A Design Framework and AI System for Affective Destination Image Generation to Influence Tourists' Emotional Response

Thuc Nguyen
Deakin University, thuc.nguyen@deakin.edu.au

Lemai Nguyen
Deakin University, Lemai.nguyen@deakin.edu.au

Huy Quan Vu
Deakin University, q.vu@deakin.edu.au

Jacob Cybulski
Deakin University, jacob.cybulski@deakin.edu.au

A Design Framework and AI System for Affective Destination Image Generation to Influence Tourists' Emotional Response

Research-in-progress

Thuc Nguyen
Deakin Business School
Deakin University
Melbourne, Australia
Email: thucn@deakin.edu.au

Lemai Nguyen
Deakin Business School
Deakin University
Melbourne, Australia
Email: lemai.nguyen@deakin.edu.au

Huy Quan Vu
Deakin Business School
Deakin University
Melbourne, Australia
Email: q.vu@deakin.edu.au

Jacob Cybulski
School of IT
Deakin University
Melbourne, Australia
Email: jacob.cybulski@deakin.edu.au

Abstract

Affective destination images have received considerable attention from tourism marketing researchers as evidence suggests that affective components in destination images affect tourists' emotional responses, which in turn influence their behavioural intentions toward the destination. Therefore, tourism practitioners seek solutions to influence the emotional effects of affective destination images for B2C communication. This paper presents a design science research project to develop an AI system to assist practitioners in generating affective destination images that potentially trigger the desired emotional responses of tourists. By leveraging knowledge and techniques from NeuroIS, this paper also proposes a framework of scientific experiments to assess how the generated affective destination images by the AI system affect tourists' emotional experiences.

Keywords Emotion AI, machine learning, design science research, NeuroIS, tourism destination image

1 Introduction

Travel agencies have been spending significantly on mass media advertising to attract tourists and maximise their market share (Huete Alcocer and López Ruiz 2020). Tourism marketing researchers demonstrated that content in an advertisement significantly influences the behavioural intentions and decision-making processes of tourists toward a destination (Walters et al. 2012). Researchers demonstrated that the affective component is a critical factor in destination decision (Tucker 2009). Bastiaansen et al. (2020) conclude that destination marketing in tourism aims to enhance the affective component of the destination image, which in turn affects the destination decision of tourists. Thus, an appropriate emotion measurement in tourism marketing research and a system capable of generating affective marketing materials to influence the emotional responses of tourists toward a destination are desirable.

The concept of emotional processes is a controversial topic that has been debated for millenniums. NeuroIS, which is a subfield of Information Systems research emphasising emotion-related phenomena research using tools and techniques from neuroscience, has a better opportunity to clarify the concept as it is a young and open-minded multi-discipline allowing new and innovative concepts to change old and rigid constructs in other established and conservative fields (vom Brocke et al. 2020). Based on the neuroscience theories and techniques, NeuroIS offers knowledge of the complex interplay between information system (IS) development, user's emotional responses, decision-making processes and behavioural intentions (Dimoka et al. 2011). NeuroIS researchers also study the impacts of images on emotions and behaviours, as well as provide a theoretical framework to understand and measure emotions in B2C communication (Gregor et al. 2014). However, while human emotion recognition and emotion analysis in communication media are well-studied in IS (Abbasi and Chen 2008), there is limited IS research into generating affective images to influence the consumer's emotions and behavioural intentions. Therefore, this study undertakes the challenge of developing a novel conceptual framework of an AI system to automate the generation of affective images that trigger desired emotional responses and behavioural intentions of consumers for B2C communication in the context of tourism destination marketing.

The study will make significant theoretical and practical contributions to IS. From the theoretical perspective, it proposes a novel conceptual framework of sentiment synthesis in visual communication by integrating AI techniques and NeuroIS concepts. As a co-product of the framework, an AI system will be developed to automate the generation of affective images to trigger desired emotional responses and behavioural intentions of consumers. From a practical perspective, the study proposes a design process including a sophisticated architecture and rigorous evaluation of artefacts associated with emotional content in B2C communication for tourism destination marketing.

2 Literature Review

2.1 Visual Sentiment Analysis

Visual sentiment analysis is a relatively new research area that aims to classify images based on the invoked emotion. Based on the correlation between the textual metadata of the image and the visual content, Siersdorfer et al. (2010) classified images into ‘positive’ and ‘negative’ sentiments. Visual features can be classified into three levels of semantics (Ortis et al. 2020), which are: (1) Low-level feature: colours, textures, brightness, contrast; (2) Mid-level features: the principle of art theories, presented object, aesthetic features, textual metadata; (3) High-level features: the content, presented facial expression, latent correlation among visual, textual and sentiment views.

In contrast to the earlier works based on image features, She et al. (2020) applied Convolutional Neural Networks (CNN) and transfer learning techniques for visual sentiment analysis using only images to determine which emotion they will evoke. Outcome of visual sentiment analysis can be presented in either a limited number of classification classes or values in a 2-dimensional space. For classification, researchers have considered binary classification that labels emotion as either positive or negative class (You et al. 2015). A multiclass classification that classifies emotion into one of three or more classes has also been considered. Xu et al. (2014) classified emotion into five polarities levels, while Peng et al. (2015) and You et al. (2016) classified emotion into seven and eight categories of emotion respectively. However, the classification method only allows emotion to be classified into a few emotional states. On the other hand, Kim et al. (2017) suggested representing emotion using the valence and arousal dimensional model. Valence represents the pleasure state while arousal represents the excitement state in the 2-dimensional space. The literature review demonstrates the lack of agreement on the standard of input (features) or output of visual sentiment analysis.

2.2 Visual Sentiment Synthesis

Visual sentiment synthesis refers to the generation of affective images, in other words, to generate images including emotional content that may trigger specific emotional responses from viewers. Through a search of the literature, only three studies examine visual sentiment synthesis that do not include humans as an object in images. Alvarez-Melis and Amores (2017) propose a model to generate emotional art photos using conditional GAN. Park and Lee (2021) propose a framework to generate emotional landscape images by modifying the GAN model to include an emotional residual unit (ERU) and optimise an affective feature matching loss. The authors use valence and arousal to represent emotion instead of emotional categories in the study by Alvarez-Melis and Amores (2017). Goetschalckx et al. (2019) demonstrate the capability of GAN in generating images with desired cognitive characteristics of images such as memorability, aesthetics and valence level of emotion. The proposed model allows users to increase or decrease their cognitive characteristics while generating images.

Another approach to evoke different emotions from images that can use the images supplied by the users is style transfer and colour transfer. Style transfer and colour transfer aim to transfer the style and colour of a style-reference image to the input image (Gatys et al. 2016). He et al. (2015) proposed a colour transfer framework based on colour combinations and a predefined colour emotion model. However, the predefined colour emotion model assigned colour weight using the Gaussian component, which might not correctly detect a significant colour to the human visual system. Kim et al. (2022) presented a photorealistic style transfer technique that allows the user to control which style to transfer to an input image using a reference image.

3 Research Methodology

This paper adopted a design science research methodology that focuses on creating and evaluating IT artifacts to solve related information issues in organisations, ensures research controls over the requisite technology development and participant feedback in the evaluation process (Hevner 2007). Figure 1 (Hevner, 2007, p. 88) presented the adopted design science research cycles of the paper.

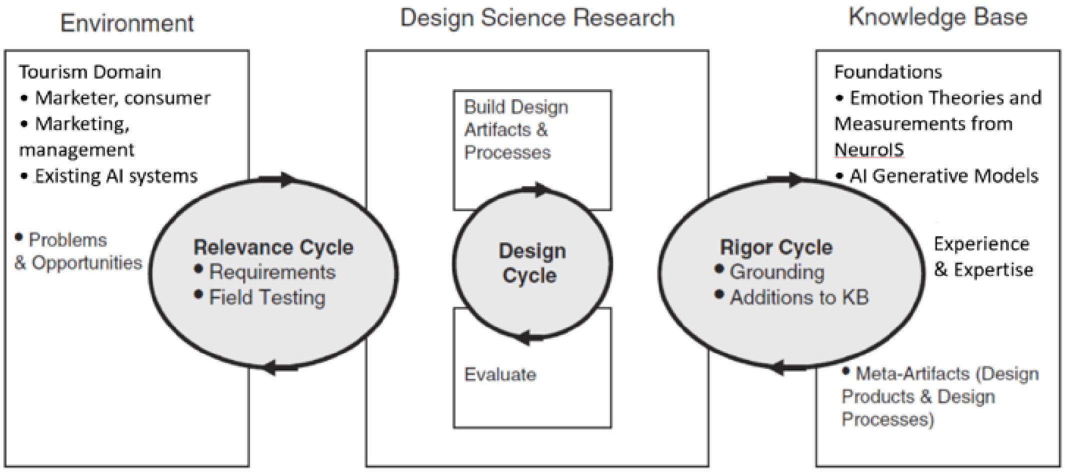

Figure 1 Design science research Cycles (Adapted from Hevner, 2007, p. 88)

3.1 Relevance Cycle

Explicating problem and defining requirements: During the initial phase of DSRM, the real-world problems from tourism marketing and requirements to develop an artefact to address the problem are defined. The problems experienced by tourism marketing researchers and practitioners have been identified in Section 2.1, in which tourism marketers have been seeking a system to support the development of affective destination images to increase advertising effectiveness. After addressing the problems, an artefact to solve the problem is identified, and requirements to develop the artefact are outlined. An AI system capable of generating affective destination images is identified as the artefact for this paper by comparing the results against related works (e.g., Kim et al. 2017). The functional requirements for the artefact, which are to detect emotion in images and to automatically generate affective tourism destination images, are elicited through research literature. Subsequently, the identified problems and the addressed requirements are verified by tourism marketing practitioners using informed arguments as proposed by Venable et al. (2012).

Evaluation: After the artefact is developed, field testing will be undertaken by tourism marketing practitioners. The main goal of field testing is to determine how the developed artefact solves the problem. In particular, how the designed artefact supports the practitioners in generating affective destination images. Results of field testing determine if an additional iteration of the cycle is required. A new iteration will start with the requirements derived from the field testing.

3.2 Rigour Cycle

Grounding theories and methods: To ensure rigorous results from DSRM, it is necessary to relate all the activities and cycles in DSRM to the existing knowledge base. Relevant knowledge including theoretical foundations and conceptual models from tourism marketing, NeuroIS, AI and Emotion AI research areas has been reviewed and applied to inform the relevance and design cycles to develop and evaluate an AI system capable of generating affective destination images for the development of tourism marketing material.

Contribution to knowledge: This paper provides theoretical and practical contributions to knowledge. Based on the design science research contribution types through activities of DSRM proposed by Gregor and Hevner (2013), this paper offers three levels of contributions: (1) Instantiation: AI system capable of visual sentiment synthesis to influence emotional response in tourism; (2) Design Theory: Design guideline of AI system capable of generating affective destination image; (3) Theoretical Construct: a conceptual framework to enhance B2C communication in tourism marketing by generating affective destination image to trigger specific emotional responses of tourists.

3.3 Design Cycle

3.3.1 AI System Development

Data collection: The study uses existing emotional image databases with emotion labels (i.e., emotion categories; valence and arousal values). Through preliminary experiments, most images in the collected data contain objects such as people, animals or vehicles. For developing an AI model to generate affective landscape images, we exclude images containing objects that significantly affect emotional responses. We aim to have an even distribution of emotion values in the final dataset.

Data labelling using visual sentiment analysis: The valence and arousal values of the collected image are obtained using the emotion predictor model trained by Kim et al. (2017).

Visual sentiment synthesis model development: The current study aims to explore different emotion theories while generating affective landscape images. Two deep learning models based on generative adversarial networks (GAN) (Goodfellow et al. 2014) will be developed. The first generative model follows the approach by Alvarez-Melis and Amores (2017) to generate affective landscape images labelled with categorical emotion states (e.g., happiness, sadness or anger). The second generative model adopted the techniques proposed by Park and Lee (2021), which embedded emotional residual unit (ERU) to improve the performance of the GAN model while training with affective images. As results from the generative models are in 128x128 pixel resolution due to computational power limitation, the Super-Resolution model (Ledig et al. 2017) is used for upscaling the images to 512x512 pixel resolution, which is more suitable for viewing on a large screen in the evaluation process. A style transfer model based on Kim et al. (2022) is also developed to transfer the style of a style-reference image to the input. Outputs from GAN models and style transfer will be compared to determine an appropriate approach for tourism marketing practitioners.

3.3.2 AI System Evaluation

Participants: For the evaluation in the design cycle, approximately 45 healthy participants aged over 18 and located in Melbourne, Australia will be recruited for the study. The number of participants is approximately equal to the number of participants in studies that involved physiological measurements of emotion in tourism (e.g. Li et al. 2018 - 38 participants).

Measures: The generated materials (referred to as stimuli in psychological research) will be evaluated in B2C communication, in which participants will view the generated stimuli while their physiological states are recorded to determine their emotional responses. The emotion measures approach follows the nomological networks proposed by Gregor et al. (2014). Objective emotional responses are captured using physiological measures, while subjective emotional responses are provided by participants through self-reports. The cognitive processes (e.g., visual perception, memory and decision-making) and behavioural intentions (e.g., attitudes toward stimuli and destination, and visit intention) are reported by participants after viewing the stimuli. The collected data such as subjective and objective emotional responses, self-reports of cognitive processes and behavioural intention are analysed (see Figure 2).

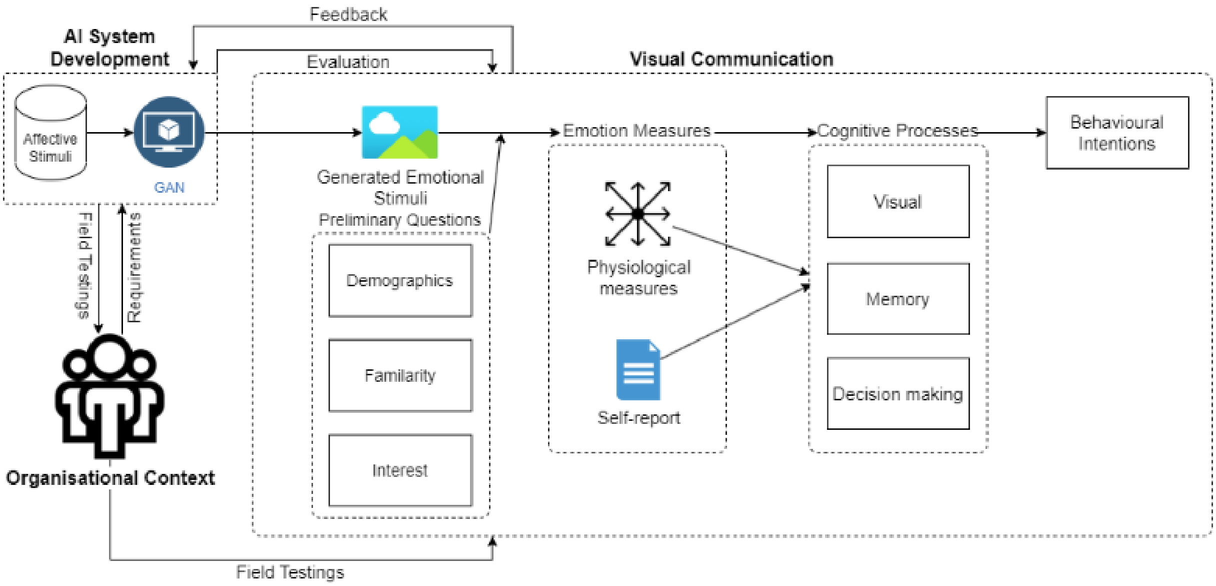

Figure 2 Design science research activities of this study

4 Current progress

We have identified the problem and motivation and defined requirements for this research. More than 20 image datasets with emotion labelled by emotion categories and dimensional emotion space have been collected. In the initial phase, we experimented with the GAN models to generate affective images. However, the generated affective destination image is not appropriate to use in the marketing materials (see images on the left in Figure 3) as tourists (or consumers in general) require a realistic image of a product or service that they are deciding to buy/use. We also experimented with visual sentiment analysis (or emotion detection) using only emotion category labels, yet, the method did not provide a required level of emotion granularity for analysis. Thus, we have moved on to another design cycle.

Figure 3 Sample results of GAN from Park and Lee (2021) vs. Style Transfer. Valence represents the pleasure state while arousal represents the excitement state in the 2-dimensional space

For the current design cycle, we are experimenting with style transfer and visual sentiment analysis using dimensional space as a preferred emotional model. First, a style transfer model was developed to transfer the style of a style-reference image to the input. In the right-hand side of Figure 3, the look and feel of the images from the second column (style) were transferred to the images from the first column (input) to generate the images in the last column (output). Next, an emotion detection model outputting the detected emotion from the image was implemented. The output from the model is emotion represented in 2-dimensional space (valence and arousal). The implemented emotion detection from this paper outperformed the emotion detection by Kim (2017) in terms of the Mean Absolute Error

(MAE) metric. Given the extracted emotional values from the input, style-reference and output images, we are analysing the result to determine which style image is appropriate for which input image and vice versa to generate the optimal image capable of evoking emotional responses of tourists.

5 Discussion and Next Steps

This paper aims to explore the capability of AI in generating affective destination images to influence tourists' emotional responses and scientifically evaluate the emotional responses using knowledge, tools and techniques from NeuroIS. We have experimented with different techniques of visual sentiment analysis and visual sentiment synthesis. The next steps include the evaluation of the designed artefact outlined in section 3.3.2. If the artifact satisfies the functional requirements, field testing will be undertaken by tourism marketing practitioners. The main goal of field testing is to determine how the developed artefact solves the problem identified in Section 3.1.

6 References

Abbasi, A., and Chen, H. 2008. "CyberGate: A Design Framework and System for Text Analysis of Computer-Mediated Communication," *MIS Quarterly* (32:4), Management Information Systems Research Center, University of Minnesota, pp. 811–837. (https://doi.org/10.2307/25148873).

Alvarez-Melis, D., and Amores, J. 2017. *The Emotional GAN: Priming Adversarial Generation of Art with Emotion*, p. 4.

Bastiaansen, M., Straatman, S., Mitas, O., Stekelenburg, J., and Jansen, S. 2020. "Emotion Measurement in Tourism Destination Marketing: A Comparative Electroencephalographic and Behavioral Study," *Journal of Travel Research*, p. 004728752098114. (https://doi.org/10.1177/0047287520981149).

vom Brocke, J., Hevner, A., Léger, P. M., Walla, P., and Riedl, R. 2020. "Advancing a NeuroIS Research Agenda with Four Areas of Societal Contributions," *European Journal of Information Systems* (29:1), Taylor & Francis, pp. 9–24. (https://doi.org/10.1080/0960085X.2019.1708218).

Dimoka, A., Pavlou, P. A., and Davis, F. D. 2011. "NeuroIS: The Potential of Cognitive Neuroscience for Information Systems Research," *Information Systems Research* (22:4), INFORMS, pp. 687–702.

Gatys, L. A., Ecker, A. S., and Bethge, M. 2016. "Image Style Transfer Using Convolutional Neural Networks," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June, pp. 2414–2423. (https://doi.org/10.1109/CVPR.2016.265).

Goetschalckx, L., Andonian, A., Oliva, A., and Isola, P. 2019. "GANalyze: Toward Visual Definitions of Cognitive Image Properties," *ArXiv:1906.10112 [Cs]*. (http://arxiv.org/abs/1906.10112).

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. 2014. "Generative Adversarial Networks," *ArXiv:1406.2661 [Cs, Stat]*. (http://arxiv.org/abs/1406.2661).

Gregor, S., and Hevner, A. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37), pp. 337–356. (https://doi.org/10.25300/MISQ/2013/37.2.01).

Gregor, S., Lin, A. C. H., Gedeon, T., Riaz, A., and Zhu, D. 2014. "Neuroscience and a Nomological Network for the Understanding and Assessment of Emotions in Information Systems Research," *Journal of Management Information Systems* (30:4), pp. 13–48. (https://doi.org/10.2753/MIS0742-1222300402).

He, L., Qi, H., and Zaretzki, R. 2015. "Image Color Transfer to Evoke Different Emotions Based on Color Combinations," *Signal, Image and Video Processing* (9:8), pp. 1965–1973. (https://doi.org/10.1007/s11760-014-0691-y).

Hevner, A. 2007. "A Three Cycle View of Design Science Research," *Scandinavian Journal of Information Systems* (19:2). (https://aisel.aisnet.org/sjis/vol19/iss2/4).

Huete Alcocer, N., and López Ruiz, V. R. 2020. "The Role of Destination Image in Tourist Satisfaction: The Case of a Heritage Site," *Economic Research-Ekonomska Istraživanja* (33:1), Routledge, pp. 2444–2461. (https://doi.org/10.1080/1331677X.2019.1654399).

Katsurai, M., and Satoh, S. 2016. "Image Sentiment Analysis Using Latent Correlations among Visual, Textual, and Sentiment Views," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March, pp. 2837–2841. (https://doi.org/10.1109/ICASSP.2016.7472195).

Kim, H.-R., Kim, Y.-S., Kim, S. J., and Lee, I.-K. 2017. *Building Emotional Machines: Recognizing Image Emotions through Deep Neural Networks*, arXiv. (https://doi.org/10.48550/ARXIV.1705.07543).

Kim, Sunwoo, Kim, Soohyun, and Kim, Seungryong. 2022. *Deep Translation Prior: Test-Time Training for Photorealistic Style Transfer*, arXiv. (https://doi.org/10.48550/arXiv.2112.06150).

Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W. 2017. "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," *ArXiv:1609.04802 [Cs, Stat]*. (http://arxiv.org/abs/1609.04802).

Li, S., Walters, G., Packer, J., and Scott, N. 2018. "Using Skin Conductance and Facial Electromyography to Measure Emotional Responses to Tourism Advertising," *Current Issues in Tourism* (21:15), Routledge, pp. 1761–1783. (https://doi.org/10.1080/13683500.2016.1223023).

Ortis, A., Farinella, G. M., and Battiato, S. 2020. "Survey on Visual Sentiment Analysis," *IET Image Processing* (14:8), pp. 1440–1456. (https://doi.org/10.1049/iet-ipr.2019.1270).

Park, C., and Lee, I.-K. 2021. "Emotional Landscape Image Generation Using Generative Adversarial Networks," in *Computer Vision – ACCV 2020* (Vol. 12625), Lecture Notes in Computer Science, H. Ishikawa, C.-L. Liu, T. Pajdla, and J. Shi (eds.), Cham: Springer International Publishing, pp. 573–590. (https://doi.org/10.1007/978-3-030-69538-5_35).

Peng, K.-C., Chen, T., Sadovnik, A., and Gallagher, A. 2015. "A Mixed Bag of Emotions: Model, Predict, and Transfer Emotion Distributions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June, pp. 860–868. (https://doi.org/10.1109/CVPR.2015.7298687).

She, D., Yang, J., Cheng, M., Lai, Y., Rosin, P. L., and Wang, L. 2020. "WSCNet: Weakly Supervised Coupled Networks for Visual Sentiment Classification and Detection," *IEEE Transactions on Multimedia* (22:5), pp. 1358–1371. (https://doi.org/10.1109/TMM.2019.2939744).

Siersdorfer, S., Minack, E., Deng, F., and Hare, J. 2010. "Analyzing and Predicting Sentiment of Images on the Social Web," in *Proceedings of the 18th ACM International Conference on Multimedia*, MM '10, New York, NY, USA: Association for Computing Machinery, October 25, pp. 715–718. (https://doi.org/10.1145/1873951.1874060).

Tucker, H. 2009. "Recognizing Emotion and Its Postcolonial Potentialities: Discomfort and Shame in a Tourism Encounter in Turkey," *Tourism Geographies* (11:4), Routledge, pp. 444–461. (https://doi.org/10.1080/14616680903262612).

Venable, J., Pries-Heje, J., and Baskerville, R. 2012. "A Comprehensive Framework for Evaluation in Design Science Research," in *Design Science Research in Information Systems. Advances in Theory and Practice*, Lecture Notes in Computer Science, K. Peffers, M. Rothenberger, and B. Kuechler (eds.), Berlin, Heidelberg: Springer, pp. 423–438. (https://doi.org/10.1007/978-3-642-29863-9_31).

Walters, G., Sparks, B., and Herington, C. 2012. "The Impact of Consumption Vision and Emotion on the Tourism Consumer's Decision Behavior," *Journal of Hospitality & Tourism Research* (36:3), SAGE Publications Inc, pp. 366–389. (https://doi.org/10.1177/1096348010390815).

Xu, C., Cetintas, S., Lee, K.-C., and Li, L.-J. 2014. "Visual Sentiment Prediction with Deep Convolutional Neural Networks," *ArXiv:1411.5731 [Cs, Stat]*. (http://arxiv.org/abs/1411.5731).

You, Q., Luo, J., Jin, H., and Yang, J. 2015. "Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks," *ArXiv:1509.06041 [Cs]*. (http://arxiv.org/abs/1509.06041).

You, Q., Luo, J., Jin, H., and Yang, J. 2016. "Building a Large Scale Dataset for Image Emotion Recognition: The Fine Print and The Benchmark," *ArXiv:1605.02677 [Cs]*. (http://arxiv.org/abs/1605.02677).

Copyright