Instagram hashtags as image annotation metadata

Stamatios Giannoulakis and Nicolas Tsapatsoulis

Dept. of Communication and Internet Studies
Cyprus University of Technology
30, Arch. Kyprianos str., CY-3036, Limassol, Cyprus
{s.giannoulakis,nicolas.tsapatsoulis}@cut.ac.cy

Abstract. Image tagging is an essential step for developing automatic image annotation methods that are based on the learning by example paradigm. However, manual image annotation, even for creating training sets for machine learning algorithms, requires hard effort and contains human judgment errors and subjectivity. Thus, alternative ways for automatically creating training examples, i.e., pairs of images and tags, are pursued. In this work we investigate whether tags accompanying photos in social media and especially the Instagram hashtags, provide a form of image annotation. If such a claim is proved then Instagram could be a very rich source of training data, easily collectable automatically, for the development of automatic image annotation techniques. Our hypothesis is that Instagram hashtags, and especially those provided by the photo owner / creator, express more accurately the content of a photo compared to the tags assigned to a photo during explicit image annotation processes like crowdsourcing. In this context, we explore the descriptive power of hashtags by examining whether other users would use the same, with the owner, hashtags to annotate an image. For this purpose a set of 30 randomly chosen, from Instagram, images were used as a dataset for our research. Then, one to four hashtags, considered as the most descriptive ones for the image in question, were selected among the hashtags used by the image owner. Three online questionnaires with ten images each were distributed to experiment participants in order to choose the best suitable hashtag for every image according to their interpretation. Results show that an average of 55% of the participants hashtag choices coincide with those suggested by the photo owners; thus, an initial evidence towards our hypothesis confirmation can be claimed.

Key words: Instagram, hashtags, image tagging, image retrieval, machine learning

1 Introduction

On average 350 million photos are uploaded to Facebook per day [9] while an average of 70 million photos are shared every day in Instagram [10]. Thus, the total amount of images in the Web nowadays is huge and continuously increasing. Locating and retrieving these images is very challenging not only in terms of effectiveness (retrieve the right image according to the user needs / queries) and efficiency (execution time) but also in terms of visibility (being locatable). Contemporary search engines retrieve images in a text-based manner since the majority of end users are familiar with text-based queries for retrieving web pages and digital documents. In text-based image retrieval images must be somehow related with specific keywords or textual description. This kind of textual description is, usually, obtained from the web page, or the document, containing the corresponding images and includes HTML alternative text, the file names of the images, captions, metadata tags and surrounding text [15]. However, images in social media, which constitute the great majority of images in the Web, cannot effectively indexed (extract relevant text description) with pure web-based techniques mainly because the user pages in social media do not follow the classic web-page structure. As a result a new research field dealing with this problem emerged: Automatic Image Annotation (AIA) [6], that is the process of extracting low-level features from an image and assigning one or more semantic concepts to it [11].

A large category of AIA involves machine learning techniques and is based on the learning by example paradigm [19]. Training examples that are used for automatic image annotation are pairs of images and related tags. Different models and machine learning techniques were developed to build the so called 'visual models', that is, models that capture the correlation between image features and textual words from the training examples. Visual models are then fed with image features extracted from unseen images to predict their tagging [16]. Assuming that good visual models can be achieved, image retrieval using the training by example paradigm provides a promising alternative to text-based methods (since it does not require explicit annotation of all images in the collection, but only a small set of properly annotated images) [17]. Nevertheless, the first important step to create effective visual models is to use good training examples (pairs of images and annotations). In this context automatic creation of training examples via crawling is highly desirable because it addresses the scalability (models for new concepts) and adaptability (modification of training models) issues.

According to a survey of Pew Research Internet Project, Instagram is the fastest growing social network among the adult Americans who use the Internet [9]. Instagram is a free application for mobile devices, which offer user the possibility to upload, edit and share with other pictures and very short video. The term Instagram is a combination of two words, from the word instant used to market old cameras and the gram comes from telegram from the snapshots people were taking[1]. Instagram launched on 6 October 2010 and rapidly gained popularity, managed to have 300 million active users per month and 70 million pictures per day [10] being shared. In January 2011 Instagram added hashtags [1] and from 27 April 2015 users are able to use emoji as hashtags[2]. Hashtags are tags or words prepended with '#' used to indicate the content of the picture, allowing users to search for pictures and increase visibility. Emoji are pictograms,

[1] Instagram: FAQ, https://instagram.com/about/faq/#
[2] Instagram: Our Story, https://instagram.com/press/

which are connected with emotions a user wants to give to the picture. The story of hashtags began with IRC (Internet Relay Chat) where the users used it in order to categorize items into groups. The first who used hashtags in Twitter was designer Chris Messina who asked from his followers how they felt about using the pound sign to group conversations [2]. Thus, we can conclude that the role of hashtags was traditionally to organize knowledge and facilitate access and enable retrieval of information. On the other hand, we know that users extend the function of hashtagging beyond findability and give hashtags a metacommunicative use. The metacommunicative function can be grouped into the following codes: Emphasizing, Iterating, Critiquing, Identifying, and Rallying. Emphasizing is used to give emphasis or call attention, Critiquing in order to express judgment or verdict, Identifying to refer to the author of the post, Iterating to express humor and Rallying to bring awareness or support to a cause [4]. In a research on the tags of a set of 2700 pictures it was measured that approximately 10% of these photos were related with emotion words not directly related with their visual [3].

It should be evident from the above discussion that Instagram provides a rich forum for automatically creating training sets for AIA. It contains a huge amount of images which are commented through hashtags by their creators / owners. Thus, if we assume that it is the owner who can better expresses the real visual content or meaning of an image then choosing among the hashtags for assigning tags to images is much more safe than traditional text-based indexing approaches. This is extremely important in training sets where pairs of images and tags have to be carefully selected because they affect the effectiveness of tag predicting models. However, Intsagram hashtags are used not only to describe the visual content of an image but also serve other functions falling under the metacommunicative use. In this work we are trying to check the extent to which hashtags are indeed related with the actual content of an image and the percentage of hashtags that are relevant to Instagram photo compared to those referring to metacommunicative use.

2 Related work

To the best of our knowledge this is the first work that examines the appropriateness of Instagram photo-hashtag pairs for creating training sets for AIA. However, this problem falls under a more broad category dealing with the quality of manual image annotation. Nowak & Ruger [14] investigated the reliability of image annotation via crowdsourcing. They tried first to explore to which extent several sets of expert annotations differ from each other and then to investigate whether non-expert annotations are reliable. The researchers selected a subset of 99 images from the MIR Flickr Image Dataset. The aforementioned set of images was annotated by 11 expert annotators from the Fraunhofer IDMT research staff using 53 concepts. The same set of images was distributed over the online marketplace Amazon Mechanical Turk in order get non-expert annotations. The consistency among expert annotators proved to be very high. The same also proved between the expert and non-expert groups. Thus, the conclusion was that crowdsourced annotation is as accurate as experts' annotation. Since the cost of crowdsourcing annotation is far more cheeper and efficient than experts' annotation this conclusion opened up new ways towards AIA. The importance of crowdsourcing annotation lead to several research efforts which further examine the quality of crowdsourced data. In crowdsourcing annotation the participants expose different behavior during the annotation task. There are many reasons for the aforementioned behavior including the level of expertise, low-attention / low-concentration when they perform the task and there is always the bad intent of the annotators. Examples of annotators with bad intention can be spammers, dishonest or try to manipulate the system by answering in an unrelated or nonsense way [12]. In a research about crowdsourcing annotators' consistency Theodosiou *et. al.* [17] used both vocabulary keywords and free keywords to check whether guided annotation (as assumed by the use of structured vocabulary) would increase annotation consistency. The researchers concluded that, indeed, by combing free keywords and vocabulary keywords annotation consistency increases compared to the use of free keywords alone.

Even the crowdsourced data are far more difficult to collect than data crawled from the Web. This was recognized by Tsikrika *et. al* [20] who examined the quality of clickthrough data for training concept detectors in images. They showed that clickthrough data, if properly filtered, would be used for AIA. The problem with clickthrough data is that they express the interpretation of end users rather than this of the image creators / owners, and, thus, they are highly subjective. Despite that, the use of clickthrough data for developing AIA models is an attractive approach and Microsoft Research announced, for a second year in a row, a challenge based on data provided by Bing search engine[3].

In order to examine the image retrieval from social media and especially the diversification of image retrieval results, Ionescu *et. al* [8] compared experts and crowdsourcing annotation. The results showed that in the crowdsourcing annotation the inter-rater agreement were a slightly lower than expert annotators. Our assumption is that the owners' annotation data (in our case Instagram hashtags) are more close to experts' annotation compared to that of crowdsourcing since the latter expresses the end-users' perspective. Furthermore, web-crawled data are far more easier to collect than crowdsourcing ones. Among the web-crawled data the ones collected from Instagram are much more accurate (in terms of descriptive value) compared to those used in traditional web-document indexing (keyword extraction from web-pages) while they are richer than those collected via clickthroughs.

3 Methodology

In order to derive concrete results, in our study we followed a hybrid methodology combing a set up from social science research with a strict mathematical

[3] http://research.microsoft.com/en-us/projects/irc/

framework which is common in natural sciences. We decided to define clear research questions and properly select the participants of the experiment rather than randomly choosing among ordinary users of social media. We consider that in order to assess the descriptive value of Instagram hashtags of the photo owners / creators we need users that are familiar both with the social media and the use of metadata in digital content. Librarians would be ideal for this purpose. They use social networks daily and one of their main tasks is to organize knowledge and annotate electronic resources, so we can say they are, in some respect, experts in image annotation. Moreover, undergraduate and postgraduate university students are also good candidates for the population group because social media are highly popular among students as we can conclude from the survey of Pew Research Internet Project [5].

3.1 Aim of the research

The present study has two goals. The first is to investigate whether Instagram hashtags accompanying images can be used as image tags so as to create image-tag pairs for training machine learning approaches for AIA. The second is to provide a rough estimation on the percentage of Instagram hashtags that describe the visual content of accompanying images. Towards this end we have designed a questionnaire which includes a total number of 30 randomly selected pictures from 18 Instagram users. Owners' hashtags surrounding these images were automatically crawled using the Beautiful Soup[4] library of Python. Then we chose, manually, 1 to 4 hashtags for each picture, which, according to our interpretation, better describe its visual content. As already stated not all hashtags are intended to describe image on increasing its findability. The selected hashtags along with irrelevant hashtags are presented below each picture, in the questionnaire, and the participants are asked to choose among them the ones that better describe the shown photo. Fig. 1 presents an example. Among the eight choices given only three are hashtags the owner of the photo used in the Instagram. If participants' choices coincide with the hashtags the owner gave we have a good indication that these hashtags are, indeed, related with the visual content of the picture (since what the participants see is the context-free picture without any sort of metadata).

3.2 Data collection

The data for this survey were gathered using an online questionnaire designed with the aid of Surveymonkey[5]. The 30 pictures were divided in three separate questionnaires containing 10 pictures each. The aim was to reduce the time required to fill in the questionnaire and avoid fatigue effects. The choices given below each picture are either 4 or 8 depending on the number of hashtags the owner used. If only one hashtag of the owner was present then the choices given

[4] http://www.crummy.com/software/BeautifulSoup/bs4/doc/
[5] http://www.surveymonkey.com/

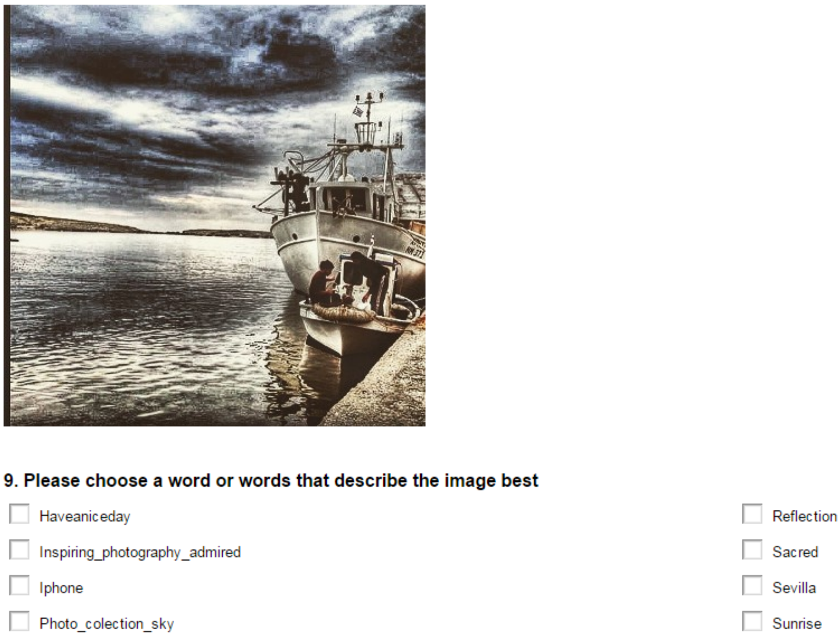

Fig. 1. An example of an image interpretation multiple choice question

to the participants were four (including the hashtag of the owner); otherwise the participants were given eight options to select from. This rule was applied in order to keep a minimum chance level higher than or equal to 25%. In any case, participants were not aware that any of the given choices were related in any respect with the picture; thus, they were free to select as many of them as they wished according to their interpretation of the shown photo.

Initially, the three online questionnaires were distributed by electronic mail to three experts in order to evaluate them. The results of the evaluation assisted the creation of a more appropriate version which was, then, distributed by electronic mail to librarians of the library of Cyprus University of Technology and to undergraduate and postgraduate students of the department of Communication & Internet Studies. It is important to mention, here, that each participant group, librarians, undergraduate and postgraduates students was randomly divided so as to distribute the three questionnaires equally in each group. The survey was conducted between March, 19th and March 31st, 2015. A total of 39 questionnaires were collected and used for analysis.

3.3 Mathematical Formulation

Let us denote by P^i the *i-th* participant (i=*1,...,*N_P) of a total of N_P participants ($N_P = 39$ in this study as already mentioned above). We also denote with I^j the *j-th* image (j=*1,...,*N_I) in the image dataset where N_I is the total number of images (in our case $N_I = 30$). By set $H = \{h_1, h_2, ..., h_{N_H}\}$ we define the set of hashtags the owners / creators used to tag the images in set I while N_H is the total number of tags used for this purpose.

In order be able to conclude on the research questions defined earlier we must use some effectiveness measures. For this purpose we modified the well known Recall, Precision and F-measures [7] to fit in with the current experiment. In particular we define the participant's P^i recall value, R_{ij}, for image I^j as the proportion of owner's hashtags, for this image, that were selected by P^i in the questionnaire. In a mathematically formal way this is given by:

$$R_{ij} = \frac{||\mathbf{T}_{jc} \cap \mathbf{T}_{ji}||}{||\mathbf{T}_{jc}||} \tag{1}$$

where $\mathbf{T}_{jc}$ is the set of distinct hashtags assigned to image I^j by the image owner, $\mathbf{T}_{ji}$ is the set of distinct hashtags the participant P^i assigned to image I^j (based on the choices presented to him/her in the questionnaire), $\cap$ is the set intersection operation and $||\Omega||$ denotes the cardinality of set Ω.

Aggregating the R_{ij} across all participants and all images we define per image and per participant recall values, R_j and R_i, respectively:

$$R_j = \frac{1}{N_P} \sum_{i=1}^{N_P} R_{ij} \tag{2}$$

$$R_i = \frac{1}{N_I} \sum_{j=1}^{N_I} R_{ij} \tag{3}$$

In a similar manner we define per image (see eq. 5) and per participant precision (see eq. 6), i.e., the proportion of a participant's choices that coincide with owner's hashtags, and F-measure (harmonic mean of recall and precision) as follows:

$$P_{ij} = \frac{||\mathbf{T}_{jc} \cap \mathbf{T}_{ji}||}{||\mathbf{T}_{ji}||} \tag{4}$$

(precision of participant's P^i choices for j-th image)

$$P_j = \frac{1}{N_P} \sum_{i=1}^{N_P} P_{ij} \tag{5}$$

$$P_i = \frac{1}{N_I} \sum_{j=1}^{N_I} P_{ij} \tag{6}$$

$$F_j = \frac{2 \cdot R_j \cdot P_j}{R_j + P_j} \tag{7}$$

$$F_i = \frac{2 \cdot R_i \cdot P_i}{R_i + P_i} \tag{8}$$

Let us now assume an index of hashtags $\boldsymbol{V}$ in which all the hashtag choices presented to the participants though the questionnaire images are concatenated. That is, if in the questionnaire the participants are asked to choose between 8 hashtags in the first image then these hashtags are the first 8 entries of vector $\boldsymbol{V}$. The available hashtag choices for the second image of the questionnaire will follow, then that of the third image and so on. Note that in index V the same hashtag may appear more than once and in different position indicating a particular choice for a specific image. If we denote with '1' the hashtags chosen by a specific participant and with '0' the hashtags not chosen then a participant P^ican be represented by a binary vector $\boldsymbol{P^i}$, with length equal to that of index $\boldsymbol{V}$, denoting his / her 'profile'. In a similar way we can define the creators / owners vector, say $\boldsymbol{C}$ in which the hashtags used by the photo owners are represented with ones and hashtags not used by zeros. Obviously, the vector $\boldsymbol{C}$ does not correspond to a specific user profile but to the aggregated profile of all photo owners. The similarity of images' interpretation between photo owners / creators and each one of the participants can be, then, estimated by any vector comparison metric. Because both vectors $\boldsymbol{C}$ and $\boldsymbol{P^i}$ are binary ones the choose of Hamming distance [13] is evident. The aforementioned distance was introduced by Richard Hamming, is implied only at two equal strings and gives the number of positions at which corresponding symbols differ [13].

Thus, the similarity $S(C, P^i)$ between the choices a participant P^i made in order to characterize the images in the questionnaire with the actual hashtags the owners used, is given by:

$$S(C, P^i) = 1 - \frac{h(\boldsymbol{C}, \boldsymbol{P^i})}{L} \tag{9}$$

where $h(\boldsymbol{C}, \boldsymbol{P^i})$ is the Hamming distance of vectors $\boldsymbol{C}$ and $\boldsymbol{P^i}$ and L is the corresponding vector space dimension (i.e., the length of vectors $\boldsymbol{C}$ and $\boldsymbol{P^i}$ and index $\boldsymbol{V}$).

Table 1. Per participant Recall, Precision and F-measure value statistics

	Mean	St. Dev.	Minimum	Maximum
Recall	0.55	0.15	0.23	0.78
Precision	0.67	0.12	0.47	1.00
F-measure	0.56	0.12	0.33	0.77

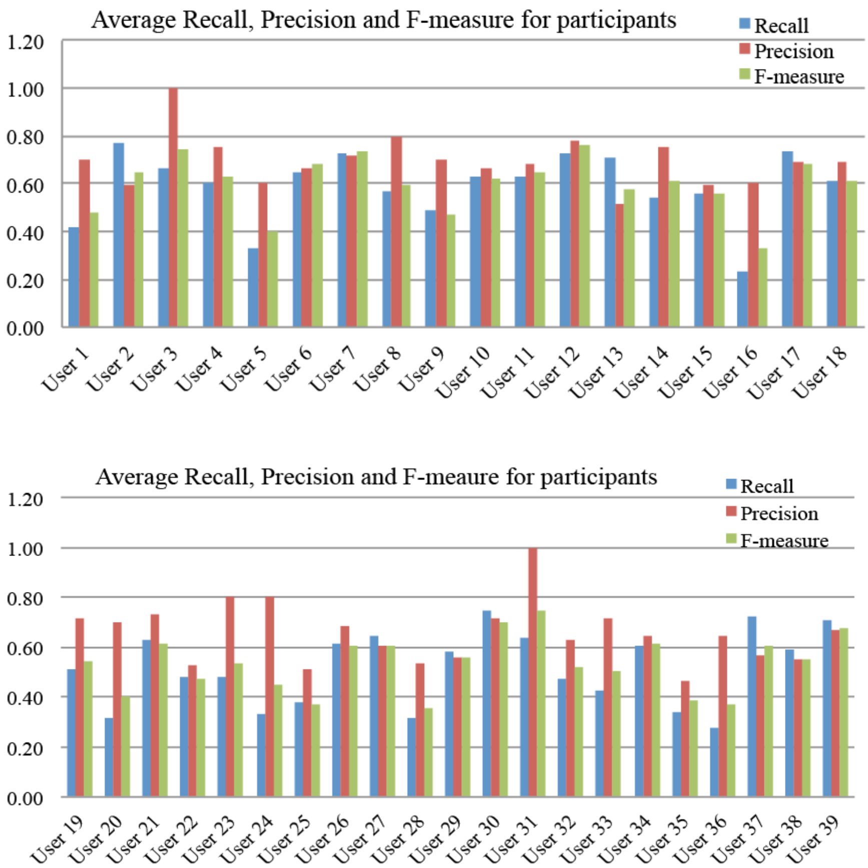

Fig. 2. Average hashtags' recall, precision and F-measure per participant

4 Experimental Results and Discussion

The data of the 39 filled in questionnaires were analyzed with aid of SPSS[6], MS Excel[7] and the MATLAB[8] platform using the metrics defined in the previous section. Fig.2 shows the per participants' Recall (eq. 3), Precision (eq. 6) and F-measure (eq. 8) of the 10 pictures each participant had to interpret in the experiment. As already explained not all participants evaluated all images; thus, the computations were done using the subsets of images shown to the participants according to the questionnaire they war given. In practice, this means that N_I in equations 3, 6, 8 was equal to ten (the number of images in each questionnaire).

Some basic statistics of the per participant Recall, Precision and F-measure are shown in Table 1. We see there that the recall performance per participant is 0.55 ± 0.15 with the extreme values being 0.23 (minimum) and 0.78 (maximum). Thus, the conclusion is that at least one out two hashtags used by the owner in Instagram images is relevant to image content since other users consider it descriptive as well. The variation in performance, among users, is rather low indicating that in the experiment there were no spammers or users with dishonest behavior. The per participant precision is significantly higher (0.67 ± 0.12) than recall, showing the tendency of people to use as few as possible keywords to describe an image. This is in agreement with the generic behavior of Web users who use, on average, one to three keywords when searching for information through search engines. Of course we do not know whether this is an intrinsic human tendency or a behavior cultivated by the way search engines work (the fewer the keywords given the more the results presented to the user). Furthermore, the high precision values indicate also that the participants did not answer (chose hashtags for the shown images) randomly.

Overall, with the aid of Fig.2 and Table 1 we can conclude on both research questions set in this study. Given that the participants in our experiments can be seen as experts (librarians and students of internet and communication studies) we can claim that around 55% of the Instagram hashtags that accompany images relevant to the actual content of the images and can be used for training purposes. By pointing out that on average only 40% of the (owner's) Instagram image hashtags are relevant to the images close to which they appear we can state that on average 22% ($0.55 \cdot 0.4$) of Instagram hashtags are related with the visual content of images.

Figure 3 shows the dissimilarity of image interpretation between each one of the participants and the photo owners with the aid of (normalized) Hamming distance and the mathematical formulation presented in the previous section. By normalized we mean that the Hamming distance is divided by the length of the strings compared (in our case total number of choices presented to the users in the 10 questions of the questionnaire they took). As we see in Table 2 the average

[6] http://www-01.ibm.com/software/analytics/spss/
[7] https://products.office.com/en-us/excel
[8] http://www.mathworks.com/products/matlab/

Table 2. Statistics of normalized Hamming distance between participants and photo owners in image interpretation

Mean	St. Dev.	Minimum	Maximum
0.245	0.048	0.138	0.375

normalized Hamming distance between the photo owners and the participants is 0.245 ± 0.048. This means that there is less than 25% disagreement (only one out of four hashtag choices between image owners and participants differ); thus, we can confirm, once again, that the participants do not answer at random or in any dishonest manner. By looking at the extreme values in Fig. 3 we see that only two users (those with ids 22 and 25) show somehow low performance (high dissimilarity with the interpretation of picture owners) but even in these cases no random or dishonest behavior can be justified since the average Hamming distance is well below 40%. On the other hand, the users with ids 3 and 31 present an excellent performance which indicates that even perfect matching between owners and participants is not impossible; this means that the hashtags given by the owners to the photos are indeed related with the visual content of images (i.e., what the images actually show and not, for instance, context or emotional information).

Table 3. Per image Recall, Precision and F-measure value statistics

	Mean	St. Dev.	Minimum	Maximum
Recall	0.55	0.18	0.15	0.91
Precision	0.67	0.22	0.22	1.00
F-measure	0.56	0.17	0.17	0.90

In Fig. 4 we present the per image Recall (eq.2), Precision (eq.5), and F-measure (eq.7) values while in Table 3 are shown summary statistics for those values. The basic aim of this analysis is to check whether the difficulty of interpreting images depends on their visual content. ComparingTables 1 and 3 we observe that the variation of Recall, Precision, and F-measure across images is higher than that across participants. The same also holds for the extreme values. Thus, we can conclude that image content affects interpretability. On the other hand, in Fig. 5 we show the images with the lowest recall and precision scores (from left to right images with ids 2, 20 and 28). In a first glance it does not seem that these images present abstract concepts, which are, generally, difficult to interpret. Thus, probably the owners hashtags for these image might be irrelevant with their visual content causing a different interpretation by the experiment participants.

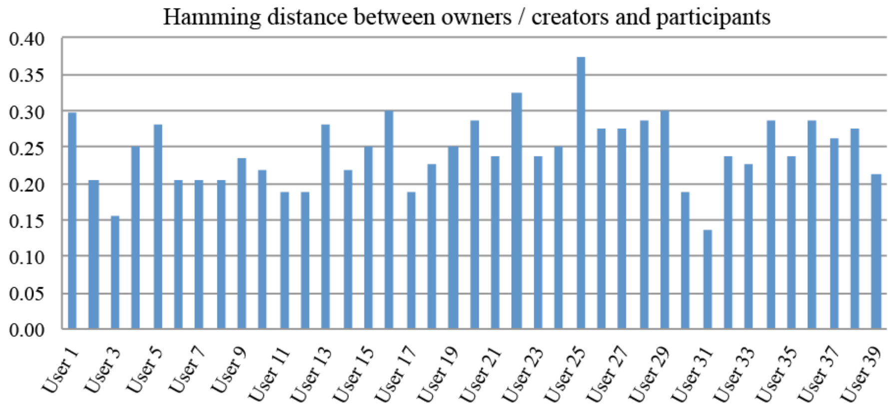

Fig. 3. The hamming distance between participants and image owners / creators

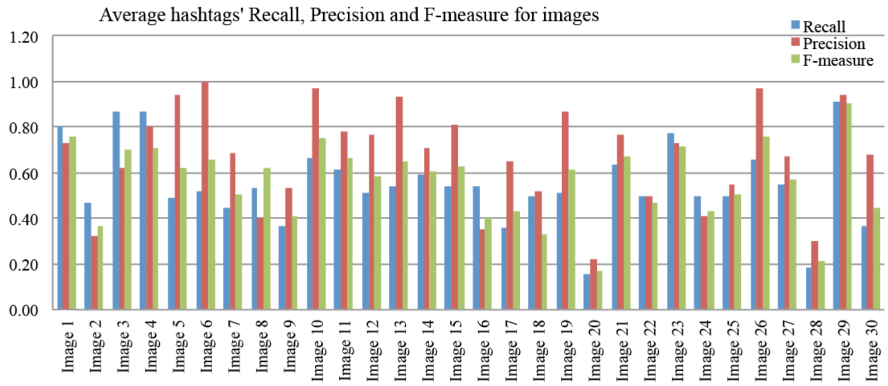

Fig. 4. Average hashtags' recall, precision and F-measure per image

Fig. 5. Difficult to interpret images

In the last part of our analysis we deal with the recall values of the hashtags. Our assumption is that abstract concepts should have lower recall values than concepts referring to tangible objects. Figure 6 presents the recall values for all owners' hashtags (the set H mentioned in Section 3.3). It is clear that abstract concepts tend to have low recall values, as expected, however, the three out of four concepts that have zero recall refer to places (Florida, Chile, Indochina). This lead us to the conclusion that out of context interpretation of images is, in some cases, problematic. Nevertheless, the difficulty of interpretation in this case does not necessarily mean that the hashtag used by the owner is inappropriate for characterizing the particular image. By saying so we mean that the pair image-hashtag is still a good training example.

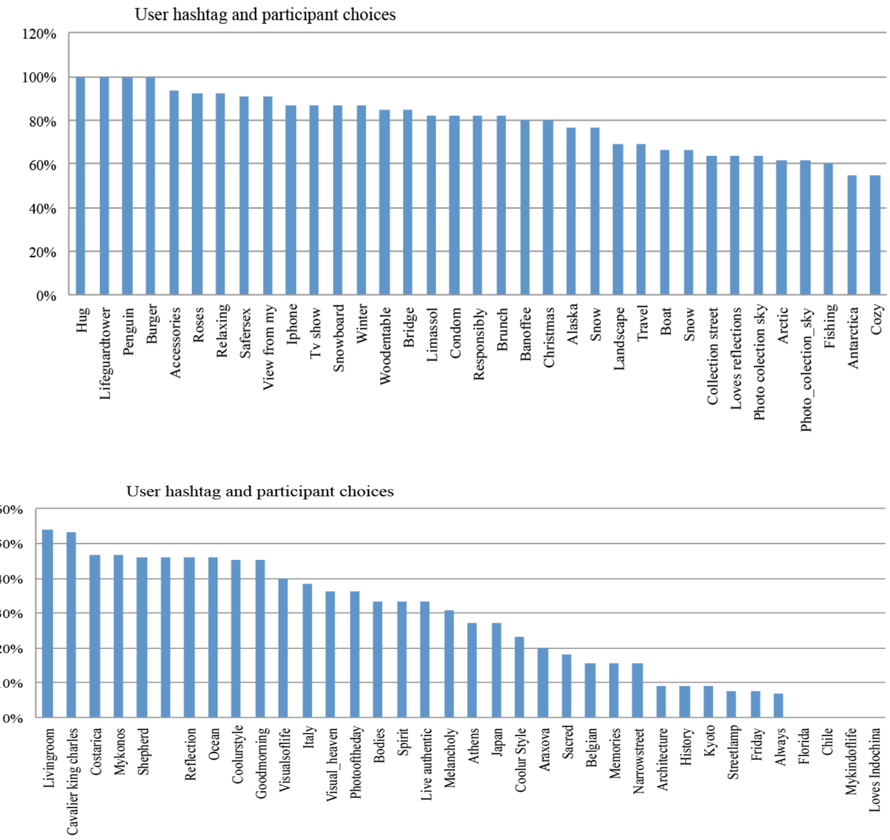

Fig. 6. Percentage of the participants that chose each of the owner / creator hashtags

5 Conclusion

In this paper we have presented our study about the descriptive value of Instagram hashtags as metadata for the images they accompany. By measuring if the participants would choose the same hashtags with the image creator / owner we found that in the 55% of the chosen hashtags participants and owners agree that the suggested hashtags can describe the visual content of an image. Moreover, we have an indication that almost 30% of the Instagram hashtag datasets are appropriate for use in training examples (image - tag pairs) for machine learning algorithms. The results show also that an important portion of image hashtags in Instagram are not directly related with the concept depicted by the image. We have found also that both the image content and the context in which an image resides affect interpretability. However, as we explained, this does not necessarily implies that the pairs images - difficult to interpret tags are invalid for training purposes.

Further research, including larger participants' and images' population, has to be done before the indications extracted in the current study become solid conclusions. For this purpose we have already started an automated process for dynamically creating online image interpretation questions with the aid of a MySQL database schema and PHP programming. This will allow larger participation in the experiment and automatic analysis of the results.

Another action that can be taken in the future is to check the validity of image-hashtag pairs for training visual concept models (see [19]) in practice. This will lead us to a second, practical, stage of investigation and will allow comparison of the theoretical findings of this study with practical issues faced during training. However, we must be aware that, in machine learning, good training examples must be properly processed to extract appropriate, for learning, low level features; this is by no means an easy task.

References

1. Baranovic, M.: What #hashtags mean to mobile photography. Online: `http://connect.dpreview.com/post/1256293279/hastag-photography` (2013).
2. Bennett, S.: The history of hashtags in social media marketing. Online: `http://www.adweek.com/socialtimes/history-hashtag-social-marketing/501237` (2014).
3. Carmean, D. M., Morris, M. E.: Selfie examinations?: applying computer vision, hashtag scraping and sentiment analysis to finding and interpreting selfies. Online: `http://nebula.wsimg.com/27bab6eda0e75b69fcab8a5cdc4e22af?AccessKeyId=A6A4DAF733A0F616E396&disposition=0`.
4. Daer, A.R., Hoffman, R., Goodman, S.: Rhetorical Functions of Hashtag Forms Across Social Media Applications. Communication Design Quarterly Review, 3(1), pp. 12-16 (2014).
5. Duggan, M., Ellison, N.B., Lampe, C., Lenhart, A., Madden, M.: Social Media Update 2014. Online: `http://www.pewinternet.org/2015/01/09/social-media-update-2014/`.
6. Hanbury, A.: A survey of methods for image annotation. Journal of Visual Languages & Computing, 19 (5), pp. 617-627 (2008).

7. Hersh, W.: Information Retrieval: A Health and Biomedical Perspective. Springer-Verlag New York (2009).
8. Ionescu, B., Popescu, A., Radu, A.-L., Mller, H.: Result diversification in social image retrieval: a benchmarking framework. Multimedia Tools and Applications, Nov. 2014, DOI: 10.1007/s11042-014-2369-4 (2014).
9. Internet.org: A Focus on Efficiency. A white paper from Facebook, Ericsson and Qualcomm. Online: `https://fbcdn-dragon-a.akamaihd.net/hphotos-ak-prn1/851575_520797877991079_393255490_n.pdf` (2013)
10. Intsagram: Introducing Layout from Instagram. Online: `http://blog.instagram.com/post/114416360957/layout-from-instagram` (2015)
11. Jin, C., Jin, S.-W.: Automatic image annotation using feature selection based on improving quantum particle swarm optimization. Signal Processing, 109, pp. 172-181 (2015).
12. Kara, Y.E., Genc, G., Aran, O., Akarun, L.: Modeling annotator behaviors for crowd labeling. Neurocomputing, 160, pp. 141-156 (2015).
13. Kulshrestha, A.: On the Hamming Distance Between Base-N Representations of Whole Numbers. Canadian Young Scientist Journal, 2012, pp. 14-17 (2012).
14. Nowak, S., Ruger, S.: How reliable are annotations via crowdsourcing? A study about inter-annotator agreement for multi-label image annotation. In: Proceedings of the International conference on Multimedia Information Retrieval, pp. 557-566 (2010).
15. Ntalianis, K., Tsapatsoulis, N., Doulamis, A., Matsatsinis, N.: Automatic annotation of image databases based on implicit crowdsourcing, visual concept modeling and evolution. Multimedia Tools and Applications, 69(2), pp. 397-421 (2014).
16. Snoek, C.G., Worring, M.: Concept-based video retrieval. Foundations and Trends in Information Retrieval 2(4), pp. 215-322 (2009)
17. Theodosiou, Z., Tsapatsoulis, N.: Crowdsourcing annotation: Modelling keywords using low level features. In: Proceedings of the 5th International Conference on Internet Multimedia Systems Architecture and Application (IEEE IMSAA 11), pp. 14. Bangalore, India (2011)
18. Theodosiou, Z., Georgiou, O., Tsapatsoulis, N.: Evaluating Annotators Consistency with the Aid of an Innovative Database Schema. In: Proceedings of the 6th International Workshop on Semantic Media Adaptation and Personalization (SMAP 2011), pp. 74-78, Luxenbourg, December 2011 (2011).
19. Theodosiou, Z., Tsapatsoulis, N.: Image retrieval using keywords: The machine learning perspective. In: E. Spyrou, D. Iakovides, P. Mylonas (eds.) Semantic Multimedia Analysis and Processing. CRC Press / Taylor & Francis (2014).
20. Tsikrika, T., Diou, C., de Vries, A. P., Delopoulos, A.: Reliability and effectiveness of clickthrough data for automatic image annotation. Multimedia Tools and Applications, 55(1), pp. 27-52 (2011).