

Colleen Nugent

Ryan Muther

Jeffrey Sternberg

Lauren Bergnes Sell

Molly Nebiolo

HIST 7370

DPLA Photographs workset - Group Post #1

How do neural networks see digital pictures?

What are the algorithms good at and bad at?

What kinds of things should we be thinking about pulling out?

Determining exactly what the neural network recognizes in the Digital Public Library of America (hereafter DPLA) is a complex question with an equally complex answer. At the surface level, the network seems to be good at recognizing clothing, settings, and props and less adept at recognizing people based on faces or body structure. For example, there was a photo of two ballet dancers in costume on stage, and the top match was stage curtain. Its ability to reliably recognize objects, however, is dependent on a number of factors about the images it is given to classify. Images with extremely low or high depth of field, for example, are difficult for the network to classify accurately. Very flat images are often categorized as things like binder, menu, website, book jacket. Equally, images with a lot of depth are often described with the labels "CRT Screen, screen" or "monitor." The network performs best when objects are fully visible in the middle ground of an image, rather than the foreground or background. Similar issues occur as a consequence of the network's limited understanding of scale. In one instance, an image of children playing on an old metal slide is interpreted by the algorithm as that of a suspension bridge, a tow truck, or even a plane wing, perhaps due to the lighting or angle. The neural network algorithms are not good at recognizing people, but instead highlight the props or locations of the person's surroundings, as mentioned above. The network also seems restrained by certain criteria, or vocabulary, to label objects. However, it could be that photographs used for this post are restricted to certain themes, which is why certain keywords are present multiple times.

The primary benefit of applying image classification on the DPLA is that it allows archivists and historians to very quickly narrow down sets of photos based on vague parameters, which probably took hours or days prior to the availability of these resources. Once a collection of pictures is made, then the user could draw parallels or differences from the actualities that exist in the photographs that all have, for example, men with mustaches in them. Metadata like dates, places, types of people would be the information people could then process on their own or using other computational methods. I wouldn't, however, just leave it to the network to formulate my digital archive in its entirety for me. It can be used to cull references from a large collection of images, but is not exhaustive, leaving out all images that do not fall in the middle ground of the photo, as noted above. Machine is not yet better than human. Even if it were, I still don't think I'd trust a purely algorithmically curated archive, if such a thing can even exist.

Overall Picture Examples:

Most Likely: A Toaster

Most Likely: Vacuum Cleaner

- This can be seen here: of a black and white photo of a buddha > statue,
- DPLA ID: *c5a787b593de314c7ed3618e7a2e3d6c*

--

26.90% book jacket, dust cover, dust jacket,
dust wrapper ----

5.06% pickelhaube 4.83% milk can 3.88%
web site, website, internet site, site 3.65%
bulletproof vest 3.59% punching bag,
punch bag, punching ball, punchball
2.86% pedestal, plinth, footstall

51.37% window screen ----
----- 16.63%
window shade 2.17% screen,
CRT screen 1.55% park bench
1.41% monitor 1.08% picket
fence, paling 0.96% suspension
bridge
