

This is the accepted manuscript of the article:

Liu, X., Nicolau, J. L., Law, R., & Li, C. (2022). Applying image recognition techniques to visual information mining in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, DOI: <https://doi-org.ezproxy.lib.vt.edu/10.1108/IJCHM-03-2022-0362>

Applying image recognition techniques to mine visual information in hospitality and tourism

Purpose

This study aims to provide a critical reflection of the application of image recognition techniques to mine visual information in hospitality and tourism.

Design/methodology/approach

The study starts by reviewing the progress of image recognition and the advantages of convolutional neural network-based image recognition. Thereafter, it continues to explain and exemplify the mechanisms and functions of two relevant applications of image recognition: object recognition and facial recognition. The study ends by providing the theoretical and practical implications and potential directions for future research.

Findings

After presenting different potential applications and comparing the use of image recognition with traditional manual methods, the main findings of this critical reflection revolve around the feasibility of the techniques described.

Originality

Visual information plays a critical role in online travel agencies and peer-to-peer accommodation platforms from the side of sellers and buyers. However, extant studies have relied heavily on traditional manual identification with small samples and subjective judgment. With the development of deep learning and computer vision techniques, current researchers are able to extract various types of visual information from large-scale datasets with high accuracy and efficiency. This study is the first to offer an outlook of image recognition techniques to mine visual information in hospitality and tourism.

Practical implications

Knowledge of how to extract valuable visual information from large-scale user-generated photos to infer the online behavior of consumers and service providers and its influence on purchase decision and firm performance is crucial for business practice in hospitality and tourism.

Keywords: Visual information; Deep learning; Image recognition; Hospitality management

1. Introduction

In the hospitality and tourism industry, consumers face uncertainties and risks during online bookings (e.g., hotels and peer-to-peer accommodations) owing to the feature of experience goods, the quality of which is difficult to evaluate before check-in. Thus, adequate information is required for consumers to make informed booking decisions. With the rapid development of information technology, the Internet, and smart devices, sharing people's experiences on popular review sites in hospitality (e.g., TripAdvisor, Booking, and Yelp) has become a common practice and form of entertainment among consumers (buyers). Among all types of information sources (e.g., ratings, texts, and photos), visual information plays an increasingly important role in the big data era because it is intuitive, easy to digest, and requires minimal cognitive effort. As the carrier of visual information, uploading photos in online sharing environments is widely adopted by hotel booking platforms and consumers (Lo *et al.*, 2011). Popular online booking platforms, such as TripAdvisor and Airbnb, allow and encourage sellers (hosts) and buyers (guests/consumers) of accommodation products/services to upload photos to convey visual information.

For sellers of accommodation products/services, visual information embedded in photos provides consumers with the vivid characteristics and features of their properties, such as the location and facilities of hotels, the dining environment and restaurant dishes. Compared with textual descriptions, visual information delivered from photos is perceived as more persuasive and trustworthy by consumers when pre-evaluating the quality of products on webpages (Zhang *et al.*, 2017). Accordingly, visual information plays an important role in promoting the popularity of listings on peer-to-peer accommodation platforms (Yan *et al.*, 2021). In particular, the visual information of sellers appears in their profile images (e.g., age, gender, and attractiveness) and becomes the initial information source of trustworthiness cues that potential consumers see when searching peer-to-peer accommodation products (Barnes, 2021). From the perspective of buyers (consumers), visual information delivered from consumer-generated photos can reveal the tangible and intangible attributes of products or services, which are perceived as more objective and credible by peer consumers compared with photos uploaded by sellers (Negri and Vigolo, 2015). Moreover, information derived from the profile images of consumers indicates their personal characteristics, which has the potential to predict their online behavior (Li *et al.*, 2021).

However, the majority of studies in hospitality and tourism involving image recognition were manually conducted, which limits the sample size, identification accuracy, and variety. To overcome the limitations of slow speed, time-consuming, and recognition bias in the manual coding of consumer-generated photos, machine learning models were adopted to extract visual information. For example, support vector machine (SVM) can be trained with general photo features to classify the hotel photos and label each photo (Trpkovski *et al.*, 2018). Nevertheless, such a method requires professional knowledge of designing specific image features (e.g., brightness, colorfulness, etc.) and can only recognize specific visual information. Thus, the

generalization and robustness of such models are extremely limited.

With the rapid development of deep learning techniques, artificial neural network (ANN) analyses emerged and were adopted in hospitality and tourism studies (Battiti *et al.*, 2021; Essien and Chukwukelu, 2022; Doborjeh *et al.*, 2022). However, traditional ANN analyses are not practical for processing images because they generate numerous parameters, thereby leading to low speed and over fitting. This task must be conducted efficiently. As stated below, the use of pictorial representations helps firms show their product/service attributes and customers know “what” they will find once in the destination. Accordingly, from a managerial viewpoint, the knowledge provided by the analysis of photos is critical for decision making, and from an academic perspective, the identification of the behavioral traits of people regarding the use of pictures when traveling offers crucial research avenues.

The emergence of convolutional neural network (CNN) solved this problem. Different from traditional ANN, CNN can accurately extract high-dimensional features from images and significantly reduce the parameters. The boost of central processing units (CPUs) and graphic processing units (GPUs) has considerably promoted the application and development of CNN. Inspired by the structure of CNN, a range of state-of-the-art image recognition models have been developed. Many of these models, such as AlexNet, VGGNet, GoogLeNet, or ResNet, have achieved excellent performance in the task of image recognition. In particular, ResNet (invented by Microsoft) reached outstanding performance, with its accuracy markedly higher than that of the human eye.

2. Information asymmetry and visual information

Information asymmetry is a prevalent characteristic of markets. Conceptualized as the contrast between information that consumers and firms possess on products and services offered to the market, information asymmetry has a critical impact on quality. That is, firms, as providers of products or services—and their quality—are well aware of the level of quality of their products or services. However, customers lack this level of knowledge. If customers are unable to distinguish between high- and low-quality products and services, then beyond their potential dissatisfaction, the so-called adverse selection phenomenon may occur when low-quality products/services displace those with high quality. For example, if hotels are unable to show that their quality is greater than that of their competitors, then the lower prices of the latter will drive the former out of the market. In this context, high-quality hotels will need to implement certain actions to demonstrate that their products and services possess high quality, thereby reducing the uncertainty that customers derive from the alluded information asymmetry. Beyond traditional devices to diminish this uncertainty—either internal, such as warranties (through which the firm assures customers that everything will go according to plan or else they will get their money back), or external, such as quality awards (through which a third party certifies the “promised” quality)—the hospitality and tourism industry has developed its own mechanisms to fight information asymmetry and help customers avoid uncertainty.

The potential effect of information asymmetry is not a minor issue, given that the hospitality and tourism industry is highly information-sensitive and customers cannot assess the quality of services until they have arrived and are physically in the destination. Moreover, not only are individuals unable to evaluate quality prior to arriving at destinations, but they need to get and stay there with the higher involvement that “getting and staying there” entails compared with other industries. For other industries in which products are physically transported to where customers live, customers do not need to move far away from their usual places of residence. In the current consumption scenario, in addition to typical, well-established uncertainties that are characteristic of the hospitality and tourism industry (e.g., whether customers will need to endure flight delays or whether food in a hotel does not have the expected quality), new elements that bring about new uncertainties emerge. Paradigmatic among these new uncertainties is the new model that has given rise to peer-to-peer accommodation. Although hotel processes are generally known, regardless of the quality of a property, customers know that firms must comply with some basic standards, which should be found according to expectations. For peer-to-peer accommodation, knowing whether or not the host of a unit will be friendly is an uncertainty-generating component. In this context, customers have a special need to reduce the uncertainty derived from the consumption of hospitality and tourism services.

The use of photos in brochures and/or websites has been a strategy widely applied in the hospitality and tourism industry. Firms in this industry utilize images to show that “they really have what they advertise they have,” specifically, to make those elements—which customers cannot see until they are in person before those elements—become tangible. With the advent of online reviews, the use of photos is no longer restricted to firms alone; customers can also post photos of their own experiences. These experiences that reflect what firms offer—usually, even in real time—can help other prospective customers reduce uncertainty and make informed booking decisions. By using the two dimensions of “user” (who generates the photos) and “content” (content of photos), Table 1 shows the four possible combinations: (i) information on services generated by buyers, (ii) information on services generated by sellers, (iii) information on individuals generated by buyers, and (iv) information on individuals generated by sellers. Buyer-generated service information refers to photos posted by customers during or after their experiences, and buyer-generated individual information shows photos of the buyers themselves. Seller-generated service information refers to photos published by firms that show the products and services they offer, and seller-generated individual information presents photos of the sellers themselves.

Table 1. Visual information categorized by user and content

Different user-generated photos reveal distinct visual information. From the perspective of content, service-related photos usually contain visual information about specific scenes, like indoor, bedroom, and bathroom. Each particular object can also

be observed in greater detail, for example, the normal facilities in a hotel like TV and bed and the dinnerware in the context of restaurants. For individual-related photos, the personal image of sellers or buyers can show their appearance. It is feasible to detect demographic characteristics like gender, age, and race as well as the facial expression or emotion (e.g., happy, angry, and sad) from their profile images. Figure 1 presents the conceptual framework of this study. In particular, the visual information embedded in user-generated photos can contribute to the marketing of sellers, the decision making of buyers, and the operation of platforms in hospitality and tourism in various online contexts (e.g., online travel agencies like Expedia, peer-to-peer accommodation platforms like Airbnb, and review sites like Yelp).

Figure 1. Conceptual framework

3. Progress of image recognition

Given the increasing importance of visual information in academia and business practice in hospitality, extensive studies have been performed on this research direction (Lyu *et al.*, 2022). To make a comprehensive review of related studies, we searched keywords like “visual information”, “image recognition”, and “facial recognition” in Google Scholar, and then picked out those keywords mainly published in hospitality and tourism journals. From the perspective of research methods, extant studies have been mainly conducted by interview, questionnaire, or experiment. Owing to the burden of manual recognition and identification of photos, samples used in the extant studies often have fewer than 1,000 observations.

Filieri (2016) used the data of 38 interviews as basis and found that visual information is often utilized by consumers to evaluate review trustworthiness. In particular, review photos (message trustworthiness) and profile photos (source trustworthiness) can affect perceived review trustworthiness. Ert *et al.* (2016) and Fleischer (2017) used facial characteristics obtained by the manual recognition of consumers’ profile photos (e.g., age, gender, and smile expression) to infer the trustworthiness and attractiveness of the hosts. Kuo *et al.* (2015) conducted an experiment with 240 consumers and found that brand trust is low for upscale hotels when consumers encounter misleading hotel website photos. Bufquin *et al.* (2020) found that photos on hotel websites can enhance the enjoyment and booking intention of consumers. Bigne *et al.* (2020) leveraged eye-tracking technology and questionnaire and observed that consumers focus on visual information at the expense of concentrating on textual information. Findings from a conjoint study showed that a seller’s negative facial expression or absence of facial image on Airbnb decreases the likelihood of renting (Fagerstrøm *et al.*, 2017).

It is noticeable that some online crowdsourcing platforms like Amazon Mechanical Turk were also employed by scholars to recruit participants to manually evaluate user-generated photos (Zhang *et al.*, 2017). For example, by analyzing the profile photos of hosts of 1,020 properties in New York City, Jaeger *et al.* (2019) showed that hosts tend to charge higher prices if their profile photos are perceived as

more attractive-looking rather than trustworthy-looking; hosts with intense smiles also charge higher. However, this approach is not applicable with large-scale data.

Owing to the rapid development of the Internet, mobile devices, artificial intelligence, and deep learning, CNN-based models possess considerable advantages in processing large-scale images efficiently and accurately, particularly for recognizing the numerous online photos generated on accommodation platforms. By virtue of CNN-based models, research has shown that visual information embedded in seller- and buyer-generated photos plays an increasingly important role in marketing (hosts and hotel managers), decision making (consumers), and the operation of accommodation platforms (peer-to-peer accommodation and hotel booking platforms).

As one of the first studies to apply CNN-based image recognition models in hospitality and tourism, Zhang *et al.* (2017) evaluated the aesthetic quality for more than 510,000 photos of Airbnb-listed properties and revealed that the improvement of image quality can bring additional revenue to hosts. Barnes (2021) and Barnes and Kirshner (2021) also showed that the trust cues of hosts' facial images can induce overvaluation and lead to a 5% increase of their listing's prices. Using Google's Cloud Vision API, which integrates multiple CNN-based models, Lin *et al.* (2021) showed that photos from social media and survey share major similarities in key image characteristics. Moreover, Gambetti and Han (2022) proposed a neural food aesthetic assessment model and found that the aesthetic scores of photos differ across restaurants with different rating levels and cuisine types.

4. Potential implications of CNN-based image recognition in hospitality and tourism

Knowing how to extract valuable visual information from large-scale user-generated photos to infer consumer online behavior and its influence on firm performance is a promising research direction and is crucial for business practices in hospitality. This study provides two illustrations (i.e., *object detection* and *facial recognition*) to show how to extract visual information using CNN-based models and discusses the potential research ideas and directions thereafter for future research. Three widely used public standard datasets (i.e., MS COCO, LFWA, and FER2013) are introduced as the sample data in the following subsections to illustrate *object detection* and *facial recognition*. MS COCO contains 328,000 images with 2.5 million labeled instances and 80 objects categories. This dataset received extensive application in category detection and instance segmentation. LFWA has 13,233 images of 5,749 identities. Each image in LFWA is annotated with forty face attributes and five key points. FER2013 contains more than 30,000 facial images with seven expressions (i.e., Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral). These two datasets are widely used in facial recognition.

4.1 Object recognition in hospitality and tourism

Object recognition can be applied to extract visual information embedded in consumer-generated photos uploaded in their comments on products/services and hotel-generated photos displaying their products or services. CNN-based models are available for identifying specific objects or scenes of photos. Figure 2 illustrates the general image recognition process of object detection with YOLO v3, a state-of-the-art object detection model that is widely applied to detect specific objects (e.g., bed, TV, cup, etc.) of photos. It works as follows. First, standard image datasets on objects (e.g., MS COCO, etc.) are separated into training and testing sets. Second, the training set is used to train the CNN-based deep learning model and the testing set is utilized to verify the performance of the trained model. Finally, after good performance is achieved, the trained model can be used to conduct object detection on the research samples. Trained with Microsoft COCO dataset, YOLO v3 can process images at 30 FPS and has a mean average precision (mAP) of 57.9%; for a single object like a person, the accuracy can be 96.42%. Specific objects detected from consumer- and hotel-generated photos provide additional information, apart from the textual information mined from review texts extensively used in extant studies.

Figure 2. Object recognition process

Future research may consider the following directions. (1) *Difference between visual information detected from consumer-generated photos in upscale hotels and budget hotels and between negative and positive reviews.* Research has shown that users are likely to post photos of upscale hotels or when they had positive experiences (An *et al.*, 2020). By analyzing large-scale consumer-generated photos using CNN-based models, we can observe the items consumers are concerned with the most when checking in upscale hotels versus budget hotels as well as their criticism (praise) in negative (positive) reviews. (2) *Knowing whether visual information richness reduces textual information richness.* Consumers may reduce their effort in writing reviews if they provide abundant visual information. However, the supply of visual information may also induce textual information inputs. (3) *Factors that affect consumers' willingness to share visual information.* Uploading photos is voluntary and involves privacy disclosure. Thus, knowing the motivations that lead consumers to share visual information online is critical for a platform's operations. The way extrinsic motivation (e.g., reward points) or intrinsic motivation (e.g., self-determination theory) functions in this process should be an important research direction as well.

4.2 Facial recognition in hospitality and tourism

On peer-to-peer accommodation platforms, trust and the characteristics (e.g., gender, age, race, etc.) of hosts and guests are crucial to conclude transactions. Guests and hosts would infer each other's characteristics from their profile photos before making decisions. Therefore, visual information extracted from profile photos is of immense value for researchers in exploring the preferences of guests and hosts and

their potential selection biases. Figure 3 illustrates the recognition process of facial characteristics through the profile photos of users with Xception. The human face image dataset, such as LFWA or FER2013, can be used to train CNN-based deep learning models. By the same token as in Figure 2, the dataset should be separated into a training set, which is used to train the CNN-based deep learning model, and a testing set, which is utilized to verify the performance of the trained model. After good performance is achieved, the trained model can be used to perform facial recognition on the research samples. Outputs of facial recognition include gender, age, race (e.g., white, black, and Asian), attractiveness (Li *et al.*, 2022), and facial expression (e.g., happy, neutral, and sad) (Li *et al.*, 2021). Trained with FER2013, the Xception model can reach good performance in the classification of facial expression; the rates of accuracy are as follows: happy (92.38%), neutral (86.19%), and sad (84.31%).

Figure 3. Facial recognition process

Future research may consider the following avenues. (1) *Effects of visual information disclosed by the profile photos of guests on their sharing intention and rating behavior.* Facial information extracted from the profile photos of guests provides potential cues to infer their personality (e.g., outgoing, friendly), thereby possibly affecting their online sharing intention and rating behavior. (2) *Effects of visual information disclosed by the profile photos of hosts on pricing and online sales.* Facial information extracted from the profile photos of hosts provides additional information for potential guests to infer the trustworthiness and attractiveness of hosts on peer-to-peer accommodation platforms. This information has the potential to affect the pricing of hosts' listings and the online sales of their listings. (3) *Using visual information to infer the ethnic or cultural backgrounds of people and their potential impact on people's online behavior.* To reduce anonymity and build trust, peer-to-peer accommodation platforms encourage hosts and guests to use their real photos as profile photos. Thus, visual information reflected in profile photos generally conveys their personal and even privacy information. Some CNN-based deep learning models can recognize the ethnic groups of human faces. Thus, people's nationalities or cultures could be feasibly inferred through their profile photos. Compared with usernames or the linguistic styles of posts, visual information extracted from profile photos is more direct and accurate in identifying people's backgrounds. Thus, research on the impact of ethnic or cultural background on the online behavior of users is doable. (4) *Anti-discrimination policy-making.* In October 2018, Airbnb took action to fight against discrimination and stated as follows: "Moving forward, rather than displaying a potential guest's profile photo before the booking is accepted, hosts will receive a guest's photo in the booking process only after they've accepted the booking request." Profile images are a direct sign providing racial-related information of hosts or guests. Therefore, whether and when to display their profile photos to each other is critical in reducing racial discrimination on peer-to-peer accommodation platforms. (5) *Fine-tuning the deep learning models with a real research sample*

dataset should be valuable in improving recognition accuracy and reliability.

5. Conclusion and Discussion

5.1. Concluding comments

From the review conducted, we conclude this critical reflection to go on to describe potential theoretical and practical implications derived and future research avenues. Information asymmetry—the different levels of information that consumers and firms possess on products and services offered to the market—is a particularly critical component in the hospitality and tourism industry. If firms within this industry cannot demonstrate their superior quality, then low-quality competitors could push them out of the market. In the current environment, in which the prevalence of visual components is considerably common, the hospitality and tourism industry has rolled out its own devices to help customers diminish information asymmetry and the subsequent uncertainty. These tools are particularly relevant to the hospitality and tourism industry because of how highly sensitive to information this industry is and the fact that customers are unable to evaluate the quality of service prior to arriving at the destination. The quality of tangible elements (e.g., hotel rooms) and intangible components (e.g., friendliness of peer-to-peer accommodation hosts) must be considered as factors inducing uncertainty on the part of customers.

In the preceding context, the use of photos plays a critical role. We are not just referring to the traditional role that images play when firms publish their promotional content; more importantly, we refer to the content that other customers have posted (i.e., customers who share their own experiences). Evidently, these visual elements are clues for prospective customers so they can make informed decisions. In addition, when customers post pictures of themselves, they are allowing service providers to know who they are, with all the knowledge derived that it involves.

We use the literature review conducted as basis to state that an increasing number of applications are beginning to flourish. The use of CNN avoids the issue of excessive parameters with the subsequent low-speed and over-fitting problems. Moreover, two illustrations are provided. The first one shows the issue of *object recognition*, through which pictorial elements can be extracted from the photos. Hence, images are treated not only as a single unit but also as a multi-component unit from which the individual component can be obtained. The other presents the issue of *facial recognition*, through which the characteristics of individuals can be extracted, that is, individuals who can be customers but also hosts in the case of peer-to-peer accommodations.

5.2. Theoretical implications

Several theoretical implications can be derived from the analysis of visual elements included in online reviews. First, further knowledge of the factors that prompt individuals to share specific types of visual information would advance the literature on

this matter. Beyond the factors that lead people to enhance their willingness to post photos is their motivation to share specific images. Knowing these motivations would enable the researcher to have a comprehensive understanding of the decision process that customers follow. Second, the market has evolved considerably since the Nobel laureate George Akerlof postulated the principle of information asymmetry, and mechanisms involved in the reduction of such information asymmetry have developed as well. Along this line, visual elements provided by customers represent an alternative to the old strategy of publishing “official” images on the part of firms. The addition of this customer-generated visual component provides relevant facets to the paradigm created by the existence of information asymmetry. Third, in terms of the effectiveness of photos in attracting prospective customers—even if such effectiveness has fundamental managerial implications (as outlined later)—its analysis entails discovering the way people think and the manner by which they process information. On the basis of the tenets of information processing theory, new developments can be attained when visual elements along with texts are combined and individuals face both to arrive at their decisions. And fourth, when analyzing firm performance, the potential existence of nonlinearities in the relationship between specific determinant factors and critical performance metrics (e.g., trustworthiness, intention to book, and satisfaction) would contribute to the literature. If the effects of photos (measured through, for example, the number of images or specific elements that appear in photos) have an inverted U-shaped effect on, for example, intention to book, then some type of saturation point will be reached. From a theoretical perspective, uncovering these types of nonlinearities is relevant because the influence of photos could be linked to theories that do not expect to have a linear effect.

5.3. Practical implications

Knowledge of how to extract valuable visual information from large-scale user-generated photos to infer the online behavior of consumers and service providers and its influence on purchase decision and firm performance is crucial for business practice in hospitality. The following practical implications can be proposed: First, if managers know the factors that prompt customers to post photos, then enticing actions could be developed on their part to incentivize customers to share tangible, visual components of their experiences of the service. For example, knowing whether internal or external motivation or whether monetary or non-monetary incentives work would help managers design their enticing actions. Another example is offering a discount—or a goodbye gift—if customers commit to upload some images of their stay at the property. Second, if service providers know which strategy is best to reduce information asymmetries, then they would be in a better position to allocate resources to available actions, either traditional actions (e.g., warrants or quality certificates) or visual-related tactics. Rather than merely distributing a promotional budget among different activities, managers could aptly select the one that is optimal. This situation is particularly relevant because all the aforementioned actions come at a cost: receiving a quality certificate implies a huge amount of investment and considerable paperwork.

Third, knowledge of factors that increase the effectiveness of visual elements in reviews would assist service providers in their task to identify reviews expected to exert a greater impact than others with different characteristics. However, identifying high-impact reviews is no minor task because taking special care of these reviews can enhance the potential of firms. Fourth, depending on the performance metric that managers decide to prioritize, the emphasis on different measures could be placed according to the goals pursued. In particular, emphasis on trustworthiness would allow reducing information asymmetries, on intention to book would allow firms to boost sales, and on customer satisfaction would permit the detection of areas of improvement when disconfirmation happens. And fifth, finding a balance between textual and visual elements (e.g., photos) would help managers determine the “optimal” review. With the current technologies, analysts can ascertain the ideal review that has the best impact overall. If this information is known by managers, they can point customers in the right direction as to how to post reviews. For example, if a certain number of photos are posted, and according to their content an “optimal” balance between photos and text can be determined, then managers would know the extent by which visual information should be complemented by textual information.

5.4. Future research lines

From the taxonomy presented in Table 1, several future research avenues can be opened. A first basic research line would consist of analyzing individuals’ willingness to disclose information. A first step is the basic decision on whether customers are predisposed to post information during and/or after consumption. Although knowing the determinants that lead customers to break this first barrier that prompts them to publish information is essential, examining the factors that explain the types of information that customers post is critical. Accordingly, the following questions should be answered: Why do some customers just post text and others go further and post photos? Why are some customers more prepared than others to post photos of their experiences by sharing images of the products and services they have enjoyed? When are they more predisposed than others to post photos of themselves? Given that this research avenue entails two decisions—whether to post and which content to post—both outcomes should be examined simultaneously because the two decisions are interdependent. A second research avenue would consist of looking into which types of photos are more impactful among those posted by customers: images with tangible elements of the service or images included in the profile of the customer. Another research thread could focus on the identification of the determinants that explain the effectiveness of photos. Relevant to this research line is the conceptualization of the performance, in which photos are effective, because it can be seen from multiple perspectives. It can be about the effectiveness in increasing customers’ trustworthiness (with the consequent reduction in uncertainty on their part). It can also be in the intention to book (with the consequent improvement in revenues on the part of firms) or in customer satisfaction (which would allow comparisons pre- and post-consumption if expectations prior to consumption were identified). A fourth

research avenue would consist of the analysis of the potential moderating effect of the posted text on the posted photo would certainly enrich the comprehensiveness of the current knowledge on this matter. Finally, beyond the quantitative relationship between photos and texts, the qualitative relationship can also unearth relevant outcomes. With current CNN-based models, analysts can identify objects portrayed in the posted photos. Key attributes described in the text can be extracted by using aspect extraction techniques. Thus, the link between the aspects obtained from the text and the objects identified in the pictures can help find degrees of homogeneity in the overall review. In particular, the examination of the outcomes of the visual and textual components could show the coherence of the review, such as whether both components present similar information and if emphasis is placed on the same elements.

References

- An, Q., Ma, Y., Du, Q., Xiang, Z., and Fan, W. (2020), "Role of user-generated photos in online hotel reviews: An analytical approach", *Journal of Hospitality and Tourism Management*, Vol. 45, pp. 633-640.
- Trpkovski, A., Vu, H. Q., Li, G., Wang, H., and Law, R. (2018), "Automatic hotel photo quality assessment based on visual features", *Information and Communication Technologies in Tourism*, pp. 394-406, Springer.
- Barnes, S. J. (2021), "Understanding the overvaluation of facial trustworthiness in Airbnb host images", *International Journal of Information Management*, Vol. 55, pp. 102265.
- Barnes, S. J., and Kirshner, S. N. (2021), "Understanding the impact of host facial characteristics on Airbnb pricing: Integrating facial image analytics into tourism research", *Tourism Management*, Vol. 83, pp. 104235.
- Battiti, R., Brunato, M., and Battiti, F. (2021), "RoomTetris in room committing: Why the role of minimum-length-of-stay requirements should be revisited", *International Journal of Contemporary Hospitality Management*, Vol. 33 No. 11, pp. 4017-4034.
- Bigne, E., Chatzipanagiotou, K., and Ruiz, C. (2020), "Pictorial content, sequence of conflicting online reviews and consumer decision-making: The stimulus-organism-response model revisited", *Journal of Business Research*, Vol. 115, pp. 403-416.
- Bufquin, D., Park, J.-Y., Back, R. M., Nutta, M. W. W., and Zhang, T. (2020), "Effects of hotel website photographs and length of textual descriptions on viewers' emotions and behavioral intentions", *International Journal of Hospitality Management*, Vol. 87, pp. 102378.
- Doborjeh, Z., Hemmington, N., Doborjeh, M., and Kasabov, N. (2022), "Artificial intelligence: A systematic review of methods and applications in hospitality and tourism", *International Journal of Contemporary Hospitality Management*, Vol. 34 No. 3, pp. 1154-1176.
- Ert, E., Fleischer, A., and Magen, N. (2016), "Trust and reputation in the sharing economy: the role of personal photos in Airbnb", *Tourism Management*, Vol.

55, pp. 62-73.

- Ert, E., and Fleischer, A. (2017), "What in a photo makes you trust a person online? A structural equation modeling approach", *SSRN*, available at: <http://dx.doi.org/10.2139/ssrn.2992043>.
- Essien, A., and Chukwukelu, G. (2022), "Deep learning in hospitality and tourism: A research framework agenda for future research", *International Journal of Contemporary Hospitality Management*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/IJCHM-09-2021-1176.
- Filieri, R. (2016), "What makes an online consumer review trustworthy?", *Annals of Tourism Research*, Vol. 58, pp. 46-64.
- Fagerström, A., Pawar, S., Sigurdsson, V., Foxall, G. R., and Yani-de-Soriano, M. (2017), "That personal profile image might jeopardize your rental opportunity! On the relative impact of the seller's facial expressions upon buying behavior on Airbnb", *Computers in Human Behavior*, Vol. 72, pp. 123-131.
- Gambetti, A., and Han, Q. (2022), "Camera eats first: Exploring food aesthetics portrayed on social media using deep learning". *International Journal of Contemporary Hospitality Management*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/IJCHM-09-2021-1206.
- Jaeger, B., Slegers, W. W. A., Evans, A. M., Stel, M., and van Beest, I. (2019), "The effects of facial attractiveness and trustworthiness in online peer-to-peer markets", *Journal of Economic Psychology*, Vol. 75, pp. 102125.
- Kuo, P.-J., Zhang, L., and Cranage, D. A. (2015), "What you get is not what you saw: Exploring the impacts of misleading hotel website photos", *International Journal of Contemporary Hospitality Management*, Vol. 27 No. 6, pp. 1301-1319.
- Li, C., Ye, Q., Nicolau, J. L., and Liu, X. (2021), "Smiley guests post long reviews!", *International Journal of Hospitality Management*, Vol. 96, 102963.
- Li, Y., Peng, L., Ma, S., and Zhou, X. (2022), "Beauty premium or beauty penalty in sharing accommodation situations based on lay theories", *International Journal of Contemporary Hospitality Management*, Vol. 34 No. 3, pp. 929-950.
- Lo, I. S., McKercher, B., Lo, A., Cheung, C., and Law, R. (2011), "Tourism and online photography", *Tourism Management*, Vol. 32 No. 4, pp. 725-731.
- Lyu, J., Khan, A., Bibi, S., Chan, J. H., and Qi, X. (2022), "Big data in action: An overview of big data studies in tourism and hospitality literature", *Journal of Hospitality and Tourism Management*, Vol. 51, pp. 346-360.
- Lin, M. S., Liang, Y., Xue, J. X., Pan, B., and Schroeder, A. (2021), "Destination image through social media analytics and survey method". *International Journal of Contemporary Hospitality Management*, Vol. 33 No.6, pp. 2219-2238.
- Negri, F., and Vigolo, V. (2015), "Hotel attributes and visual image: a comparison between website and user-generated photos", *Information and Communication Technologies in Tourism 2015*, pp. 621-633.
- Yan, R., Zhang, K. Z., and Gong, X. (2021), "Listing popularity on the peer-to-peer

accommodation platform: the heuristic-systematic and uncertainty reduction perspectives”, *International Journal of Contemporary Hospitality Management*, Vol. 33 No. 12, pp. 4161-4181.

Zhang, S., Lee, D., Singh, P. V., and Srinivasan, K. (2017), “What makes a good image? Airbnb demand analytics leveraging interpretable image features”, *SSRN*, <http://dx.doi.org/10.2139/ssrn.2976021>.

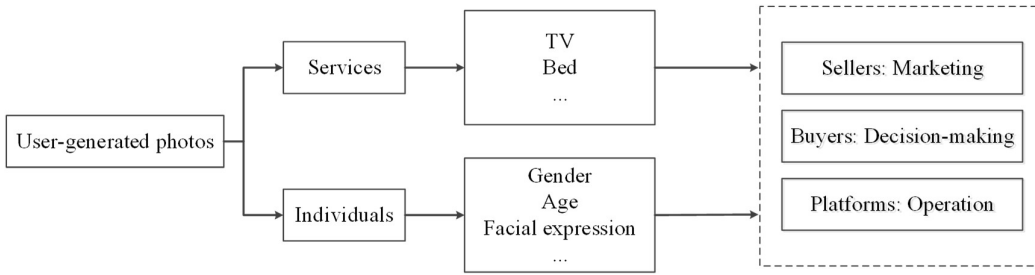


Figure 1. Conceptual framework

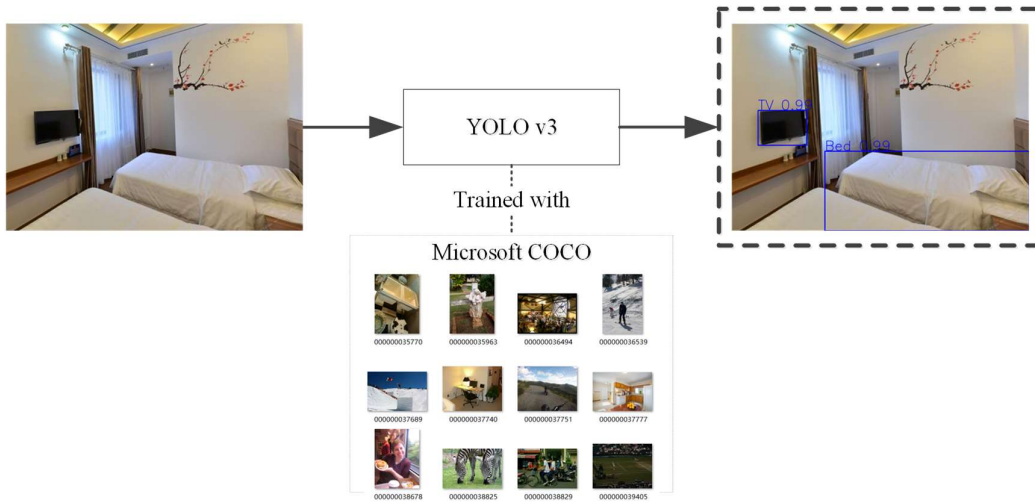


Figure 2. Object recognition process

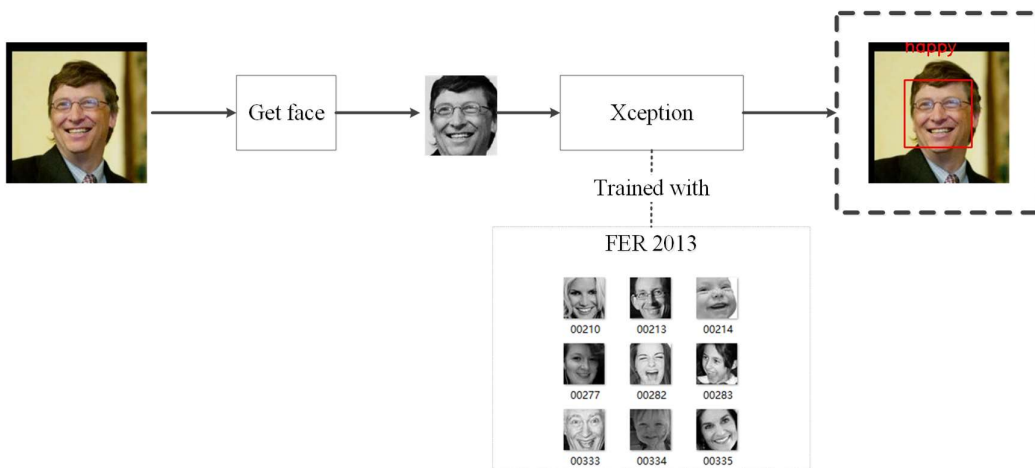


Figure 3. Facial recognition process

Table 1. Visual information categorized by user and content

		Contents of photos	
		Services	Individuals
User-generated photos	Buyer-generated	Buyer-generated service information	Buyer-generated individual information
	Seller-generated	Seller-generated service information	Seller-generated individual information