Disclaimer

If a student turns on their microphone or camera or uses the public chat feature, this constitutes consent for the student's video image or sound audio to be uploaded with the office hour or tutorial on university approved platforms such as D2L. If the student wishes to ensure that their questions/faces/voices are not recorded in the video, they should instead use private chat to ask questions.

Convolutional Neural Networks

A Gentle Introduction

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

January 2022

UNIVERSITY OF CALGARY

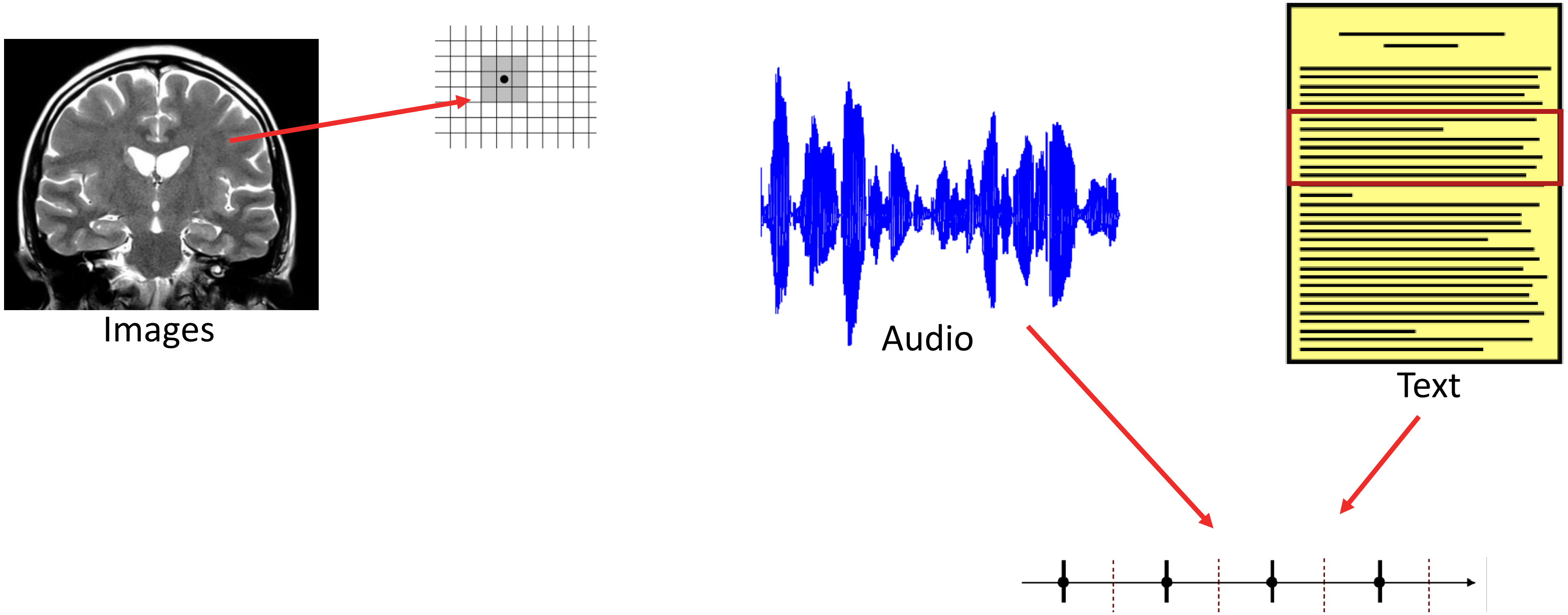
Outline

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

Learning Goals

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

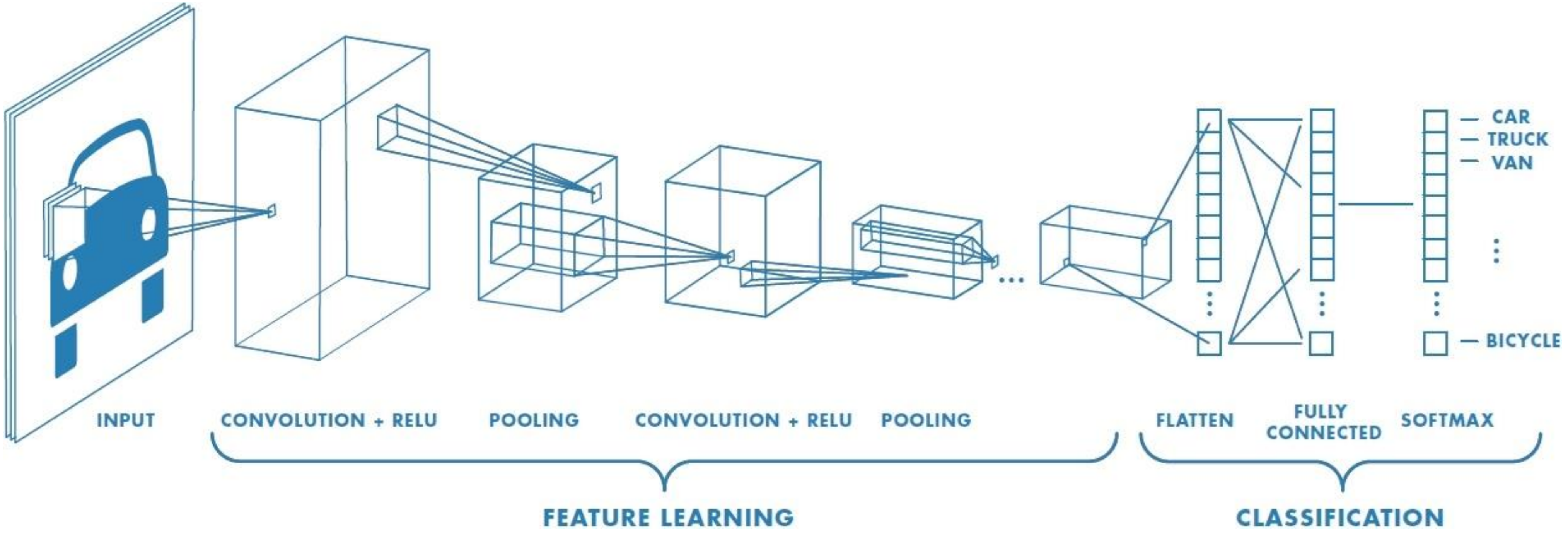
Data – Euclidean Domains



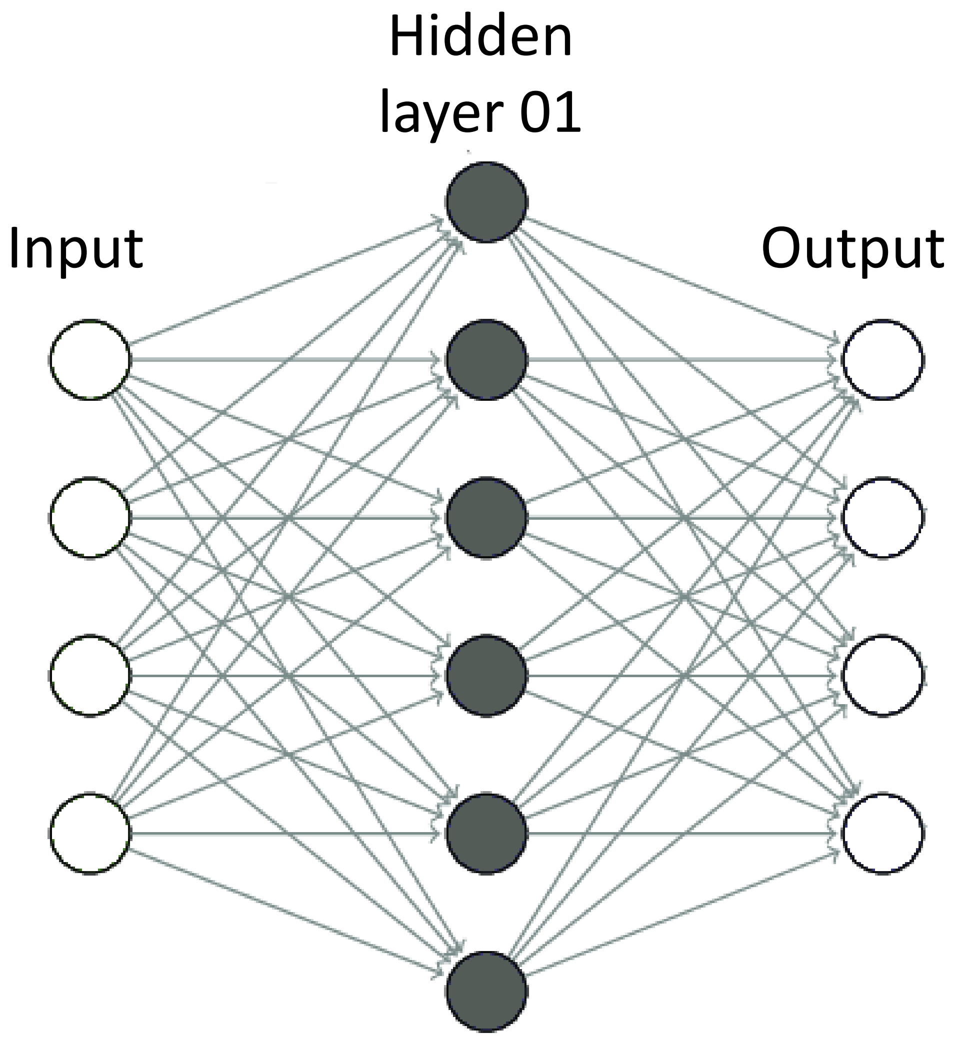
- Images, audio, text among others all have regular structures in a Euclidean space
  - Convolutions are well-defined operations that can be computed efficiently in these structures

Convolutional Neural Networks

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



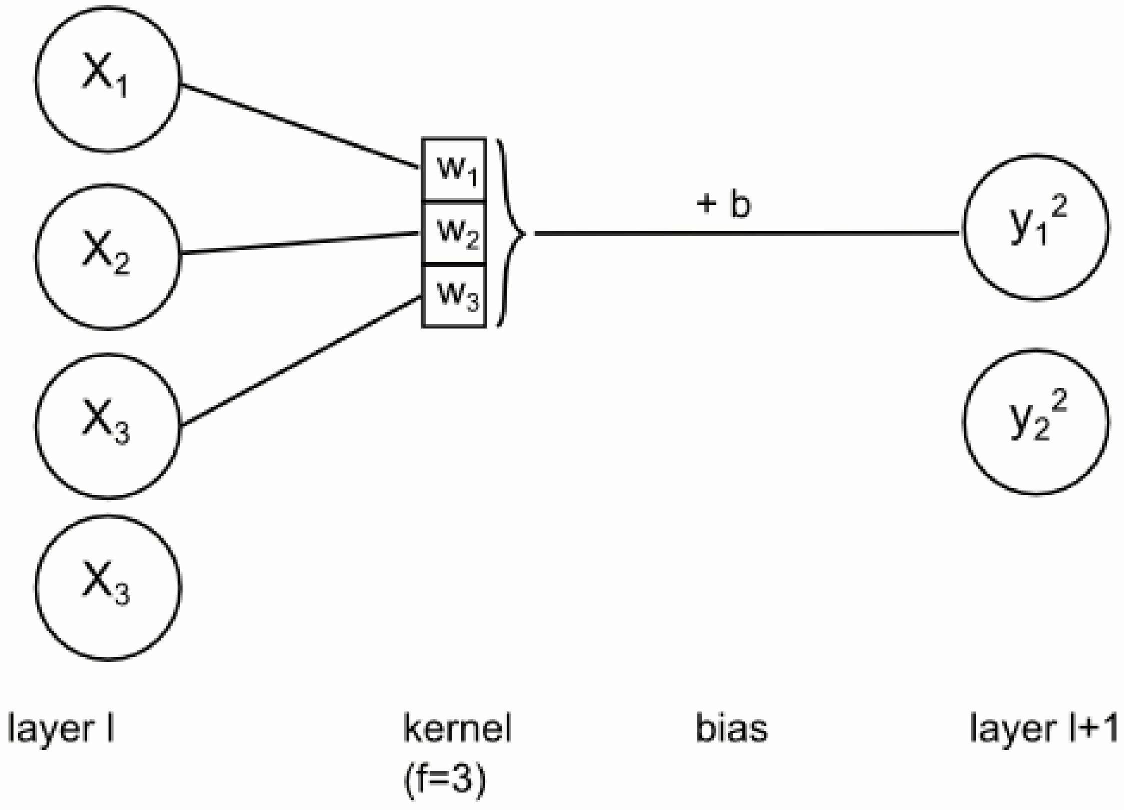
Fully Connected Neural Network – Global Property



- 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)

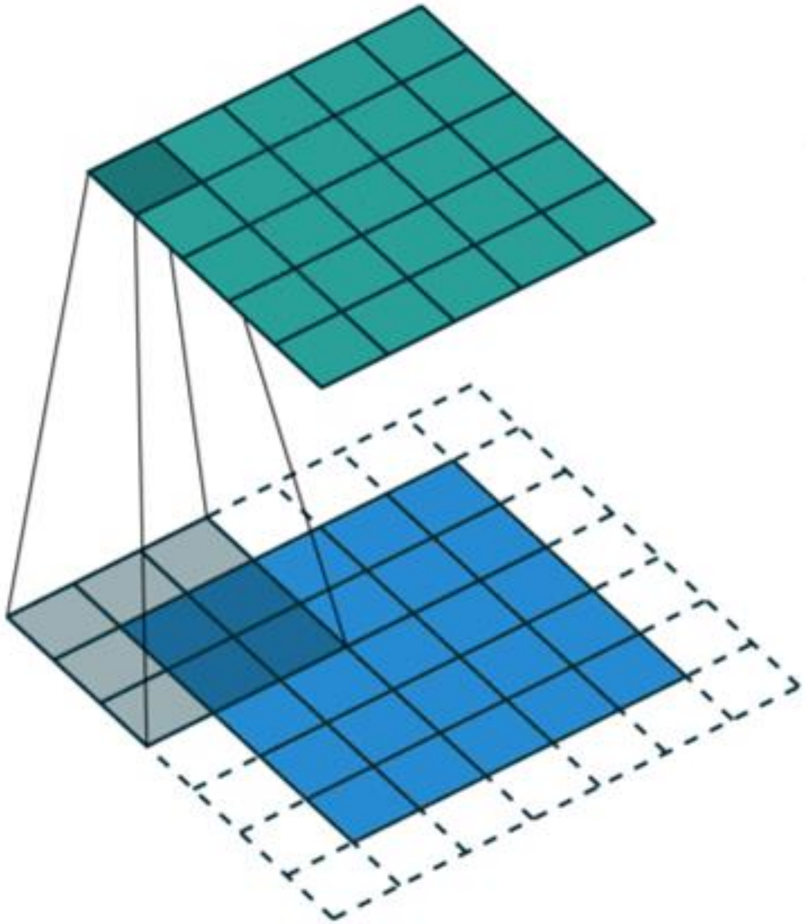
1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

4		

The bias term is added after the convolution

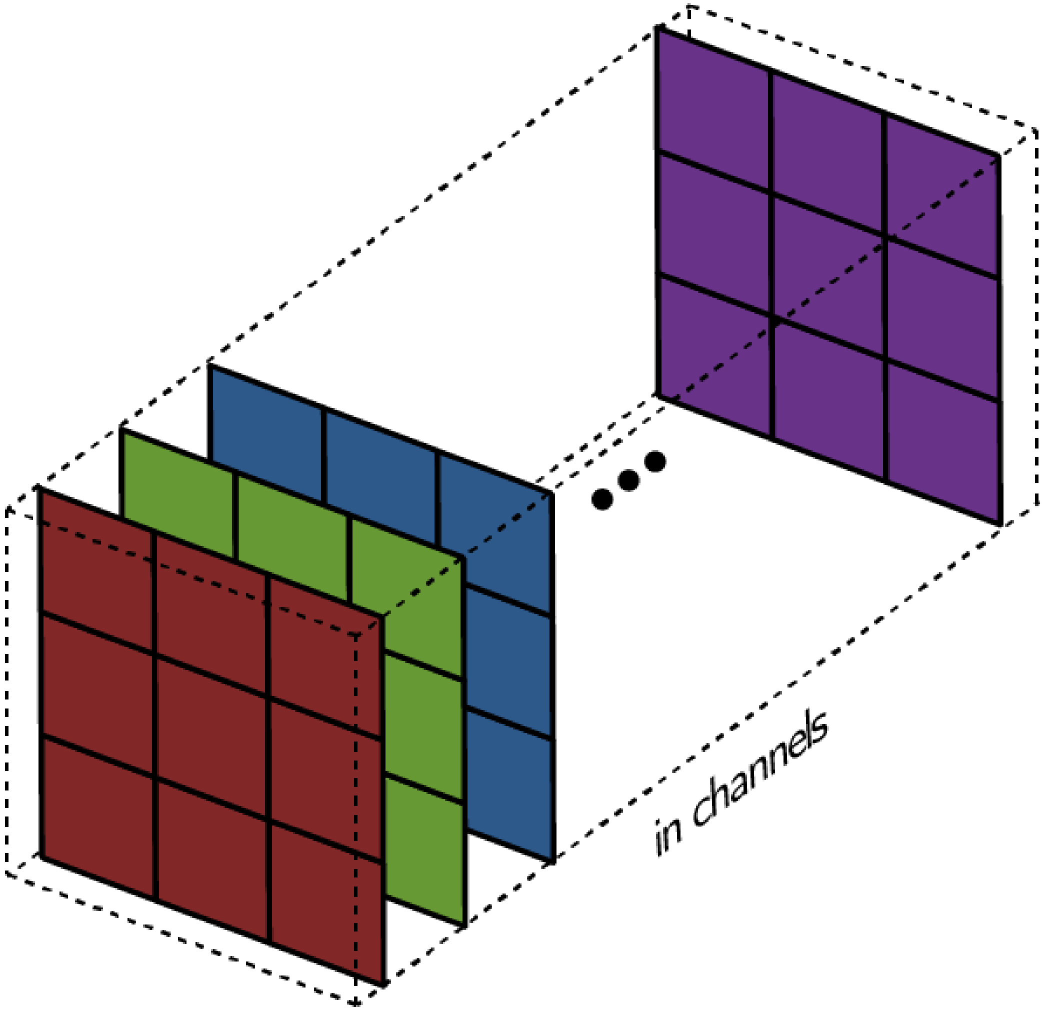
Convolution (padding)

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



UNIVERSITY OF CALGARY

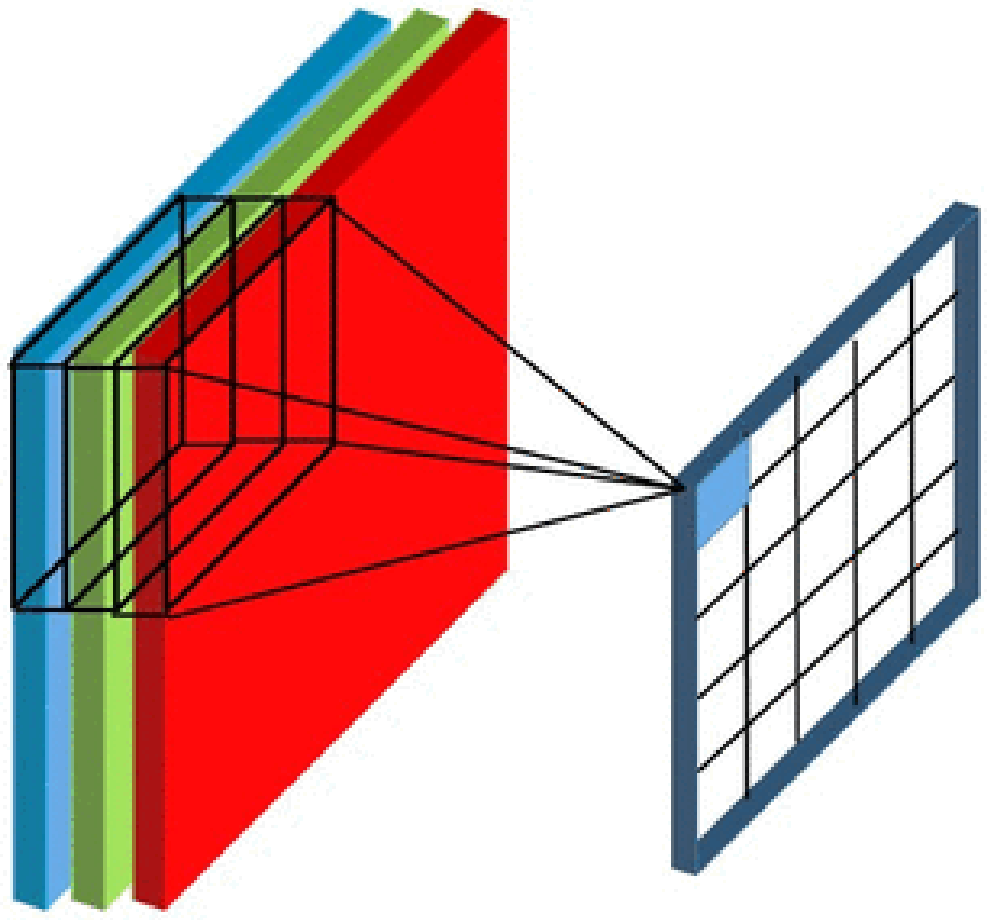
N-channel Images/Signals



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

UNIVERSITY OF CALGARY

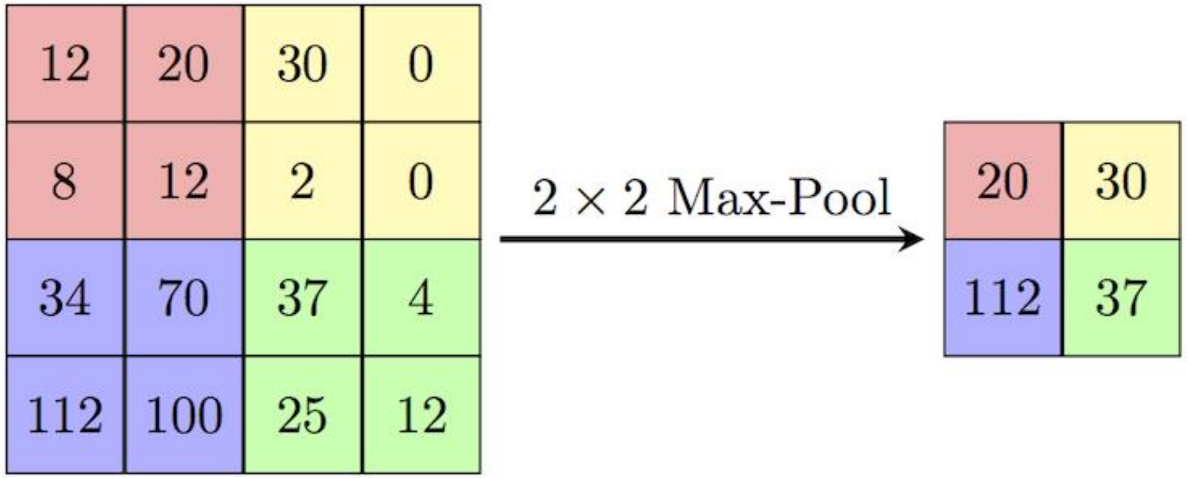
Convolution – multi-channel input



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

Max-Pooling

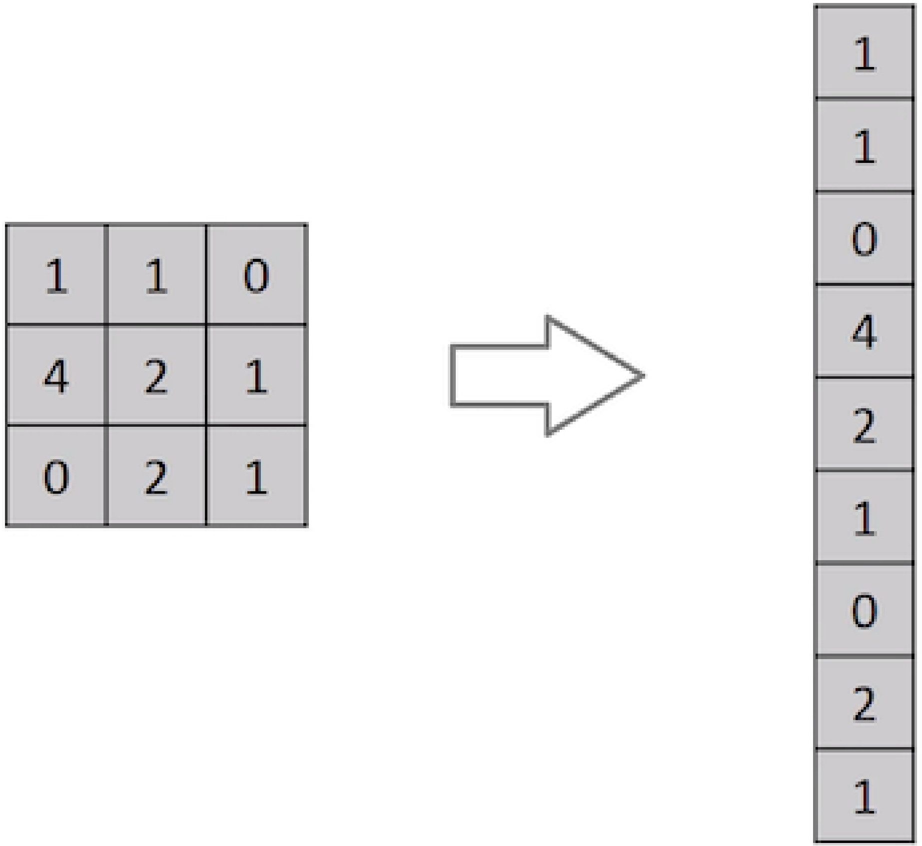
- Non-linear operation
- Reduce dimensionality and computational cost
- After a max-pooling, the number of filters in the subsequent layer is increased



UNIVERSITY OF CALGARY

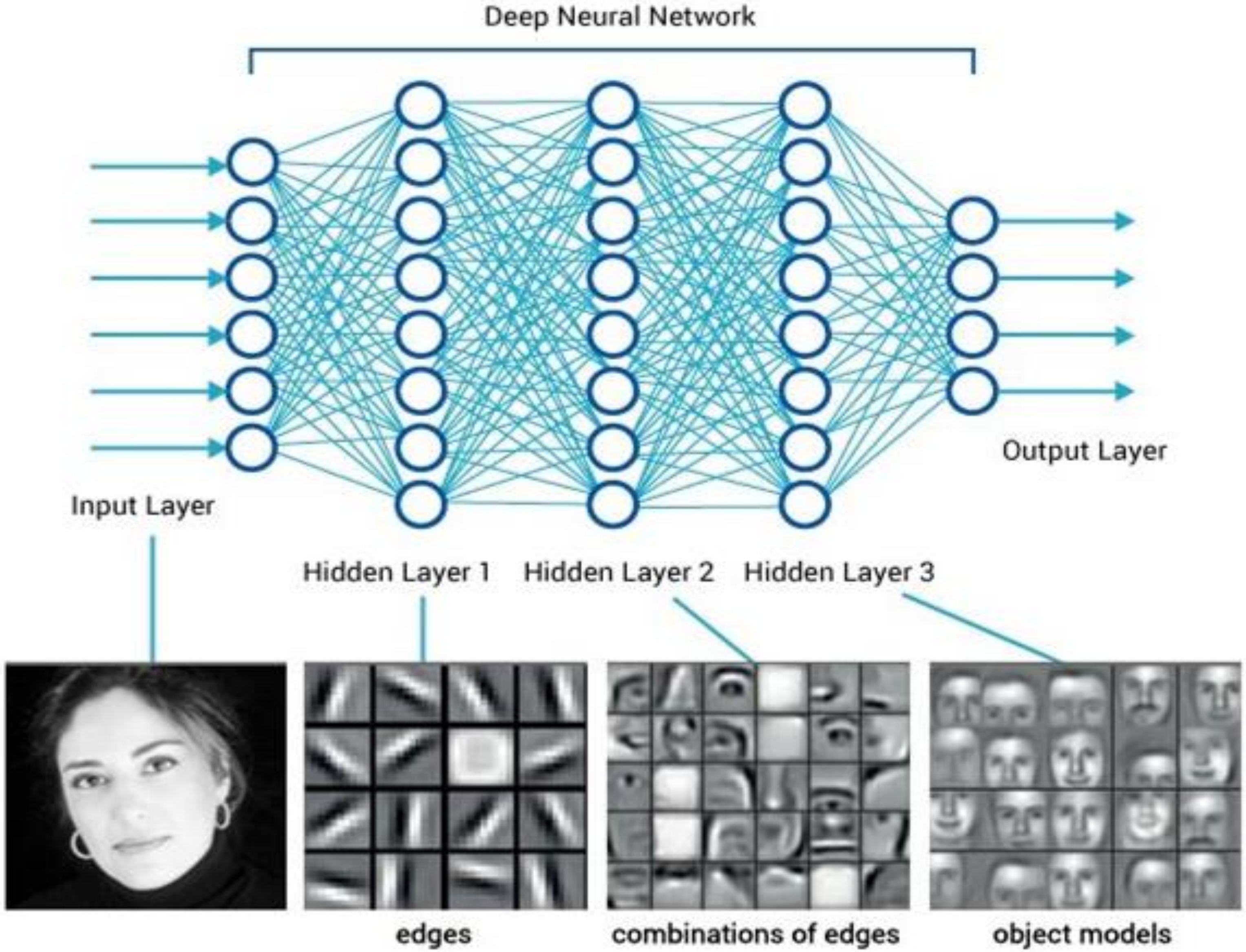
Flatten

- The flatten operation is applied before the fully connected layer



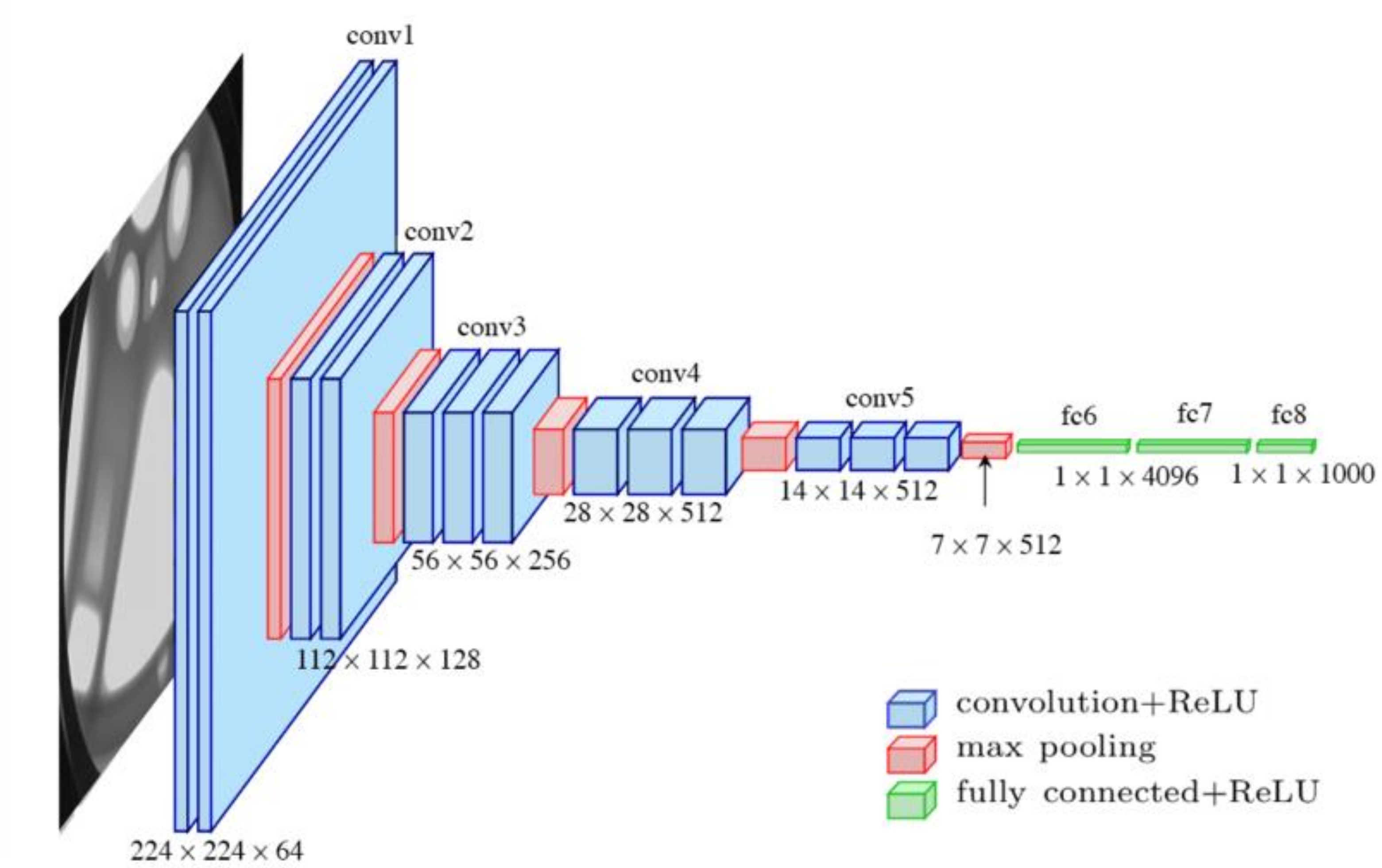
UNIVERSITY OF CALGARY

CNN Hierarchy of Concepts



UNIVERSITY OF CALGARY

VGG- 16 Architecture



UNIVERSITY OF CALGARY

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

UNIVERSITY OF CALGARY

Thank you!

UNIVERSITY OF CALGARY