Developing a ML pipeline to detect centrioles in human cells

Leo Burgy
2021-11-15

Goals

- Command line program: multidimensional* image → cells described in terms of centrioles.

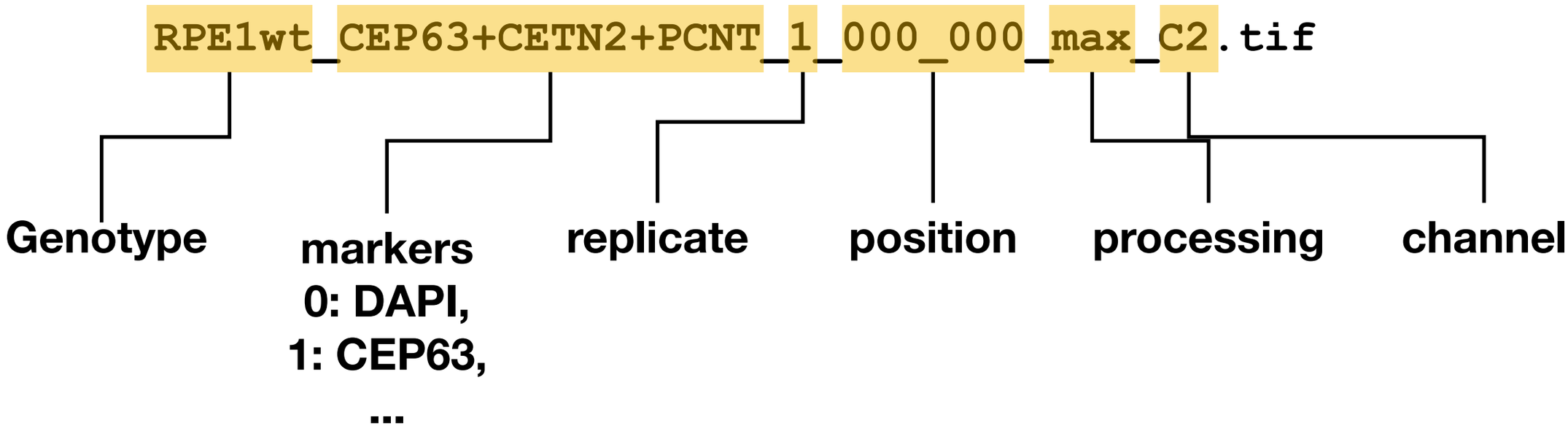
- Conservative approach: Precision and recall

* DAPI + centriole marker; DAPI + all centriole markers

Dataset description

- 3 experiments
- multiple markers for centrioles
- 1 marker for DAPI (nucleus, ~cell)
- z-max projection
- Annotation for every centriole marker, for every experiment.
- https://www.dropbox.com/sh/5042baoljv2vg24/AAB1aqWVag_klErKZarvyykta?dl=0
- https://journals.biologists.com/jcs/article/132/4/jcs228833/57571/Centriole-assembly-at-a-glance

Naming Convention



Strategies

split-combine learning: detect objects in parallel, then associate them

one-pass learning: learn the relationship directly.

A) Combine information from several channels

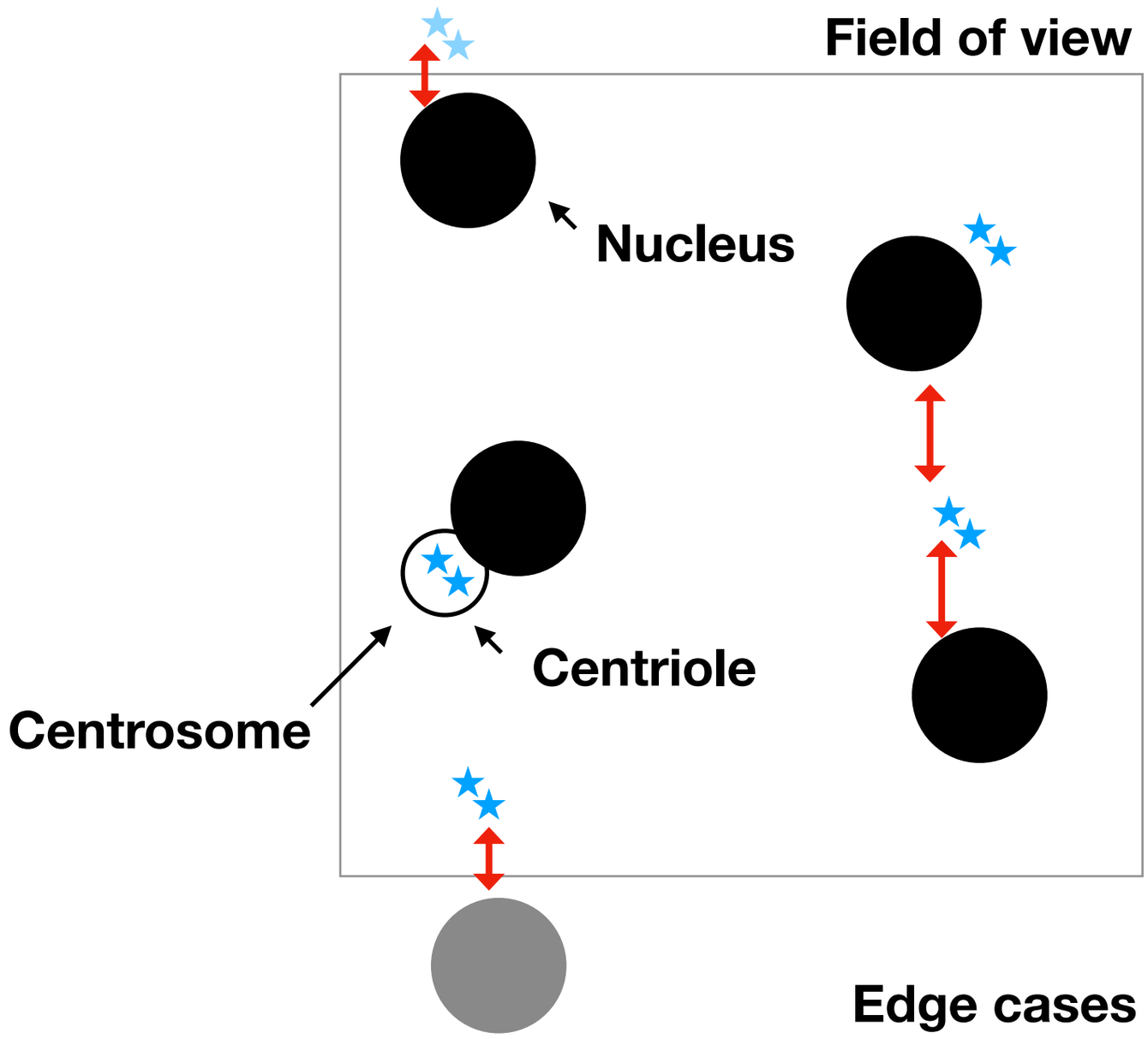
- https://www.nature.com/articles/s42256-021-00379-y

- Find an embedding

B) Assign the centrioles to cell

- Segment the nuclei on images (Labelbox)

- How to handle edge cases? Objects on the margin, excluding the associated nucleus/ centrioles.



(Re)view the performances

- Precision? Recall?

- Visualise the predictions: write full annotation on images; use napari, generate crops?