Selected Papers of #AoIR2021:
The 22nd Annual Conference of the
Association of Internet Researchers
Virtual Event / 13-16 Oct 2021

DEEP NEURAL NETWORKS FOR SOCIAL VISUALS: STUDYING CLIMATE COMMUNICATION ON YOUTUBE

Matteo Magnani
InfoLab, Uppsala University

Alexandra Segerberg
Uppsala University

Visual politics is becoming increasingly prominent, in particular in online environments that are primed for visual content. Visuals are persuasive, affective, and may convey different information than text. As visuals become digitally augmented—replicated, modified and shared online—it is increasingly important to study what they convey and do: their content, social negotiation, and the role of platforms in their prominence and spread.

However, research on visual politics has typically employed qualitative methods in small scale settings (e.g., focused on national newspapers in selected western countries), with approaches that do not easily expand to complex media ecologies and to questions with a global aspect, in which it matters how communication on an issue (e.g., climate change) or a communicative style (e.g., populism) is characterised, perceived and distributed across particular populations or across different parts of the world. A stumbling block has been the inability to automate the analysis of visual communication on the basis of actual images. Now, advances in deep learning in computer vision and image analysis open a potential path to large-scale analysis of visual politics on the basis of actual visual content. Promising work is starting to develop this potential in a variety of directions (see for example, Casas & Webb Williams 2019; Joo et al., 2019; Peng 2021; Zannettou et al, 2019).

Harnessing deep neural networks in this context has the potential to make important contributions to understanding contemporary political communication. Nevertheless, it can also easily lead to misleading results. It is challenging to robustly integrate deep learning into the study of politics, and a particularly difficult area concerns the analysis of social visuals. In contrast to other application areas for neural networks more focused on object detection, visuals of social and political scenes are semantically rich and can convey complex ideas and narratives about power and politics (Brennen et al., 2021).

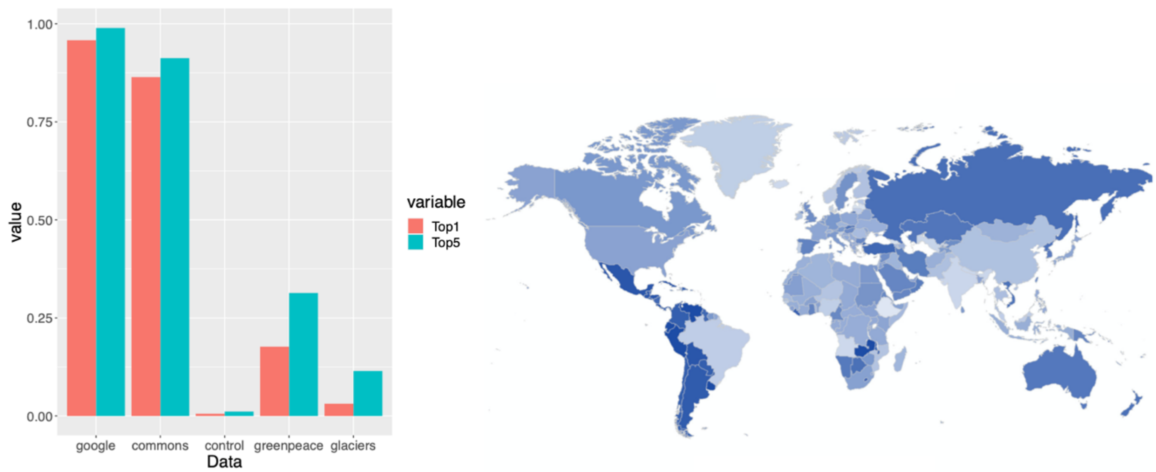
Suggested Citation (APA): Magnani, M., Segerberg, A. (2021, October). Deep neural networks for social visuals: studying climate communication on YouTube. Paper presented at AoIR 2021: The 22nd Annual Conference of the Association of Internet Researchers. Virtual Event: AoIR. Retrieved from http://spir.aoir.org.

This paper examines validity conditions for integrating deep neural network tools in the study of digitally augmented social visuals. We argue that the complexity of social visuals needs to be reflected in the validation process and its communication. With "social visuals" we refer to the social and political themes present in visuals, which we in this context infer empirically from objects (denotational content of the images) and labels (annotations recognised by the deep neural networks). A key validity concern in neural network research focuses on designing the structure of networks that can accurately replicate the relation between some input data and an output classification. In this case, the data is assumed to be unproblematic and the focus is on building a valid classification tool. A second key validity concern focuses on the challenges of obtaining good data (e.g., correct and unbiased). In this case, even an accurate tool can give invalid results when it is executed on *wrong data*. Here we go beyond this dichotomy and focus on the interdependency between data and tool. We note that *good data* can be defined as such only with respect to it being representative of the themes (classes) that we want to identify. However, when we analyse social visuals, the operationalisation of the themes may develop while we are preparing the final analysis, as the processing of new content may lead to the revision of our analytical frame – so that a final definition of good data is not available until the end of the process, which itself relies on a tool which needs good data to be trained. Some new themes may even emerge during the analysis. In addition, the operationalisation may develop not just because of our interaction with the data, as they would in a small-scale, qualitative analysis, but also because of our interactions with the tool and the specific way in which it mediates our analysis. Here, the final empirical operationalisation is an artifact produced by this hybrid methodology involving the iterated interaction between researcher and computational tool. An important implication is that the conventional approach of performance assessment – i.e., counting errors – while a necessary validation baseline, is potentially misleading if used on its own. At the same time, the validation of the tool can also be used to better understand the tool itself and help the researcher to refine the analytical process.

We explore our argument experimentally in the context of an ongoing study that asks whether the tendency to frame climate change as a remote problem is a global trend on YouTube. While it does not exercise conventional editorial control, its rankings matter for the visibility of contents on the site. This platform is a common source of science information, and has been criticized for amplifying science skeptics (although the evidence is mixed) (Avaaz 2020; see also Allgaier 2019, Bounegru et al. 2020). Iconic (but distancing) climate themes such as the *polar bear* in arctic landscapes and focus on *elite* people and events (e.g., the COP summits) involve tough cases of social visuals.

For example, in Figure 1a we show the accuracy of a pre-trained network used to identify polar bears in YouTube video frames. A first analysis of performance on plausible datasets (google search, wikimedia commons, control data without polar bears) indicated a high accuracy of our classifier. However, testing on a different source (Greenpeace campaign pictures about polar bears) highlighted that the first datasets were capturing only a specific instance of the polar bear theme. The qualitative analysis of all errors also allowed us to generate hypotheses about different types of *polarbearness* (animals vs. activism vs. arctic environments). Adding one more dataset, we could corroborate our hypothesis about surroundings, showing that the network was recognizing pictures of glaciers as polar bears.

A recurrent point throughout is the imperative that the uncertainties of the process are systematically communicated, not to undermine the quality of the results but to make them more precise. We use as starting point a possible representation of the distribution of YouTube *polarbearness* around the world, shown in Figure 1b. We use this example to discuss how to qualitatively and quantitatively communicate validity questions.



(a) (b)

Figure 1: Two steps in a large-scale analysis of visual politics: (a) tool validation and explanation (top-1 and top-5 recall of the resnet50 network on five test datasets); (b) presentation of results (relative frequency of top-50-relevant YouTube videos showing polar bears in the wild, with search term "climate change" translated in the official languages of each region)

References

Allgaier, J. (2019). Science and environmental communication on YouTube: Strategically distorted communications in online videos on climate change and climate engineering, *Frontiers in Communication*, 4:36.

Avaaz (2020). Why is YouTube Broadcasting Misinformation to Millions? YouTube is driving its users to climate misinformation and the world's most trusted brands are paying for it. *Avaaz report* 15/01/2020.

Brennen, J., Simon, F., & Nielsen, R. (2021). Beyond (Mis)Representation: Visuals in COVID-19 Misinformation, *The International Journal of Press/Politics*, 26(1): 277–299.

Bounegru, L., De Pryck, K., Venturini, T., & Mauri, M. (2020). "We only have 12 years": YouTube and the IPCC report on global warming of 1.5°C. *First Monday*, 25(2):1-20.

Casas, A., & Webb Williams, N. (2019). Images that Matter: Online Protests and the Mobilizing Role of Pictures. *Political Research Quarterly* 72(2): 360 - 375.

Joo, J., Bucy, E., & Seidel, C. (2019). Automated Coding of Televised Leader Displays: Detecting Nonverbal Political Behavior with Computer Vision and Deep Learning, *International Journal of Communication*, 13: 4044-4066.

Peng, Y. (2021). What Makes Politicians' Instagram Posts Popular? Analyzing Social Media Strategies of Candidates and Office Holders with Computer Vision, *The International Journal of Press/Politics*, 26(1): 143–166.

Zannettou, S., Caulfield, T. Bradlyn, B., De Cristofaro, E., Stringhini, G., Blackburn. J. (2020). Characterizing the Use of Images in State-Sponsored Information Warfare Operations by Russian Trolls on Twitter. *14th International AAAI Conference on Web and Social Media* (ICWSM).