Image Classification Using CNNs

Presented by

BENOUAKLIL Hodhaifa

BOUKACEM Younes

Presentation Plan

01 INTRODUCTION

02 WHY CNN?

03 CNN ARCHITECTURE

04 PRACTICAL EXAMPLE

01 INTRODUCTION

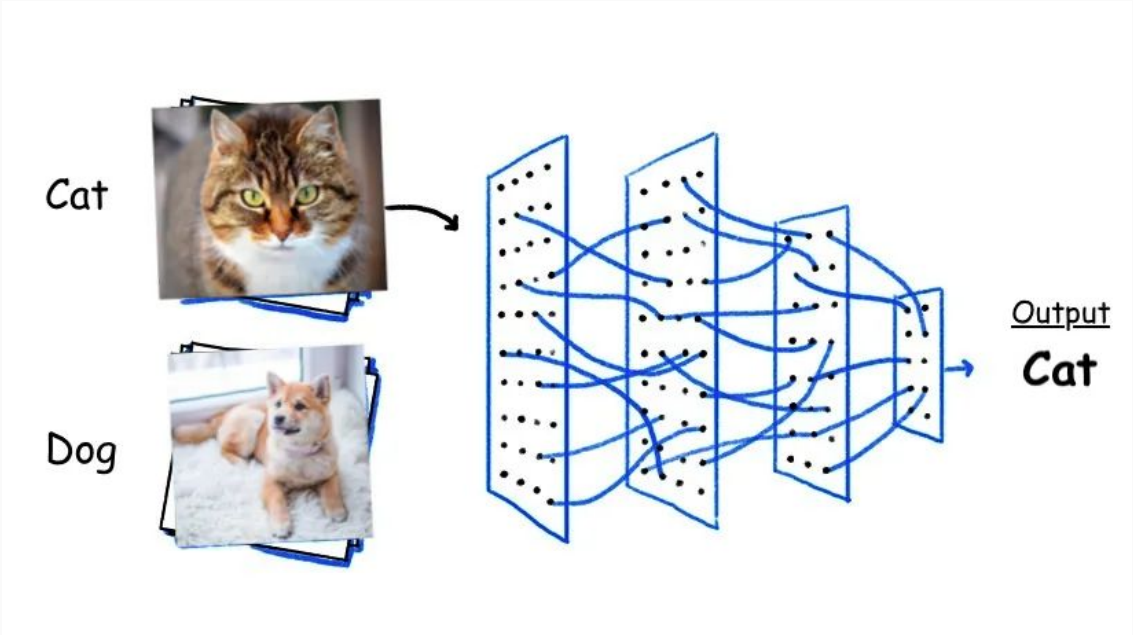
What is Image Classification?

*“ Image classification is a **computer vision** task where the goal is to **categorize** an entire image into a specific **class** or **label**”*

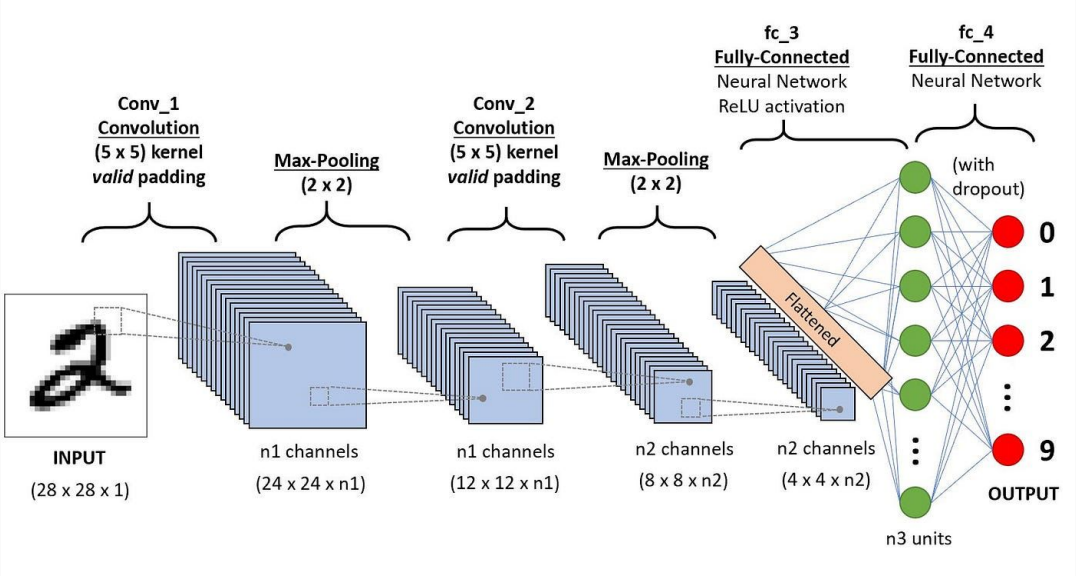
For Example

- Identifying whether an image contains a cat or a dog
- Recognizing handwritten digits
- Classifying medical images (e.g., detecting tumors)
- Identifying plant species from leaf images

Example



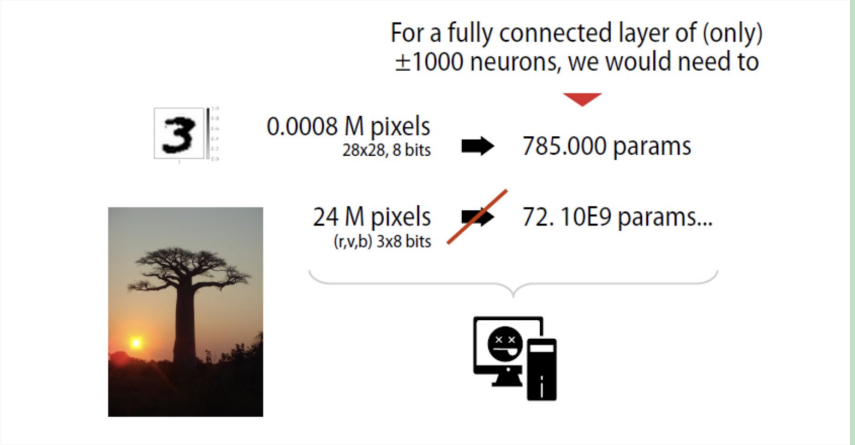
What are Convolutional Neural Networks (CNNs)?



02 WHY CNN?

Limitations of Traditional Neural Networks

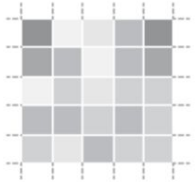
- Spatial information loss
- High computational complexity
- Large number of parameters
- Poor scaling with image size

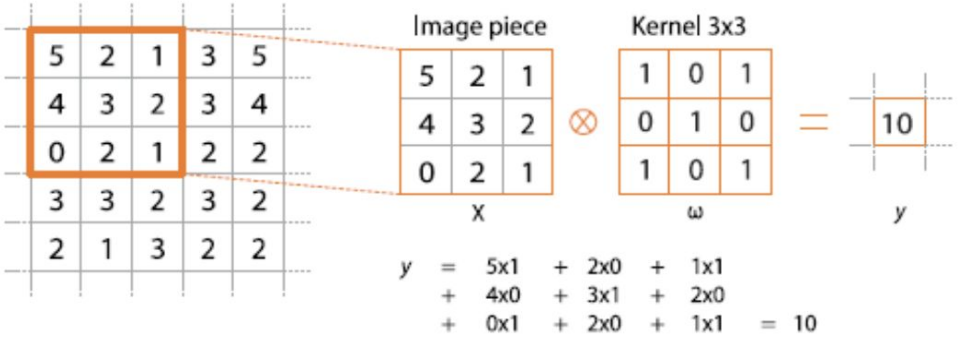


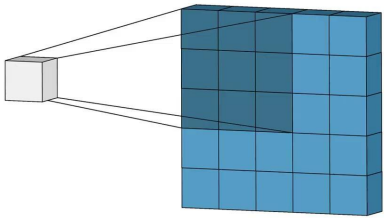
03 CNN ARCHITECTURE

Convolution Operation

By Jan Kroon, from Pexels.com

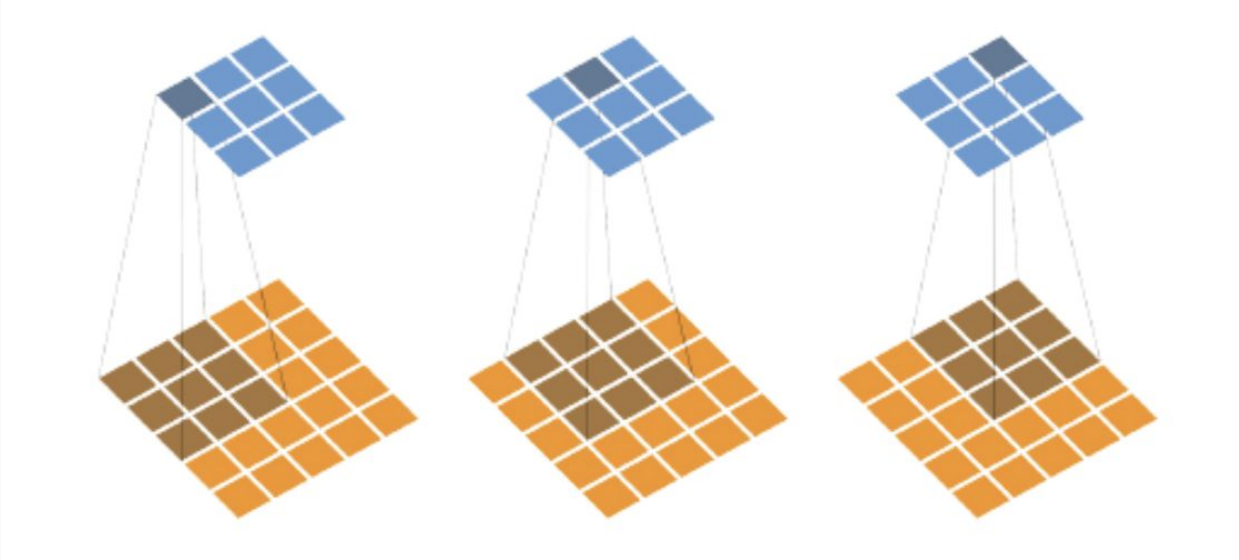






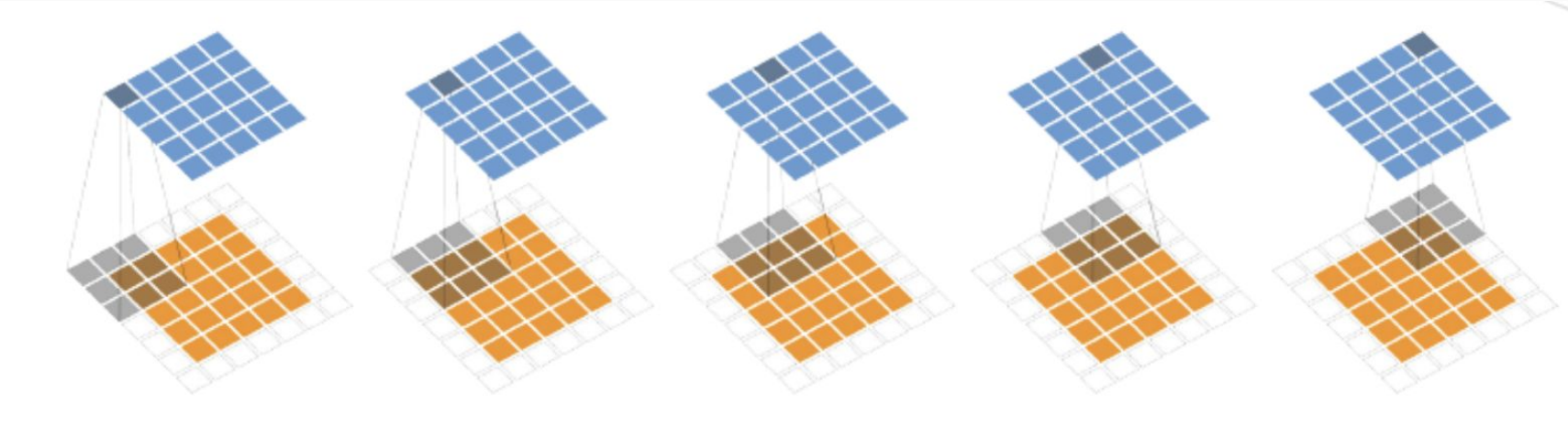
Padding

Valid padding (no padding):



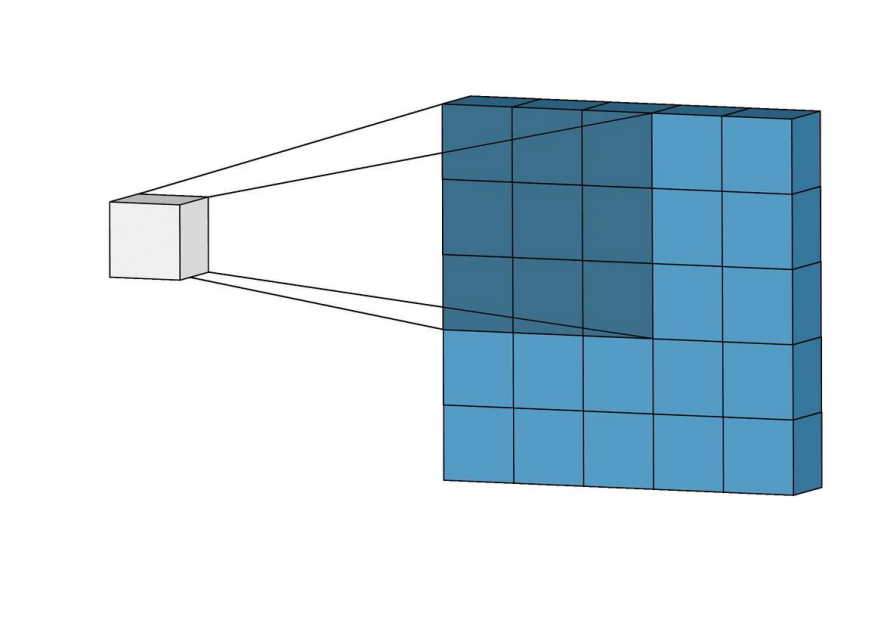
Padding

Same padding (output size = input size):



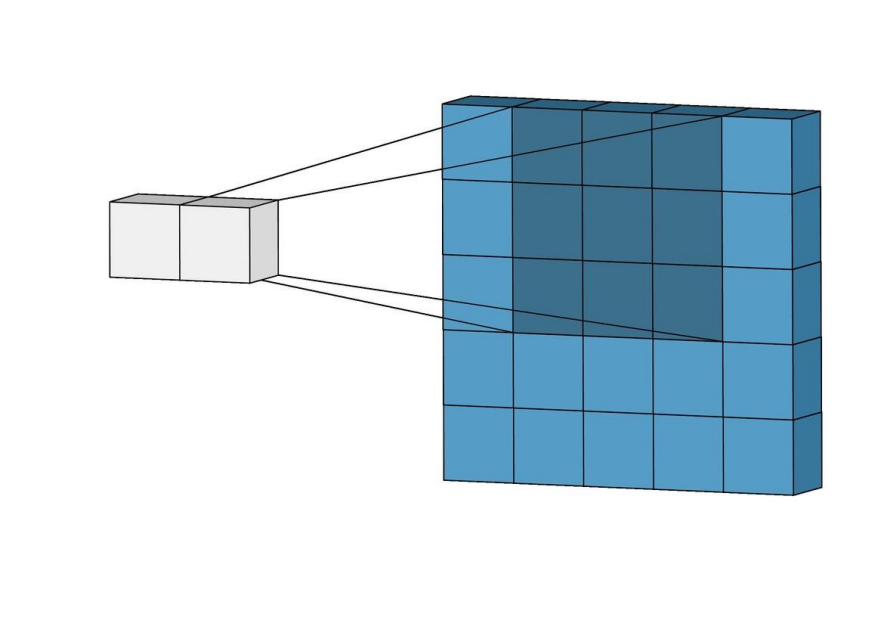
Stride

Stride = 1:



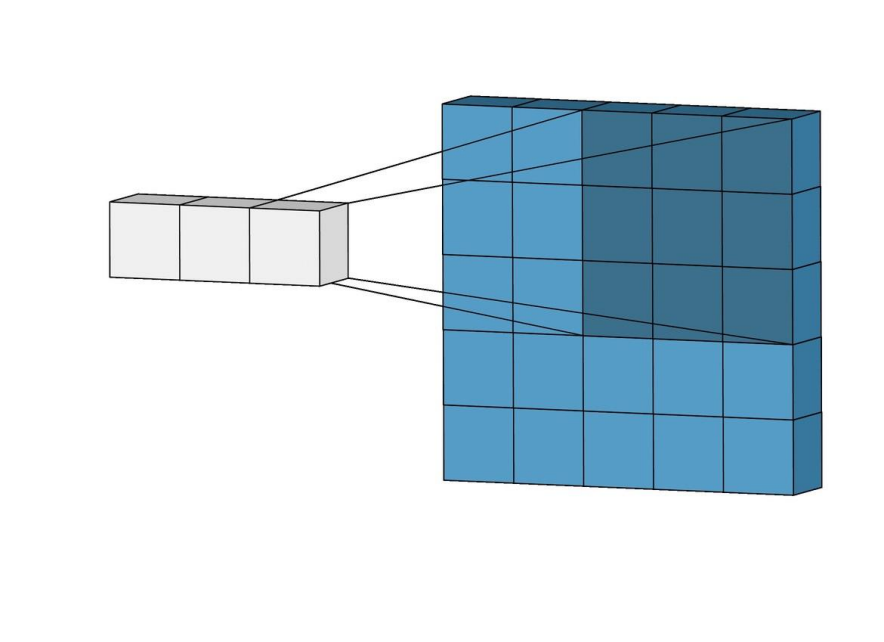
Stride

Stride = 1:



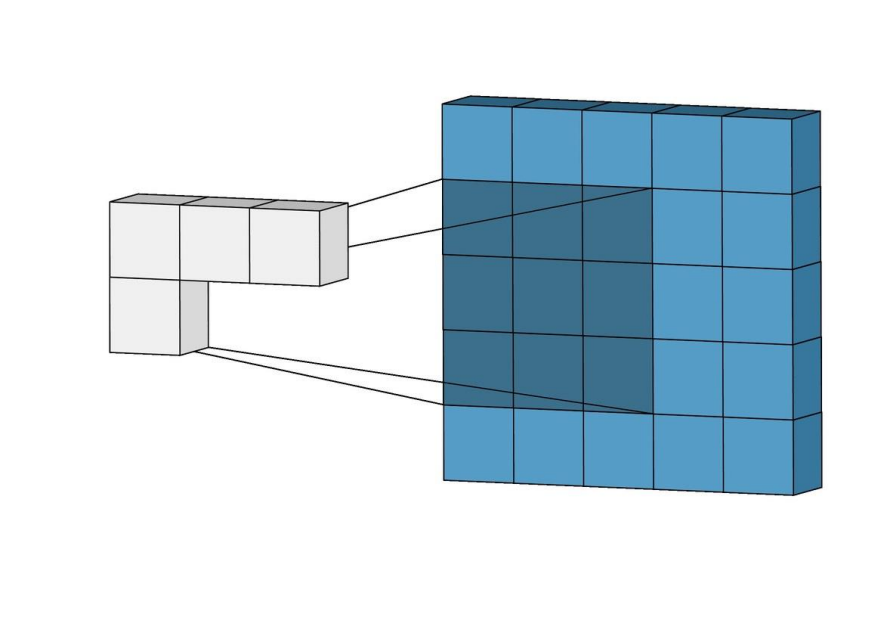
Stride

Stride = 1:



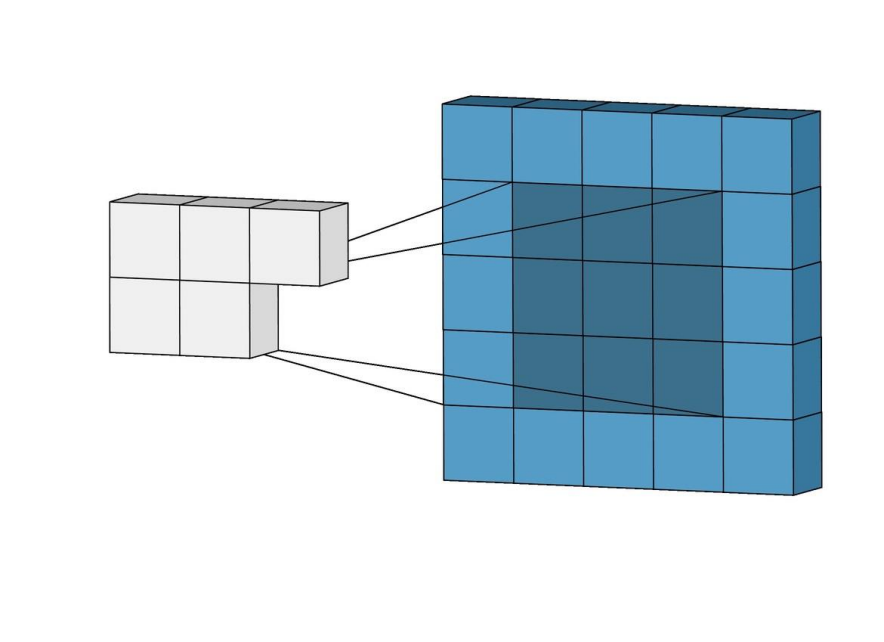
Stride

Stride = 1:



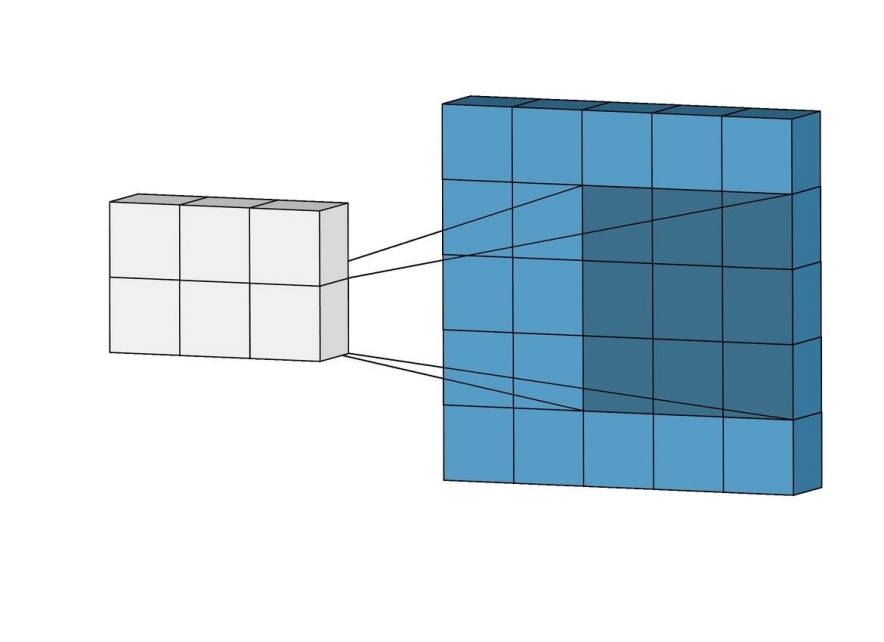
Stride

Stride = 1:



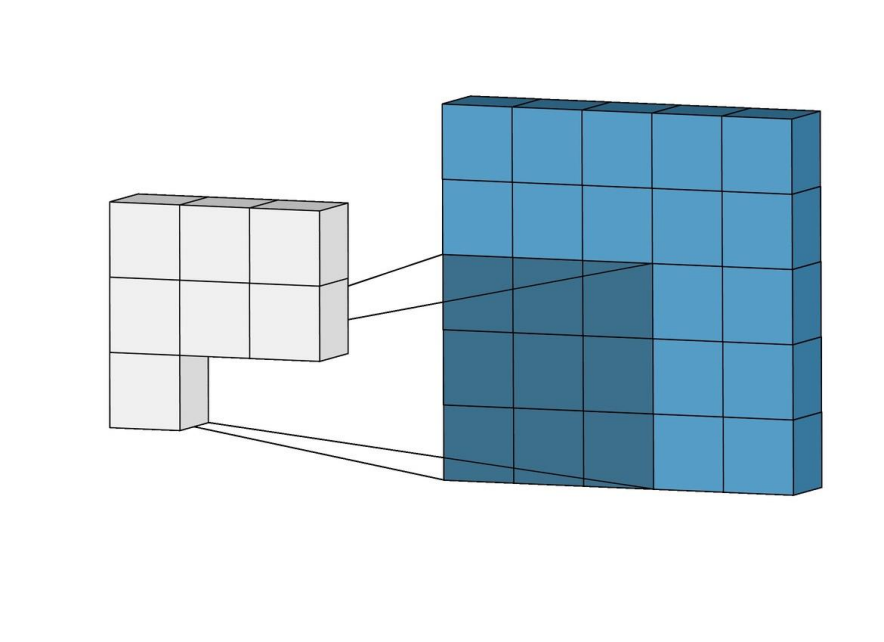
Stride

Stride = 1:



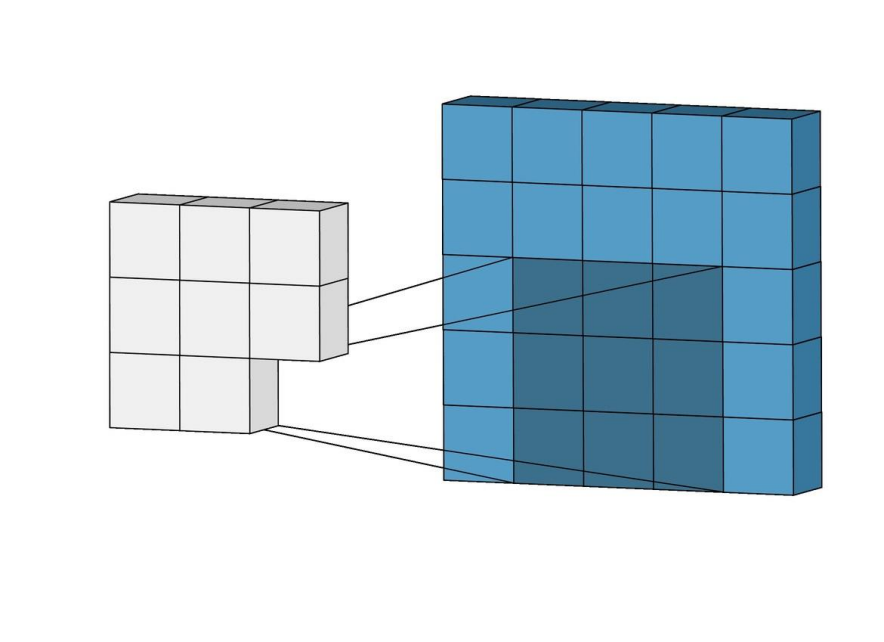
Stride

Stride = 1:



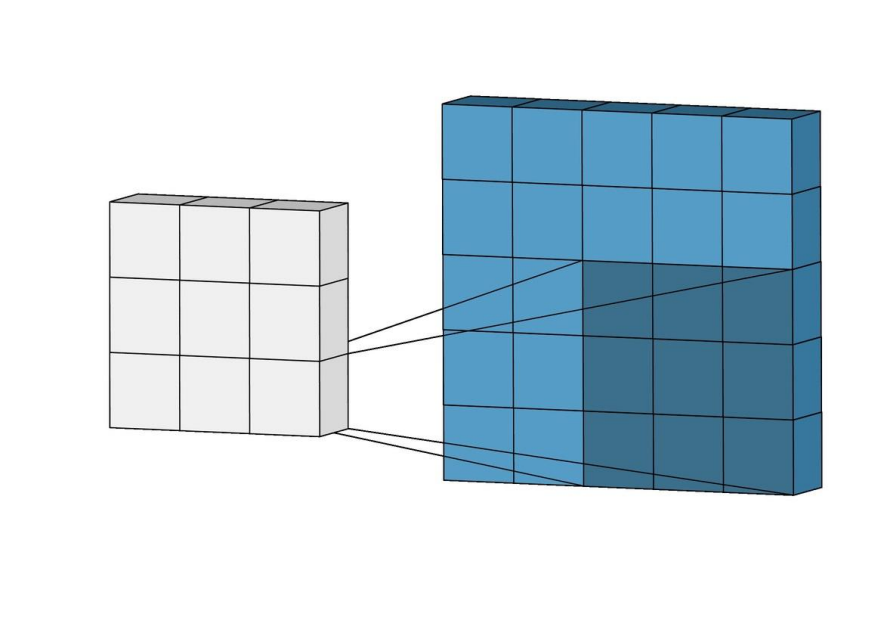
Stride

Stride = 1:



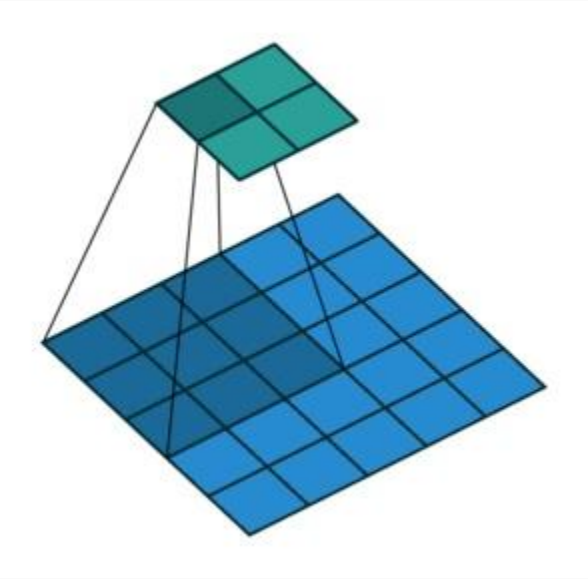
Stride

Stride = 1:



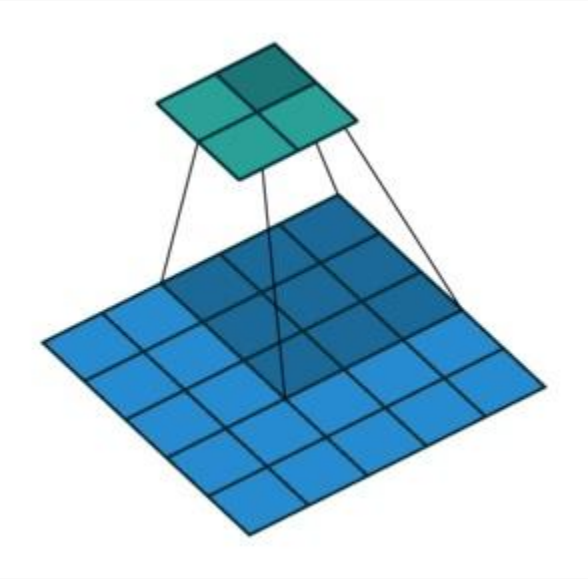
Stride

Stride = 2:



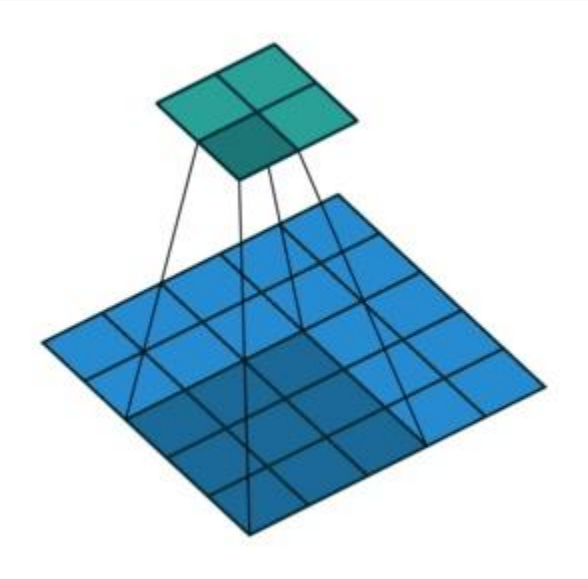
Stride

Stride = 2:



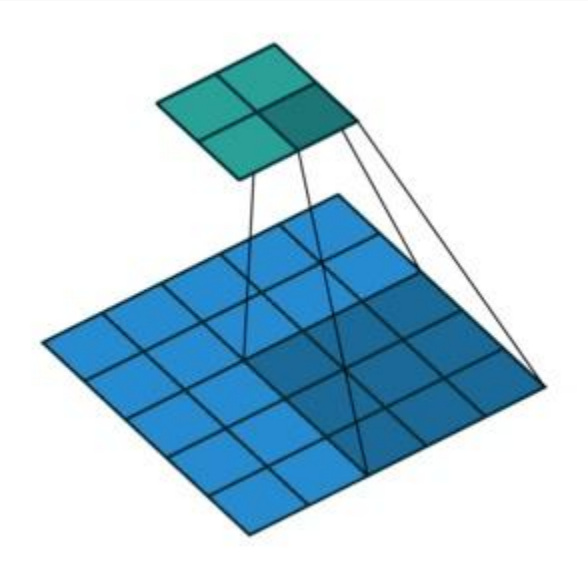
Stride

Stride = 2:



Stride

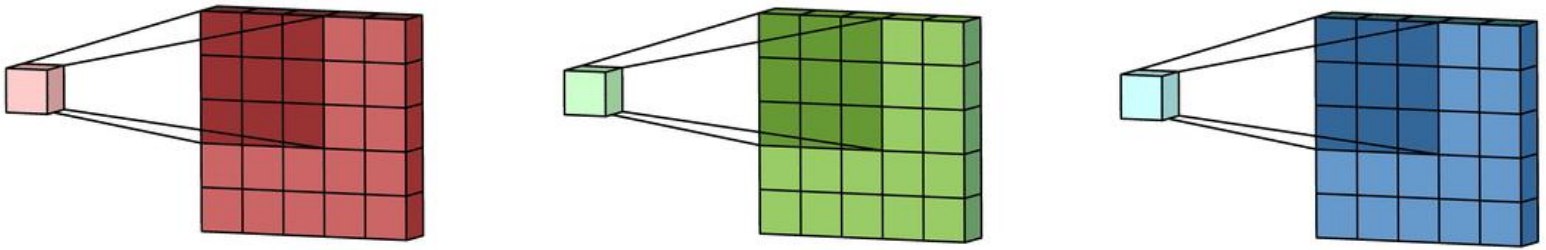
Stride = 2:



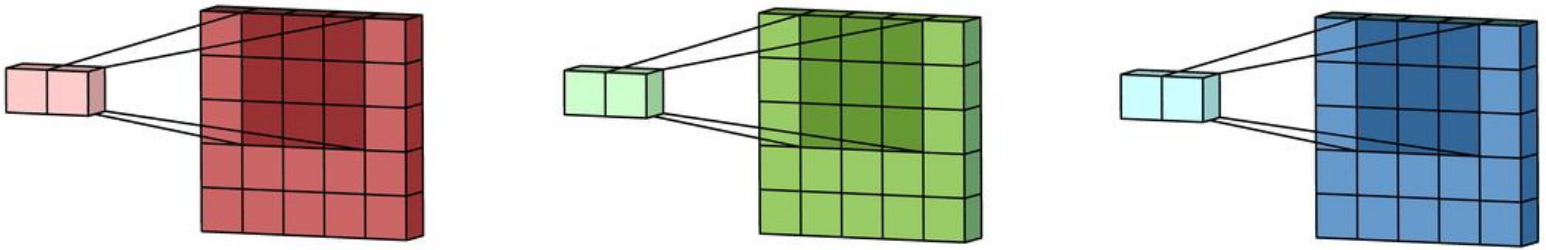
Stride

- Step size for filter movement
- Controls output size
- Impact on feature detection:
  - Larger stride → Less detail / faster training
  - Smaller stride → More detail / slower training

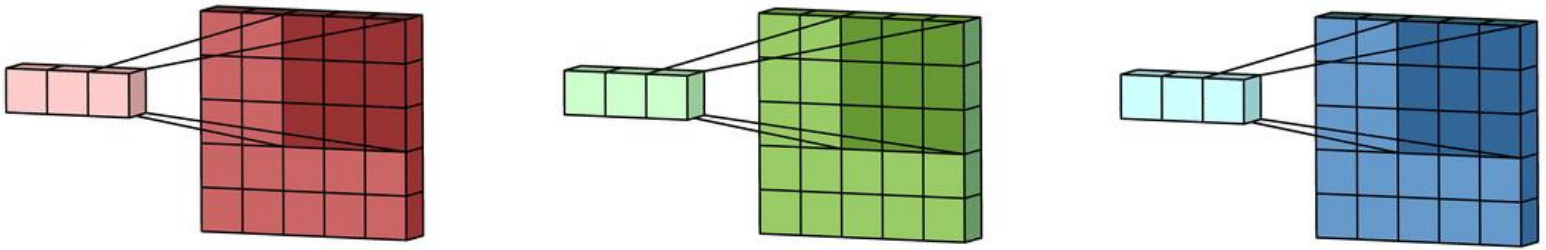
Multi-channel (RGB for ex)



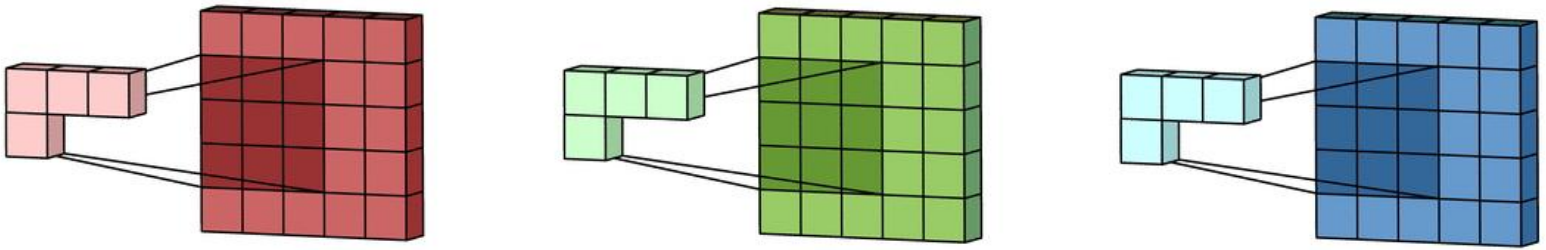
Multi-channel (RGB for ex)



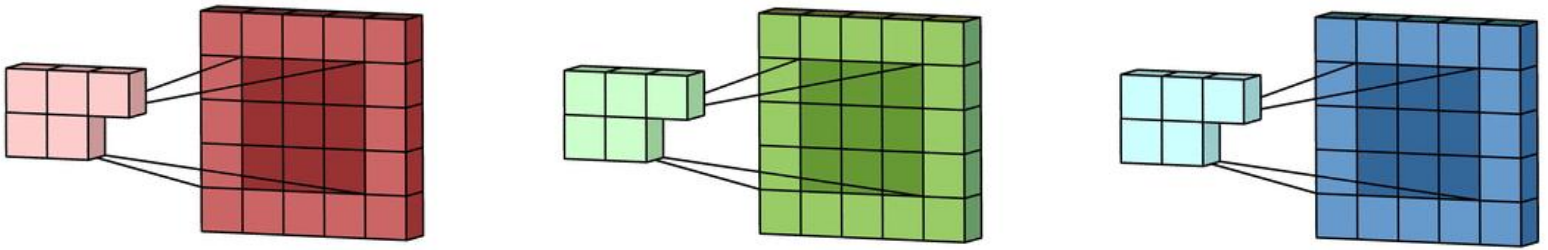
Multi-channel (RGB for ex)



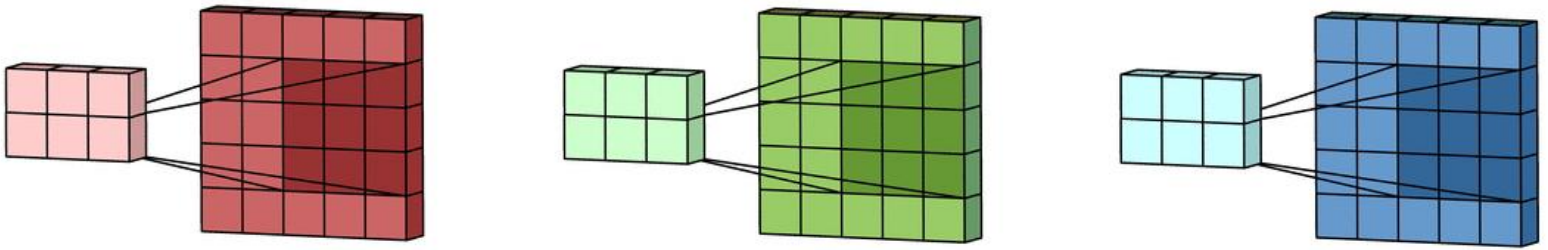
Multi-channel (RGB for ex)



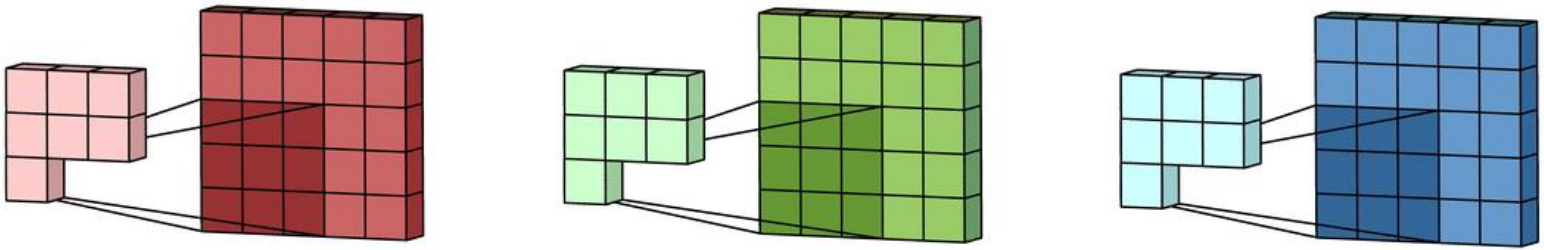
Multi-channel (RGB for ex)



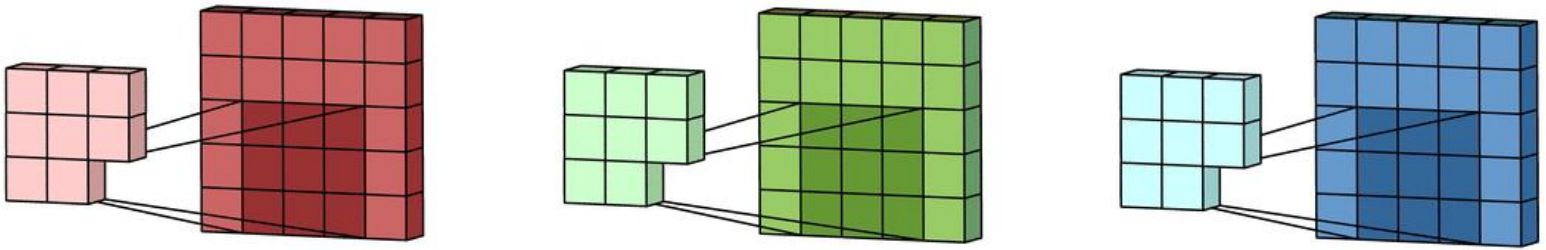
Multi-channel (RGB for ex)



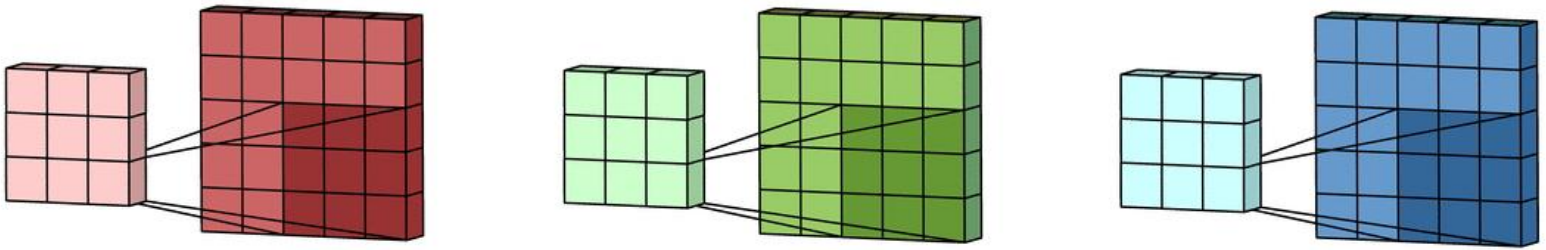
Multi-channel (RGB for ex)



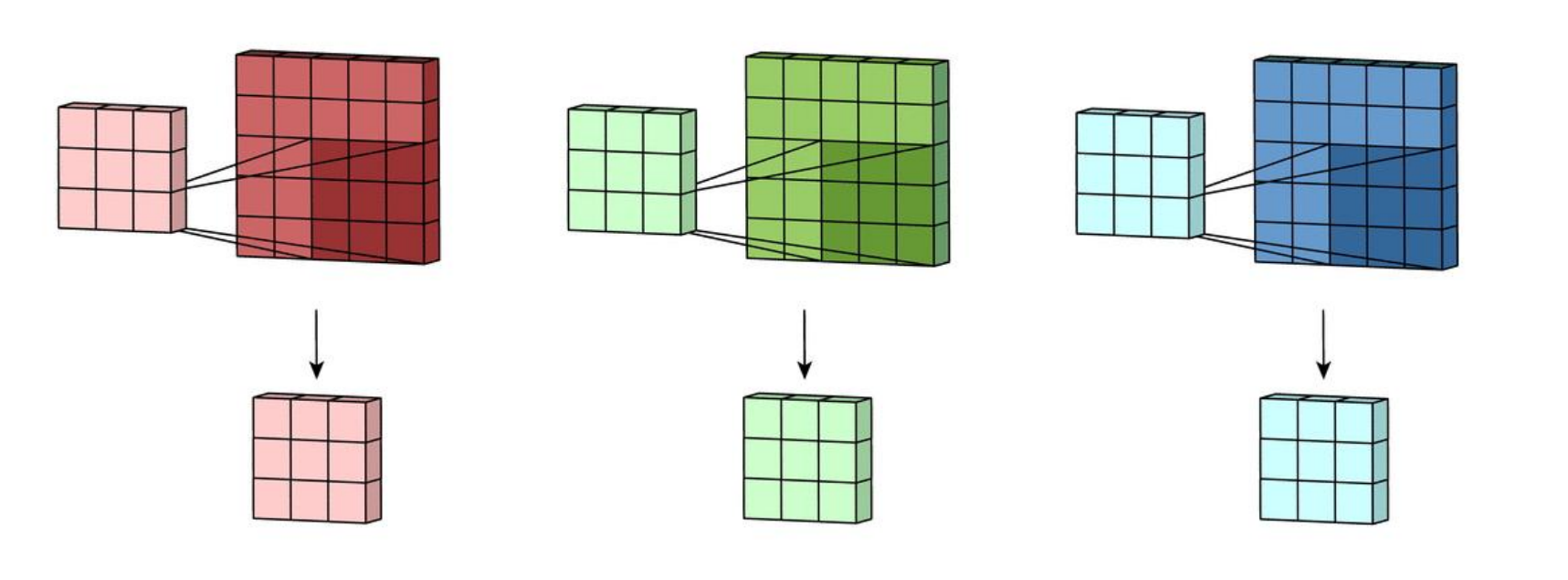
Multi-channel (RGB for ex)



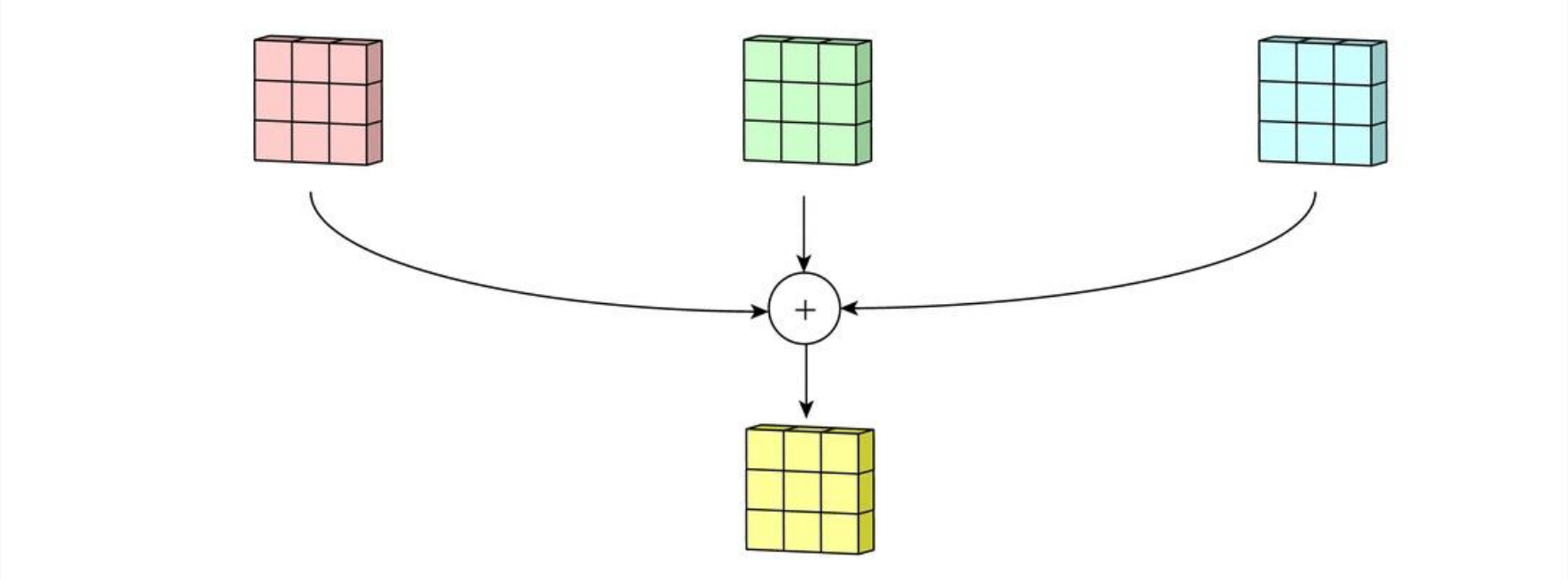
Multi-channel (RGB for ex)



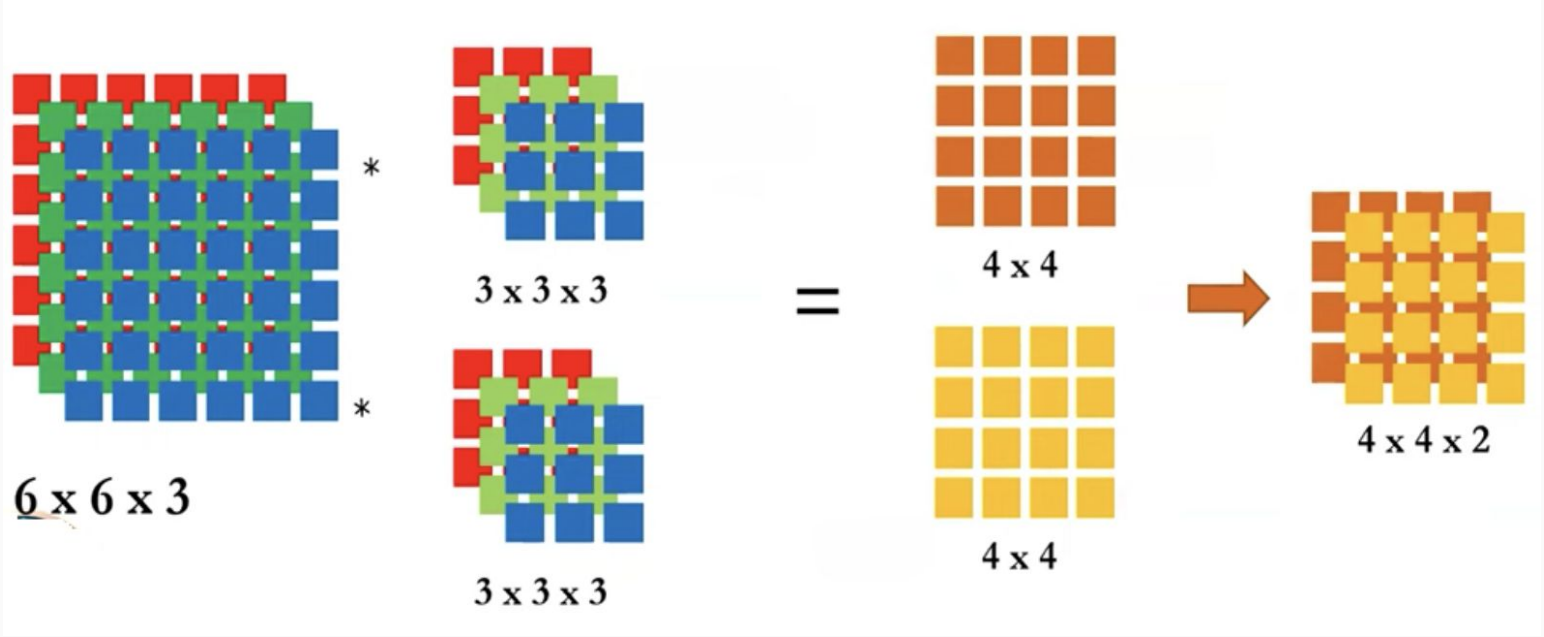
Multi-channel (RGB for ex)



Multi-channel (RGB for ex)



Multiple Filters

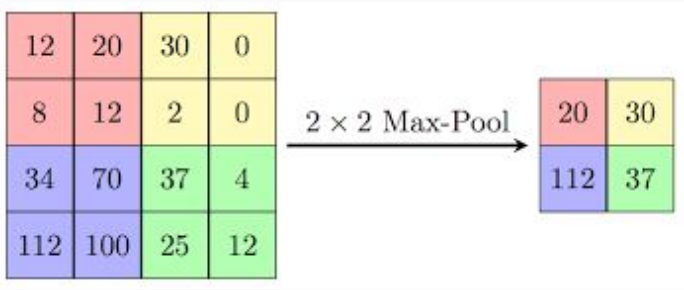


Multiple Filters

- Each filter detects different features
- Creates multiple feature maps
- Parallel processing
- Feature diversity

Pooling Layers

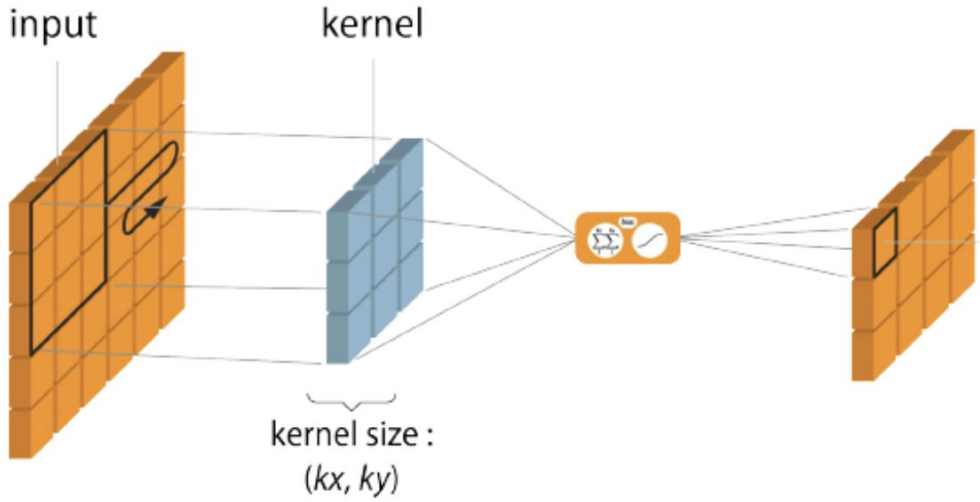
1. Max Pooling
   - Takes maximum value
   - Preserves strong features
2. Average Pooling
   - Takes average value
   - Preserves feature distribution



Pooling Layers

- Dimensionality reduction
- Computational efficiency
- Spatial invariance

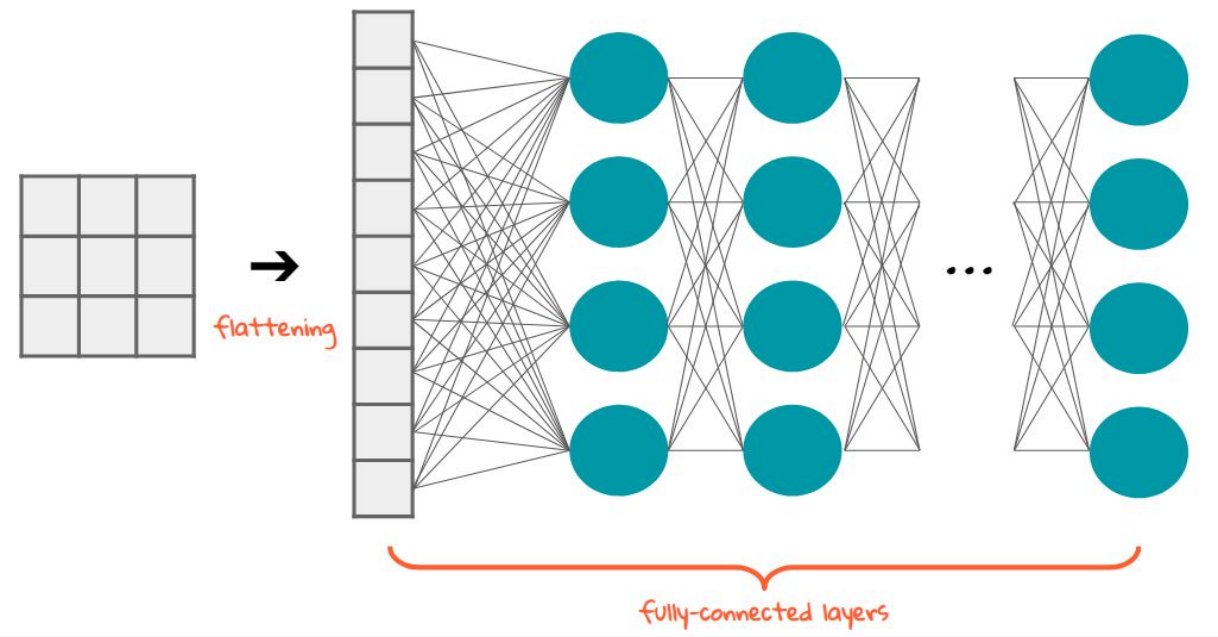
Activation functions



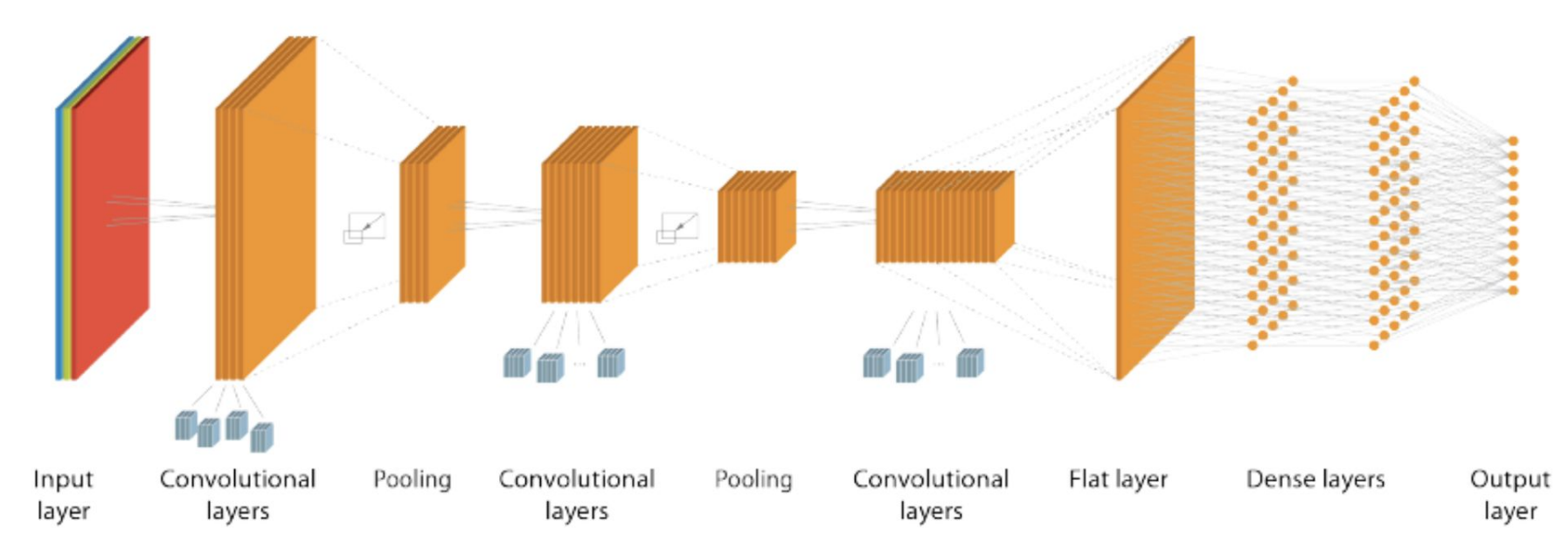
$$y = \sigma(\sum \text{input} \otimes \text{kernel} + b)$$

Where :
σ : is an activation function
b : is a bias

Fully Connected Layer



Typical CNN Structure



05

PRACTICAL EXAMPLE

MNIST Classification

Notebook available here:
https://www.kaggle.com/code/hodhaifabenouaklil/mnist-classification

THANKS FOR YOUR ATTENTION

Do you have any questions?