



Image Analysis

Tim B. Dyrby

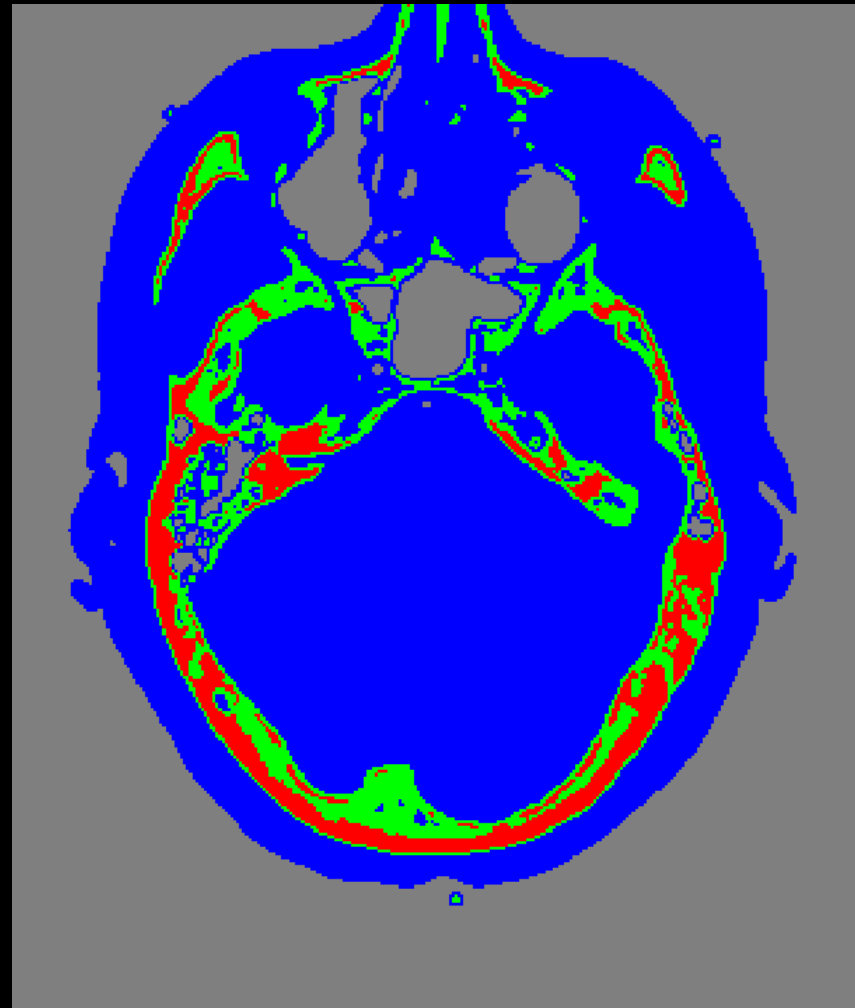
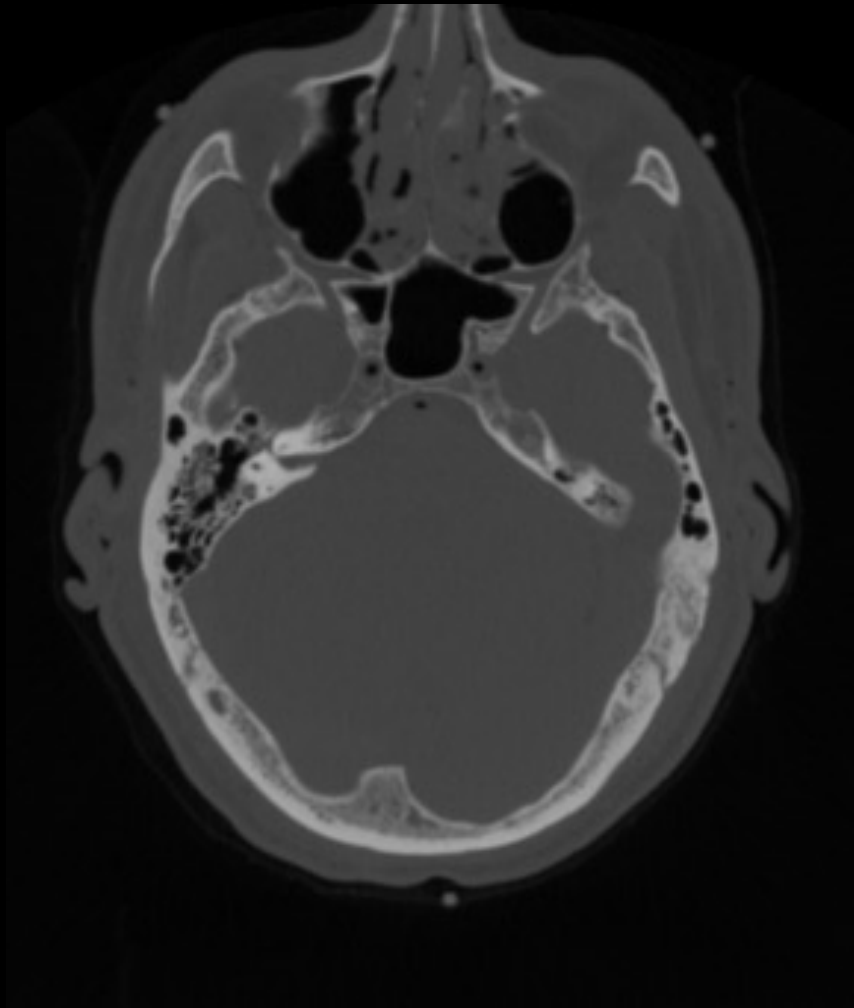
Rasmus R. Paulsen

DTU Compute

tbdy@dtu.dk

<http://www.compute.dtu.dk/courses/02502>

Lecture 8 – Pixel Classification and advanced segmentation





What can you do after today?

- Describe the concept of pixel classification
- Compute the pixel value ranges in a minimum distance classifier
- Implement and use a minimum distance classifier
- Approximate a pixel value histogram using a Gaussian distribution
- Implement and use a parametric classifier
- Decide if a minimum distance or a parametric classifier is appropriate based on the training data
- Explain the concept of Bayesian classification
- Implement and use the linear discriminant analysis (LDA) classifier
- Decide where to place a decision boundary
- Understand the use of linear vs non-line hyperplanes for segmentation



Go to www.menti.com and use the code 59 42 89 7

Quiz 0: What is advanced segmentation?

0	0	0	0
To separate colours?	Use methods that mimics the human brain?	It just some vectors pointing in a space?	To draw linear and non-linear hyper plans in space

Classification

- Take a measurement and put it into a class

Measurement

Classes



Wheels: 2

HP: 50

Weight: 200

Classifier



- Bike
- Truck
- Car
- Motorbike
- Train
- Bus

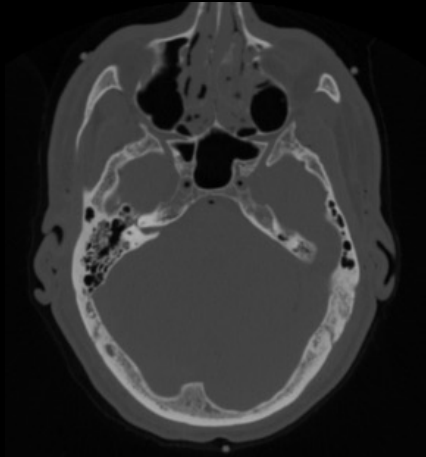


General Classification

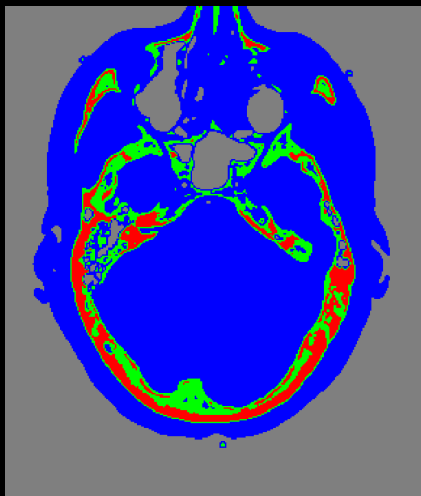
- Multi-dimensional measurement
- Pre-defined classes
 - Can also be found automatically – can be very difficult!

Pixel Classification

CT scan of human head



Pixel wise segmentation



Four Class labels

Background

Soft-Tissue

Trabecular Bone

Hard Bone

- Classify each pixel
 - Independent of neighbours
- Also called labelling
 - Put a label on each pixel
- We look at the pixel value and assign them a label
- Labels already defined



Quiz 1: Two class pixel classification?

Background and object

- A) Median filter
- ☒ B) Threshold
- C) Brightness
- D) Morphological Erosion
- E) BLOB analysis



Pixel Classification – formal definition

Pixel value (the measurement) $v \in R$

k classes

$$C = c_1, \dots, c_k$$

Classification rule

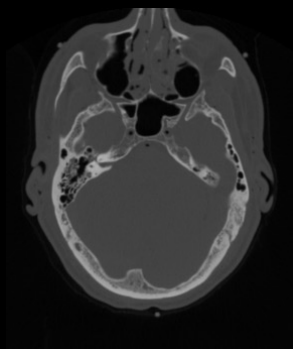
$$c: R \longrightarrow \{c_1, \dots, c_k\}$$

Pixel Classification – example

Pixel value $v \in [0,255]$

Set of 4 classes $C = \{\text{background, soft-tissue, trabeculae, bone}\}$

Classification rule $c: v \rightarrow \{\text{background, soft – tissue, trabeculae, bone}\}$



How do we construct a classification rule?

Pixel classification rule

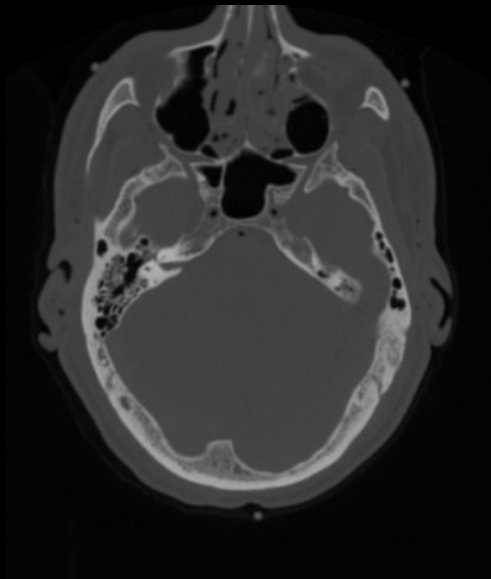
$c: v \rightarrow \{\text{background, soft - tissue, trabeculae, bone}\}$

background

trabeculae

soft-tissue

bone



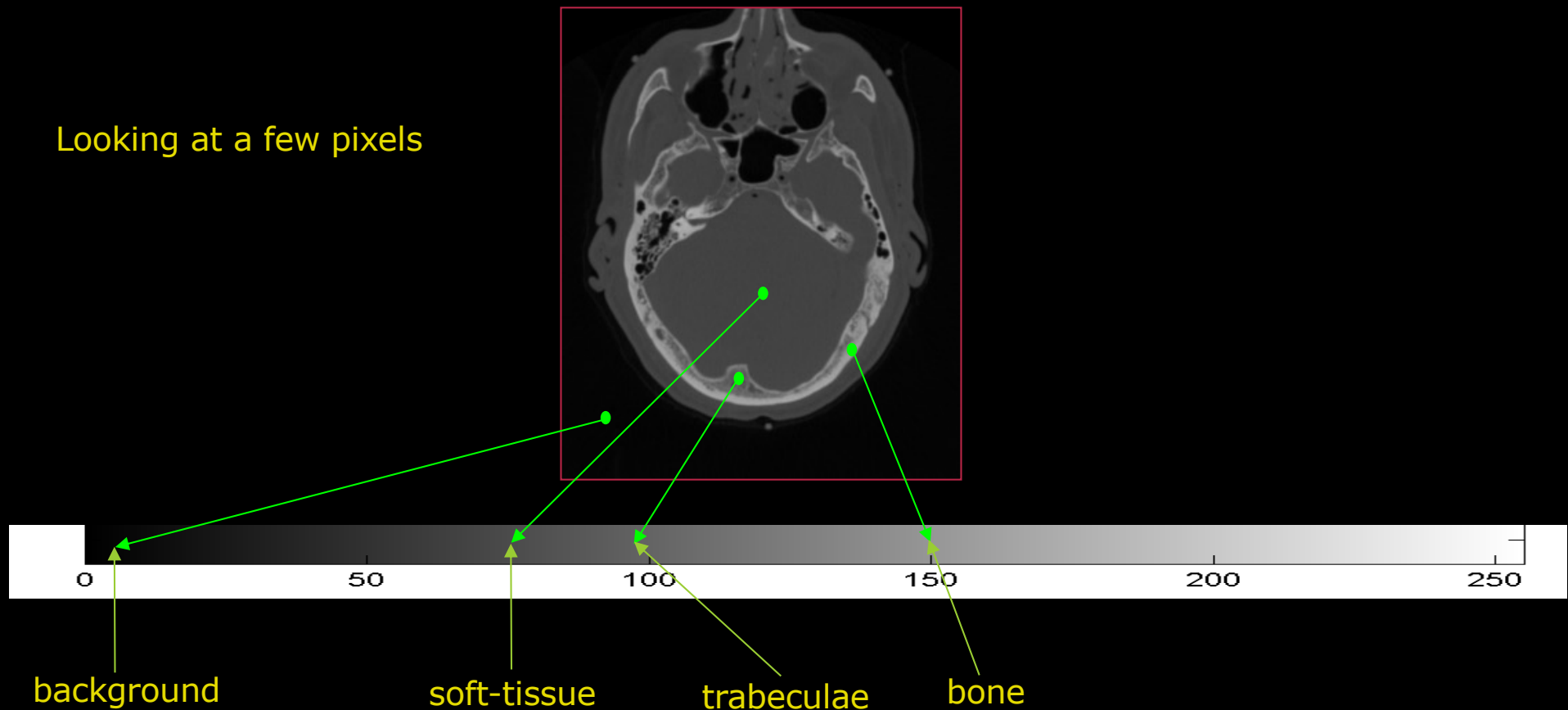
How do we do this?



Pixel classification rule – manual inspection

$c: v \rightarrow \{\text{background, soft – tissue, trabeculae, bone}\}$

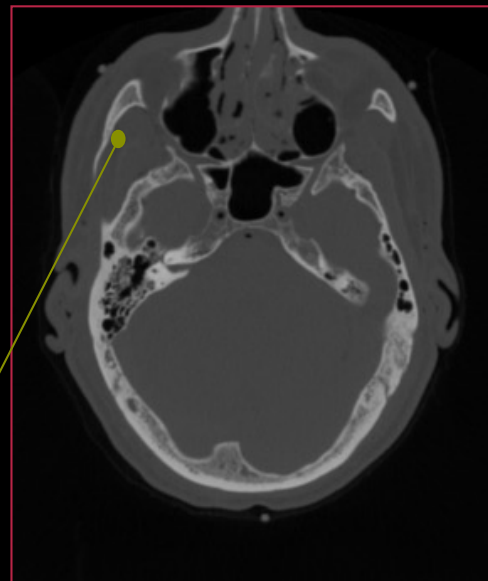
Looking at a few pixels



Pixel classification rule – manual inspection

$c: v \rightarrow \{\text{background, soft – tissue, trabeculae, bone}\}$

Looking at some few pixels



New pixel – where do we put it?



background

soft-tissue

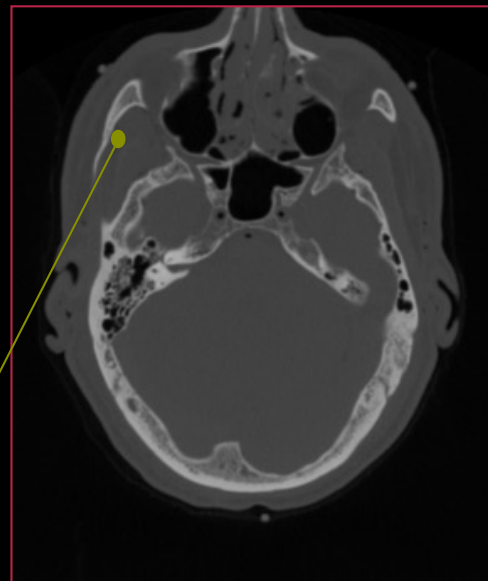
trabeculae

bone

Pixel classification rule – manual inspection

$c: v \rightarrow \{\text{background, soft – tissue, trabeculae, bone}\}$

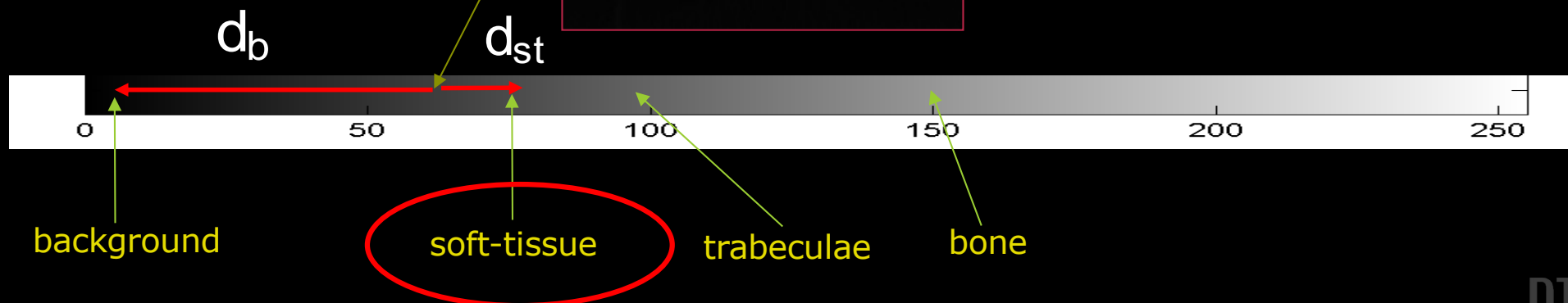
Looking at some few pixels



New pixel – where do we put it?

- Measure the “distance” to the other classes
- Select the closest class

Minimum distance classification

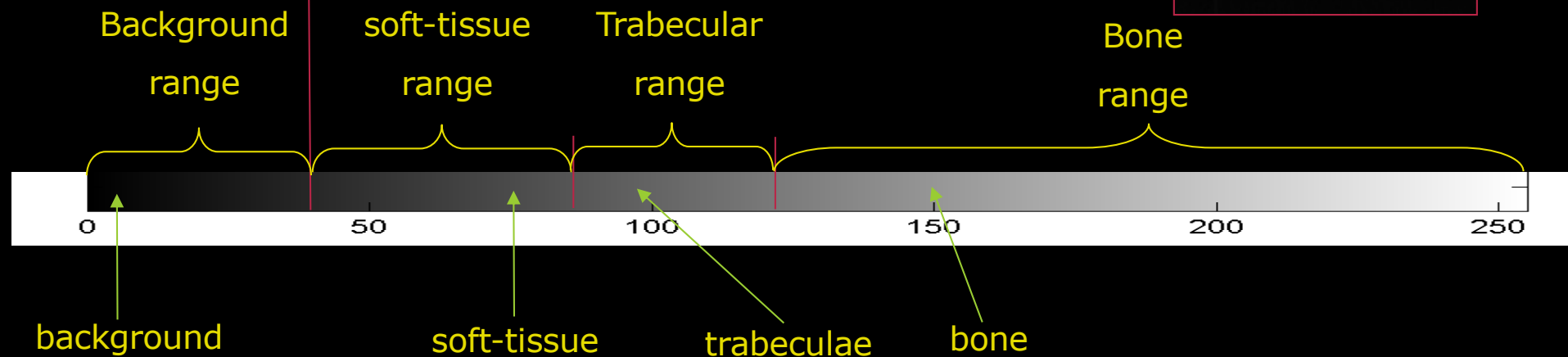


Pixel classification rule

Minimum Distance Classification

The possible pixel values are divided into ranges

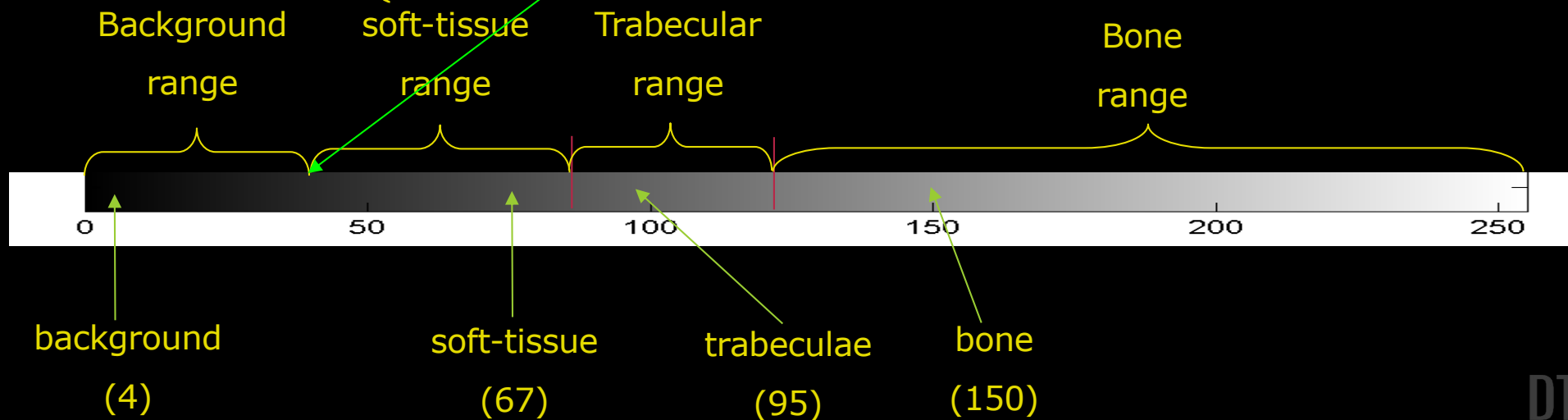
Here the distance to "background" is equal to "soft-tissue"



Pixel classification rule

Minimum Distance Classification

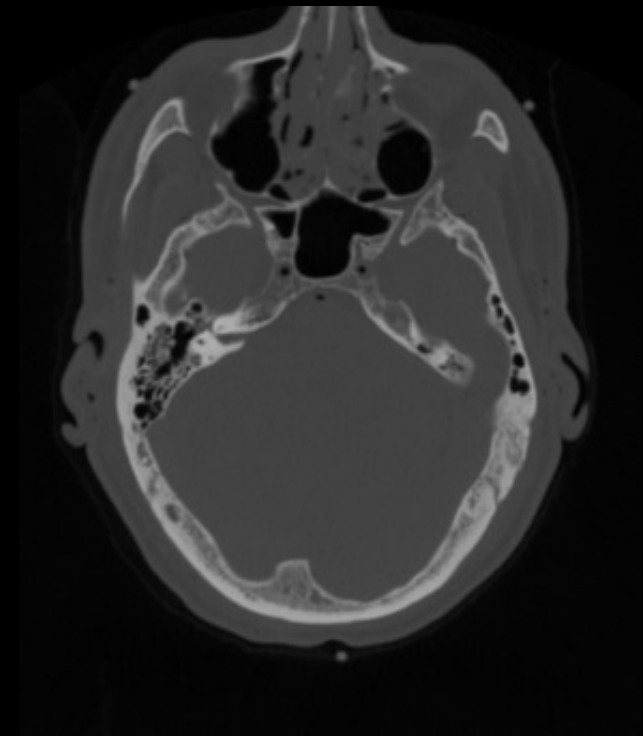
$$c(v) = \begin{cases} \text{background, if } v \leq (4 + 67)/2 \\ \text{soft - tissue, if } \frac{(4 + 67)}{2} < v \leq \frac{67 + 95}{2} \\ \text{trabeculae, if } \frac{67 + 95}{2} < v \leq \frac{95 + 150}{2} \\ \text{bone, if } v > \frac{95 + 150}{2} \end{cases}$$



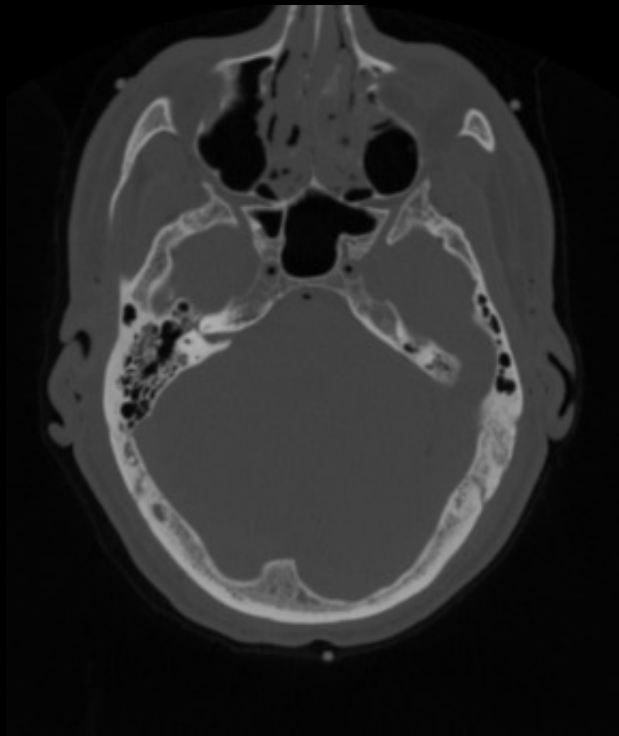
Pixel classification rule

■ For all pixel in the image do

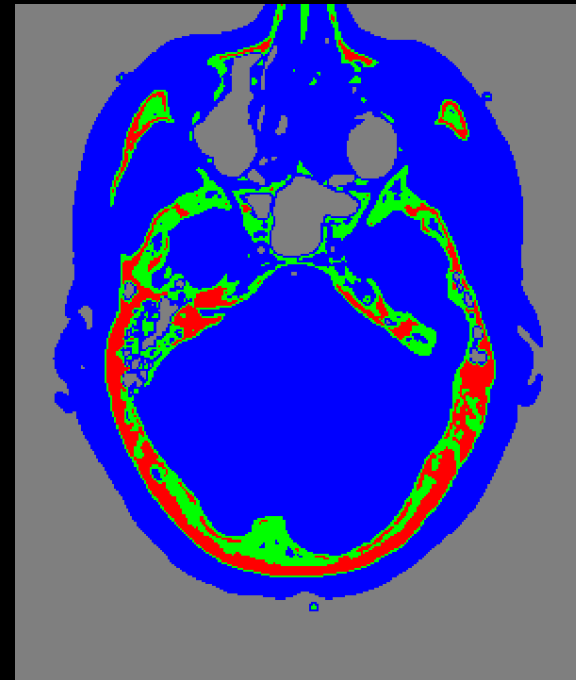
$$c(v) = \begin{cases} \text{background, if } v \leq (4 + 67)/2 \\ \text{soft - tissue, if } \frac{(4 + 67)}{2} < v \leq \frac{67 + 95}{2} \\ \text{trabeculae, if } \frac{67 + 95}{2} < v \leq \frac{95 + 150}{2} \\ \text{bone, if } v > \frac{95 + 150}{2} \end{cases}$$



Pixel Classification example

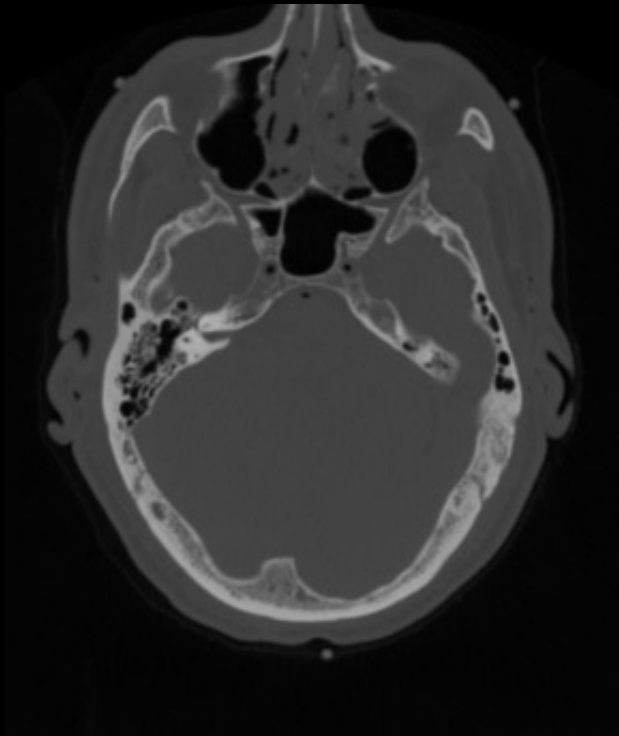


CT scan of human head



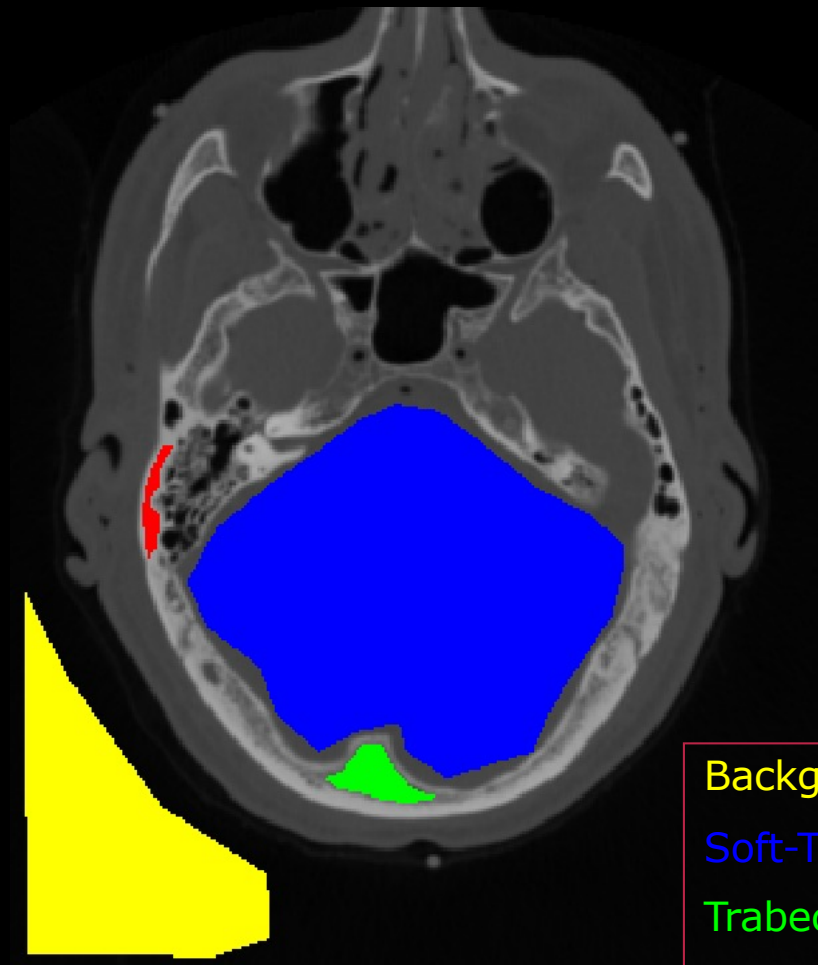
Background
Soft-Tissue
Trabecular Bone
Hard Bone

Better range selection



- Guessing range values is not a good idea
- Better to use “training data”
- Start by selecting representative regions from an image
- *Annotation*
 - To mark points, regions, lines or other significant structures

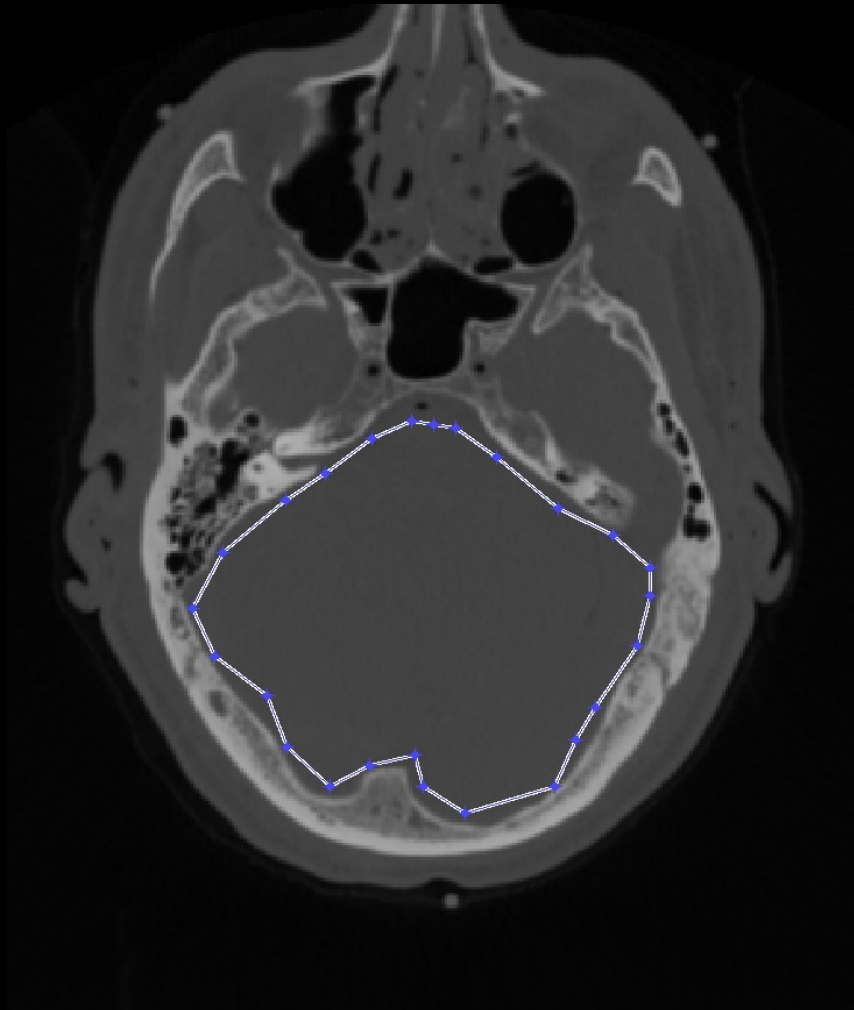
Classifier training - annotation



- An “expert” is asked how many different tissue types that are possible
- Then the expert is asked to mark representative regions of the selected tissue types

Background
Soft-Tissue
Trabecular Bone
Hard Bone

Classifier training – region selection

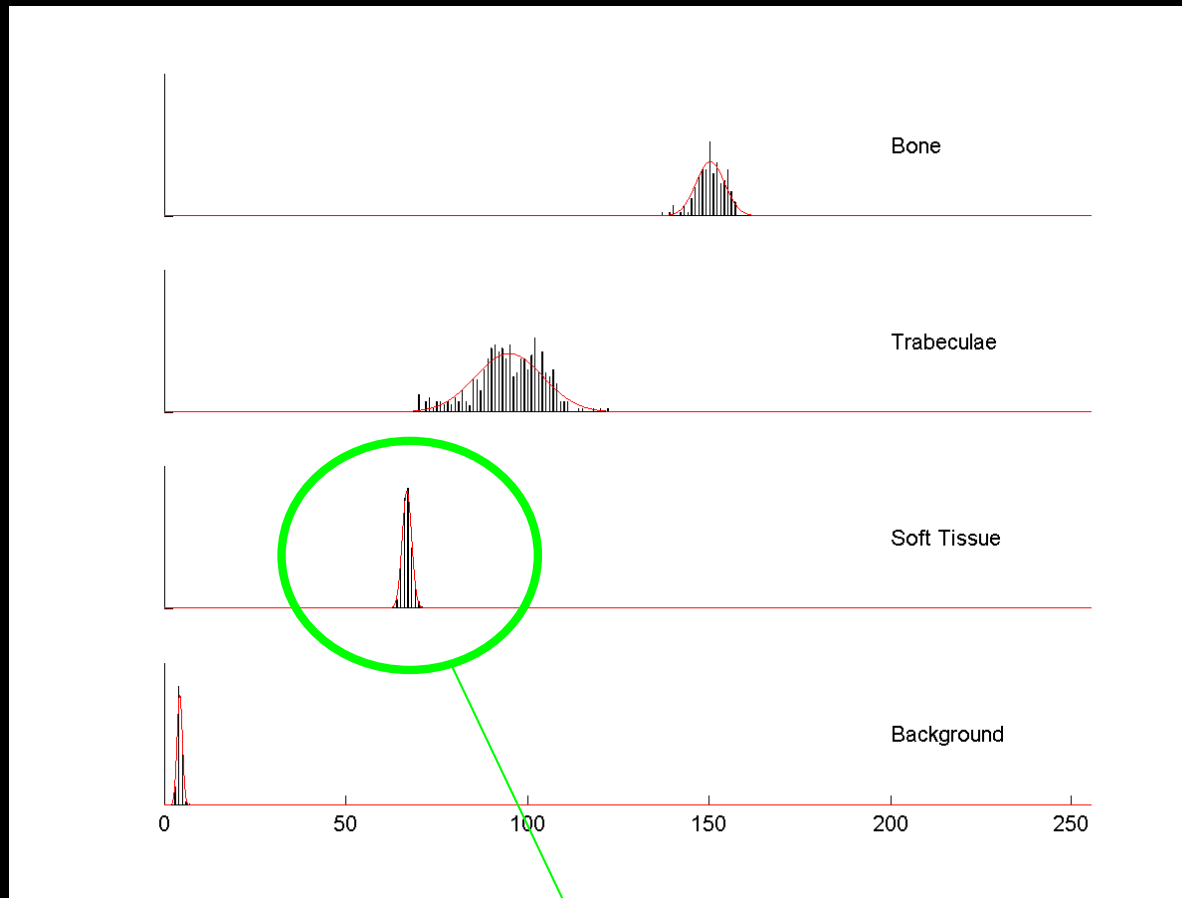


- Many tools exist
- Python module `roipoly`
 - Select closed regions using a piecewise polygon

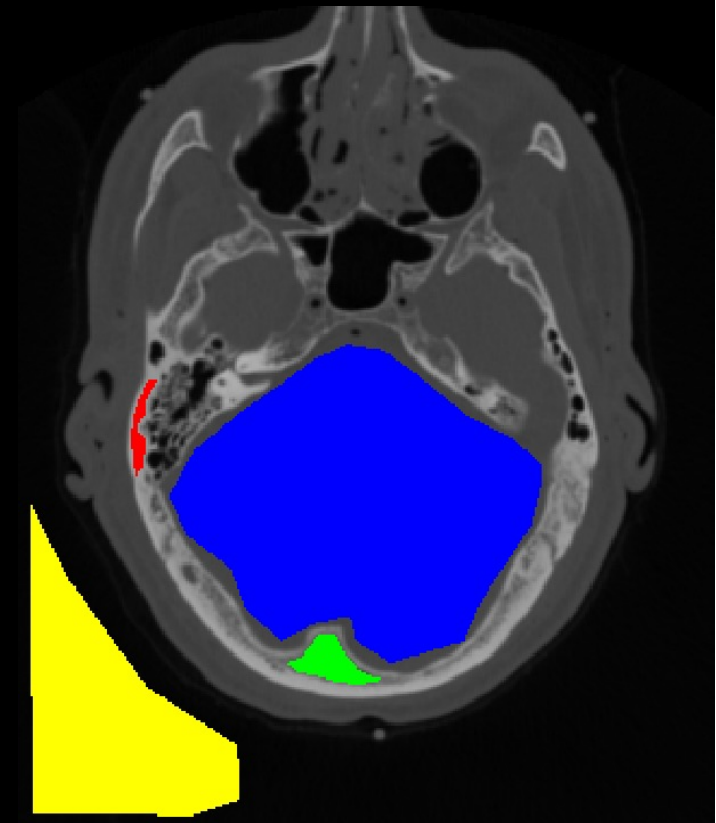
Training is only done once!

Optimally, the training can be used on many pictures that contains the same tissue types

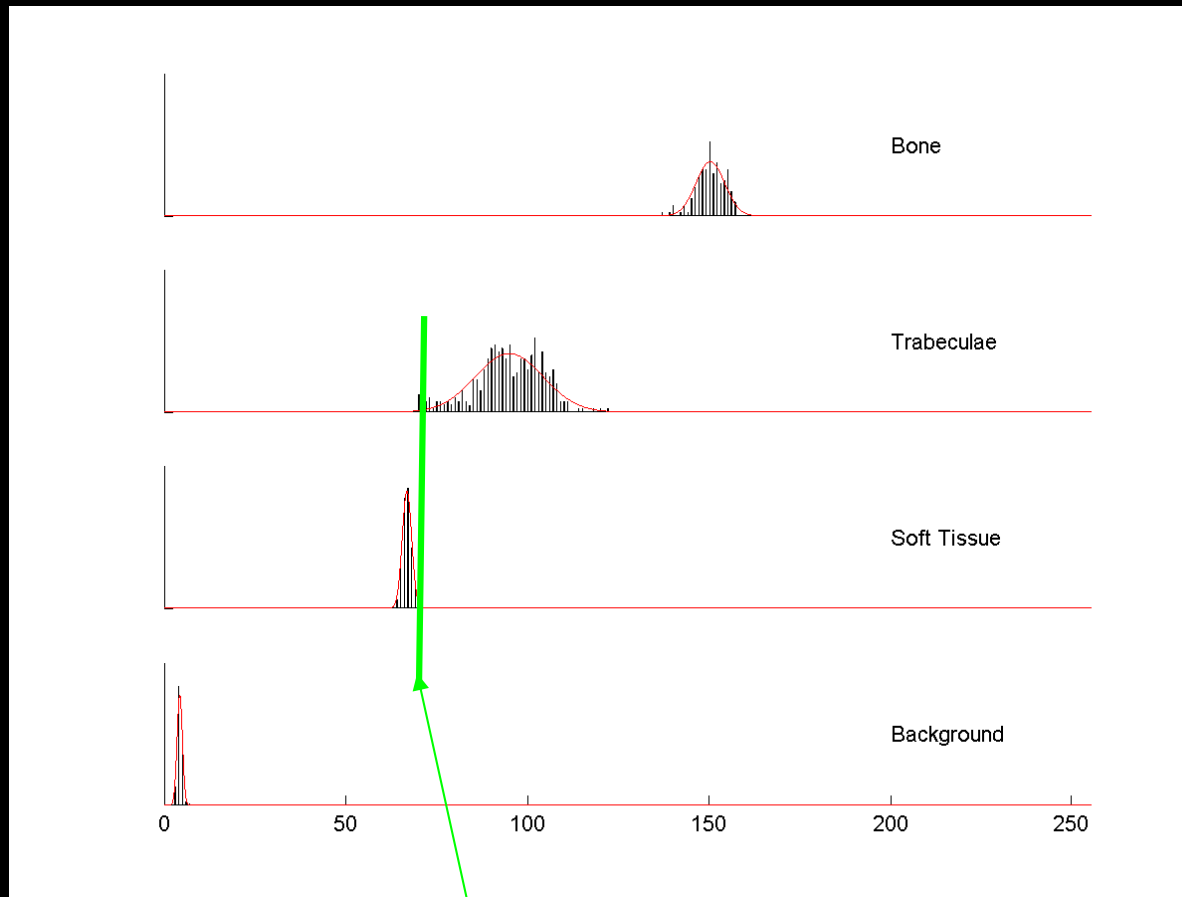
Initial analysis - histograms



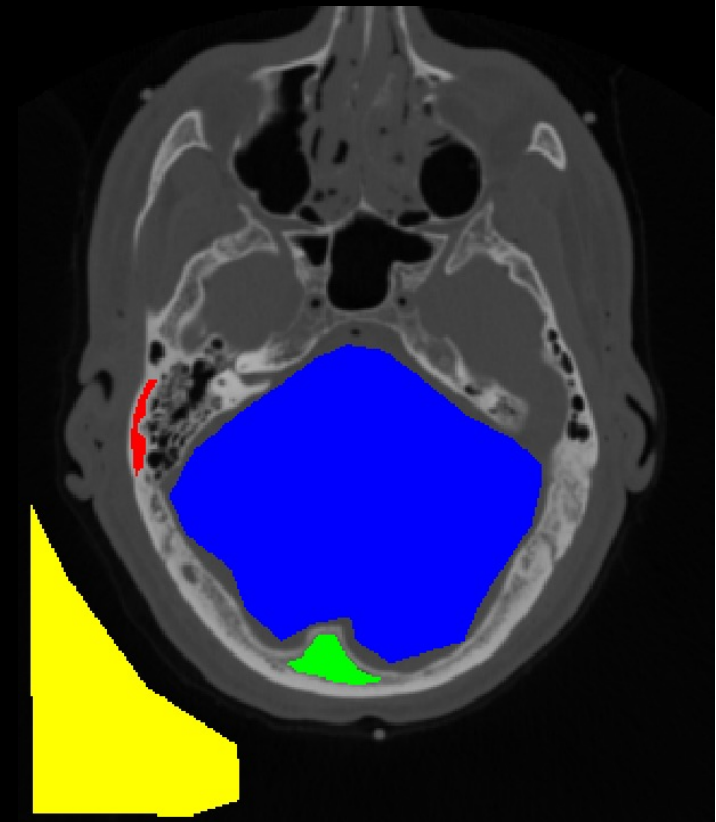
Gaussian



Initial analysis - histograms

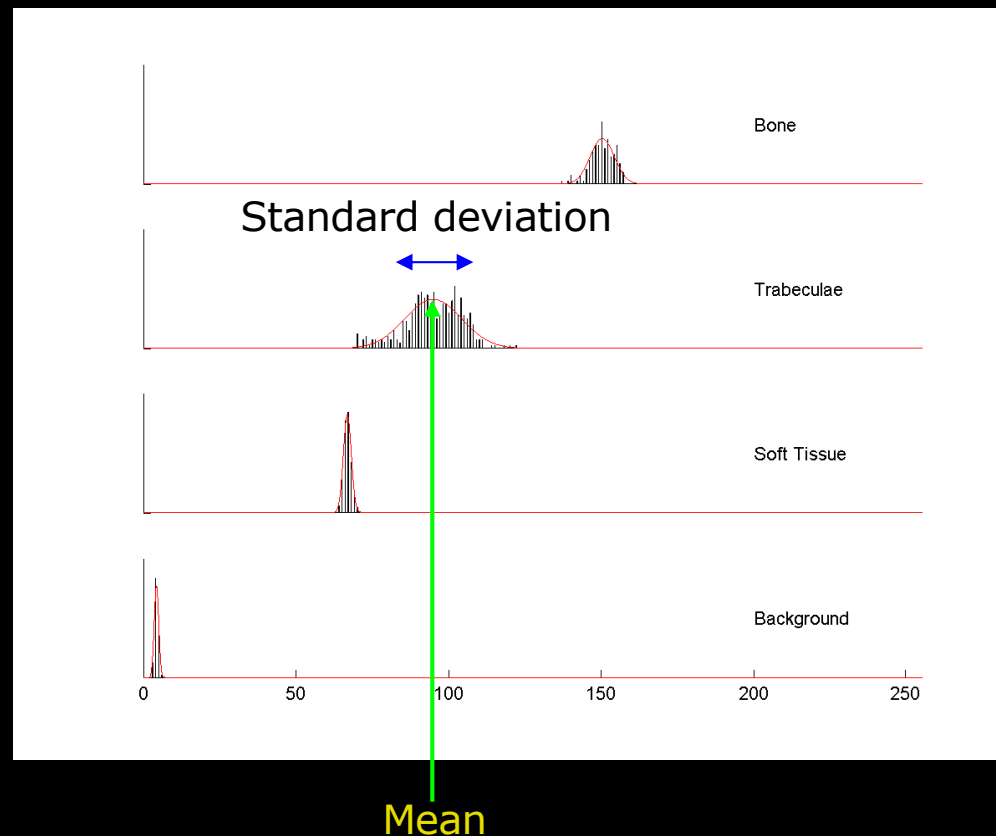


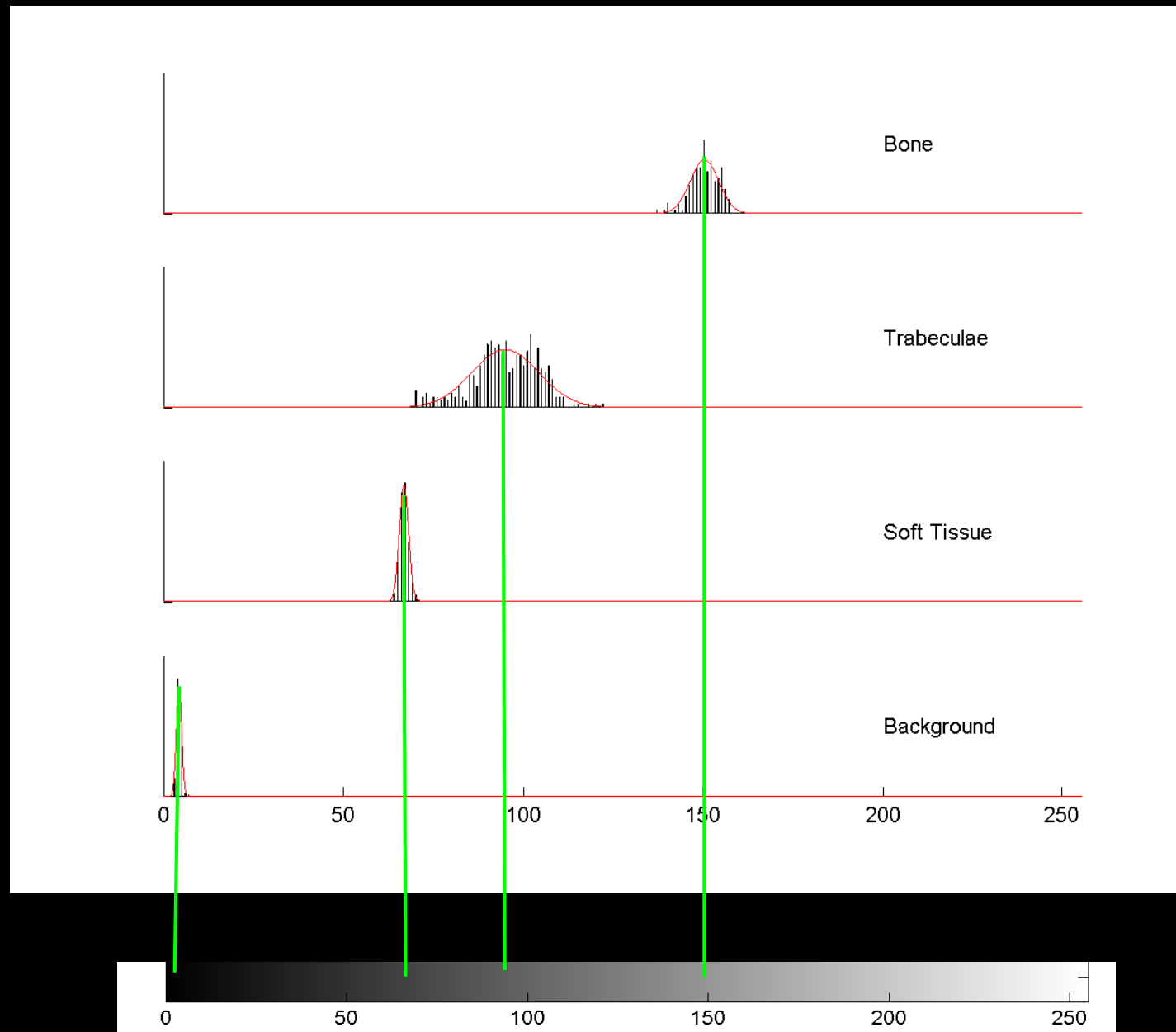
Class separation



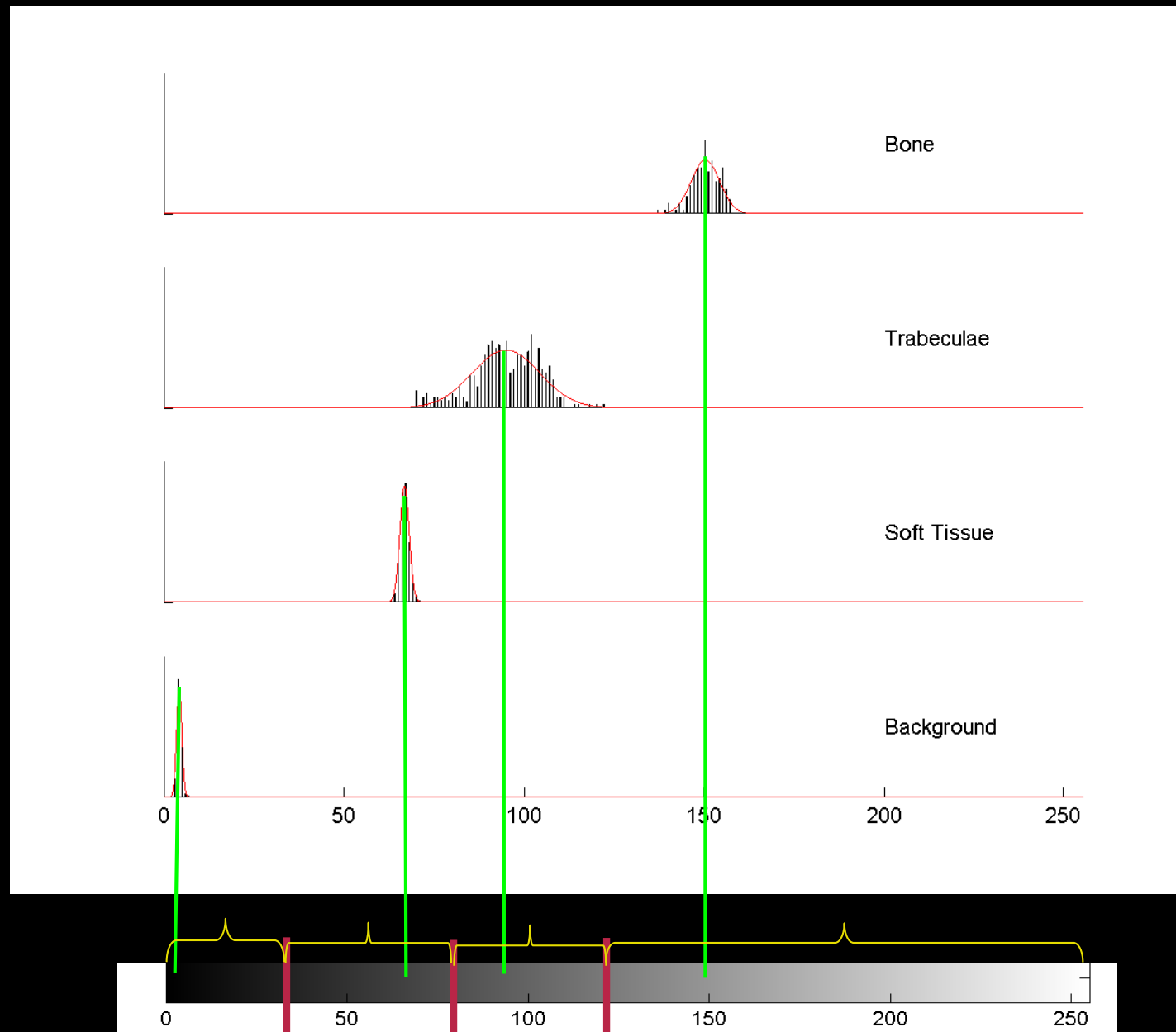
Simple pixel statistics

- Calculate the mean and the standard deviation of each class





Minimum distance classification



Any objections?

The pixel value ranges are not always in good correspondence with the histograms?

Quiz 2: Minimum distance classification

- A) Background
- B) Soft tissue
- ☒ C) Fat
- D) Bone
- E) None of the above

Solution:

Green: $(6+4+9+5)/4=6$

Blue: $(132+130+134+133)/4=132,25$

Yellow: $(178+175+176+174)/4=175,75$

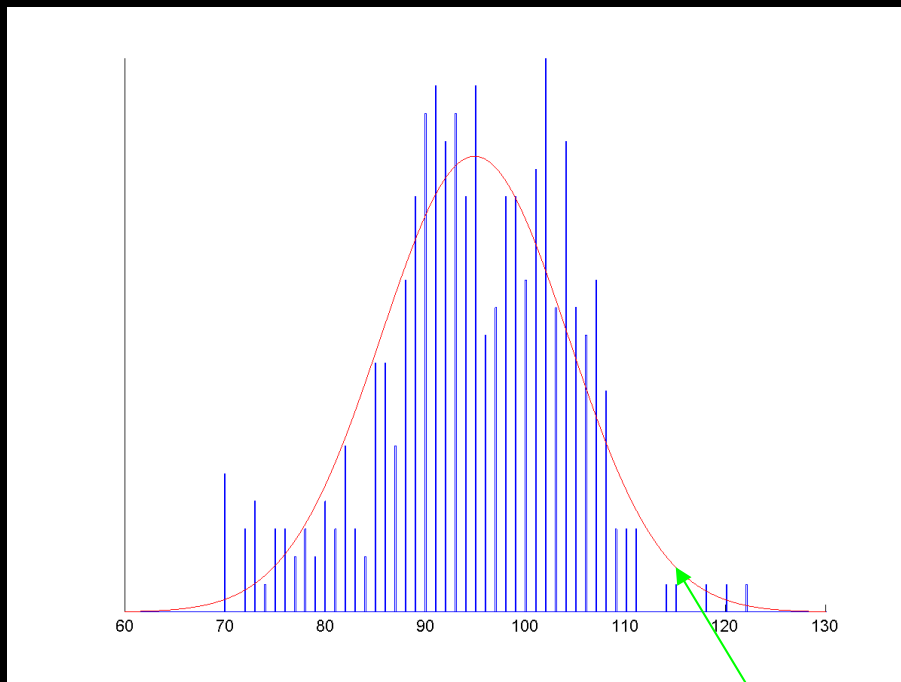
Purple: $(222+220+219+221)/4=220$

Blue: 158 is closest to 175,75 (yellow)= fat

To make a pixel classification an expert has selected representative regions in the image. They contain background (green), soft tissue (blue), fat (yellow), and bone (purple). The goal is to classify the pixel marked with a light blue circle. Using a minimum distance classifier it is classified as?

5	6	5	81	180	182	222	220
8	9	4	108	181	175	219	221
7	8	132	130	148	182	174	223
58	231	134	133	61	173	178	175
44	250	181	130	117	101	176	174
5	6	7	204	246	94	86	175
156	158	6	7	7	252	173	230
157	161	7	6	6	10	35	227

Parametric classification



Trabecular bone

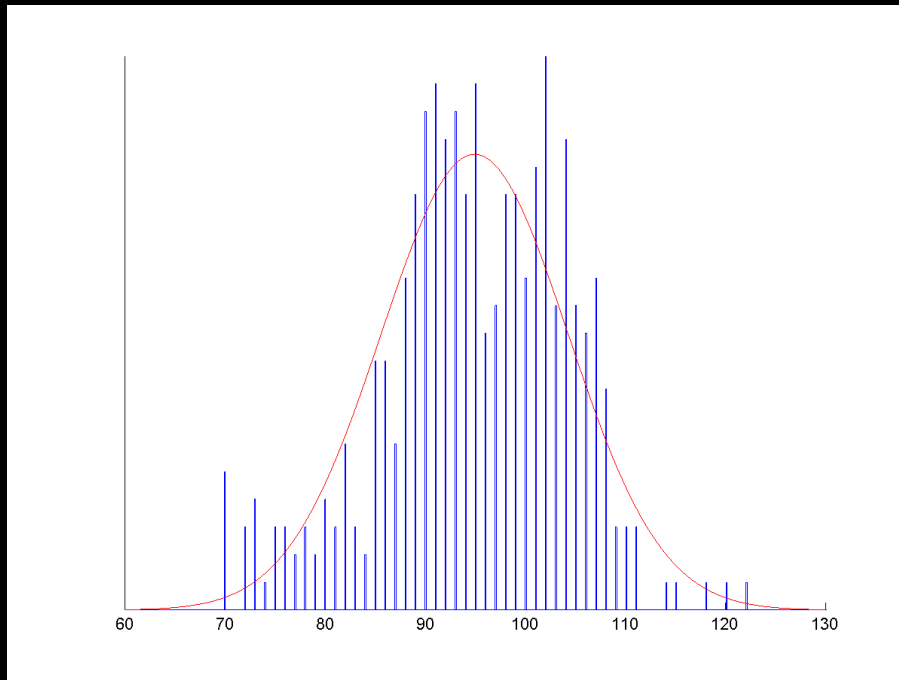
Only two values needed

- Describe the histogram using a few parameters
- Assume a “model” describing the signal values
- Model: Gaussian/Normal distribution

- The mean μ
- Standard deviation σ
- $\mathcal{N}(\mu, \sigma)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

Parametric classification



Trabecular bone

Training pixel values
(Belonging to one class) v_1, v_2, \dots, v_n

Estimated mean
$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n v_i$$

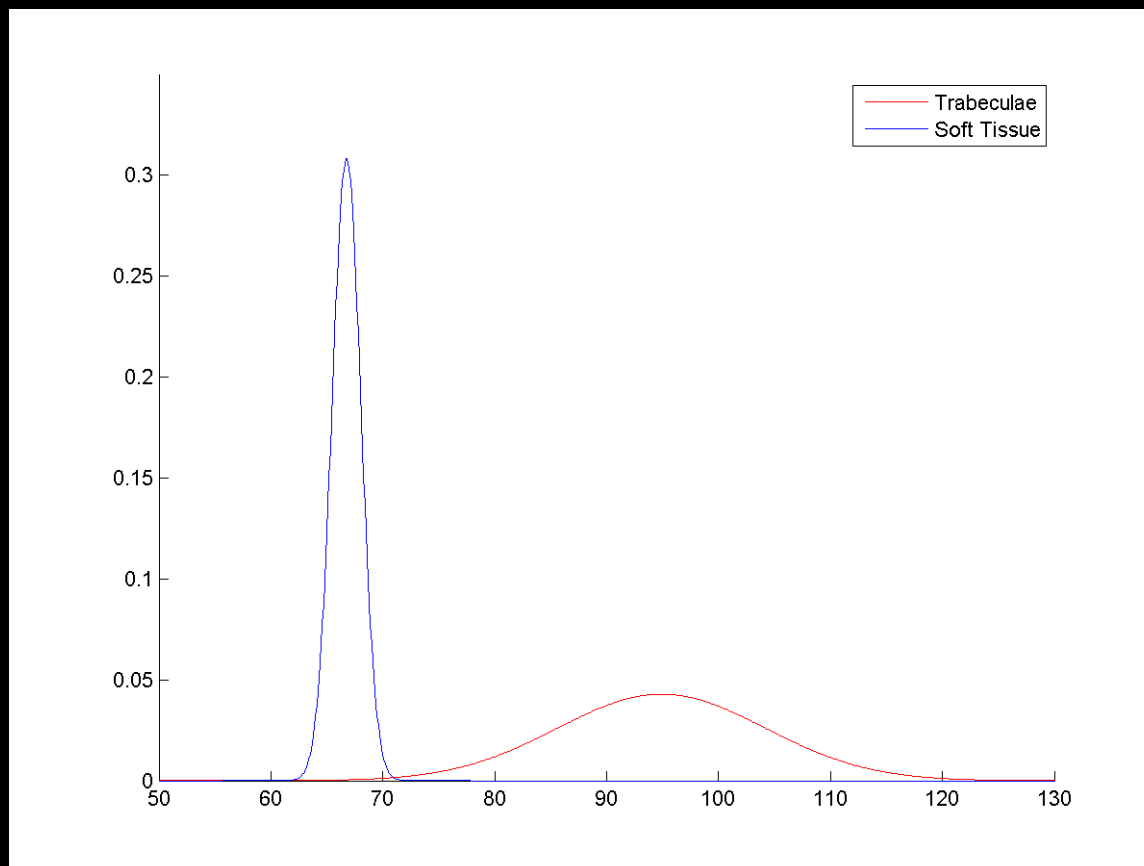
Estimated standard deviation
$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (v_i - \hat{\mu})^2}$$

The "signal model" is a Gaussian distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

Parametric classification

- Fit a Gaussian to the training pixels for all classes



What do we see here?

What is the difference between the two classes?

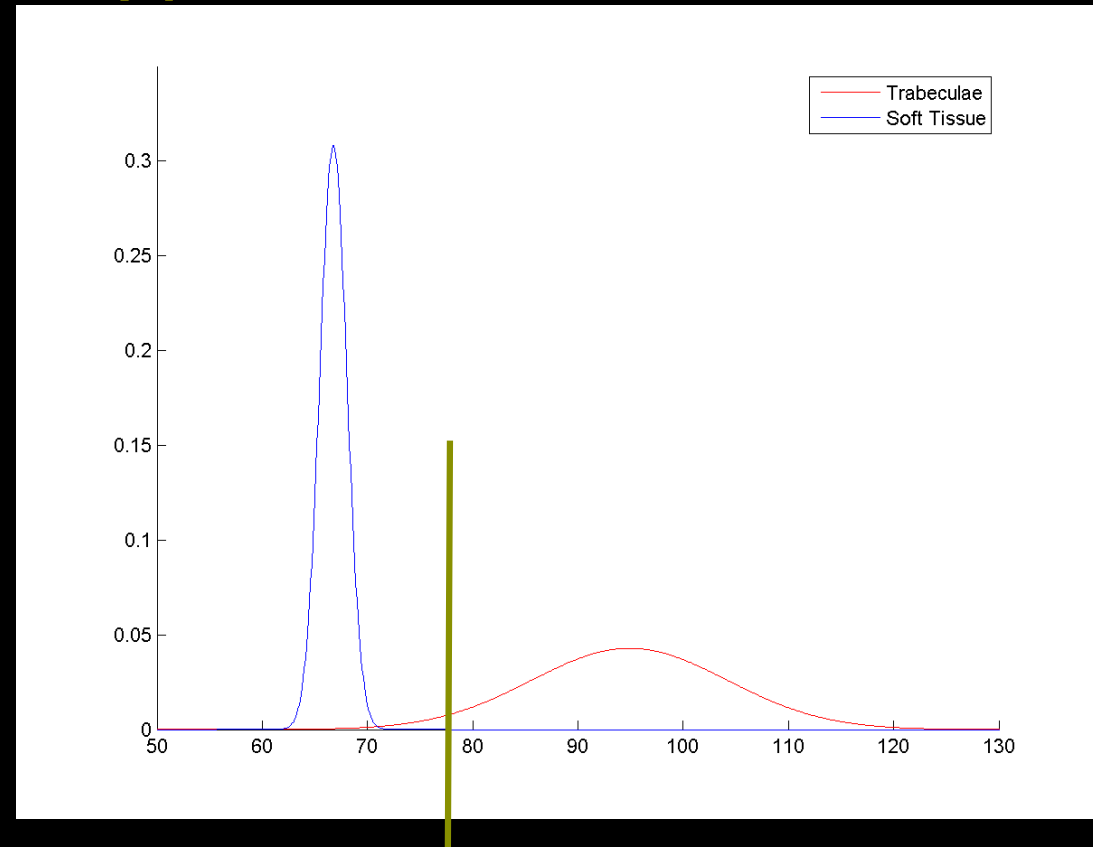
Trabeculae has much higher variation in the pixel values

Quiz 3: Two tissue types – minimum distance

$v = 78$

Which tissue class?

- A) Trabeculae
- B) Soft-tissue**



$v = 78$

Solution: Minimum distance classifier

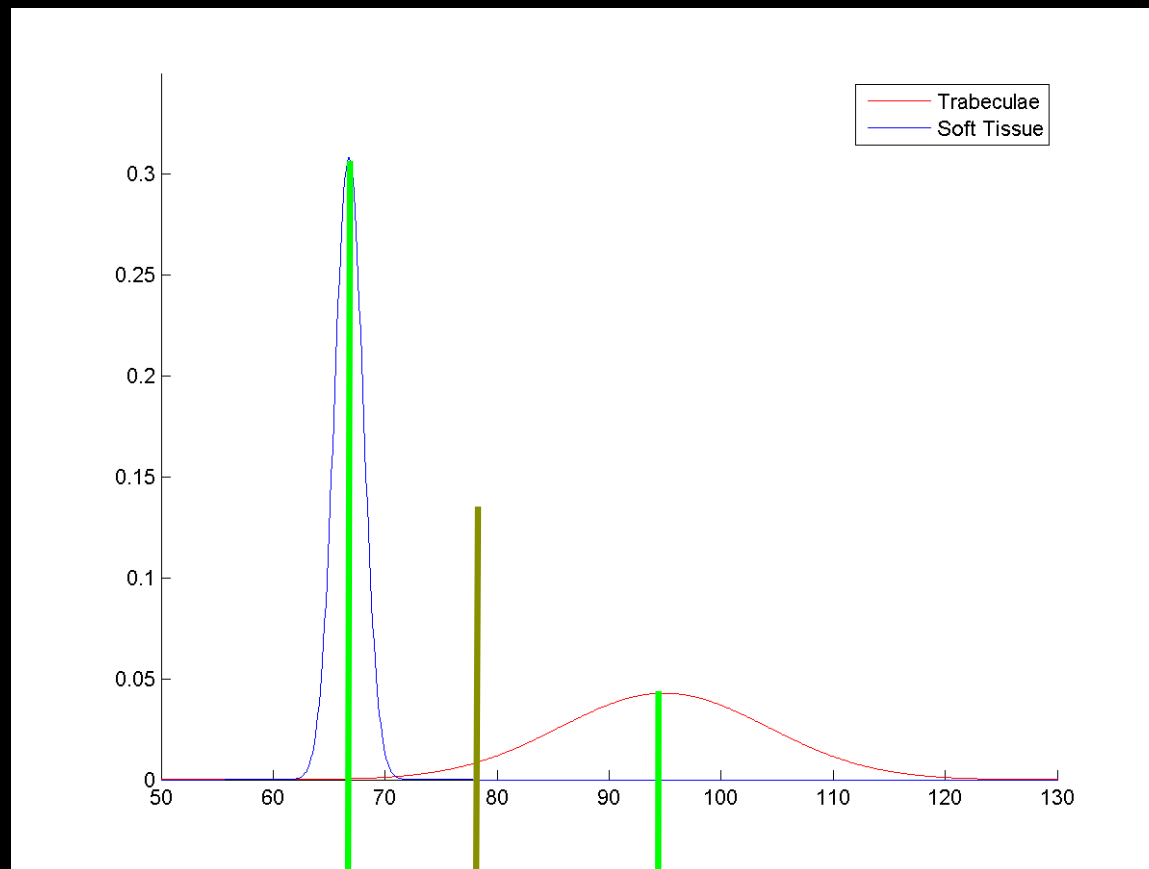
First we find the threshold, T :

B: mean(Soft Tissue)=68 and A: mean(Trabeculae)= 95

$$T = (95+68)/2 = 81,5$$

Then we classify/segment $v=78$: A if $v > 81,5$ or B if $v < 81,5$

Parametric classification



$v = 78$

- New pixel with value 78
 - Is it soft-tissue or trabecular bone?
- Minimum distance classifier?
 - Soft-tissue
- Is that fair?
 - Soft-tissue Gaussian says “Extremely low probability that this pixel is soft-tissue”

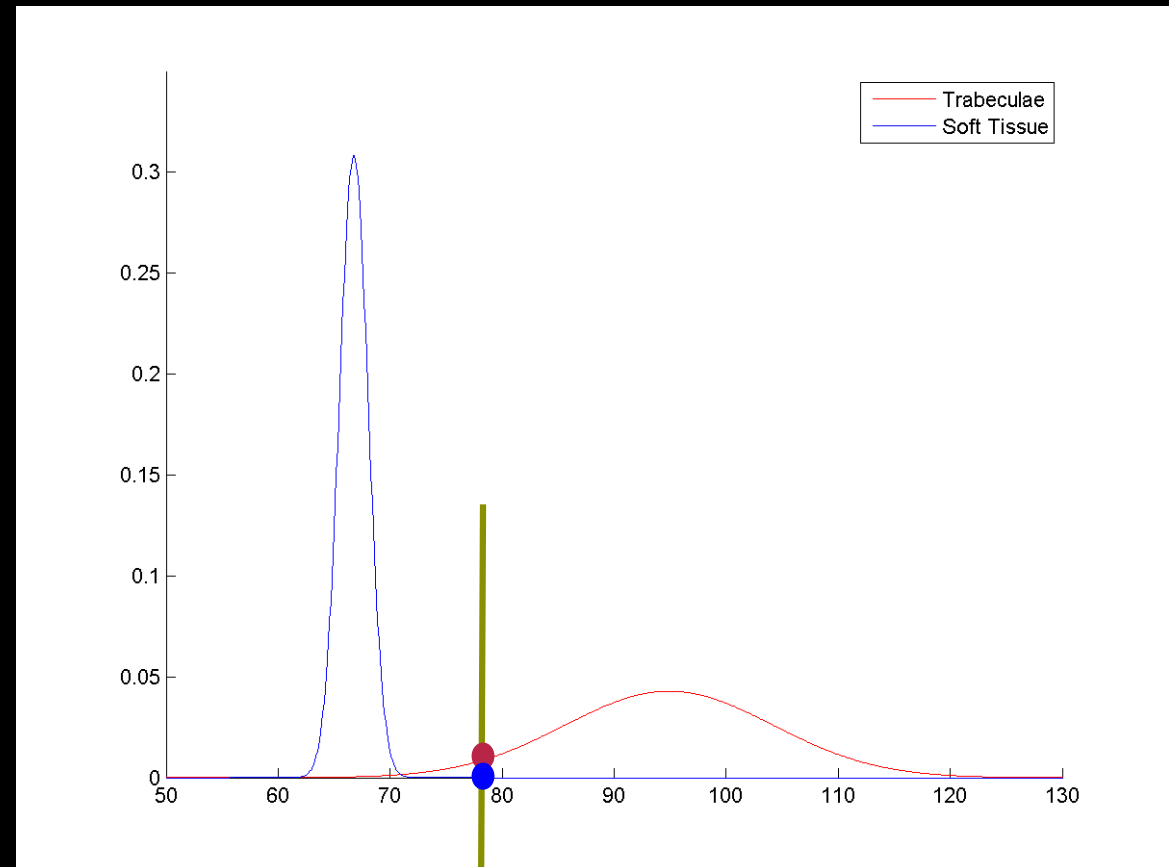
Quiz 4: Two tissue types – parametric classification

Which tissue class?

- A) Trabeculae**
- B) Soft-tissue

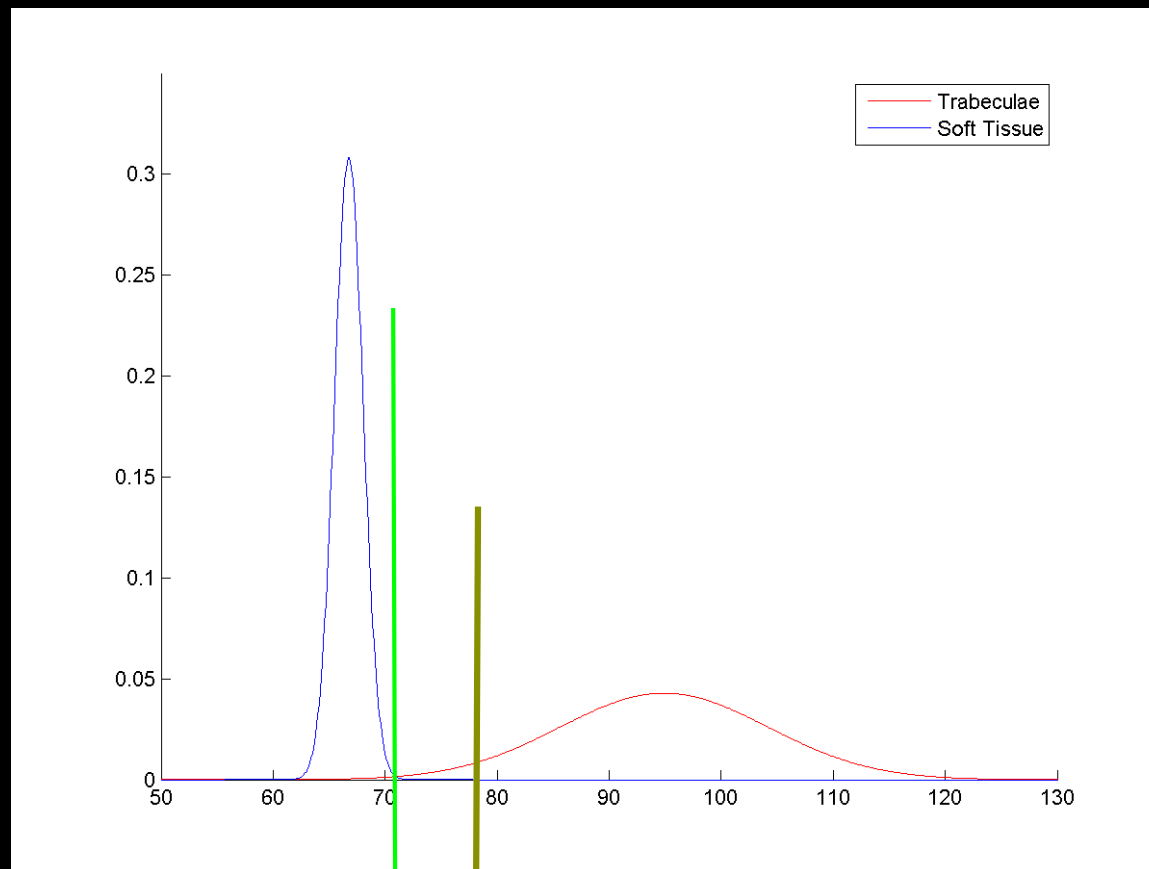
Solution:

The A distribution (red) is higher than B (blue) at $v=78$



$v = 78$

Parametric classification – repeat the question



$v = 78$

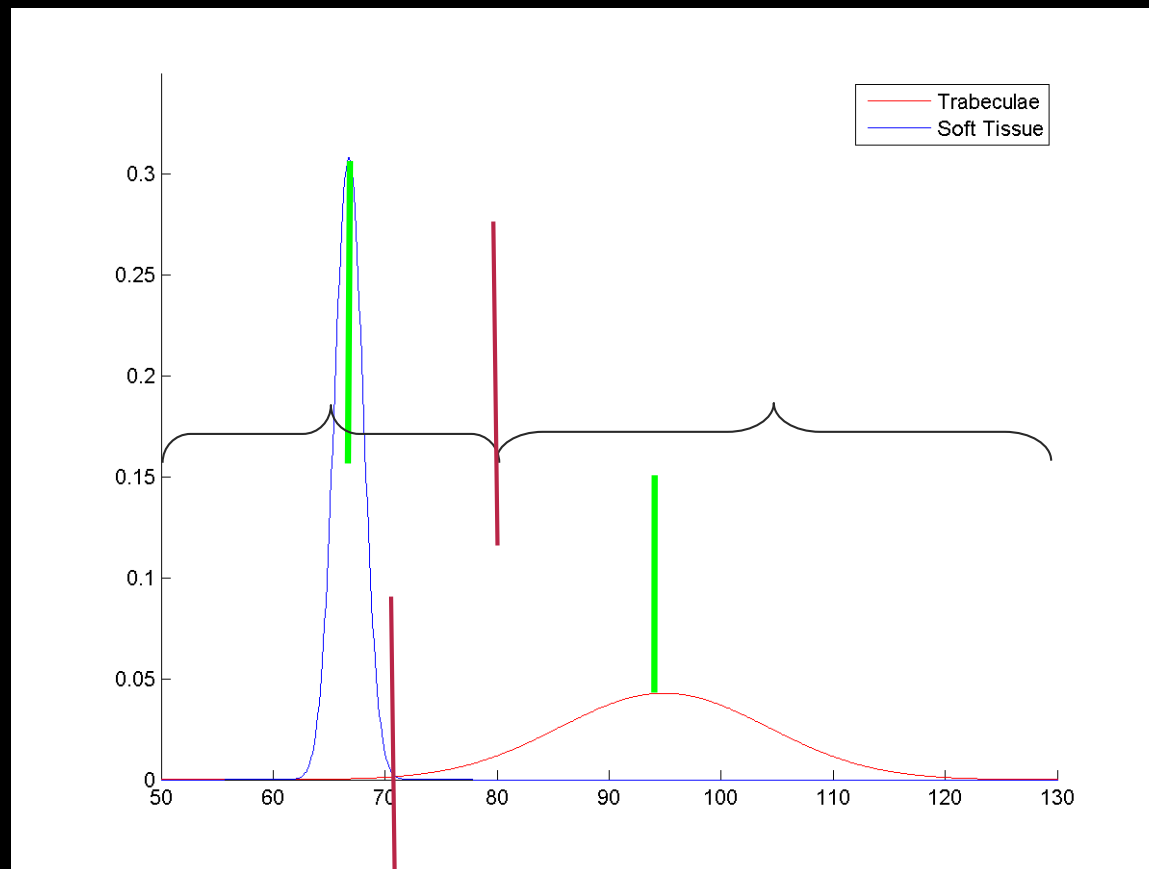
- New pixel with value 78

- Is it soft-tissue or trabecular bone?
- Most probably trabecular bone

- Where should we set the limit?

- Where the two Gaussians cross!

Parametric classification – ranges



Soft-tissue

Trabecular bone

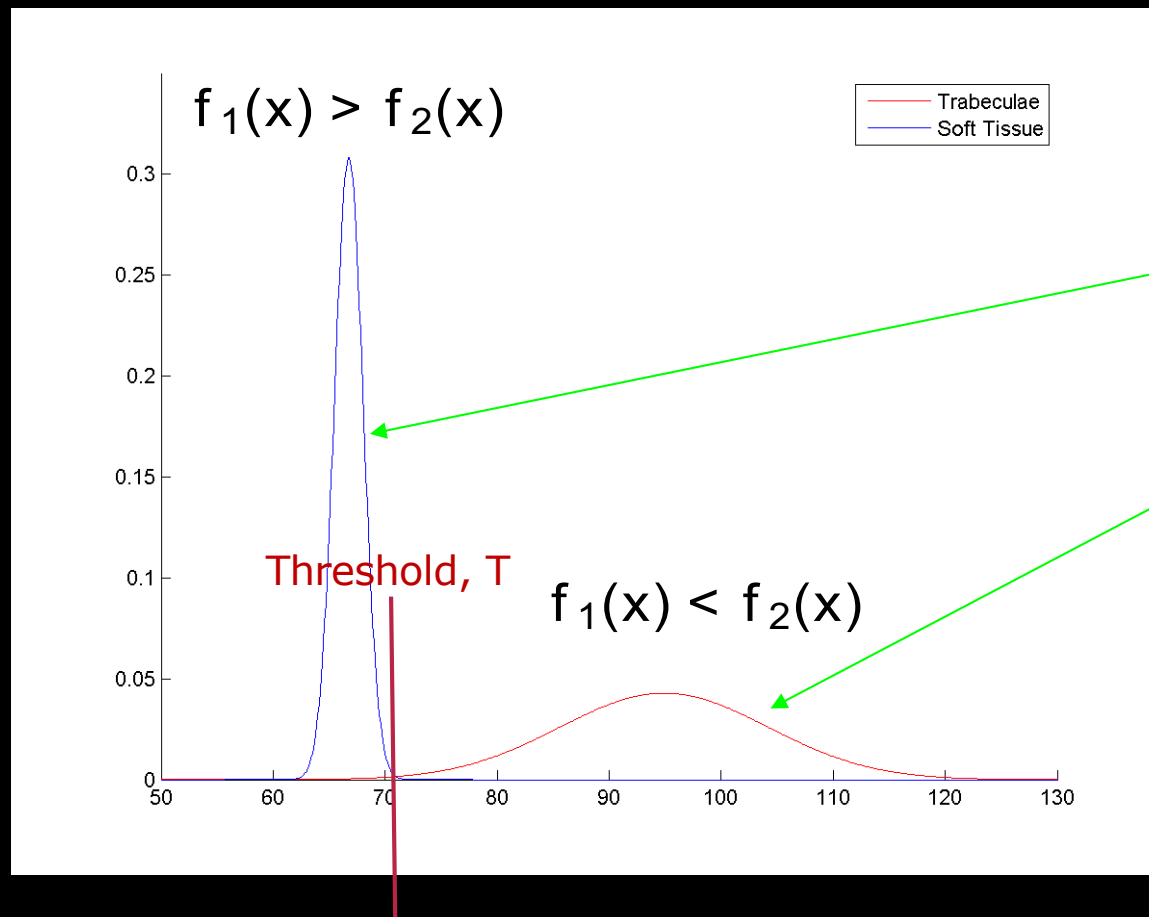
- The pixel value ranges depends on
 - The mean
 - The standard deviation
- Compared to the minimum distance classifier
 - Only the average



Parametric classification – how to

- Select training pixels for each class
- Fit Gaussians ($\mathcal{N}(\mu_i, \sigma_i)$) to each class
- Use Gaussians to determine pixel value ranges

Parametric classifier - ranges



- We want to compute where they cross

$$f_1(x) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left(-\frac{(x - \mu_1)^2}{2\sigma_1^2} \right)$$

$$f_2(x) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp \left(-\frac{(x - \mu_2)^2}{2\sigma_2^2} \right)$$

Create a lookup table:

- Run through all 256 possible pixel values
- Check which Gaussian is the highest
- Store the [value, class] in the table

Alternatively – analytic solution

The two Gaussians

$$\frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left(-\frac{(v - \mu_1)^2}{2\sigma_1^2} \right) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp \left(-\frac{(v - \mu_2)^2}{2\sigma_2^2} \right)$$

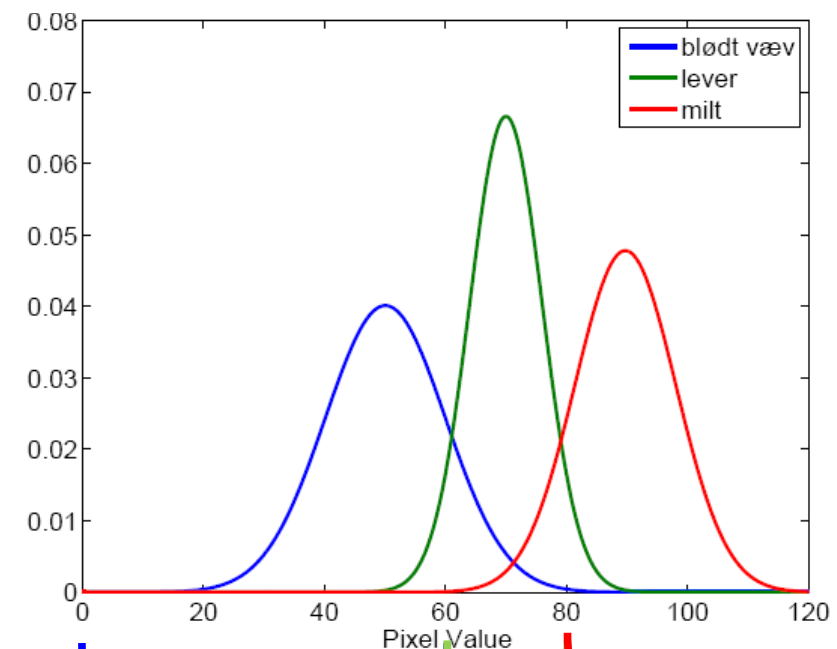
Intercept at

$$v = \frac{\sigma_1^2 \mu_2 - \sigma_2^2 \mu_1 \pm \sqrt{-\sigma_1^2 \sigma_2^2 \left(2 \mu_2 \mu_1 - \mu_2^2 - 2 \sigma_2^2 \ln \left(\frac{\sigma_2}{\sigma_1} \right) - \mu_1^2 + 2 \sigma_1^2 \ln \left(\frac{\sigma_2}{\sigma_1} \right) \right)}}{-\sigma_2^2 + \sigma_1^2}$$

Quiz 5: Class ranges

- A) [0,45],]45, 75],]75,255]
- B) [40,60],]60,100],]100,140]
- C) [0, 60],]60,80],]80,255]**
- D) [0,60],]60,100],]100,255]
- E) [0,75],[75,100],]100,255]

An expert have chosen representative regions in an image that contains soft tissue, liver and spleen. The image pixel minimum and maximum values are 0 and 255. To make a parametric classification, the histograms are parameterized using Gaussian distributions as seen in the image. What are the class ranges?



Solution:

Thomas Bayes

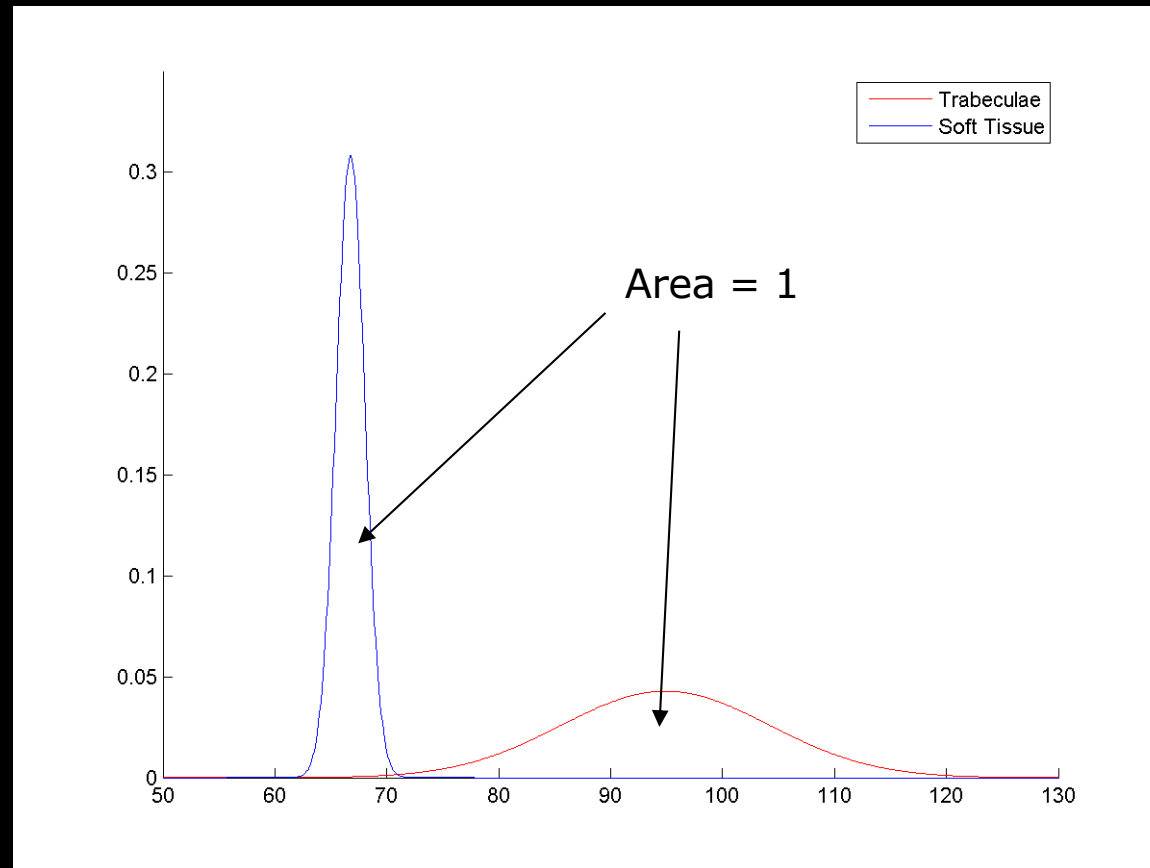


Wikipedia

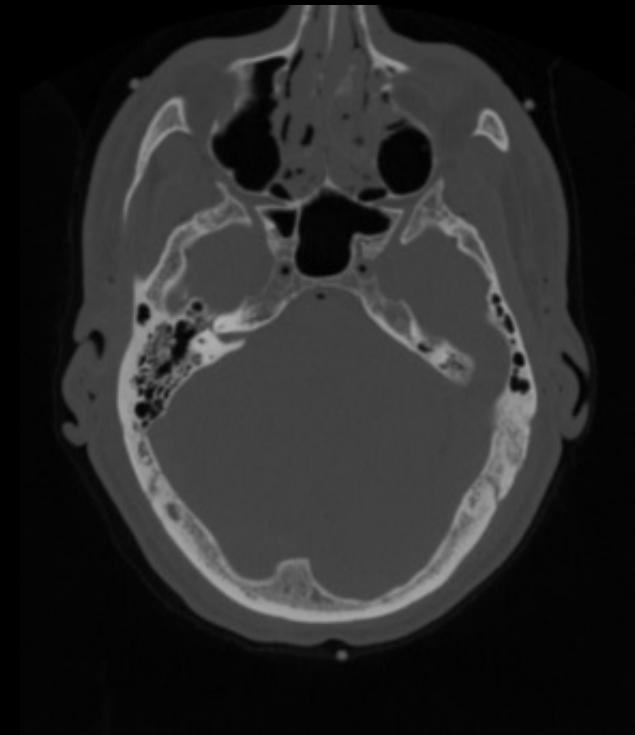
- 1702-1761
- English mathematician and Presbyterian minister
- Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

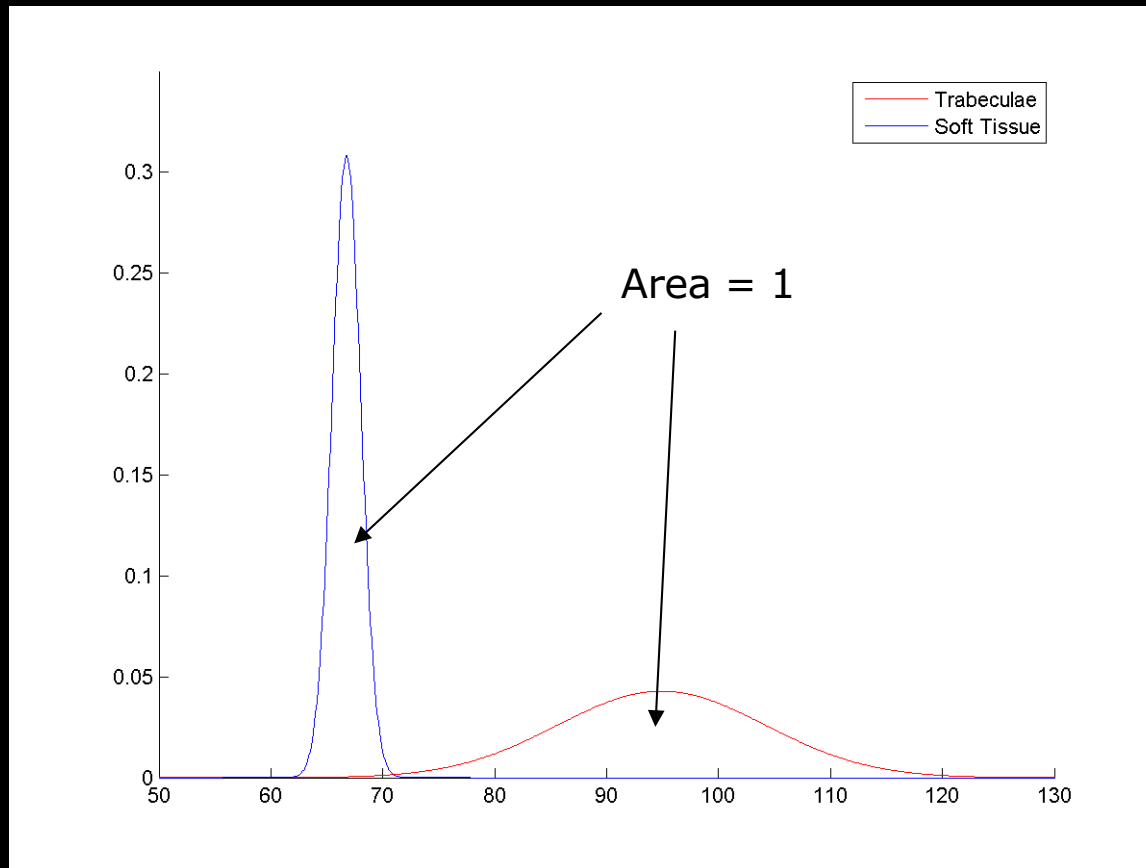
Bayesian Classification



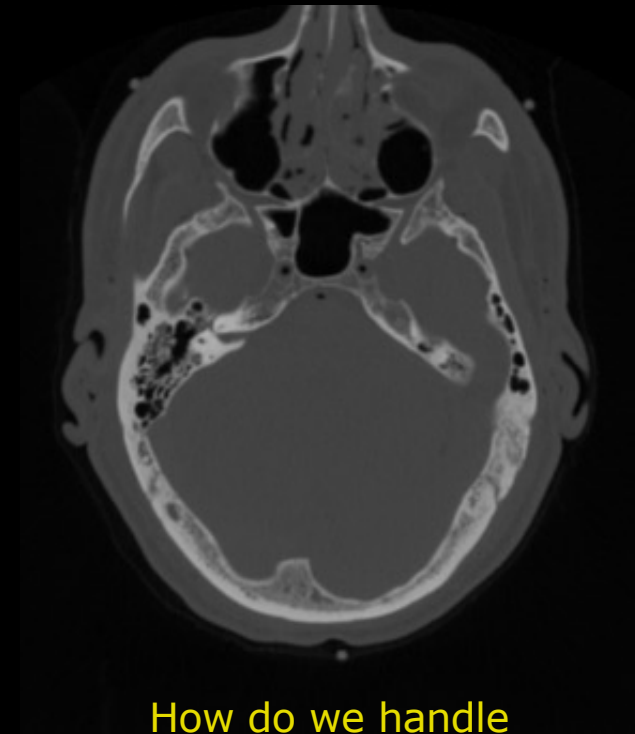
Pure parametric classifier
assumes equal amount of
different tissue types



Bayesian Classification



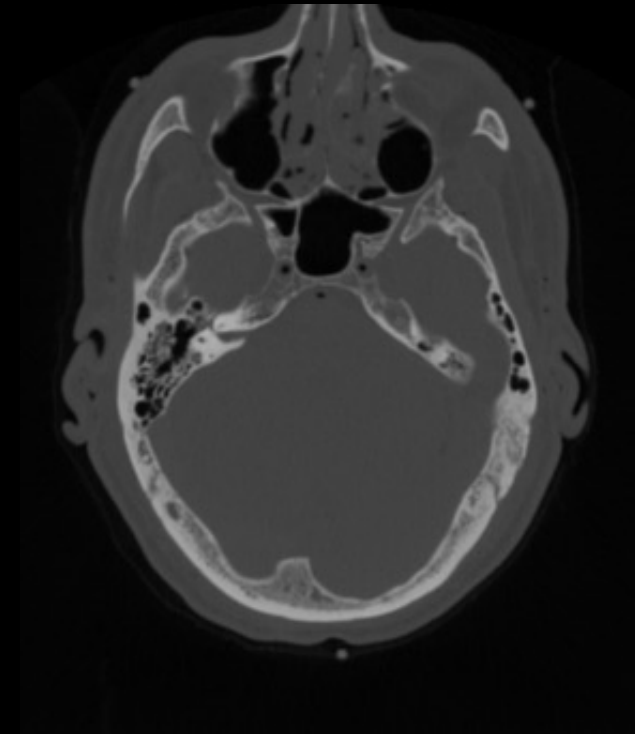
But much more soft-tissue than trabecular bone



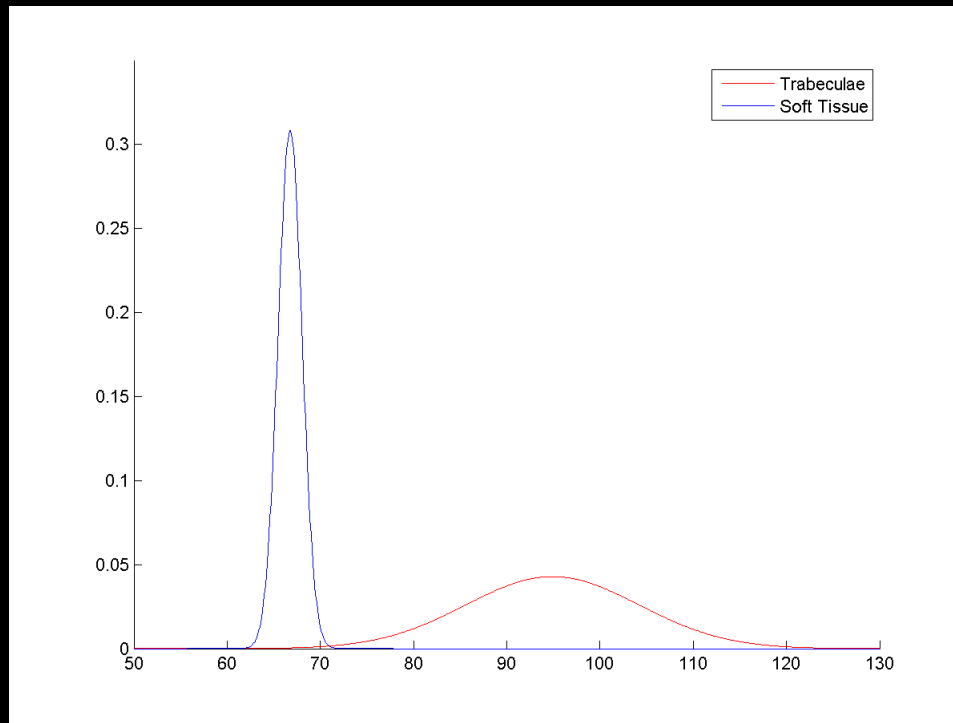
How do we handle that?

Bayesian Classification

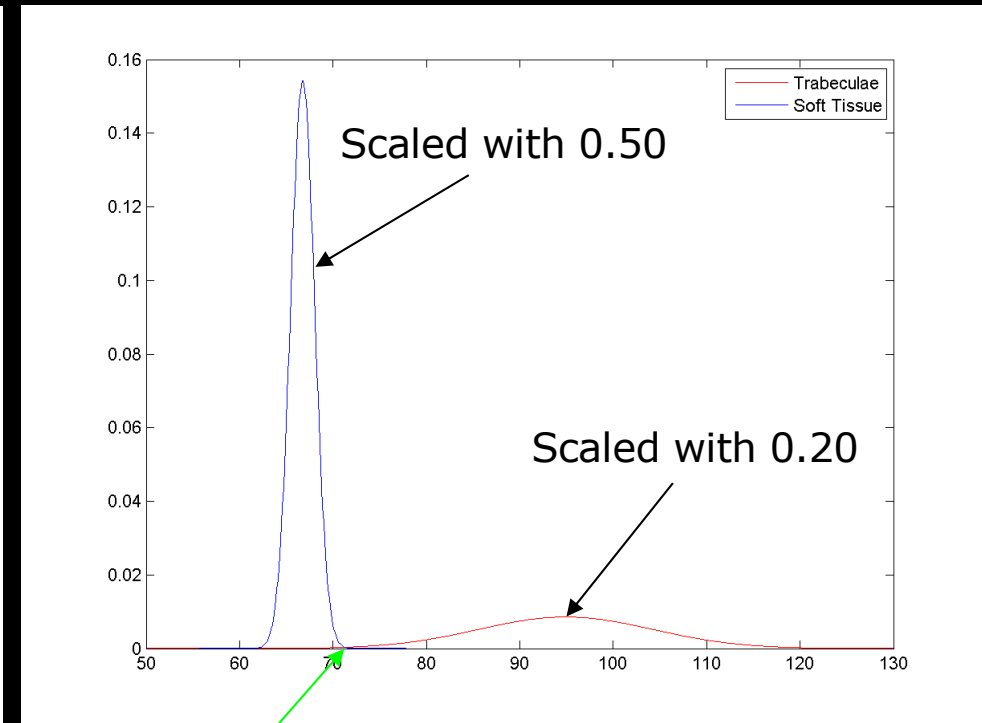
- An expert tells us that a CT scan of a head contains
 - 20% Trabecular bone
 - 50% Soft-tissue
- Picking a random pixel in the image
 - 20% Chance that it is trabecular bone
 - 50% Chance that it is soft-tissue
- How to use that?



Bayesian Classification – histogram scaling



Parametric classifier



Bayesian classifier

Little change in class border
(sometimes significant changes)



Formal definition

- The *posterior probability*
- Given a pixel value v
 - What is the probability that the pixel belongs to class C_i

Example: If the pixel value is 78, what is the probability that the pixel is bone

$$P(c_i|v) = \frac{P(v|c_i)P(c_i)}{P(v)}$$



Formal definition

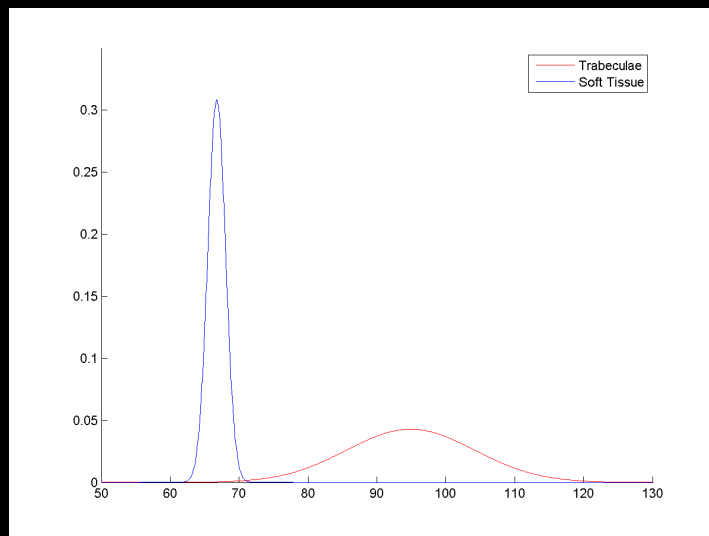
- The *a priori probability* (what is known from before)

Example: From general biology it is known that 20% of a brain CT scan is trabecular bone. Therefore $P(\text{trabecular}) = 0.20$

$$P(c_i|v) = \frac{P(v|c_i)P(c_i)}{P(v)}$$

Formal definition

- The *class conditional probability* also called the *likelihood*
- Given a class, what is the probability of a pixel with value v ?



Example: If we consider class = soft-tissue.
What is the probability that the pixel value is 78?

Very low

$$P(c_i|v) = \frac{P(v|c_i)P(c_i)}{P(v)}$$



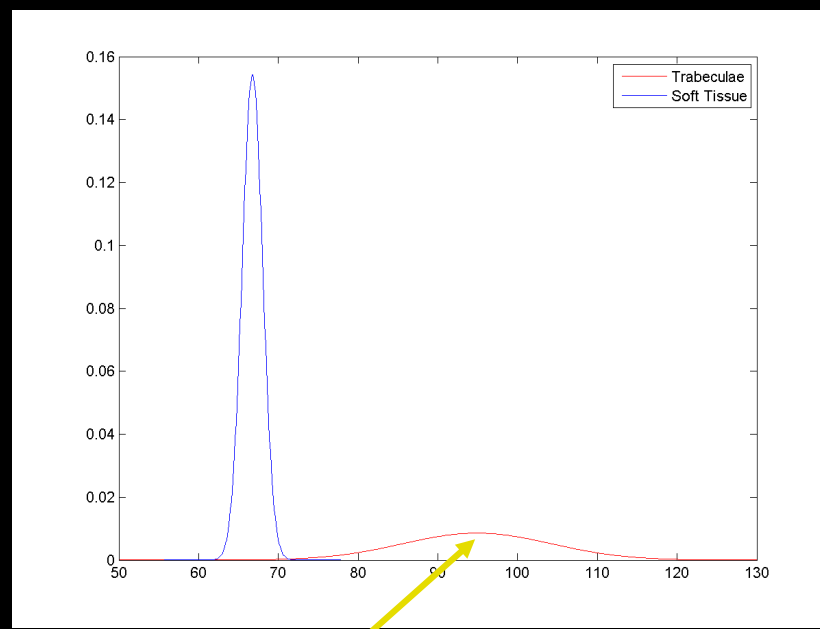
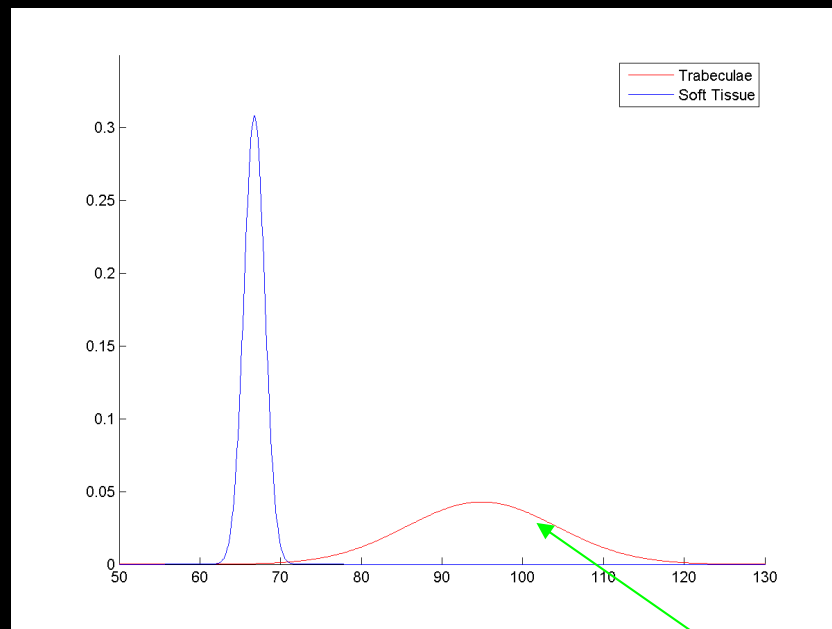
Formal definition

- The *model evidence* or *marginal probability*
- It is basically a normalisation factor: $P(v) = \sum_i P(v|c_i)P(c_i)$

Constant – ignored from now on

$$P(c_i|v) = \frac{P(v|c_i)P(c_i)}{P(v)}$$

Formal definition – sum up



$$P(c_i|v) = \frac{P(v|c_i)P(c_i)}{P(v)}$$

; C_i = trabeculae



Bayesian classification – how to

- Select training pixels for each class
- Fit Gaussians to each class
- Ask an expert for the prior probabilities (how much there normally is in total of each type)
- For each pixel in the image
 - Compute $P(c_i|v)$ for each class (the *a posterior probability*)
 - Select the class with the highest $P(c_i|v)$

$$P(c_i|v) = \frac{P(v|c_i)P(c_i)}{P(v)}$$



When to use Bayesian classification

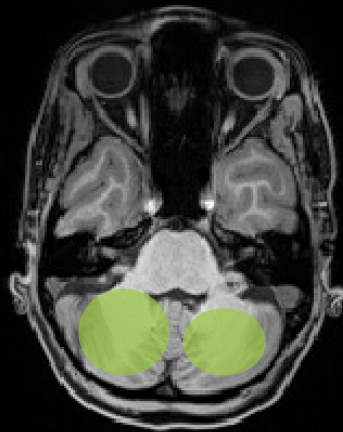
- The parametric classifier is good when there are approximately the same amount of all type of tissues
- Use Bayesian classification if there are very little or very much of some types
- A more general formulation for segmentation
 - especially when going to a higher dimensional feature space

High dimensional feature space

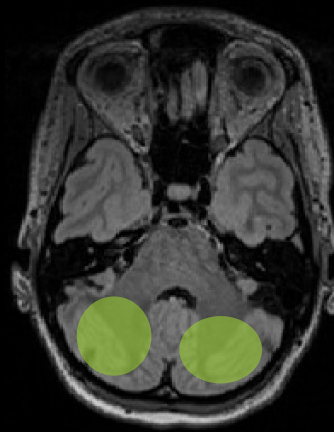
CT



MRI – T1w

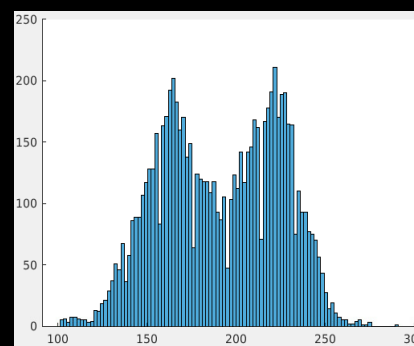
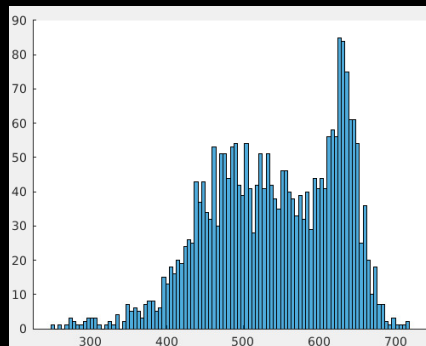
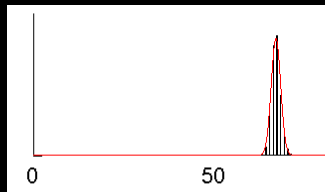


MRI – T2w



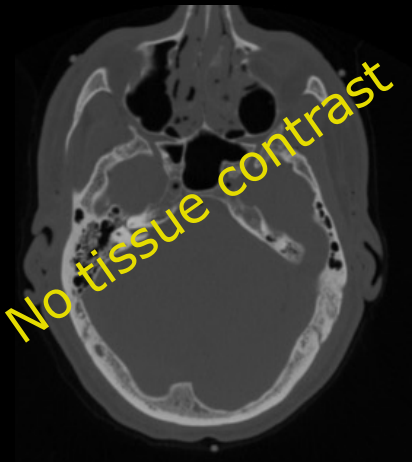
■ Combine different feature inputs to **improve** segmentation

- Different image modalities e.g. CT vs MRI
- Subject groups
 - Healthy vs disease
- Different angles of object e.g. cars



High dimensional feature space

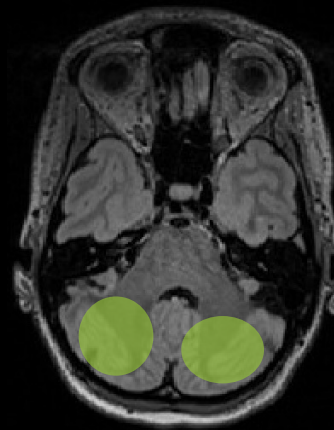
CT



MRI – T1w



MRI – T2w

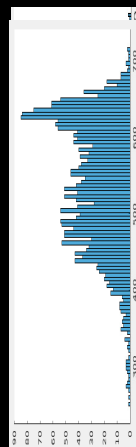


■ Feature space:

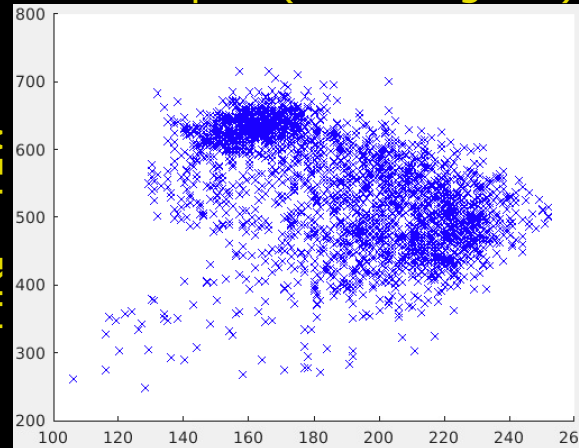
- 1D is a histogram
- 2D is a scatterplot i.e. 2D histogram
- >2D is bit more complicated to show

scatter plot (2D histogram)

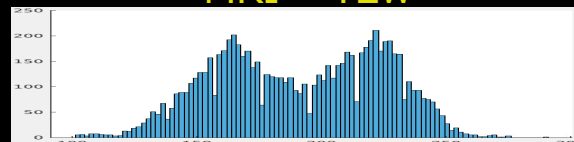
1D histogram



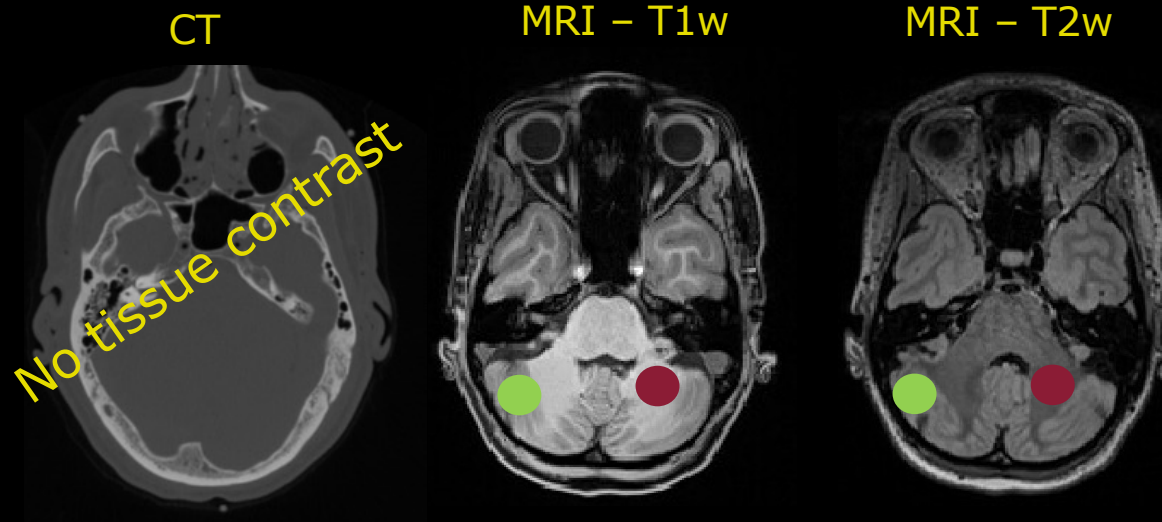
MRI – T1w



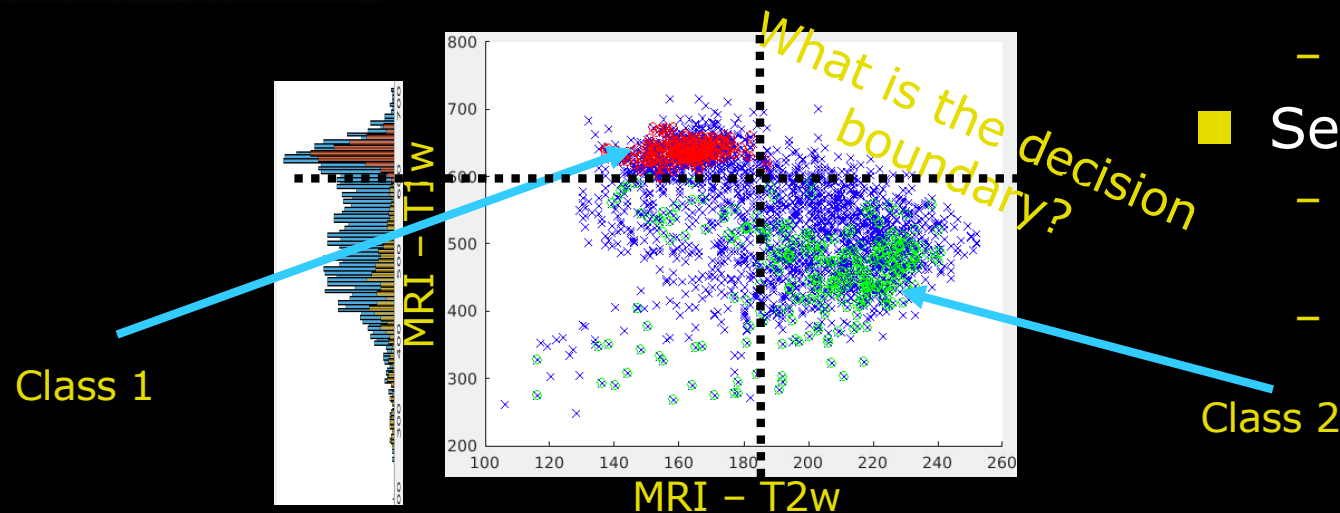
MRI – T2w



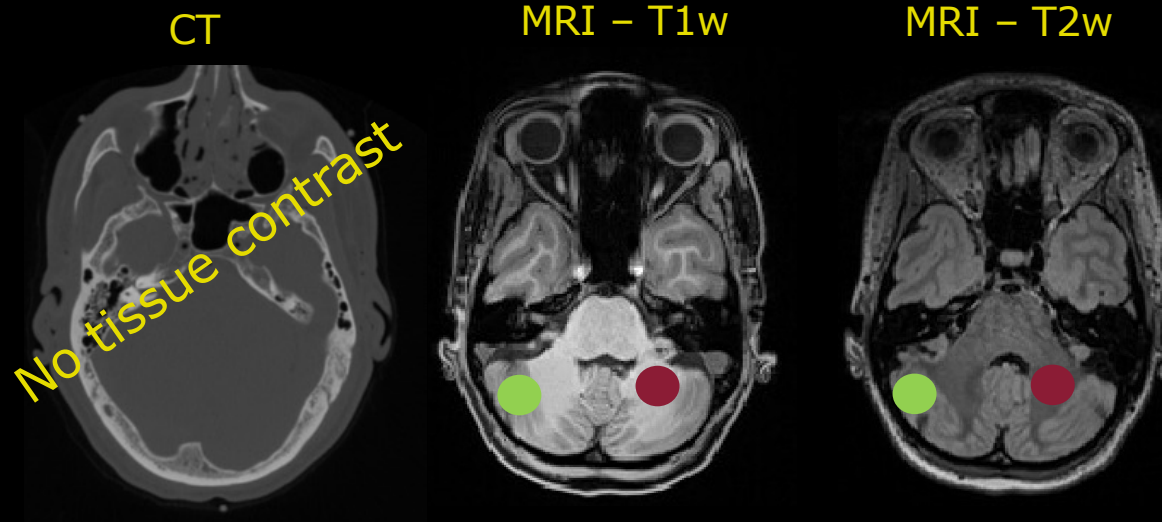
High dimensional feature space



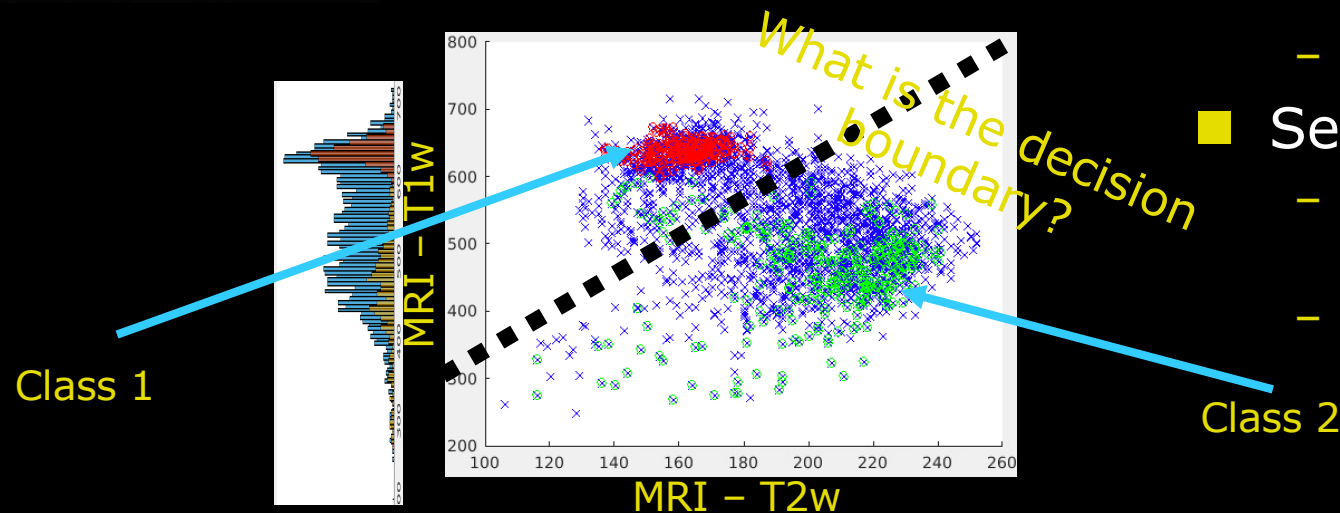
- Segmentation with more feature inputs
- To train our **classifier** model with class examples
 - Draw tissue specific regions for each class
 - **Class 1** and **Class 2**
 - Tissue **type 1** and **type 2**
- Segmentation:
 - Define the threshold for the decision boundary?
 - 1D vs 2D



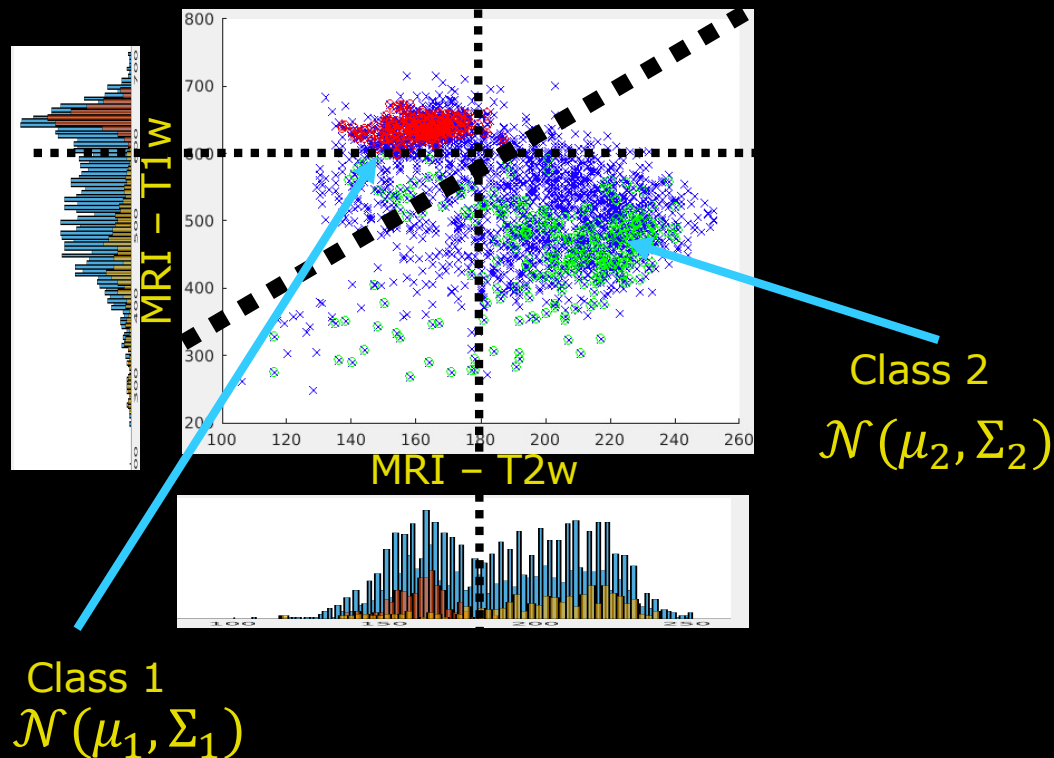
High dimensional feature space



- Segmentation with more feature inputs
- To train our **classifier** model with class examples
 - Draw tissue specific regions for each class
 - **Class 1** and **Class 2**
 - Tissue **type 1** and **type 2**
- Segmentation:
 - Define the threshold for the decision boundary?
 - 1D vs 2D



Decision boundary: Define a model



- 2D feature space
 - Better class separation vs 1D?
- Model assumption
 - Type of distribution?
- Intensity histograms looks Gaussian-like, or?
 - We assume Gaussian distributions: $\mathcal{N}(\mu_i, \Sigma_i)$
- Use Bayes theorem
 - Probability of belonging to C2:

$$\frac{P(C2|\mathbf{x})}{P(C1|\mathbf{x})} > T$$

- Decision boundary
 - A hyperplane for $T=1$:
 - $P(C2|\mathbf{x}) = P(C1|\mathbf{x})$

Decision boundary: Train a model

- We wish to use Bayes:

$$\frac{P(C2|\mathbf{x})}{P(C1|\mathbf{x})} > T$$

- The posterior probability

$$- P(C_i|\mathbf{x}) = P(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)P_{C_i}$$

- The likelihood: A Gaussian model

$$P(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = K_i \exp((\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i))$$

- Data points:

$$\blacksquare \mathbf{x}_i = [x_1, x_2]^T$$

- Training set:

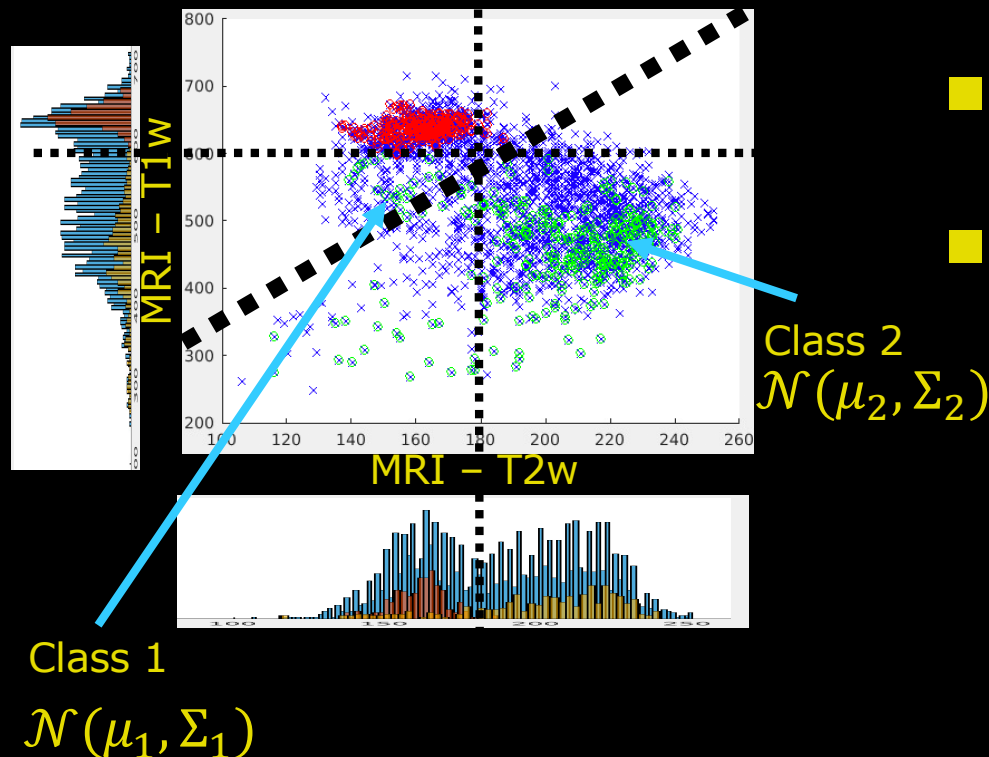
$$\blacksquare \mathbf{t}_{\mathbf{x} \in C1} = 0 \text{ and } \mathbf{t}_{\mathbf{x} \in C2} = 1$$

- The class mean- parameter

$$\blacksquare \boldsymbol{\mu}_i = \frac{1}{N} \sum_{n \in C_i} \mathbf{x}_n$$

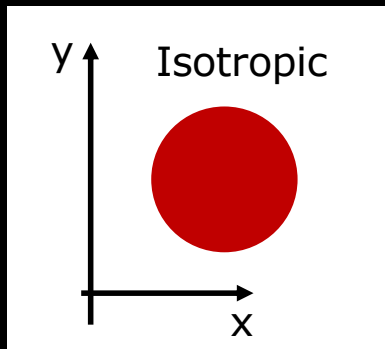
- The covariance matrix-parameter

$$\blacksquare \boldsymbol{\Sigma}_i = (\mathbf{x} - \boldsymbol{\mu}_i)^T (\mathbf{x} - \boldsymbol{\mu}_i)$$



- What about the prior probability $P(C_i)$?

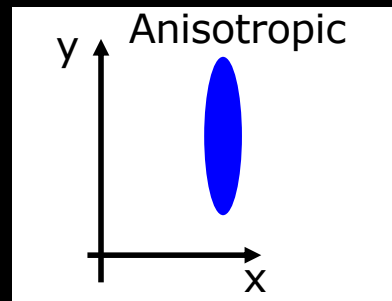
Gaussian in 2D: The covariance matrix



Rotational invariant

$$\Sigma = \begin{bmatrix} \sigma_{xx}^2 & 0 \\ 0 & \sigma_{yy}^2 \end{bmatrix}$$

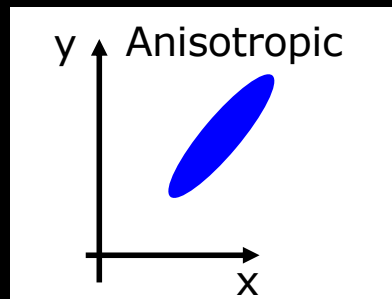
$\sigma_{xx} = \sigma_{yy}$



Aligned with coordinate system

$$\Sigma = \begin{bmatrix} \sigma_{xx}^2 & 0 \\ 0 & \sigma_{yy}^2 \end{bmatrix}$$

$\sigma_{xx}^2 \neq \sigma_{yy}^2$



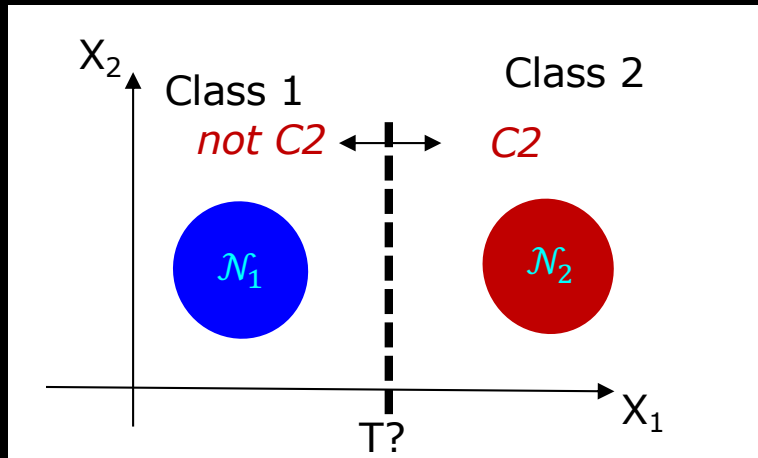
Not aligned with coordinate system

$$\Sigma = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{bmatrix}$$

QUICK REFRESH:

- The covariance matrix:
 $\Sigma_i = (\mathbf{x} - \boldsymbol{\mu}_i)^T (\mathbf{x} - \boldsymbol{\mu}_i)$
- Expresses the orientation of anisotropic variance in relation to coordinate system

The linear discriminant classifier



- Classifier: If \mathbf{x} belongs to C_2 :

$$\frac{P(C2|\mathbf{x})}{P(C1|\mathbf{x})} > T$$

- Take the logarithm

$$\ln(P(C2|\mathbf{x})) - \ln(P(C1|\mathbf{x})) > \ln(T)$$

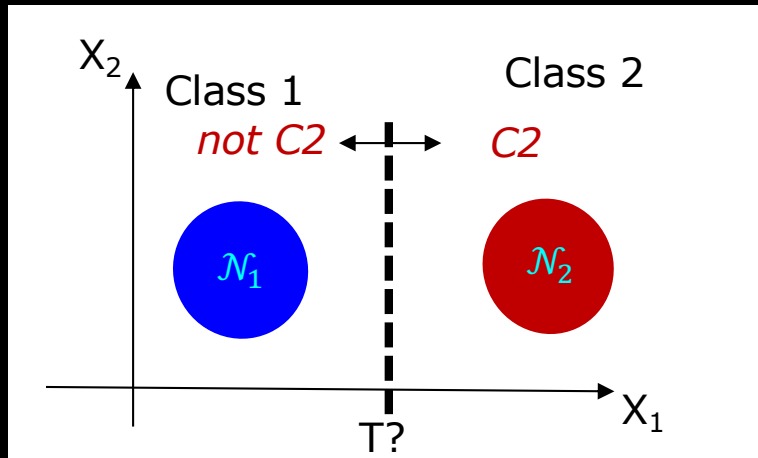
$$\mathcal{N}_1(\mu_1, \Sigma_1) \quad \mathcal{N}_2(\mu_2, \Sigma_2)$$

Inspiration derive:

https://en.wikipedia.org/wiki/Linear_discriminant_analysis

<https://people.revoledu.com/kardi/tutorial/LDA/LDA%20Formula.htm>

The linear discriminant classifier



- Classifier: If \mathbf{x} belongs to C_2 :

$$\frac{P(C2|\mathbf{x})}{P(C1|\mathbf{x})} > T$$

- Take the logarithm

$$\ln(P(C2|\mathbf{x})) - \ln(P(C1|\mathbf{x})) > \ln(T)$$

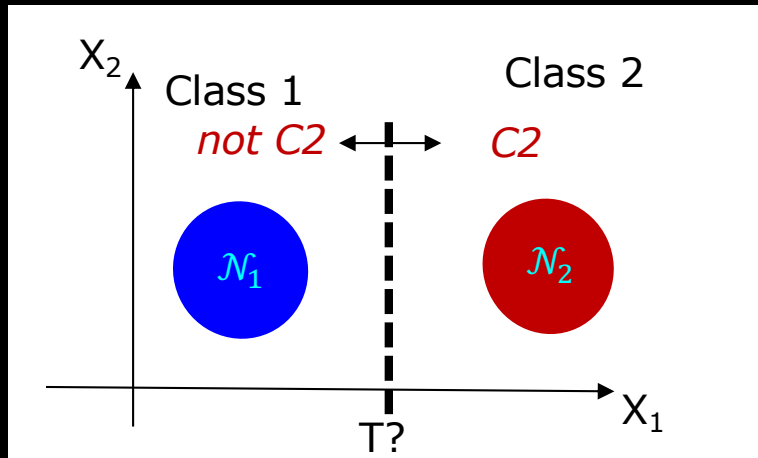
- Where the log-posterior probability for C_i :

$$\ln(P(C_i|\mathbf{x})) = \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) + \ln(K_i) + \ln(P_i)$$

- P_i is the prior probability for class C_i

$$\mathcal{N}_1(\mu_1, \Sigma_1) \quad \mathcal{N}_2(\mu_2, \Sigma_2)$$

The linear discriminant classifier



- Classifier: If \mathbf{x} belongs to C_2 :

$$\frac{P(C2|\mathbf{x})}{P(C1|\mathbf{x})} > T$$

- Take the logarithm

$$\ln(P(C2|\mathbf{x})) - \ln(P(C1|\mathbf{x})) > \ln(T)$$

- Where the log-posterior probability for C_i :

$$\ln(P(C_i|\mathbf{x})) = \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) + \ln(K_i) + \ln(P_i)$$

- P_i is the prior probability for class C_i

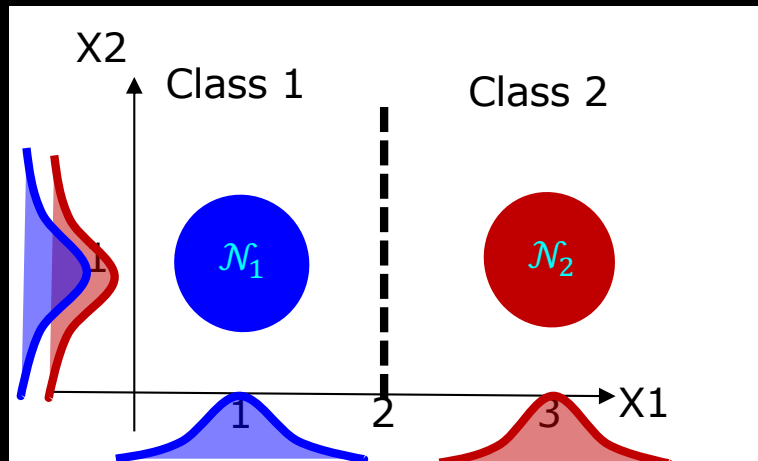
$$\mathcal{N}_1(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \quad \mathcal{N}_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$$

- Assuming homoscedasticity ($\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_0$) and isotropic covariance matrix we have **the Linear Discriminant Analysis (LDA) classifier model**:

$$\ln \frac{P_2}{P_1} + \frac{1}{2}(\boldsymbol{\mu}_2 + \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) - \mathbf{x}^T \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) > \ln(T)$$

- We train the classifier with examples obtained from the two distributions N_1 and N_2

Quiz 6 - The LDA classifier



■ Linear Discriminant Analysis (LDA):

$$\ln \frac{P_2}{P_1} + \frac{1}{2} (\mu_2 + \mu_1)^T \Sigma_0^{-1} (\mu_2 - \mu_1) - x^T \Sigma_0^{-1} (\mu_2 - \mu_1) > \ln(T)$$

Where:

$$\Sigma_1 = \Sigma_2 = \Sigma_0 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Prior probabilities: $P_1 = P_2 = 0,5$

Which data points are placed on the hyperplane for $P(C_2 | x) = P(C_1 | x)$?

A) $[0,5]^T$

B) $[1,7]^T$

C) $[3,3]^T$

D) $[2,0]^T$

E) $[0,7]^T$

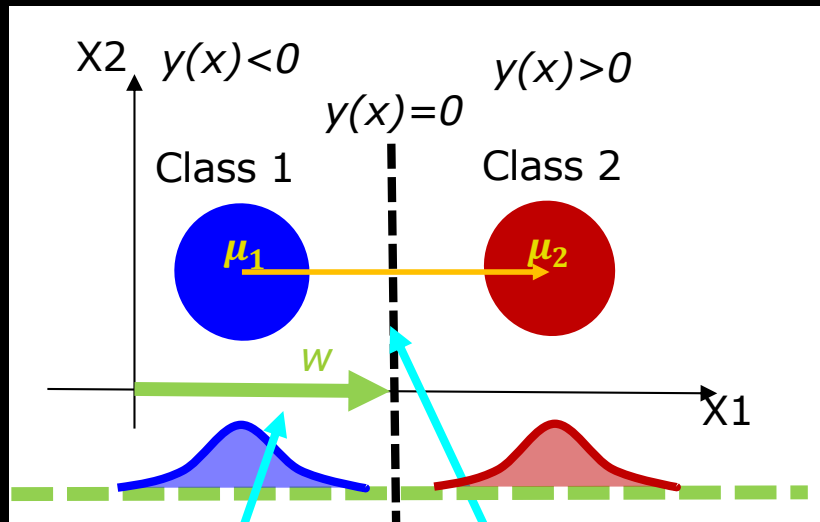
Solution – We see that when $T=1 \Rightarrow \ln(1)=0$ is the decision boundary which is placed only along X_1 i.e. a solution in 1D:

$$\ln \frac{P_2}{P_1} + \frac{1}{2} (\mu_2 + \mu_1) \frac{(\mu_2 - \mu_1)}{\sigma_0} = x_1 \frac{(\mu_2 - \mu_1)}{\sigma_0}$$

$$\ln \frac{0,5}{0,5} + \frac{1}{2} (3 + 1) \frac{(3-1)}{2} = x_1 \frac{(3-1)}{2}$$

$x_1 = 2$ & $x_2 = \text{all values}$

Projections in the feature space



decision boundary

- w projects the class mean direction i.e. the weight vector
- w is normal to the hyperplane of the decision boundary for $y_i(x)=0$
- $x^T w$ is a dot product i.e. x and c are projected onto w ($a^T b = \|a\| \|b\| \cos(\theta)$)

- General formulation of a classifier
 - A projection of data points in relation to the decision boundary

- The LDA function for C_2 :

$$\underbrace{\ln \frac{P_1}{P_2}}_C + \underbrace{\frac{1}{2} (\mu_2 + \mu_1)^T \Sigma_0^{-1} (\mu_2 - \mu_1)}_W - \underbrace{x^T \Sigma_0^{-1} (\mu_2 - \mu_1)}_W > \ln T$$

W_0

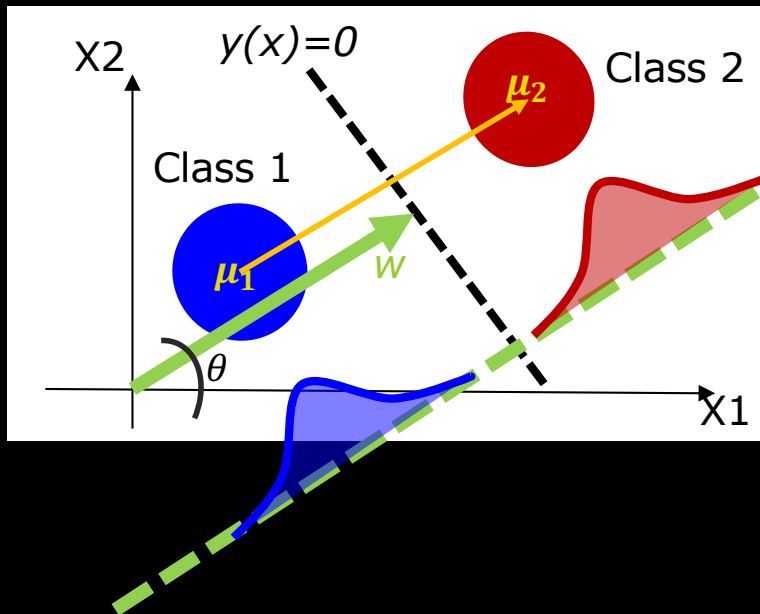
- The linear discriminant function

$$y_{C \in 2}(x) = x^T w + w_0$$

-where negative w_0 is the threshold

- x is assigned to C_2 if $y_{C \in 2}(x) > 0$

Projections in the feature space



- General formulation of a classifier
 - A projection of data points in relation to the decision boundary

- The LDA function for C_2 :

$$\underbrace{\ln \frac{p_1}{p_2}}_C + \underbrace{\frac{1}{2}(\mu_2 + \mu_1)^T \Sigma_0^{-1}(\mu_2 - \mu_1)}_W - \underbrace{x^T \Sigma_0^{-1}(\mu_2 - \mu_1)}_W > \ln T$$

W_0

- w projects the class mean direction i.e. the weight vector
- w is normal to the hyperplane of the decision boundary $y_i(x)=0$
- $x^T w$ is a dot product i.e. x and c are projected onto w ($a^T b = \|a\| \|b\| \cos(\theta)$)

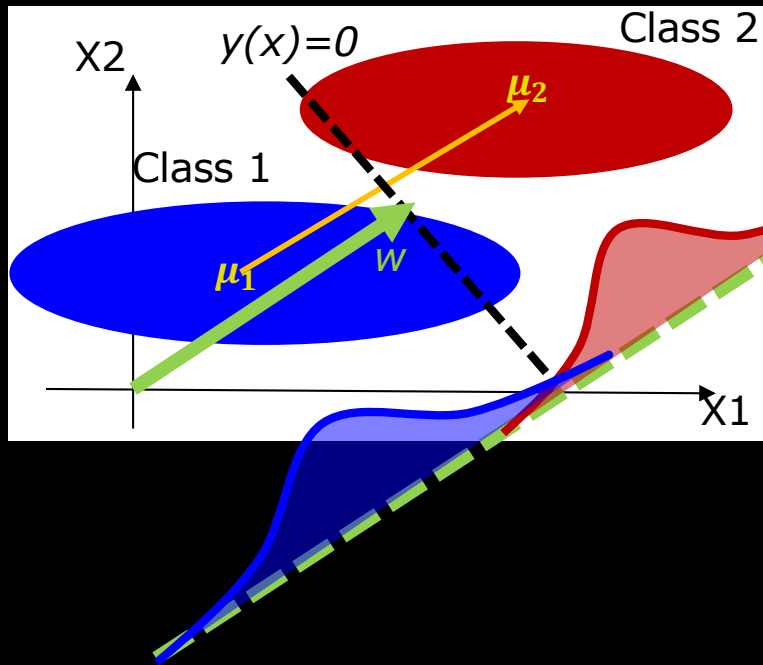
- The linear discriminant function

$$y_{C \in 2}(x) = x^T w + w_0$$

-where negative w_0 is the threshold

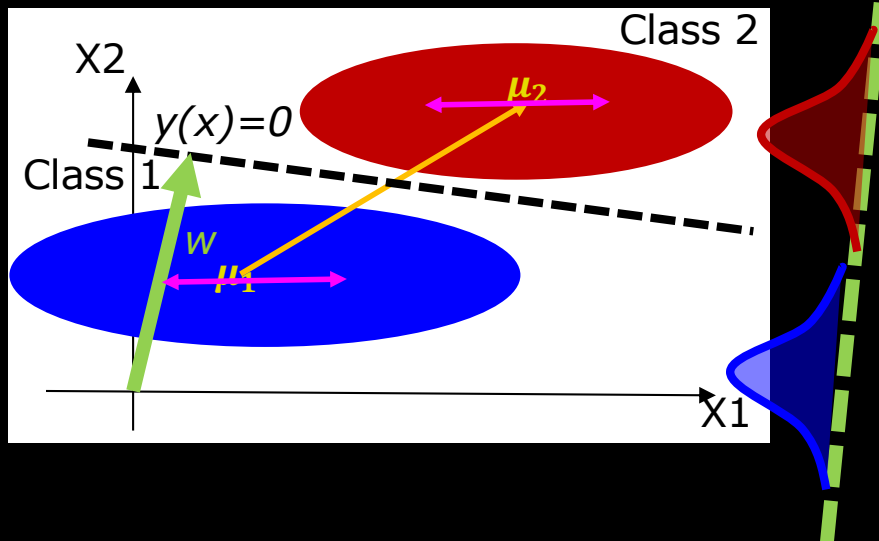
- x is assigned to C_2 if $y_{C \in 2}(x) > 0$

Projections in the feature space



- If the covariance is *anisotropic* and have different class variances
 - The LDA classifier does not ensure an optimal class separation!
 - LDA only separate the class means
- To improve the separation
 - We need to change the model hence the **weight vector, W**

Projections in the feature space



Optimal class separation:

- The *weight vector*, w , now accounts for both class means and variances

■ Fisher's LDA:

- Uses: *between-class (means) covariance*:

$$S_B = (\mu_2 - \mu_1)^T (\mu_2 - \mu_1)$$

- and: optimise *(total) within-class covariance*

$$S_W = \Sigma_1 + \Sigma_2$$

■ Find projection w using a cost function:

$$J(w) = \frac{w^T S_B w}{w^T S_W w}$$

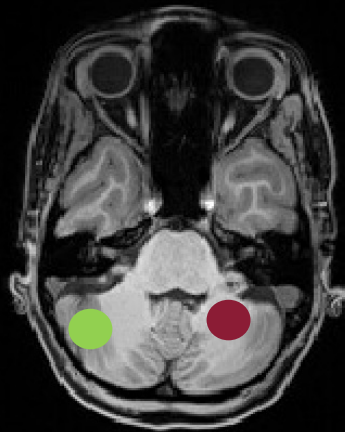
- differentiate: $\frac{\partial J(w)}{\partial w} = 0$

- which gives (simple solution):

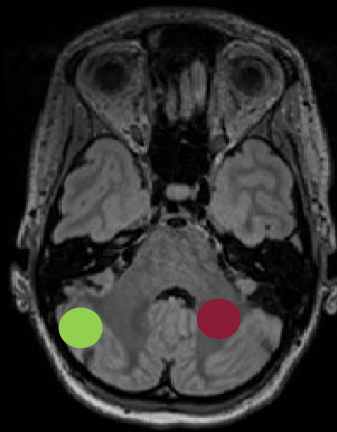
$$w \propto S_W^{-1} (\mu_2 - \mu_1)$$

Segmentation of brain data using LDA

MRI – T1w

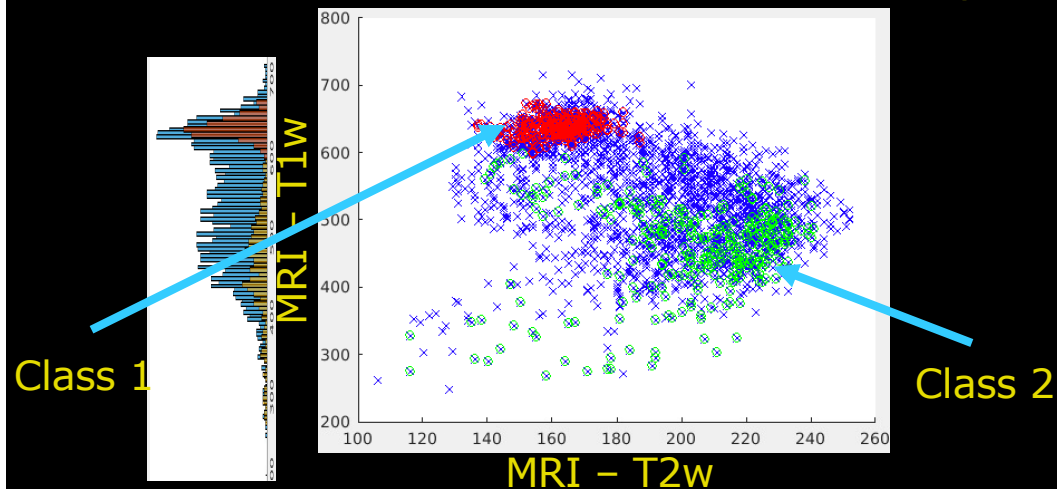


MRI – T2w



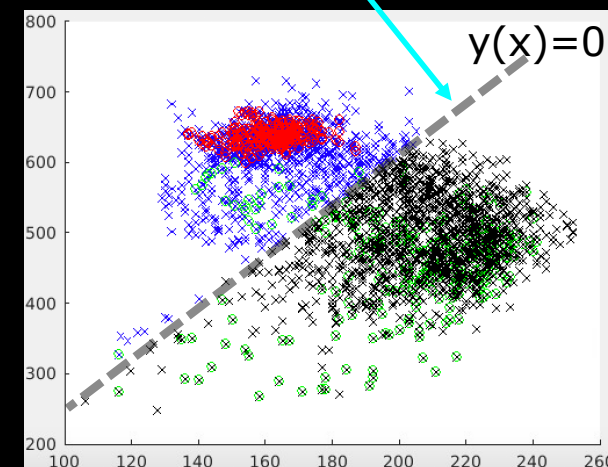
■ Fisher's LDA

Decision boundary?

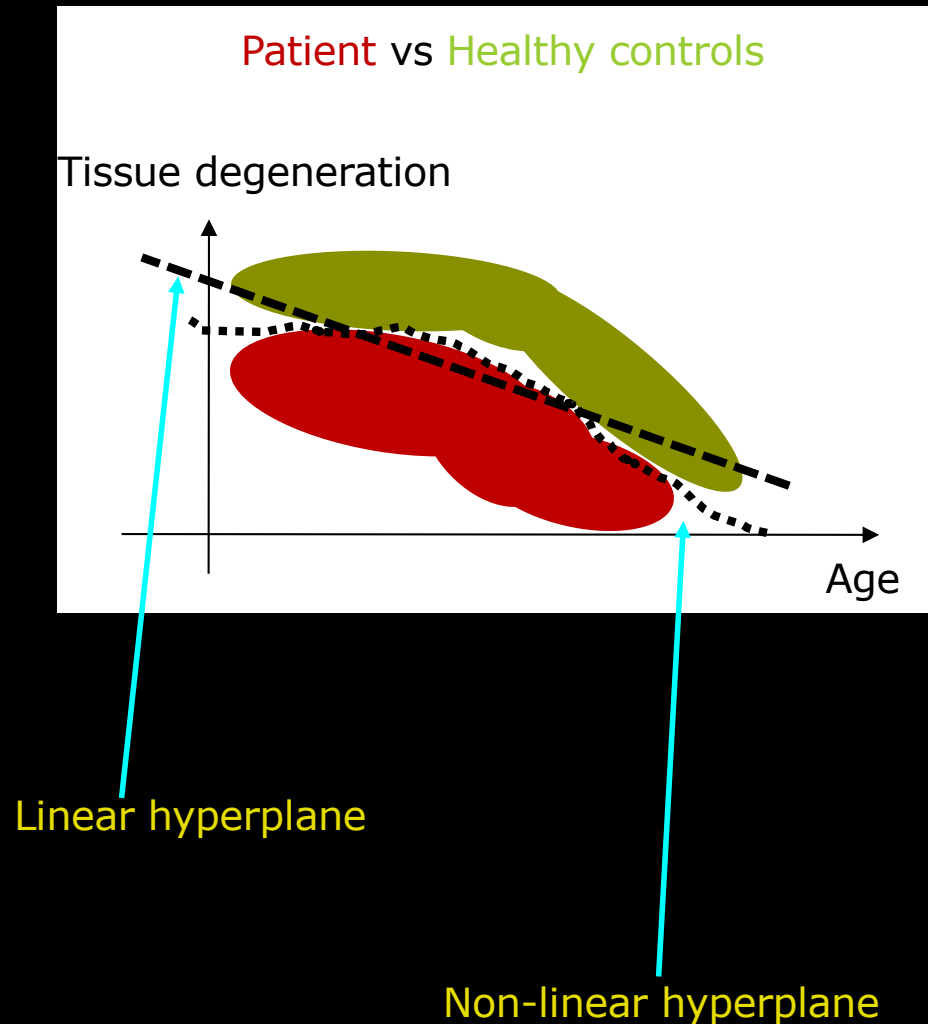


Decision boundary ($T=1$)

Segmentation result: Fisher's LDA

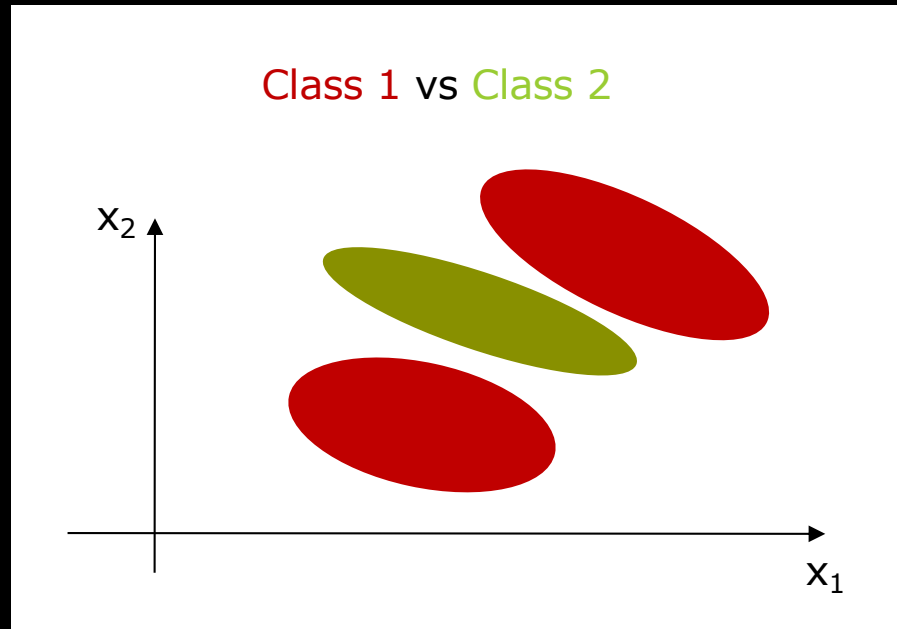


Limitations of LDA



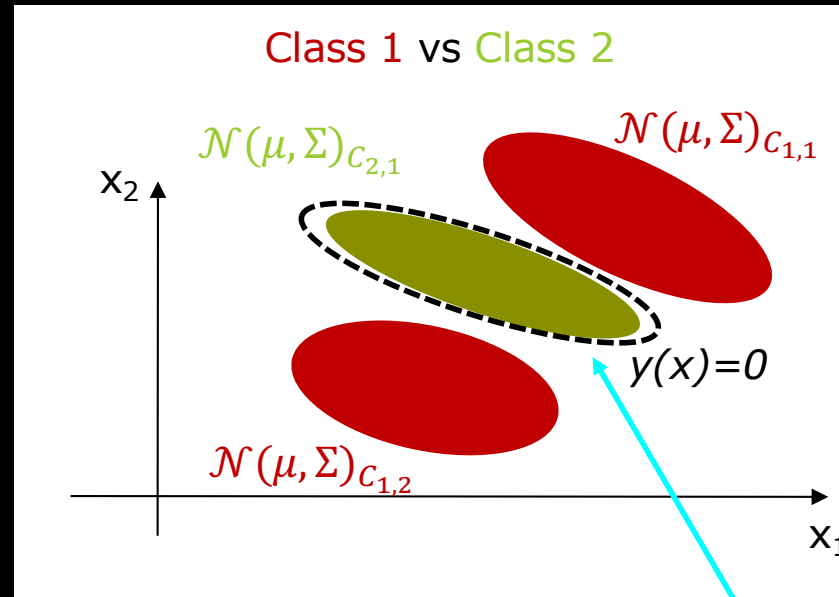
- Linear discriminant analysis (LDA)
 - Only linear hyperplanes
- Non-linear hyperplanes?
- Example:
 - I wish to make a classifier
 - Features (2D):
 - Age vs. Tissue degeneration
- Classes
 - Healthy controls vs Patient

Limitations of LDA



- One class can be separated
 - A non-linear problem

Non-linear Hyperplanes



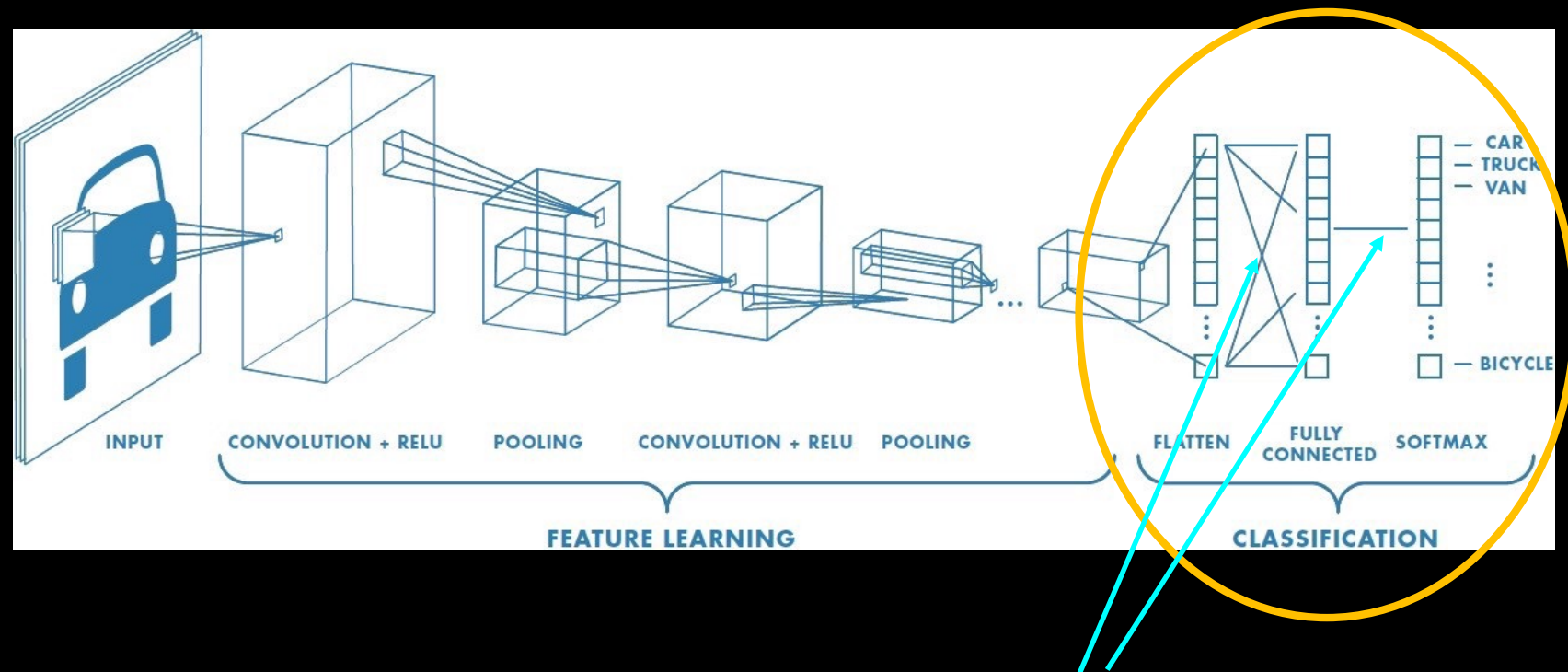
- Class 1: $\mathcal{N}(\mu, \Sigma)_{c_{1,1}} + \mathcal{N}(\mu, \Sigma)_{c_{1,2}}$
 - Class 2: $\mathcal{N}(\mu, \Sigma)_{c_{2,1}}$
- Non-linear hyperplane

Non-linear classifiers
(Machine learning):
Example:

- Gaussian Mixture Model
 - Each class is modelled using a number of Gauss distributions e.g. class 1
- Again use Bayes theorem also for Gaussian Mixture Model
- Optimisation:
 - We derive $\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} = 0$ for a Gaussian mixture model
 - Iterative optimisation algorithm is used to find \mathbf{w}

Segmentation - Non-linear Hyperplanes

- Convolutional neural network and classification



Weights can be non-linear sigmoid functions: $y_k = \phi(x, w, w_0)$



What did you learn today?

- Describe the concept of pixel classification
- Compute the pixel value ranges in a minimum distance classifier
- Implement and use a minimum distance classifier
- Approximate a pixel value histogram using a Gaussian distribution
- Implement and use a parametric classifier
- Decide if a minimum distance or a parametric classifier is appropriate based on the training data
- Explain the concept of Bayesian classification
- Implement and use the linear discriminant analysis (LDA) classifier
- Decide where to place a decision boundary
- Understand the use of linear vs non-line hyperplanes for segmentation



Lecture 9 – Industry presentations

JLIVision
Videometer
TrackMan
Visiopharm
FOSS Analytics
Claas E-systems
Radiobotics



Teaching – the speed of the lecture

- A) Come oooooon! I am so bored
- B) I can easily follow and knit my sweater
- C) The speed is fine
- D) I need to concentrate a lot to follow
- E) Hey! Wait! You are too fast