Introduction to Computer Vision

Agenda

- Computer Vision Quiz
- Problems with ANNs for Images
- Convolutional Neural Networks (CNNs)
- Components of CNNs

Let's begin the discussion by answering a few questions on computer vision and convolutional neural networks (CNNs)

Computer Vision Quiz

ANNs are not invariant to which of the following transformations?

A. Translation

B. Rotation

C. Scaling

D. Translation followed by Rotation

Computer Vision Quiz

ANNs are not invariant to which of the following transformations?

A. Translation

B. Rotation

C. Scaling

D. Translation followed by Rotation

Problems with ANNs for Images

ANNs may struggle with recognizing objects that are translated, rotated, and/or scaled within an image

Each of the pairs below (T and T’ for translation, R and R’ for rotation, S and S’ for scaling) are different sets of inputs to the ANN - the outputs will also be different

But if the task is to identify a triangle in an image, then each pair should give the same output.

Computer Vision Quiz

How do CNNs address the problem of translation invariance?

A. By moving filters across the entire image to detect patterns at different locations.

B. By randomly selecting regions of the image for feature extraction.

C. By resizing the image.

D. By adding brightness to the image.

Computer Vision Quiz

How do CNNs address the problem of translation invariance?

A. By moving filters across the entire image to detect patterns at different locations.

B. By randomly selecting regions of the image for feature extraction.

C. By resizing the image.

D. By adding brightness to the image.

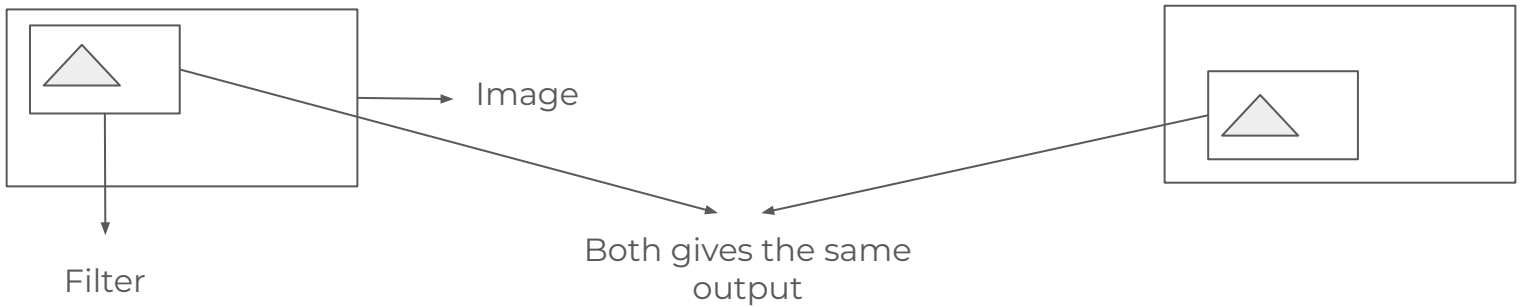
CNN Filters

A CNN filter scans an image to find features like edges or textures

Adapts its parameters to recognize particular patterns during training

CNNs address the problem of translation invariance by moving filters across the entire image through a sliding mechanism

Allows the network to detect patterns or features at different spatial locations



Computer Vision Quiz

What is the output of convolving a 2x2 filter of all ones with any 2x2 input matrix of real values?

A. The product of all elements in the input matrix.

B. The mean of all elements in the input matrix

C. The sum of all elements in the input matrix.

D. The maximum value of all elements in the input matrix.

Computer Vision Quiz

What is the output of convolving a 2x2 filter of all ones with any 2x2 input matrix of real values?

A. The product of all elements in the input matrix.

B. The mean of all elements in the input matrix

C. The sum of all elements in the input matrix.

D. The maximum value of all elements in the input matrix.

Convolution Operation

Refers to the process of applying a filter to the input data by computing the dot product between the filter and the local regions of the input

a	b
c	d

*

1	1
1	1

= a * 1 + b * 1 + c * 1 + d * 1

= a + b + c + d

Thus, we get the sum of all elements of the input matrix

Computer Vision Quiz

Great Learning
POWER AHEAD

Which operation applied to an image helps in preserving its dimensionality after convolution?

A. Translation

B. Rotation

C. Increasing the brightness

D. Padding

Computer Vision Quiz

Which operation applied to an image helps in preserving its dimensionality after convolution?

A. Translation

B. Rotation

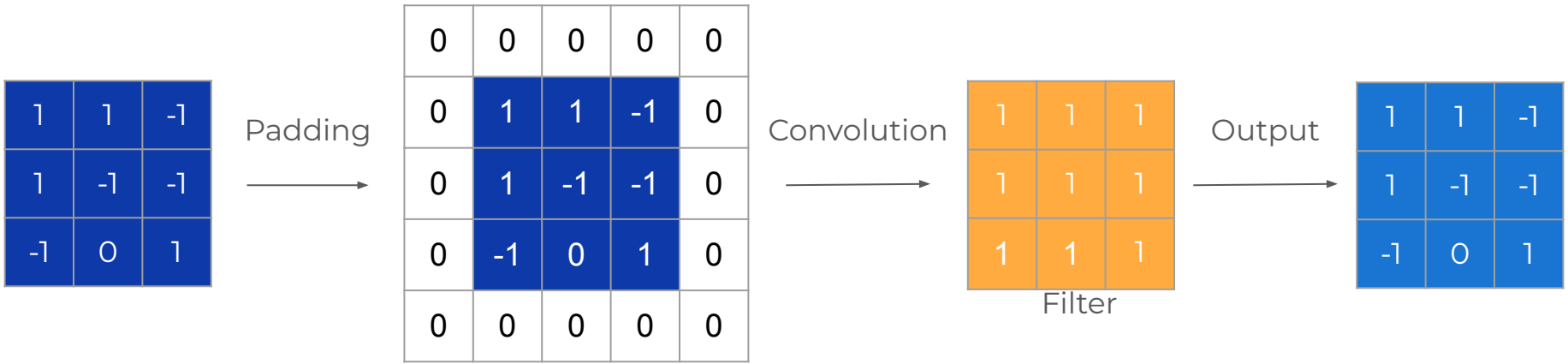
C. Increasing the brightness

D. Padding

Padding

Refers to adding extra border pixels around the input data before applying the convolution operation

Helps preserve spatial dimensions and information at the edges of the input, ensuring that the output feature map has the desired size



Convolution with a padding of 1

Computer Vision Quiz

In convolutional neural networks (CNNs), what effect does increasing the stride value have on the output feature map size?

A. It increases the size of the output feature map.

B. It has no effect on the size of the output feature map.

C. It reduces the number of channels in the output feature map.

D. It decreases the size of the output feature map.

Computer Vision Quiz

In convolutional neural networks (CNNs), what effect does increasing the stride value have on the output feature map size?

A. It increases the size of the output feature map.

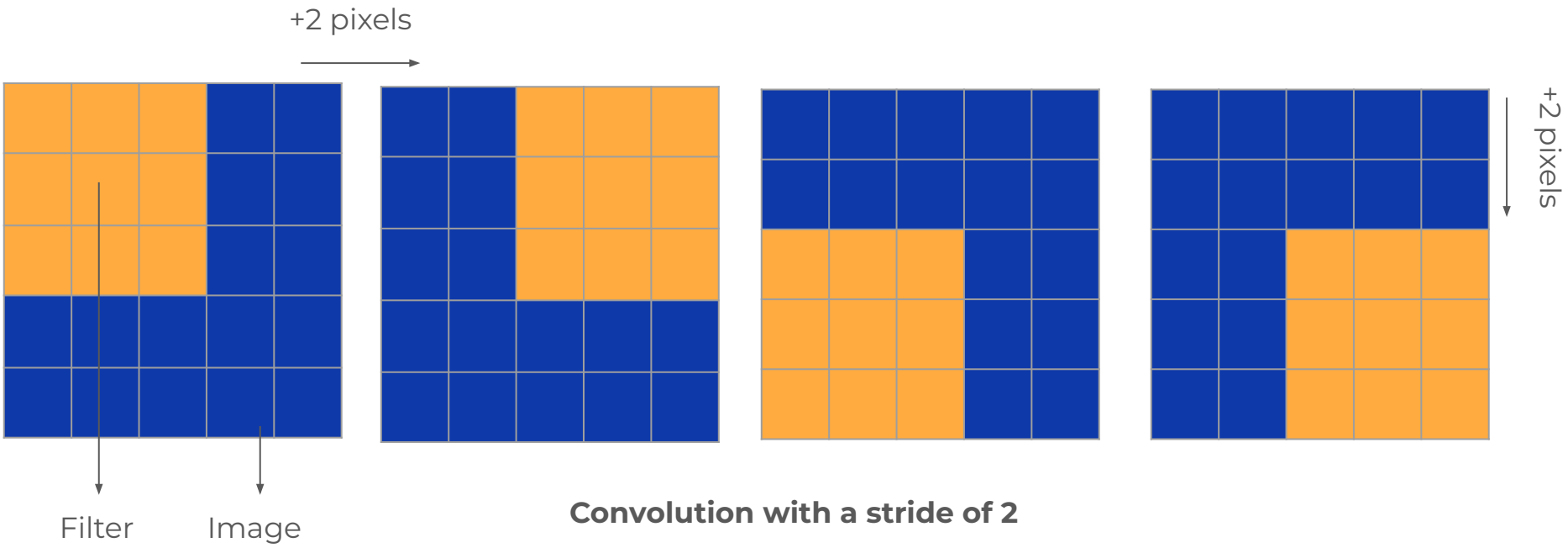
B. It has no effect on the size of the output feature map.

C. It reduces the number of channels in the output feature map.

D. It decreases the size of the output feature map.

Stride

Refers to the number of pixels by which the filter is moved across the input data during the convolution operation



Convolution with a stride of 2

Computer Vision Quiz

Why do we need a pooling layer in Convolutional Neural Networks (CNNs)?

A. To increase the number of parameters.

B. To reduce the spatial dimensions of the feature maps, reducing computational complexity.

C. To increase the spatial dimensions of the feature maps, reducing computational complexity.

D. To scale the input.

Computer Vision Quiz

Great Learning
POWER AHEAD

Why do we need a pooling layer in Convolutional Neural Networks (CNNs)?

A. To increase the number of parameters.

B. To reduce the spatial dimensions of the feature maps, reducing computational complexity.

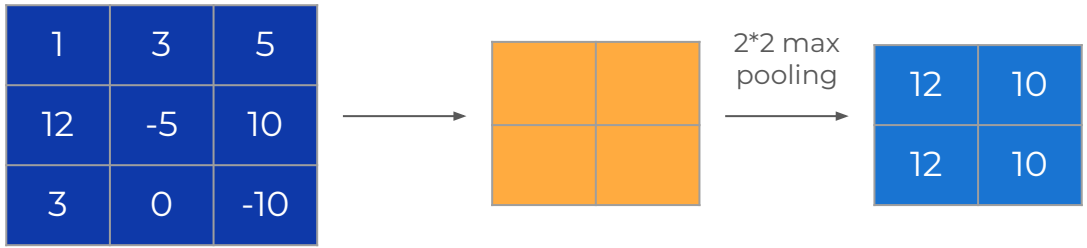
C. To increase the spatial dimensions of the feature maps, reducing computational complexity.

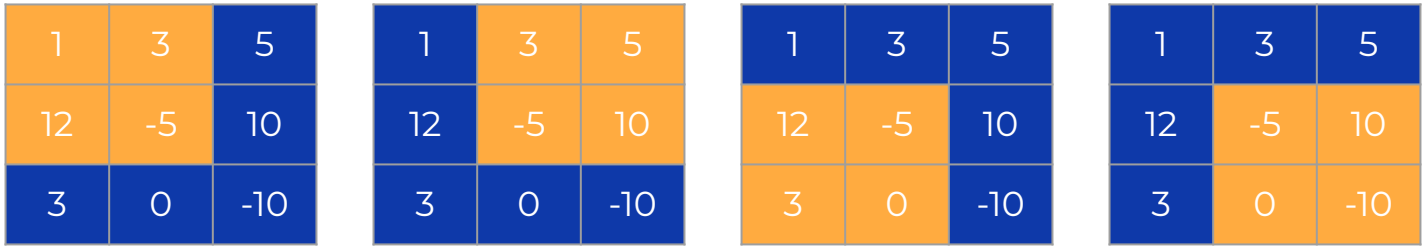
D. To scale the input.

Pooling

Decreases the spatial dimensions of feature maps

Reduction in size aids in managing computational complexity, facilitating faster training





Taking the maximum value in each of the regions.

Computer Vision Quiz

What is the total number of learnable parameters in a Max Pooling layer ?

A. Depends on the kernel size.

B. 1

C. Depends on the size of the input feature map

D. 0

Computer Vision Quiz

What is the total number of learnable parameters in a Max Pooling layer ?

A. Depends on the kernel size.

B. 1

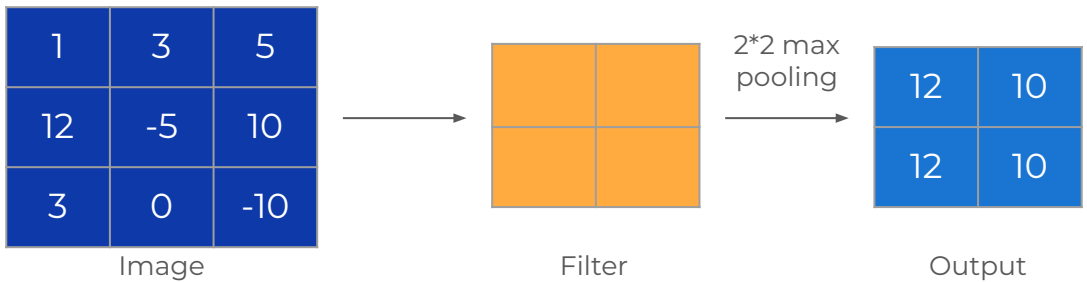
C. Depends on the size of the input feature map

D. 0

Max Pooling

Operates by selecting the maximum value within each pooling region, without any trainable weights or biases.

As we are performing an operation to find the maximum value, there are **no learnable parameters**



Computer Vision Quiz

Great Learning
POWER AHEAD

Consider the following code snippet for defining a CNN:

```
model = Sequential()
model.add(Conv2D(filters=1, kernel_size=(2, 2), input_shape=(3, 3, 1)))
model.add(Flatten())
model.add(Dense(units=1))
```

Given this model architecture with a 3x3 input image, how many learnable parameters are there in the entire model?

A. 10

B. 8

C. 9

D. 5

Computer Vision Quiz

Great Learning
POWER AHEAD

Consider the following code snippet for defining a CNN:

```
model = Sequential()
model.add(Conv2D(filters=1, kernel_size=(2, 2), input_shape=(3, 3, 1)))
model.add(Flatten())
model.add(Dense(units=1))
```

Given this model architecture with a 3x3 input image, how many learnable parameters are there in the entire model?

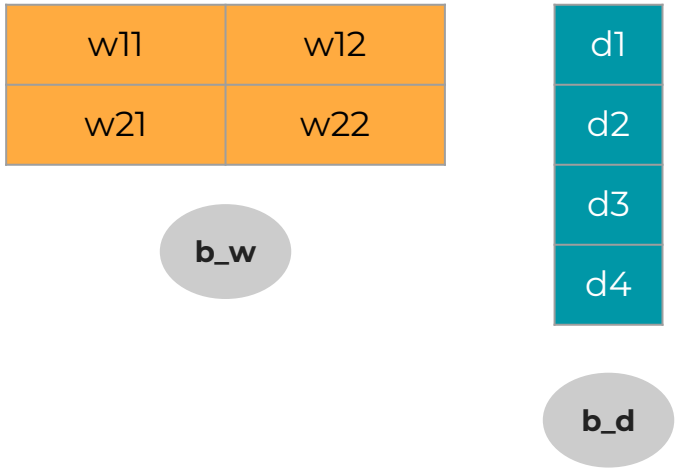
A. 10

B. 8

C. 9

D. 5

CNN Parameters



Parameter	Parameter Description	Count
wij	Weights of the CNN layer	4
b_w	Bias of the CNN layer	1
d_i	Weights of the Dense layer	4
b_d	Bias of the Dense layer	1

Thus, we have 10 learnable parameters in this neural network

Great Learning

Happy Learning !