OBSERVATÓR!O2016: exploring Rio-2016 image dataset through Deep Learning and visualization techniques

This poster focus on OBSERVATÓR!O2016, a web-based platform for collecting, structuring and visualizing the online response to Rio-2016 from content shared on Twitter. This project was developed at the Vision and Computer Graphics Laboratory (http://www.visgraf.impa.br) 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://visgraf.impa.br/dl_rio2016) regarding the application of a digital method to explore large sets of images. The goal is to highlight how we use *Deep Learning* applications - such as automatic image classification - and visualization techniques to enhance discoverability and expression of subject features within an extensive image collection.

Method: *Deep Learning* and mosaic visualization

OBSERVATÓR!O2016 collected around 1 million tweets from April 18th to August 25th, 2016. In other to gather different perspectives on the debate about the Olympics we created seven Twitter search queries. Data presentation were in eight interactive visualizations:

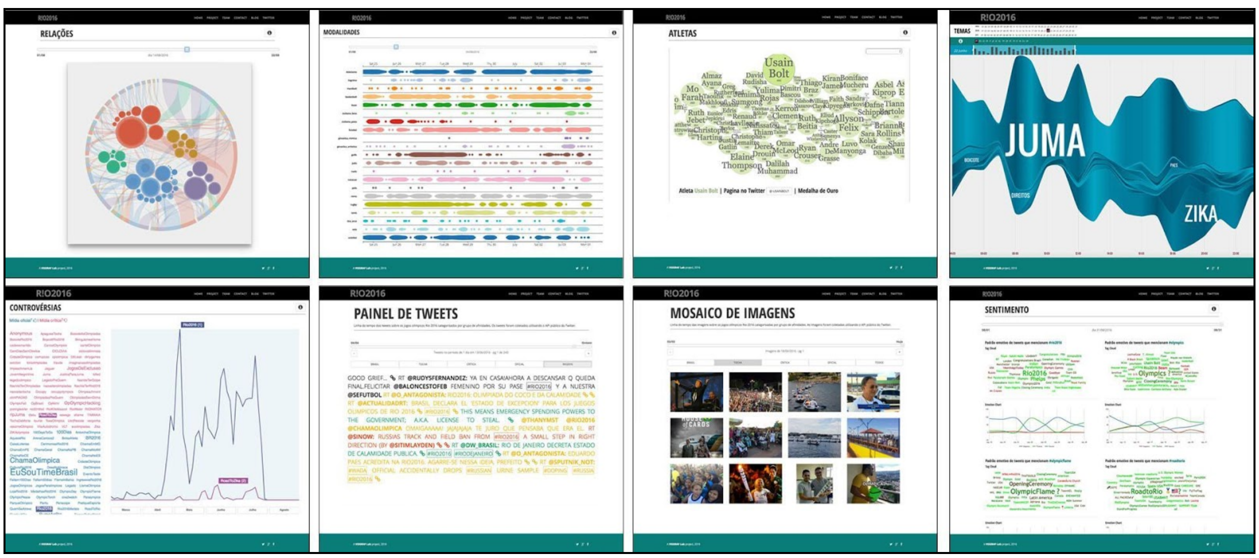

Figure 1. Visualizations

Approximately 180,000 of the collected tweets included unique images. The production of large sets of digital images inaugurates new avenues for researchers interested in the human creative practice. In this sense, the investigation of Lev Manovich on digital methods to study visual culture is quite relevant. Aware of that, we explored Rio-2016 images through *Deep Learning* approaches. As the investigation is currently ongoing, we'll report the research process of a single task, which resulted in the *Torch Mosaic* visualization.

During pre-Olympics, it became evident that many of the images gathered by our query scripts were related to the Olympic torch relay and depicted the iconic object. Part of these images were accompanied by texts that mentioned the torch, but not all. In addition, some tweets mentioning the torch relay incorporated images that didn't depict the object. In other words, text analysis wasn't enough to detect a set of images containing the torch.

Thus, we referred to a *Deep Learning* approach to recognize the Olympic torch in our database. The field of visual pattern recognition has been recently improved by the efficient performance of *Convolutional Neural Networks* (CNN). In 2012, the work of Krizhevsky et al. on training a deep convolutional neural network to classify the 1.2 million images in the *ImageNet* LSVRC-2010 contest into 1000 different classes, had a substancial impact on the computer vision community.

More recently, and thanks to Google, computer vision tasks such as image classification have become more accessible and applicable. That's because the company released last year their open source software library *TensorFlow.* This library runs code for image classification on *Inception-v3* CNN model, which can be retrained on a distinct image classification task (this quality is referred to as *transfer learning*). By creating a set of training images, it's possible to update the parameters of the model and use it to recognize a new image category. That said, we retrained the network by showing it a sample of 100 manually labeled images containing the torch. Finally, the retrained network ran over our database and returned a set of images with their corresponding confidence score for the new category.

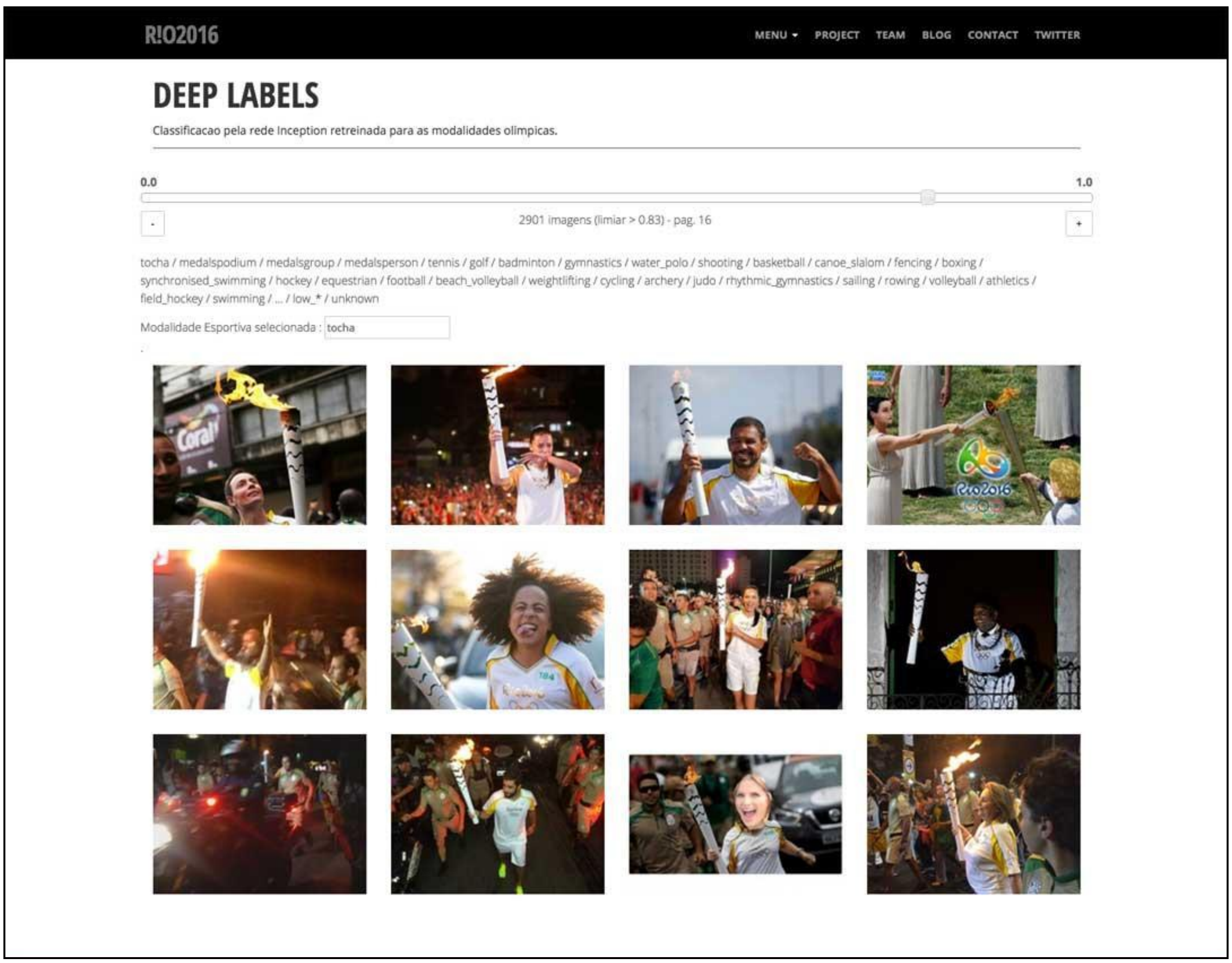

Figure 2. Confidence score over 85%

By June 25th, 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 visualization that can be zoomed and panned (http://lvelho.impa.br/dl_rio2016/mosaico.html). The mosaic idea is that, given an image (target image), another image (mosaic) is automatically build up from several smaller images (tile images). To implement the mosaic we used a web-based viewer for high-resolution zoomable images called *OpenSeadragon*.

Figure 3. The target image

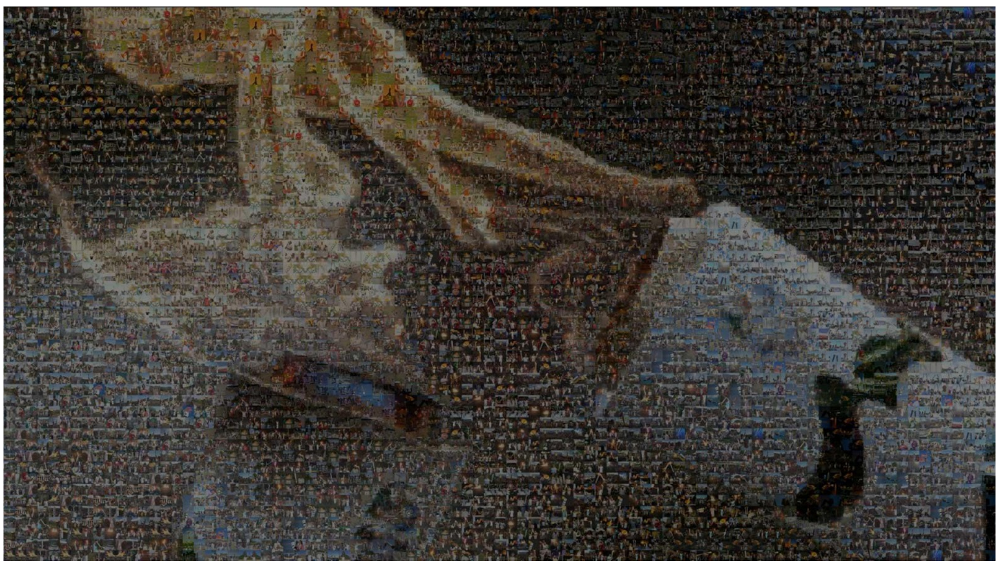

Figure 4. The mosaic

Figure 5. Tile images

The organization nature of the mosaic visualization is mainly aesthetic. Nevertheless, zooming and panning the mosaic, allows the user to explore a wide variety of views and discover image details and surprises, such as the spoof picture of *Fofão*, a Brazilian fictional character, carrying the torch.

For DH2017 poster session, we expect to present the results for another subject feature - the sporting disciplines - and visualisation technique - the *videosphere* - we are working on at the moment. Besides, we envision to discuss possible scenarios with participants in which *Deep Learning* models could be applied to help image collections become more discoverable and expressive.

Bibliography

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

Manovich, L (2012). "How to Compare One Million Images?”. In Berry, D. (ed), *Understanding Digital Humanities.* Palgrave, pp. 249-278.

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