Imperial College London

ACSE 2020

# Reminder CNNs

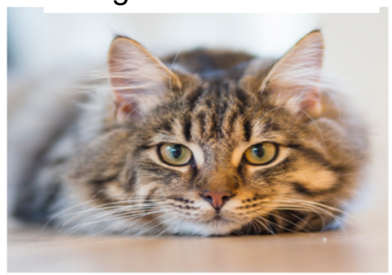
Olivier Dubrule, Lluís Guasch

# Reminder CNNs

Imperial College London

## Computer Vision Topics

Image Classification



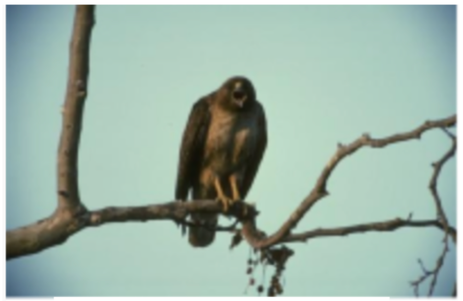
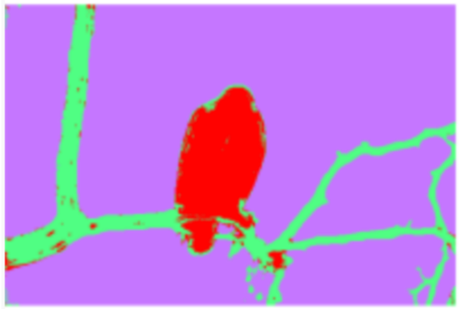
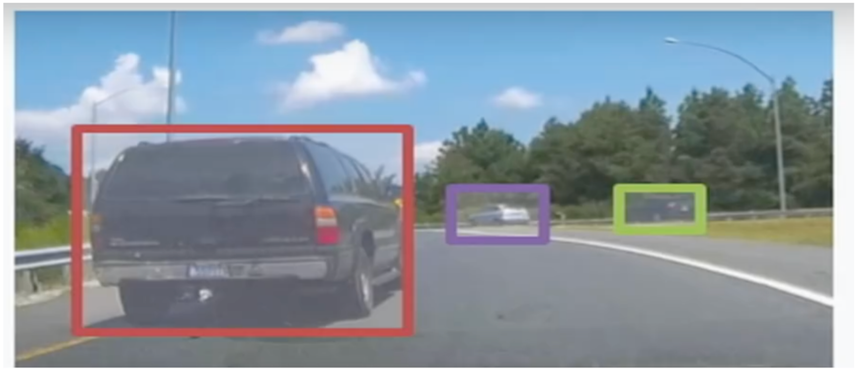


Image Segmentation





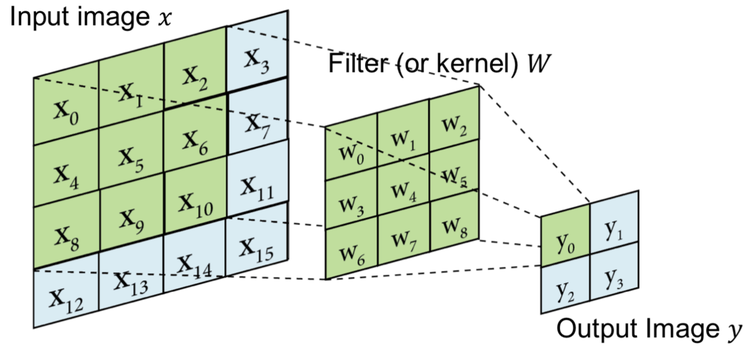
Object Localization (1 object) or Detection (several objects)

# Reminder CNNs

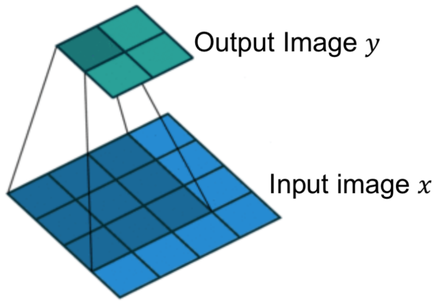
Imperial College London

## Idea Behind Convolutional Neural Networks:

*Parametrize the Filter and Optimize its Coefficients  Based on the Objective!*



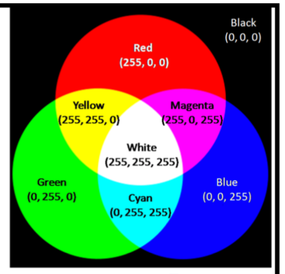

$$W * x = y$$



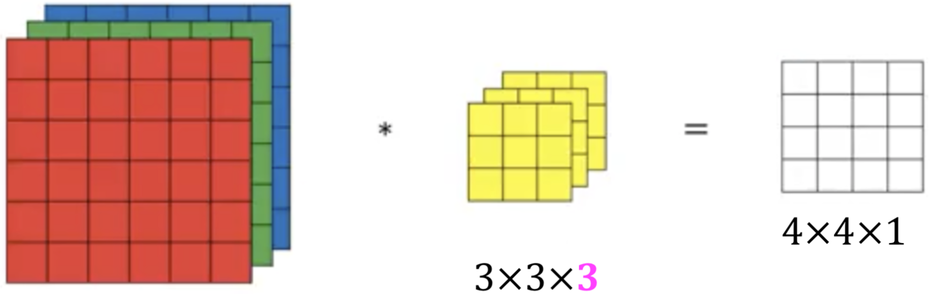

$$y_0 = w_0x_0+w_1x_1+w_2x_2+w_3x_4+w_4x_5+w_5x_6+w_6x_8+w_7x_9+w_8x_{10}$$

# Reminder CNNs

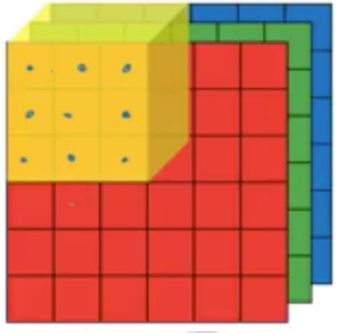
Imperial College London




## If the Two-Dimensional image has 3 RGB channels…




**27 operations each time (28 if bias term)**



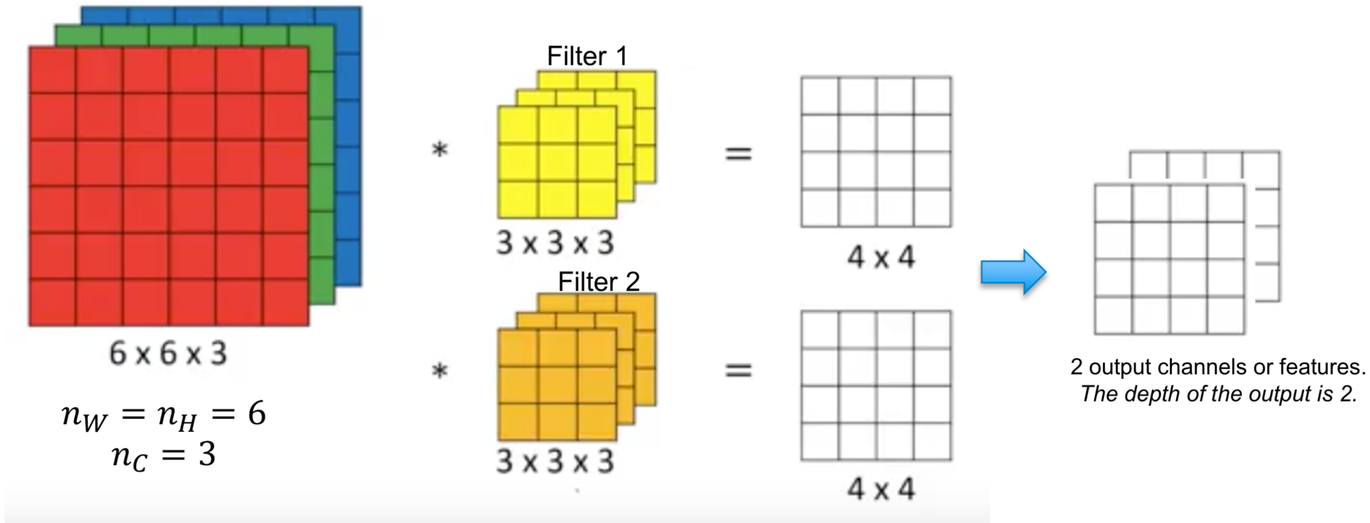
*Each slice of the filter can detect different features in each of the 3 colour channels, for instance vertical edges in Red, horizontal edges in Green and vertical edges in Blue!*

Andrew Ng https://www.youtube.com/watch?v=KTB_OFoAQcc&list=PLkDaE6sCZn6Gl29AoE31iwdVwSG-KnDzF&t=0s&index=7

# Reminder CNNs

Imperial College London

## If Different Filters are Applied …




**Exercise:** *if my input Image is 50x50 with 10 channels, and if I apply 25 filters each of size 3x3x10 with a stride of 1, no padding, what is the size and number of channels of the output image?*

**Answer:** *size is 48x48. with 25 channels*

Andrew Ng https://www.youtube.com/watch?v=KTB_OFoAQcc&list=PLkDaE6sCZn6Gl29AoE31iwdVwSG-KnDzF&t=0s&index=7

# Reminder CNNs

Imperial College London

## Example of Max Pooling Layer (for Downsampling)


4 x 4


2 x 2

Stride $s$ and Filter size $f$ are the main hyperparameters for pooling.

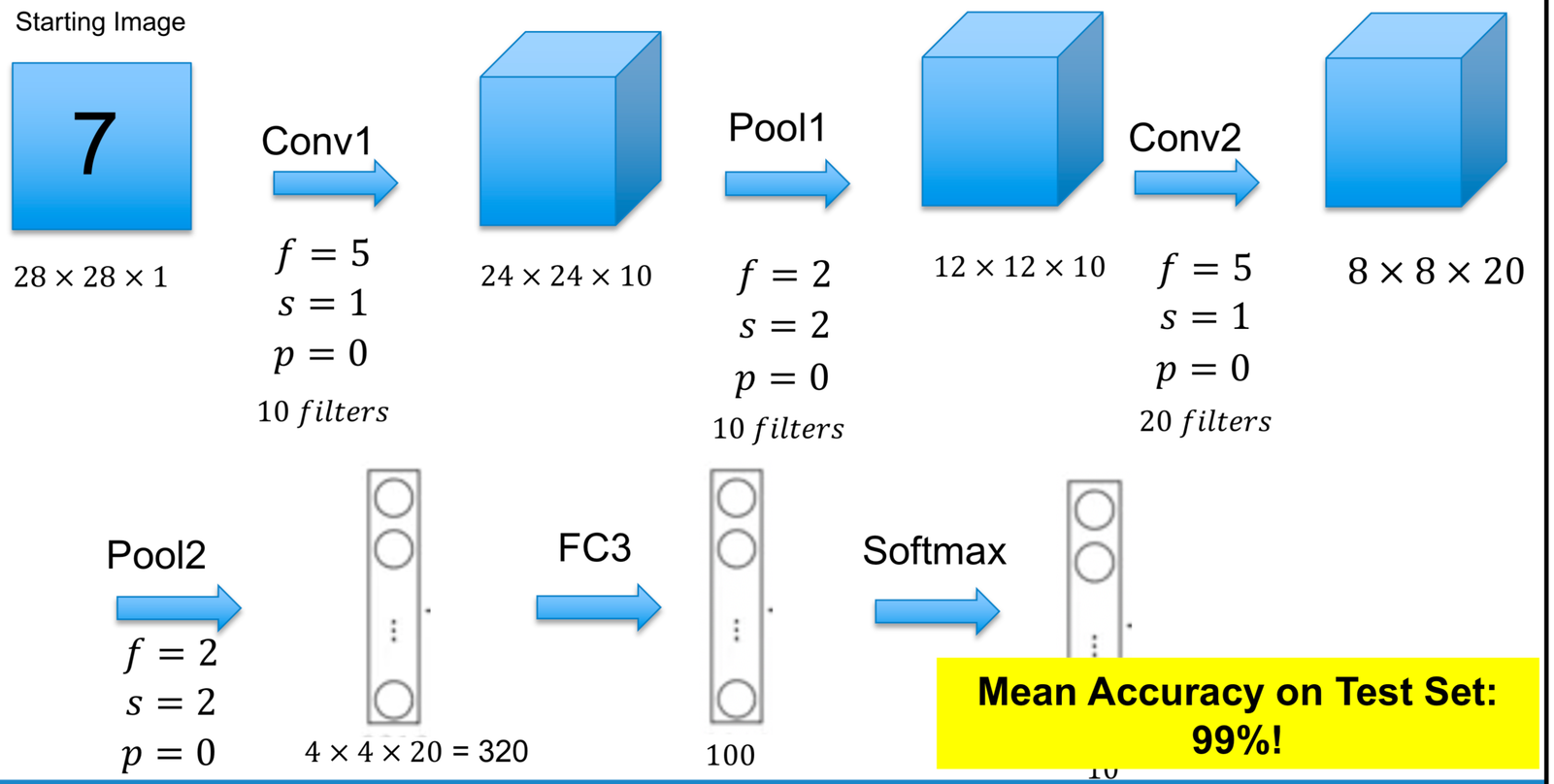
Here $s = 2$ and $f = 2$ Usually $f = s$

*Pooling makes the input representations (feature dimension) smaller and more manageable for the next layer, Max Pooling is used because it may be interesting to keep the high values for the activation of the next layer as they may characterize some important features. Pooling reduces the number of parameters and computations in the network, therefore controlling overfitting.*

# Reminder CNNs

Imperial College London

## MNIST CNN Example (The LeNet-5 Architecture from LeCun)



# Reminder CNNs

Imperial College London

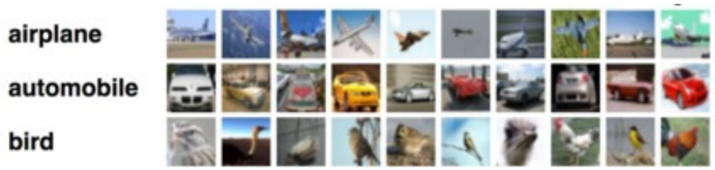
## Some of the Best-Known Datasets

Available on `torchvision.datasets`

***Very deep neural networks work best when trained on very large datasets!***

- **MNIST: Handwritten digits, 60000 Training Images, 10000 Test Images**
- **CIFAR-10 / CIFAR-100: 50k Training, 10k Test Images of 10 (CIFAR-10) or 100 (CIFAR-100) classes**

  **Color Images are 32x32, Task: Classification https://www.cs.toronto.edu/~kriz/cifar.html**



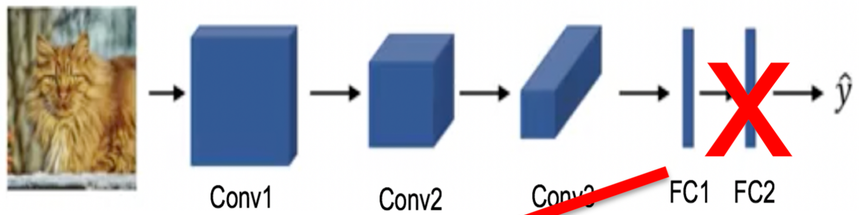

- **Imagenet: > 15 Million Images in 20,000 classes! https://en.wikipedia.org/wiki/ImageNet**

# Reminder CNNs

Imperial College London

## Example of Transfer Learning Approach

Take existing trained network such as Inception v3 model, trained on ImageNet dataset for differentiating between 1,000 different classes of images.




Re-train last layer of the network on images of interest, such as pictures of carbonate cores .

Rudstone Sample

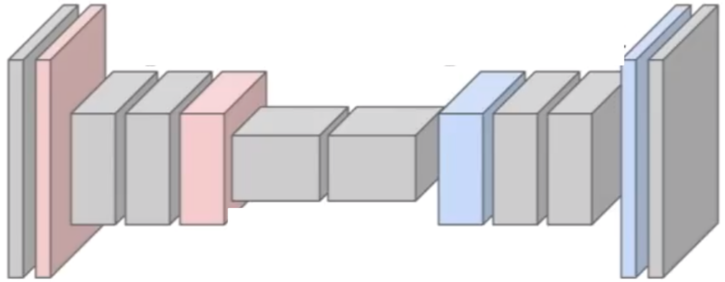
From Sharinia Kanagandran and Cedric John, Imperial College

# Reminder CNNs

Imperial College London

## Semantic Segmentation with Down- and Up-Sampling

Input:
3 x H x W



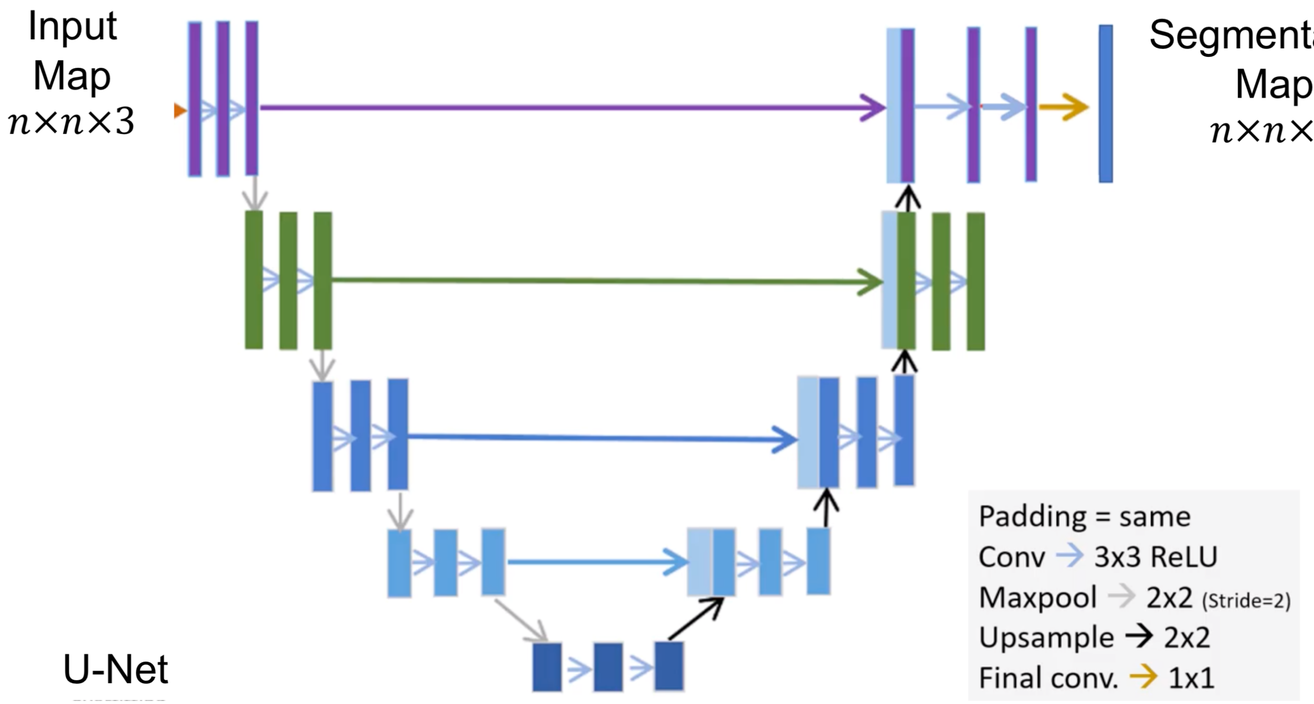
Predictions:
H x W

https://www.youtube.com/watch?v=azM57JuQpQI

# Reminder CNNs

Imperial College London

## Segmentation Model: U-Net (Ronneberger et al, 2015)




U-Net

https://www.youtube.com/watch?v=azM57JuQpQI