Feature Extraction in Python

Bio-image Analysis, Bio-stats, Programming & Machine Learning for Comp. Bio.

02 May 2023

Allyson Quinn Ryan, PhD

Contributors:

Allyson Quinn Ryan

Robert Haase

Carl D. Modes

Mara Lampert

object classification &

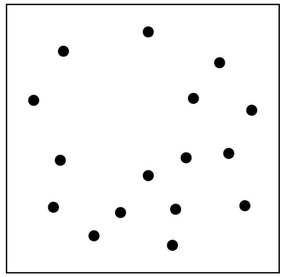
science communication

network complexity & systems biophysics

https://focalplane.biologists.com/author/marabuuu/

Topology: scale-free invariant features.

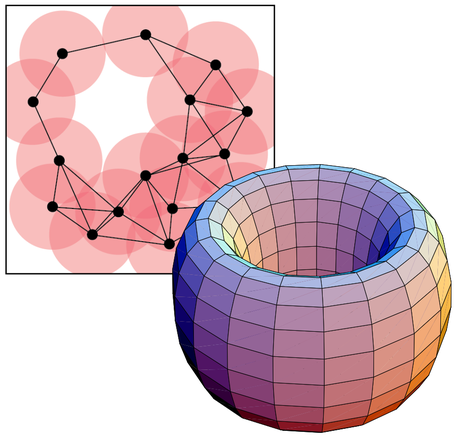
algebraic topology



persistent homology

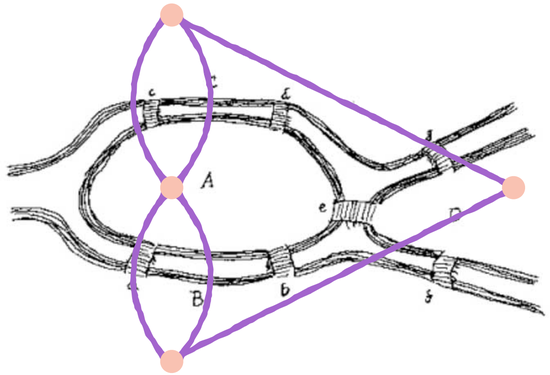
Topology: scale-free invariant features.

algebraic topology



persistent homology
polygonal meshes

network topology



neighbourhood analysis
collaborations

digital topology

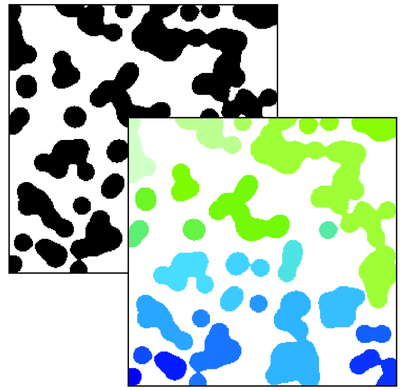


image segmentation
connected components

(Wolfram Alpha; 'Solutio problematis ad geometriam situs pertinentis' — L. Euler, 1736)

What is an image feature?

a quantification or relationship that describes your system

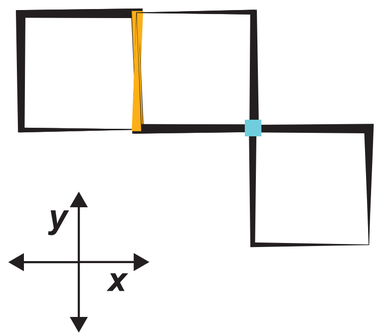
What do we need to consider first?

neighbourhoods, structuring elements & feature categories

Neighbourhoods & Structuring Elements

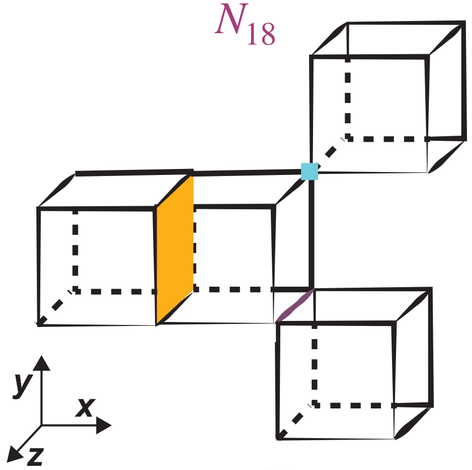
2D

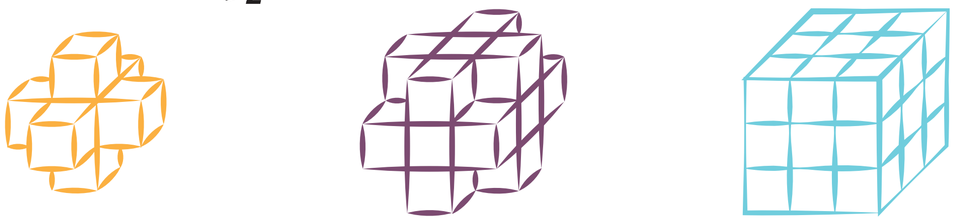
N_4 N_8



3D

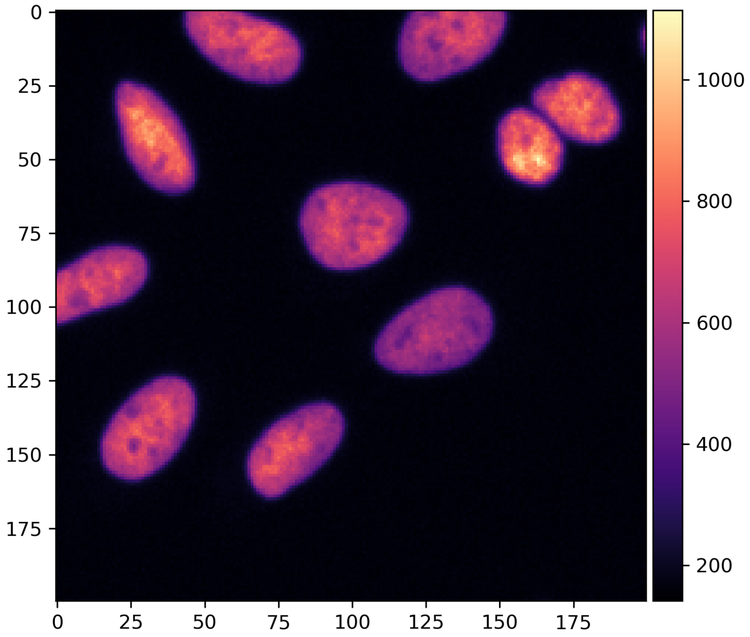
N_6 N_{18} N_{26}



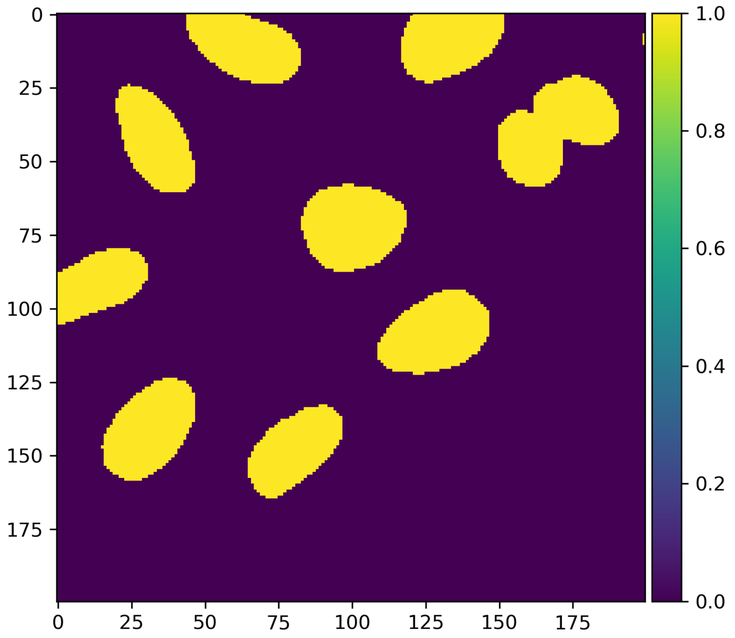


structuring elements are also often referred to as kernels

Intensity: all foreground objects



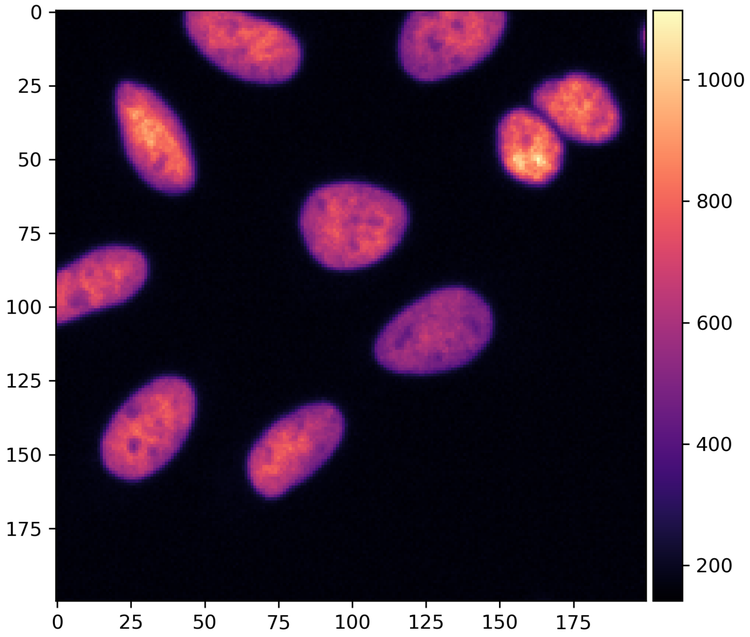
retrieve binary image



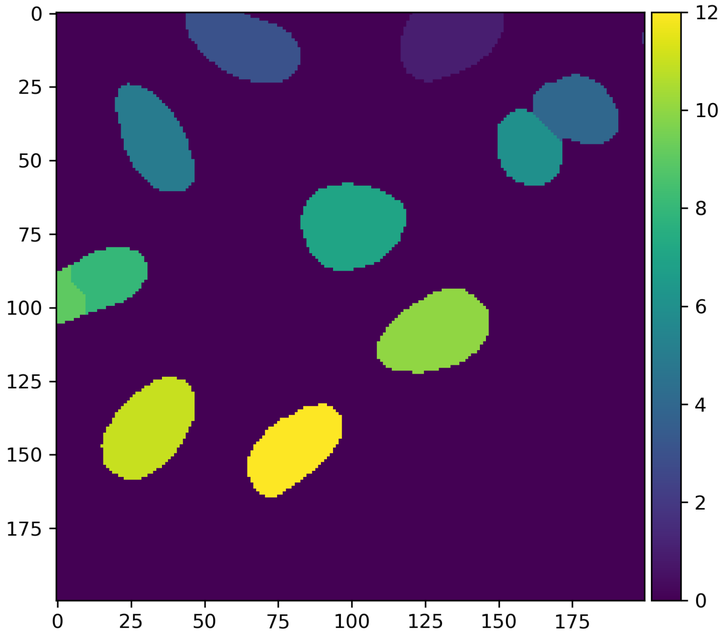
Index raw input image

mean, minimum, maximum, standard deviation, etc.

Intensity: individual objects



retrieve label image

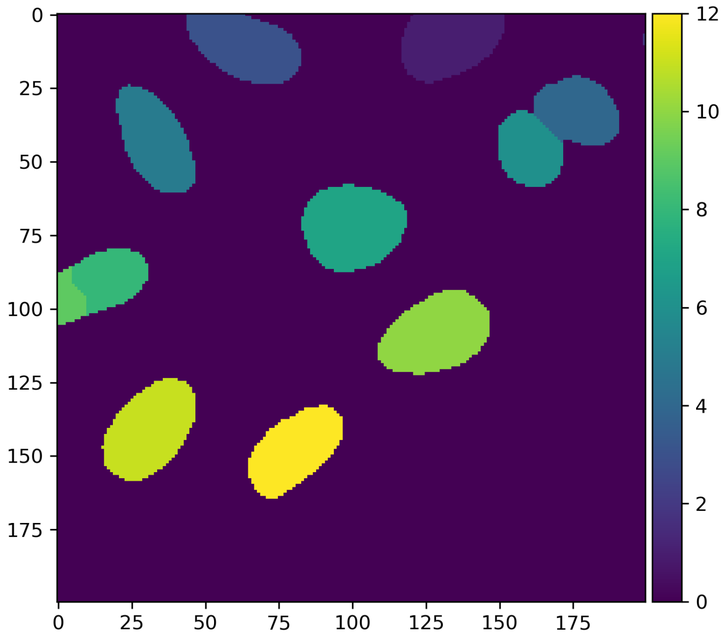


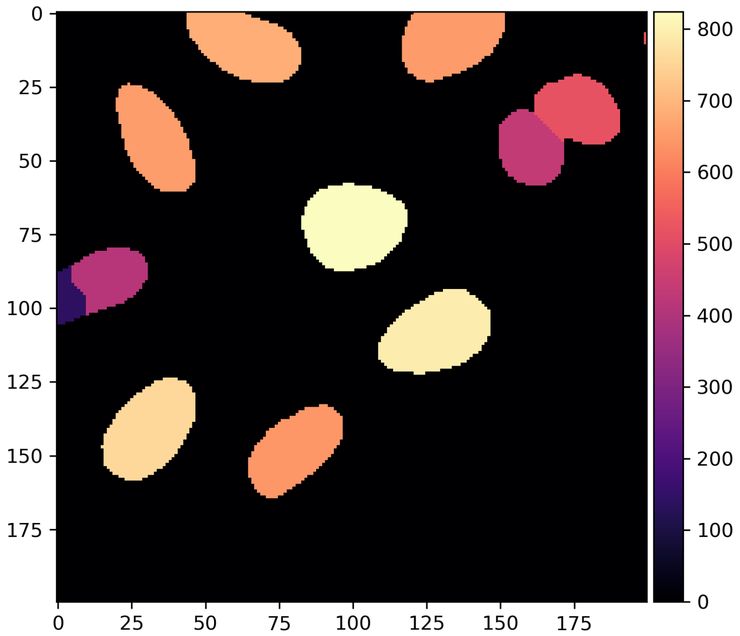
`labeled_comprehension`

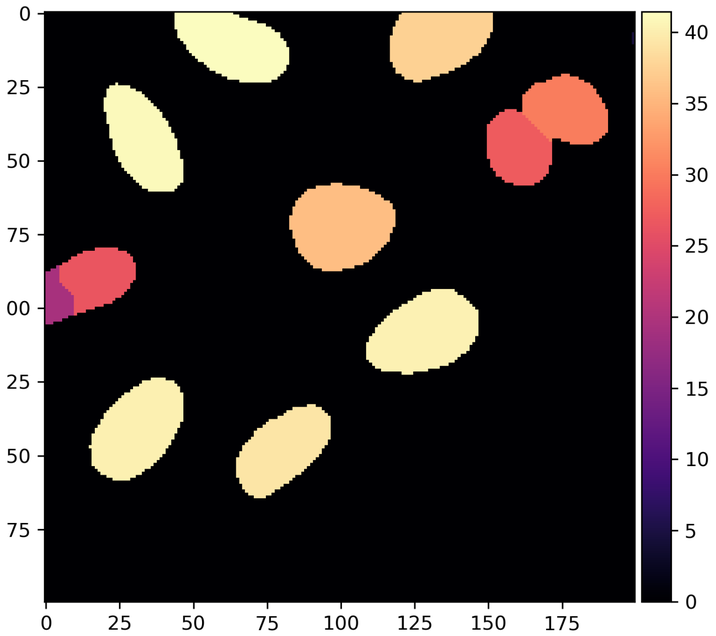
Index raw input image

mean, minimum, maximum, standard deviation, etc.

Size, Shape & Measurement Robustness







Discretisation impacts measurements!

size

2D: area & perimeter

3D: volume & surface area

shape

axis length, circularity,

aspect ratio, eccentricity, etc.

Perimeter Calculation

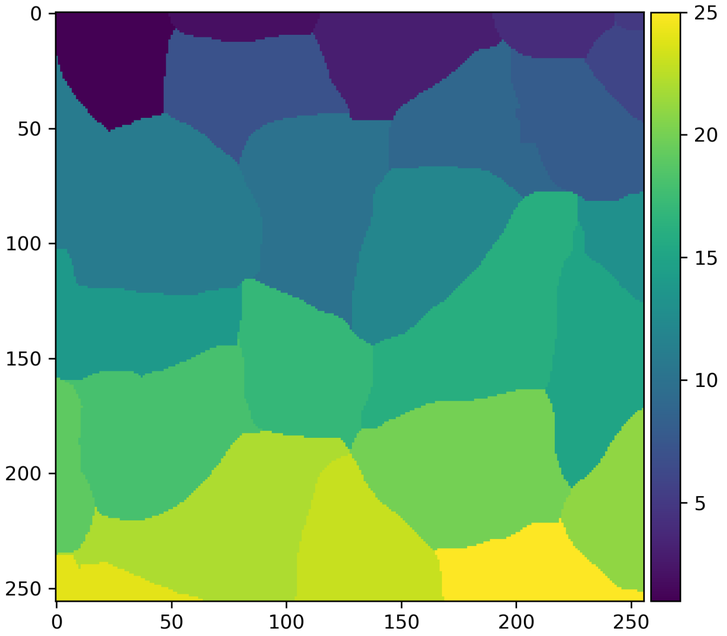
sphericity

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

roundness

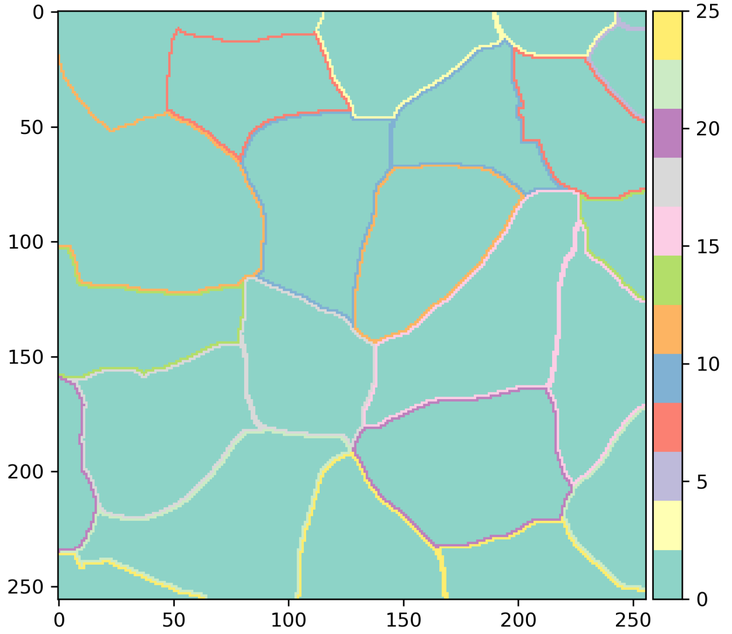
$$R = \frac{4\pi A}{P^2}$$

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



extract boundaries

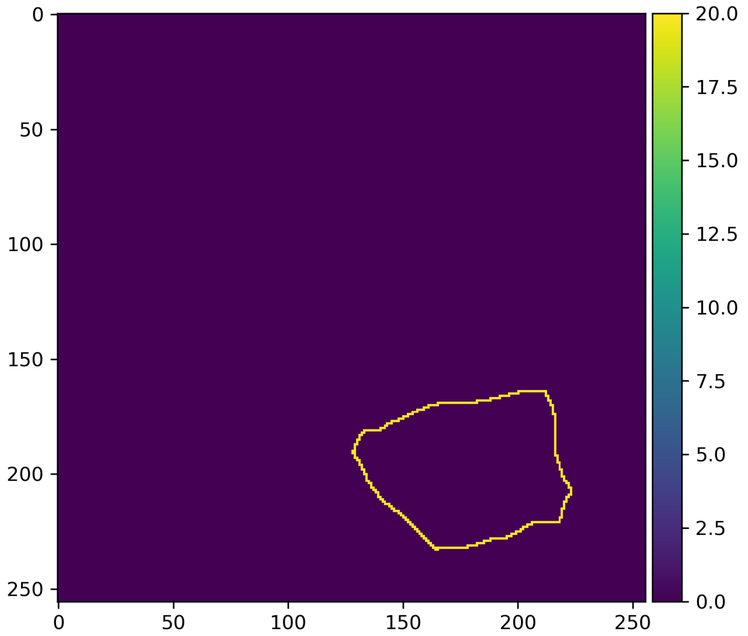
`find_boundaries`



Index segmentation

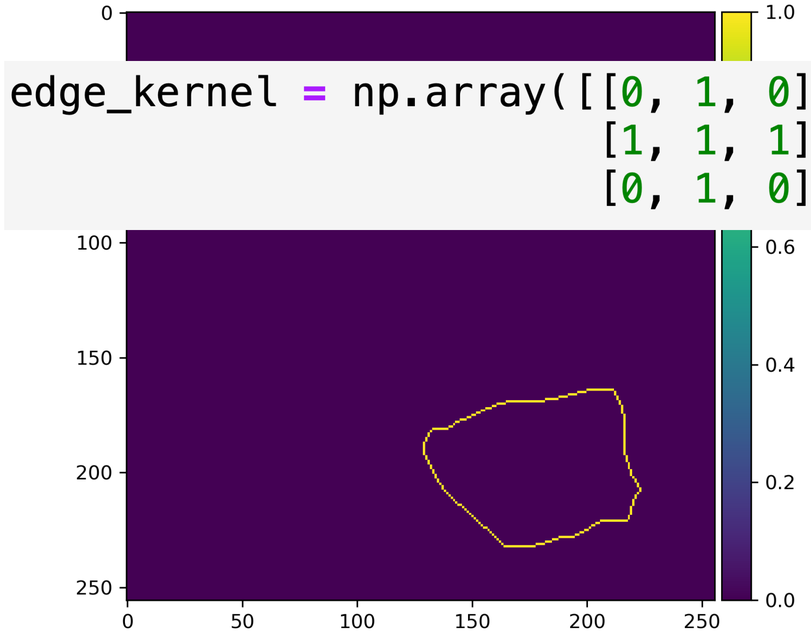
`cmap = 'Set3'`

Perimeter Calculation

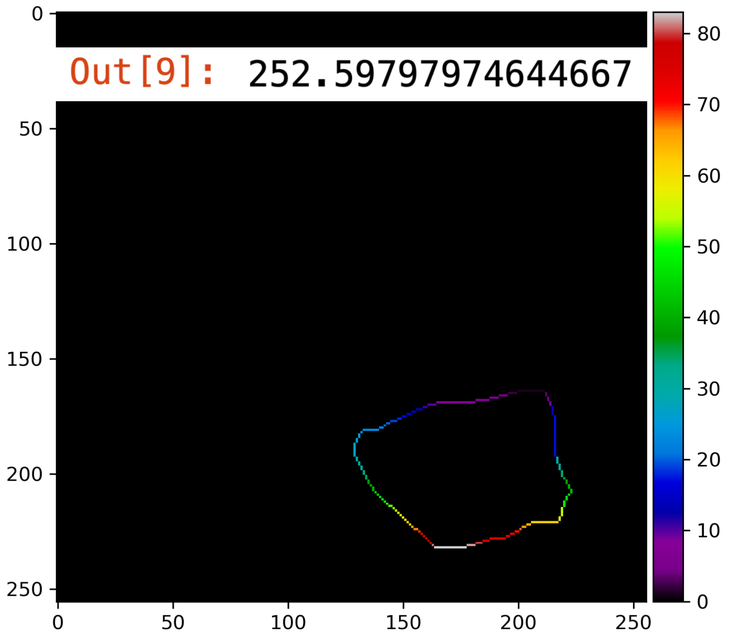


thin boundary

`skeletonize`




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

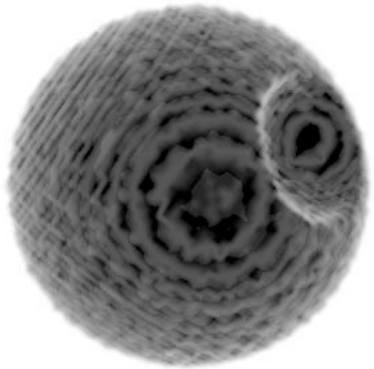


separate N_4 and N_8

boundary components

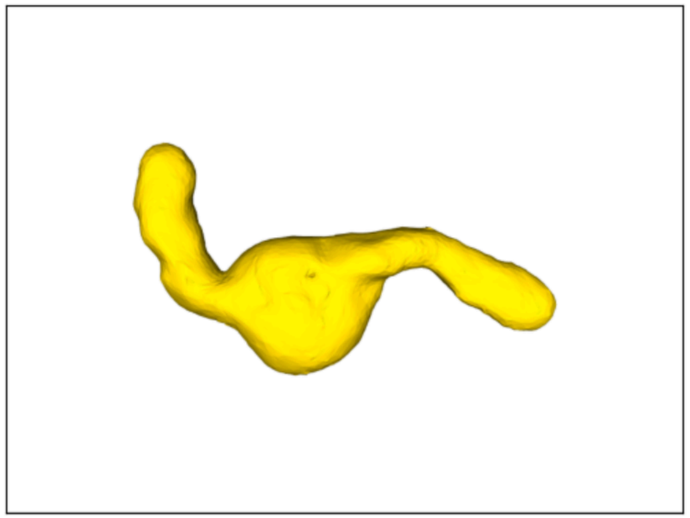
$\sqrt{2}$	1	$\sqrt{2}$
1		1
$\sqrt{2}$	1	$\sqrt{2}$

What if your object has a dent?



Meshes are concavity friendly tools.

Vedo enables quick feature extraction.



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

sphericity

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

range (0,1)

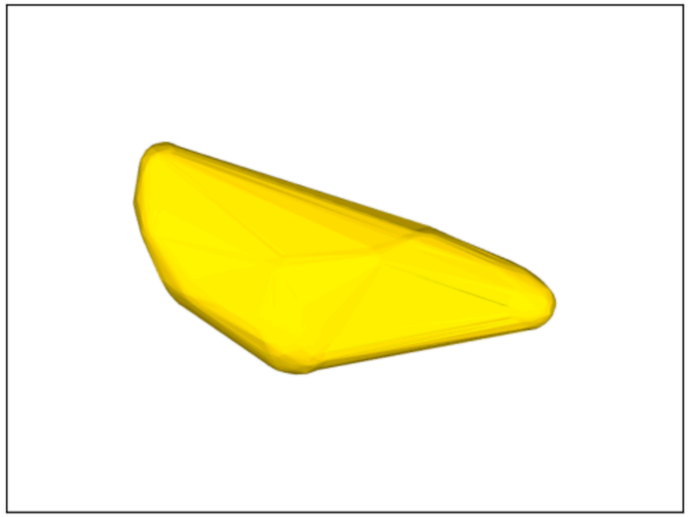
```
mesh_volume = mesh_gastruloid.volume()
```

```
mesh_surface_area = mesh_gastruloid.area()
```

```
sphericity = sphericity(mesh_volume, mesh_surface_area)
sphericity
```

```
0.5903099128683961
```

Convex Hulls as comparative objects



nppas.SurfaceTuple

origin (z/y/x)	[0. 0. 0.]
center of mass(z/y/x)	56.991,308.755,432.733
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	224.553
number of vertices	818
number of faces	1632

solidity

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

range (0,1)

related structures are the Delaunay triangulation & the Voronoi diagram

↓

positional features

Is there interest in positional features?

- centroid
- centre of mass

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

- bounding box

- number of neighbours
- giant component fraction

$$\mathbb{E}[k^2] - 2\mathbb{E}[k] > 0 \; or \; G'(1) = 1$$

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

Keep an eye on **FocalPlane** for Mara's upcoming feature extraction blog!

https://focalplane.biologists.com/