How Machines See the World: Understanding Image Labelling

Carloalberto Treccani

School of Creative Media, City University of Hong Kong

carloalberto.t@my.cityu.edu.hk

Abstract

Michael Baxandall, in *Painting and Experience in 15th Century Italy* (1988), shows the existence of a series of rules that painters were advised to follow. These "guidelines" explained, for instance, how each different figure or hand position painted, within that specific cultural context, represented a different concept. These rules helped the painter maintain relevance in that historical and cultural context. [1]

Today, more than 500 years after the Renaissance Italy described by Baxandall, companies all around the world are trying to teach Machines and Algorithms (M/A) to see and understand what they see (image recognition). However, this process of signification, simple for a human being, is still complex for M/A. Therefore hundreds of thousands of workers, therefore, are hired, through crowdsourcing platform, in order to *label* what they see. An image of a house appears on the monitor and the worker then attributes the "house" label to that image. These images are then categorized by the received label, or semantic area, and then collected in databases which are used to train M/A. [2]

However, this labelling process produces a series of problems. The workers are paid in pennies per image labelled and work in precarious working conditions without any labor protection. [3] Sometimes the *annotators* are required to label unknown scenes or objects (e.g., objects and tools in a physics laboratory) even when they lack the competence or knowledge. Moreover, if the employer considers their work unsatisfactory, payment can be denied without any explanation. [4] All these different reasons often cause insufficient and confusing labelling. Yet these "low quality" labels are determining the way M/A understands the world.

Furthermore, every time we make a click on internet, on social media we are not only conveying some information, but also engaging and establishing a pedagogical process. We are not only viewers and users, instead, we are teaching M/A how to look at the world. [5]

Given this context, I would like to address some questions as follows: What are the consequences of a learning process that is confused, inaccurate, and qualitatively poor, in this unprecedented historical moment where there are more M/A than human beings examining and trying to create sense of what they see? What are the implications of this low quality work, which does not appear today as an image but instead as labelled data, which in turn contributes to fully defining the visual experience of these M/A? [6]

References

1. Michael Baxandall, *Painting and Experience in 15th Century Italy* (Oxford: Oxford University Press, 1988).

2. Alexander Sorokin & David Forsyth, "Utility Data Annotation with Amazon Mechanical Turk". *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops* (2008). doi:10.1109/cvprw.2008.4562953.

3. Nicholas Malevé, "The politics of image search - A conversation with Sebastian Schmieg" [part I and II], The photographers gallery website, accessed February 15, 2018. https://unthinking.photography/themes/intervie ws/interview.

4. Treccani Carloalberto, "How Machines See the World: Understanding Image Annotation",

NECSUS_European Journal of Media Studies, no. Spring 2018_#Resolution, accessed October 12, 2018. https://necsus-ejms.org/how-machines-see-the-world-understanding-image-annotation/.

5. Nicholas Malevé, "The politics of image search - A conversation with Sebastian Schmieg" [part I and II], The photographers gallery website, accessed February 15, 2018. https://unthinking.photography/themes/interviews/interview.

6. Trevor Paglen, "Invisible Images (Your Pictures Are Looking at You)", The new inquiry website, accessed May 01, 2017. https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/.

Biography

Carloalberto Treccani is a PhD candidate at the School of Creative Media, City University of Hong Kong, and an artist. His research investigates how machine vision is affecting human vision. More broadly he is interested in how technology affects human perceptions and emotions. His artworks have been exhibited in group and solo exhibitions and commissioned by galleries and institutions.