Postgraduate course
Universitat de Valencia 2020

Introduction to Machine Learning for physicists

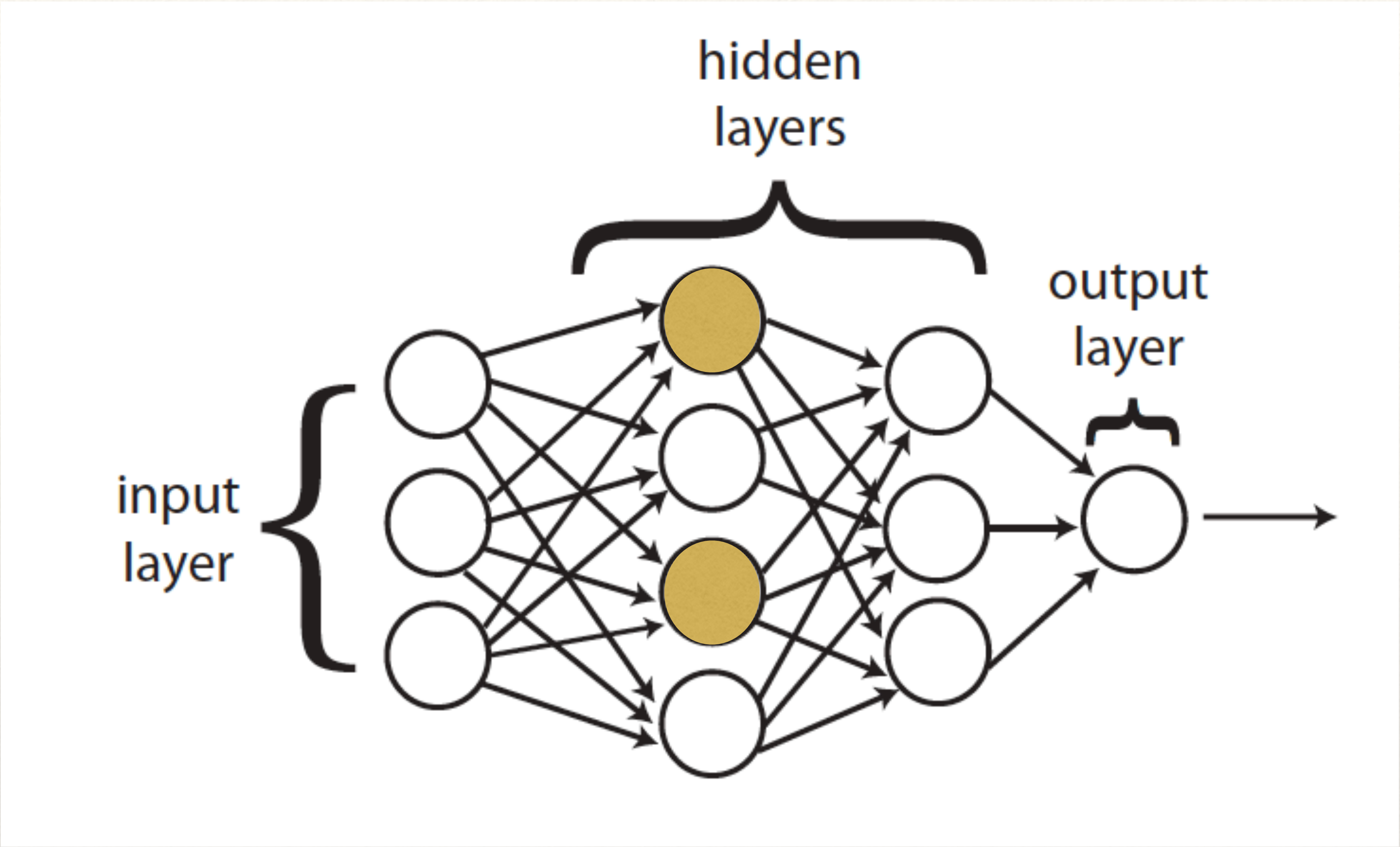
Veronica Sanz (UV/IFIC)

LECTURE 4
SHORT Intro SUSY

Today

You'll be going over the CNN notebook, learning the syntax of keras,
the different parameters in Sequential etc
but some of you may finish early
additional task: Supersymmetry search
using LogReg, Trees, and DNNs

Why are we doing this?



Neural Networks can model **complexity**

They have a high degree of expressivity

/exhibit high representational power

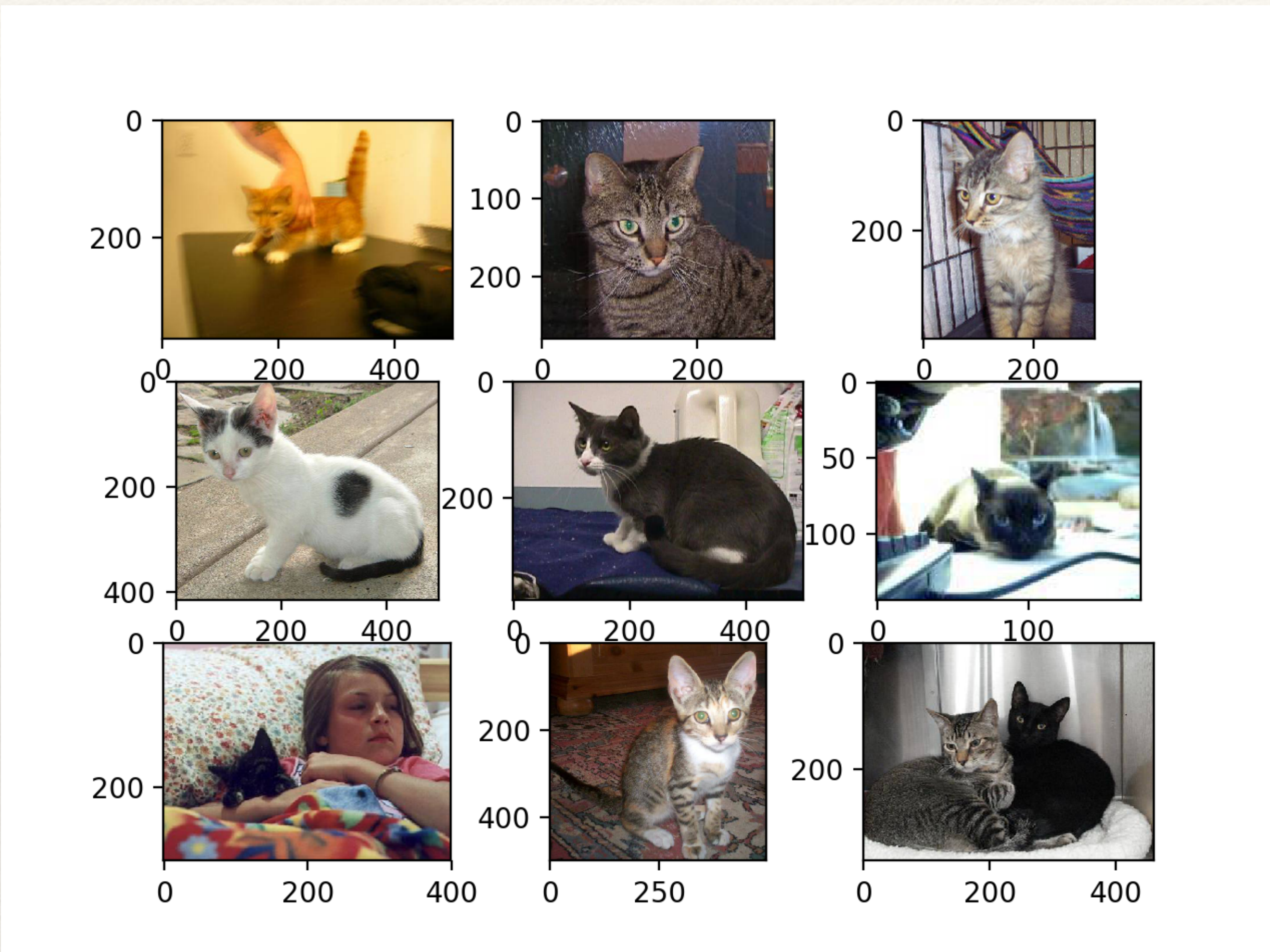
More hidden layers=> more complex features

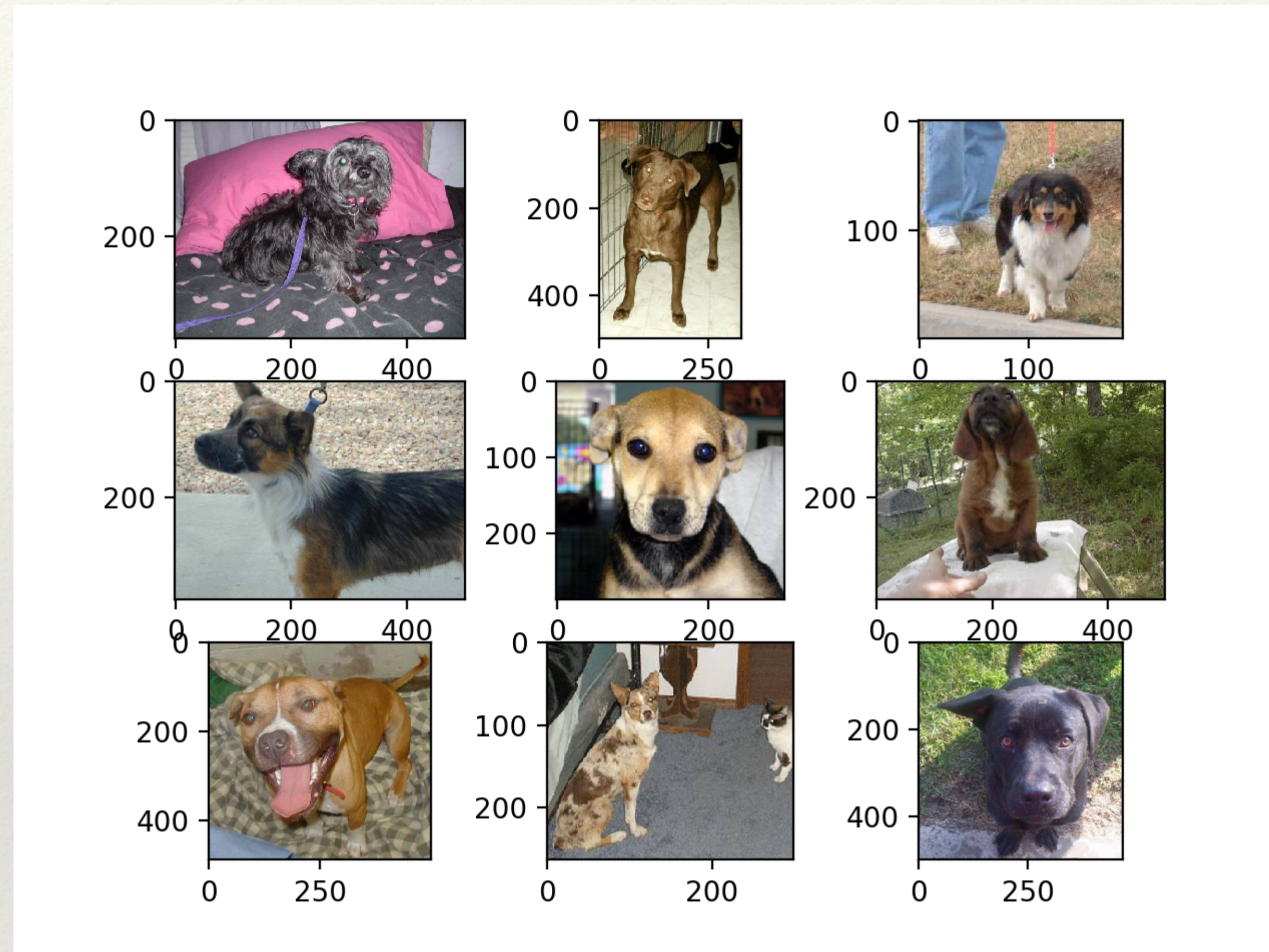
Deep learning, deep NN

Complex features

images, speech : are complex

For example: cats / dogs



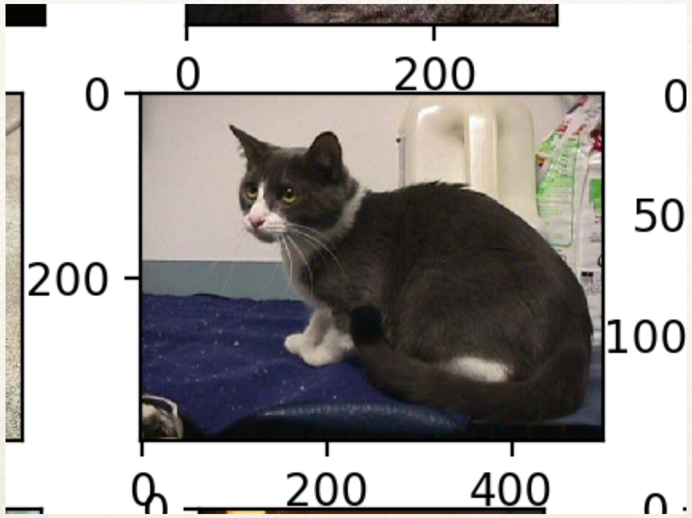


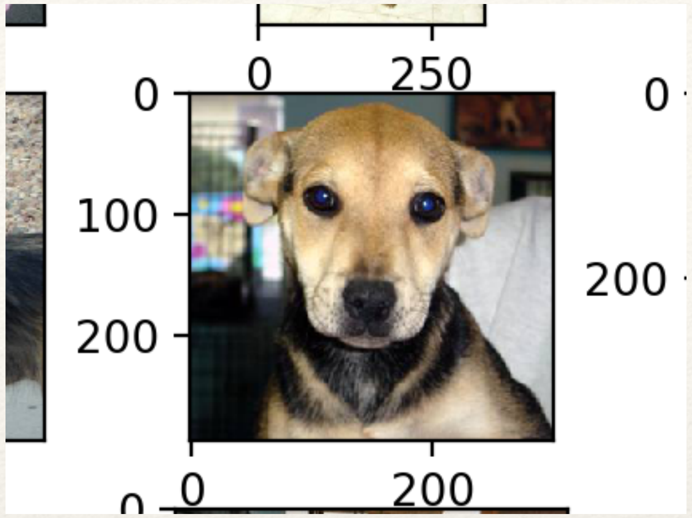
you can distinguish these cats and dogs, right? but how?
would you be able to write a code which classifies them with ~ 100% accuracy? well, a NN can learn to do this!

Convolutional Neural Networks

CNNs

Complex features are often *local*



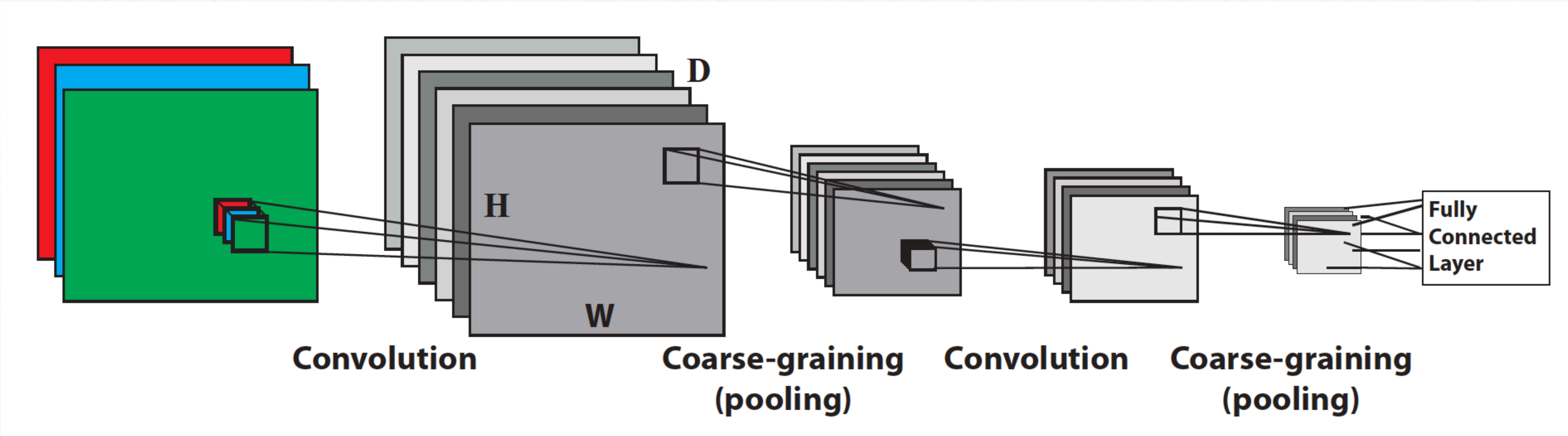


Apart from shape and color,
we know a cat is a cat because there are relations among their features, e.g. the position of the eyes/ ears respect to the head centre, independently of where in the image the cat is

Locality and **translational invariance** must end up playing a role in the identification task

Convolutional Neural Network (CNN)
a type of NN architecture designed to exploit these two characteristics

CNNs

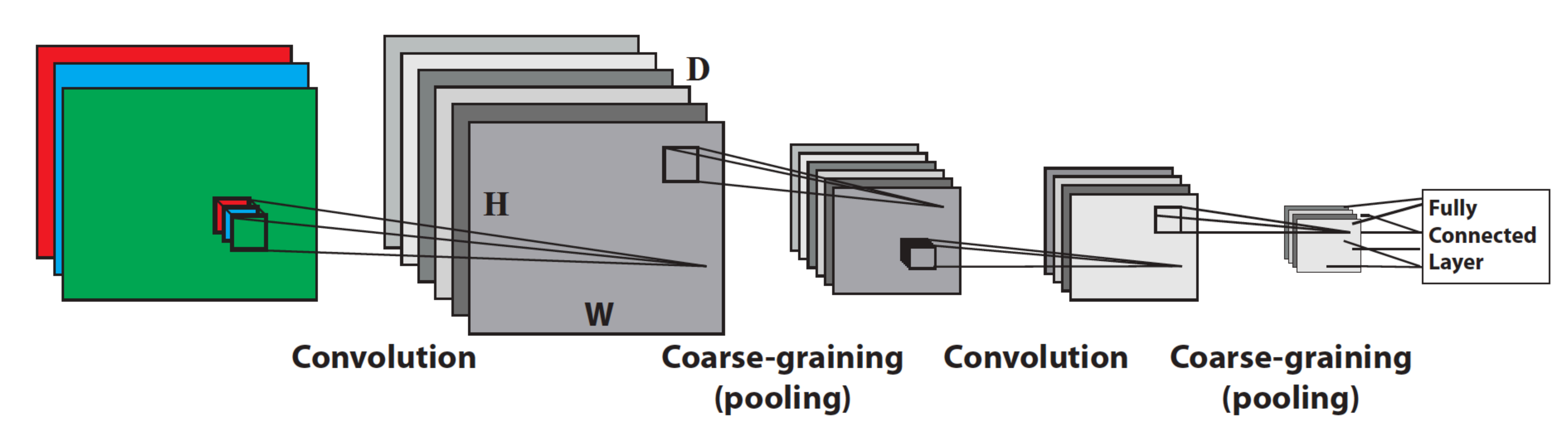


Two types of basic layers

Convolution layer: Height, Width and Depth (e.g. RGB channels)
Convolution= operation to reduce information while maintaining spatial relations (locality and translation properties)

Pooling: Take areas of the image and reduce them. Example, max-pooling would take 2X2 neurons and replace by a single neuron with input the max of the 4

CNNs



Why do we do this?

Too much superfluous information in an image

Need to transform the image and capture the essentials while maintaining spatial relations

As we advance in the layers, the CNN is transforming the original image into something more and more abstract

In physics, translationally invariant systems can be parametrised by wave number and functional form (sin, cos) whereas an arbitrary system would be *much more complex*

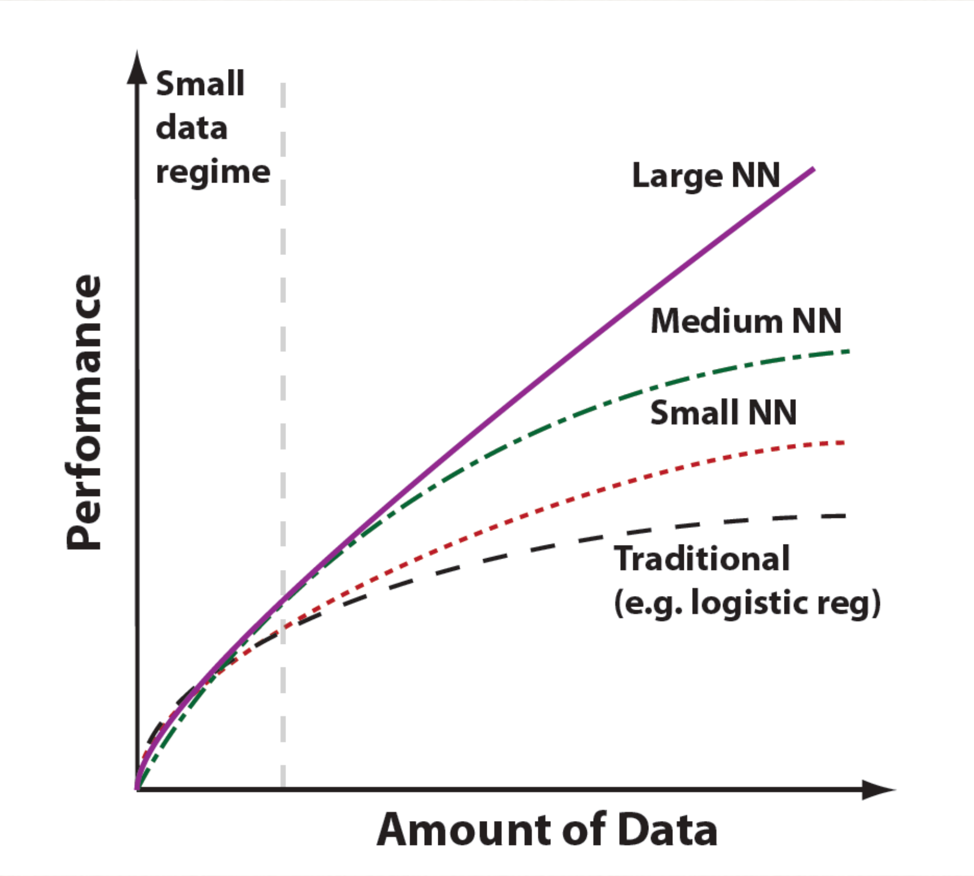
Why are NNs so good at learning?

Good at learning: ability to learn with little *domain knowledge*

That's something physicists (as humans) are good at

(Physics -> other things)

DNNs are good at this too, they are able to take large streams of data and learn features with little guidance, work like *black boxes*



Good at handling large amounts of data: needle in a haystack

The NN structure (layers, 0/1 gates) allows a high representation power with moderate computational demands, e.g. allows parallelisation, use of GPUs…

It scales better than other learning methods (like SVMs)

Today: NNs and CNNs

We will take a standard dataset, MNIST

Build and train a NN
to become better at recognising hand-written numbers
This is a *supervised* ML problem
(we know the true labels)
we train on a large sample (60K) images

We will build a fully connected NN,
a Convolutional Neural Network,
and use Data Augmentation
Our precision will go from 96% till 99%
Link to Google Colab notebook on Neural Networks

TOMORROW: *Unsupervised*