Impact of Quality Of Images On Users Behavior On Social Media

Yazid Bounab
Faculty of ITEE, CMVS
University of Oulu, Oulu, Finland
yazid.bounab@oulu.fi

Mourad Oussalah
Faculty of ITEE, CMVS
University of Oulu, Oulu, Finland
mourad.oussalah@oulu.fi

Djamila Romaissa Beddiar
Faculty of ITEE, CMVS
University Of Oulu, Oulu, Finland
ad_beddiar@esi.dz

Abstract—The last decade has witnessed one of the fastest growth in Web 2.0 technology with the diversity of channels, structure and content of user generated content. Such data can be textual, visual or multimedia. This created valuable source of data for users, professionals and research community. This paper explores potential correlation that may exist between images and the associated textual content. Especially, we explore the influence of the image quality on the nature of the users' content. For this purpose, Brisque image quality index has been performed. The quantification between the image quality and the users' textual generated content is evaluated using both the number of comments generated by users and overall sentiment score associated to the given image. The results of such analysis confirms the hypothesis of positive correlation between the quality of posted image and the associated user's content in the above sense. A new tourism dataset Tourism48 [1] constructed from IMGUR web service was also introduced and made public.

Index Terms—social media, Image captioning, image description, sentiments.

I. INTRODUCTION

Internet has brought a lot of revolutionary technologies during the past two decades, which made people's life easier and more comfortable in terms of daily routine activities like shopping, spreading the news, communications, online learning, booking appointments and purchasing services. Indeed, the emergence of web 2.0 technology in the first decade of the twenty-first century fostered the development of social media and other interactive, crowd-based communication tools, this enables individuals and communities to create and share content that removed all kinds of space and time communication barriers. This fostered the development of online communities, in almost every aspect of our daily activities, empowering a cultural shift where social media is intertwined with a commercial logic that promoted online activities and creativity. For instance, the sharing of travel experiences and adventures on social media became ubiquitous, enabling random people to instantaneously follow, share, rate and discuss various types of content. Sites such as YouTube, Trip Advisory and Expedia freely provide users a forum to share and rate their travel experiences so that anyone who comments about a destination/product becomes part of the marketing process.

Nowadays, the use of visual user-posted content through video sharing sites such as YouTube is playing an important role in both shaping a destinations and in counteracting any negative perceptions, by connecting directly with the tourist community.

This empowered social media to have a tremendous impact on tourism industry and altered the landscape of marketing in the leisure and hospitality industry where consumers engage with social networking sites to research trips, share their experiences of a particular tourist attraction site, hotel, restaurant or airline. For instance, TripAdvisor had a wide reaching effect on customer-booking where many travelers determine their travel plans based on users' reviews and social media shares, which renders online customer service crucial for building a positive tourism product reputation. As a result, many agencies have shifted their focus from in-person to online experiences to adapt to the new technology and market trends. This fostered the development of the culture of urban exploration where urban explorers identify, visit and document popular, unknown and derelict man-made or nature-made structures [1]. This raises the question of the veracity and completeness of the information supplied by the users in their posts. On the other hand, there is an empirical evidence that suggests a strong link between the destination image (e.g., as highlighted in user's posts) and tourist behavior. For instance, [2] found that "consumer's attitude towards a product (and product purchase) is influenced by the matching of the product's user-image with the consumer's self-concept". This was tested and applied to evaluate tourist' attitude towards a destination as tourists have stereo-typic images of different destinations. Furthermore, [3] questions the importance of the image quality on the received user's feedbacks. This raises the importance of the image/video layout as posted by the users as it may have a strong influence and can dictate consumer's opinion and behavior.

This paper aims to shed the light on the impact of image layout / quality on the nature of user's social media dataset. For this purpose, we use IMGUR [4], an online image sharing community platform that hosts millions of commented images/galleries in several topics and subjects and focusing on tourism related products. The database contains both selected image galleries from 48 countries and the associated users' comments for each image. The quality of the images varies greatly across galleries, which opens up the door to test the

---

following hypothesis

Hypothesis P: *Media posters are impacted by the layout of the visual post.*

The second section of this paper highlights the collected dataset. The third section details the implemented BRISQUE image quality metric while the fourth section emphasizes the data analysis and correlation analysis in order to validate or refute the aforementioned hypothesis. Finally, conclusion and perspective work are provided in Section 6.

II. DATA COLLECTION

The data collection task is guided by two main principles: the quality of the images/videos and the availability of relevant users' comments on each image/video. The first requirement is needed to enable the subsequent image processing task performed by the image captioning software as bad quality image will lead to either void outcome or non-satisfactory result. The second principle is motivated by the study design plan, which relies on user's generated content in order to validate or refute the hypotheses highlighted in previous section. The first attempt in collecting the images/videos with users' comments is performed using Youtube social media platform. Nevertheless, the rate-limit of the Youtube API prevented us from collecting a high quality dataset in terms of user's generated content (title, comments/ replies, annotations...). Therefore, we used **IMGUR** web service, which is an online image sharing community repository founded by Alan Schaaf in 2009 [4]. The availability of an API that allows us to automate the search operation and monitor the quality of the user's generated content played a key-role in our choice. We named the collected dataset "Tourism48". It contains data associated to 48 different countries. Each country has 10 galleries. Each gallery is associated with an album of pictures. In our case, we restricted to only one picture per gallery because we wanted to have the comments/replies associated to one picture only. Otherwise, if the album contains more than one picture, an extra processing step and data association reasoning would be required to find out to which picture a given comment/reply is associated to. Besides, a manual checking is performed in order to monitor the quality of the user's generated content at each gallery so that pictures that contain large proportion of unrelated comments are discarded. Using the IMGUR API, both the image and all the associated user's comments/replies are extracted and saved in separate files **gallery_id.jpg** that represent the image and **Comments.json** for textual comments.

In the second phase of the data collection, we want to explore the content of each image in terms of key patterns it contains (e.g., sea, mountain, street, car, table, face,...) to comprehend the visual content of the gallery. For this purpose, we used an automatic image captioning system; namely, Google Cloud Vision [5] because of its proven efficiency elsewhere. Strictly speaking, Google Cloud Vision API enables developers to build a metadata on image catalogue and, therefore, generate textual description of the image content. It can easily classify images into thousands of predefined categories and detect individual objects, faces, among other labels contained in the images. For this purpose, we feed each of the downloaded image to Google cloud vision service to get what we call the visual characters of that image (annotations/labels) and save them into a third file **google.json**. Google Cloud Vision (GCV) for image annotation has different output entities (labels, faces, web, locations, Landmarks, properties, etc...). In our study, we were only interested to Label and Web-entities because they encompass most of the textual description of objects identified in the image/video. The output of the GCV is organized in a dictionary structure where the keys are labels and the values are confidence scores of the recognized objects. The preprocessing step is carried out on the content of the dictionary to remove uncommon characters, URLs and mentions often occurring in user's generated content. Besides, we organized the collected database so that the image with its associated labels (from Google Cloud Vision) and user's textual comments can be accessed and queried separately. Finally, images, user's comments and GCV labels were merged to form a single enlarged dictionary that will be used during the matching phase. Figure 1 shows a real sample of one gallery in Tourism48 dataset.
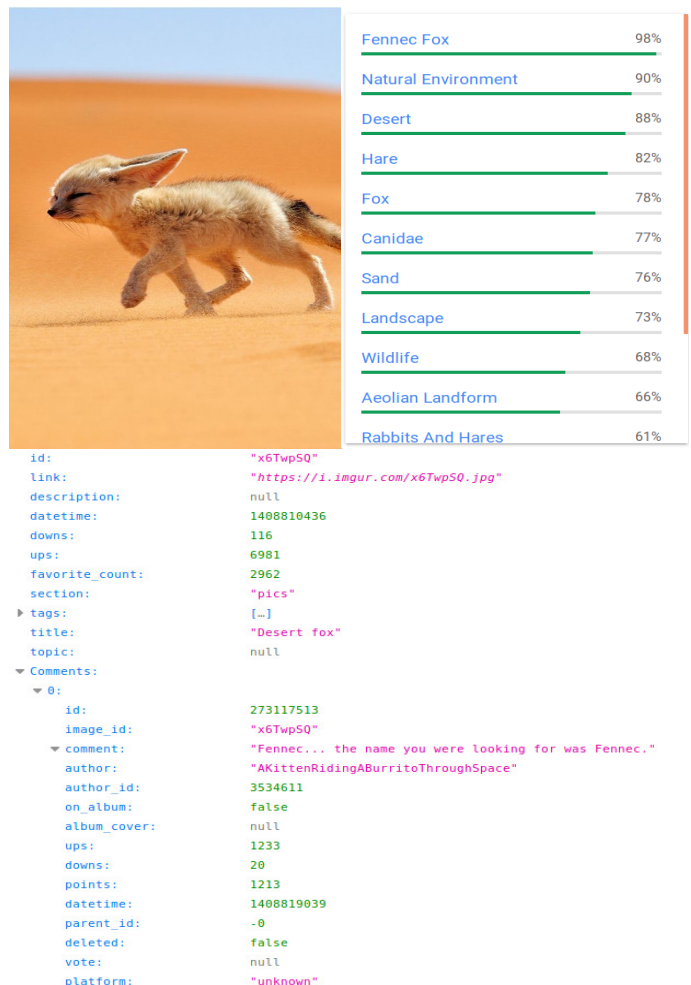

Fig. 1. Tourism48 dataset gallery sample

Left-hand side of the figure 1 shows an image of desert fox taken in desert where the right-hand side shows the associated labels outputted by the GCV. Particularly, it is easy to acknowledge the ability of the GCV to identify the key descriptors of the image content (Fox, Natural Environment, Desert, Sand). The figure also highlights the various attributes of the users' generated content. One may notice, for instance, the file includes a title attribute containing "Desert", "Fox", which fully match with GCV's output. Nevertheless, the attributes related to Topic or Tag were void in this example. User's comment in this example is also relevant as it points to type of the Fox "Fennec", which overlaps with GCV's output as well. A statistical summary of the collected dataset is highlighted in TableI. The latter summarizes the statistics associated with both the attributes of the images and the user's generated content in terms of textual data.

TABLE I
TOURISM48 DATASET SUMMARY

Dataset Summary				
Dataset Size	254.9 MB			
Number of images	480 image			
Image/gallery	01 image			
Image		Min	Max	Avg	
	Height	199	5113	1171.32	
	Width	232	7666	1378.3	
	Size KB	7.9	3789.52	432.38	
	Total	Avg	Std	Skew	Kurt
Comments Number	37360	77.83	76.29	2.43	9.41
Comments (Words)	355429	9.27	6.75	6.75	0.06
Emoticons	11.0	0.0	0.01	58.25	3391.36
Links/Urls	2027	0.06	0.23	4.31	19.83
Mentions	213	0.01	0.07	13.67	193.74
Symbols	427908	11.18	8.13	0.97	0.84

The table above highlights the variety of the gathered images in terms of resolution and size as well as the variation of user's generated content in terms of size, number of comments associated to the image, presence of special characters (emoticons, links/URLs, Mentions, Symbols).

III. IMAGE QUALITY

With the advance of a wide range of handheld devices which can capture, compress, process, store and share a variety of audiovisual content, a tremendous amount of visual data is making its way to consumers. For that, it is important to afford many resources and efforts to ensure that the end user could benefit from good quality images or videos. It is to note, image quality affects the user's behavior. According to Cornell [6], poor pictures may negatively impact on the user's experience, website conversion ratio, the time one stays on the website, and trust/credibility of the web content. Image quality assessment (IQA) aims to quantitatively represent the human perception of quality in a scalar or vector representation [7]. In other words, the main goal of IQA is to automatically predict perceived image quality. Methods of IQA are mainly split into two distinct categories: (1) Reference-based evaluation and (2) No-reference based evaluation.

A. Reference-based Image Quality Assessment

Reference-based evaluation requires a high-quality image to be considered as a reference. The differences between this reference and the distorted images are evaluated and the quality scores are therefore calculated. These algorithms are used to measure the ability of image processing techniques such as image compression or transmission in compromising the image quality. We can quote among reference-based evaluators the Structural Similarity Index (SSIM) [8].

B. Referenceless Image Quality Assessment

In many real-world scenarios, where image quality computation is undertaken, reference information is generally unavailable. In this case, another category of IQA is used. No-reference IQA, called also Blind or Referenceless IQA does not require a reference image to evaluate another image quality. To predict the image quality, the only information needed is the distorted image whose quality is being assessed.

[9] introduced a Blind/referenceless Image Spatial Quality Evaluator BRISQUE. It is purely spatial, does not require a mapping to a different co-ordinate domain (unlike many other blind methods based on the image modeling in wavelet or DCT domains) and relies only on image pixels to calculate features. BRISQUE quantifies possible losses of "naturalness" in the image based on the hypothesis that MSCN (mean subtracted contrast normalized) coefficients have characteristic statistical properties that are changed by the presence of distortion. Quantifying these changes may help to predict the type of distortion affecting an image.

Given an image $I(i,j)$, first compute locally normalized luminescence that refers to MSCN. It is computed as shown in Equation 1 where $i \in 1,2...M$, $j \in 1,2...N$ and $C = 1$ or $= 1/255$ if $I(i, j)$ domain is [0, 255] or [0, 1] respectively. M and N are the image height and width respectively.

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + C} \quad (1)$$

The local mean could be obtained using Equation 2 and in practice it is calculated by just applying a Gaussian filter to the image. w is a Gaussian kernel of size (K, L). Similarly, the local deviation is calculated using Equation 3.

$$\mu(i,j) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} I_{k,l}(i,j) \quad (2)$$

$$\sigma(i,j) = \sqrt{\sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l}(I_{k,l}(i,j) - \mu(i,j))} \quad (3)$$

The MSCN coefficients are proven to be distributed as a Generalized Gaussian Distribution (GGD) for a broader spectrum of the distorted image. Equation 4 illustrates the GGD density function where α controls the shape of the distribution and σ^2 controls the variance.

$$f(x; \alpha, \sigma^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp(-\left(\frac{|x|}{\beta}\right)^{\alpha}) \quad (4)$$

$$\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}} \qquad (5)$$

and $\Gamma(.)$ is the gamma function for $a > 0$

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \qquad (6)$$

Since the signs of adjacent coefficients exhibit a regular structure that can be disturbed due to a distortion, it is also important to model the statistical relationships between neighboring pixels. For that, [9] model this structure using the empirical distributions of pairwise products of neighboring MSCN coefficients along four orientations: horizontal H, vertical V, main-diagonal D1 and secondary-diagonal D2 as follows.

$$H(i,j) = \hat{I}(i,j)\hat{I}(i,j+1) \qquad (7)$$

$$V(i,j) = \hat{I}(i,j)\hat{I}(i+1,j) \qquad (8)$$

$$D1(i,j) = \hat{I}(i,j)\hat{I}(i+1,j+1) \qquad (9)$$

$$D1(i,j) = \hat{I}(i,j)\hat{I}(i+1,j-1) \qquad (10)$$

Also, the GGD found to be not good enough to fit to the empirical histograms of coefficient products. Thus, instead of fitting these coefficients to GGD, they are fit to an Asymmetric Generalized Gaussian Distribution (AGGD) model whose density function is given by the Equation 11. v controls the shape of the distribution, while σ_l^2 and σ_r^2 control the spread on each side. *side* could be either r or l and the mean η of AGGD is given in Equation 13.

$$f(x; v, \sigma_l^2, \sigma_r^2) = \begin{cases} \frac{v}{(\beta_l + \beta_r)\Gamma(\frac{1}{v})} e^{(-(\frac{-x}{\beta_l})^v)} & x < 0 \\ \frac{v}{(\beta_l + \beta_r)\Gamma(\frac{1}{v})} e^{(-(\frac{x}{\beta_r})^v)} & x >= 0 \end{cases} \qquad (11)$$

Where

$$\beta_{side} = \sigma_{side} \sqrt{\frac{\Gamma(\frac{1}{v})}{\Gamma(\frac{3}{v})}} \qquad (12)$$

$$\eta = (\beta_r - \beta_l) \frac{\Gamma(\frac{2}{v})}{\Gamma(\frac{1}{v})} \qquad (13)$$

Finally, shape and variance features are extracted by fitting GGD to MSCN coefficients and shape, while, mean, left variance and right variance are extracted for each pairwise products fitting to the AGGD, to identify distortions. This yields a holistic measure of quality, which is summarized in Algorithm 1 where the BRISQUE coefficient calculus for a given image is detailed.

Algorithm 1 Calculate BRISQUE Coefficient

1: $GivenImage : I(i,j)$
2: $C : constant$
3: $W : GaussianFilter$
4:
5: $//Calculate local mean :$
6: $\mu(i,j) \leftarrow W(I(i,j))$
7:
8: $//Calculate\ local\ deviation\ :$
9:
10: $\sigma(i,j) \leftarrow \sqrt{W(I(i,j) - \mu(i,j))}$
11: $MSCN\ Coefficients \leftarrow (\mu, \sigma)$
12:
13: $//Calculate\ Pairwise\ products\ Empirical\ Distributions \hat{I}$

14: $H(i,j) \leftarrow \hat{I}(i,j)\hat{I}(i,j+1)$
15: $V(i,j) \leftarrow \hat{I}(i,j)\hat{I}(i+1,j)$
16: $D1(i,j) \leftarrow \hat{I}(i,j)\hat{I}(i+1,j+1)$
17: $D2(i,j) \leftarrow \hat{I}(i,j)\hat{I}(i+1,j-1)$
18:
19: $//Apply\ GGD\ to\ MSCN\ Coefficients$
20: $GGD\ (MSCN\ Coefficients)$
21:
22: $/*Apply\ AGGD\ to\ Empirical\ histograms$
23: $of\ coefficient\ products * /$
24: $AGGD\ (Coef\_pair\_products) =0$

IV. ANALYSIS

To test the hypothesis of whether users are impacted by the layout of a given image. We check whether the quality of that image has an impact on either the user tendency of engagement through textual posts associated with that image or positive sentiment as gathered by the users' posts. This follows the intuition that a good quality image would attract more interest from the user's community, which is then materialized by a higher number of posts associated to the image or positive sentiment, in this purpose, a process for image quality assessment is required. this study employed the following three variables:

- V_1: Number of comments per image.
- V_2: Average sentiment score per image.
- V_3: BRISQUE quality index value per image.

The quantification of the first two variables is made through a combination of simple natural language processing tasks and statistics. In this respect, V_1 turns out to be a simple count of number of comments associated to a given image as recorded in Tourism48 database. V_2 relies on the application of sentiment analysis on the user's textual comments and recording the average sentiment score over the sentiment scores of individual posts pertaining to the same image. For this purpose, we utilize SentiStrength package [10]. The latter provides both positive and negative sentiment scores as integer values in the range [4,-4] interval. An arithmetic of the positive and negative sentiment provides the average sentiment for the

given utterance (post). Next, individual sentiment score per posts are average to yield a global sentiment measure for the whole underlined image, while V_3 is directly outputted by the BRISQUE quality index module for each inputted image. Strictly speaking, BRISQUE uses scene statistics of locally normalized luminance coefficients to quantify possible losses of "naturalness" in the image due to the presence of distortions, yielding a holistic measure of quality score. A smaller BRISQUE score indicates a better perceptual quality. BRISQUE scores can then be used to create a binary partition of a (relatively) good quality images.

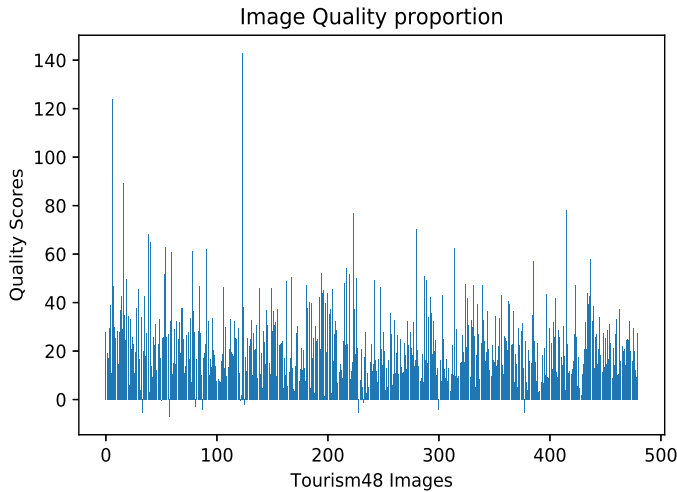Figure 2 shows the proportion of quality of image in Tourism48 dataset using BRISQUE Algorithm.

Fig. 2. Quality of Image using BRISQUE on Tourism48 dataset

A quick examination of the figure indicates the need for further statistical analysis in order to validate or refute the hypothesis P.

Table I shows the overall statistics of the three variables: Number of comments per image, average positive sentiment score per image and BRISQUE quality score per image in terms of mean value, standard deviation, Minimum and Maximum values, Kurtosis and Skewness.

Now in order to test the validity of Hypothesis P, a statistical test is carried out. For this purpose, we use BRISQUE score to generate a partition according to the mean value of BRISQUE scores leading to:

P_1: Set of images I such that BRISQUE(I) > E0
P_2: Set of images I such that BRISQUE(I) < E0
where E0 stands for the mean value of the BRISQUE image score.

Since the population P_1 and P_2 have different size, we use the non-parametric statistical test Wilcoxon rank-sum test [11], [12] to evaluate the null hypothesis that the two independent samples of Number of comments (resp. positive sentiments) in P_1 and P_2 comes from distribution with equal medians. The alternative hypothesis is that the two samples do not have same median distribution, and therefore, validate that one sample has a significantly larger median than the other. The results of this statistical test are shown in Table II and Table III.

Wilcoxon rank-sum test for Number of Comments samples in P_1 and P_2		
	Population P_1	Population P_2
Mean	70.80	92.45
Standard deviation	13.85	19.89
Observations	204	276
Hypothesized equal median	h=1	
P-value	0.03	
Rank Sum statistics value	203.78	
Z-score	-1.23	

The reading of Wilcoxon rank-sum test results indicates clearly that Hypothesis can only be validated when using the number of comments each image received from the user (H-value = 1 and p-value less than 5 percent), while the sentiment score test cannot refute the hypothesis of equal median populations at 5 percent significance level. This can also be explained by the high variability and instability characterizing the sentiment of the posts because of the lack of clear statements and expressions that would support positiveness due to inherent structure of human language where the neutrality is a dominant behaviour. On the other hand, this also can be partly attributed to inherent limitations of the SentiStrength package employed to capture the sentiment score of utterances.

V. CONCLUSION

In the digital era, social media has tremendously impacted tourism industry and altered the landscape of marketing in the leisure and hospitality industry where consumers constantly engage with social networking sites to share their experiences, thoughts and, subsequently, find interesting opportunities raised by the user community instead of professionals, which, ultimately fosters the development of the culture of urban exploration. Therefore, comprehending users' behaviours using user's generated content plays a key-role for tourism industry stakeholders. This paper contributes to this overreaching goal where a new tourism database; namely, Tourism48, containing galleries of attraction from 48 different countries together with their associated user comments, has been constructed using IMGUR web service. Next, in order to test the hypothesis how the image quality and layout impacts users' generated content, an original analysis has been carried out. This involves the

Wilcoxon rank-sum test for Sentiment scores of samples in P_1 and P_2		
	Population P_1	Population P_2
Mean	0.125	0.107
Standard deviation	0.214	0.187
Observations	204	276
Hypothesized equal median	h=0	
P-value	0.4	
Rank Sum statistics value	50317	
Z-score	0.83	

use of BRISQUE index for automatically quantifying image quality. This is then correlated to two variables generated from textual inputs: Number of users' comments and Average sentiment score. The analysis carried out using Wilcoxon rank-sum statistical test indicates the tenacity of the hypothesis when considering the Number of comments attribute. In other words, the higher the quality of the posted images, the higher the number of posts assigned to the image. However, the sentiment score does not provide enough evidence to support or refute the underlined hypothesis. The results highlighted in this paper will pave the way for efficient reuse of the growing user - generated content in tourism related sources in order to strengthen the tourism industry and benefit from the increasing urban exploration culture.

VI. ACKNOWLEDGMENT

REFERENCES

[1] S. Lash, "Urry, j.(1994) economies of signs and space," *Theory, Culture & Society. Sage, London*, 1994.

[2] M. J. Sirgy and C. Su, "The ethics of consumer sovereignty in an age of high tech," *Journal of Business Ethics*, vol. 28, no. 1, pp. 1–14, 2000.

[3] P. Turner, S. Turner, and F. Carroll, "The tourist gaze: Towards contextualised virtual environments," in *Spaces, spatiality and technology*, pp. 281–297, Springer, 2005.

[4] A. Schaaf, "Imgur," 2009.

[5] Google, "Web service, cloud computing," 2008.

[6] X. Min, G. Zhai, K. Gu, Y. Liu, and X. Yang, "Blind image quality estimation via distortion aggravation," *IEEE Transactions on Broadcasting*, vol. 64, no. 2, pp. 508–517, 2018.

[7] H. Maître, *From photon to pixel: the digital camera handbook*. John Wiley & Sons, 2017.

[8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[9] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on image processing*, vol. 21, no. 12, pp. 4695–4708, 2012.

[10] A. Rabab'Ah, M. Al-Ayyoub, Y. Jararweh, and M. Al-Kabi, "Evaluating sentistrength for arabic sentiment analysis"," in *Proc. 2016 7th International Conference on Computer Science and Information Technology (CSIT*, p. 1–6.

[11] B. Rosner, R. J. Glynn, and M.-L. Ting Lee, "Incorporation of clustering effects for the wilcoxon rank sum test: a large-sample approach," *Biometrics*, vol. 59, no. 4, pp. 1089–1098, 2003.

[12] J. Perolat, I. Couso, K. Loquin, and O. Strauss, "Generalizing the wilcoxon rank-sum test for interval data," *International Journal of Approximate Reasoning*, vol. 56, pp. 108–121, 2015.