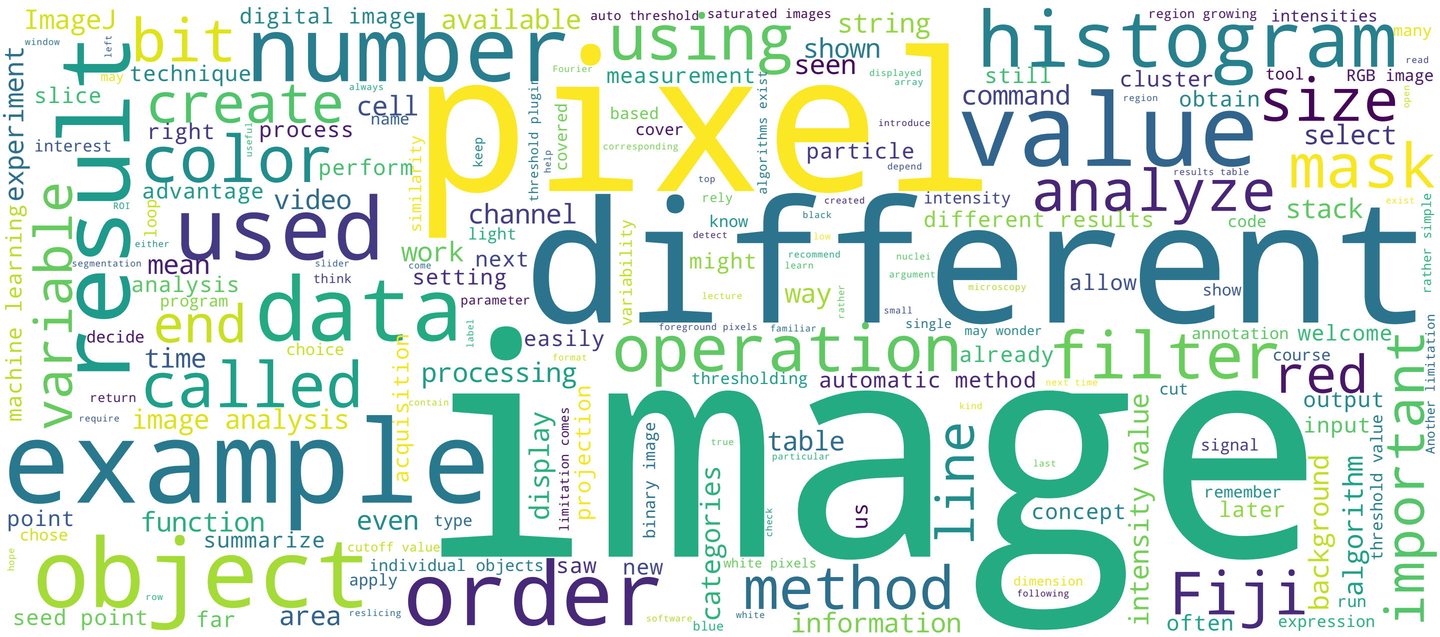
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Segmentation Techniques

Image Processing & Analysis for Life Scientists

Olivier Burri, Romain Guiet & Arne Seitz

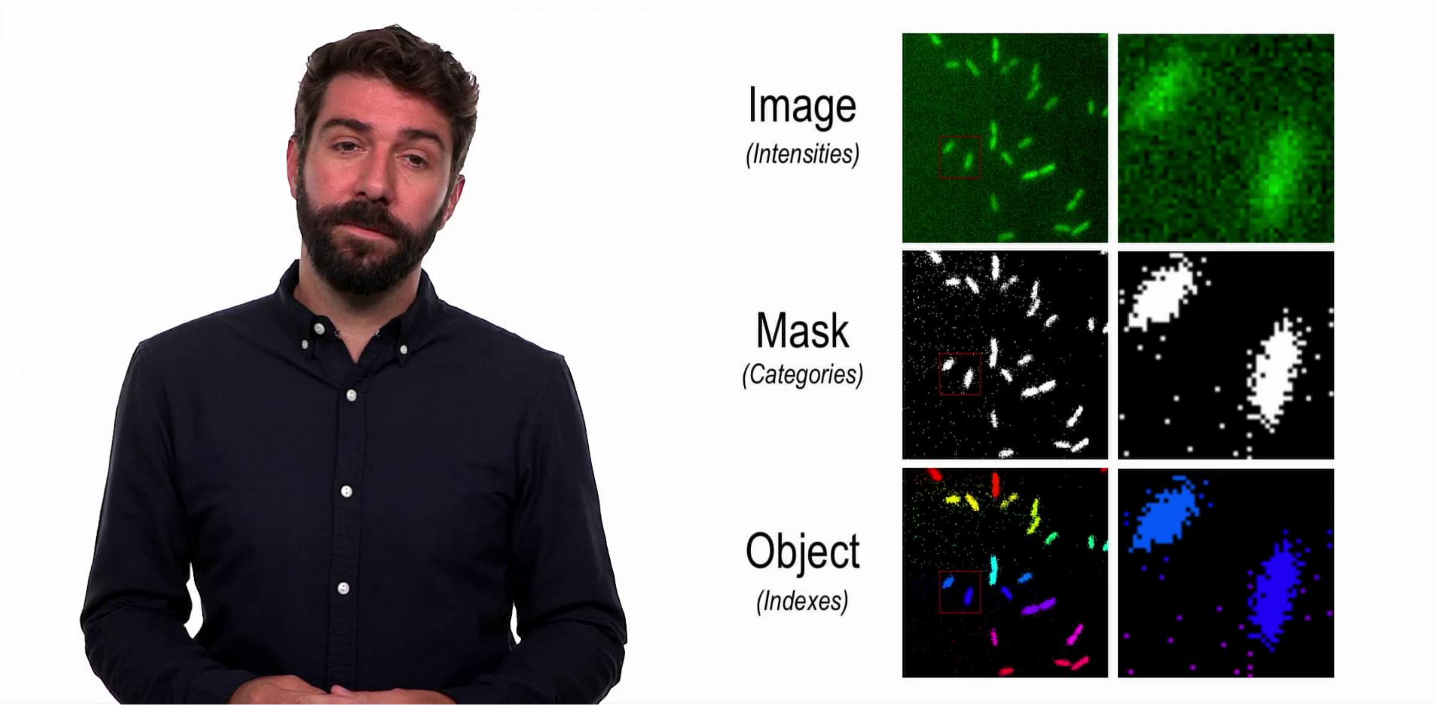


Search MOOC

Video

EPFL

Segmentation Techniques



Hi! And welcome to this new lesson. Today we'll learn a bit more about Segmentation Techniques. Image segmentation is a process of assigning a label to every pixel in an image. Such pixels with the same label share certain characteristics. You are now familiar with the images and their intensities. We'll see how to create binary mask. And later, how to define individual objects.

Notes

Summary

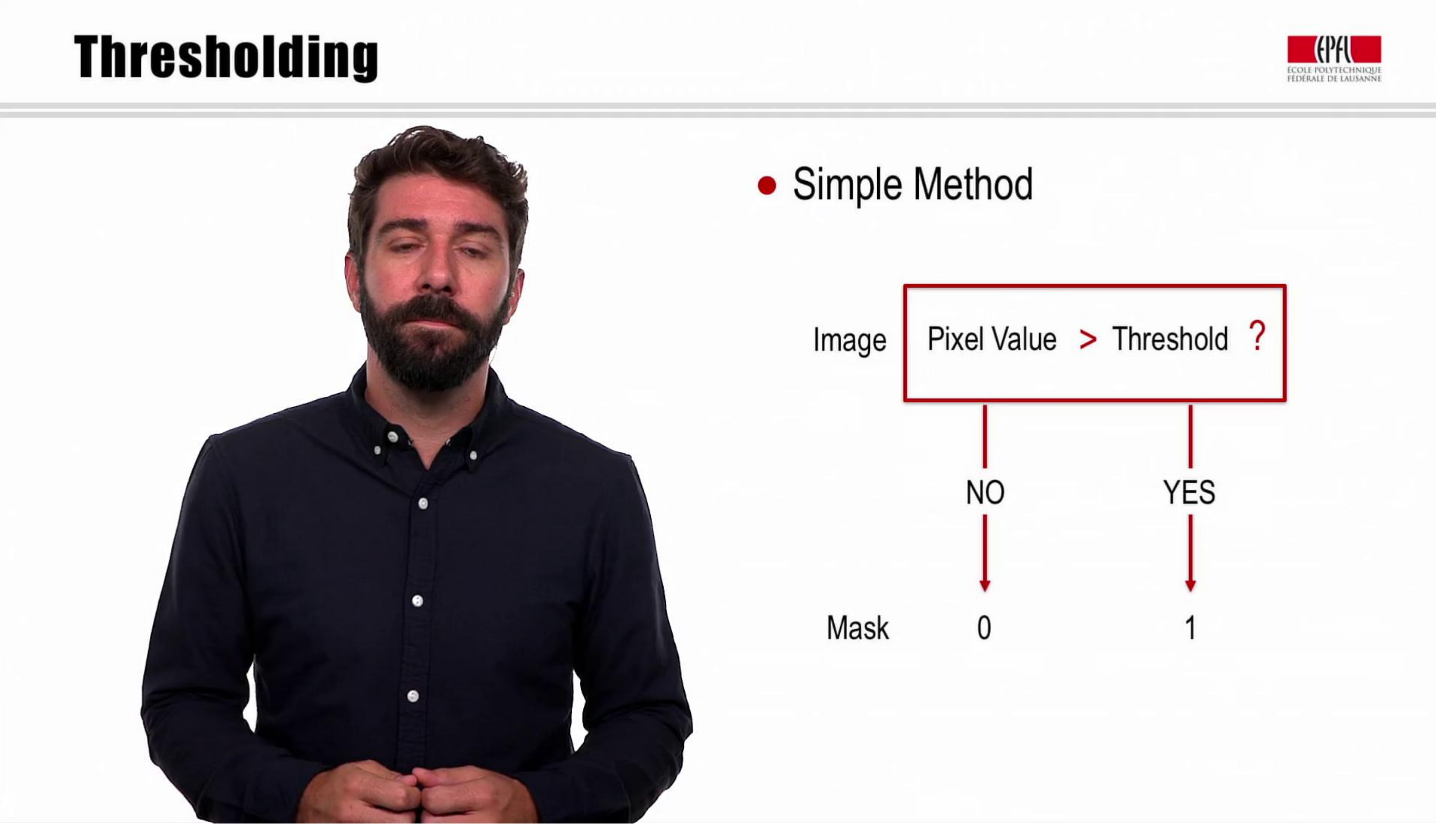
0m 06s

There are different techniques used to achieve segmentation, like thresholding, clustering, region growing, and now, famous, machine learning.

Notes

Summary

0m 31s

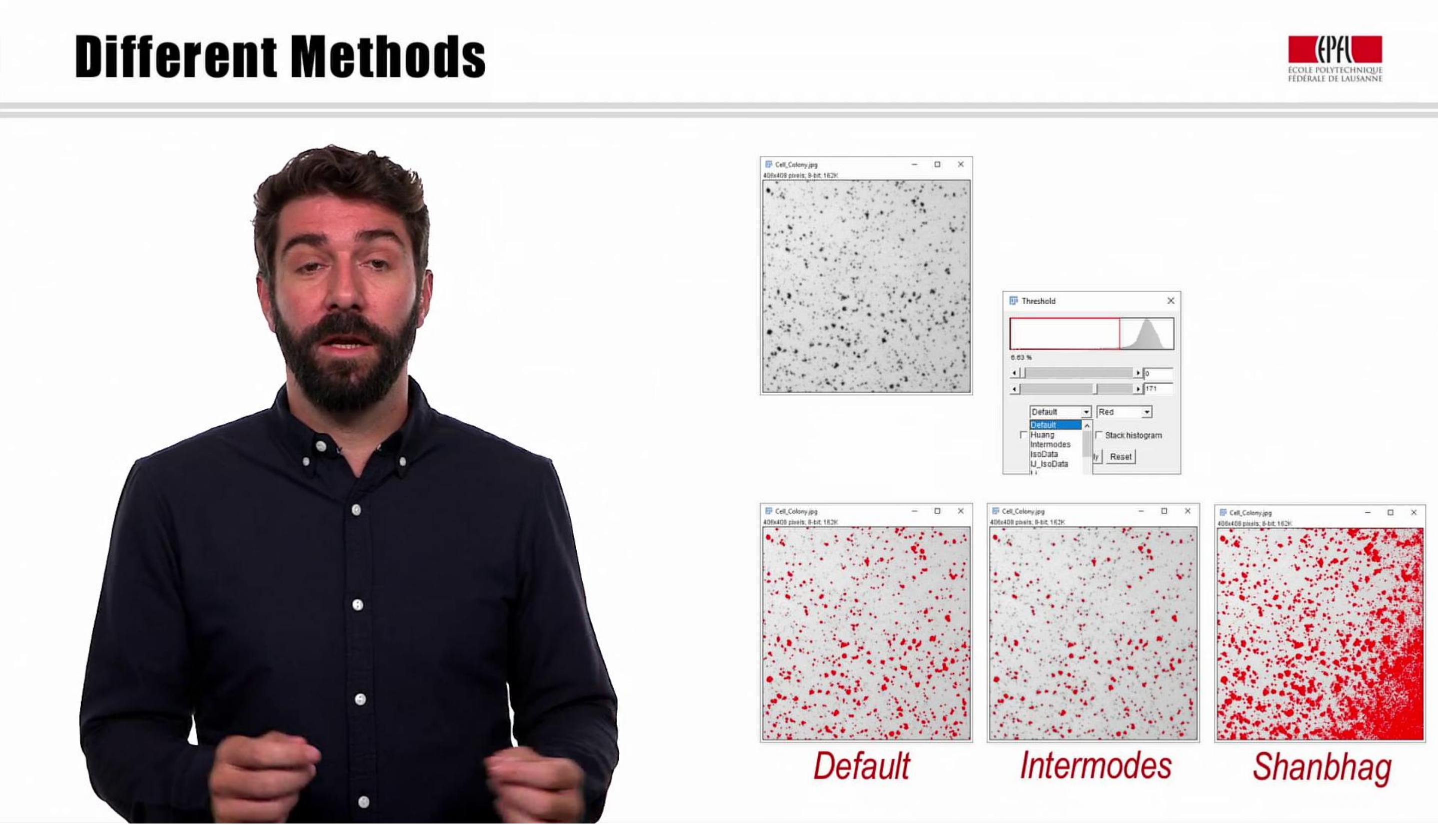


Thresholding is a rather simple method. Considering an image, we analyze each pixel individually and check if its intensity value is above or bellow a certain threshold value. This allows us to categorize each pixel as zero for background, or 1 for signal; and to create an binary image. Also called the mask.

Notes

Summary

0m 40s



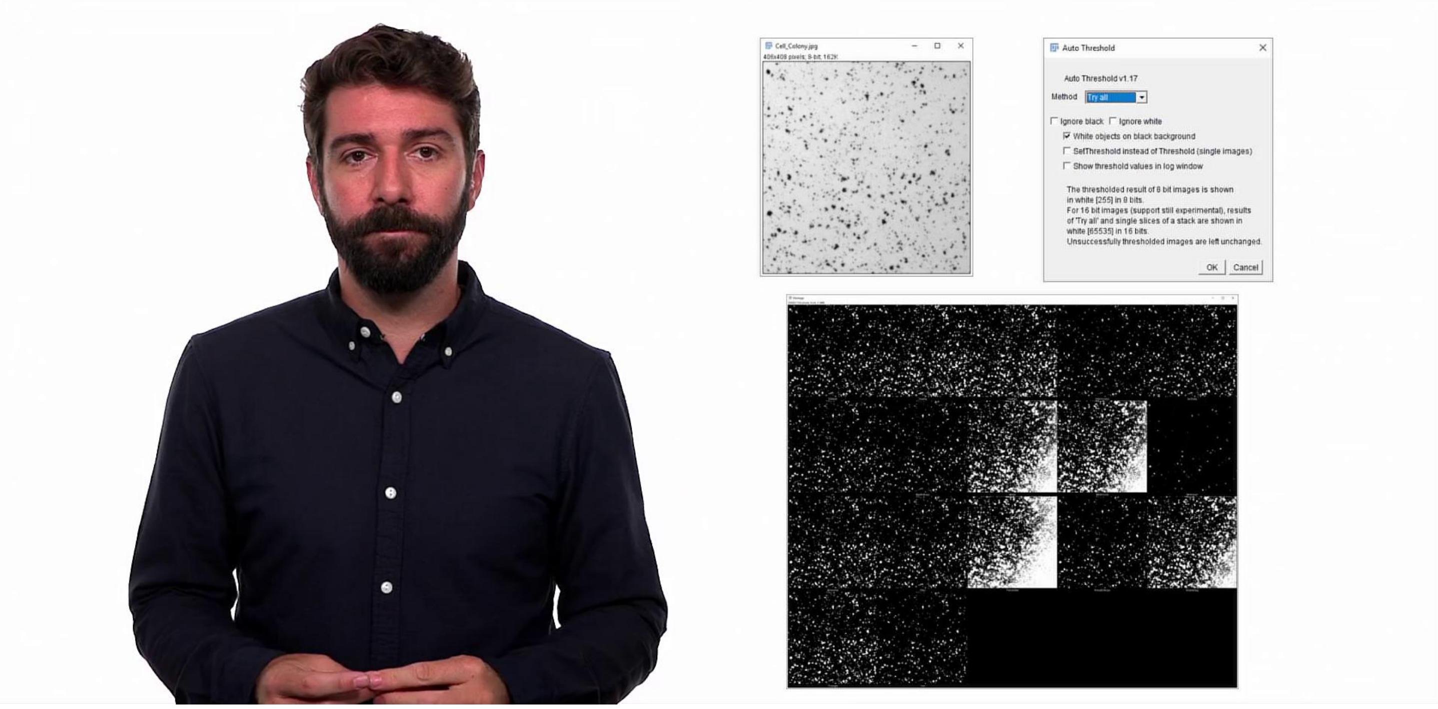
In ImageJ you can use a threshold plugin. That allows you to set manually the threshold value using the slider, and have a preview. Or you can select an automatic method. To generate the mask you just need to press the apply button. There are different automatic methods available in Fiji giving different results. This is because they rely on different algorithms that will analyze the histogram of the image, and then decide of the cut off value. You may wonder how to decide which method you should use. The easiest way is to try all of them, and look at what the results look like.

Notes

Summary

1m 03s

How to choose ?

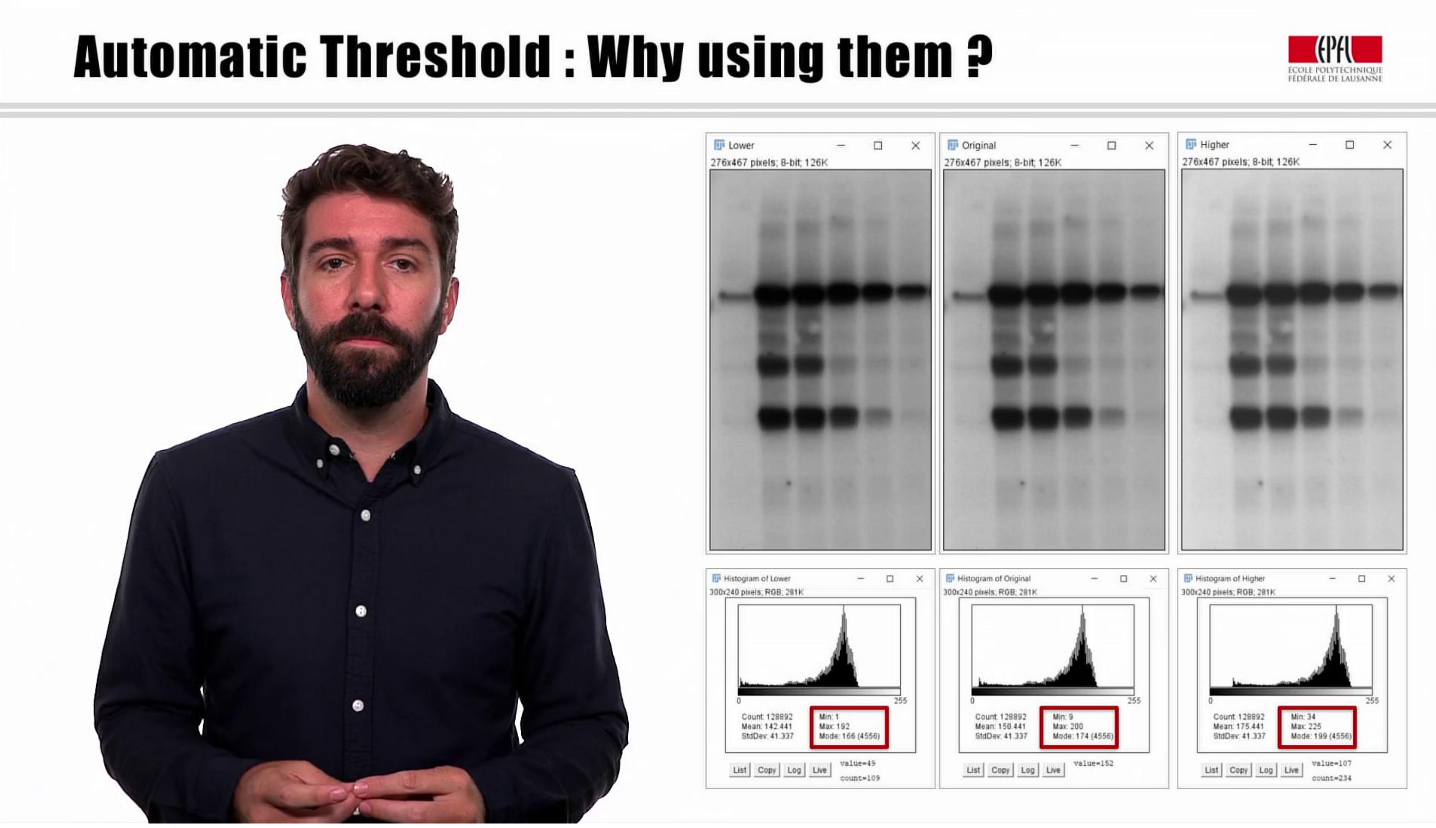


Using the auto threshold plugin you can try the different methods and get an overview of the results. It seems we have 3 big families of methods. Some are setting the cutoff value very low, and you get a lot of white pixels. Others are setting the cut off at much higher value. And you get very few white pixels. Finally you have some intermediate results. This might seem pretty arbitrary. We could read the publication for each method, and chose one that best matches your sample for example. But so far, the golden standard, when performing image analysis, is still the human mind. If it looks right to you, then can try it. But the method you chose should work the same way for all the images in your experiment. This will make your analysis reproducible. And we recommend you that you always test your choice on several images of yours.

Notes

Summary

1m 43s

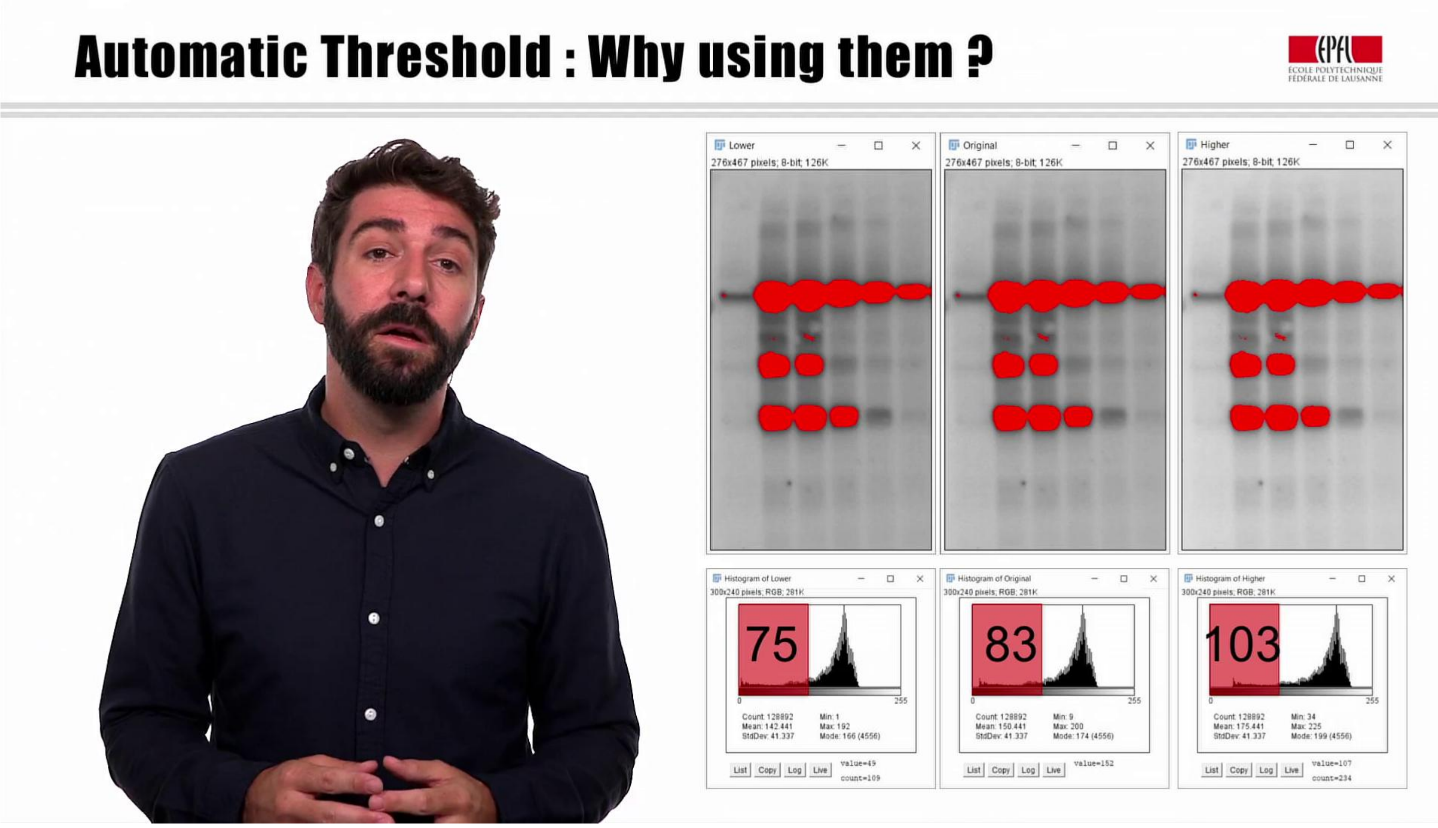


You may wonder why we should use automatic threshold. Let's take an example! Here we have an image of a western blot. And we want to compare 3 different acquisitions of the same same sample. If we look at the histogram of these images, you can see that they have different minimum, maximum, and mode. Indeed even if the acquisitions system in parameter are the same, you can still have some variations, like flickering of the lamp for example.

Notes

Summary

2m 42s



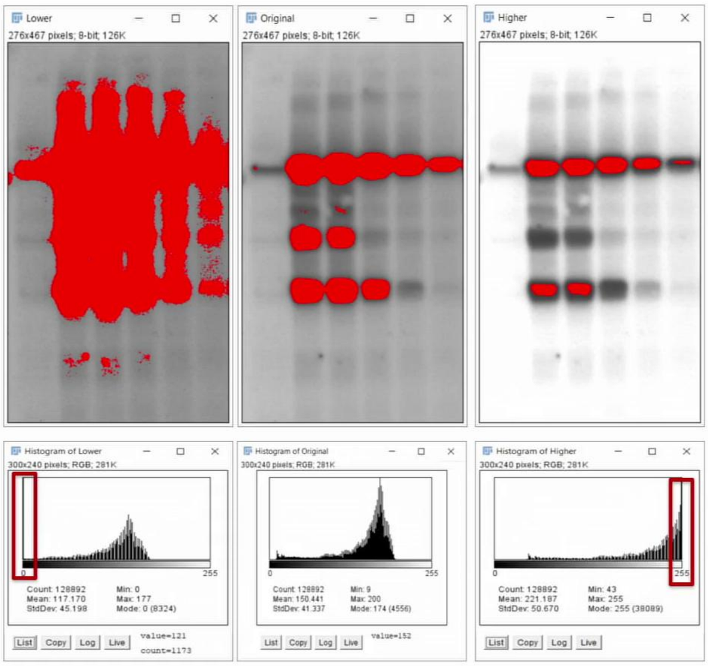
If we manually set a single value at the threshold for the 3 different images, you can easily see here in red that we get 3 different results. The area is different Now if we use one of the available method of Fiji we see that it's the same red area for all 3 images. The algorithm analyzes the histogram and adapts it to the value of the cutoff. So it selects the same area.

Notes

Summary

3m 14s

Automatic Threshold : Limitations

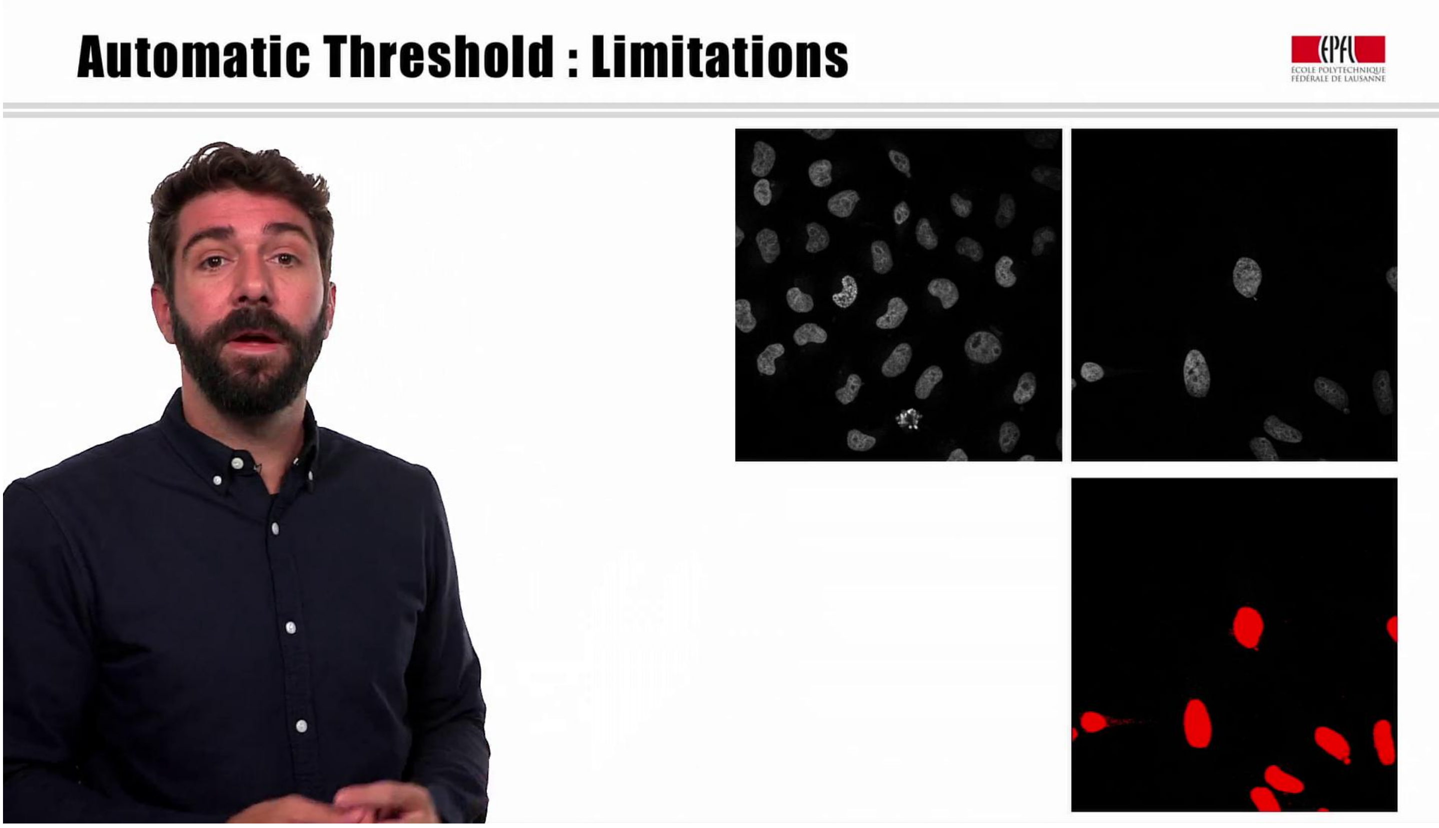


So you see it's a robust method to analyze scientifically some images. Now let's have a look to the limitations of the technique. Let's re-use the previous example but now with saturated images. You remember, too many 0, or too many 255 value. You see that now the auto threshold fails to analyze the histogram in a robust way and gives very different results for each image. That's why we insist so often on how it is important not to saturate your images. Saturated images break the mathematical models that are used by automatic methods. And even though our brain has no problem coping with this issue, computers cannot.

Notes

Summary

3m 44s



Another limitation comes from the variability of your images. In this experiment the different fields of view have different cell density. Then even if one can find a special method that detects the nuclei pretty well, it may totally fail in the next image.

Notes

Summary

4m 34s

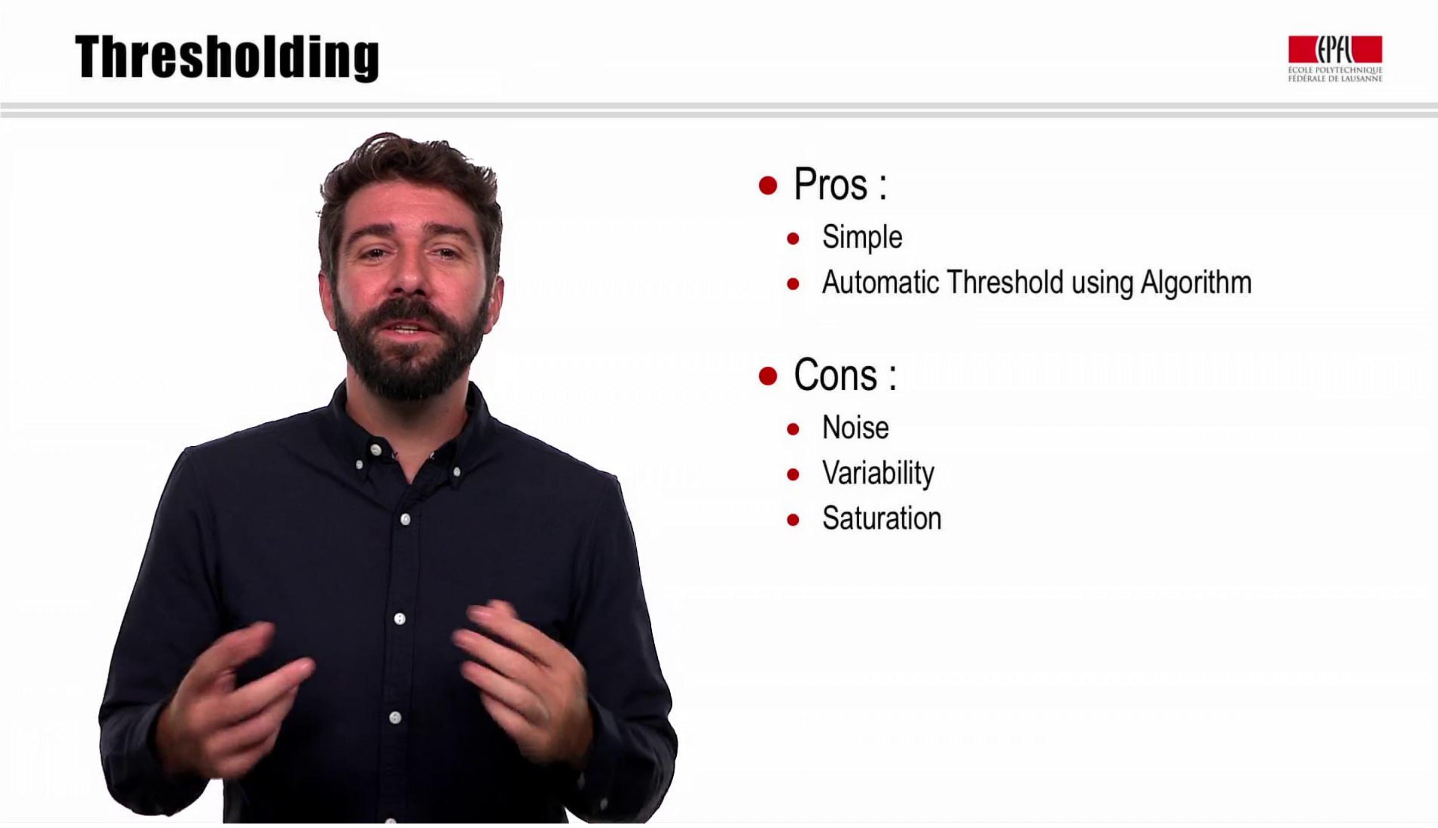
Automatic Threshold : Limitations

Because it has a very different histogram. Finally, if you have noisy images this thresholding could highlight this noisy dusty looking pixel.

Notes

Summary

4m 51s



Ok! Let's summarize what we have seen about thresholding. It's a rather simple technique, and some algorithms exist to analyze the histogram and to make an automatic elections of the cutoff value. But noise can create dust in the resulting image. And the variability between your images may restrain your choice of methods. Finally, you should always avoid saturation in your images.

Notes

Summary

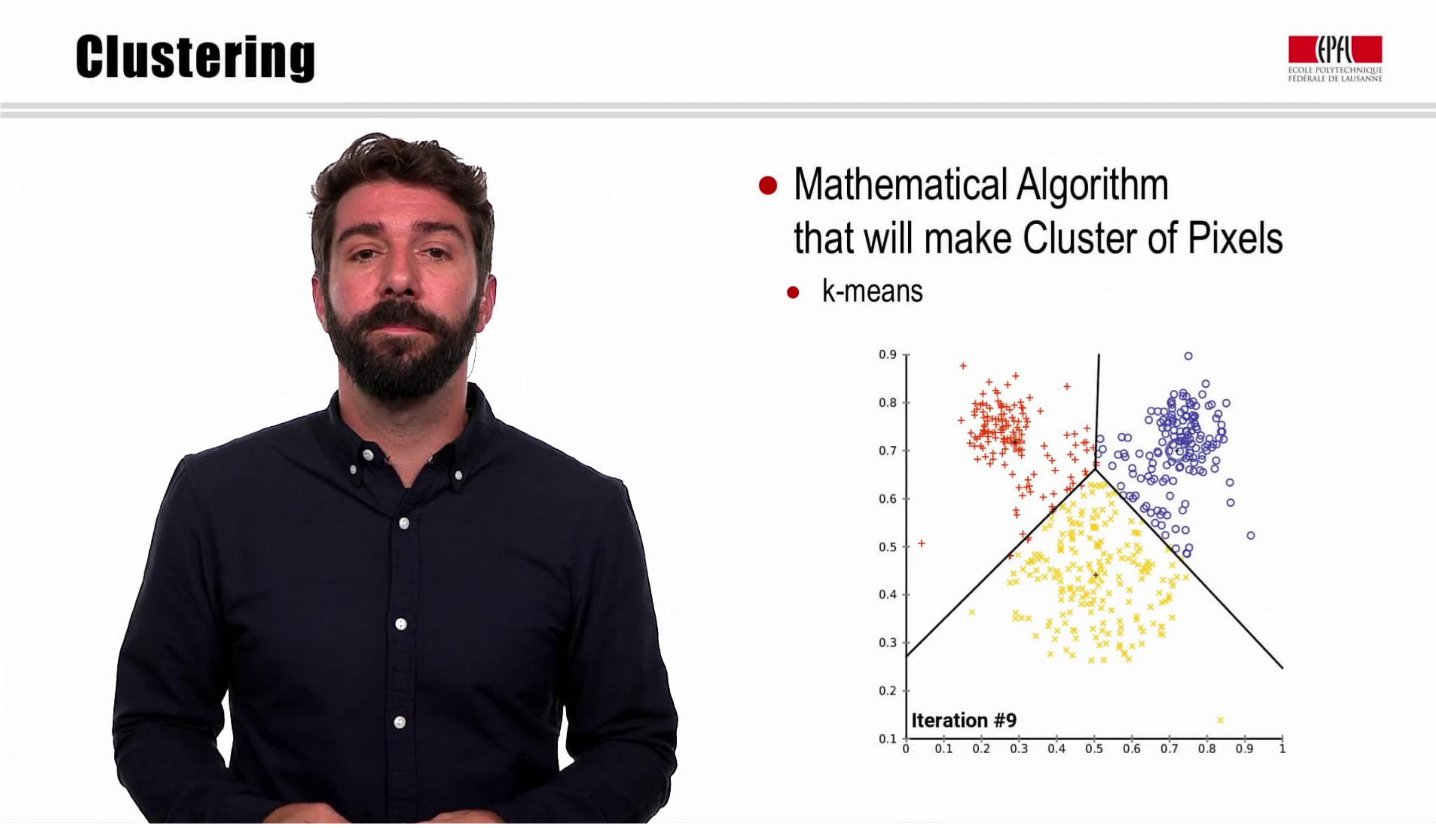
5m 02s

The description of the thresholding technique was quite long. And I'm sorry about it. And I will go a bit faster on the other techniques. Let's continue with the Clustering.

Notes

Summary

5m 31s

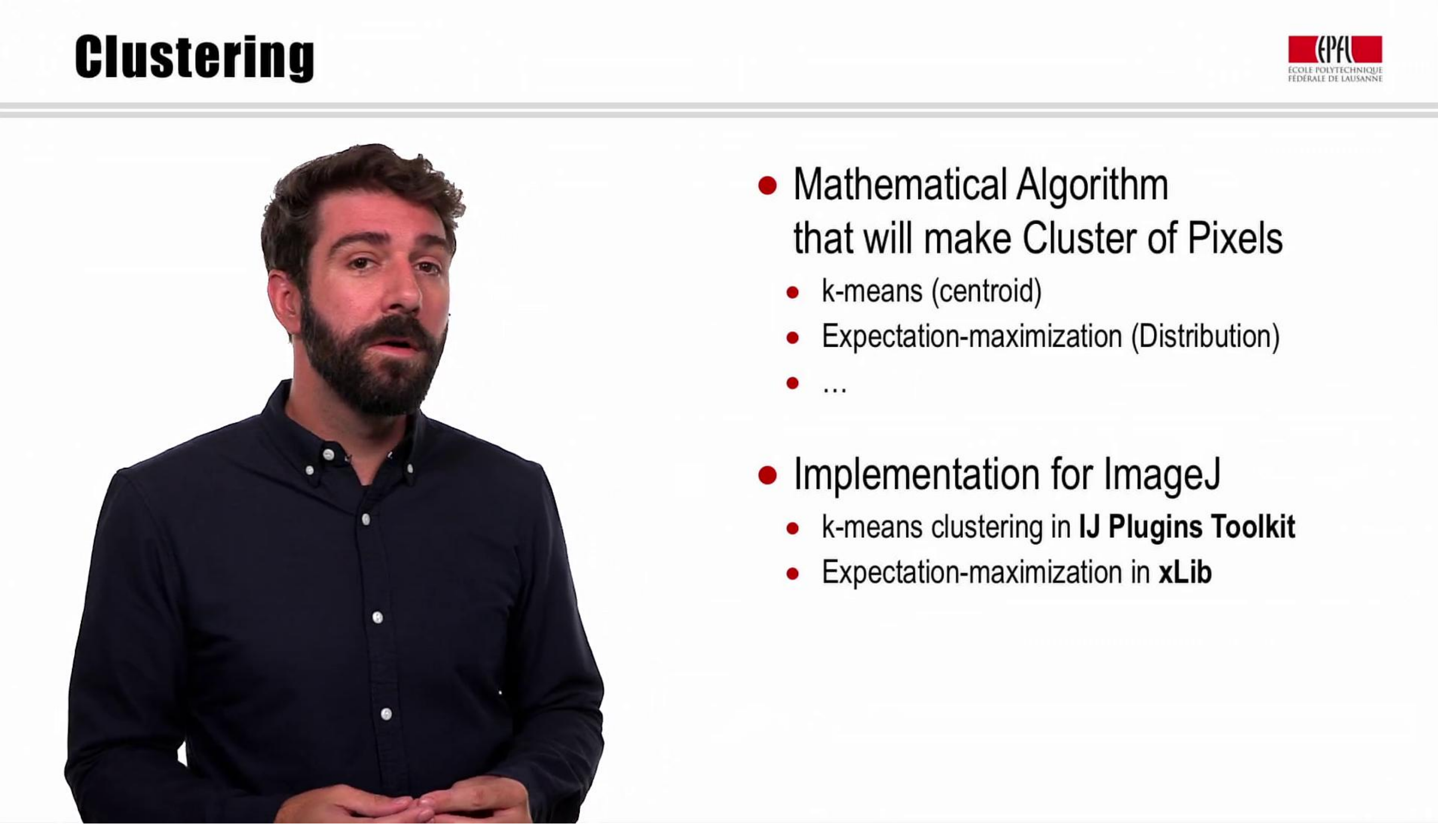


They rely again on mathematical algorithms that create clusters of pixels. You may have already heard about one of them: The K-means clustering. The K initial means (in this case 3) there are a number of cluster we want to have in the image. Then K cluster are created by associating every observation within the nearest mean. The centroid of each of the K cluster becomes the new mean. The last 2 steps repeat until we reach a convergence.

Notes

Summary

5m 43s

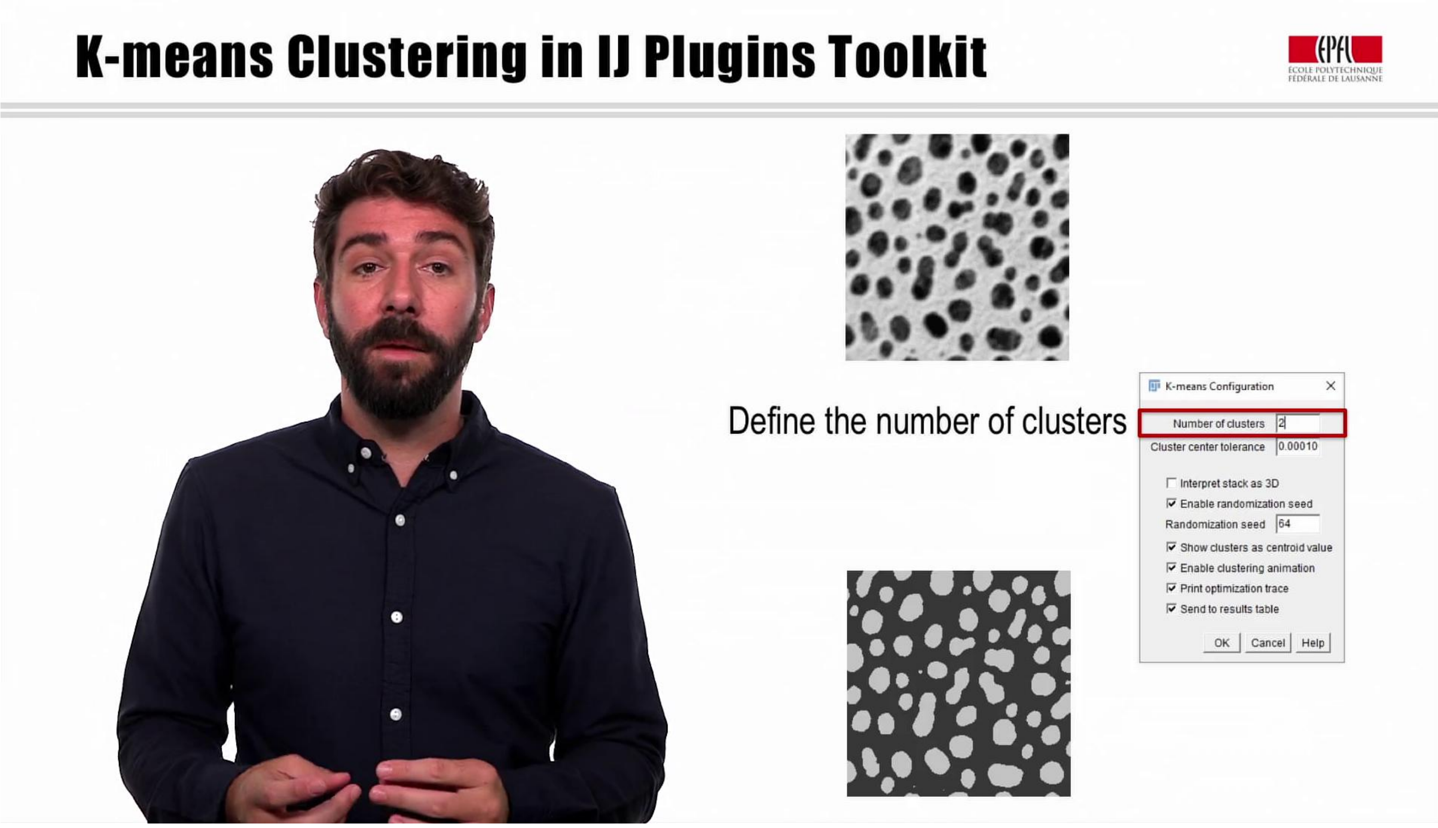


Of course other algorithms exist, like expectation maximization. And in Fiji you will find corresponding implementation.

Notes

Summary

6m 13s

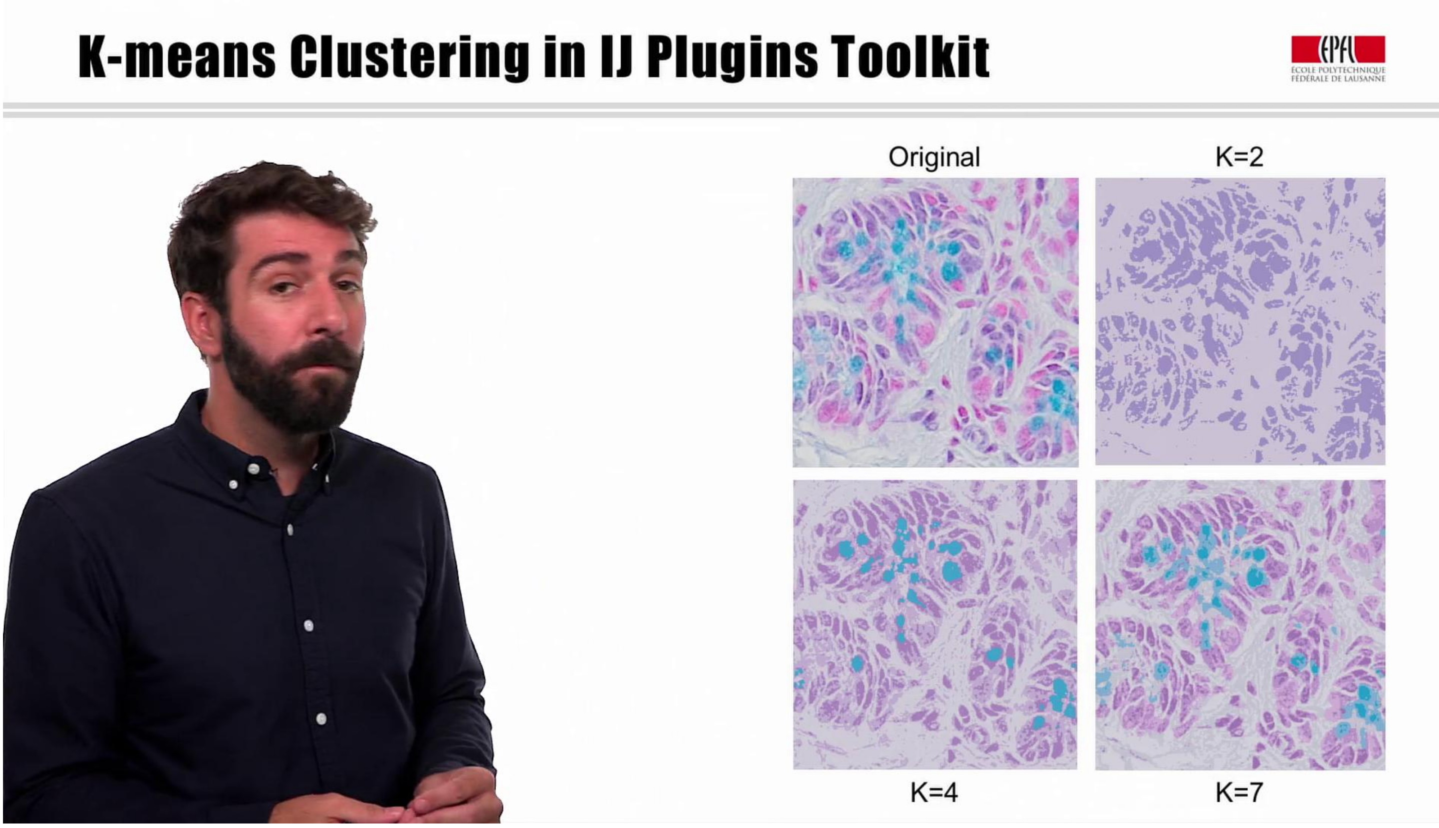


If we take the famous Blobs image as an example, we will have to define the number of clusters. One inerresting thing is that we can also use k-mean on RGB images.

Notes

Summary

6m 25s



And there you see what results we can expect when you look for different k-numbers on the same image.

Notes

Summary

6m 37s

Clustering

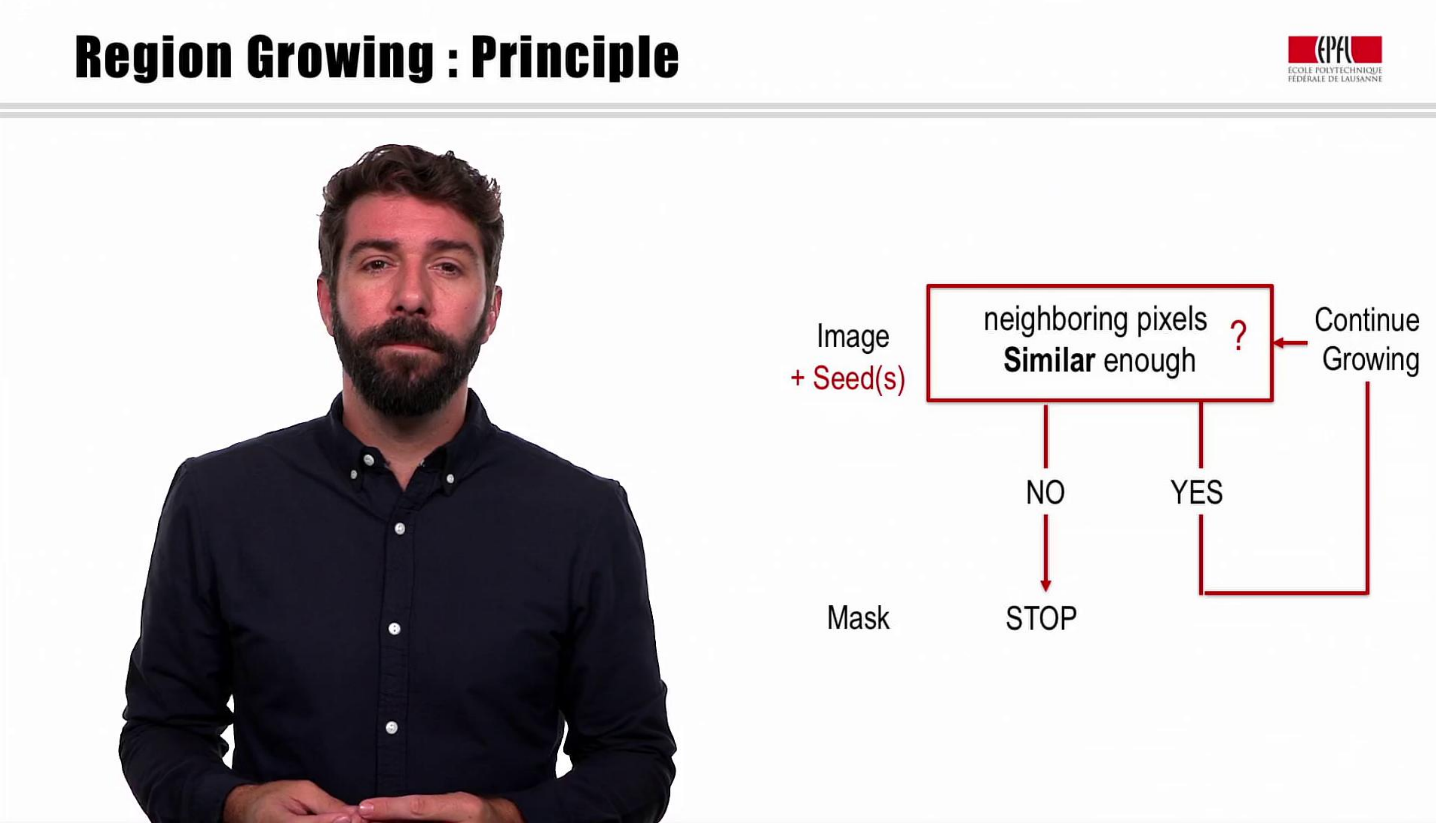
- **Pros**
  - More than 2 categories
  - Works on RGB image
- **Cons**
  - Needs apriori knowledge (categories number)
  - Initial Centroids
  - Local Minima

So the advantages of clustering techniques are that you can have more than 2 categories; and that most implementations let you work with RGB images. Nevertheless it requires some a priori kwoledge: the number of categories. And the results could depend on the number of initial centroids. Finally, because the seed points are initialized randomly, the algorithm can fall into local minima.

Notes

Summary

6m 44s



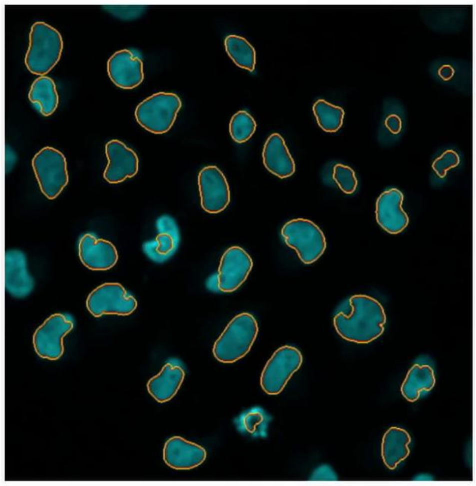
So let's have a look at Region Growing. The method requires some seed points. And the algorithms will look at the similarity of the neighboring pixels. If they are similar enough the region will continue expanding or retracting. Otherwise it stops. Finally, when the convergence criteria is met, or there are no other pixels to analyze, you get the resulting mask.

Notes

Summary

7m 10s

Region Growing



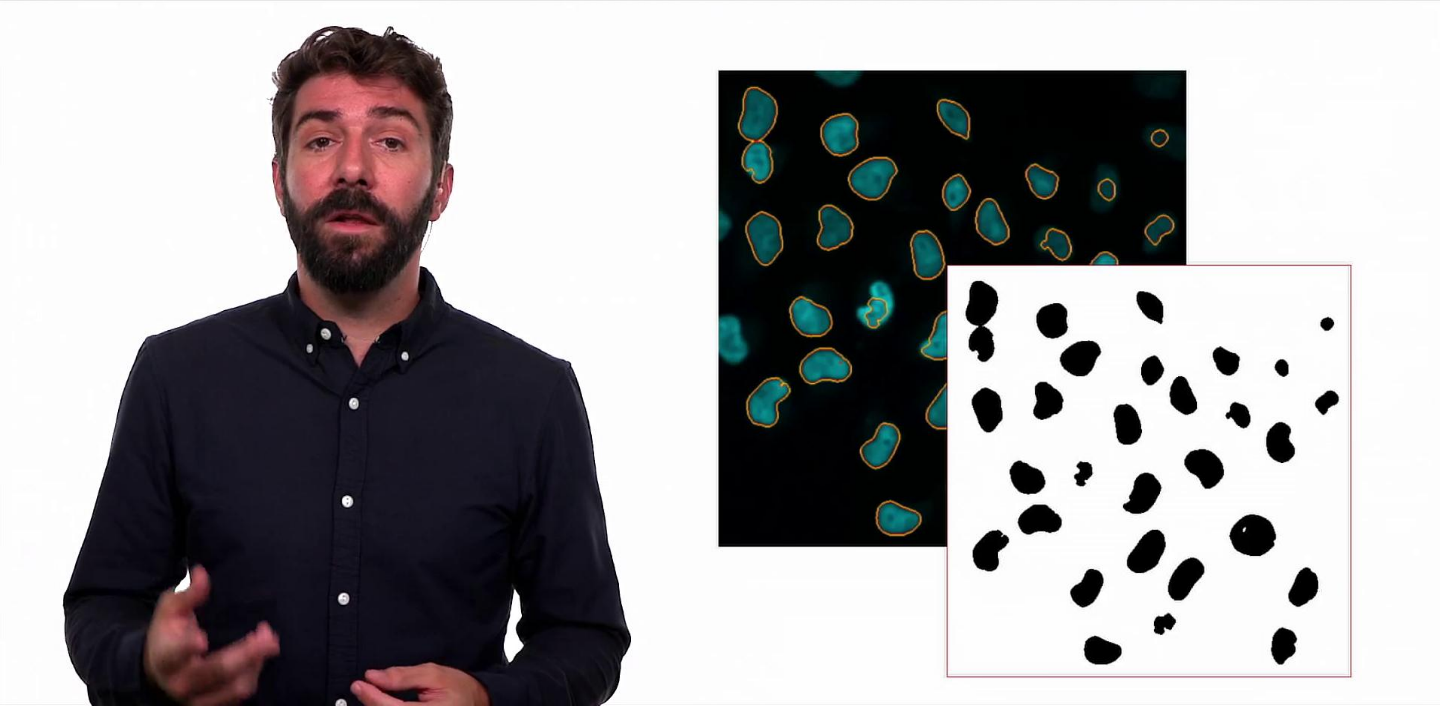
You can see on this example video that seed points are growing fast at the beginning, and then converge to an equilibrium.

Notes

Summary

7m 38s

Region Growing

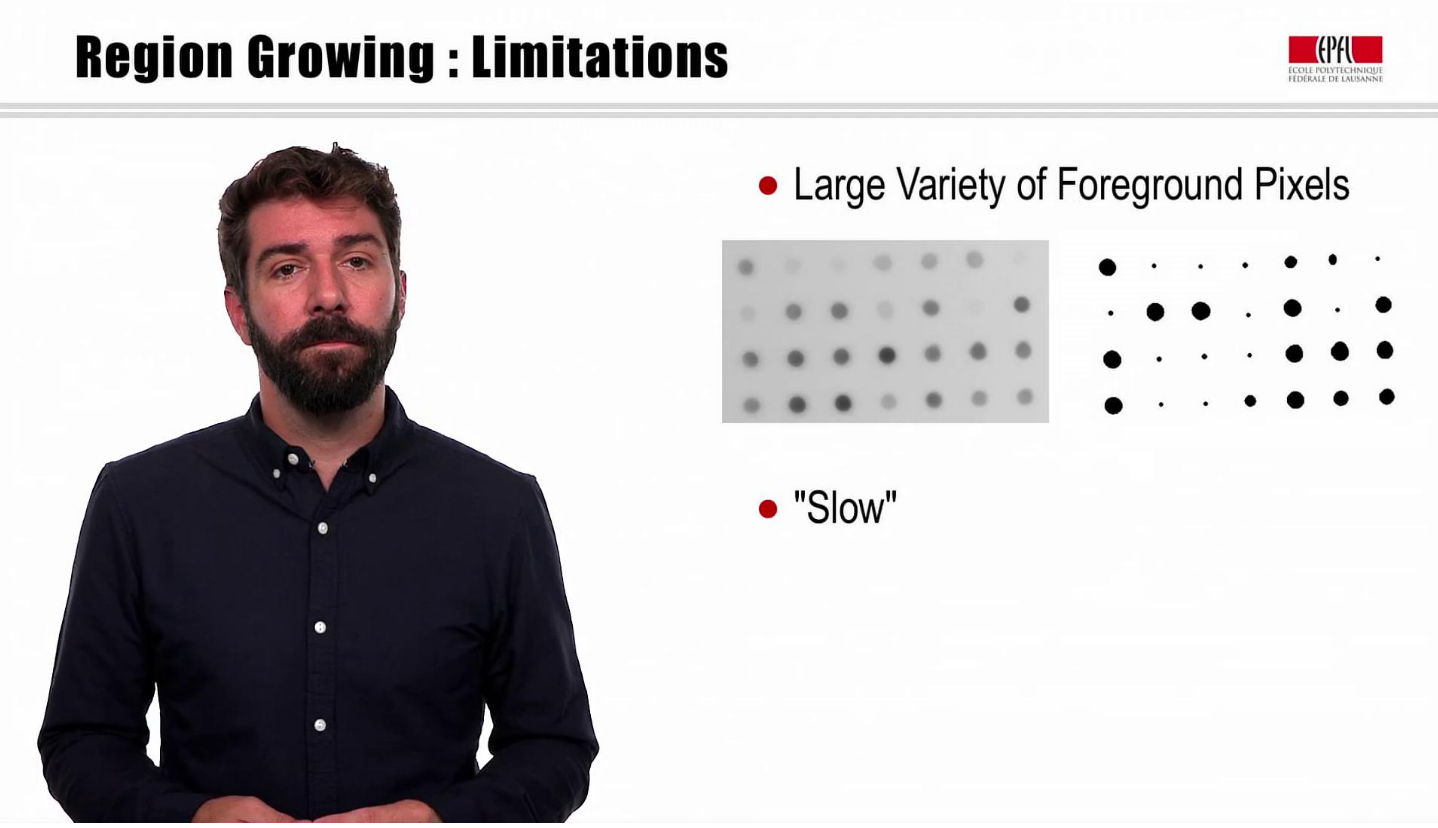


If we wait a bit longer we get the final output.

Notes

Summary

7m 46s

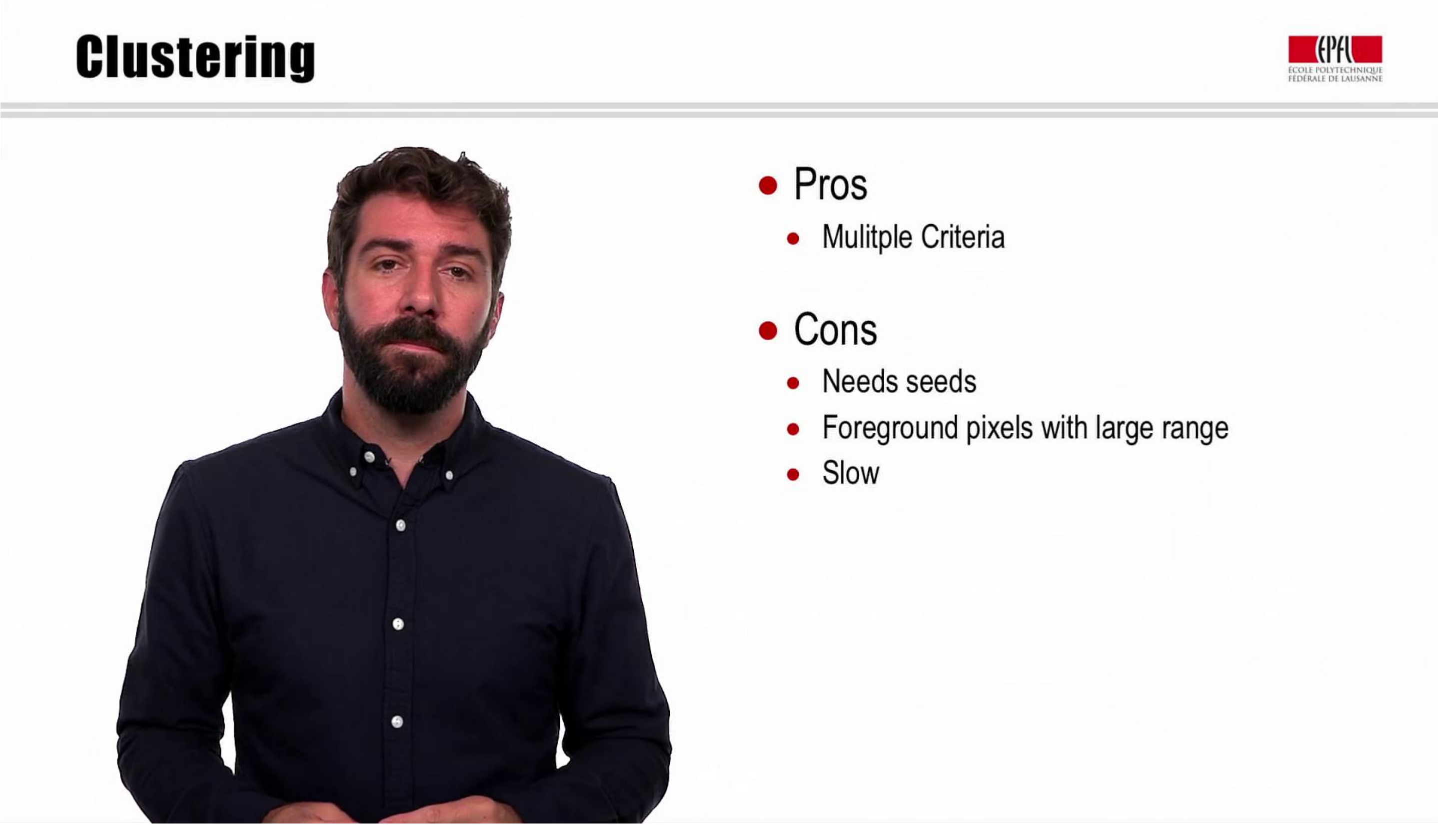


One limitation comes from images where the pixels of interest, the foreground pixels, have a large range of value. Another limitation is the speed of the processing.

Notes

Summary

7m 50s

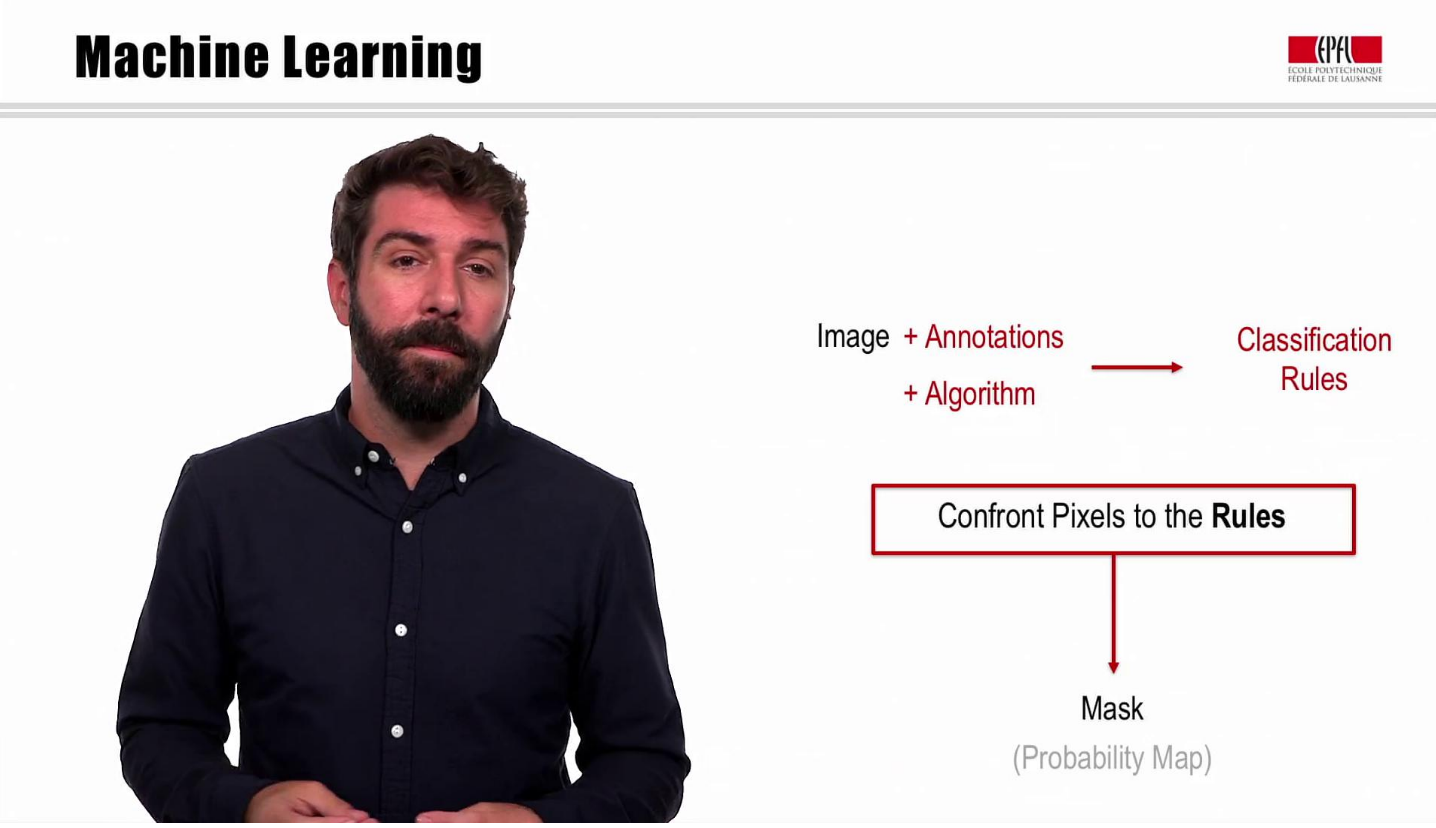


So one advantage is that you have multiple criteria to define the similarity of the pixels. On the other hand you need seeds. There is a limitation ranged on foreground pixels, and it's pretty slow: seconds to minutes.

Notes

Summary

8m 02s



Let's finish with a quick look at Machine Learning. On top of your images you will need some annotations and an algorithm. It will create some classification rules based on your annotations. This is the teaching learning part. Then the processing will just be a classification of each pixel, using the defined rules to create a mask. Or we often refer to it as Probability Map.

Notes

Summary

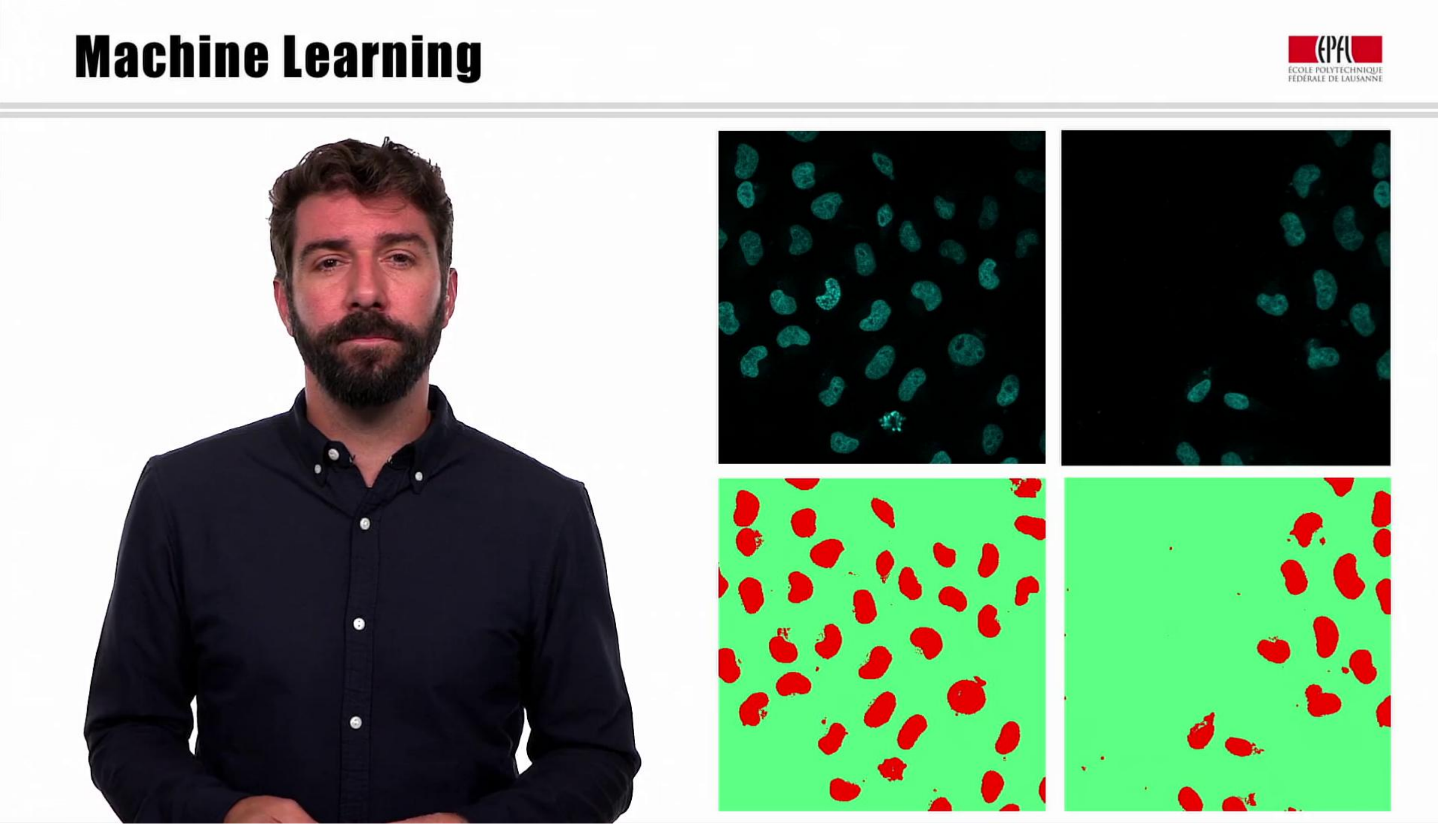
8m 18s

Here you can see annotations manually drawn where the red and the green lines define the pixel belonging to eather the nuclei or the background. Doing so on different parts of the image, and on different images, you will try to teach the program how to recognize these different categories. And it will give you the ouput with the classification of the overall pixel in the image.

Notes

Summary

8m 49s



Then you can load another image, re-apply the same rules called the classifier; and get a new result.

Notes

Summary

9m 18s

Machine Learning

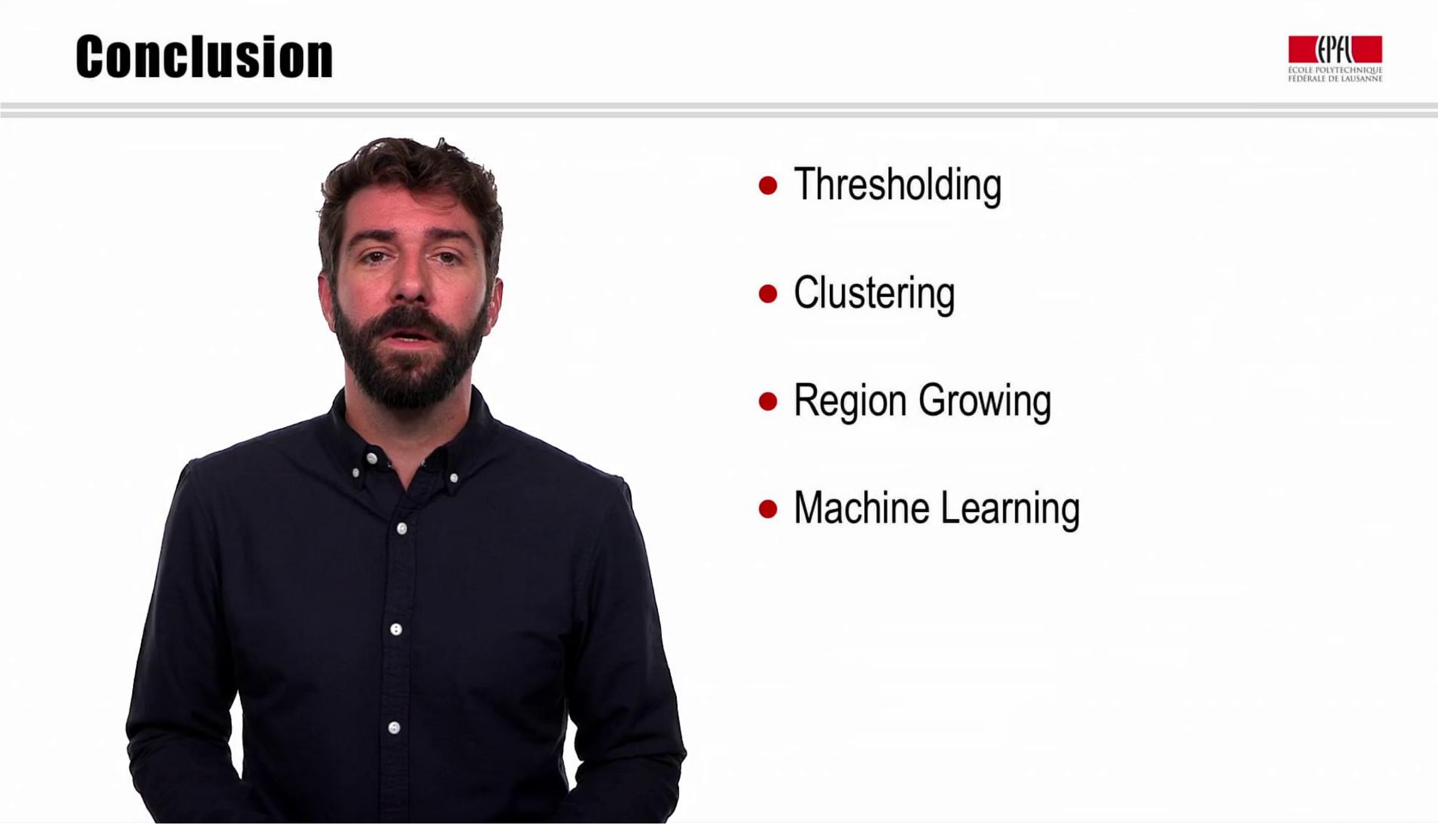
- Pros
  - Complex Classification
- Cons
  - Needs «Good» Annotations
  - Robust Imaging Condition

Machine learning can help solving complex classification problems (like in brightfield microscopy for example). But it's very delicate and needs good manual annotations and robust imaging conditions.

Notes

Summary

9m 25s



This is the end of the lesson! We covered several ways in which we can transform image data into masks that define categories for each pixel of the image. Later we'll see how we can use this mask to obtain individual objects, and extract meaningful statistics from them. See you next time!

Notes

Summary

9m 40s