

# Medical Image Analysis Assignment 2

Casper Bresdahl

## 1 Introduction

This week we will look at where segmentation can be used in medical image analysis, look at the risks and benefits of using dilation, then look into segmentation from the Random walker and Graph cut algorithm and end by looking at PCA.

## 2 Medical image segmentation

Segmentation is widely used in medical image analysis where knowledge of an object's shape, size or volume is needed. It could be there is a suspicion that a patient's heart is enlarged, then finding the shape of the heart from an x-ray image would allow for easy measurement of its size. However, this would require someone to draw the shape of the heart from that x-ray scan. With the development of segmentation algorithms this can now be automated. Medical image segmentation can also be used for imaged guided procedures, where a surgeon would have to operate with great precision, for example when inserting pronator screws. Medical image segmentation is also, as touched upon before, more broadly used for computer aided diagnosis where perhaps the shape of a patient's organs is measured and used for a diagnosis. Other examples would be to apply thresholding on a CT image to extract the bones, to apply Region growing in mammograms in order to extract a potential lesion from its background, and to separate different tissues from each other based on EM results [3].

## 3 Dilation and erosion

One must employ care when using the morphological operators dilation and erosion. As they are not inverses of each other, the effect of applying one and then the other is not the same as applying them in reverse order. An example of this can be seen in Figure 1. In the example we also see the advantage of erosion and dilation as the figure is now closed, i.e. we removed the hole.

In a real scenario one might want to close holes in an initial segmentation of microvasculature (Figure 2a) or in a brain (Figure 2b), but if one uses erosion on the microvasculature they risk disconnecting the vessels or even worse removing tips of the vessels. In the brain segmentation if dilation is used one risks the brain folds collapse and the brain becoming one big blob. Thus, care must be taken when segmenting thin objects not to make erosions, and care must be taken not to make dilations when details in a segmentation spatially lie very close to other

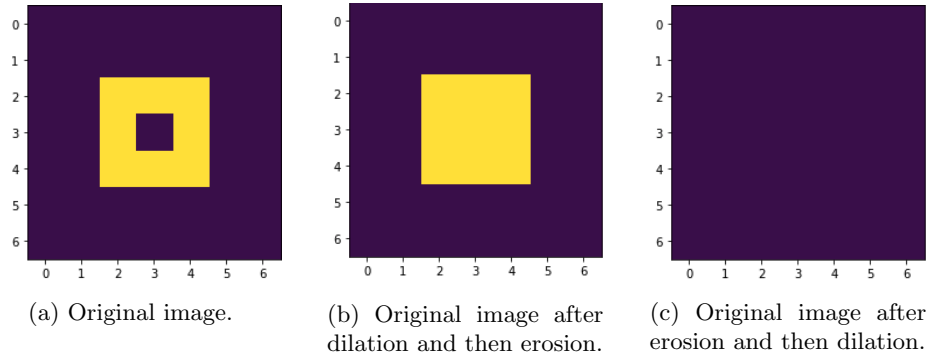
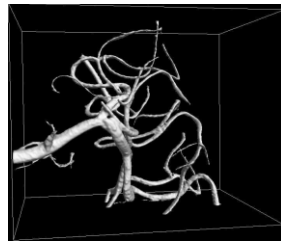
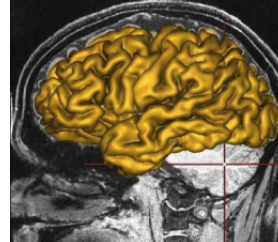


Fig. 1: Example of applying dilation and then erosion is not the same as applying erosion and then dilation.

details in the segmentation. However, when care *is* taken, one can close the holes and remove leaks from their segmentation using dilation and erosion, and thus obtain an improved segmentation.



(a) Initial segmentation of microvasculature.



(b) Initial segmentation of a brain.

Fig. 2: Examples of where care must be taken with dilation and erosion. Images taken from slides.

## 4 Random walker

Random walker is an algorithm which given seeds denoting different object classes has a 'walker' starting at a pixel location walking around the image with probabilities of where he is going proportional to the similarity between the pixel he is in and the neighbouring pixels. When he meets a seed, the class of the seed is recorded and after  $x$  trials, the walker will begin in the next pixel of the image. Each pixel then gets the class its walkers reached most times out of the  $x$  trials. This gives a segmentation of the whole image into any number

of classes, but requires seeds are placed in the image beforehand. This describes an iterative solution for the algorithm, but a solution can be reached more efficiently solving the diffusion equation.

If we for instance want to segment the *cell* image we need to first create an 'image' of seeds. This has been provided for us, but it only has one class (cell or not cell). We can thus use connected component decomposition to separate all connected points of seeds into their own class. We can then run the Random walker algorithm and we get a segmentation of each individual cell. This can be seen in Figure 3. As seen in the overlaid image we correctly find each of the cells we made seeds for, however, the quality of the segmentation is questionable as many cells are jagged, with some even having 90 degree angles. We also see quite a few leaks, and a few holes. However, the quality could be improved by using one or more erosions and then dilating the cells again, as this would remove the leaks and the rough edges and close the holes. We can also use Random walker for 3D segmentation, which can be seen in Figure 4, where we segment a mandible.

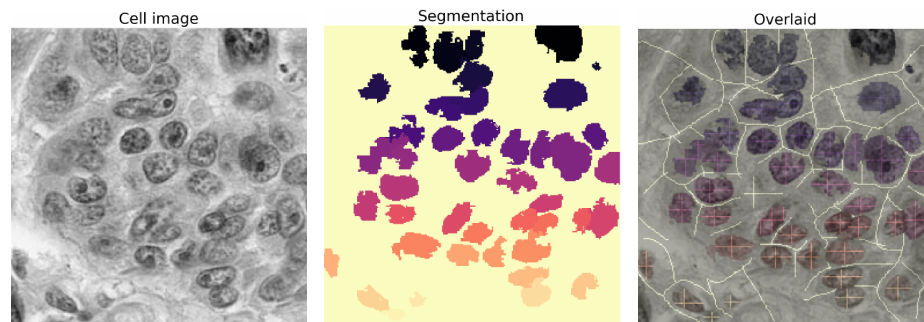


Fig. 3: Original image, the final segmentation and markers, segmentation and original image overlaid.

Another algorithm which can be used for segmentation / to find object boundaries is Graph cut. Here the image is made to a graph representation and the more similar neighbouring pixels are, the more flow we allow between them. A max flow problem is then solved and a min cut is made to find the object boundary. As we are solving a max flow problem, this approach gives us an analytical solution to our segmentation problem opposed to the non-deterministic approach of using Random walker. This is a big advantage as the solution to the problem becomes consistent, however, it also requires it is possible to make a graph representation of the image and that it is possible to come up with a cost function. And even if this is possible, it requires that the graph representation can be contained in memory to be solved efficiently. Another issue could be boundaries are not 'clear' and a high amount of flow would go through what is supposed to be our boundaries, making it hard to make a min cut at an exact

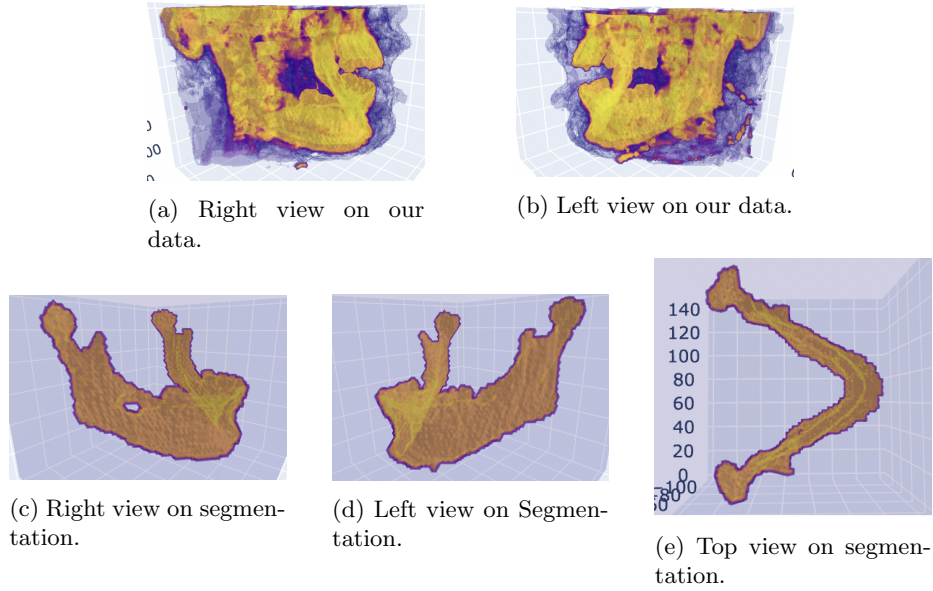


Fig. 4: Result of 3D segmentation of the mandible.

location. These issues are not present in the Random walker.

In conclusion, the Graph cut algorithm is preferred as it gives us an analytical solution, however it is not always possible to employ the Graph cut algorithm due to memory usage, cost function formulation or vague boundaries, but the Random walker algorithm solves these issues by being non-deterministic.

## 5 PCA

PCA is a procedure which captures the directions of greatest data variability. That is, having a high dimensional dataset and performing an PCA would give us the eigenvectors and eigenvalues explaining the variance in the data. We can then choose to keep the eigenvectors that explain, say, 98% of the variance in the data, and because the first few components / eigenvectors explain the vast majority of the variance, this will often lead to a big dimensionality reduction of the data.

In medical image analysis we can obtain such a high dimensionality dataset if we describe a shape by several data points / landmarks. If we were to describe a lung field, this could be done with 100 points in 2D giving us a 200 dimensional dataset for each lung field we describe. If we have several lung fields described in this manner, we can now normalize them using a Procrustes analysis and then perform PCA. In the data we are given, we have 124 lung fields described by 94 points. After performing PCA and retaining 98% of the variance we find 20 components are enough. We can now visualize how adjusting the magnitude of

the principal components affects our lungs. In Figure 5 we see how the first three principal components affects our lung shape. These are the components which describes most of our variance in the data, and thus the components which gives the biggest change in shape.

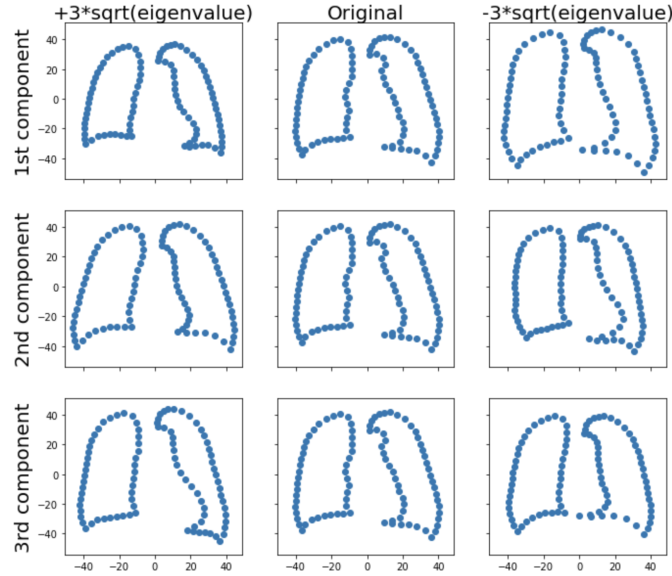


Fig. 5: How the first three principal components affects the shape of our lung.

Using the 20 components describing 98% of the variance in our lung fields we can now fit a mean lung shape, the lung found by taking the average shape of our lungs, to a never-before-seen lung and because we fit it using the components we avoid fitting to noise. We can then make an error measure of how well this active shape model / our mean lung fits our new lung, and use this to determine if we are seeing a healthy or sick lung / a normal or abnormal lung. Fitting our active shape model also gives a good boundary segmentation of the new lung. Even just using the mean lung shape would give a decent initial segmentation of a new lung field in most cases.

## 6 Conclusion

In conclusion we have seen where segmentation can be used in medical image analysis, we have seen the risks of dilation and erosion and explained the benefits of them, we have concluded Graph cut segmentations are preferred, but also have limited applications and Random walker alleviates these constraints by being non-deterministic, and finally we have looked into how PCA can be used for segmentation.

## 7 Code collaboration

The code for this assignment has been developed in collaboration with: Marcus Hansen, Ulrik Larsen and Sarah Hansen

## References

1. A.P. Dhawan, Medical Image Analysis, 2nd edition, 2011, John Wiley
2. A. Maier, et al., Medical Imaging Systems, 2018, Springer Open
3. A. Norouzi, et. al., 'Medical Image Segmentation Methods, Algorithms, and Applications', IETE Technical Review, Volume 31, Issue 3, 23 June 2014.