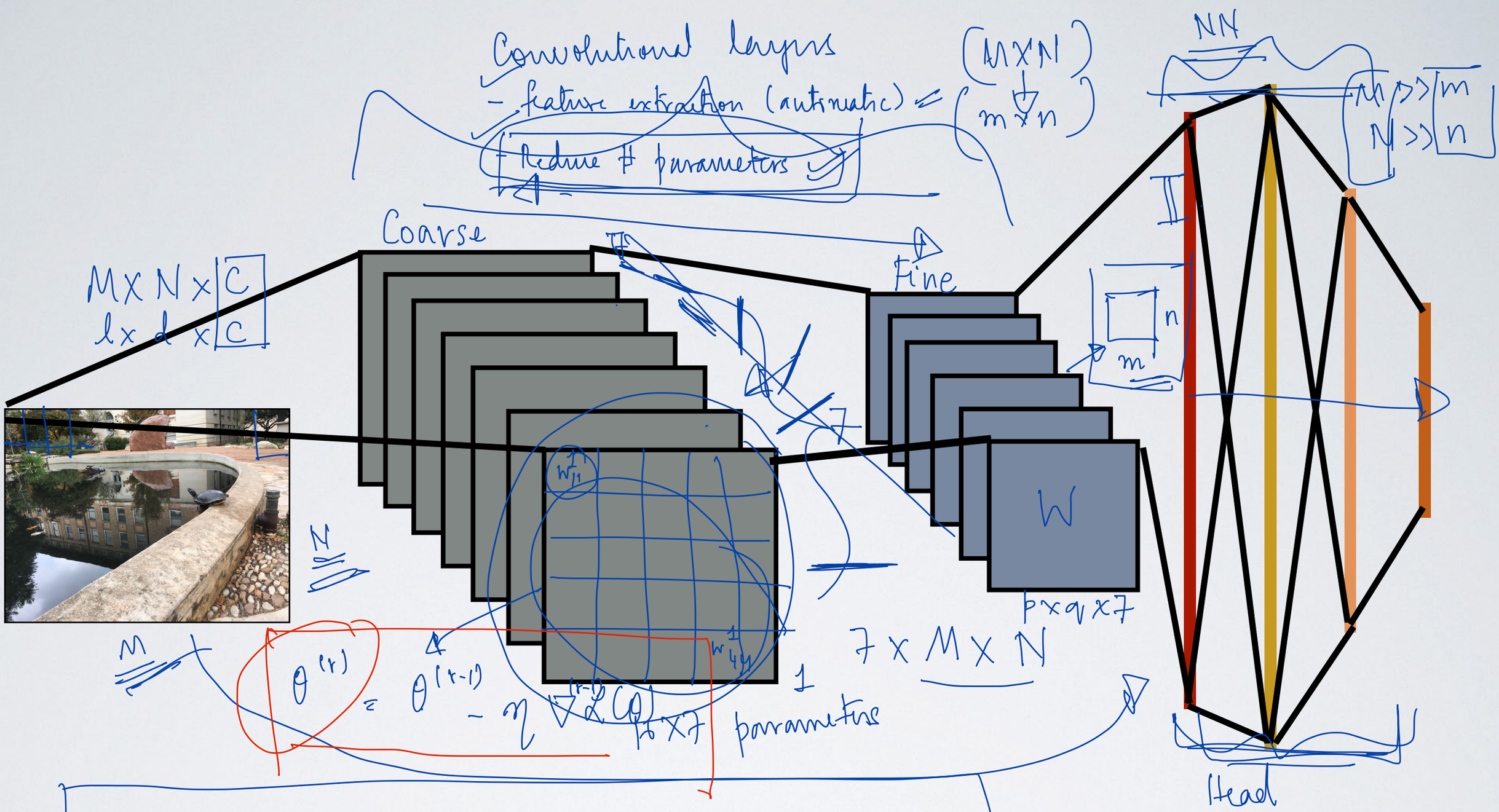


CONVOLUTIONAL NEURAL NETWORKS - I

Sumohana S. Channappayya

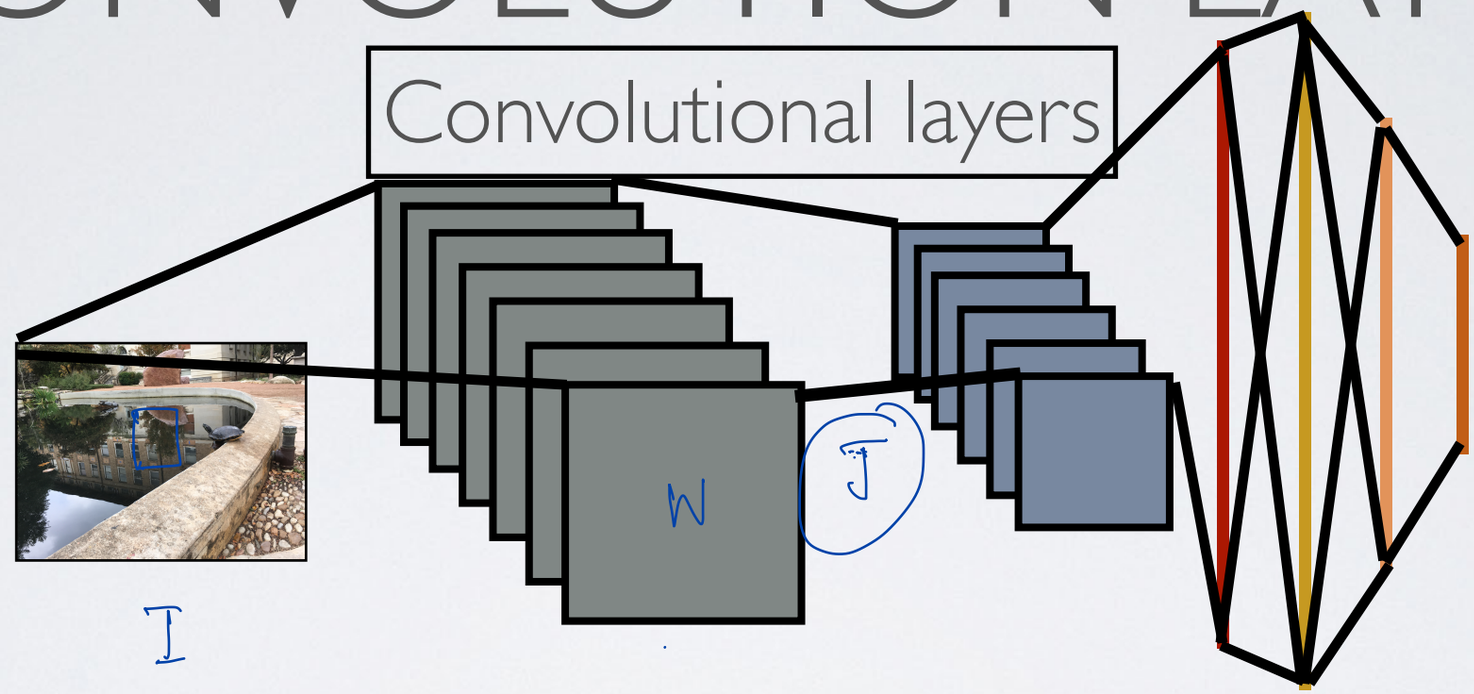
OUTLINE

- Motivation
- Building blocks
  - Convolutional layer/non-linearity layer
  - Pooling layer
  - Fully connected layer
- Training
  - Loss function
  - Gradient descent and learning rate
  - Data handling and stopping condition
- Generalisation

MOTIVATION

- Human Visual System (HVS)
- Dealing with 2D data
- Parameter reduction

BUILDING BLOCK: CONVOLUTION LAYER



- *Linear component* of the network
- Parameter *sharing*
- *Local* connectivity

$$J[p,q] = \sigma\left(\sum_m \sum_n w[m,n]\, I[p-m, q-n]\right)$$

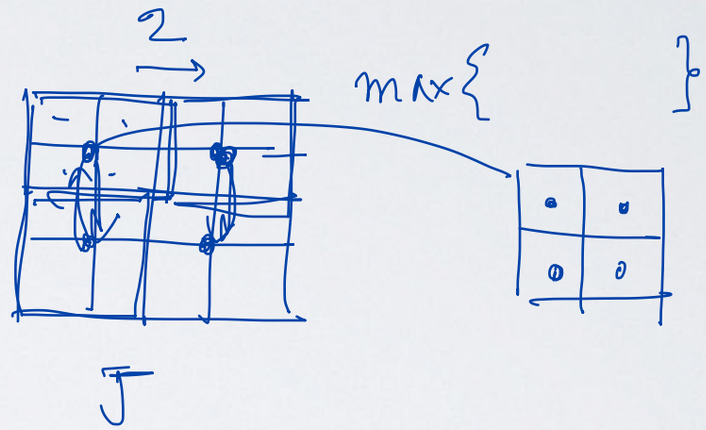
(NL) Conv (linear)

BUILDING BLOCK: NON-LINEARITY LAYER

- Fundamental to NN
  - Allows for *modelling complex input/output relations*
  - Inspired by the *neuron*
  - Examples: sigmoid, tanh, ReLU etc.
  - Also called *Activation Layer*

BUILDING BLOCK: POOLING LAYER

- Pooling of output



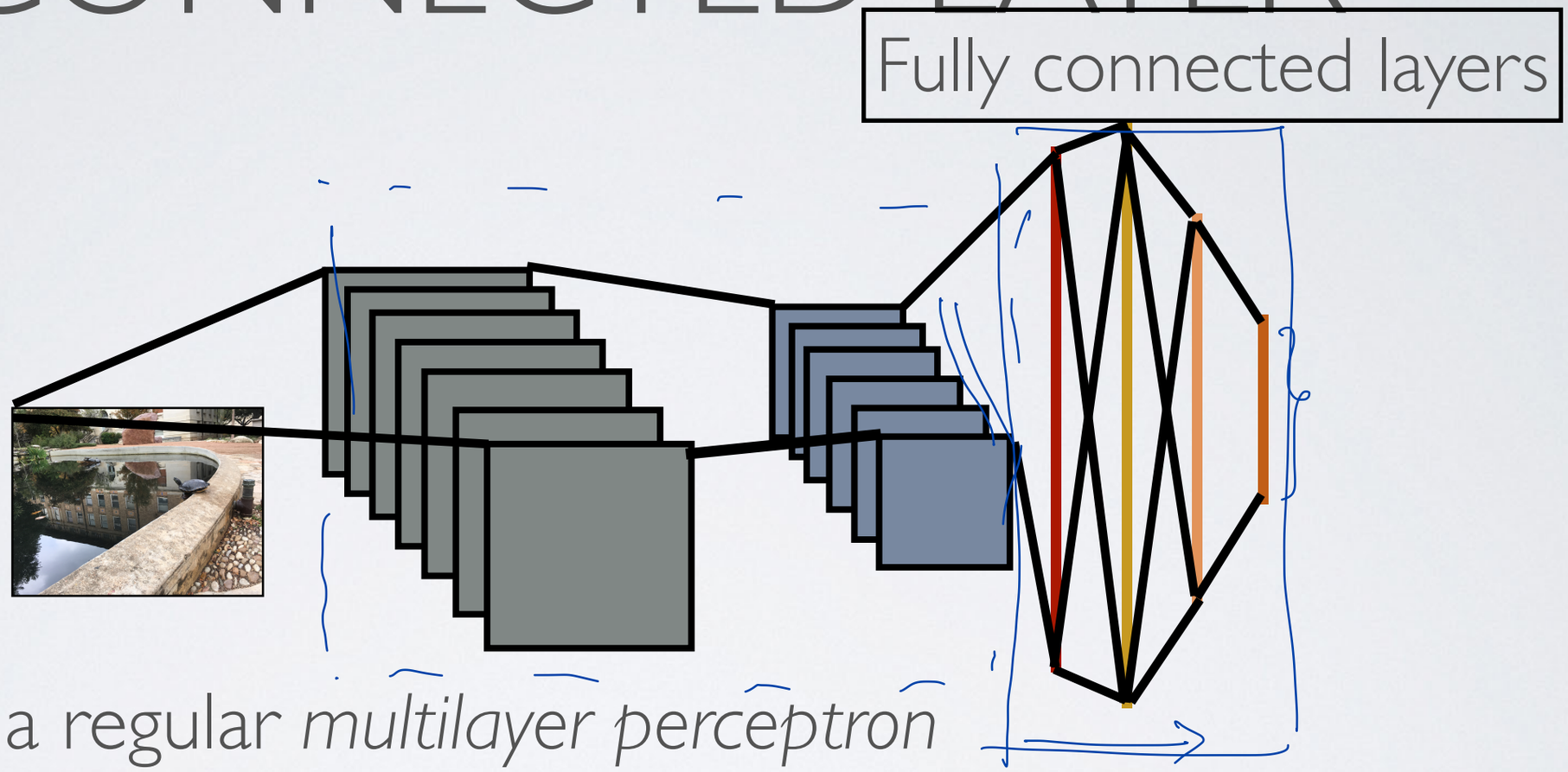
  - *Dimensionality reduction*

- Control *overfitting*

- *Invariant to small translation*

- Examples: max, l2 norm, average etc.

BUILDING BLOCK: FULLY CONNECTED LAYER



- Essentially, a regular *multilayer perceptron*
- Input is a vector formed by *flattening* pooling layer output
- Output is vector whose size equals *number of classes*

TRAINING: LOSS FUNCTION

- Standard loss functions such as *softmax* or *cross entropy*
- Loss functions are *non-convex* leading to *locally optimal* solutions

TRAINING: GRADIENT DESCENT BASED METHODS

- *Several* optimisation methods can be applied:
  - Gradient descent, Stochastic gradient descent
  - Momentum, Nesterov momentum
  - AdaGrad, RMSProp, Adam
- *Stopping* condition based on *training/validation error*
- Choice of *initialisation* important
- *Backpropagation* of forward loss used for gradient computation

TRAINING: PARAMETERS AND DATA

- Similar rules as in neural network training:
  - Weights typically *initialised randomly*
  - *Pre-trained weights* also used commonly
  - Feed data in *small batches (mini batches)*
  - *Epoch* is one forward and backward pass of all training data points

GENERALISATION

- Any model's performance is measured on *previously unseen inputs*
- This is also known as *generalisation*
- Typically, the dataset is divided into *training, validation and test set*
- Generalisation performance is reported on the *test set*

SUMMARY

- CNNs form the *fundamental building blocks* of modern machine learning models
- Allow for *feature learning*
- *Extremely successful* in solving several machine learning problems: image recognition, video analysis, natural language processing, drug discovery, visual system modelling etc.
- *Reason for the deep learning revolution*

REFERENCES

- https://www.deeplearningbook.org

- https://cs231n.github.io/convolutional-networks/