Image User Profiling with Knowledge Graph and Computer Vision

Vincent Lully[1,2(✉)], Philippe Laublet[1], Milan Stankovic[1,2],
and Filip Radulovic[2]

[1] Sorbonne Université, 28 rue Serpente, 75006 Paris, France
{vincent.lully,philippe.laublet}@sorbonne-universite.fr
[2] Sépage, 38 avenue de l'Opéra, 75002 Paris, France
{milstan,filip}@sepage.fr

Abstract. In this paper, we explore the synergy between knowledge graph technologies and computer vision tools for image user profiling. We propose two image user profiling approaches which map an image to knowledge graph entities representing the interests of a user who appreciates the image. The first one maps an image to entities which correspond to the objects appearing in the image. The second one maps to entities which are depicted by visually similar images and which exist in the conceptual scope of the dataset within which further personalisation tasks are conducted. A demo configured with a real and recent commercial travel domain dataset is given at ESWC 2018.

Keywords: Image · User profiling · Knowledge graph · Computer vision Personalisation · Travel · Recommender system

1 Introduction

Recent research efforts show several interesting convergence points between knowledge graph and computer vision, such as improving object detection with external knowledge graphs [1], scene description with triples [2], knowledge graph completion with visual features [3] and visuo-semantic search [4]. In this paper, we explore their synergy for image user profiling which has not been sufficiently studied so far.

Since several years, knowledge graphs have been leveraged to conduct user profiling through semantic analysis of text and to improve content-based recommendation approaches by providing structured metadata [5–10]. Today, a tremendous amount of multimedia data are available on the web and are being produced continuously. Modern websites should be equipped with systems which can understand users' interests through their interactions with multimedia data and adapt the services accordingly in order to provide a better user experience.

Our main contribution is two novel image user profiling approaches:

- The first one maps an image to entities which correspond to the objects appearing in the image.
- The second one maps to entities which are depicted by visually similar images and which exist in the conceptual scope of the dataset within which further personalisation tasks are conducted.

In the rest of the paper, we discuss some related work on image user profiling in Sect. 2, we then present our two image user profiling approaches in Sect. 3, Sect. 4 describes the demonstration given at ESWC 2018 and Sect. 5 concludes the paper.
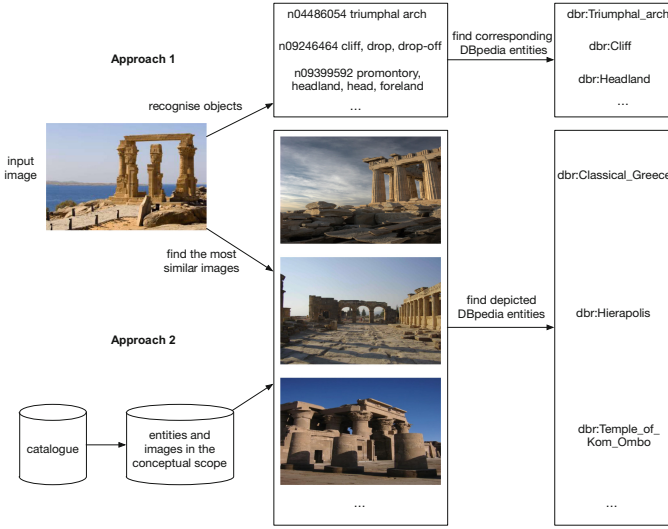
2 Related Work

We present some existing approaches which create user profiles from images. In [11], the authors try to detect demographic attributes of individual users and group types from the photos posted on photo sharing sites. In [12], the authors derive users' personalities from pictures posted on Instagram. In [13], the authors introduce a picture-based user elicitation and recommendation method for tourism products. The system creates a user profile which consists of 7 traveller types accompanied with a matching degree. A very similar tool is presented in [14] which maps photos to 17 tourist types. In [15], the photos are mapped to several pre-defined categories such as "leisure", "art" and "culture". Different from these existing approaches, the approaches that we propose map images to knowledge graph entities. This choice has been motivated by existing work which has proven the advantages of such semantic user profiling per se and in personalisation systems [5–10].

3 Two Novel Image User Profiling Approaches

In this section, we present two image user profiling approaches which map an input image to knowledge graph entities. A user profile contains top-n entities representing things of interest to a user who appreciates the input image. In this paper, we use DBpedia, knowing that other similar large-scale knowledge graphs like Wikidata can also be used.

The first approach consists of mapping an image to entities which correspond to the objects appearing in the image. There are two main steps: object detection and entity linking. For object detection, we use a computer vision tool named "Inception-V3" [16]. Inception-V3 is a convolutional neural network model trained for the ImageNet Large Visual Recognition Challenge using the data from 2012. The model tries to classify entire images into 1000 classes which are WordNet synsets like "gazelle" and "patio, terrace". At the entity linking step, we map the 1000 synsets to corresponding DBpedia entities. We are completely aware that this is a very basic and obvious approach. We still present it because we did not find it in the state of the art.

Fig. 1. Example illustrating the two proposed image user profiling approaches

The second approach consists of mapping an image to entities which are depicted by visually similar images and which exist in the conceptual scope of the catalogue within which further personalisation tasks are conducted. The conceptual scope is a new notion that we propose in our work. We assume that the user profiling is not an end in itself but should serve further personalisation tasks. The created user profiles should be useful for these further tasks. Given a catalogue of items, we currently consider that its conceptual scope consists of all knowledge graph entities which directly appear in the catalogue. The entities can be obtained by two means: direct item linking and item description linking as presented in [17]. To compute the visual similarity between images, we rely on the penultimate layer outputted by Inception-V3 which is a 2048-dimensional vector. The similarity between two images is determined by the Euclidean distance between their vectors.
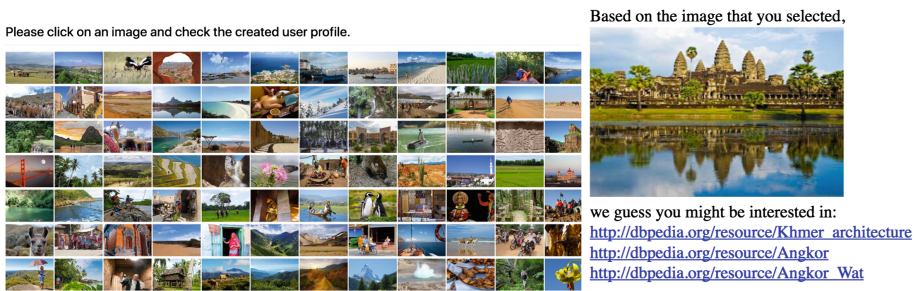
Our second approach requires the following steps:

1. We constitute the conceptual scope of the catalogue.
2. We retrieve the images depicting the entities in the conceptual scope (linked by the property "foaf:depiction").
3. We compute pairwise visual similarity between the input image and the depicting images with the method explained above.
4. We retain the *n* most similar depicting images and thereafter the entities linked to them.

In Fig. 1, we give an example to illustrate the proposed approaches.

4 Demonstration Given at ESWC 2018

In the demonstration given at ESWC 2018, we showcase the second approach which is more advanced. We configured it with a real and recent commercial catalogue of a popular French travel agency. The catalogue contains 1,357 tours which take place in more than 136 countries and regions. The tours are depicted by 11,614 distinct images. The conceptual scope is obtained by item description linking and contains 13,109 DBpedia entities. We provide a Web interface where users can select an image that he/she is interested in and are then shown the profile (top-3 DBpedia entities) corresponding to the selected image (Fig. 2).

Fig. 2. Screenshots of the demo given at ESWC 2018

5 Conclusion

In this paper, we explored the synergy between knowledge graph technologies and computer vision tools. We proposed two novel image user profiling approaches which map an image to knowledge graph entities representing the interests of a user who appreciates the image. We described the demonstration given at ESWC 2018 which is configured with a real and recent travel domain dataset. As future work, we plan to evaluate our profiling approaches and to apply them in personalisation tasks.

References

1. Fang, Y., Kuan, K., Lin, J., Tan, C., Chandrasekhar, V.: Object detection meets knowledge graphs. In: IJCAI, pp. 1661–1667 (2017)
2. Baier, S., Ma, Y., Tresp, V.: Improving visual relationship detection using semantic modeling of scene descriptions. In: d'Amato, C., et al. (eds.) ISWC 2017. LNCS, vol. 10587, pp. 53–68. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68288-4_4
3. Thoma, S., Rettinger, A., Both, F.: Towards holistic concept representations: embedding relational knowledge, visual attributes, and distributional word semantics. In: d'Amato, C., et al. (eds.) ISWC 2017. LNCS, vol. 10587, pp. 694–710. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68288-4_41

4. Ferrada, S., Bustos, B., Hogan, A.: IMGpedia: a linked dataset with content-based analysis of Wikimedia images. In: d'Amato, C., et al. (eds.) ISWC 2017. LNCS, vol. 10588, pp. 84–93. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68204-4_8

5. Di Noia, T., Mirizzi, R., Ostuni, V.C., Romito, D., Zanker, M.: Linked open data to support content-based recommender systems. In: Proceedings of the 8th International Conference on Semantic Systems, pp. 1–8. ACM, September 2012

6. Lu, C., Stankovic, M., Radulovic, F., Laublet, P.: Crowdsourced affinity: a matter of fact or experience. In: Blomqvist, E., Maynard, D., Gangemi, A., Hoekstra, R., Hitzler, P., Hartig, O. (eds.) ESWC 2017. LNCS, vol. 10249, pp. 554–570. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58068-5_34

7. Nguyen, P.T., Tomeo, P., Di Noia, T., Di Sciascio, E.: Content-based recommendations via DBpedia and Freebase: a case study in the music domain. In: Arenas, M., et al. (eds.) ISWC 2015. LNCS, vol. 9366, pp. 605–621. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-25007-6_35

8. Piao, G., Breslin, J.G.: Exploring dynamics and semantics of user interests for user modeling on Twitter for link recommendations. In: Proceedings of the 12th International Conference on Semantic Systems, pp. 81–88. ACM, September 2016

9. Ristoski, P., Paulheim, H.: RDF2Vec: RDF graph embeddings for data mining. In: Groth, P., et al. (eds.) ISWC 2016. LNCS, vol. 9981, pp. 498–514. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46523-4_30

10. Kapanipathi, P., Jain, P., Venkataramani, C., Sheth, A.: User interests identification on Twitter using a hierarchical knowledge base. In: Presutti, V., d'Amato, C., Gandon, F., d'Aquin, M., Staab, S., Tordai, A. (eds.) ESWC 2014. LNCS, vol. 8465, pp. 99–113. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07443-6_8

11. Chen, Y.Y., Cheng, A.J., Hsu, W.H.: Travel recommendation by mining people attributes and travel group types from community-contributed photos. IEEE Trans. Multimedia **15**(6), 1283–1295 (2013)

12. Ferwerda, B., Schedl, M., Tkalcic, M.: Using instagram picture features to predict users' personality. In: Tian, Q., Sebe, N., Qi, G.J., Huet, B., Hong, R., Liu, X. (eds.) MMM 2016. LNCS, vol. 9516, pp. 850–861. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-27671-7_71

13. Neidhardt, J., Seyfang, L., Schuster, R., Werthner, H.: A picture-based approach to recommender systems. Inf. Technol. Tour. **15**(1), 49–69 (2015)

14. Berger, H., Denk, M., Dittenbach, M., Pesenhofer, A., Merkl, D.: Photo-based user profiling for tourism recommender systems. In: Psaila, G., Wagner, R. (eds.) EC-Web 2007. LNCS, vol. 4655, pp. 46–55. Springer, Berlin, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74563-1_5

15. Linaza, M.T., Agirregoikoa, A., Garcia, A., Torres, J.I., Aranburu, K.: Image-based travel recommender system for small tourist destinations. In: Law, R., Fuchs, M., Ricci, F. (eds.) Information and Communication Technologies in Tourism 2011, pp. 1–12. Springer, Vienna (2011). https://doi.org/10.1007/978-3-7091-0503-0_1

16. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016)

17. Di Noia, T., Ostuni, V.C.: Recommender systems and Linked open data. In: Faber, W., Paschke, A. (eds.) Reasoning Web 2015. LNCS, vol. 9203, pp. 88–113. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-21768-0_4