How to train your Neural Network

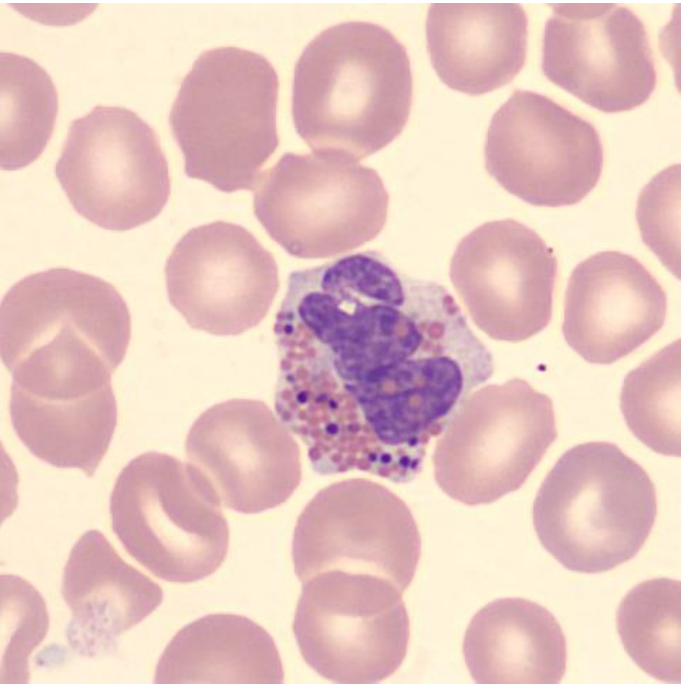
CELLAVISION

Artificial Intelligence in pop culture

’I need your bike, your boots and your DM system!’

CELLAVISION

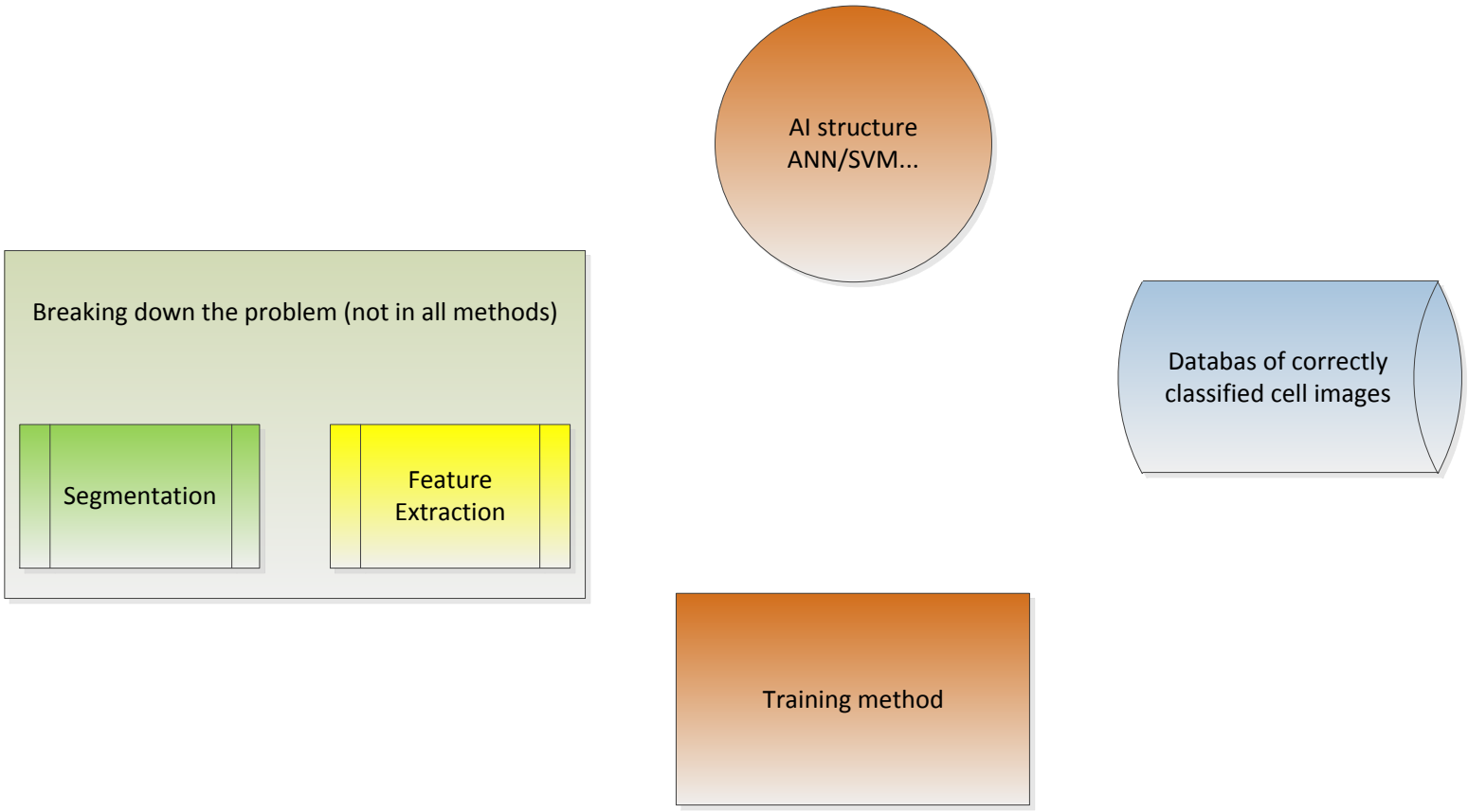
Artificial Intelligence at CellaVision



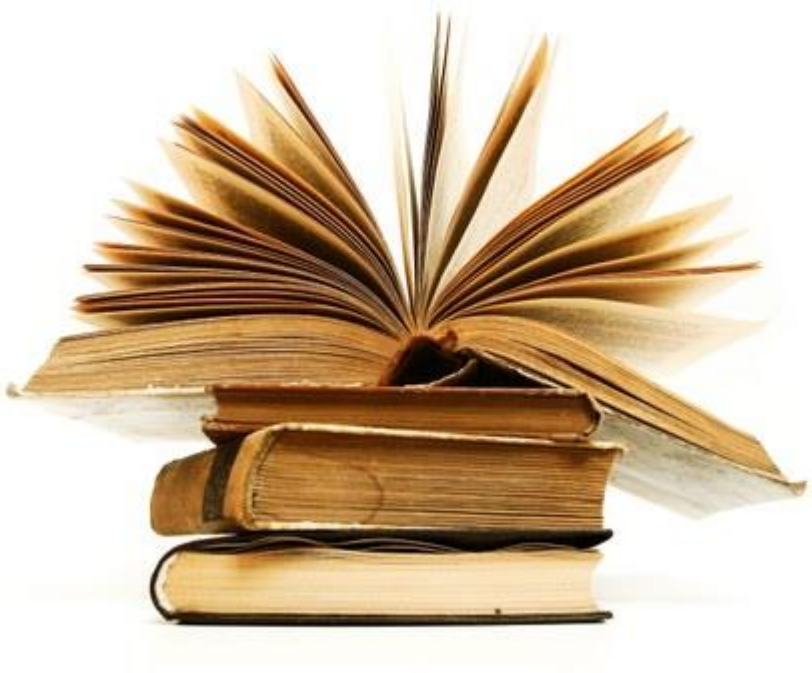
’Suggest class for this WBC.’

CELLAVISION

Tools


AI structure
ANN/SVM...
Breaking down the problem (not in all methods)
Segmentation
Feature
Extraction
Databas of correctly
classified cell images
Training method

How to select training material for your Neural Network



Expert database

- Collect images.
- For each cell in these images assign a correct class.
- This is done by human experts.
- Enough examples of each cell type.
- Cells from different stainings.

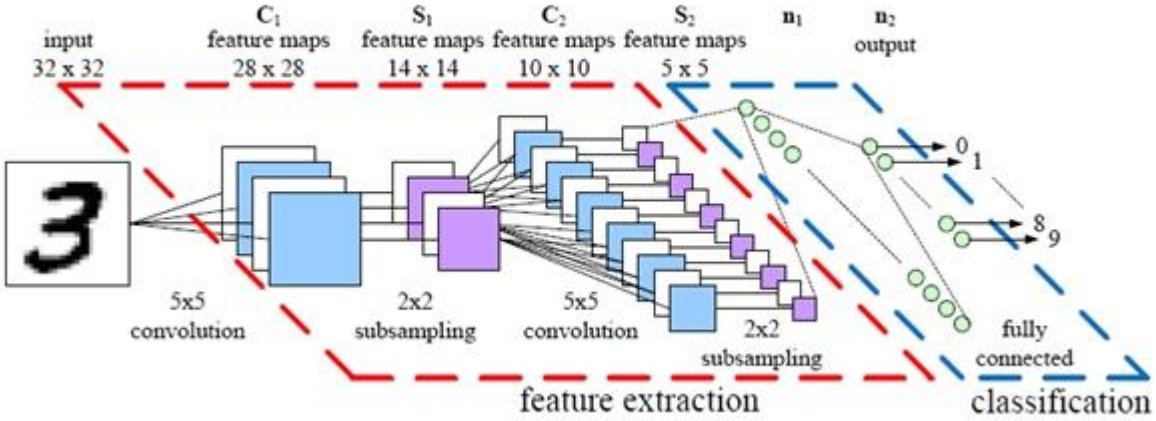
CELLAVISION

Complexity of the problem

- Each image ~250x250 pixels.
  - 187500 integers [0,255] representing the image.

CELLAVISION

Features

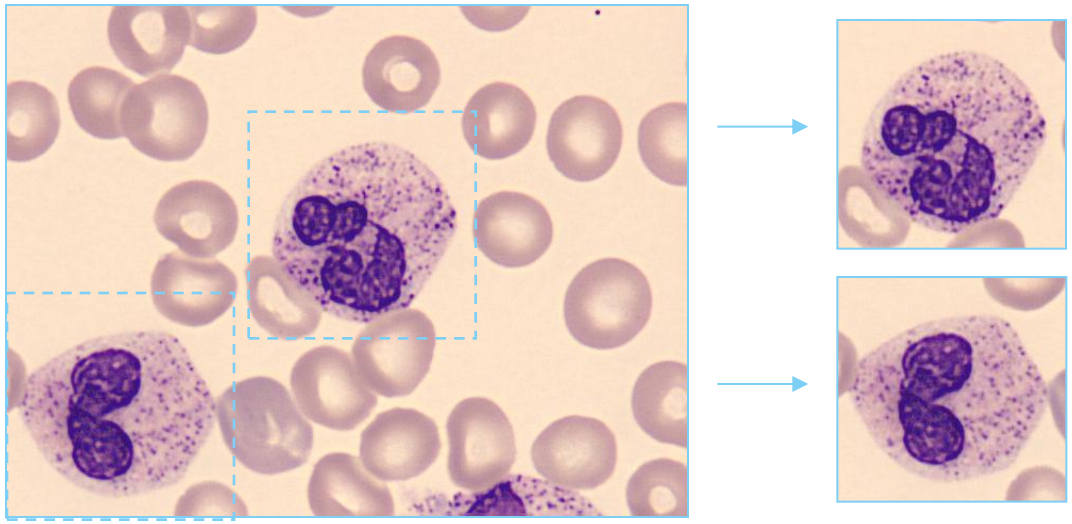


- Reduce amount of input to the Neural Network.
  - Let the network find its own features – current research area.
  - Manually design features - traditional. Used at CellaVision.

CELLAVISION

Segmentation

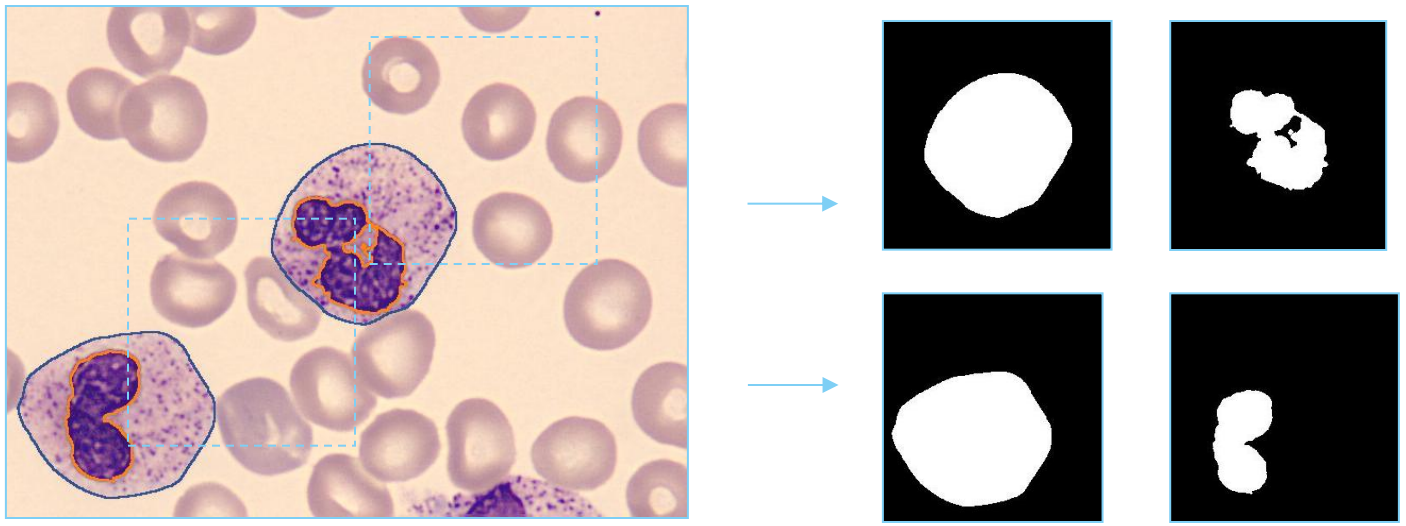
Find the cells on the image



Segmentation

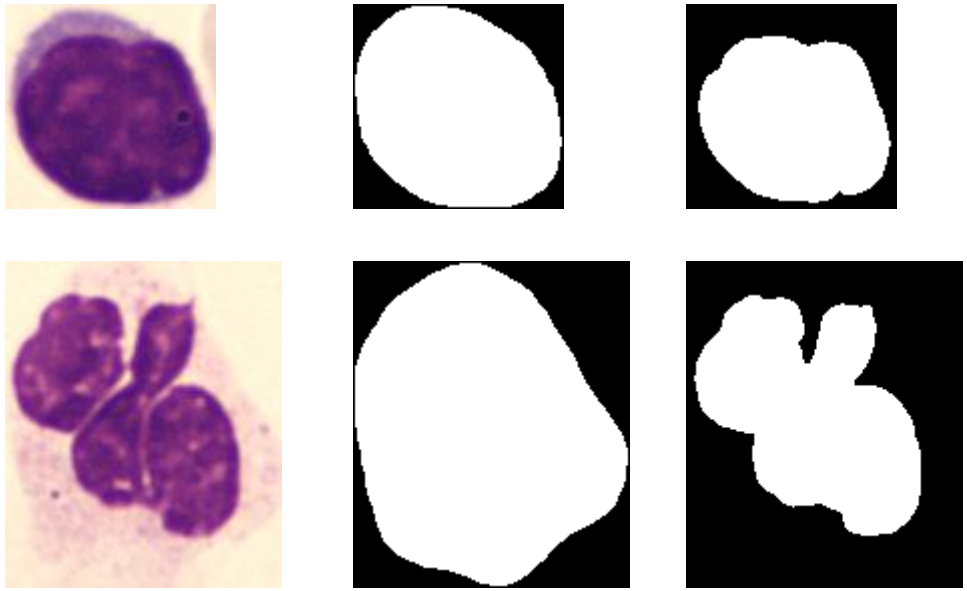
Find nucleus and cytoplasm

- Threholding
- Active contours



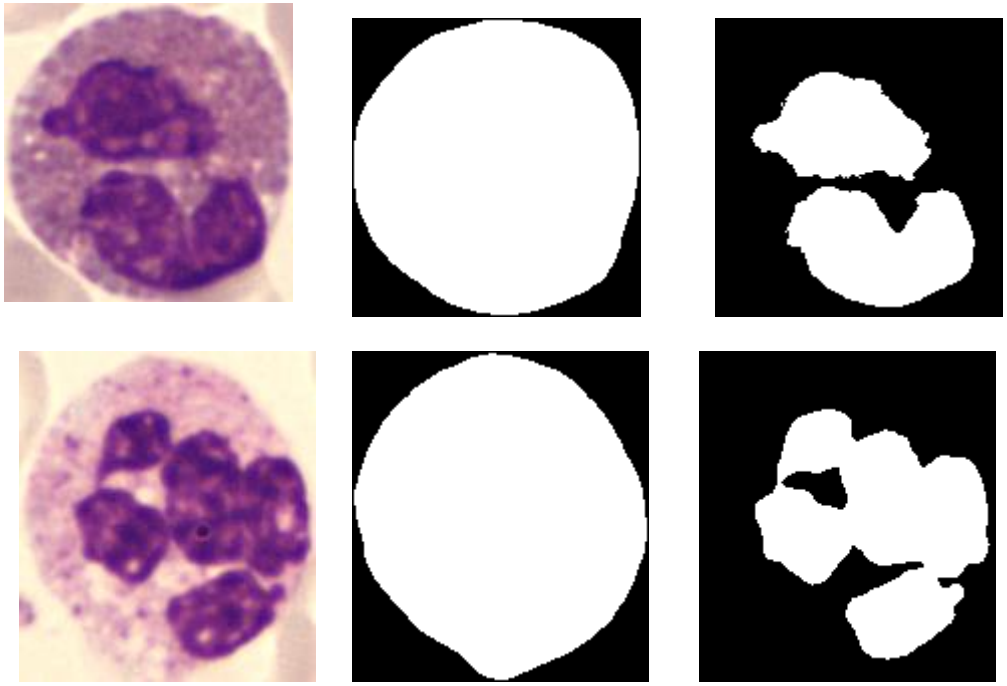
Shape

E.g. circumfence, area, roundness, nucleus part of entire cell area, curvature of contour etc.



Colour

E.g. mean intensity of the colour channels in the nucleus and cytoplasm and their relative strength to each other.



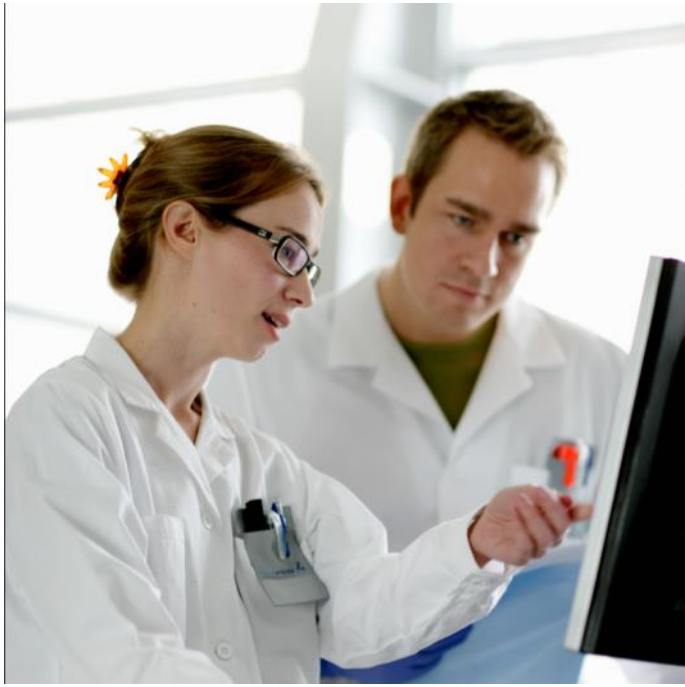
Texture

Granularity, invariant moments, fractal dimension, wavelets.

CELLAVISION

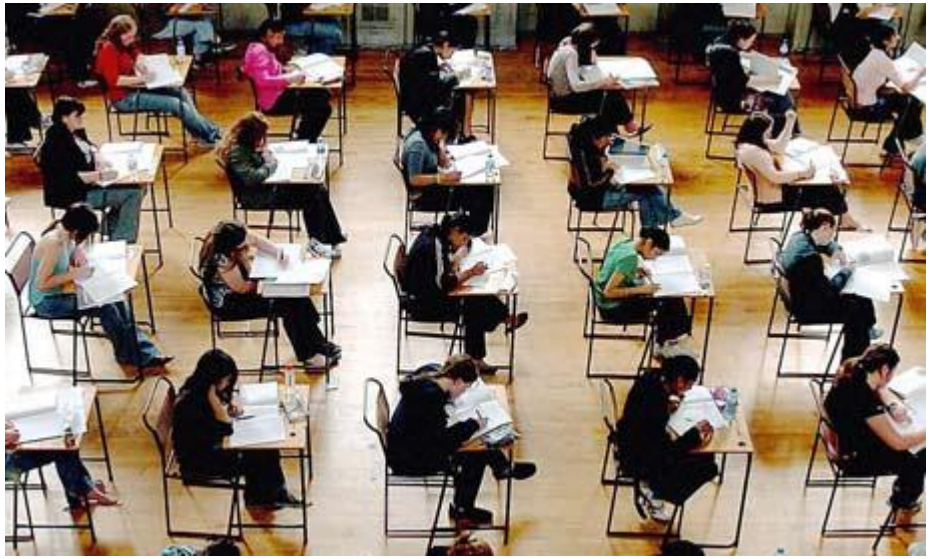
New reduced problem

- Given a set of some hundred featuers determine class.

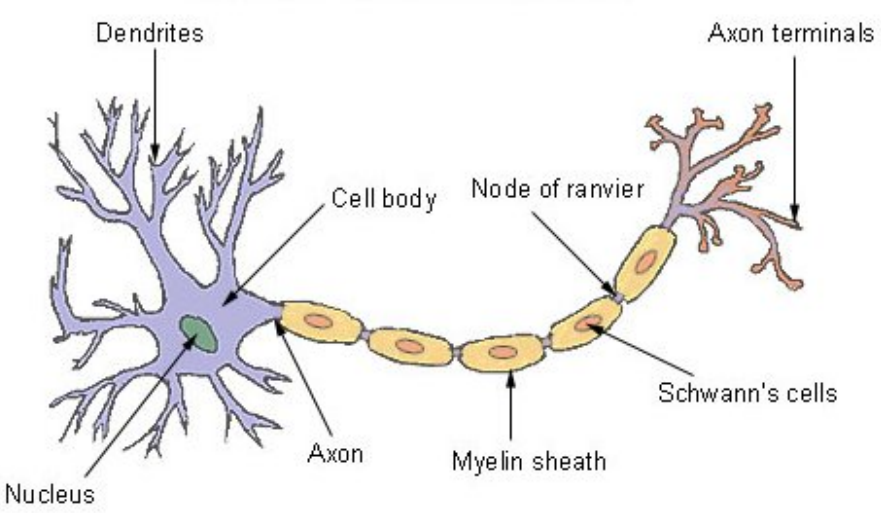


CELLAVISION

How to select your Neural Network



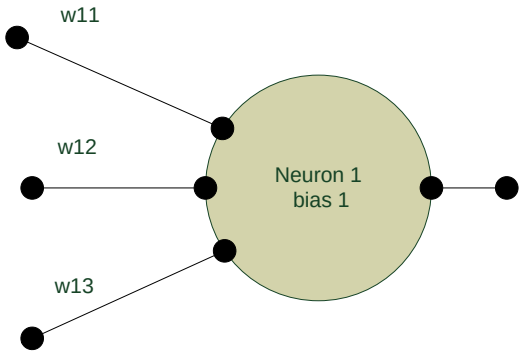
Neurons in biology

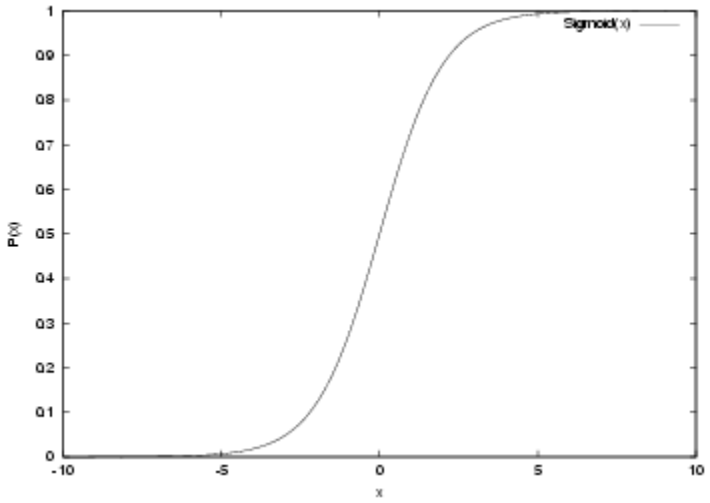


- Inputs from the dendrites
- If inputs strong enough the neuron ’fires’
- A signal is sent through the axon
- Communications with other neurons dendrites via synapses

CELLAVISION

Mathematical representation

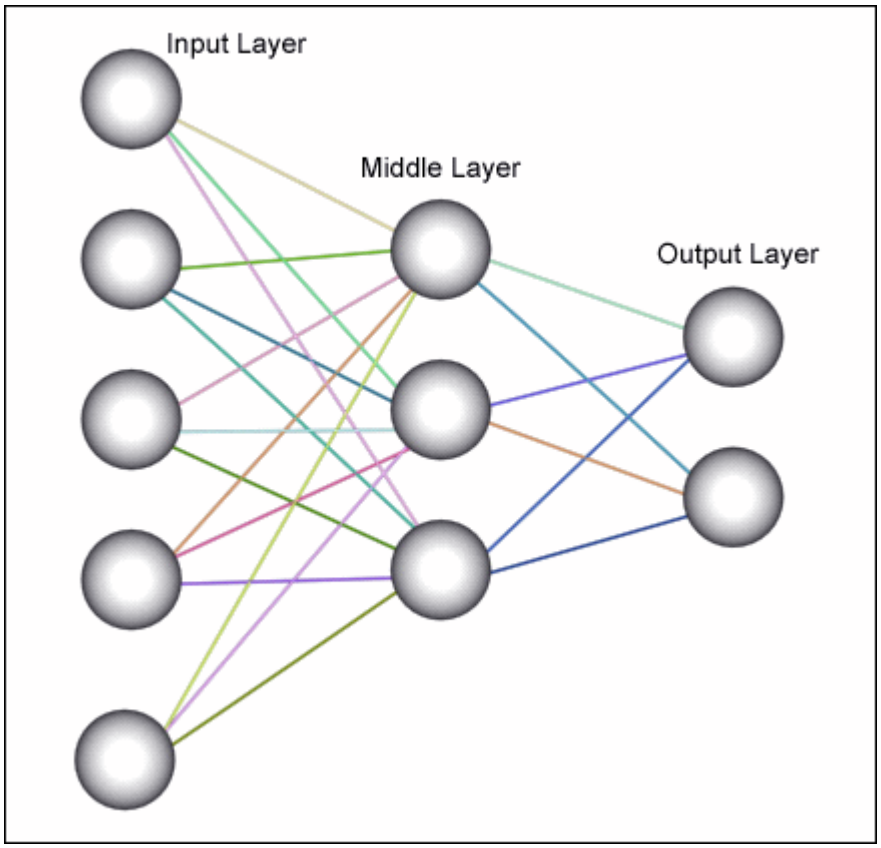




- Output = *f*(w11*input1+w12*input2+w13*input3+bias1)
- *f* as above
- Since *f* is almost constant 0 for small values, almost constant 1 for large and grows rapidly close to zero it mimics the on/off behavior of a neuron.

CELLAVISION

A common way of constucting a neural net



- Multi layer Neural Net
- Input layer
- One or more hidden layers
- Output layer

CELLAVISION

How to use multi layers ANN

- Input layer, here some numbers representing what we want to classify goes. These are called ’features’.
- Hidden layer, huh? 
- Output layer. Assign each output neuron to one of the classes.
- #input neurons = #features
- #hidden neurons ????
- #output neurons = #classes

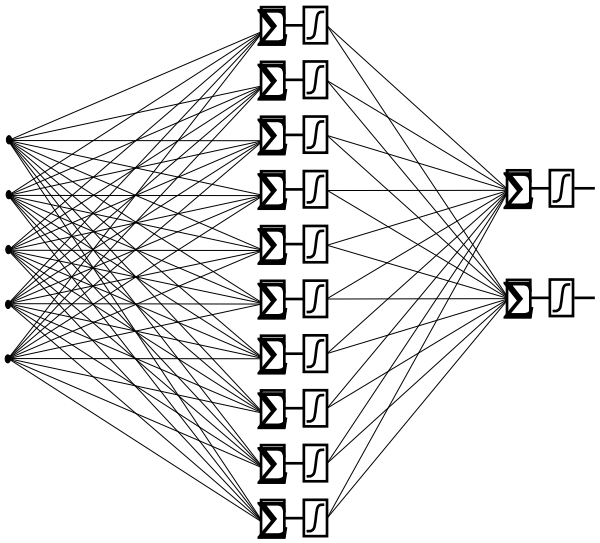
CELLAVISION

How to classify

- Show the features to the Neural Network.
- Look at the outputs. Ideally neuron for the correct class ’fires’ (=1) the others remain silent (=0).
- In practice. Output neuron with largest output = correct class. Output strength is a measure of how certain the classification is (confidence).
- Only one tiny problem left. How do the net know which class is right?

How to select your training method

Training of Neural Networks

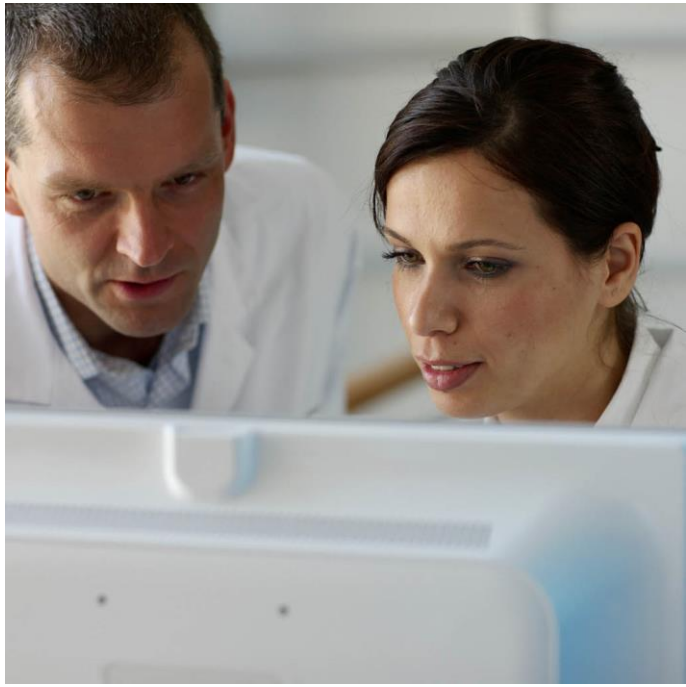


- The nets memories or experience is stored in the weights. (w11,w12 etc.)
- By setting the weights right the net can classify.
- There is a method called back-propagation which can be used to set the weights given a number of examples with known class and known feature values.

CELLAVISION

Back propagation

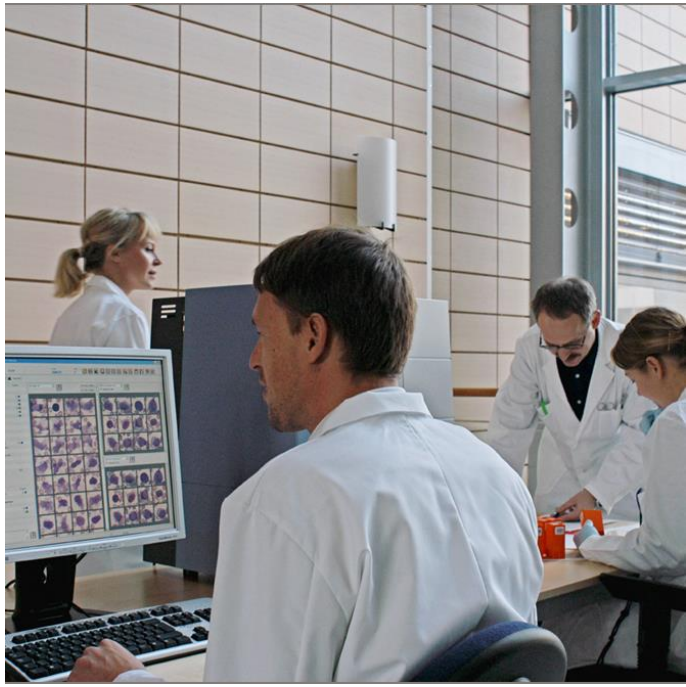
- The difference between the output for an known example and the expected (correct) output is used to update the weights by letting these differences propagate from the outputs back towards the inputs.
- Works if *f* can be differentiated in a simple way.



CELLAVISION

Hidden layer(s)

- The more hidden nodes the more complex problems the net can solve.
- A net with one hidden layer can be just as complex as one with several given the nodes in that layer are increased.
- Most common is one hidden layer.



CELLAVISION

Over training

- If the number of weights in a Neural Net is larger than the number of examples the net can learn the examples ’by heart’.
- This is over training. The net will not be very good at classifying new examples.

How to counter over-training

- Split up dataset into training examples and test examples.
- Stop training process when the change between the steps is small enough.
- More training examples than connections.

CELLAVISION

Optimal brain damage

- Remove connections with small effect (pruning) on classification and train again. One such method is called optimal brain damage.
- The goal is to get complex nets with many nodes but still not to many connections.

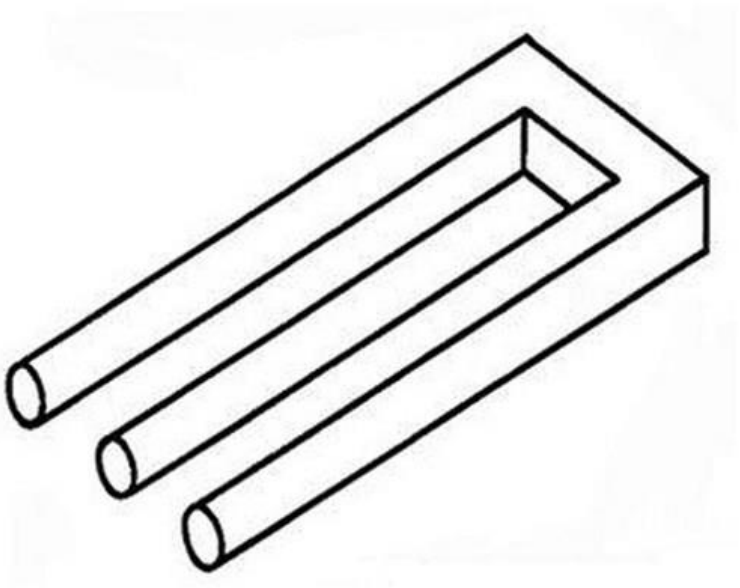
CELLAVISION

Multiple expert systems

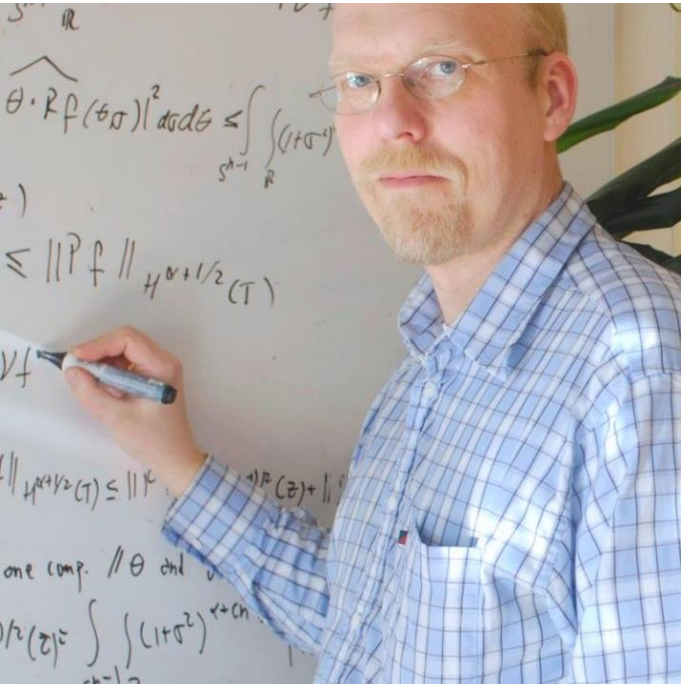
- The classification can be improved by letting several Neural Networks collaborate.

CELLAVISION

How to confuse your Neural Network



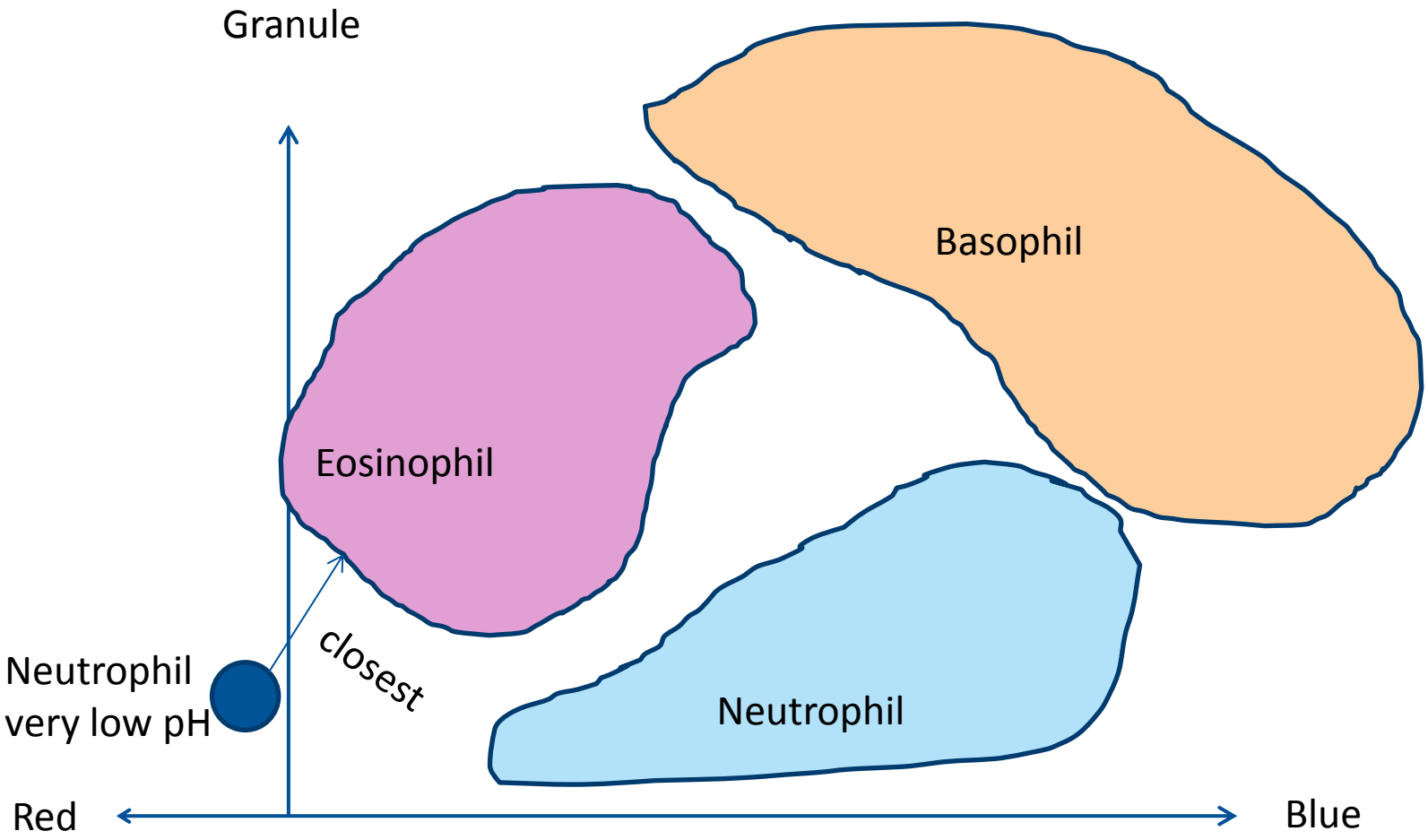
Problems leading to misclassifications



- Neural Nets can only classify cells that are generalizations of cells in the database. New staining, unfamiliar pH etc. are problematic.
- Neural Nets can only classify cells based on the features. Cannot differentiate between cells where the differences is not described by the features.

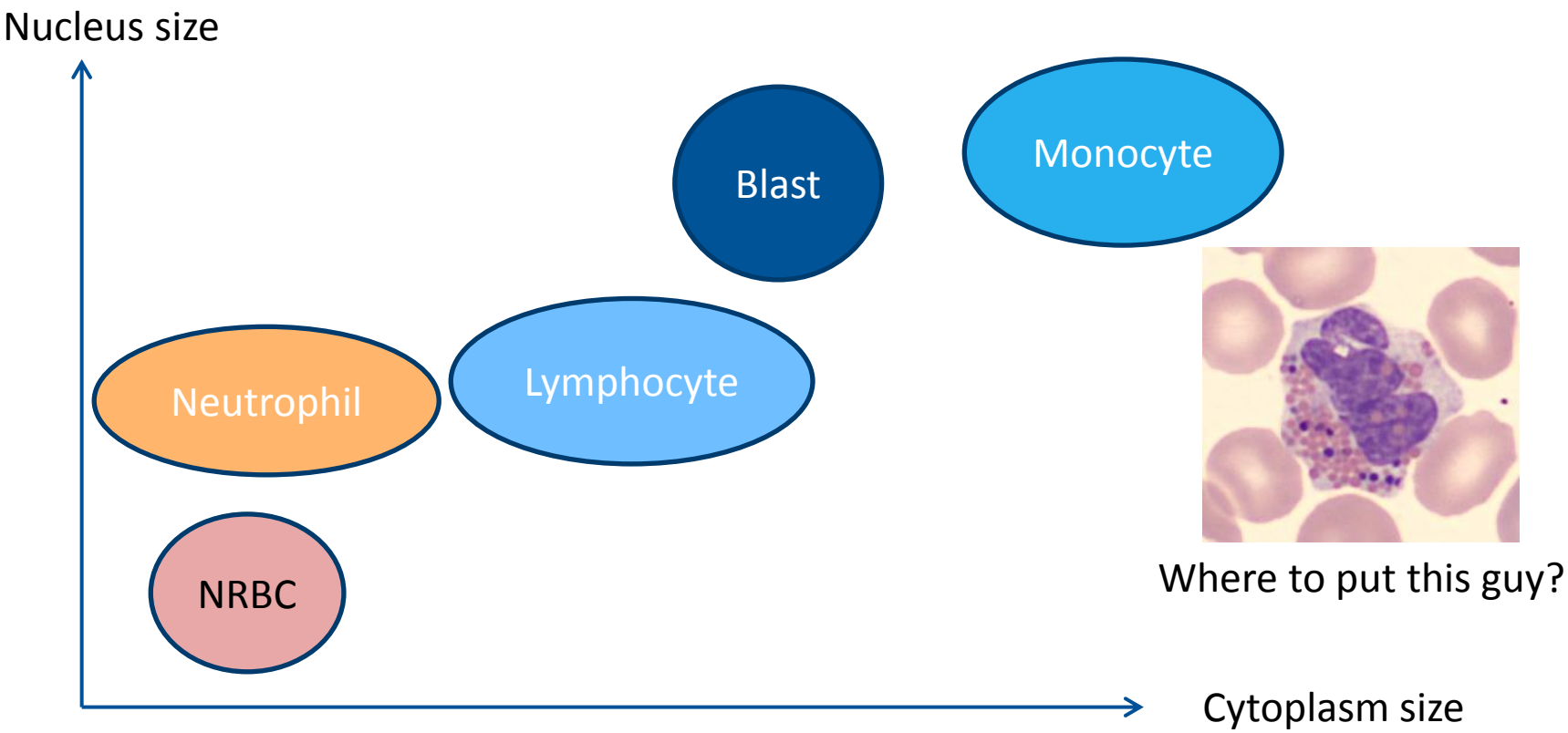
CELLAVISION

Example data outside database set



CELLAVISION

Data not described by features



CELLAVISION

Sarah Connor?

Terminator only given one feature, targets name.

CELLAVISION

Conclusions

Things to consider when training your Neural Network

- Input data
  - Database of cells
  - Segmentation
  - Features
- Network structure
  - Size of Neural Network
  - Pruning
  - Multiple experts
- Training method

CELLAVISION