Convolutional Neural Networks

A Gentle Introduction

Roberto Souza
Assistant Professor
Electrical and Computer Engineering
Schulich School of Engineering

February 2021

UNIVERSITY OF CALGARY

Outline

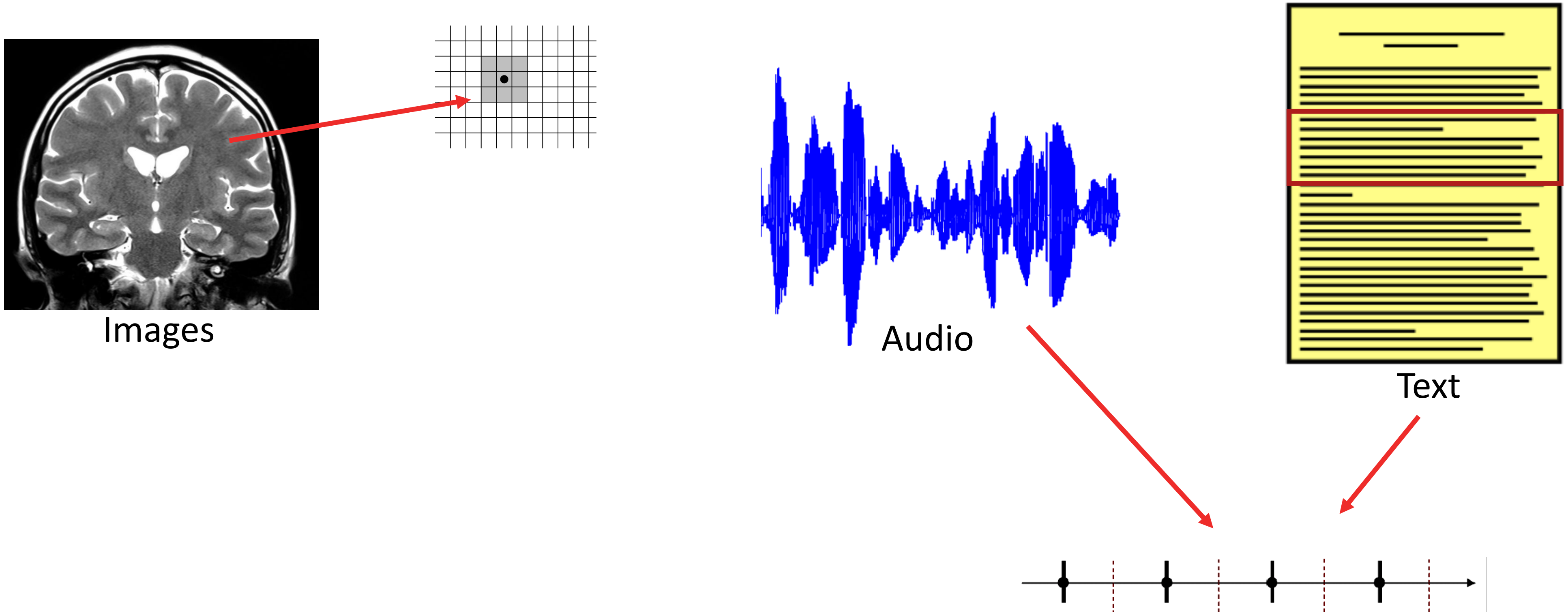
- Learning Goals
- Convolutional Neural Networks (CNNs)
  - Basic Operations
  - Properties
  - Computing number of parameters
- Summary

UNIVERSITY OF CALGARY

Learning Goals

- Understand how CNNs work and when to apply them
- Compute the number of parameters of your model

Data – Euclidean Domains

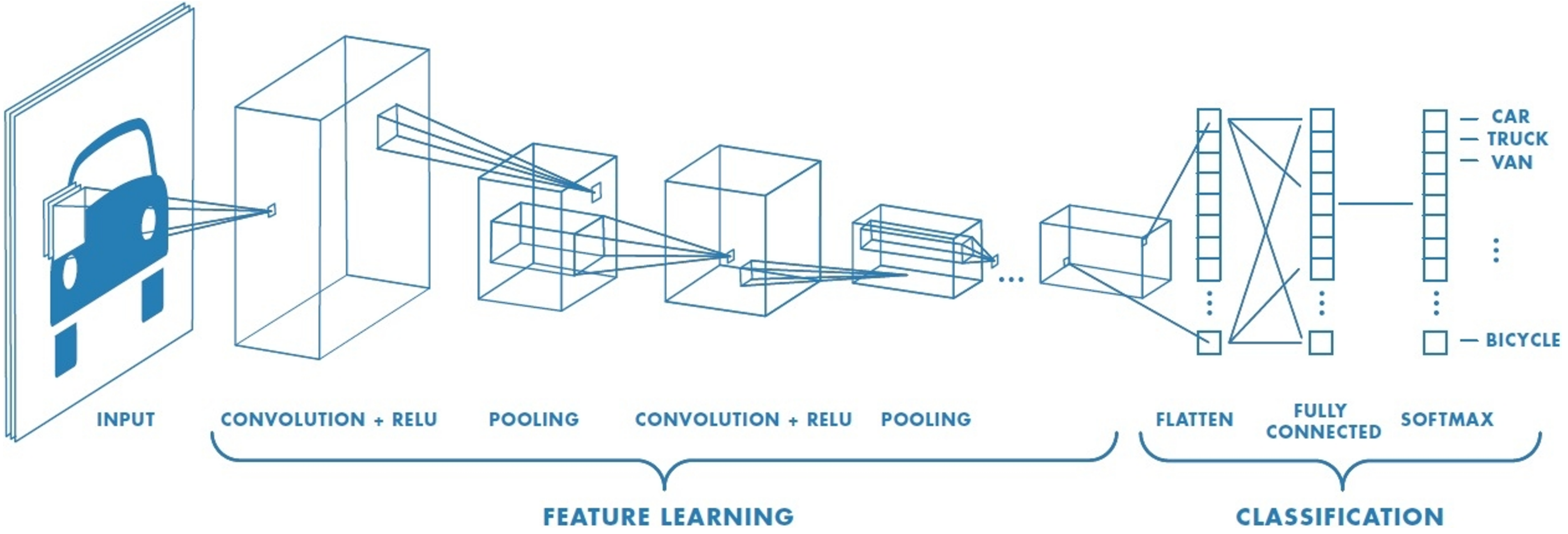


- Images, audio, text among others all have regular structures in a Euclidean space
    - Convolutions are well-defined operations that can be computed efficiently in these structures

UNIVERSITY OF CALGARY

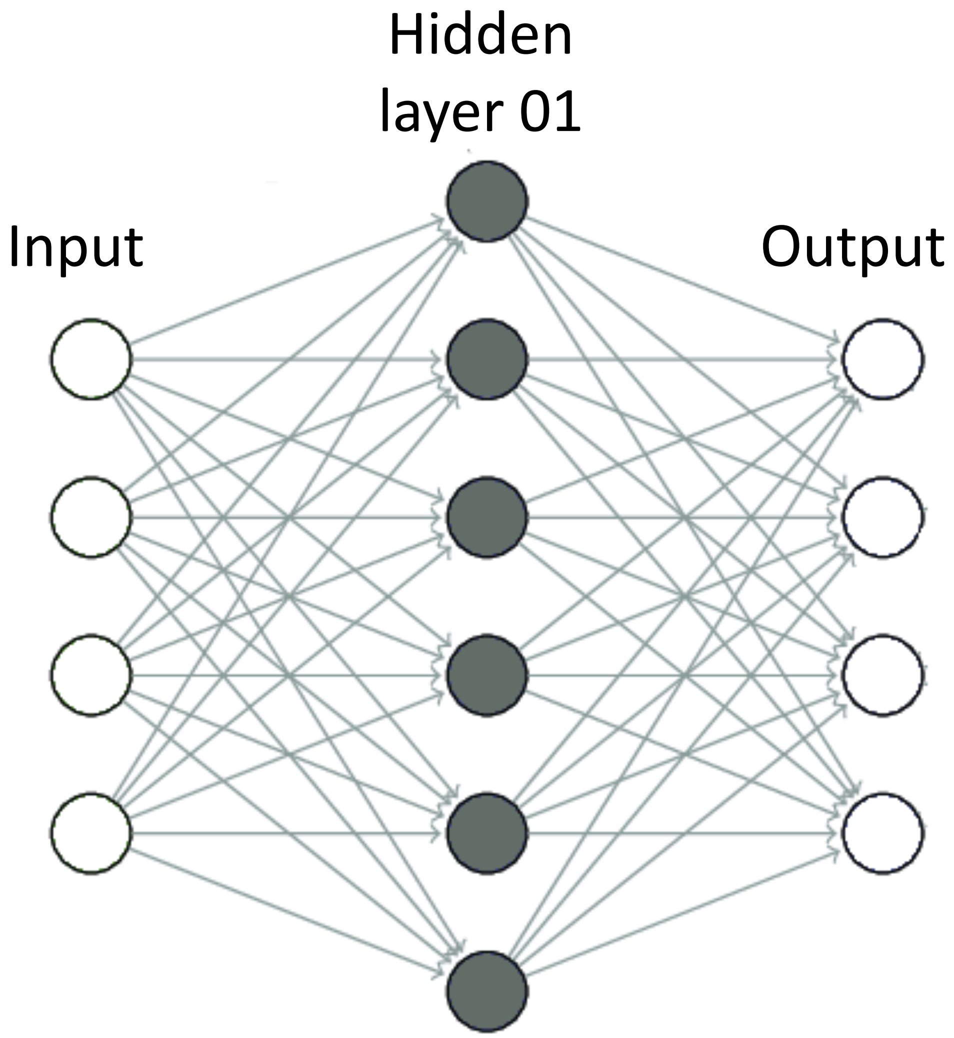
Convolutional Neural Networks

- Convolutional layers learn features
- Connected layers perform classification
- Fewer trainable parameters than fully connected networks



UNIVERSITY OF CALGARY

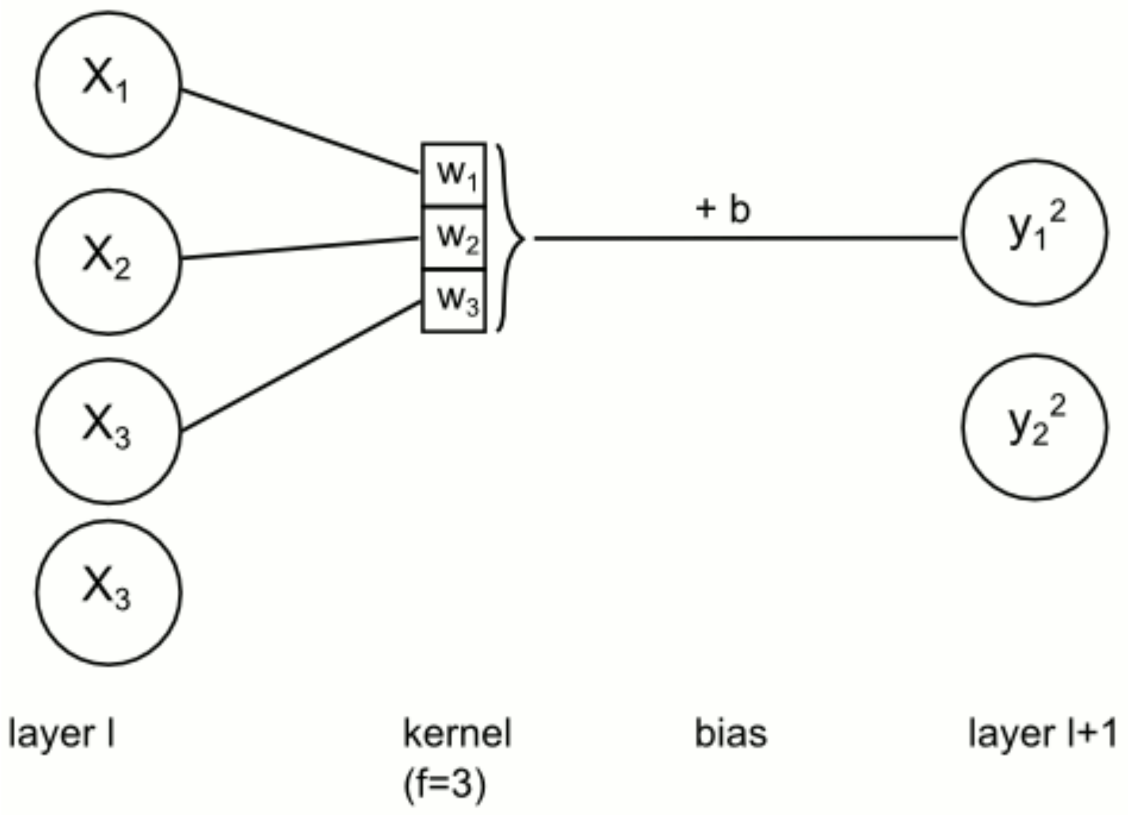
Fully Connected Neural Network – Global Property



- The fully connected layer can lead to an explosion in the number of parameters

- Imagine your input is a 256 x 256 image and your layer has 10 outputs, how many parameters would the model have?
  - 256x256x10 + 10 = 655,370

Convolutional Neural Network – Weight Sharing



- Convolutional neural networks share weights across inputs (*i.e.*, connection **sparsity**)
- Convolutions leverage local correlations (*i.e.*, **locality**)
- Imagine your input is a 256 x 256 image, your convolution size is 3 and your layer has 10 filters, how many parameters would the model have?
  - 3 x 10 + 10 = 40

UNIVERSITY OF CALGARY

Convolution (single-channel input)

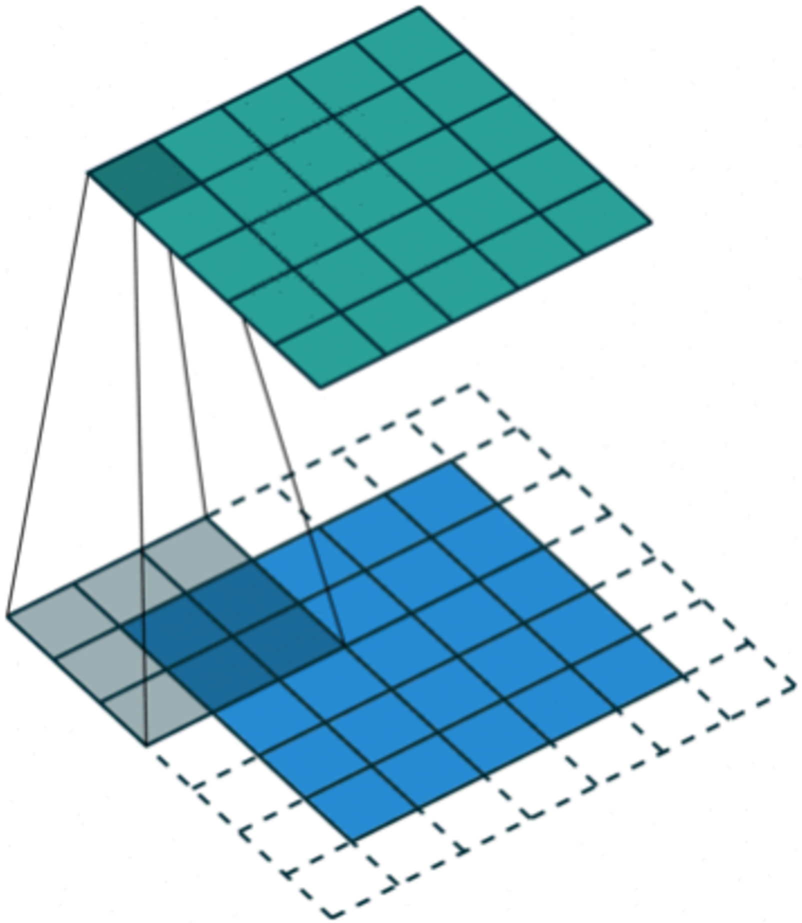
1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

4		

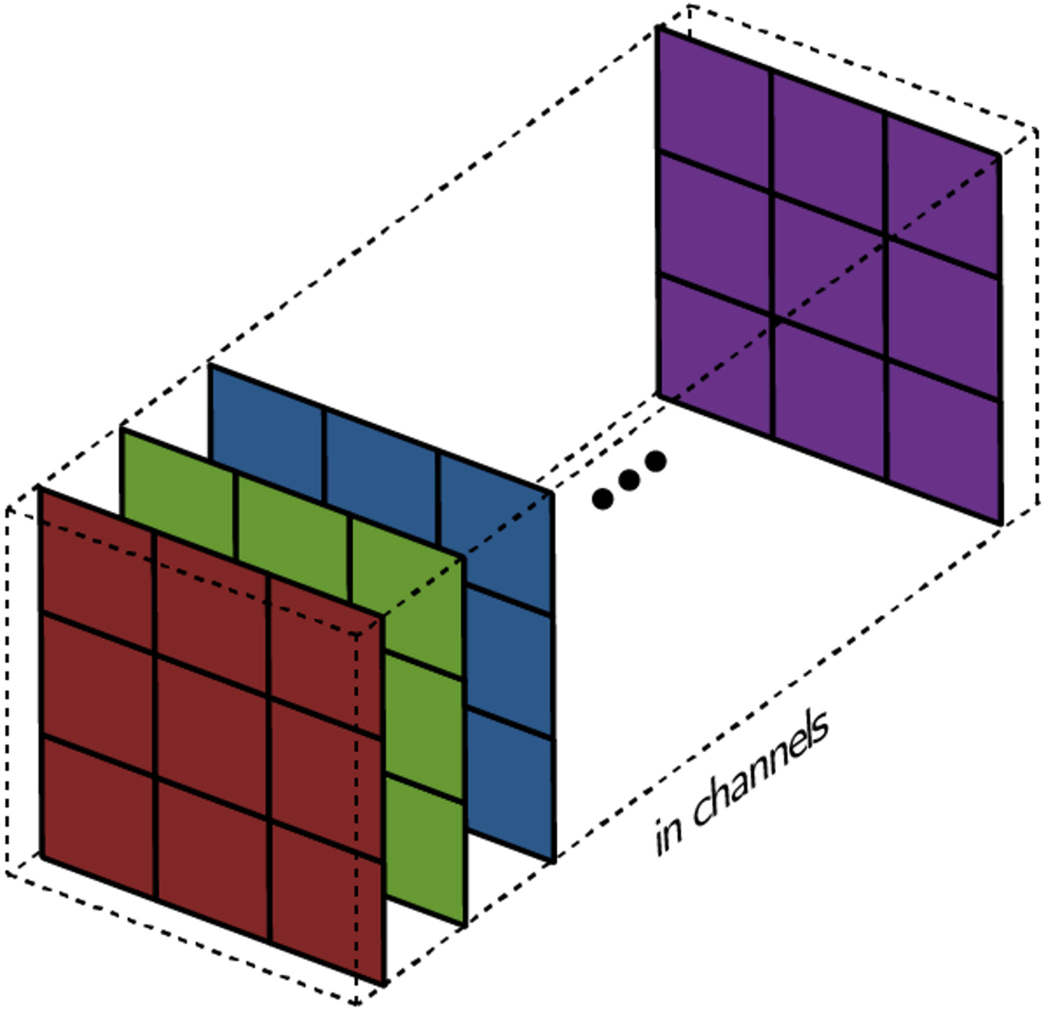
The bias term is added after the convolution

Convolution (padding)

- Image can be padded prior to convolution to preserve its dimensions

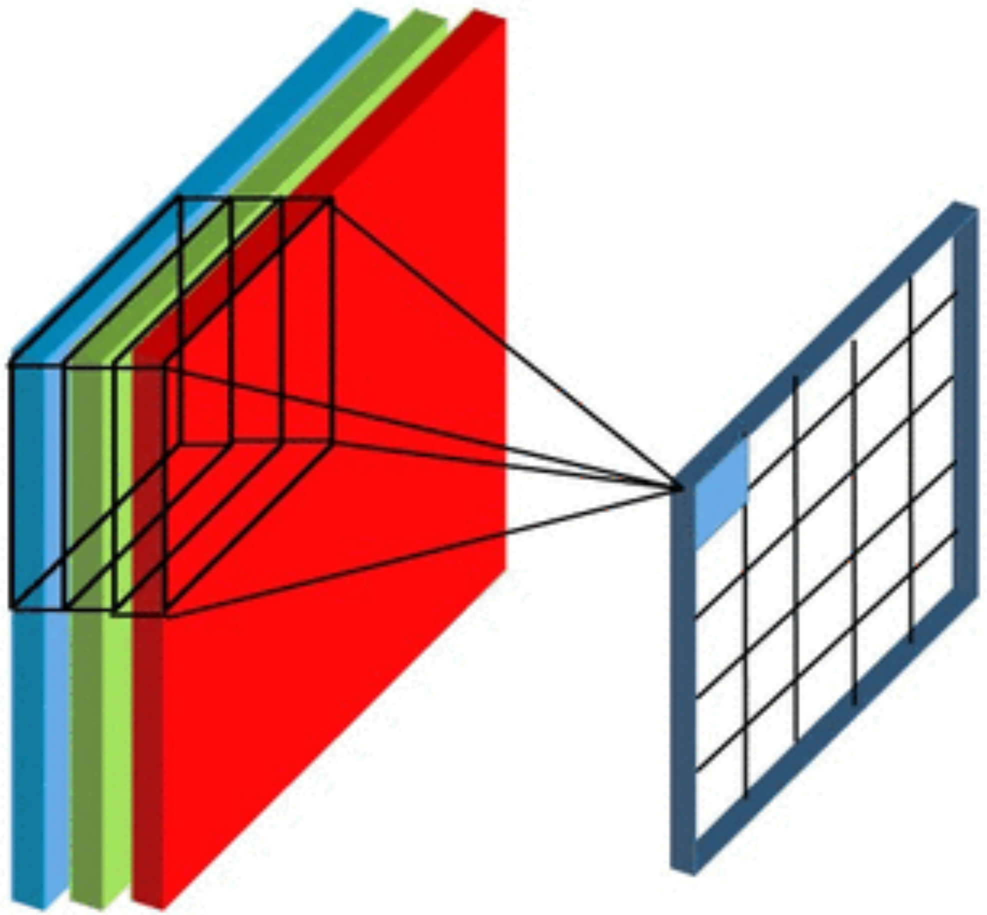


N-channel Images/Signals



- Results of convolutions are stacked resulting in n-channel images/signals

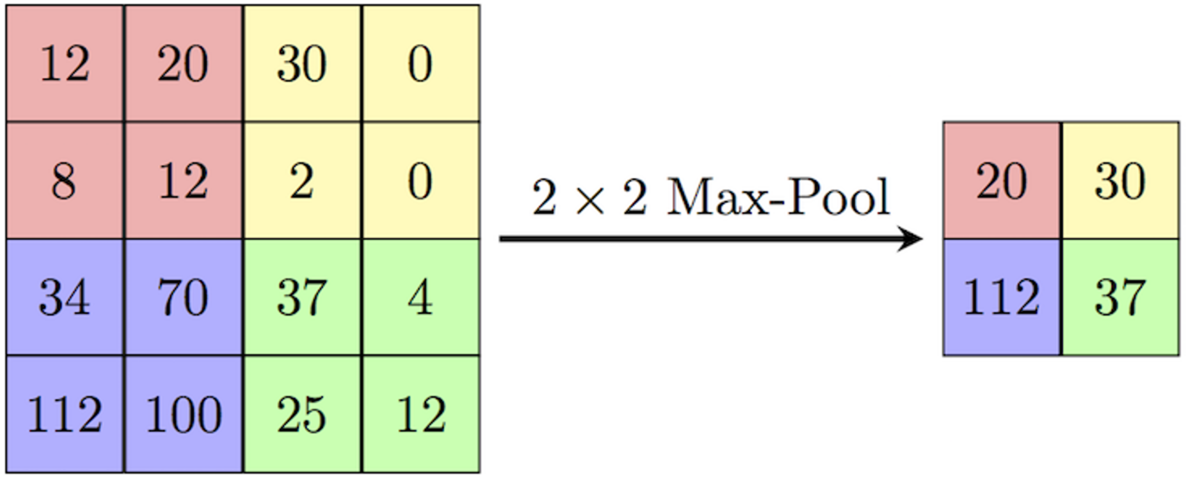
Convolution – multi-channel input



- The convolutions encompass the channels of the input
- A W1 x W2 convolution is actually a W1 x W2 x nchannels convolution

Max-Pooling

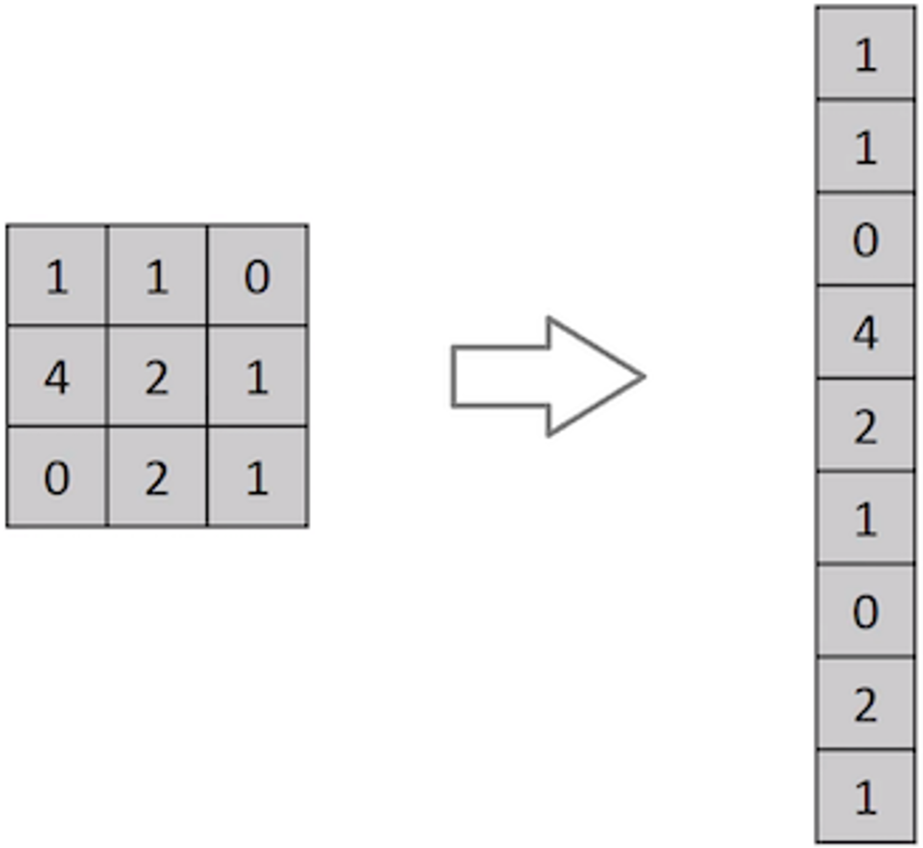
- Non-linear operation
- Reduce dimensionality and computational cost
- After a max-pooling, the number of filters in the subsequent layer is increased



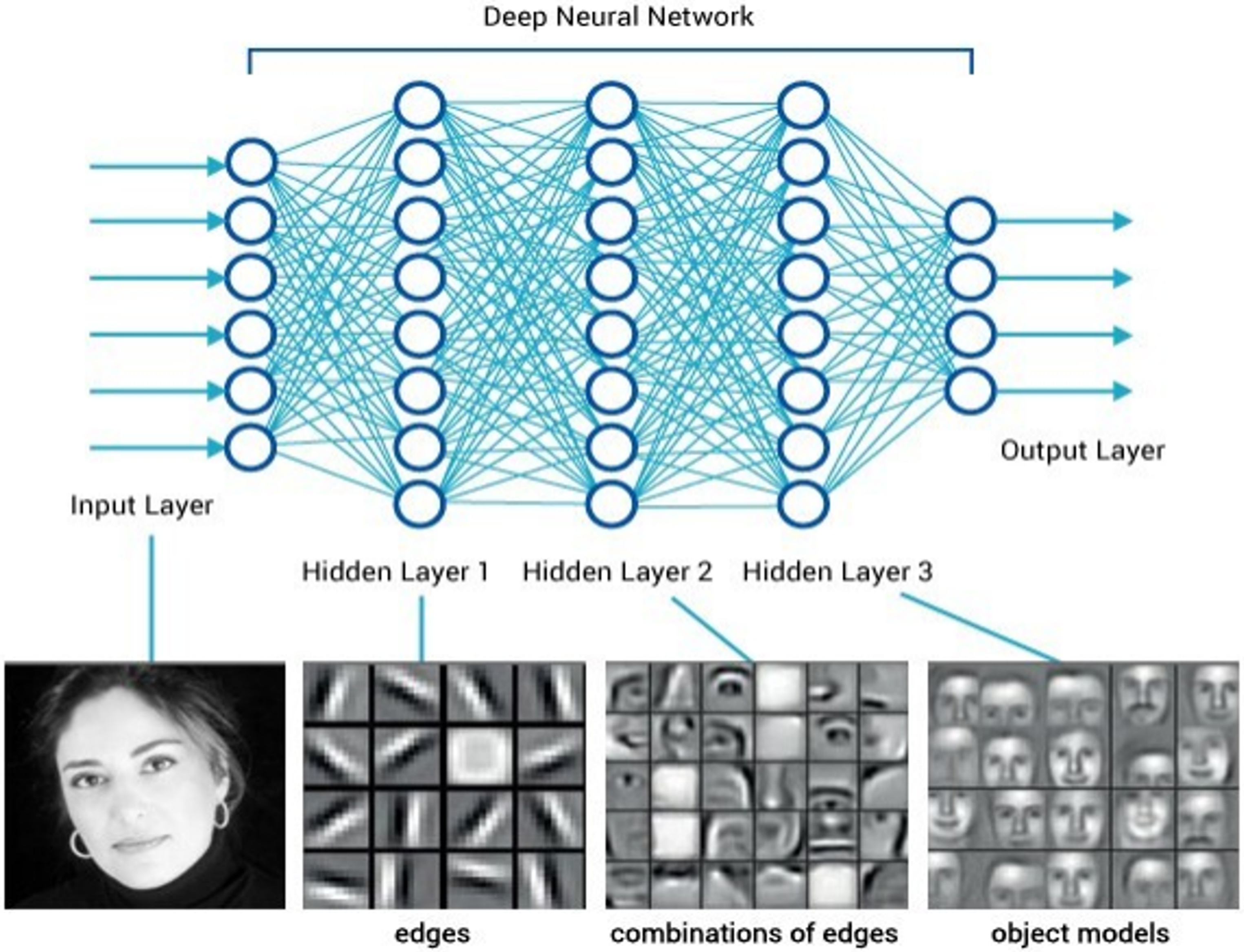
UNIVERSITY OF CALGARY

Flatten

- The flatten operation is applied before the fully connected layer

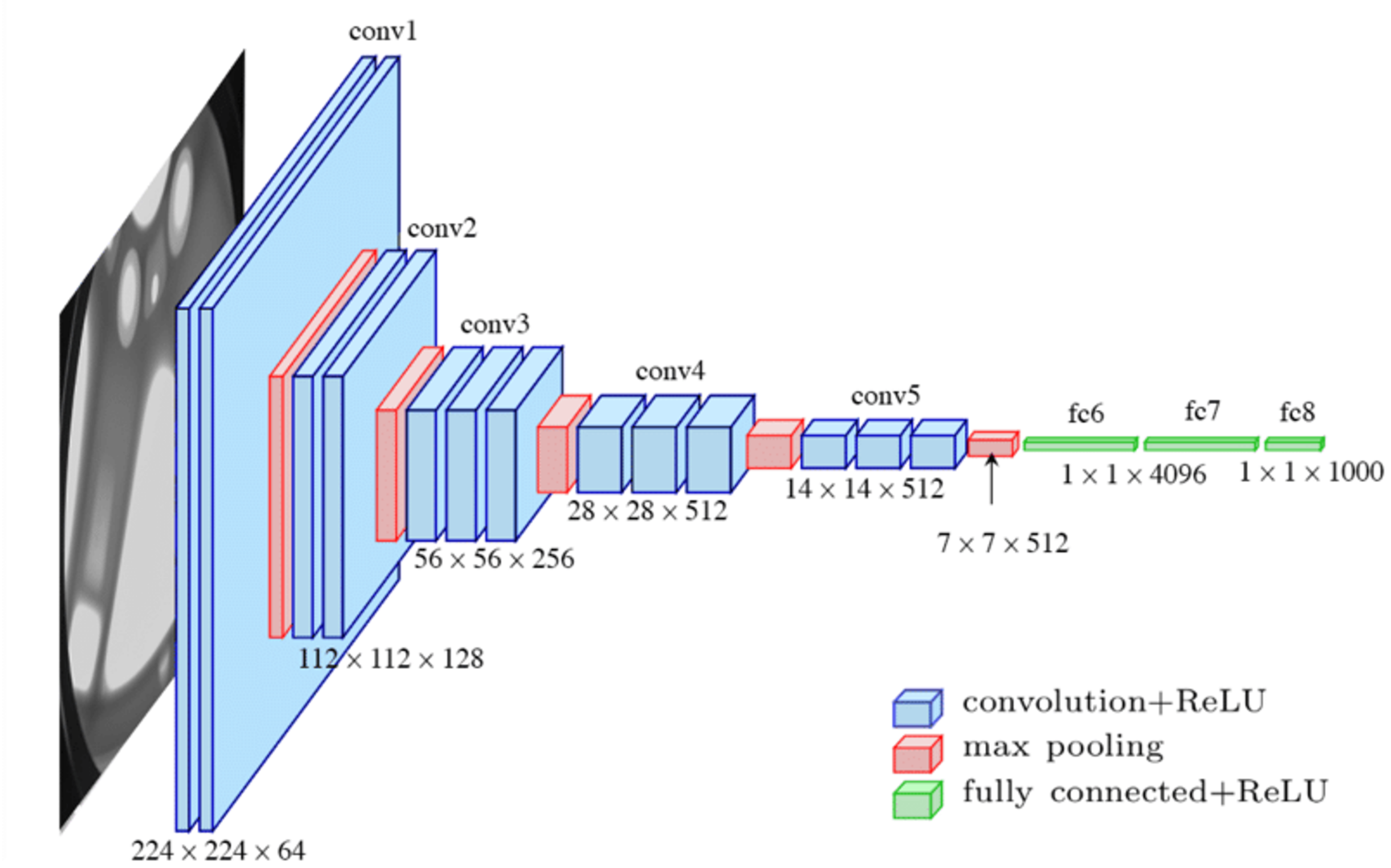


CNN Hierarchy of Concepts



UNIVERSITY OF CALGARY

VGG- 16 Architecture



Summary

- CNNs share weights and have sparse connections
- They depend on local correlations to operate
- The basic operations are convolutions and max-pooling layers
- Implicit hierarchy of concepts

Thank you!

UNIVERSITY OF CALGARY