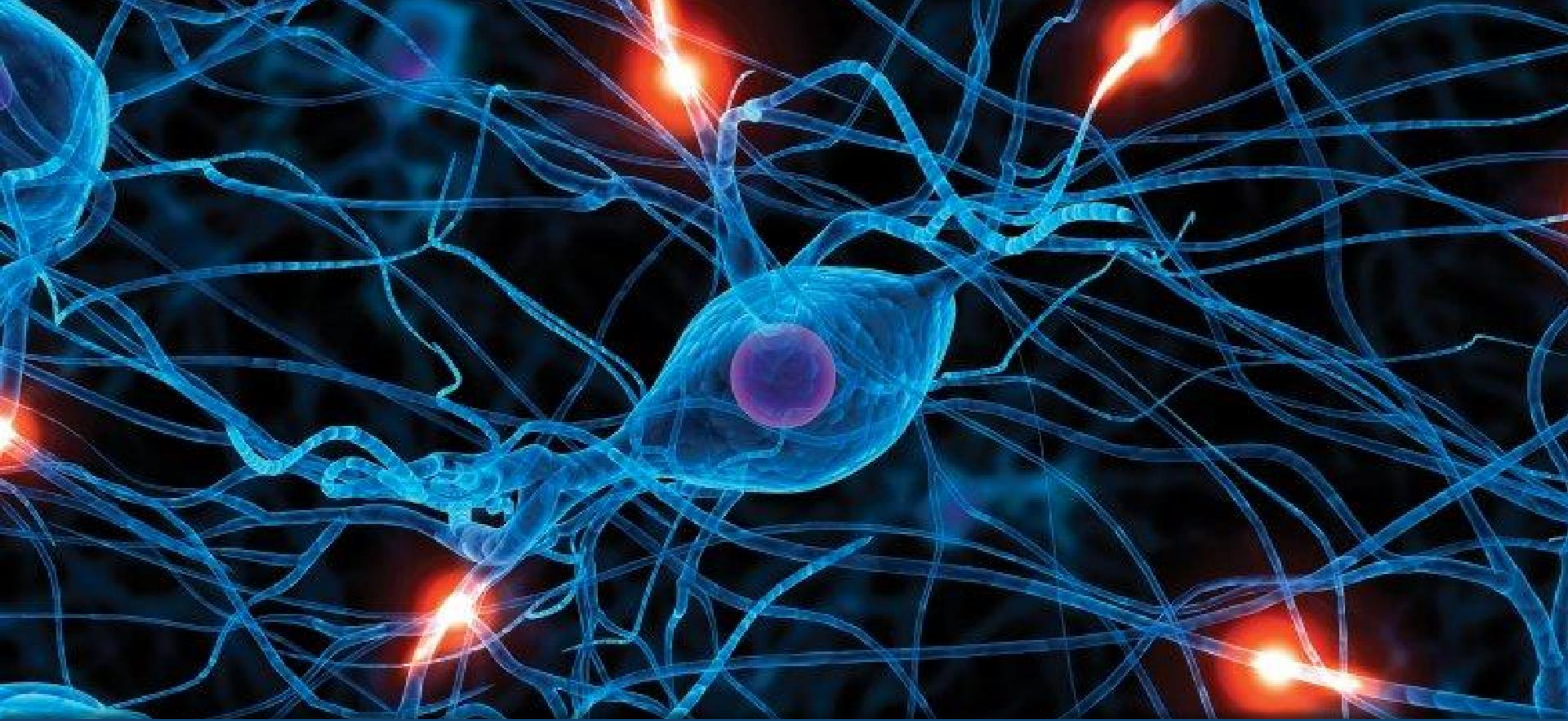


An Introduction to Deep Learning

Labeeb Khan

Special Thanks:

Lukas Masuch

@lukasmasuch

+lukasmasuch

Lead Software Engineer:
Machine Intelligence, SAP

The Big Players

Companies

facebook

Microsoft

YAHOO!

Google

IBM

NVIDIA

Bai du 百度

The Big Players

Startups

vicarious
enlitic
clarifai
SKYMIND
SIGNALSENSE
nervana
nnaisense
cortica™ In Every Image
ersatz labs
Numenta
sentient
OpenAI
MetaMind

DEEPMIND
AlchemyAPI™ An IBM Company
wit.ai
DNNresearch

Acquired

Machine Learning - Basics

Learning Approaches

Supervised Learning: Learning with a labeled training set
Example: email spam detector with training set of already labeled emails

Unsupervised Learning: Discovering patterns in unlabeled data
Example: cluster similar documents based on the text content

Reinforcement Learning: learning based on feedback or reward
Example: learn to play chess by winning or losing

What is Deep Learning?

Part of the machine learning field of learning representations of data. Exceptional effective at learning patterns.

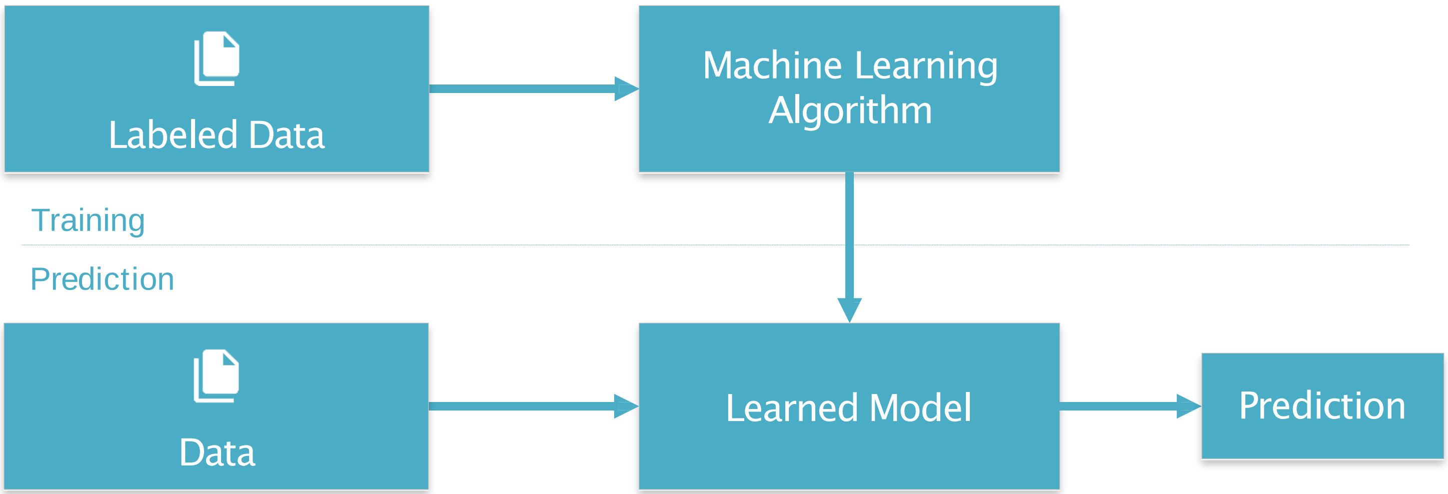
Utilizes learning algorithms that derive meaning out of data by using a hierarchy of multiple layers that mimic the neural networks of our brain.

If you provide the system tons of information, it begins to understand it and respond in useful ways.

Machine Learning - Basics

Introduction

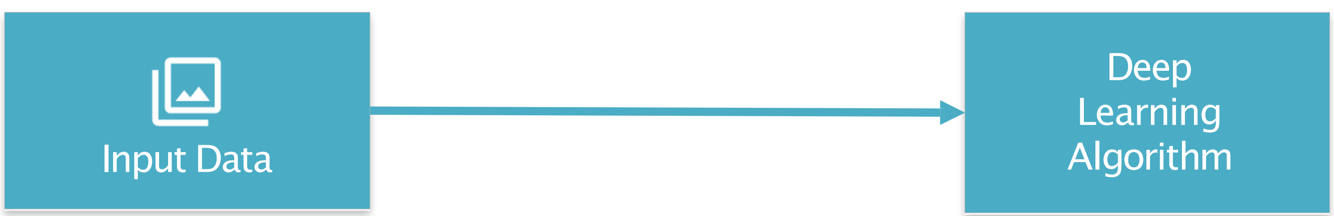
Machine Learning is a type of Artificial Intelligence that provides computers with the ability to learn without being explicitly programmed.



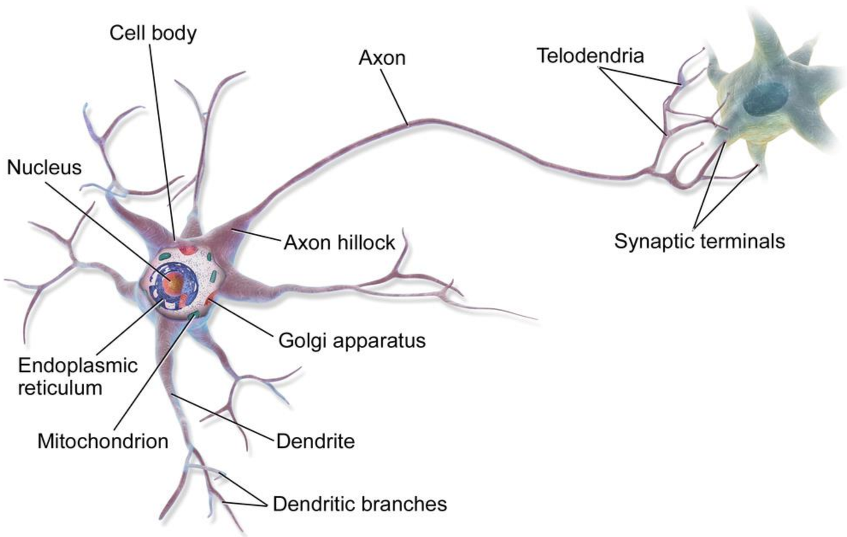
Provides various techniques that can learn from and make predictions on data

Deep Learning - Basics

No more feature engineering



Inspired by the Brain



- Humans have ~100 billion neurons
- Each neuron contains a cell body, dendrites, axon connected to ~10,000 other neurons

Our neurons pass signals to each other via 1000 trillion synaptic connections, which is approximately a 1 trillion bit per second processor (125,000 MB/s).

1

One learning algorithm hypothesis: all significant mental algorithms are learned except for the learning and reward machinery itself.

Our Natural System

What is it good at?

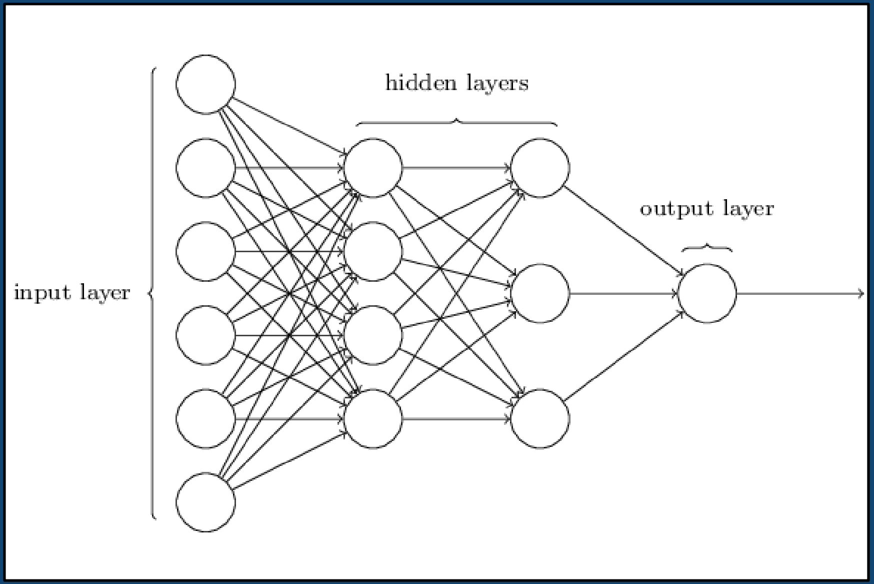
Good at:

- Vision
- Hearing
- Speech Recognition & Speaking
- Driving
- Playing Games
- Natural Language Understanding

Not good at:

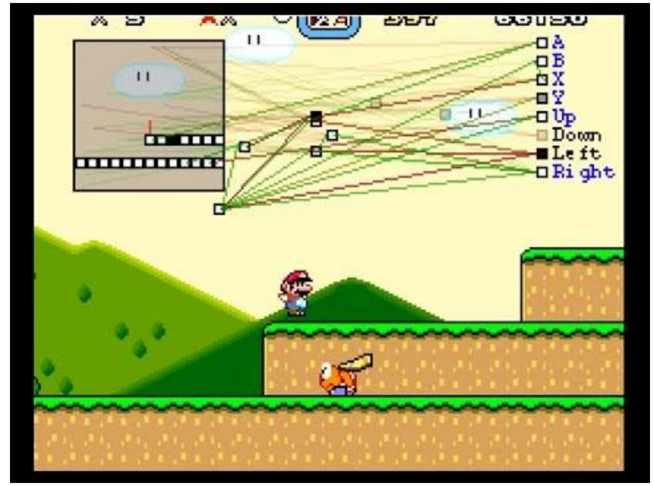
- Multiplying 2 numbers
- Memorizing a phone number

Feedforward Neural Networks
Architecture

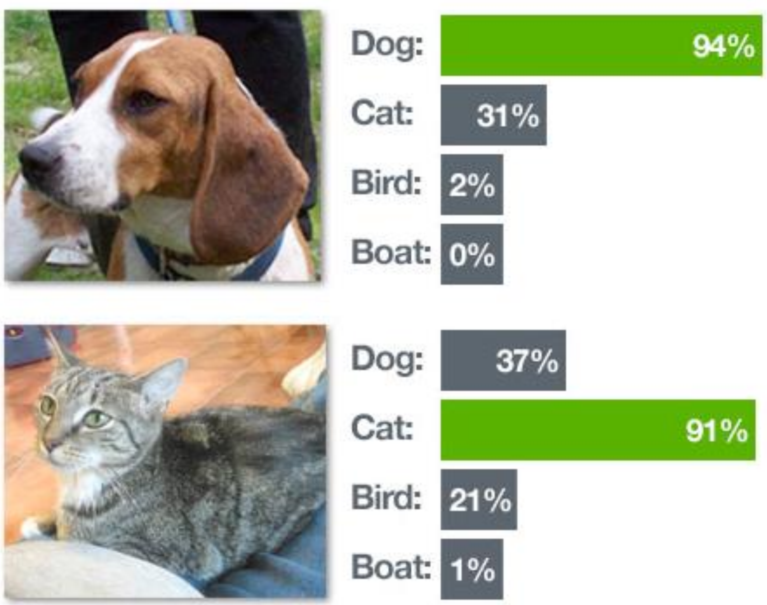


Feedforward Networks - Applications

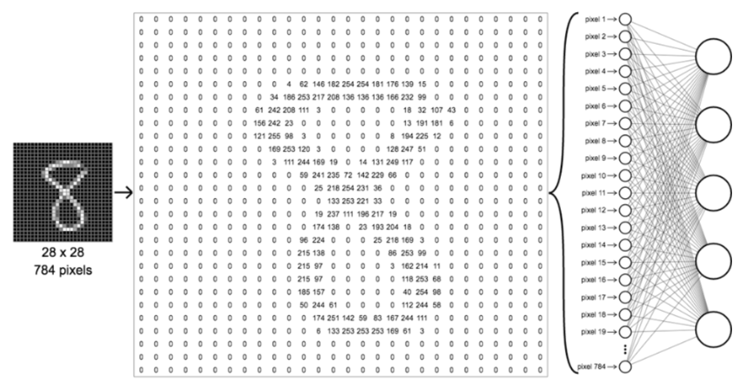
Game AI
Mario Neural Network



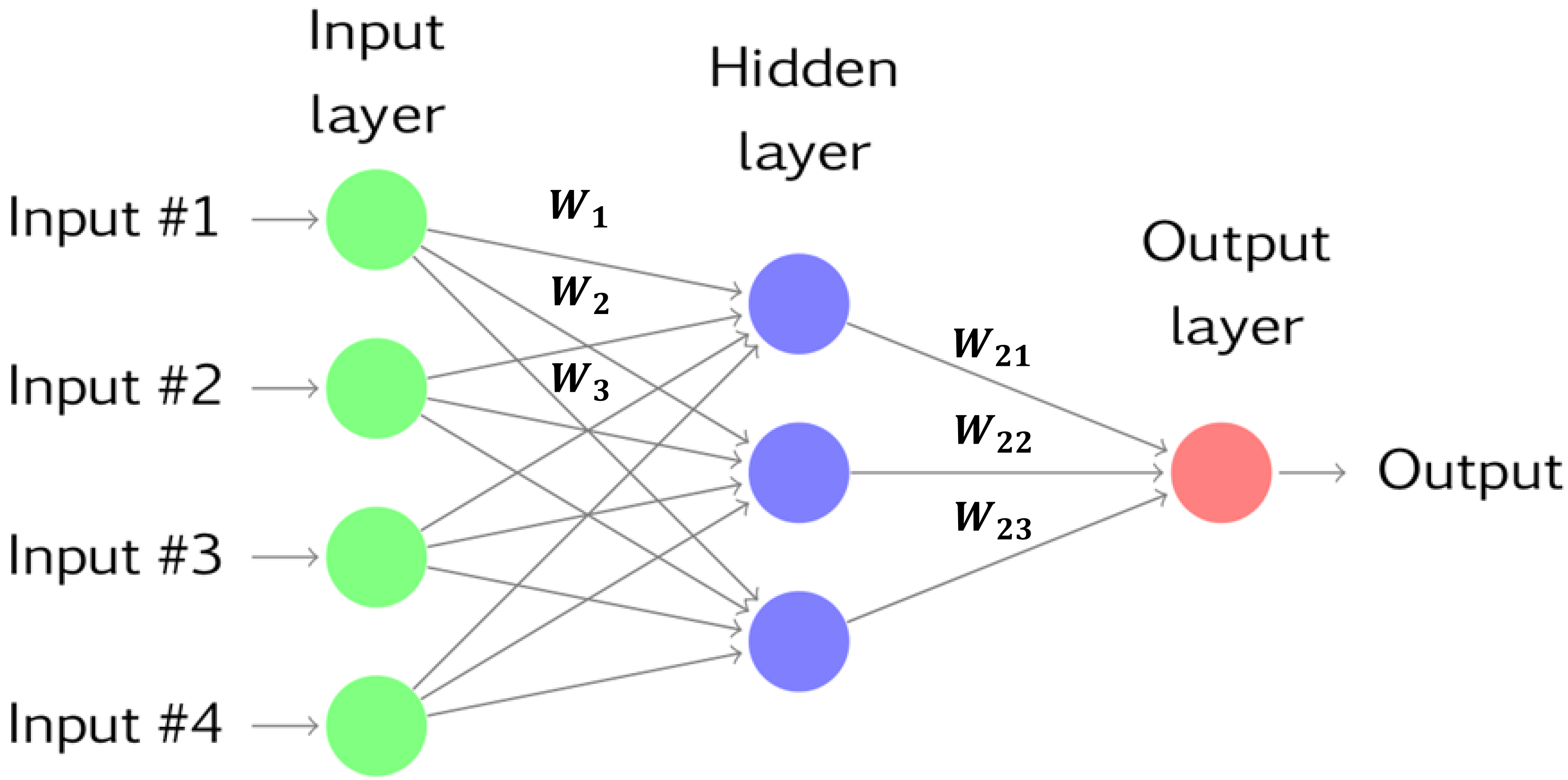
Animal Recognition



Digit Recognition



Network Architecture - Introduction



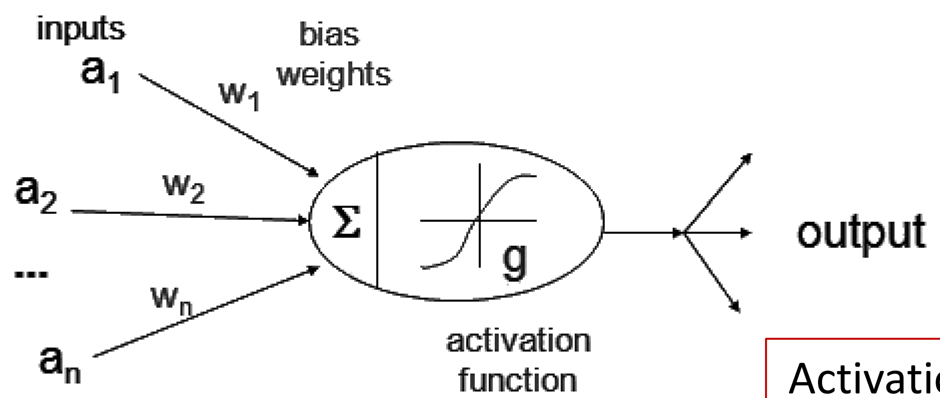
1. Inputs are mapped to a hidden layer
2. Weights are initialized randomly
3. Output / Prediction is made

Network Architecture – Sigmoid Activation Function

1. Each neuron utilizes an activation function

2. Calculates a weighted sum of inputs

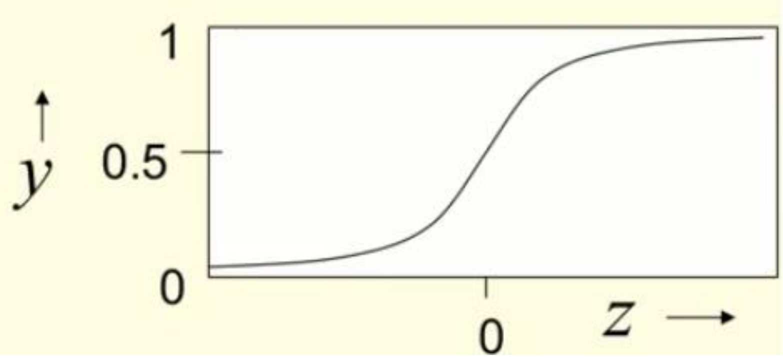
3. Decides weather to “fire” or not



Activation Function

$$z = b + \sum_i x_i w_i \qquad y = \frac{1}{1 + e^{-z}}$$

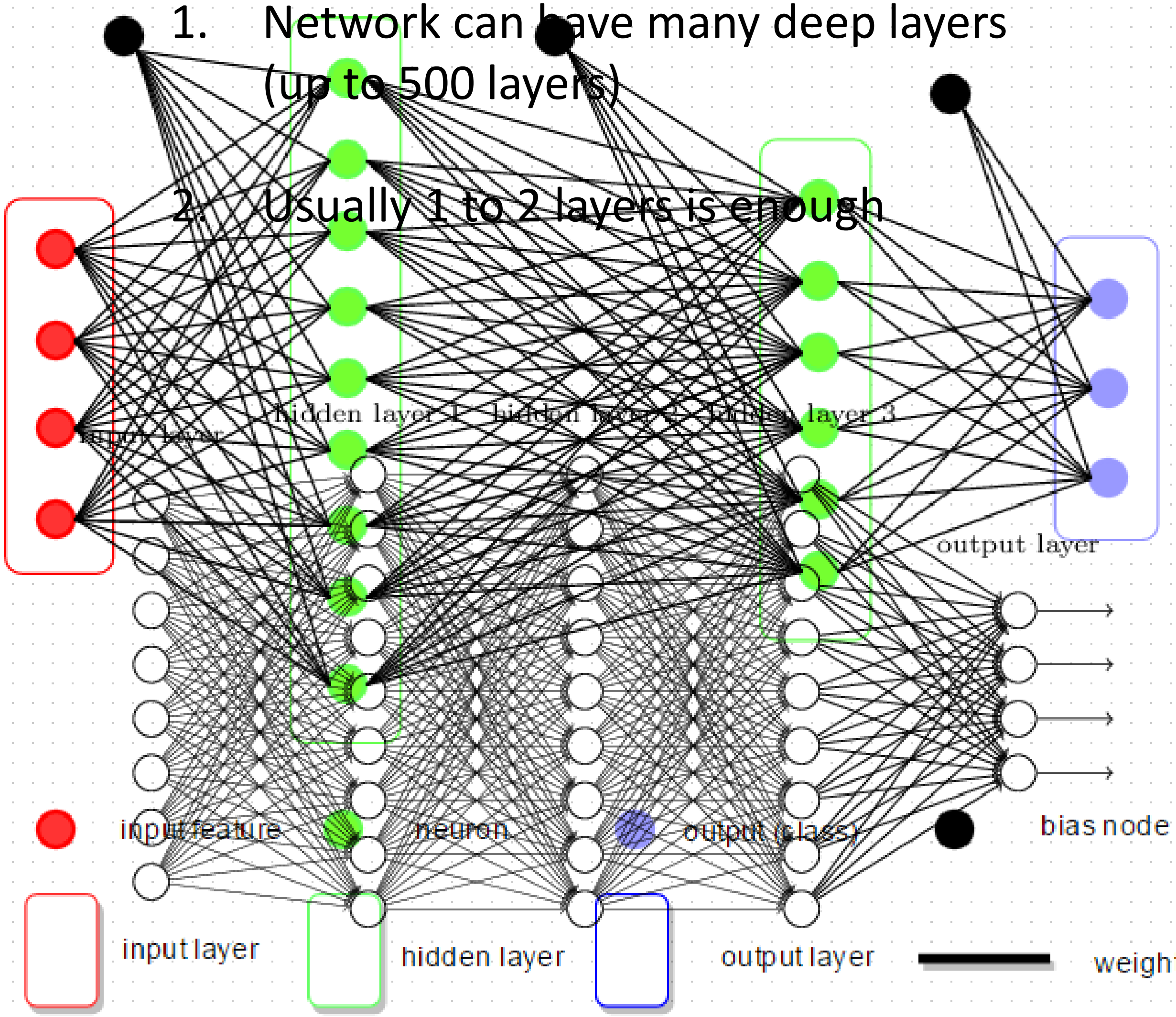
Weighted sum of inputs



Network Architecture – Many Layers

A 3-layers fully connected neural network (DNN)

1. Network can have many deep layers (up to 500 layers)
2. Usually 1 to 2 layers is enough



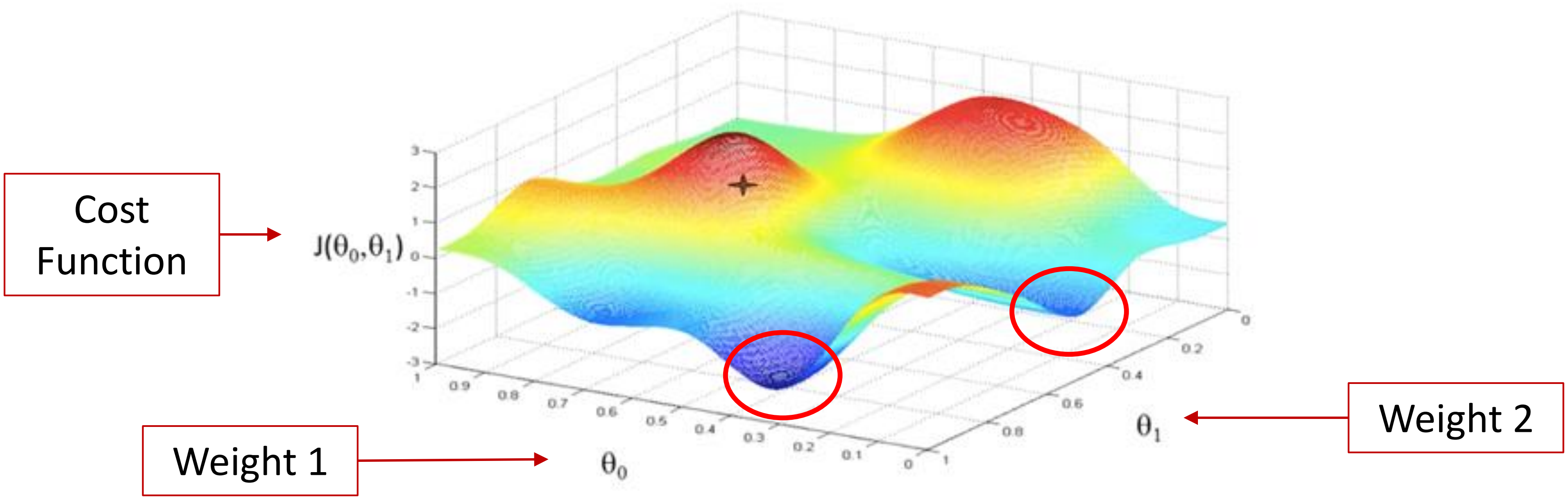
Network Architecture – Optimizing the Cost Function

1. Each network prediction on the **training data** contains an associated **error**, or “**cost**”

2. Plotting each error with an associated weight gives us a Cost Function
(this is abstract, not seen by the network)

For the network to “**learn**” the problem:

We must find a set of weights that globally minimize the cost function

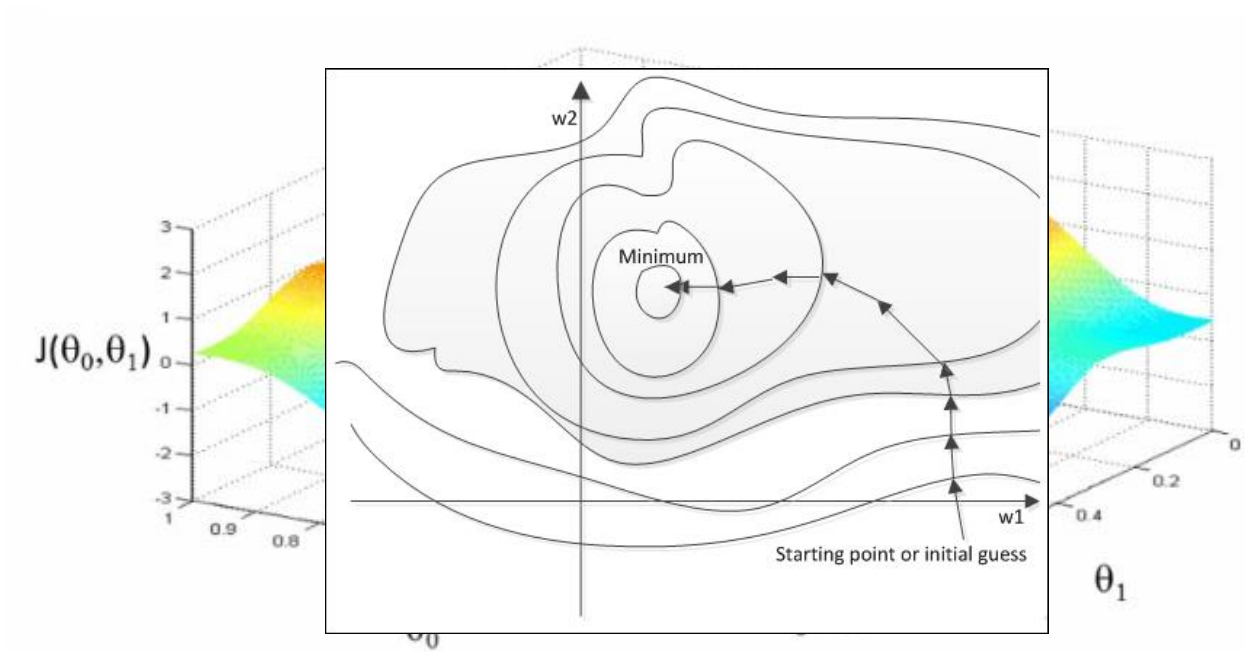


Network Architecture – Backpropagation and Gradient Descent

Backpropagation: Backward propagation of errors using **Gradient Descent**

Gradient Descent: Calculates the change in error with respect to each network weight

Learning Rate: Speed and quality at which the network learns



Old Weight

Gradient

Repeat until convergence {

$$\theta_j \leftarrow \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

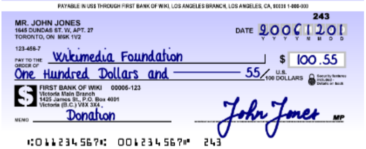
}

New Weight

Learning Rate

Feedforward Networks – Applications

Cheque Recognition



Medical Diagnosis

House AI

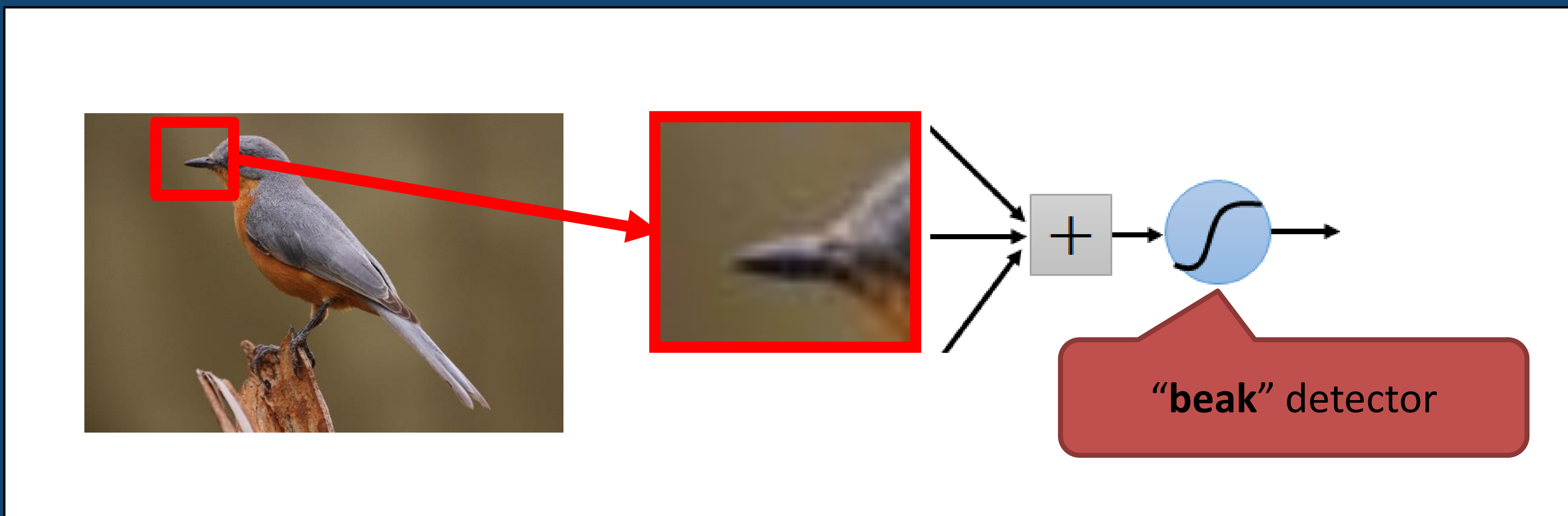
Feedforward Architecture – Problems with Image Processing

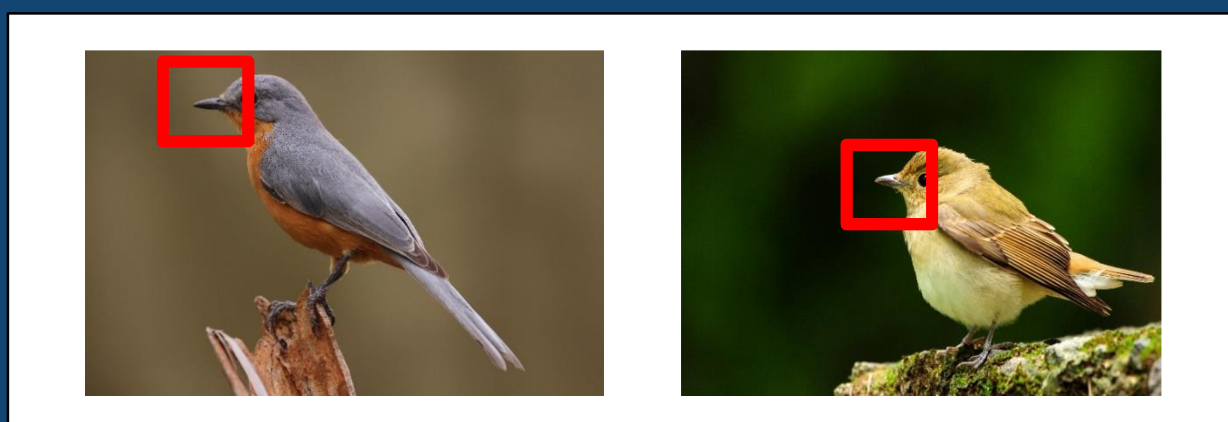
Image Processing & Vision:

- Some patterns appear in different places, these cannot be compressed with a feedforward network!
- Some patterns are much smaller than the whole image
- Feedforward networks map pixels to a hidden layer, images can be of different sizes!

Convoluted Neural Networks (CNN)

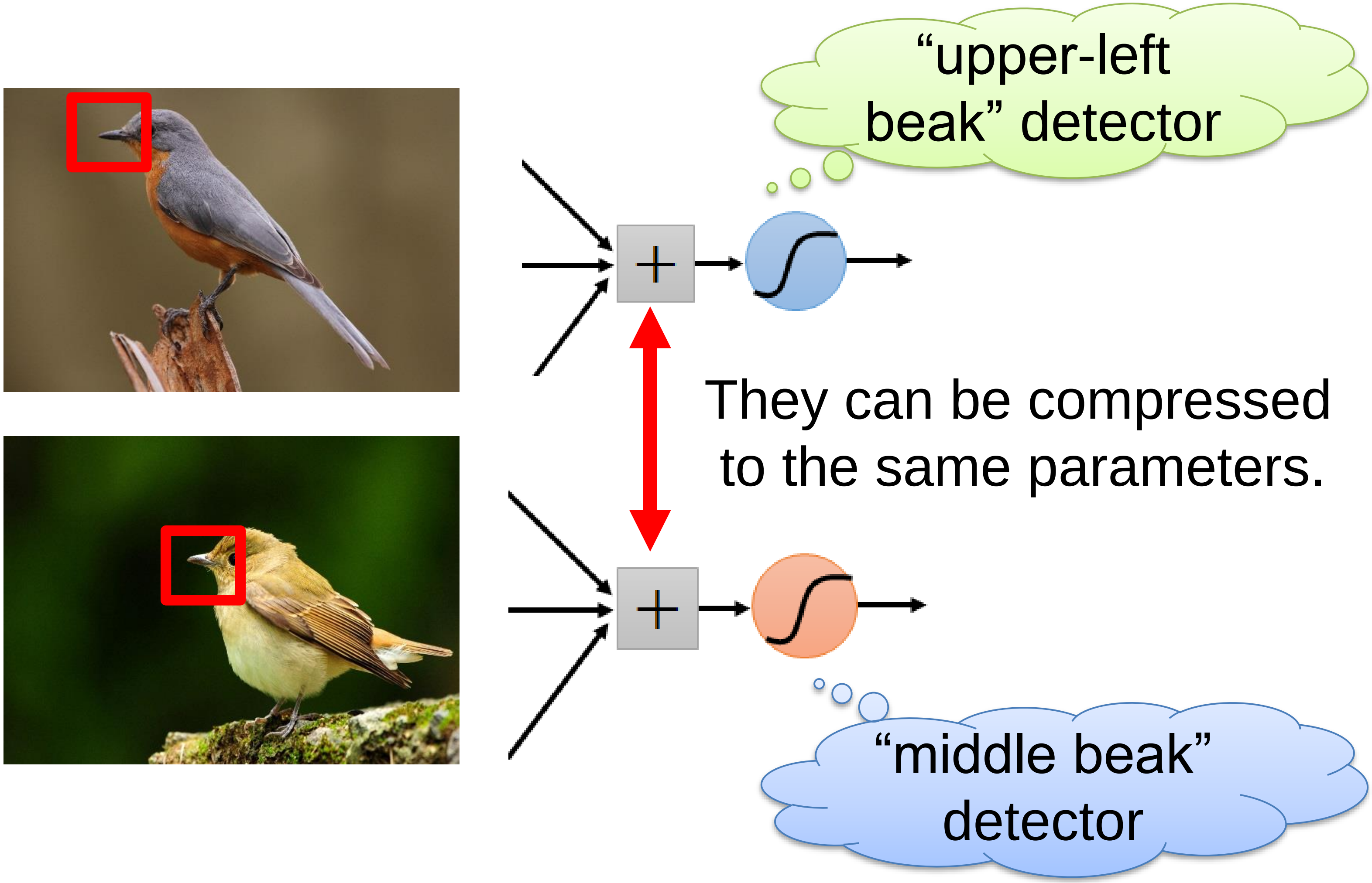
Architecture





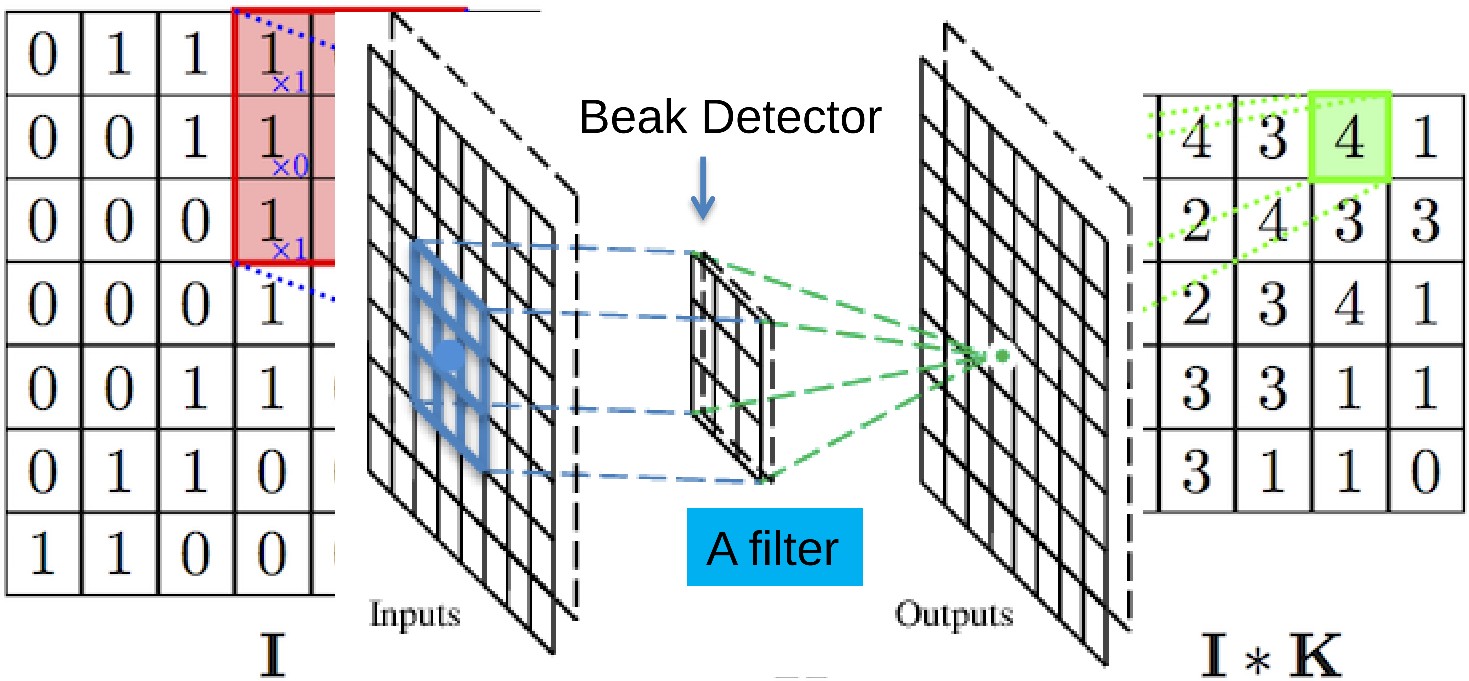
Convoluted Neural Networks

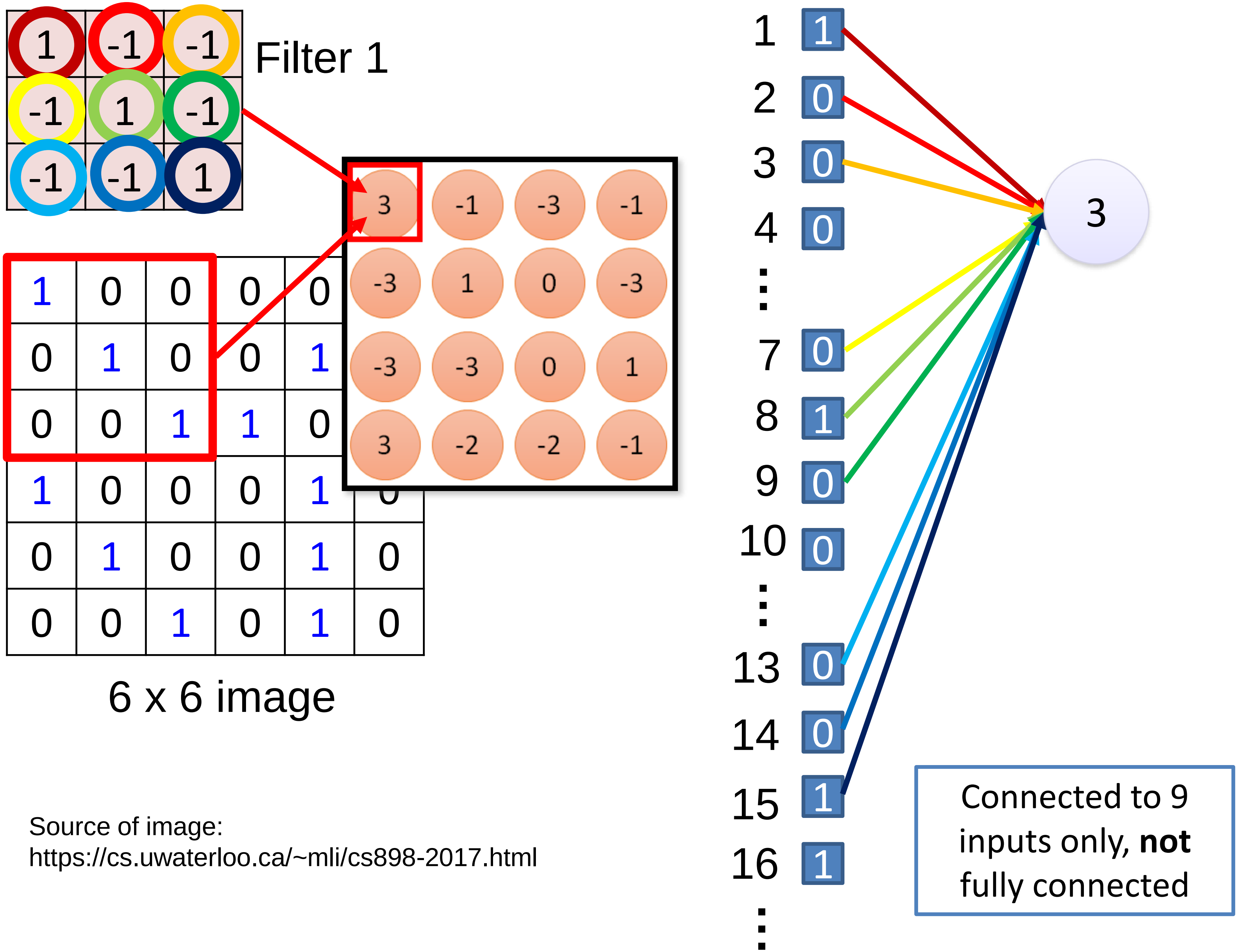
- Some patterns appear in different places, these can be compressed!



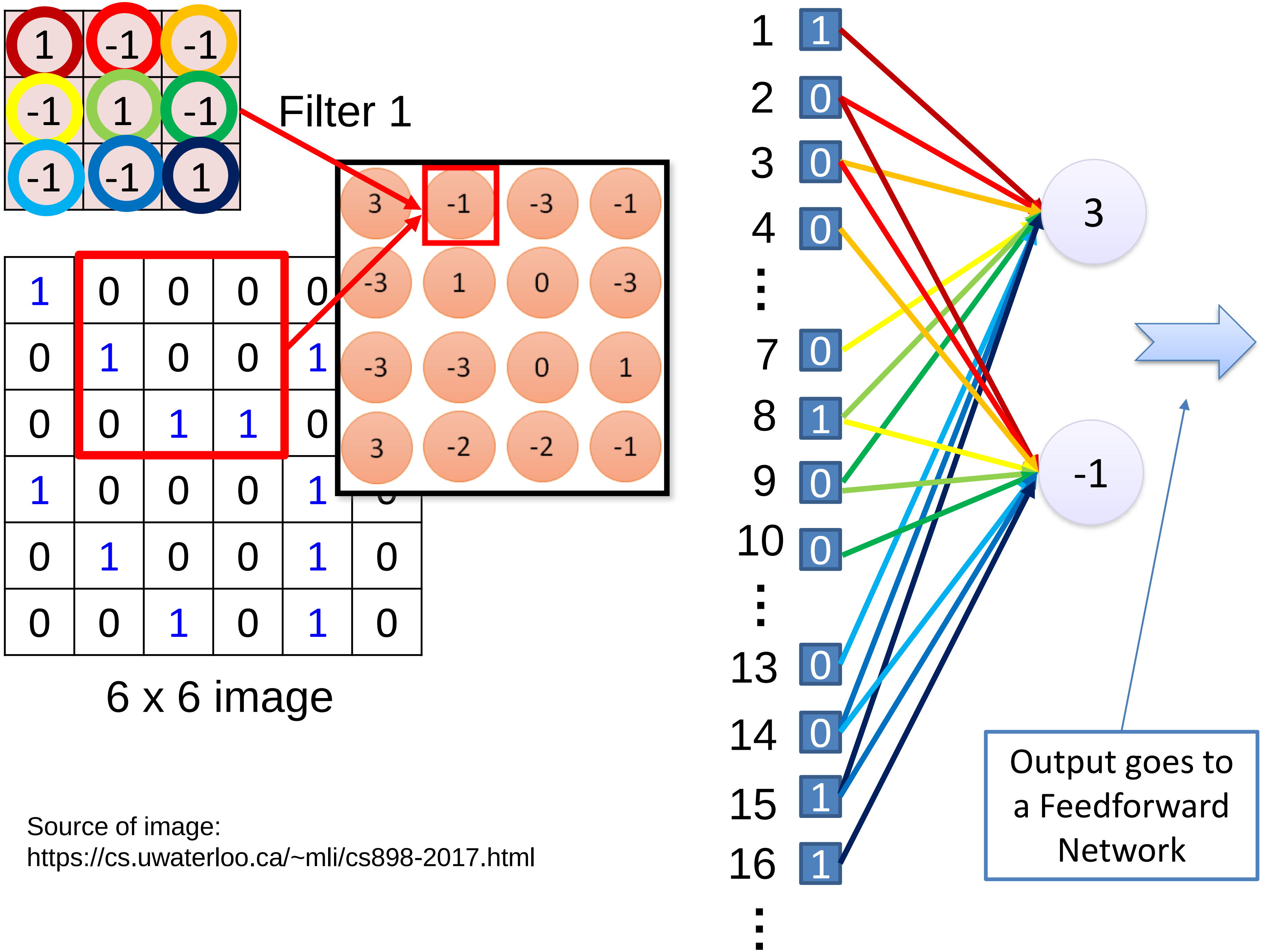
CNN Network Architecture – Convolutional Layer

- A neural network with convolutional layers. The convolutional layers are generated by filters that do convolutional operations



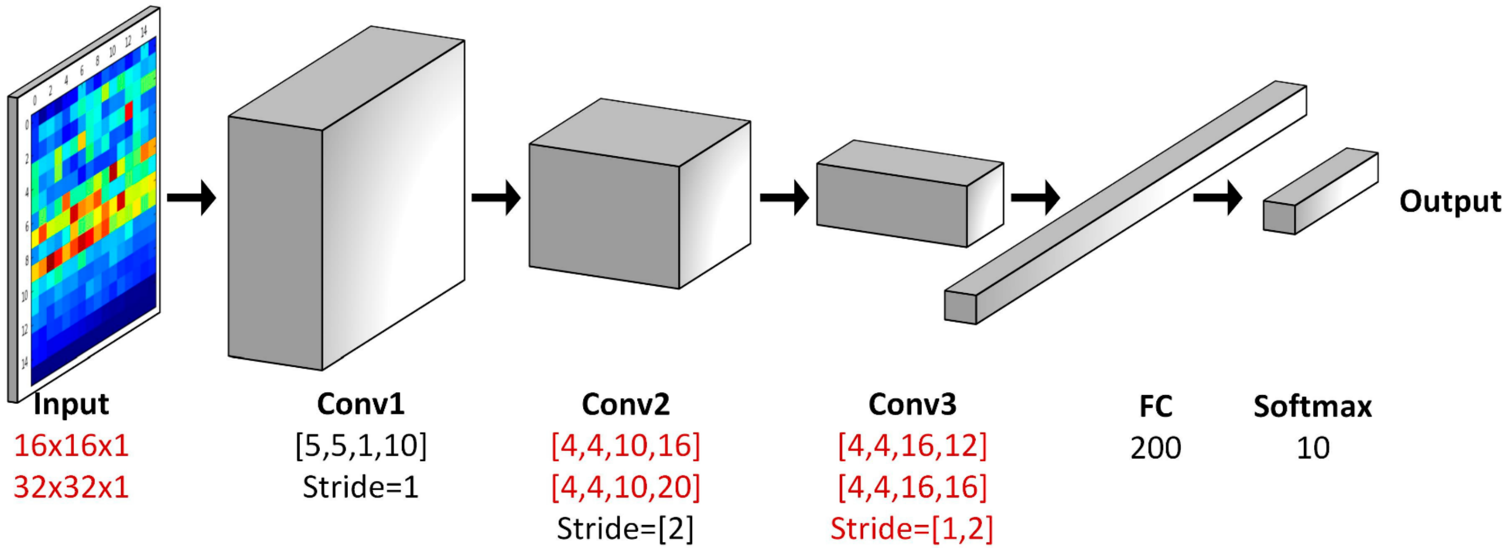


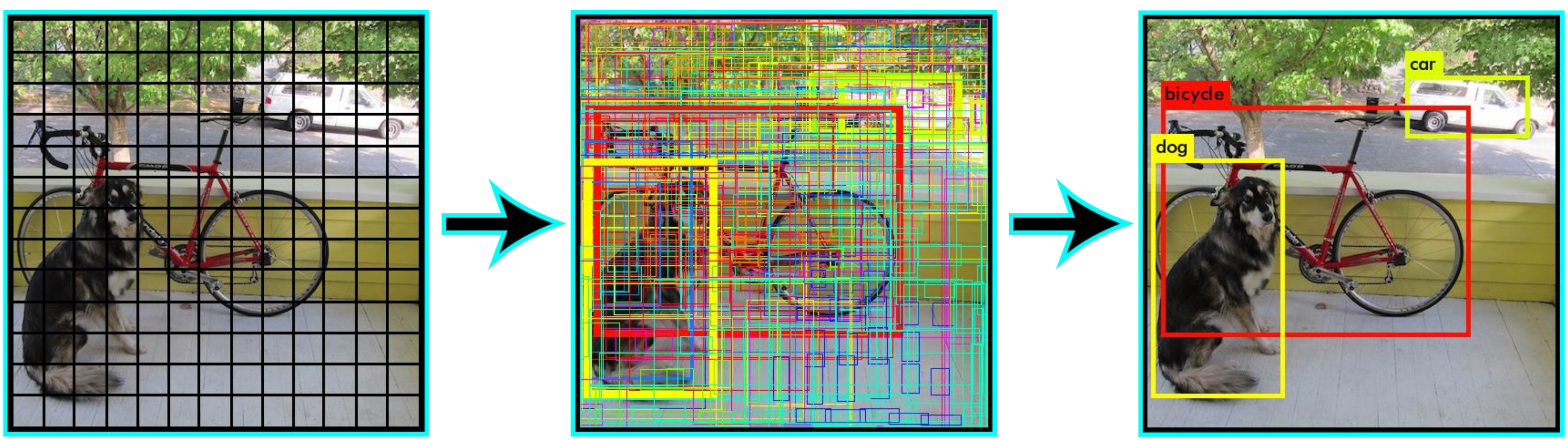
Source of image:
https://cs.uwaterloo.ca/~mli/cs898-2017.html



Source of image:
https://cs.uwaterloo.ca/~mli/cs898-2017.html

CNN Network Architecture – Process

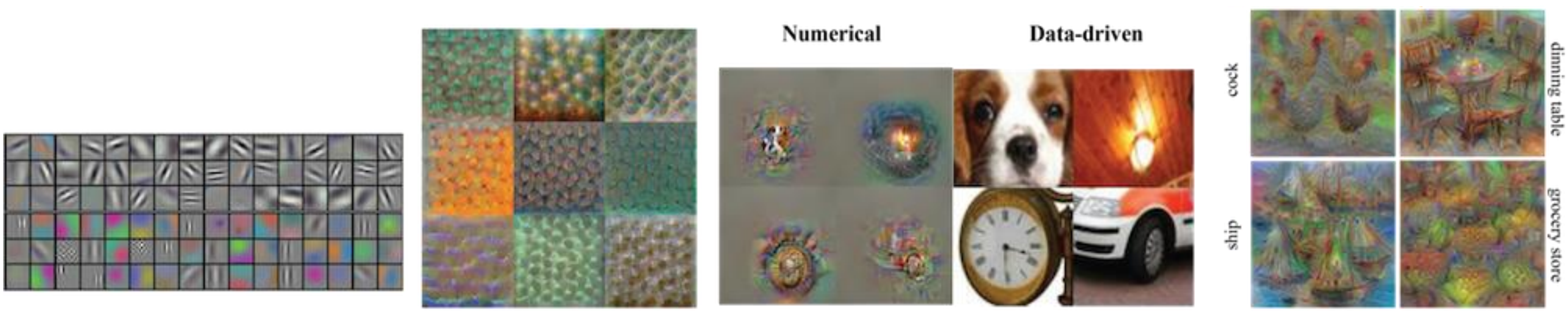




CNN Network Architecture – Hierarchical Representation

A convoluted neural network consists of a hierarchy of layers, whereby each layer transforms the input data into more abstract representations (e.g. edge -> nose -> face). The output layer combines those features to make predictions.

First Layer Representation Second Layer Representation Third Layer Representation

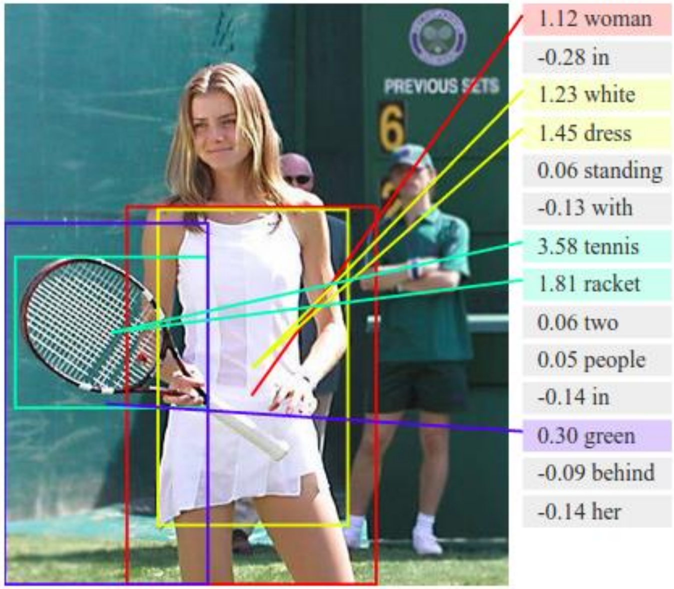


Conv 1: Edge+Blob Conv 3: Texture Conv 5: Object Parts Fc8: Object Classes

CNN Network Architecture – Examples

Alpha GO:

- Fully-connected feedforward network can be used
- But CNN performs much better



19 x 19 matrix

Black: 1

white: -1

none: 0

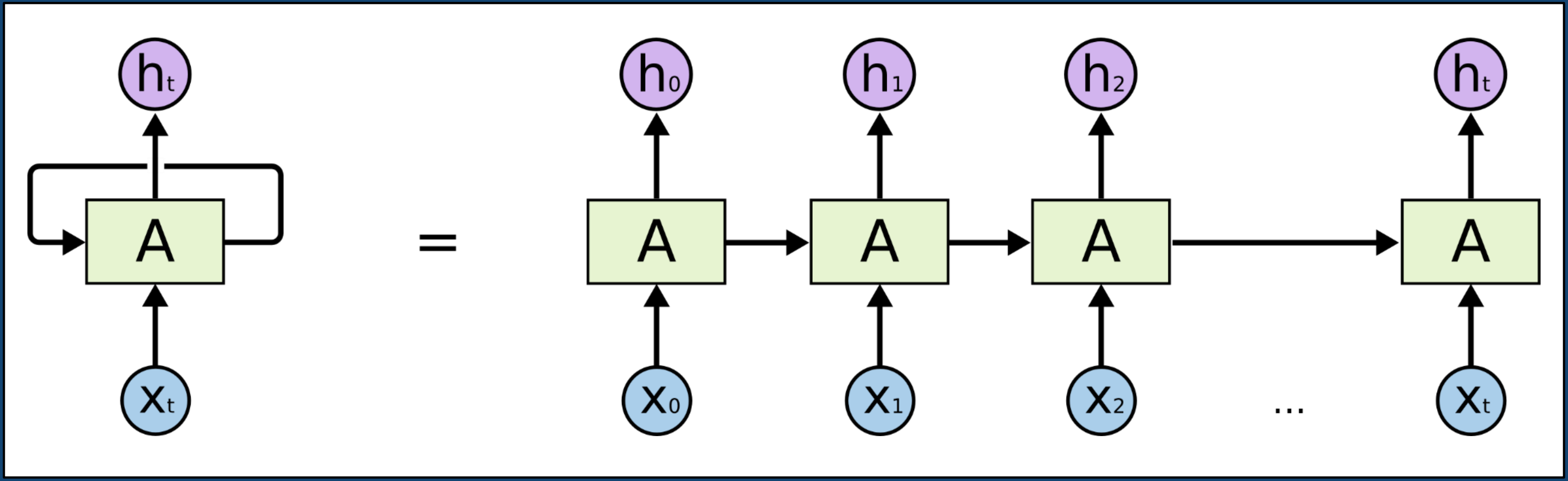
Neural Network

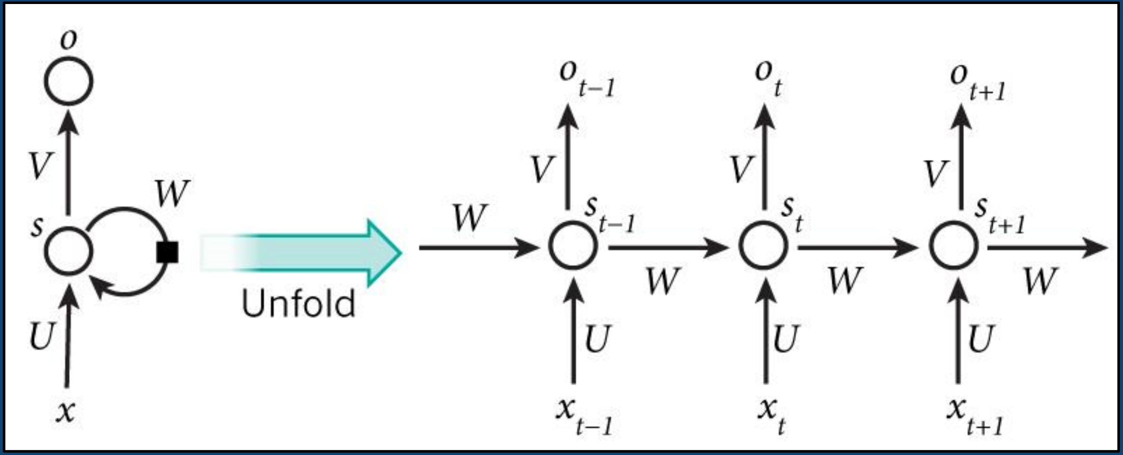
Next move

(19 x 19 positions)

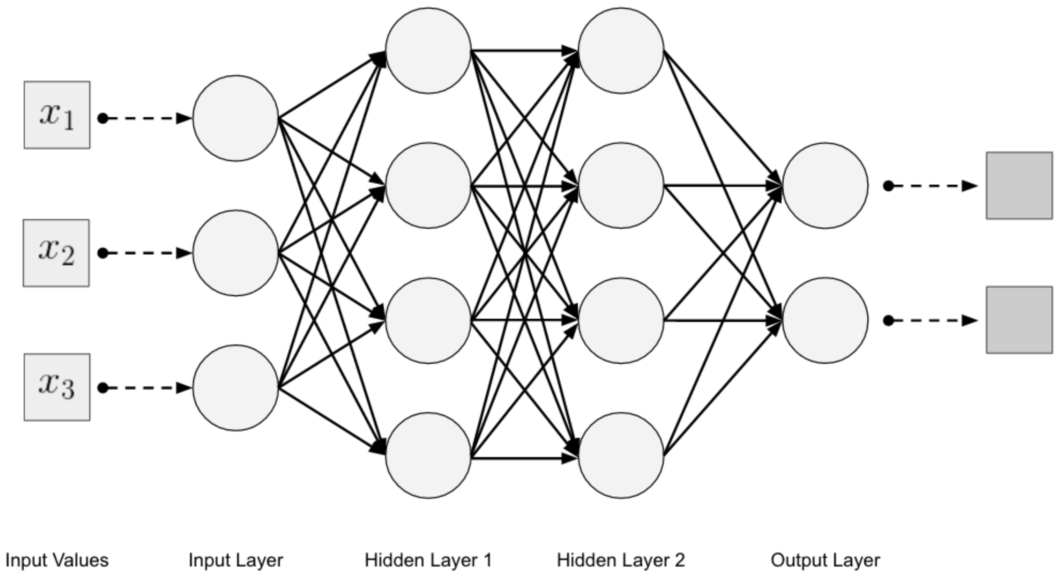
Recurrent Neural Networks (RNN)

Architecture





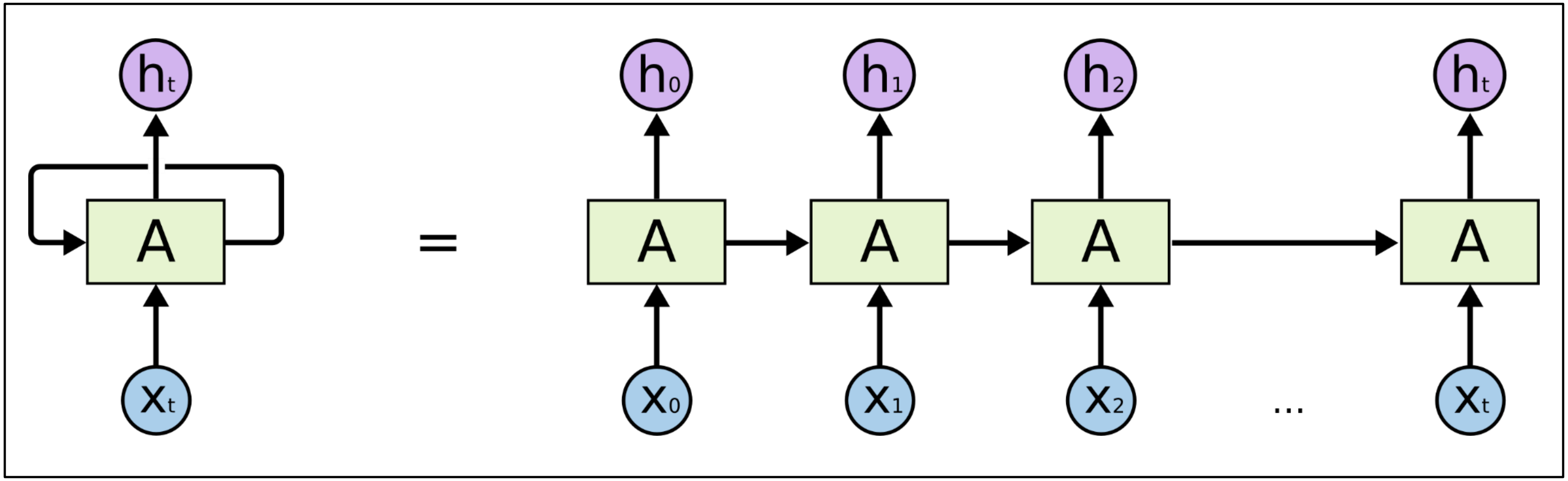
Recurrent Neural Networks - Introduction

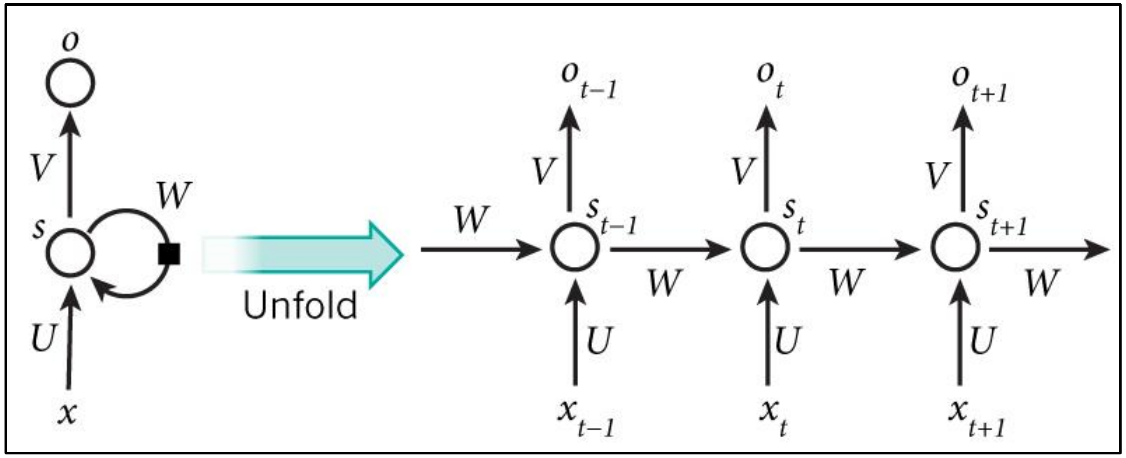


- If input amount: $x_1, x_2, x_3, \ldots, x_n$, is large and ***increasing*** (large n), the network would become too large and is unable to train

- We will now input one x_i at a time, and re-use the same network weights

Recurrent Neural Networks – Model Representations

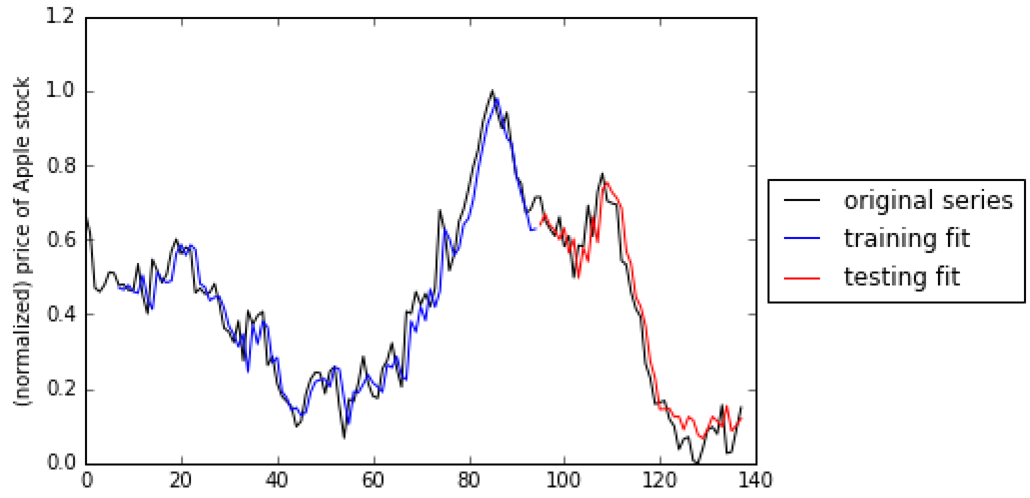


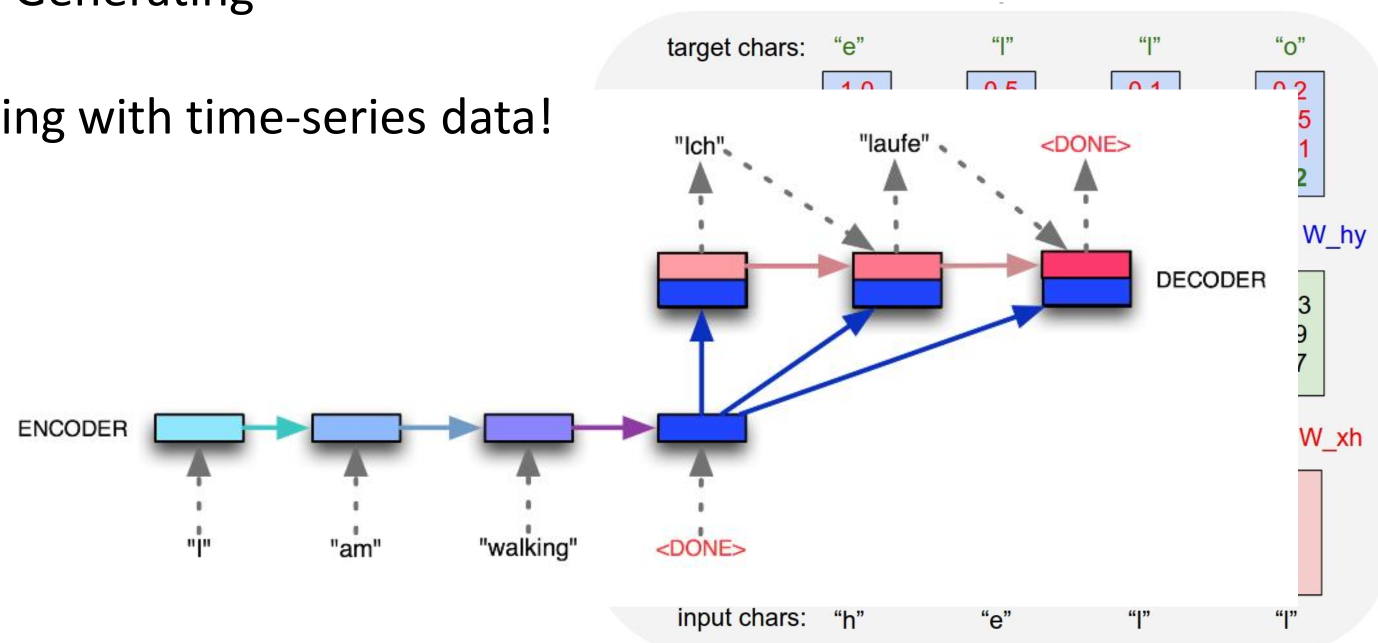


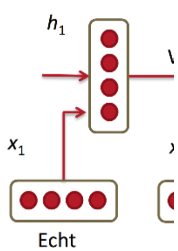
Recurrent Neural Networks – Application

- Time Series Predictions
  - Stock prices
- Natural Language Processing
- Translation
- Speech Recognition
- Video Processing
- Music Generating

- Anything with time-series data!

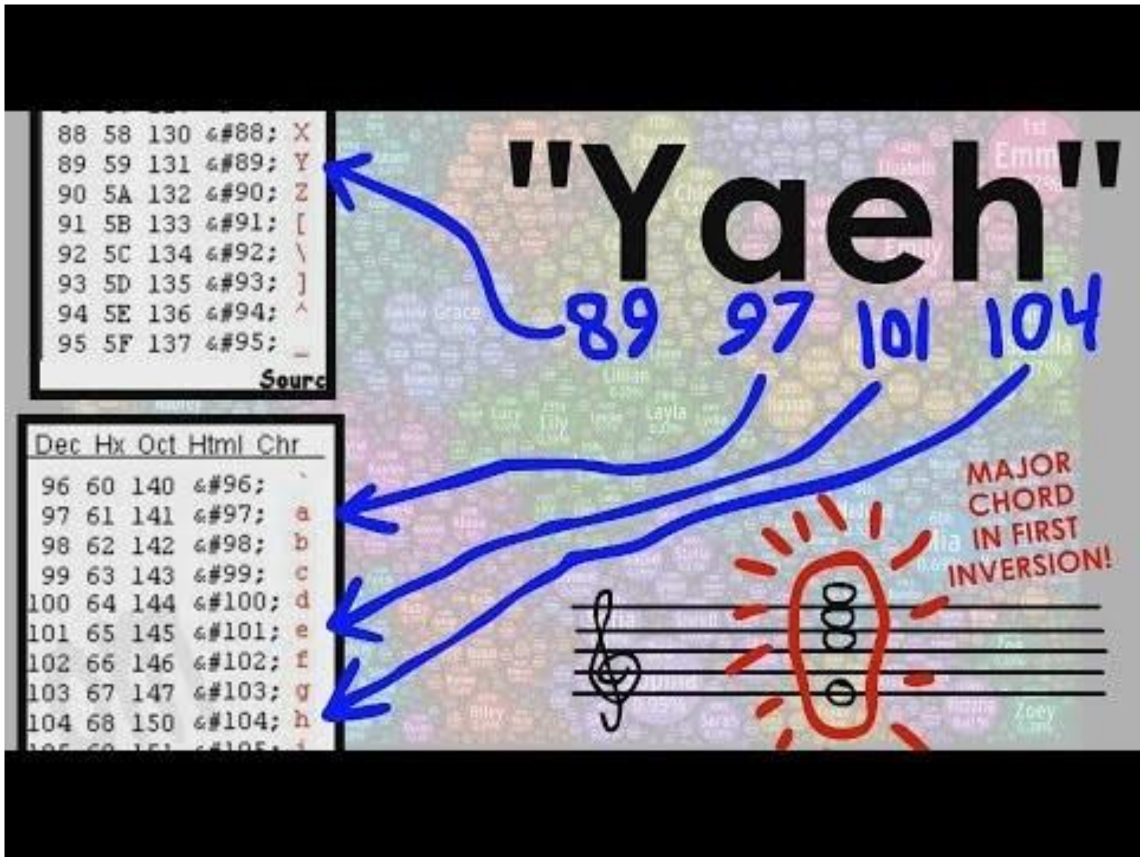






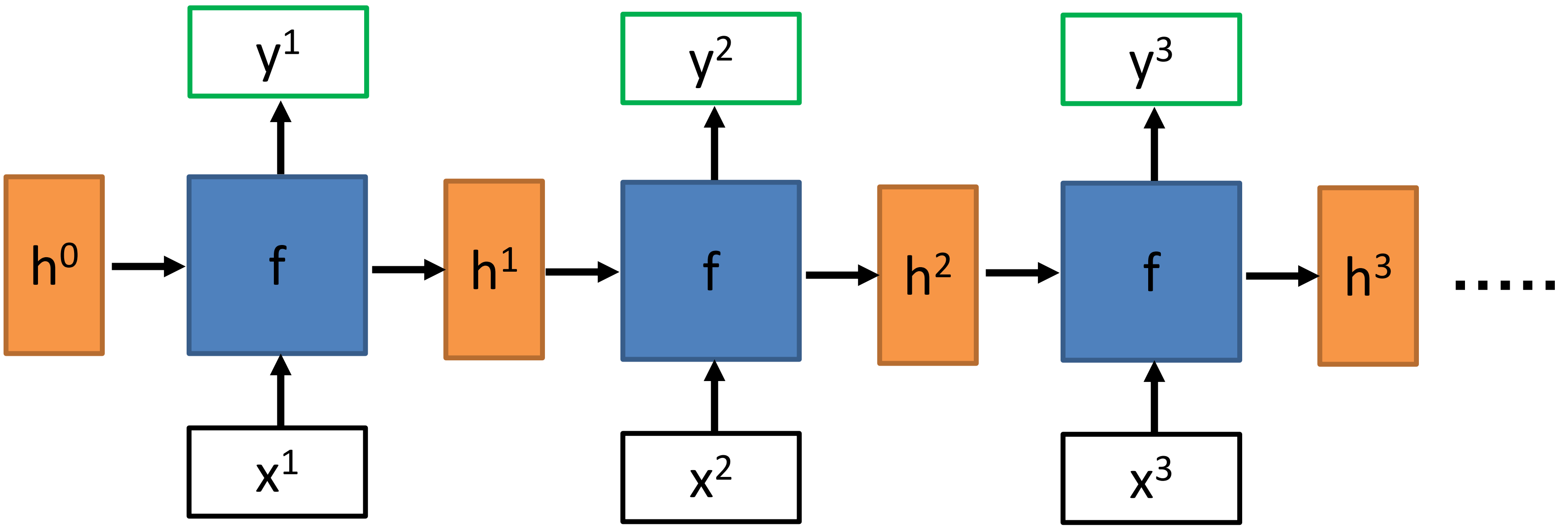
Recurrent Neural Networks – Application

Music Generating

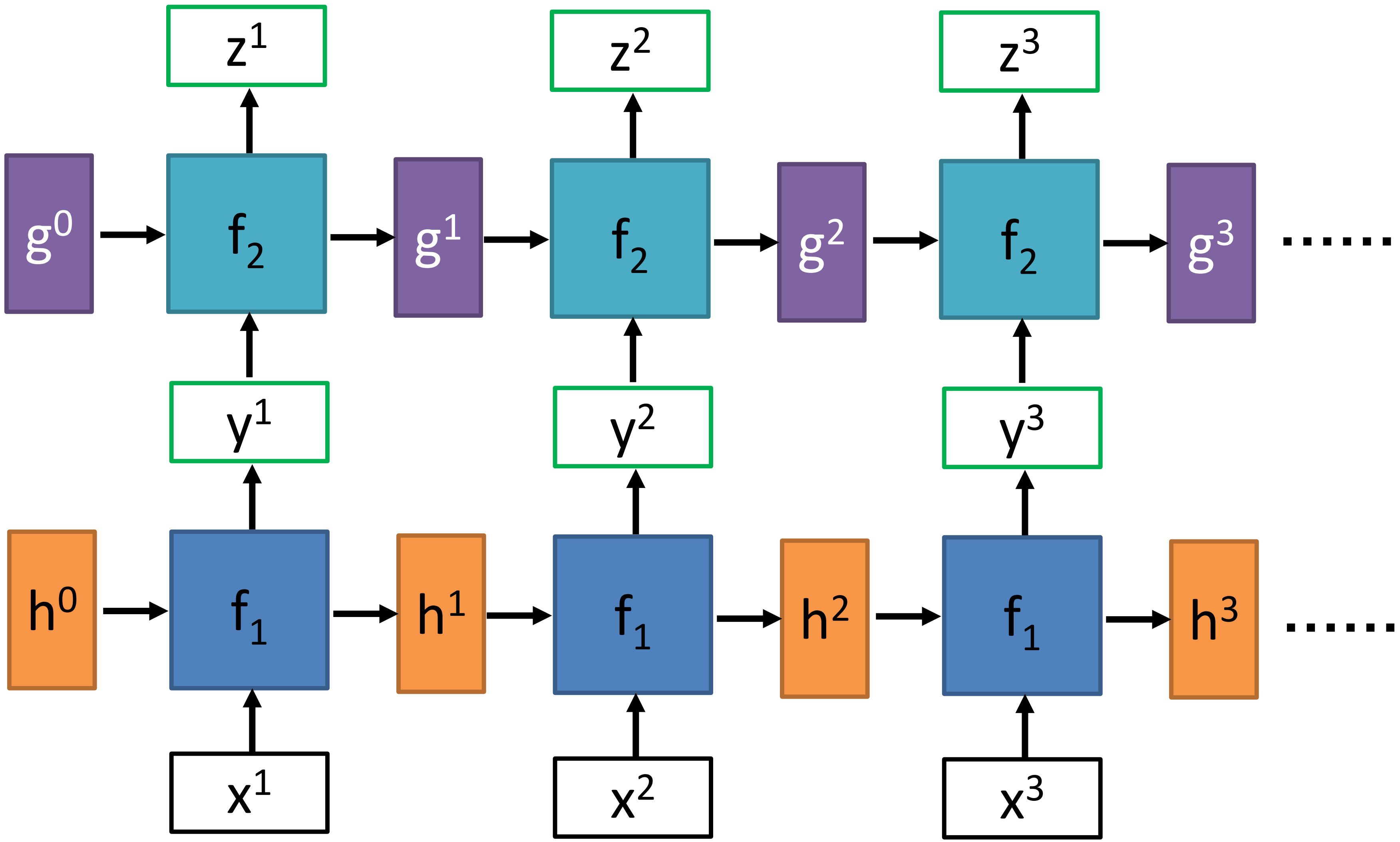


Recurrent Neural Networks – Architecture

- We can apply the same function f to an unbounded number of inputs x_i

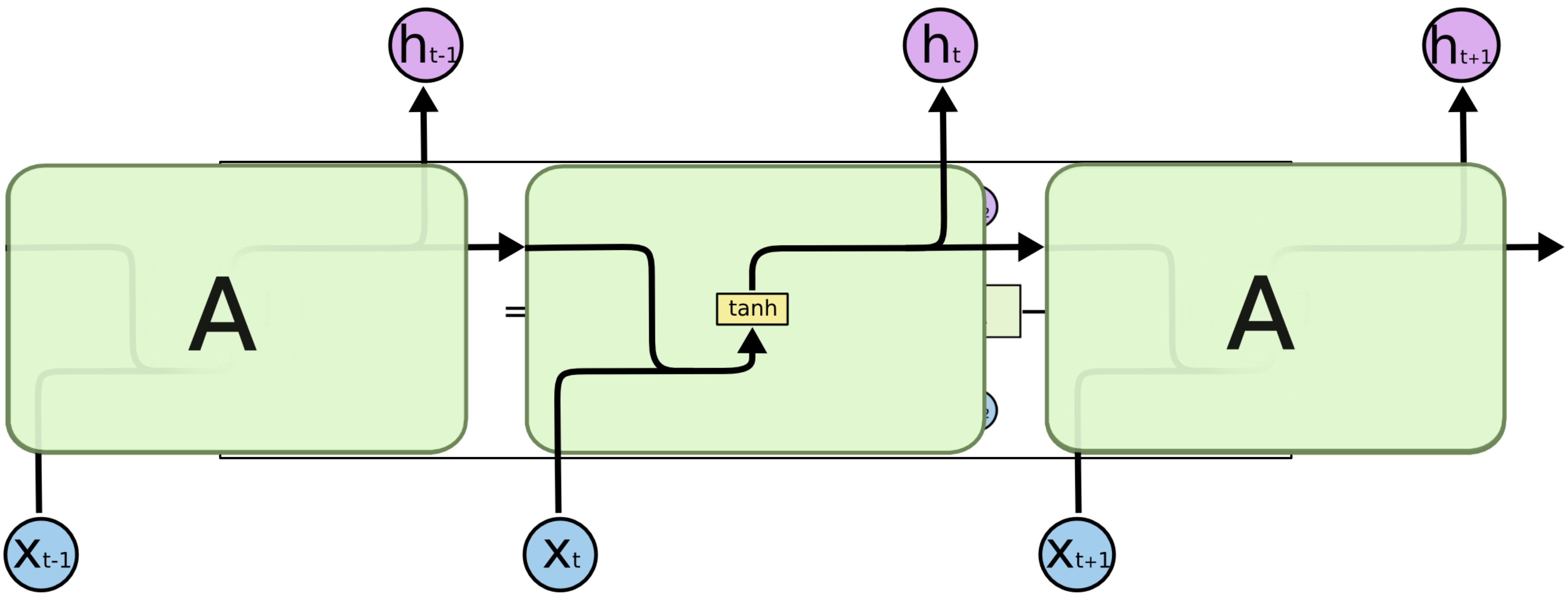


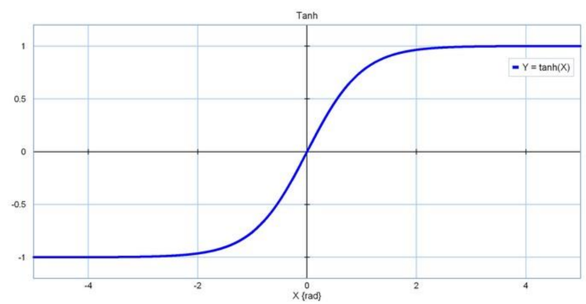
Recurrent Neural Networks – Deep RNN



Recurrent Neural Networks – Naïve RNN

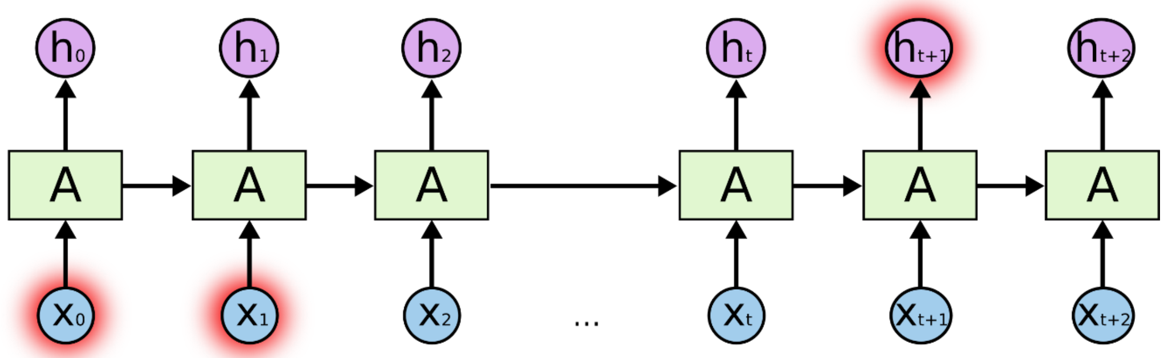
- Single tanh(x) layer as the activation function



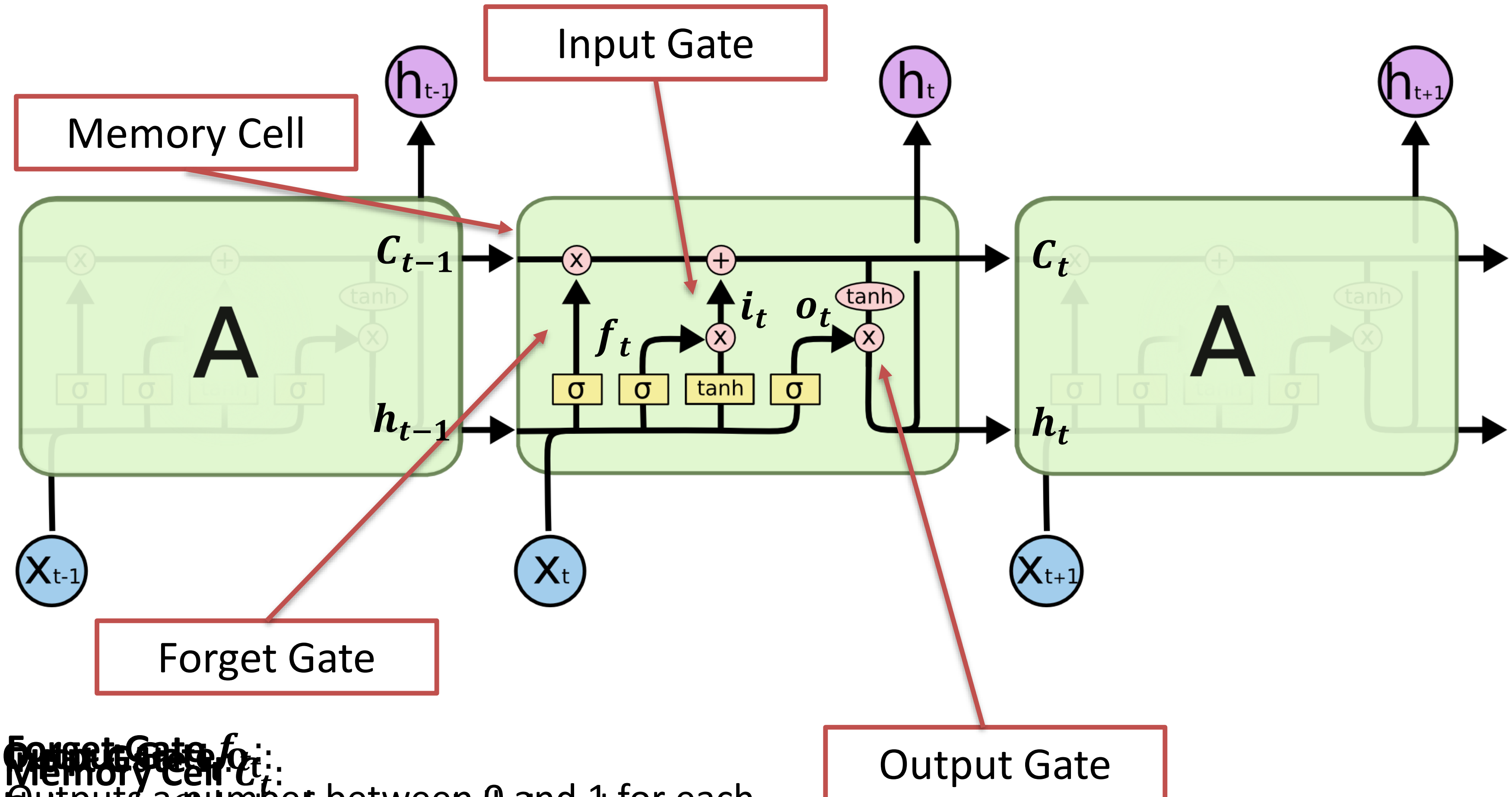


Recurrent Neural Networks – Naïve RNN Criticism

- For time series data, old information tends to be forgotten
- For a distant relationship of unknown length, we wish to have a "memory" to it



Recurrent Neural Networks – LSTM (Long Short-Term Memory)



Forget Gate f_t:
Outputs a number between 0 and 1 for each element in C_{t-1}. A 1 represents to "completely keep this" while a 0 represents to "completely forget this"

Memory Cell C_t:
The cell state C_t changed slowly and it is very easy for information to flow along it unchanged.

Recurrent Neural Networks – LSTM + CNN

- Self driving!

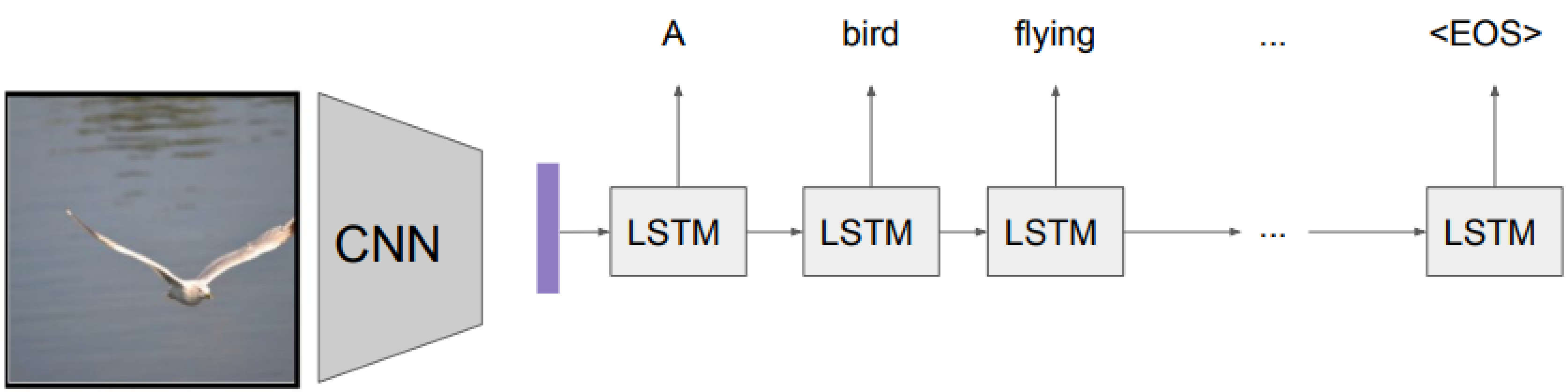
- Convolute an image for object recognition (CNN), and recur (LSTM) over a series of images/frames (video)

Recurrent Neural Networks – Image Captioning

- Neural Image Caption Generator generates fitting natural-language captions only based on the pixels by combining a vision CNN and a language-generating RNN

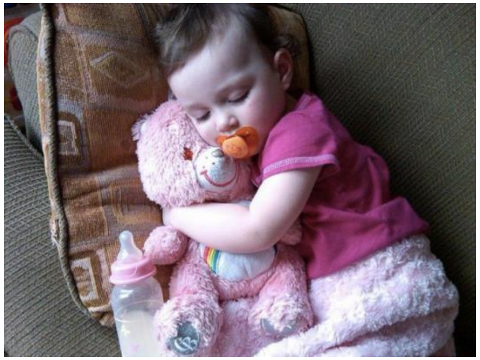
E.g.: Image Captioning

A bird flying over a body of water



Recurrent Neural Networks – Image Captioning Examples

- Examples (success and failure)



A close up of a child holding a stuffed animal

Two pizzas sitting on top of a stove top oven

A man flying through the air while riding a skateboard

Recurrent Neural Networks – Image Captioning Examples

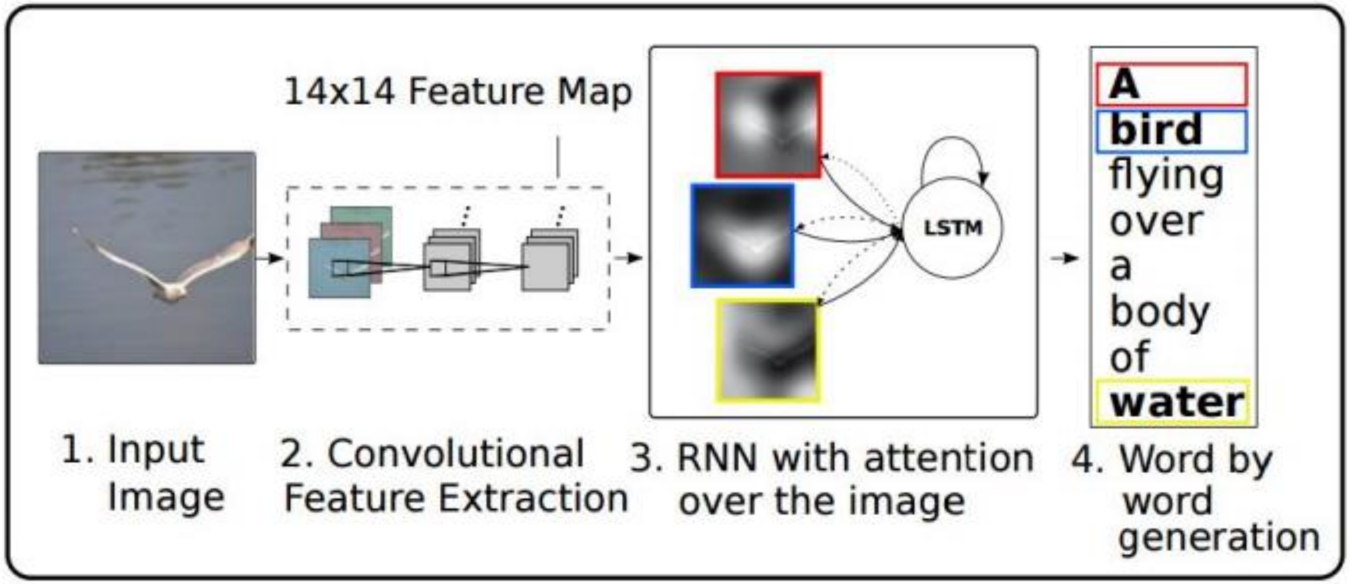
- Examples (success and failure)

Describes without errors	Describes with minor errors	Somewhat related to the image

A person riding a motorcycle on a dirt road.	Two dogs play in the grass.	A skateboarder does a trick on a ramp.
A group of young people playing a game of frisbee.	Two hockey players are fighting over the puck.	A little girl in a pink hat is blowing bubbles.

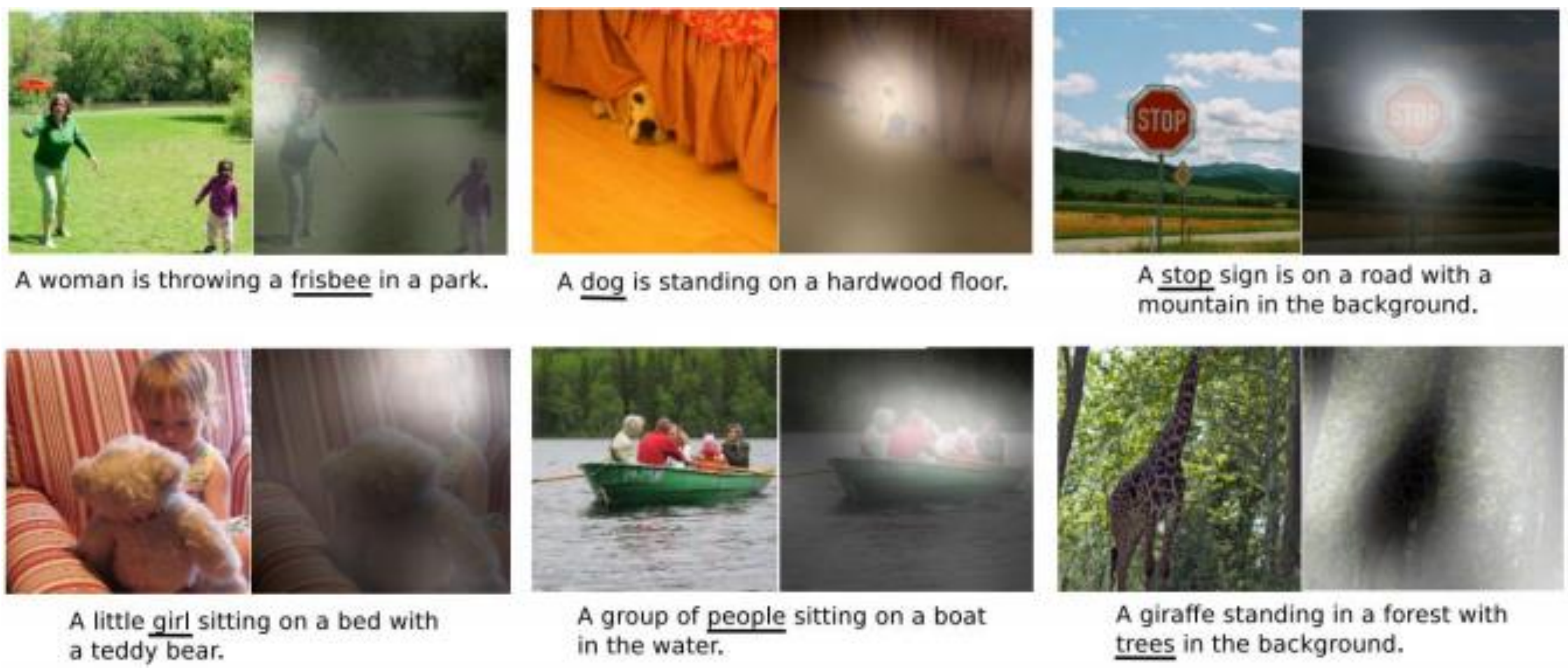
Recurrent Neural Networks – Attention Mechanism

- CNN + LSTM can provide 'attention' to an area of an image / video



Recurrent Neural Networks – Attention Mechanism Examples

- CNN + LSTM can provide ‘attention’ to an area of an image / video



Generative Adversarial Networks (GANs) – 2014

Architecture

Generative Adversarial Networks – Introduction

- First introduced by Ian Goodfellow et al. in 2014

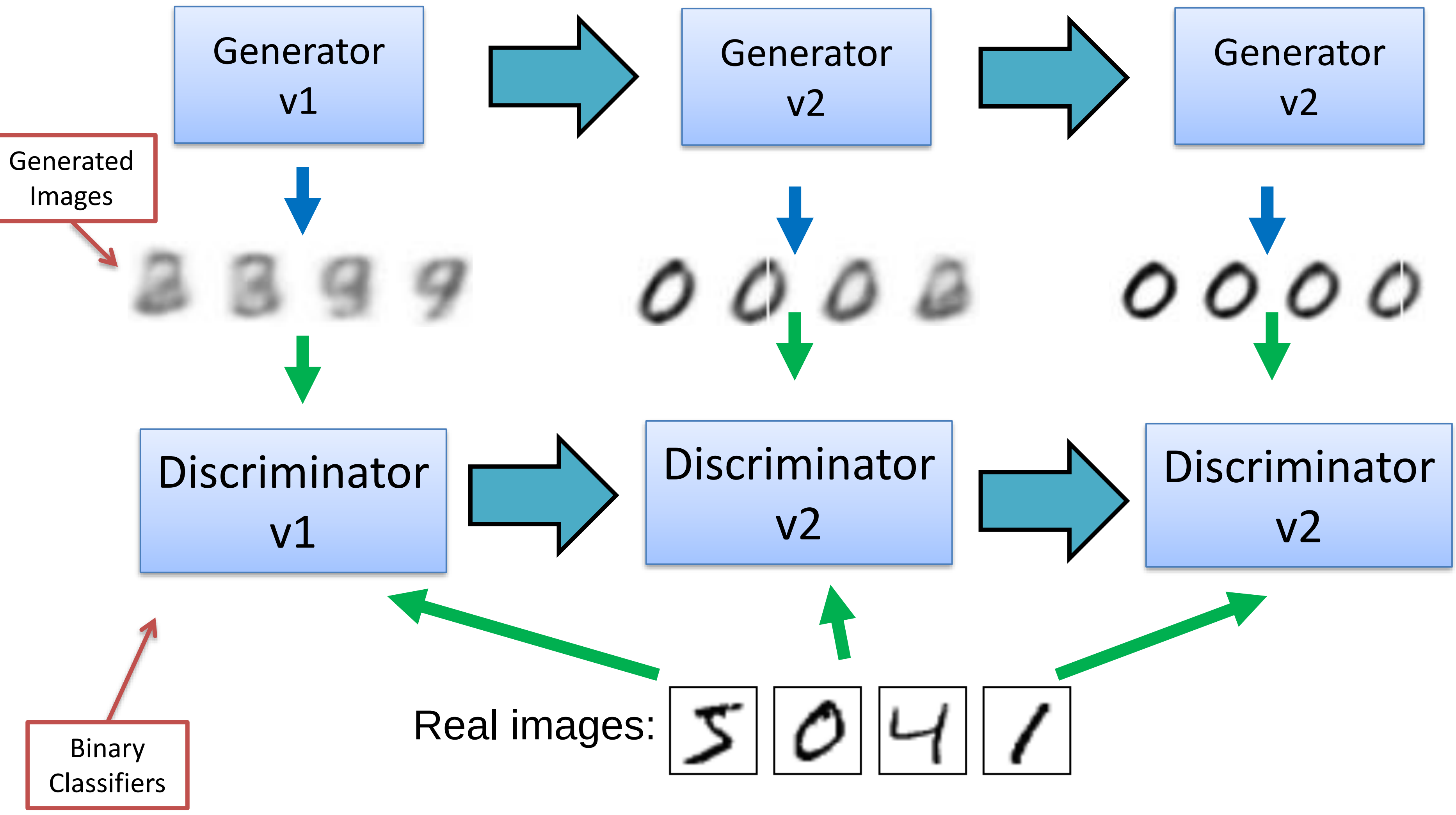
- GANs have been used to generate images, videos, poems, and some simple conversation

Generator:
- Generates candidates/images (from a probability distribution)
- It's objective is to 'fool' the discriminator by producing novel synthesized instances that appear to come from the true data

Discriminator:
- Evaluates the generated images to see if they come from the true data or not

Backpropagation applied to both networks:
- Generator to produce better images
- Discriminator to be more skilled at evaluating generated images

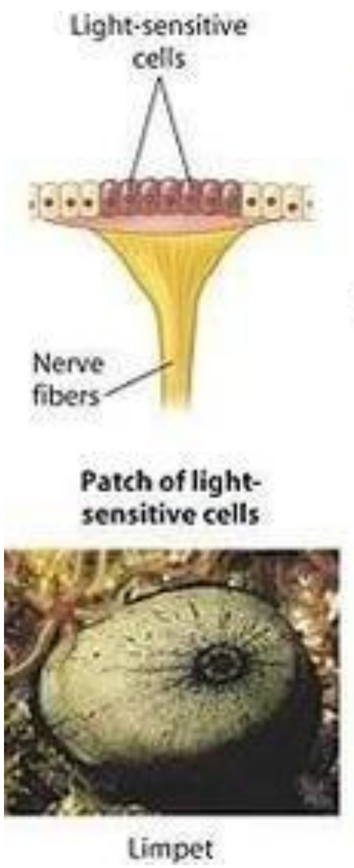
Generative Adversarial Networks – Training a Generator



Generative Adversarial Networks – Training a Generator

50,000
Rounds

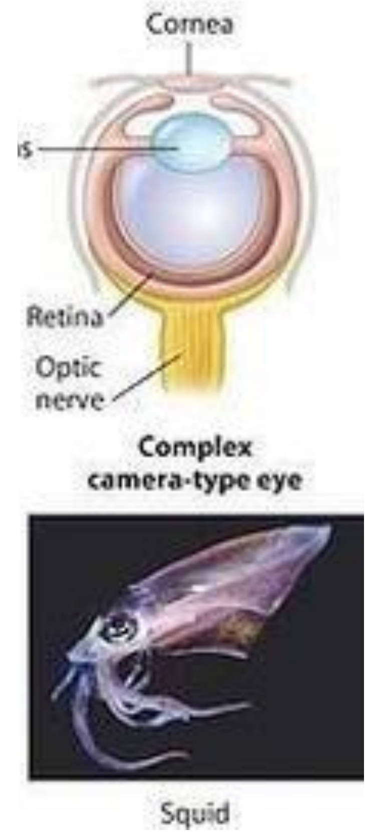
Generative Adversarial Networks – Evolution as a GAN



Time

Genetic Offspring = Generator

Predator / Prey = Discriminator

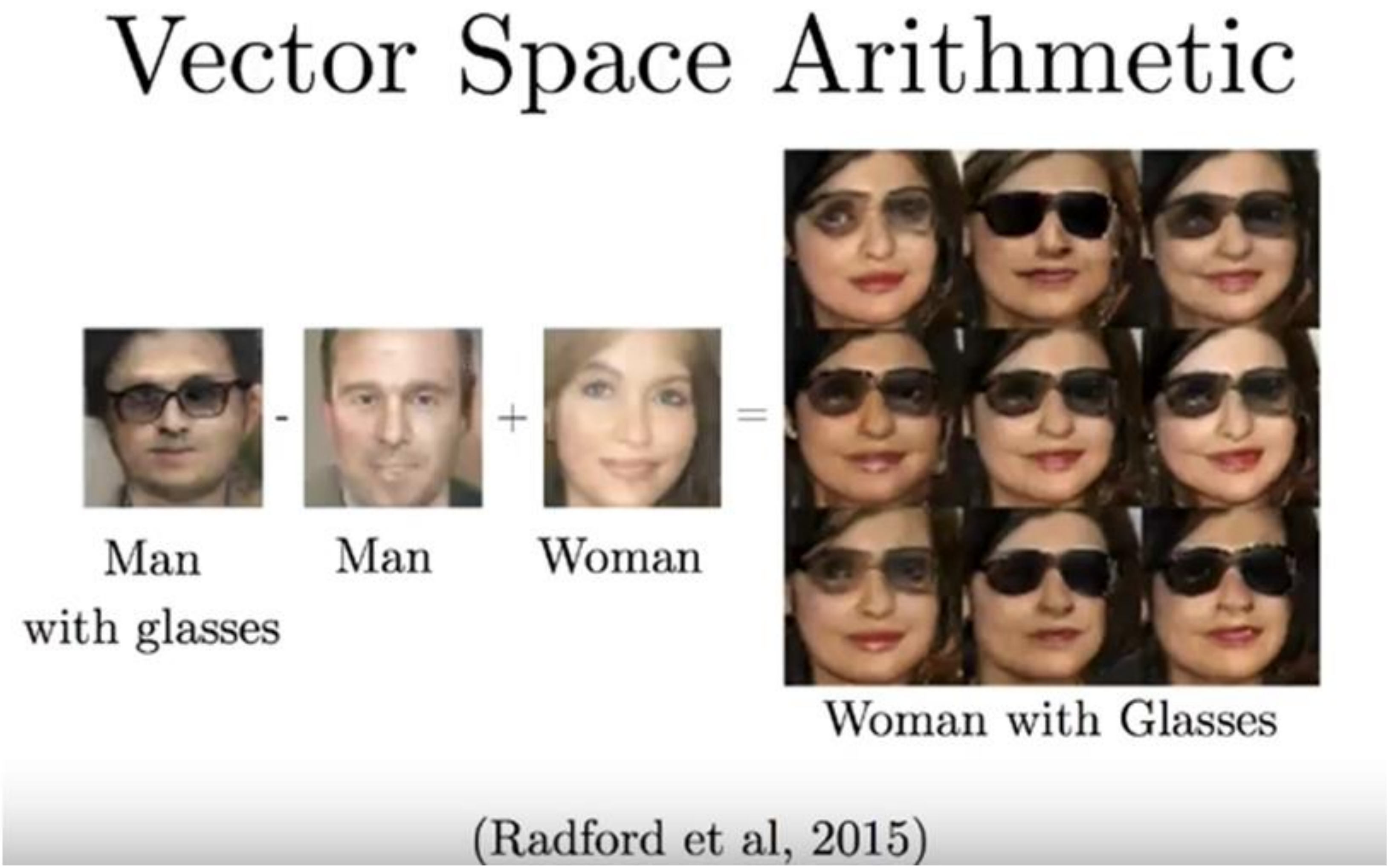


Generative Adversarial Networks – Image Generating Examples

DCGANs for LSUN Bedrooms

(Radford et al 2015)

Generative Adversarial Networks – Vector Arithmetic



Generative Adversarial Networks – Text to Image

Caption	Image
a pitcher is about to throw the ball to the batter	
a group of people on skis stand in the snow	
a man in a wet suit riding a surfboard on a wave	

Generative Adversarial Networks – Text to Image

Caption	**Image**
this flower has white petals and a yellow stamen	
the center is yellow surrounded by wavy dark purple petals	
this flower has lots of small round pink petals