Imperial College London

AI in Medicine I

AI/Machine Learning for Imaging – Image Segmentation

Prof Julia Schnabel

Chair for Computational Imaging and AI in Medicine
School of Computation, Information & Technology – Department of Computer Science
Theoretical Tutorial & Practical: Cosmin Bercea and Felix Meissen

Slide contributions: Dr Ben Glocker, Imperial College London, UK

Overview

- Introduction to segmentation
  - *why do we do this?*
- Challenges for segmentation
  - *why is it so hard?*
- Evaluation of image segmentation
  - *how good is it?*
- Segmentation Algorithms & Techniques
  - *how does it work?*
- Summary
  - *take home messages*

Introduction to Segmentation

Common Image Analysis Tasks

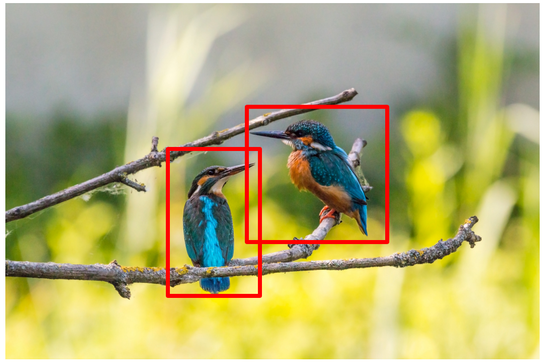
Image Classification

Output: Category (e.g., “bird”)

Object Detection

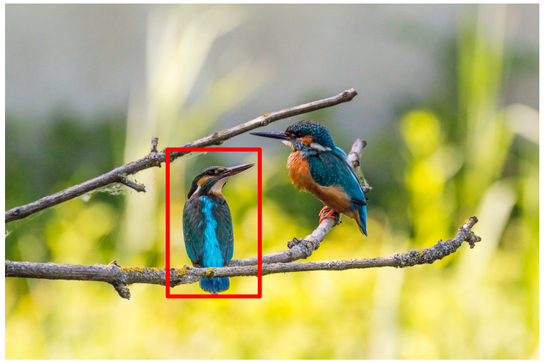
Output: Coordinates (e.g., centroid)

Object Localisation



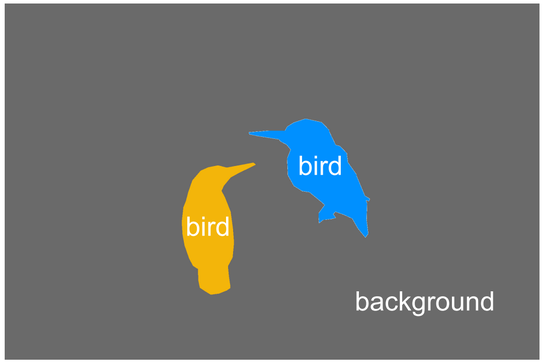
Output: Coordinates (e.g., bounding box)

Object Recognition



Output: Category (e.g., “kingfisher”)

Semantic Segmentation

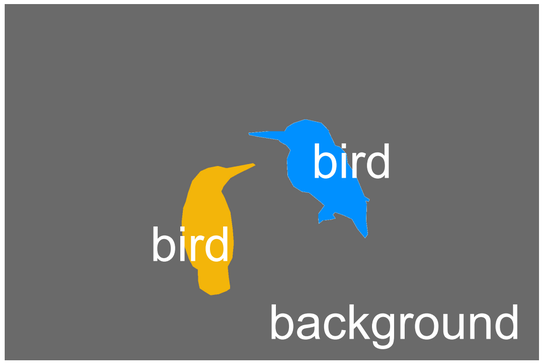


Output: Label map

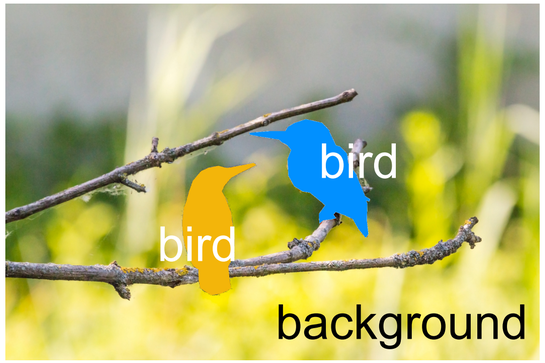
Image Captioning

Output: Text
(e.g., “two birds sitting on a branch”)

Semantic Segmentation



Output: Label map

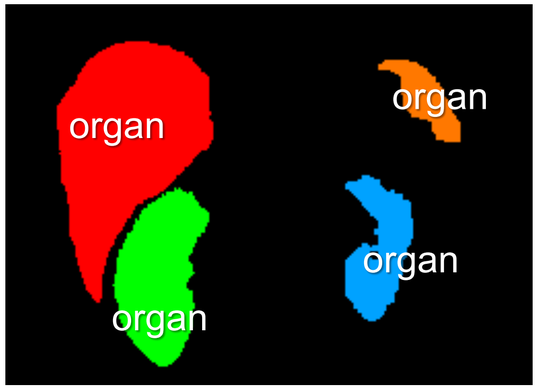


Output: Label map

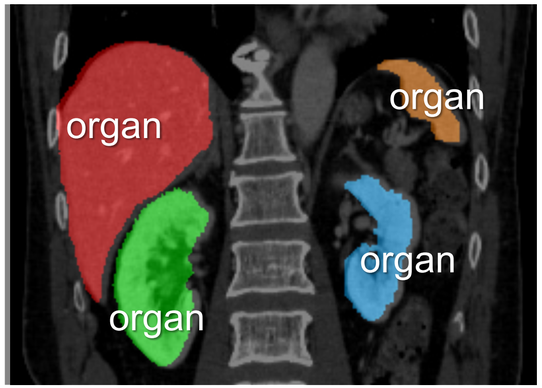
In semantic segmentation, a segmented region is also assigned a **semantic meaning** (*and not just some clustering into coherent regions)*

Semantic Segmentation

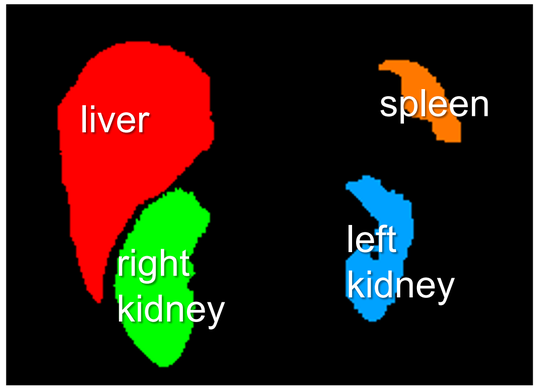
Original image



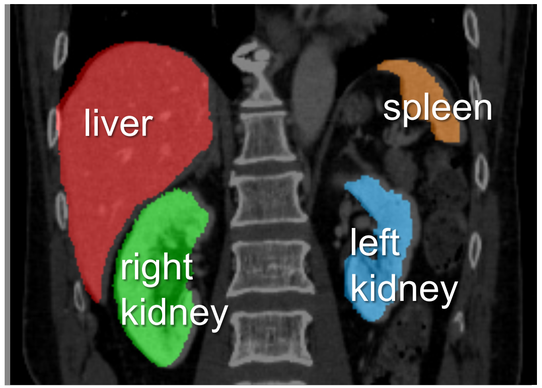
Output: Label map



Output: Label map



Semantic segmentation



Semantic segmentation

Segmentation in Medical Imaging

- **Image segmentation is useful for...**
  - conducting **quantitative analyses**, e.g. measuring the volume of the ventricular cavity.
  - determining the precise **location and extent** of an organ or a certain type of tissue, e.g. a tumour, for treatment such as radiation therapy.
  - creating **3D models** used for **simulation**, e.g. generating a model of an abdominal aortic aneurysm for simulating stress/strain distributions.

Challenges for Medical Image Segmentation

- There are several effects in medical imaging that may hamper the application of segmentation algorithms
  - noise,
  - partial volume effects,
  - intensity inhomogeneities,
  - anisotropic resolution,
  - imaging artifacts,
  - limited contrast,
  - morphological variability,
  - and many more …

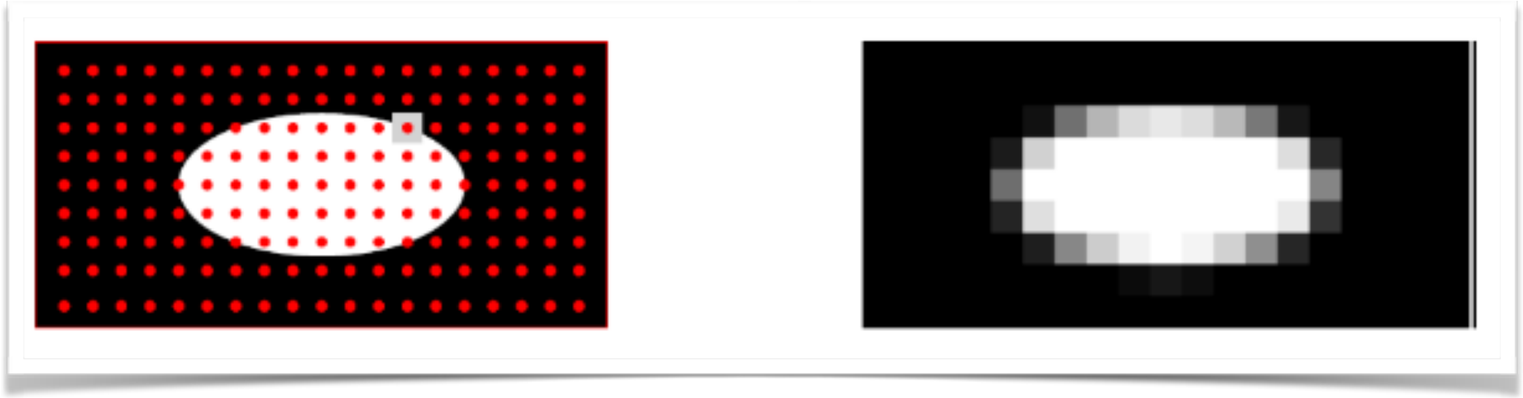
Noise

- Example: Transcranial 2D ultrasound image of the brain

Not to be confused with speckle, attenuation and reflection

Partial Volume Effects

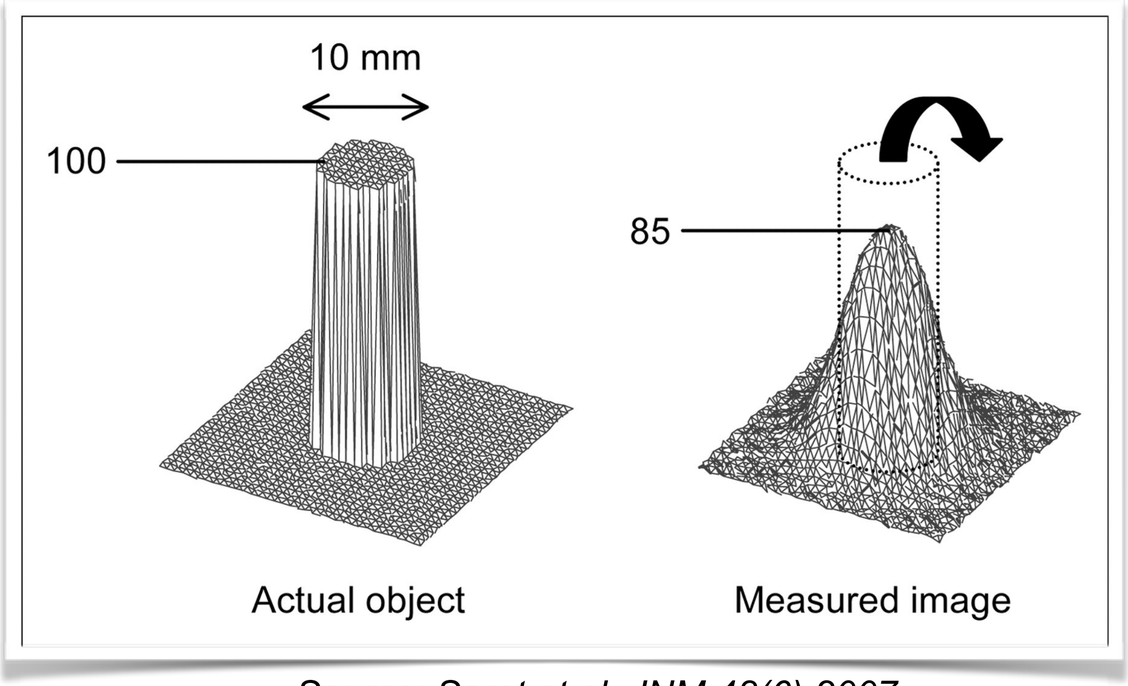
- Due to the coarse, **discrete sampling** (red dots), the resulting image shows **partial volume effects** at the boundary of the white “organ”



- Both ”tissue" types (black and white) contribute to the intensity value of the generated image (right) due to the relatively large influence area of each pixel (gray squares in right image)

Partial Volume Effects

- The point spread function (PSF) describes **the response of an imaging system to a point source or point object**.



Source: Soret et al. JNM 48(6) 2007

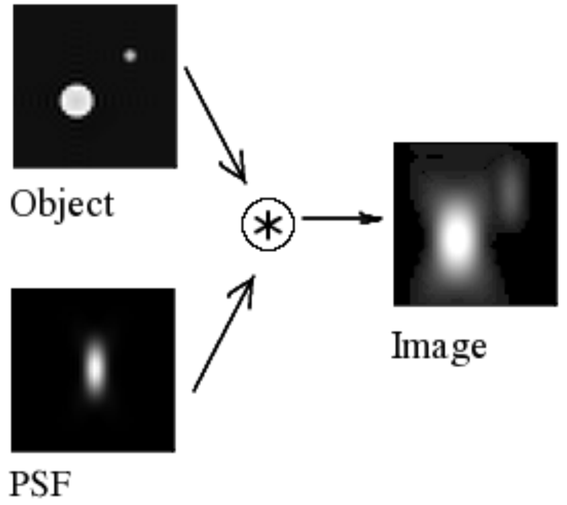
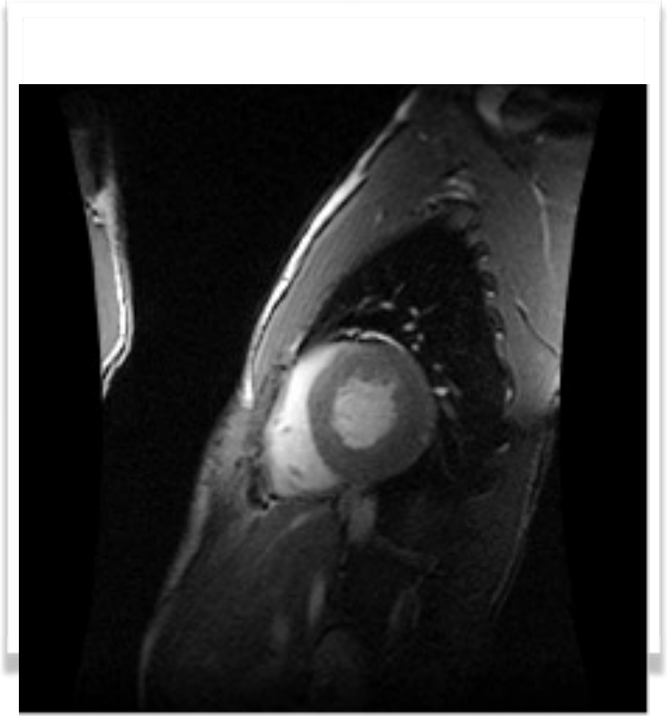


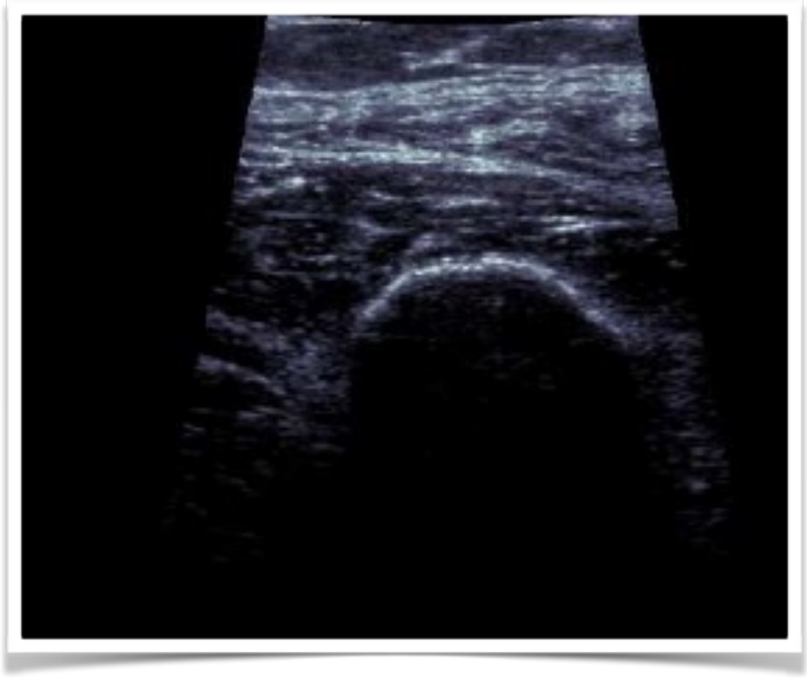
Image formation in a confocal microscope: central longitudinal (XZ) slice. The 3D acquired distribution arises from the convolution of the real light sources with the **PSF**. *Source: Wikipedia*

Intensity Inhomogeneities

- In particular, typical for MRI and ultrasound imaging data



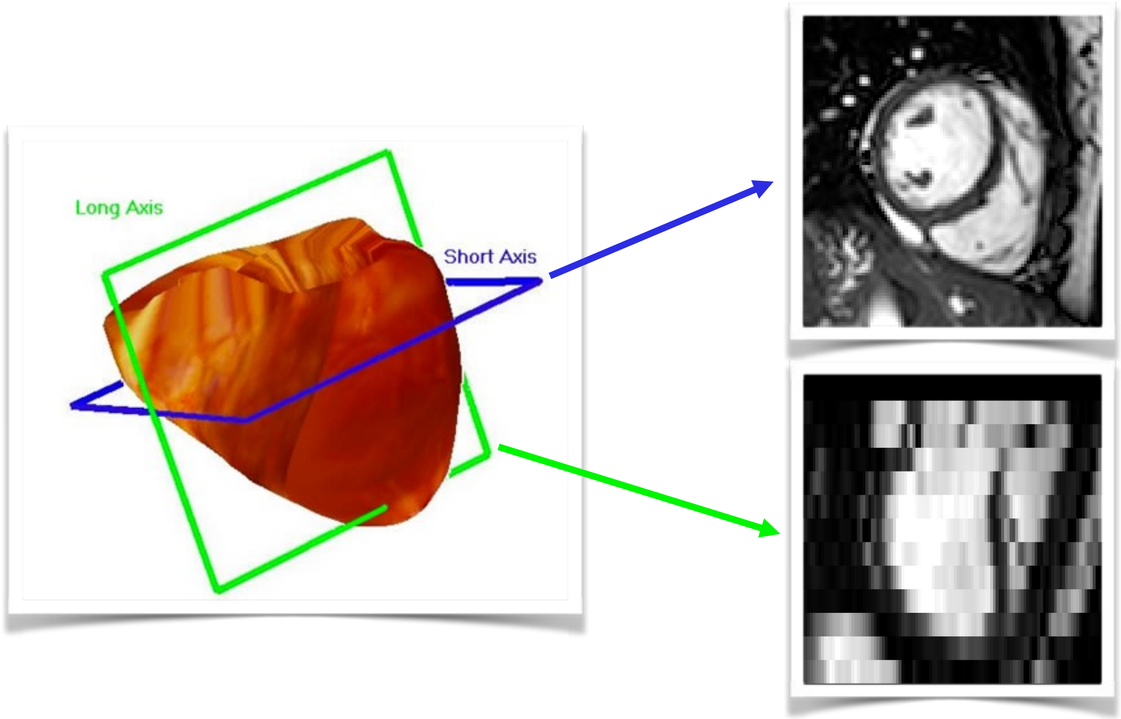
Magnetic field inhomogeneity



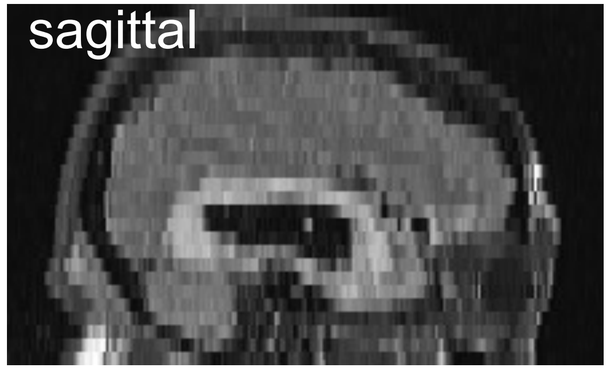
Attenuation

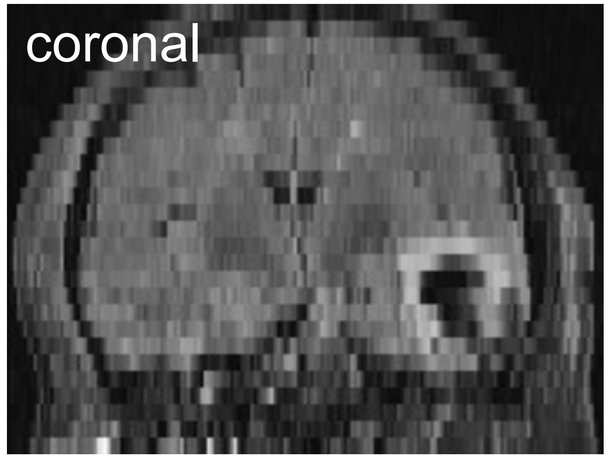
Anisotropic Resolution

- Almost all 3D imaging modalities, such as CT or MRI suffer from this problem
  - For example, (dynamic) MRI acquisition of short axis slices of the left ventricle may have intra-slice resolution of 1.3mm and an inter-slice resolution of 8mm

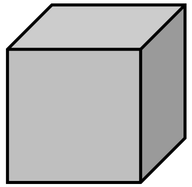


Anisotropic Resolution

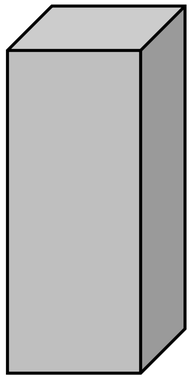




Voxel size: 0.7 x 0.7 x 5mm



isotropic



anisotropic

(also called “Manhattan voxels)

Imaging Artifacts

CT image showing streak artifacts caused by metal implants

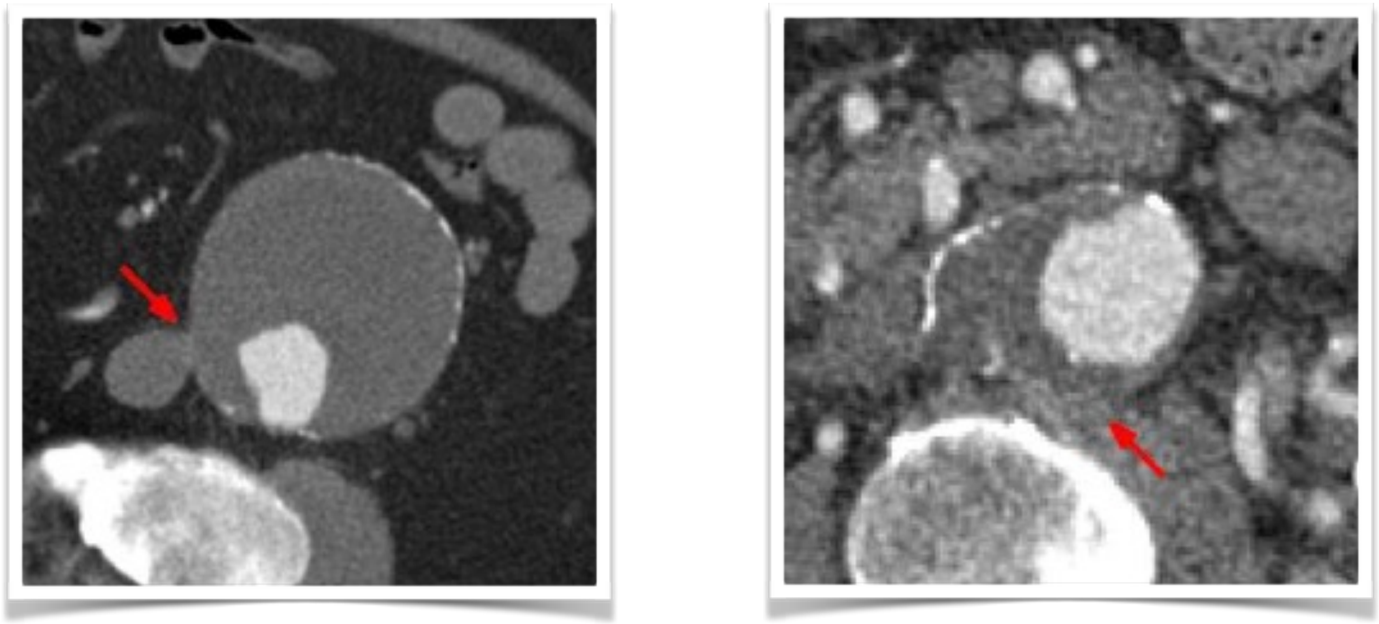
(dental filling)

Source: American Journal of Roentgenology. 2004;182: 532-532.

T1-weighted MR image showing susceptibility artifacts caused by iron oxide particles suspended in the beeswax dressing in the hair of the patient.

Limited Contrast

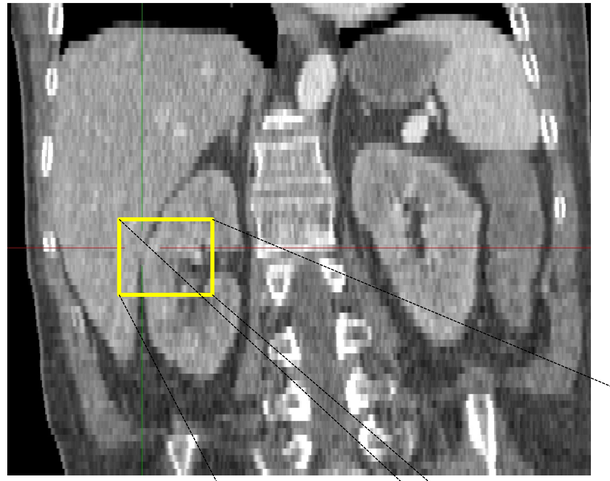
- Different tissues can have similar physical properties and thus similar intensity values



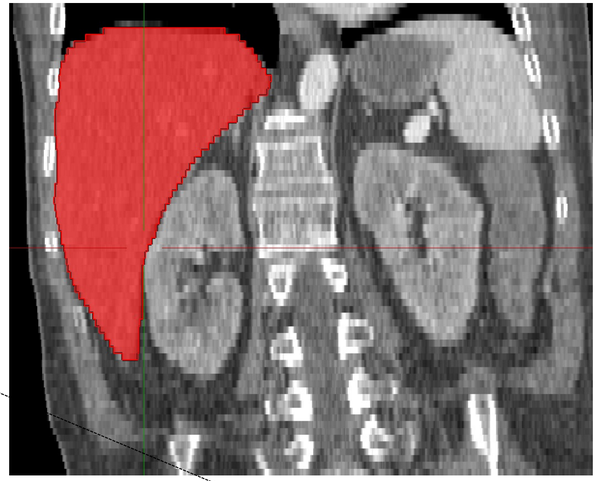
Example: Thrombus in the wall of abdominal aortic aneurysms is often hard to distinguish from the surrounding tissue (modality: CTA).

- Purely intensity-based algorithms are prone to fail or “leak” into adjacent tissue

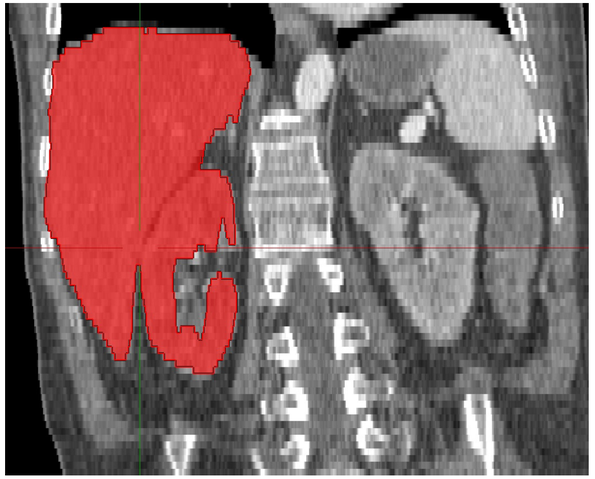
Limited Contrast



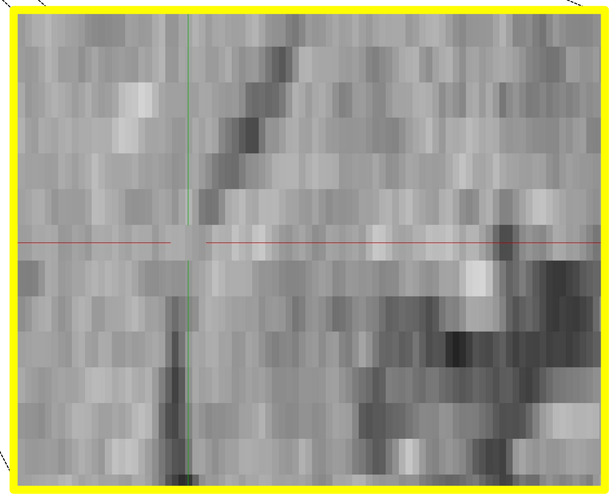
CT abdominal image



manual liver segmentation

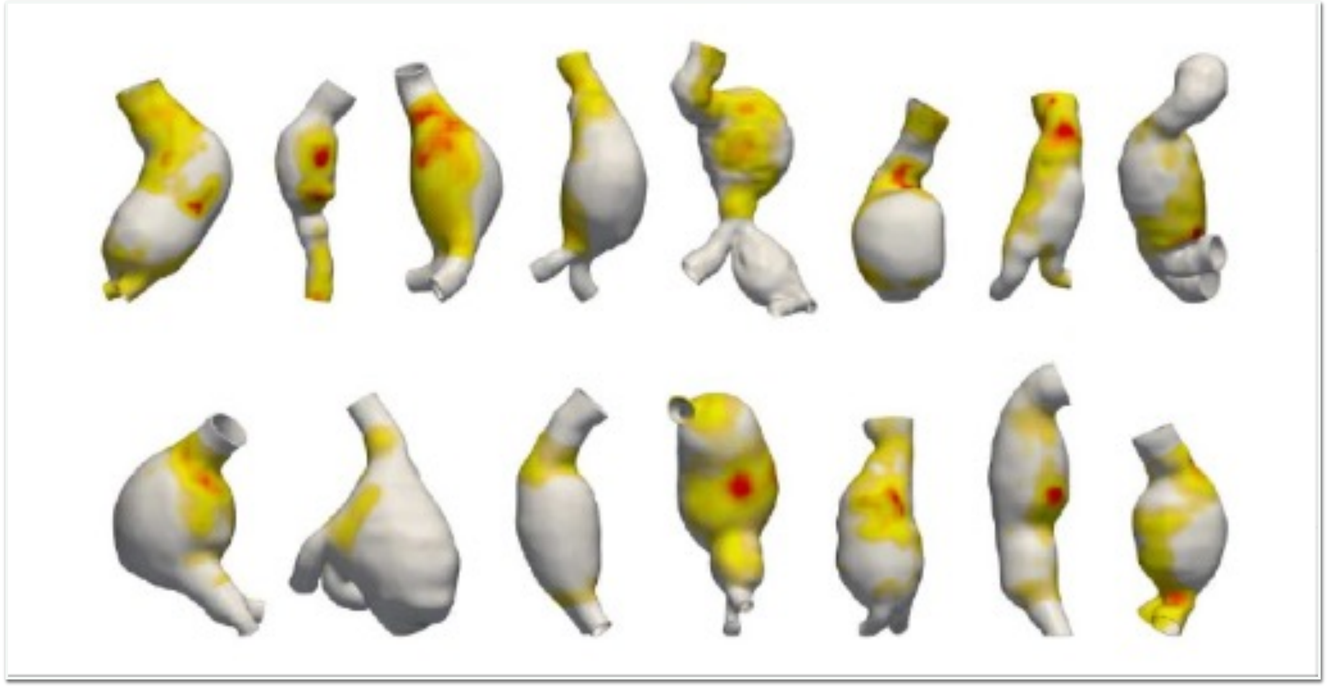


region growing with leakage into kidney



Morphological Variability

- Morphological variability makes it hard to incorporate meaningful prior information or useful shape models



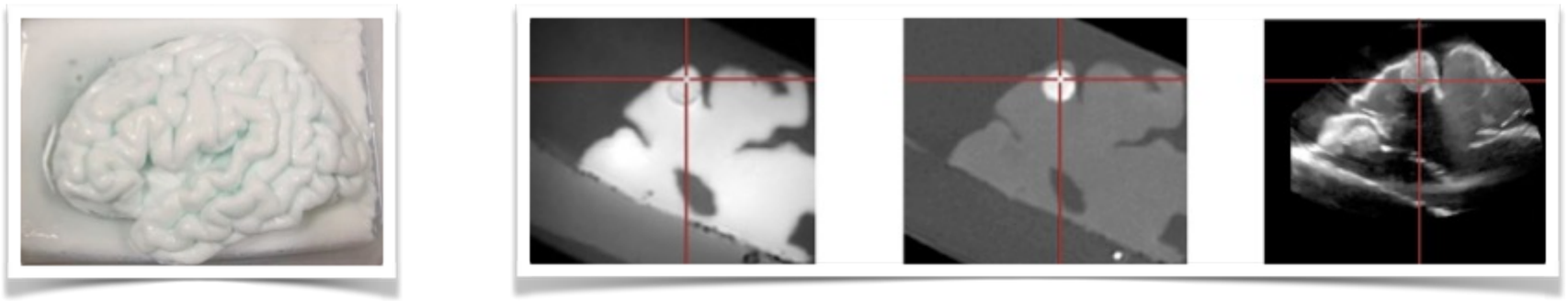
Example: A collection of abdominal aortic aneurysms acquired with PET-CT and colored by FDG-18 uptake values.

Evaluating Image Segmentation

- Ground truth
- Gold standard
- Performance measures *(some you have already come across in the image classification lecture!)*
  - precision/recall
  - sensitivity/specificity
  - accuracy
  - Segmentation overlap
  - Surface distance measures

Ground Truth

- Reference or standard against a method can be compared, e.g. the optimal transformation, or a true segmentation boundary
- In practice, it is not available!
- Instead, one makes one up
  - Synthetic or simulated phantoms (e.g., computer phantoms)
  - Physical phantom (e.g., gel phantoms)

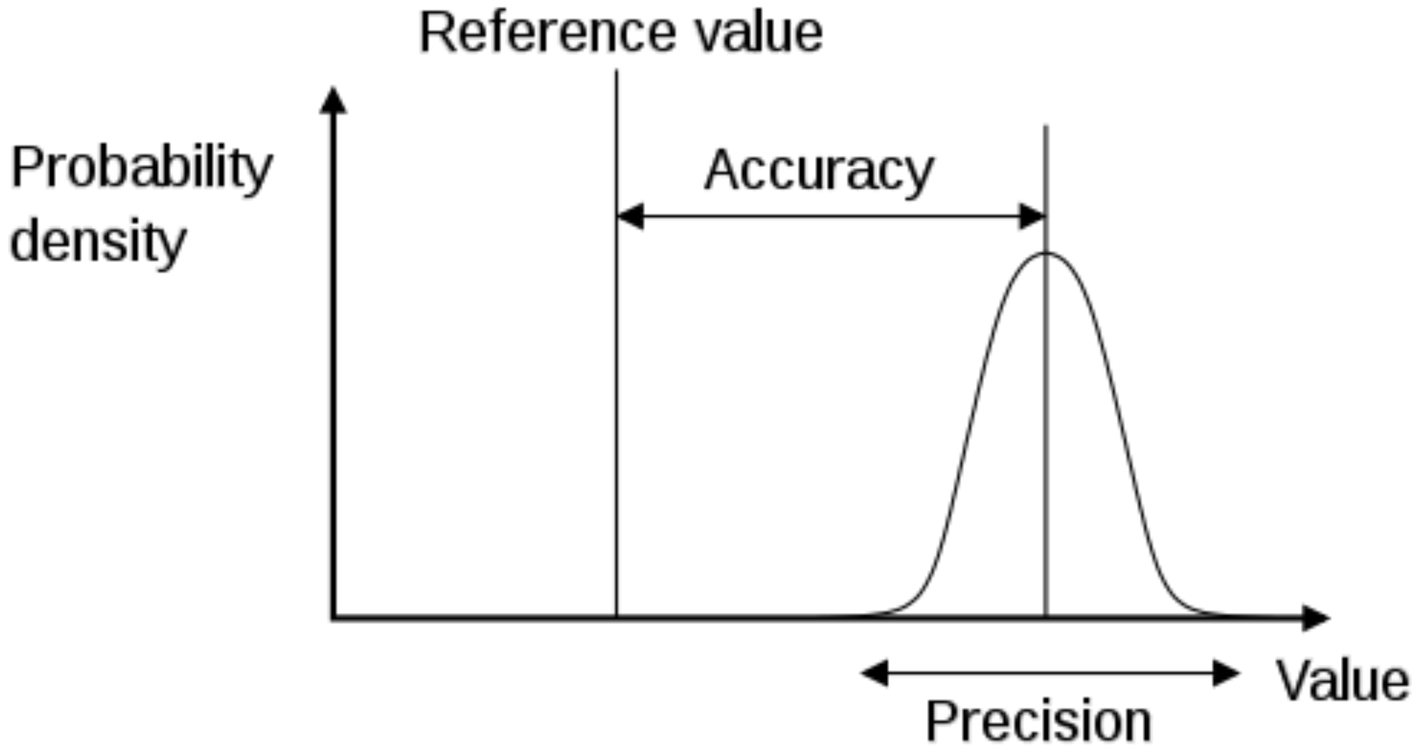


PVA mould of a brain model, scanned with MR,CT and US.

Gold Standard

- Expert (often referred to as *gold* standard)
  - Manual segmentation by human observer (e.g. experienced clinician)
- Disadvantage
  - Requires training and is tedious and time-consuming
  - Intra-observer variability (disagreement between same observer on different occasions)
  - Inter-observer variability (disagreement between observers)
- Remedy
  - Human observer can perform segmentation repeatedly
  - A number of experts can perform segmentations
  - Agreement or disagreement can be quantified

How to Assess Performance?



How to Assess Performance?

Precision

- is a description of **random errors**, or a measure of **statistical variability**
- the **repeatability**, or r**eproducibility** of the measurement

Accuracy *has two definitions:*

- More commonly: description of **systematic errors**, a measure of **statistical bias**; as these cause a difference between a result and a "true" value, ISO also calls this **trueness**.
- Alternatively: ISO defines accuracy as describing a combination of both types of observational error above (**random** and **systematic**), so high accuracy requires both high precision and high trueness

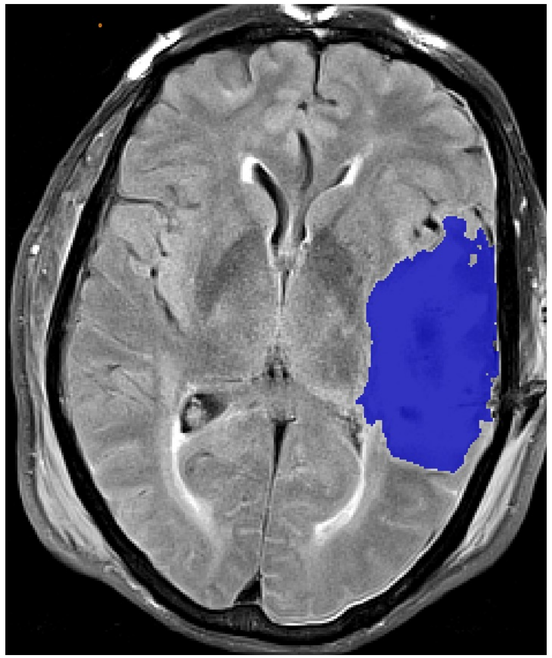
Robustness

- refers to the **degradation in performance** with respect to varying noise levels or other imaging artefacts

Performance Measures

- Assume we have an automated segmentation algorithm, how do we assess ‘how good it is’?
- There are several ways of quantifying segmentation performance

MR image



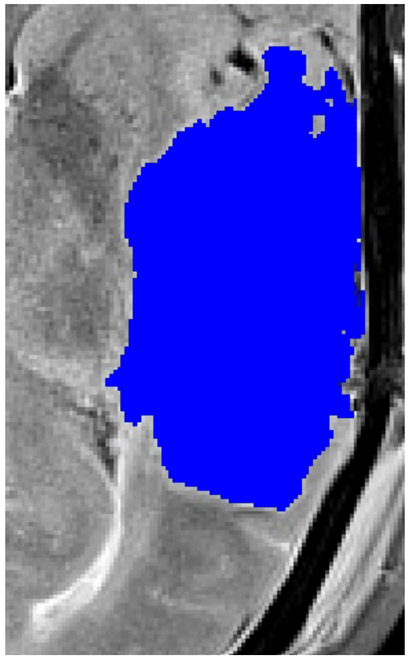
Gold standard

Auto Segmentation

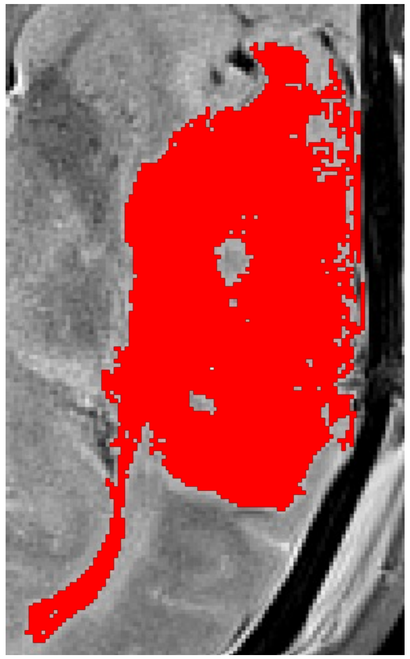
Performance Measures

- Assume we have an automated segmentation algorithm, how do we assess ‘how good it is’?
- There are several ways of quantifying segmentation performance

MR image



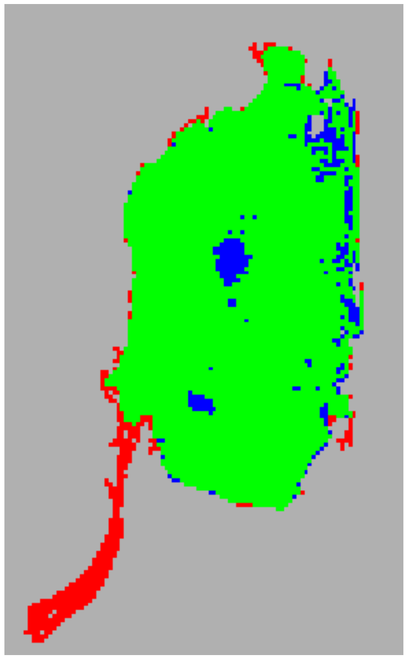
Gold standard



Auto Segmentation

Performance Measures

- Measure performance of an automated method in terms of **agreement** of its result with a reference gold standard

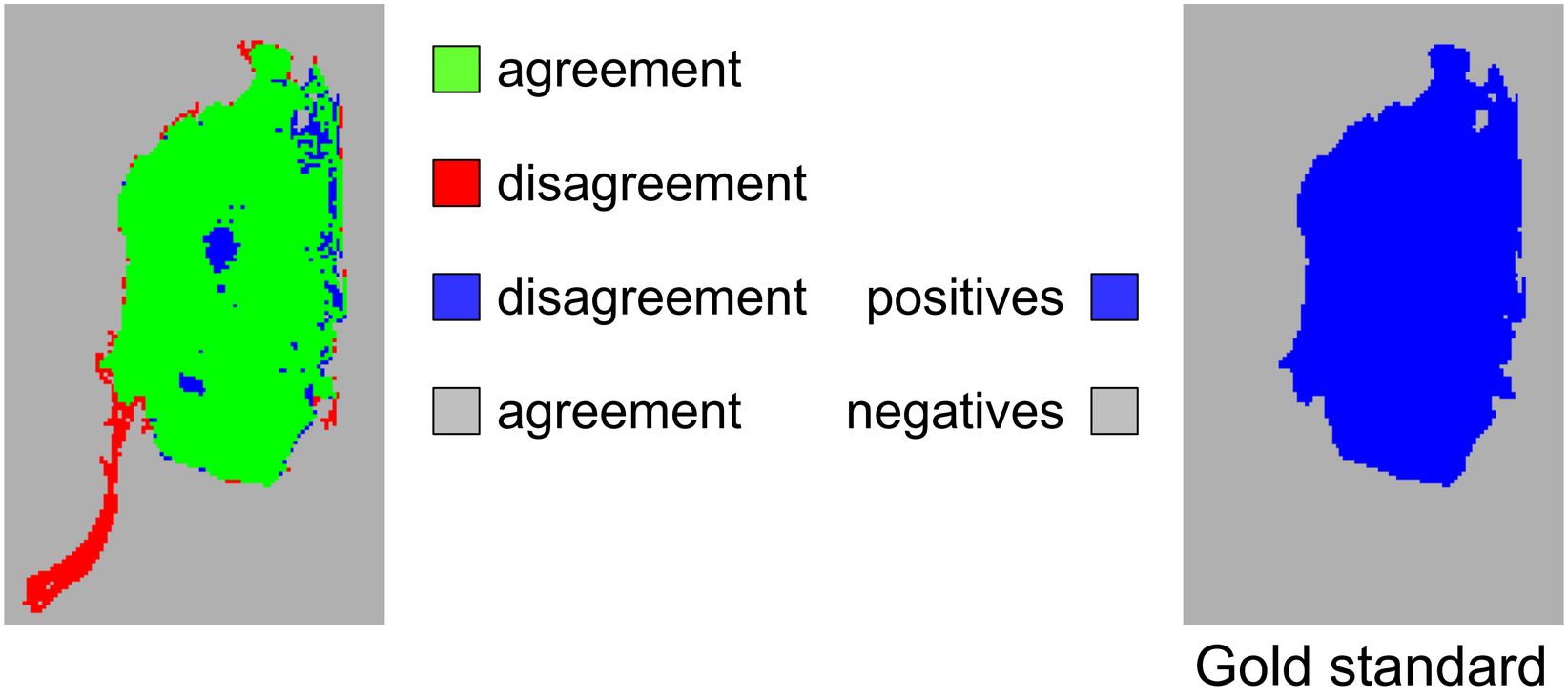


Gold standard

Auto Segmentation

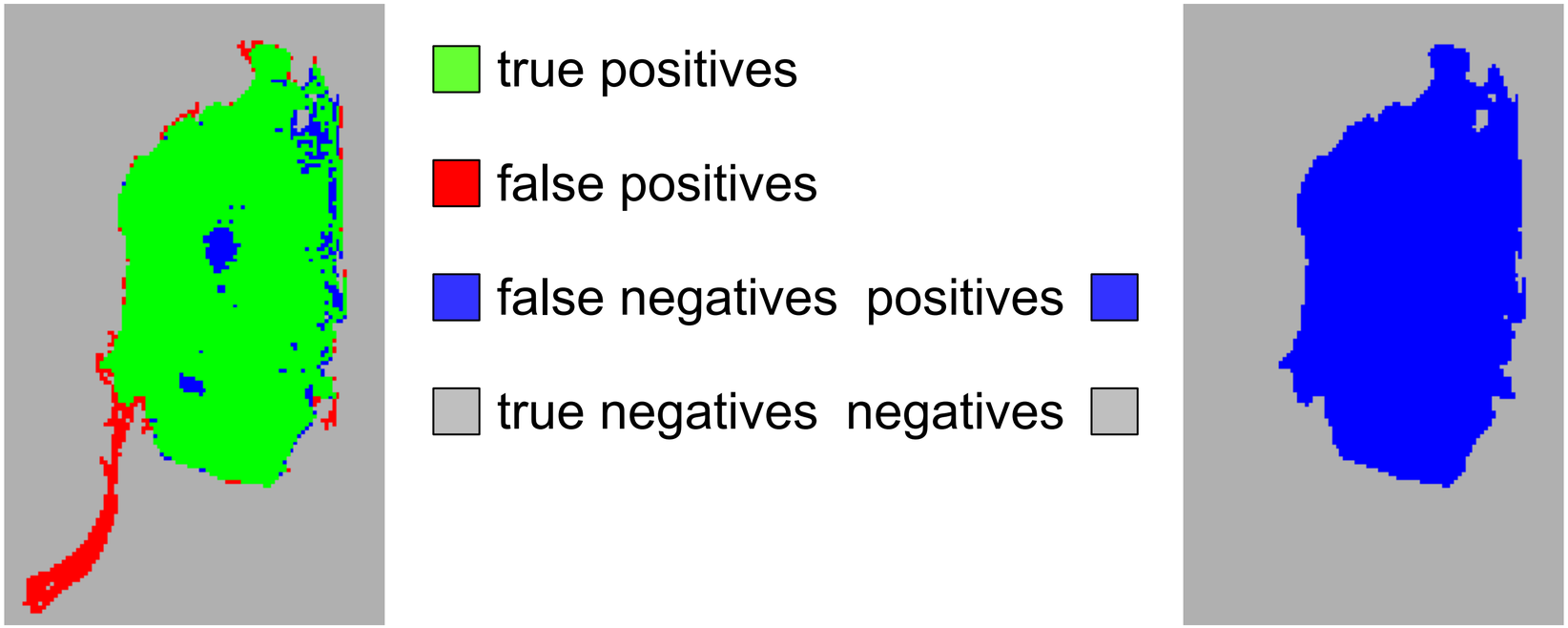
Performance Measures

- Measure performance of an automated method in terms of **agreement** of its result with a reference gold standard



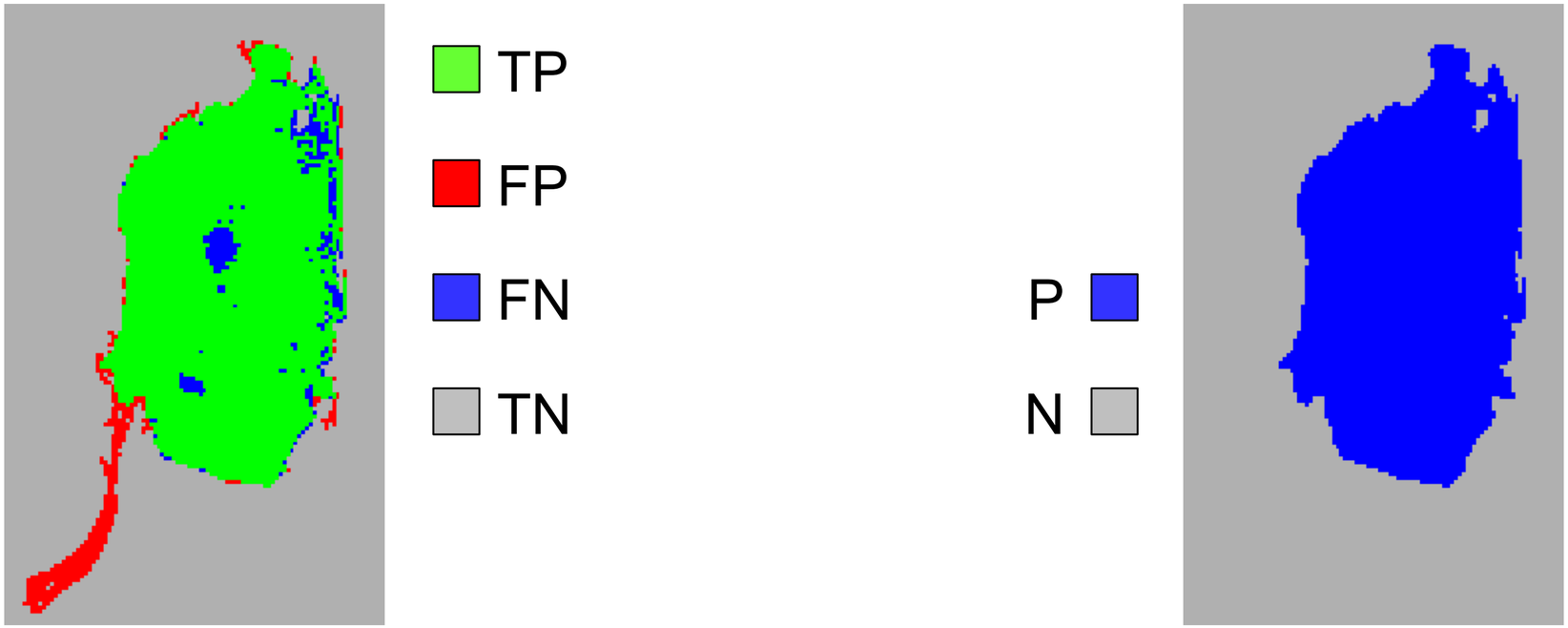
Performance Measures

- Measure performance of an automated method in terms of **agreement** of its result with a reference gold standard



Performance Measures

- Measure performance of an automated method in terms of **agreement** of its result with a reference gold standard



Confusion Matrix

condition positive (P)
the number of real positive cases in the data
condition negatives (N)
the number of real negative cases in the data

true positive (TP)
eqv. with hit
true negative (TN)
eqv. with correct rejection
false positive (FP)
eqv. with false alarm, Type I error
false negative (FN)
eqv. with miss, Type II error

		True condition	
	Total population	Condition positive	Condition negative
Predicted condition	Predicted condition positive	**True positive**	**False positive**, Type I error
	Predicted condition negative	**False negative**, Type II error	**True negative**

Accuracy, Precision, Recall, ...

accuracy

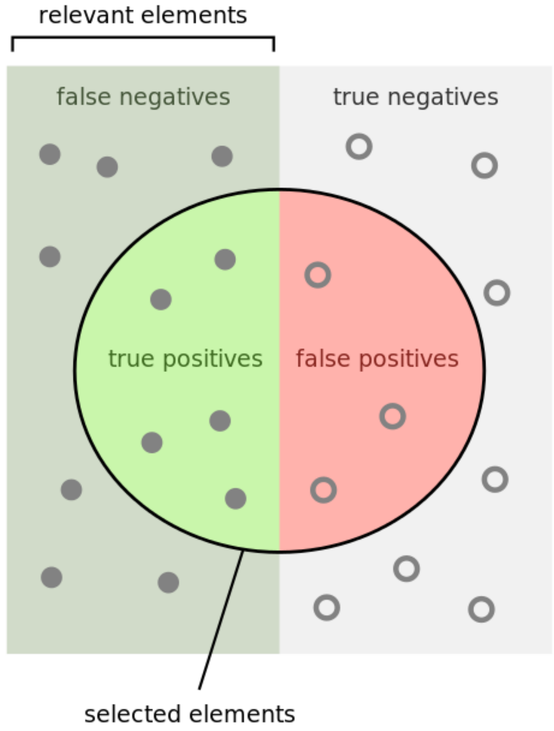
$$ACC = \frac{TP+TN}{P+N},\ P = TP + FN,\ N = TN + FP$$

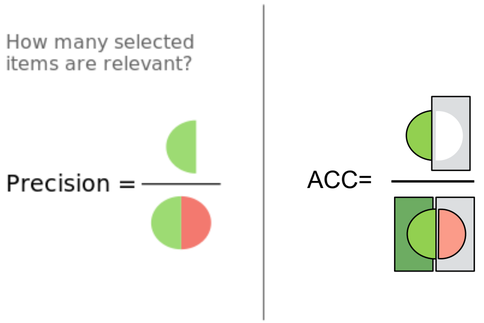
e.g., if TP=P and TN=N, then ACC=1
or if TP=0 and TN=0, then ACC=0

precision or **positive predictive value**

$$PPV = \frac{TP}{TP + FP}$$

e.g., if FP=0 then PPV=1
or if TP=0 then PPV=0





Accuracy, Precision, Recall, ...

recall, sensitivity, **hit rate** or **true positive rate**

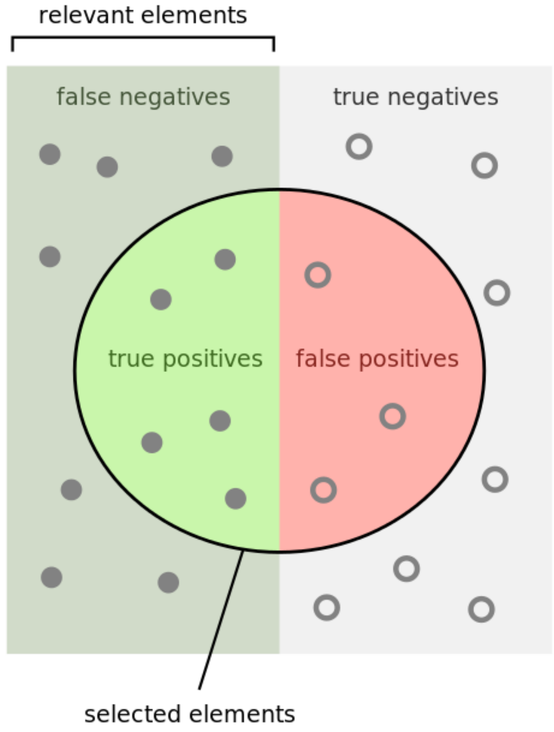
$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$$

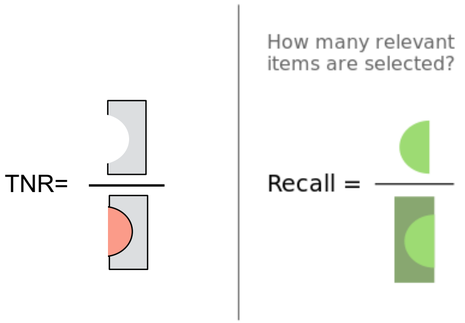
e.g., if TP=P, then TPR=1
or if TP=0, then TPR=0

specificity or **true negative rate**

$$TNR = \frac{TN}{N} = \frac{TN}{TN + FP}$$

e.g., if TN=N then TNR=1
or if TN=0 then TNR=0



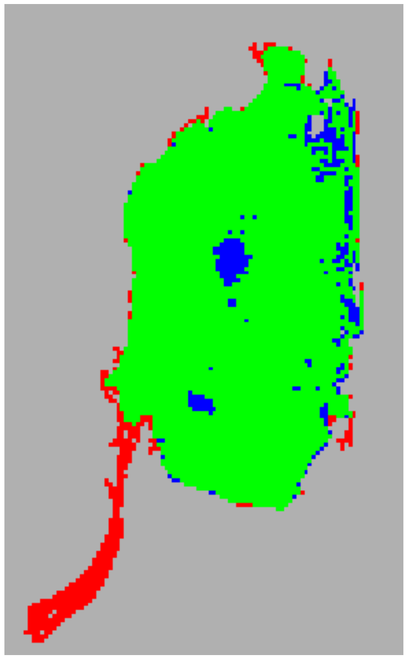


F1 Score

precision $PPV = \frac{TP}{TP+FP}$ **recall** $TPR = \frac{TP}{TP+FN}$

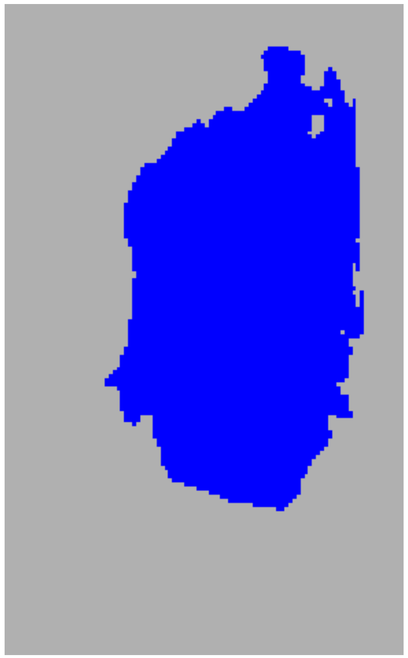
F1 score
is the harmonic mean of precision and recall

$$F_1 = 2\frac{PPV \cdot TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$$



TP
FP
FN
TN

P
N

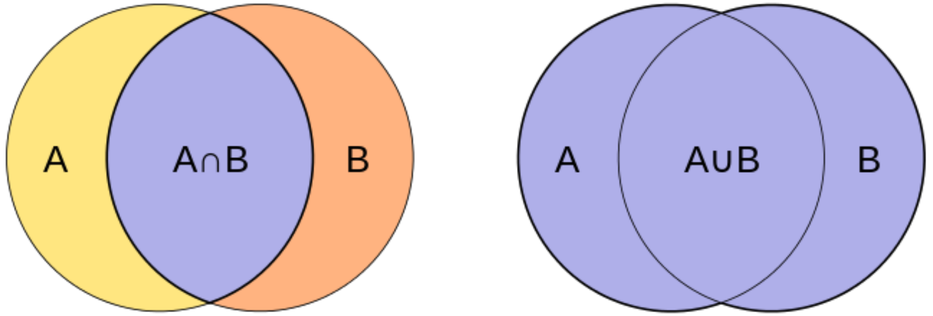


Gold standard

Overlap

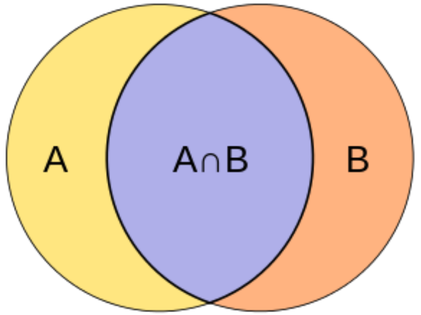
- Measure agreement by looking at overlap of reference and predicted segmentation (e.g., auto segmentation)
- **Jaccard Index**
  - aka **Jaccard Similarity Coefficient**
  - aka **Tanimoto Coefficient**
  - aka **Intersection over Union (IoU)**

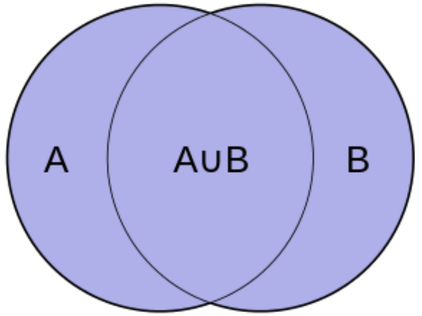
$$JI = \frac{|A \cap B|}{|A \cup B|}$$

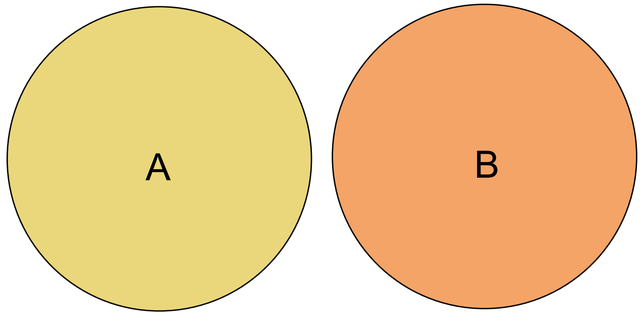


Overlap

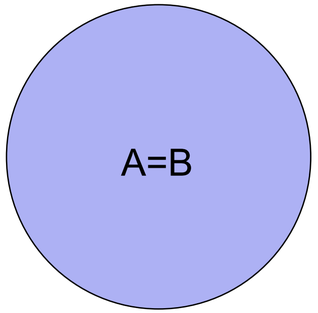
- Measure agreement by looking at overlap of reference and predicted segmentation (e.g., auto segmentation)
- **Jaccard Index**
  - JI=0 for no overlap
  - JI=1 for perfect overlap







$|A \cap B| = 0$

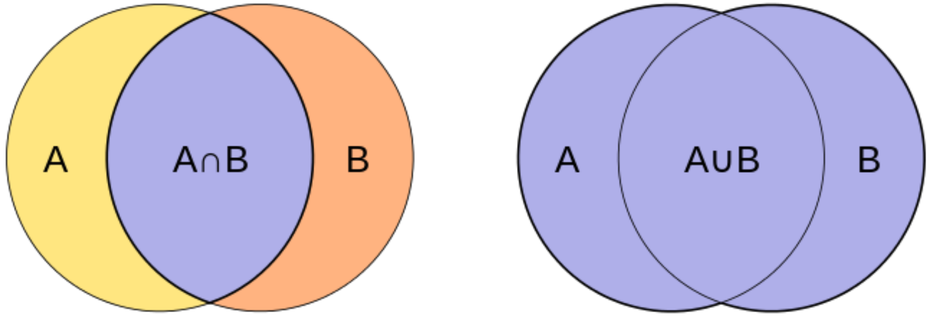


$|A \cap B| = |A \cup B|$

Overlap

- Measure agreement by looking at overlap of reference and predicted segmentation (e.g., auto segmentation)
- **Dice Coefficient**
  - aka **Sørensen Index**
  - aka **Dice Similarity Coefficient (DSC)**
  - aka **intersection over mean volume**

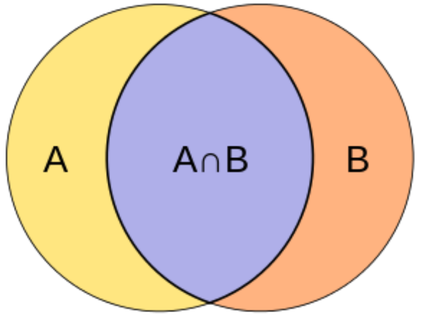
$$DSC = \frac{|A \cap B|}{(|A| + |B|) \,/\, 2} = \frac{2|A \cap B|}{|A| + |B|}$$

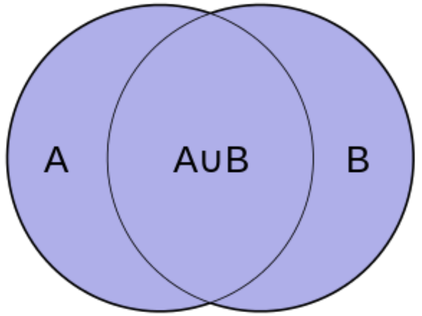


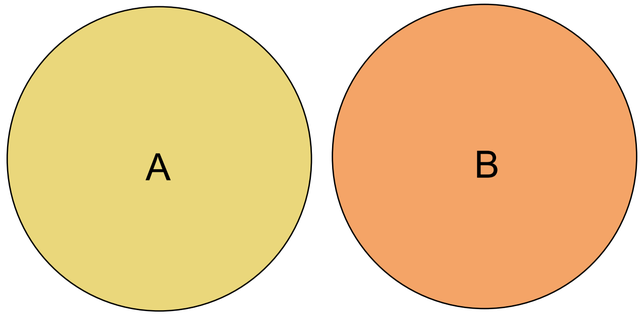
Overlap

- Measure agreement by looking at overlap of reference and predicted segmentation (e.g., auto segmentation)
- **Dice Coefficient**
  - DSC=0 for no overlap
  - DSC=1 for perfect overlap

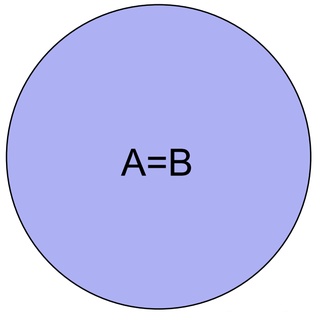
$$DSC = \frac{|A \cap B|}{(|A| + |B|) \,/\, 2} = \frac{2|A \cap B|}{|A| + |B|}$$







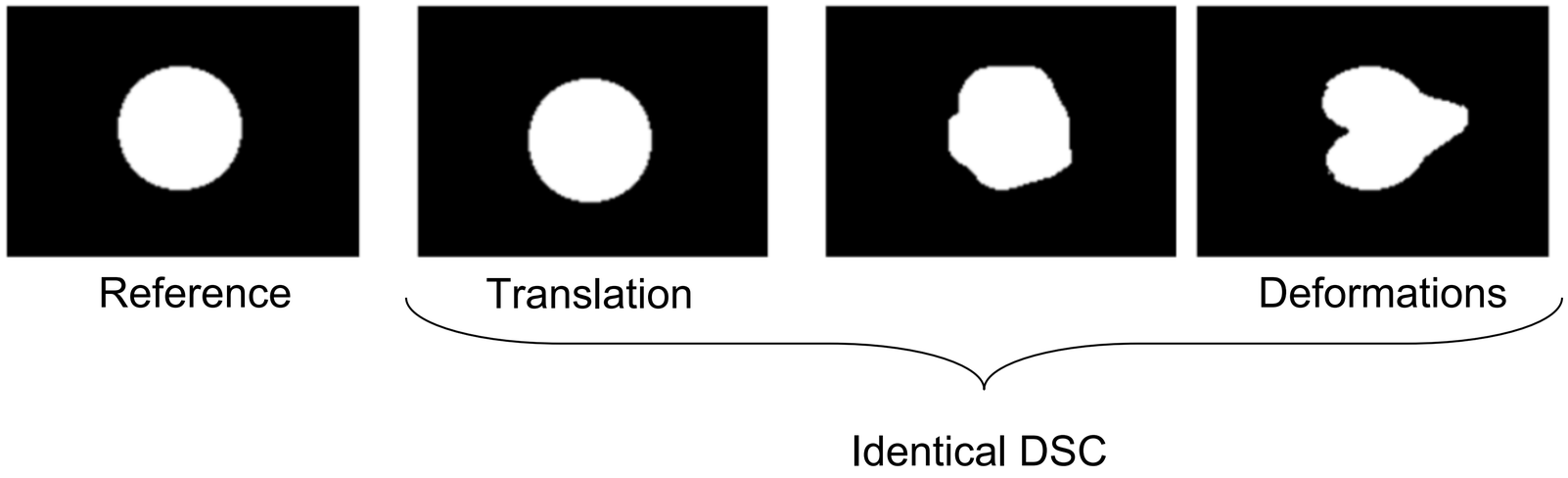
$$|A \cap B| = 0$$



$$|A \cap B| = \frac{|A| + |B|}{2}$$

Overlap - Limitations

- A limitation of most overlap measures is that they are sensitive to overall volume size and shape:



See Konukoglu et al. 2012 IEEE TMI

Other Measures

Surface distance measures

- **Hausdorff distance**

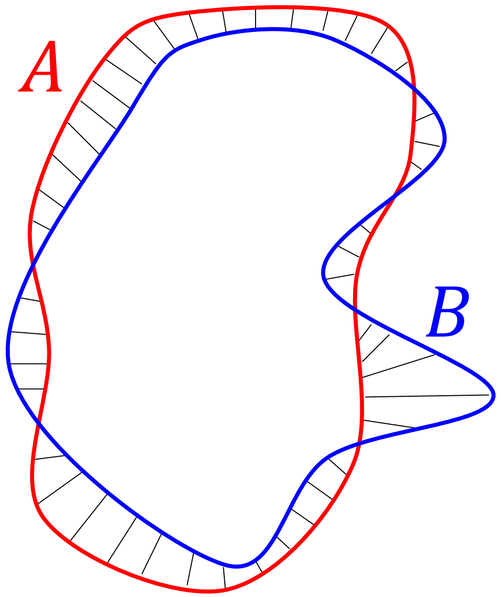
$$HD = \max\big(h(A,B), h(B,A)\big)$$

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

- **(symmetric) average surface distance**

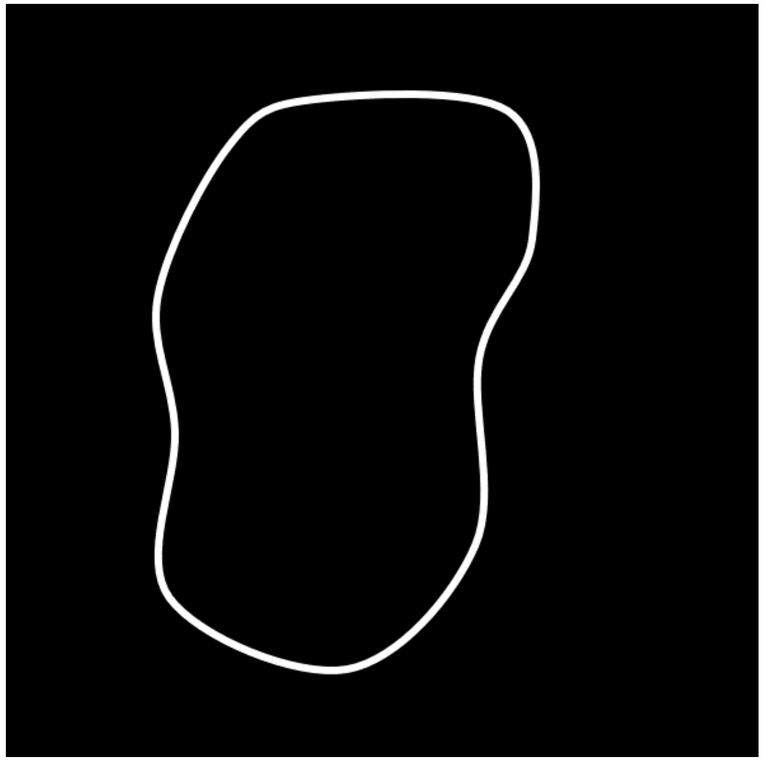
$$ASD = \frac{d(A,B)+d(B,A)}{2}$$

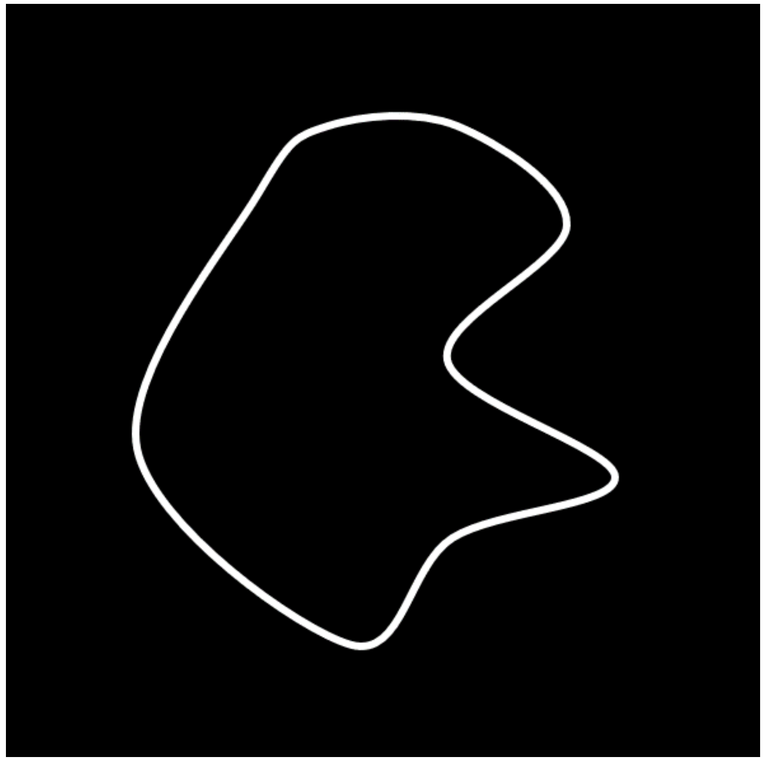
$$d(A,B) = \frac{1}{N}\sum_{a\in A} \min_{b\in B} \|a-b\|$$



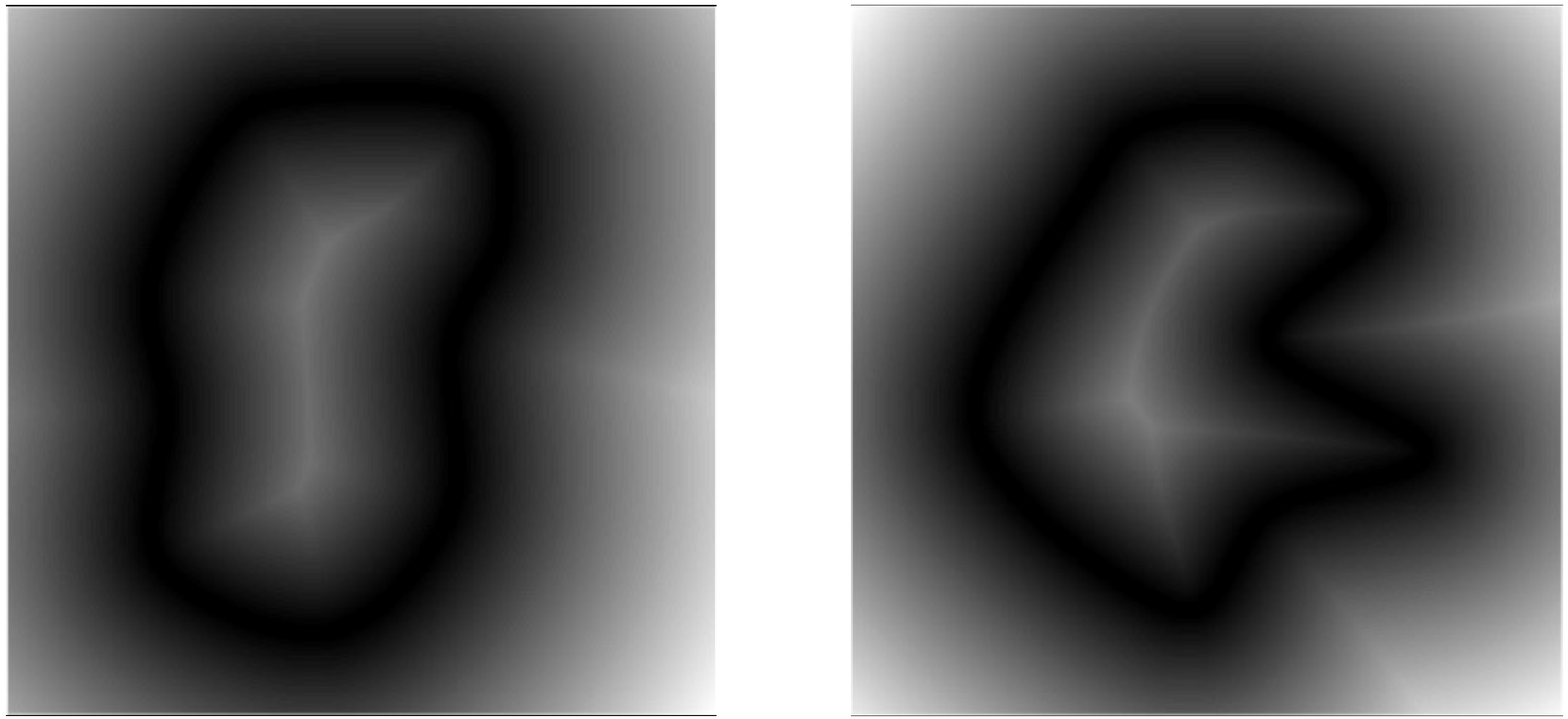
Taha and Hanbury 2015

Surface Distance



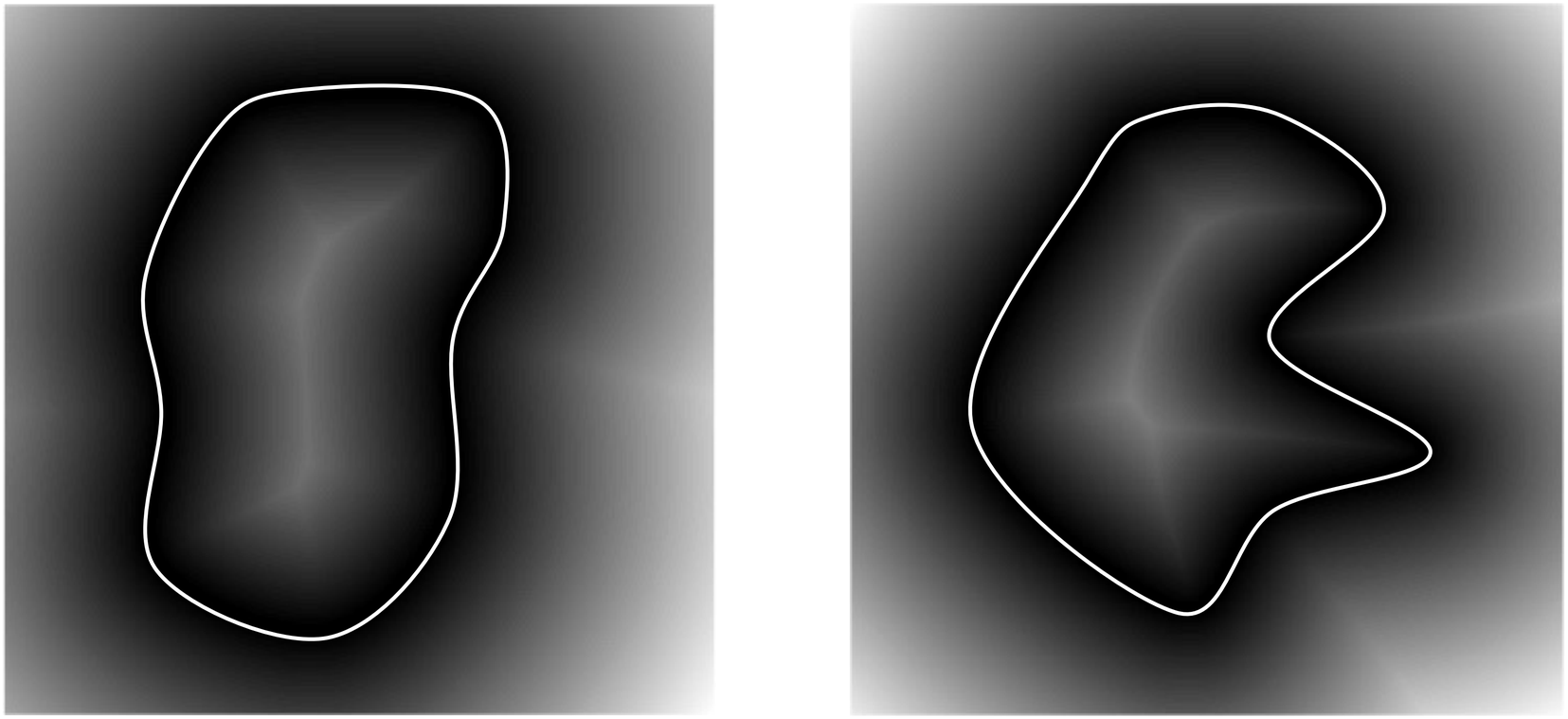


Surface Distance



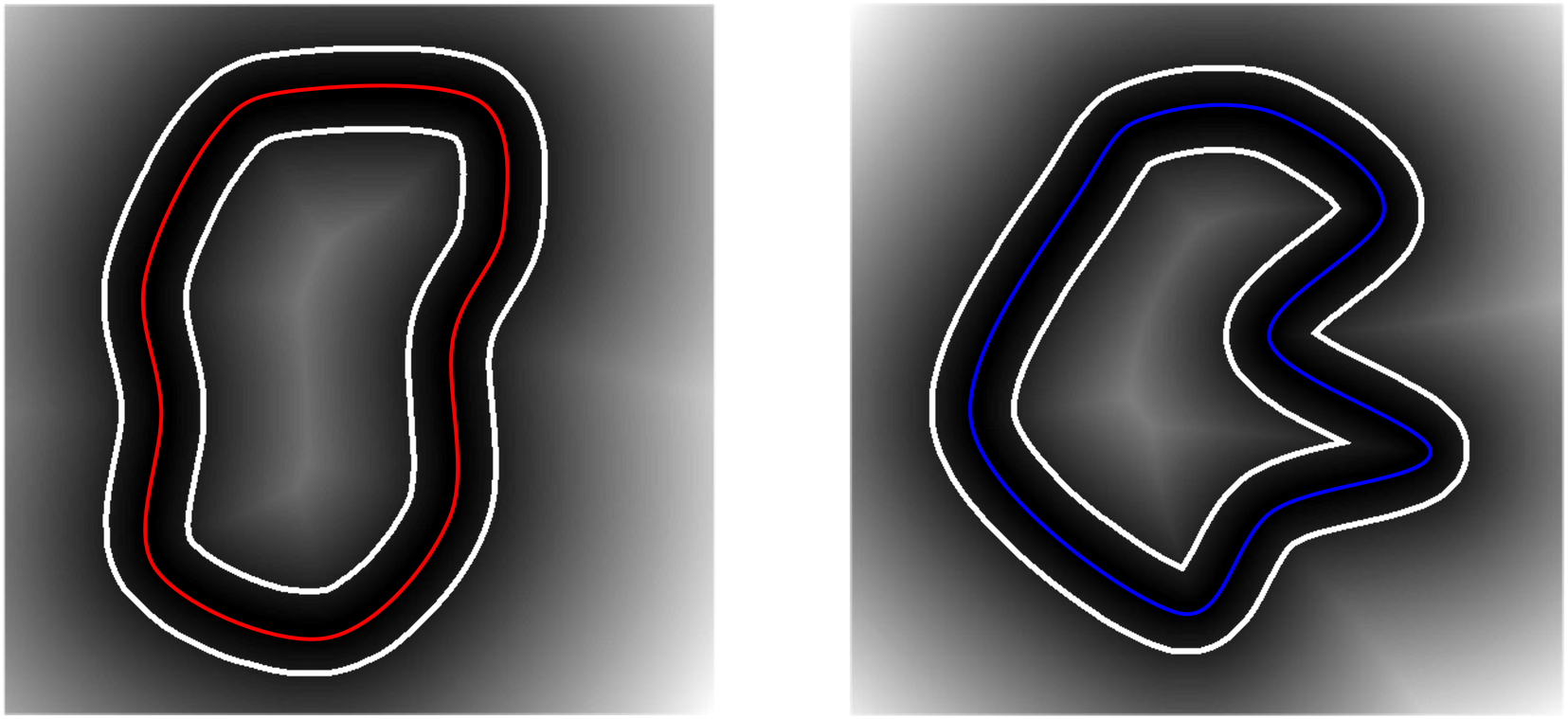
Euclidean Distance Maps

Surface Distance



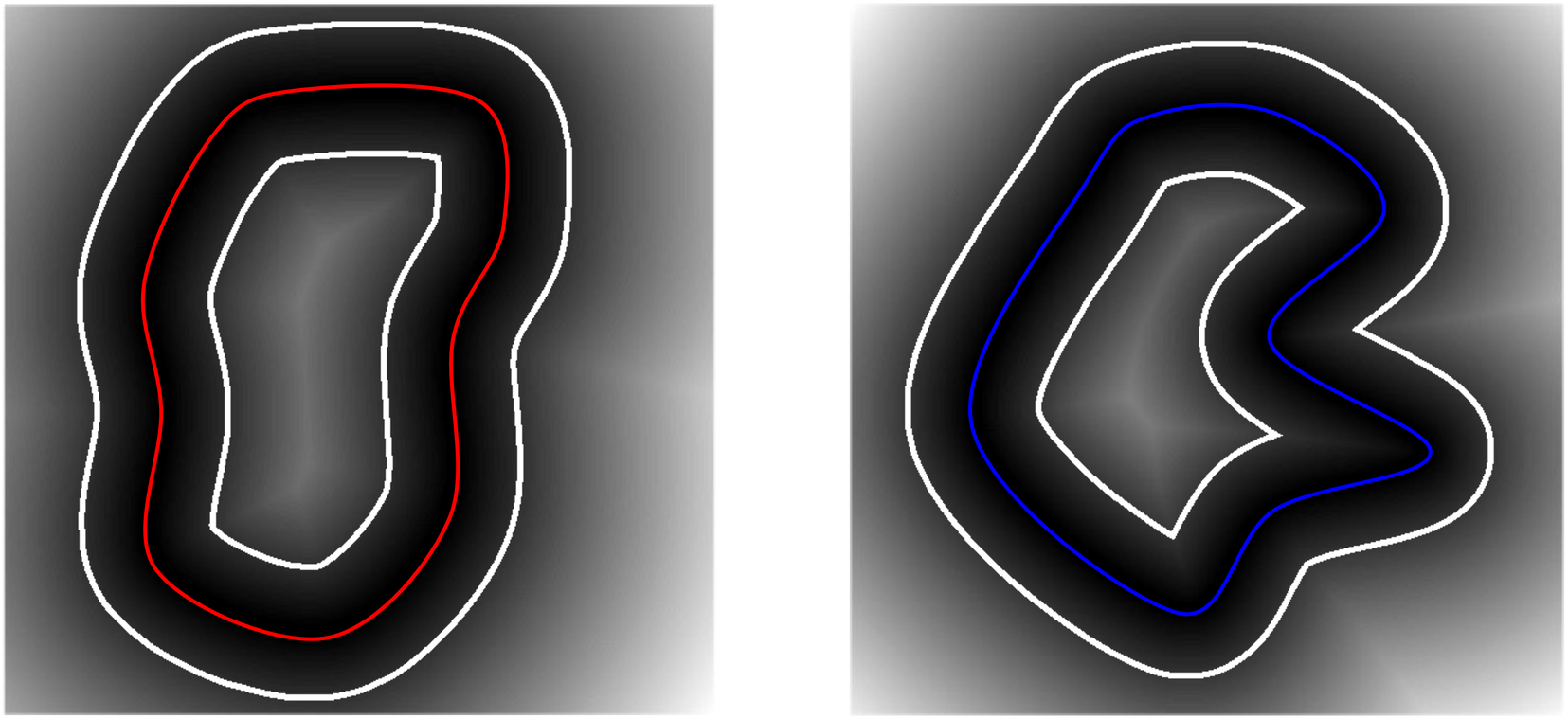
Zero-Level Surface

Surface Distance



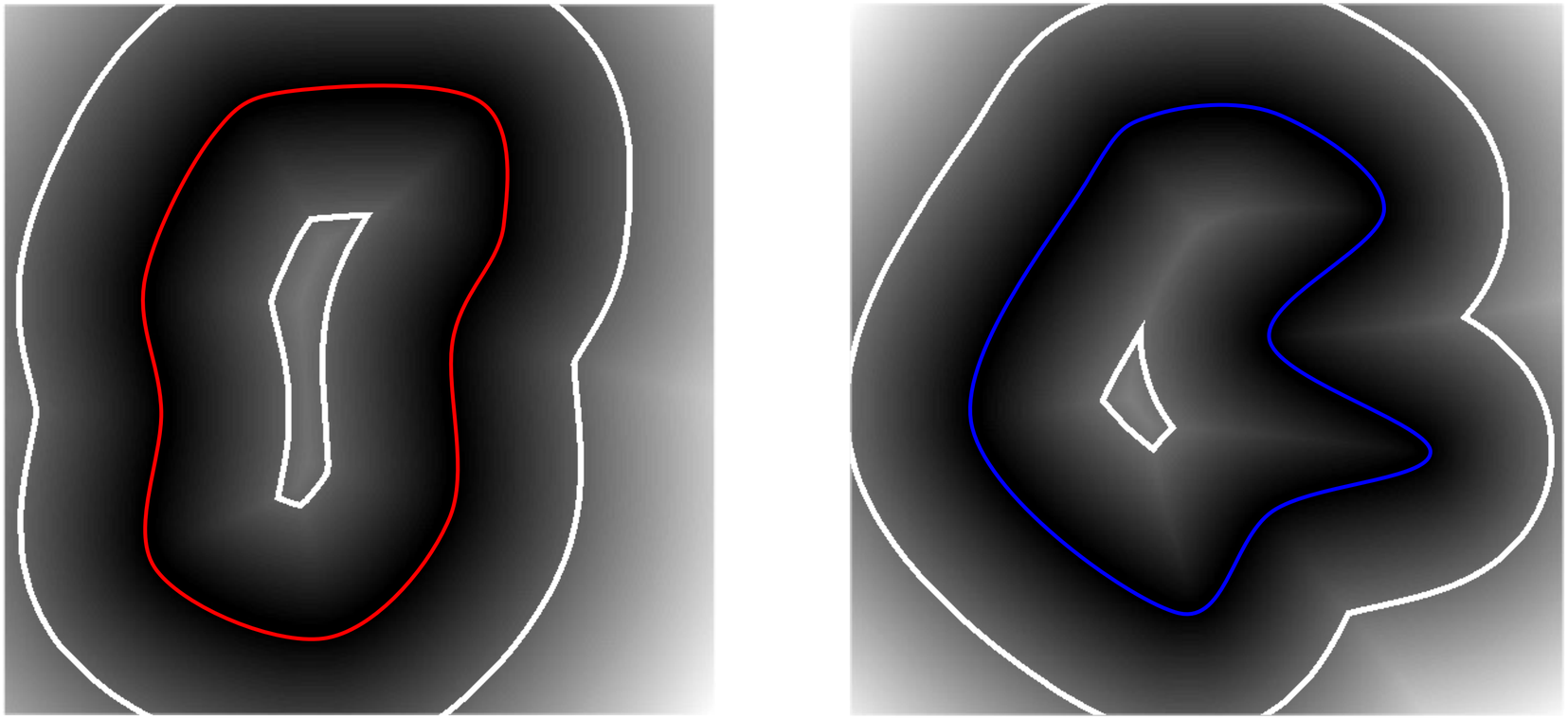
30-Pixels Distance Surface

Surface Distance



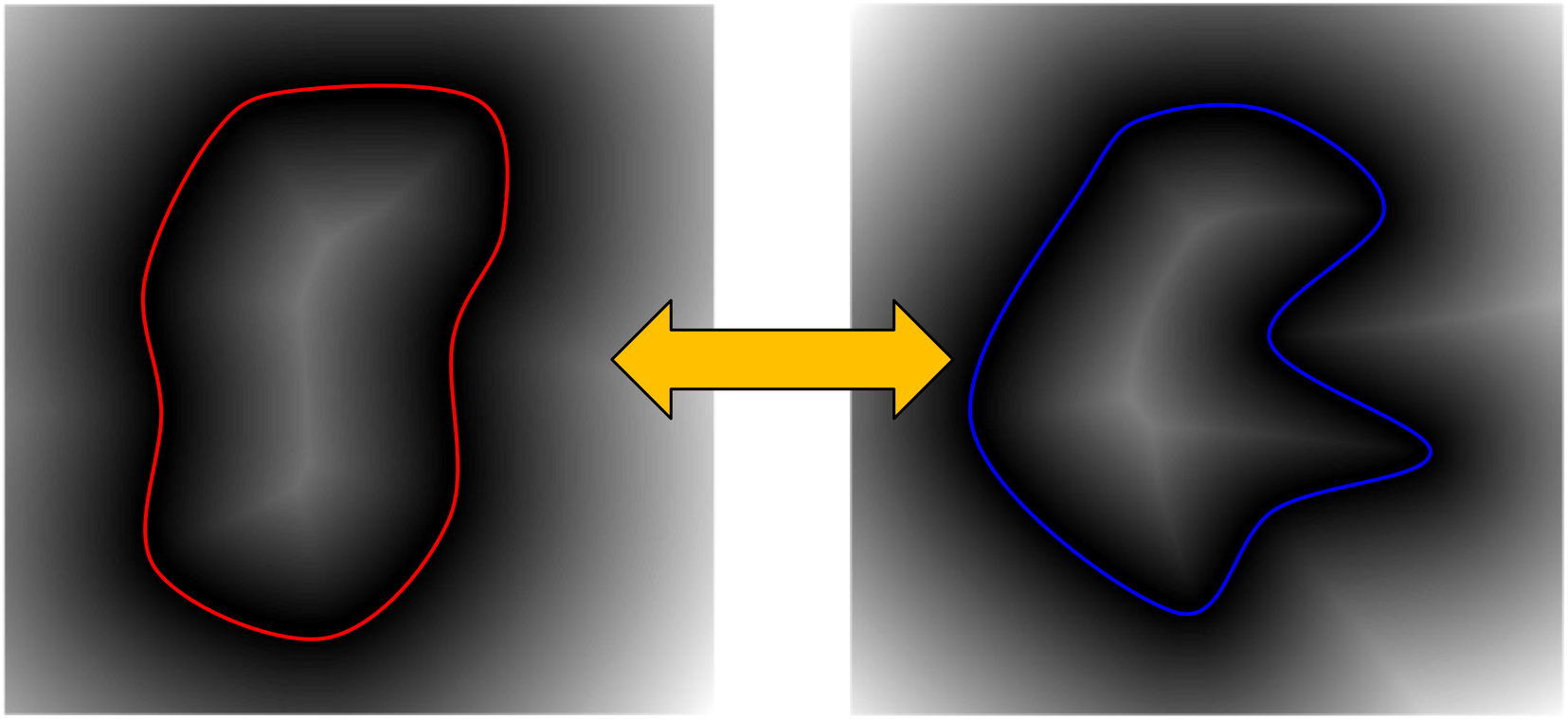
50-Pixels Distance Surface

Surface Distance



100-Pixels Distance Surface

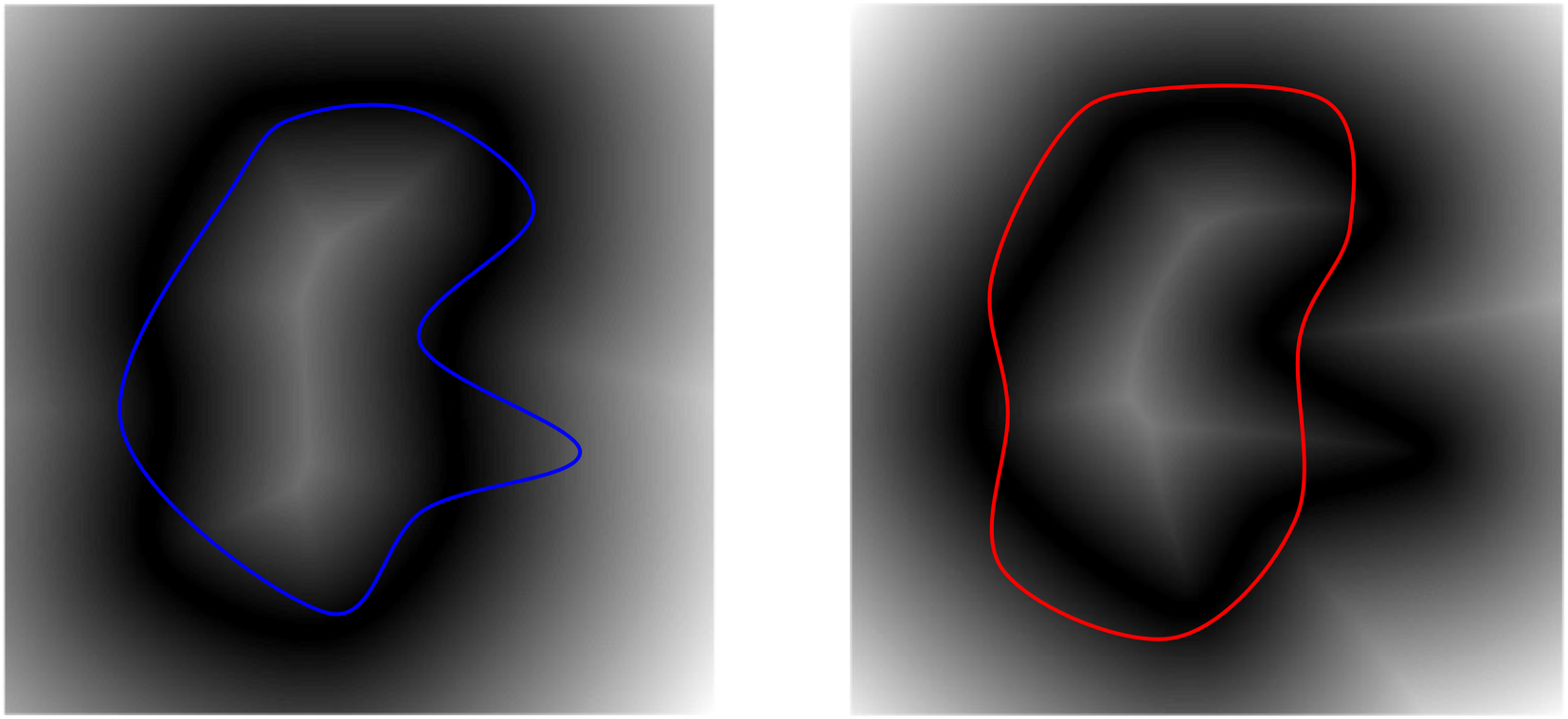
Surface Distance



Euclidean Distance Maps

Surface Distance

Sum up distances along pixels on the boundaries of one contour overlaid on the distance map of the other



Euclidean Distance Maps

Imperial College London

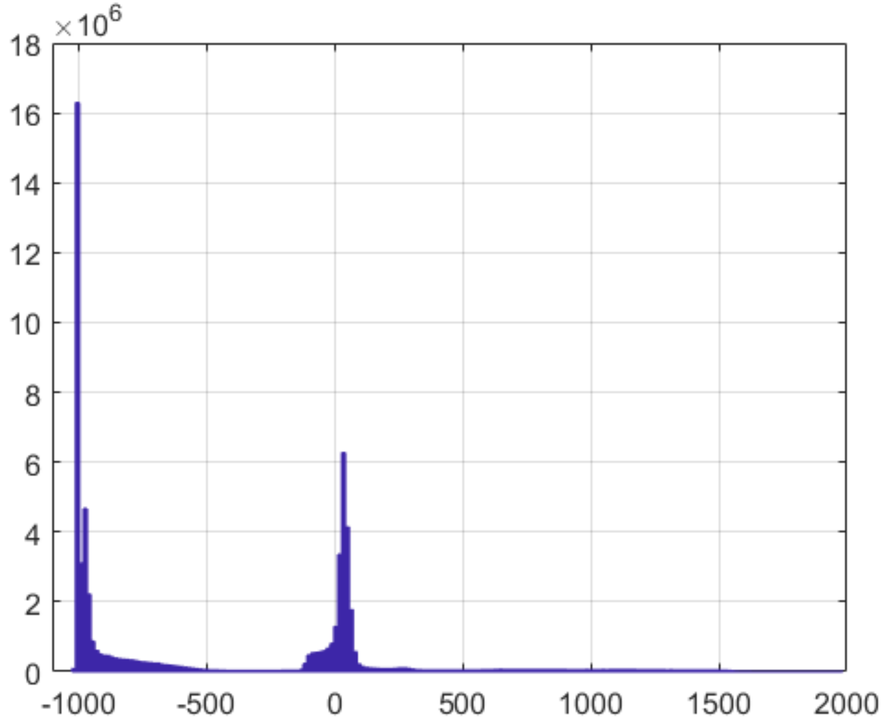
Segmentation Algorithms & Techniques

Segmentation Algorithms & Techniques

- Classic techniques (just a quick whistle-stop tour):
  - Intensity-based segmentation (e.g. thresholding)
  - Region-based segmentation (e.g. region growing)
  - Graph-based segmentation (e.g. graph cuts)
  - Atlas-based segmentation (using **image registration**)
- **Learning-based segmentation**
  - Random forests (only briefly, for completeness)
  - **Deep learning based methods (e.g. Convolutional neural networks)**

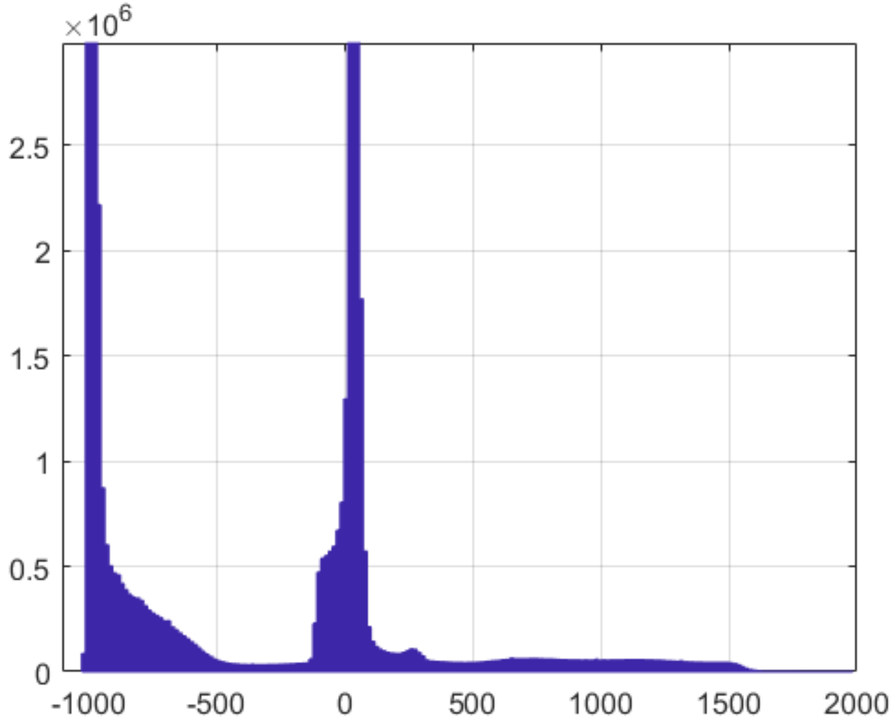
Simple Thresholding

- Select a threshold on the intensity range



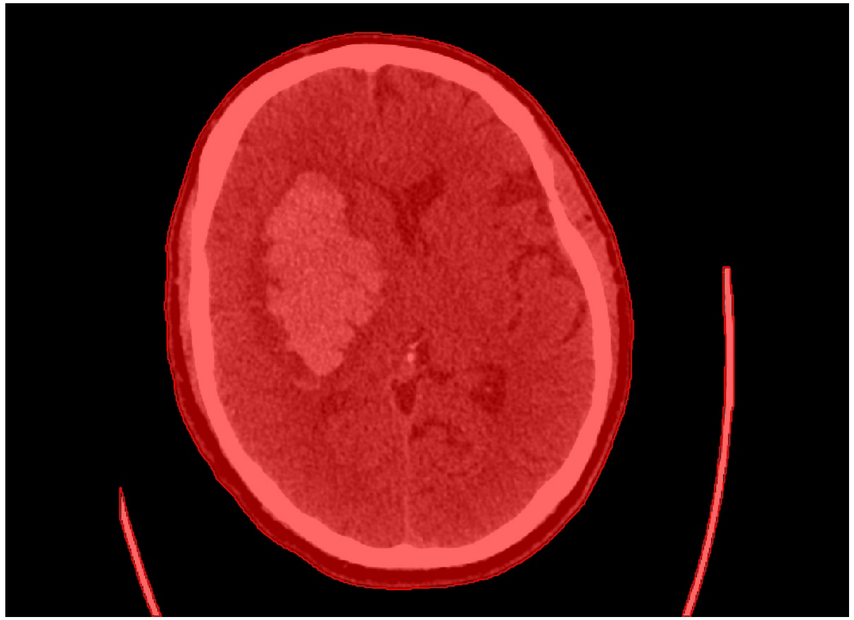
Simple Thresholding

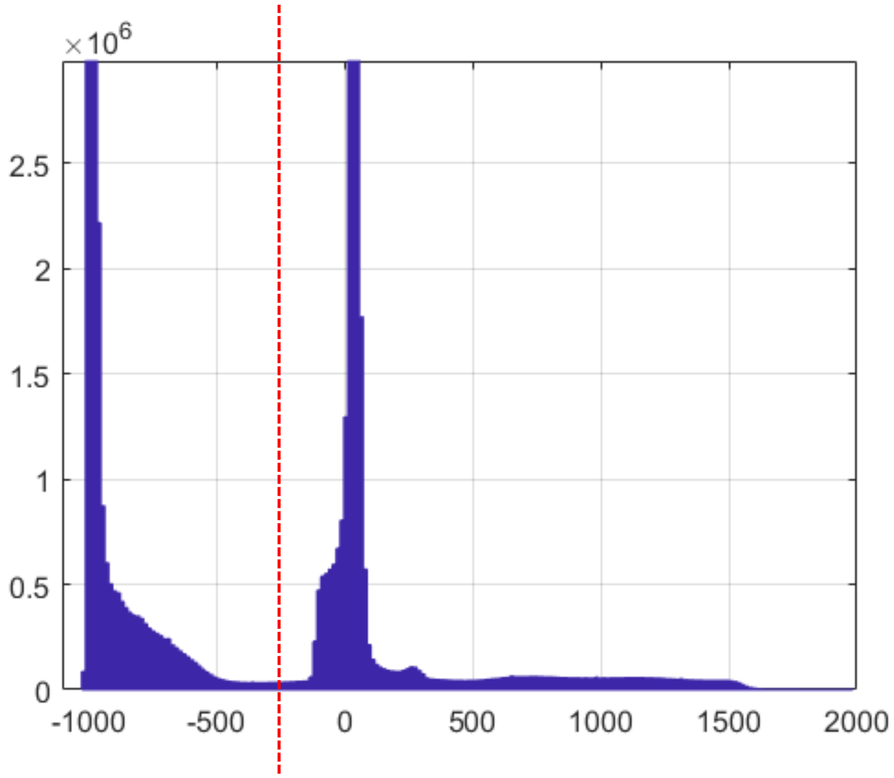
- Select a threshold on the intensity range



Simple Thresholding

- Select a threshold on the intensity range

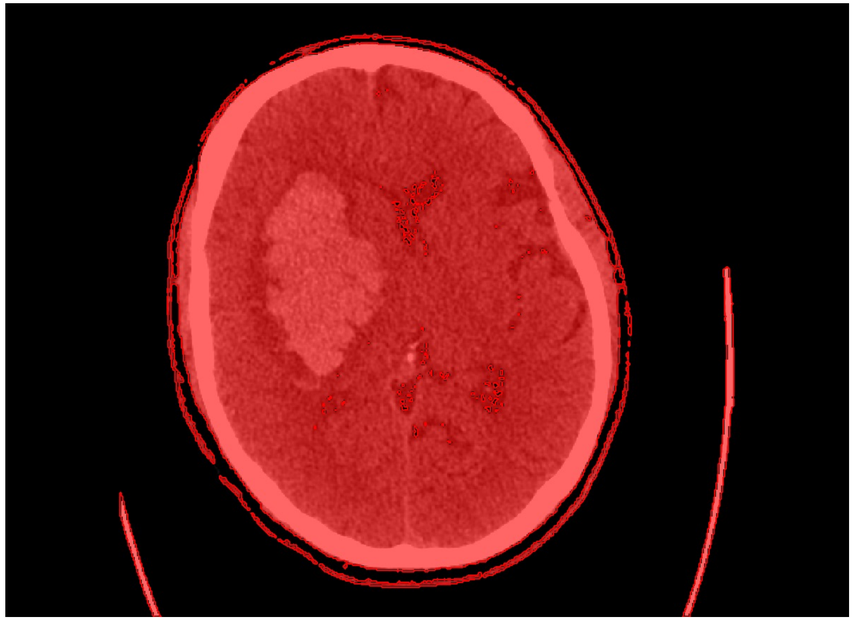


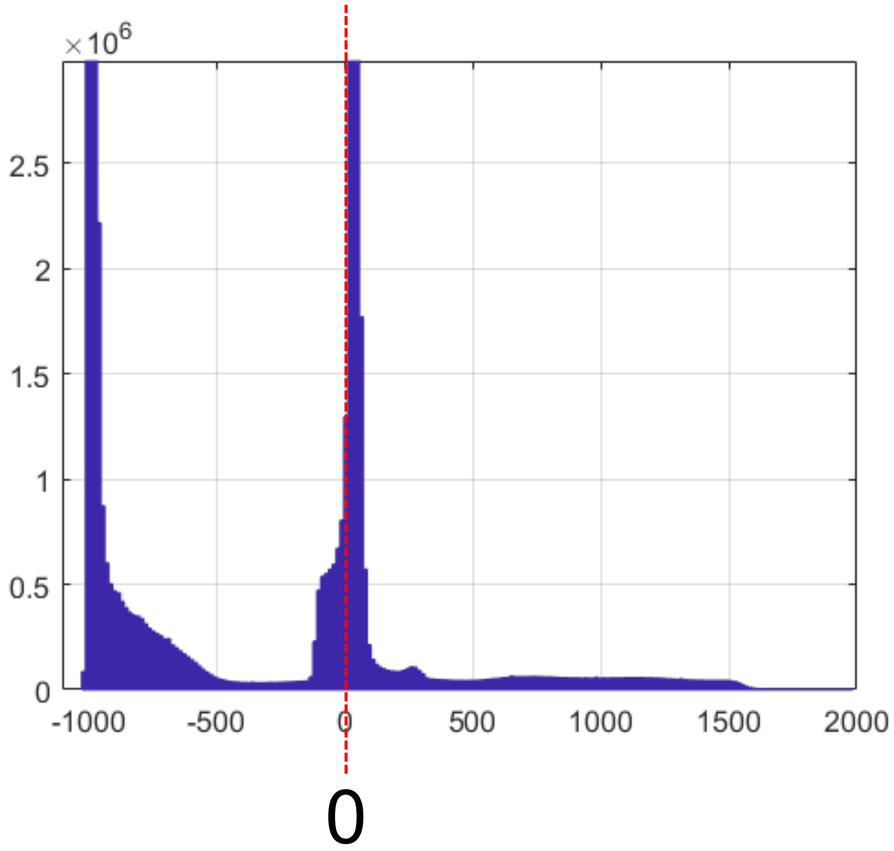


-250

Simple Thresholding

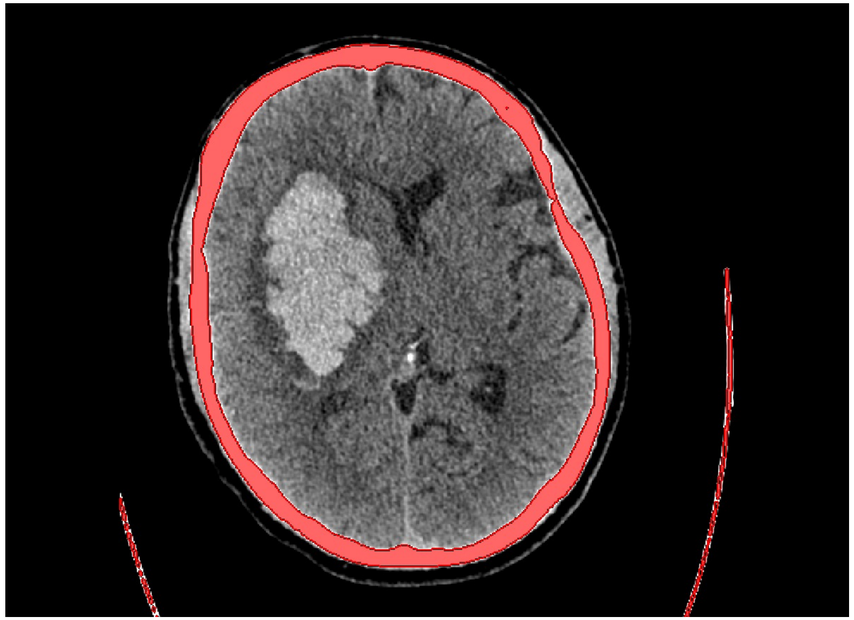
- Select a threshold on the intensity range

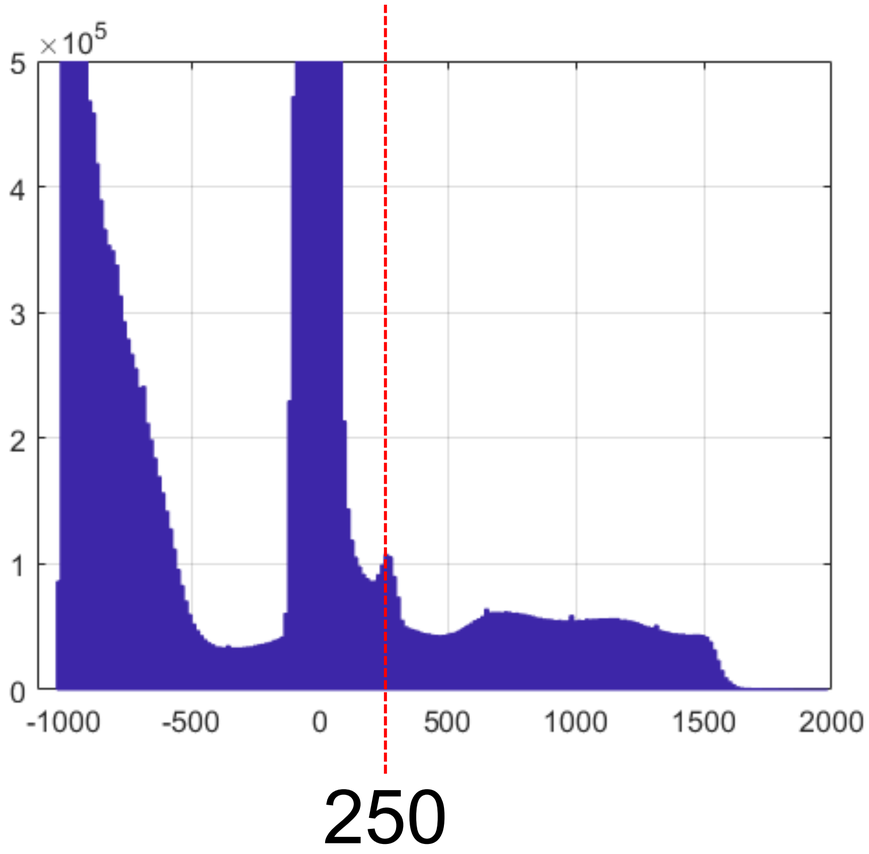




Simple Thresholding

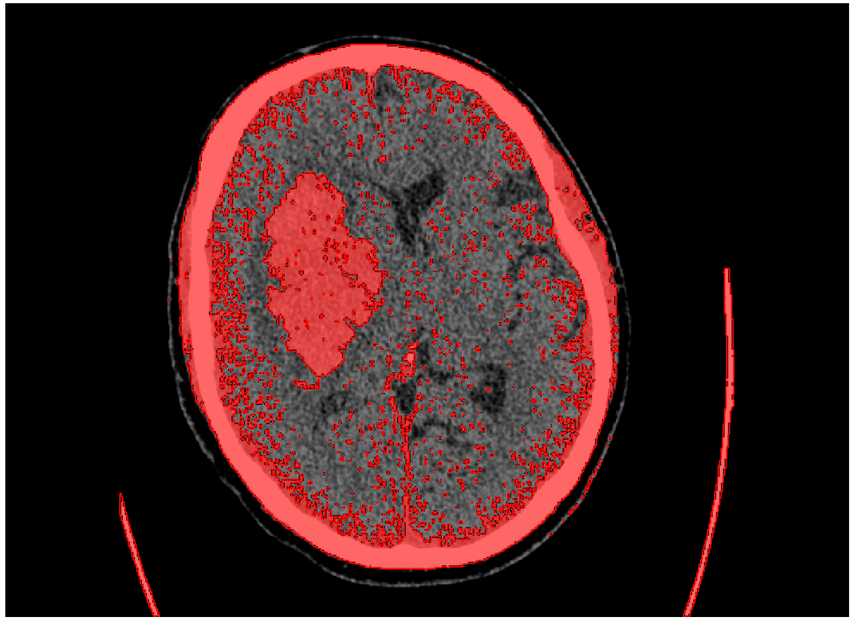
- Select a threshold on the intensity range

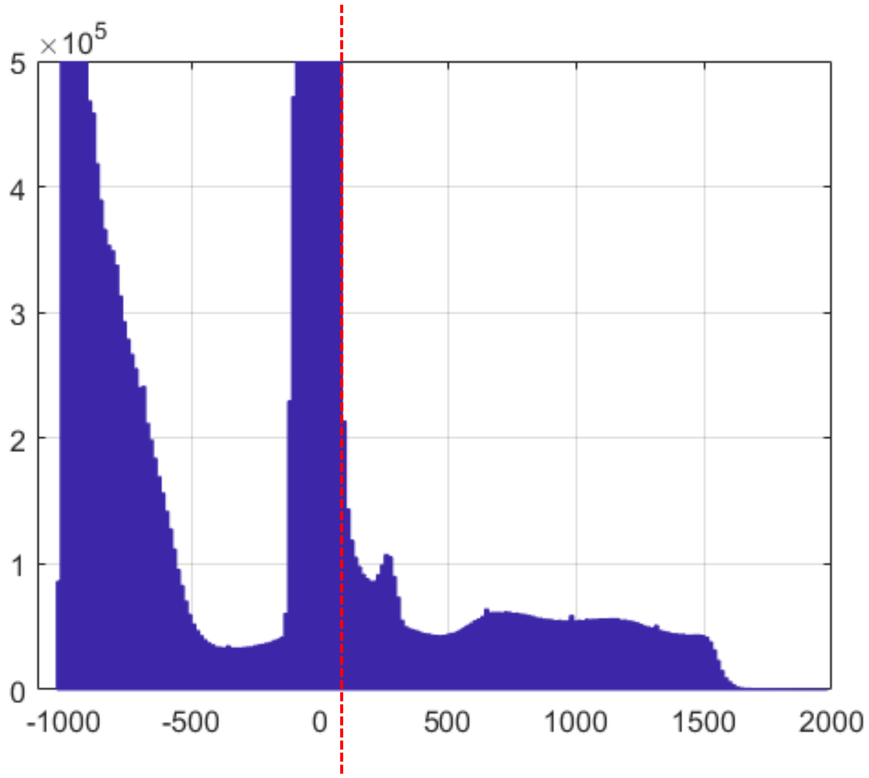




Simple Thresholding

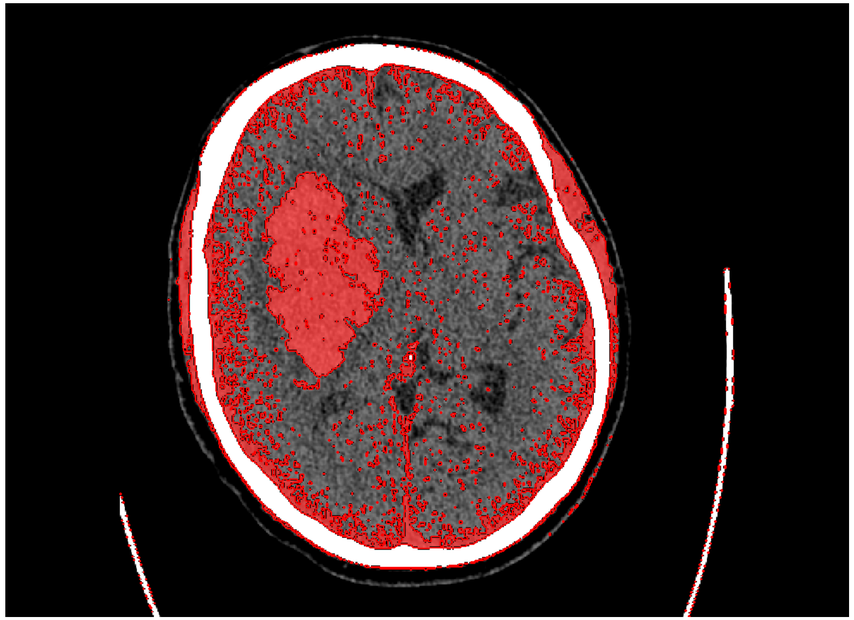
- Select a threshold on the intensity range

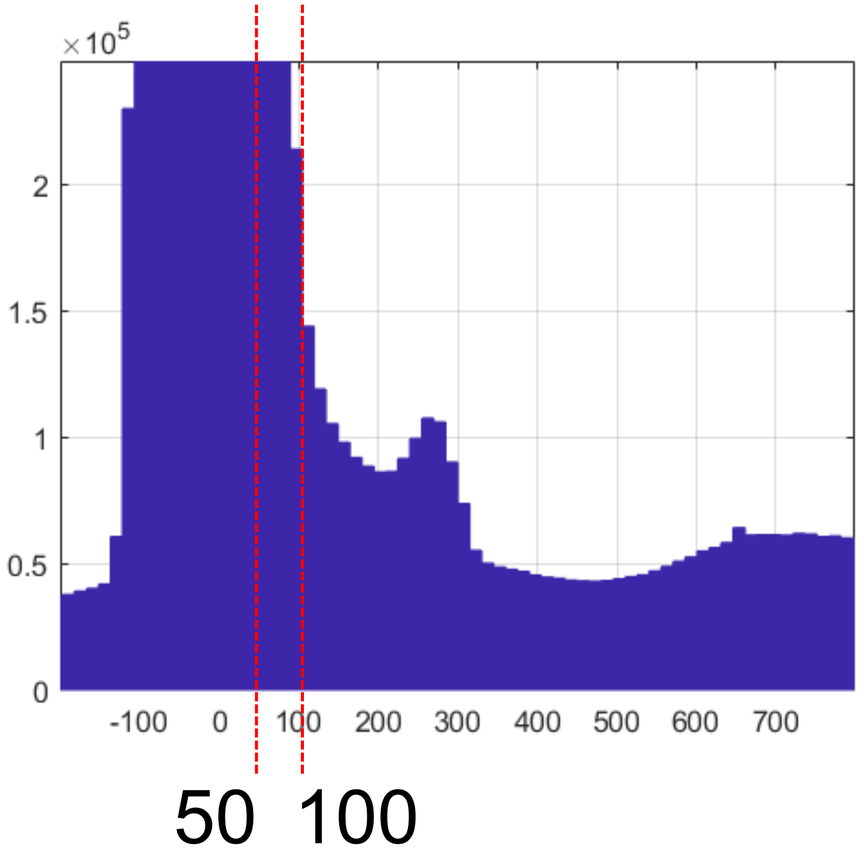




UL Thresholding

- Select a lower and upper threshold





Thresholding

Advantages

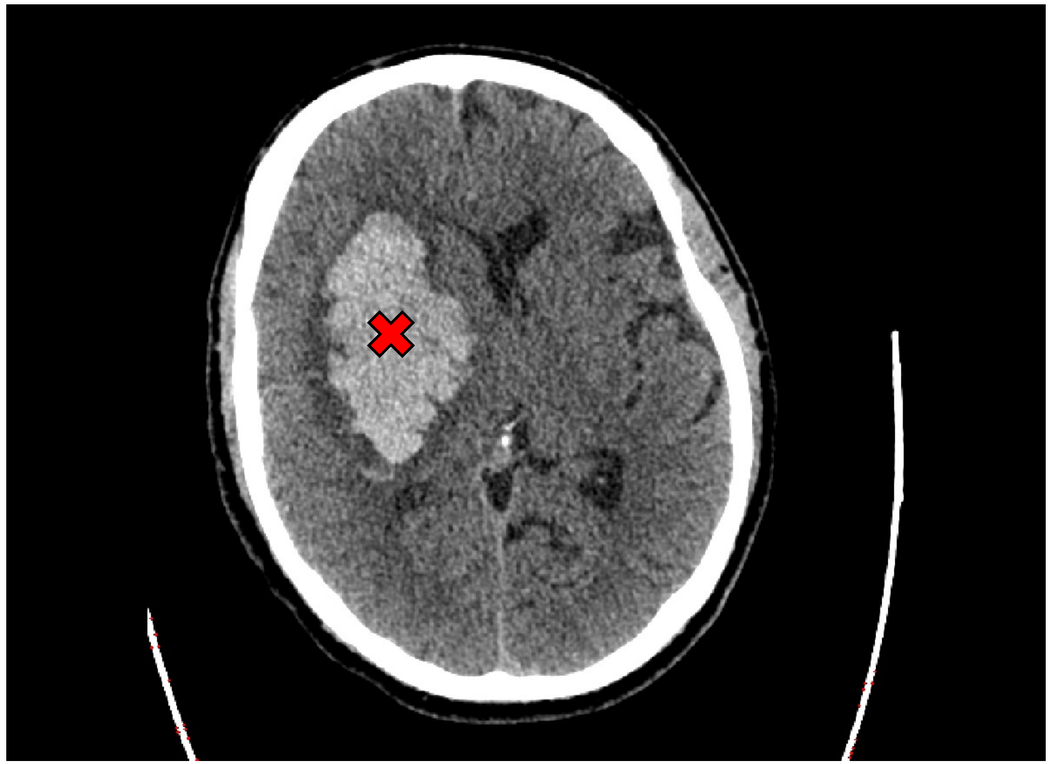
- simple
- fast

Disadvantages

- regions must be homogeneous and distinct
- difficulty in finding consistent thresholds across images
- leakages, isolated pixels and ‘rough’ boundaries likely

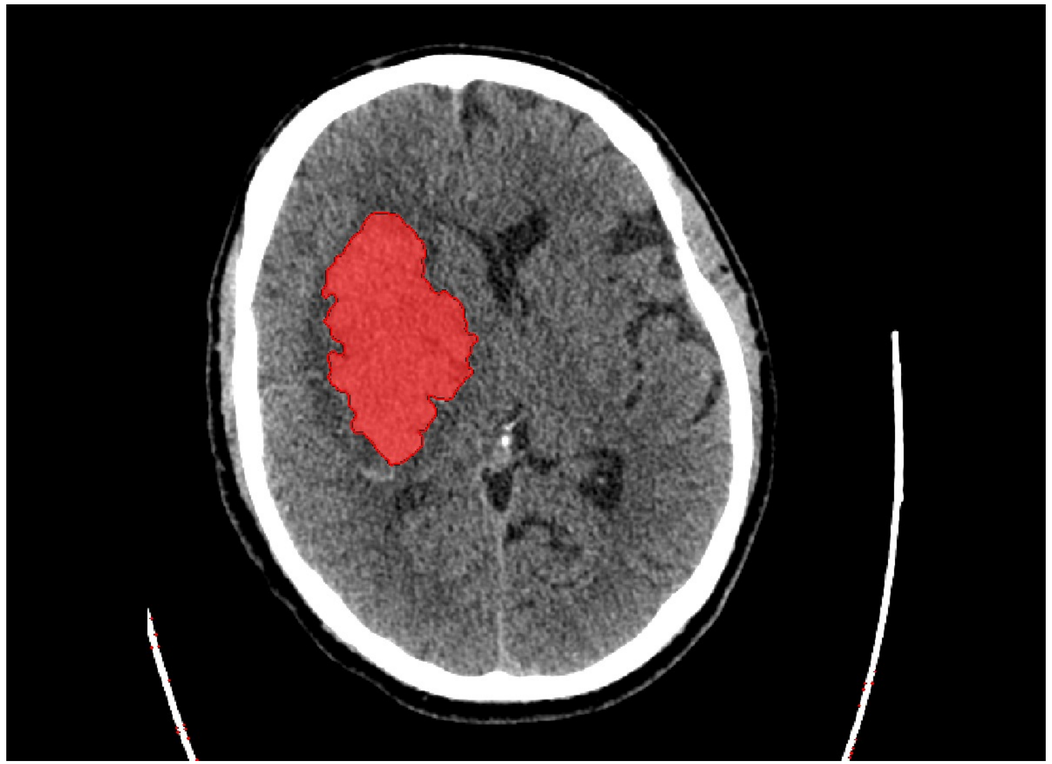
Region Growing

- Start from (user selected) seed point(s), and grow a region according to an intensity threshold



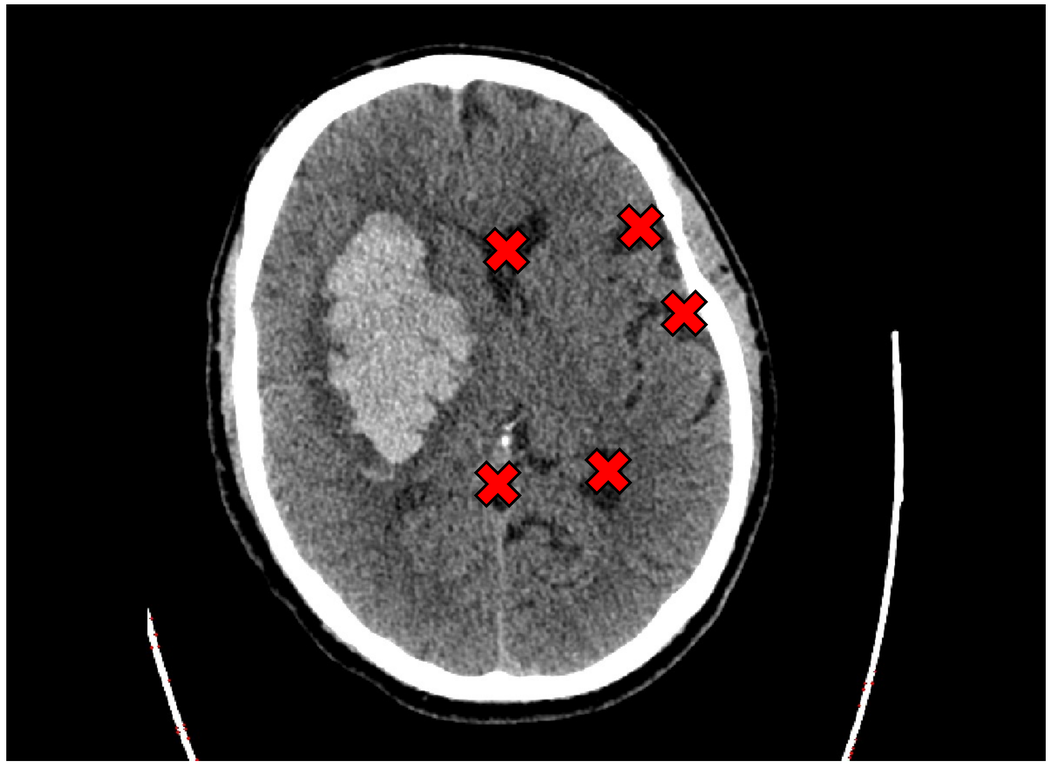
Region Growing

- Start from (user selected) seed point(s), and grow a region according to an intensity threshold



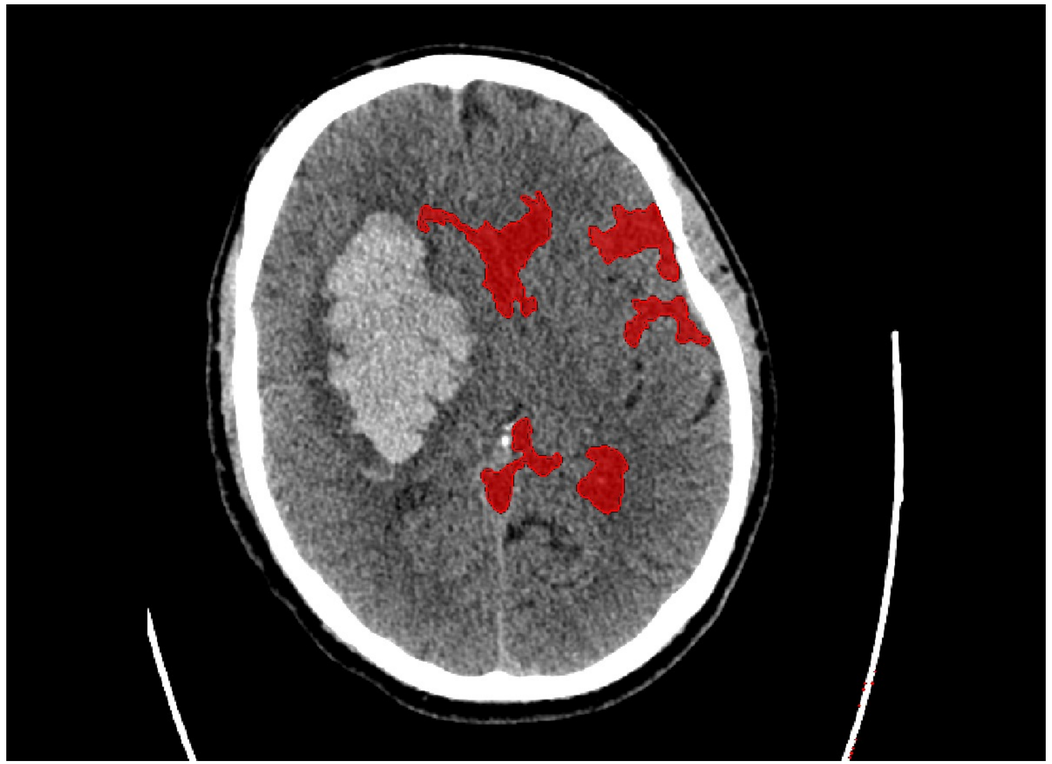
Region Growing

- Start from (user selected) seed point(s), and grow a region according to an intensity threshold



Region Growing

- Start from (user selected) seed point(s), and grow a region according to an intensity threshold

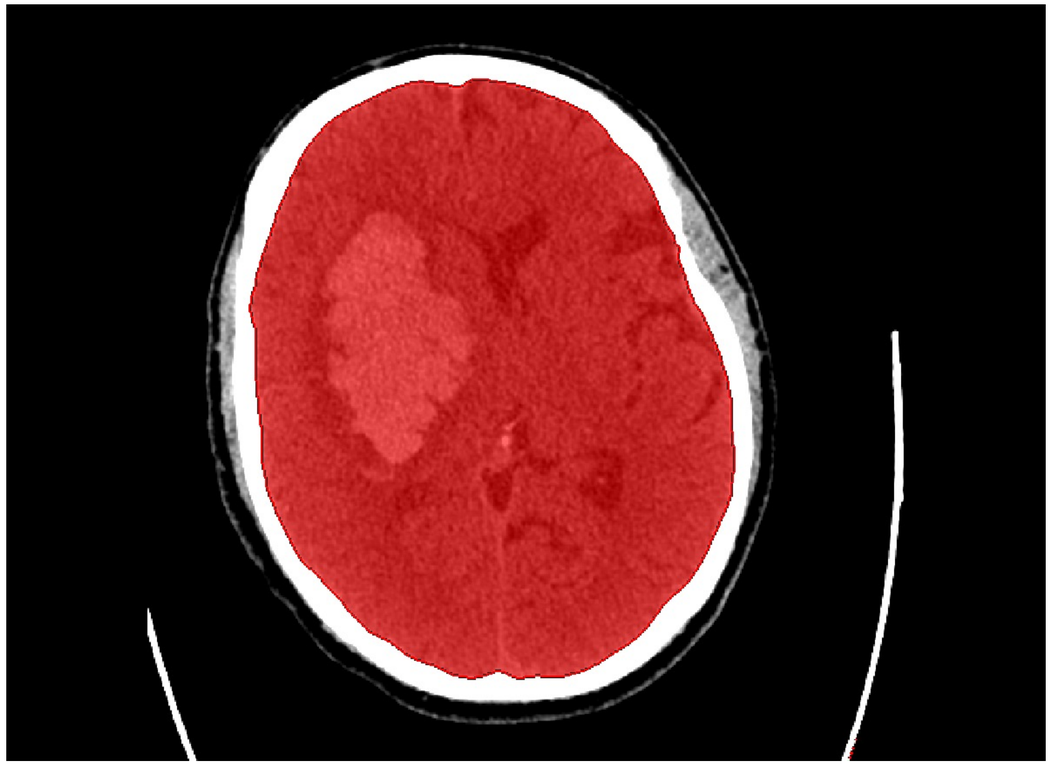


Region Growing

- Start from (user selected) seed point(s), and grow a region according to an intensity threshold

Region Growing

- Start from (user selected) seed point(s), and grow a region according to an intensity threshold



Region Growing

Advantages

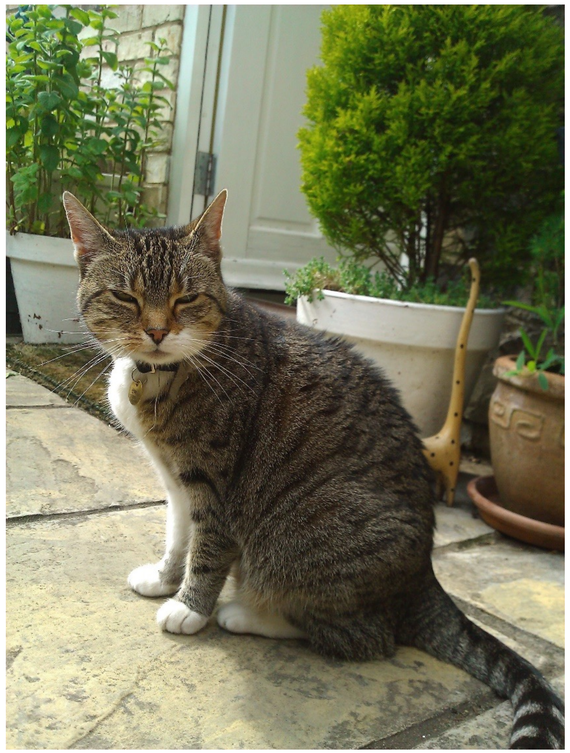
- relatively fast
- yields connected region (from a seed point)

Disadvantages

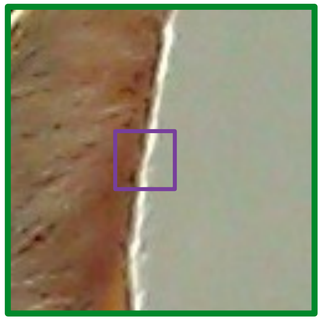
- regions must be homogeneous
- leakages and ‘rough’ boundaries likely
- requires (user) input for seed points

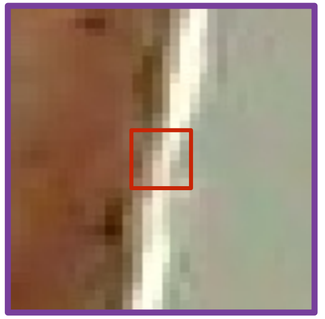
Graph Cuts

- Segmentation based on max-flow/min-cut algorithm

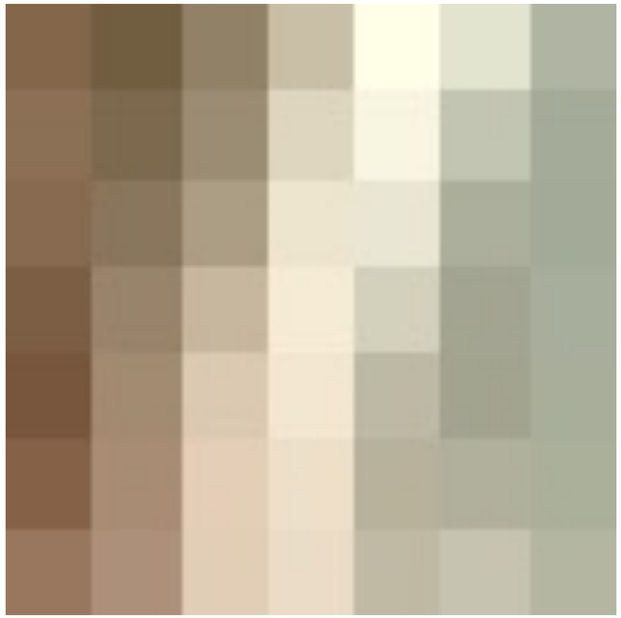


Graph Cuts

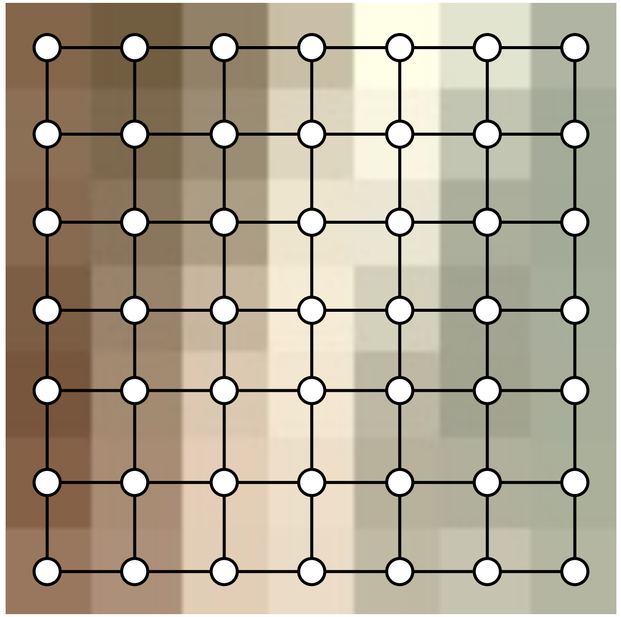




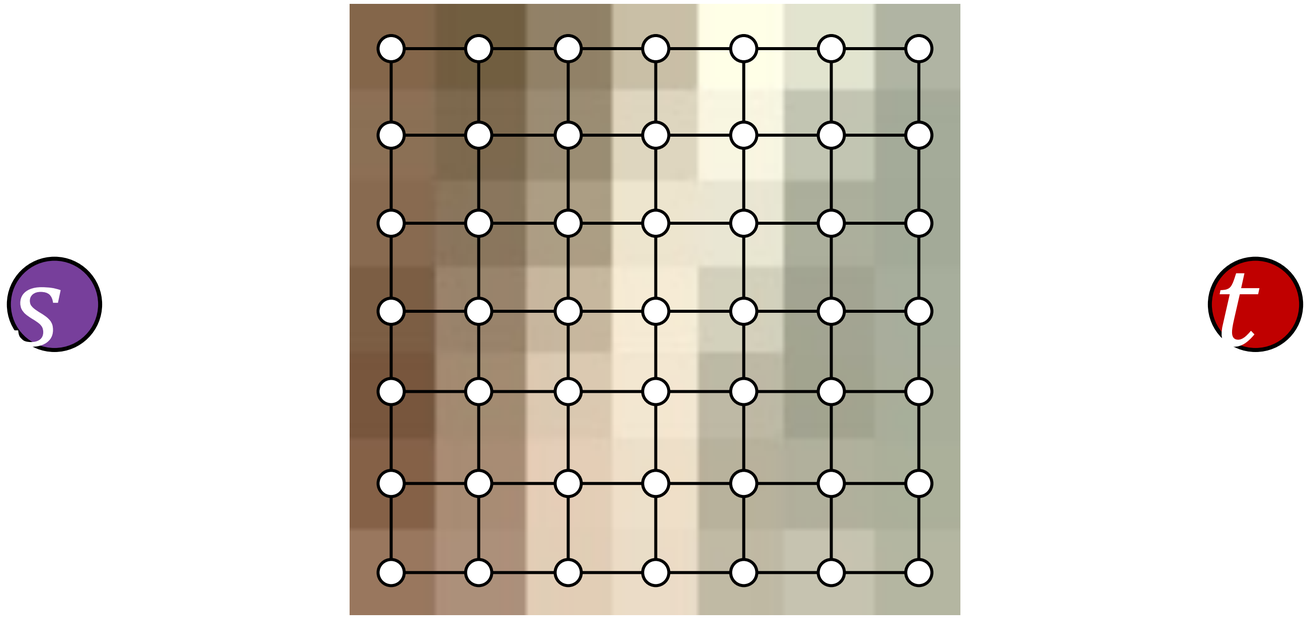
Graph Cuts



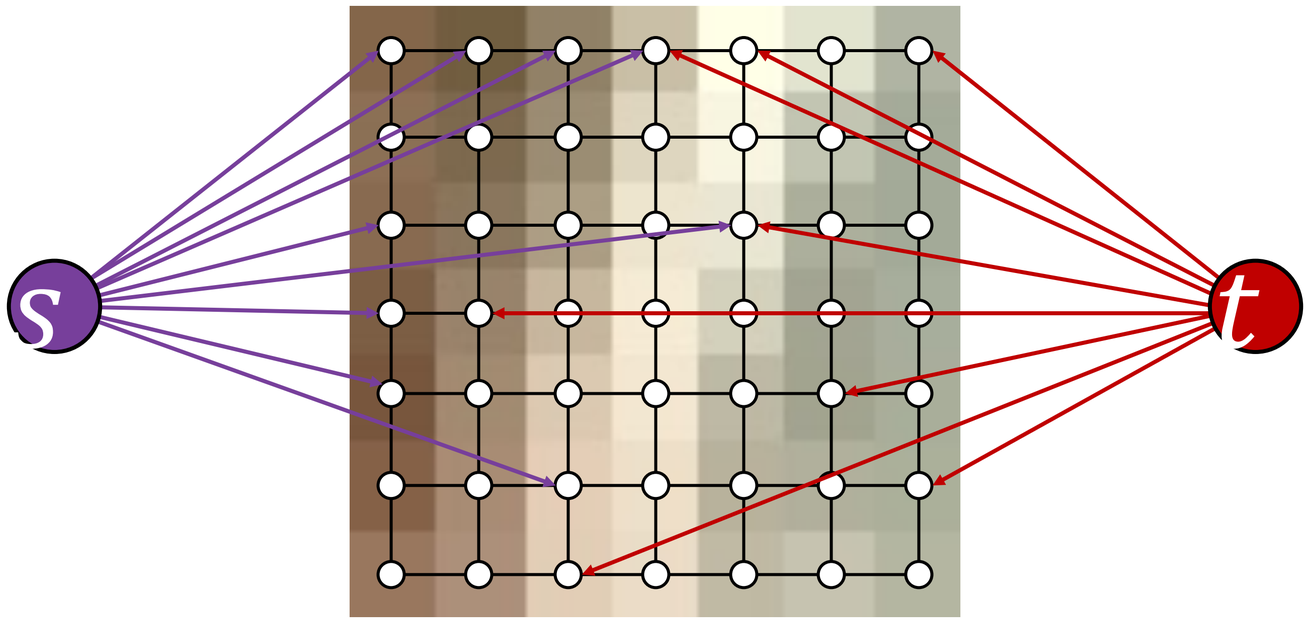
Graph Cuts



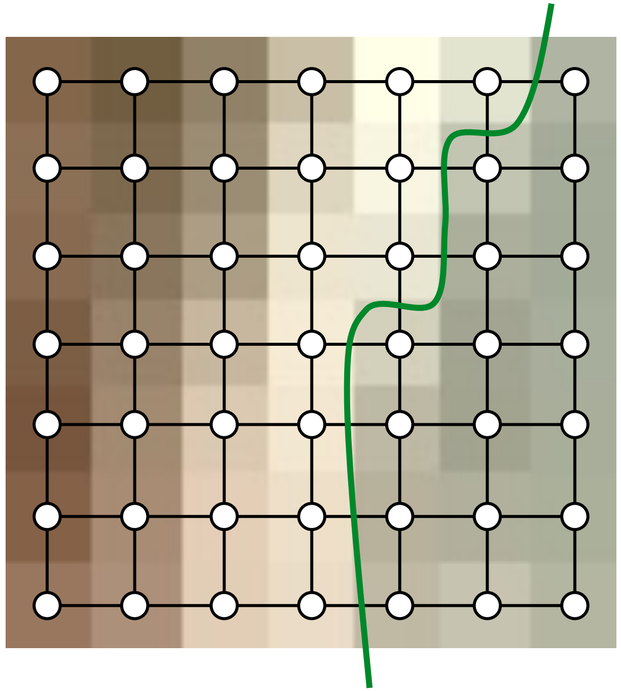
Graph Cuts



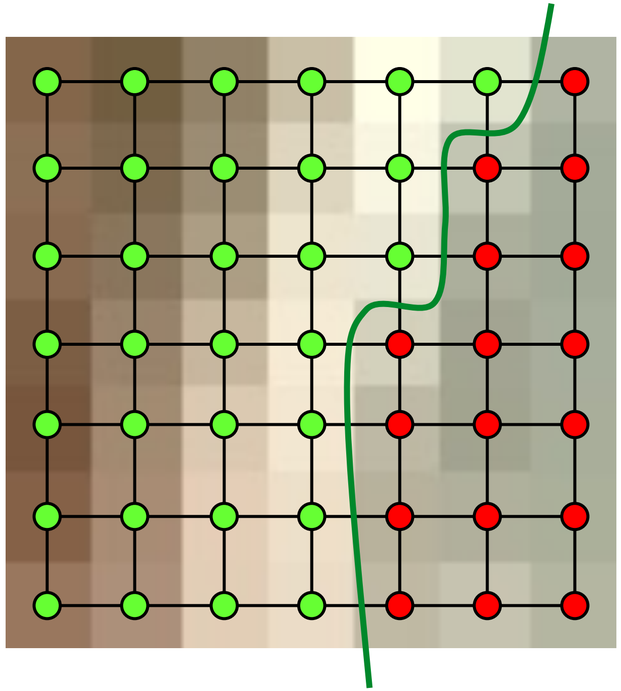
Graph Cuts



Graph Cuts

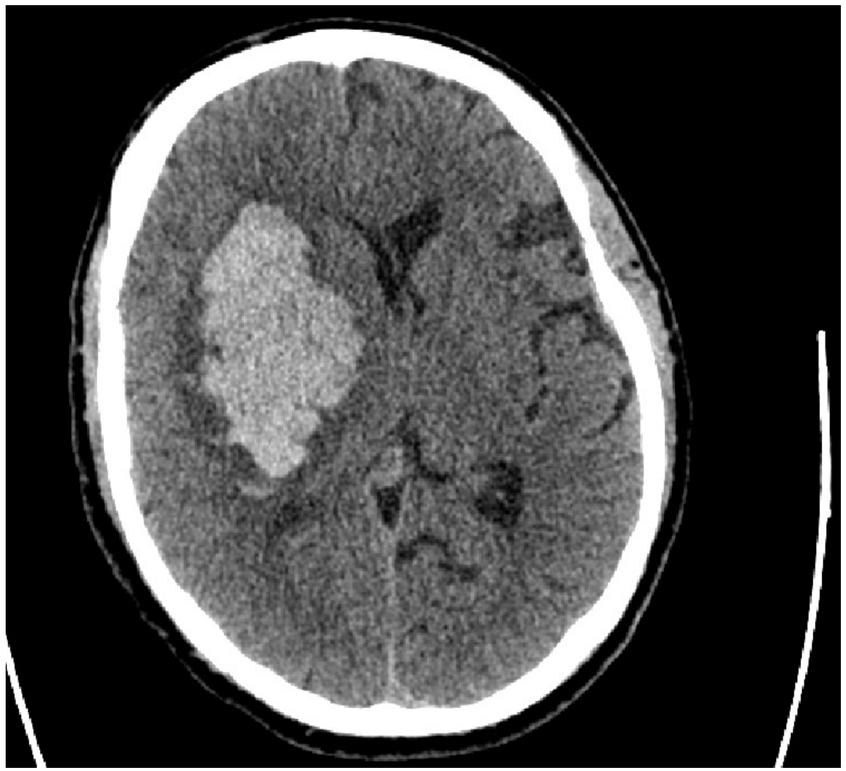


Graph Cuts



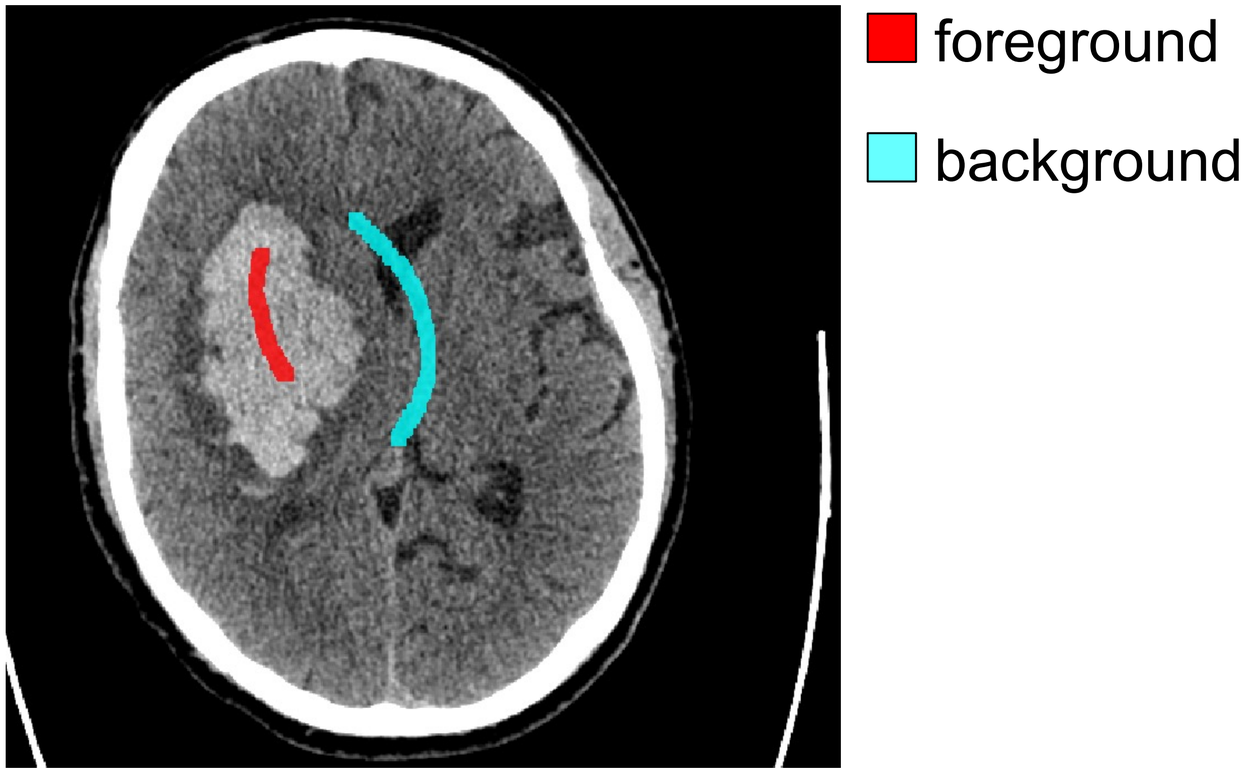
Graph Cuts – Example

- Input: (user) seed brushes



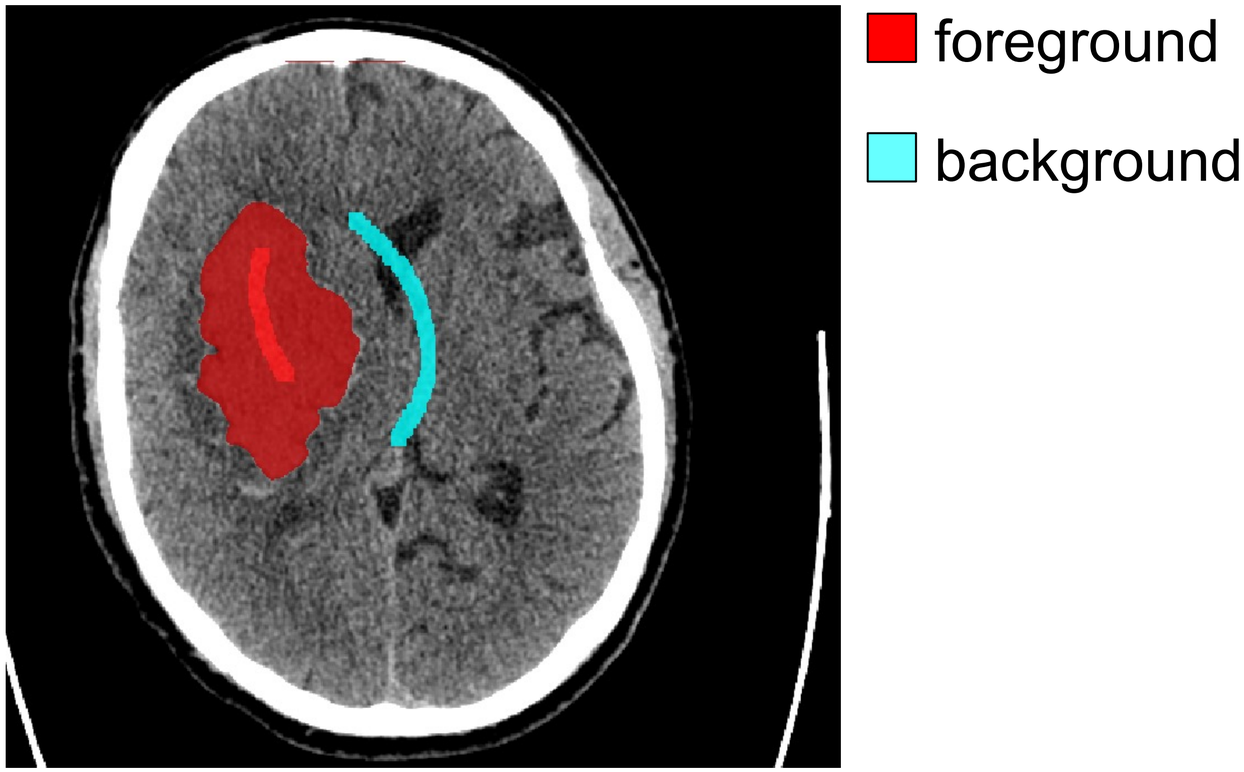
Graph Cuts – Example

- Input: (user) seed brushes



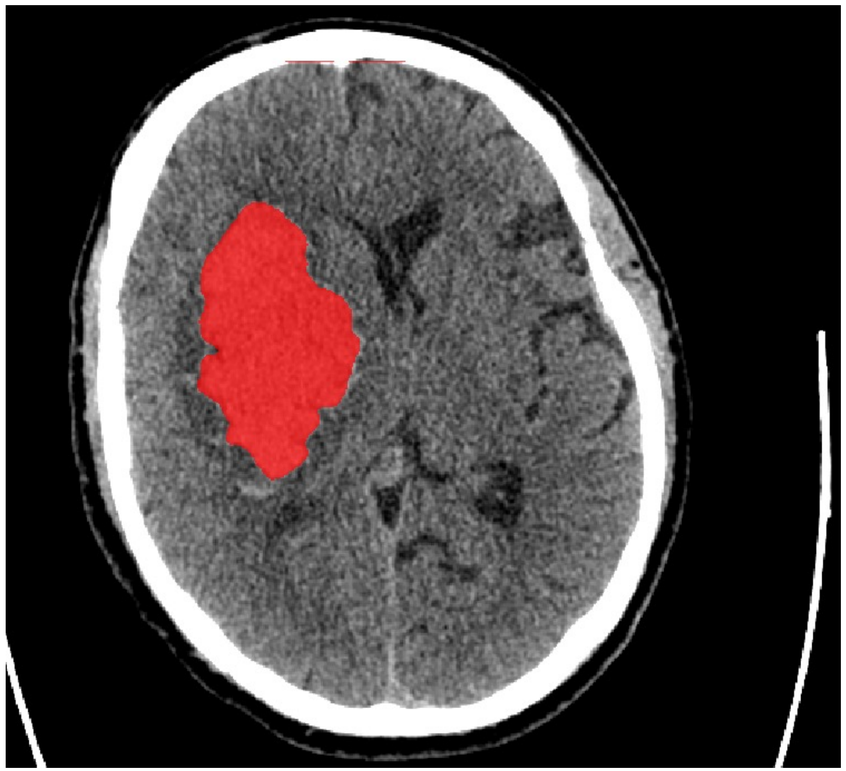
Graph Cuts – Example

- Computation of graph cut



Graph Cuts – Example

- Output: segmentation

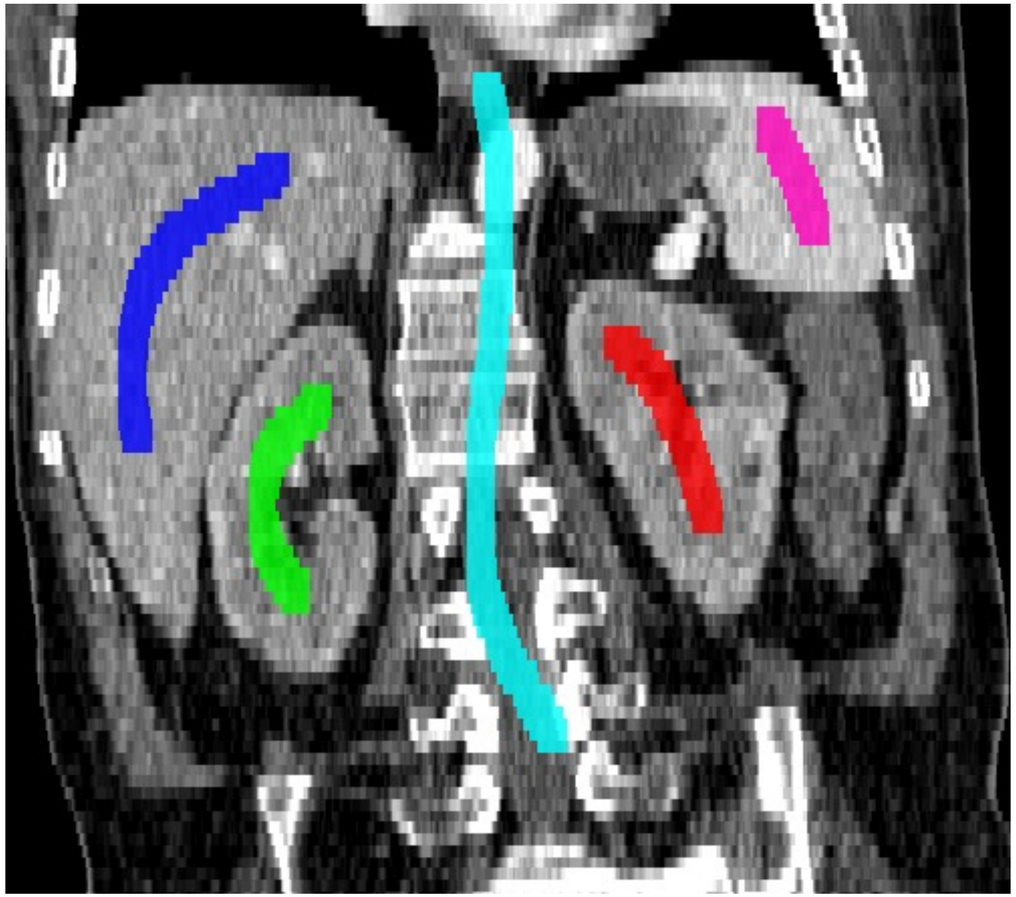


Multi-Label Graph Cuts – Example

- Segment multiple organs simultaneously

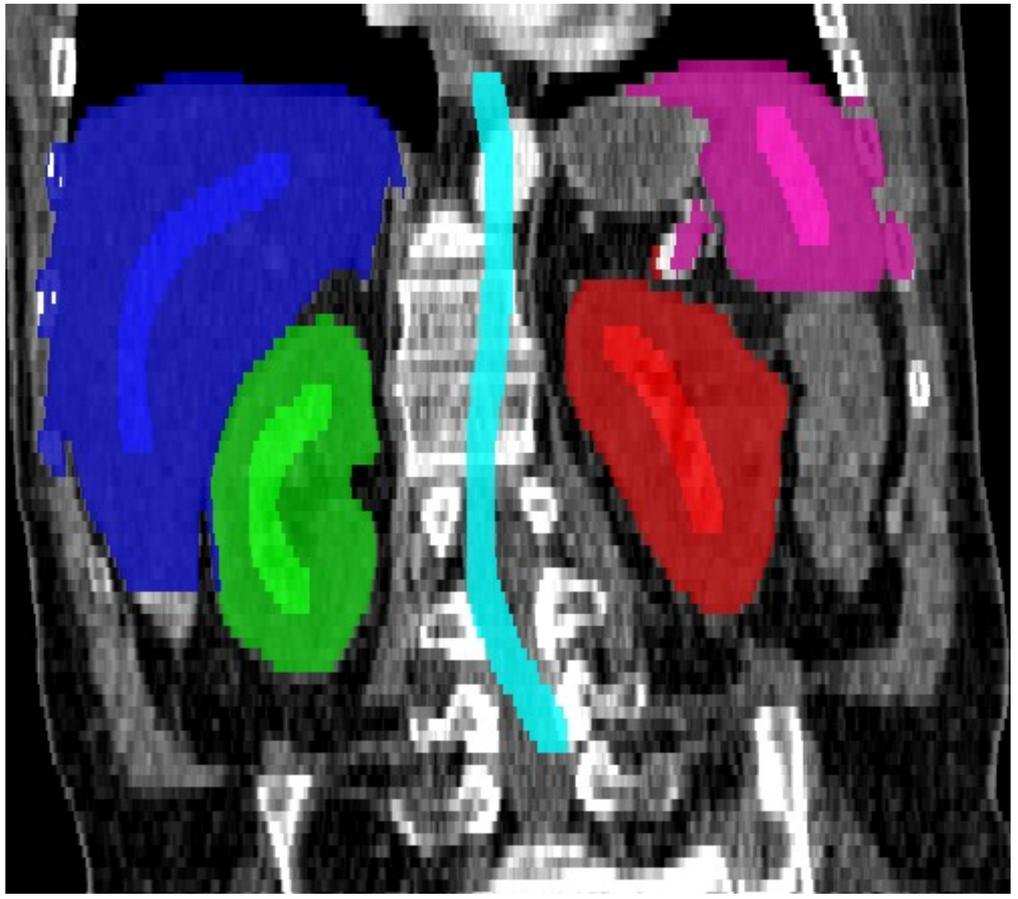
Multi-Label Graph Cuts – Example

- Segment multiple organs simultaneously



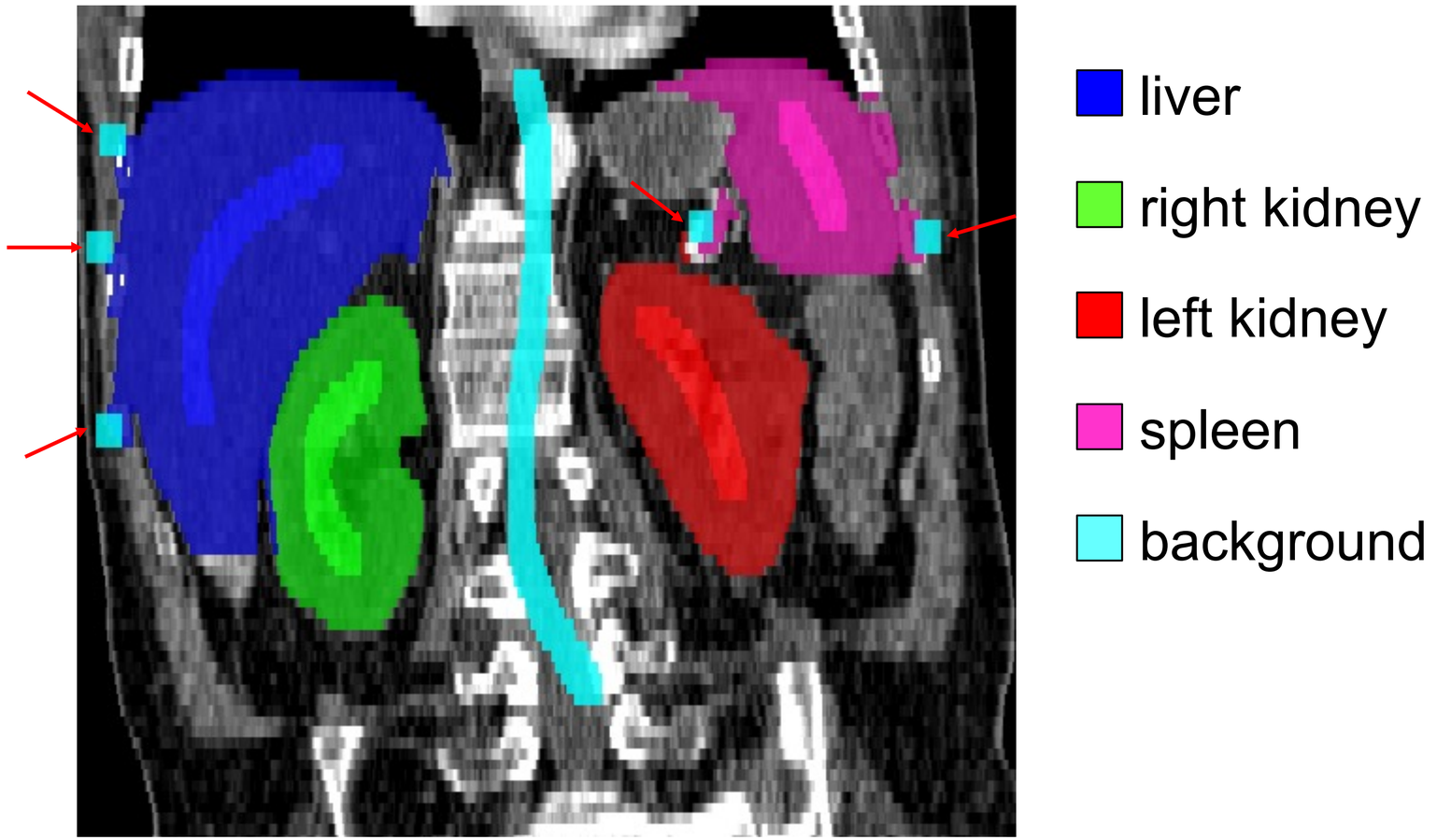
Multi-Label Graph Cuts – Example

- Interactive corrections



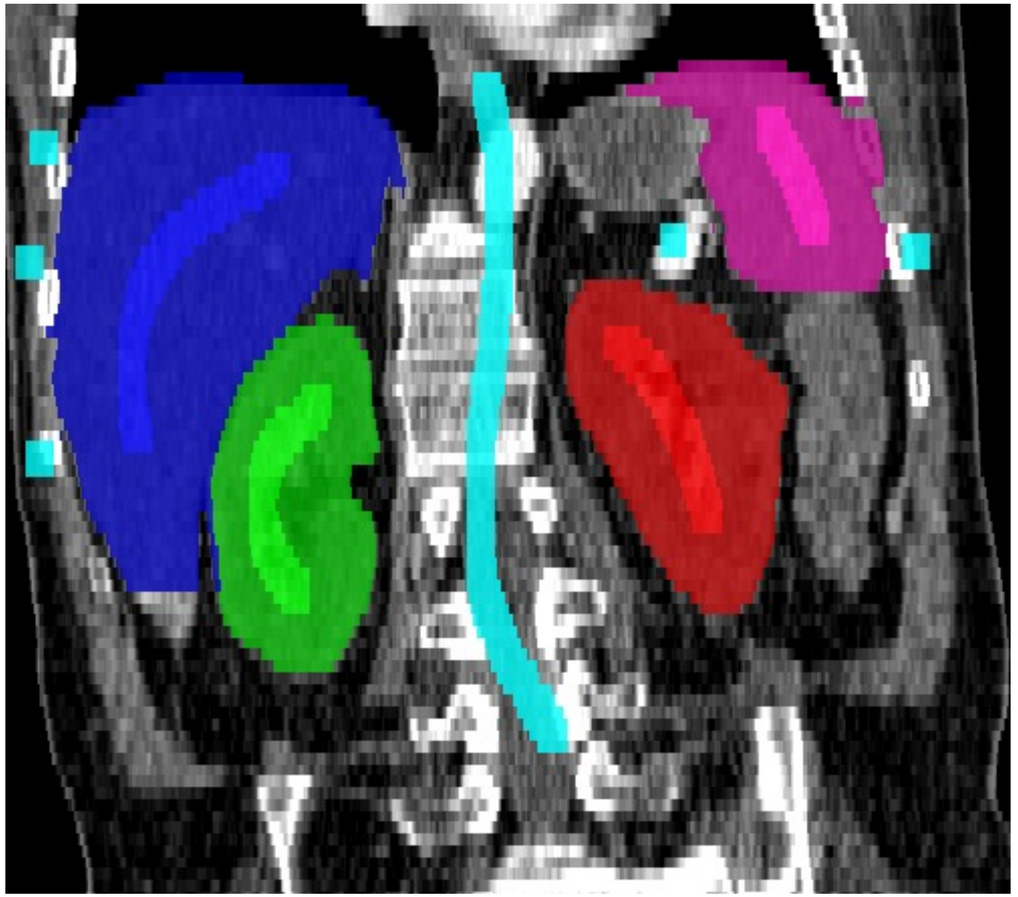
Multi-Label Graph Cuts – Example

- Interactive corrections



Multi-Label Graph Cuts – Example

- Interactive corrections



Graph Cuts

Advantages

- accurate
- reasonably efficient, interactive

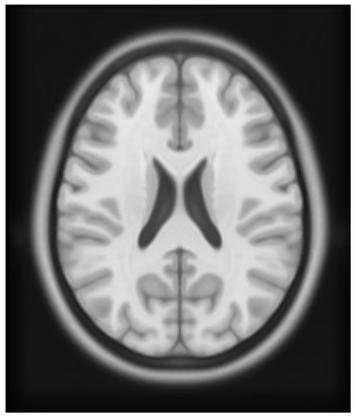
Disadvantages

- semi-automatic, requires user input
- difficult to select tuning parameters

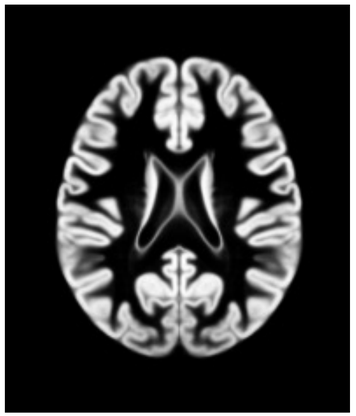
Atlas-Based Segmentation

What is an atlas?

MNI ICBM 152 Nonlinear atlases (2009)

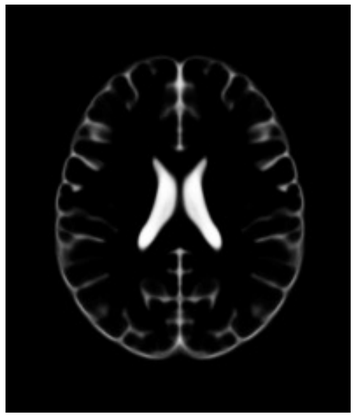


T1 MRI



grey matter

white matter

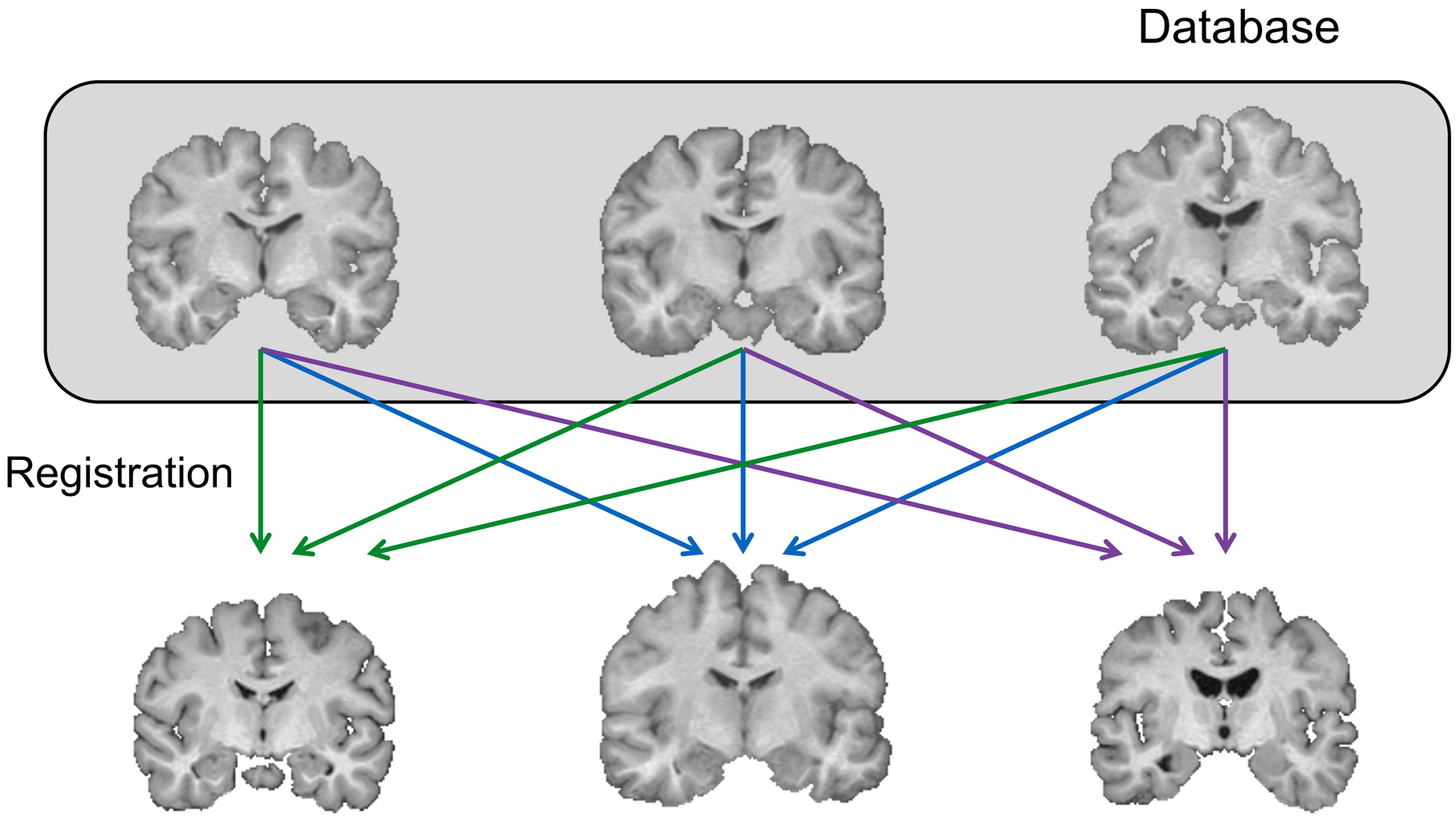


CSF

- An map or chart of the anatomy
- Atlases usually have
  - geometric information about points, curves or surfaces, or
  - label information about voxels (anatomical regions or function)
- Atlases are usually constructed from example data
  - single subjects
  - populations of subjects, e.g. by averaging to produce probabilistic atlases

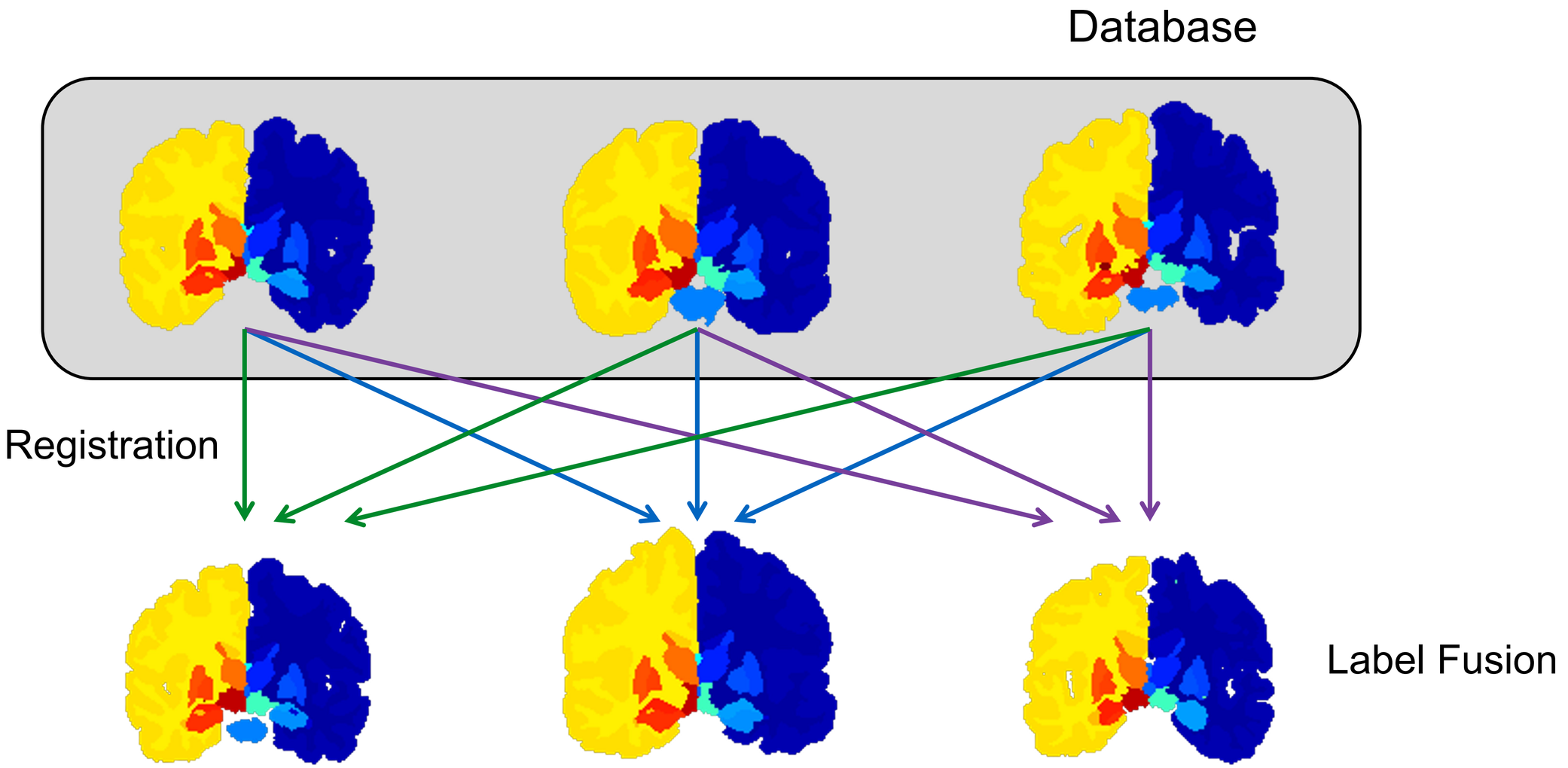
Segmentation using Registration

Label propagation

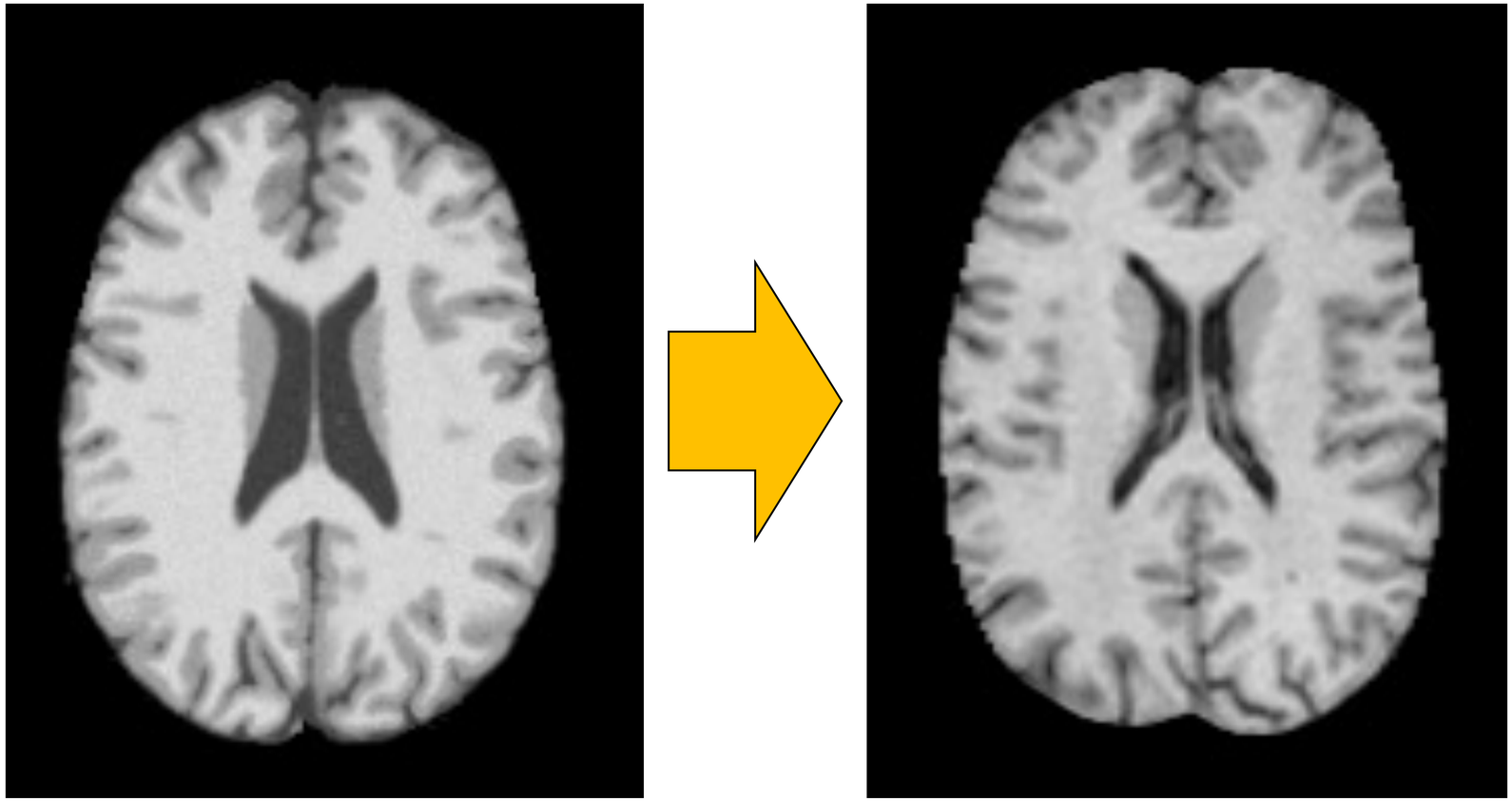


Segmentation using Registration

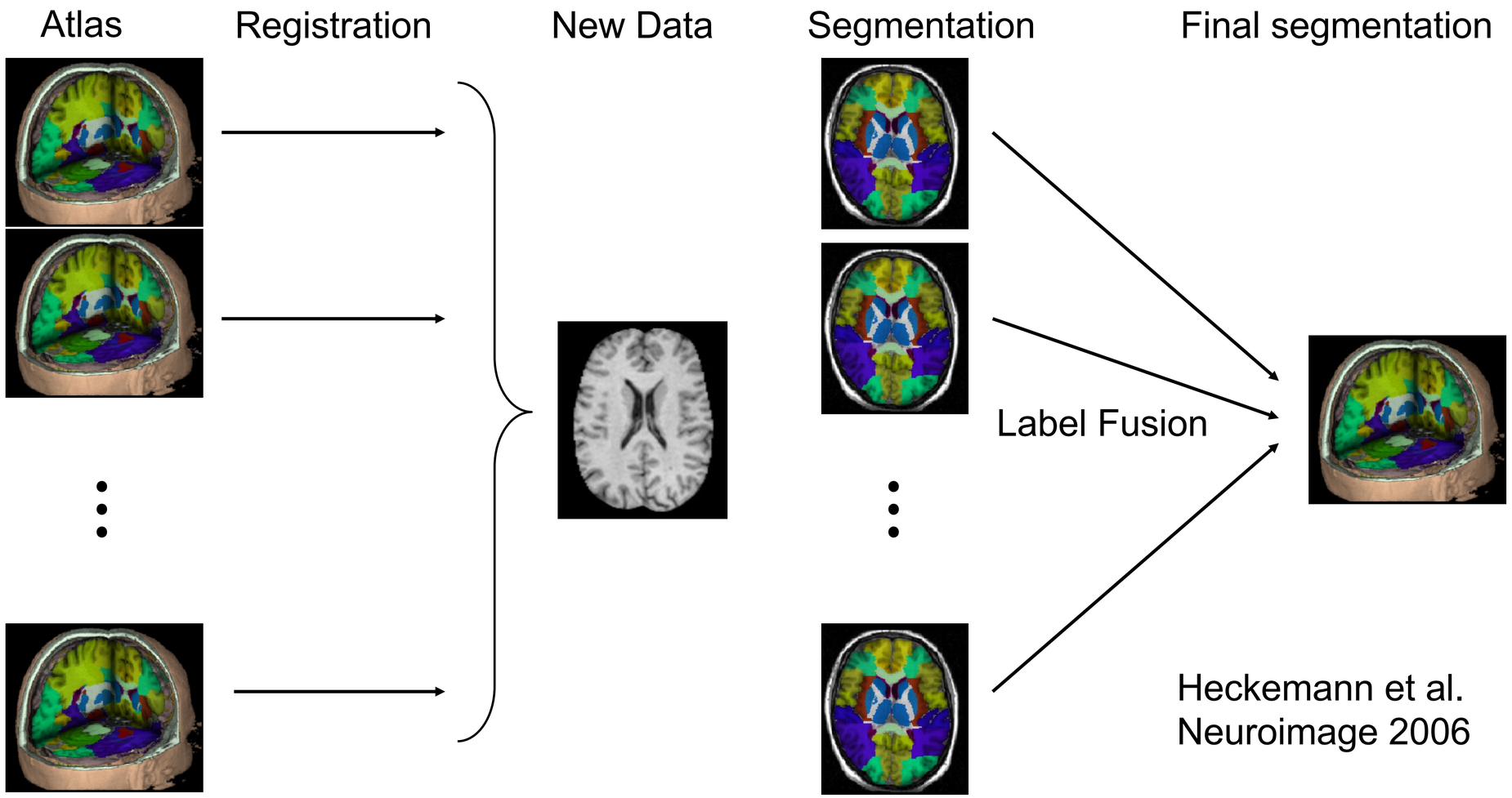
Label propagation



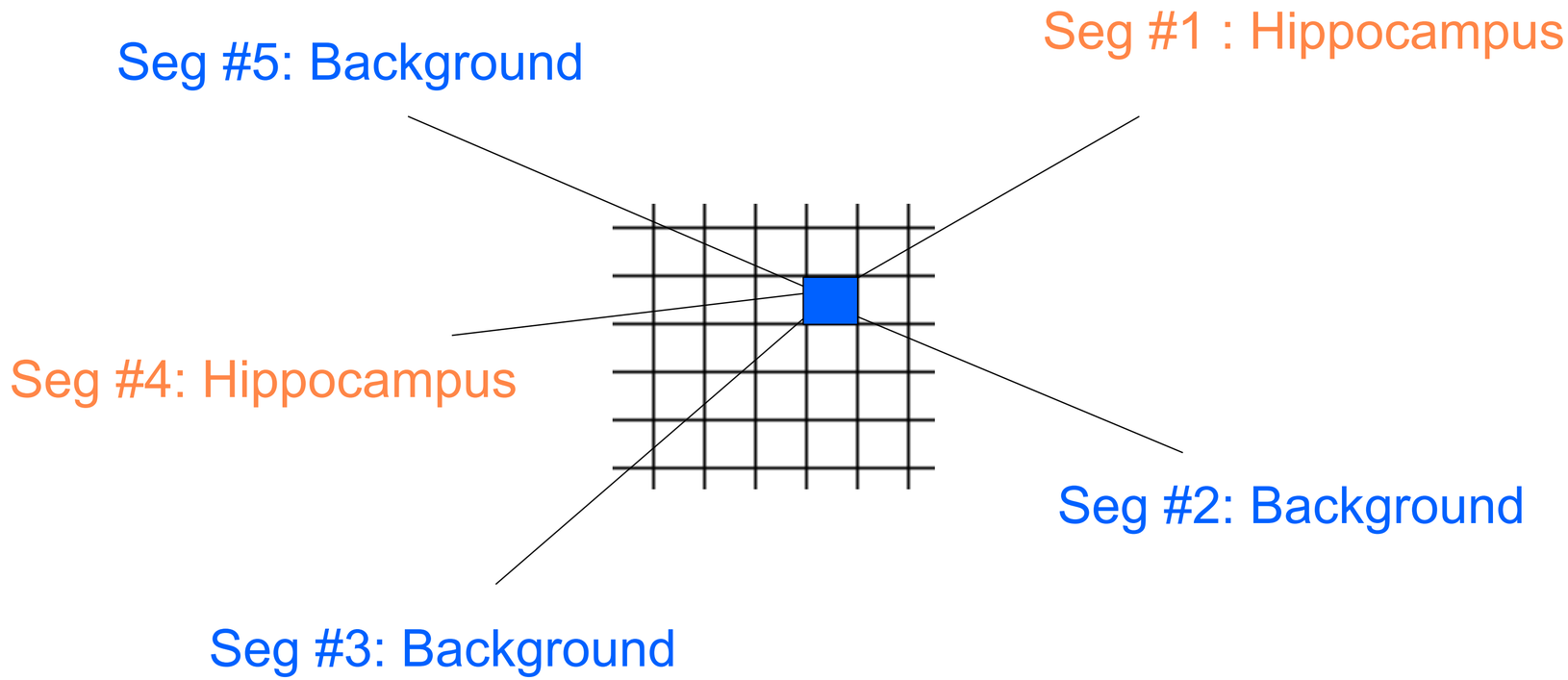
Segmentation using Registration



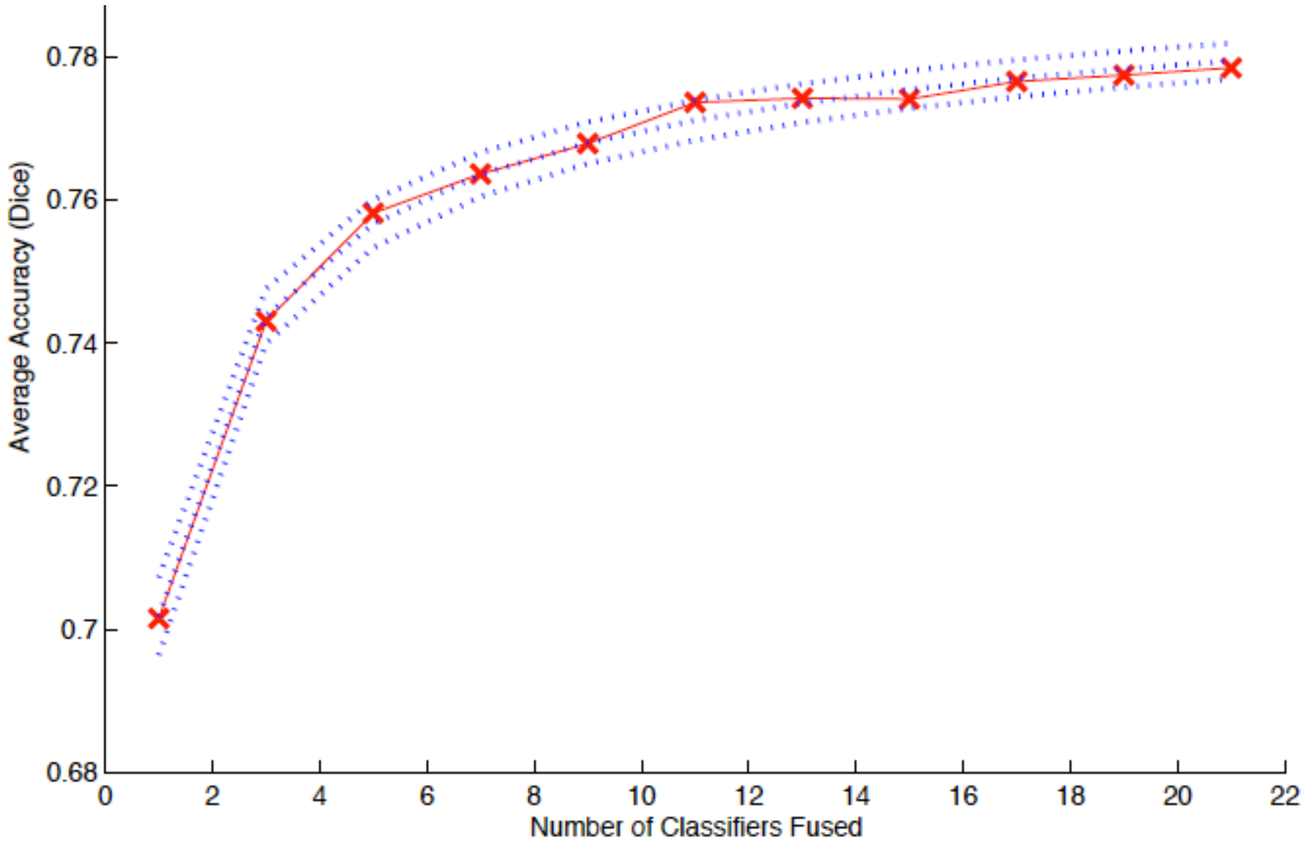
Multi-Atlas Label Propagation



Label Fusion: Simple Majority Voting



Effect of Number of Atlases



Heckemann et al.
Neuroimage 2006

Multi-Atlas Label Propagation

Advantages

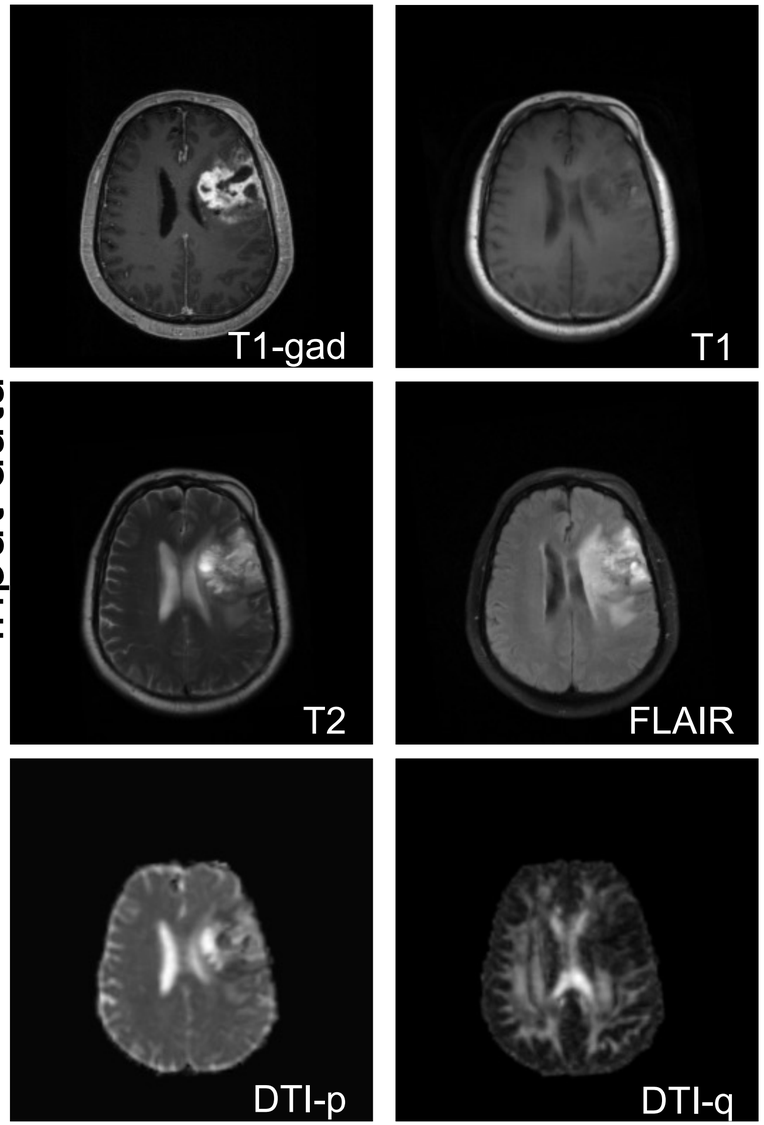
- robust and accurate (like ensembles)
- yields plausible segmentations
- fully automatic

Disadvantages

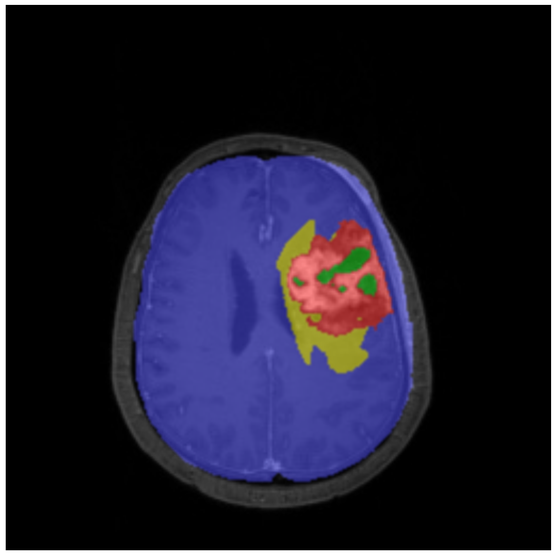
- computationally expensive
- cannot deal well with abnormalities
- not suitable for tumour segmentation

Random Forests

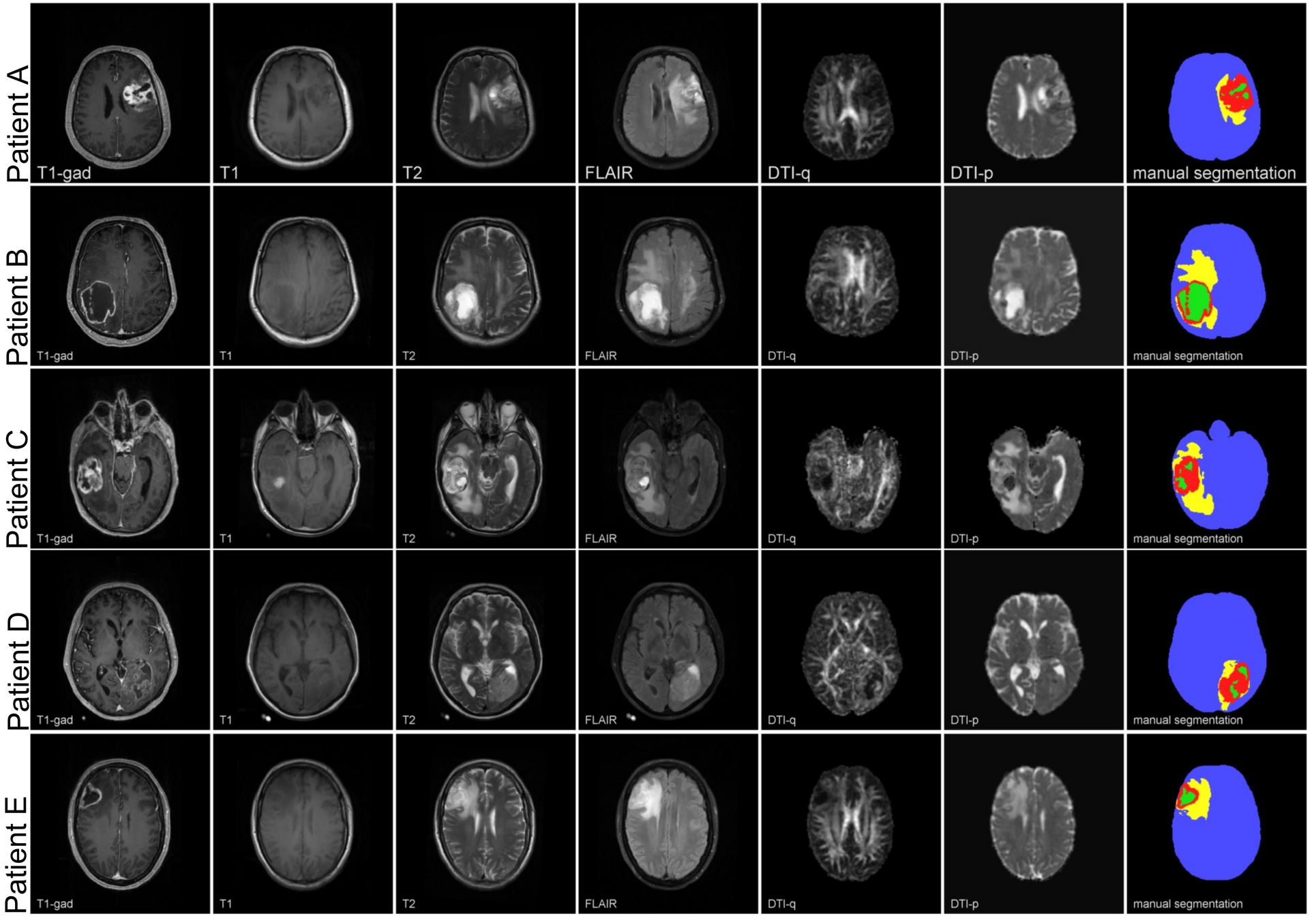
Multi-channel 3D MRI input data



Segmentation of tumorous tissues:

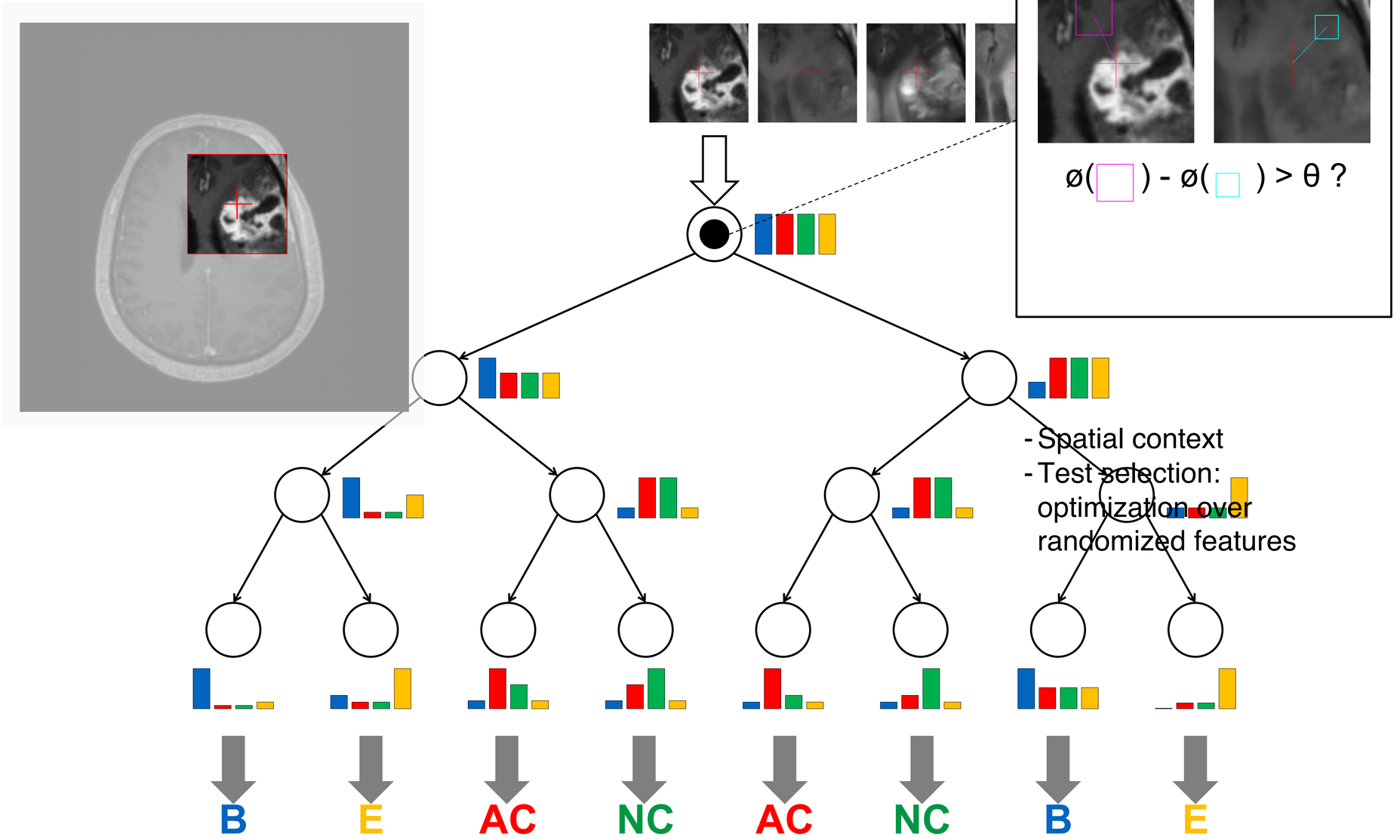


- Active cells
- Necrotic core
- Edema
- Background


Patient A
T1-gad
T1
T2
FLAIR
DTI-q
DTI-p
manual segmentation
Patient B
T1-gad
T1
T2
FLAIR
DTI-q
DTI-p
manual segmentation
Patient C
T1-gad
T1
T2
FLAIR
DTI-q
DTI-p
manual segmentation
Patient D
T1-gad
T1
T2
FLAIR
DTI-q
DTI-p
manual segmentation
Patient E
T1-gad
T1
T2
FLAIR
DTI-q
DTI-p
manual segmentation

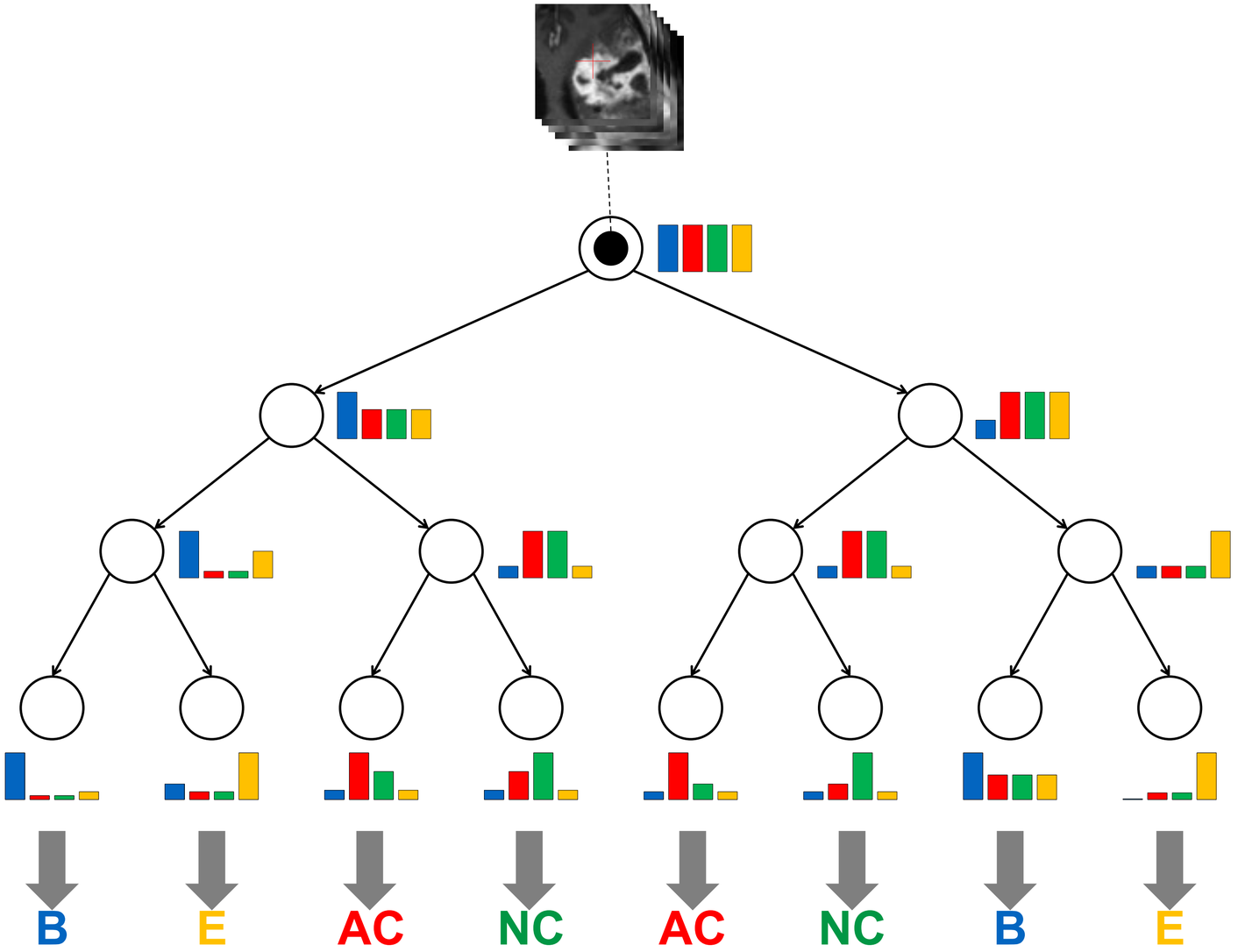
Training: Growing the Trees

Determine tests to distinguish between tissue classes



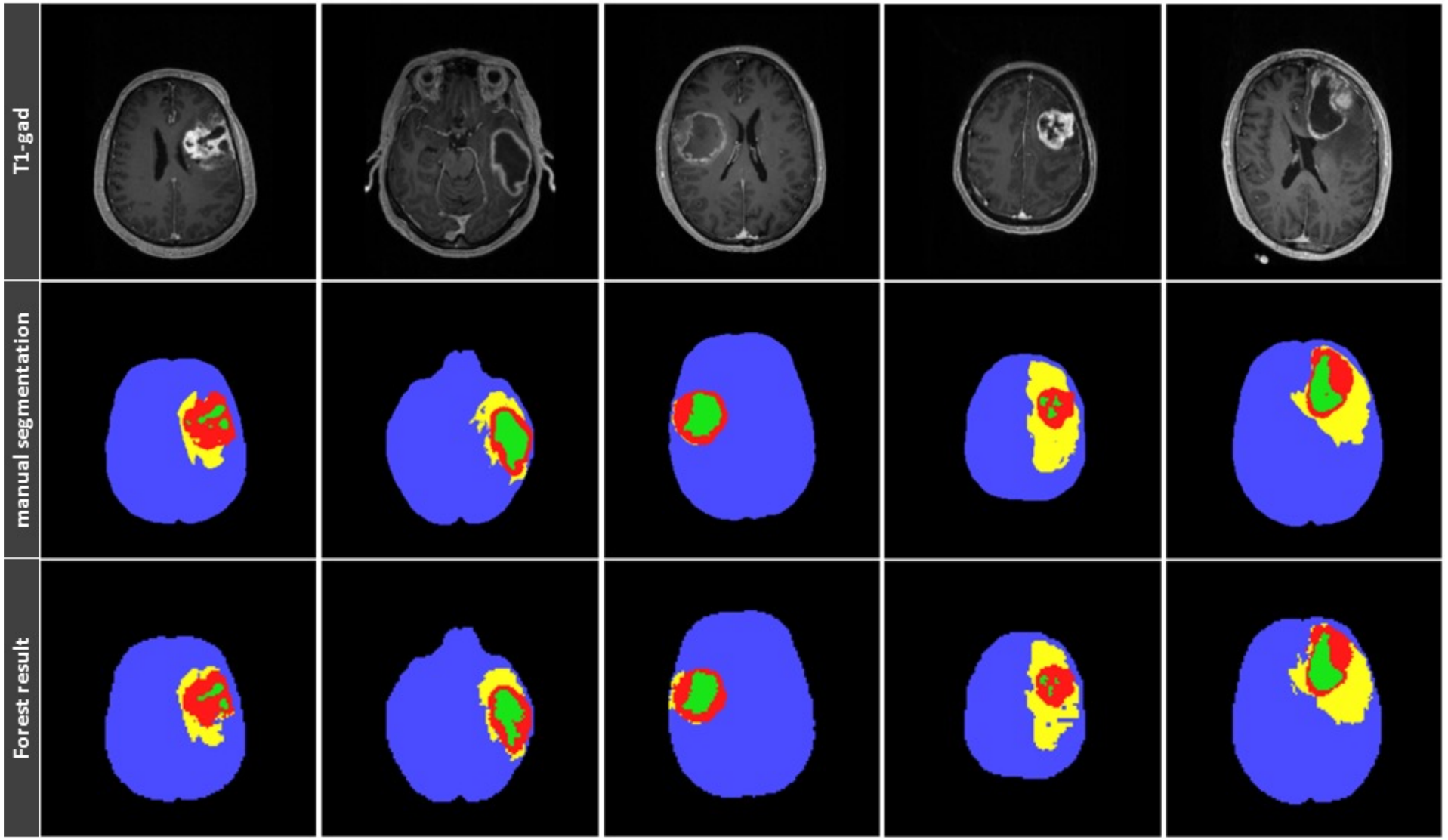
Testing: Traversing the Trees

Apply the learned tests to classify tissue points



Visual Results

https://www.doc.ic.ac.uk/~bglocker/pdfs/zikic2012miccai.pdf
https://www.doc.ic.ac.uk/~bglocker/pdfs/zikic2012brats.pdf



Random Forests

Advantages

- ensemble classifiers are robust and accurate
- computationally efficient
- fully automatic

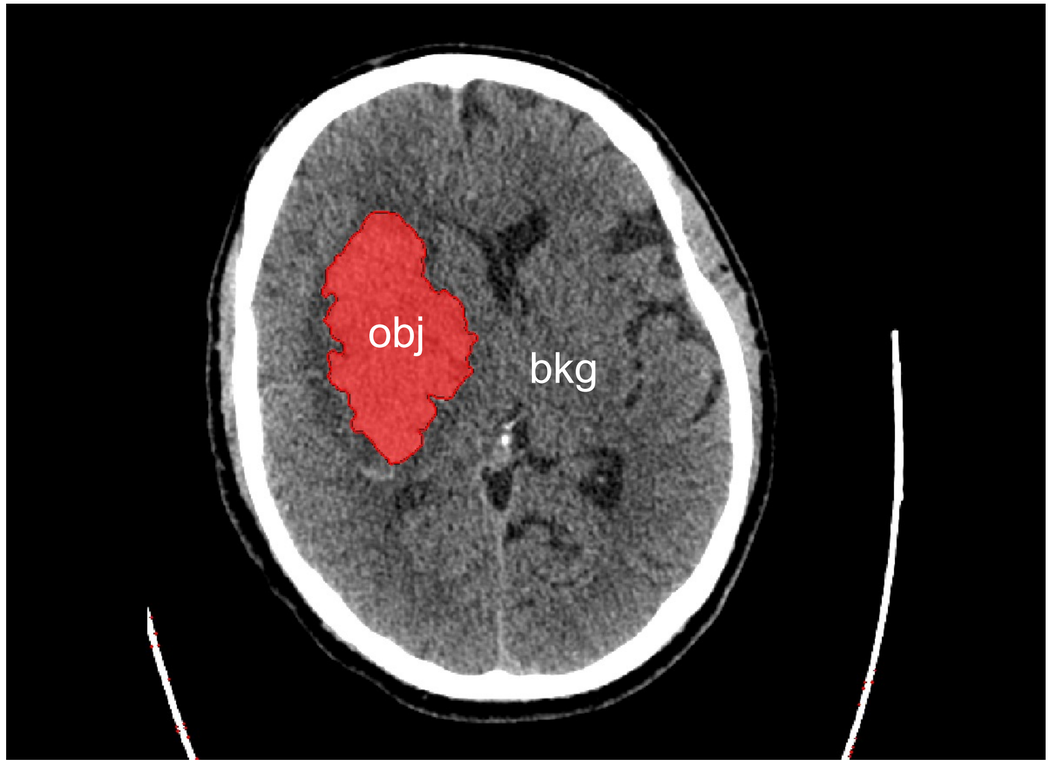
Disadvantages

- shallow model, no hierarchical features
- no guarantees on connectedness

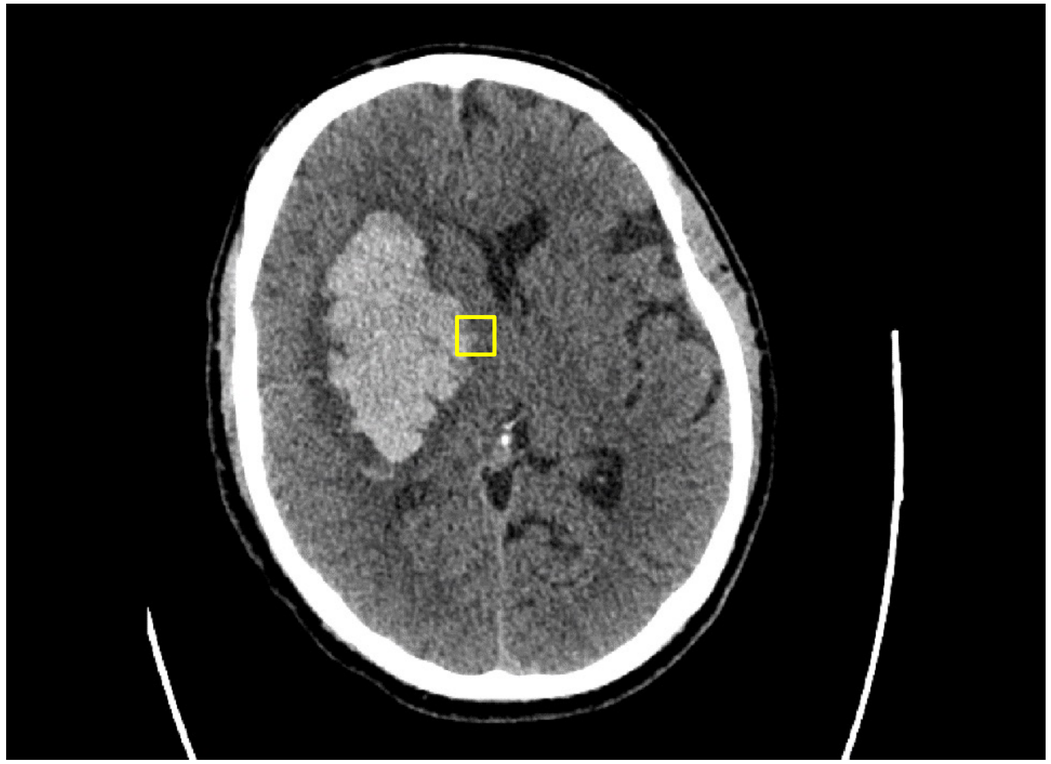
Deep learning for image segmentation…

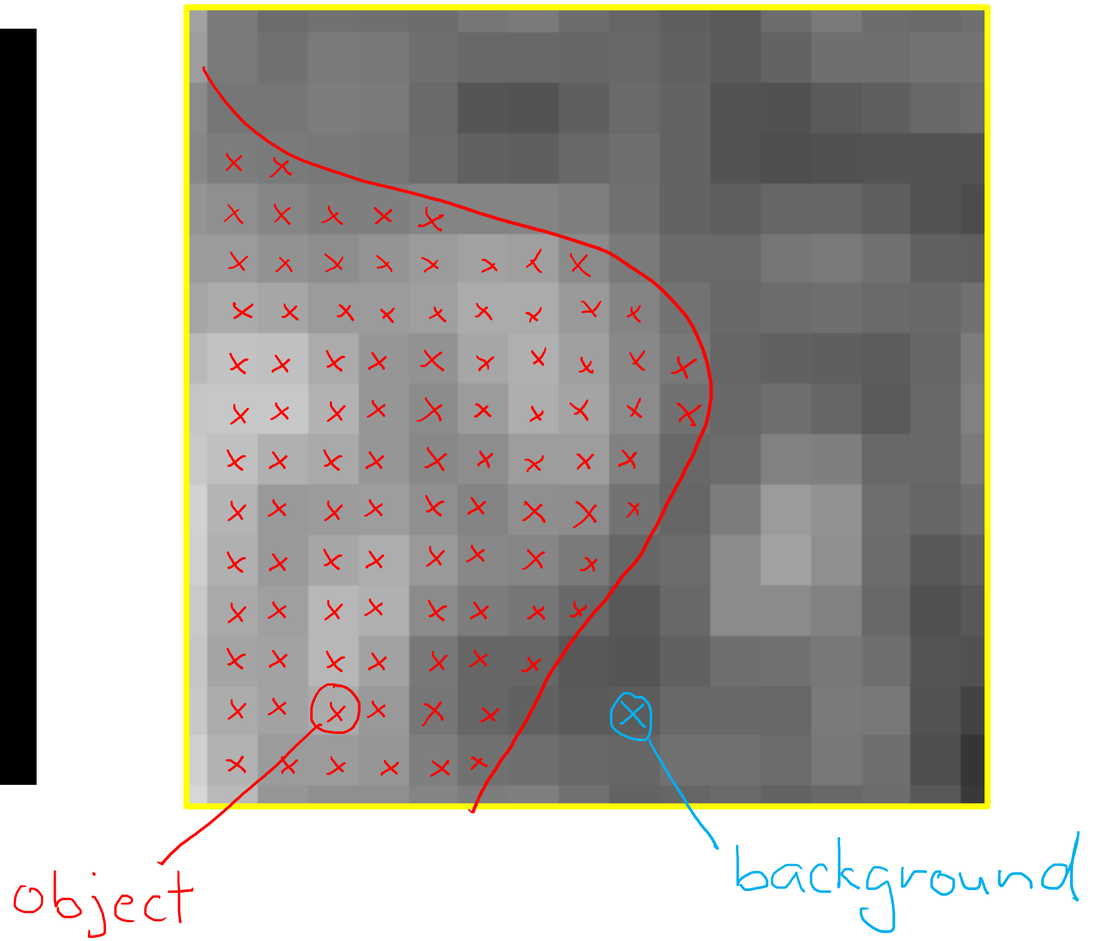
Segmentation via Dense Classification

Segmentation via Dense Classification



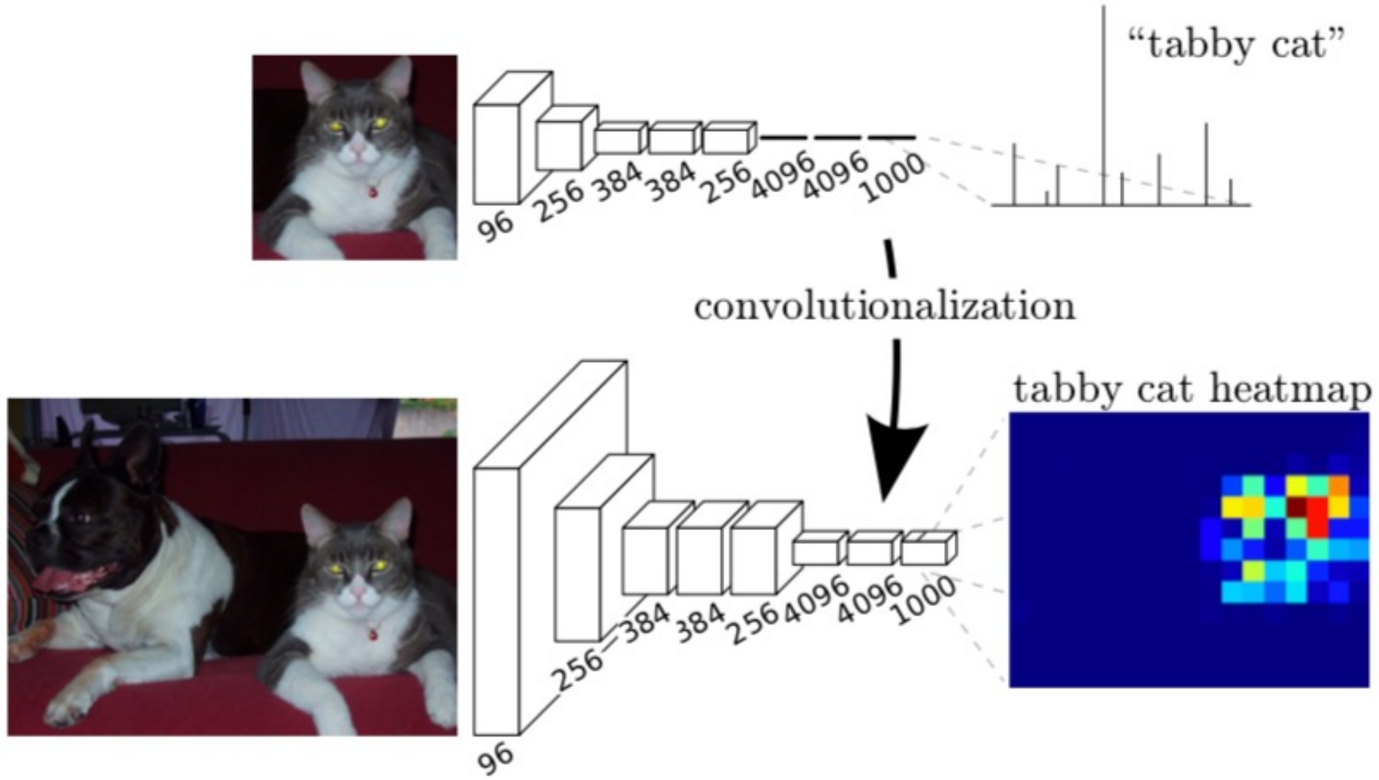
Segmentation via Dense Classification





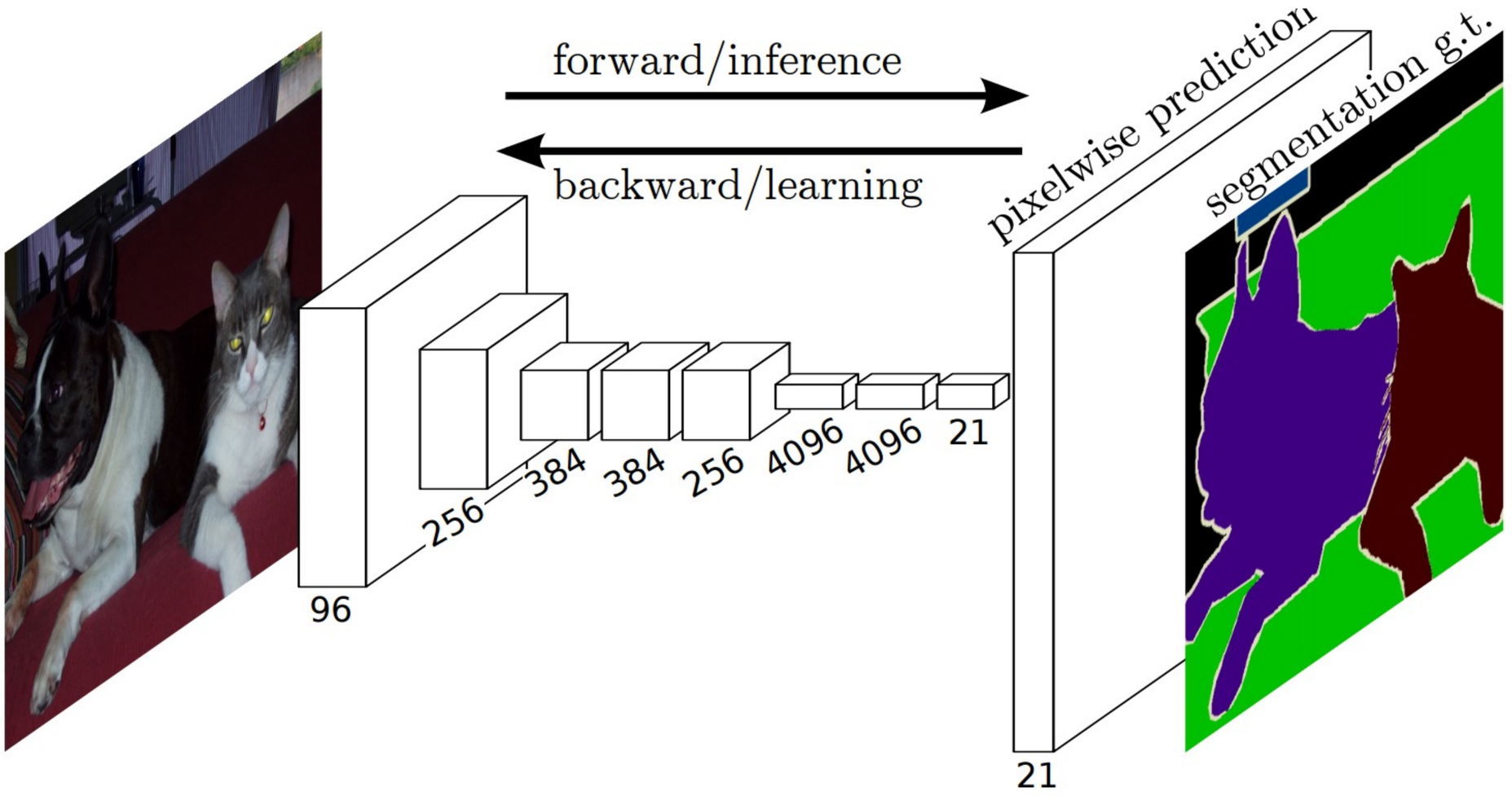
FCNs for image segmentation

- From classifier to dense Fully Convolutional Networks:

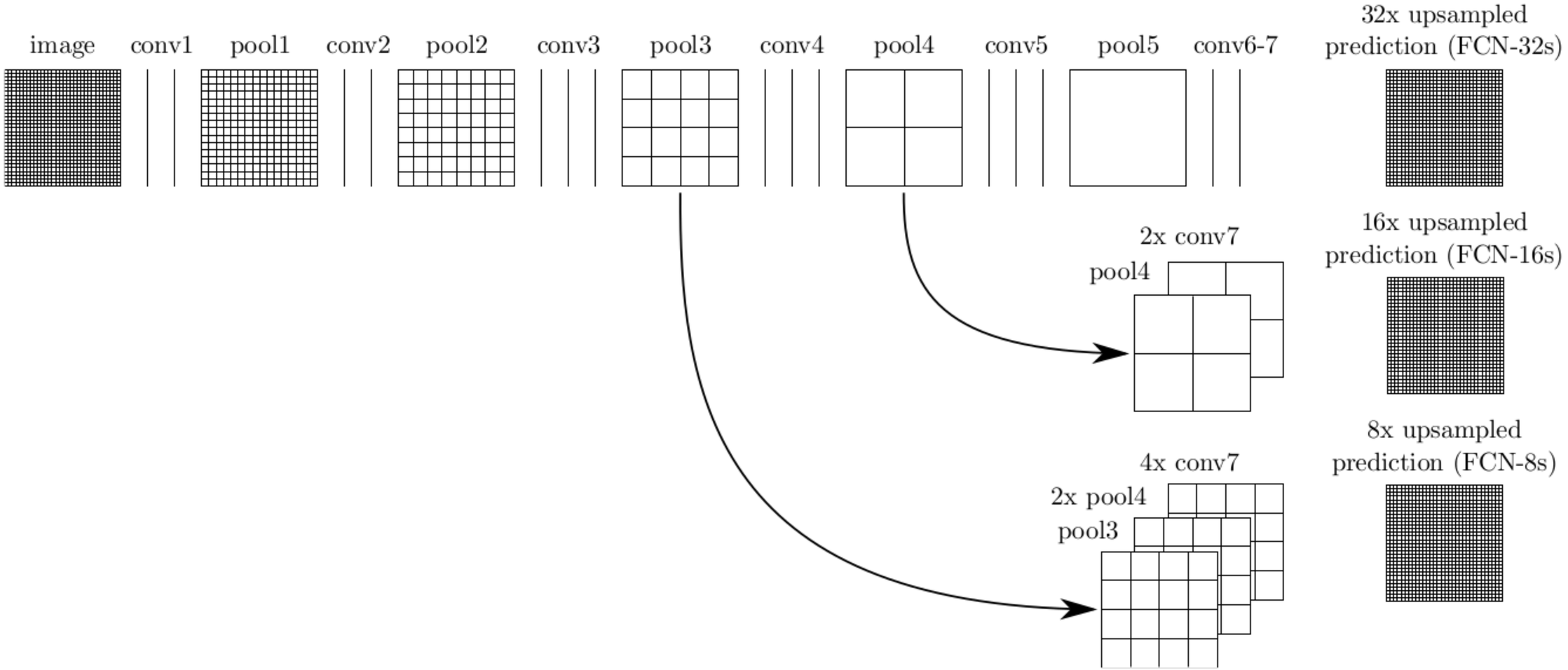


Long, J., Shelhamer, E., Darrell, T. Fully Convolutional Networks for Semantic Segmentation. 2014.

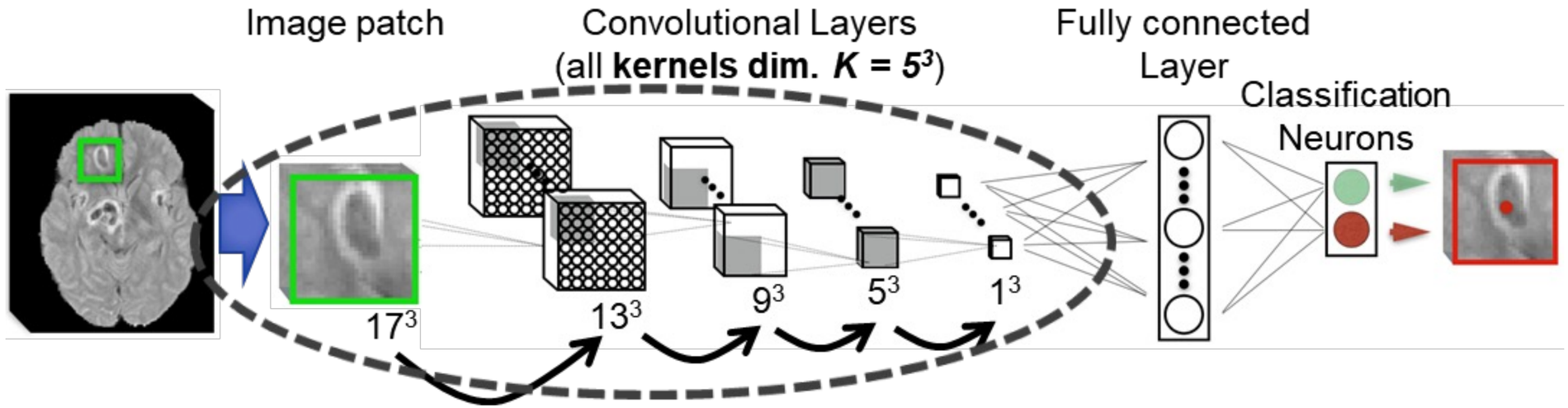
FCNs for image segmentation

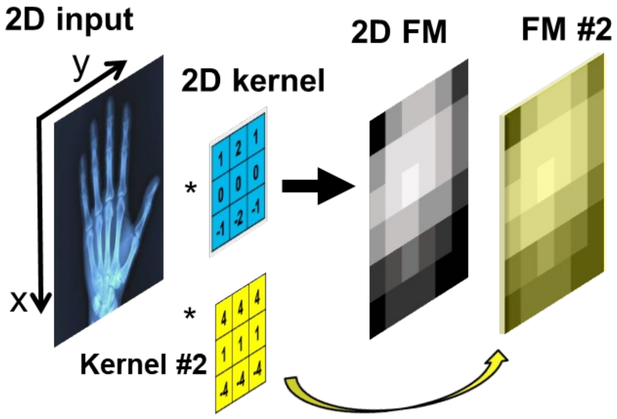


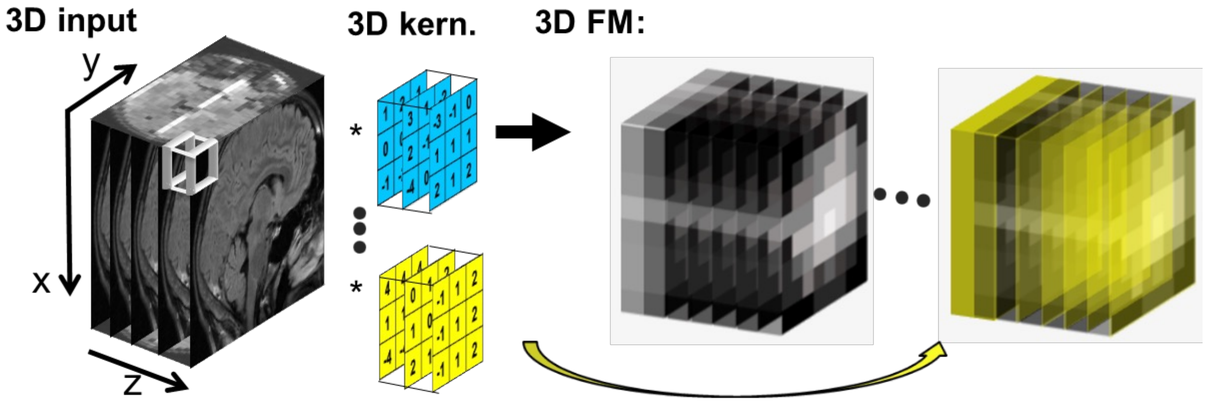
FCNs for image segmentation



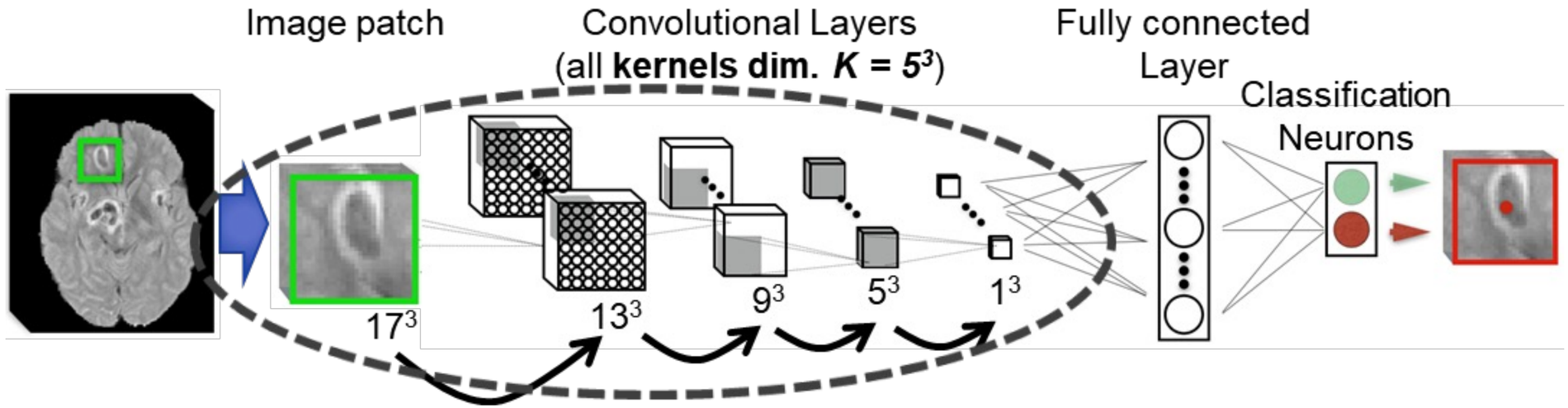
Segmentation via Dense Classification

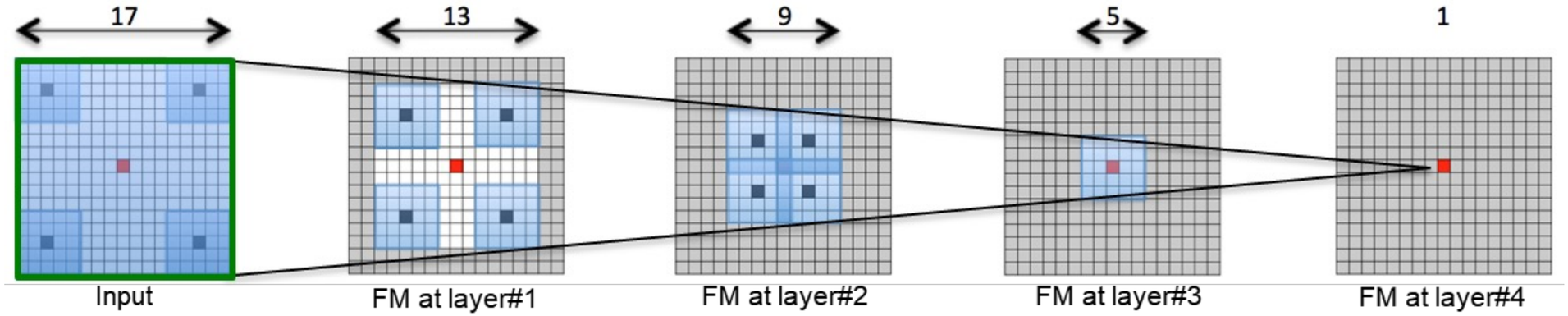






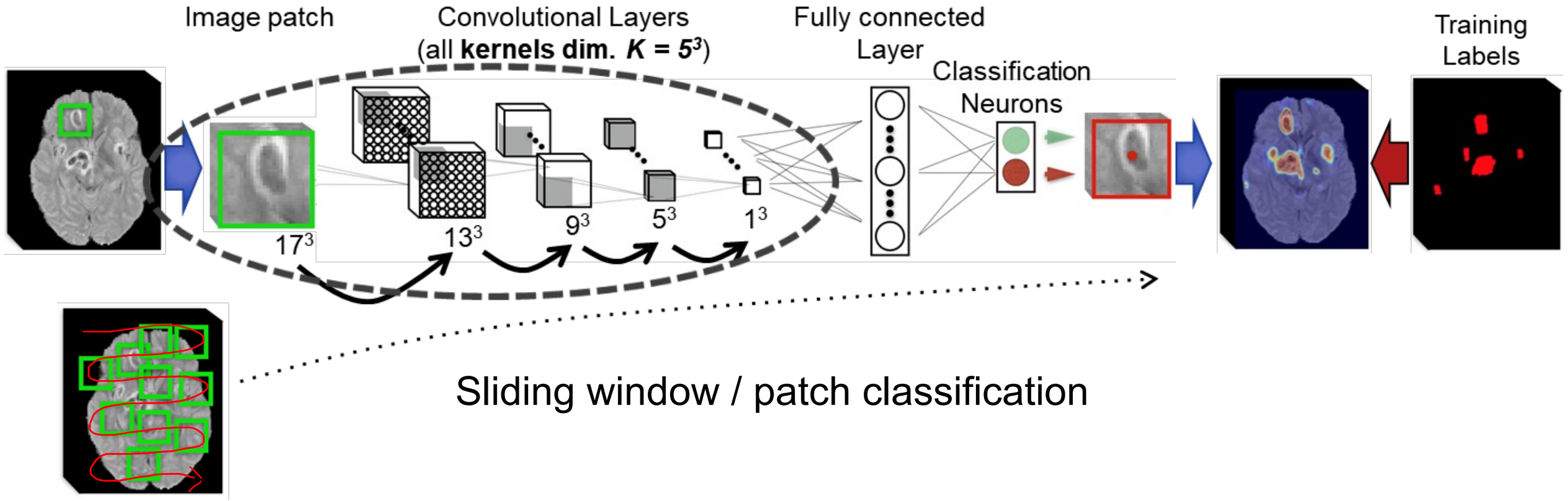
Segmentation via Dense Classification



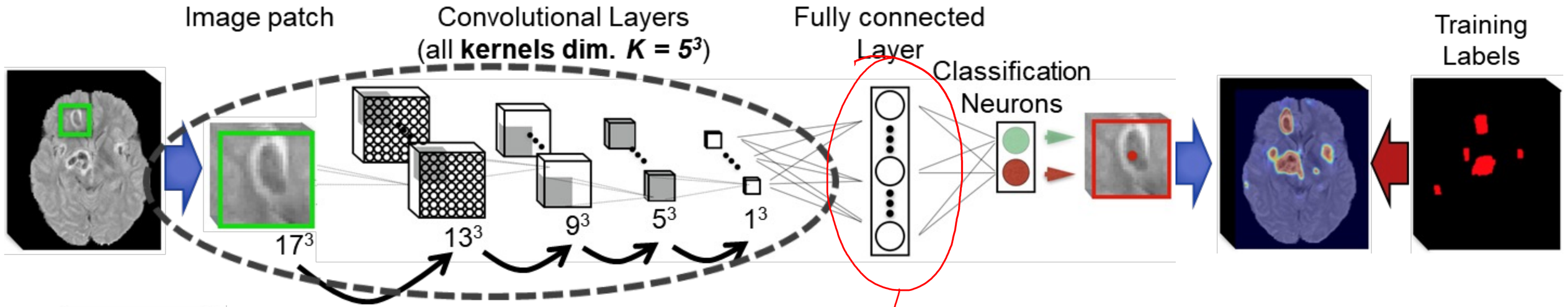


Predicted pixel | Conv. kernel | Out of bounds for valid conv.

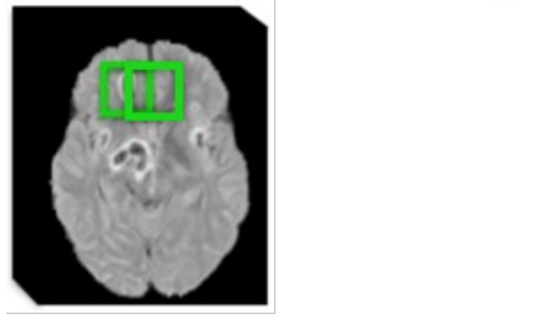
Segmentation via Dense Classification

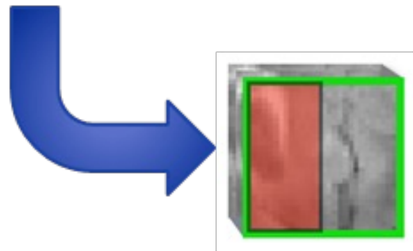


Segmentation via Dense Classification



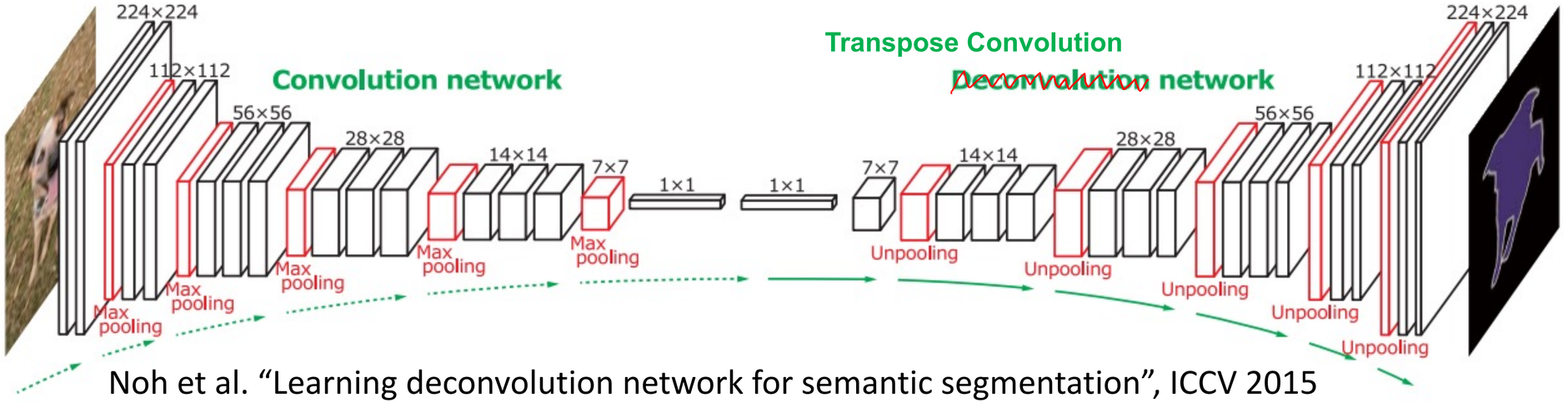
This is the crux



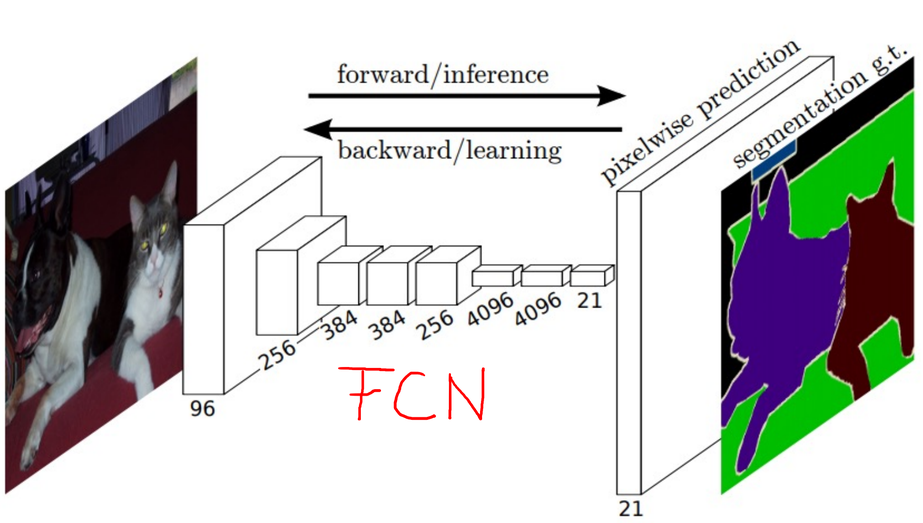


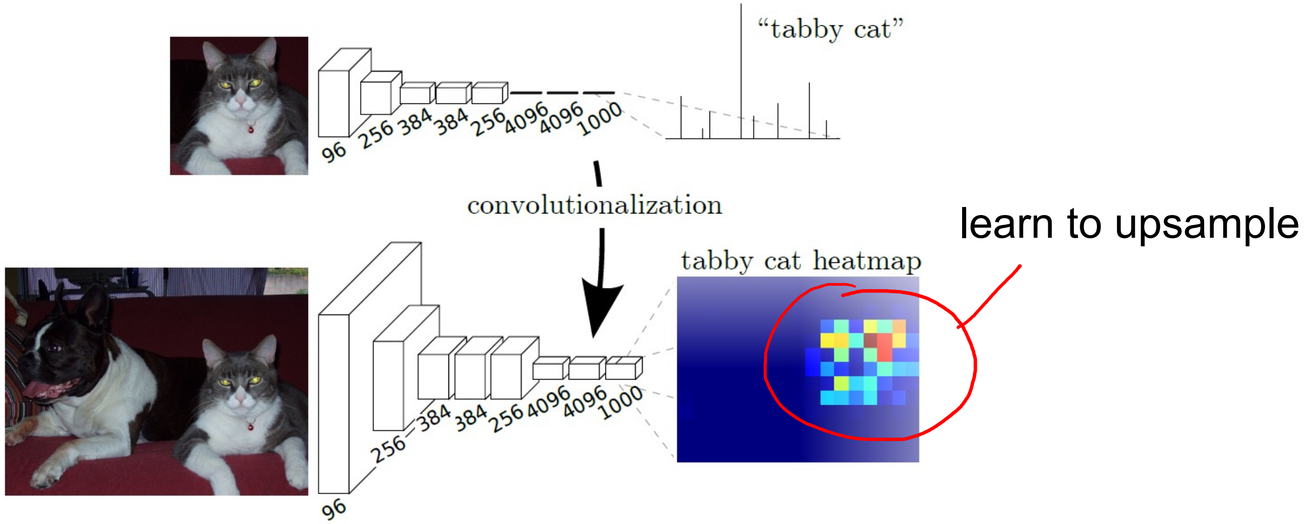
Very inefficient. Redundant computations!

Encoder-Decoder Networks



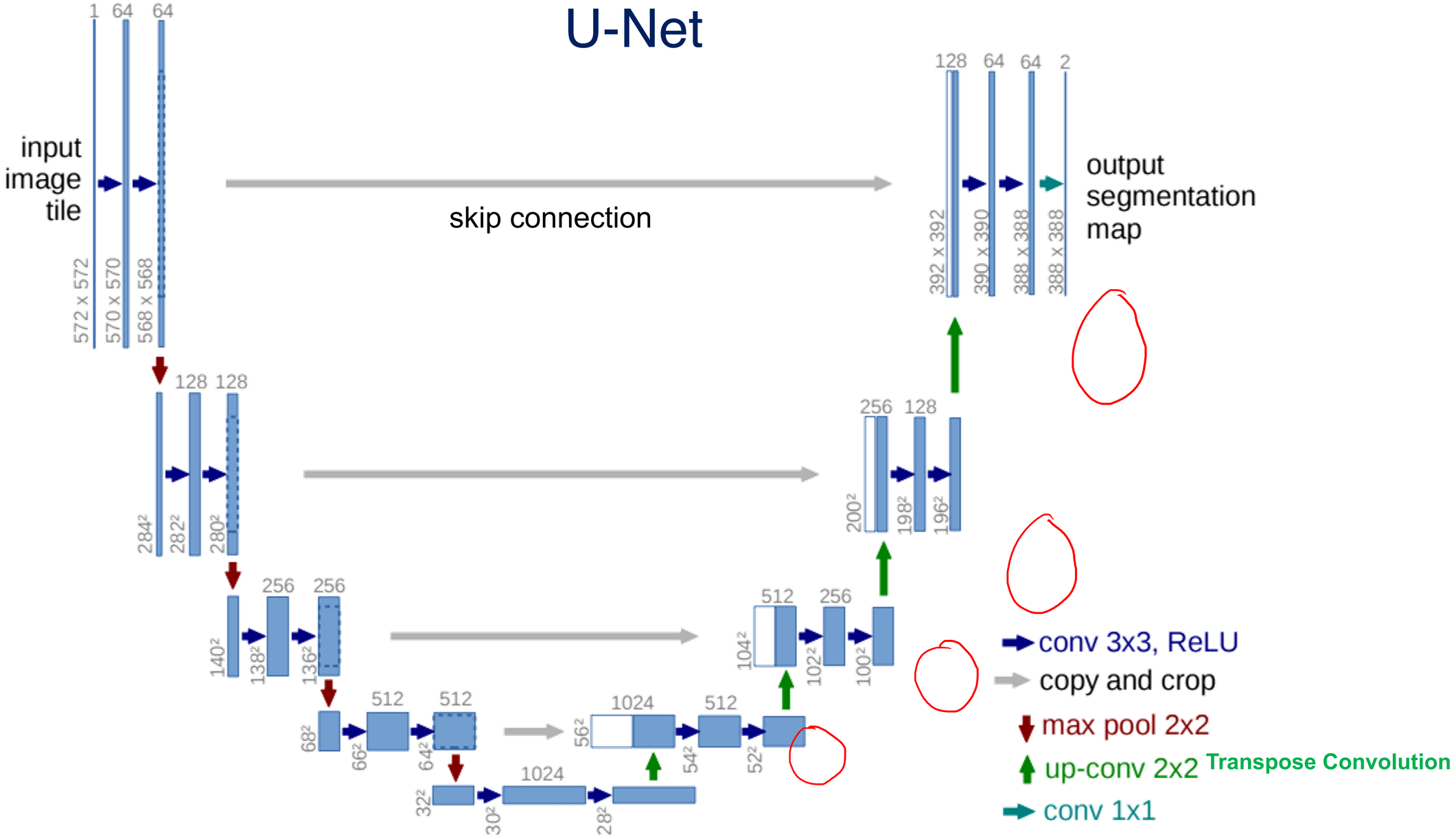
Noh et al. “Learning deconvolution network for semantic segmentation”, ICCV 2015





Long et al. “Fully convolutional networks for semantic segmentation”, CVPR 2015

U-Net



Ronneberger et al. “U-Net: Convolutional Networks for Biomedical Image Segmentation”, MICCAI 2015

Upsampling

Low Res. FM in

5	3	1
0	2	4
4	3	5

0	0	0	0	0	0	0	0
0	5	0	3	0	1	0	0
0	0	0	0	0	0	0	0
0	0	0	2	0	4	0	0
0	0	0	0	0	0	0	0
0	4	0	3	0	5	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

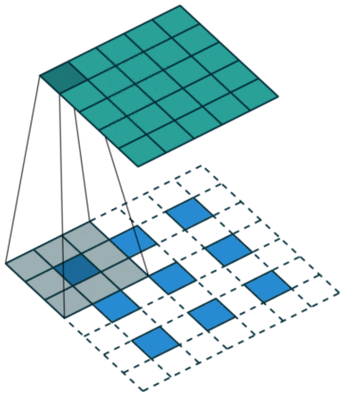
*

0.25	0.5	0.25
0.5	1	0.5
0.25	0.5	0.25

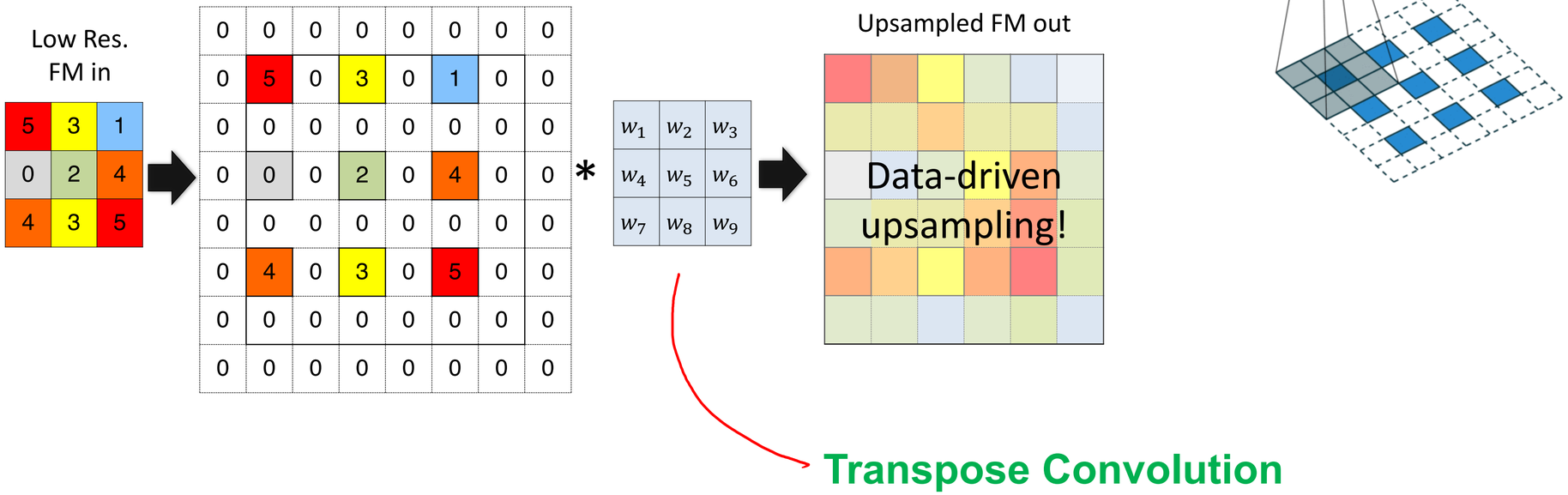
Upsampled FM out

5	4	3	2	1	0.5
2.5	2.5	3.5	2.5	2.5	1.3
0	1	2	3	4	2
2	2.3	2.5	3.5	4.5	2.3
4	3.5	3	4	5	2.5
2	1.8	1.5	2	2.5	1.3

let's learn these weights

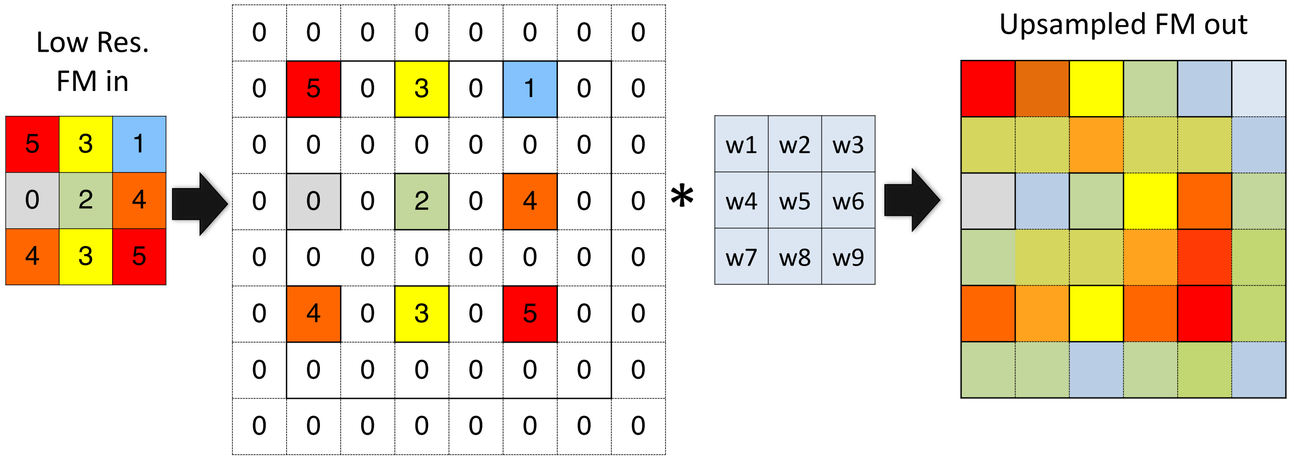


Upsampling

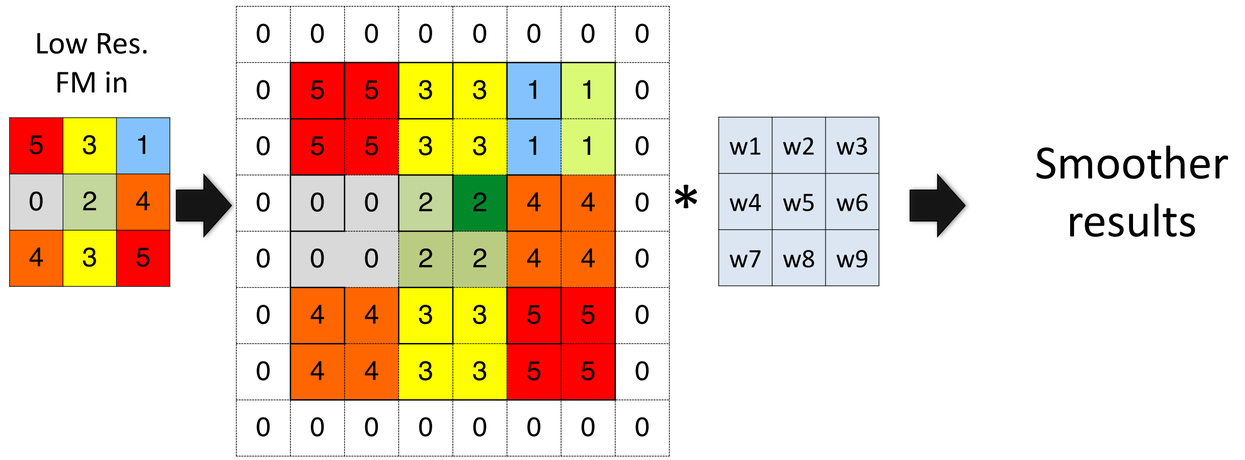

Low Res.
FM in
5 3 1
0 2 4
4 3 5
0 0 0 0 0 0 0 0
0 5 0 3 0 1 0 0
0 0 0 0 0 0 0 0
0 0 0 2 0 4 0 0
0 0 0 0 0 0 0 0
0 4 0 3 0 5 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
*
w_1 w_2 w_3
w_4 w_5 w_6
w_7 w_8 w_9
Upsampled FM out
Data-driven upsampling!
Transpose Convolution

Upsampling

Transpose convolution with **zero-fill / “bed of nails”**

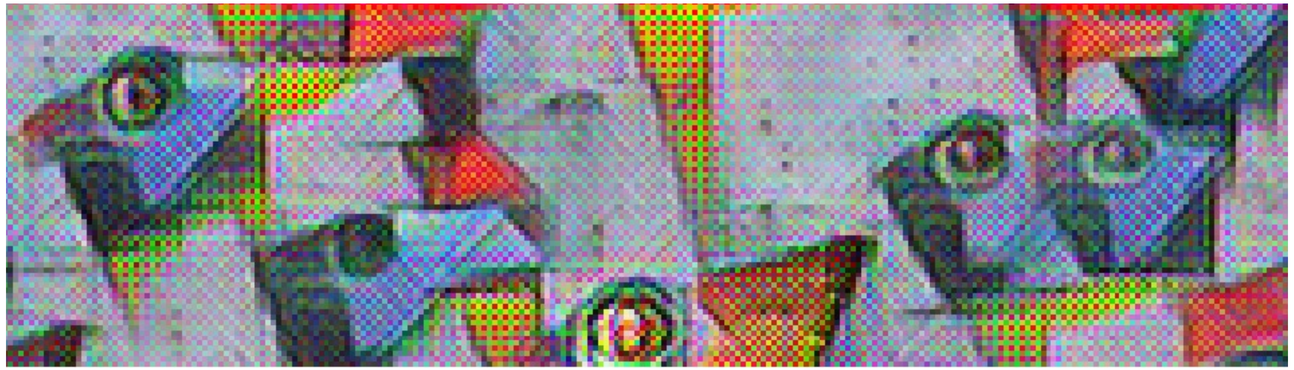


Transpose convolution with **nearest-neighbour-fill**



Upsampling Artefacts

Transpose convolution with **zero-fill / “bed of nails”**



Transpose convolution with **nearest-neighbour-fill**

From: http://distill.pub/2016/deconv-checkerboard/

Padding

"Valid" Convolution

Dim. D^{in} x D^{in}

	4	5	5	4	1
4	6	6	6	5	3
5	6	7	6	5	3
4	5	5	4	4	1
1	2	3	2	1	
	1	1	1		

*

Dim. K x K

1	2	1
0	0	0
-1	-2	-1

→

	-7	-5	-8
3	5	6	6
16	5	16	15
16	15	14	12

Output dimensions:
$D^{out} = D^{in} - (K-1)$

+ Convenient
- Synthetic content
Use other padding?
(e.g., mirroring)

"Same" Convolution

0	0	0	0	0	0	0	0
0		4	5	5	4	1	0
0	4	6	6	6	5	3	0
0	5	6	7	6	5	3	0
0	4	5	5	4	4	1	0
0	1	2	3	2	1		0
0		1	1	1			0
0	0	0	0	0	0	0	0

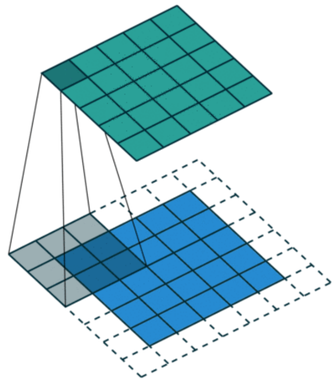
*

1	2	1
0	0	0
-1	-2	-1

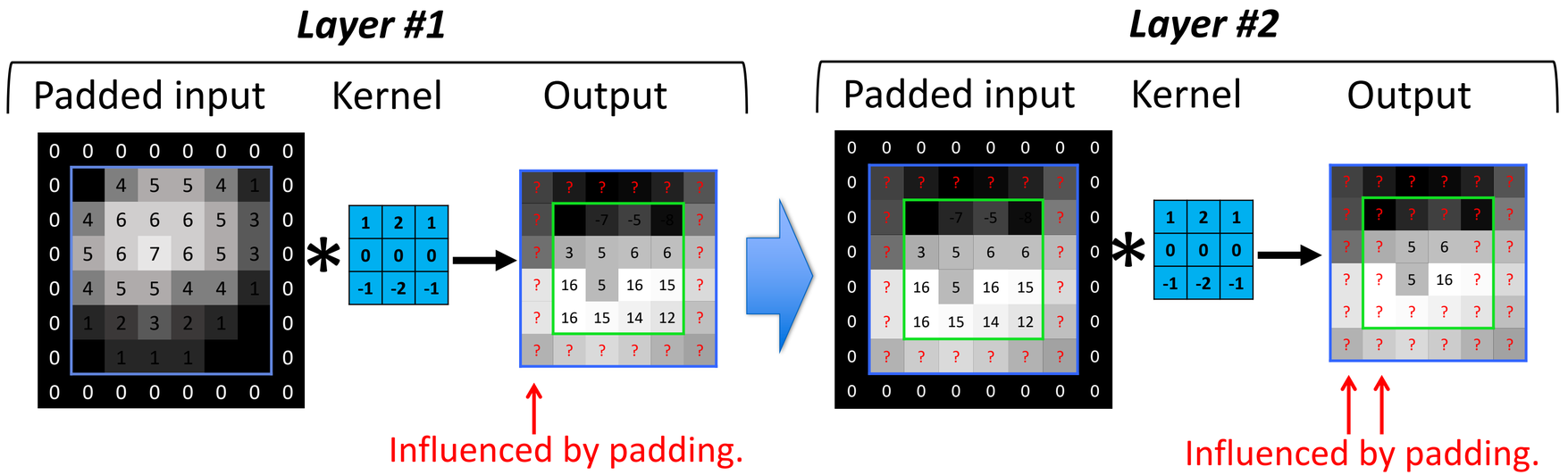
→

Influenced by padding.

?	?	?	?	?	?
?		-7	-5	-8	?
?	3	5	6	6	?
?	16	5	16	15	?
?	16	15	14	12	?
?	?	?	?	?	?

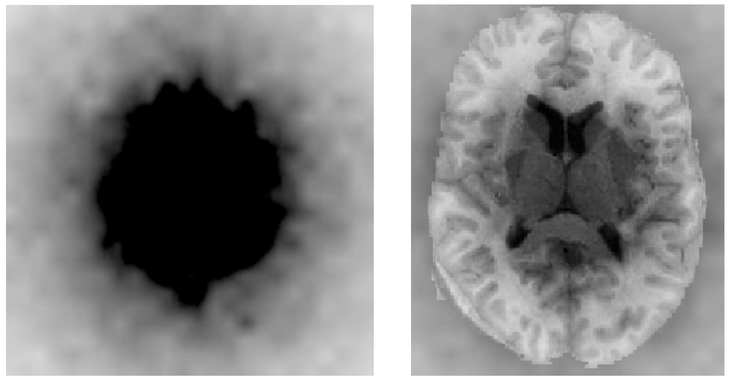


Boundary Effects

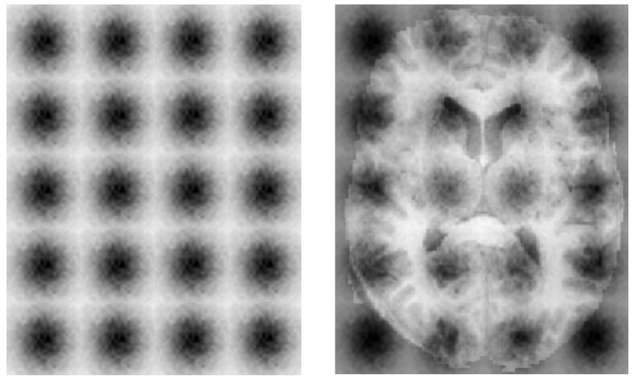


Patterns/artifacts that may show up due to padding (exaggerated for clarity):
(Segmentation or regression. Not classification)

If processing whole image at testing.

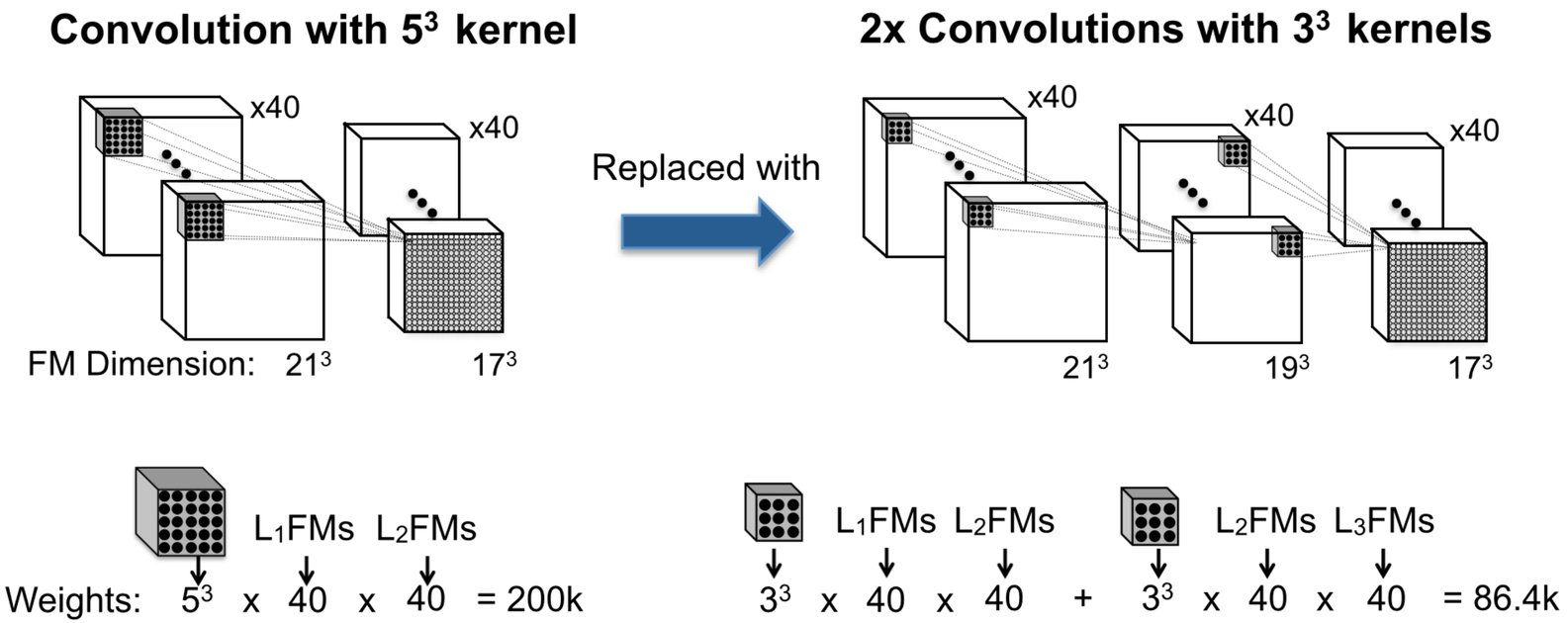


If tiling the image at testing.



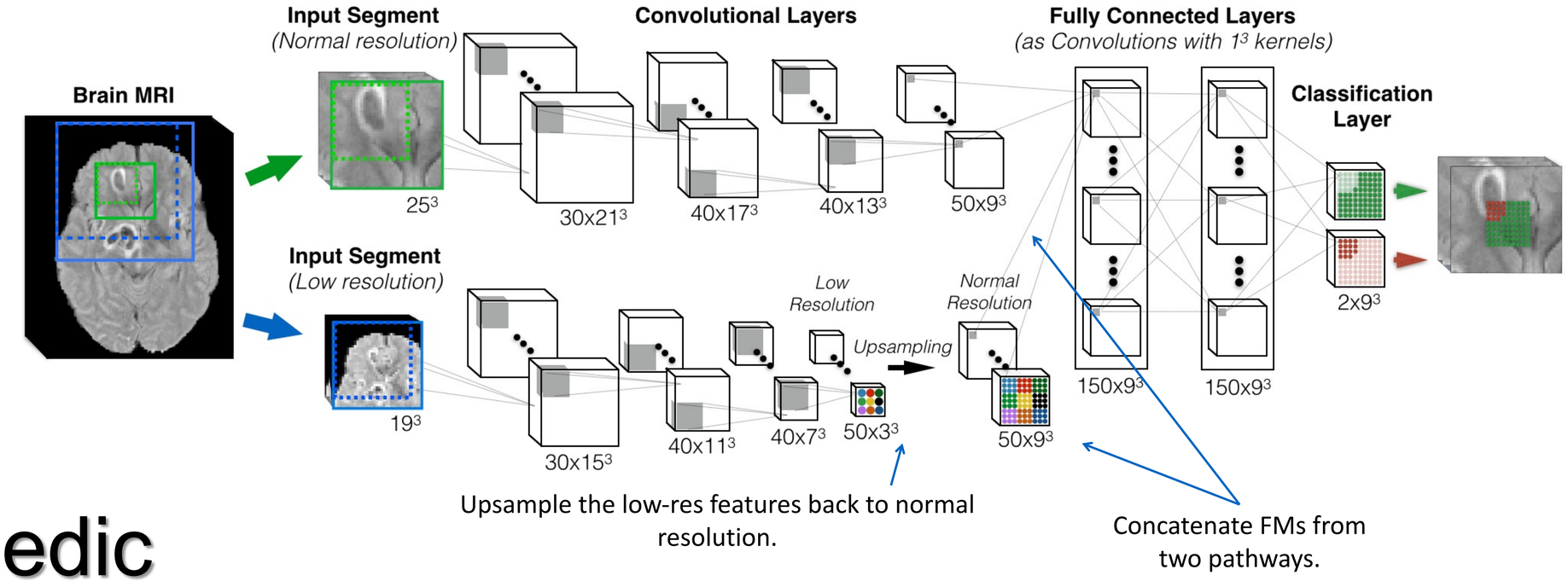
Going Deeper

- Deeper networks can represent more complex functions
- Just adding more layers is inefficient (too many parameters)
- **Idea:** Use only layers with small kernels [Simonyan et al. 2014]



Multi-Scale Processing

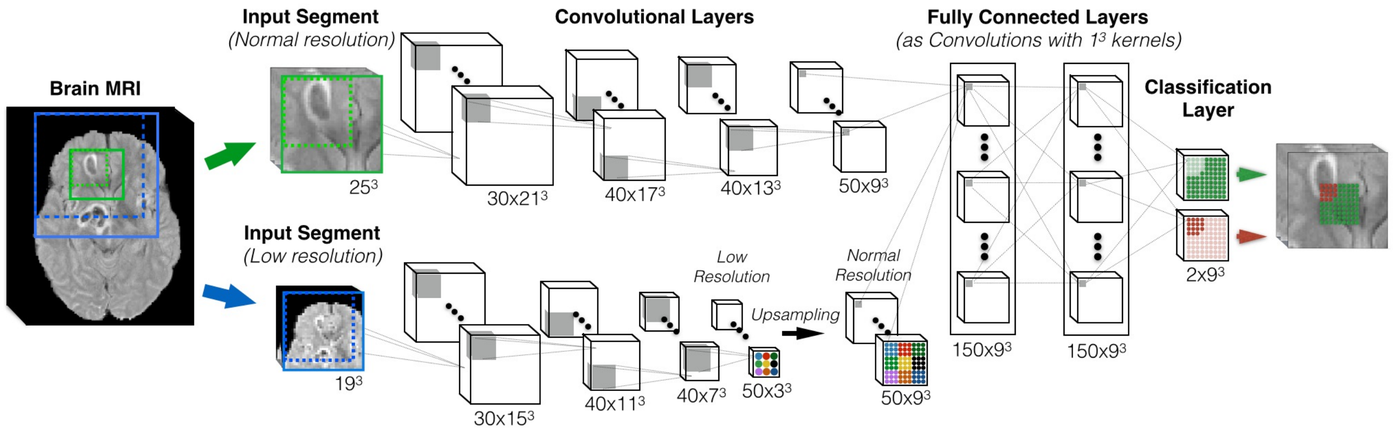
- How can we make the network to “see” more context
- **Idea:** Add more pathways which process downsampled images



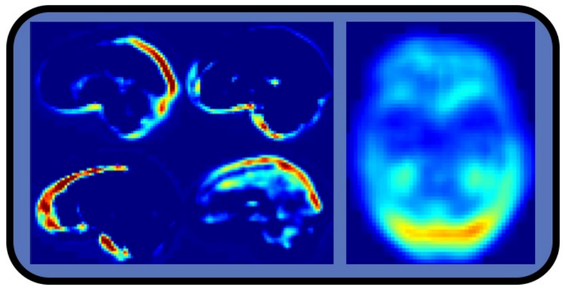
DeepMedic

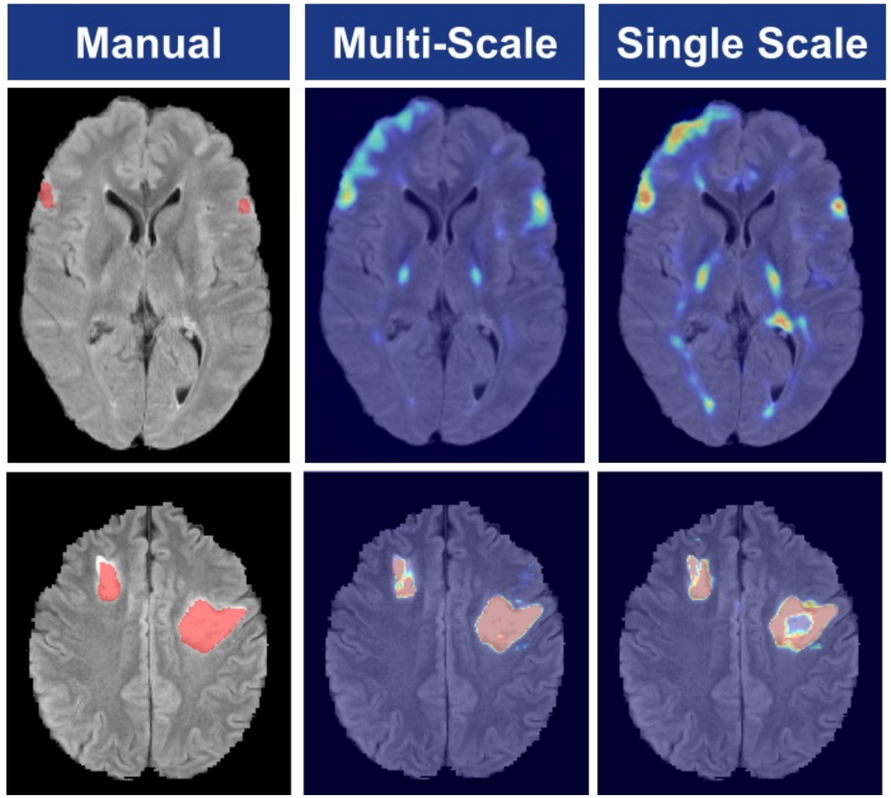
Kamnitsas et al, “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation”, MedIA, 2017
Source code: https://github.com/deepmedic/deepmedic

Multi-Scale Processing

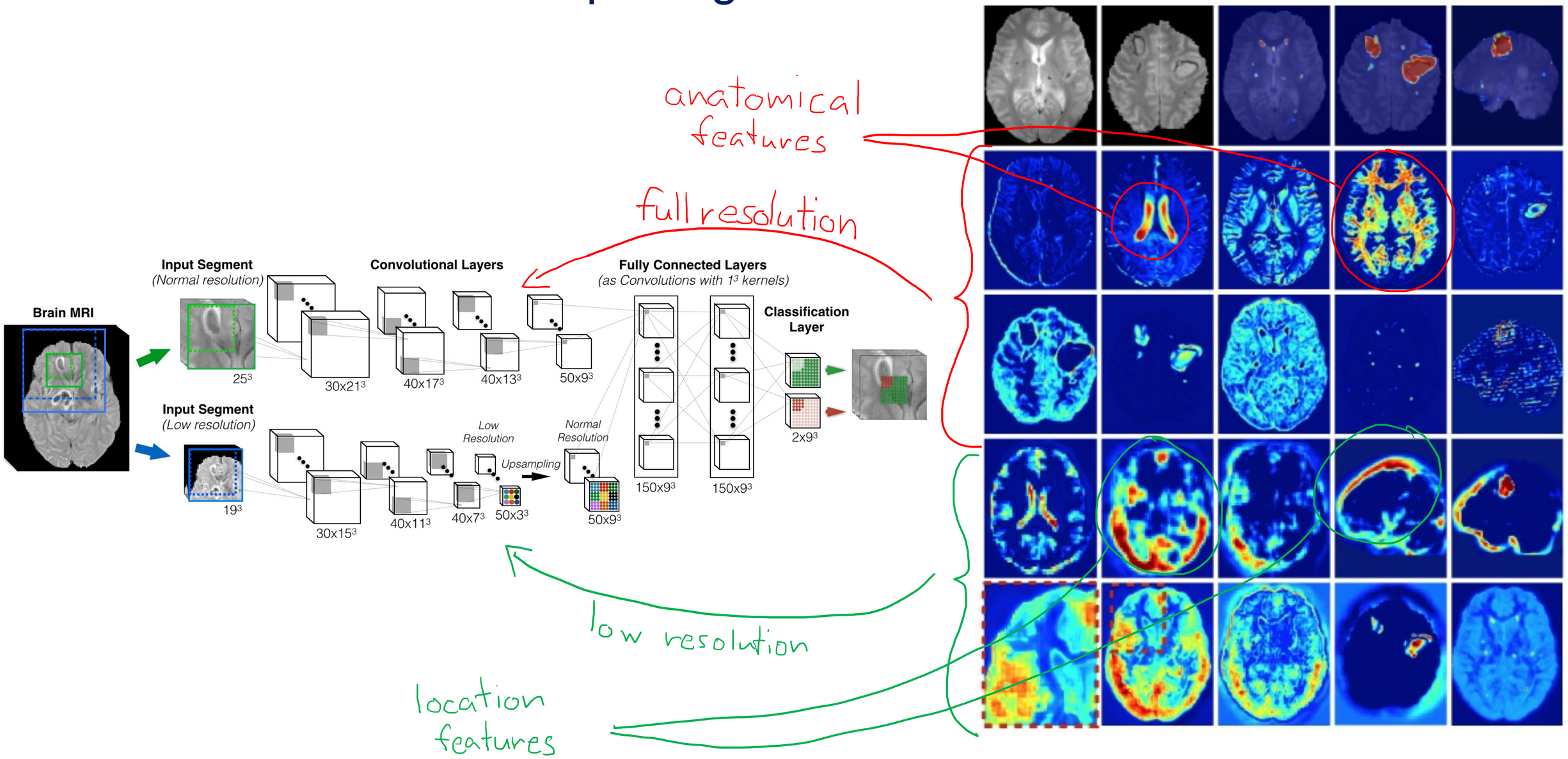


Additional spatial information enhances network's localization capabilities



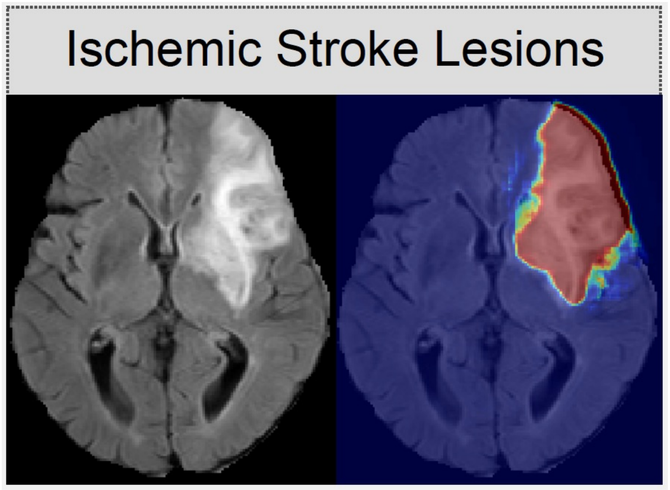


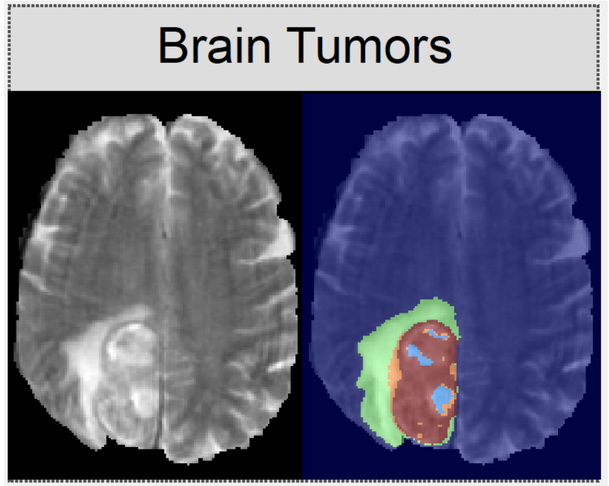
Interpreting Feature Maps

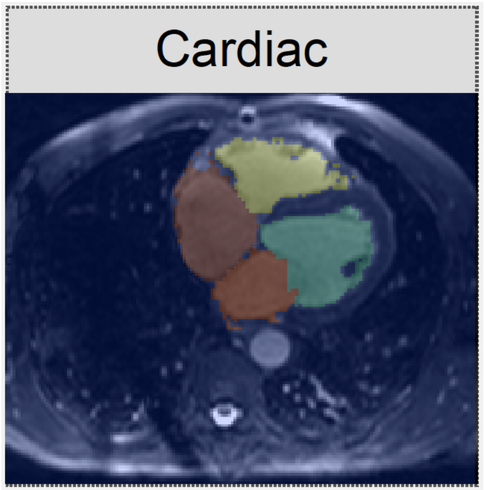


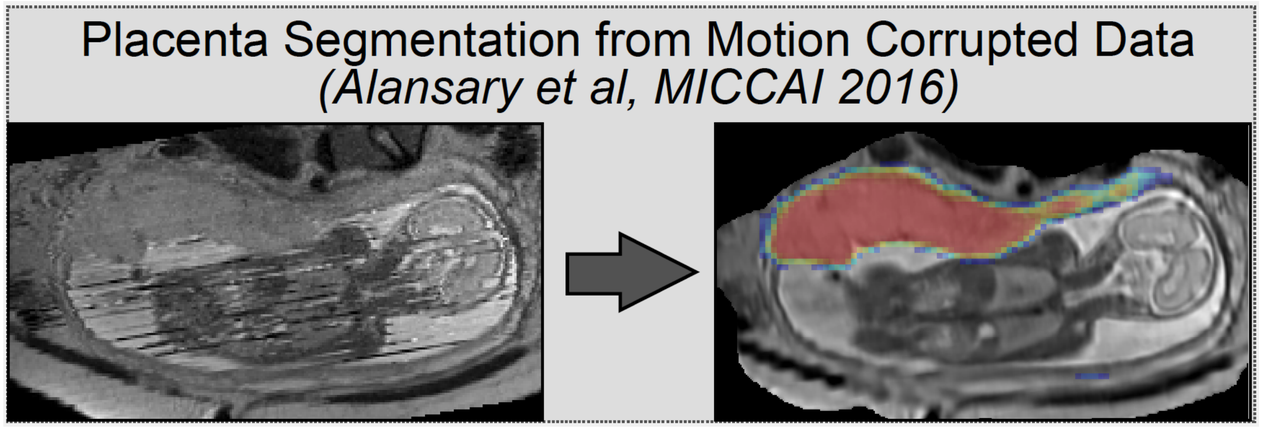
Kamnitsas et al, “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation”, MedIA, 2017

DeepMedic: A Versatile Segmenter









[1] Kamnitsas et al, “Multi-scale 3D convolutional neural networks for lesion segmentation in Brain MRI”, MICCAI–ISLES, 2015.
[2] Kamnitsas et al, “Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation”, MedIA, 2017.
[3] Alansary et al “Fast fully automatic of the human placenta from motion corrupted data”, MICCAI 2016.
[4] Giannakidis et al, “Fast fully automatic segmentation of the severely abnormal human right ventricle…”, SITIS 2016.
[5] Gordon et al, “Automatic 3D ultrasound segmentation of the first trimester placenta using deep learning”, ISBI 2017.

Recap: Common Image Analysis Tasks

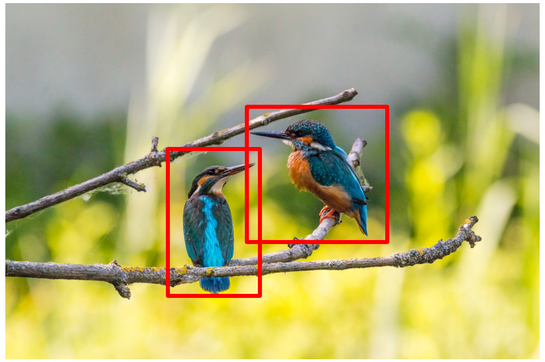
Image Classification

Output: Category (e.g., “bird”)

Object Detection

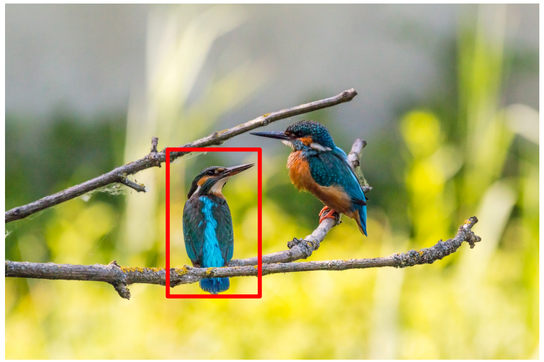
Output: Coordinates (e.g., centroid)

Object Localisation



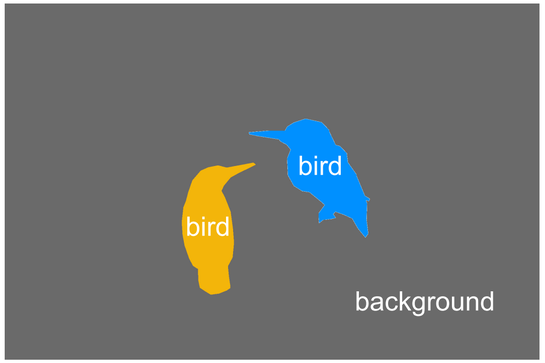
Output: Coordinates (e.g., bounding box)

Object Recognition



Output: Category (e.g., “kingfisher”)

Semantic Segmentation



Output: Label map

Image Captioning

Output: Text
(e.g., “two birds sitting on a branch”)

Summary

- **Medical Image Segmentation is hard**!
- Medical images are highly variable, especially in presence of disease
  – *We will talk about domain generalisation and data augmentation later in this course*
- **Don't forget to evaluate** – use many measures and not just the best
  – Tendency is just to report on DSC which however doesn't tell full story
  – *Beware that DSC makes a good loss function, and so is not objective evaluation criterion*
- **Deep learning based segmentation has taken over the field**:
  – **U-Net** and derivates most popular/successful techniques
  – Large, well annotated databases are needed for supervised methods – *we will talk about weakly supervised methods later in course*
  – Multi-task methods can help to further improve the segmentation performance – e.g. combine localisation, detection or classification with segmentation task
  – Some postprocessing for clean-up is still needed – *so do also look into some of the more traditional methods for that*

(At least for now!)