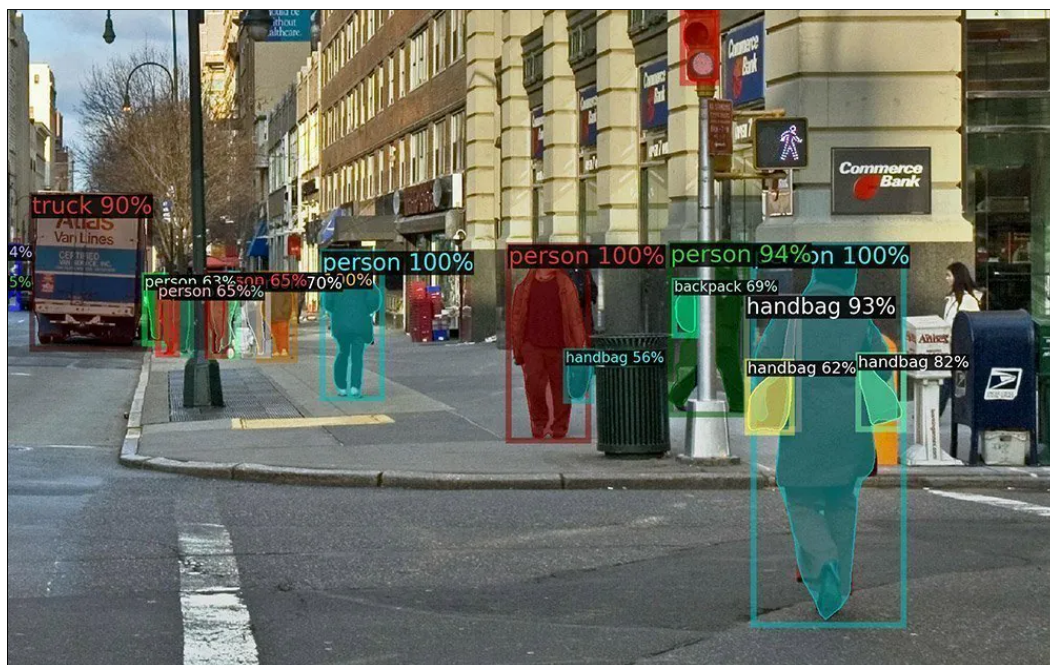


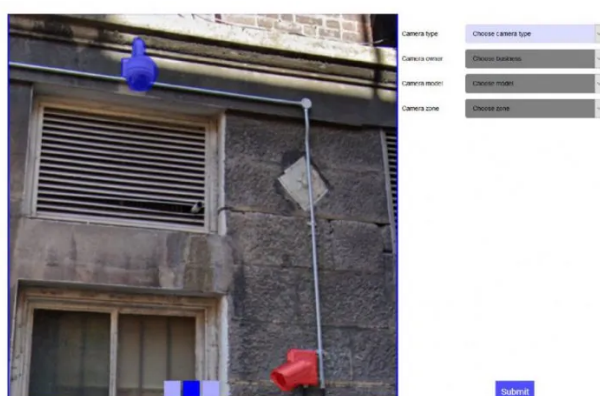
A Browser-Based Image Annotation Tool for Computer Vision Datasets

Originally published at <https://www.unite.ai/a-browser-based-image-annotation-tool-for-computer-vision-datasets/>



Researchers from Finland have developed a browser-based image labeling tool intended to improve the ease and speed of tedious image annotation processes for computer vision datasets. Installed as an OS-agnostic extension for the most popular browser engines, the new tool enables users to 'annotate while freely browsing', rather than needing to put a labeling session into the context of a dedicated set-up, or to run client-side code and other special circumstances.

Entitled [BRIMA](#) (Low-Overhead BRowser-only IMage Annotation tool), the system was developed at the University of Jyväskylä. It removes the need to scrape and compile datasets into local or remote directories, and can be configured to derive useful data from the various data parameters available on any public-facing platform.



BRIMA in action. Source: <https://arxiv.org/pdf/2107.06351.pdf>

In this way BRIMA (which will be presented at [ICIP 2021](#), when the [code](#) will also be made available) obviates the potential obstacles that can arise when automated web-scraping systems are blocked via IP ranges or other methods, and impeded from gathering data – a scenario that is set to become more common as IP protection

comes increasingly into focus, as it has [recently done](#) with Microsoft's AI-driven code generation tool, Copilot.

Since BRIMA is intended solely for human-based annotation, its usage is also less likely to trigger other kinds of roadblocks, such as CAPTCHA challenges, or other automated systems intended to block data-gathering algorithms.

Adaptive Data-Gathering Capabilities

BRIMA is implemented via a Firefox add-on or Chrome extension on Windows, OSX or Linux, and can be configured to ingest salient data based on data points that a particular platform may choose to expose. For instance, when annotating images in Google Street View, the system can account for the orientation and viewpoint of the lens, and register the exact geo-location of the object specified under attention by the user.



BRIMA was tested in September of 2020 by its creators, during collaboration on a crowdsourced initiative to generate an object detection dataset for CCTV objects (video surveillance cameras mounted in public spaces, or viewable from public spaces).

The system is comprised of a lightweight JavaScript client-side installation in the form of the browser extension, and a server-side aspect which receives and compiles the annotation data. Reference implementations of the server-side installation were written in Python and PHP with Flask and Swagger/OpenAPI, but the researchers emphasize that the central processing architecture can easily be ported to other languages and configurations.

The browser extension and the server communicate via RESTful API requests and HTTP/XHR, with the client-side data sent home in a JSON format that's compatible with MS COCO. This means that the data is immediately usable with a variety of the most popular object detection frameworks, including diverse back-ends to TensorFlow, such as Facebook's [Detectron2](#), and [CenterMask2](#).

Project-Specific Tooling

Despite the generic nature of BRIMA, it can be configured into highly specific data-gathering configurations, including the imposition of drop-down menus and other kinds of contextual input related to a particular domain. In the image below we see that a drop-down menu relating to camera information has been written into BRIMA, so that a group of annotators can provide detailed and project-relevant information.



This additional tooling can be configured locally. The extension also features easy installation and configurable keyboard shortcuts, along with color-coded UI elements.

The work builds on a number of attempts in recent years to improve the facility of image annotation for web-obtained or public-facing data. The PhotoStuff tool, supported by DARPA, offers online annotation via a dedicated web portal, and can be run on the semantic web or as a standalone application; in 2004 UC Berkeley proposed [Photo Annotation on a Camera Phone](#), which heavily leveraged metadata, due to the limitations of network coverage and the viewport limitations of the era; MIT's 2005 [LabelMe](#) project also approached browser-based annotation, with a reliance on MATLAB tools;

Since its release in 2015, the FOSS Python/QT framework [Labellmg](#) has gained popularity in crowdsourced annotation efforts, with a dedicated local installation. However, the BRIMA researchers observe that Labellmg centers on PascalVOC and YOLO standards, does not support MS COCO JSON format, and eschews polygonal outlining tools in favor of simple rectangular capture regions (which will require subsequent segmentation).

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More