Neural Networks

Declan Groves

June 30 2016

Outline

What is a Neural Network?

Activations

Convolutional Neural Networks

1 What is a Neural Network?

2 Activations

3 Convolutional Neural Networks

Overview

Neural Networks

Declan Groves

What is a Neural Network?

Activations

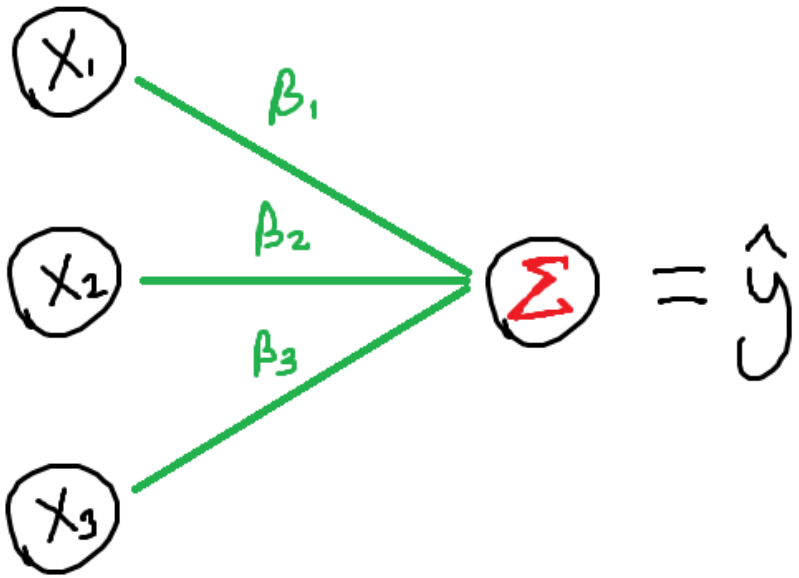
Convolutional Neural Networks

- Hypest ML
- Good at unstructured probelms
- Suboptimal at structured problems

History

- Around since the 1950s
- Resurgence in 1970s
- Resurgence in late 2000s

A graphical linear model



A graphical linear model

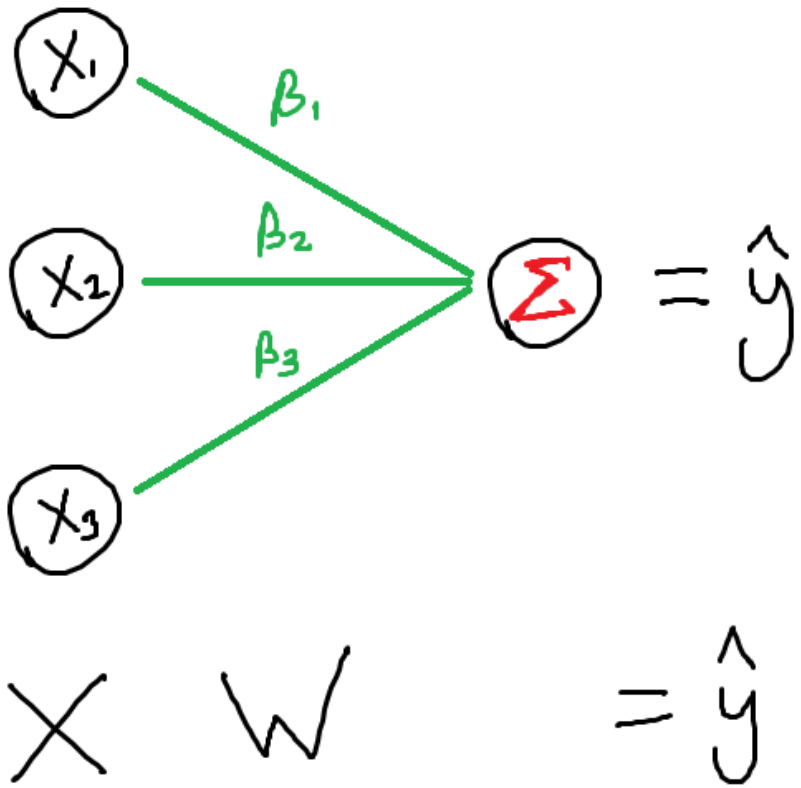
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



A graphical linear model

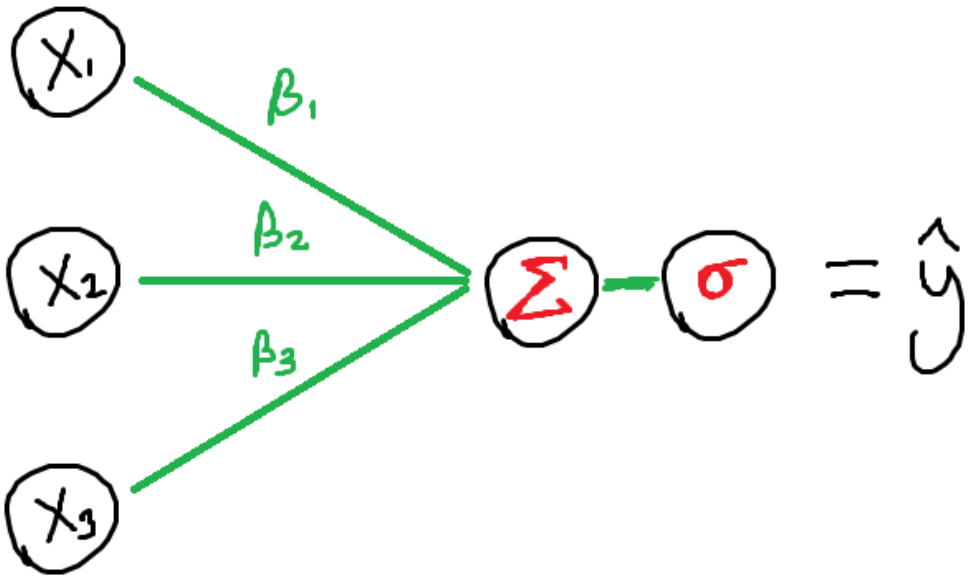
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



$$\sigma(X \quad W) \quad = \hat{y}$$

A graphical linear model

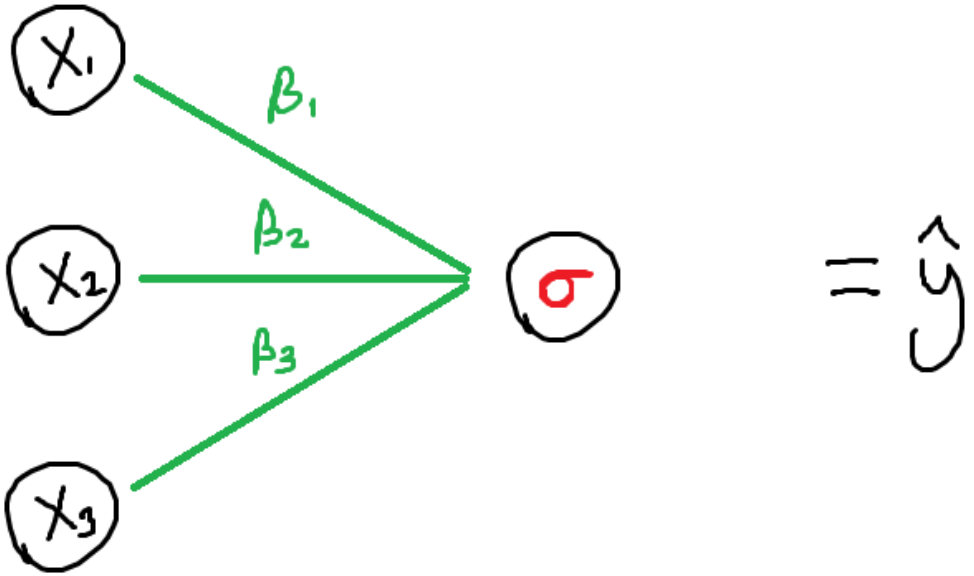
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



$$\sigma(X \quad W) \quad = \hat{y}$$

A simple neural network

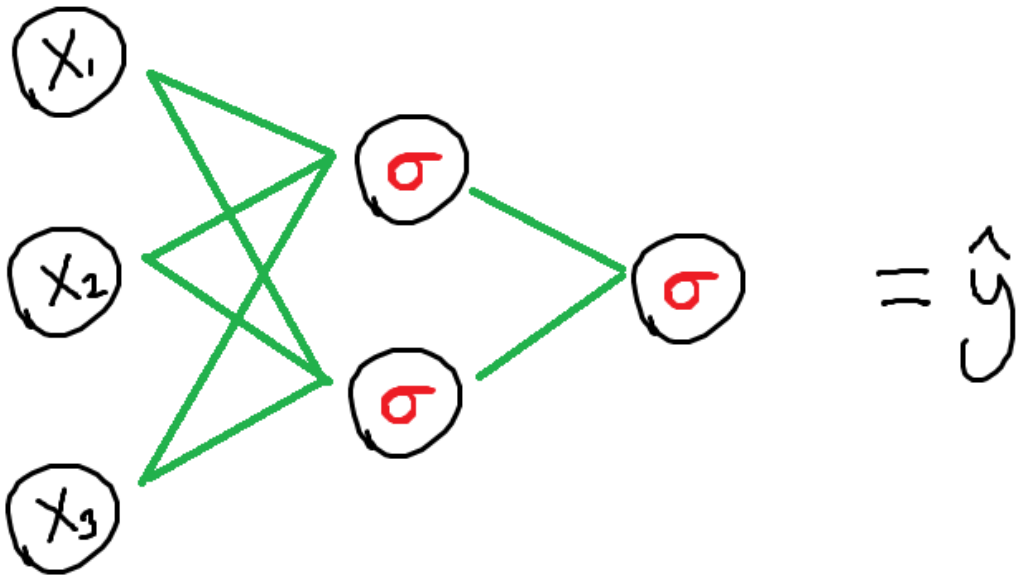
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



A simple neural network

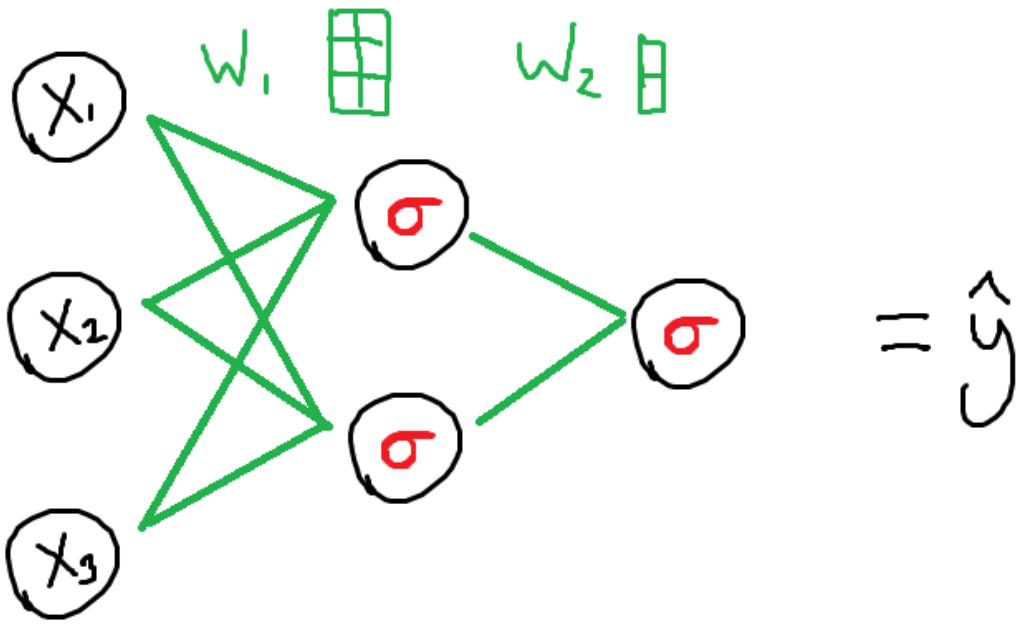
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



A simple neural network

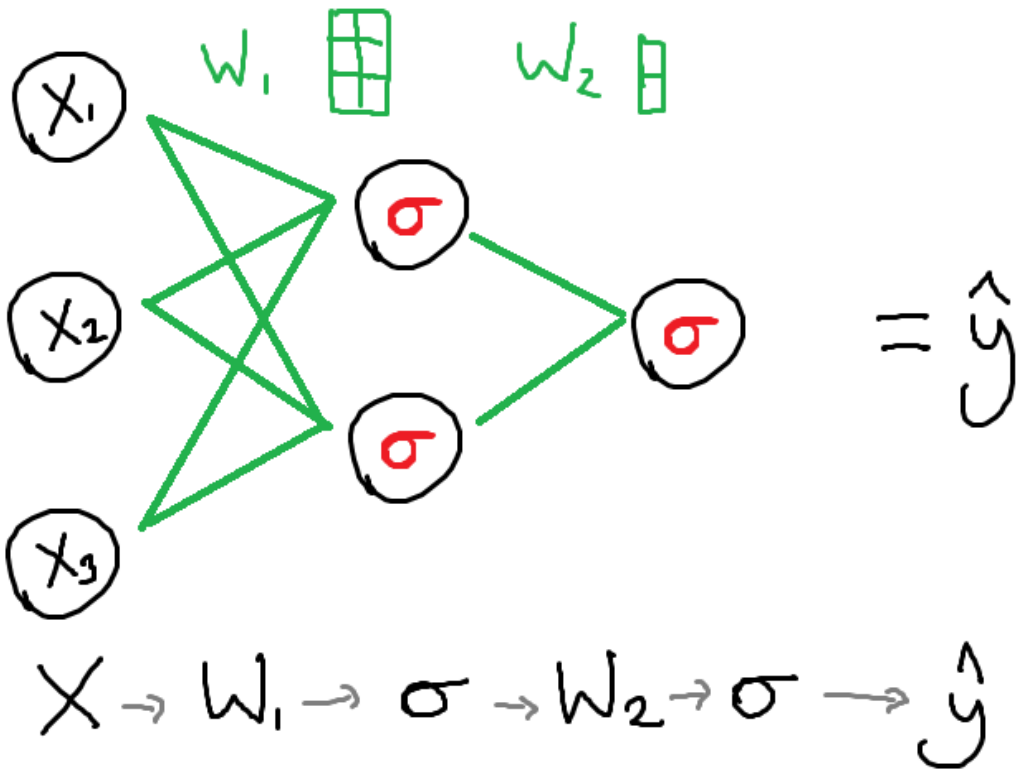
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



A simple neural network

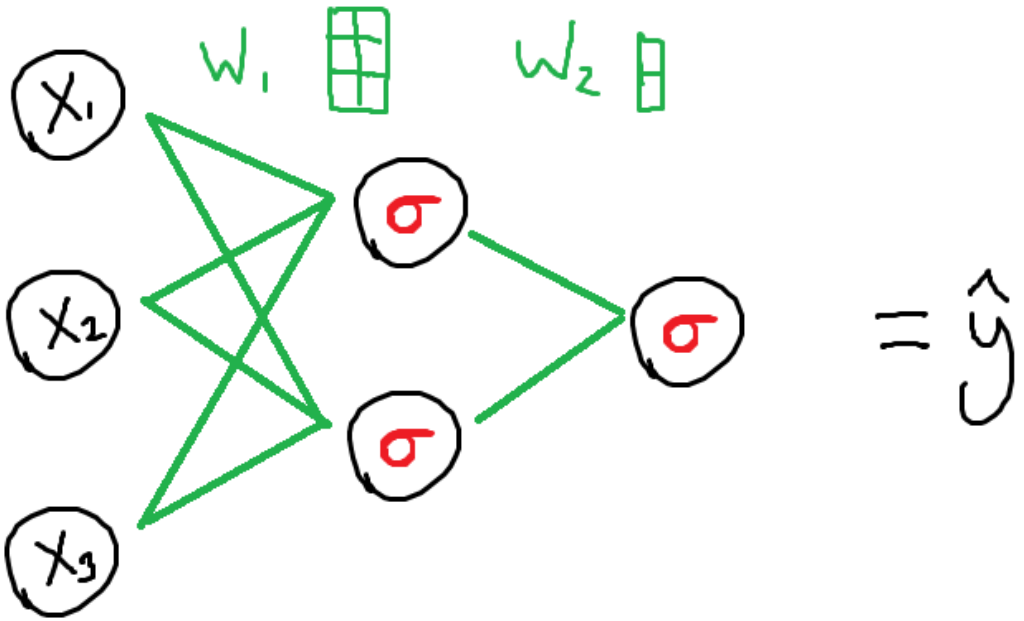
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



$$X \rightarrow W_1 \rightarrow \sigma \rightarrow W_2 \rightarrow \sigma \rightarrow \hat{y}$$

$$\sigma(\sigma(XW_1)W_2) = \hat{y}$$

A multilayer multinomial classifier

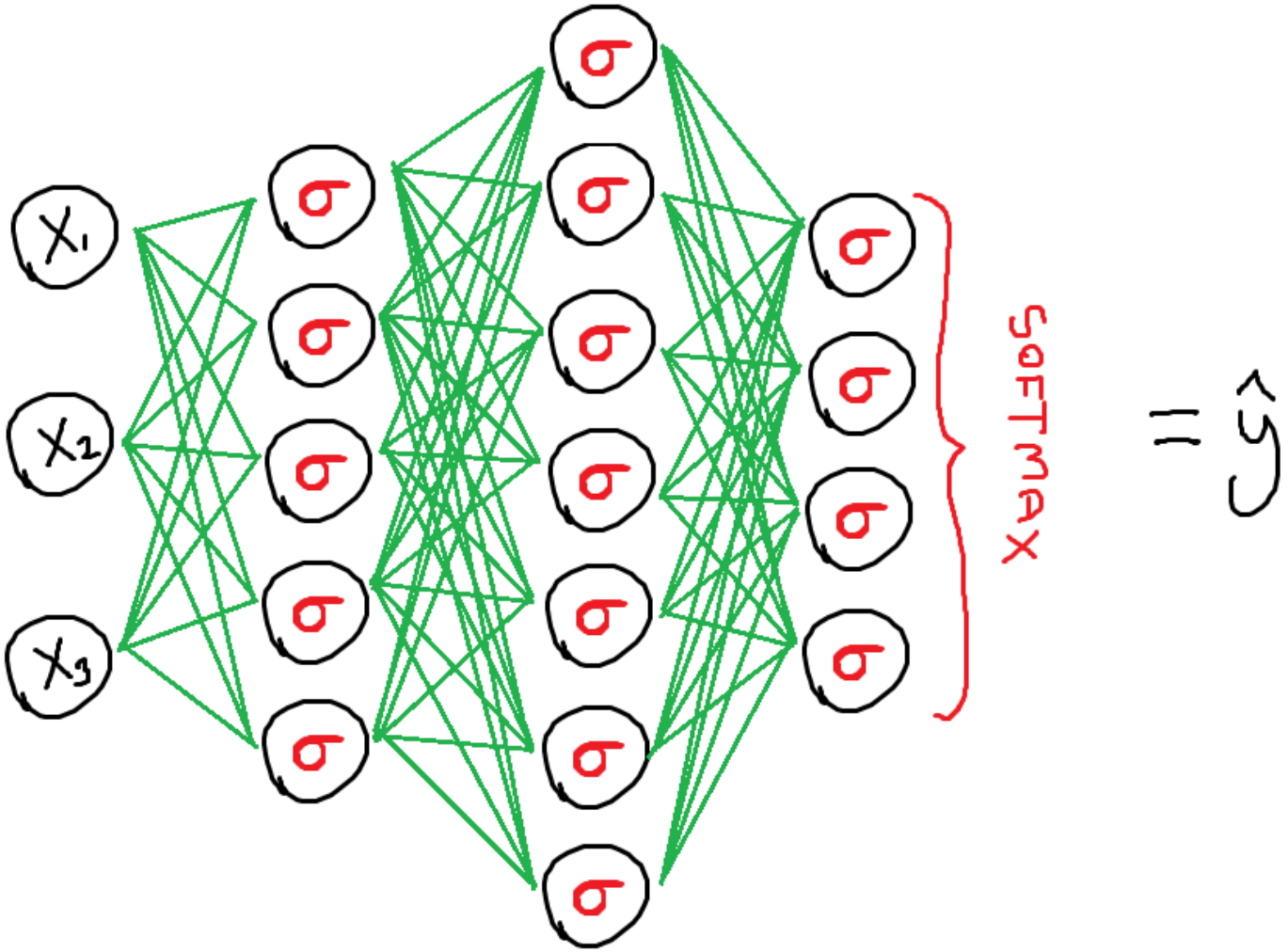
Neural Networks

Declan Groves

What is a Neural Network?

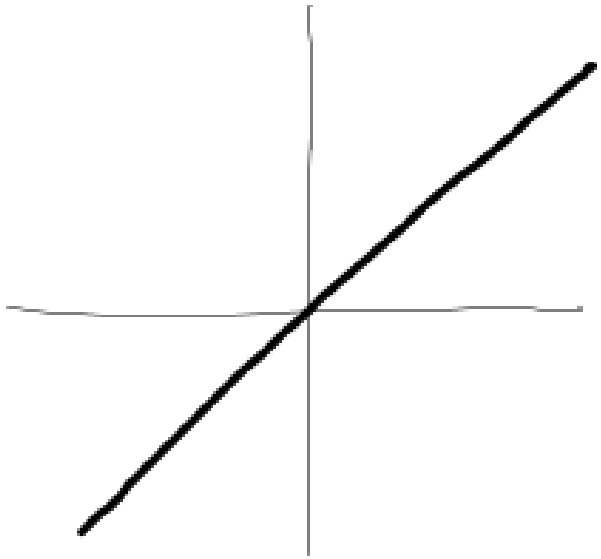
Activations

Convolutional Neural Networks



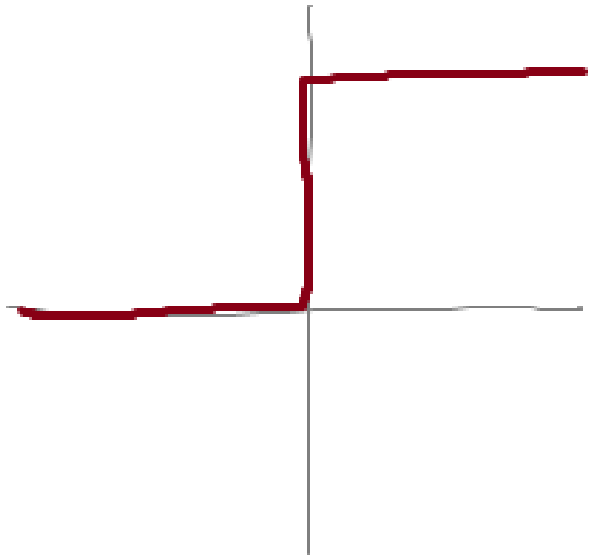
Linear

- Useless



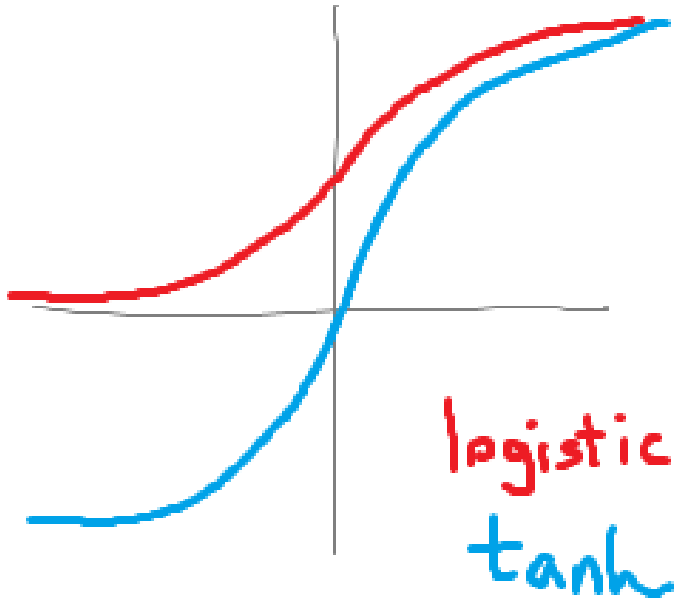
Threshold

- Similar to biological neuron
- No gradient



Sigmoid

- *tanh* preferred
- Gradients can vanish



ReLU

Neural Networks

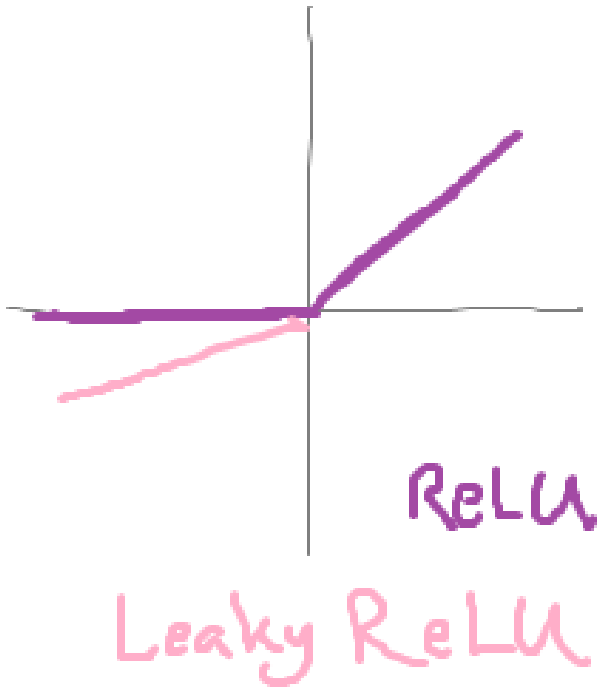
Declan Groves

What is a Neural Network?

Activations

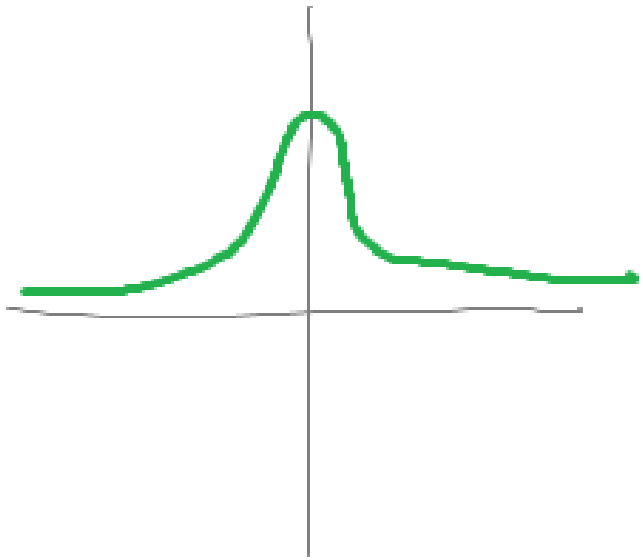
Convolutional Neural Networks

- No vanishing gradient
- Cheap to compute
- Can explode and die
- Popular with CNNs



Radial basis functions

- Gaussian + others
- Train very quickly
- Good at interpolation



Convolutional Neural Networks

Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks

- Image recognition killer

Conceptual structure

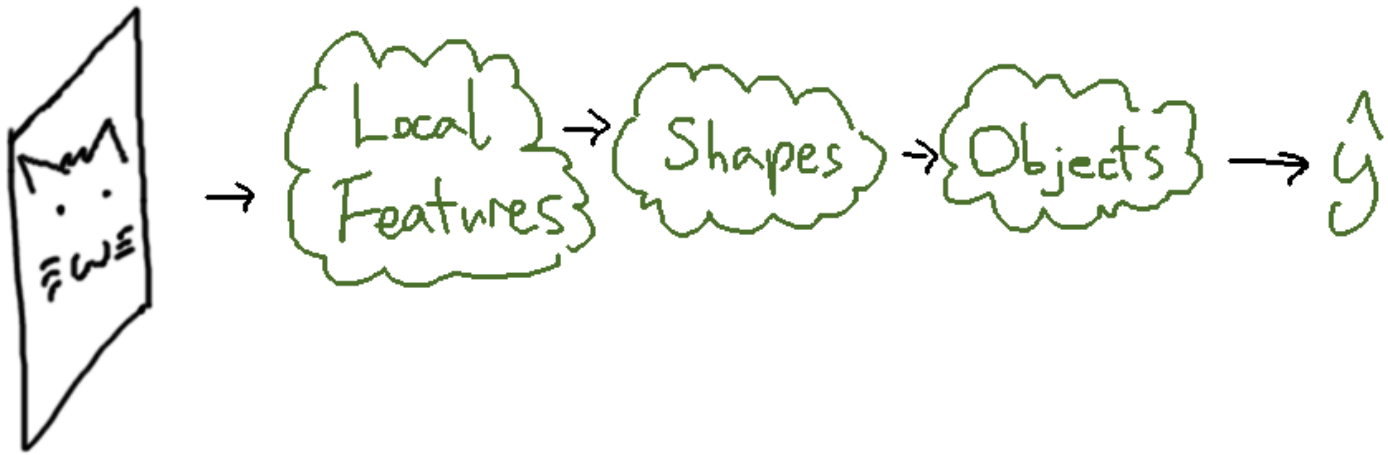
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



MLP $\rightarrow$ too many weights!

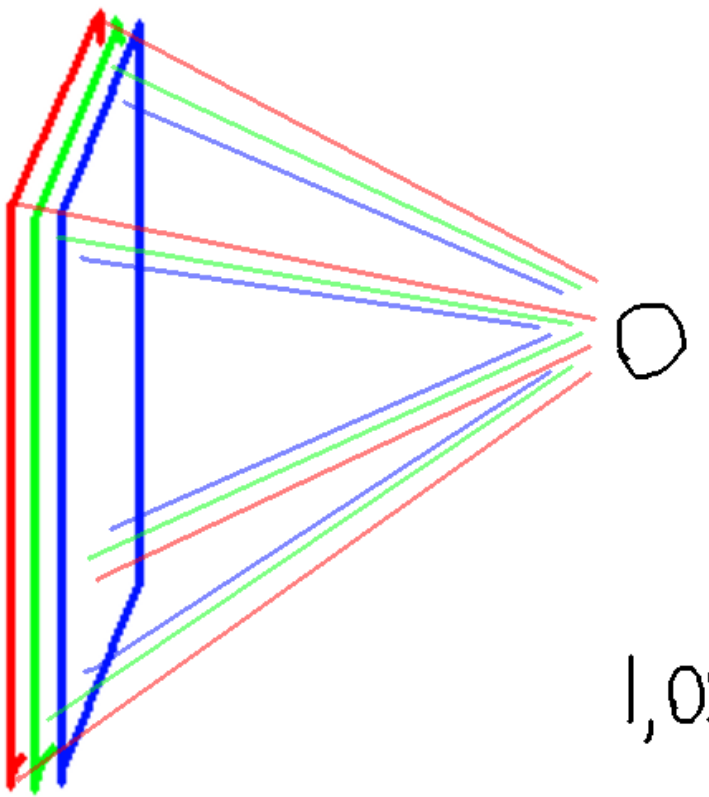
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks

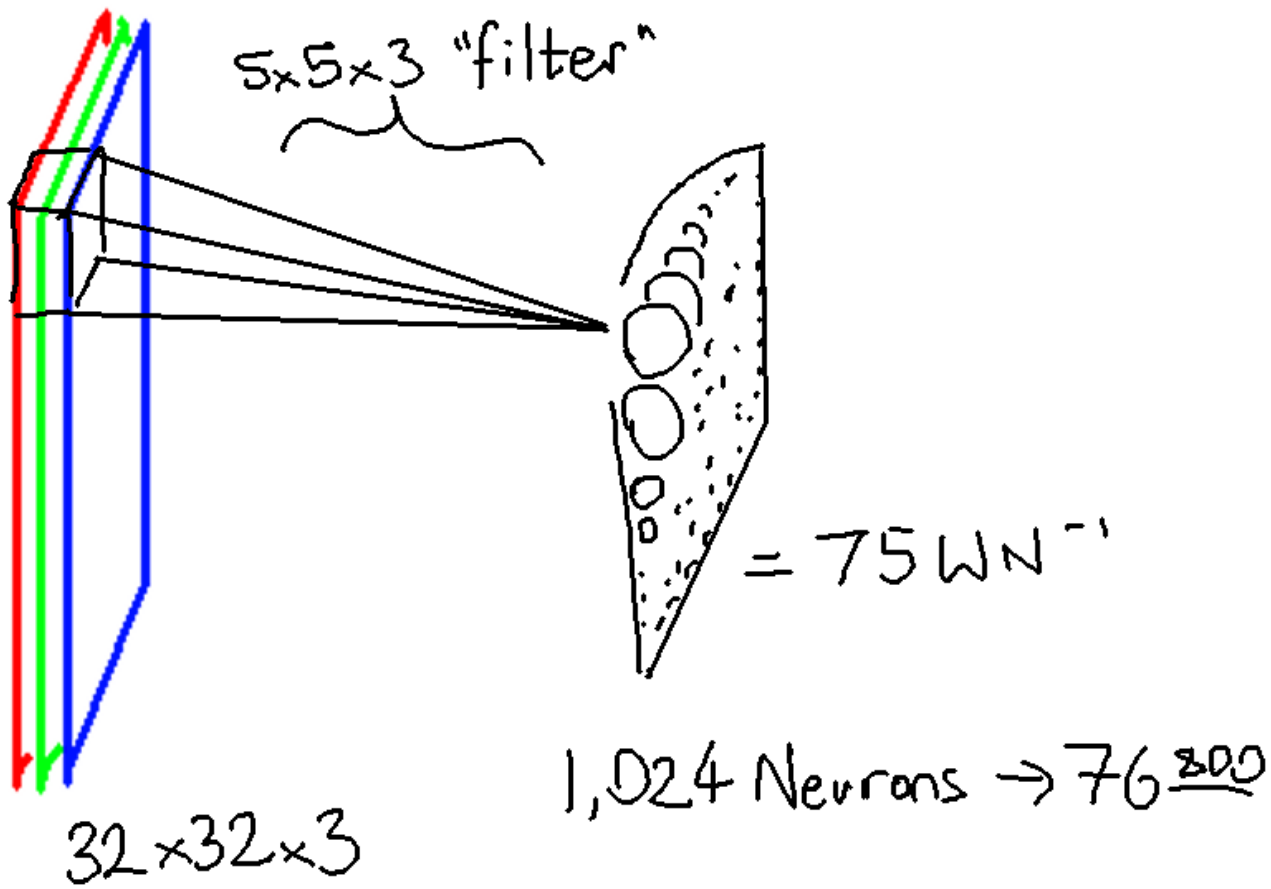


$= 3,072$ weights N^{-1}

1,024 Neurons

$\rightarrow$ >3M weights

Local connectivity



Local connectivity + convolution

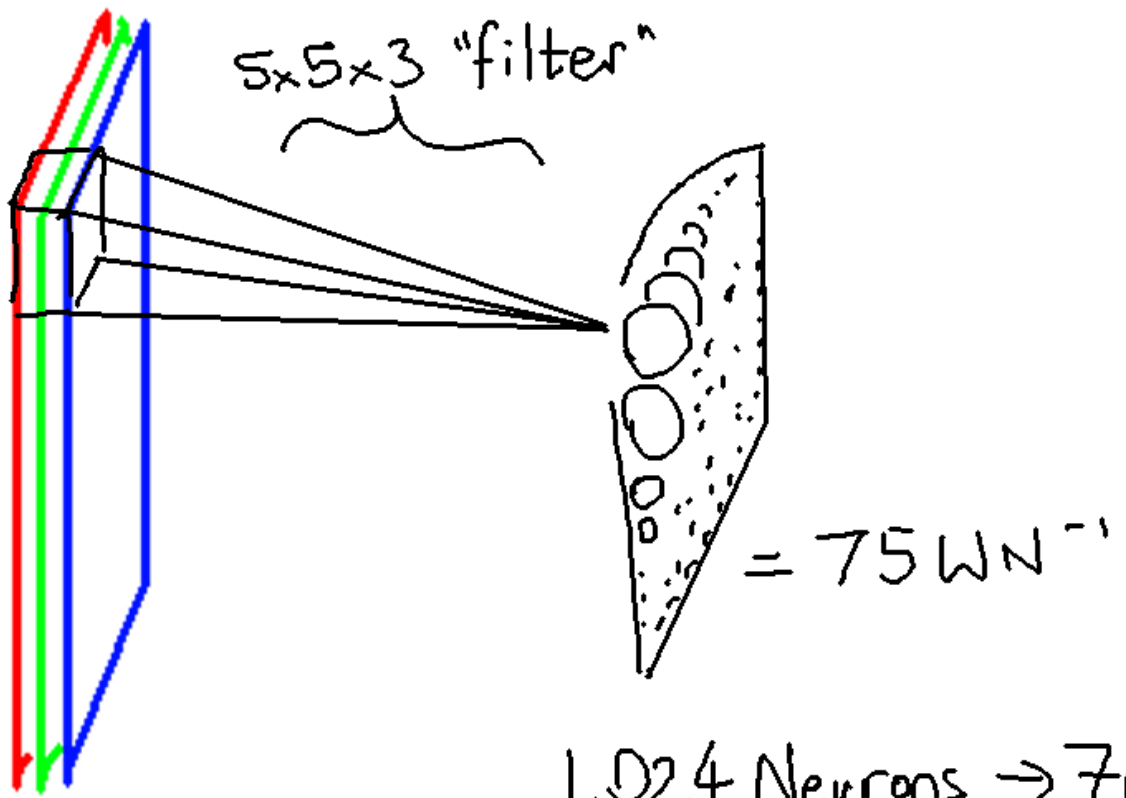
Neural Networks

Declan Groves

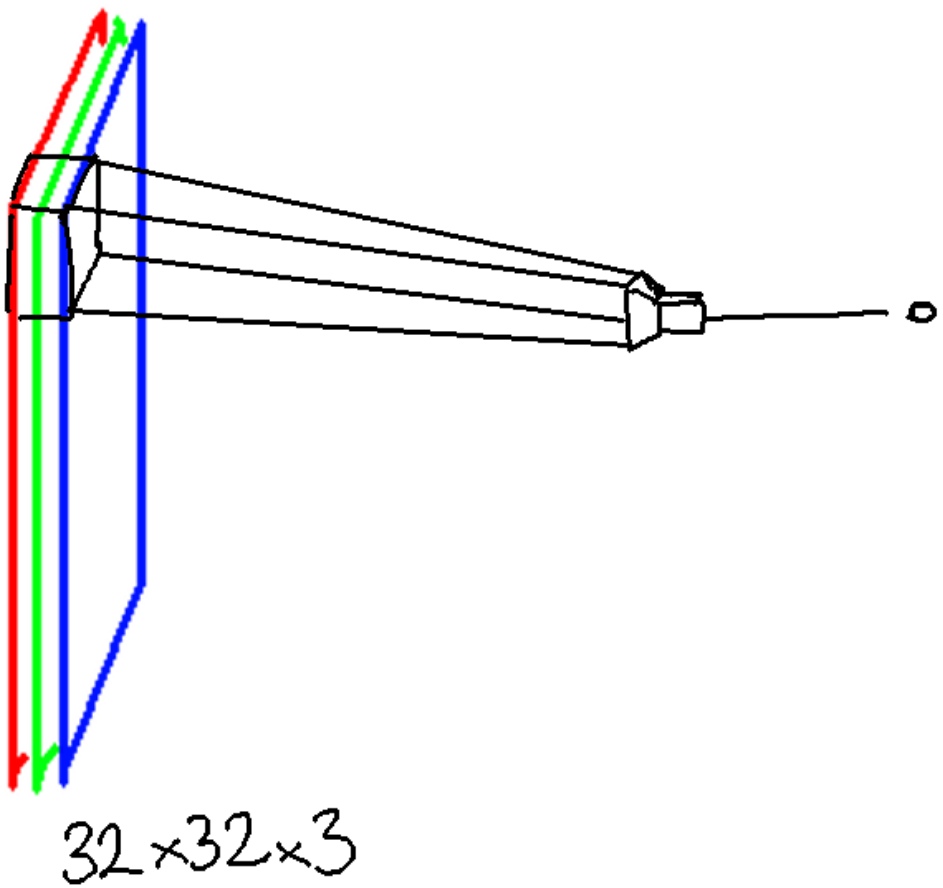
What is a Neural Network?

Activations

Convolutional Neural Networks



Local connectivity + convolution



Local connectivity + convolution

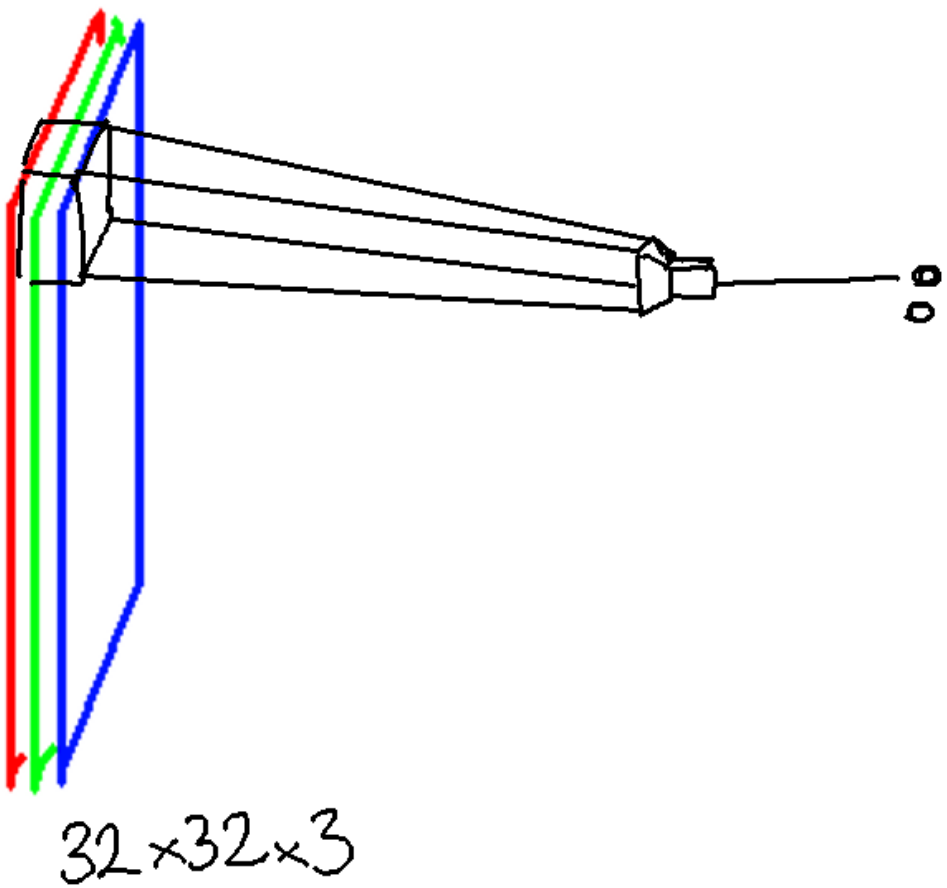
Neural Networks

Declan Groves

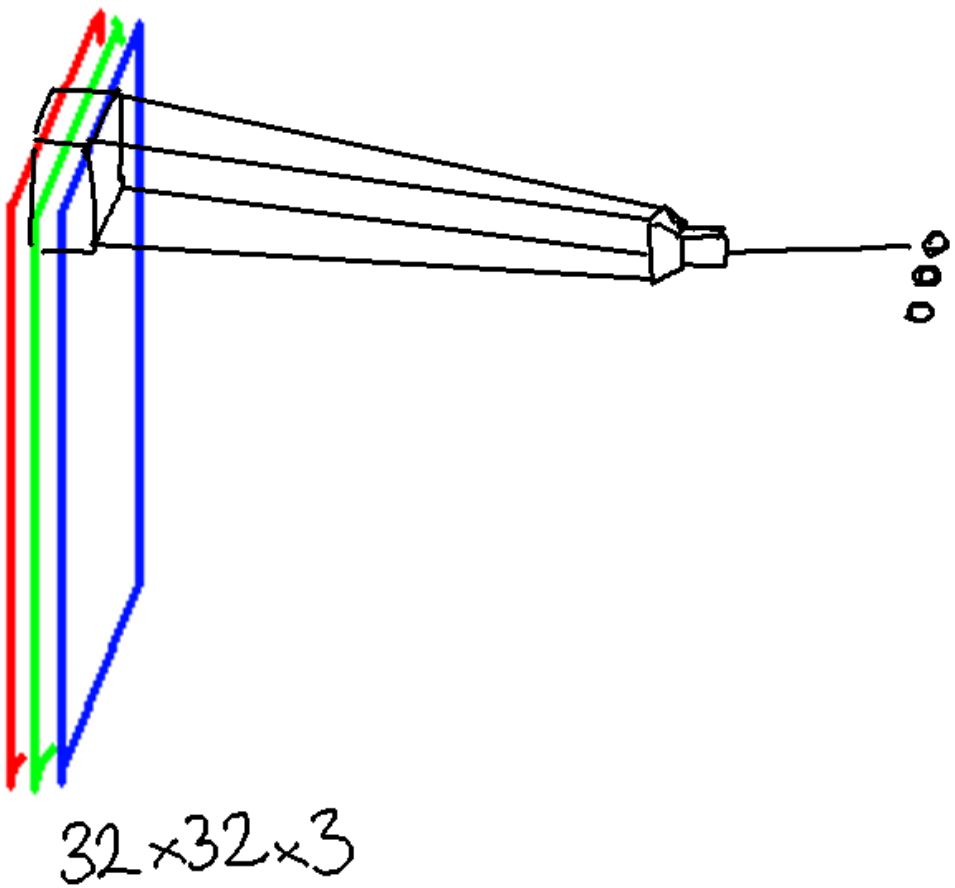
What is a Neural Network?

Activations

Convolutional Neural Networks



Local connectivity + convolution



Local connectivity + convolution

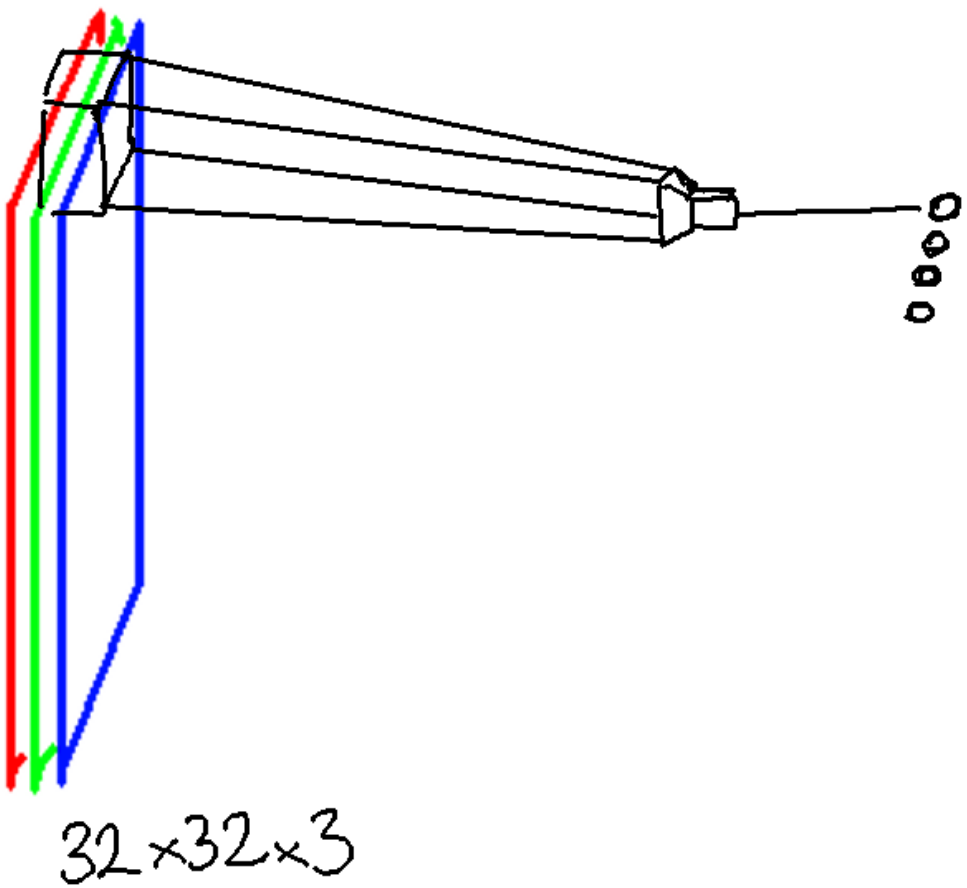
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



Local connectivity + convolution

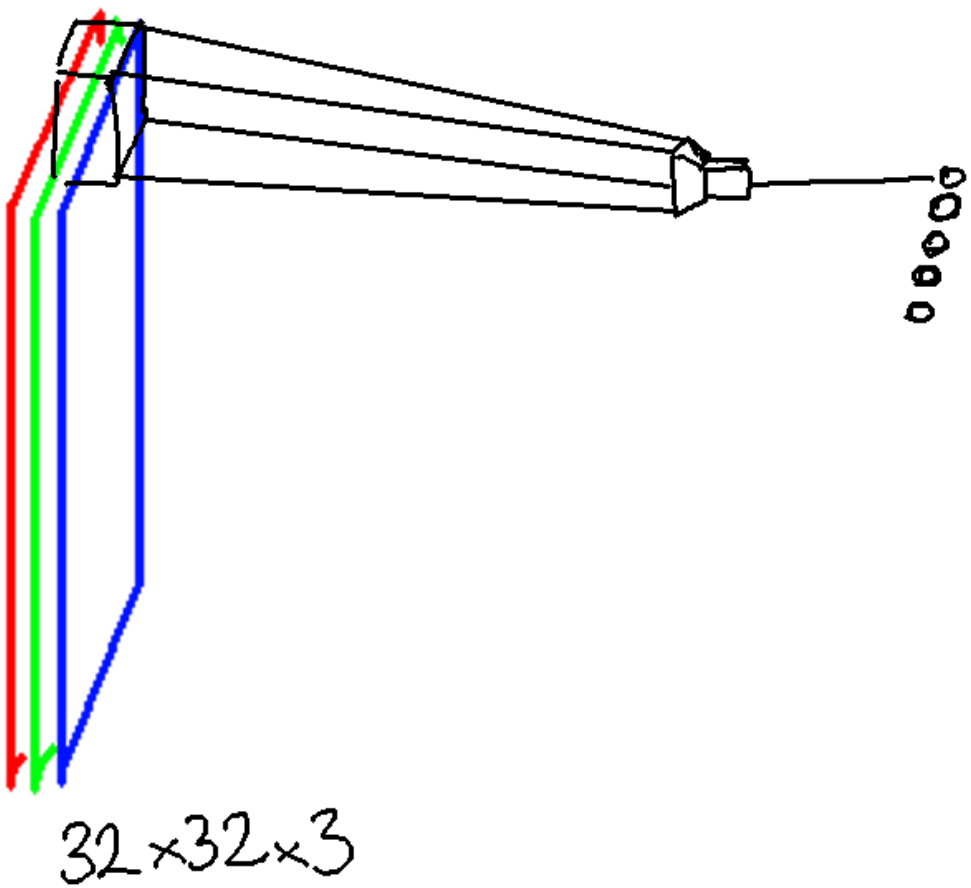
Neural Networks

Declan Groves

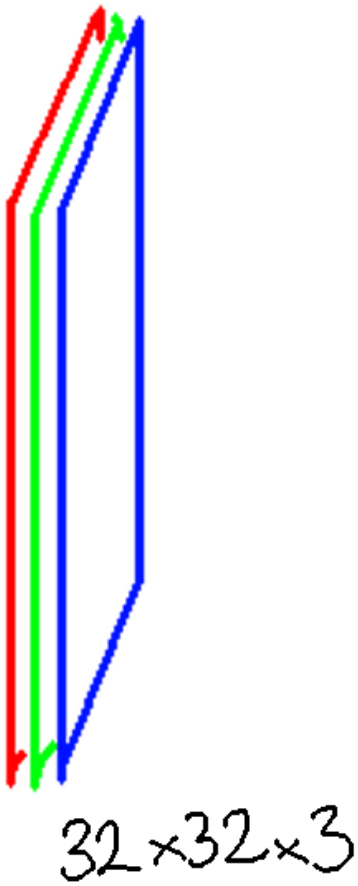
What is a Neural Network?

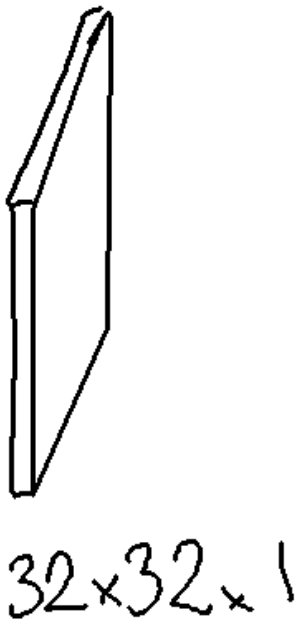
Activations

Convolutional Neural Networks

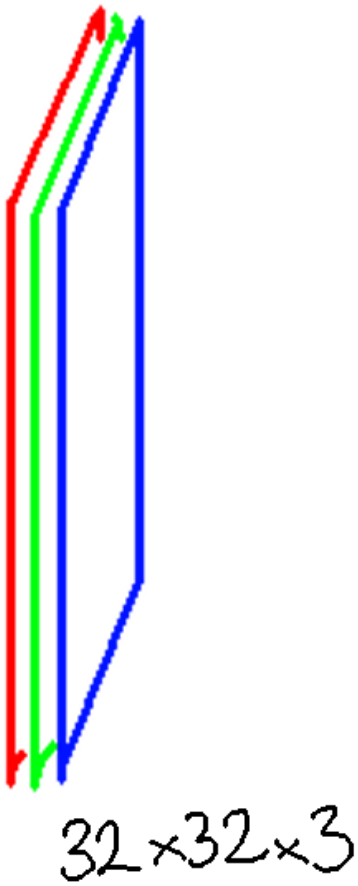


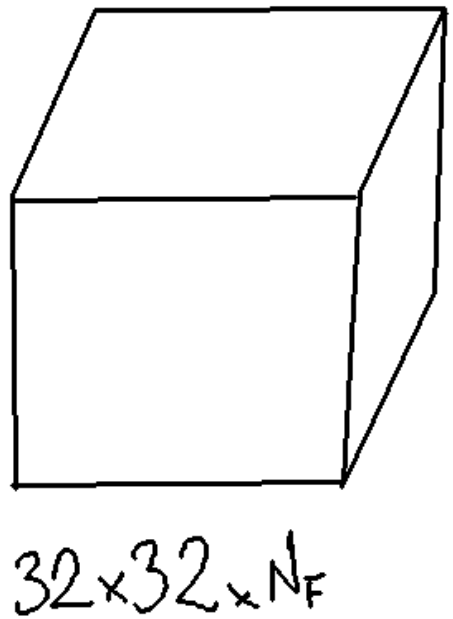
One filter has a surface



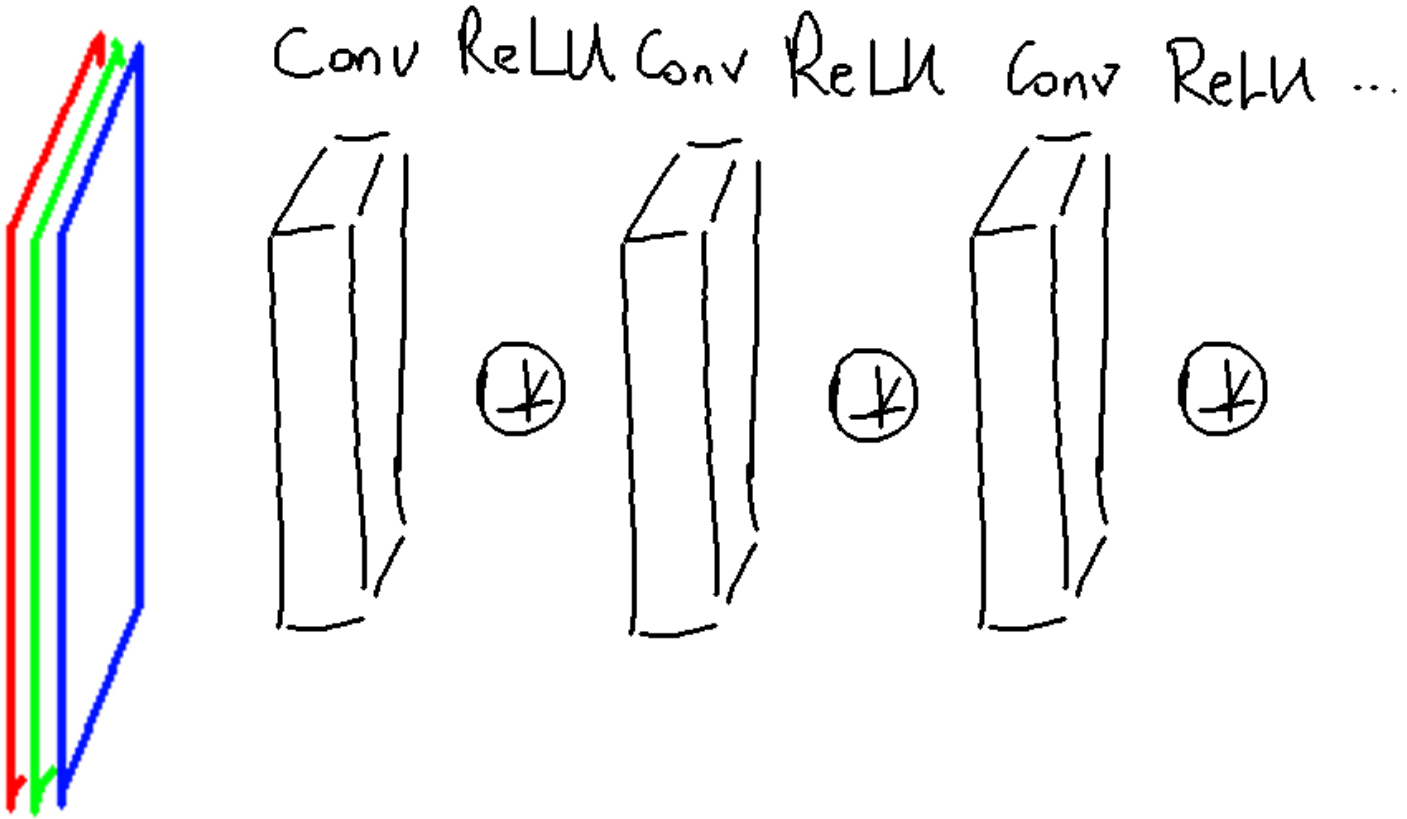


Many filters form a volume





Typical processing structure



AlexNet's filters

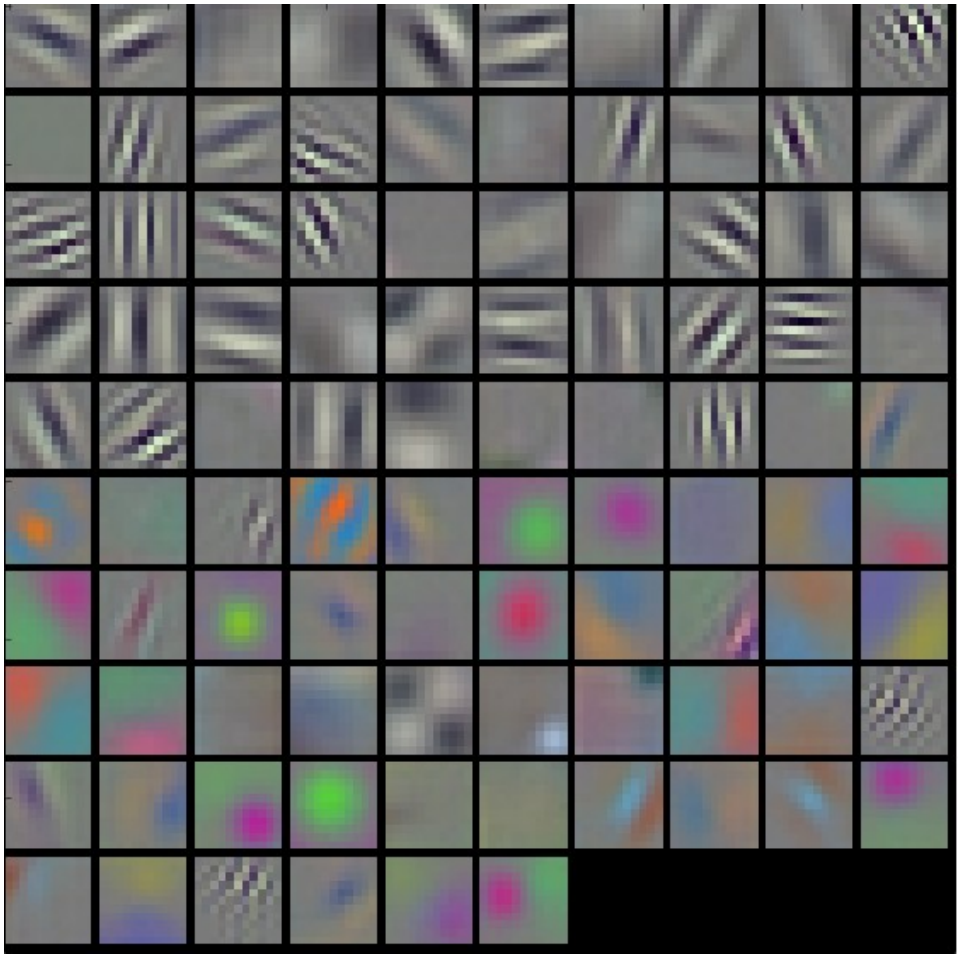
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



Pooling

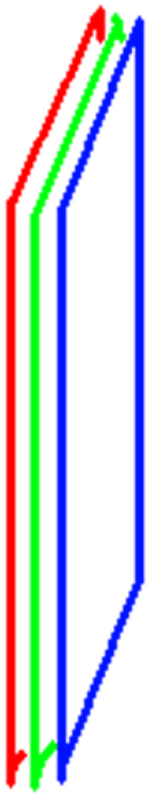
Neural Networks

Declan Groves

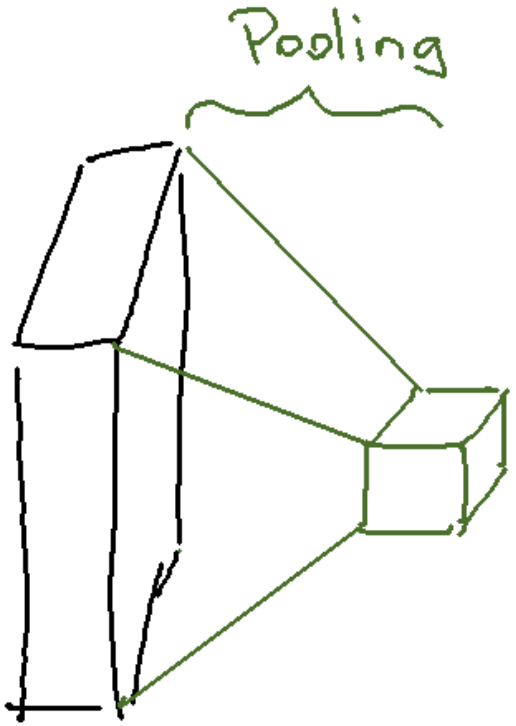
What is a Neural Network?

Activations

Convolutional Neural Networks



$\{Conv, ReLU\}_M$



$\{X, Y, 3\} \rightarrow \{X, Y, K\} \rightarrow \{<X, <Y, K\}$

Pooling

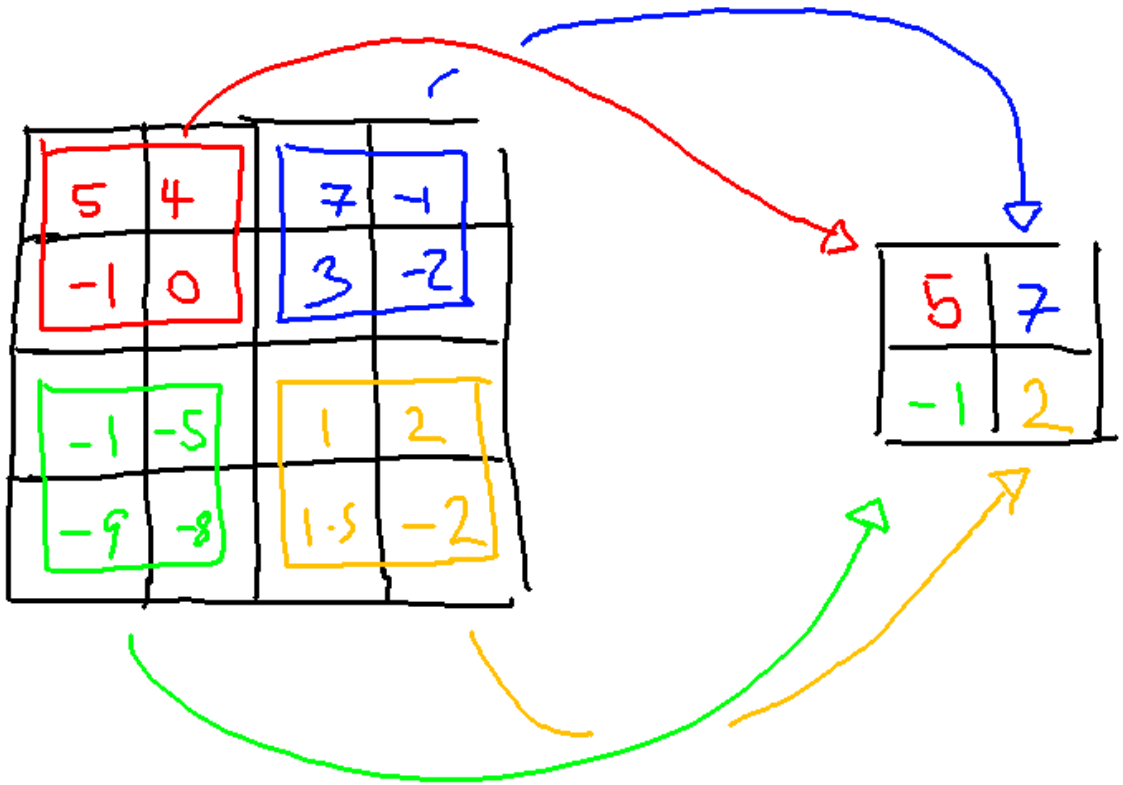
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



Typical structure of a CNN

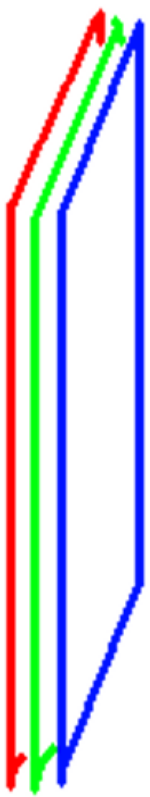
Neural Networks

Declan Groves

What is a Neural Network?

Activations

Convolutional Neural Networks



Standard Convnet

$$\{\{\text{Conv, ReLU}\}_M \text{ Pool}\}_N \quad \{\text{FC}\}_K \rightarrow \hat{y}$$