Where Do Consumers Look When Viewing Fashion Advertisements? Saliency Based Models for Visual Attention Prediction

Seung-Hee Lee, Ying Chen, Ali Mahdi, and Jun Qin, Southern Illinois University, USA
Yuli Liang, Texas State University, USA

Keywords: Prediction, visual attention, advertisements, saliency modeling, salience algorithms

Conceptual Framework. Advertising is a key marketing tool and is costly to the company, so creating ads as effectively as possible to attract potential customers is essential for business success (Simmonds et al., 2020). Many industries, including the fashion industry, use magazine ads to deliver brand images to consumers, increase brand reputation, and encourage consumers to buy their brand products. An understanding of consumers' visual attention to advertising is vital because there is a robust relationship between consumers' visual attention and memory (Pieters et al., 2007). Most marketing communication devices work through the sense of sight. Advertisement content including brand logos, colors, packaging, letters and fonts all attract visual attention to create consumer awareness of a brand (Simmonds et al., 2020).

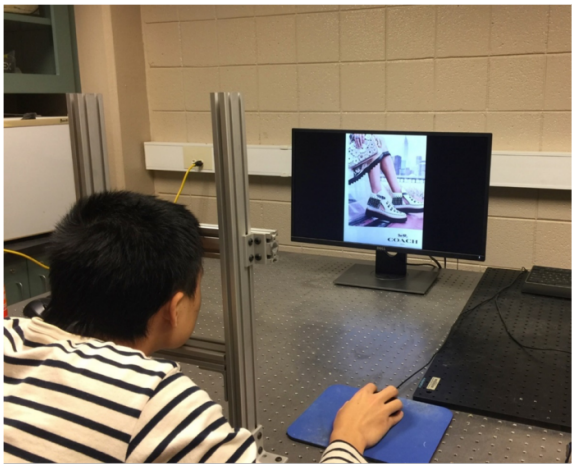
Human beings tend to minimize their neural resources by using eye, head, and body movement to shift their visual attention and gaze behavior toward more informative image spatial locations (Mahdi et al., 2017). In computer vision, a saliency map is a 2D topological map that indicates visual attention priorities using a numerical scale. A higher visual attention priority indicates the object of interest is irregular or rare in its surroundings. Recently, a diversity of saliency models has become a popular technical term for human vision attention study in fields such as communication and electronic engineering (e.g., Garnett et al., 2014; Mahdi et al., 2019; Tsiami et al., 2019). However, there is little research that applies saliency modeling to the fashion area. Thus, it is meaningful to investigate whether saliency modeling can be applied to the fashion area.

Purposes of the study. This multidisciplinary study were 1) to explore where consumers look at fashion advertisements using saliency based models, and 2) to investigate which saliency model(s) are the most effective tools in predicting consumers' visual attention when viewing fashion advertisements.

Method. Seventy college students were recruited to participate in the study at a mid-western university in US. Participants were from different majors such as fashion and electronic engineering. One hundred fashion images were selected from fashion magazines (e.g., *Vogue*) and divided into four image groups (25 images in each group). Each participant was seated about 1 meter away from the screen, viewed 25 fashion images with 5 seconds per image, and selected locations on each image that most attracted their attention by clicking a mouse on the screen. The mouse clicking points were saved as human attention data on the 100 fashion images. The experimental setup is shown in Figure 1. Moreover, six saliency models (Ground-truth, AW5, COV, DeepFeat, GBVS, ISEEL, LDS) were applied to create saliency maps (Erdem & Erdem, 2013; Garcia-Diaz et al, 2009; Harel et al, 2007; Mahdi et al., 2019; Tavakoli et al, 2017) on the

100 fashion images and compared with experimental human attention data maps using three evaluation metrics, an integral of the area under the receiver operator characteristics curve (AUC), Pearson's correlation coefficient (CC), and KL divergence.

Results. The results showed that all six saliency maps achieved a high agreement with the human attention data on fashion images of advertisements (as shown in Figure 2). It indicates that six saliency models can be used to predict where human subjects looked on the fashion advertisement images. Figure 3 shows quantitative evaluation of the performance of six saliency models on fashion advertisement images. The DeepFeat model consistently and significantly outperformed the other five saliency models. This means that the DeepFeat saliency model has the most effective performance in predicting consumers' attention when viewing fashion advertisements among six saliency models.



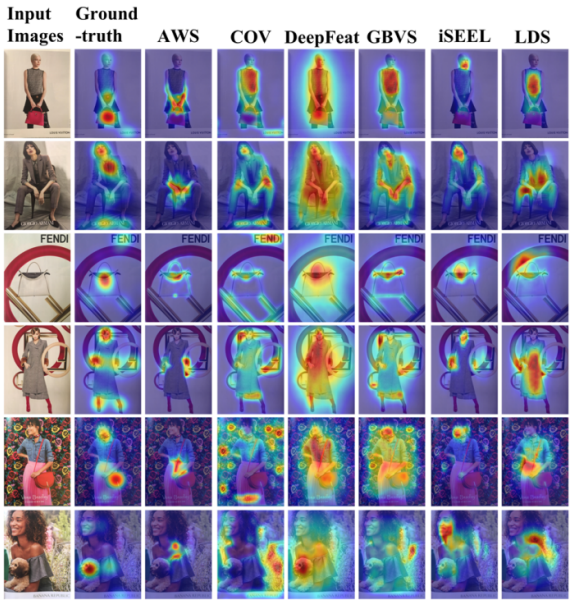


Figure 1 (Left): A photograph of the experimental setup of the human subject study.
Figure 2 (Right): Six representative fashion images and their experimental results. Ground-truth are experimentally recorded human attention data when viewing fashion advertisement images. AWS, COV, DeepFeat, GBVS, iSEEL, and LDS are saliency maps generated by six state-of-the-art saliency algorithms.

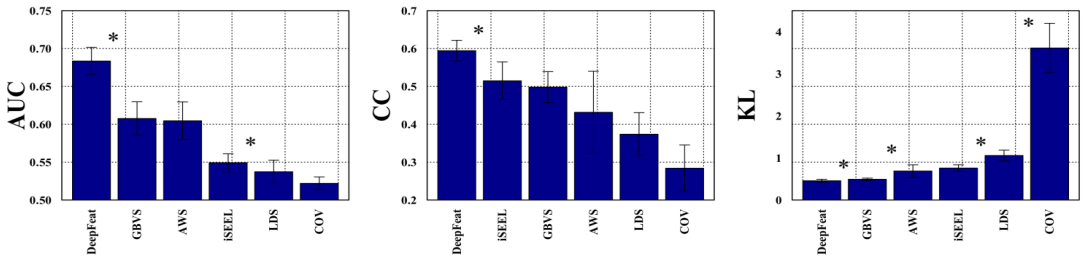


Figure 3: Averaged AUC, CC, and KL scores of six saliency models for evaluation of human attention on 100 fashion images. A * indicates two consecutive models are significantly different using t-test at $p > 0:05$. Models that are not consecutive have a larger probability to achieve statistical significance.

Discussion/implications. This study has important implications for marketing practitioners. The results indicate that the saliency algorithms can be a very useful tool for analysis and prediction of consumers' visual attention when viewing fashion advertisements. The saliency maps of images show where consumers are likely to focus their attention. This research can help fashion marketers better understand young consumers' visual attention when viewing fashion magazine advertisements based on saliency algorithms and saliency maps so that they can designate the most effective layout for their advertising marketing strategies. Further study is recommended to expand this research to examine other fashion advertisement types such as fashion web advertisements with more variety of samples.

References

Erdem, E., & Erdem, A. (2013). Visual saliency estimation by nonlinearly integrating features using region covariances. *Journal of vision*, 13(4), 11-11.

Garcia-Diaz, A., Fdez-Vidal, X. R., Pardo, X. M., & Dosil, R. (2009). Decorrelation and distinctiveness provide with human-like saliency. In *International Conference on Advanced Concepts for Intelligent Vision Systems* pp. 343-354.

Garnett, B. R., Buelow, R., Franko, D. L., Becker, C., Rodgers, R. F., & Austin, S. B. (2014). The importance of campaign saliency as a predictor of attitude and behavior change: a pilot evaluation of social marketing campaign fat talk free week. *Health Communication*, *29*(10), 984-995.

Harel, J., Koch, C., & Perona, P. (2007). Graph-based visual saliency. In *Advances in neural information processing systems,* pp. 545-552.

Mahdi, A., Su M., Schlesinger M., & Qin J. (2017). A Comparison Study of Saliency Models for Fixation Prediction on Infants and Adults. *IEEE Transactions on Cognitive and Developmental Systems*, *10*(3), 485-498.

Mahdi, A., Qin, J., & Crosby, G. (2019). Deepfeat: a bottom-up and top-down saliency model based on deep features of convolutional neural nets. *IEEE Transactions on Cognitive and Developmental Systems*, *12*(1), 54-63.

Pieters, R., Wedel, M., & Zhang, J. (2007). Optimal feature advertising design under competitive clutter. *Management Science*, *53*(11), 1815-1828.

Simmonds, L., Bellman, S., Kennedy, R., Nenycz-Thiel, M., & Bogomolova, S. (2020). Moderating effects of prior brand usage on visual attention to video advertising and recall: An eye-tracking investigation. *Journal of Business Research*, *111*, 241-248.

Tavakoli, H. R., Borji, A., Laaksonen, J., & Rahtu, E. (2017). Exploiting inter-image similarity and ensemble of extreme learners for fixation prediction using deep features. *Neurocomputing*, *244*, 10-18.

Tsiami, A., Koutras, P., Katsamanis, A., Vatakis, A., & Maragos, P. (2019). A behaviorally inspired fusion approach for computational audiovisual saliency modeling. *Signal Processing: Image Communication*, *76*, 186-200.