UT Southwestern Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

# Image processing with Python

[web] portal.biohpc.swmed.edu
[email] biohpc-help@utsouthwestern.edu



*Updated for 2023-09-20*

# Python image processing workflows

1

webGUI/ webGPU

Create conda env and install python libraries

2


OnDemand Jupyter

Activate the conda env and run the workflow

https://portal.biohpc.swmed.edu/media/filer_public/18/86/18864d7a-28ca-4ae5-be84-3e94b7c3bc4b/software_installation_2023_09_13.pdf

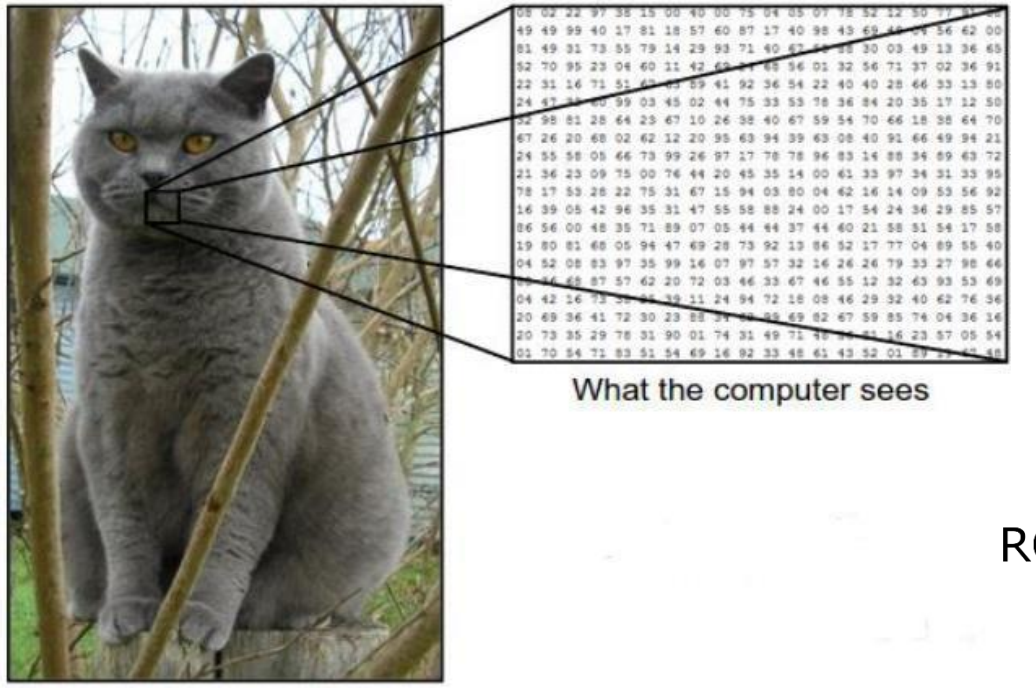
# Python Libraries + Modules used in this training

**Docs:**

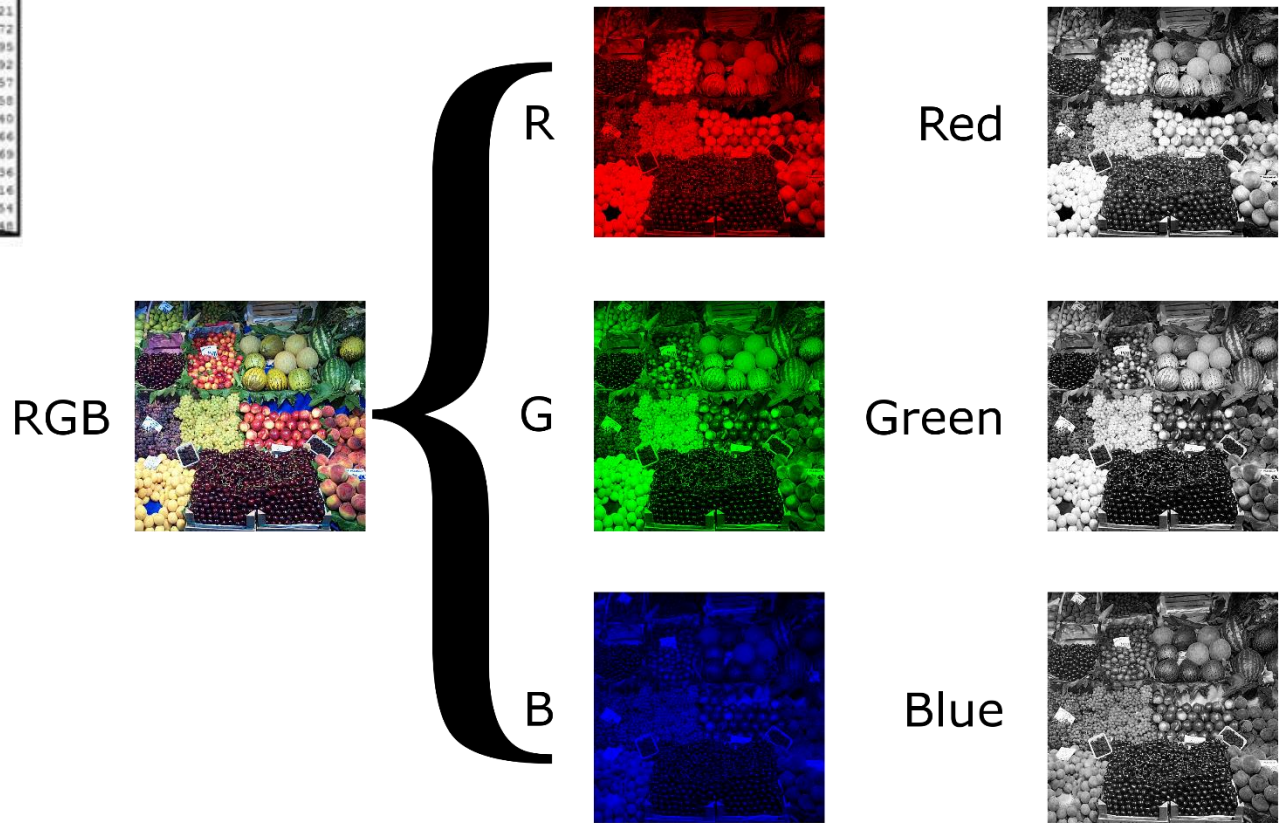
- os : https://docs.python.org/3/library/os.html
- matplotlib : https://matplotlib.org/
- scipy :
  - General : https://docs.scipy.org
  - ndimage : https://docs.scipy.org/doc/scipy/reference/ndimage.html
- skimage : https://scikit-image.org/
- sklearn : https://scikit-learn.org/stable/
- numpy :
  - General : https://numpy.org/doc/stable/index.html
  - ndarrays : https://numpy.org/doc/stable/reference/arrays.ndarray.html#id1

**Already installed in Jupyter/JupyterLab OnDemand.**

# How a digital image is stored on a computer



What the computer sees




Source - http://cs231n.github.io/

# Images as Arrays

**Different Python libraries have different array implementations**

- array
- **numpy**
  - ndarray
- **openCV**
  - cv::Mat

**Common data types for image pixels:**

- **bool** (binary)- [0,1]
- **int8** (signed integer 8 bit) – numbers in the range: [-128 : 127]
- **float** (double-precision floating point) – Decimal numbers (e.g. 2.2251e-308, 0.4, 0.33...)
- **uint8** (unsigned 8-bit) – [0,255]
- **uint16** (unsigned 16-bit) – [0,65535]

# Python Array Indexing

- Python starts counting indexes from 0, and arranges coordinates like C does (row-major)
- Array elements can be access in two ways:
  - By forward indexing
  - By backward indexing

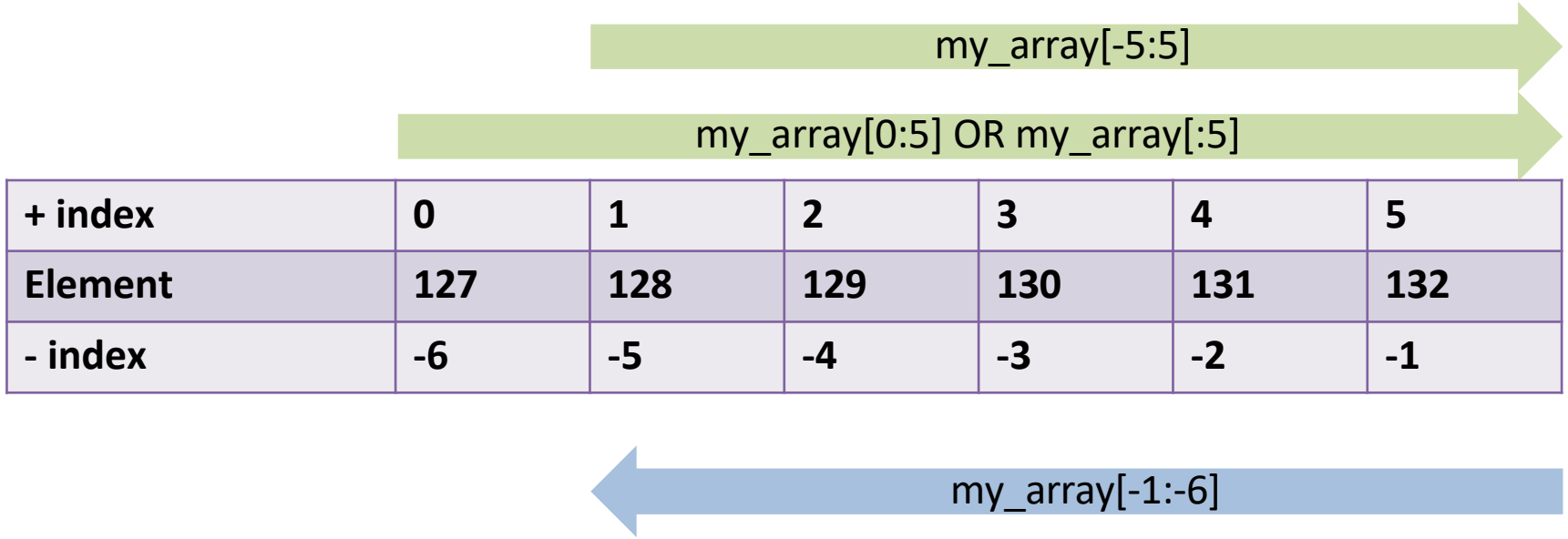
```
my_array = numpy.array([127, 128, 129, 130, 131, 132],
dtype=np.int8)
```

| + index | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Element | 127 | 128 | 129 | 130 | 131 | 132 |
| - index | -6 | -5 | -4 | -3 | -2 | -1 |

**Slice indexes are defined by [Start:Stop] or [Start:Stop:Step] (Stop not included)**

# Python Array Indexing

```
my_array = numpy.array([127, 128, 129, 130, 131, 132], dtype=np.int8)
```

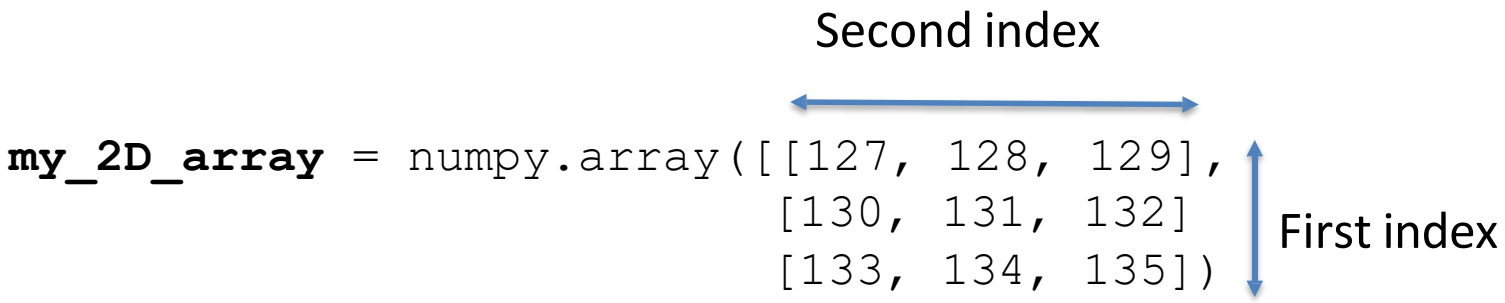



| + index | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Element | 127 | 128 | 129 | 130 | 131 | 132 |
| - index | -6 | -5 | -4 | -3 | -2 | -1 |

# Multi-dimensional arrays – numpy arrays

Python counts in ‘row-major’ ordering, and orders dimensions like C does.

- Multidimensional arrays are ‘lists of lists’
- This is in fact how elements are stored in memory

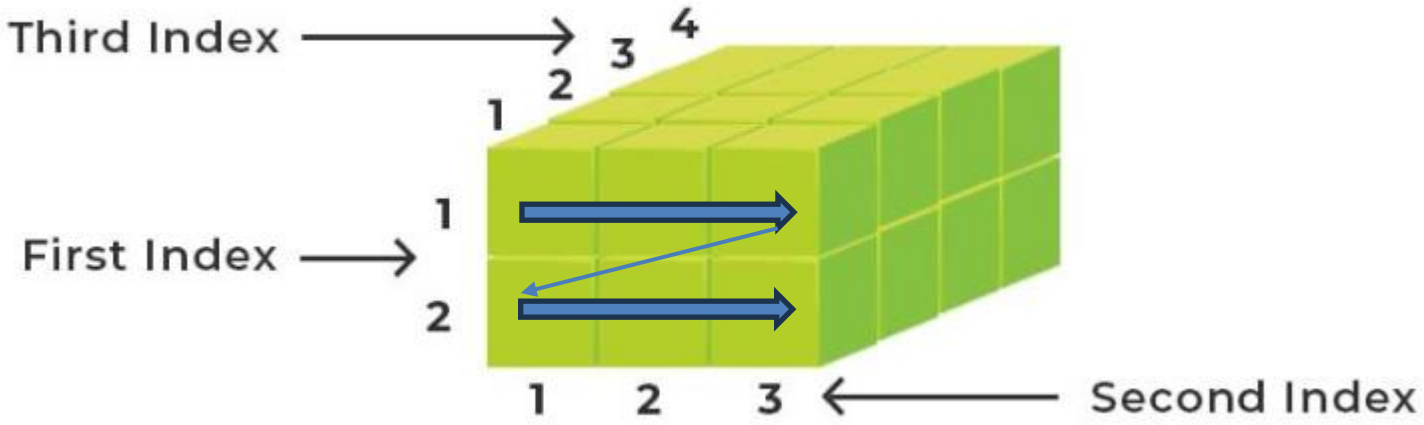



```
my_2D_array[1][0:1] = [130, 131]
my_2D_array[-1][:] = [133, 134, 135]
```

# Multi-dimensional arrays – numpy arrays

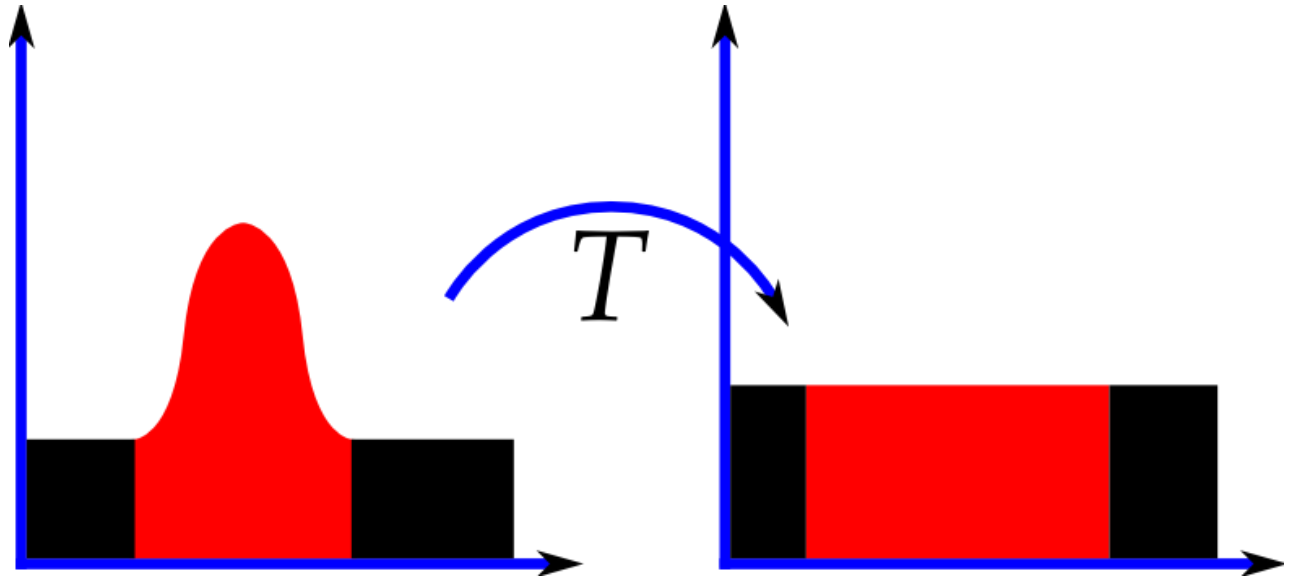
Python counts in ‘row-major’ ordering, and orders dimensions like C does.

- Multidimensional arrays are ‘lists of lists’
- This is in fact how elements are stored in memory



# Intensity enhancement

- Contrast stretching
- Histogram equalization
- Adaptive equalization

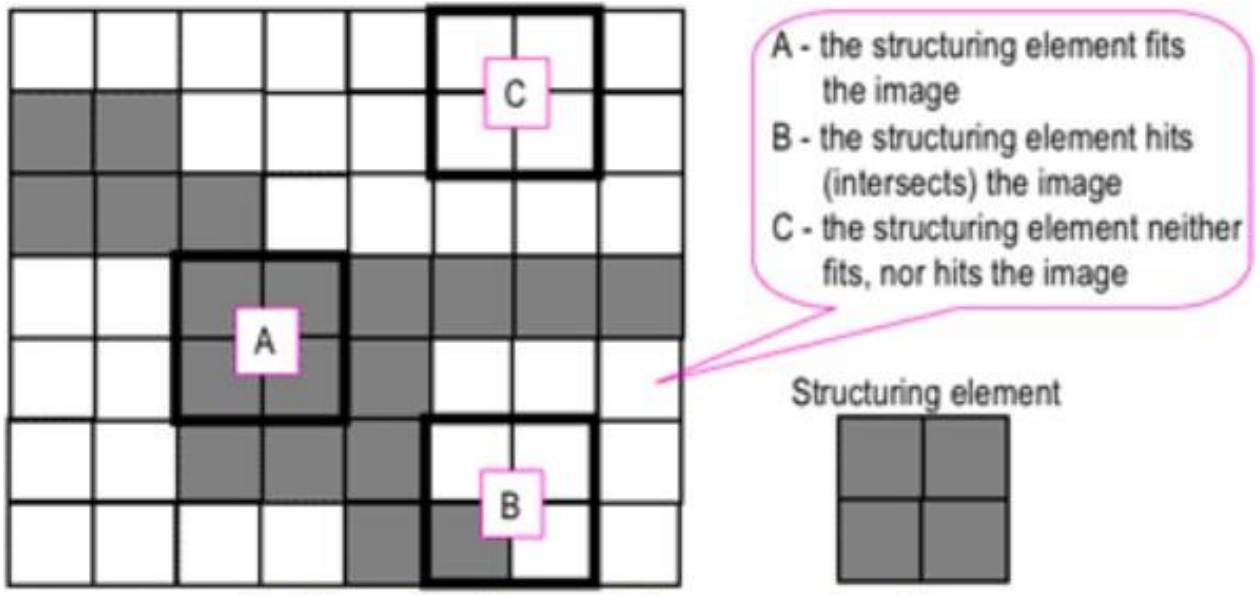



https://en.wikipedia.org/wiki/Histogram_equalization

# Morphological operations: Structuring element

The structuring element is a small binary image or matrix such that:

- The matrix dimensions specify the size of the structuring element.
- The pattern of ones and zeros specifies the shape of the structuring element.




Probing of an image with a structuring element
(white and grey pixels have zero and non-zero values, respectively).

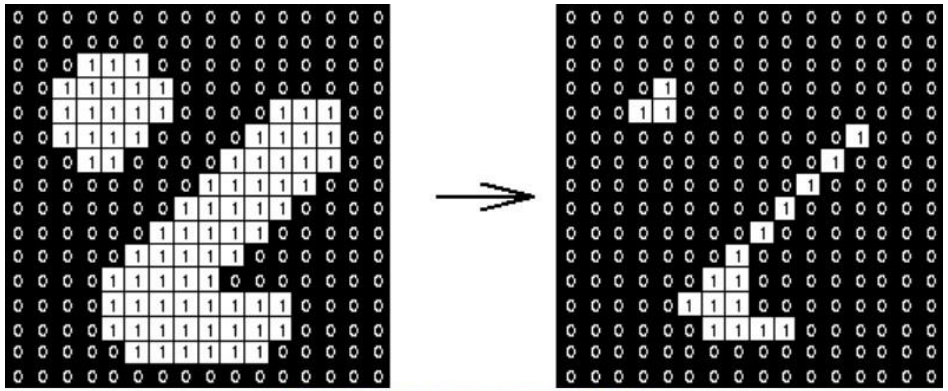
https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm

# Mathematical Morphology - Grayscale

- Grayscale images can be treated similarly, but with a slightly modified interpretation of ‘hit or miss’
- Dilation will result in a pixel taking on the max value defined by the moving window of the strel.
- Erosion will result in a pixel taking on the min value defined by the moving window of the strel.

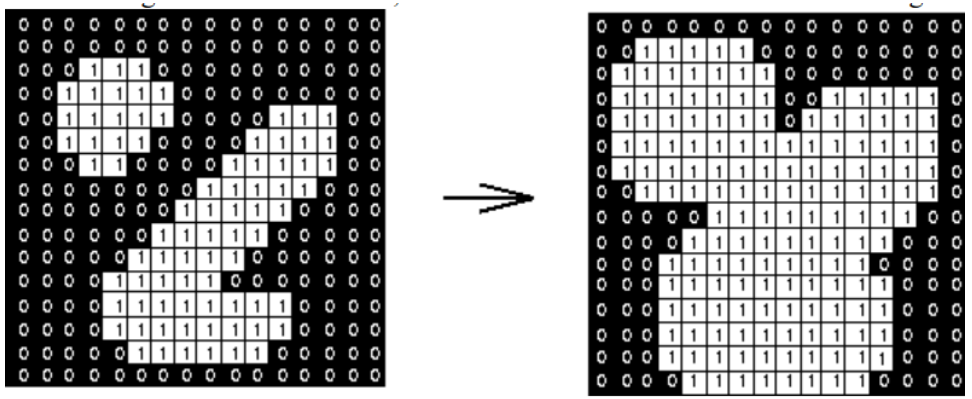
# Morphological operations: Dilation and Erosion

**Erosion**:



Erosion: a 3×3 square structuring element
(www.cs.princeton.edu/~pshilane/class/mosaic/).

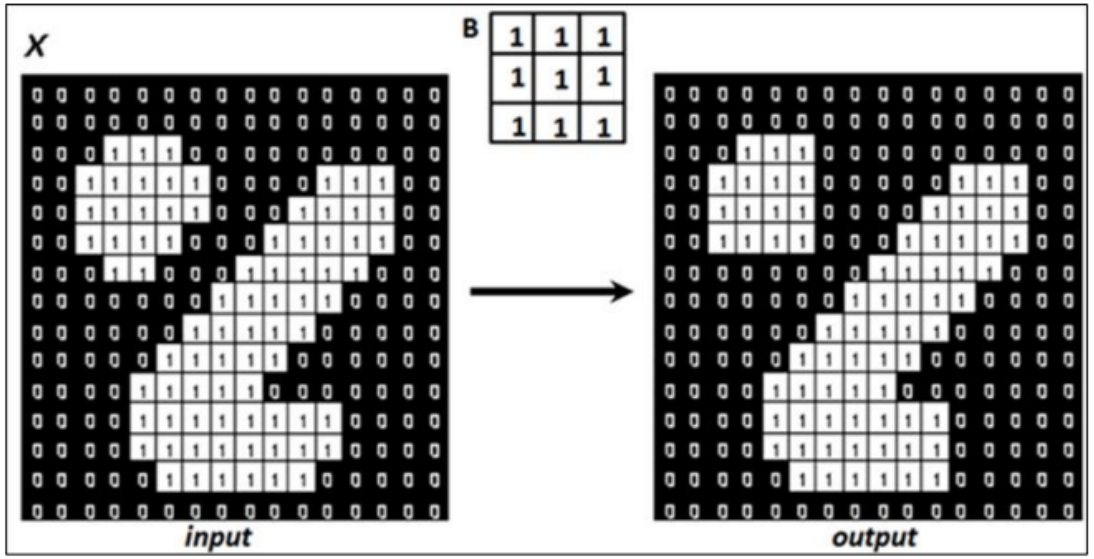
**Dilation:**



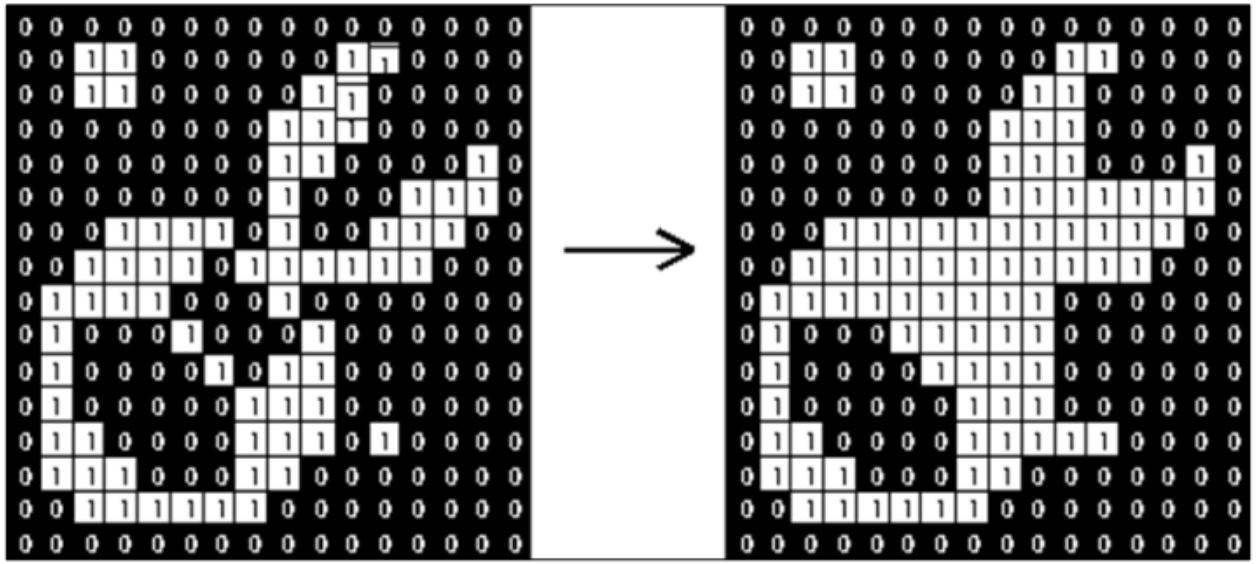
Dilation: a 3×3 square structuring element
(www.cs.princeton.edu/~pshilane/class/mosaic/).

# Morphological operations: Open and Close

Opening: erosion followed by a dilation




Closing: dilation followed by a erosion



- Opening an image smoothes its contour, fractures narrow isthmuses, making items more separated
- Closing fills in small gaps/holes and brings items closer, also smoothes the contour

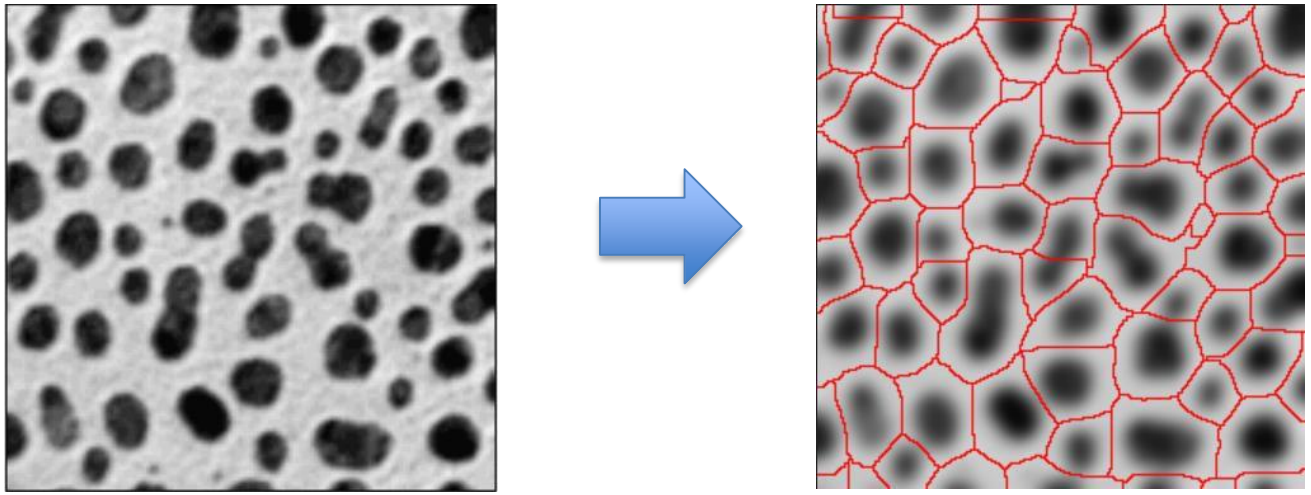
# Segmentation – Separating an image into parts

Most basic: Manual thresholding

- Bright or dark background with a dark or bright foreground, respectively.
- Choose a cutoff value, threshold.
- Using global thresholds may miss important elements

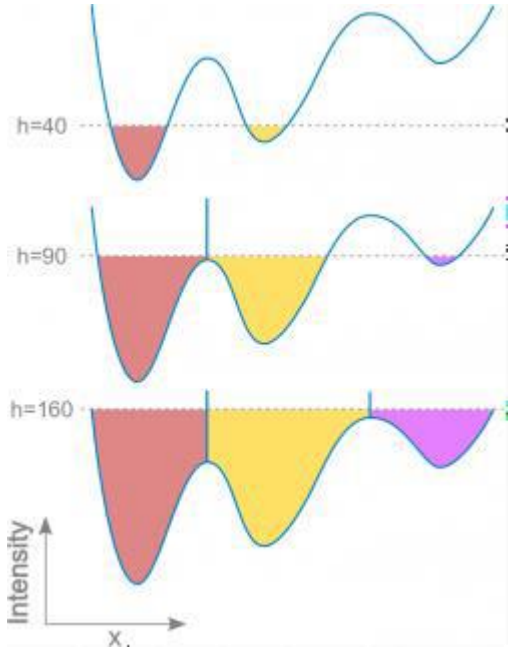
More complex:

- Adaptive thresholds
- Morphological segmentation
- Clustering
- Machine learning methods



# Morphological Segmentation: Watershed method

- Consider grey levels as altitudes
- Identify local minima
- Flood basins starting from minima
- Separate the basins by a “dam” → the watershed




**Steps for performing the watershed method:**

1. Segment objects of interest
2. Convert the mask into an intensity profile using the distance transform
3. Run the watershed algorithm

# More useful materials

- Previous image processing using Matlab slides (More intro on filters) https://portal.biohpc.swmed.edu/media/filer_public/61/34/6134df89-c5b8-4efd-9f60-fbc1c5005bb0/training_matlab_2022_10_19.pdf

- Image Processing with Python https://datacarpentry.org/image-processing/