Google Research Awards Proposal

1. Overview

Title: **Mapping Graffiti from Google Street View Data**

Faculty: **Ricardo Matsumura Araujo**
Center for Technological Development - Federal University of Pelotas (UFPel)
Rua Santa Cruz, 1796/1002, Pelotas, RS, 96015710, Brazil
ricardo@inf.ufpel.edu.br +55 53 9102-1171

2. Proposal

Abstract

We aim at using Google Street View data to automatically map graffitis in a given city. We propose to use Convolutional Neural Networks to detect and locate graffitis in Street View images. The extracted features will be used by an unsupervised learning algorithm in order to cluster graffiti styles and provide more useful sociological information on the resulting map.

Problem Statement and Research Goals

Graffiti are writing and drawings painted on surfaces, such as walls and doors, often used to express political and social messages and create group identification (tagging) [6]. Individually they can provide information about groups, ideas and the general *zeitgest* of a local population. However, as an urban phenomena, the geographic relationship between multiple graffiti can provide valuable information to better understand the social dynamics of a city, which can help with better urban planning, law enforcement and shed light into sociological and anthropological behaviors [7,9].

The central problem in this project is that of mapping different graffiti in urban areas from readily available geo-localized photographs of city streets. Our main goal is to test and develop machine learning algorithms that are able to use Google Street View data, which provide panoramic views along streets of several cities in the world, in order to extract graffiti information to be overlayed over digital maps, such as Google Maps. The information to be extracted is primarily the presence or absence of graffiti but we also aim at grouping graffiti according to their style or content in order to present a more detailed map allowing for group identification and tagging analysis.

Work Description and Expected Outcomes

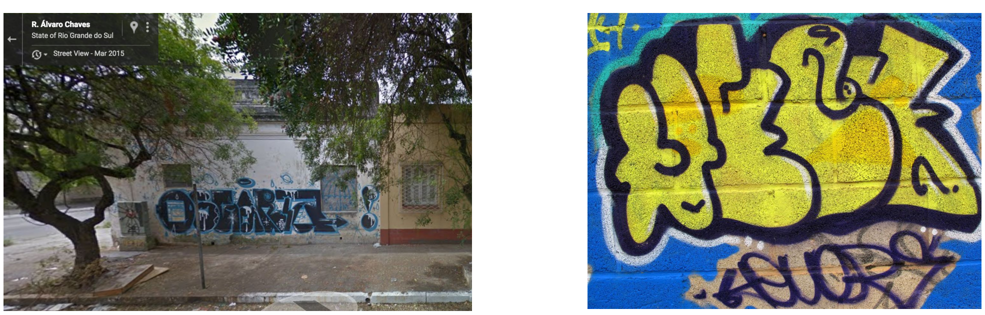
To identify the presence of graffiti we treat the problem as a boolean classification problem with images as inputs. We propose to use Convolutional Neural Networks (CNN), as they provide the current state-of-the-art technique for many such tasks [2]. In order to do so, we need large training, validation and test sets consisting of positive examples (images containing graffiti - e.g. Fig. 1a) and negative examples (images without graffiti).

We will use Street View Services from Google Maps API in order to obtain the base images. We plan on collecting images from two major cities in Brazil where there is prior knowledge on most likely locations to obtain graffiti images. These images will be labeled using crowdsourcing systems (e.g. Mechanical Turk), with additional validation by the researchers. However, this approach will likely return a relatively small number of positive examples. To attain a higher number of such examples, we plan on experimenting with expanding the data set with tagged images from other sources, such as Yahoo's Flickr (Fig. 1b). While these alternative sources provide images that can be considerably different from the base images, we hypothesize they are useful to build the feature extraction layers of CNNs.

Validation and test sets will be exclusively composed of images from Street View. We will experiment with different network architectures and also plan on using pre-trained CNNs, such as AlexNet and GoogLeNet trained on Imagenet, by using only the feature extracting layers feeding another model trained using our data sets. The performance of the networks will be tracked using several metrics, in particular specificity, sensitivity and AUC. Care will be taken so as not to overfit the model, in particular by the use of a hold-out test set and possibly nested cross-validation.

The second problem is to cluster graffiti by their feature similarities. In order to do so, we will first extract the graffiti from their surroundings in the images, by creating a bounding box around each one. Several techniques can be used (e.g. [4]) and we will experiment with them to obtain the best possible results. With the extracted regions and after possibly some pre-processing stage to clean noise and artifacts, clustering algorithms will be applied over features of these images. K-means will be used as the baseline algorithm and we will compare it to more informative approaches, in particular Gaussian mixture models [5]. As for the features to be used, a promising approach is to use features learned by the CNN in the previous problem.

We expect at the end of this project to have effective models able to identify and cluster graffiti in images. Minimally these models will work well in the selected cities, but we hope to be able to make the models generalise to any city and graffiti style. By doing so we will have contributed to the growing Deep Learning literature with a novel practical application and ways to solve this particular problem. Moreover, these models will be used to build a prototype web service that, given a city, creates an overlay over the city's map with the information extracted from Street View images of this city.



(a) (b)

Figure 1: Example of images containing graffiti: (a) extracted from Google's Street View and (b) extracted from a search in Yahoo's Flickr

Related Work

Geo-localization of graffiti is typically performed using reports from concerned citizens [7] or crowdsourcing efforts [8,9], which introduce bias from the reporters and severely limits scalability and applicability to different regions and countries.

Current attempts at recognizing graffiti automatically from images are limited to matching known patterns (e.g. gang signatures [3]) or identifying text-based graffiti [1,4]. As far as our knowledge goes, there are no attempts at exploiting geo-localized images to identify graffiti in a more general level, nor attempts at grouping them by stylistic features.

References

[1] A. Parra, B. Zhao, J. Kim, and E. J. Delp, "Recognition, segmentation and retrieval of gang graffiti images on a mobile device" *2013 IEEE Int. Conf. Technol. Homel. Secur. HST 2013*, pp. 178–183, 2013.

[2] S. Sudholt and G. A. Fink, "PHOCNet : A Deep Convolutional Neural Network for Word Spotting in Handwritten Documents," 2016.

[3] W. Tong, J.-E. Lee, R. Jin, and A. Jain, "Gang and Moniker Identification by Graffiti Matching," in *Proceedings of the 3rd international ACM workshop on Multimedia in forensics and intelligence*, 2011, pp. 1–6.

[4] C. Yang, P. C. Wong, W. Ribarsky, and J. Fan, "Efficient graffiti image retrieval" *Proc. 2nd ACM Int. Conf. Multimed. Retr. - ICMR '12*, vol. 1, no. c, p. 1, 2012.

[5] Verbeek JJ, Vlassis N, Kröse B. Efficient greedy learning of Gaussian mixture models. Neural computation. 2003;15(2):469-85.

[6] Ferrell J. Crimes of style: Urban graffiti and the politics of criminality. New York: Garland; 1993.

[7] Haworth B, Bruce E, Iveson K. Spatio-temporal analysis of graffiti occurrence in an inner-city urban environment. Applied Geography. 2013 Mar 31;38:53-63.

[8] Burghardt M, Hertlein F, Hinterleitner B, Lehenmeier C, Spröd T. A Crowdsourced Approach for the Documentation and Transcription of Graffiti in Public Restrooms. InISI 2015 (pp. 160-164).

[9] Megler V, Banis D, Chang H. Spatial analysis of graffiti in San Francisco. Applied Geography. 2014 Oct 31;54:63-73.

3. Data Policy

All developed code and data sets will be made available at GitHub or another appropriate public repository. All papers derived from this work will be published preferentially in open-access journals.