Extracting theory from black boxes: Using machine vision APIs in communication research

Araujo, T.; Lock, I.; van de Velde, B.

Published in:
IEEE 14th International Conference on eScience

DOI:
10.1109/eScience.2018.00068

2018 IEEE 14th International Conference on e-Science

Extracting Theory from Black Boxes: Using Machine Vision APIs in Communication Research

Theo Araujo
Amsterdam School of Communication Research (ASCoR)
University of Amsterdam
Amsterdam, Netherlands
t.b.araujo@uva.nl

Irina Lock
Amsterdam School of Communication Research (ASCoR)
University of Amsterdam
Amsterdam, Netherlands
I.J.Lock@uva.nl

Bob van de Velde
Informatics Institute
University of Amsterdam
Amsterdam, Netherlands
r.n.vandevelde@uva.nl

***Abstract*— The increasing volume of images published digitally requires social science and communication researchers to employ methods able to perform visual content analysis at a large scale. Ongoing advances in machine vision and the ability to automatically detect objects, concepts and features in images provide a promising opportunity to address this challenge, yet it is often not feasible for social science researchers to develop their own custom classifier given the volume of images, resources and technical expertise needed. We therefore propose a research protocol with which existing pre-trained (commercial) models can be used for theory-building purposes despite their black box approach.**

Keywords — Machine Vision, Theory-driven research, Commercial APIs, Large-scale Content Analysis

I. Introduction

The increasing volume of images published online requires communication and social science researchers to employ methods able to perform visual content analysis at a large scale. Ongoing advances in machine vision, especially the ability to detect objects, concepts and features in images, provide a promising opportunity to address this challenge. Communication research using these new methods, however, is still scarce and, given the volume of images needed for algorithm training and the resources and technical expertise necessary [1], it is often not feasible for social science researchers to develop their own custom classifier. Moreover, such a specific tool would only be applicable to one narrow research purpose, which for social science research is disproportionate to the investment needed for development.

Therefore, it seems efficient to refer to adapt machine vision APIs either offered commercially or as pre-trained models. Such pre-trained models can be deployed for communication research with supposedly little to no manual coding. However, researchers face a difficulty: they often cannot access the classification and basic labeling structure of these classifiers, and thus have to cope with results from a black box. In addition, many of these models were not trained with social research purposes in mind, and may therefore not fit directly with the theories and concepts relevant to communication researchers, either because of task (e.g. recognizing common objects), domain (e.g.: website pictures) and genre (e.g.: holiday pictures) differences. We use Corporate Social Responsibility (CSR) communication in corporate websites as a case study.

This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperation (HPC Cloud).

II. Objectives and Related work

To overcome these issues, we propose a research protocol which enables social scientists to use pre-trained models for their research. In this sense, our work complements earlier research that currently makes use of pre-trained models (e.g., [2], in the context of brand-related social media posts) by proposing a protocol to be used when interpreting the output of pre-trained machine vision models. Moreover, our proposal also extends corporate communication research in corporate websites, which has focused primarily on text [3], or employed manual content analysis for the categorization of images [4]. This opens, therefore, the door to other modes and multimodal forms of communication.

III. Protocol and Approaches

We propose a research protocol that involves the development of a theory-driven codebook encompassing the (latent and manifest) variables that are interesting for the study. After the development of this codebook, a random sample of images is manually annotated by human coders. This manually annotated sample is then considered as the golden standard that should be achieved by automated labelling. Given the wide variability and complexity of (latent and manifest) variables that may be theoretically interesting for social science researchers, we propose three approaches to use the output of pre-trained models as an input for a classification task having the manually annotated sample as a target: (1) Expert-driven approach (in which an expert associates the labels provided by the pre-trained models with the categories in the codebook), (2) Supervised models (in which supervised machine learning algorithms are used, with the labels provided by the pre-trained models as features, and the manual annotations as targets), and (3) Unsupervised models (in which, for example, LDA models are used with the labels provided by the pre-trained models as input).

IV. Preliminary Results

A random sample of 900 images extracted from the corporate websites of top 25 European companies was manually coded for key CSR topics (People, Planet or Profit), and pre-trained models from Clarifai, Google and Microsoft were tested. The expert-driven approach had average F1-scores above 0.7 (People, Planet) or 0.6 (Profit), whereas the supervised approach had average F1-scores above 0.8 (People, Planet) or 0.7 (Profit). When used as input for supervised classifications, LDA topic models scored in average better than the expert or the supervised approach, while having lower F1-scores and validity when attempting to predict the three categories at the same time.

978-1-5386-9156-4/18/$31.00 ©2018 IEEE
DOI 10.1109/eScience.2018.00068

REFERENCES

[1] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015.

[2] M. Mazloom, R. Rietveld, S. Rudinac, M. Worring, and W. van Dolen, "Multimodal Popularity Prediction of Brand-related Social Media Posts," 2016, pp. 197–201.

[3] D. L. Gill, S. J. Dickinson, and A. Scharl, "Communicating sustainability: A web content analysis of North American, Asian and European firms," J. Commun. Manag., vol. 12, no. 3, pp. 243–262, Jul. 2008.

[4] P. W. Braddy, A. W. Meade, and C. M. Kroustalis, "Organizational recruitment website effects on viewers' perceptions of organizational culture," J. Bus. Psychol., vol. 20, no. 4, pp. 525–543, Jun. 2006.