SEEING HISTORY: ANALYZING LARGE-SCALE HISTORICAL VISUAL DATASETS USING DEEP NEURAL NETWORKS

Melvin Wevers, DHLAB, KNAW HUC, Amsterdam, the Netherlands
Thomas Smits, Radboud University, Nijmegen, the Netherlands

Scholars are increasingly applying computational methods to analyze the visual aspects of large-scale digitized visual datasets.[1] Inspiring examples are the work of Seguin on visual pattern discovery in large databases of paintings and Moretti's and Impett's large-scale analysis of body postures in Aby Warburg's Atlas Mnemosyne.[2] In our paper, we will present two datasets of historical images and accompanying metadata harvested from Dutch digitized newspapers and reflect on ways to improve existing neural networks for historical research. We will discuss how large historical visual datasets can be used for historical research using neural networks. We will do this by describing two case studies, and will end our paper by arguing for the need for a benchmarked dataset with historical visual material.

The sets were produced during two researcher-in-residence projects at the National Library of the Netherlands. The first set (SIAMESET) consists of more than 400,000 advertisements published in two major national newspapers between 1945 and 1995.[3] Using the penultimate layer in a Convolutional Neural Network (CNN), 2,048 visual aspects were abstracted from these images, which can be used to group images together.[4] The second dataset (CHRONIC) includes about 313,000 classified images from newspapers published between 1860 and 1922.[5] The images were classified using a pipeline that consists of three classifiers. The first one detects images with faces, the second categorizes images according to eight different categories (buildings, cartoons, chess problems, crowds, logos, schematics, sheet music, and weather reports), and the last one sorts images as either photographs or drawings.[6]

We developed two tools to query these datasets. The first tool (SIAMESE) offers exploratory search in the advertisement dataset, which enables users to find images sharing a degree of visual similarity and can be used to detect visual trends in large visual datasets.[7] For example, you can input an advertisement with a particular design, or perhaps a shape or object,

---

[1] Roeland Ordelman et al., "How to Integrate Audiovisual Material in Digital Humanities Research" (Digital Humanities 2014, Lausanne, 2014).
[2] Benoit Seguin, Isabella di Lenardo, and Frédéric Kaplan, "Tracking Transmission of Details in Paintings" (Digital Humanities 2017, Montreal, Canada, 2017), https://dh2017.adho.org/abstracts/562/562.pdf; Leonardo Impett and Franco Moretti, "Totentanz. Operationalizing Aby Warburg's Pathosformeln" (Stanford Literary Lab, 2017).
[3] Melvin Wevers and Juliette Lonij, "SIAMESET," KB Lab, 2017, http://lab.kb.nl/dataset/siameset.
[4] Isabella di Lenardo, Benoit Seguin, and Frédéric Kaplan, "Visual Patterns Discovery in Large Databases of Paintings." (Digital Humanities 2016, Krakow, Poland, 2016).
[5] Thomas Smits and Willem-Jan Faber, "CHRONIC," KB Lab, 2018, http://lab.kb.nl/dataset/chronic-classified-historical-newspaper-images.
[6] Adam Geitgey, *Face_recognition: The World's Simplest Facial Recognition Api for Python and the Command Line*, Python, 2018, https://github.com/ageitgey/face_recognition; Jeff Donahue et al., "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition," *ArXiv:1310.1531 [Cs]*, October 5, 2013, http://arxiv.org/abs/1310.1531 We would like to thank Leonardo Impett for his help on developing the classifier for photographs and illustrations. https://github.com/KBNLresearch/CHRONIC/tree/master/drawing_or_photo.
[7] Juliette Lonij and Melvin Wevers, "SIAMESE," KB Lab, 2017, http://lab.kb.nl/tool/siamese.

and the algorithm then looks for similar images. The second one (CHRONREADER) enables users to find images in the second set by searching for specific (combinations) of visual categories and keywords.[8] For example, images of 'buildings' with 'faces' and the keyword 'protest' in the text.

Our paper discusses several challenges and possibilities of computer vision techniques for historical research. Most CNN's are trained on contemporary materials (ImageNet).[9] As a result, these networks perform well in recognizing the categories of the ImageNet challenge. However, the fact that they were trained on contemporary data can cause problems when working with historical images. For example, detecting bicycles works relatively well because the design of the bicycle has remained more or less similar during the last century, while trains are much more difficult since they have changed significantly over the years. Also, models trained on ImageNet have difficulties detecting objects in illustrations, which are often used in newspapers. They are regularly classified within the uninformative category 'cartoon.' In short, we will discuss how to improve these models and argue for the development and benchmarking of datasets with visual historical material.

[8] Thomas Smits and Willem-Jan Faber, "CHRONReader," KB Lab, 2018, http://lab.kb.nl/tool/chronreader.
[9] Olga Russakovsky et al., "Imagenet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision* 115, no. 3 (2015): 211–252.