**DISCUSSION AND CONCLUSION**

In this paper, we have presented an innovative methodology, based on the HITS algorithm and the principles of collective intelligence, for the identification of Instagram hashtags that describe the visual content of the images they are associated with. We have empirically shown that the application of a two-step HITS algorithm in a crowdtagging context provides an easy and effective way to locate pairs of Instagram images and hashtags that can be used as training sets for content-based image retrieval systems in the learning by example paradigm. As a proof of concept, we have used 25 000 evaluations (500 annotations for each one of 50 images) collected from the Figure-eight crowdsourcing platform to create a bipartite graph composed of users (annotators) and the tags they selected to describe the 50 images. The hub scores of the HITS

algorithm applied to this graph, called hereby full bipartite graph, give us a measure of the reliability of the annotators. The aforementioned approach is based on the findings of Theodosiou et al. [39], in which the reliability of annotators is better approximated if we consider all the annotations they have performed rather than the subset of gold test questions. In the second step, a weighted bipartite graph for each image is composed in the same way as the full bipartite graph. The weights of these graphs are the hub scores computed in the previous step. By thresholding the authority scores of the per image graphs, obtained by the application of the HITS algorithm on the weighted graphs, we can rank and then effectively locate the hashtags that are relevant to their visual content as per the annotators evaluation.

Some important findings of this paper are briefly summarized here. The first refers to the value of crowdtagging itself. In several studies before, we found that the crowd can substitute the experts in the evaluation of images with respect to relevant tags. However, even with a large number of annotators (499 in our case), it seems that a perfect agreement between annotators and experts cannot be achieved. In particular, it was found that from the 145 different tags suggested for the 50 images used in this paper by the two experts, only 135 were also identified by the 499 annotators. This leads to a maximum achievable recall value equal to 0.931. Thus, in subjective evaluation tasks, such as those referring to the identification of tags that are related to the visual content of images, no perfect agreement between the experts and the crowd should be expected

A second finding is that crowdtagging of images can be effectively modeled through user–tag bipartite graphs, one per image. Thresholding the authority score of the HITS algorithm applied on these graphs is a robust way to identify the tags that characterize the visual content of the corresponding images. Getting the top ranked tags based on the authority score is an alternative solution, but, with a little bit lower effectiveness.

A final remark of this paper refers to the importance of using weighted user–tag bipartite graphs for the crowdtagged images. It appears that weighting the bipartite graphs with the hub scores of the annotators provides the best results. However, even in the case that the reliability metric of the crowdsourcing platform itself (the \_trust variable of Figure-eight in our case) is used to weight the bipartite graphs, the results are not significantly worse. We are a little bit reluctant to generalize this conclusion because, in this paper, we have used too many annotations (499) per image. Thus, one of our future tests will involve a more typical image crowdtagging scenario in which much more images will be used and much fewer (typically less than five) annotations per image will be considered. In that case, only partial coannotation of the same images by the same annotators will take place in contrast to this paper in which all annotators annotated all images.

We are currently working to check, in practice, that the image–hashtags pairs mined from the Instagram through the approach described in this paper can be used, indeed, for a large-scale AIA in a content-based image retrieval scenario as proposed by Theodosiou and Tsapatsoulis