# Annotations protocol

Last updated: 14 July 2023

*For the interactive version, please refer to the gdoc here.*

# VGG Visual Image Annotator

## General Notes

The web UI works best with a limited number of images, so try to work with mutually exclusive subsamples (say tens of images per session).

## Task description

Annotators are required to perform polygon annotation, outlining the boundaries of nuclei and soma/cytoplasm based on the provided fluorescence microscopy images. The annotations should accurately represent the contours of these structures according to the descriptions provided below. The two biological structures are highlighted by means of three functional markers: c-FOS (green) for nuclei, CTb (yellow) for cytoplasm, and Orx (red) for cytoplasm.

### Background

In this context, background refers to all areas except the nuclei and soma/cytoplasm of interest. It encompasses void areas (typically black), tissue portions with faint staining and artifacts. The fainted background often exhibits a mostly uniform texture and intensity, although it may not always be perfectly homogeneous. Artifacts are stains that resemble marked cells in terms of intensity or texture, but with different shapes or dimensions (i.e., too small or too big to be a real cell or cell-nuclei). These can be technical artifacts caused by fluorophore accumulations or biological artifacts such as stripes or stains along tissue borders and filaments.

### Green fluorochrome, cell nuclei (c-Fos)

Nuclei appear as circular-shaped objects with prominent staining. The borders of nuclei may vary in definition depending on their position relative to the focal plane.

### Yellow fluorochrome, cell soma (CTb)

For the soma/cytoplasm, larger stains are observed. The shape is typically rounded, with some heterogeneity on the exact form of the contours. The edges of soma/cytoplasm are sharper compared to nuclei. In certain areas, cells may appear crowded with overlapping stains from adjacent cells. Staining in the soma/cytoplasm is homogeneous, intense, and usually clearly distinguishable from the background.

### Red fluorochrome, cell soma (different cytoplasmic targets)

Similarly, for the soma/cytoplasm, objects may be larger and exhibit elongated or triangular shapes. The contrast between stained soma/cytoplasm and the background is typically clear. However, crowded areas with overlapping objects are more common. In such cases, cell contours can be determined by observing the gradient of staining intensity. Contextual information, such as darker areas and prior knowledge of soma shape and dimension, can assist in identifying cell boundaries, i.e., distinguishing a single, larger stain from multiple cells.

### Borderline cases

“Borderline cases” refer to stained structures that are compatible with the shape of cells but have intensities close to the fainted background. Recognizing such cases relies on subjective judgment and depends on the annotator's interpretation. As a general guideline, a conservative approach should be taken when evaluating borderline cases.

Marked structures are defined as neurons or neuronal nuclei. Neurons’ shape is defined by a cell body or soma, the core section, and different branches (called axon and dendrites). The soma, which contains the nucleus and many other specialized organelles, is relatively large, typically round-shaped with defined edges; however, it often appears elongated or triangle-shaped with extensions emanating from vertices.
The staining of the soma, obtained by detecting cytoplasmic proteins, is usually homogeneous, rather intense, and clearly distinguishable from the background staining (i.e., see below). The branches, at least the first part, are clearly visible and it is often possible to identify the nucleus as a circle less intensely stained.
On the other hand, by using markers specific for nuclear proteins, it is possible to exclusively highlight cellular nuclei: they will appear round shaped and clearly stained, with borders more or less defined depending on their position with respect to the actual focal plane. Often, nuclei are also provided with one or more unstained nucleoli.
The term “background” refers to the non-specific staining that, although faint, is shown by nervous parenchyma. Generally, it is quite uniform (although not perfectly homogeneous), both in structure and in staining intensity. It should not be confused with areas where the biological tissue is lacking, such as tissue fractures of the sample or brain ventricles for instance: in these cases, the staining is totally absent.
The color itself is not important since it is chosen by the experimenter and it depends on the technical needs that characterize the specific experiment.
It is frequent to observe stained structures that follow the above-described criteria but have a staining intensity very similar to that of the "background": in this case, normally, the neuron is considered "non stained". The precise and specific threshold that allows distinguishing an intensity considered as "positively stained" from one that is considered as background cannot be objectively and univocally defined in a manual analysis but is subjected to the mere experimenter decision. Therefore, the definition of "borderline" cases as positively or negatively stained represents a highly subjective aspect of the procedure and heavily depends on the experimenter's evaluation.

## Collection design

In order to assess potential operator effects, we keep track of who annotated what. These metadata will be anonymized and shared together with the dataset.
To avoid spurious effects, the assignment of images to annotators has to be random.
In practice:

1. assess annotators expertise and availability
2. determine number of images to annotate per operator, accounting for possible limitations (e.g. operator X has not sufficient expertise for dataset Y)
3. random samples image buckets for each annotator
4. perform annotations and register annotator and annotations mode in the metadata*

This process can be done once at the beginning, or repeated multiple times (at each annotation session) depending on convenience.

## Metadata*

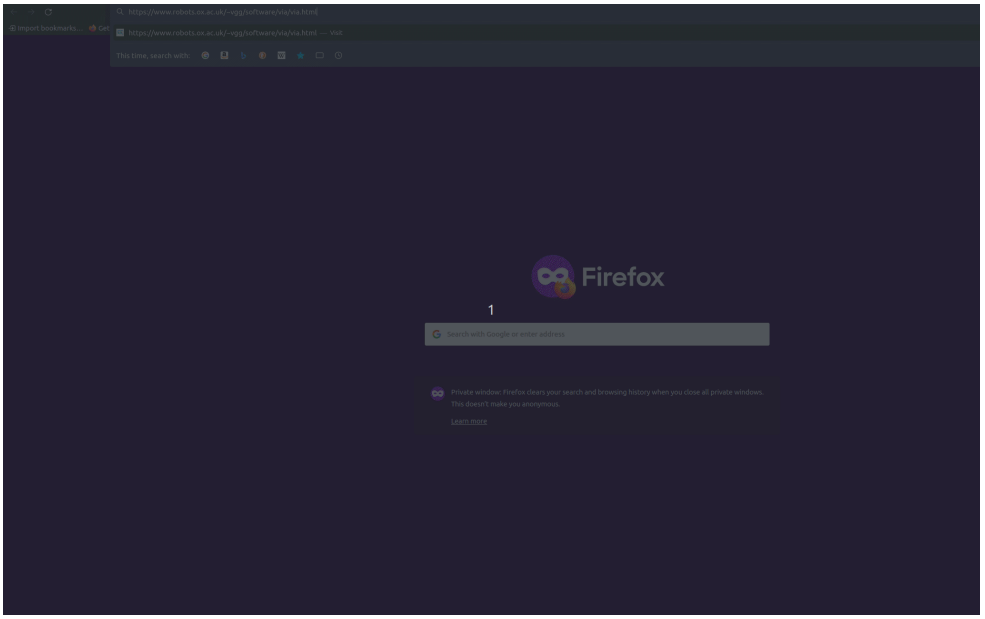
Alongside annotations, we keep track of collection metadata about:

- **annotator:** the list of annotators is tracked in *sheet:Annotations metadata* of metadata. During the collection, the operator has to fill in the corresponding drop-down menu in *sheet:Dataset metadata*
- **annotation type:**
  - *only manual* means that all objects are added manually
  - *semi-automatic* means that the annotations started from “pre-labels” and the annotator added/deleted objects
- **additional comments:** this is a free text field to write down any relevant comment, e.g. issues, annotation session, time, … (*not mandatory*)

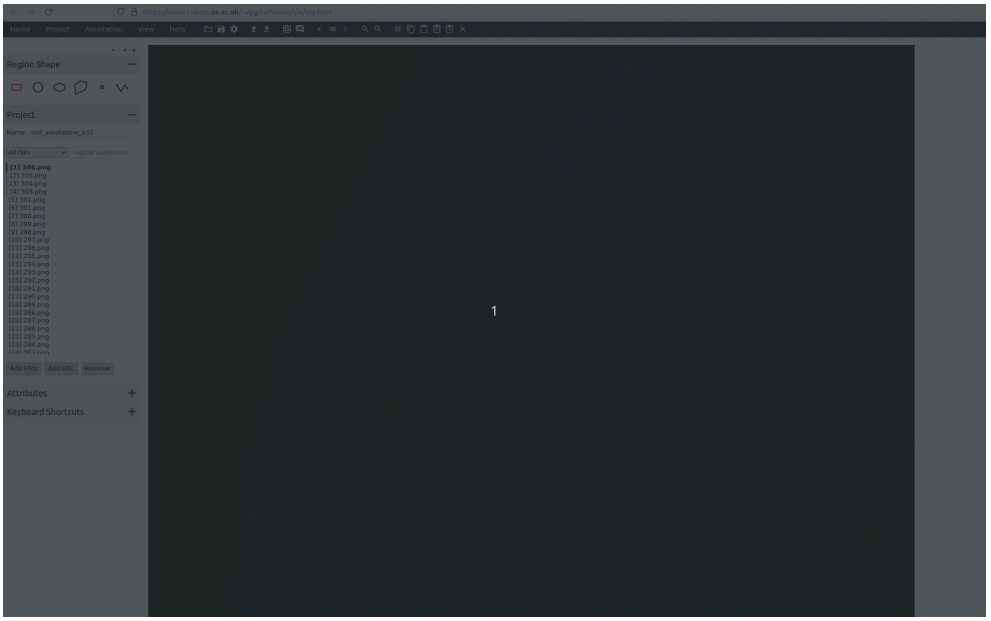
## Protocol

1. Go to: https://www.robots.ox.ac.uk/~vgg/software/via/via.html

2. Click on `Add Files` on the left menu and import all the images to annotate
   *NOTE: it may take few seconds to display them in case they are a lot*

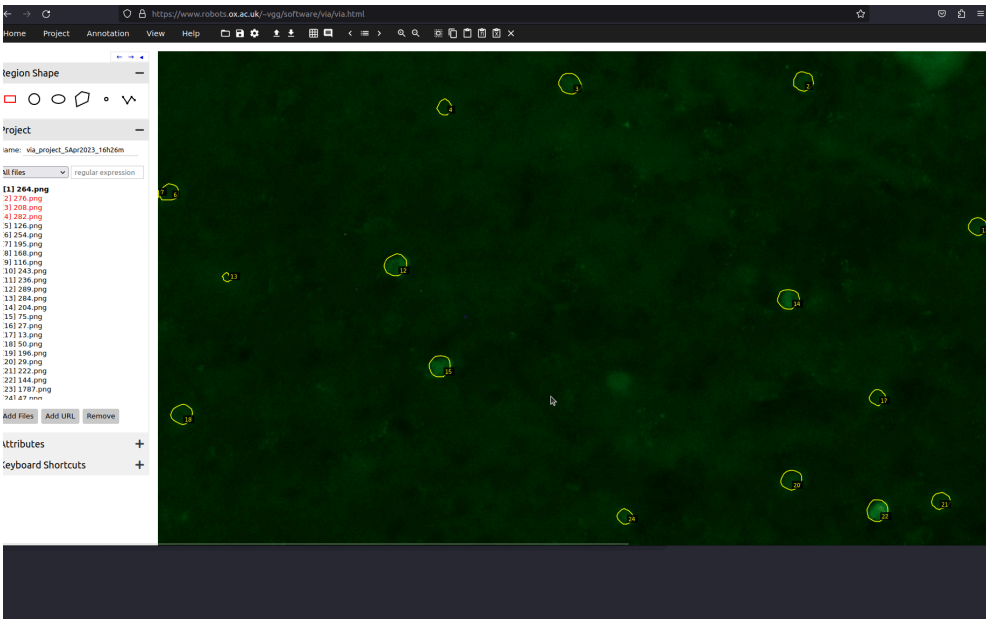



3. Click on the `Annotation` button in the top menu, then select `Import annotations (from csv)`. Search for the annotations csv file in your filesystem and click load it. You will find an import summary at the bottom of the page, and the pre-annotated regions (cells) will appear as colored polygons overlapped on the images.

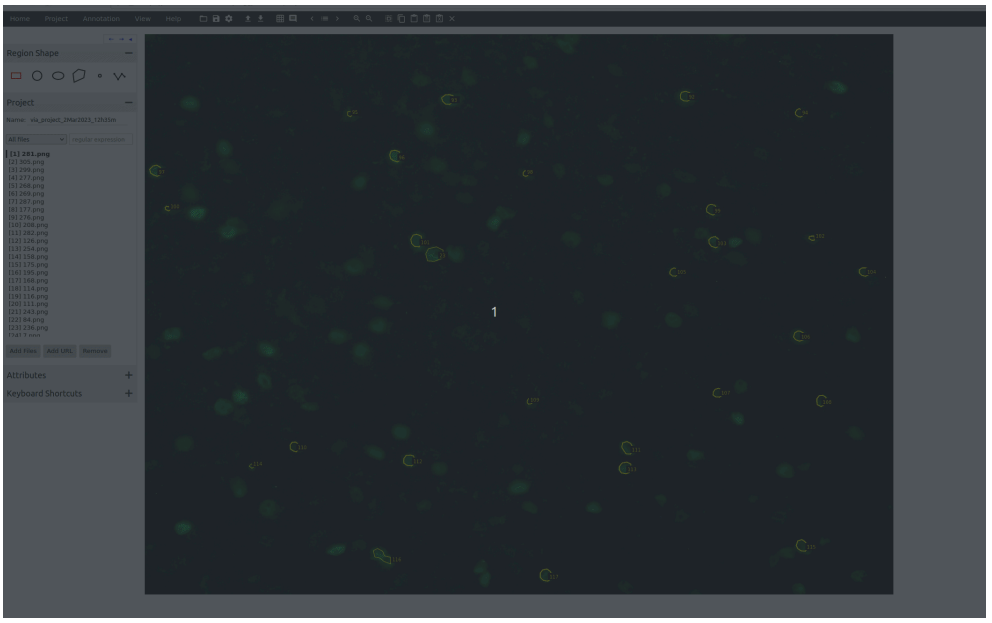



4. First, double check current annotations:
   a. you can zoom in/out with *Ctrl +/-*
   b. you can move right/left with *Maiusc + scroll down/up*
   c. Select one region by clicking on it (it highlights in red)
      *NOTE: if that doesn't work, try pressing Esc button before*
   d. Edit region points by dragging them around
   e. Delete region points with *Ctrl + click*

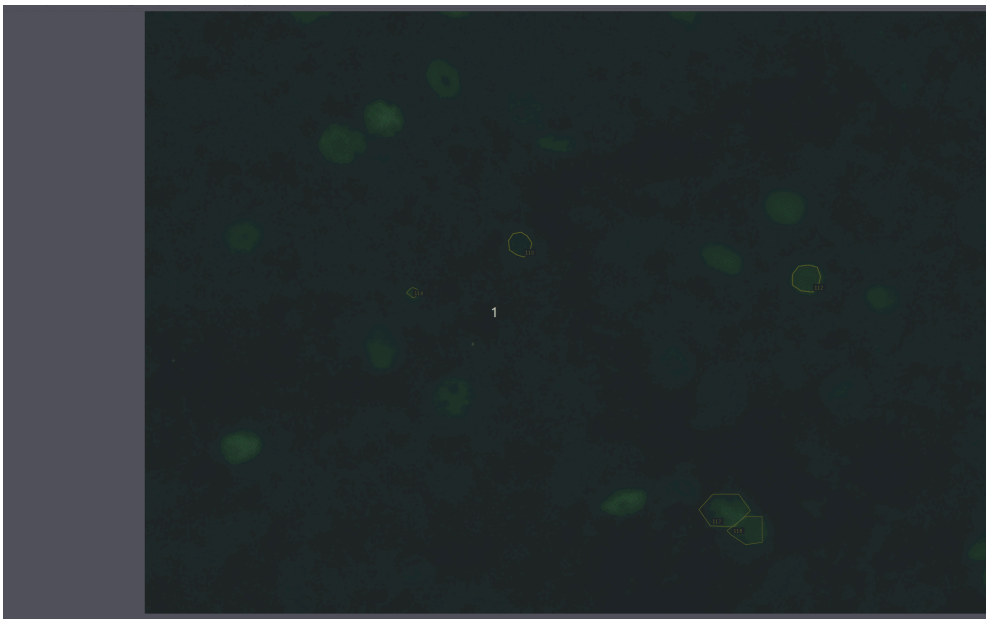
f. Delete region by hitting `*d*` key when the region is selected




5. Then, add new regions if necessary
    a. click on the polygon icon on the top-left corner
    b. set up the field of view and start positioning points by clicking
    c. click enter after the last point to let the UI complete the shape



d. you can also move/edit regions if needed



6. Finally, you can preview and export the annotations from the `Annotation` menu at the top
   NOTE: currently, our code only supports csv annotations