How to train your Neural Network
Artificial Intelligence in pop culture

’I need your bike, your boots and your DM system!’
 Artificial Intelligence at CellaVision

’Suggest class for this WBC.’
Breaking down the problem (not in all methods)

Segmentation

Feature Extraction

AI structure
ANN/SVM...

Databas of correctly classified cell images

Training method

Tools
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.
Complexity of the problem

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

Find the cells on the image
Segmentation

Find nucleus and cytoplasm

- Threholding
- Active contours
Shape

E.g. circumference, 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.
New reduced problem

• Given a set of some hundred features determine class.
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
• Output = \( f(w_11*\text{input}_1+w_12*\text{input}_2+w_13*\text{input}_3+\text{bias}_1) \)
• \( f \) 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.
A common way of constructing a neural net

- Multi layer Neural Net
- Input layer
- One or more hidden layers
- Output layer
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
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. \((w_{11}, w_{12} \text{ etc.})\)
- 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.
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.
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.
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.
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 too many connections.
Multiple expert systems

- The classification can be improved by letting several Neural Networks collaborate.
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.
Example data outside database set

Granule

- Eosinophil
- Neutrophil

Neutrophil very low pH

Red

Closest

Blue
Data not described by features

Nucleus size

Neutrophil

Lymphocyte

NRBC

Blast

Monocyte

Cytoplasm size

Where to put this guy?
Sarah Connor?

Terminator only given one feature, targets name.
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