University of
South Australia

# COMP 2019

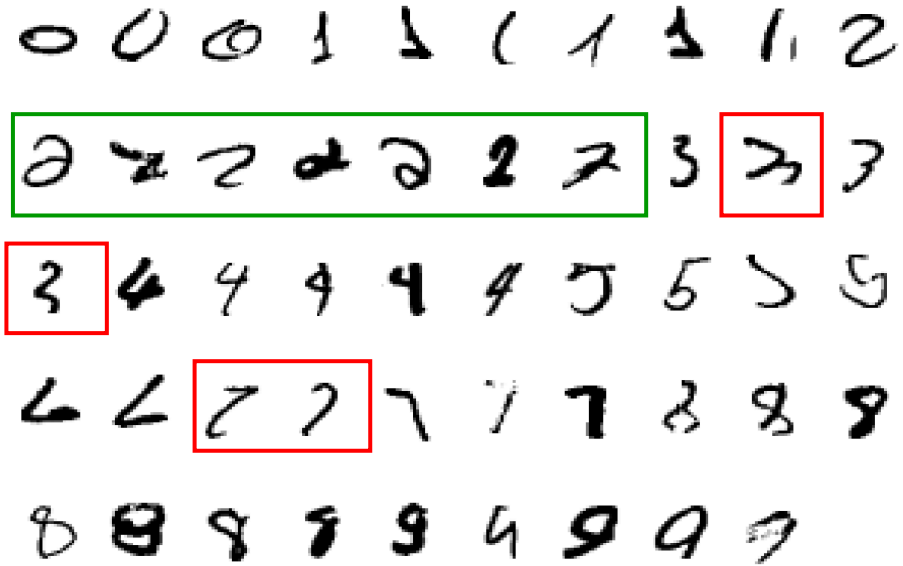
## Week 9

## Neural Nets and Computer Vision

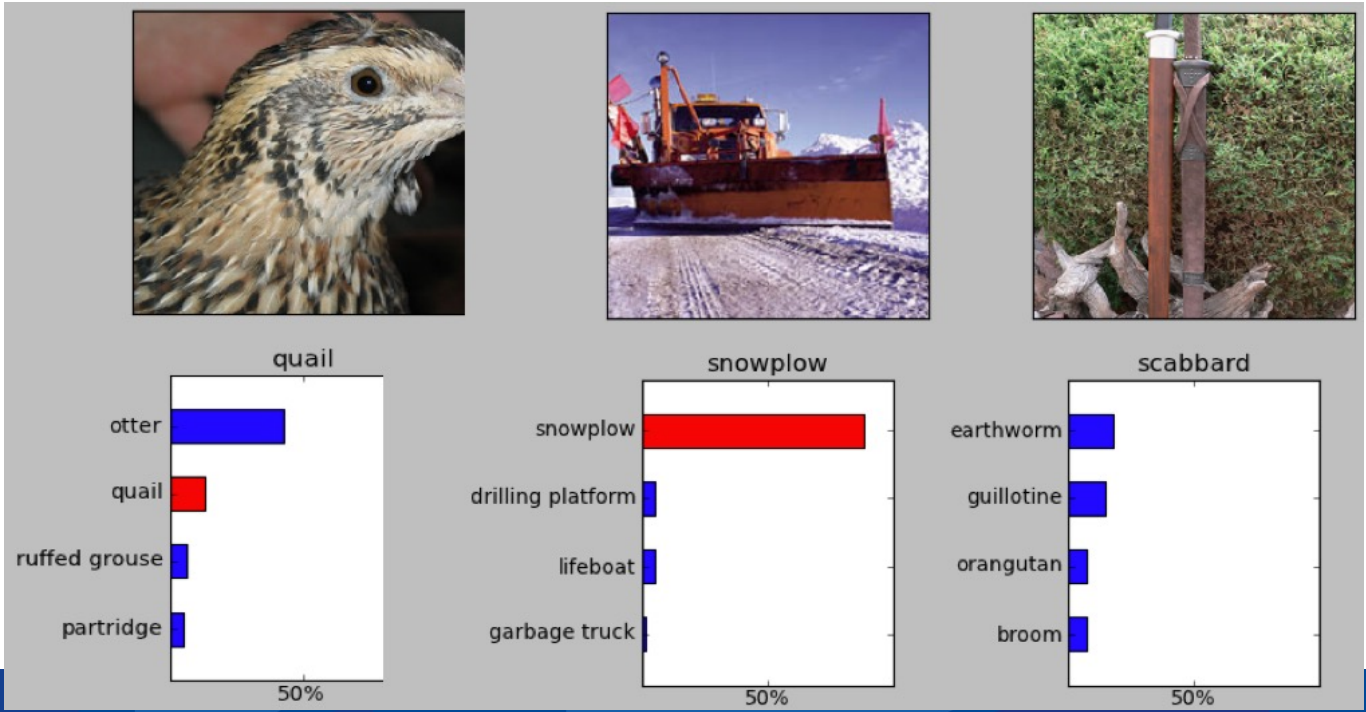
# Learning Objectives

- Understand the principles of Neural Computing (CO3)
- Explain how Deep Neural Nets work (CO3)
- Explain how Deep Learning is applied to solve Computer Vision tasks (CO3)

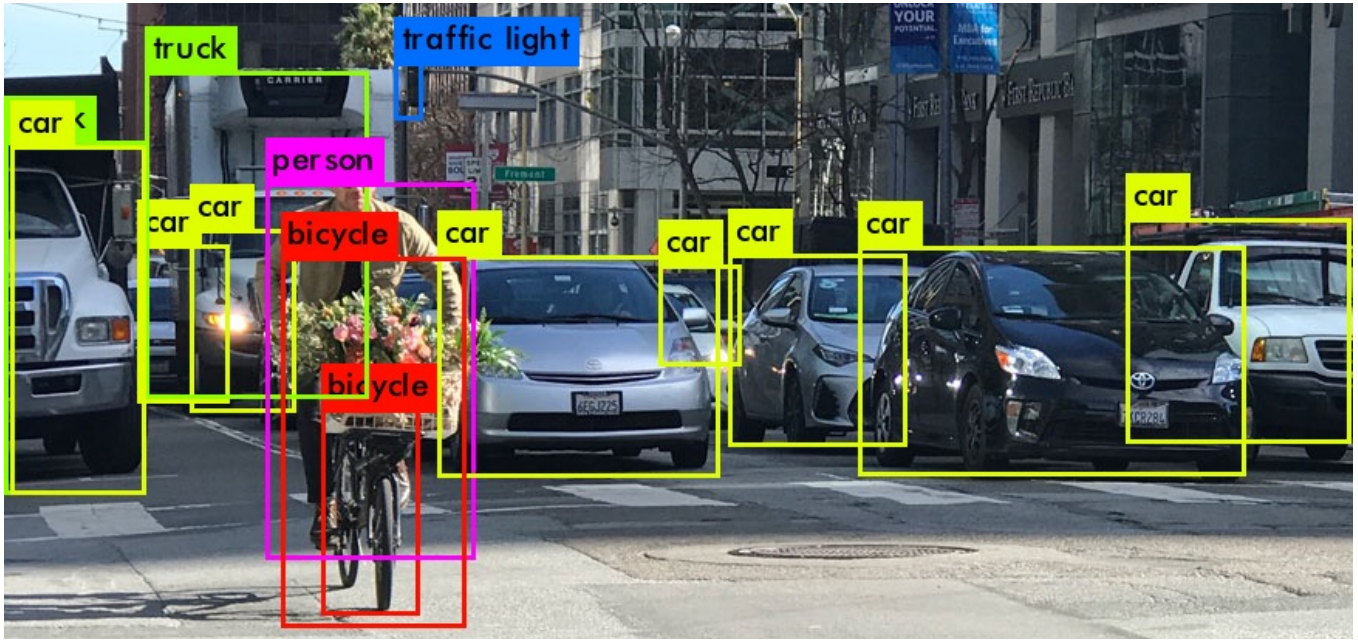
# Handwriting Recognition



# Object Classification

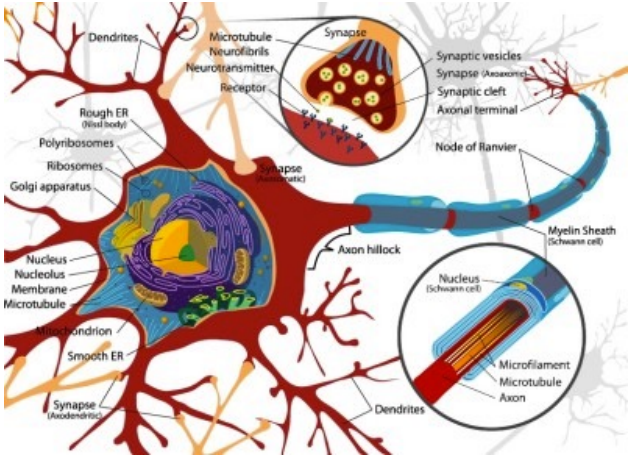


# Object Tracking



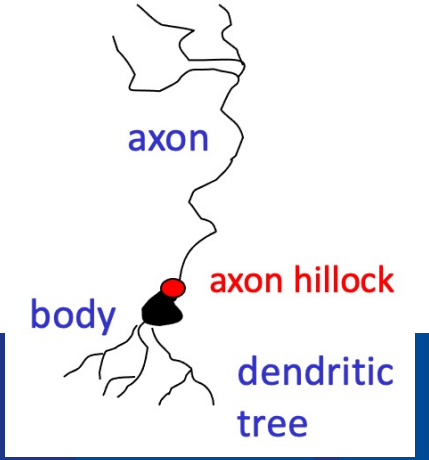

https://ai-pool.com/a/s/yolov3-and-yolov4-in-object-detection

# Neuron

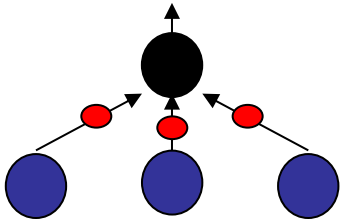



http://en.wikipedia.org/wiki/Neuron

- Simplified physical structure:
  - One axon that branches
  - Dendritic tree that collects input from other neurons
- Axons contact dendritic trees at synapses
  - A spike of activity in the axon causes charge to be injected into the post-synaptic neuron
  - Effectiveness of the synapses can be changed
- Spike generation:
  - An axon hillock generates a spike whenever enough charge has flowed in from dendritic tree



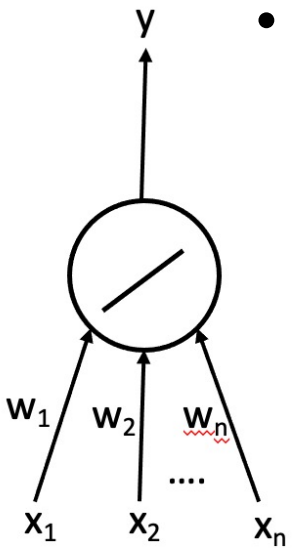
# The Brain



- Each neuron receives inputs from other neurons
- The effect of each input on the neuron is controlled by a synaptic weight (positive or negative)
- The weights adapt so that the network learns
- We have about $10^{11}$ neurons each with about $10^4$ weights
- Massive parallelism, much better than a computer

- Computer models are different from how the actual brain works
  - Real numbers instead of spikes, structure of connections, etc

# Linear Neuron




- Output y is a weighted linear combination of the inputs $x_1,\ldots,x_n$
  - Weights $w_1,\ldots,w_n$ can be positive or negative

$$y = b + \sum_i x_i w_i$$

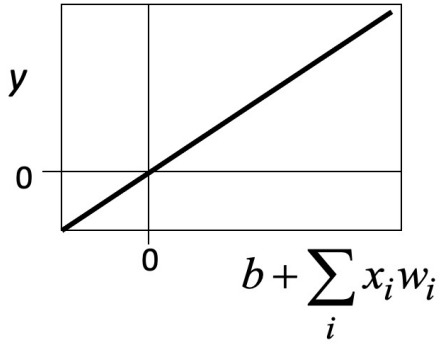
bias

$i^{th}$ input

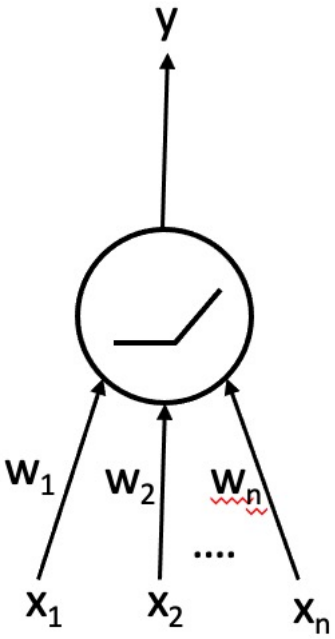
output

index over input connections

weight on $i^{th}$ input



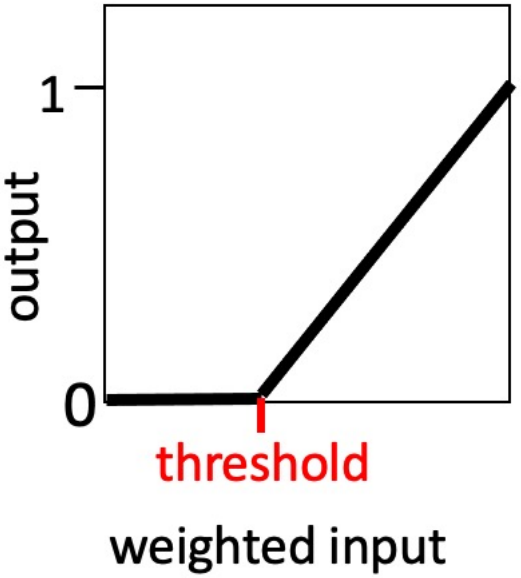
# ReLU



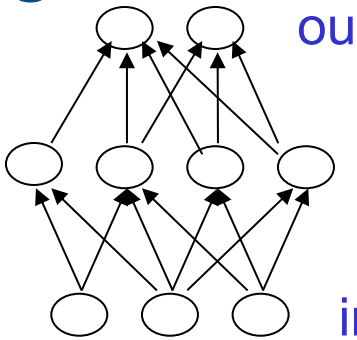

$$z = b + \sum_i x_i w_i$$

Outputs

- z if z > 0
- 0 otherwise



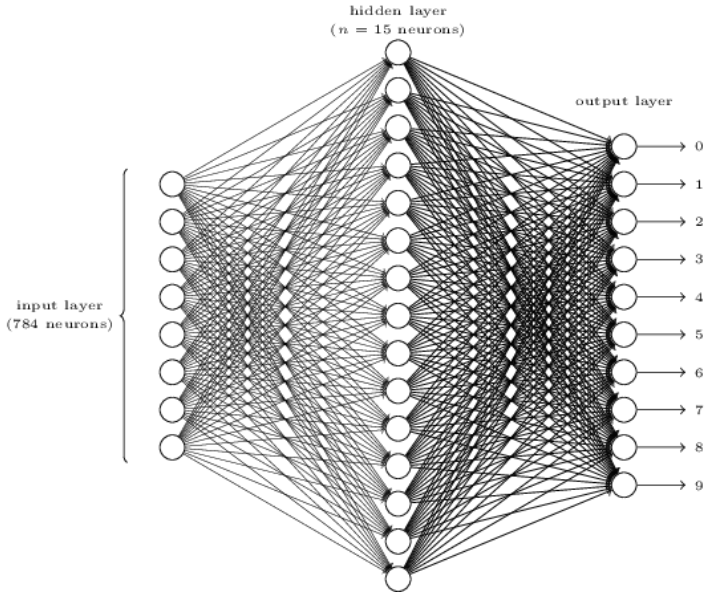
# Feed Forward Architecture




- Layered
  - First (bottom) layer is the input
  - Last (top) layer is the output
  - The layer(s) in-between are called “hidden”
  - If there are >1 hidden layers, the network is called “Deep”
- The output of the neurons in each layer are a non-linear function of the inputs in the layer below
  - No connections within layers

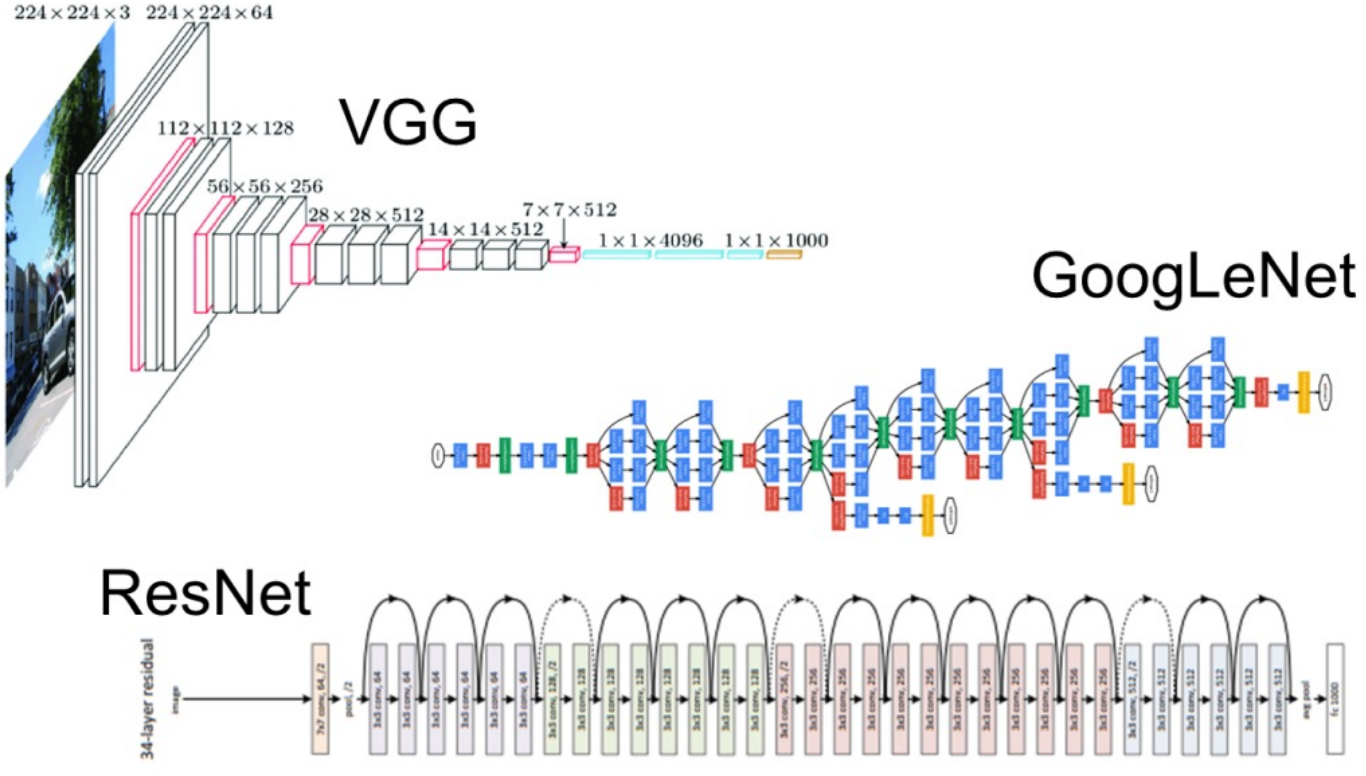
# Neural Network Output

- Regression: predict a real number
  - One output unit
  - The price of a stock next week
- Classification: predict a class label
  - Two classes
    - » One output unit
    - » Is the tumor carcinogenic (yes/no)?
  - Multiple classes
    - » Multiple output units, one per class
    - » Which digit is shown in the image? (0,…,9)



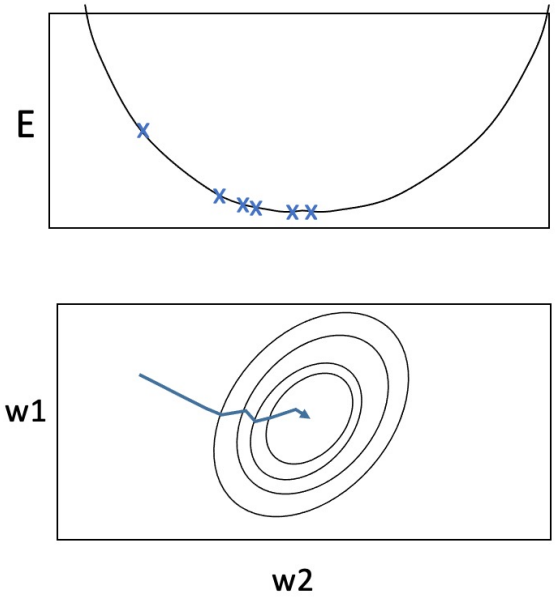

http://neuralnetworksanddeeplearning.com/chap1.html

# Deep Neural Nets



# Training

- Learning as error minimization
  1. Assign (small) initial weights at random
  2. Determine the error the net makes on the training samples
  3. Change the weights a little to reduce that error
  4. Repeat 2&3 until error is small enough

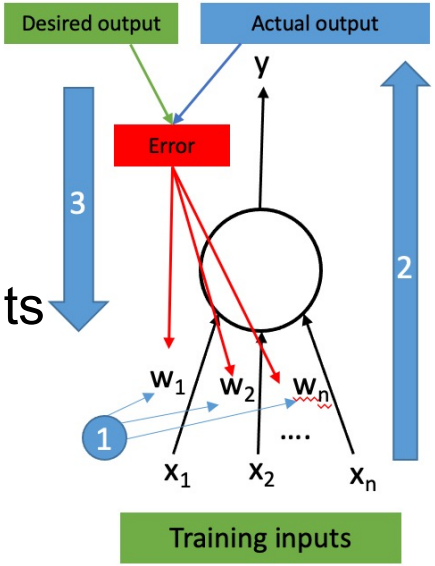


# Backpropagation

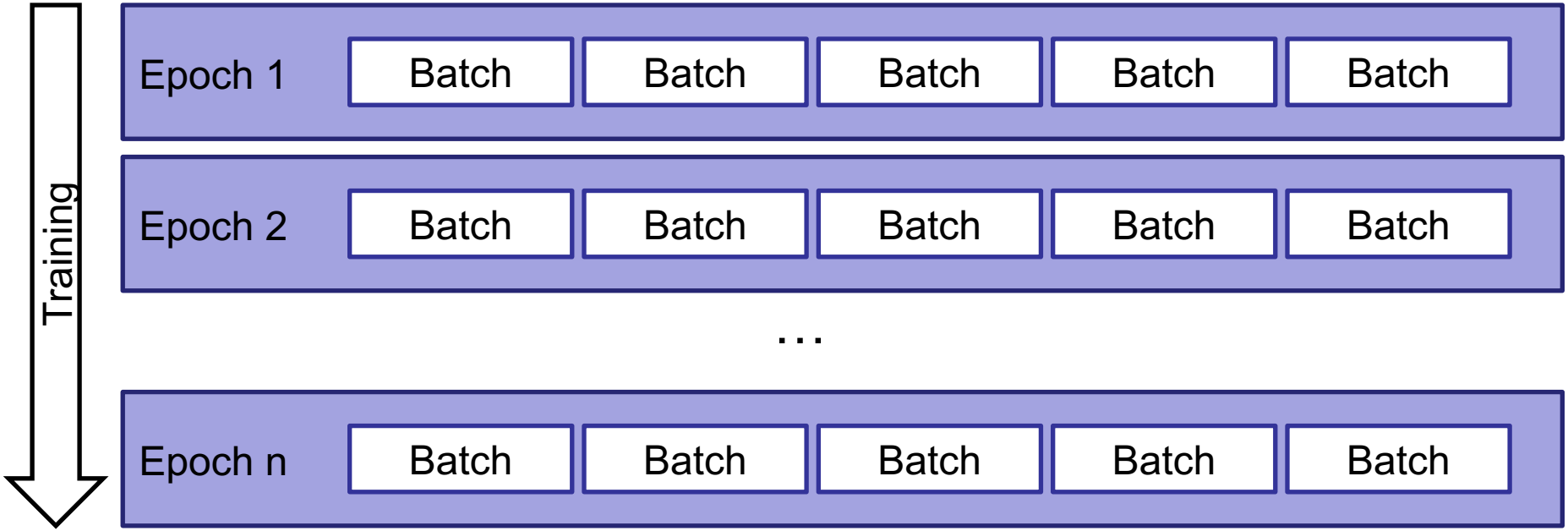
- **Loss function** measures how well the net is doing
- Defines an error surface as a function of the weights

- Find the minimum on that surface by following the direction of steepest descent (**gradient descent**)
- **Gradient**: a vector of partial derivatives w.r.t. the weights

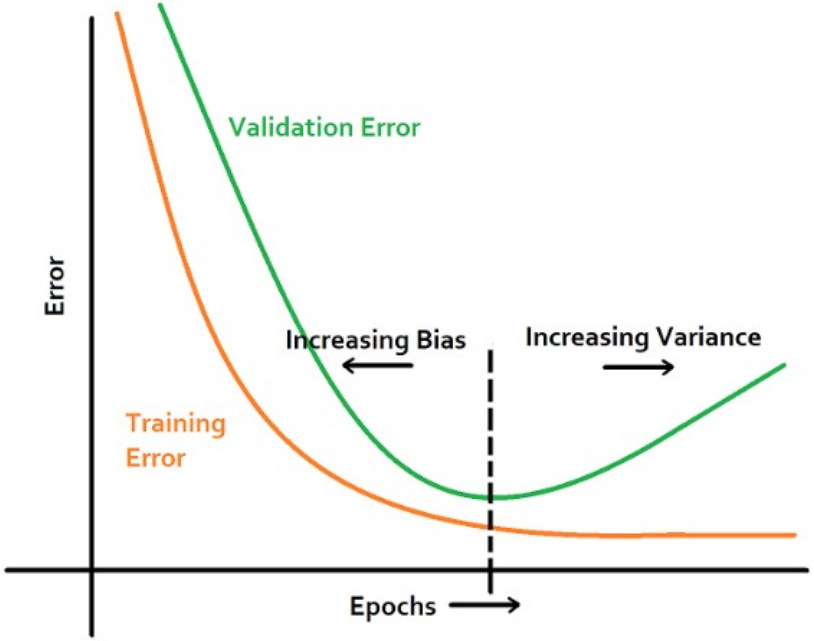
- **Backpropagation**: change each weight proportional to the gradient



# Batches and Epochs



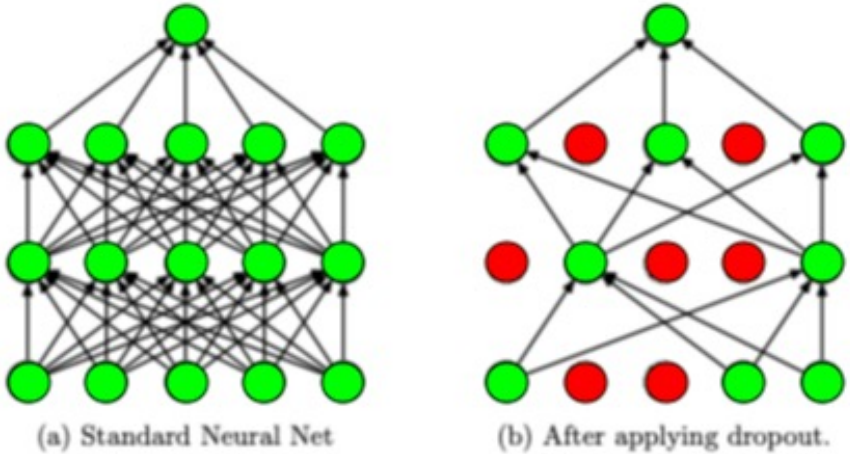
# Preventing Overfitting






https://medium.com/@rahuljain13101999/why-early-stopping-works-as-regularization-b9f0a6c2772

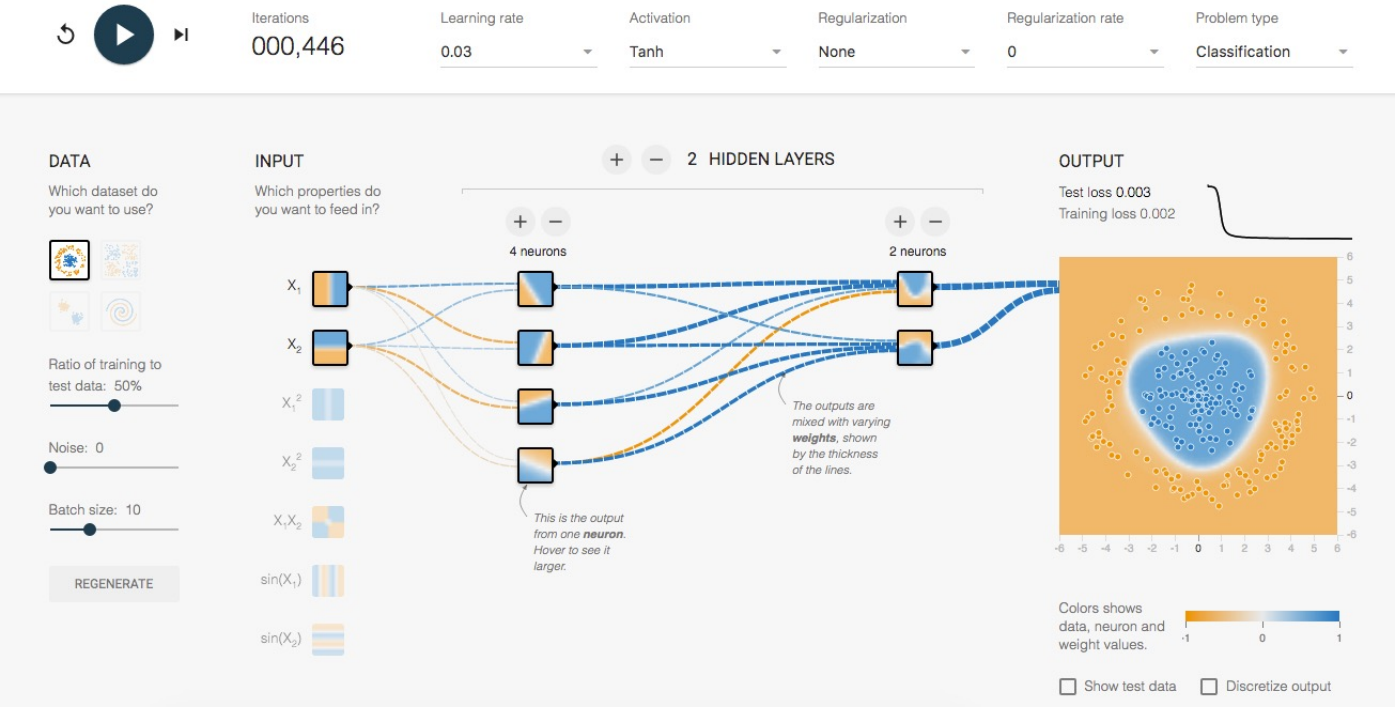
# Preventing Overfitting: Regularisation



(a) Standard Neural Net

(b) After applying dropout.

https://medium.com/analytics-vidhya/a-simple-introduction-to-dropout-regularization-with-code-5279489dda1e

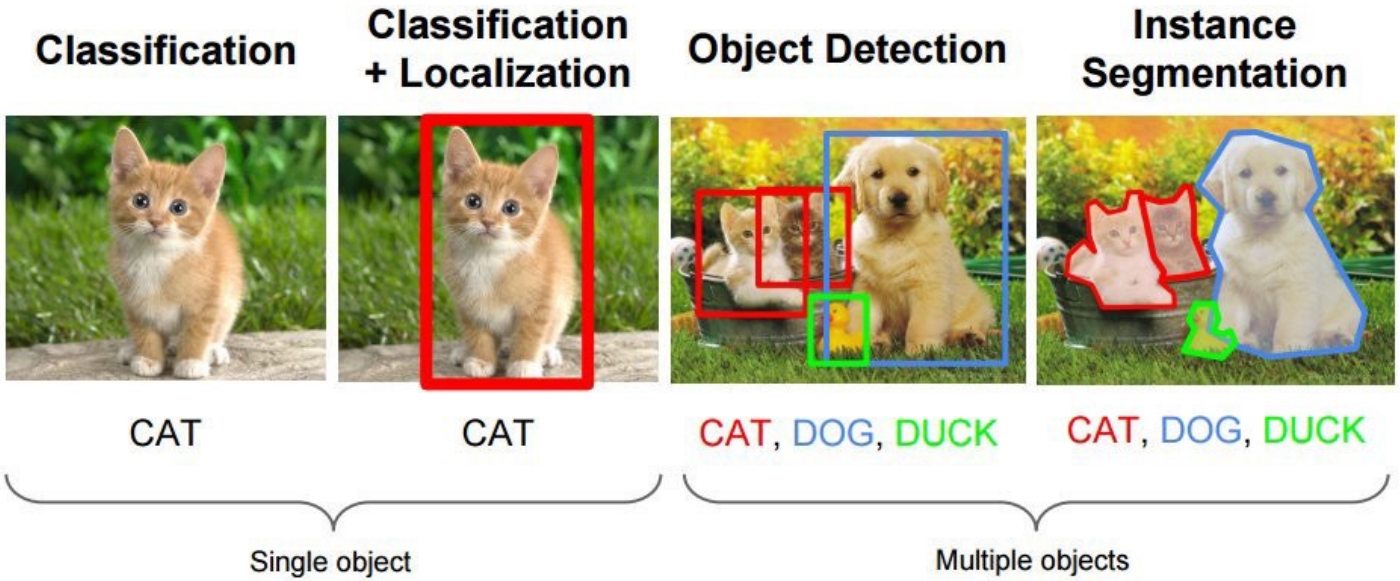



http://playground.tensorflow.org/

# Factors Affecting Results

- Network Architecture
- Number of epochs to train
- Batch size
- learning rate
- Learning rate schedule/decay
- Dropout & regularisation
- Early stopping conditions

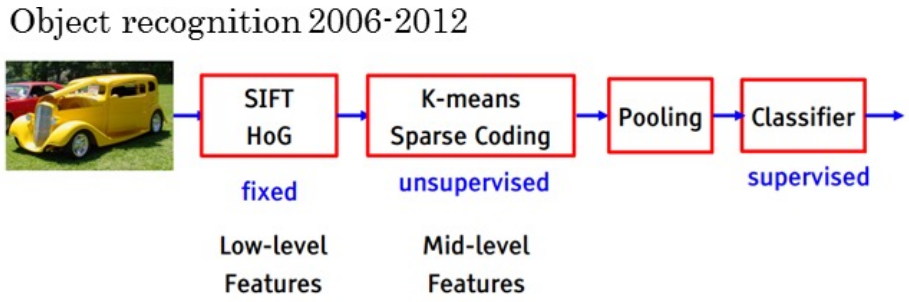
# Computer Vision





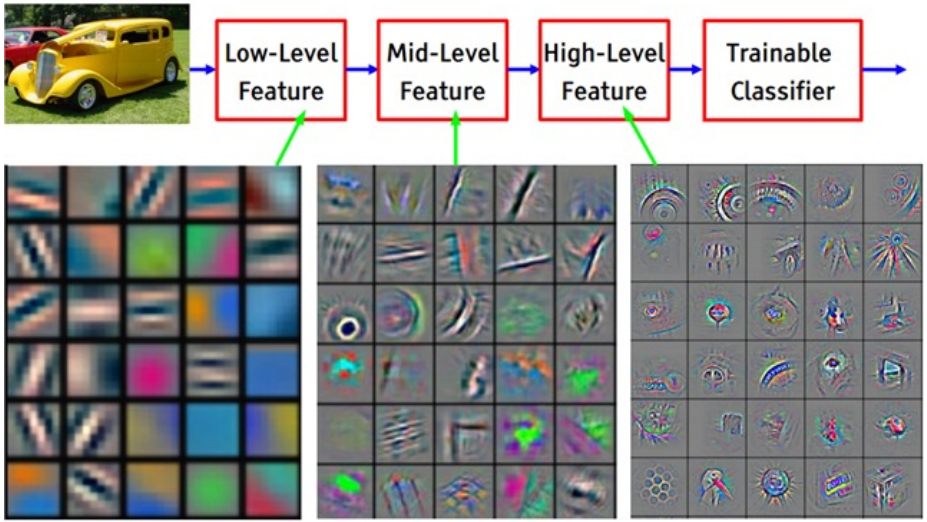

https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e

# Computer Vision Pipeline (until 2012)



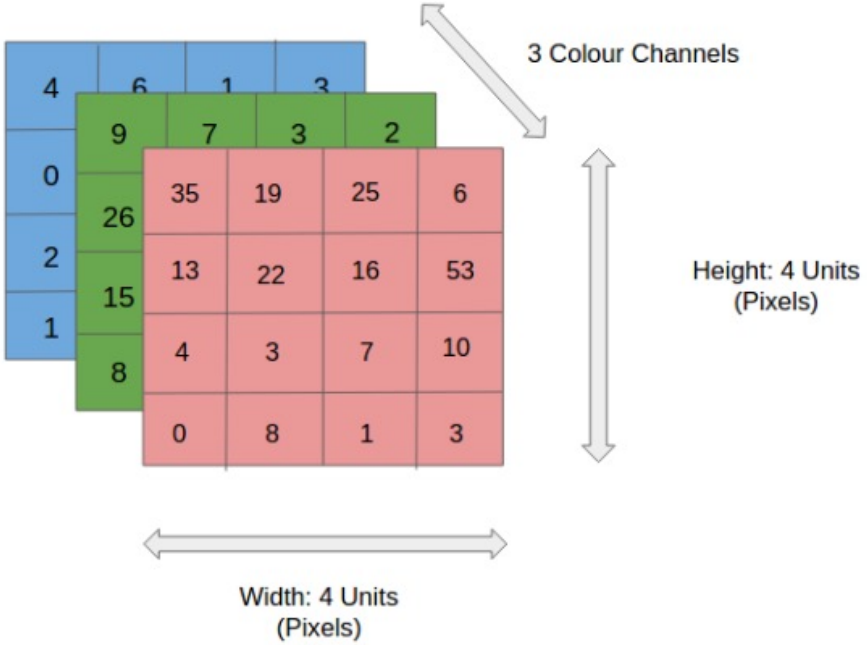

http://www.rsipvision.com/exploring-deep-learning/

# Computer Vision Pipeline (now)



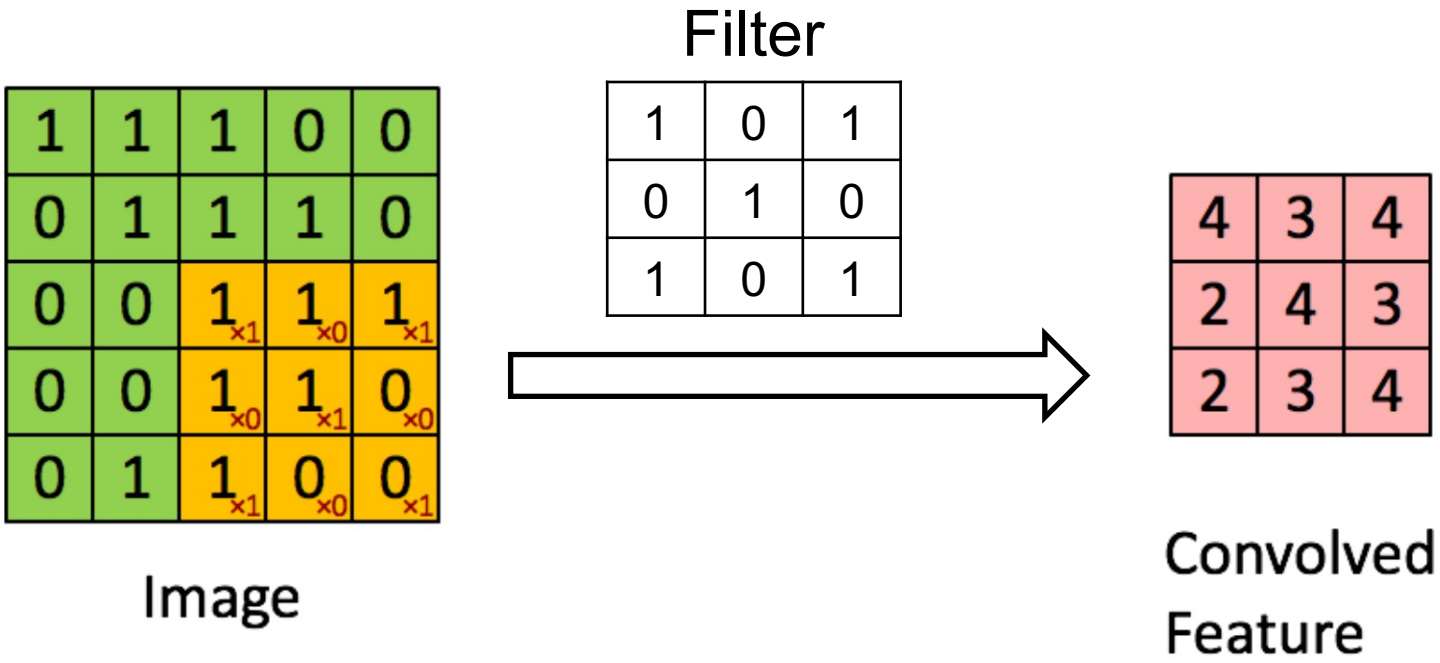

http://www.rsipvision.com/exploring-deep-learning/

# Image Representation




https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

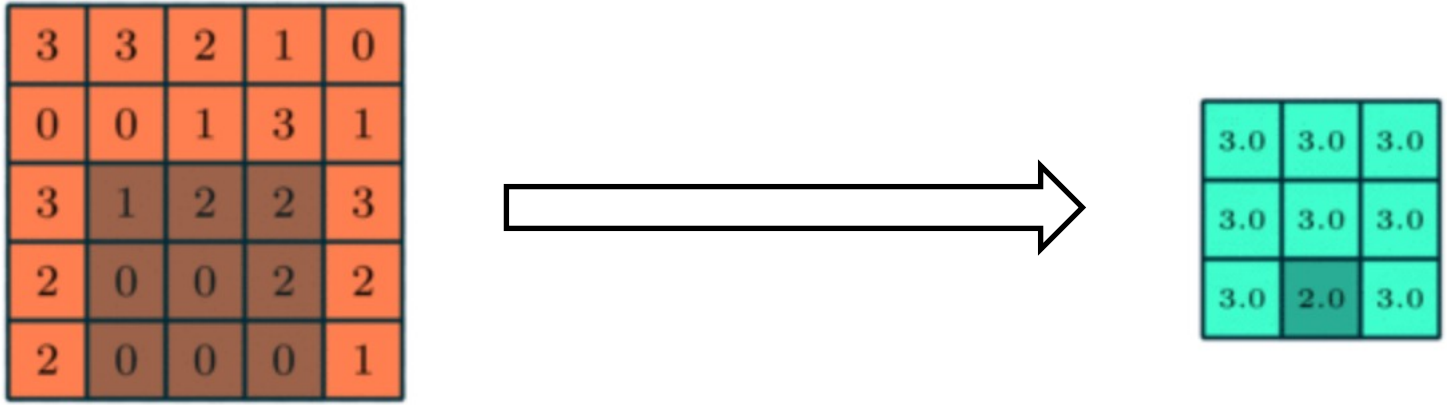
# Convolution





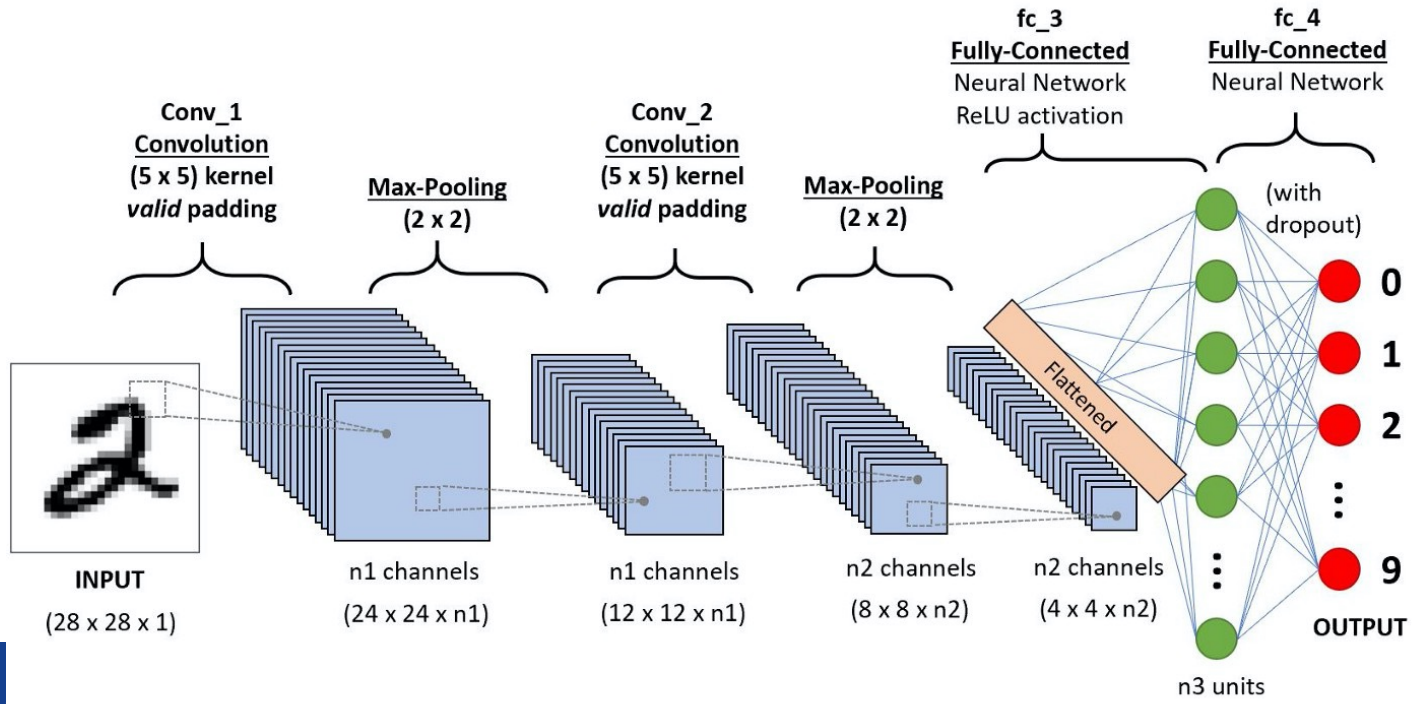

https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

# Max Pooling (3x3)




https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

# CNN: Handwriting Recognition



# Summary

- Neural nets are inspired by how the brain works
- Deep learning is at the core of computer vision, language processing, and reinforcement learning approaches
- Trained on huge datasets using error minimisation
- CNNs have revolutionised image and video processing
- Creating effective neural architectures is still a “black art”
- NNs are complex - use simpler models if possible

University of
South Australia

# Questions?