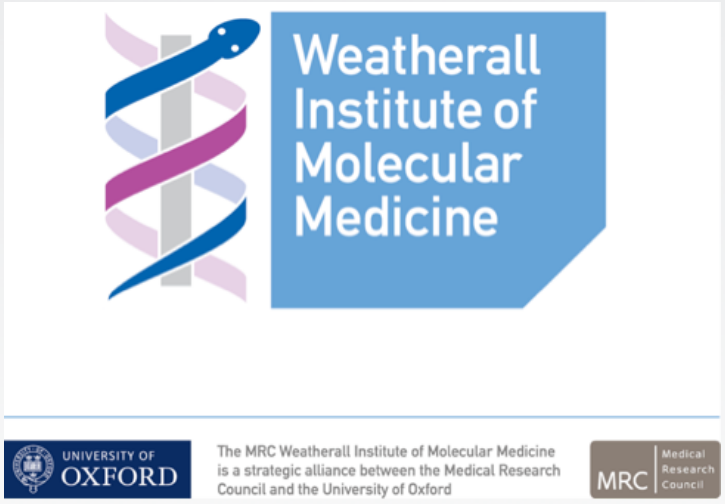
Nanoscopy Oxford

# Pattern recognition, machine learning and feature detection

by Dominic Waithe

- Background
- Machine learning in microscopy
- Supervised machine learning
- The Problem: Organoids
- Solution: Random Forest + features
- Features
- Random Forest
- Conclusions

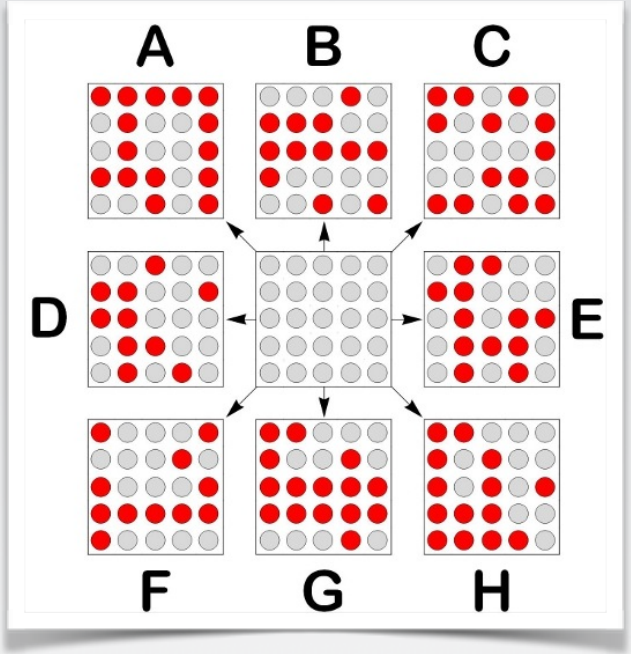


# Machine learning recap

**Machine learning** is a scientific discipline that deals with the construction and study of algorithms that can learn from data without being explicitly programmed.

Source: http://en.wikipedia.org/wiki/Machine_learning

# Pattern Recognition




**Pattern recognition** is a branch of machine learning that focuses on the recognition of patterns and regularities in data.

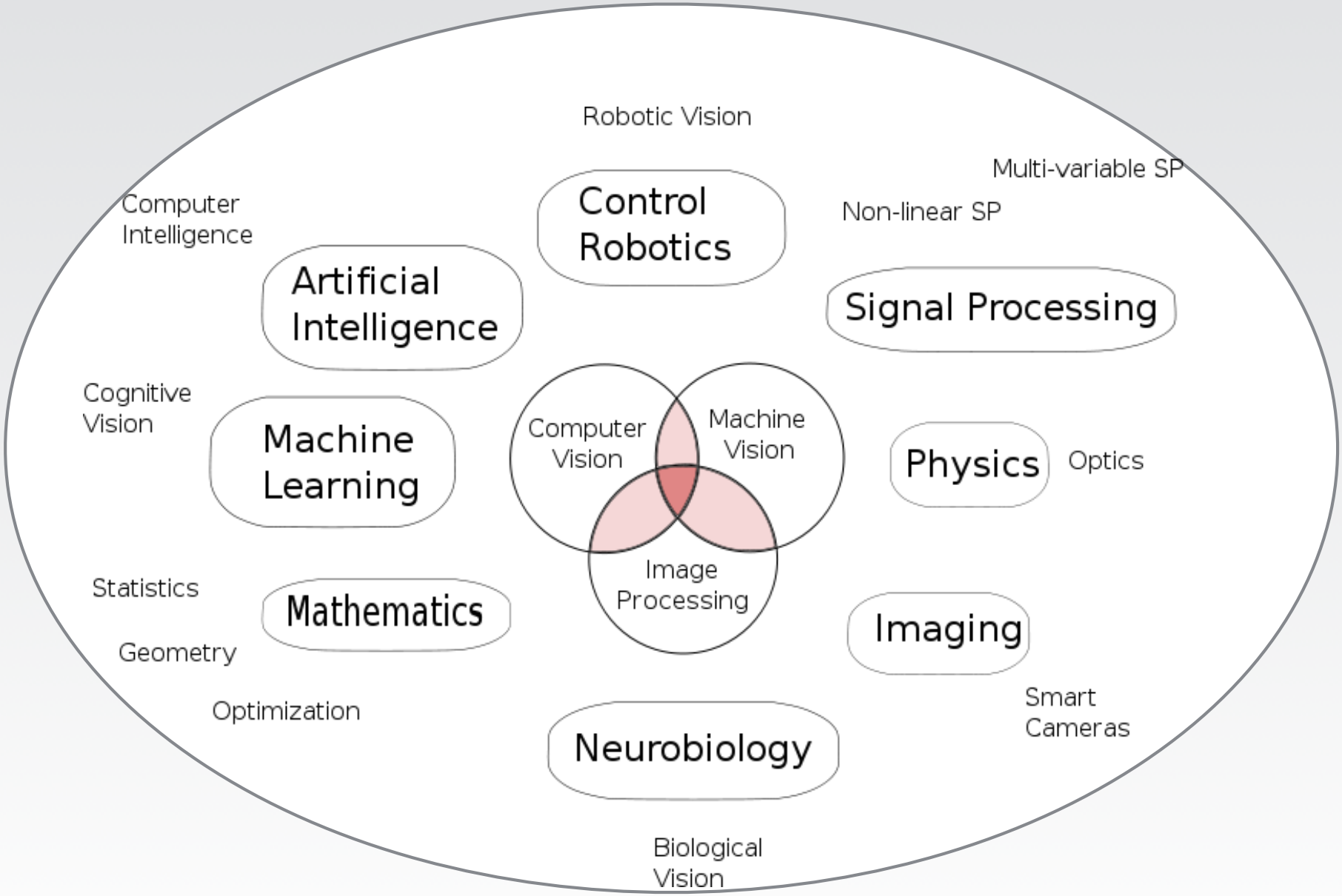
Source: http://en.wikipedia.org/wiki/Pattern_recognition

# Image Processing

**Image processing** is any form of signal processing for which the input is an image, such as a photograph or video frame. Includes feature detection.

Source:

# Many overlaps.




Source: http://en.wikipedia.org/wiki/Computer_vision

# Microscopy and computer science

Machine learning and pattern recognition is becoming more and more important in the imaging sciences.

Students from all backgrounds will have some interaction with these algorithms in the future.


Arteta, C., V. Lempitsky, J. A. Noble and A. Zisserman (2012). Learning to detect cells using non-overlapping extremal regions. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012, Springer: 348-356.

Eliceiri, K. W., M. R. Berthold, I. G. Goldberg, L. Ibáñez, B. Manjunath, M. E. Martone, R. F. Murphy, H. Peng, A. L. Plant and B. Roysam (2012). "Biological imaging software tools." Nature methods **9**(7): 697-710.

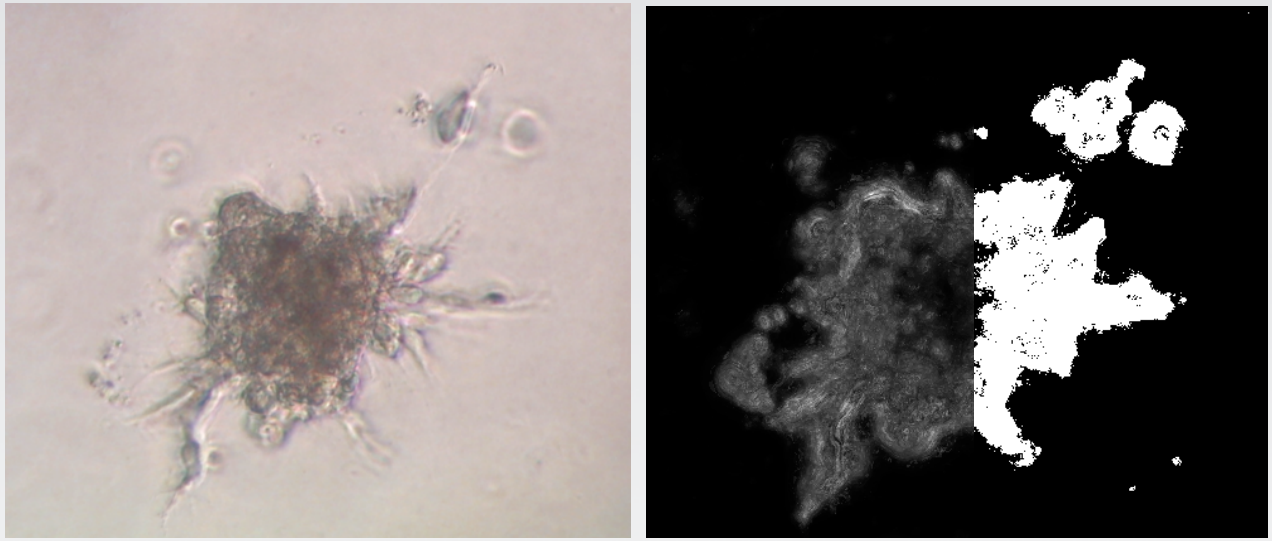
Huh, S., D. F. E. Ker, R. Bise, M. Chen and T. Kanade (2011). "Automated mitosis detection of stem cell populations in phase-contrast microscopy images." Medical Imaging, IEEE Transactions on **30**(3): 586-596.

Kanade, T., Z. Yin, R. Bise, S. Huh, S. Eom, M. F. Sandbothe and M. Chen (2011). Cell image analysis: Algorithms, system and applications. Applications of Computer Vision (WACV), 2011 IEEE Workshop on, IEEE.

Lee, H., R. Grosse, R. Ranganath and A. Y. Ng (2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. Proceedings of the 26th Annual International Conference on Machine Learning, ACM.


Source:

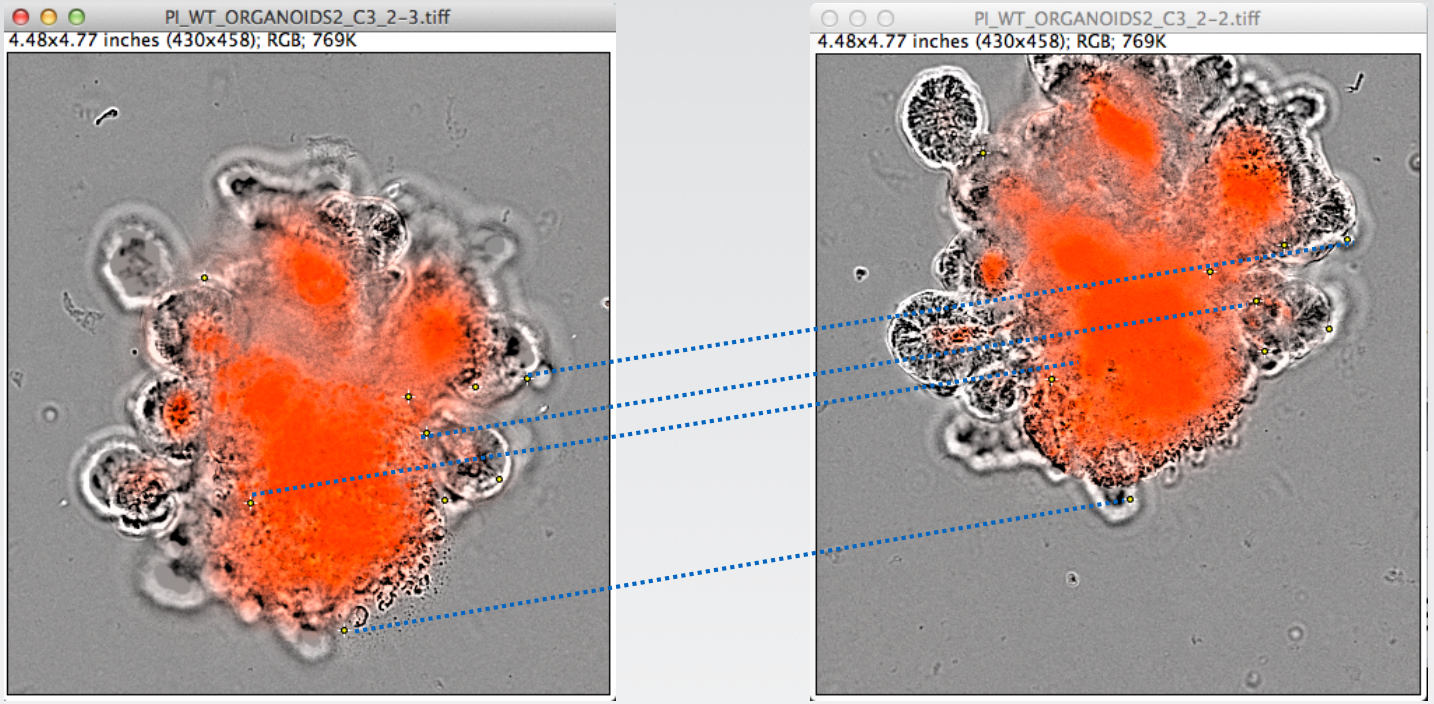
# Label-free segmentation



Pattern recognition techniques allow for segmentation of complex distributions in label-free imaging modalities (e.g. photography, DIC, phase-contrast).

Source: David Favara

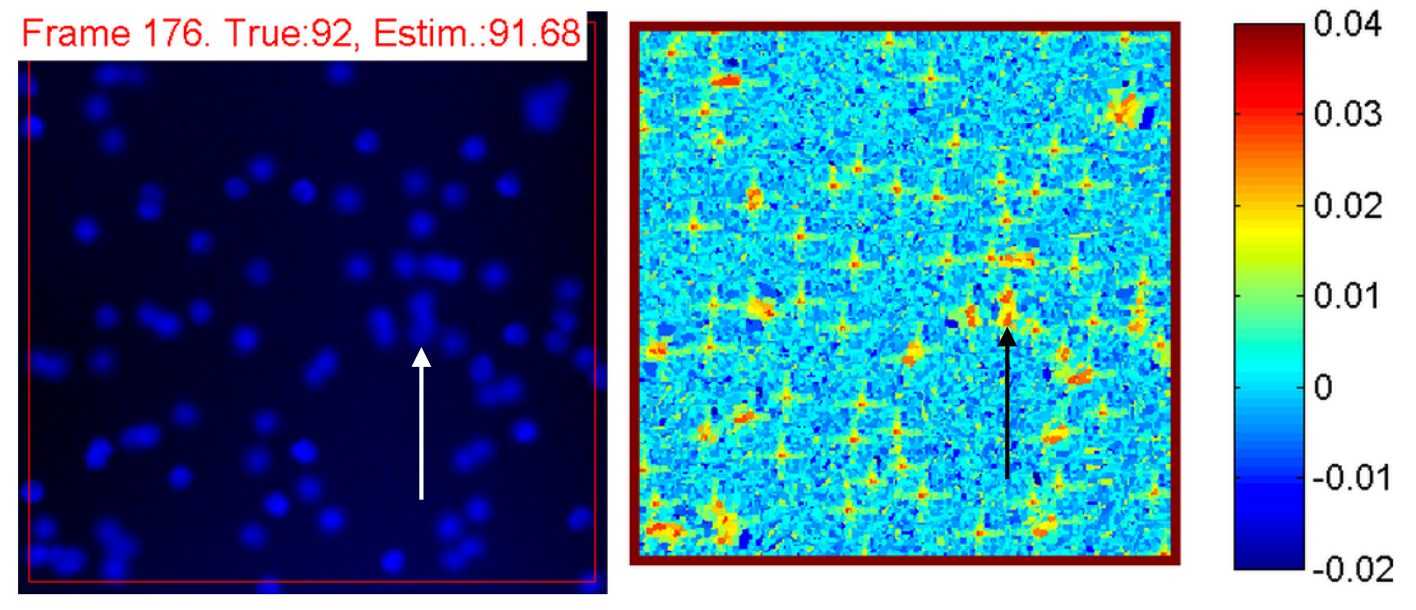
# Examples: Registration, feature detection.




MOPS and SIFT + other feature detectiors are able to recognise the same specific points in different images. (This is computer vision/IP more than ML).

Source: http://www.cs.bath.ac.uk/brown/mops/mops.html http://www.cs.bath.ac.uk/brown/mops/mops.html Daniele Muraro

# Counting objects without segmentation




Sometimes objects are occluded or vary subtly in appearance between instances.

Source: Learning to Count Objects in Images,Victor Lempitsky,Andrew Zisserman.

- Background
- Machine learning in microscopy
- Supervised machine learning
- The Problem: Organoids
- Solution: Random Forest + features
- Features
- Random Forest
- Conclusions

Source:

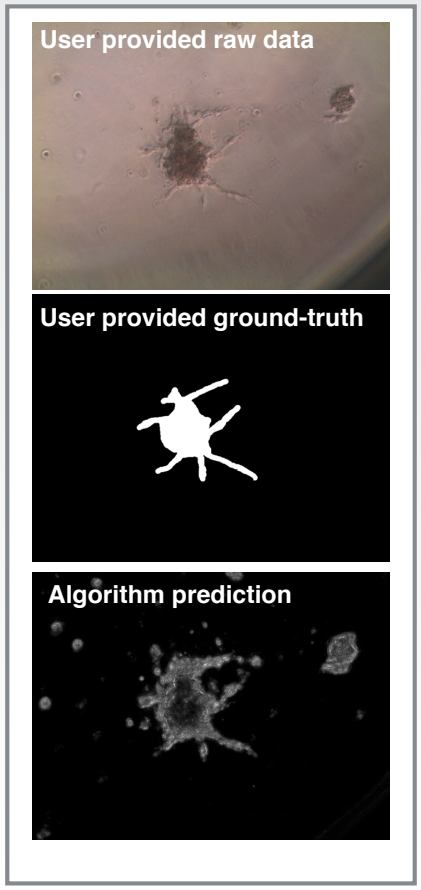
# Today's lecture:

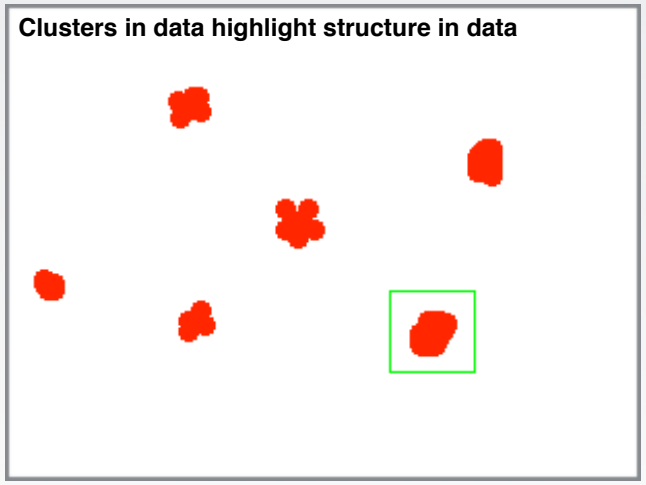
Machine Learning

Supervised

Unsupervised

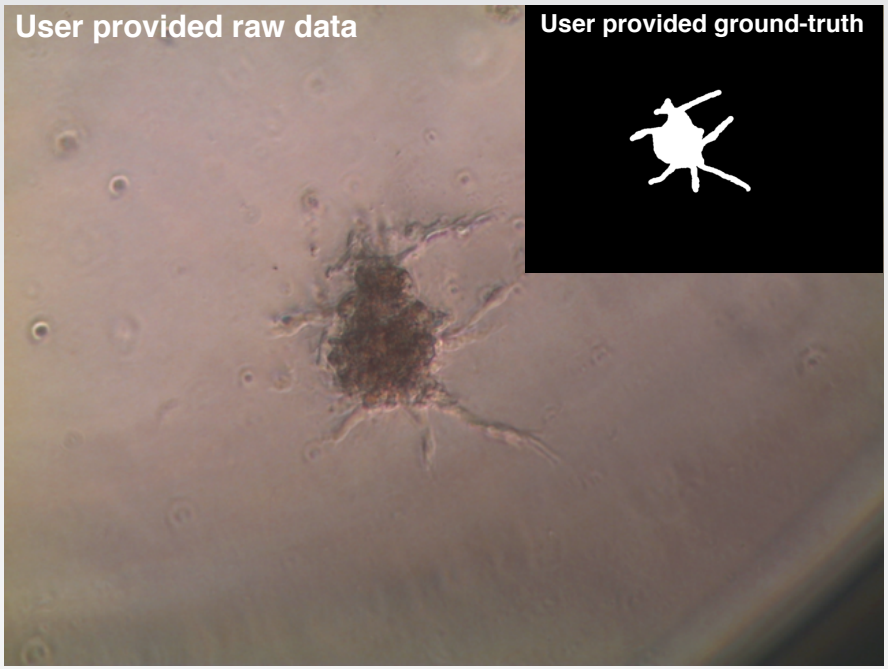
reinforcement learning

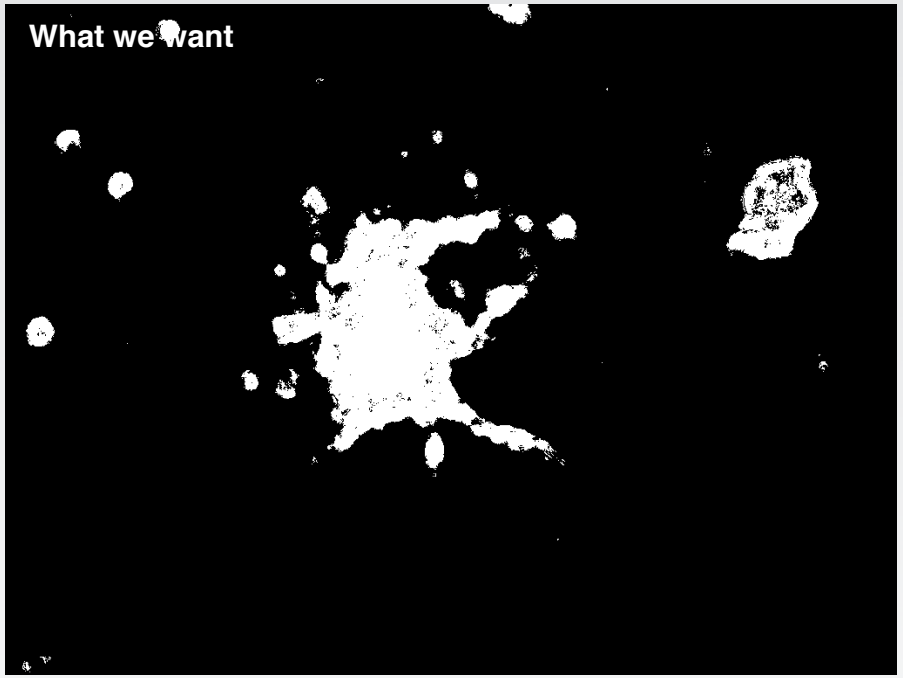


# Supervised machine learning example

Organoids are a small 3D sections of organs which can be grown artificially and which are becoming a very important aspect of regenerative medicine.




Organoids are dense 3D structures that are difficult to stain and so label-free techniques are important to help segment and classify these structures.

$g : \mathcal{X} \to \mathcal{Y}$

**We want to find the function which maps from X to Y.**

$x \in \mathcal{X}$ **X is our input data.**

$y \in \mathcal{Y}$ **Y is our output labels.**

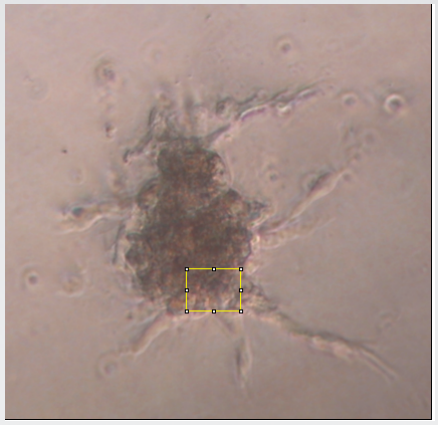
$\mathbf{D} = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ **training data**

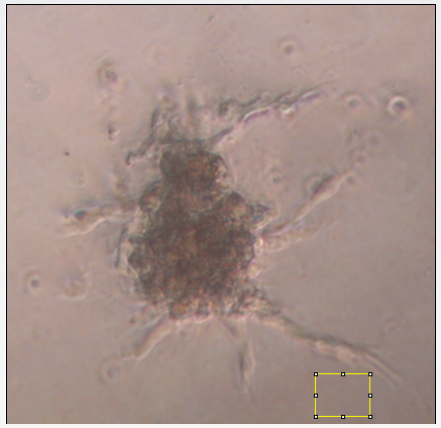
$h : \mathcal{X} \to \mathcal{Y}$

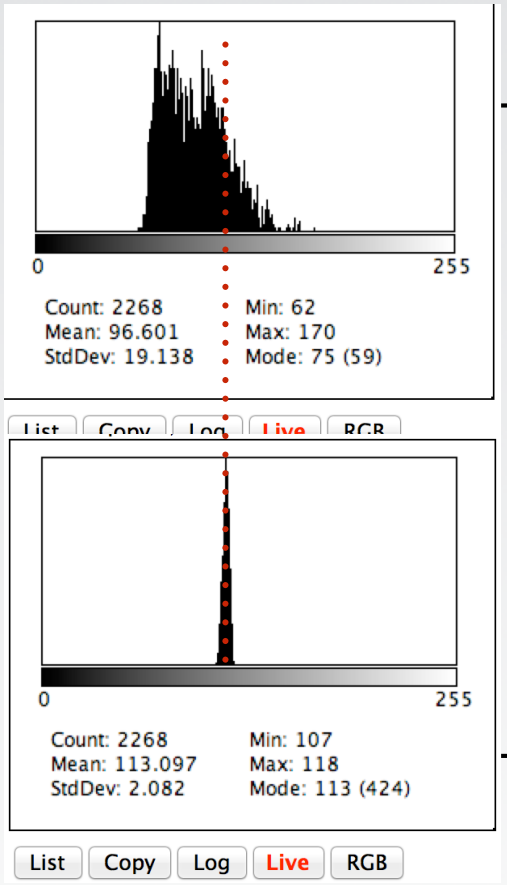
**h is our approximated version of g**

Source: Kerry Grens (December 24, 2013). "2013's Big Advances in Science". The Scientist. Retrieved 26 December 2013.

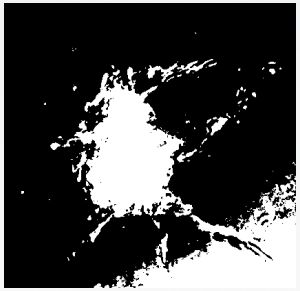
# Can we use Global thresholding of the histogram?







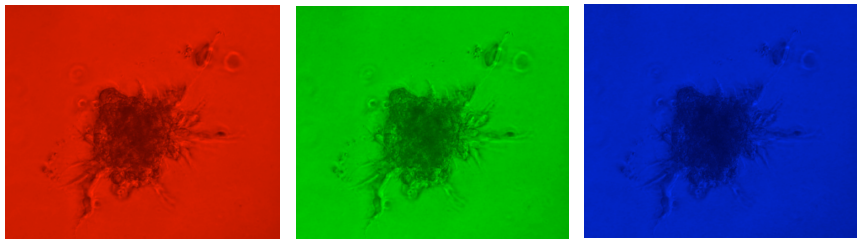

**Limited to one channel pixels are too similarly distributed making thresholding not so good**



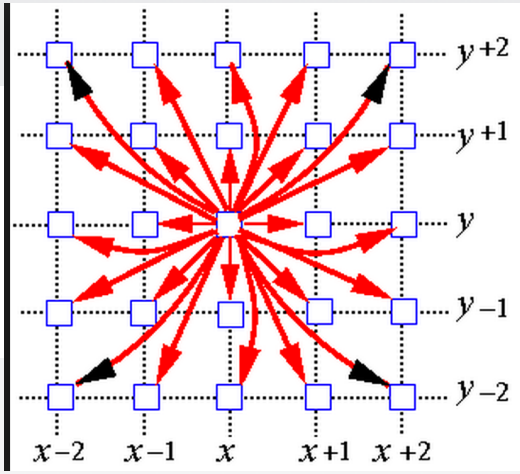
Source:

# Many image-processing and machine learning solutions for this problem

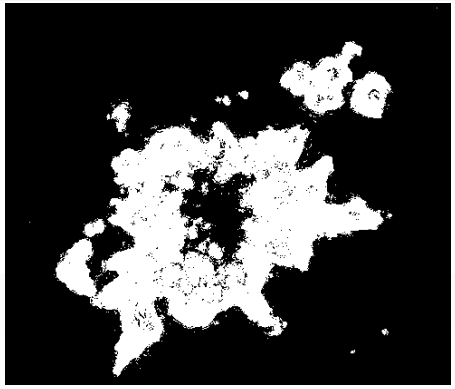
Include multi-channel information (RGB).



Include spatial information to inform the classifier.




Organise information and provide output.



Source: https://www.cs.auckland.ac.nz/courses/compsci708s1c/lectures/Glect-html/topic4c708FSC.htm

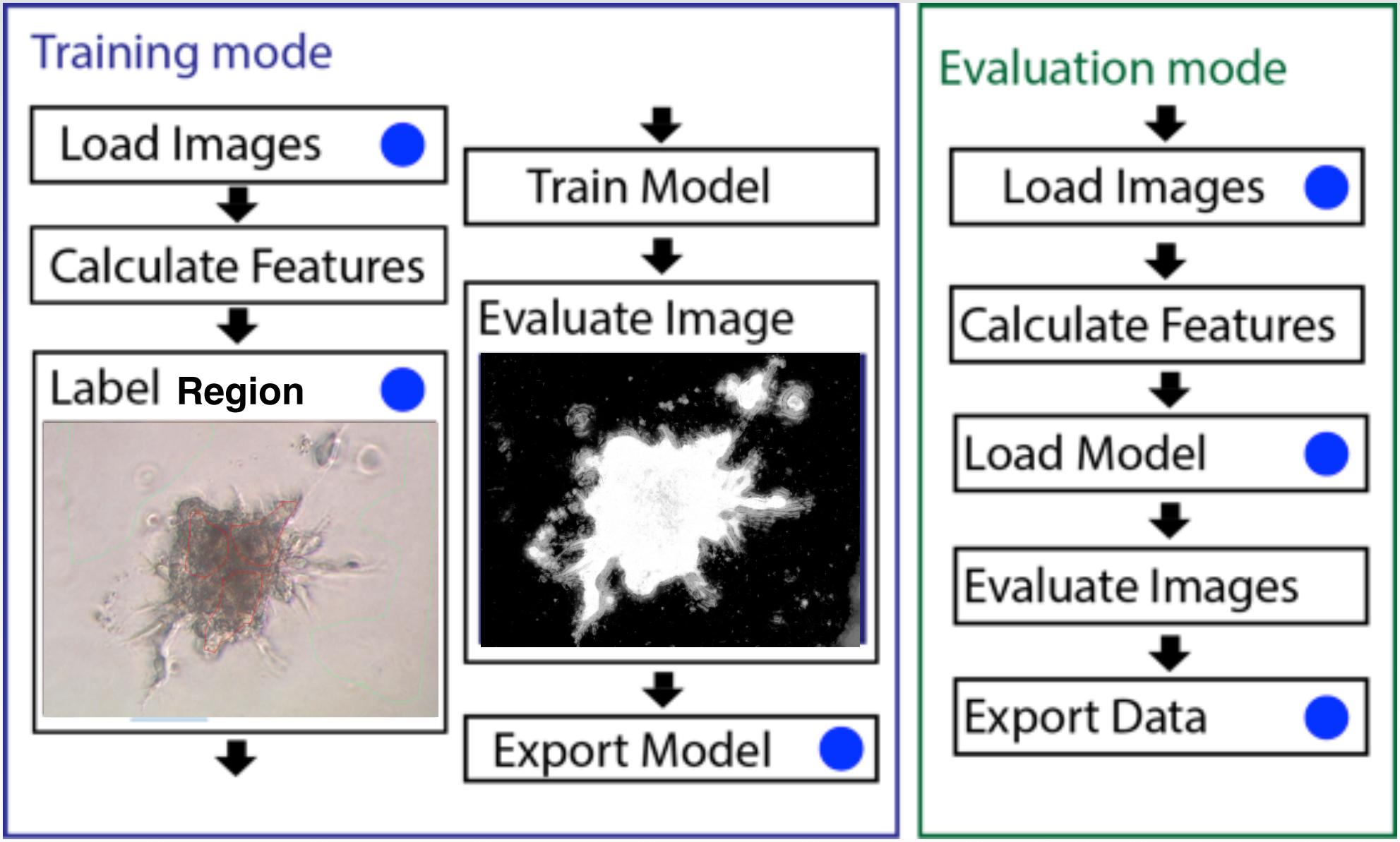
# Options available - many!

**Feature calculation:**
Covariant detectors
Histogram of Oriented Gradients
Scale Invariant Feature Transform (SIFT)
Local Intensity Order Pattern (LIOP)
Maximally Stable Extremal Regions (MSER)
Image distance transform

**Statistical methods (learning frameworks)**
Gaussian Mixture Models
k-means
Agglomerative Information Bottleneck (AIB)
Quick shift
SLIC
Support Vector Machine (SVM)
Forest of kd-trees

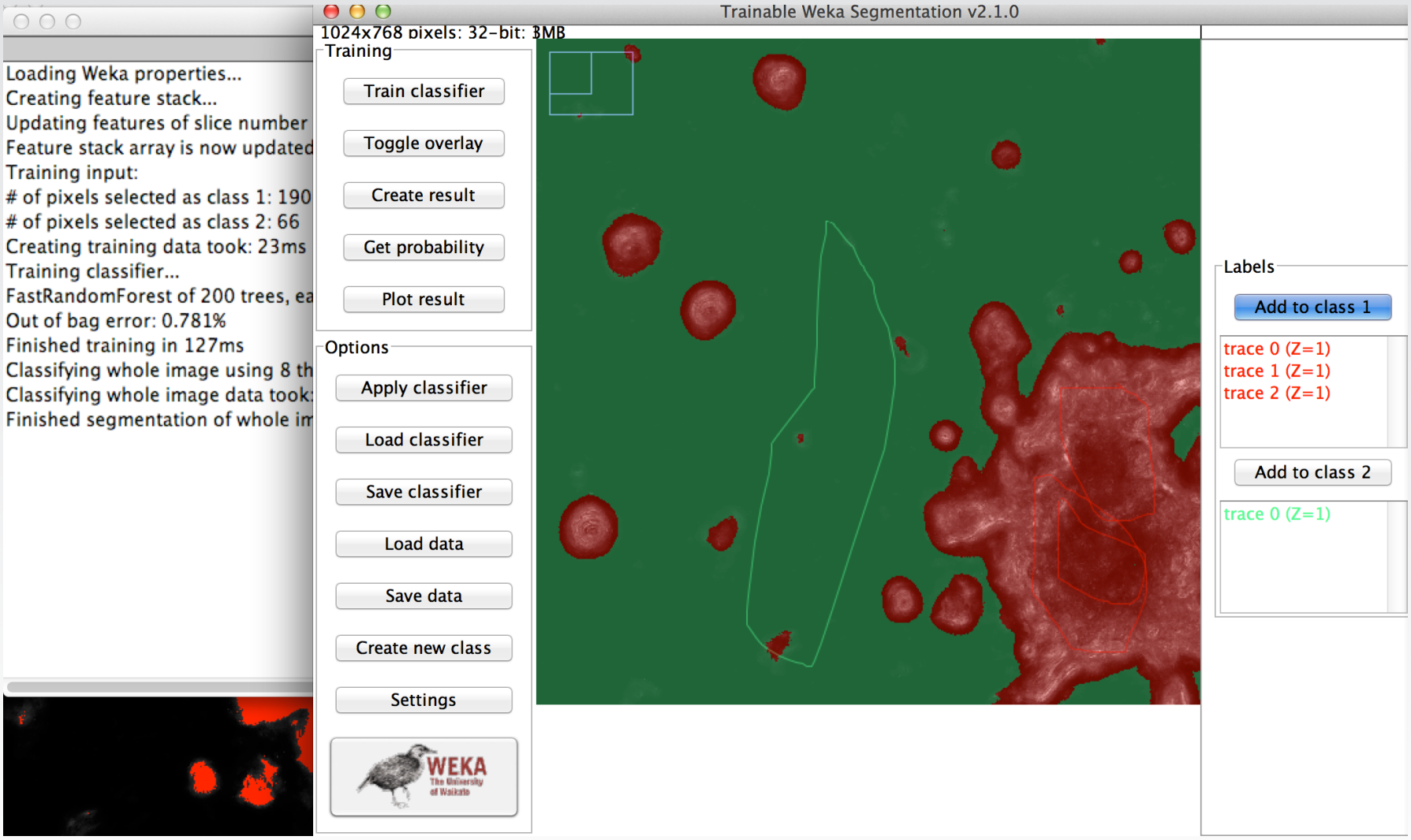
Source: http://www.vlfeat.org/overview/tut.html

# Typical work-flow for supervised learning




Source:

# Today we focus on the Trainable Weka Segmentation as an exempla




Source: Fiji WEKA

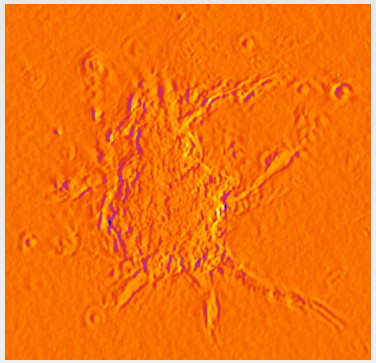
- Background
- Machine learning in microscopy
- Supervised machine learning
- The Problem: Organoids
- Solution: Random Forest + features
- Features
- Random Forest
- Conclusions

Source:

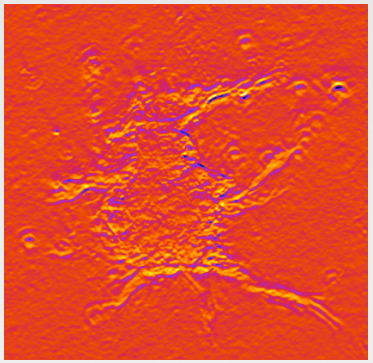
# Feature calculation

The WEKA plugin uses a popular means of feature calculation ( there are many others) which is to filter the input image with a bank of geometric features). Each filter exposes details in the image which can be used in the classification

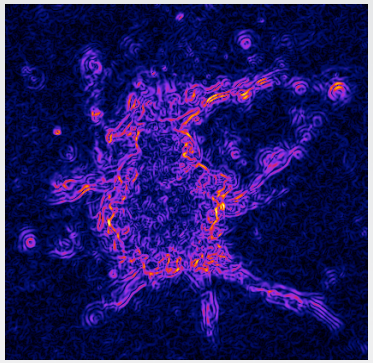
derivative-X



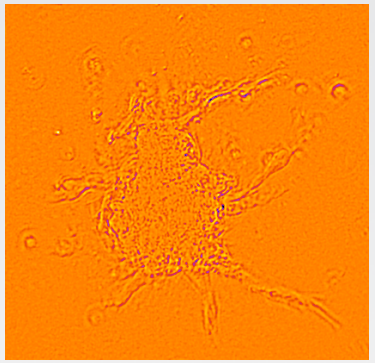
derivative-Y



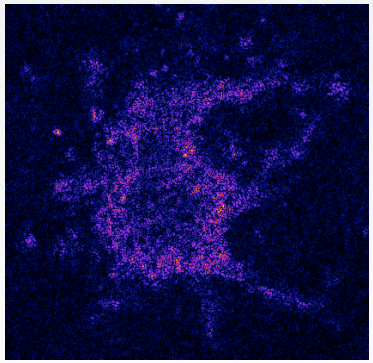
Edge mag.



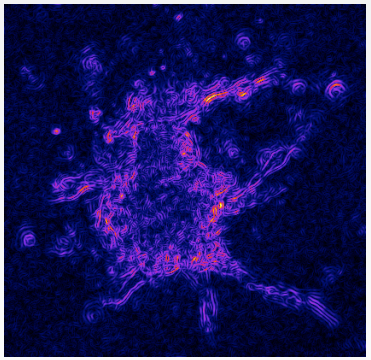
Laplacian



Hessian min



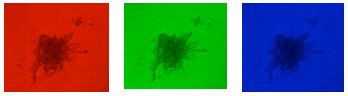
Hessian max



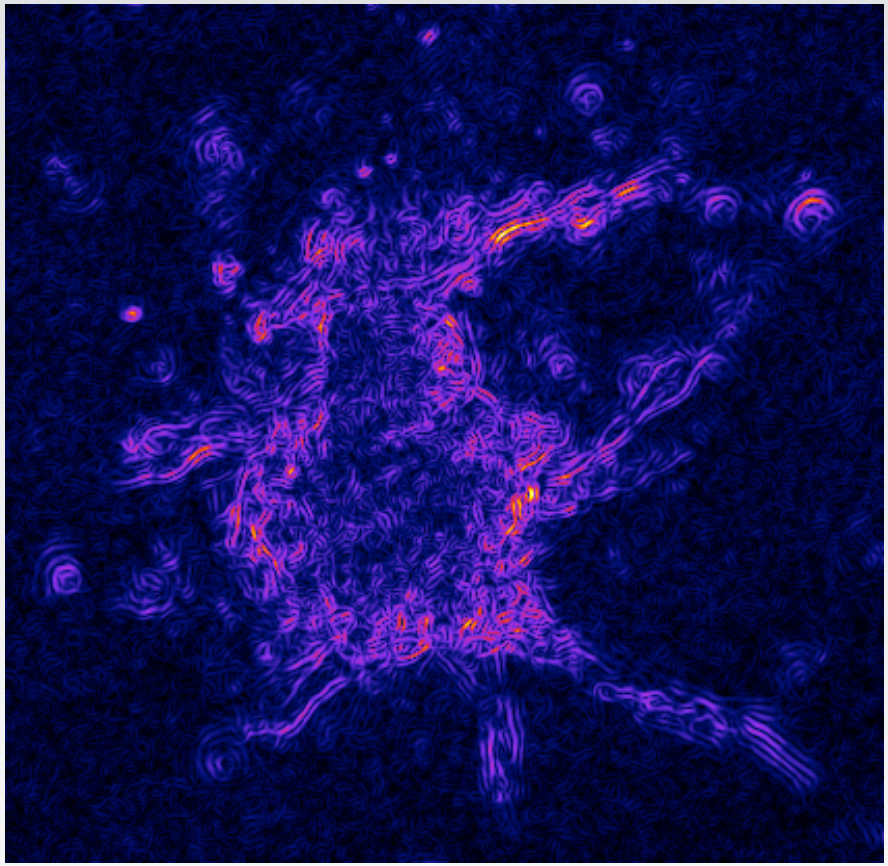
Source:FeatureJ

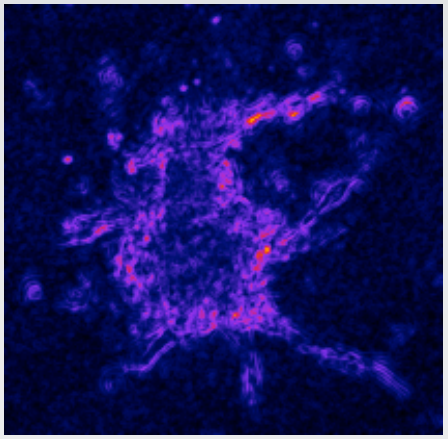
Each pixel in each colour has a value for each one of these image interpretations

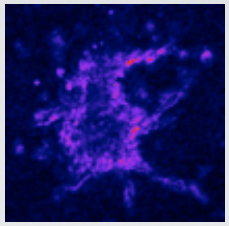
e.g. 3 x 6



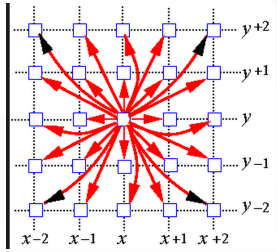
# Feature Calculation different scales







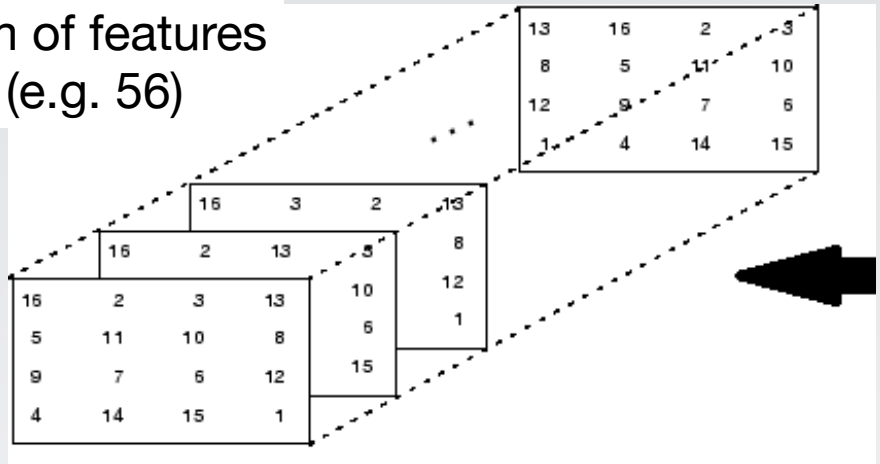
Not only are different Geometric filters used, but at different scales. By blurring the image and applying the different filters you get a sense of the greater neighbourhood. More than 3 different scales can be used 3 x 6 x 3




Source:

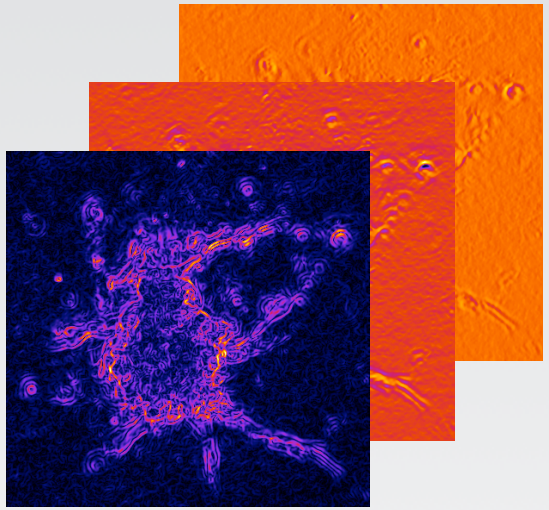
# Feature calculation- matrix

num of features
(e.g. 56)



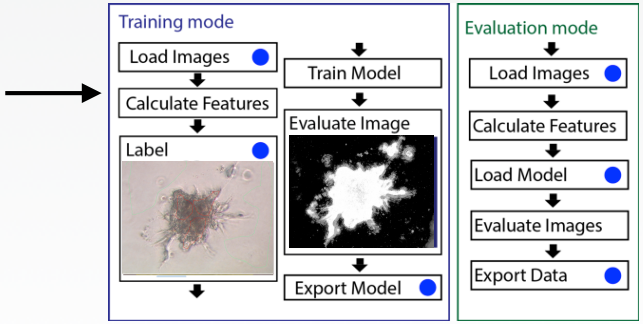

$im_{height}$

$im_{width}$



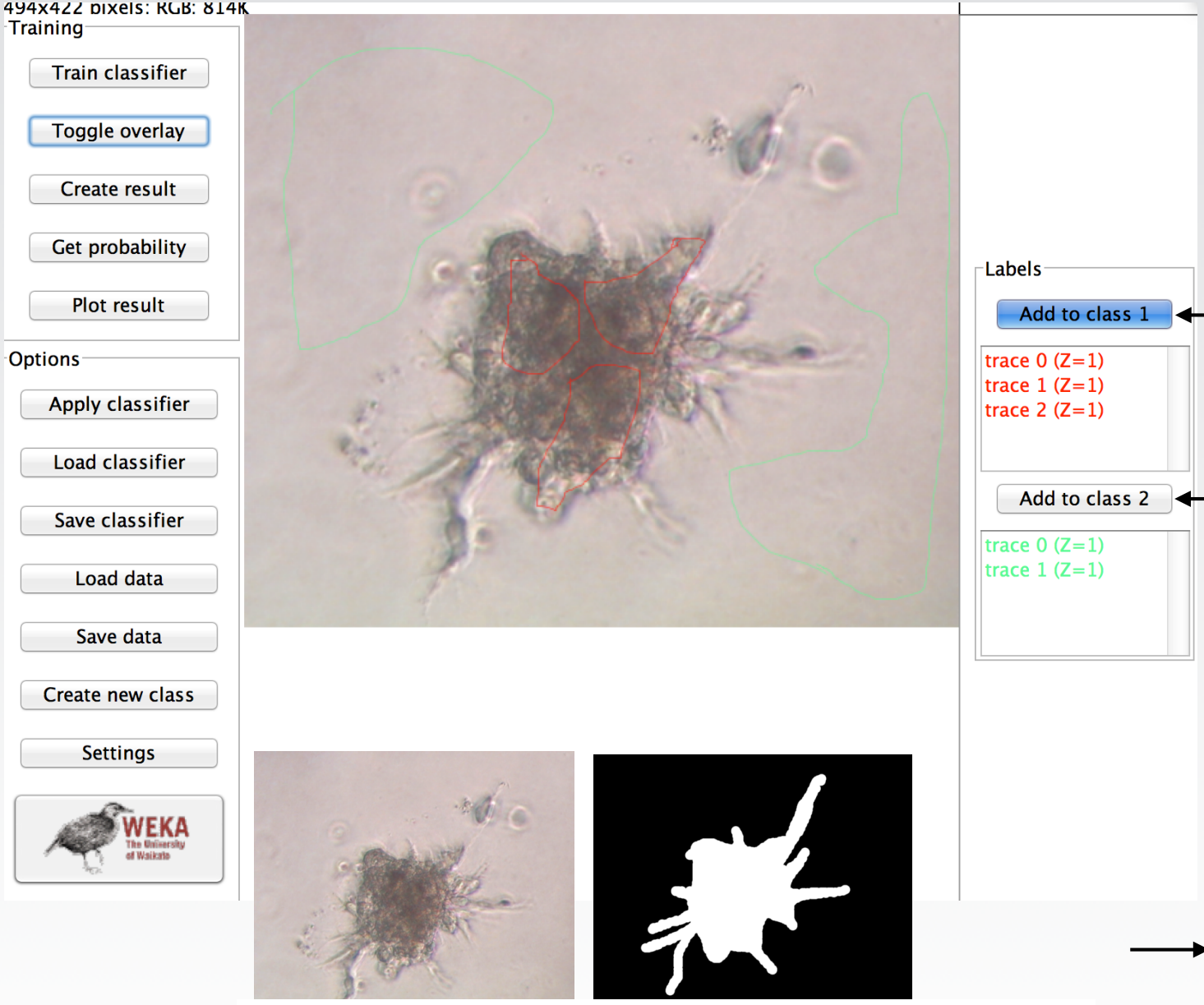
The result is a matrix which has 3-dimensions. and is equivalent in size to $im_{height}$ x $im_{width}$ x 56 (no. of features). For each image.

For each pixel, there is a feature vector.




Source:

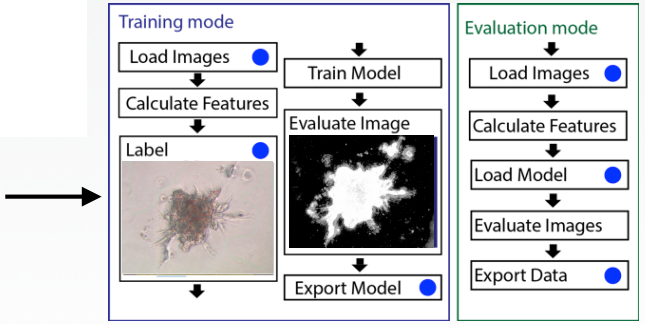
# Labelling- generation of training data




foreground (cell)

background






Source:

Matched training data $\mathbf{D} = \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_n, y_n)\}$

- Background
- Machine learning in microscopy
- Supervised machine learning
- The Problem: Organoids
- Solution: Random Forest + features
- Features
- Random Forest
- Conclusions

Source:

# We now want to learn our mapping

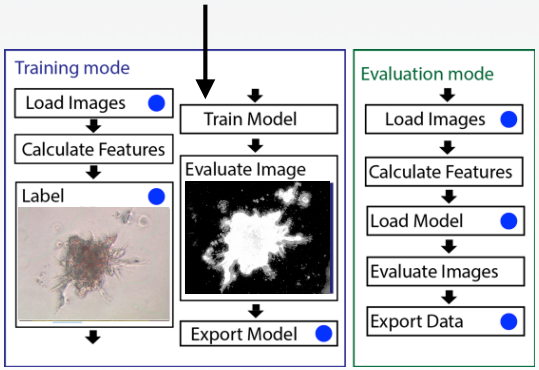
Features for each image +

Matched training data $\mathbf{D} = \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_n, y_n)\}$

$$g : \mathcal{X} \rightarrow \mathcal{Y}$$

**We want to find the function which maps from X to Y.**

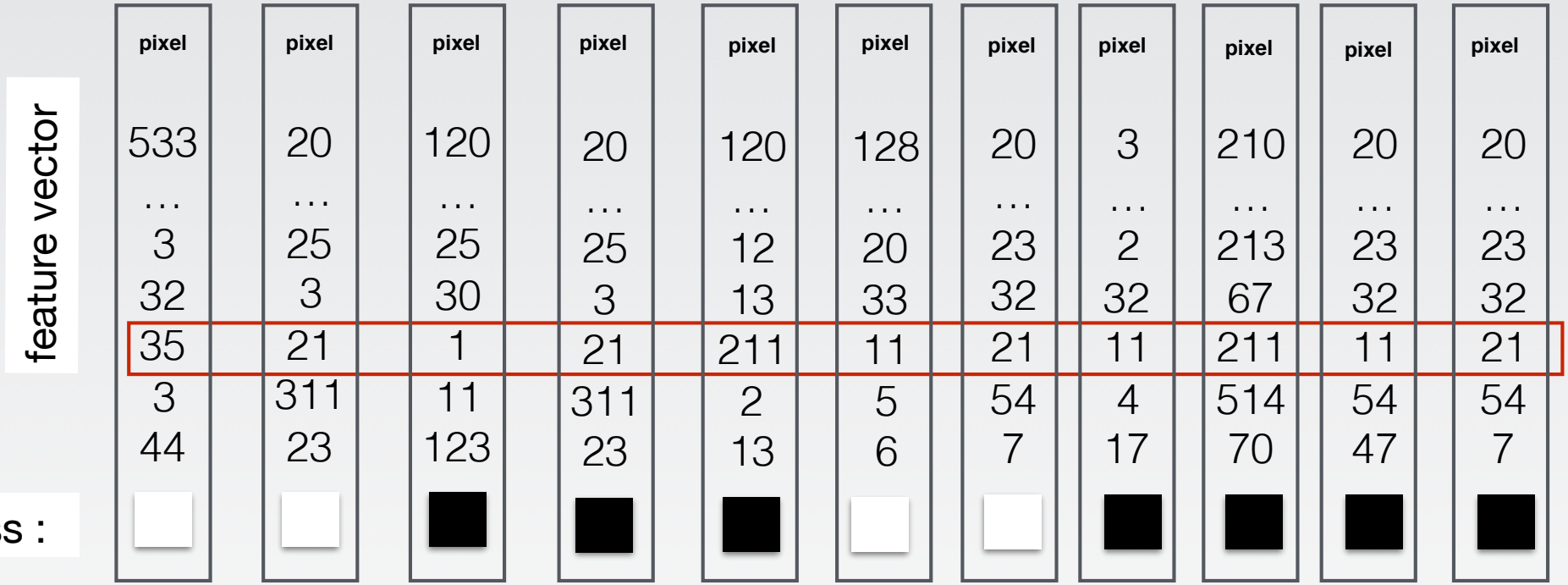
The WEKA system uses a learning framework called a Random Forest.




Source:

# Training-decision trees.

Input data is taken and each feature vector is associated with its training label




repeat until you get high purity.

4 | 7

1) Choose a feature (randomly)

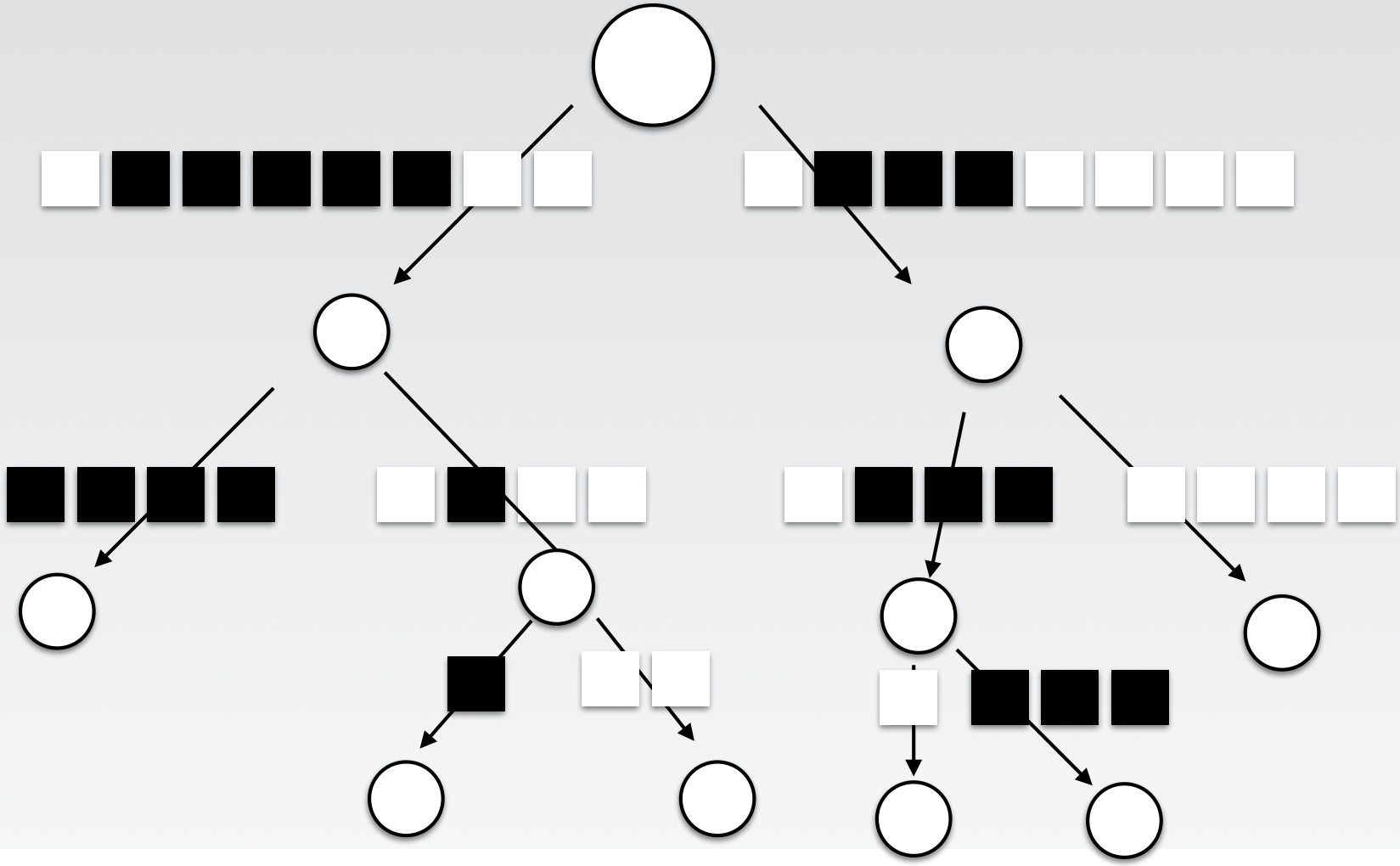
2) Choose a random threshold e.g. 34

3) Assess purity of output

3 5 | 1 2

< 34 | > 34

# Decision Trees

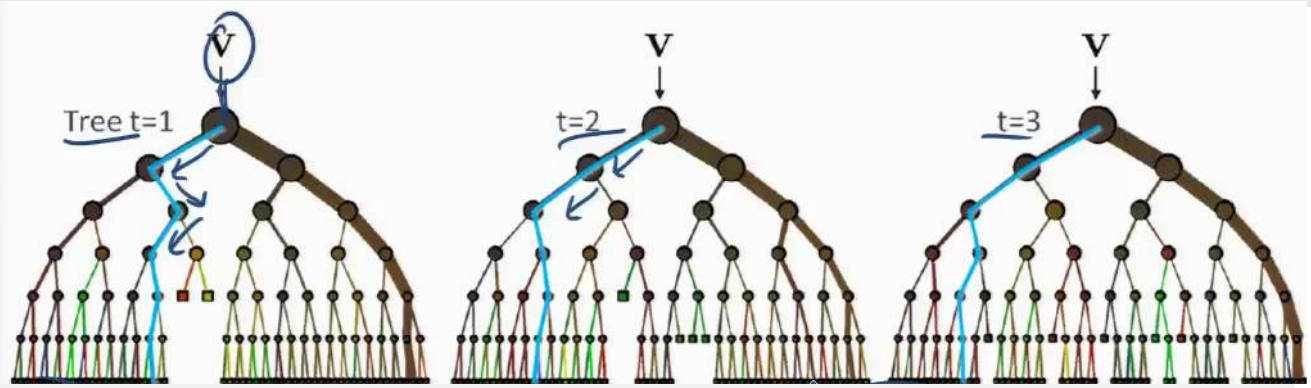


Random forests are comprised of decision trees. Each tree is trained to recognise pixels through there feature representation and arrange them into bins related to the classification (background or foreground).

Source: Leo Breiman and Adele Cutler, http://www.sph.umich.edu/scr/research/project.cfm?deptID=3&projectID=266&groupID=75

# Not just a tree but a forest

We generate not just one  tree but multiple trees. Why? To produce a framework with is more reliable and less fragile.




$$h : \mathcal{X} \rightarrow \mathcal{Y}$$

**h is our approximated version of g**

The trees are similar but we sample the input data slightly differently (bootstrapping) and they are generated freshly each time.

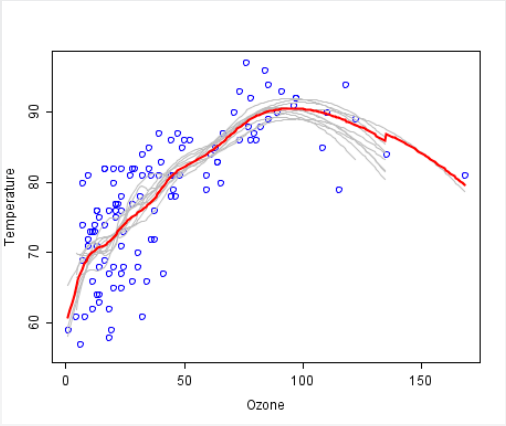
Each tree has slightly different data and also varies due to the generation procedure. For a pixel the classification comes from the average, or dominant output of each of the trees.

Source:

# Bagging (Bootstrap aggregation)

For each tree, the training data is collected through bootstrapping.

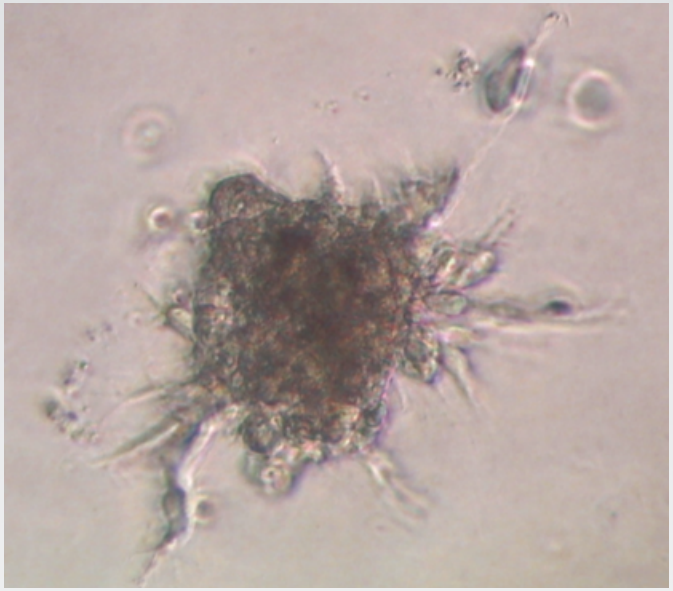
This means sampling from the original data pixels (feature vector + class) uniformly and with replacement. This results in 37.8% replicates.

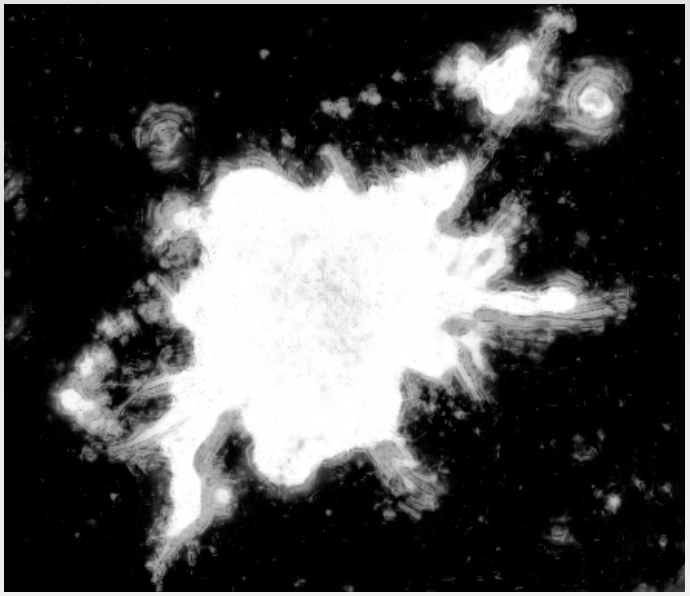



Bootstrap aggregation reduces the effect of outliers on data and reduces the effect that a poor tree initialisation may have on the output.

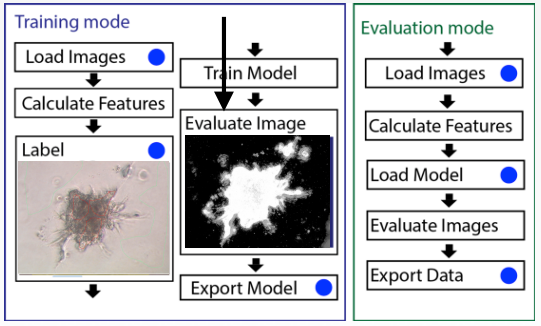
Source: Breiman ; http://en.wikipedia.org/wiki/Bootstrap_aggregating

# Output



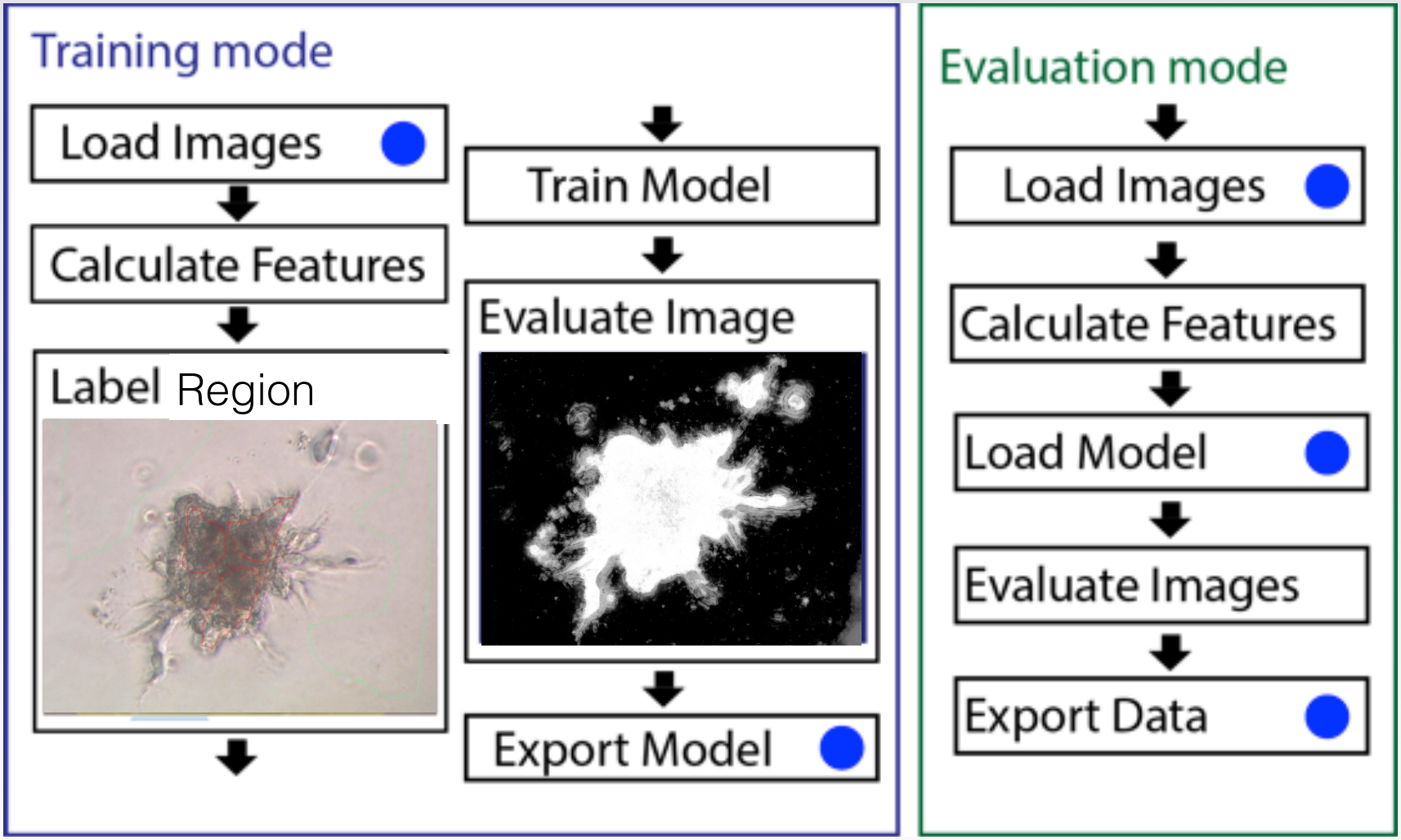


Finally we get an output from the algorithm which gives us the average response form each of the trees for each pixel. We can then threshold this and obtain an accurate representation of the cell.




Source:

# Output




Source:

# WEKA is a java library, python and Matlab?

http://scikit-learn.org/stable/modules/ensemble.html

http://uk.mathworks.com/help/stats/treebagger.html


Feature detection and other learning frameworks
http://www.vlfeat.org/overview/tut.html


Why use Random Forests over other stuff:

-very accurate.

-runs on large databases efficiently (fast).

-fast to train, can be parallelised.

-Can measure certainty of prediction.

-Can handle feature vectors which are large (100-1000's).

Source:

# Conclusion

Machine learning, pattern recognition and image processing intermingle.

Becoming more and more important in the imaging sciences.

Used when simple techniques are not sufficient. Where multiple aspects of the image are useful for completing the task.

Many different approaches, most supervised machine learning follow the same paradigm. Image loading, feature calculation, training, evaluation.

Geometric filter banks and Random Forests are a powerful combination of techniques for describing and classifying image pixels.

Source: