DeepTags: Integration of Various VGI Sources Towards Enhanced Data Quality

Ahmed Loai Ali
University of Bremen
Bremen, Germany
Bremen Spatial Cognition Center
loai@informatik.uni-bremen.de

Rami Al-Salman
trivago GmbH
Düsseldorf, Germany
rami.alsalman@trivago.com

ABSTRACT

During the last 10 years, various volunteered geographic information (VGI) sources have been developed under the authority of crowdsourcing communities. The resulting data showed a great role in various application domains (e.g., map provision). However, several research addressed the questionable quality of VGI. They argued for developing intuitive strategies to integrate and to enrich data sources with each others. In OpenStreetMap (OSM), many contributions have abstract level of details or incomplete information that limit their use. However, the missing information could be available in other voluntary data sources. For instance, Flickr and Mapillary, contain geotagged photos that depict geographic features. Visual inspection of these photos can be used to enrich the OSM data. This paper investigates plausible utilization of geotagged photos of Flickr to enrich the OSM data. We applied deep learning (DL) algorithms on the photos to generate tags of various related concepts. The resulting tags can be used either to assure and enrich existent contributions or to create new contributions. The paper presents the abstract description of the proposed approach, as well as the encountered challenges. The findings show the potential role of DL in enhancing OSM data quality.

Keywords

VGI quality, OpenStreetMap, Flickr, deep learning

1. INTRODUCTION

Ubiquity of location aware devices and advanced information and communication technologies facilitate the way of geographic data collection. Following the crowdsourcing paradigm, public collect and develop volunteered geographic information (VGI); VGI is a particular kind of user generated content (UGC), such as, Wikipedia[1], where the content is related to geographic properties. Under authorities of crowdsourcing communities several geographic contents have been developed serving different application domains. For example, Wikimapia[2], OpenStreetMap[3] (OSM), Geowiki[4], and Mapillary[5] are examples of voluntary mapping projects. In addition, geotagged contents at online platforms like Flickr[6] and Twitter[7] represent another format of VGI. During the last decade, GIScience research showed the role of VGI as a potential source of information. It can act either as solely or as complementary data source. However, the research pointed out issues related to data quality of VGI as well.

We are interested in VGI mapping projects (i.e., collaborative mapping) particularly the OSM project. Currently, the OSM data covers most of the world and is supported by millions of contributors. These contributors have various characteristics. These varieties results in rich data sources, however, with heterogeneous data quality. Completeness is one problematic issue of data quality in the OSM. Many contributed features have lack of details, while others have less attractions of contributions, and consequently, they provide incomplete information. However, the missing information can be founded in other VGI sources.

This paper focuses on the integration of various VGI sources towards enhancing data quality. For example, in the OSM data there exist many contributions with an abstract level of details (LOD), such as, a *building*, a *water body*, a *beach*, etc. However, what is the type of the building? Which kind of water body it is? Are sailing activities valid on that beach? All these details might be available in another VGI resource (e.g., geotagged photos). Thus, this paper proposes an approach to enhance the OSM data quality by integrating geotagged photos from Flickr.

The proposed approach applies deep learning (DL), which is a branch of machine learning that shows promising findings when it is employed on VGI [11, 30]. In our approach DL is utilized to generate tags from a given geotagged photo. These tags can be utilized to enhance the quality of the OSM contributions by various scenarios. The proposed approach demonstrates two enhancement scenarios of the OSM data quality. So far, the finding indicates the feasibility of our approach regardless the implementation challenges.

[1] `https://www.wikipedia.org/`

[2] `http://wikimapia.org/`
[3] `http://www.openstreetmap.org/`
[4] `http://geo-wiki.org/`
[5] `https://www.mapillary.com/`
[6] `https://www.flickr.com/`
[7] `https://twitter.com/`

The reminder of this paper is organized as follow: Section 2 presents related work of our approach. The proposed approach description, illustration, and implementation challenges are argued in Section 3. Finally, Section 4 includes conclusions and points out to further developments.

2. RELATED WORK

This section gives insights about related work regarding the questionable data quality of VGI and the potential role of geotagged photos in various applications (Section 2.1). As well, it presents the applications of deep learning to enhance VGI quality (Section 2.2).

2.1 Data Quality in VGI

Various research investigate the data quality of OSM from multiple views: biased contributions among urban and rural areas [13], the increase attraction of particular features (e.g., highways) over the others, inconsistent level of details (LOD) [14, 29, 32], and problematic classification [2, 3]. Otherwise, researchers propose motivation procedures [12], integration strategies [24], and semantic interpretations solutions [6] to tackle those issues. Moreover, others call for standard procedures of data collection in [22] toward enhanced data quality of VGI.

From other side, research analyzes the quality of geotagged photos in Flickr, Panoramia, and different sources as well [4, 5]. They show potential role of these photos to explore places [15], to make sense of the context [17], and to support disaster relief [20]. The work in [26] presents a comprehensive review of the challenges of using Flickr photos to perform auto-tagging, to extract knowledge, and to develop applications of traveling and humanities.

2.2 Deep Learning

DL shows promising results in several areas including natural language processing [9] and computer vision [8]. DL in particular shows a great success in images recognition area (e.g., images classification) [1, 31, 19, 16]. For performing several tasks such as images classification, the DL systems are relying on Convolutional Neural Networks (CNNs) [21] to achieve very close performance to humans [1]. For example, Bauidu, and Google have developed images classifiers that have 5.98% and 6.67% error rates respectively [1, 27], where humans achieved 5.1% error rate.

Images classification based on CNNs have successfully been applied to VGI domain. DeepOSM[8] and the work proposed in [25] show DL satellite images classification systems with 80% and 69% accuracy rates respectively. They use the developed classifiers to enrich the OSM data, e.g., missing object labels such as “parking”. Recently, DeepVGI [7] has been proposed as a research project to improve data quality of VGI, particularly the OSM. DeepVGI aims to train classification models by using big datasets of the OSM, where such models can be used to classify objects such as buildings, grass lands, etc, and that can be used to enrich OSM data.

So far, there are no works that apply DL to target the integration of various VGI sources. Limited studies investigate knowledge extraction from geotagged photos, however, without intension to employ the extracted knowledge to enrich voluntary mapping projects. In contrast to others, we utilize DL to integrate various resources of VGI. This work aims to increase the level of granularity in OSM, while others enriched the OSM only on a higher level of abstraction.

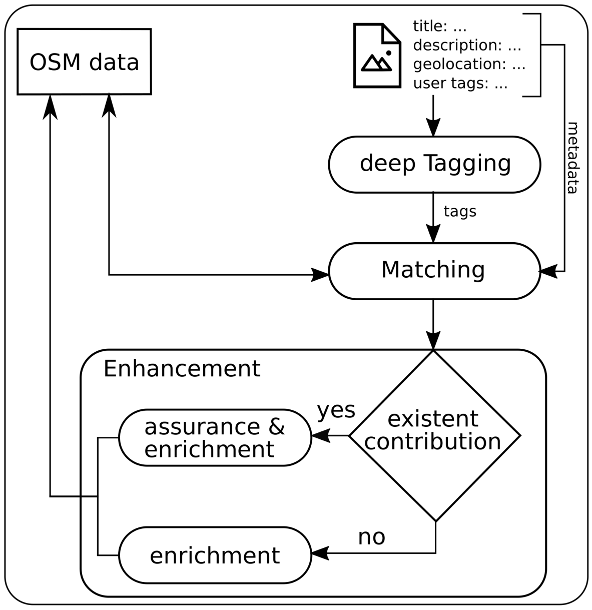


Figure 1: The proposed TME approach that consists of three processes: deep Tagging, Matching, and Enhancement

3. TOWARD ENHANCED DATA QUALITY

The approach aims to enhance OSM data quality by two ways: 1) ensure the quality of the existent contributions; and 2) enrich the OSM data either by adding more information to the existent contributions or by creating new contributions. We apply DL algorithms (i.e., deep tagging) on geotagged photos from Flickr. The resulting tags are utilized to enhance the OSM data quality. The proposed approach description, enhancement scenarios of data quality, and the encountered implementation challenges are described in the following subsections.

3.1 Tagging-Matching-Enhancement (TME)

In this work, the geotagged photos are photos that are assigned with a geographic location by either an explicit or implicit way. Figure 1 illustrates the proposed approach, which consists of three processes as follows: Tagging, Matching, and Enhancement.

- Deep Tagging: assume a given geotagged photo, for instance, from Flickr, we use CNNs as a powerful DL-based classifier to extract tags of the photos [23]. Particularly, we use Google Inception v3 [28] as a mature variant of CNNs that is trained by using 9 millions images dataset called OpenImages [18]. The extracted tags describe the photos from different concepts and with various levels of abstractions.

[8]DeepOSM (`https://github.com/trailbehind/DeepOSM`)

- Matching: taking into account the location of a given photo and the generated tags, matching process aims to identify a geographic identical entity in the OSM, such as, a building or a beach that is depicted by the photo. The process might find out non-existence of matched entities in the OSM.

- Enhancement: our approach proposes two scenarios; first the assurance scenario, when a matched entity is found, the process acts to ensure the existent entity; and second the enrichment scenario, when the process either adds additional information (i.e., increase level of granularity) of an existent entity or creates a new entity in case of non-existence, if that is possible. This entire process is illustrated by examples in Section 3.2.

3.2 Proposed Enhancement Scenarios

This section exemplifies the potential scenarios to enhance the OSM data quality. The proposed enhancement depend on whether a matched entity is identified or not as well as on the ability to extract useful knowledge from deep tagging process.

3.2.1 OSM Assurance Scenario

Figure 2 illustrates the assurance scenario. For example, the Hyde Park in London is one of the most attractions for locals and tourists as well. In OSM, the park is represented by an areal entity associated with tags of `name = Hyde Park, leisure = park, and access = yes` (see Figure 2a). This information may be not sufficient for tourists, who might look for additional pictorial information. The limited LOD on the OSM might be unable to answer questions like *How does this park look like? Is the park appropriate for jogging?* or *What are the scenic views in this park?*

In contrast, the place has large number of attraction for photographers, which is indicated by large availability of geotagged photos on Flickr, as seen in Figure 2b. On one hand, applying deep tagging on such photos in Figure 2c results in “park”, “landscape”, “nature”, “tree”, “grass”, and “footways” tags. The tags can be utilized to ensure existence of the OSM contribution as well as to provide further LOD (e.g., `landuse/landcover = grass`). On the other hand, the given photos can be linked to the OSM entity to add pictorial information and to ensure the contributed entity on OSM.

3.2.2 OSM Enrichment Scenario

Figure 3 describes the enrichment scenarios. The Coney Island Beach & Boardwalk is one of best boardwalks in New York (see Figure 3). The place is mapped on OSM with simple tags as follow: `name = Coney Island Beach, natural = beach, and ele = 1` (indicates the elevation above sea level) (see Figure 3a). However, these tags might provide limited information about the place. For example: *Is it public or private beach*? *Which sport activities (swimming, surfing, etc.) could people do there*? *Are dogs allowed on the beach or not*?, and for elderly people who would like to enjoy walking there, *are benches and seats facilities are available alongside the boardwalk*?

At the same moment, the place attracts thousands of tourists and local people. They voluntarily upload their personal photos on web-shared image platform. For instance, Flickr has about 125,000 geotagged photos within the location of the Coney Island Beach as indicated in Flickr explore world API on Figure 3b. Hence, the integration with such data might be one possible solution to enrich OSM data by increasing the LOD.

One proposed scenario is to enrich the OSM by adding additional information to an existent contribution. For example, applying deep tagging on the left photo in Figure 3c will result in tags like “dog”, “beach”, and “recreation”, while applying it on the right photo of Figure 3c generates tags like “sailboat”, “sailing”, “water sports”, and “boat”. Furthermore, deep tagging of other photos results in tags of “swimming”, “bikini”, “people”, and “sand”. By these tags, first we assure the existence of the contributed entity on Figure 3a; second, we can increase the LOD of this entity by adding tags like `dogs = yes`, `sport = swimming`, `sport = sailing`, and `surface = sand`.

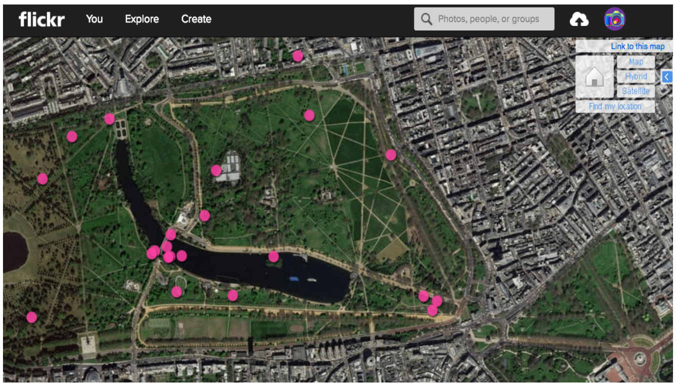
Another proposed enrichment scenario is when the matching process can not find a counterpart entity (i.e., where some features in photos are not mapped on the OSM). Our approach proposes to create new contributions to OSM (when it is possible depending on the availability of sufficient photos with precise geographic location). For example, applying the deep tagging process on photos in Figure 3d results in tags like “seat”, “chair”, “bench”, “wood”, and “wooden”, while the matching process and manual visual inspection can not identify such of these features on the OSM. Based on photos’ geolocation, we can create a new contribution, like a node, with tags of `amenity = bench` and `material = wood`.

3.3 Discussion and Challenges

The previous scenarios illustrate the potential improvements of the OSM data quality. This paper investigates only utilizing geotagged photos from Flickr as an example. However, with increase availability of geotagged photos on various platforms (e.g., Microblogs), the OSM data quality can be improved toward professional use. In tourism applications, the integration can be utilized to enrich the OSM data especially in the places where there are no active mapper communities. In uncommon situations, tourists like to have spatial and pictorial information together of attraction places. The integrated information could be provided with a particular degree of uncertainty. This paper presents a preliminary discussions on feasibility of the proposed approach. Although our approach sounds straightforward, it has multiple challenges regarding its implementation.

It is commonly known that not all public geotagged photos contain useful information. Most of these photos includes persons, and hence, they are subjective to privacy issues. Therefore, an adapted filtering preprocess is needed to find out useful photos. In deep tagging process, not all the resulting tags might be useful. Some tags represent high level of abstraction, such as, “water”, “human”, “natural”. They might describe a given photo from various concepts; and thus, semantic interpretation of these tags might not be a trivial task. Moreover, temporal aspects must be consider, as a photo might reflect a temporal feature that has been changed (removed/modified) with time.

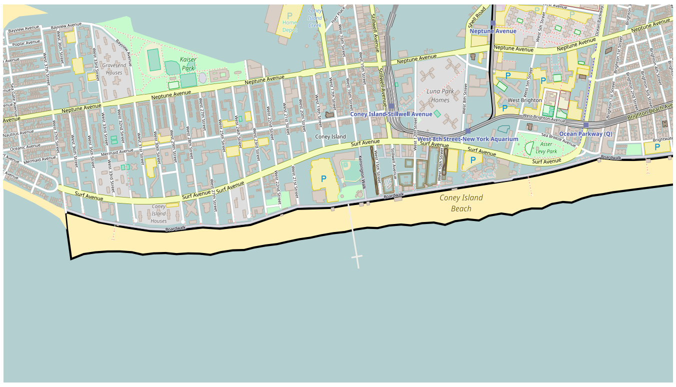
(a) The Hyde Park in London on the OSM with tags of: name = Hyde Park, leisure = park, and access = yes



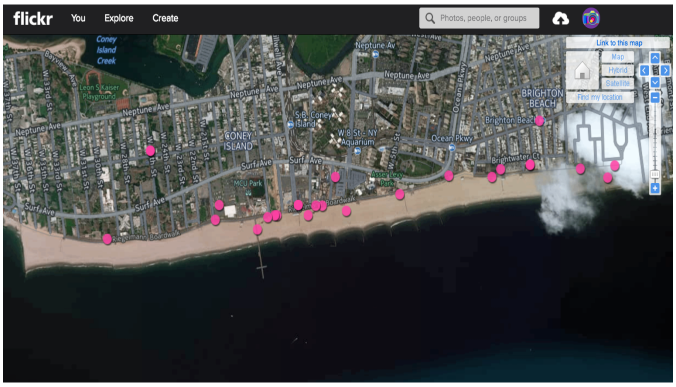
(b) The Hyde Park in London on Flickr web platform includes around 87,000 geotagged photos related to this park

(c) Applying deep tagging on these photos results in “park”, “landscape”, “nature”, “tree”, “grass”, and “footways” tags

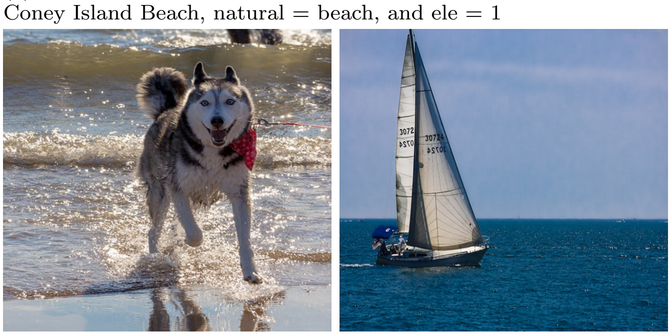
Figure 2: Assurance scenario of the OSM data by applying the TME approach



(a) The Coney Island Beach at NY on OSM with tags: name = Coney Island Beach, natural = beach, and ele = 1



(b) The Coney Island Beach at NY on Flickr web platform includes around 125,00 geotagged photos



(c) Applying deep tagging on these photos results in “dog”, “beach”, “recreation”, “sailboat”, and “water sports” tags

(d) Applying deep tagging on these photos results in “bench”, “chair”, “fence”, “wood”, and “seat” tags

Figure 3: Enriching scenarios of the OSM data by applying the TME approach

The matching process has many difficulties; It depends on the positional accuracy of the assigned geotags to the photos. For instance, the geotag attribute of Flickr's photos is described by a numerical scale (1 to 16), from region∼3, to city∼6, to street∼16 levels. Qualitative and quantitative aspects should be considered, such as, distance, direction, topology, and orientation. Qualitative Spatial Reasoning (QSR) [10] might have a primary role to tackle these aspects.

Otherwise the entire enhancement process can not be carried out by an automatic way, as automatic contributions are discouraged from the OSM community. Therefore, we might need volunteers to investigate and revise the proposed enhancement scenarios. One possibility is to employ the "Gamification" approach to attract contributors to validate the proposed enhancement scenarios.

4. CONCLUSIONS

This paper proposed the integration between various VGI sources toward improved data quality. We concerned with voluntary mapping projects particularly the OSM project. Although the OSM represents a potential data source for various applications, some applications can not make use of it due to its questionable data quality. In OSM, incompleteness and limited LOD restrict the beneficial use of data. However, the problem can be tackled by integration with other available VGI resource.

We proposed an approach that utilizes geotagged photos from Flickr to enhance the OSM data. The approach is called TME, which stands for *Tagging*, *Matching*, and *Enhancement*. In TME, we utilized deep learning (DL) to carry out what we called deep tagging process. Given a geotagged photo, DL is applied to generate tags describing the given scene from various perspectives. Based on the given photo geolocation and the generated tags, the matching process acts to find counterpart geographic features in the OSM. Finally, in the enhancement process TME either ensure or enrich the OSM data. In the assurance scenario, we confirm the existence contributions in the OSM by linking them to additional pictorial information. Whereas in the enrichment scenarios additional information or a new contribution is added to the OSM.

This paper presented the preliminary idea and discussions of the TME approach. The paper included illustrations of the potential enhancement scenarios of the OSM. In future work, we will search for possible solutions for the aforementioned implementation challenges. Subsequently, we plan to implement the approach on concrete case studies. Finally, an intuitive GUI (e.g., web service) will be developed to visualize the results and to serve the validation process.

Aknowledgment

This work is partially funded by the German Research Foundation (DFG) through the projects Human-centered Relational Feature Classification for VGI (FR 806/17) in the Priority Programme 1894 of VGIScience (`https://vgiscience.org/`).

5. REFERENCES

[1] Deep image: Scaling up image recognition. *CoRR abs/1501.02876* (2015).

[2] Ali, A. L. Tackling the thematic accuracy of areal features in openstreetmap. In *European Handbook of Crowdsourced Geographic Information*. Ubiquity Press, 2016, p. 113.

[3] Ali, A. L., Sirilertworakul, N., Zipf, A., and Mobasheri, A. Guided classification system for conceptual overlapping classes in openstreetmap. *ISPRS International Journal of Geo-Information 5*, 6 (2016), 87.

[4] Andrienko, G., Andrienko, N., Bak, P., Kisilevich, S., and Keim, D. Analysis of community-contributed space-and time-referenced data (example of flickr and panoramio photos). In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology (VAST 2009)* (2009), IEEE, pp. 213–214.

[5] Antoniou, V., Morley, J., and Haklay, M. Web 2.0 geotagged photos: Assessing the spatial dimension of the phenomenon. *Geomatica 64*, 1 (2010), 99–110.

[6] Ballatore, A., Bertolotto, M., and Wilson, D. C. Geographic knowledge extraction and semantic similarity in openstreetmap. *Knowledge and Information Systems 37*, 1 (2013), 61–81.

[7] Chen, J., and Zipf, A. Deepvgi: Deep learning with volunteered geographic information. In *Int. World Wide Web Conference 2017 (WWW2017)* (2017).

[8] Ciregan, D., Meier, U., and Schmidhuber, J. Multi-column deep neural networks for image classification. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (2012), pp. 3642–3649.

[9] Conneau, A., Schwenk, H., Barrault, L., and LeCun, Y. Very deep convolutional networks for natural language processing. *CoRR abs/1606.01781* (2016).

[10] Freksa, C. Qualitative spatial reasoning. In *Cognitive and linguistic aspects of geographic space*. Springer, 1991, pp. 361–372.

[11] GIScience Research Group. DeepVGI – Deep Learning Volunteered Geographic Information - Combining OpenStreetMap, MapSwipe and Remote Sensing, 2017. Last checked: 24-02-2017.

[12] Haklay, M., and Budhathoki, N. Openstreetmap–overview and motivational factors. *Proceedings of communication au Horizon Infrastructure Challenge Theme Day, Université de Nottinghan* (2010).

[13] Hecht, B., and Stephens, M. A tale of cities: Urban biases in Volunteered Geographic Information. In *Proceedings of the 8th International Conference on Weblogs and Social Media (ICWSM)* (Michigan, USA, 2014).

[14] Hecht, R., Kunze, C., and Hahmann, S. Measuring completeness of building footprints in openstreetmap over space and time. *ISPRS International Journal of Geo-Information 2*, 4 (2013), 1066–1091.

[15] Hollenstein, L., and Purves, R. Exploring place through user-generated content: Using flickr tags to describe city cores. *Journal of Spatial Information Science*, 1 (2014), 21–48.

[16] Karpathy, A., and Li, F. Deep visual-semantic alignments for generating image descriptions. *CoRR abs/1412.2306* (2014).

[17] Kennedy, L., Naaman, M., Ahern, S., Nair, R., and Rattenbury, T. How flickr helps us make sense of the world: context and content in community-contributed media collections. In *Proceedings of the 15th ACM international conference on Multimedia* (2007), ACM, pp. 631–640.

[18] Krasin, I., Duerig, T., Alldrin, N., Veit, A., Abu-El-Haija, S., Belongie, S., Cai, D., Feng, Z., Ferrari, V., Gomes, V., Gupta, A., Narayanan, D., Sun, C., Chechik, G., and Murphy, K. Openimages: A public dataset for large-scale multi-label and multi-class image classification. *Dataset available from https://github.com/openimages* (2016).

[19] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, p. 2012.

[20] Liu, S. B., Palen, L., Sutton, J., Hughes, A. L., and Vieweg, S. In search of the bigger picture: The emergent role of on-line photo sharing in times of disaster. In *Proceedings of the Information Systems for Crisis Response and Management Conference (ISCRAM)* (2008).

[21] Matsugu, M., Mori, K., Mitari, Y., and Kaneda, Y. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks 16* (2003), 555 – 559. Advances in Neural Networks Research: {IJCNN} '03.

[22] Mooney, P., Minghini, M., Laakso, M., Antoniou, V., Olteanu-Raimond, A.-M., and Skopeliti, A. Towards a protocol for the collection of vgi vector data. *ISPRS International Journal of Geo-Information 5*, 11 (2016).

[23] Murthy, V. N., Maji, S., and Manmatha, R. Automatic image annotation using deep learning representations. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval* (New York, NY, USA, 2015), ICMR '15, ACM, pp. 603–606.

[24] Pourabdollah, A., Morley, J., Feldman, S., and Jackson, M. Towards an authoritative OpenStreetMap: conflating osm and os opendata national maps' road network. *ISPRS International Journal of Geo-Information 2*, 3 (2013), 704–728.

[25] Roemheld, L. Humanitarian mapping with deep learnin.

[26] Spyrou, E., and Mylonas, P. A survey on flickr multimedia research challenges. *Engineering Applications of Artificial Intelligence 51* (2016), 71–91.

[27] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S. E., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. Going deeper with convolutions. *CoRR abs/1409.4842* (2014).

[28] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. Rethinking the inception architecture for computer vision. *CoRR abs/1512.00567* (2015).

[29] Touya, G. What is the level of detail of openstreetmap. In *Workshop on role of volunteered geographic information in advancing science: quality and credibility. Columbus (Ohio), USA* (2012).

[30] TrailBehind, I. Deeposm. Last checked: 24-02-2017.

[31] Wu, J., Yu, Y., Huang, C., and Yu, K. Deep multiple instance learning for image classification and auto-annotation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015), pp. 3460–3469.

[32] Zielstra, D., and Zipf, A. A comparative study of proprietary geodata and volunteered geographic information for germany. In *13th AGILE international conference on geographic information science* (2010), vol. 2010.