Yeast Cells Segmentation & Classification

Student: Nizar Cherkaoui

Advisor: Dr. Chia-Ling Tsai (Computer Science Dept.)

Advisor: Dr. Eric Muller (Biology Dept.)

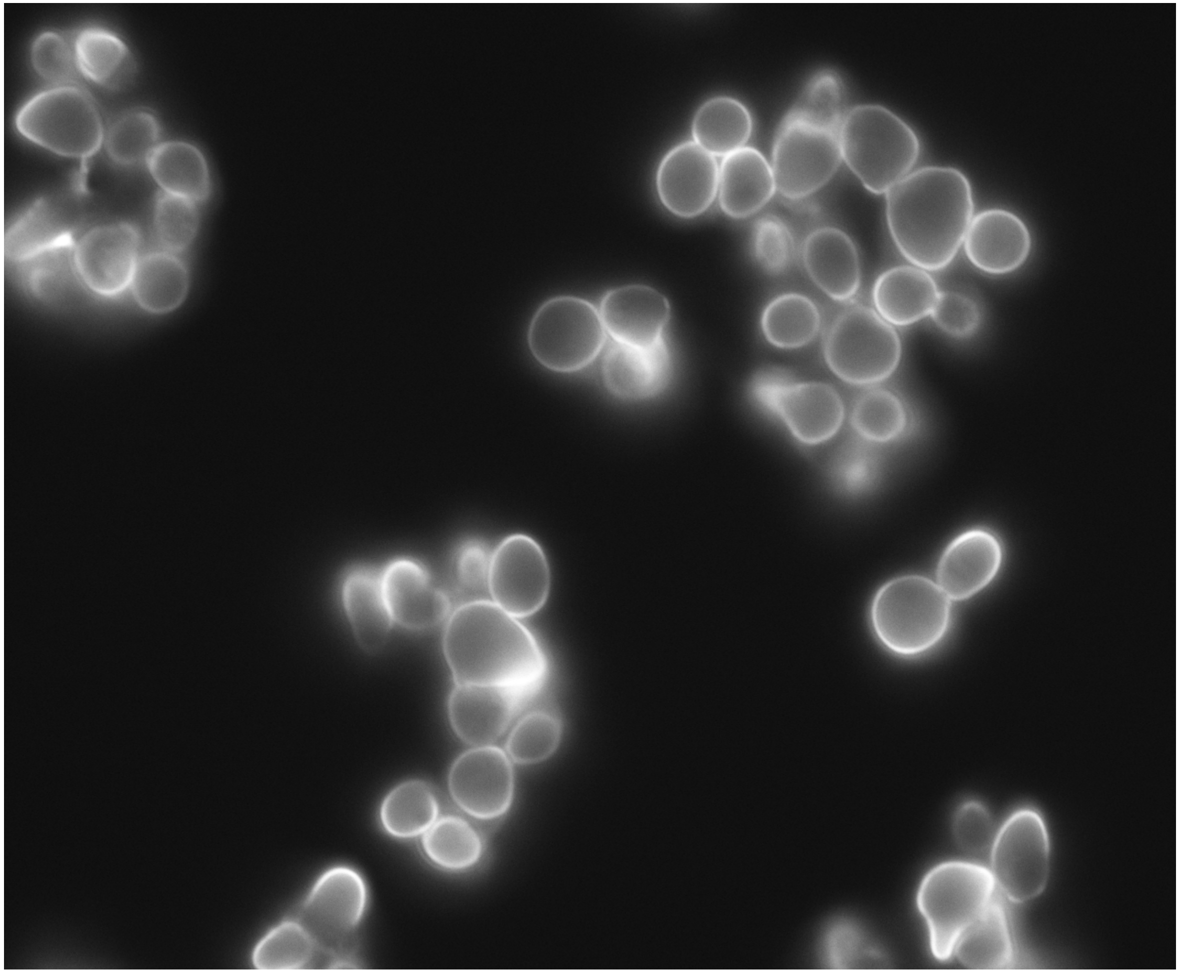
Outline

- Introduction
- Foreground Extraction
- Blob Segmentation and Labeling
- Classification

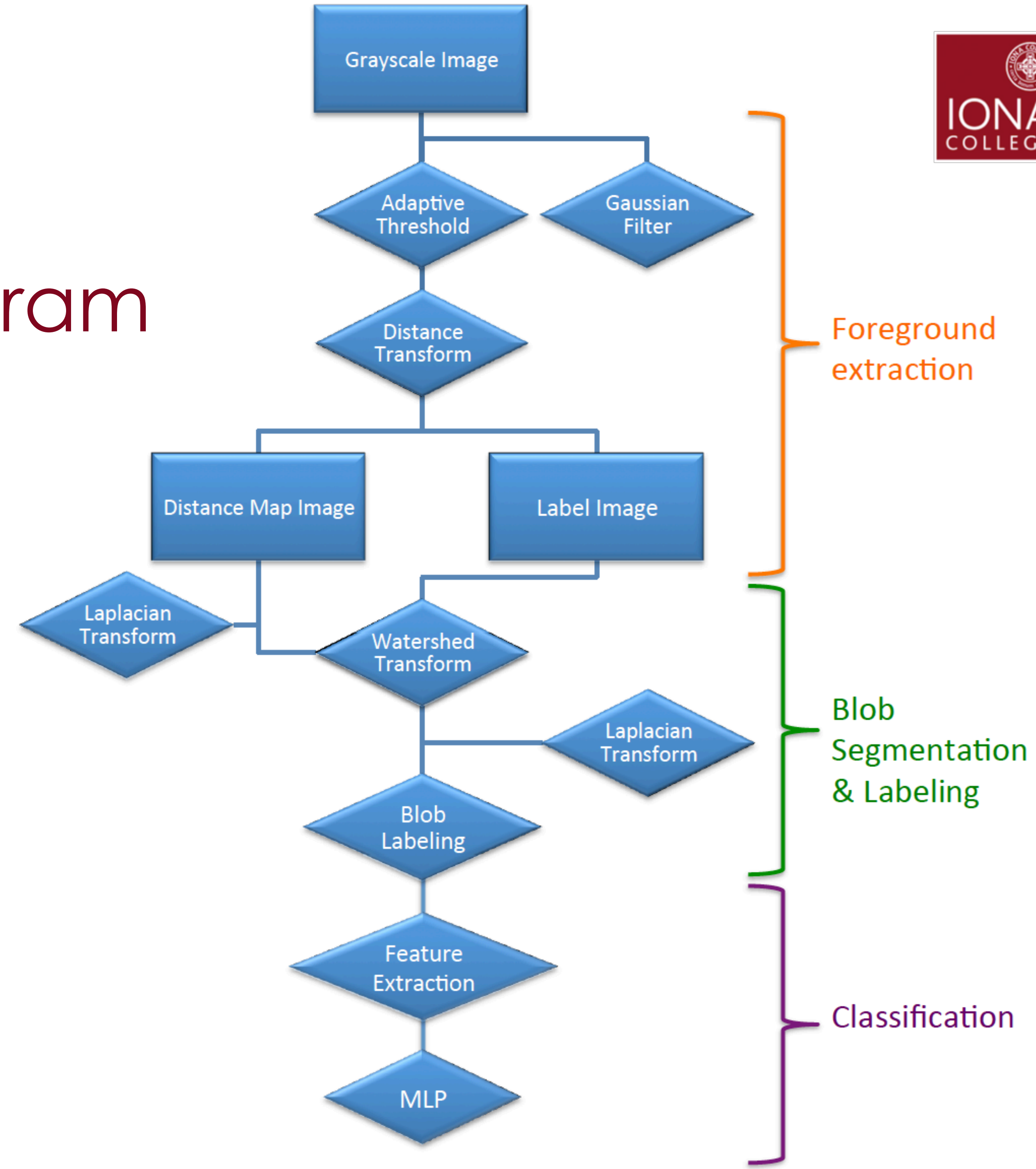
Introduction

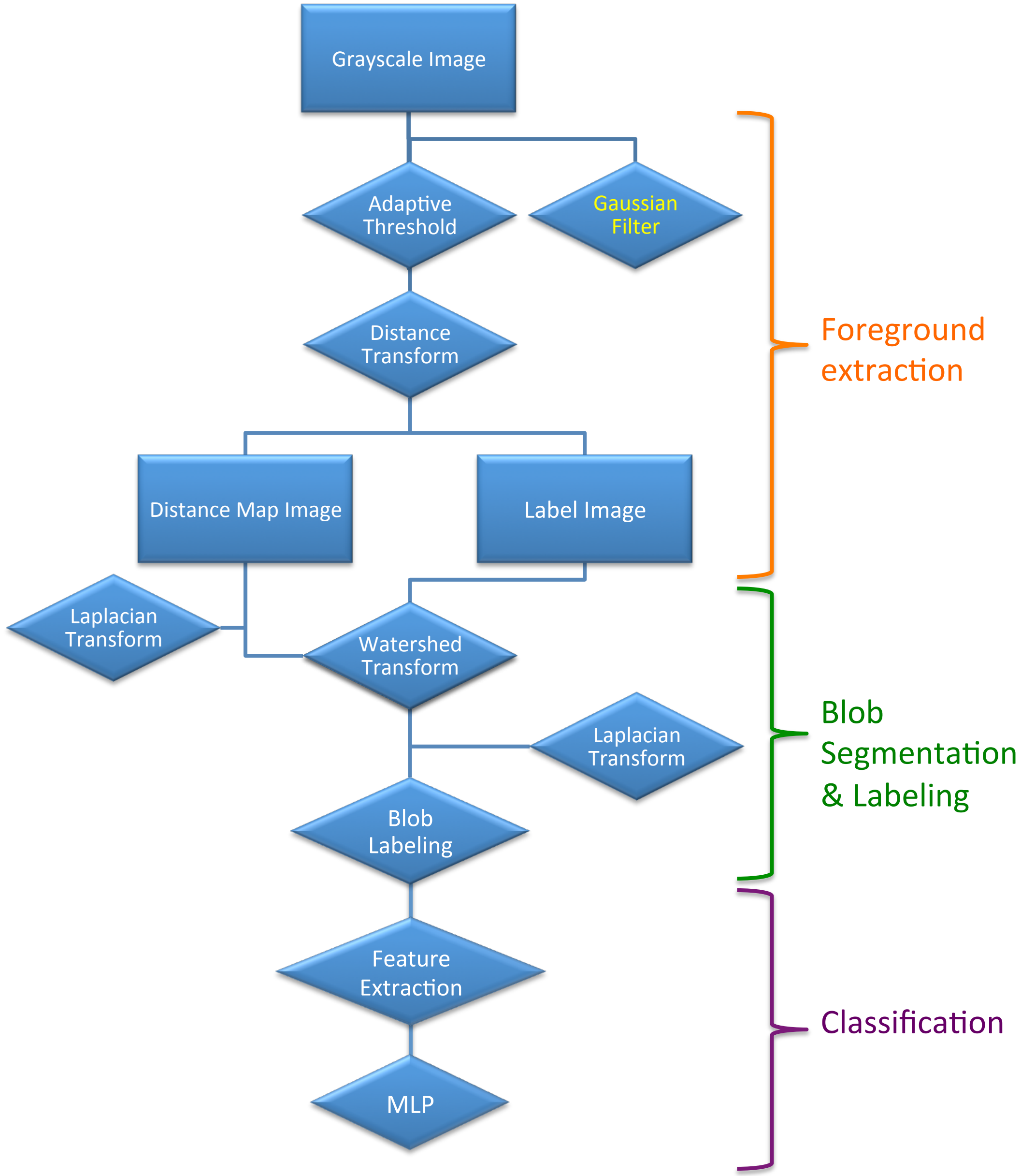
- Image segmentation in Biology has been used in the analysis of microscopy based images.
- The aim is to automatically process huge amount of image samples and produce useful data.
- The challenge lies in:
  - Splitting apart cells that are densely packed and overlap each other.
  - Generate features that can be useful for classification

Introduction



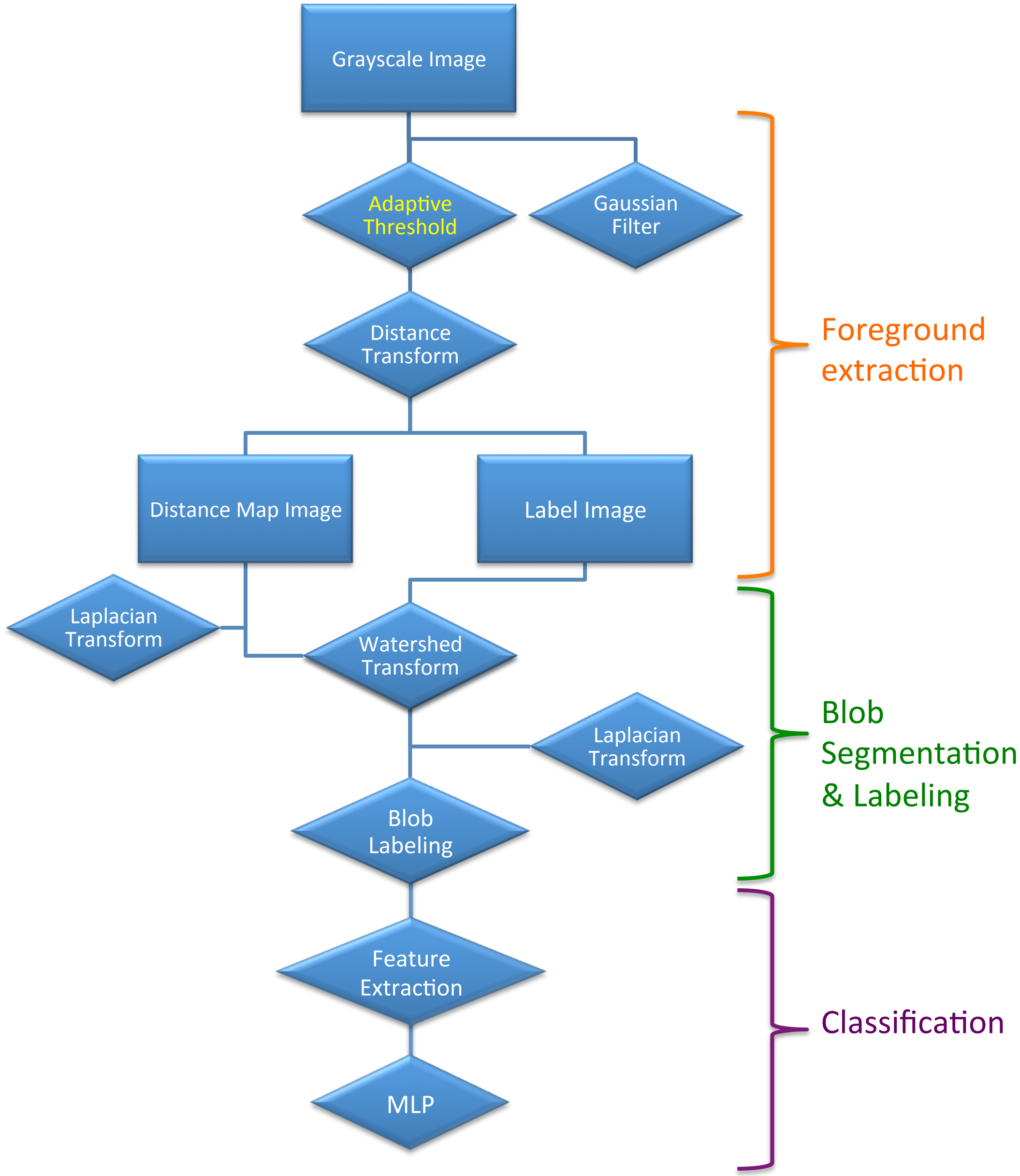
Flow Diagram




Grayscale Image
Adaptive Threshold
Gaussian Filter
Distance Transform
Distance Map Image
Label Image
Laplacian Transform
Watershed Transform
Laplacian Transform
Blob Labeling
Feature Extraction
MLP
Foreground extraction
Blob Segmentation & Labeling
Classification

Gaussian Smoothing

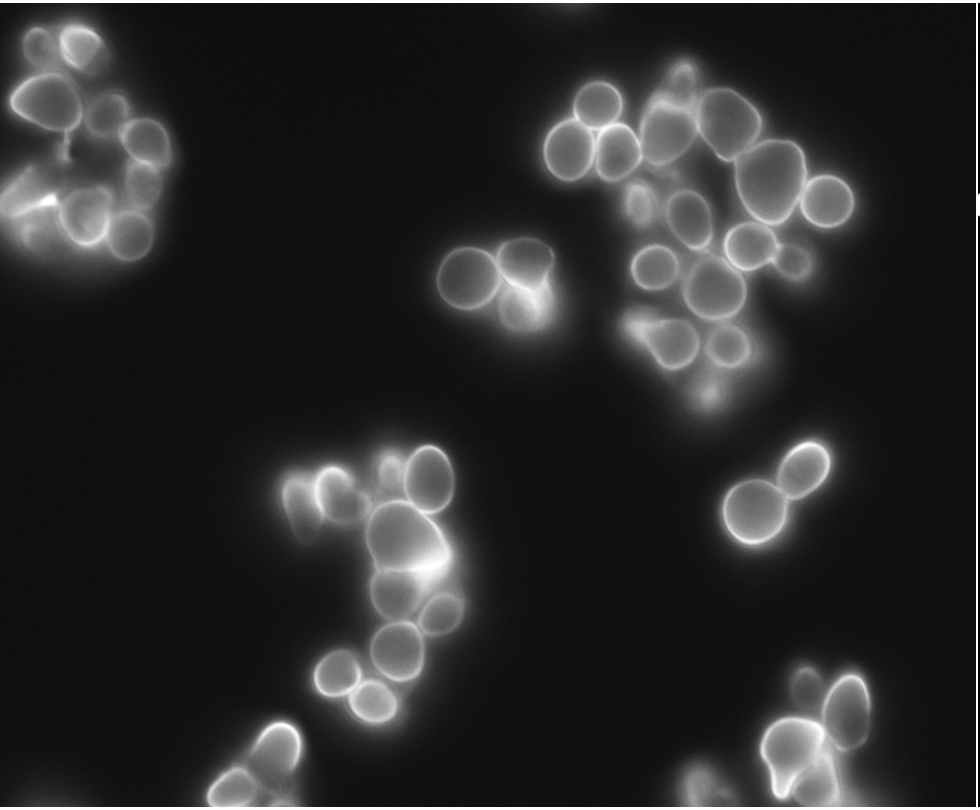
- Images have noise (unwanted information)
  - The intensity measured at a pixel is the "true" intensity plus noise.
  - nearby pixels usually have similar "true" intensities.
  - Smoothing the image reduces noise
- To perform smoothing a Gaussian filter is applied to the image.


Grayscale Image
Adaptive Threshold
Gaussian Filter
Distance Transform
Distance Map Image
Label Image
Laplacian Transform
Watershed Transform
Laplacian Transform
Blob Labeling
Feature Extraction
MLP
Foreground extraction
Blob Segmentation & Labeling
Classification

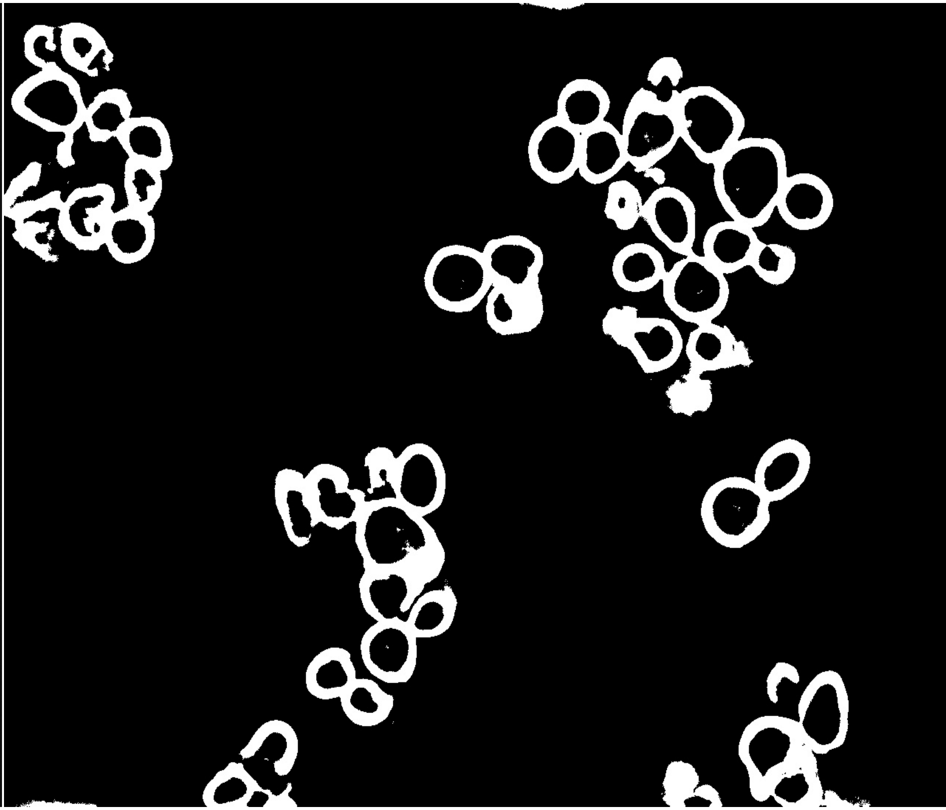
Adaptive Threshold

- To separate foreground objects with background.
- It is used to create a binary image
  - Object pixels are given value of '1'
  - Background pixels are given value of '0'
- Pixels are colored black or white depending on the pixel's label.
- Pixels are processed in a Gaussian window instead of individually.

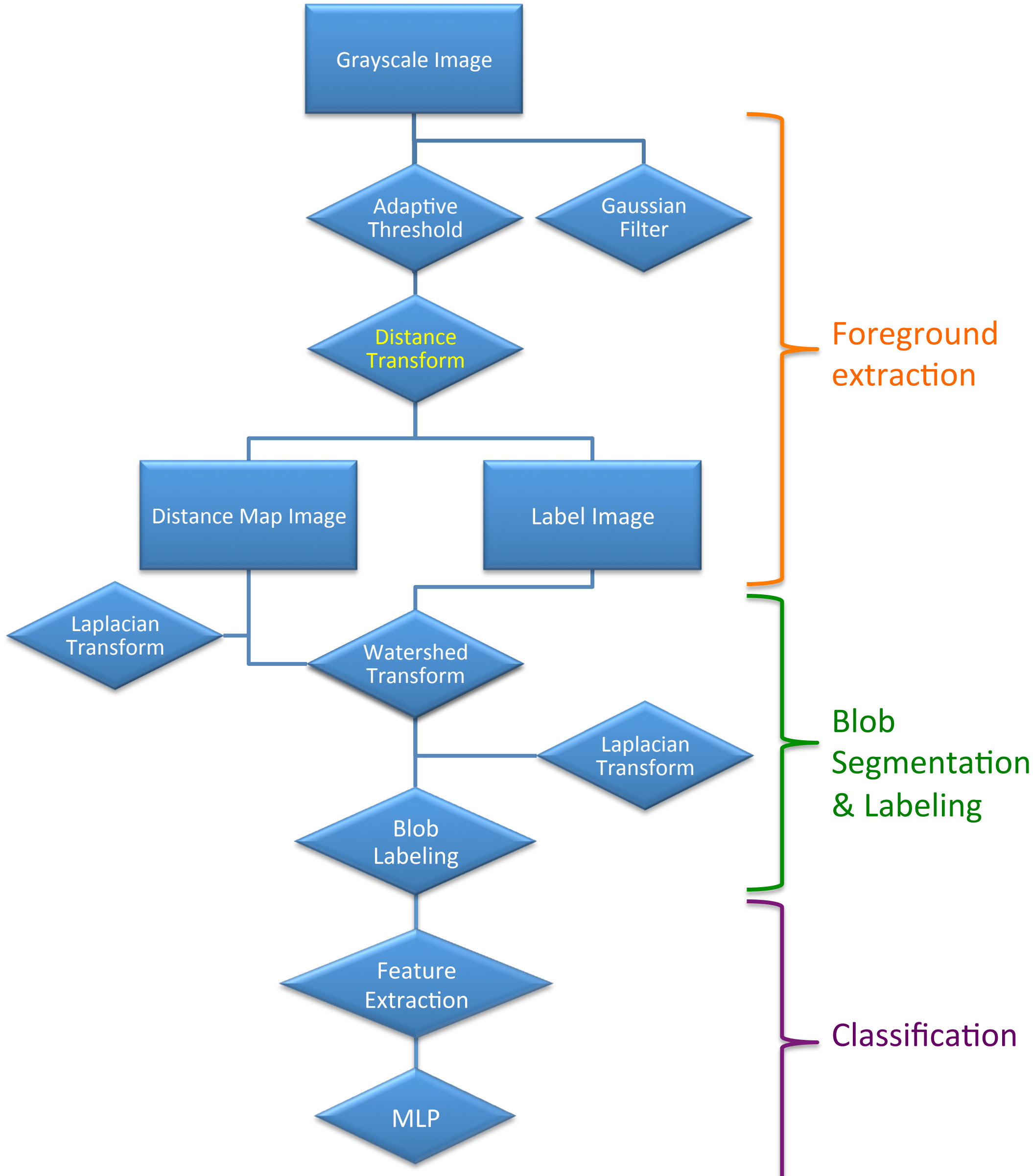
Adaptive Threshold



Original Image



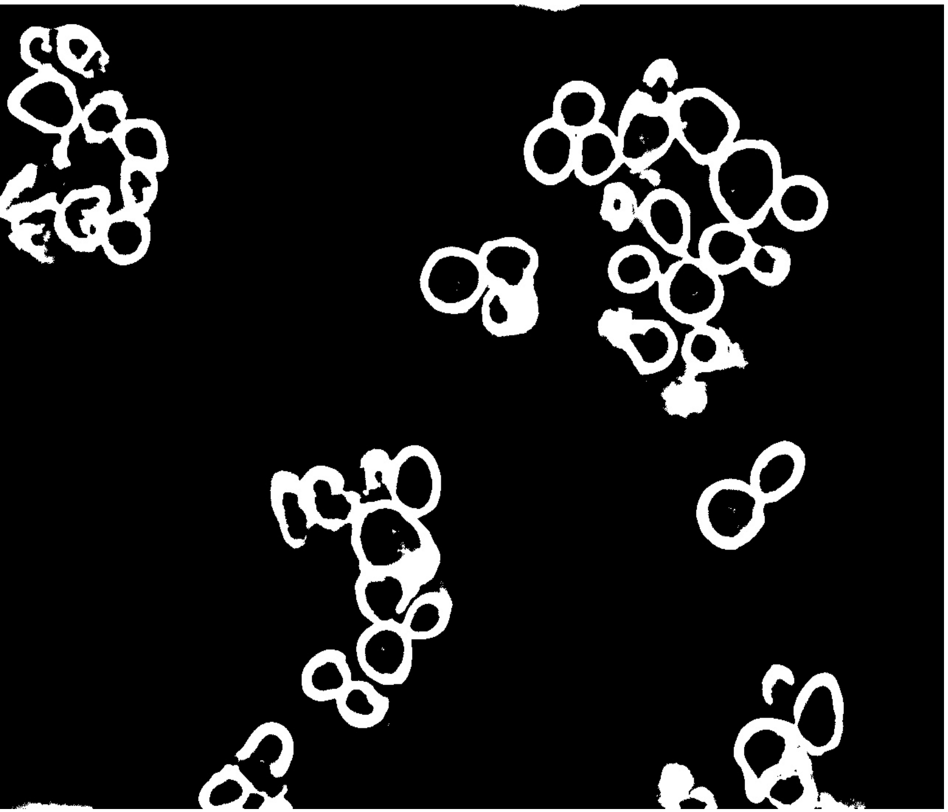
Adaptive Threshold


Grayscale Image
Adaptive Threshold
Gaussian Filter
Distance Transform
Distance Map Image
Label Image
Laplacian Transform
Watershed Transform
Laplacian Transform
Blob Labeling
Feature Extraction
MLP
Foreground extraction
Blob Segmentation & Labeling
Classification

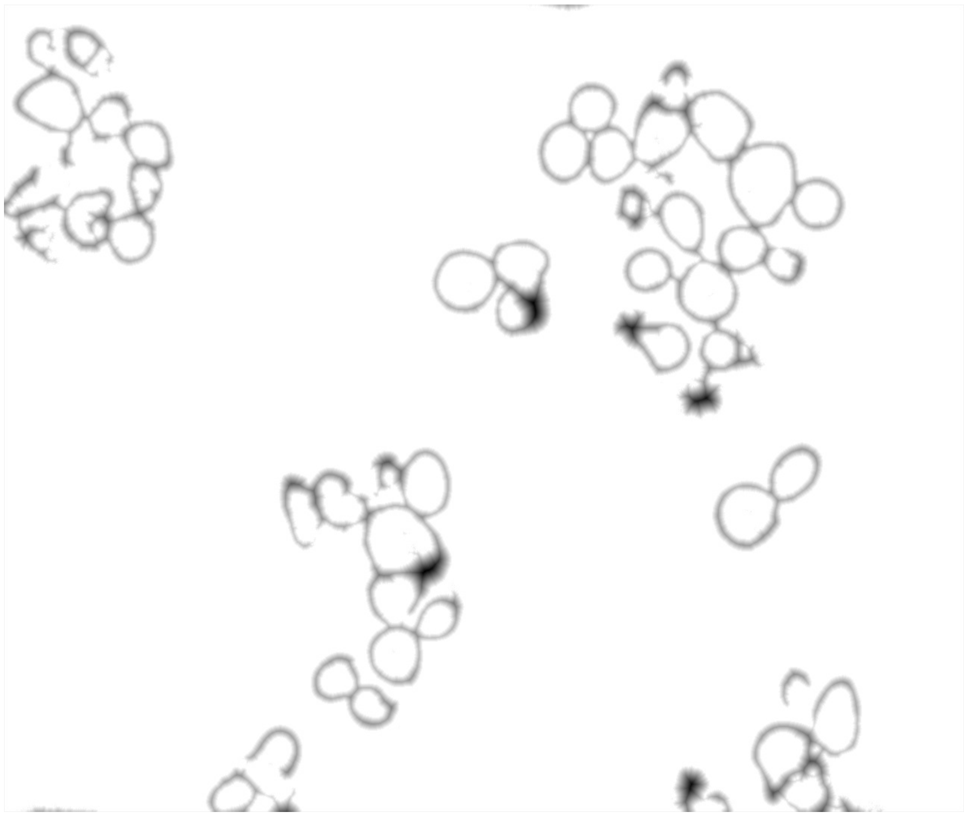
Distance Transform

- The function calculates the approximated distance from every binary image pixel to the nearest background pixel.
  - For background pixels the function sets the zero distance.
  - For foreground pixels the function sets an intensity proportional to the distance calculated.

Distance Transform



Adaptive Threshold

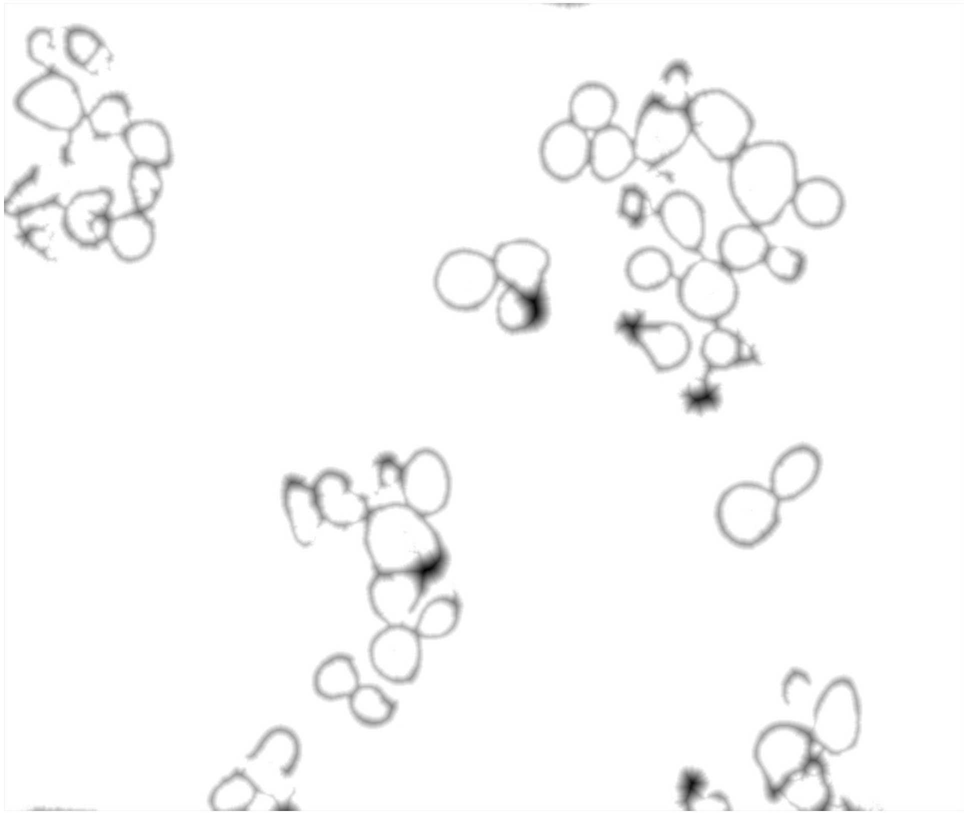


Distance Transform

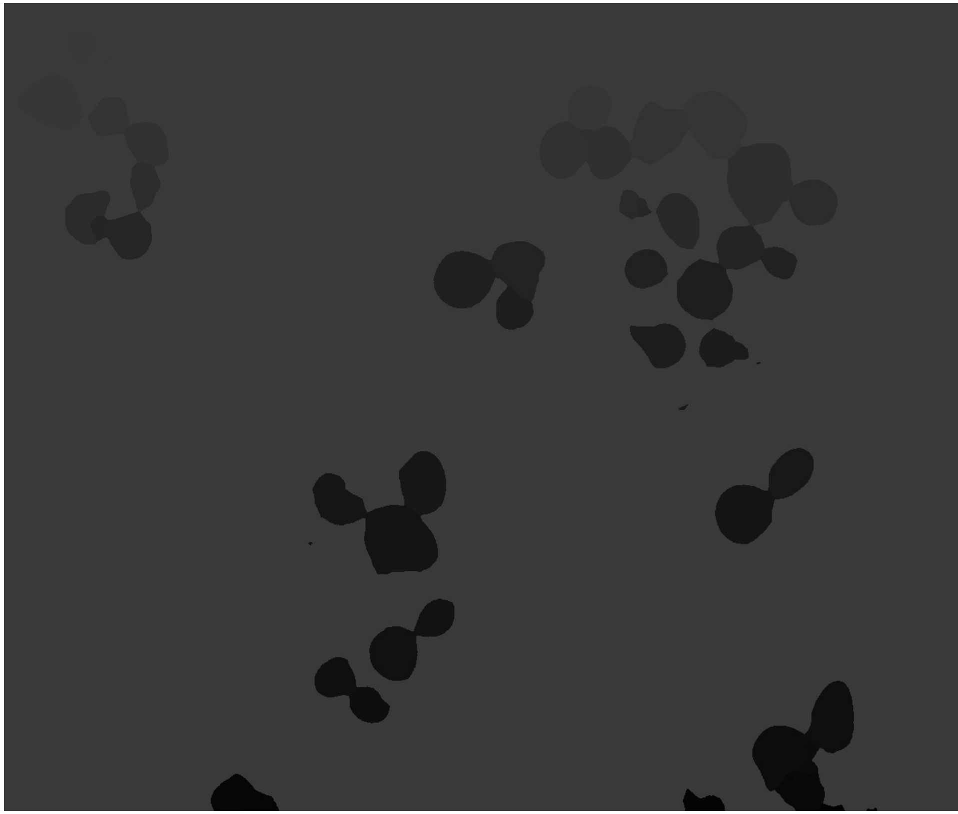
Distance Transform

- The distance Transform also generates an image of Labels.
  - For every foreground object the function finds the nearest connected component consisting of background pixels and uniquely labels it.
  - Using the label image, a gray level image is created containing the labels starting at 1, translated into gray scale intensities.
- The image generated is the result of Foreground and Background Extraction.

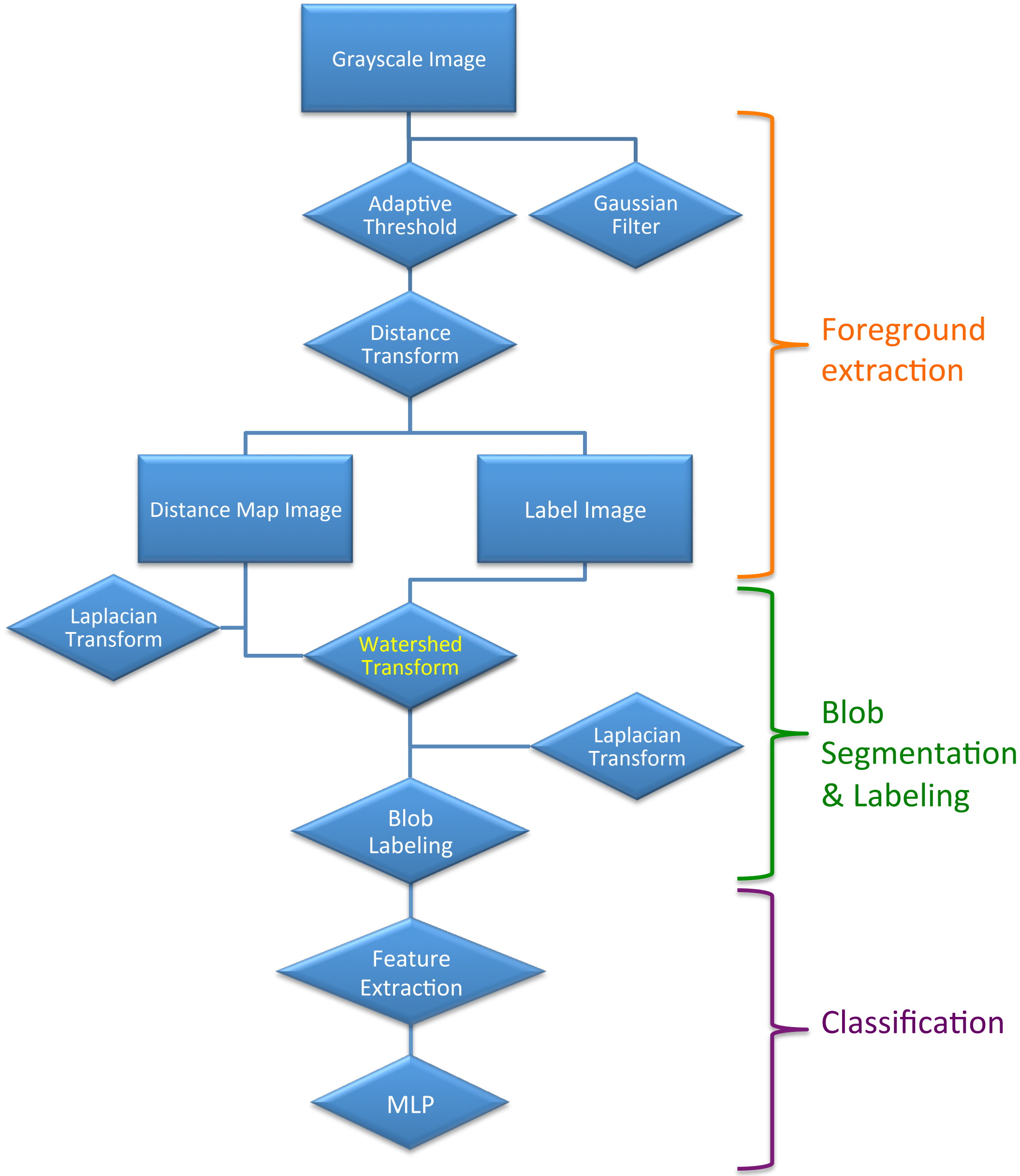
Distance Transform



Distance Transform



Label Image


Grayscale Image
Adaptive Threshold
Gaussian Filter
Distance Transform
Distance Map Image
Label Image
Laplacian Transform
Watershed Transform
Laplacian Transform
Blob Labeling
Feature Extraction
MLP
Foreground extraction
Blob Segmentation & Labeling
Classification

Marker-based Watershed

- Enhancement of the watershed algorithm
- Treats the input image as a topographic surface where:
  - Dark pixels are high.
  - Light pixels are low.
- Consists of flooding the topographic surface from predefined set of markers.
  - Each foreground object (cell) must have only one set of connected markers.

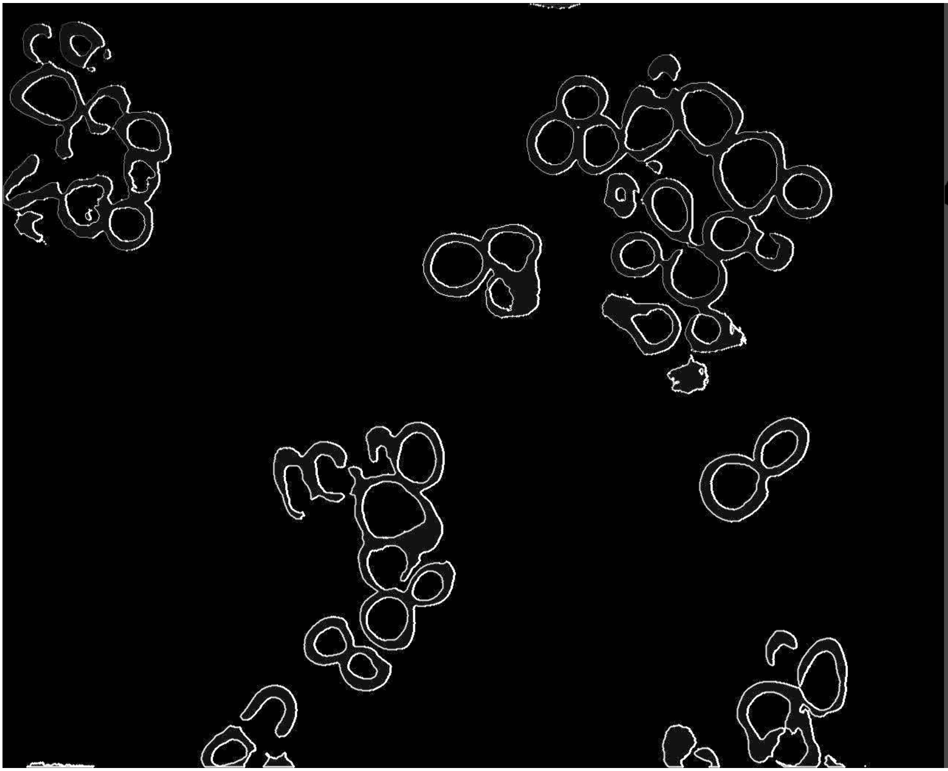
Marker-based Watershed

- To generate the watershed segmentation, pixels that certainly belong to the foreground and pixels that certainly belong to the background have to be identified.
  - The other pixels, that is the ones for which the labeling is unknown, have to be assigned a value 0.
- Foreground and Background pixels were previously labeled from 1 to 255.

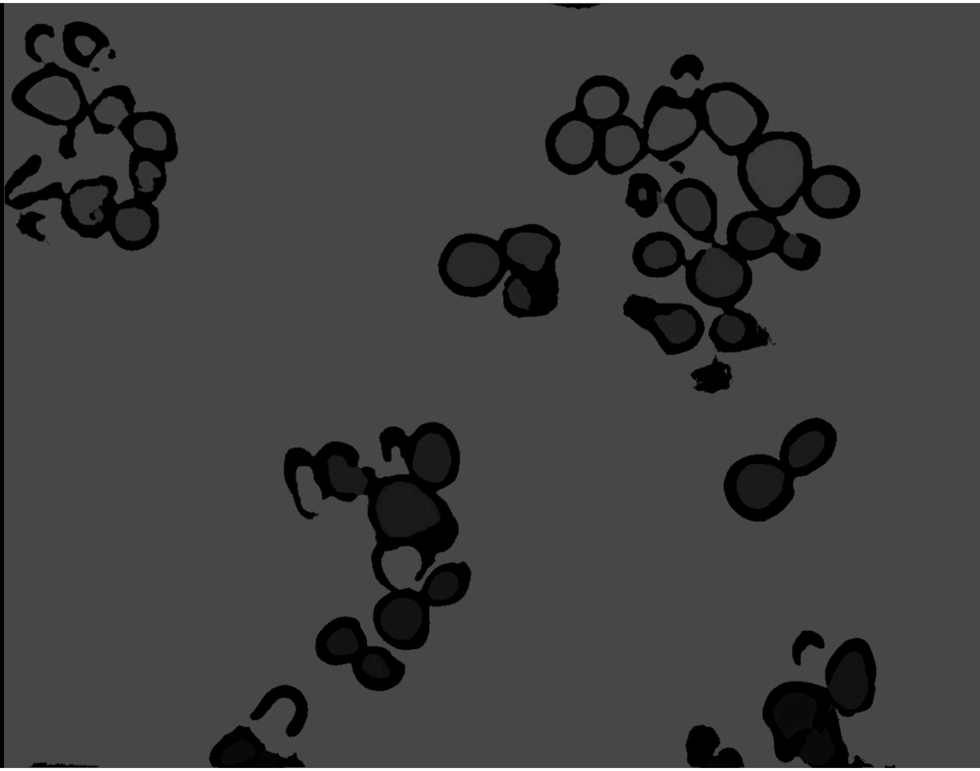
Marker-based Watershed

- In our case, The unknown pixels are those that represent the cell boundaries.
- The idea is to extract the boundaries from the Threshold image and to redraw them on the Foreground and Background labeled image.

Marker-based Watershed

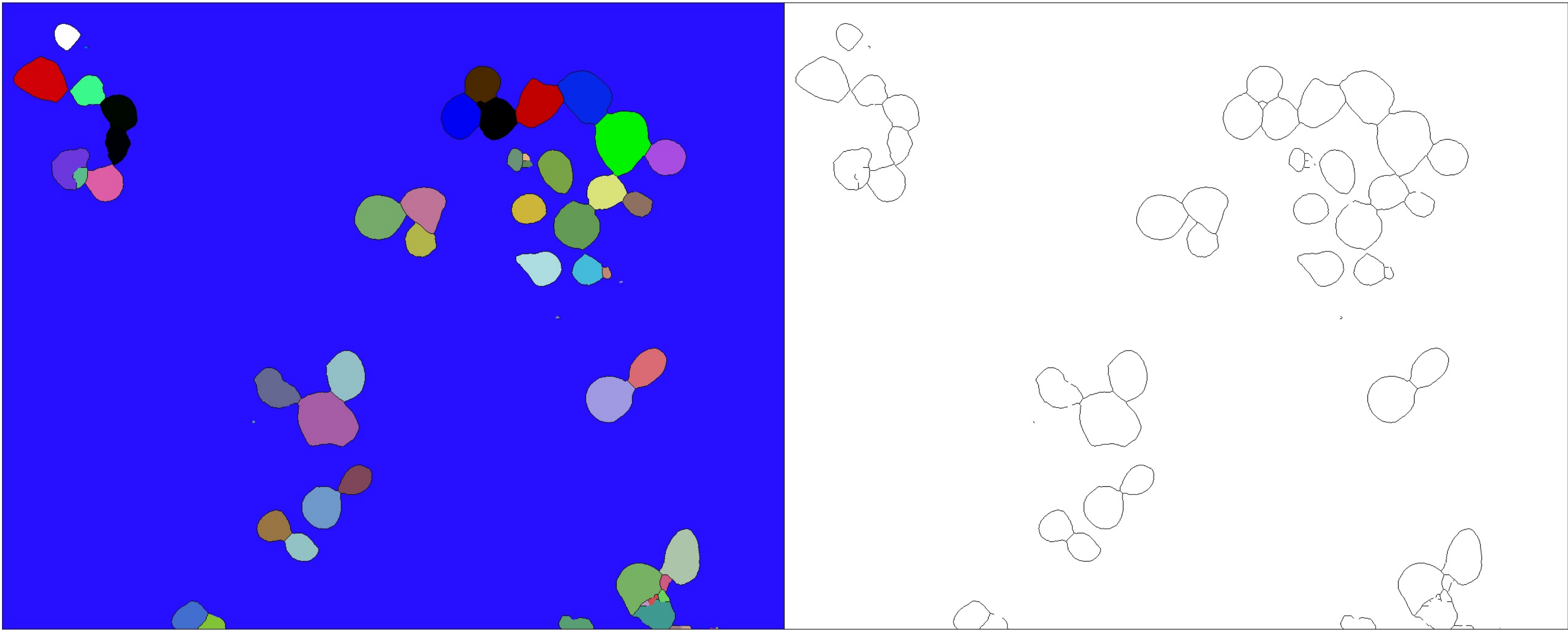


Contour Extraction



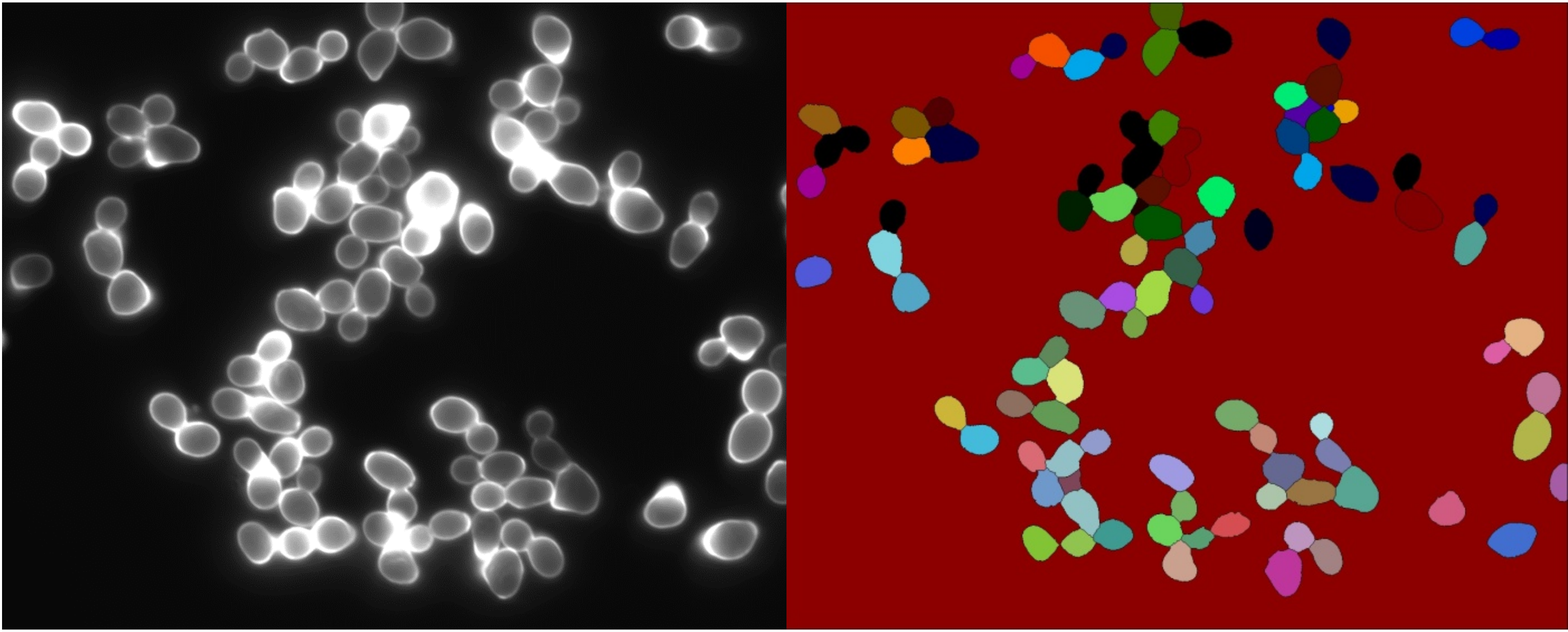
Final Marker Image

Marker-based Watershed

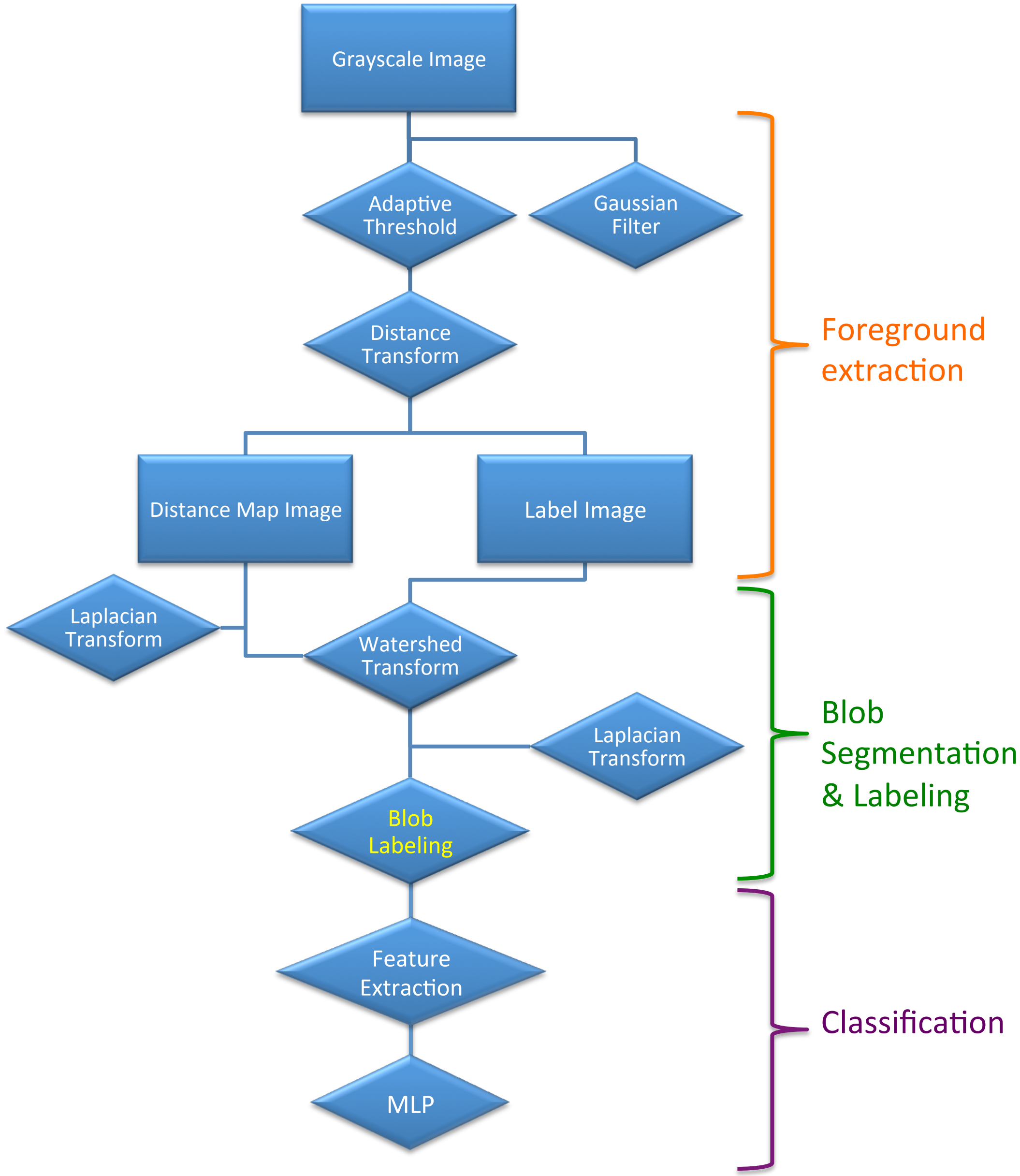


Watershed Segmentation

Marker-based Watershed



Watershed Segmentation



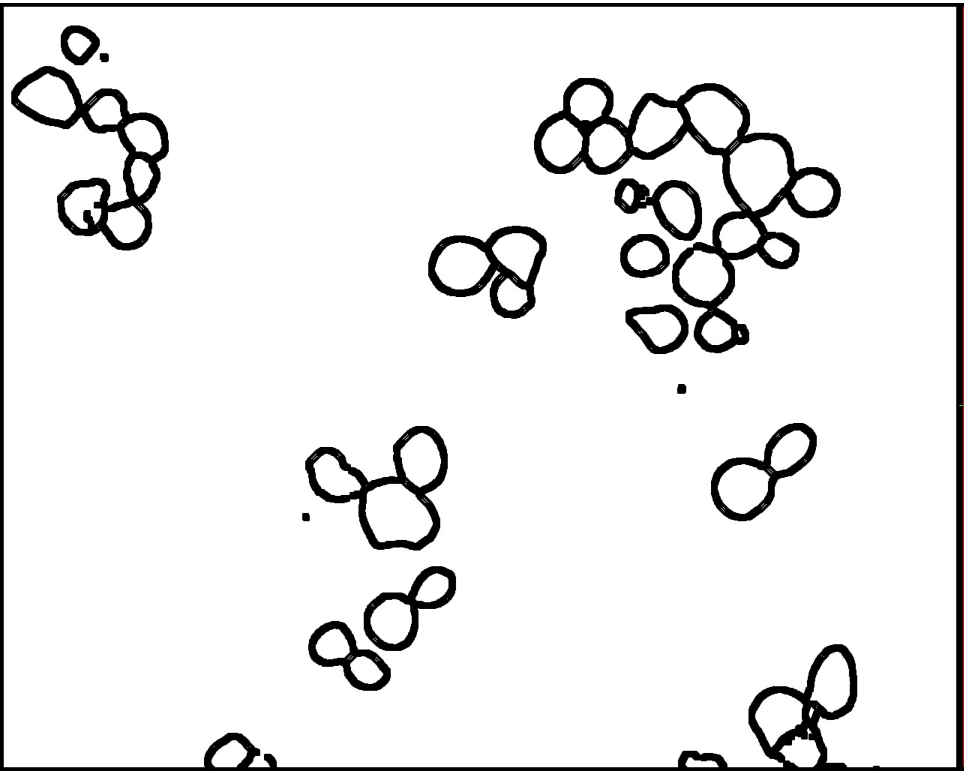
Grayscale Image
Adaptive Threshold
Gaussian Filter
Distance Transform
Distance Map Image
Label Image
Laplacian Transform
Watershed Transform
Laplacian Transform
Blob Labeling
Feature Extraction
MLP
Foreground extraction
Blob Segmentation & Labeling
Classification

Blob Labeling

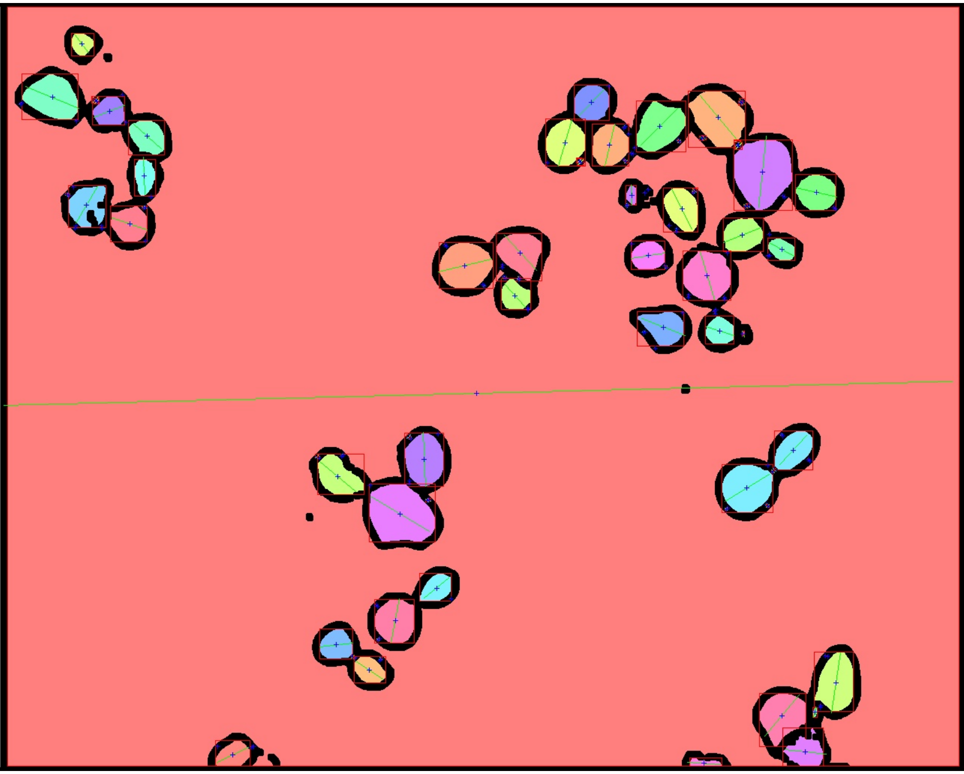
- Region-based component labeling using contour tracing of the segmented blobs.

  - Region-based methods cover more pixels than Edge-based methods.
    - More information available to characterize the blobs.
    - Noisy edges have less negative effects on the feature set.

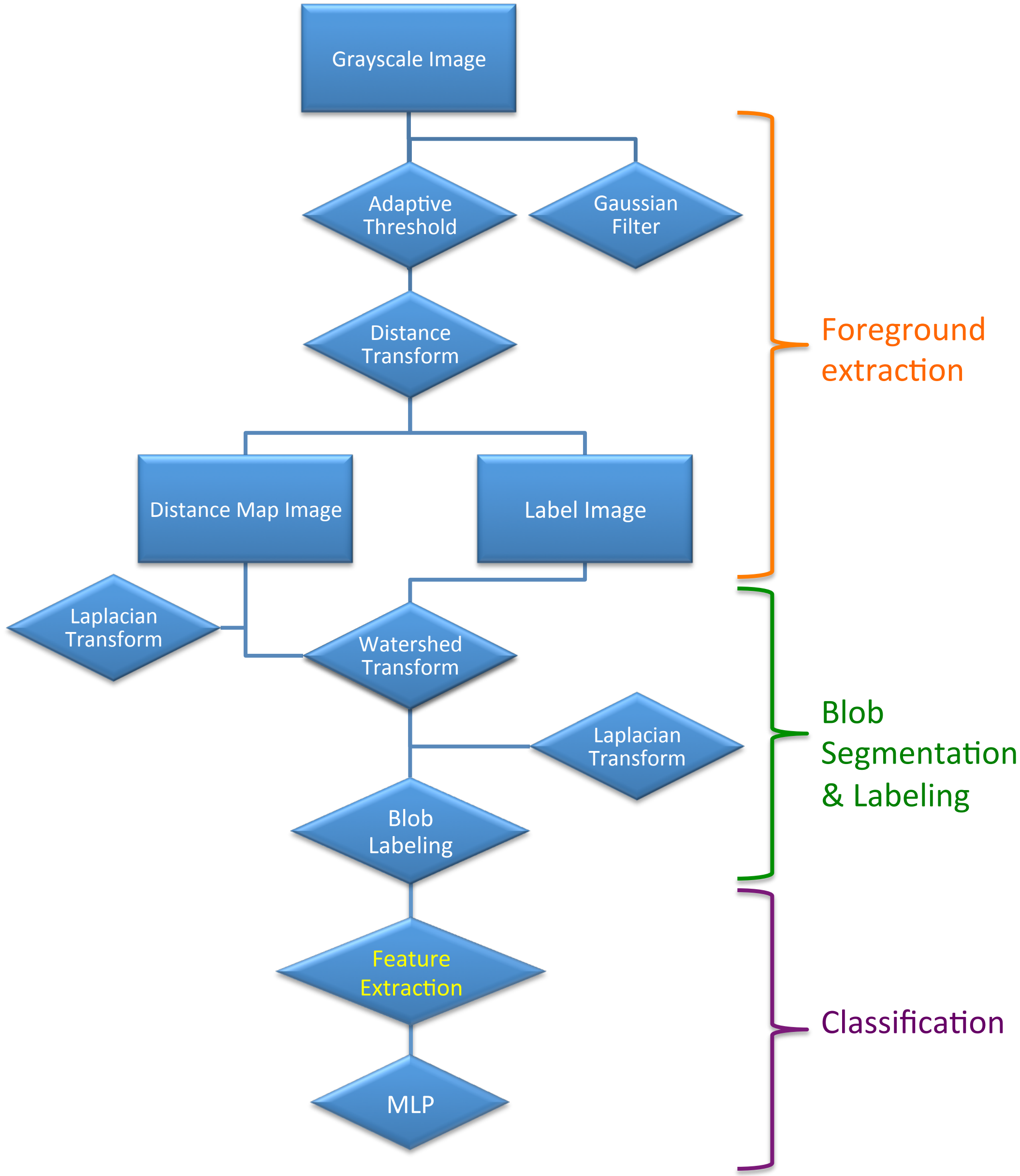
Blob Labeling



Watershed Segmentation



Blob Labeling


Grayscale Image
Adaptive Threshold
Gaussian Filter
Distance Transform
Distance Map Image
Label Image
Laplacian Transform
Watershed Transform
Laplacian Transform
Blob Labeling
Feature Extraction
MLP
Foreground extraction
Blob Segmentation & Labeling
Classification

Theory of Moments

- Geometrical moment of order $i+j$ for a two-dimensional discrete function $I(x,y)$ is computed by using the formula:

$$M_{ij} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} x^i y^j I(x,y)$$

Where M, N are Image Dimensions

- A simple property derived from moments is the area M_{00} .

Theory of Moments

- To be translation invariant the moments are centralized.

$$\mu_{ij} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \overline{x})^i (y - \overline{y})^j I(x, y)$$

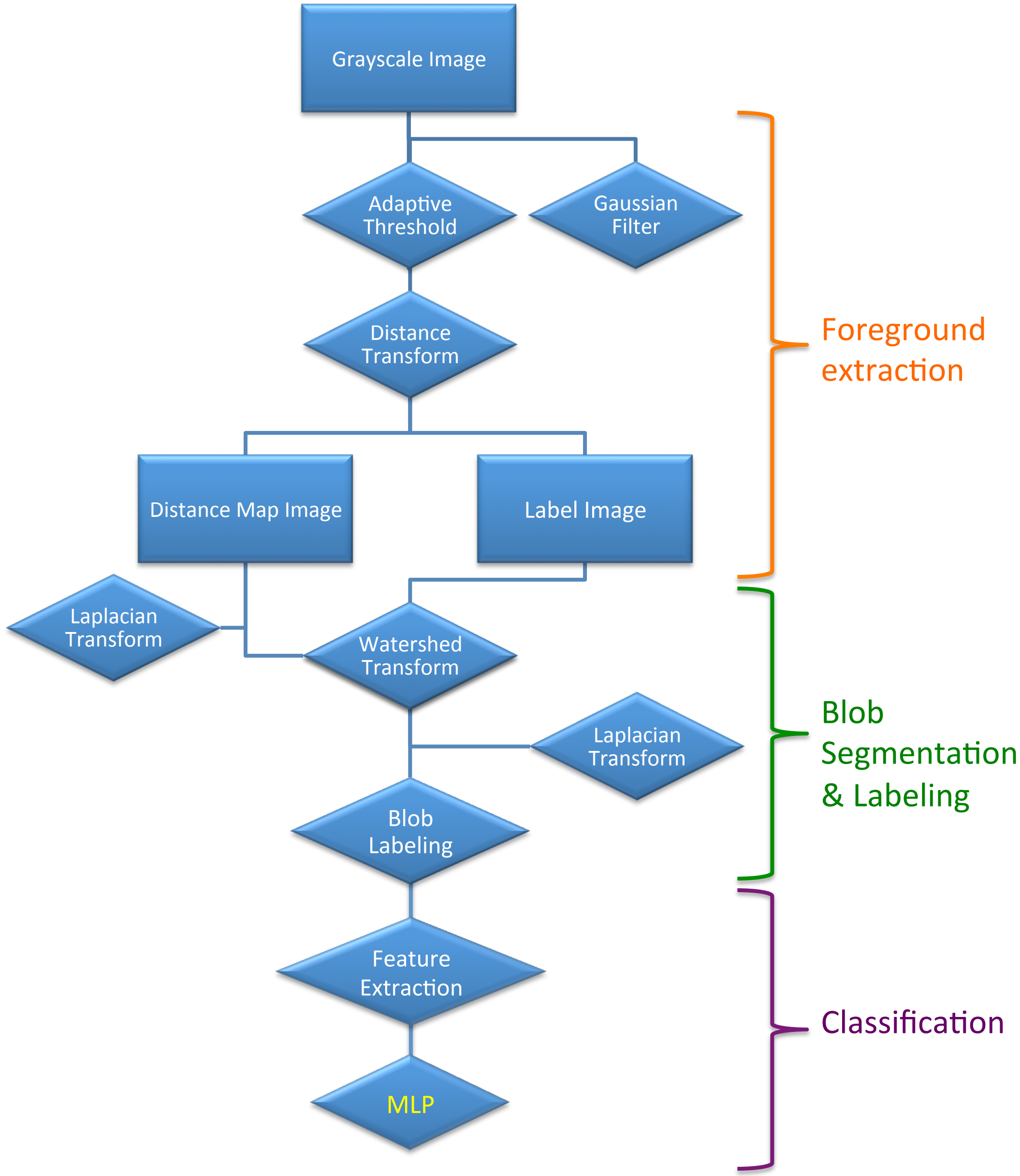
Where $(\overline{x}, \overline{y})$ *is the Blob Center of Mass*

- To be scale invariant the moments are normalized.

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{(1+\frac{i+j}{2})}}$$

Theory of Moments

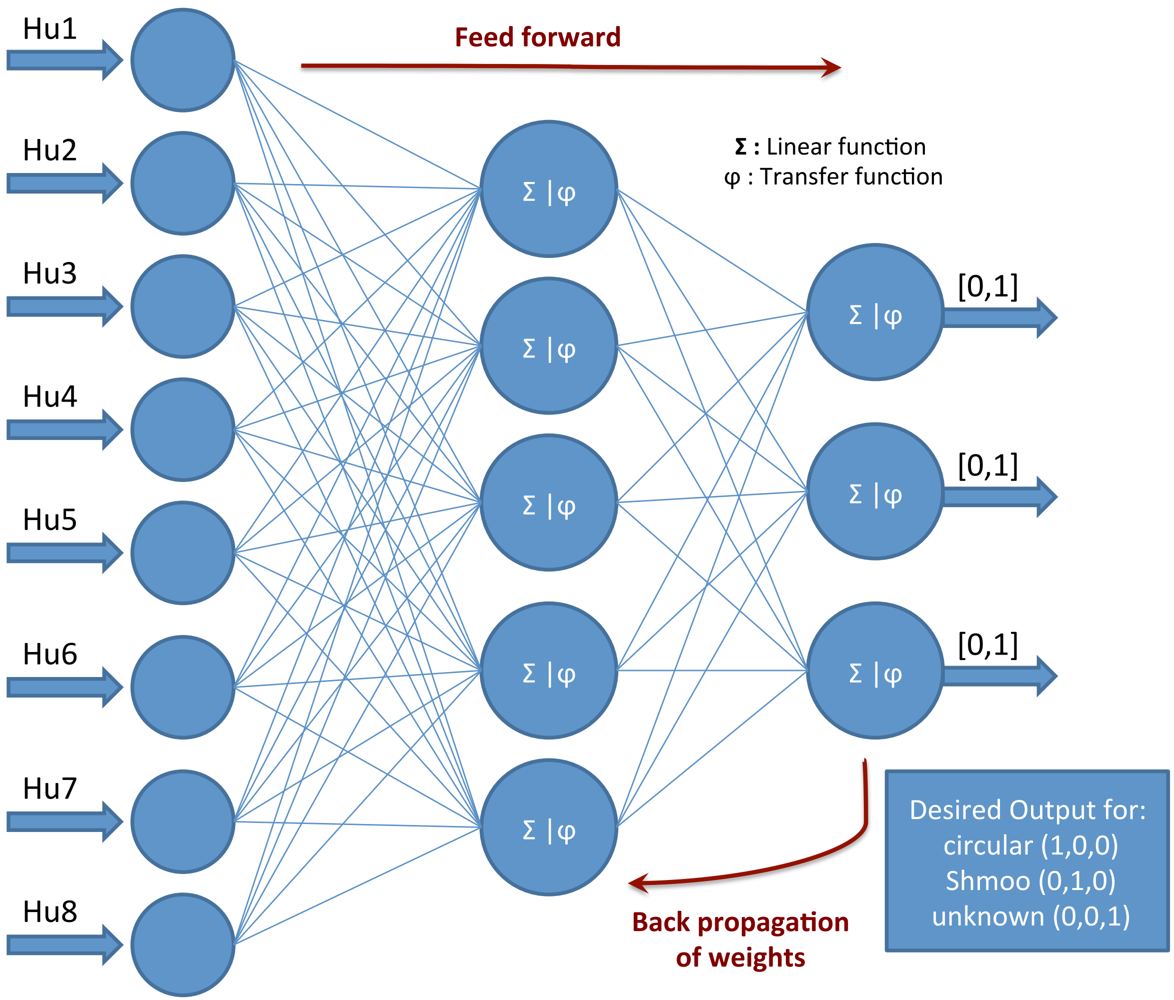
- Using normalized central moments up to the third order, *Hu* introduced eight moments invariant to translation, scale and orientation.
- The eight Hu invariant moments were used as our feature set.


Grayscale Image
Adaptive Threshold
Gaussian Filter
Distance Transform
Distance Map Image
Label Image
Laplacian Transform
Watershed Transform
Laplacian Transform
Blob Labeling
Feature Extraction
MLP
Foreground extraction
Blob Segmentation & Labeling
Classification

Multi-Layer Perceptron

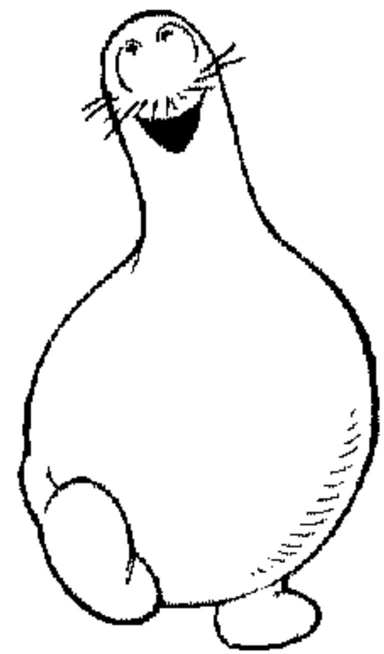
- Artificial Neural Network with interconnected simple computational elements called neurons.
  - Neurons are grouped into layers (input, hidden, output).
  - Useful in mapping an input vector to different classes by optimizing the weights associated with each neuron in the hidden and output layer.
- Optimization of the weights is done using the back propagation learning algorithm.
  - By comparing the actual output with the desired output.

Multi-Layer Perceptron



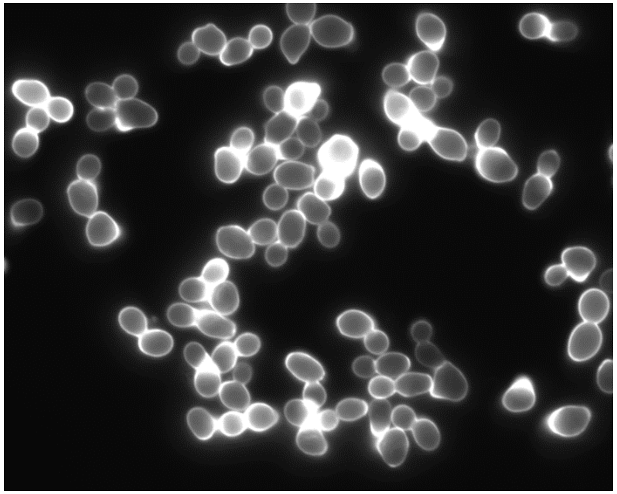
Data Set

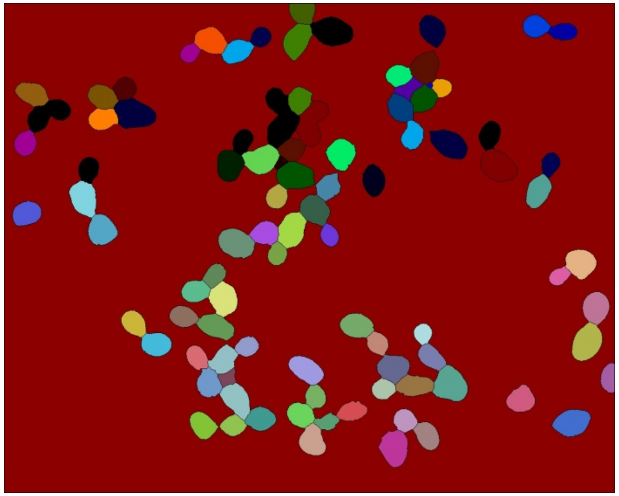
- Three types of cells were extracted from a set of 20 images to test our classifier:
  - Normal (circular): 26 cells.
  - Shmoo : 24 cells.
  - Noisy (unclassified) : 17 cells.
    - Total number : 67 cells.



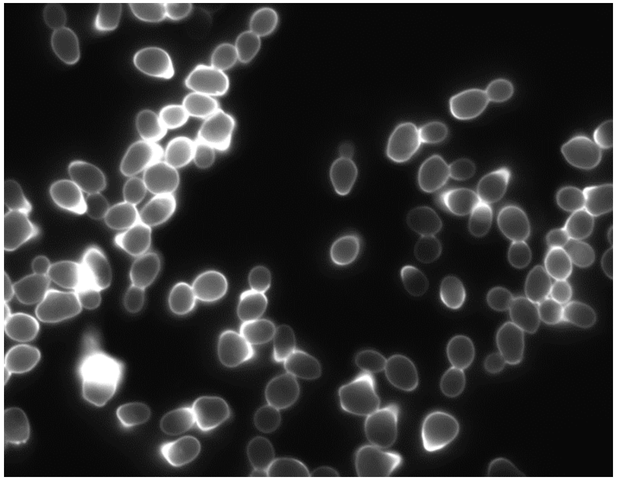
Shmoo Cell

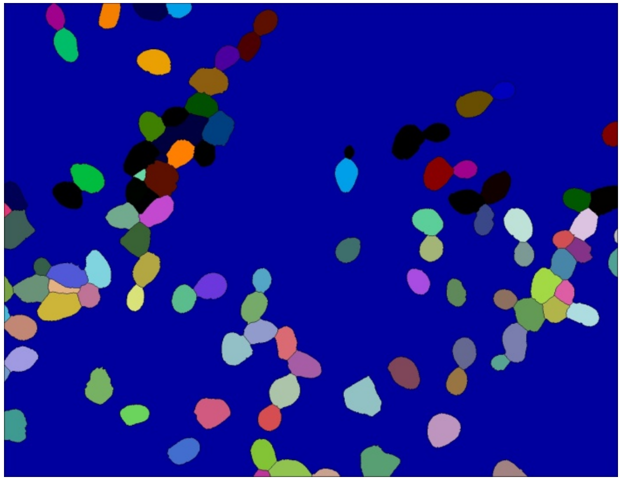
Segmentation Results





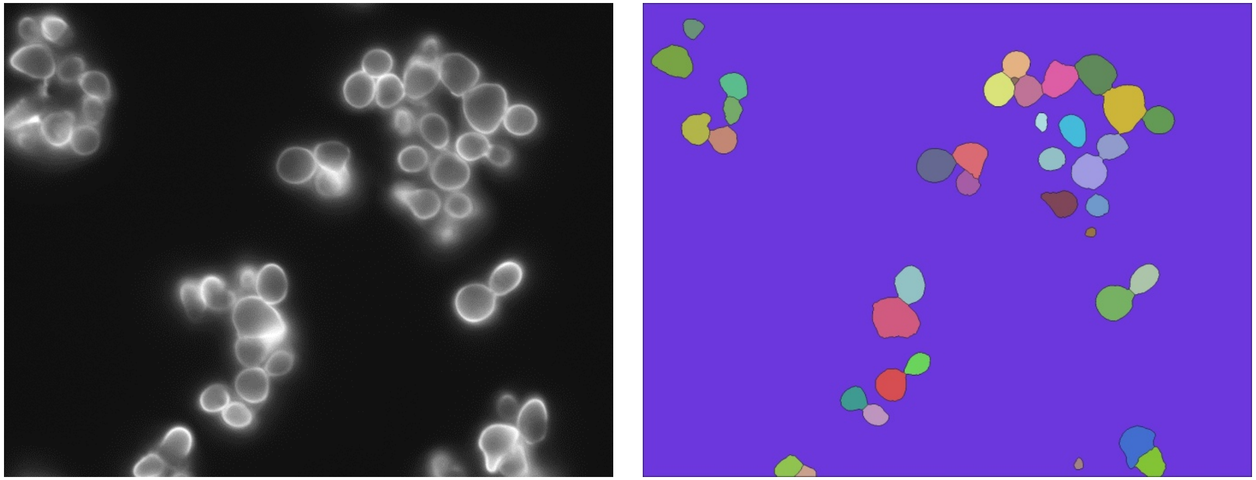
96.77%



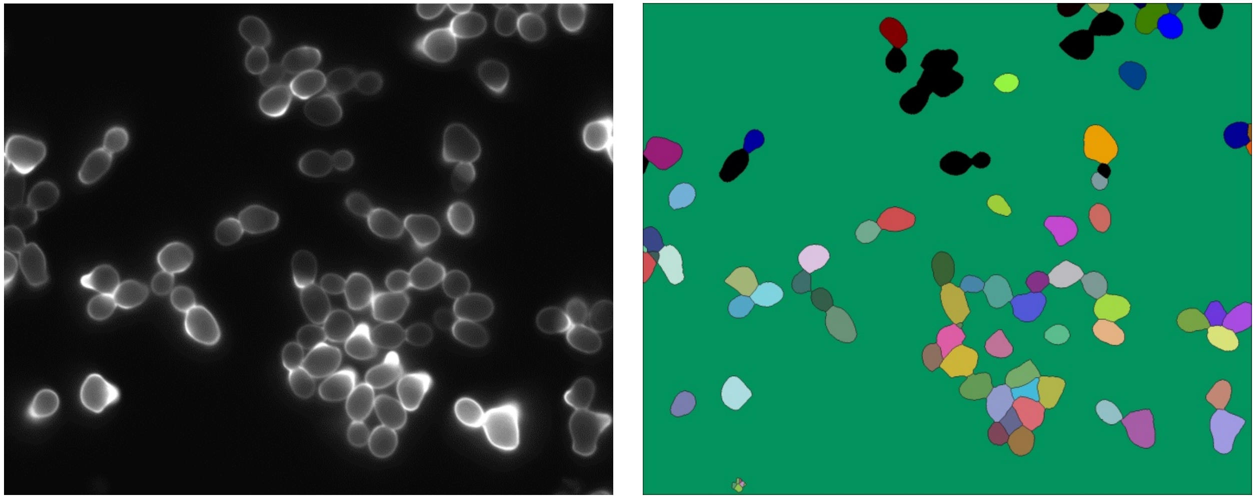


95.01%

Segmentation Results



90.24%

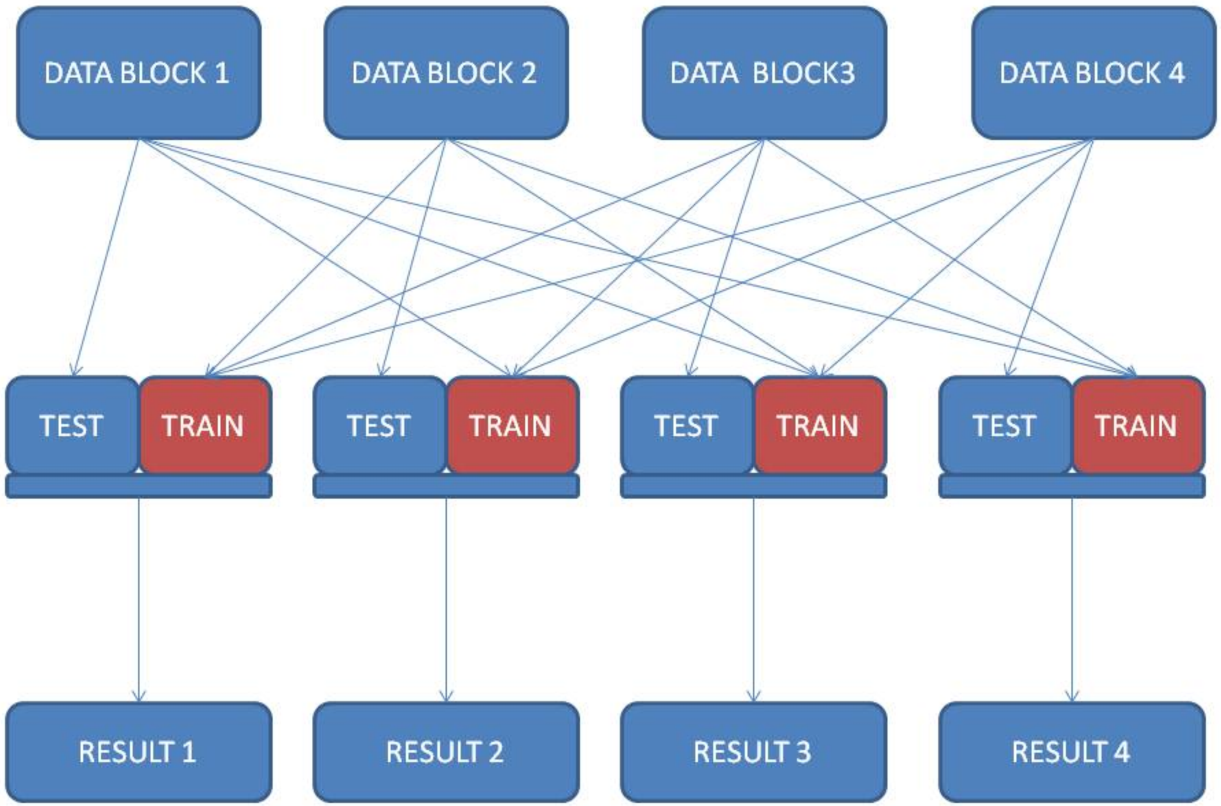


89.21%

Classification Results

- The cells extracted were evenly distributed to four data blocks.
  - Data Block 1 : 6 Normal, 6 Shmoo, 2 Unclassified.
  - Data Block 2 : 6 Normal, 6 Shmoo, 5 Unclassified.
  - Data Block 3 : 7 Normal, 6 Shmoo, 5 Unclassified.
  - Data Block 4 : 7 Normal, 6 Shmoo, 5 Unclassified.

Classification Results



Normal class	100%	100%	100%	86%
Shmoo class	100%	84%	100%	100%

Future Works

- Get more data to further test the method.
- Make the algorithm able to process different kind of images.

Questions?