Introduction to Image Segmentation

Dr. Ruisheng Su, Assistant Professor

Department of Biomedical Engineering, Medical Image Analysis group

IMAG/e TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Overview & course schedule

Modules	Date	Topic
Registration	April 24 (Thursday)	Course introduction, geometrical transformations
	April 28 (Monday)	Point-based image registration
	May 1 (Thursday)	Intensity-based image registration
Segmentation	May 8 (Thursday)	Introduction and evaluation metrics for image segmentation
	May 15 (Thursday)	Segmentation in feature space
	May 19 (Monday)	Segmentation using graph-cuts
	May 22 (Thursday)	Statistical shape models
Deep learning for MIA	May 26 (Monday)	Convolutional neural networks
	June 2 (Monday)	Deep learning applications (registration)
	June 5 (Thursday)	Guest lecture by **Danny Ruijters** (principal scientist @ Philips, full professor @ TU/e)
	June 10 (Tuesday)	Deep learning applications (segmentation)
	June 12 (Thursday)	Unsupervised deep learning for medical image analysis

Overview & course schedule

Modules	Date	Topic
Registration	April 24 (Thursday)	Course introduction, geometrical transformations
	April 28 (Monday)	Point-based image registration
	May 1 (Thursday)	Intensity-based image registration
Segmentation	May 8 (Thursday)	Introduction and evaluation metrics for image segmentation
	May 15 (Thursday)	Segmentation in feature space
	May 19 (Monday)	Segmentation using graph-cuts
	May 22 (Thursday)	Statistical shape models
Deep learning for MIA	May 26 (Monday)	Convolutional neural networks
	June 2 (Monday)	Deep learning applications (registration)
	June 5 (Thursday)	Guest lecture by **Danny Ruijters** (principal scientist @ Philips, full professor @ TU/e)
	June 10 (Tuesday)	Deep learning applications (segmentation)
	June 12 (Thursday)	Unsupervised deep learning for medical image analysis

Outline

- Recap of previous lecture
- Introduction to medical image segmentation
- Recap of previously learned segmentation technique
- Evaluation of segmentation accuracy

Outline

- Recap of previous lecture
- Introduction to medical image segmentation
- Recap of previously learned segmentation technique
- Evaluation of segmentation accuracy

Recap of learning objectives (previous lecture)

The student can:

- **explain** and **implement** three important intensity-based image **similarity metrics**, namely sum of squared differences, normalized cross correlation, and mutual information.
- **select** the correct image similarity metric for an image registration task based on the assumptions of these metrics.

- explain how the **joint probability mass function (p.m.f.)** can be used to measure the similarity between two images if we consider the image intensities as random variables.
- **interpret joint histograms** to judge whether two images are well aligned.

- describe the numerical **procedure** to register two images by maximizing a similarity function (gradient ascent/descent).
- explain the effect of the **learning rate** and the **initialization** of the parameters on the optimalization process when using gradient ascent/descent.

Recap

Quiz: You want to register two **mono-modal** images that differ only by a small translation and scaling.

Which **intensity-based** similarity metric is most appropriate?

A. Mutual Information (MI)
B. Sum of Squared Differences (SSD)
C. Normalized Cross-Correlation (NCC)
D. Fiducial Registration Error (FRE)

Recap

Quiz: When registering **multi-modal** images (e.g. CT $\leftrightarrow$ MRI), which metric best handles **arbitrary** intensity relationships?

A. Sum of Squared Differences (SSD)

B. Normalized Cross-Correlation (NCC)

C. Mutual Information (MI)

D. Mean Absolute Difference (MAD)

Recap

Quiz: A joint histogram of two well-aligned images typically shows:

A. A uniform cloud of points across all bins
B. A tight diagonal ridge
C. Two clusters symmetric about the diagonal
D. Completely empty except one bin

Recap

Quiz: In a gradient-based optimization of the registration parameters, if you choose a **learning rate** that is too large, you are most likely to:

A. Converge faster to the global optimum
B. Get stuck in a shallow local optimum
C. Overshoot the optimum and diverge
D. Make the similarity metric independent of initialization

Recap

Quiz: Which initialization strategy is **least** likely to help gradient-based registration find the correct alignment?

A. Starting from an approximate transform obtained by landmark matching
B. Using the identity transform (no displacement) when images are already close
C. Sampling hundreds of random initial guesses and choosing the best
D. Initializing with a large, arbitrary rotation/translation

Outline

- Recap of previous lecture
- Introduction to medical image segmentation
- Recap of previously learned segmentation technique
- Evaluation of segmentation accuracy

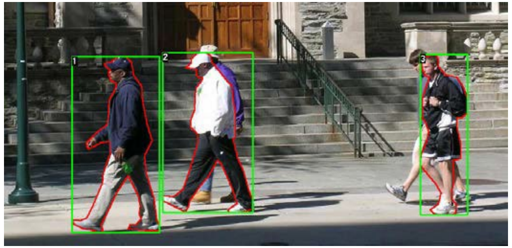
Intended learning outcomes

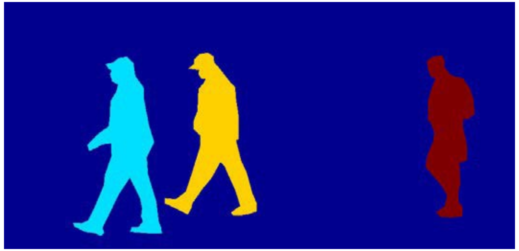
The student can:

- Describe what problems arise with using a **threshold-based** segmentation method
- Calculate how many parameters are needed for **K Gaussians** in p dimensions
- Recognize **covariance** matrices and identify their properties

- Select a suitable evaluation **metric** for a given medical image analysis task
- calculate the **accuracy**, **Dice score** and **Hausdorff distance**, given an image segmentation
- interpret the results and assess the quality of the validation methods used in medical image analysis research papers

Segmentation examples





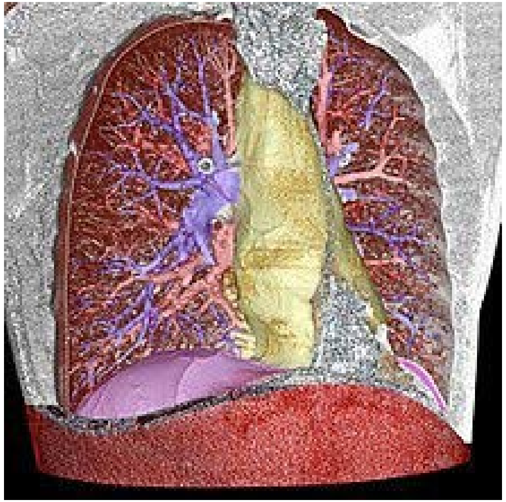
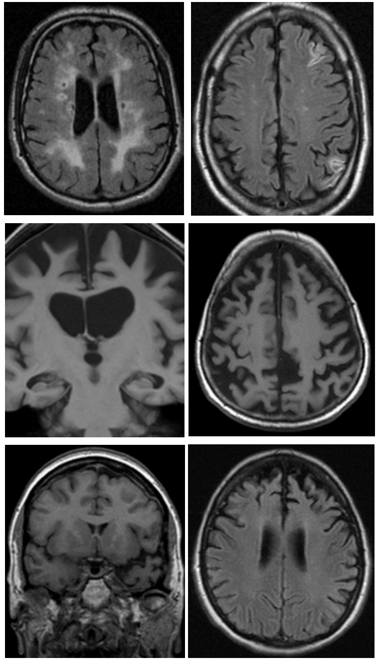


Image source: 1). Pedestrians: https://www.cis.upenn.edu/~jshi/ped_html/, 2). Lung image: https://en.wikipedia.org/wiki/Image_segmentation

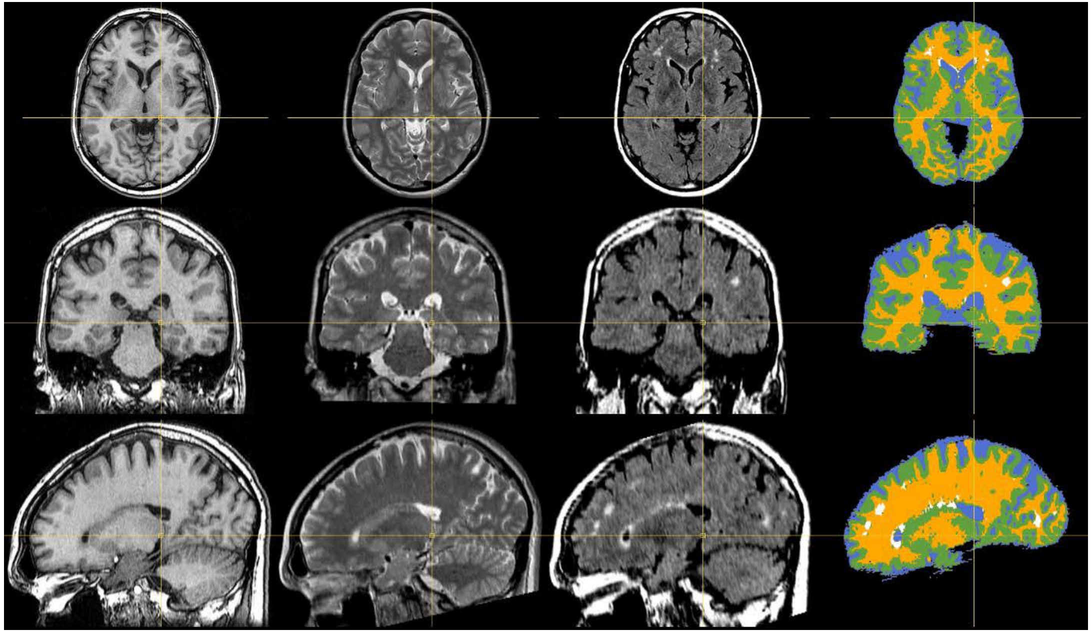
Example: dementia diagnosis

- Symptoms (such as memory loss) may be difficult to detect
- Symptoms similar between different types of dementia (frontotemporal, Alzheimer..)
- MR scans show differences in atrophy between types
- **How much tissue there is in different parts of the brain?**



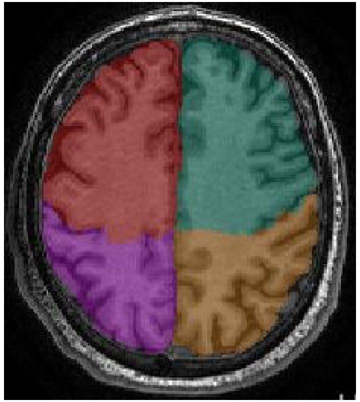
Step 1: tissue segmentation

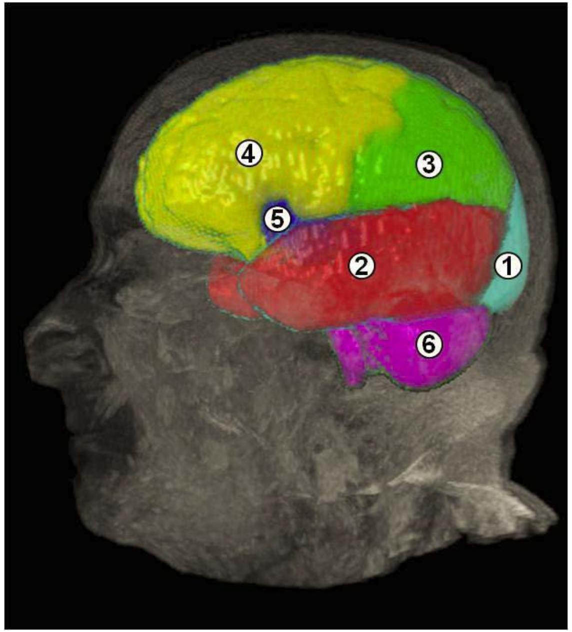
Each pixel is classified as white matter, gray matter, cerebrospinal fluid (CSF) or background



Step 2: segmentation of anatomical structures

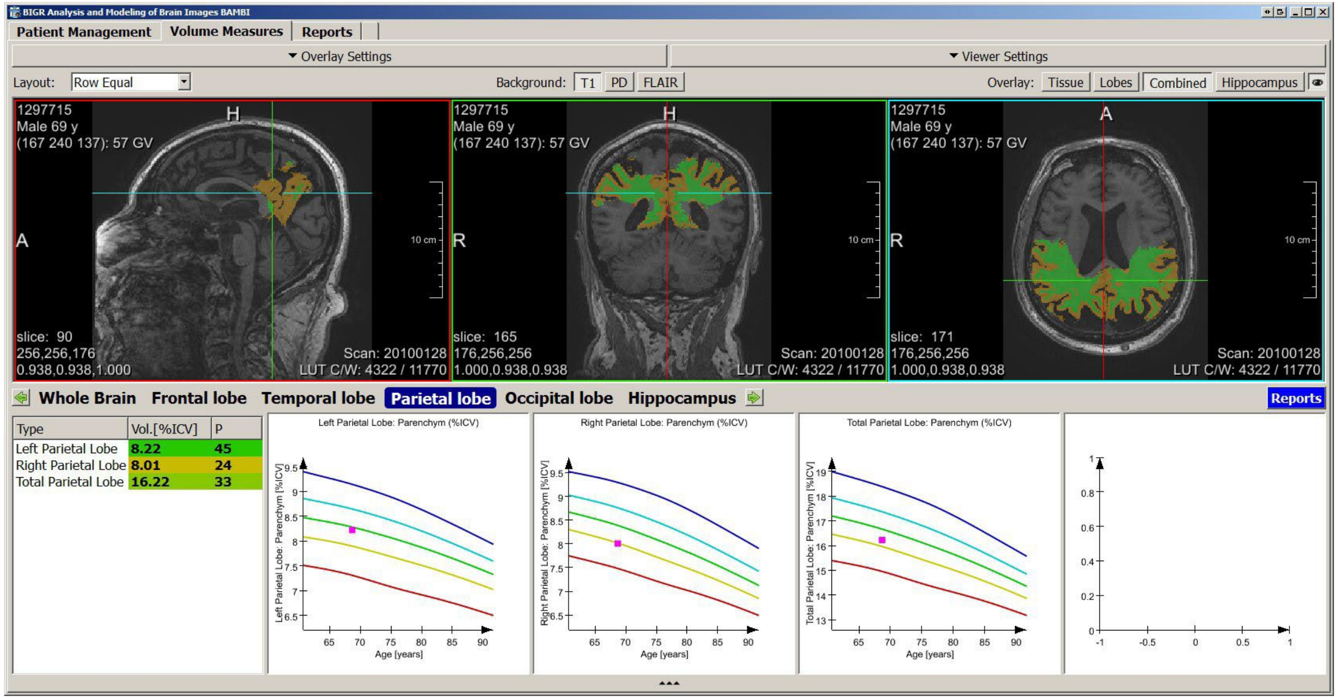
Each pixel is classified as being in one of the lobes (or smaller structures like hippocampus)



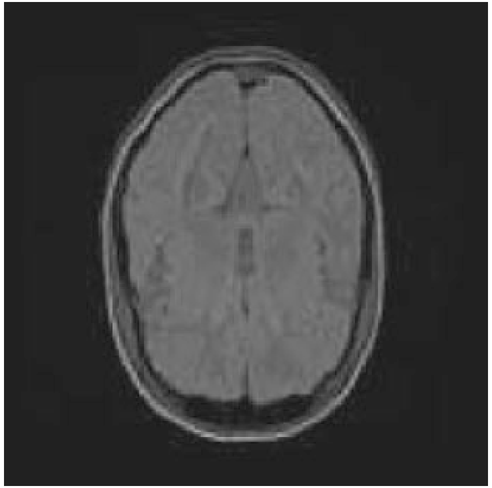


Step 3: Analysis

Use step 1 & step 2 to compare measurements to reference values



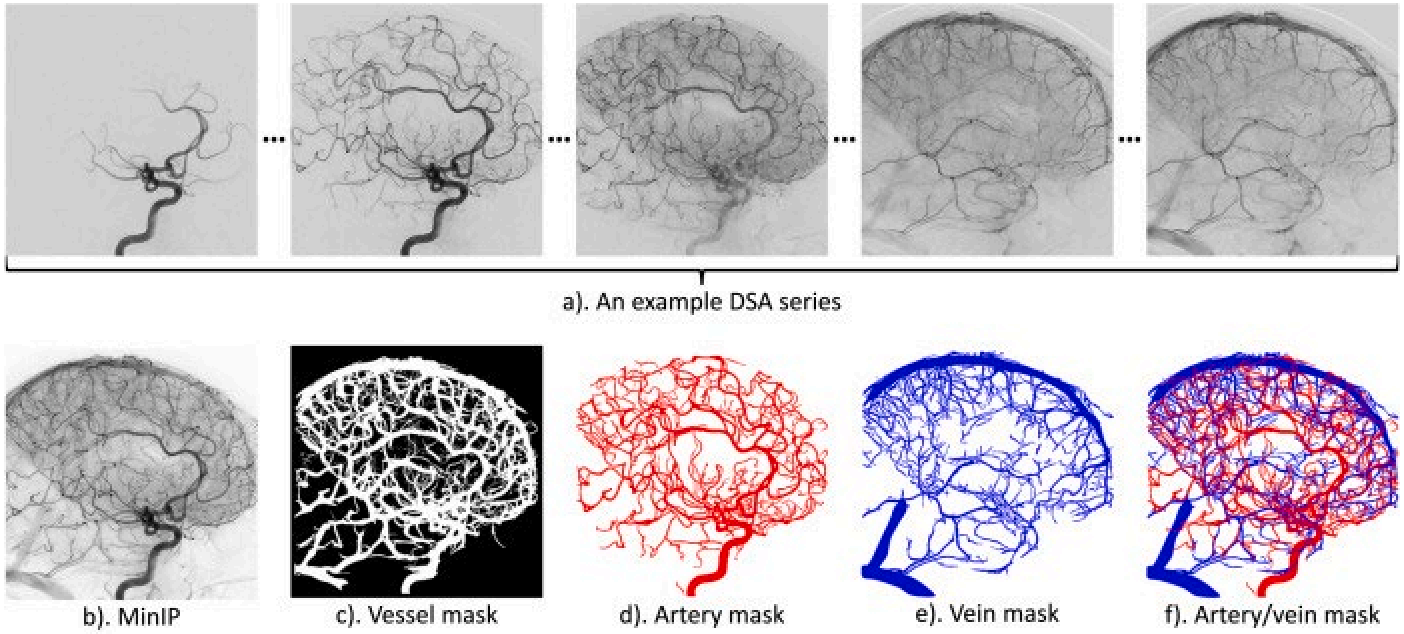
Example: lesion segmentation



Read Toennies chapter 1.3 for background about this problem

Images: Toennies, Guide to Medical Image Analysis (2012).

Example: brain vessel segmentation



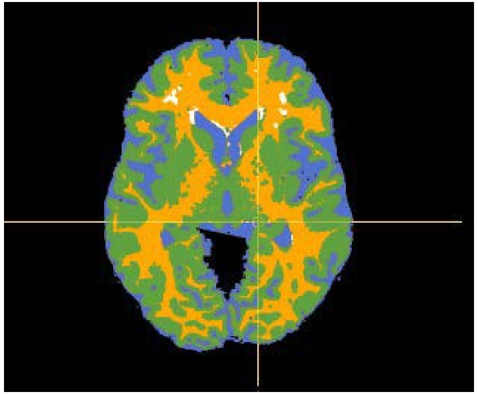
Source: Su, R., van der Sluijs, P. M., Chen, Y., Cornelissen, S., van den Broek, R., ... & van Walsum, T. (2024). CAVE: Cerebral artery–vein segmentation in digital subtraction angiography. Computerized Medical Imaging and Graphics, 115, 102392.

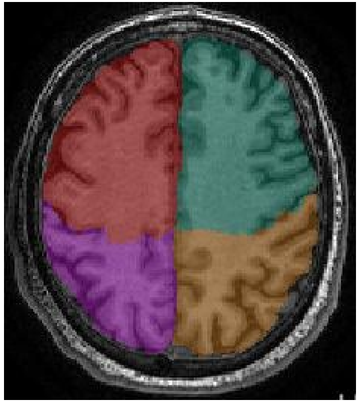
Purposes of segmentation

- Avoid manual delineation
- Insightful representation of image data
- Quantification: derive information required for diagnosis, therapy, monitoring, decision support, image-based interventions.
- Standardized way of working: objective and reproducible

Medical image segmentation

Different tasks have different characteristics,
which influence which model & optimization is more suitable.





Images: Biomedical Imaging Group Rotterdam & Toennies 2012

Medical image segmentation - overview

- Threshold-based segmentation
- Edge-based segmentation
- Region-based segmentation

8BB050

- Segmentation in feature space (clustering & classifiers)
- Graph-based segmentation
- Shape- and deformable-model based segmentation
- Learning-based segmentation

8BE030 (this course)

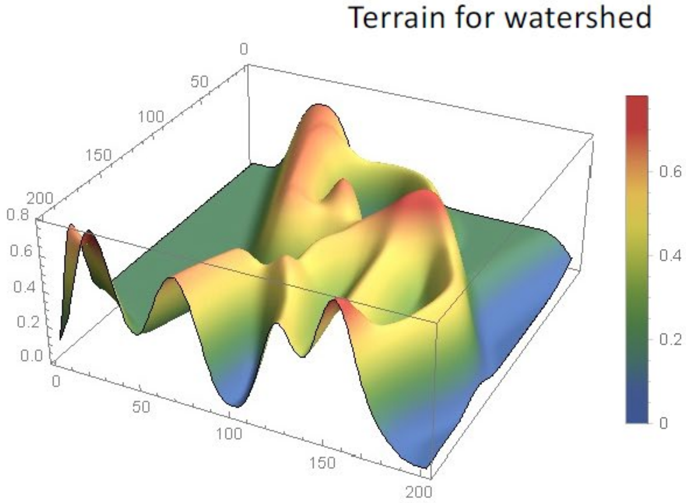
- Atlas-based segmentation

MSc

Outline

- Recap of previous lecture
- Introduction to medical image segmentation
- Recap of previously learned segmentation technique
- Evaluation of segmentation accuracy
- Segmentation in the feature space (Clustering techniques)

Recap – edge-based watershed segmentation



What is a common problem with this method?

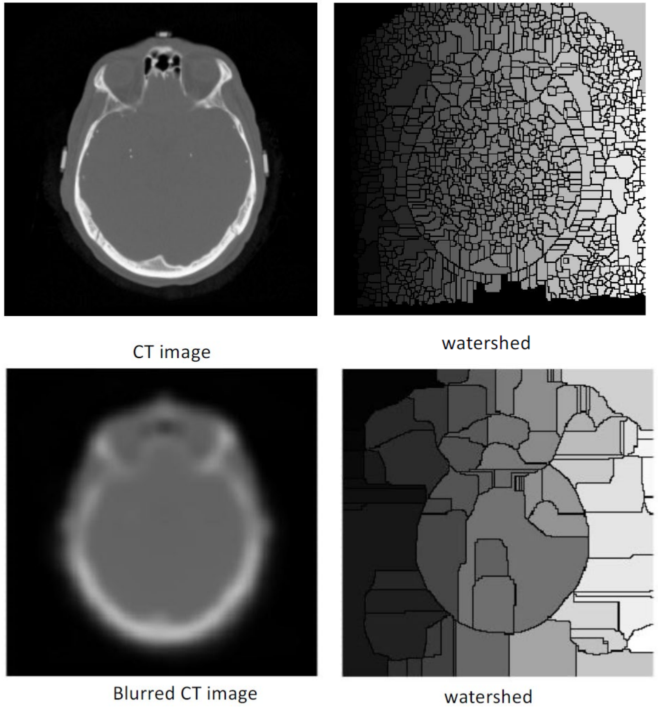
And how can we address it?

Recap – edge-based watershed segmentation

Over segmentation

Possible solutions:

- Gaussian blur

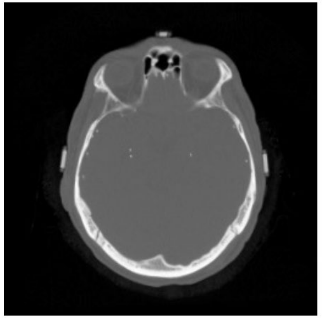


Recap – edge-based watershed segmentation

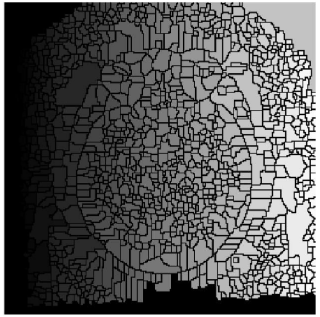
Over segmentation

Possible solutions:

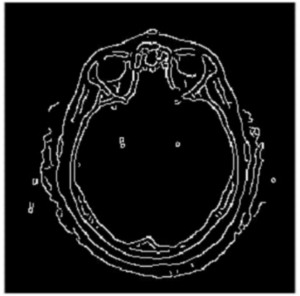
- Gaussian blur
- Remove weak edges



CT image



watershed

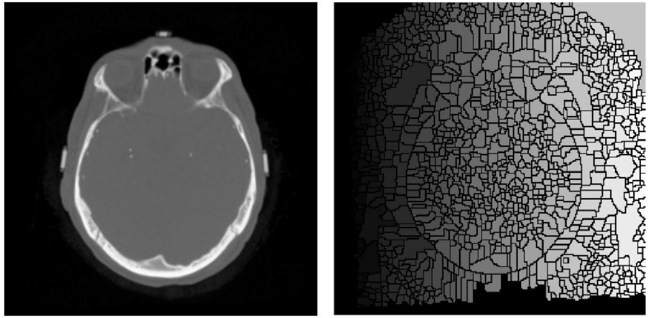


Edges CT image

Weak edges removed

Recap – edge-based watershed segmentation

Over segmentation

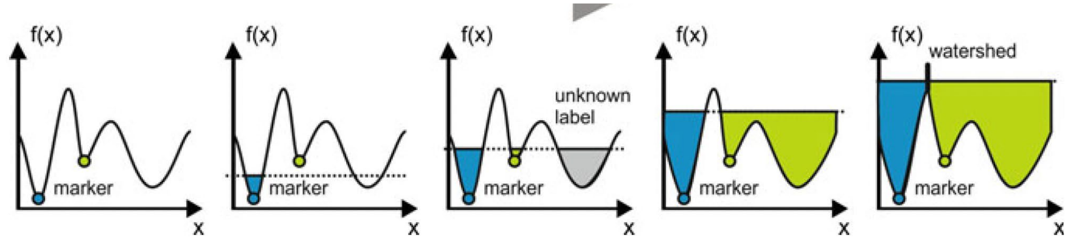


CT image

watershed

Possible solutions:

- Gaussian blur
- Remove weak edges
- Supply markers as prior knowledge

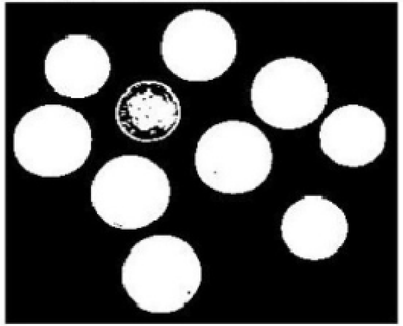


Recap – thresholding-based segmentation

- Histogram extrema (valley finding)
- Otsu's thresholding (variance-minimization)
- Gaussian mixture intersection

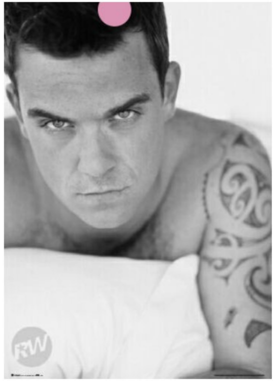
Original image.

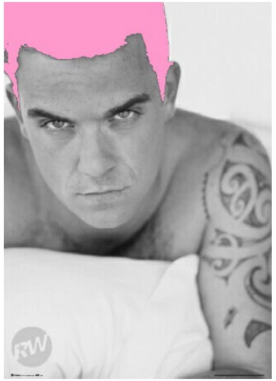
Image segmentation with Otsu thresholding.

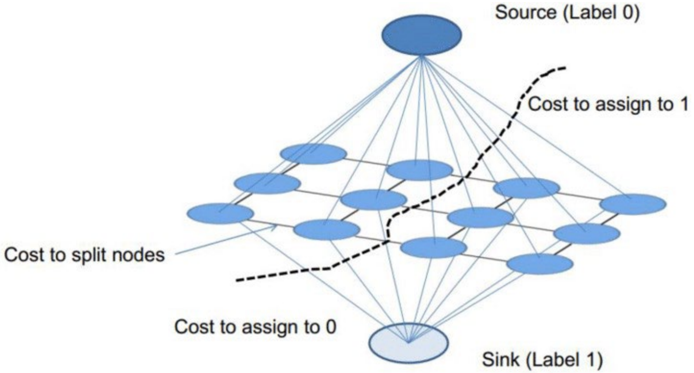


Recap –region-based segmentation

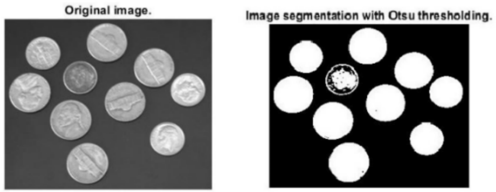
- **Assumption**: little intensity variation within the object region
- Merging, e.g., region-growing
- Splitting, e.g., graph-cut



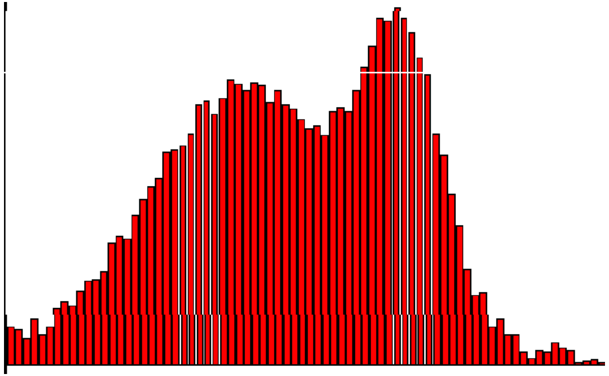


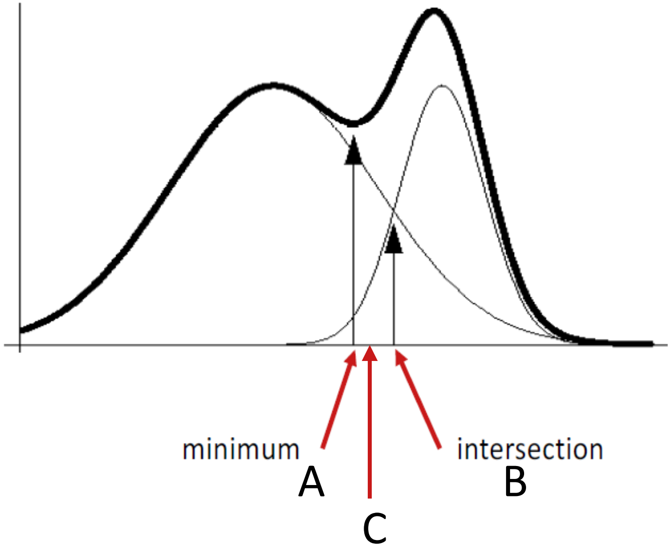


Recap – thresholding-based segmentation

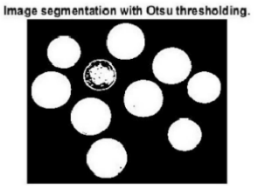


1. Histogram extrema (valley finding)

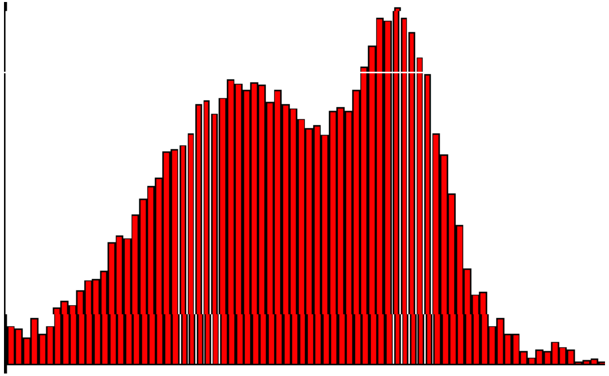


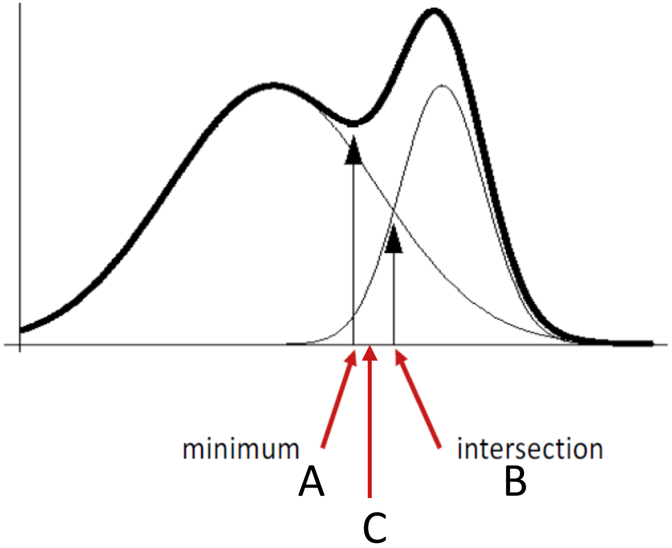


Recap – thresholding-based segmentation



2. Otsu's thresholding (variance minimization)

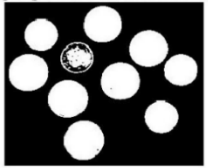




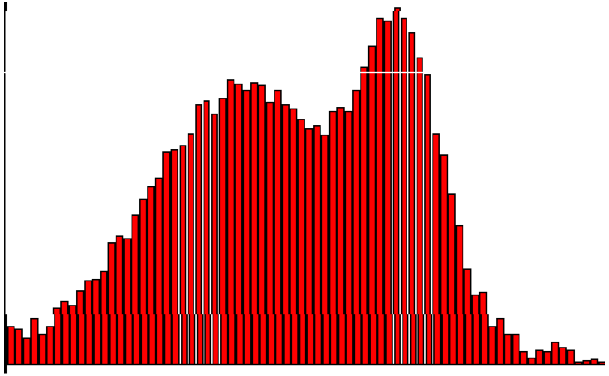
Recap – thresholding-based segmentation

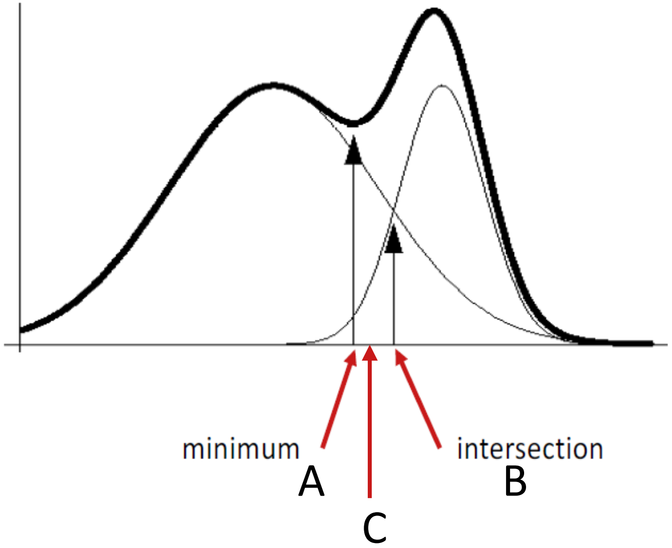
Original image.

Image segmentation with Otsu thresholding.



3. Gaussian mixture intersection





Recap - Gaussian mixture intersection

Model histogram by a weighted sum of two Gaussians, g_0 and g_1.

Probability of pixel value v:

$$\begin{aligned} h(v) &\approx p_0 G_0(v) + p_1 G_1(v) \\ &= (1 - p_1) G_0(v) + p_1 G_1(v) \end{aligned}$$

P_i indicates the probability that a pixel belongs to segment *i*.

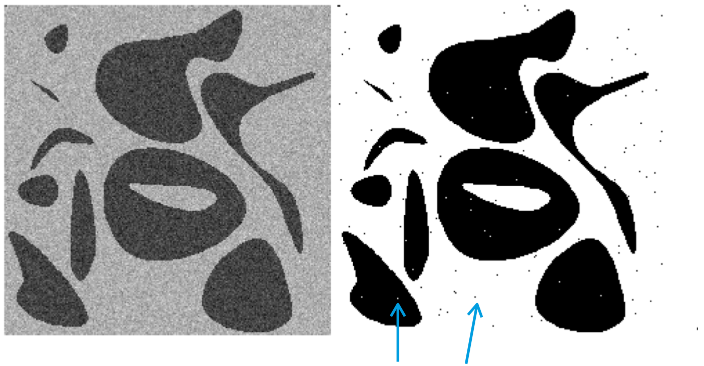
Method has **five** parameters to find (two means, two variances, P_1)

Q: How many parameters for tissue segmentation (4 classes total)?

Slide from 8BB50

Recap – thresholding-based segmentation

- Our vision uses context to assign a pixel to a segment
- Thresholding is insensitive to context → noisy segmentation



Small segments due to noise

Image context

- The smallest scale in digital images are individual pixels
- By thresholding with pixel intensity, we do not consider the larger context of each pixel
- How can our algorithms “see” context at different scales?

Overview - segmentation in feature space

- Adding context
  - Convolution with Gaussians
  - Covariance between intensity and other features
  - Optimal thresholding with 2 or more features

- Segmentation by clustering and classifiers (next time)

Image context

- We are good in observing a scene at multiple scales
- When watching an object from further away the small details disappear and larger patterns become more visible

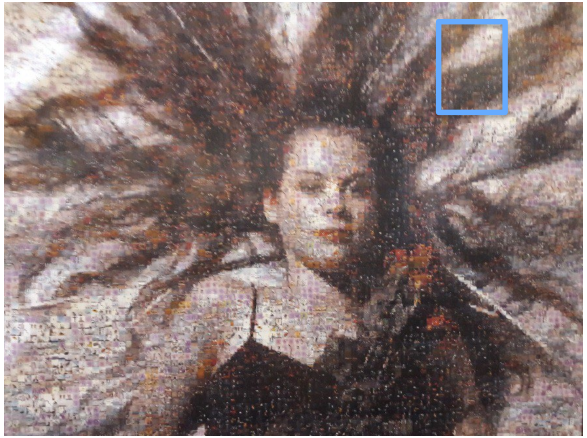


Image context - convolution with a Gaussian

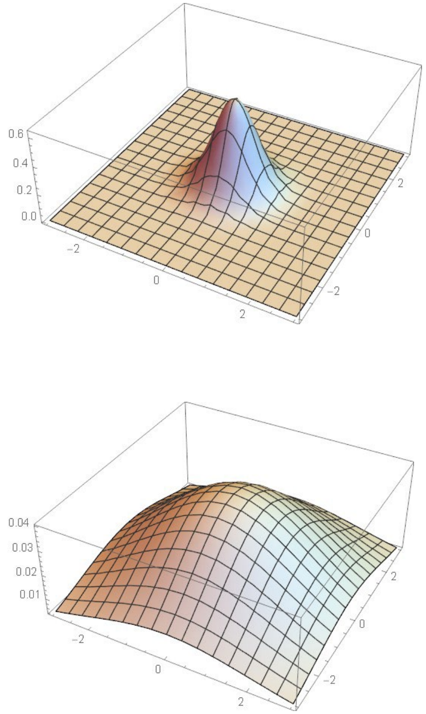
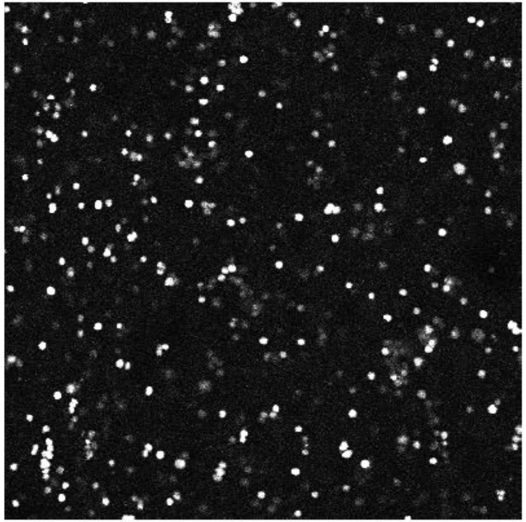
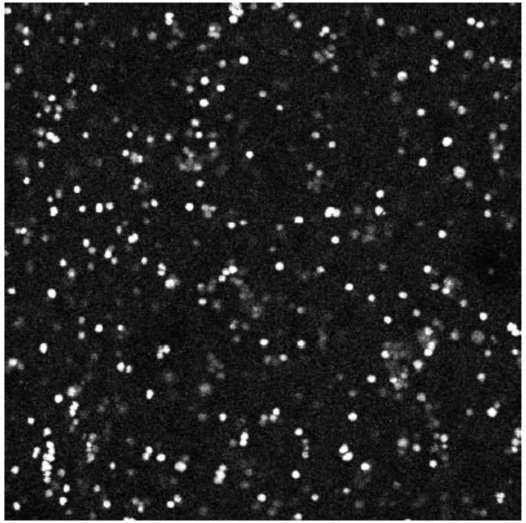


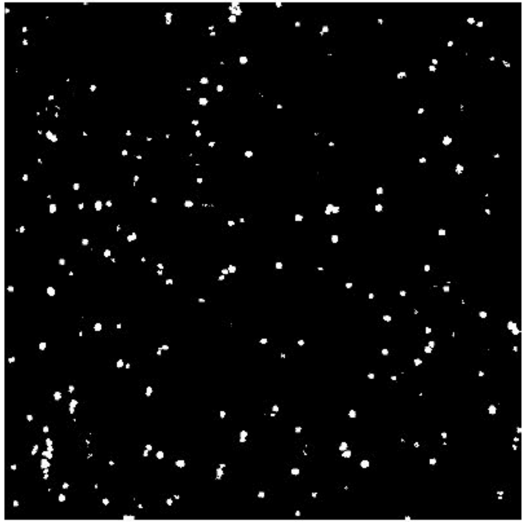
Image context - convolution with a Gaussian



Cell image



Filtered with Gaussian filter



Thresholded

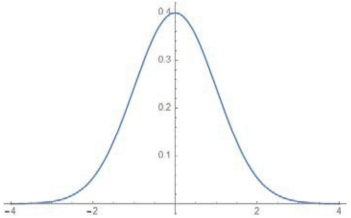
Living cell count after filtering: 107 (*vs.* 113)

Slide from 8BB50

- Convolution with a Gaussian kernel/filter (blurring)

  –Gaussian:

$$g(x,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$



  – Convolution of image f(x) with Gaussian:

$$(g_\sigma * f)(x) = \int g_\sigma(u) f(x-u)\, du$$

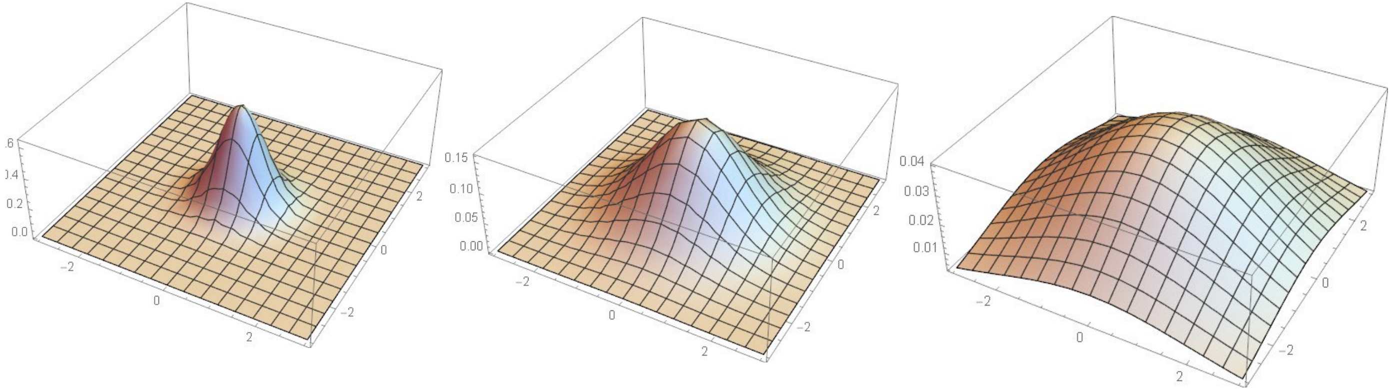
  – Discrete in practice: pixel value becomes a weighted average of pixels around it (nearby pixels get higher weight)

- We denote the Gaussian scale space of an image $f(x,y)$ as

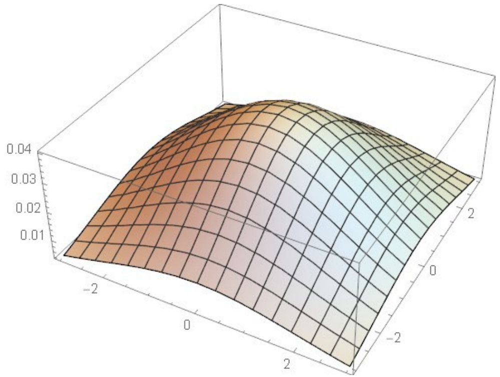
$$f_\sigma(x,y) = g_\sigma(x,y) * f(x,y)$$

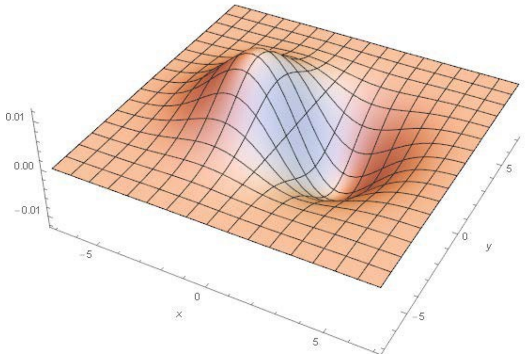
- An image at scale σ is an image convolved with a Gaussian of scale σ

$\sigma = 0.5$ $\sigma = 1$ $\sigma = 2$

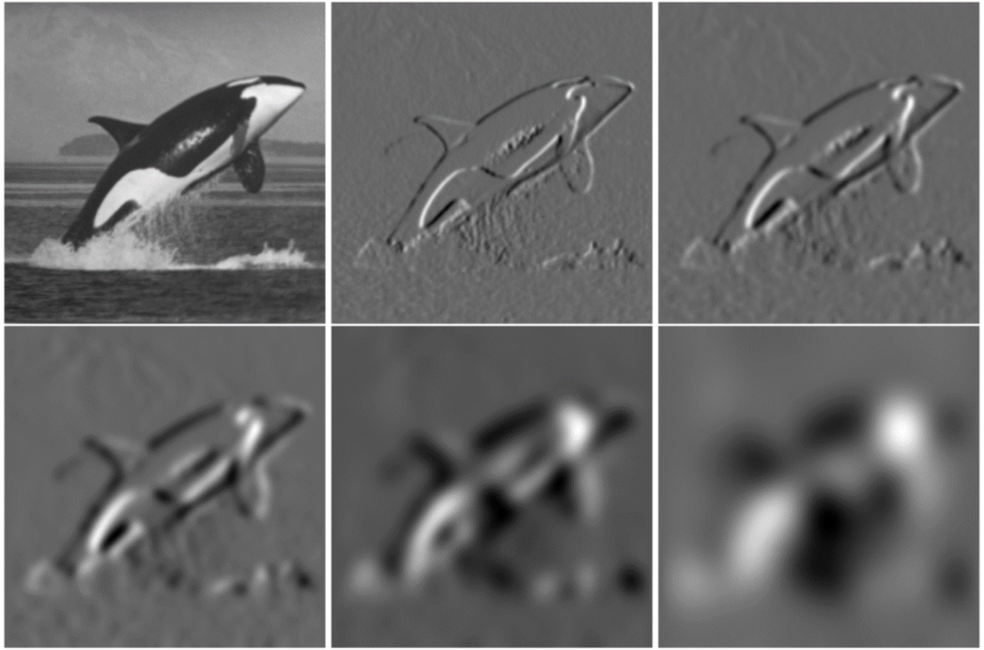


- Not only Gaussians, but also derivatives



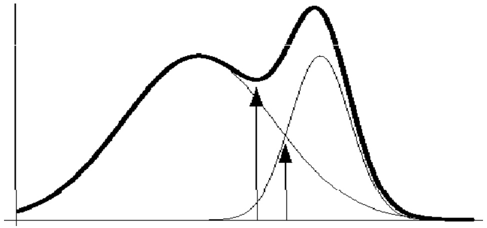


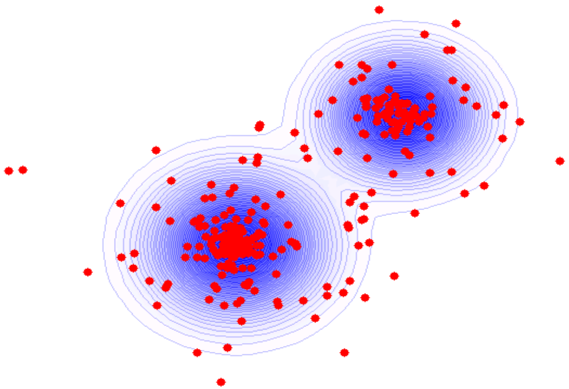
- Gaussian x-derivatives of an image with increasing scales



Adding context to optimal thresholding

- E.g. add “average intensity of neighborhood” (intensity after blur)
- 2D feature space
- Idea: extend the previous method to 2D, i.e. approximate by sum of 2D Gaussians
- Q: How many parameters now?



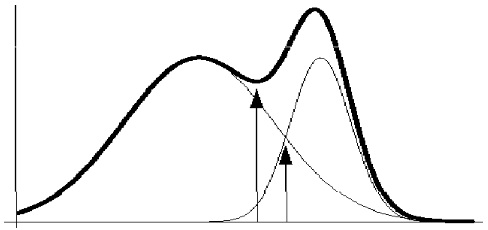


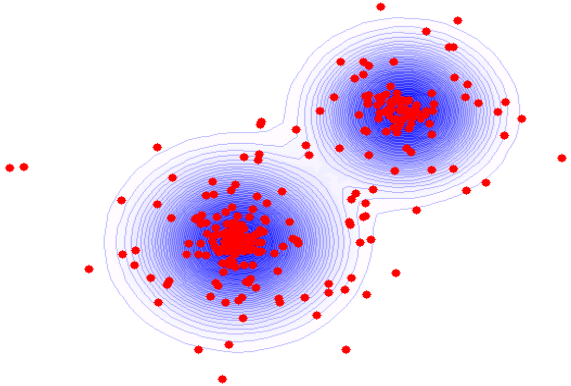
- Model = Mixture of K Gaussians

- We need to find:
  - (K-1) class probabilities *p*
  - K means
  - K **covariance matrices**

- It might seem that for 2 Gaussians in 2D, we still need 5 parameters:
  - 1 class probability
  - 2 means
  - 2 covariance matrices

  But it's a bit more complex – for this we look at the covariance matrix

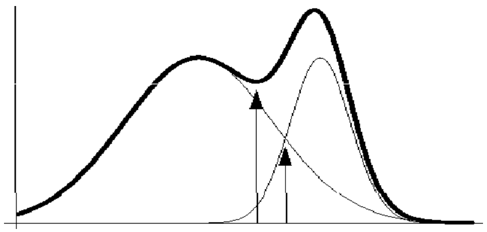




Modeling 1D data

- With 1 feature (such as intensity), how the data “behaves” can be characterized with the **variance**

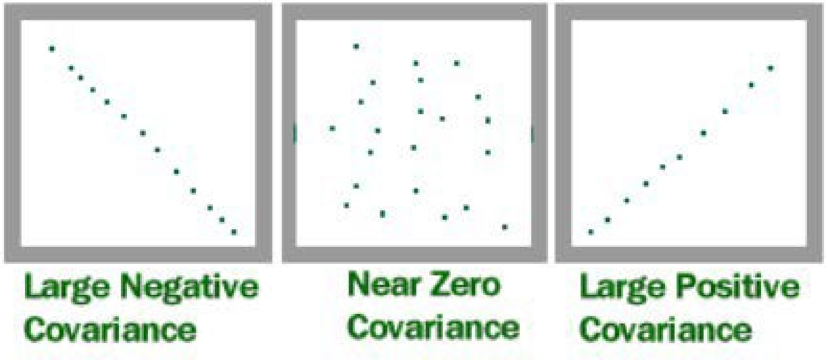
- Variance = “on average, how different is my data compared to the mean”



$$\operatorname{Var}(X) = \operatorname{E}\left[(X-\mu)^2\right].$$

Modelling 2D data

- With more features, each feature still has a variance
- But the variances are not enough to model 2D data – we also need covariance



https://www.statisticshowto.datasciencecentral.com/covariance/

Covariance

- Covariance generalizes the variance
- “On average, how much do the these two variables differ together”

- Covariance is an unnormalized version of correlation

$$\operatorname{Var}(X) = \operatorname{E}\left[(X - \mu)^2\right].$$

$$\Sigma_{ij} = \operatorname{cov}(X_i, X_j) = \operatorname{E}[(X_i - \mu_i)(X_j - \mu_j)] = \operatorname{E}[X_i X_j] - \mu_i \mu_j,$$

Covariance

- You can summarize variance & covariance in a covariance matrix
- $corr\ (X) = \begin{bmatrix} 3 & 3/2 \\ 3/2 & 2 \end{bmatrix}$
- The **variances** are 3 and 2
- The **covariance** is 1.5
- The Gaussian on the right has this covariance matrix

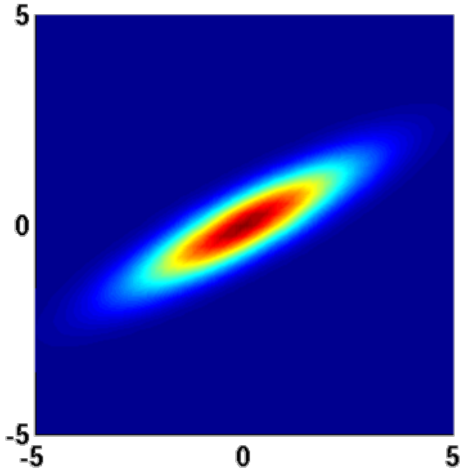
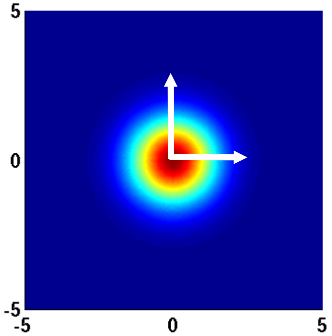


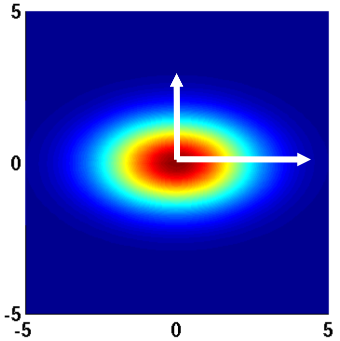
Image by David Tax

IMAG/e TU/e

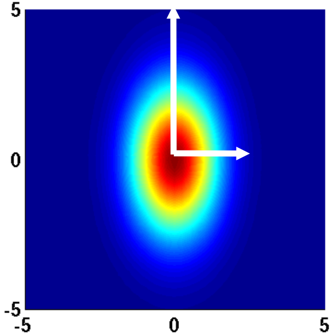
Covariance examples



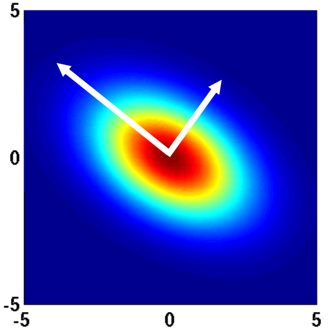
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



$$\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$



$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$



$$\begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}$$

Image by David Tax

Covariance matrix properties

There are a few properties of a covariance matrix:

- Variance can only be positive
- Covariance can be positive, zero or negative

- Matrix is symmetric

- Matrix is "positive semidefinite",
  - for any vector X, $X^T \Sigma X \geq 0$
  - It enforces nonnegativity on every possible direction of the multivariate data, which is consistence with the variance > 0

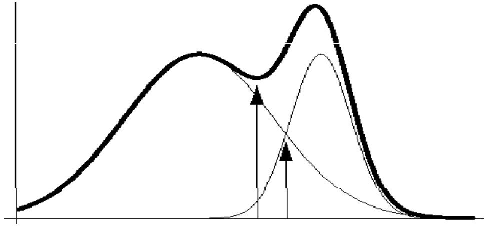
Covariance

Further reading on covariance matrices:

- Blog post by VisionDummy
- Chapter on eigenvalues and eigenvectors → Essence of Linear Algebra by 3Blue1Brown (Youtube)

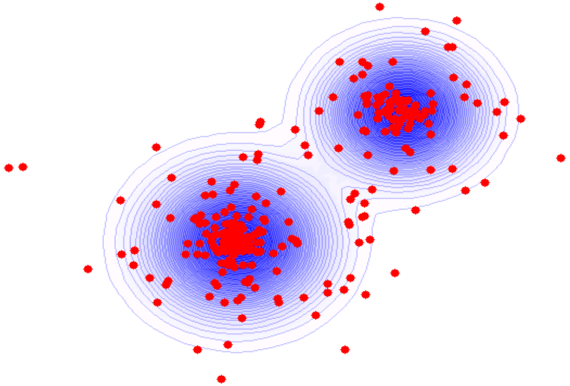
...Back to **Optimal thresholding – Mixture of Gaussians**

- We need to find:
  - (K-1) class probabilities p
  - K means μ
  - K **covariance matrices $\boldsymbol{\sigma}$**



- How many parameters for 1 Gaussian in 2D?

- How many parameters for 2 Gaussians in 2D?



Optimal thresholding – Mixture of Gaussians

Per Gaussian need to estimate:

- The mean is already 2 parameters:
  - mean of feature 1
  - mean of feature 2

- The covariance matrix is 3 independent parameters:
  - variance of feature 1
  - variance of feature 2
  - **covariance** of features 1 & 2.

Optimal thresholding – Mixture of Gaussians

Parameter type	Per Gaussian	Total for 2 Gaussians
Means	2 (one for x, one for y)	2×2=4
Covariances	3 unique entries in a symmetric 2×2 matrix (σ_x^2, σ_y^2, covariance σ_{xy})	2×3=6
Mixing weights	1 (the two weights must sum to 1)	1
Grand total		**4 + 6 + 1 = 11 parameters**

Optimal thresholding – Mixture of Gaussians

For a p-dimensional Gaussian :

- μ: vector with p elements
- Σ: matrix with 0.5 p (p+1) elements

Parameter type	Count
Means (2 components)	2p
Covariances (2 comp.)	$2 \times p(p+1)/2 = p(p+1)$
Mixing weight	1
Total	**2p+p(p+1)+1**

- This means the number of parameters increases **quadratically** with p
- A lot of data needed to estimate parameters well

Outline

- Introduction to medical image segmentation
- Recap of previously learned segmentation technique
- Evaluation of segmentation accuracy

Evaluation of segmentation

- Which approach do we choose?
- Compare to ground truth: score = evaluate_segmentation(segmentation, ground_truth)

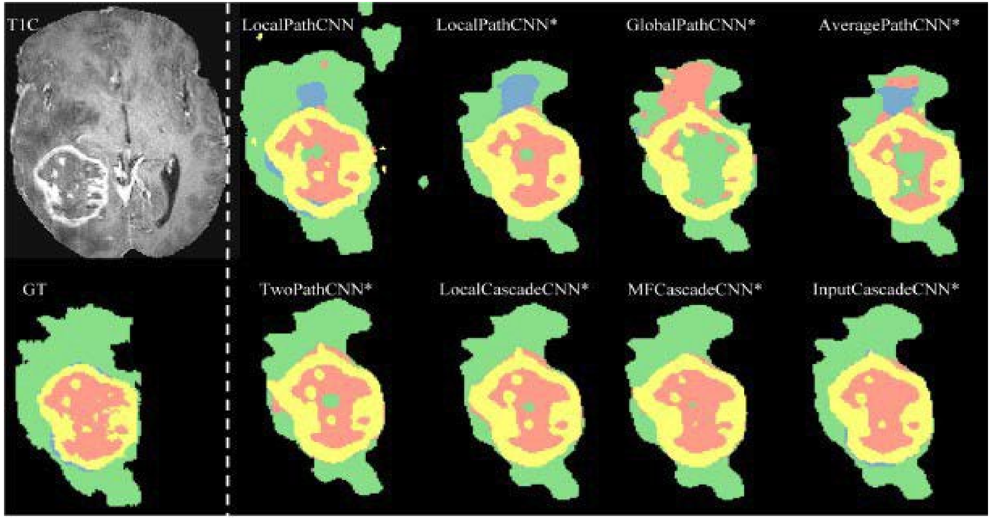


Image https://www.sciencedirect.com/science/article/pii/S1361841516300330

Evaluation metrics – medical image analysis

Important characteristics to consider when evaluating medical image analysis methods:

- **Accuracy** = deviation of results from known ground truth.
- **Precision**, **reproducibility, reliability** = extent to which equal or similar input produces equal or similar results.
- **Robustness** characterizes the change of analysis quality if conditions deviate from assumptions made for analysis (e.g., when noise level increases or if object appearance deviates from prior assumptions).
- **Efficiency** = effort necessary to achieve an analysis result.

Accurate vs Precise

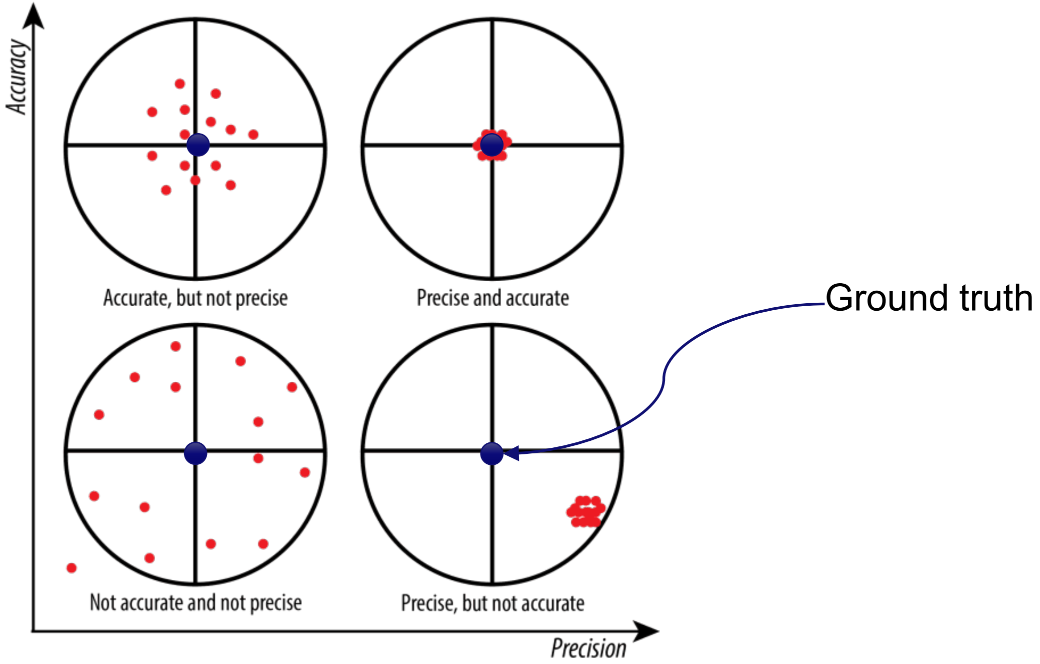


Image source: https://wp.stolaf.edu/it/gis-precision-accuracy/

Evaluation of segmentation

Many metrics available, we look at

- Accuracy
- Dice score
- Hausdorff distance

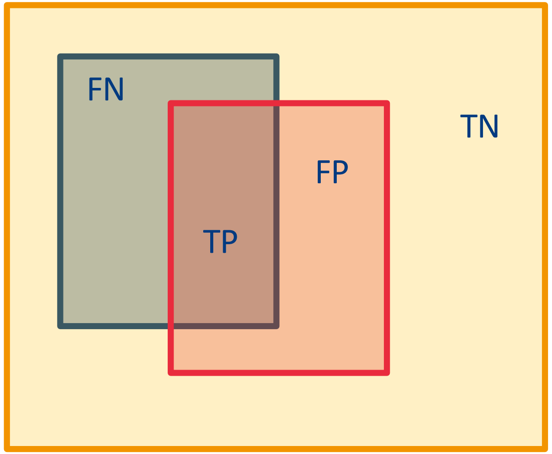
Evaluation metrics - Accuracy

"How many pixels are correct?"

Accuracy = (TP + TN) / (TP+FP+FN+TN)

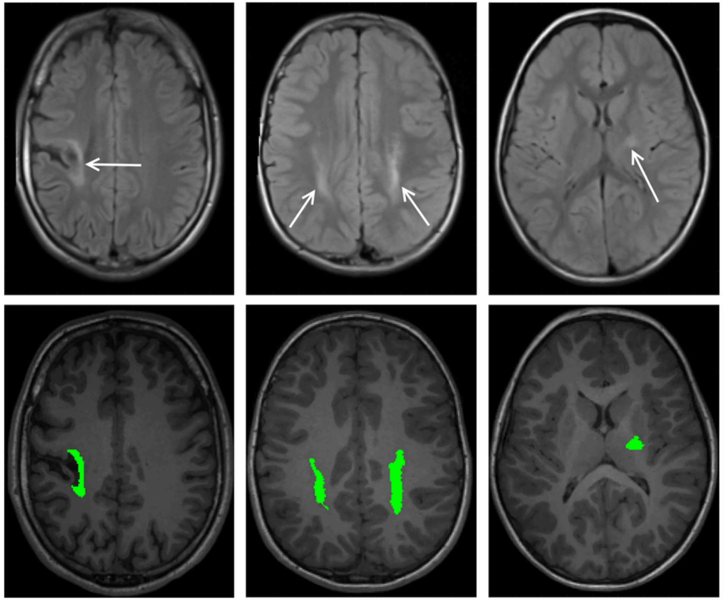
- TP = True Positive
- FP = False Positive
- FN = False Negative
- TN = True Negative

Orange = whole image
Blue = ground truth
Red= segmentation result



Evaluation metrics - Accuracy

What if the ground truth is small?


(a) (b) (c)

IMAG/e TU/e

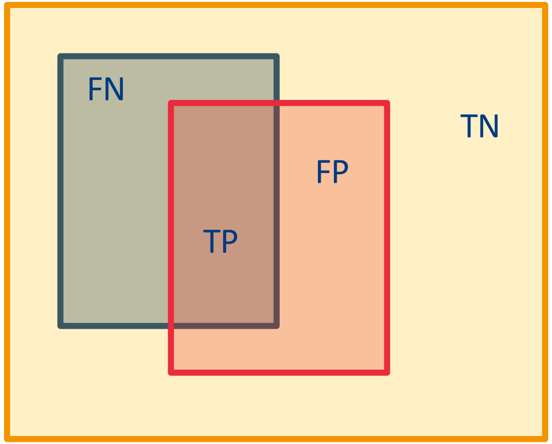
Evaluation metrics – Dice score

Measure overlap excluding TN

Sørensen–Dice index a.k.a. Dice coefficient

DSC = 2TP / (2TP + FP + FN)

Orange = whole image
Blue = ground truth
Red= segmentation result



Evaluation metrics – Dice score

Two equivalent definitions

DSC = **2**TP / (**2**TP + FP + FN)

DSC = 2 | A * B | / (| A | + | B |) for binary images A and B

| A | = size of blue ground truth

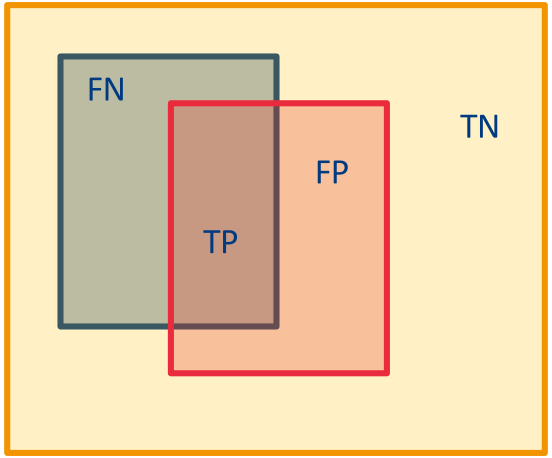
| B | = size of red result

| A * B| = size of overlap

Orange = whole image

Blue = ground truth

Red= segmentation result

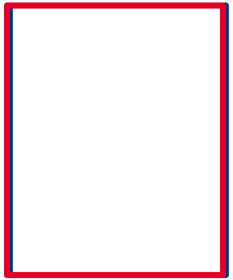


IMAG/e TU/e

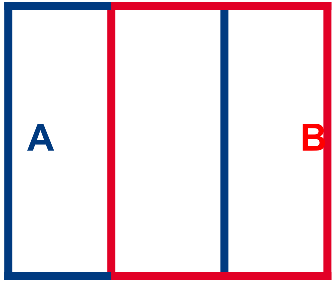
Evaluation metrics – Dice score

Dice score = 1 → full overlap

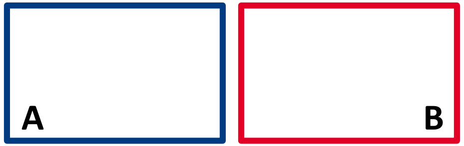
Dice score = 0 → no overlap



DSC = 1



DSC = 0.5



DSC = 0

Evaluation metrics – Dice score

Dice can be evaluated on binary vectors (0 = background, 1 = foreground)

e.g.

true_labels = [0 0 1 1]';

predicted_labels = [0 1 1 0]';

Q: What is the Dice score?

True ↓, Predicted →	1	0
1	TP	FN
0	FP	TN

True ↓, Predicted →	1	0
1	1	1
0	1	1

Evaluation metrics – Dice score

- For more than 2 classes,
  → consider each class as foreground separately

- Can average all individual scores

- Dice(A,B,C) =
  1/3 * (Dice(A vs B+C)
  + Dice (B vs A+C)
  + Dice (C vs A+B))

Dice(A vs B+C) = (2AA / (2AA + (AB + AC) + (BA + CA))

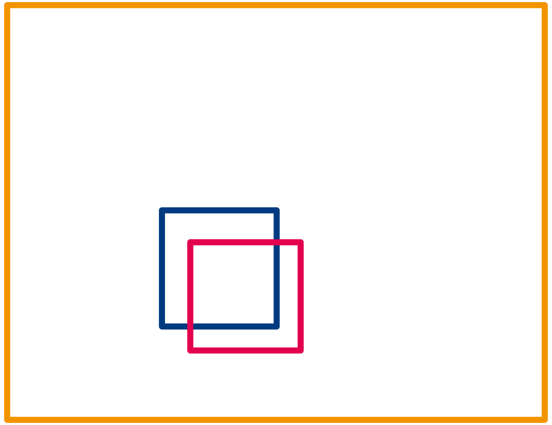
True ↓, Predicted →	A	B	C
A	AA	AB	AC
B	BA	BB	BC
C	CA	CB	CC

Accuracy vs Dice

- Both are **similarity measures** defined on TP, TN, FP, FN

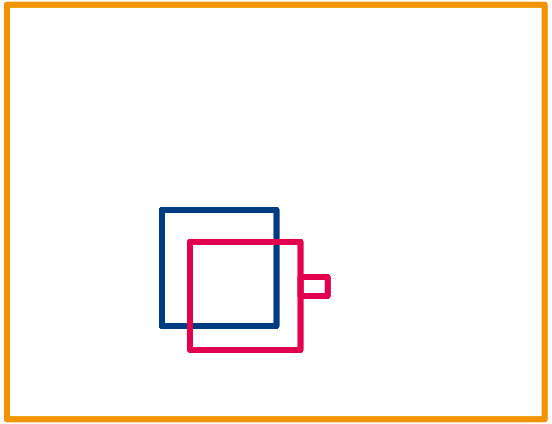
- Accuracy is not as suitable for segmentation due to TN
- Accuracy is also used for classification/diagnosis

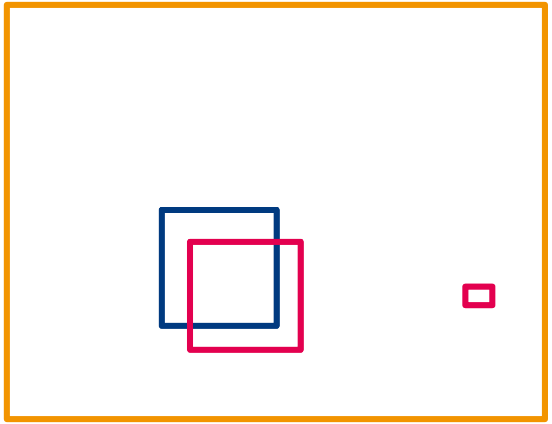
- Dice is specific for segmentation



Evaluation metrics – Dice score

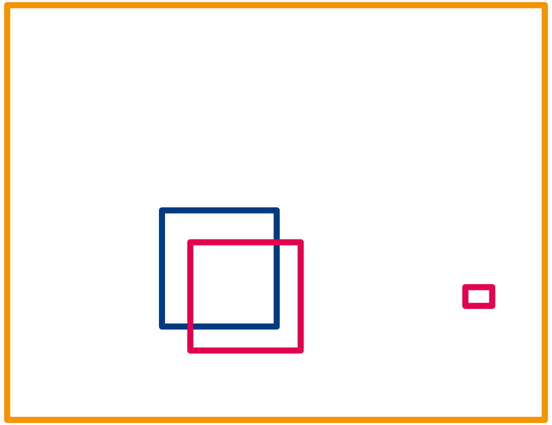
Dice (and other metrics based on TP, FP etc) are not sensitive to location





Evaluation metrics – Hausdorff distance

Hausdorff distance: compare sets of points on the boundaries

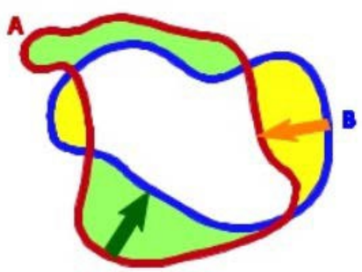


Evaluation metrics – Hausdorff distance

Hausdorff distance = maximum shortest distance between the boundary points

$$h\,(A,B) = \max_{a\in A}\min_{b\in B} d\,(a,b)$$

$$\mathrm{H}\,(A,B) = \max(h(A,B), h\,(B,A))$$

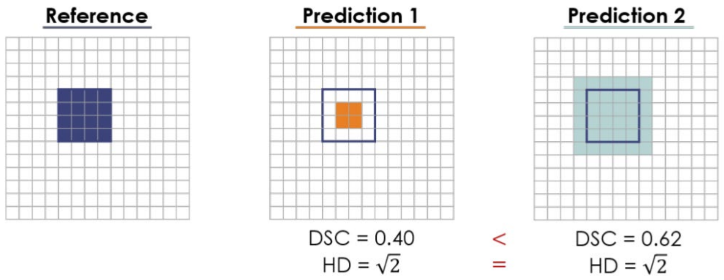


https://www.slideshare.net/UlaBac/lec14-evaluation-framework-for-medical-image-segmentation

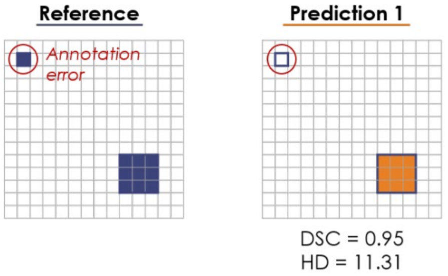
Evaluation metrics – Hausdorff distance

Hausdorff distance (HD) takes location into account and represents over- and under-segmentation equally, but is also more sensitive to outliers/errors:

Over/undersegmentation:



Outliers:



From: Common Limitations of Image Processing Metrics: A Picture Story (https://arxiv.org/pdf/2104.05642.pdf)

Evaluation – Hausdorff distance

Can you create a segmentation example with low Hausdorff distance, but low Dice?

Evaluation metrics – other metrics

- Percentile-based Hausdorf distance
- Surface-distance metrics, .e.g, Average symmetric surface distance (ASSD)

Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool

CrossMark

Abdel Aziz Taha* and Allan Hanbury

Abstract

Background: Medical Image segmentation is an important image processing step. Comparing images to evaluate the quality of segmentation is an essential part of measuring progress in this research area. Some of the challenges in evaluating medical segmentation are: metric selection, the use in the literature of multiple definitions for certain metrics, inefficiency of the metric calculation implementations leading to difficulties with large volumes, and lack of support for fuzzy segmentation by existing metrics.

Result: First we present an overview of 20 evaluation metrics selected based on a comprehensive literature review. For fuzzy segmentation, which shows the level of membership of each voxel to multiple classes, fuzzy definitions of all metrics are provided. We present a discussion about metric properties to provide a guide for selecting evaluation metrics. Finally, we propose an efficient evaluation tool implementing the 20 selected metrics. The tool is optimized to perform efficiently in terms of speed and required memory, also if the image size is extremely large as in the case of whole body MRI or CT volume segmentation. An implementation of this tool is available as an open source project.

Conclusion: We propose an efficient evaluation tool for 3D medical image segmentation using 20 evaluation metrics and provide guidelines for selecting a subset of these metrics that is suitable for the data and the segmentation task.

Keywords: Evaluation metrics, Evaluation tool, Medical volume segmentation, Metric selection

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4533825/

Ranks given by different metrics can be compared

Correlation Dice ranking & HD ranking = 0.52

	ARI	KAP	ICC	DICE	AVD	MHD	PBD	VS	MI	AUC	TPR	HD	TNR	RI	GCE	VOI	
ARI	1.00	1.00	1.00	1.00	0.95	0.93	0.91	0.81	0.80	0.75	0.74	0.52	-0.07	-0.07	-0.15	-0.15	Group 1
KAP	1.00	1.00	1.00	1.00	0.95	0.93	0.91	0.81	0.80	0.75	0.74	0.52	-0.08	-0.08	-0.16	-0.16	
ICC	1.00	1.00	1.00	1.00	0.95	0.93	0.91	0.81	0.81	0.75	0.74	0.52	-0.08	-0.09	-0.17	-0.17	
DICE	1.00	1.00	1.00	1.00	0.95	0.93	0.91	0.81	0.81	0.75	0.74	0.52	-0.08	-0.09	-0.17	-0.17	
AVD	0.95	0.95	0.95	0.95	1.00	0.93	0.86	0.76	0.67	0.70	0.69	0.70	0.07	0.08	0.00	0.00	
MHD	0.93	0.93	0.93	0.93	0.93	1.00	0.83	0.71	0.73	0.74	0.74	0.53	-0.07	-0.06	-0.13	-0.13	
PBD	0.91	0.91	0.91	0.91	0.86	0.83	1.00	0.74	0.71	0.65	0.64	0.45	-0.07	-0.09	-0.16	-0.16	
VS	0.81	0.81	0.81	0.81	0.76	0.71	0.74	1.00	0.60	0.45	0.44	0.40	-0.03	0.00	-0.08	-0.07	
MI	0.80	0.80	0.81	0.81	0.67	0.73	0.71	0.60	1.00	0.65	0.65	0.22	-0.49	-0.58	-0.64	-0.64	Group 3
AUC	0.75	0.75	0.75	0.75	0.70	0.74	0.65	0.45	0.65	1.00	1.00	0.35	-0.35	-0.14	-0.19	-0.19	
TPR	0.74	0.74	0.74	0.74	0.69	0.74	0.64	0.44	0.65	1.00	1.00	0.34	-0.36	-0.15	-0.20	-0.20	
HD	0.52	0.52	0.52	0.52	0.70	0.53	0.45	0.40	0.22	0.35	0.34	1.00	0.32	0.35	0.30	0.30	
TNR	-0.07	-0.08	-0.08	-0.08	0.07	-0.07	-0.07	-0.03	-0.49	-0.35	-0.36	0.32	1.00	0.84	0.84	0.84	Group 2
RI	-0.07	-0.08	-0.09	-0.09	0.08	-0.06	-0.09	0.00	-0.58	-0.14	-0.15	0.35	0.84	1.00	0.99	1.00	
GCE	-0.15	-0.16	-0.17	-0.17	0.00	-0.13	-0.16	-0.08	-0.64	-0.19	-0.20	0.30	0.84	0.99	1.00	1.00	
VOI	-0.15	-0.16	-0.17	-0.17	0.00	-0.13	-0.16	-0.07	-0.64	-0.19	-0.20	0.30	0.84	1.00	1.00	1.00	
	Group 1								Group 3				Group 2				

Fig. 3 The correlation between the rankings produced by 16 different metrics. The pair-wise Pearson's correlation coefficients between the rankings of 4833 medical volume segmentations produced by 16 metrics. The color intensity of each cell represents the strength of the correlation, where blue denotes direct correlation and red denotes inverse correlation

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4533825/

Summary

The student can:

- Describe what problems arise with using a threshold-based segmentation method
- Calculate how many parameters are needed for K Gaussians in p dimensions
- Recognize covariance matrices and identify their properties

- select a suitable evaluation metric for a given medical image analysis task
- calculate the accuracy, Dice score and Hausdorff distance, given an image segmentation
- interpret the results and assess the quality of the validation methods used in medical image analysis research papers

Thank you

r.su@tue.nl

Next: Segmentation in feature space