

CONSISTENCY AND RELEVANCE OF VGI AVAILABLE FROM SOCIAL NETWORKS FOR EMERGENCY MITIGATION AND MUNICIPAL MANAGEMENT

C. C. Fonte^{1,2,*}, J. Patriarca², J. Estima^{3,5}, J.-P. de Almeida^{1,2}, A. Cardoso^{2,4}

¹Department of Mathematics - Geomatic Engineering Group, Faculty of Science & Technology, University of Coimbra, Apartado 3008, 3001-501 Coimbra, Portugal - (cfonte, zepaulo)@mat.uc.pt

² Institute for Systems Engineering & Computers at Coimbra (INESCC), Rua Sílvio Lima - Edifício DEEC, S.3.4, 3030-290 Coimbra, Portugal - joaquimpatriarca@gmail.com

³ Setúbal School of Technology, Campus do IPS, Estefanilha, 2914-504 Setúbal, Portugal - jacinto.estima@gmail.com

⁴ CISUC - Department of Informatics Engineering, Faculty of Science & Technology, University of Coimbra, Rua Sílvio Lima, 3030-290 Coimbra, Portugal - alberto@dei.uc.pt

⁵ Institute for Systems Engineering & Computers, Research and Development at Lisboa (INESC-ID), Rua Alves Redol, 9 (office 435), 1000-029 Lisbon, Portugal

Commission IV, WG IV/4

KEY WORDS: Volunteered Geographic Information (VGI), Flickr, Emergency Mitigation, Municipal Management, Accuracy

ABSTRACT:

Volunteered geographical information (VGI) is an increasing source of data for many applications. In order to explore some of these sources of data, an algorithm was conceived and implemented in the *ExploringVGI* platform enabling the collection of georeferenced data from collaborative projects that provide an Application Programming Interface (API). This paper presents a preliminary study to evaluate the consistency and relevance of VGI extracted from Flickr platform for emergency mitigation and municipal management. The study carried out was based on data extraction and analysis with keywords related to emergency events (“Accident”, “Flood” and “Fire apartment”), and municipal management (“Graffiti” and “Homeless”) in four European cities (Frankfurt, Lisbon, London, and Rome). The proposed approach sets up a region of interest on a map, selects one or more keywords for the search, and carries out a search using the Flickr API. Data detected and extracted were then loaded into a database and further analysed to verify whether they were consistently obtained through consecutive searches at different locations. A statistical analysis performed on data collected for each case provided us with: the total number of data collected for each keyword and location; their relevance in terms of search goal; and the quality of the associate geolocation of the post. Results obtained illustrate the effectiveness of the approach when applied to different scenarios, which contributes to assess the role that VGI available on the Web may have in different events depending on the specific context of a geolocation/keyword(s) combination.

1. INTRODUCTION

1.1 Relevance of VGI

Recent developments in Web services, geospatial technologies, along with a population with means to collect and publish various kinds of data, have nourished an environment in which Web users may provide in their daily life, activities, and hobbies vast volumes of data – consisting of, for example, text messages, classifications and geospatial and temporal coded images collected with mobile devices (Coftas and Diosteanu, 2010). This phenomenon is also revolutionizing the whole context of geospatial data creation in such a way that is raising pertinent issues mainly related to its usability in several applications, ranging from market analysis and advertising to data collection for scientific purposes - such as, wild species observation or land-cover map validation (e.g. Bonter and Cooper, 2012; Fonte et al. 2015; Antoniou et al., 2016). Some crowdsourcing projects collect several types of data along with location information expressed by coordinates, called Volunteered Geographical Information (VGI) (Goodchild, 2007). In some cases, like Flickr, Facebook, Twitter or YouTube, those data can be downloaded and visualized over maps.

Given the availability of these large volumes of data, some authors have focused on the redefinition of some concepts in the

context of VGI creation. For instance, Budhathoki (2007) and Budhathoki et al. (2008) have maintained that VGI’s emergence led to the need for reconceptualizing the role of users, often referred to as “end-users”. They are no longer simple passive recipients of information in the context of VGI and therefore represent an extra value. Because it is widely acknowledged that those who are closer to a particular geospatial phenomenon have the richest geospatial knowledge, which should be captured and utilized, citizens should be seen as potential sources of geospatial information – provided that the means for that are understood and successfully created. Even so, a key point about usefulness of VGI is the geospatial data quality, reliability, and trust (Flanagin and Metzger, 2008; Fonte et al., 2017), this information can be considered at least supplementary to other sources (Budhathoki 2007).

Even though the ultimate goal of the work under development is the integration of several sources of data (including VGI and, for example, physical sensor data), this paper only explores the possible usage of VGI on its own to provide data potentially useful for emergency mitigation or municipal management.

1.2 Related work

Our previous work entailed the development of an application to extract and integrate into the same system data collected from

existing collaborative projects that provide an Application Programming Interface (API) for data extraction. The search in each collaborative project is made by: 1) identifying on the map a region of interest - specified through a circular buffer; 2) typing in one or more keywords - which are used through the respective collaborative project API. The obtained data are then loaded onto a database, displayed on a map, and may be further exported either in *shapefile* or *kml* format (Fonte et al., 2018).

Further to our previous work described above, this paper focuses primarily on the analysis of the usefulness of data gathered from some collaborative projects, namely Flickr, Facebook, Twitter, and/or YouTube. For this purpose, circular regions of interest were selected within different European cities. Sets of predefined keywords to identify data with potential interest in emergency situations (e.g. “Flood”, “Accident”, “Fire apartment”), or municipal management (e.g. “Graffiti”, “Homeless”), were used and searches were carried out using these keywords. For purposes of this paper, only Flickr was used as an example scenario.

2. METHODOLOGY

To investigate the problem stated above, we followed the subsequent methodology:

1. Identification of study areas in four European cities and creation of a search schema that performs multiple searches including the same area, to assess data download consistency
2. Downloading of posts/photographs within the predefined areas
3. Manual classification of the photographs content according to a given keyword in terms of usefulness
4. Evaluation of the geositional accuracy using the information visible in the downloaded photographs

The aforementioned steps are described in detail in the following sections.

2.1 A multi-circular buffer-based search

Flickr data was searched and captured using the Web application *ExploringVGI App* (<https://vgi.uc.pt>). Data obtained was then analysed in order to check whether the retrieved photographs were consistently obtained in consecutive searches over different size and placement circular buffers - covering though all together the same region of interest. Such consistency is indeed crucial to assess whether the limitations above have a real impact in data gathered when various search regions, with different sizes and geolocations, are considered.

For purposes above, an algorithm was designed in order to carry out, for each region of interest, a “multi-circular buffer” based search. In such search, a 10km-radius circular buffer is initially placed by the user in the central part of the region of interest; in turn, eight half-radius (i.e. 5km) circular buffers circumscribed to the initially set up 10km-circular buffer are then considered; those eight inner buffers are positioned within their surrounding 10km-circle aligned with the corresponding eight azimuth compass octants in such a way that they cover all together the entire 10km-circle’s area - as illustrated in Figure 1.

To accomplish the “multi-circular buffer” based search, the algorithm described and illustrated was prepared to use the “flickr.photos.search” method available through the Flickr API in order to extract all data according to the search parameters

(search area and keywords of interest). The algorithm executes the “flickr.photos.search” for all combinations of search areas versus keywords, producing a table with the Flickr posts/photographs detected within the area of interest for each combination search-area/keyword. All data were subsequently aggregated into a single table, which includes a serialization of all posts/photographs found in the survey. In addition to the publication geolocation, the algorithm identifies some indicators by extracted post related to the search process, such as:

- Identification of the buffers used in the search that returned a given publication
- Identification of the buffers containing the post/photograph geolocation
- The number of buffers that contain the corresponding post geolocation
- The number of times a post/photograph was returned by the Flickr API
- The number of times a post/photograph was returned using a search area that contains the corresponding post geolocation
- The number of times a post/photograph was returned using a search area that does not contain the corresponding post geolocation
- The number of times a post/photograph was not returned although it was located within in the search area.

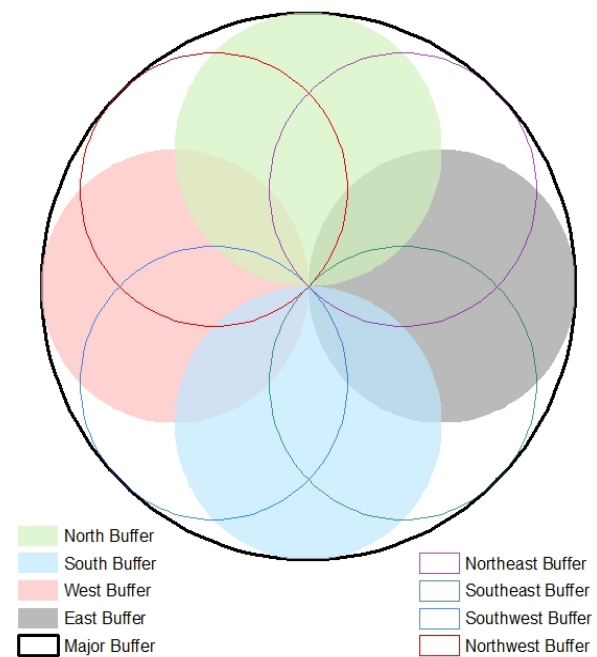


Figure 1. Relative spatial position of the multi-circular buffers considered in each single search

As a complement, the algorithm also produces a table with some indicators related to the number of posts/photographs obtained using a particular search area. For each search area, such table quantifies the following indicators.

- The number of posts/photographs obtained using a particular search buffer
- The number of posts/photographs located outside the search buffer but still returned
- The number of posts/photographs contained within the search buffer

- The number of posts/photographs contained within the search buffer but not returned when the associate search radius was used.

2.2 Tagging of captured data

Extracted photographs were analysed in terms of their relevance towards the ultimate goal of the search. As such, each contribution was tagged as “useful”, “not useful” or “maybe useful” according to the keyword used. This classification was manually performed by a group of volunteers.

2.3 Evaluation of geospatial accuracy of captured data

A sample of data with photographs classified as “useful” was then used to perform positional accuracy analysis. Those photographs were manually assigned a pair of coordinates by volunteers using local knowledge and other sources, like satellite image interpretation and Google Street View; those coordinates were then used as a reference in order to calculate the difference between the two positions (post versus reference). The difference was then used to classify the positional accuracy of the post/photograph according to five classes: 1) “exact match” for differences up to 20 meters; 2) “nearby match” for differences between 20 and 50 meters; 3) “close match” for differences between 50 and 100 meters; 4) “farther match” for differences between 100 and 300 meters; and 5) “no match” for differences over 300 meters.

2.4 Statistical analysis

A statistical analysis was applied to data collected in order to identify: 1) the number of posts/photographs extracted by each considered buffer, their relative position to the buffer (contained/not contained) and consistency of data captured when searched using different buffers; 2) the total number of data per keyword extracted for each considered study area; 3) their relevance envisaging the ultimate goal of a search (such as, identification of flooded areas or sites where graffiti can be observed); and 4) the quality of the geospatial location related to the user contribution/post.

The main aim of this analysis is to compare, for each type of situation under analysis, the quantity and quality (in terms of usefulness and positional accuracy) of the data obtained. Moreover, to identify limitations imposed by the available API of the considered web collaborative project considered, and which changes to tackle them would be useful to be implemented.

3. CASE STUDIES

3.1 Study areas & keywords

For purposes of this paper, circular regions of interest (as described in Section 2.1) were selected within four European city centres: Lisbon, Frankfurt, London, and Rome. Searches were performed taking into consideration five different keywords: “flood”, “Accident”, graffiti”, “homeless”, and “Fire apartment”. In these particular case studies, keywords in the local native language were not considered.

3.2 Results

Further to the application of the proposed methodology to the case studies described in the previous section, an analysis was carried out to obtain some indicators that could contribute to the evaluation of both consistency and relevance of the identified posts. In the next subsections, the main results are presented

contributing to the quantification of these indicators, which will be discussed at the end.

3.2.1 Retrieved posts dates and location: The total number of posts downloaded was 425 for Lisbon, 529 for Frankfurt, 744 for Rome, and 1067 for London. Figure 2 shows the temporal distribution of posts per city, where a trend of posts increase can be observed per city in the last two decades, with an exception for Frankfurt, where a very large number of contributions was obtained between 2010 and 2015.

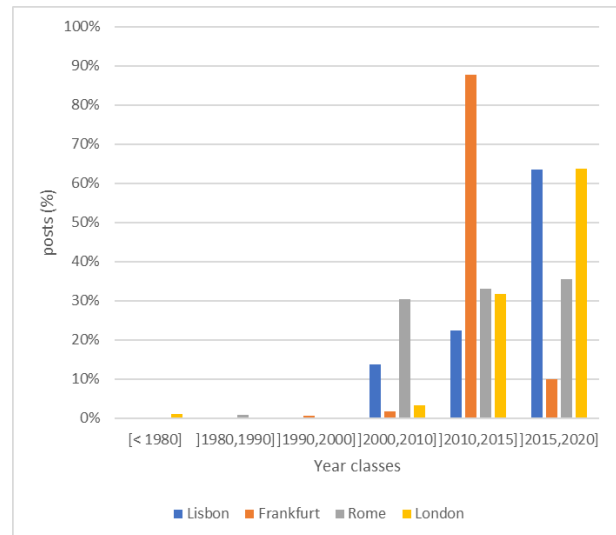


Figure 2. Temporal distribution of posts by city, with posts being grouped by classes of years

Figures 3 to 6 show the location of the extracted posts, classified by the keyword used for their extraction, for the four study areas. These Figures show that in some cases the posts for some keywords are spread all over the study area, and in other cases they are clustered in some regions. For example, for Lisbon the posts extracted with the keyword “Graffiti” are mostly grouped in the city historical region (Figure 3), while for Frankfurt (Figure 4) most of the posts extracted with the keyword “Flood” are along the river bed.



Figure 3. Position of the extracted posts for the Lisbon study area, coloured according to the keyword used in the search, and the major buffer

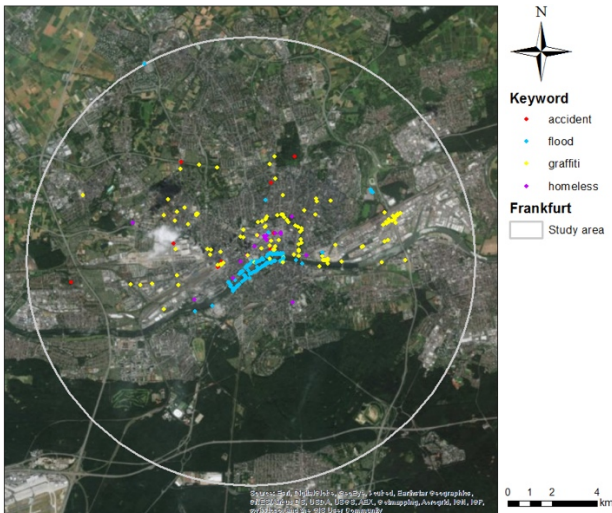


Figure 4. Position of the extracted posts for the Frankfurt study area, coloured according to the keyword used in the search, and the major buffer

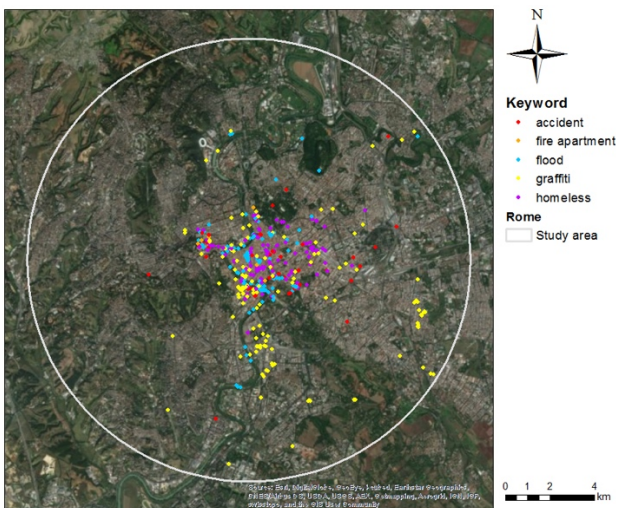


Figure 5. Position of the extracted posts for the Rome study area, coloured according to the keyword used in the search, and the major buffer

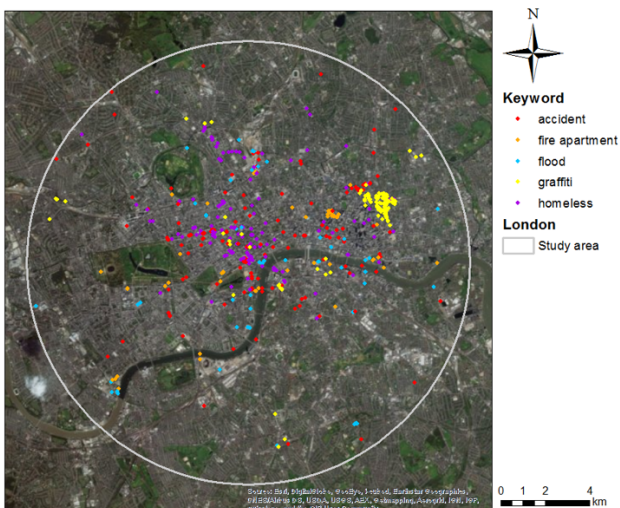


Figure 6. Position of the extracted posts for the London study area, coloured according to the keyword used in the search, and the major buffer

3.2.2 Retrieved posts related to the buffers: A first statistical analysis was performed to understand whether the different buffers were retrieving the same posts on overlapping areas, following the method described in section 2.1. Figure 7 shows the percentage of posts that were not extracted when considering their own containing buffers. The results are grouped by classes that represent the percentage of times a post was not extracted when considering the buffers that contain that post location. For example, class [0,0] contains the posts that were retried when considering all buffers containing the post. Class [0,20] contains the posts that were not retrieved when considering up to 20% of the buffers that contain their location, and so forth for the other classes shown in Figure 7. Results showed that for all cities most of the posts were retrieved by all the containing buffers (class [0-0]), even though that is not always the case. For example, for London more than 15% of the posts were not retrieved by up to 20% of the containing buffers (class [0,20]), while 5% were not retrieved when considering between 20% and 40% of the containing buffers (class [20,40]).

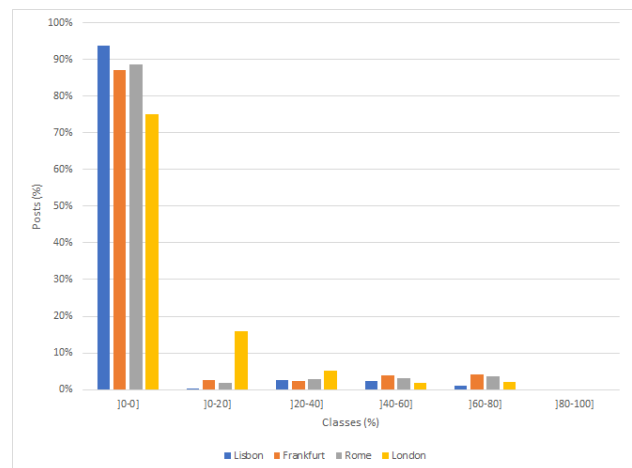


Figure 7. Percentage of posts that are positioned within a given buffer but were not retrieved when using that buffer

Table 1 shows: 1) the number of posts extracted within the area of interest (corresponding to the larger buffer); 2) the number of posts extracted when considering the larger buffer; and, 3) the number of posts that would not have been identified if the smaller buffers had not been considered - corresponding to the difference between the second and third columns. It can be seen that in all instances the number of posts/photographs captured by the larger outer buffer (larger buffer in Figure 1) is indeed different from that obtained with its inner smaller buffers and always smaller than the number of photographs captured by the smaller search buffers.

City	Posts in the areas of interest	Posts obtained using the major buffer	Posts that would be lost without the use of smaller buffers
Lisbon	425	398	27
Frankfurt	529	461	68
Rome	744	659	85
London	1067	802	265

Table 1. Comparison of the number of photographs captured by the outer major buffer (interest area) and its inner buffers

The procedure used to assess the consistency of data obtained by these buffers allowed therefore the identification of some inconsistencies in the Flickr API search tools. Indeed, even though the maximum number of posts that, according to the API description, could be extracted at each time was not reached in any of the searches (500 posts), contrary to what would have been expected, more posts in the study area were obtained when multiple smaller search buffers were used instead of a single search using the major buffer.

3.2.3 Posts analysis per keyword: Considering that this analysis is focused on images, Table 2 shows the total number of posts and the number and percentage of posts with a link to an effective photograph, grouped per keyword for all cities considered in this study. This number does not include posts where the link is not active or does not match a photograph.

From these results, it is possible to observe some non-uniformity of the total number of posts per keyword. Though a high percentage of posts do have effective photographs. “Graffiti” is the keyword with the highest number of contributions in all cities, most of which are appropriate. On the contrary, “Fire apartment” was the keyword with fewer occurrences also in all cities.

	keywords				
	Accident	Flood	Graffiti	Homeless	Fire apartment
Lisbon					
Extracted posts	41	19	273	92	0
Posts with photo	39	15	270	74	0
%	95%	79%	99%	80%	0%
Frankfurt					
Extracted posts	14	252	244	19	0
Posts with photo	12	252	242	14	0
%	86%	100%	99%	74%	0%
Rome					
Extracted posts	55	251	253	171	14
Posts with photo	47	171	247	148	14
%	85%	71%	98%	87%	100%
London					
Extracted posts	237	242	250	245	93
Posts with photo	235	67	250	245	89
%	99%	28%	100%	100%	96%

Table 2. Total number of posts, as well as number and percentage of posts with photographs per keyword for all considered cities

The subsequent analysis was performed by volunteers to evaluate the usefulness of each posted photo and classify them qualitatively as “Useful” (Yes), “Maybe useful” (Maybe) or “Not Useful” (No). As an example, Table 3 shows the number of photographs per classifier and by keyword for the city of Lisbon, classified by two different volunteers.

Comparing the classifications made by two different volunteers, they have agreed in 89% of the cases, representing the high assertiveness of the process. Some of the divergences correspond to cases where a subjective interpretation on how useful that image would be for the aim under analysis. For example, Figure 8 shows a photograph extracted when using the keyword “Graffiti”. In fact, the image shows a graffiti on a public transport vehicle. However, even though an automatic classification of this

photograph as “graffiti” would probably result in a positive classification, the actual graffiti is placed on a funicular tram, therefore the associate geolocation is meaningless in terms of municipal management as it corresponds to a moving object.

Vol.	Class	Keywords			
		Accident	Flood	Graffiti	Homeless
V1	Yes	2 (5%)	0 (0%)	224 (83%)	45 (61%)
	Maybe	4 (10%)	0 (0%)	29 (11%)	15 (20%)
	No	33 (83%)	15 (100%)	17 (6%)	14 (19%)
V2	Yes	6 (16%)	0 (0%)	244 (91%)	44 (60%)
	Maybe	4 (10%)	0 (0%)	9 (3%)	18 (24%)
	No	29 (74%)	15 (100%)	17 (6%)	12 (16%)

Table 3. Number and percentage of photographs classified as “Useful” (Yes), “Maybe useful” (Maybe) or “Not useful” (No) per keyword for Lisbon by two different volunteers V1 and V2



Figure 8. Photograph extracted using the keyword “Graffiti” in the city of Lisbon (Flickr, 2018a)

Figure 9 shows the percentage of photographs classified as “Useful”, “Maybe useful” and “Not useful” by keyword and city. It can be seen that for the keyword “Accident” the majority of photographs were classified as “Not useful” for all considered cities (between 69% and 92%). For the keyword “Flood”, the same happens to all cities except Frankfurt, where 69% of the photographs were classified as “Useful”. All photographs obtained for Lisbon corresponding to this keyword were classified as “Not useful”. For the keyword “Graffiti” the great majority of photographs were classified as “Useful” (more than 80% for all cities), while for the keyword “Homeless” around 50% of the photographs (between 42% for London and 61% for Rome) were classified as “Useful”. For this last keyword, Frankfurt and Lisbon have a relatively high percentage of photographs classified as “Maybe useful” (respectively 29% and 22%). This is mainly due to the difficulty of classifying the depicted persons as homeless or not (e.g. see Figures 10 and 11), or discriminating between beggars and real homeless people, which are not necessarily the same, depending on the ultimate aim of the analysis. For the keyword “Fire apartment” no posts were obtained for Lisbon and Frankfurt, and for Rome and London more than 90% were classified as “Not useful”.

These results show that the keyword “Graffiti” has the highest percentage of photographs classified as “Useful”, followed by the keyword “Homeless”. Posts associated with “Accident”, “Flood” and “Fire Apartment” also showed to have little relevance to the desired purpose of this analysis.

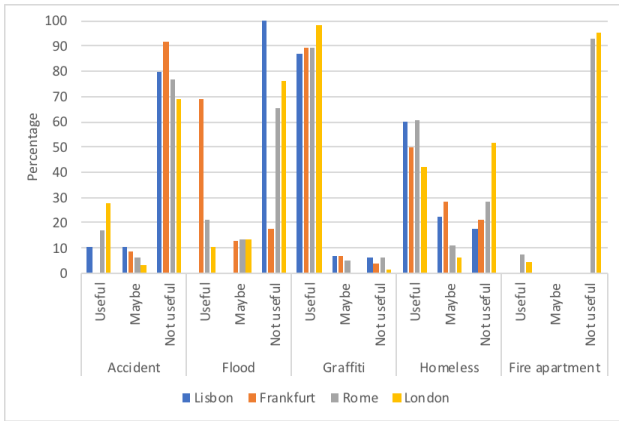


Figure 9. Percentage of photographs classified as “Useful”, “Maybe useful” and “Not useful” by keyword and city.



Figure 10. Photograph clearly classified as “Useful” regarding the keyword “Homeless” (Flickr, 2018b)



Figure 11. Photograph classified as “Maybe useful” regarding the keyword “Homeless” (Flickr, 2018c)

3.2.4 Classification regarding location: In order to assess the accuracy of the geolocation of posts (regarding the true location of the photographer when the photograph was taken), the methodology described in section 2.3 was only applied to a sample of the photographs of Lisbon. To overcome changes in

the city that would make the identification of the true location of the photographer more difficult, only photographs taken in 2018 were considered (32 photographs). Out of these, the true location of 11 could not be identified (34%), either because there were no reference points in the photograph that enabled the identification of its location (for example, see Figure 12, where only a wall is shown with some graffiti, that was not found in the vicinity and may even be located inside private property), or because using only image interpretation and Google Street View was not enough to identify the shown region in the vicinity of the post location, and probably means that its true location is in fact relatively far from the post location. Whenever a photograph was showing city landmarks, even when the post position was far away, it was still possible to identify the “true position” of the photographer.

Figure 13 shows the results obtained for the considered sample. It can be seen that 29% of the photographs to which a “true position” was assigned were equally located very close to the position of the post (between 0 to 20 m away), between 20 m and 50 m away, and between 50m and 100m away from the post. Then only 10% were between 100m and 300m away and one photograph (5%) was located at a much larger distance (718m) of the post. Figure 14 shows the photograph of this post. Its true location was only found because there is a limited number of streets with stairs in the city.



Figure 12. Photograph extracted using the keyword “Graffiti” in the city of Lisbon (Flickr, 2018d), which true location was not identified.

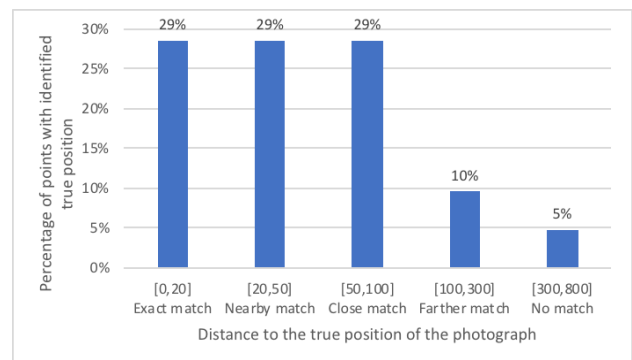


Figure 13. Percentage of posts whose distance (in meters) between their true location and the post location lies within the indicated distance intervals



Figure 14. Photograph extracted using the keyword “Graffiti” in the city of Lisbon (Flickr, 2018e), which true location was identified 713m away from the post location

4. CONCLUSIONS

The study presented in this article aims to assess the relevance and usefulness of photographs extracted from the Flickr project with one of the services available within the *ExploringVGI* App, for applications related to emergency mitigation and municipal management. The consistency of the obtained data in terms of how they were retrieved was also assessed.

The results show that the API used to extract the Flickr posts does not consistently retrieve the posts located inside each of the considered buffers. Therefore, considering additional buffers with different positions and radius enables to extract more posts than only one larger buffer, even when the maximum number of posts that can be downloaded with one search is not achieved.

Post were obtained for all the considered keywords, except for the keyword “Fire apartment” for the cities of Lisbon and Frankfurt. The classification of the usefulness of photographs is not always easy, as in some cases, the content is either not clear or allows for multiple interpretations depending on the background of the interpreter and the aim of the analysis. However, for the keywords related to emergency mitigation (“Accident”, “Flood” and “Fire apartment”) most of the posts showed to be irrelevant for that purpose. The only exception were the posts obtained for “Flood” in the city of Frankfurt, which were almost all located along the river and showed images of a flood and the affected region. On the other hand, the majority of posts obtained with the keywords related to municipal management (namely “Graffiti” and “Homeless”) were classified as “Useful”, as they actually showed either places with graffiti in the city or homeless people.

Regarding the reliability of the posts geolocation, most of those to which a “true location” could be assigned based on the analysis of the photograph, showed to be within a distance of up to 100m. Geolocation errors like those in an urban fabric may be due to the Global Navigation Satellite System (GNSS) positioning errors, as multipath is expected to occur mainly in neighbourhoods with narrow streets. However, large positioning errors are probably due to the original geolocation of photographs being obtained not by the photographic camera but using an alternative positioning method, such as the manual definition of the geolocation when data was uploaded.

ACKNOWLEDGEMENTS

This work has been partially supported by INESC Coimbra under project INESCC FR004 (ExtremeCGI - Monitoring Extreme Events integrating Crowdsourced Geo- graphic Information and Real-time Sensor Data) and the Portuguese Foundation for Science and Technology (FCT) under the project grants UID/MULTI/ 00308/2013 and UID/EEA/00066/2013.

REFERENCES

- Antoniou, V., Fonte, C., See, L., Estima, J., Arsanjani, J., Lupia, F., Minghini, M., Foody, G., Fritz, S., 2016. Investigating the Feasibility of Geo-Tagged Photographs as Sources of Land Cover Input Data. *ISPRS International Journal of Geo-Information* 5(5), pp. 64-64.
- Bonter, D., Cooper, C., 2012. Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment* 10, pp. 305–307.
- Budhathoki, N., 2007. Reconceptualization of user is essential to expand the voluntary creation and supply of spatial information. *Vespucci Workshop on Volunteered Geographic Information*. Santa Barbara (CA, USA), 13-14 December.
- Budhathoki, N., Bruce, B., Nedovic-Budic, Z., 2008. Reconceptualizing the Role of the User of Spatial Data Infrastructure. *GeoJournal* 72(3-4), pp. 149-160.
- Coftas, L., Diosteanu, A., 2010. Evaluating Accessibility in Crowdsourcing GIS. *Journal of Applied Collaborative Systems* 2(1), pp. 45-49.
- Flanagin, A., Metzger, M., 2008. The credibility of volunteered geographic information. *GeoJournal* 72(3-4), pp. 137-148.
- Flickr, 2018a. https://farm5.staticflickr.com/4200/35014143405_bab9ab0a11_b.jpg (July 2018)
- Flickr, 2018b. https://farm3.staticflickr.com/2628/3994723570_9b92d3a4ae_b.jpg (July 2018)
- Flickr, 2018c. https://farm1.staticflickr.com/34/71879201_5899839afb_b.jpg (July 2018)
- Flickr, 2018d. https://farm1.staticflickr.com/906/42045855322_088ffb1b64_b.jpg (July 2018)
- Flickr, 2018e. https://farm1.staticflickr.com/898/26407412717_bd671d2a25_b.jpg (July 2018)

Fonte, C., Bastin, L., See, L., Foody, G., Lupia, F., 2015. Usability of VGI for validation of land cover maps. *International Journal of Geographical Information Science*. 29 (7) , pp. 1269-1291.

Fonte, C., Antoniou V., Bastin L., Estima, J., Arsanjani, J. Laso-Bayas, J.C., See, L., Vatsava R., 2017. Assessing VGI Data Quality. In: *Mapping and the Citizen Sensor*. London, Ubiquity Press.

Fonte, C., Fontes, D., Cardoso, A., 2018. A web GIS-based platform to harvest georeferenced data from social networks: examples of data collection regarding disaster events. *International Journal of Online Engineering* 14(2), pp. 165-172.

Goodchild, M., 2007. Citizens as Sensors – The World of Volunteered Geography. *GeoJournal* 69(4), pp. 211-227.