How to Enhance Online Hotel Ad Effectiveness Based on Real-World Data: Mobile Eye-Tracking and Machine Learning Tell

Wen Xie, University of Houston
Ming Chen, UNC Charlotte
Zhu Han, University of Houston

ABSTRACT

Marketing researchers have adopted eye-tracking technology to extract information on consumers' eye-fixation and attention in order to fully understand their shopping behaviors under online and offline settings. One application is to examine the patterns of eye-fixation from different consumers on a certain website and to provide insights on the effectiveness of online advertisements. However, because of the complexity of data analysis and computational burden, it is quite challenging for researchers to track consumers' visual focus and investigate eye-fixation patterns in a dynamical way. In this paper, we try to solve the aforementioned challenges by utilizing the powerful computing tool of machine learning (ML) methods to investigate consumers' areas of interest (AOIs) on a hotel booking website using mobile eye-tracking technology. From the eye-tracking data collected from a field experiment, by applying the ML YOLO approach, we find that the mean average precision (mAP) is 92.35% which indicates high accuracy to identify AOIs. Furthermore, the findings show that different AOIs receive different levels of attention reflected on different eye-fixation counts and there exists consumer heterogeneity in the attention distribution amongst different AOIs. From the managerial perspective, this paper suggests that web designers can provide the online hotel ad layout according to personal preferences and make customized recommendations spontaneously based on a consumer's eye-fixation patterns and demographics. The insights provided from this research can help evaluate the effectiveness of the online hotel advertisements layout on the booking rate.

Keywords: *mobile eye-tracking, machine learning, online hotel ad layout, advertising effectiveness*

Description: *Combining machine learning and mobile eye-tracking technology, this research investigates the differences of consumers' eye-fixation counts with respect to different Areas of Interest of online hotel ads and then proposes suggestions for enhancing the advertising effectiveness.*

Introduction

The total spending on travel and tourism is reaching $1.1 trillion in 2018 (the U.S. Travel Association), which accounts for 5% of the U.S. Gross Domestic Product of the United States. Most of these expenditures happen on Online Travel Agency (OTA) websites due to its convenience. $7.5 Billion of the total spending went to hotel and lodging. The common OTAs include Expedia, Hotels, TripAdvisor, Priceline, Booking, Orbitz, etc. The online advertisement on each OTA has its unique attributes and layout. Obviously, each OTA pays much attention to the way of enhancing the effectiveness of its online advertisements because that's a key to increase the OTA's competitiveness and booking rate.

Many researchers have investigated the importance of different attributes (e.g, price, text, and image) of online tourism ads (Law and Hsu 2006; Pan, Zhang, and Law 2013). In addition, some of the previous studies try to link the perceived importance of different attributes with consumer's demographic information such as age, gender, and education

For further information contact: Wen Xie, University of Houston (wxie5@uh.edu).

level (Pan et al. 2004; Marchiori and Cantoni 2015). To measure consumers' perceived importance, many methods (e.g, Verbal protocols and self-regulated survey) were used in different studies. In recent years, the eye-tracking method stands out in finishing data collection. Compared to data collected by self-regulated survey methods, eye-tracking data is more objective and less subject to bias (Scott et al. 2017). Russo (1978) also argued that eye-tracking data has excellent validity and detail compared with other methods (e.g. Input-output analysis and Verbal protocols) in his pioneering article: "Eye-Fixation Can Save the World." Since then, eye-tracking technology has been widely used in many research domains such as Marketing, Psychology, Reading, and Engineering (Rayner 1998; Duchowski 2002; Wedel and Pieters 2008).

According to three survey articles about the use of eye-tracking in Visual Marketing (Wedel and Pieters 2008; Wedel 2013; Scott et al. 2017), many researchers have addressed the impacts of top-down factors (e.g. memory, attitudes, emotions, and goals) and bottom-up factors (e.g. spatial location, color, lines, and size of objects) on the effectiveness of different kinds of advertisements in different scenarios. However, most of the eye-tracking studies deal with fixed stimuli or employ fixed eye-tracking devices because of their specific data characteristics of easy analysis (Scott et al. 2017). Mobile eye-tracking devices are seldom used due to the specific issue such that the visual stimuli observed by participants are always changing and each participant may view different scenes that complicate the analysis of data and even make it unfeasible when handling huge eye-tracking data using professional software (e.g. Tobii Pro Lab). However, it's much closer to real-world settings when employing mobile eye-tracking devices (Kassner, Patera and Bulling 2014).

In order to conduct research based on real-world conditions, this study, to our knowledge, is the first to propose to use the machine learning (ML) method to overcome the computational burden and complexity of mobile stimuli data analysis. The proposed approach can take place of traditional professional software such as Tobii Pro Lab to analyze mobile eye-tracking data collected in dynamic scenarios. From a field experiment, the authors, using a mobile eye-tracking glass, collect participants' eye movement data when they book a hotel on the OTA Booking.com website. The ML YOLO model is employed to extract and identify our focused areas of interest (AOIs) in the scene video recorded by the eye-tracking device. After training and testing its performance, we find that each average precision of all our focused AOIs is above 90%, and the mean average precision (mAP) is 93.25% which means that the proposed model identifies the AOIs successfully. The precision rate and recall rate are high while the false positive rate and missing probability are low. Meanwhile, to investigate participants' attention distribution, we compute their eye-fixation counts on each AOI. The results show that different AOIs receive different attention, and different participants have different attention distribution. This research provides insights on evaluating the effectiveness of the online hotel ads.

Section 2 describes the online tourism relevant literature and our research questions. Section 3 presents the proposed approach to analyze our focused AOIs by the ML method. In Section 4 we show the results. Finally, we give concluding remarks in Section 5.

Relevant Literature

Previous studies related to enhancing the effectiveness of online tourism and hospitality ads using eye-tracking method mainly investigated consumer's perceived importance of attributes such as price, text, and image. Law and Hsu (2006) found that five dimensions (Reservations information, facilities information, contact information, surrounding area information, and Website management) have different perceived importance through a large-scale survey, which indicated potential different perceived attention. Pan, Zhang, and Law (2013) investigated the effects of the number of hotel options, and the position of the hotel option in the OTA search engine results page (SERP). They found that hotel options near the top of the list received more attention than those showing up later which concluded the same results with later research (Ert and Fleischer 2016). Additionally, it was found that the hotel image can alleviate the perceived information-overloaded problem when a consumer's exposed to too many hotel options. Some studies concluded that pictures attract more attention than textual material (Aicher et al. 2016; Noone and Robson 2014). However, Kong et al (2019) found that web advertisement with text-only advertisement was the most efficient. It's worthy to conduct research to make it clear about consumer's attention distribution on these typical attributes.

Except that researchers investigated the importance of different attributes using the eye-tracking method, some studies furthermore tried to analyze the relationships between consumer's demographic information and their behaviors. Pan et al. (2004) explored the determinants of ocular behavior and found that the gender of subjects influences online ocular behavior significantly. Lorigo et al. (2008) surveyed the use of eye-tracking in investigations of online search about how users view the ranked results on SERP. The results show that the rank rather than gender has a strong influence on viewing behavior and viewing patterns. Hernández-Méndez and Muñoz-Leiva (2015) carried out a mixed design with an eye-tracking method and a self-administered questionnaire. They found that participants' duration of fixation on the banner does not reflect significant differences when comparing gender, and experience level. In terms of age, it was found that young people take longer to reach the banner and view it for longer periods of time than senior people. Marchiori and Cantoni (2015) found that younger people are more likely to change their opinion about tourism destinations based on brief exposure to user-generated content (UGC) on online social media.

Proposed Research Analysis

Research Questions

The extant literature reviewed above, suggests that consumers may respond differently to the same ad. We believe that displaying personalized ads and customized recommendations that take into account customers' attention can enhance advertising effectiveness. Further research should also look to shed light on the links between advertisement response and other consumers' demographics, such as weight and marital status. The present study investigates the following three research questions in the hospitality and lodging context:

1. Do different AOIs of online hotel ads attract different levels of consumers' visual attention (operationalized as consumers' eye-fixation counts)?
2. Is there consumer heterogeneity in the attention distribution amongst different AOIs of online hotel ads?
3. How does an information-overloaded environment (too many hotel options examined by consumers) affect consumers' ability to find their ideal hotel options?

Methodology

In this study, we focus on four AOIs including hotel picture, rating, price, and text information (including hotel name, address and others together in the middle of each hotel option). To track participants' eye movement data when they book a hotel on an online website, we use Tobii Pro Glass 2 mobile eye-tracking device. The scene camera resolution is 1920 × 1080 pixels with 25 frames per second (FPS). The horizontal recording angle/visual angle of the scene camera is 82° and the vertical recording angle is 52°. After finishing eye movement data collection using the mobile eye-tracking device, we propose to utilize the powerful computational ability of the ML algorithm to identify different AOIs in the huge recorded dynamic eye-tracking data. Here to identify and track our focused AOIs, we choose YOLO (You Only Look Once) version 3 algorithm which is widely used in object identification scenarios because of its high speed and high accuracy (Redmon and Farhadi 2018).

The collected eye-tracking data include live video stream (LVS) and live data stream (LDS). The LVS recorded by the scene camera is about what the participants are looking at and the LDS is the coordinates where participants' eye gaze point is in real-time. The eye gaze data consist of fixations and saccades (Busswell 1935). Saccades are rapid, ballistic jumps of the eyes, typically lasting around 20–40 milliseconds. Fixations are moments during which the eye is relatively still, typically lasting around 200–500 milliseconds (Rayner 1998). In this study, we compute eye-fixation counts to reflect participants' attention distribution on AOIs. With the LVS, we train our YOLO model and identify the AOIs. To measure the performance of the YOLO model, the typical metrics: mAP and precision-recall curve are used, respectively, as

(1) $$Precision = \frac{True\ positive}{True\ positive\ +\ False\ positive}$$

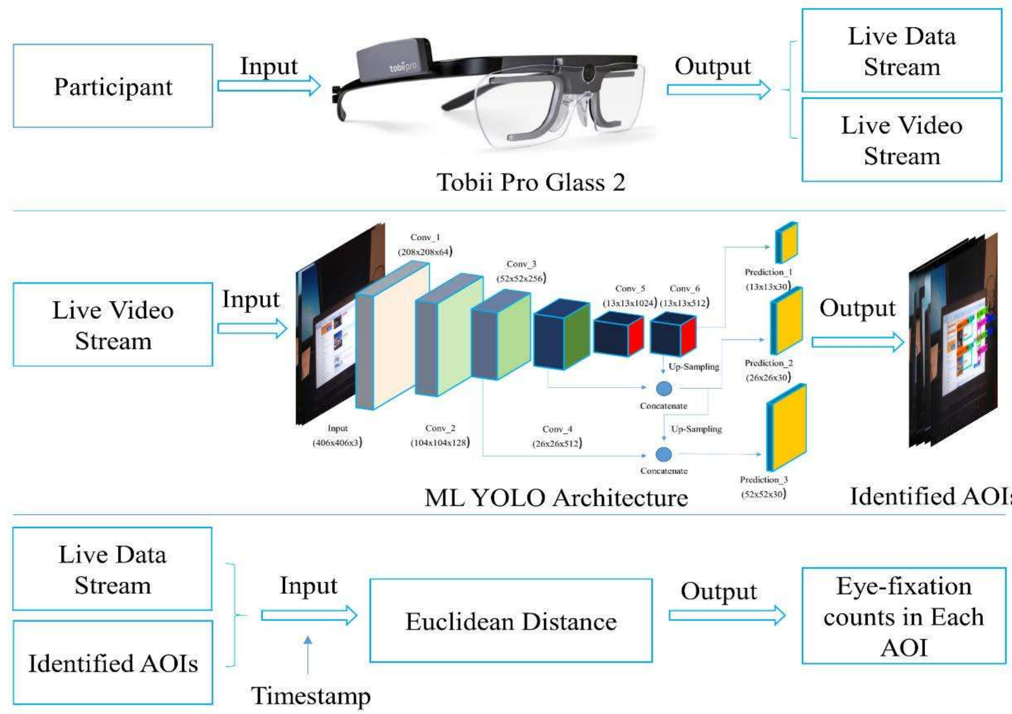
(2) $$Recall = \frac{True\ positive}{True\ positive + False\ negative}$$

After training and testing the model, the identified AOIs are synchronized and matched with the LDS according to timestamp and the Euclidean Distance. Meanwhile, eye-fixation counts in different AOIs are calculated to reflect consumers' attention on each AOI. Therefore, we know clearly what and where participants are looking at in real-time and their perceived importance of different AOIs. The whole frame of our methodology is depicted in Figure 1.

Study Design

Setup. This study is designed imaginary in a lab-setting environment. Each participant is asked to browse the SERP of hotels on the booking.com website and select one they will book. Each participant is told with four assumptions in advance. The first is that each participant is an adult. The second is that each participant will go to the University of Texas at Austin. The third is that each participant will stay there for 7 nights. The fourth is that each participant is alone. Participants are supposed not to click through each hotel option for details when browsing the hotel options. In other words, we only consider the browsing stage (Noone and Robson 2014) in this study. Besides each participant is supposed to finish the booking process in two to five minutes which may make a difference in an information-overloaded environment. Given that the previous study found that 20 hotel options seem to overwhelm the participants (Pan, Zhang and Law 2013), we choose the first 18 hotel options on SERP. Before the experiment, the SERP has been displayed on the computer which includes the first 18 hotel options. Participants will not need to search and can start to browse the results directly. The 18 hotel options are displayed by a laptop with resolution 1366*768. When participants are browsing the SERP, they can use the mouse to move the page up and down to scrutinize each hotel option.

Figure 1. The Frame of the Whole Methodology



Procedure. Participants will sign in the consent form first and then they need to fill out a demographic survey. Next, they will wear the Tobii Pro Glass 2 under an instructor's help and calibrate it using the one-point calibration procedure. After calibrating successfully, they are told to find their comfortable sitting posture sitting at the front of the laptop and then they can start to browse the hotel options. When they make a decision, participants can stop and tell the instructor. At last, they will fill out an evaluation survey about their thinking importance of each AOI.

Data Processing

Four graduate students with normal vision (visual above .8 as a standard metric (Wang, Tsai and Tang 2018)) participate in our study. After finishing data collection, the LDS of three participants is valid while all LVS is useful. To train and test the YOLO model, every five frames from each live video are exported first which is reasonable considering that the FPS of the live video is 25. Each frame is like a screenshot showing what participants are looking at. Totally, we get 2,484 frames that are randomly divided into a training dataset (2,033 frames), validation dataset (225 frames) and testing dataset (226 frames). After that, we label each frame with the four AOIs using Yolo mark software. Next, we train the YOLO model. In the initial training process, we set training parameters as epochs 50, learning rate .001 and batch-size 16. In the further training process, we set training parameters as epochs 50, initial learning rate .0001 and batch-size 16. However, it early stopped at the 66th epoch. So we continue to train it with training parameters: epochs 50, initial learning rate .0001 and batch-size 8. Again, it early stopped at the 87th epoch. The reason for the early stop is that the validation loss does not decrease in consecutive three epochs.

After training the model, we test its performance with the training dataset. We use the final model to identify AOIs in the LVS. Finally, after synchronizing the identified AOIs with LDS at each timestamp and using the Euclidean Distance to match each eye-fixation with the identified AOI, we get the participant's eye-fixation counts on each AOI and visualize them with different figures intuitively shown in the next section.

Results and Analysis

In this study, the results mainly include two parts. The first part is about the performance of the proposed ML YOLO model used to identify AOIs in eye-tracking data. The sec-

Figure 2. Train and Validation Loss Curve

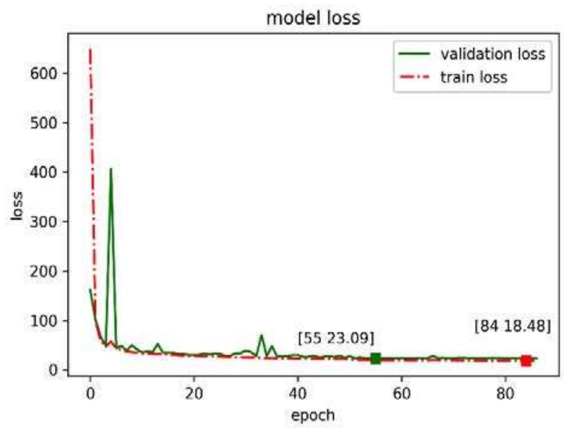


Figure3. Mean Average Precision

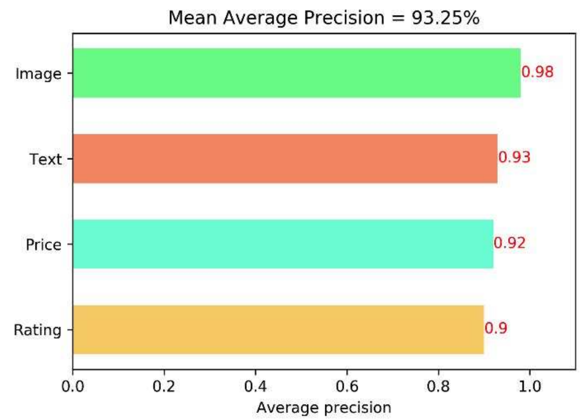


Figure 4. Average Precision of Hotel Picture

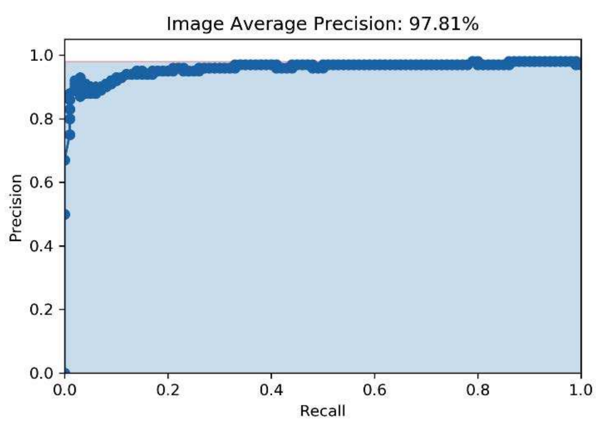
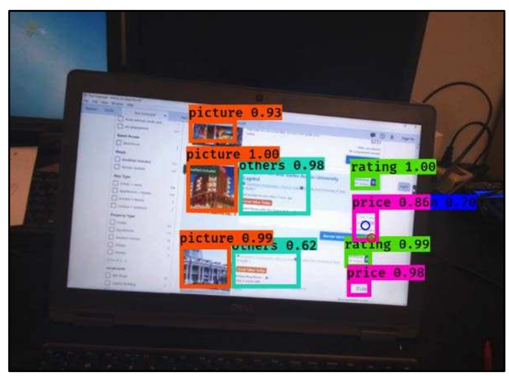


Figure 5. Screenshot of Identified AOIs



ond part is about participants' attention distribution on each AOI which is reflected by eye-fixation counts.

The train and validation loss curves of the YOLO model are displayed in Figure 2. The two losses decrease and stabilize around 18 and 23 respectively. The whole training process early stopped at the 87th epoch. The loss is used to measure the difference between the identified AOI and the true AOI. The less the loss is, the better the performance of the YOLO model is. Theoretically, train loss will decrease to 0. However, it's always a necessity to trade off the model accuracy and bias which is indicated by the train and validation loss. So, the loss curves mean a good training process.

The final model is selected to finish later procedures. The mAP and precision-recall curve of the image are displayed in Figures 3 and 4. Results show that the average precision (AP) of each AOI is above 90% and the curve in Figure 4 is close to the top-left point. Besides, the mAP is 93.25%. These indicate a high probability to identify the right AOIs and a low probability to regard the wrong AOIs as the right AOIs. In brief, the performance of our trained ML model is excellent in identifying the AOIs with both high accuracy and low missing possibility. With the trained ML YOLO model, the AOIs in the live videos are identified and synchronized with the LDS. In Figure 5, we can see rectangles with different colors, names, and probability, which shows the identified AOIs. The red circle refers to the eye-fixation where the participant is looking at, and the blue circle means which AOI we assign the eye-fixation to according to the Euclidean Distance. To show participants' attention distribution, their eye-fixation counts on each AOI are computed.

Figure 6. Attention Distribution of Participant 2.

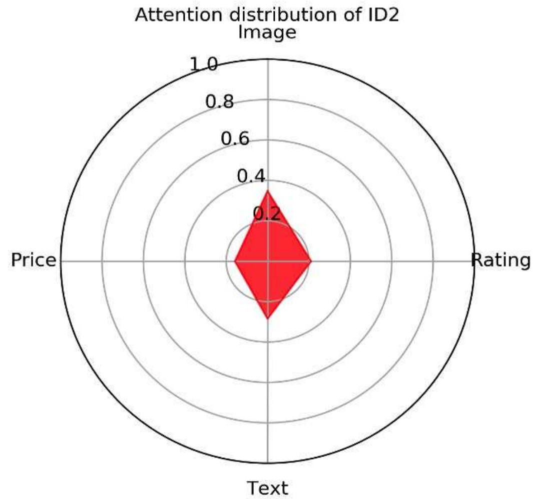


Figure 7. Attention Distribution of Participant 3

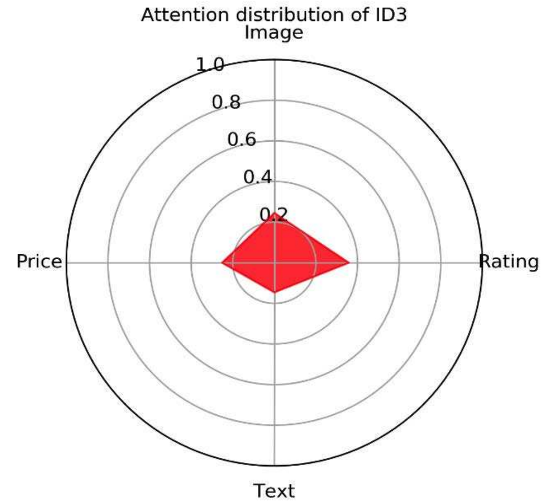


Figure 8. Attention Distribution of Participant 4

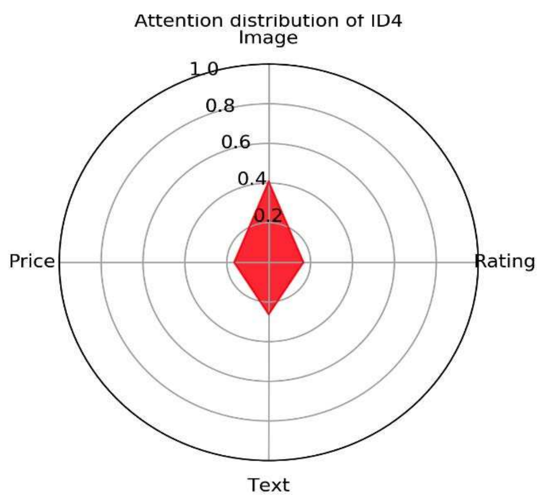
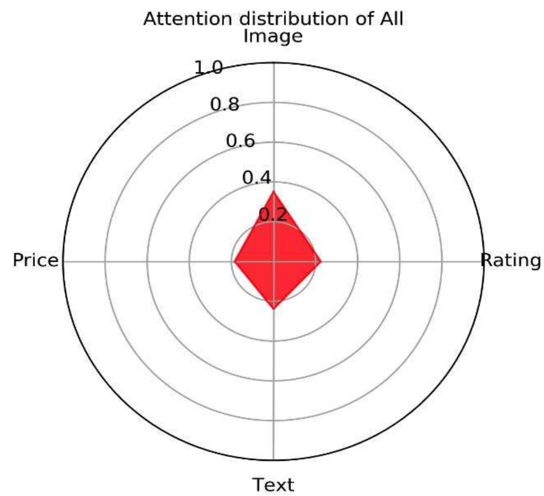
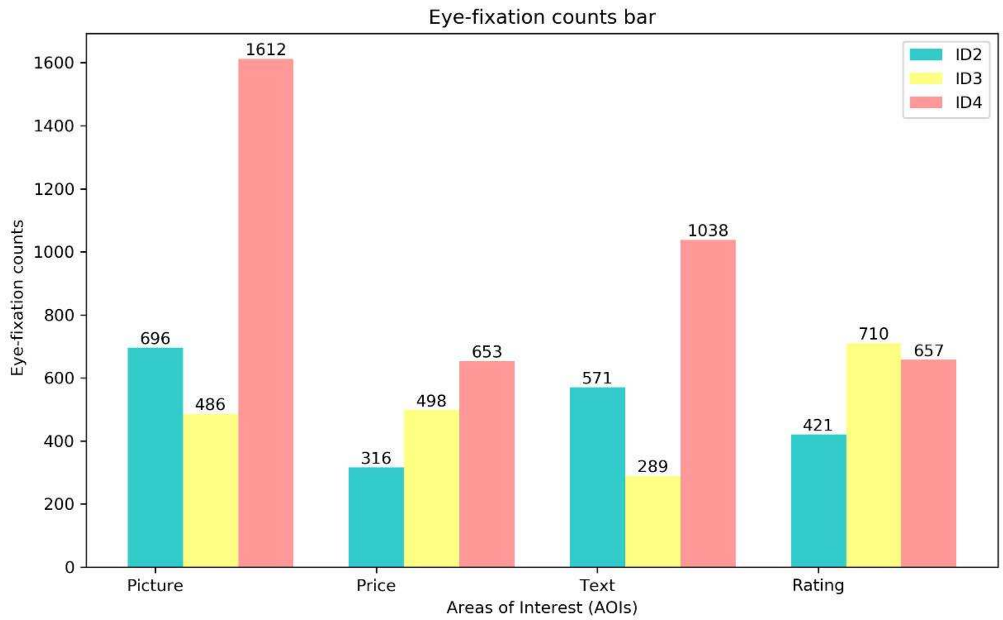


Figure 9. Total Attention Distribution



To describe participants' eye-fixation patterns on each AOI intuitively, we draw the polar plots. Figures 6, 7 and 8 show the eye-fixation patterns of each participant respectively. The eye-fixation counts are uniformed from 0 to 1. From Figures 6 and 8, we can see that these two participants pay the most attention (nearly 40%) on hotel image which is in the same line with previous researches (Aicher et al 2016; Noone and Robson 2014). However, participant 3 in Figure 7 pays the most attention to the hotel price and rating rather than the image, which indicates that he/she is trying to reduce his/her consideration set using different strategies, notably by focusing on price and rating (Pan, Zhang and Law 2013). So it's suggested that different consumers have different attention distributions on different AOIs. From Figure 9, the hotel image is found to be focused most, while the hotel price and rating receive less attention. The text information receives the least attention. Participants focus much on hotel image which can be explained by the recovery psychology. Some hotel images such as natural images can help individuals recover from stress or fatigue (Wang, Tsai, and Tang 2018). So, the result shows that different AOIs have received different levels of attention. Figure 10 shows the eye-fixation counts result. The larger the counts on the AOI is, the more attention the participant pays on the AOI. Moreover, the larger the total counts of each participant is, the more time the participant used until he/she found the ideal hotel. From Figure 10, it's

Figure 10. Eye-Fixation Counts of Each Participant



clear that participant 4 has the largest eye-fixation counts. The number of eye-fixation counts of participant 4 is about twice that of other participants, which may indicate that this participant encounters difficulty in finding the ideal hotel in the information-overloaded online environment. In summary, we have answered our three research questions based on real-world data analysis.

Conclusion and Discussion

Methodologically, we propose to employ the ML method into an online advertising setting. Our results show that the YOLO model identifies AOIs accurately and efficiently, which can tremendously reduce the workload from manual coding. To the best of our knowledge, this paper is the first to adopt the ML method into online advertising research using mobile eye-tracking study. It's the mobile eye-tracking device that makes experiments closer to the real-world setting. From the managerial perspective, the results show that different AOIs receive different attention and different consumers pay their most attention to different AOIs. As a result, the findings from our research provide suggestions on displaying the online hotel ad according to individual consumer's preferences based on their demographics and eye-fixation patterns. Different search filter parameters can be preset on the left side of the SERP such as the star rating, review score, etc. In addition, Web designers can pop up the focused AOIs or make more striking recommendations, so as to improve the effectiveness of the online ads on the booking rate.

References

Aicher J., Asiimwe F., Batchuluun B., Hauschild M., Zöhrer M., Egger R. (2016), *"Online Hotel Reviews: Rating Symbols or Text… Text or Rating Symbols? That Is the Question!"* In: Inversini A., Schegg R. (eds) Information and Communication Technologies in Tourism 2016. Springer, Cham, 369-382.

Busswell, G.T. (1935), *"How People Look at Pictures: A Study of the Psychology of Perception in Art,"* Chicago: University of Chicago Press.

Duchowski, A.T. (2002), *"A Breadth-First Survey of Eye-Tracking Applications."* Behavior Research Methods, Instruments, and Computers, 34 (4): 455–470.

Ert, E., and Fleischer, A. (2016), *"Mere Position Effect in Booking Hotels Online,"* Journal of Travel Research, 55(3), 311–321.

Hernández-Méndez, J., Muñoz-Leiva, F. (2015), *"What type of online advertising is most effective for eTourism 2.0? An eye-tracking study based on the characteristics of tourists,"* Computers in Human Behavior, 50, 618-625.

Kassner, M., Patera, W., and Bulling, A. (2014), *"Pupil: an open-source platform for pervasive eye tracking and mobile gaze-based interaction,"* In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication., 1151-1160.

Kong, S., Huang, Z., Scott, N., Zhang, Z., and Shen, Z. (2019), *"Web advertisement effectiveness evaluation:*

Attention and memory," Journal of Vacation Marketing, 25(1), 130–146.

Law, R., and Hsu, C.H.C. (2006), *"Importance of Hotel Website Dimensions and Attributes: Perceptions of Online Browsers and Online Purchasers,"* Journal of Hospitality & Tourism Research, 30 (3), 295–312.

Marchiori, E., Cantoni, L. (2015), *"The role of prior experience in the perception of a tourism destination in user-generated content,"* Journal of Destination Marketing & Management, 4(3), 194-201.

Noone, B., and Robson, S. (2014), *"Using Eye-tracking to obtain a deeper understanding of what drives online hotel choice,"* Cornell Hospitality Report, 14(18), 6–16.

Pan, B., Hembrooke, H.A., Gay, G.K., Granka, L.A., Feusner, M.K., and Newman, J.K. (2004), *"the determinants of web page viewing behavior: an eye-tracking study,"* Eye-tracking research & applications, 147-154.

Pan, B., Zhang, L., and Law, R. (2013), *"the Complex Matter of Online Hotel Choice,"* Cornell Hospitality Quarterly, 54(1), 74–83.

Russo, J.E. (1978), *"Eye Fixations Can Save the World: a Critical Evaluation and a Comparison Between Eye Fixations and Other Information Processing Methodologies,"* Advances in Consumer Research, 05, 561-570.

Rayner, K. (1998), *"Eye movements in reading and information processing: 20 years of research,"* Psychological Bulletin, 124 (3), 372-422.

Redmon, J. and Farhadi, A. YOLOv3 (2018), *"An incremental improvement,"* arXiv: 1804.02767.

Scott, N., Zhang, R., Le, D. and Moyle, B. (2017), *"A review of eye-tracking research in tourism*," Current Issues in Tourism, 22 (10), 1244-1261.

Wedel, M. and Pieters, R. (2008), *"Eye Tracking for Visual Marketing,"* Foundations and Trends in Marketing, 1 (4), 231-320.

Wedel, M. (2013), *"Attention Research in Marketing: A Review of Eye Tracking Studies,"* Robert H. Smith School Research Paper No. RHS 2460289.

Wang, T.C.; Tsai, C.L., Tang, T.W. (2018), *"Exploring Advertising Effectiveness of Tourist Hotels' Marketing Images Containing Nature and Performing Arts: An Eye-Tracking Analysis,"* Sustainability, 10, 3038.