OBSERVATÓR!O2016

Abstract

This paper focus on OBSERVATÓR!O2016, a web-based platform for collecting, structuring and visualizing the online response to Rio 2016 Olympic Games from content shared in Twitter. This project was developed at VISGRAF Laboratory and is based on two cross related research lines. First, the conception and design of OBSERVATÓR!O2016 website (http://oo.impa.br), which provides a space to explore comments and images about the Olympics through structured visualizations. Second, the ongoing deployment of a research (http://lvelho.impa.br/dl_rio2016/index.html) regarding the application of a Convolutional Neural Network (CNN) into the 180,000 images of our database. In this paper, we explore how CNN models can be used in image analysis in the context of Digital Humanities. The goal is to highlight how we are using deep learning approaches such as automatic image classification to enhance discoverability and expression of image similarities and subject features within a image collection.

Context

The Olympic Games for the first time took place in a country in South America. The realization of a worldwide sport event gathers attention from both media and online audience. It also involves multiple expectations and implications that reflect diverse impressions and moods shared on social networks. As a consequence, this sporting event becomes a social phenomena that produces a heterogeneous debate. But, how to explore, understand and communicate this debate that is spread in non-structured data streams?

Understanding the historical and social importance of Rio-2016 to Brazil, the VISGRAF Laboratory - part of Instituto de Matemática Pura e Aplicada (IMPA) - conceived the project OBSERVATÓR!O2016, a multidisciplinary research interested in collecting, structuring and visualizing comments and images about the Olympics shared on social networks, mostly from Twitter. This kind of work is set in a line of research that seeks to understand big social data and give it meaning through visualizations.

OBSERVATÓR!O2016 collected around 1 million Twitter messages, via Twitter's public API, from 18 April to 25 August 2016. Approximately 180,000 of these tweets included unique images that were also stored in our database. In other to gather different perspectives of the debate about the Olympics we carefully created seven different Twitter search queries each one with its operators. More technical information about the IT framework can be found in the documentation website (http://www.visgraf.impa.br/rio2016/). As the project unfolded and the Games draw nearer, OBSERVATÓR!O2016 processed the data stream and visualized hashtags, textual messages and images looking for a better understanding of the audience response to Rio-2016. Data presentation were in eight main interactive visualizations.

Figure 1: Visualizations present structured information (text and images) related do the audience response to Rio-2016.

After the Olympics, the project continues to expand. OBSERVATÓR!O2016 website will be a digital memory of commentary and imagery about Rio-2016, shared on Twitter. Since the data collected are stored in their entirety, future data analysis and visualizations are possible. In this sense, a new research line is ongoing and involves deep learning approaches for attending computer vision tasks. Our goal is to investigate avenues for clustering and visualizing image features of Rio-2016 collection enhancing the discoverability and expression of the collection.

Deep Learning Methods

The production and circulation of large volumes of digital images on social networks inaugurate new avenues for researches interested in understanding the human creative practice. In this sense, the work of Lev Manovich on using computational and data visualization techniques to investigate visual culture is quite notable. If, on the one hand, access to this vast imagery repertoire is facilitated by advancements in visual data gathering and storage, on the other hand, access is not always sufficient to propose analysis and synthesis problems for images. Therefore, massive collections of visual artifacts still justify the need to explore intelligent mechanisms for identifying subject and stylistic features and clustering images not only from their textual metadata, but also from computer vision techniques. Aware of that, we decided to explore Rio-2016 images dataset through deep learning approaches. As the investigation is currently ongoing, we will report the research process of a single task, which is the *Torch Mosaic*. Nevertheless, the conduction of this initial study shows it can be unfolded in other tasks that are outlined in the final session of this paper.

Torch Mosaic

During the pre-Olympics, it became evident that many images that were being shared in Twitter and gathered by our query scripts were related to a specific subject: the Olympic torch relay. As a common iconographic element, many of these images depicted the Olympic torch in different contexts and scenarios. Part of these images were accompanied by texts that mentioned the torch in some way. But not all of them. In addition, some tweets relating the torch relay incorporated images that didn't depict the torch. In other words, text analysis and metadata filtering weren't enough to detect a set of images containing the torch.

Thus, we referred to a deep learning approach to visually recognize the Olympic torch in the images of our database. The field of visual pattern recognition has been recently improved by the efficient performance of deep learning approaches using Convolutional Neural Networks (CNN). In 2012, the work of Krizhevsky et al. on training a large, deep convolutional neural network to

classify the 1.2 million images in the *ImageNet* LSVRC-2010 contest into the 1000 different classes, had an important impact on the computer vision community.

More recently, and thanks to Googles' effort, computer vision tasks such as image classification and object recognition have become relatively more accessible and applicable. That's because the company released last year their open source software library called *TensorFlow*, which assists the development of a system for machine learning. This library runs code for image classification on *Inception-v3* CNN model, which demonstrates substantial gains over *AlexNet* CNN model: a top–5 error rate of 3.46% compared to *AlexNet* top-5 error rate of 15.3% (Szegedy et al, 2015).

Besides, *Inception-v3* can be retrained on a distinct image classification task. This quality is often referred to as transfer learning, and has been a topic of particular enthusiasm in computer vision because it shortcuts a lot of work by taking a fully-trained model for a set of categories and retrains from the existing weights for new classes. By creating a set of training images to teach the network, it's possible to update the parameters of the model and use it for a specific vision task such as recognizing a new image category. That said, we retrained the network to detect the Olympic torch by showing it a sample of one hundred manually labeled images containing this iconic object. Finally, the retrained network ran over our database and returned a set of images with their corresponding confidence score for the Olympic torch category.

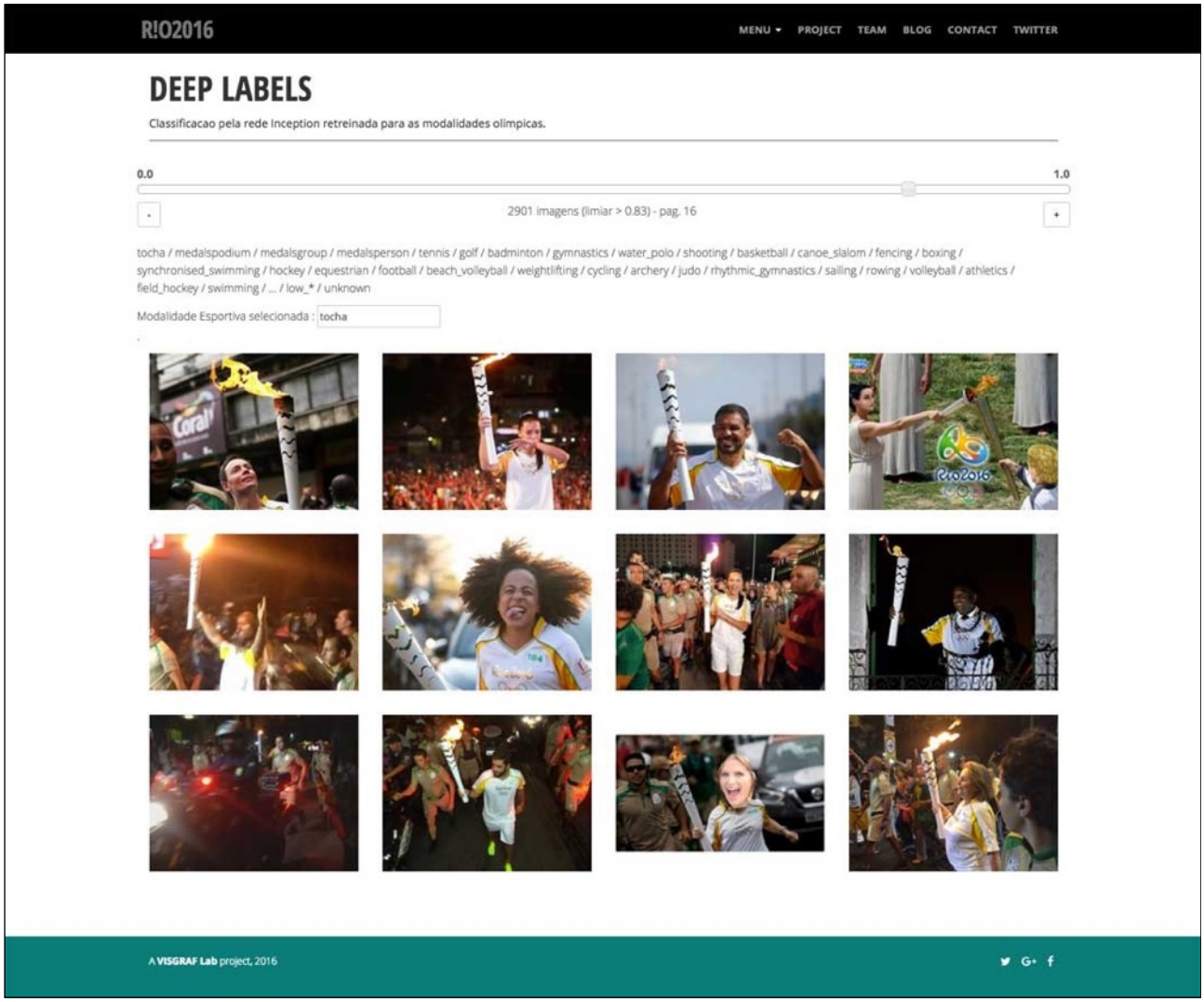

Figure 2: A subset of 12 from 2091 images with confidence score over 83% for the Olympic torch category.

Until June 25, around 1500 images with over 85% confidence score for the Olympic torch category had been classified by our network. We used them to create a mosaic image that can be zoomed

and panned (http://lvelho.impa.br/dl_rio2016/mosaico.html). The mosaic conceptual idea is that, given an image (target image), another image (mosaic) is automatically build up from several smaller images (tile images) (Finkelstein & Range, 1998). To implement the mosaic we used a web-based viewer for high-resolution zoomable images called *OpenSeadragon*.

Figure 3: The target image.

Figure 4: The mosaic.

Figure 5: Tile images.

Since the tile images are organized to compose the representation of a bigger image (the target image), we can't expect obvious visual patterns insights from this mosaic visualization technique. The nature of this visualization is mainly artistic and aesthetic. Nevertheless, zooming and panning the image, take you to a high resolution version where it's possible to explore details and discover surprises, such as the spoof picture of Fofão, a Brazilian fictional character, carrying the torch.

Figure 6: A spoof picture of the Olympic torch relay.

Perspectives

The convergence between the increasing access to visual large databases, progresses in computer vision techniques and, ultimately, the openness of documented high performance deep learning frameworks motivates the use of CNN for identifying visual patterns such as subject and stylistic features in the context of Digital Humanities. This paper reflects a initial study upon social images from Rio 2016 Olympic Games in which we are interested in enhancing discoverability and expression of the digital collection. We've already developed an analysis and representation solution for a specific subject feature - the Olympic torch relay - through deep learning and visualisation techniques. For DH2017 we expect to present the results for another Olympic subject feature - the sporting disciplines - and visualisation technique - the *videosphere* - we are working on at the moment.

Bibliography

Finkelstein, A., & Range, M. (1998). Image mosaics. In Electronic Publishing, Artistic Imaging, and Digital Typography (pp. 11-22). Springer Berlin Heidelberg.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-9).