

Feature Extraction in Python

PoL Bio-Image Analysis Training School & Symposium

29 August 2023

Allyson Quinn Ryan, PhD
BiA-PoL & Modes Group

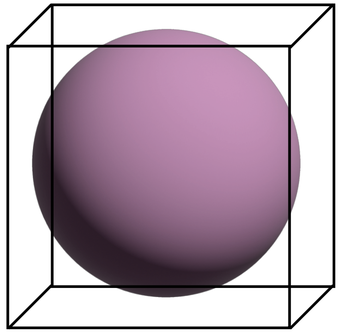
What defines a feature?

a quantification or relationship that describes your system

What do we need to consider first?

object type, neighbourhoods, structuring elements... ⟶ feature categories

Feature Classes: Size

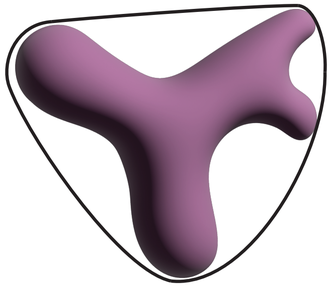


bounding box

pixel/voxel count

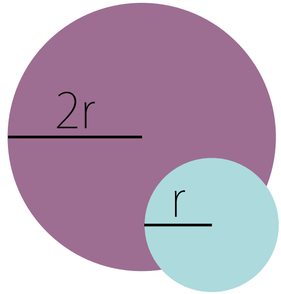
area; volume
(scaling info req)

perimeter;
surface area



convex hull

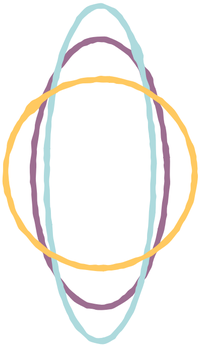
convex
area/volume



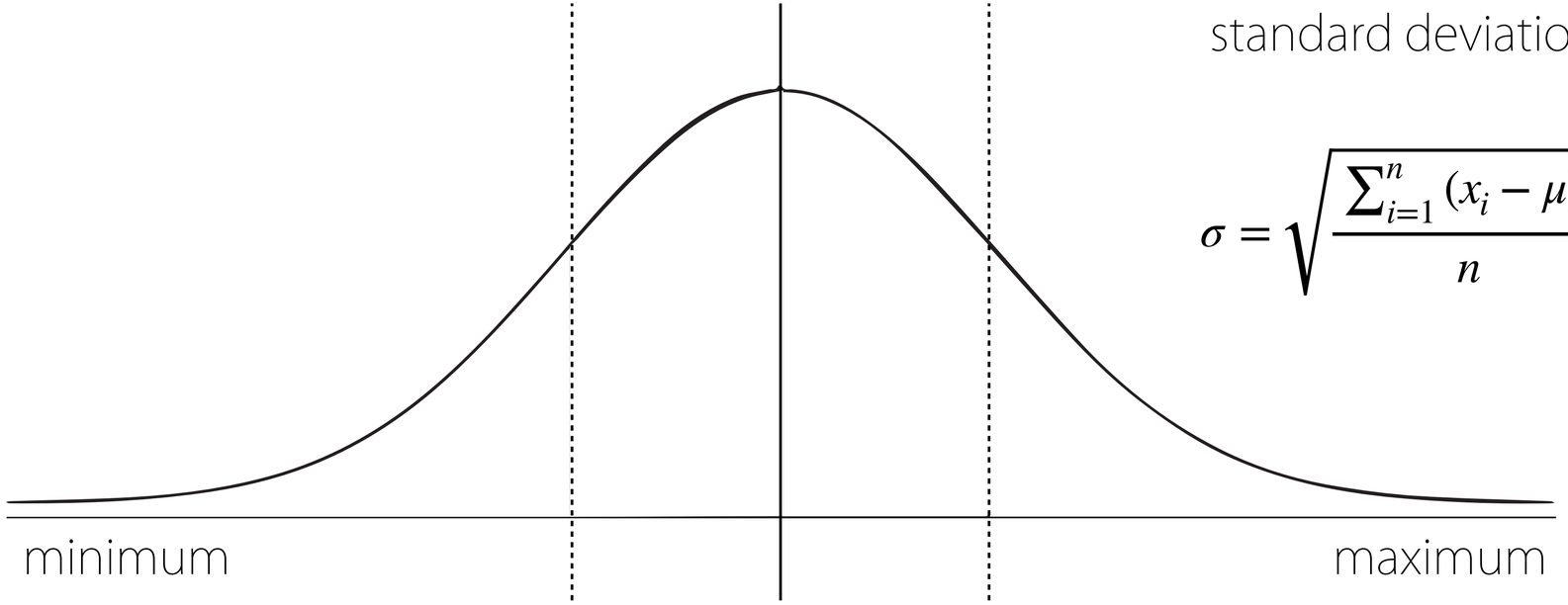
equivalent
diameter/axes

spherical fit

ellipsoidal fit



Feature Classes: Intensity

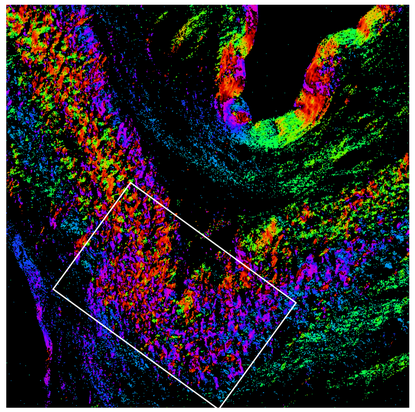


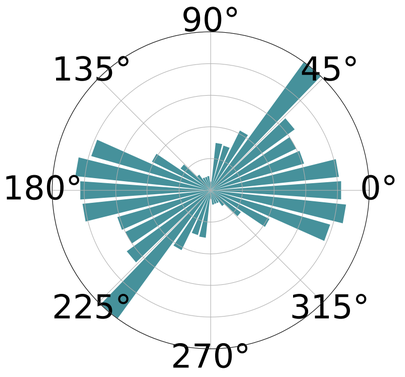
standard deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}}$$

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

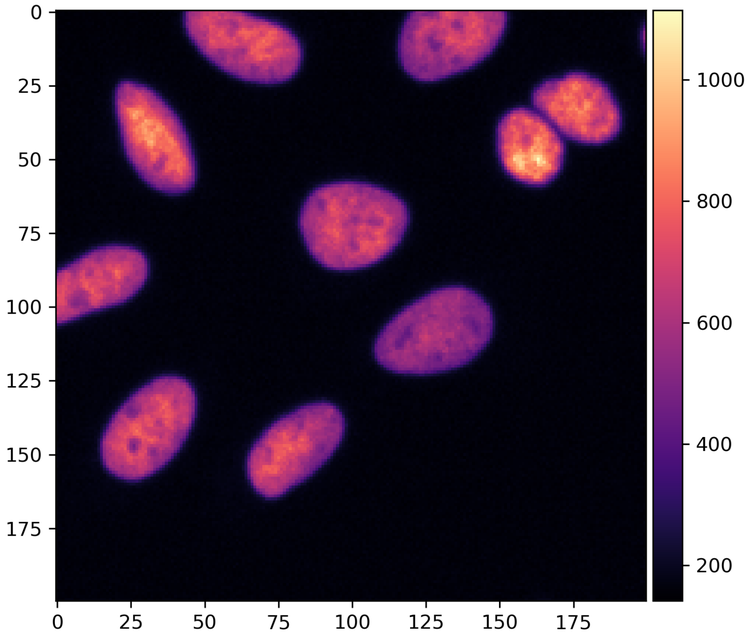
sum, median, variance, etc.



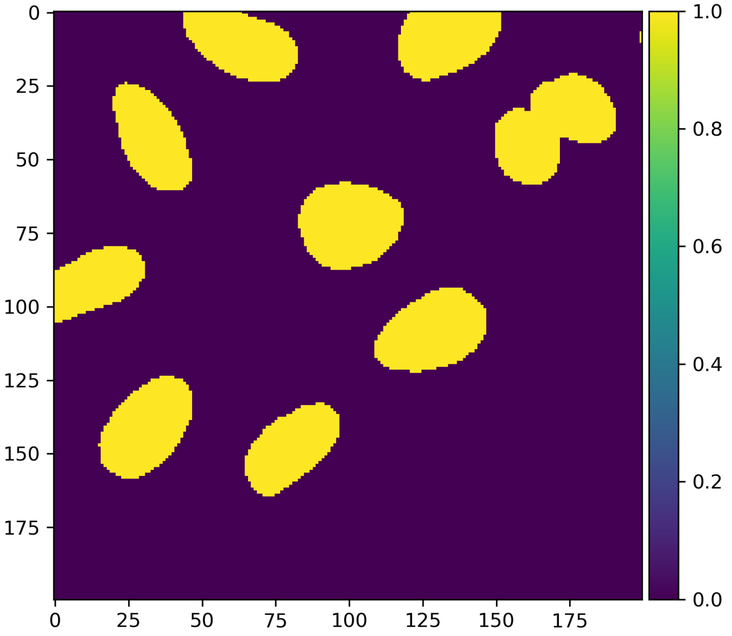


(subclass: texture)

Intensity: all foreground objects



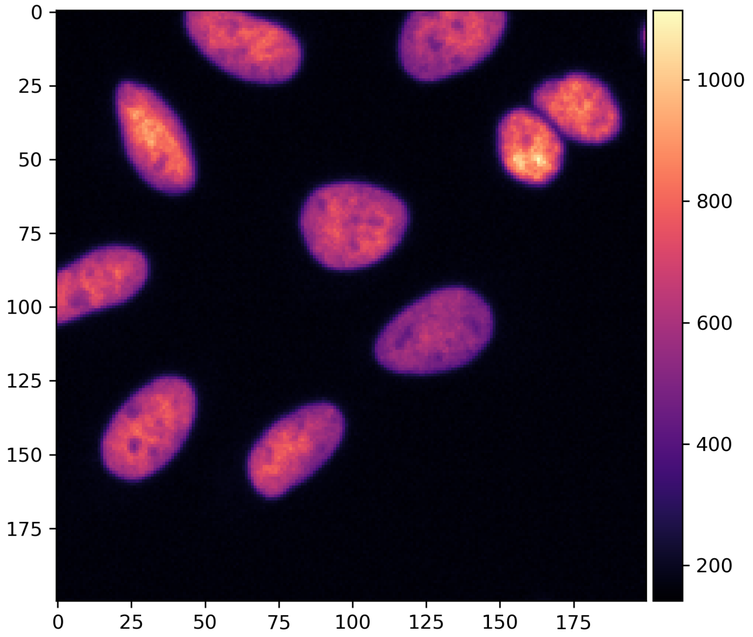
retrieve binary image



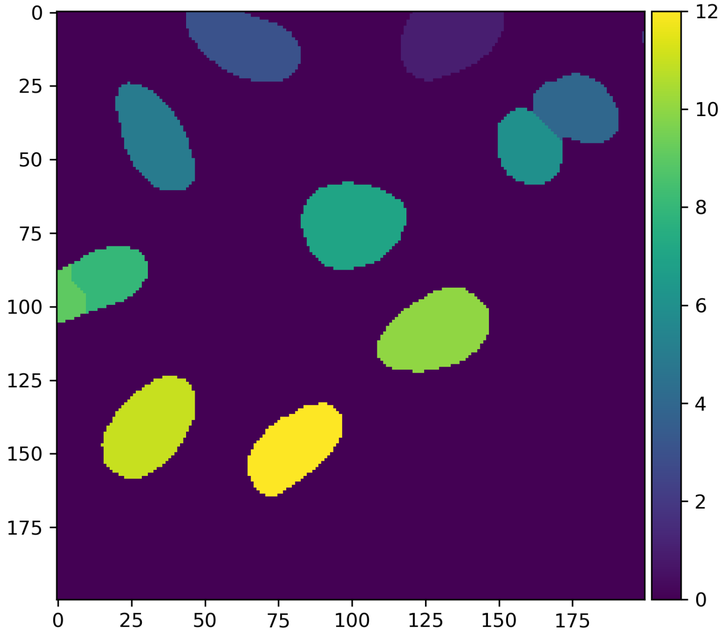
index raw input image

mean, minimum, maximum, standard deviation, etc.

Intensity: individual objects



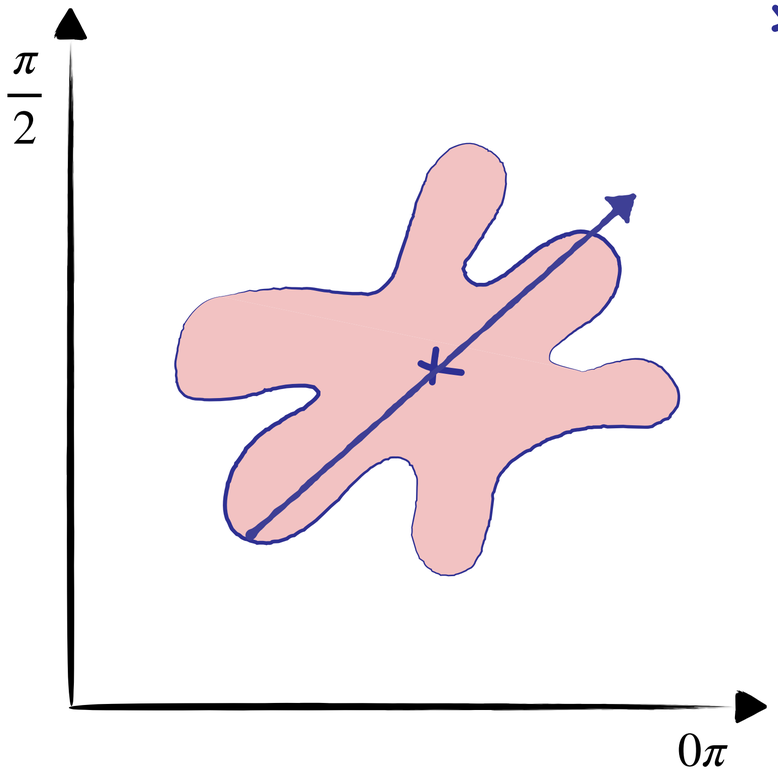
retrieve label image



Index raw input image

mean, minimum, maximum, standard deviation, etc.

Feature Classes: Position & Moments



- centroid/centre of mass/ weighted center

$$\boldsymbol{R} = \frac{1}{M}\iiint_{Q} \rho(\boldsymbol{r})\boldsymbol{r}dV$$

orientation

$$v = < x, y >$$

1st moment

centre of mass

2nd moment

intertial; variance

3rd moment

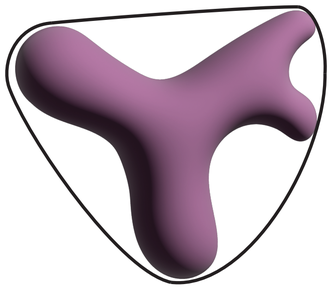
shape assymmetry ; skewness

(doodad from Mara Lampert)

Feature Classes: Shape

solidity

$$S = \frac{V_{obj}}{V_{ch}}$$

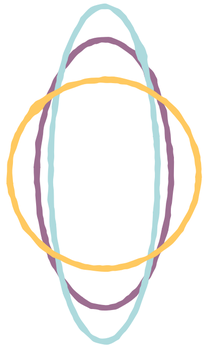


roundness

$$R = \frac{4\pi A_{obj}}{P_{ch}^2}$$

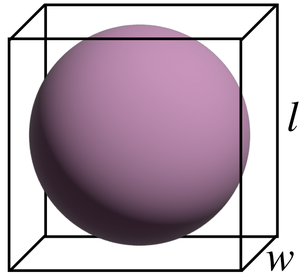
aspect ratio (eccentricity)

$$ar = \frac{l_{major}}{l_{minor}} \quad \text{OR} \quad ar = \frac{l_{minor}}{l_{major}}$$



elongation

$$f = \frac{w_{bbox}}{l_{bbox}}$$



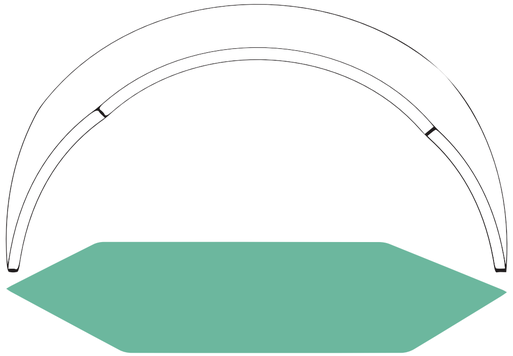
sphericity

$$\psi = \frac{\pi^{1/3}(6V)^{2/3}}{SA}$$

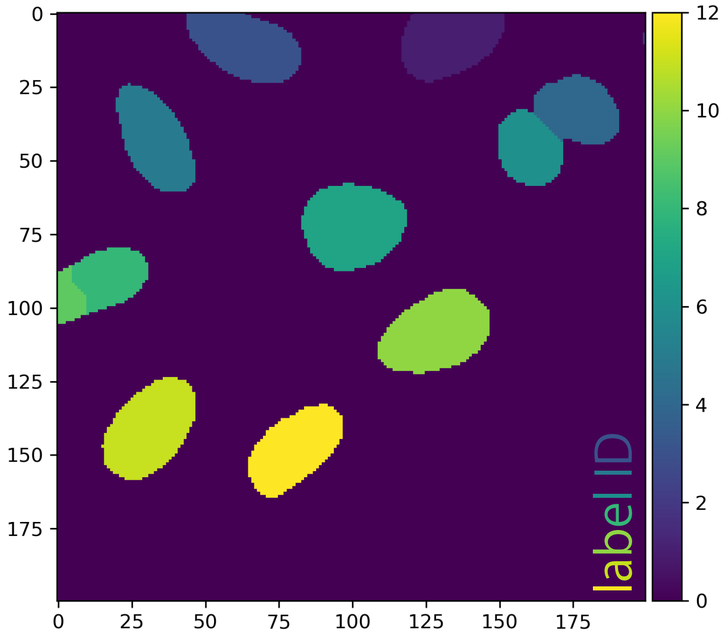
circularity

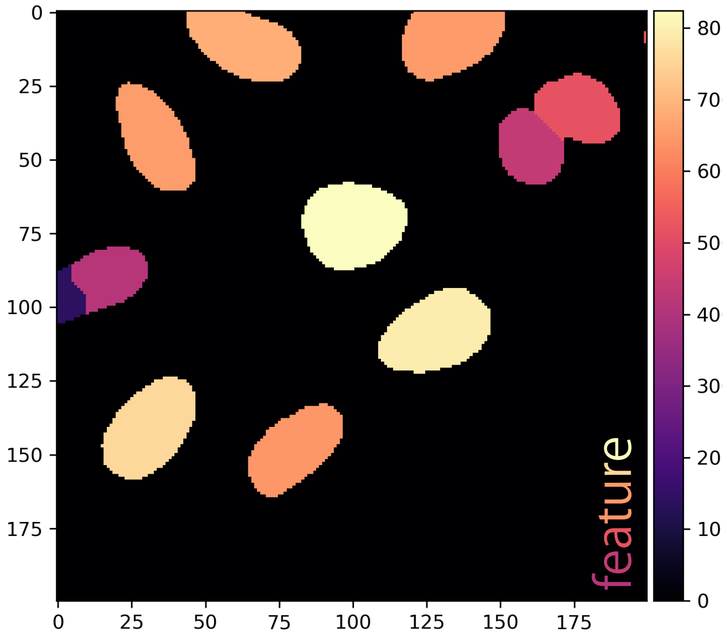
$$C = \frac{4\pi A}{P^2}$$

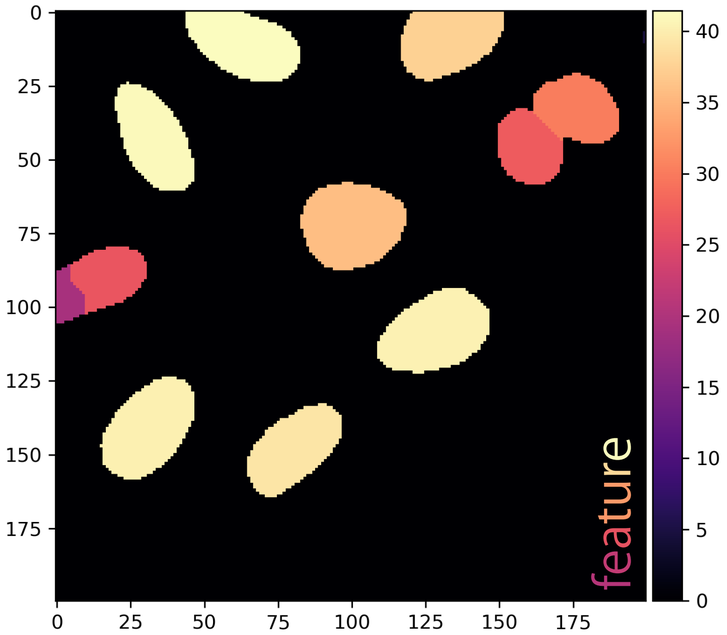
flatness (projection)



Pitfalls in measurement robustness.







Discretisation impacts measurements!

size

2D: area & perimeter

3D: volume & surface area

shape

axis length, circularity,

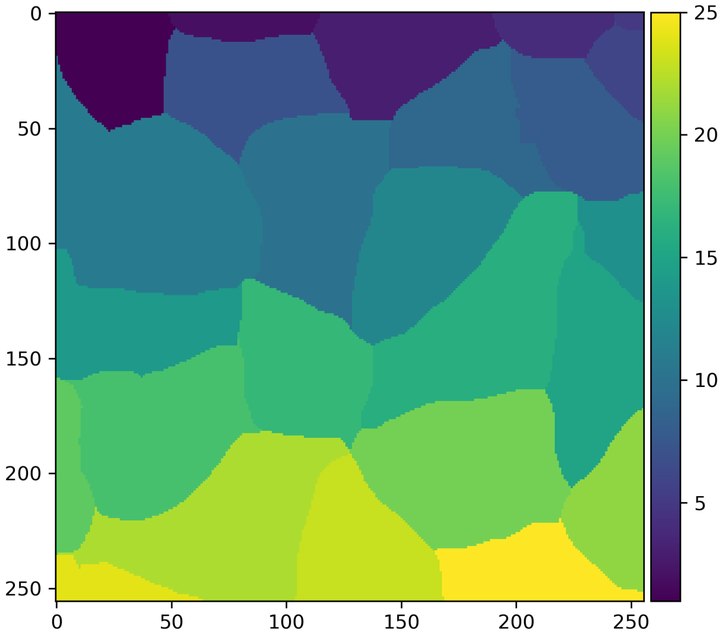
aspect ratio, eccentricity, etc.

Pitfall example: perimeter calculation

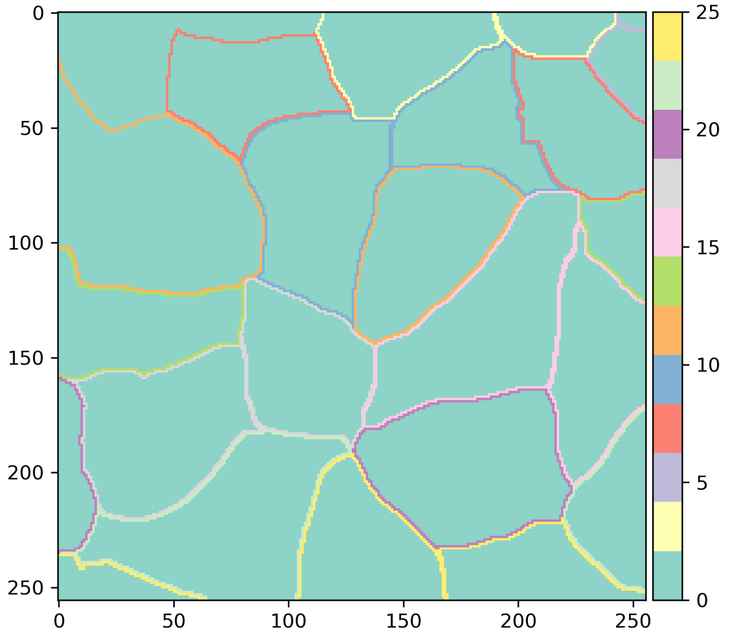
circularity

$$R = \frac{4\pi A}{P^2}$$

$\sqrt{2}$	1	$\sqrt{2}$
1		1
$\sqrt{2}$	1	$\sqrt{2}$

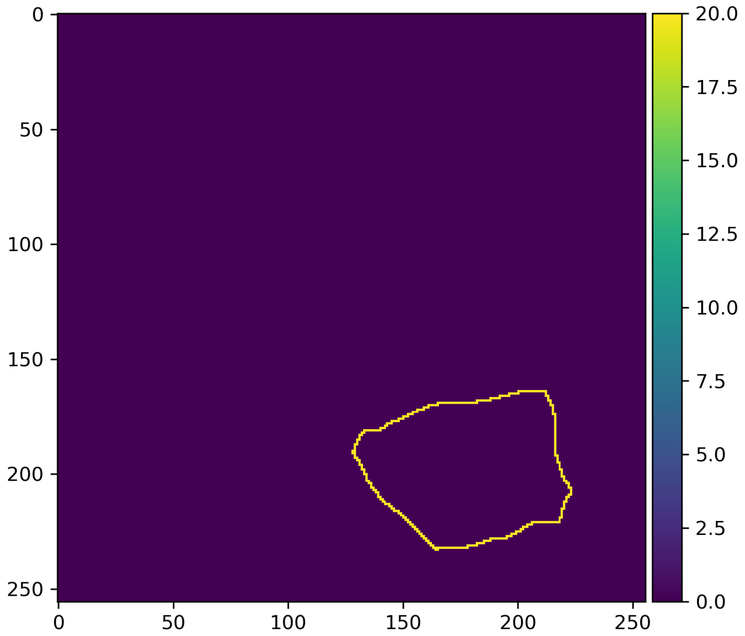


extract boundaries



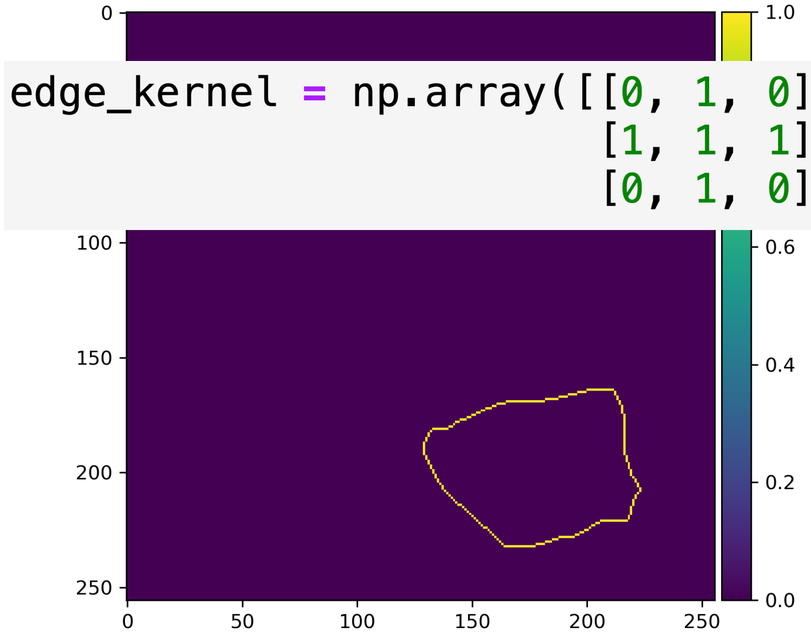
index segmentation

Pitfall example: perimeter calculation

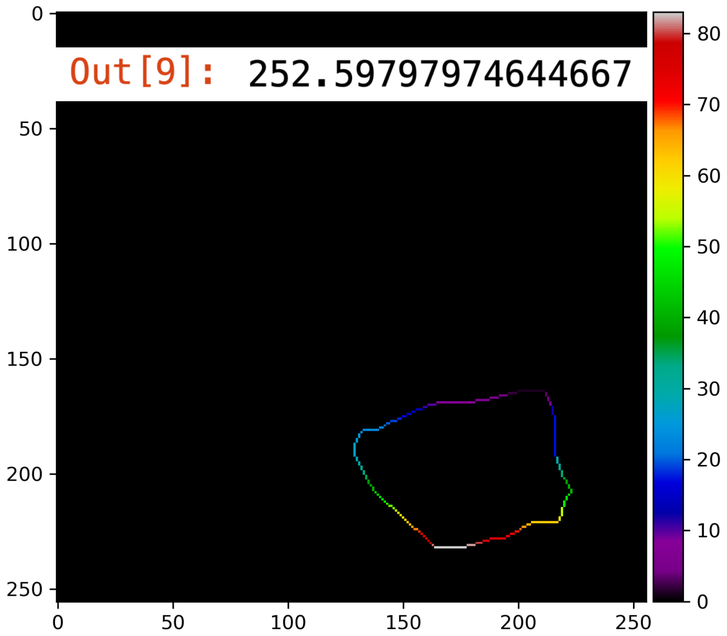


thin boundary

`skeletonize`




```
edge_kernel = np.array([[0, 1, 0],
                        [1, 1, 1],
                        [0, 1, 0]])
```

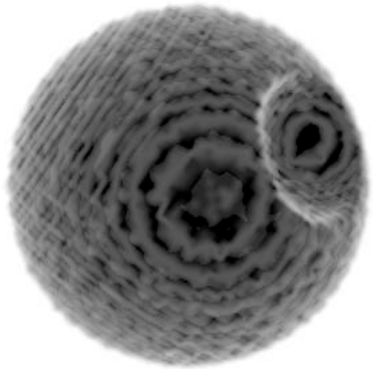


separate N_4 and N_8 boundary components

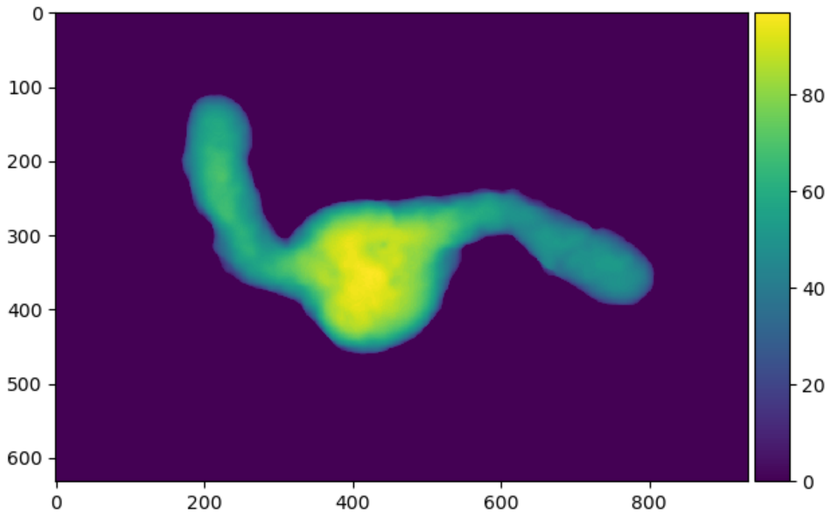
$\sqrt{2}$	1	$\sqrt{2}$
1		1
$\sqrt{2}$	1	$\sqrt{2}$

Dealing with complex shapes in 3D...

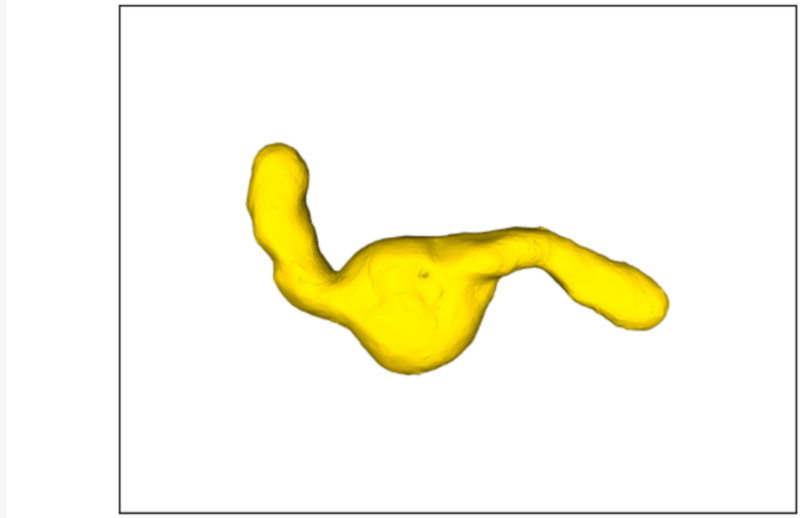
Dealing with complex shapes in 3D...



Meshes can reduce discretisation errors.



marching cubes mesh creation



nppas.SurfaceTuple

origin (z/y/x)	[0. 0. 0.]
center of mass(z/y/x)	57.605,308.700,440.973
scale(z/y/x)	1.000,1.000,1.000
bounds (z/y/x)	13.911...112.189 111.132...461.732 169.884...807.977
average size	170.787
number of vertices	3309
number of faces	6614

Important Documentation & Reading

skimage.measure offers many functions for feature extraction from images (particularly regionprops):

https://scikit-image.org/docs/dev/api/skimage.measure.html#skimage.measure.regionprops

vedo.mesh is useful for shape, size and positional feature extraction:

https://vedo.embl.es/docs/vedo/mesh.html

pyclesperanto_prototype offers many example workflows for intensity, size and positional features:

https://github.com/clEsperanto/pyclesperanto_prototype

FocalPlane feature extraction blogpost by Mara Lampert

https://focalplane.biologists.com/2023/05/03/feature-extraction-in-napari/