003

Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management*, *55*, August, 62–73. https://doi.org/10.1016/j.tourman.2016.01.013

Fairweather, J. R., & Swaffield, S. R. (2002). Visitors' and locals' experiences of Rotorua, New Zealand: An interpretative study using photographs of landscapes and Q method. *International Journal of Tourism Research*, *4*(4), 283–297. https://doi.org/10.1002/jtr.381

Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2014). Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(9), 1627–1645. https://doi.org/10.1109/TPAMI.2009.167

Gallarza, M. G., Saura, I. G., & GarciA, H. C. (2002). Destination image: Towards a conceptual framework. *Annals of Tourism Research*, *29*(1), 56–78. https://doi.org/10.1016/S0160-7383(01)00031-7

Garcıa-Floriano, A., Ferreira-Santiago, A., Camacho Nieto, O., & Yanez-Marquez, C. (2019). A machine learning approach to medical image classification: Detecting age-related macular degeneration in fundus images. *Computers & Electrical Engineering*, *75*, May, 218–229. https://doi.org/10.1016/j.compeleceng.2017.11.008

Garrod, B. (2007). A snapshot into the past: The utility of volunteer-employed photography in planning and managing heritage tourism. *Journal of Heritage Tourism*, *2*(1), 14–35. https://doi.org/10.2167/jht018.0

Garrod, B. (2009). Understanding the relationship between tourism destination imagery and tourist photography. *Journal of Travel Research*, *47*(3), 346–358. https://doi.org/10.1177/0047287508322785

Göring, S., Brand, K., & Raake, A. (2018). Extended features using machine learning techniques for photo liking prediction. In *Tenth international conference on quality of multimedia experience* (pp. 1–6), Cagliari, Italy.

Govers, R., Go, F. M., & Kumar, K. (2007). Promoting tourism destination image. *Journal of Travel Research*, *46*(1), 15–23. https://doi.org/10.1177/0047287507302374

Hadjkacem, B., Ayedi, W., Abid, M., & Snoussi, H. (2018). Multi-shot human re-identification for the security in video surveillance systems. In *27th IEEE international conference on enabling technologies: Infrastructure for collaborative enterprises* (pp. 203–208), Paris, France.

Hao, X., Wu, B., Morrison, A. M., & Wang, F. (2016). Worth thousands of words? Visual content analysis and photo interpretation of an outdoor tourism spectacular performance in Yangshuo-Guilin, China. *Anatolia*, *27*(2), 201–213. https://doi.org/10.1080/13032917.2015.1082921

Hu, J., Sun, Z., Sun, Y., & Shi, J. (2018). Accumulative image categorization: A personal photo classification method for progressive collection. *Multimedia Tools Application*, *77*(24), 32179–32211. https://doi.org/10.1007/s11042-018-6152-9

Jeong, M., & Choi, J. (2005). Effects of picture presentations on customers' behavioral intentions on the web. *Journal of Travel & Tourism Marketing*, *17*(2–3), 193–204.https://doi.org/10.1300/J073v17n02_15

Ji, Z., Tang, J., & Wu, G. (2019). Personalized recommendation of photography based on deep learning. In *25th international conference on multimedia modeling* (pp. 214–226), January 8-11, Thessaloniki, Greece.

Jing, Y., Youssefi, B., Mirhassani, M., & Muscedere, R. (2017). An efficient FPGA implementation of optical character recognition for license plate recognition. In *30th IEEE canadian conference on electrical and computer engineering, CCECE 2017* (pp. 1–4), Windsor, ON, Canada.

Kao, Y., Wang, C., & Huang, K. (2015). Visual aesthetic quality assessment with a regression model. In *IEEE international conference on image processing (ICIP)* (pp. 1583–1587), Quebec City, QC, Canada.

Kim, M. J., Lee, C. K., & Bonn, M. (2017). Obtaining a better understanding about travel-related purchase intentions among senior users of mobile social network sites. *International Journal of Information Management*, *37*(5), 484–496. https://doi.org/10.1016/j.ijinfomgt.2017.04.006

Kwortnik, R. J., Jr, & Ross, W. T., Jr. (2007). The role of positive emotions in experiential decisions. *International Journal of Research in Marketing*, *24*(4), 324–335. https://doi.org/10.1016/j.ijresmar.2007.09.002

Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research*, *75*, March, 410–423. https://doi.org/10.1016/j.annals.2019.01.014

Law, R., Rong, J., Vu, H. Q., Li, G., & Lee, H. A. (2011). Identifying changes and trends in Hong Kong outbound tourism. *Tourism Management*, *32(5)*, 1106–1114. https://doi.org/10.1016/j.tourman.2010.09.011

Lee, H., Guillet, B. D., & Law, R. (2013). An examination of the relationship between online travel agents and hotels: A case study of choice hotels international and expedia.com. *Cornell Hospitality Quarterly*, *54*(1), 95–107. https://doi.org/10.1177/1938965512454218

Lee, J., & Whaley, J. E. (2019). Determinants of dining satisfaction. *Journal of Hospitality Marketing & Management*, *28*(3), 351–378. https://doi.org/10.1080/19368623.2019.1523031

Leong, A. M. W., & Li, J. X. (2010). Interactive strategies for cultural heritage tourism and Macau city development: The tourists' perspectives. *International Journal of Asian Business and Information Management*, *1*(4), 1–19. https://doi.org/10.4018/jabim.2010100101

Leung, X. Y., & Bai, B. (2013). How motivation, opportunity, and ability impact travelers' social media involvement and revisit intention. *Journal of Travel & Tourism Marketing*, *30*(1–2), 58–77. https://doi.org/10.1080/10548408.2013.751211

Li, G., Law, R., Vu, H. Q., & Rong, J. (2013). Discovering the hotel selection preferences of Hong Kong inbound travelers using the Choquet integral. *Tourism Management*, *36*(3), 321–330. https://doi.org/10.1016/j.tourman.2012.10.017

Li, G., Law, R., Vu, H. Q., Rong, J. R., & Zhao, X. (2015). Identifying emerging hotel preferences using emerging pattern mining technique. *Tourism Management*, *46*, February, 311–321. https://doi.org/10.1016/j.tourman.2014.06.015

Liu, S., Law, R., Rong, J., Li, G., & Hall, J. (2013). Analyzing changes in hotel customers' expectations by trip mode. *International Journal of Hospitality Management*, *34*, September, 359–371. https://doi.org/10.1016/j.ijhm.2012.11.011

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoint. *International Journal of Computer Vision*, *60*(2), 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Lu, W., & Stepchenkova, S. (2015). User-generated content as a research mode in tourism and hospitality applications: Topics, methods, and software. *Journal of Hospitality Marketing & Management*, *24*(2), 119–154. https://doi.org/10.1080/19368623.2014.907758

Luo, K., Vu, H. Q., Zhao, X., Li, G., & Law., R. (2017). Analyzing museum visitor experiences based on user-generated travel photos. *Journal of Travel & Tourism Marketing*, *35*(4), 493–506. https://doi.org/10.1080/10548408.2017.1363684

Ma, J., & Takagi, H. (2012). Design of composite image filters using interactive genetic programming. In *Third International Conference on Innovations in Bio-inspired Computing & Applications* (pp. 26–28), September 26-28, Kaohsiung City, Taiwan.

Ma, Y., Xiang, Z., Du, Q., & Fan, W. (2018). Effects of user-provided photos on hotel review helpfulness: An analytical approach with deep leaning. *International Journal of Hospitality Management*, *71*, April, 120–131. https://doi.org/10.1016/j.ijhm.2017.12.008

Marine-Roig, E., & Ferrer-Rosell, B. (2018). Measuring the gap between projected and perceived destination images of Catalonia using compositional analysis. *Tourism Management*, *68*, October, 236–249. https://doi.org/10.1016/j.tourman.2018.03.020

Markwick, M. (2001). Postcards from Malta: Image, consumption, context. *Annals of Tourism Research*, *28*(2), 417–438. https://doi.org/10.1016/S0160-7383(00)00049-9

Napiorkowska, M., Petit, D., & Marti, P. (2018). Three applications of deep learning algorithms for object detection in satellite imagery. In *IEEE international geoscience and remote sensing symposium* (pp. 4839–4842), Valencia, Spain.

Negri, F., & Vigolo, V. (2015). Hotel attributes and visual image: A comparison between website and user-generated photos. In *International conference on information and communication technologies in tourism* (pp. 621–633), Lugano, Switzerland.

Nixon, M. S., Liu, X. U., Direkoglu, C., & Hurley, D. J. (2011). On using physical analogies for feature and shape extraction in computer vision. *The Computer Journal*, *54*(1), 11–25. https://doi.org/10.1093/comjnl/bxp070

Ojala, T., Pietikäinen, M., & Topi Mäenpää, T. (2002). Gray scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *24*(7), 971–987. https://doi.org/10.1109/TPAMI.2002.1017623

Ojansivu, V., & Heikkilä, J. (2008). Blur insensitive texture classification using local phase quantization. In *The 3rd International Conference on Image and Signal Processing* (pp. 236–243), July 1-3, Cherbourg-Octeville, France.

Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science*, *28*(6), 441–453. https://doi.org/10.1177/016555150202800601

Pan, S., Lee, J., & Tsai, H. (2014). Travel photos: Motivations, image dimensions, and affective qualities of places. *Tourism Management*, *40*, February, 59–69. https://doi.org/10.1016/j.tourman.2013.05.007

Picazo, P., & Moreno-Gil, S. (2019). Analysis of the projected image of tourism destinations on photographs: A literature review to prepare for the future. *Journal of Vacation Marketing*, *25*(1), 3–24. https://doi.org/10.1177/1356766717736350

Pritchard, A. (2001). Tourism and representation: A scale for measuring gendered portrayals. *Leisure Studies*, *20*(2), 79–94. https://doi.org/10.1080/02614360110068651

Russell, J. (2017). *Google's AlphaGo AI wins three-match series against the world's best go player*. *Verizon Media*, accesse on June 11, 2019, https://techcrunch.com/2017/05/24/alphago-beats-planets-best-human-go-player-ke-jie/

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, January, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

song, S.-G., & Kim, D.-Y. (2016). A pictorial analysis of destination images on Pinterest: The case of Tokyo, Kyoto, and Osaka, Japan. *Journal of Travel & Tourism Marketing*, *33*(5), 687–701. https://doi.org/10.1080/10548408.2016.1167384

Sonka, M., Hlavac, V., & &Boyle, R. (2014). *Image processing, analysis, and machine vision*. Springer.

Stepchenkova, S., & Zhan, F. (2013). Visual destination images of Peru: Comparative content analysis of DMO and user-generated photography. *Tourism Management*, *36*, June, 590–601. https://doi.org/10.1016/j.tourman.2012.08.006

Stern, B., Zinkhan, G. M., & Jaju, A. (2001). Marketing images: Construct definition, measurement issues, and theory development. *Marketing Theory*, *1*(2), 201–224. https://doi.org/10.1177/147059310100100203

Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to data mining* (2nd ed.). Pearson.

Tan, X., & Triggs, B. (2007). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, *19*(6), 1635–1650. https://doi.org/10.1109/TIP.2010.2042645

Tapachai, N., & Waryszak, R. (2000). An examination of the role of beneficial image in tourist destination selection. *Journal of Travel Research*, *39*(1), 37–44. https://doi.org/10.1177/004728750003900105

Tian, D., Zhao, X., & Shi, Z. (2012). Support vector machine with mixture of kernels for image classification. In *12th international conference on intelligent information processing*, pp. 68–76, Guilin, China.

Trabelsi, R., Jabri, I., Melgani, F., Smach, F., Conci, N., & Bouallegue, A. (2019). Indoor object recognition in RGBD images with complex-valued neural networks for visually-impaired people. *Neurocomputing*, *330*, February, 94–103. https://doi.org/10.1016/j.neucom.2018.11.032

Tran, M. T., Jeeva, A. S., & Pourabedin, Z. (2016). Social network analysis in tourism services distribution channels. *Tourism Management Perspectives*, *18*, April, 59–67. https://doi.org/10.1016/j.tmp.2016.01.003

Trpkovski, A., Vu, H. Q., Li, G., Wang, H., & Law, R. (2018). Automatic hotel photo quality assessment based on visual features. In: Stangl B., Pesonen J. (eds) *Information and communication technologies in tourism 2018* (pp. 394–406). Jonkoping, Sweden.

Tzimiropoulos, G., & Pantic, M. (2017). Fast algorithms for fitting active appearance models to unconstrained images. *International Journal of Computer Vision*, *122*(1), 17–33. https://doi.org/10.1007/s11263-016-0950-1

Xia, H., Vu, H. Q., Lan, Q., Law, R., & Li, G. (2019). Identifying hotel competitiveness based on hotel feature ratings. *Journal of Hospitality Marketing & Management*, *28*(1), 81–100. https://doi.org/10.1080/19368623.2018.1504366

Xiang, Z., Schwartz, Z., Gerdes, J. H., Jr, & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, *44*(44), 120–130. https://doi.org/10.1016/j.ijhm.2014.10.013

Yang, X., Zhou, Y., Zhu, Q., & Wu, Z. (2018). Joint graph regularized extreme learning machine for multi-label image classification. *Journal of Computational Methods in Sciences and Engineering*, *18*(1), 213–219. https://doi.org/10.3233/JCM-180783

Yao, K., Yu, D., Seide, F., Su, H., Deng, L., & Gong, Y. (2012). Adaptation of context-dependent deep neural networks for automatic speech recognition. In *IEEE spoken language technology workshop*, pp. 366–369, Miami, FL, USA.

Yip, L. L. H. (2017). *Developing the employability competency for the tourism and hotel industry in Hong Kong*. [PhD thesis], University of Nottingham.

Zhang, K., Chen, Y., & Li, C. (2020). How are tourists different? - Reading geo-tagged photos through a deep learning model. *Journal of Quality Assurance in Hospitality & Tourism*. 21(2), 234-243. https://doi.org/10.1080/1528008X.2019.1653243.

Zhang, K., Chen, Y., & Li, C. (2019). Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing. *Tourism Management*, *75*, December, 595–608. https://doi.org/10.1016/j.tourman.2019.07.002

Zhang, Y., Li, G., Muskat, B., Law, R., & Yang, Y. (2020). Group pooling for deep tourism demand forecasting. *Annals of Tourism Research*, *82*, May, 102899. https://doi.org/10.1016/j.annals.2020.102899