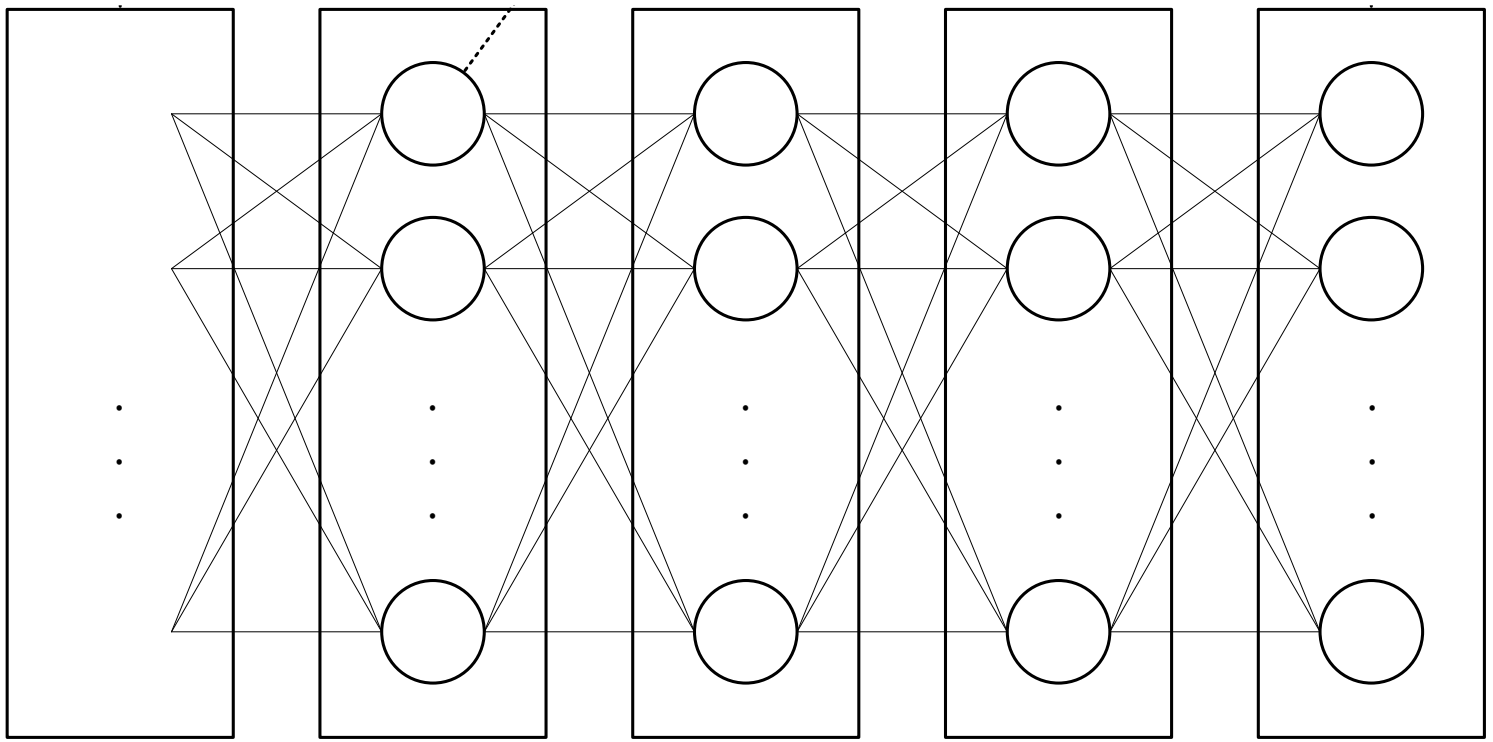
An Introduction to Convolutional Neural Networks

Problems with Fully Connected Networks

Fully Connected Network Structure

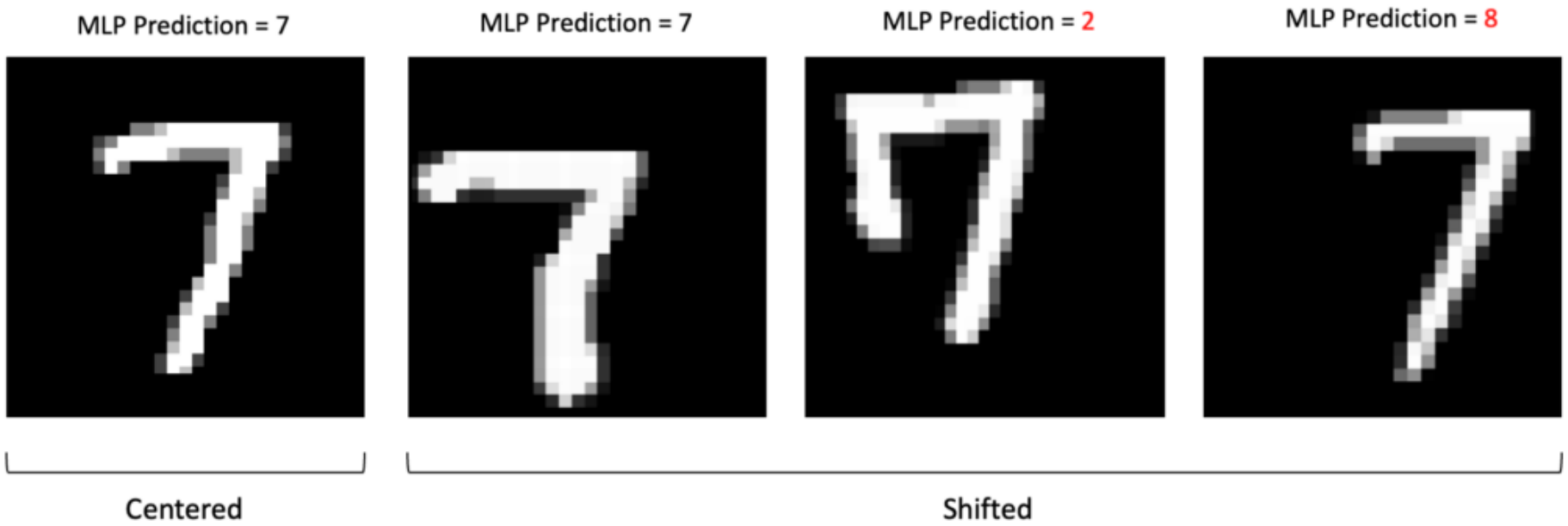


Problems with Fully Connected Networks

- As the networks grows in size, the number of parameters, especially weights, increases significantly.
- More parameters need stronger hardware for computation and storage.
- Can we reduce the number of the weights or even share them?

Problems with Fully Connected Networks

- Fully Connected Networks are not translation invariant.



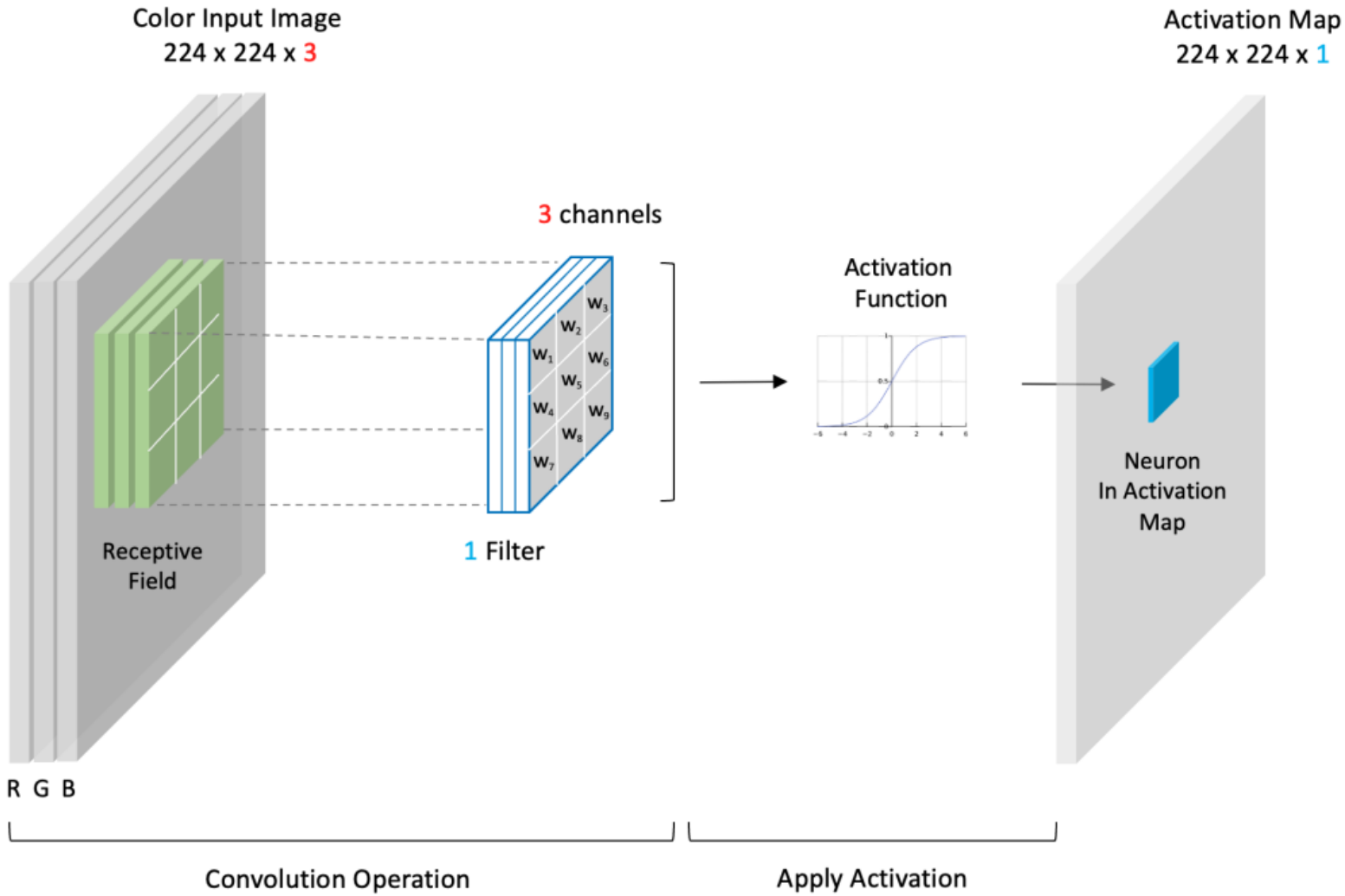
Intuitive Explanation

Case: Image Classification

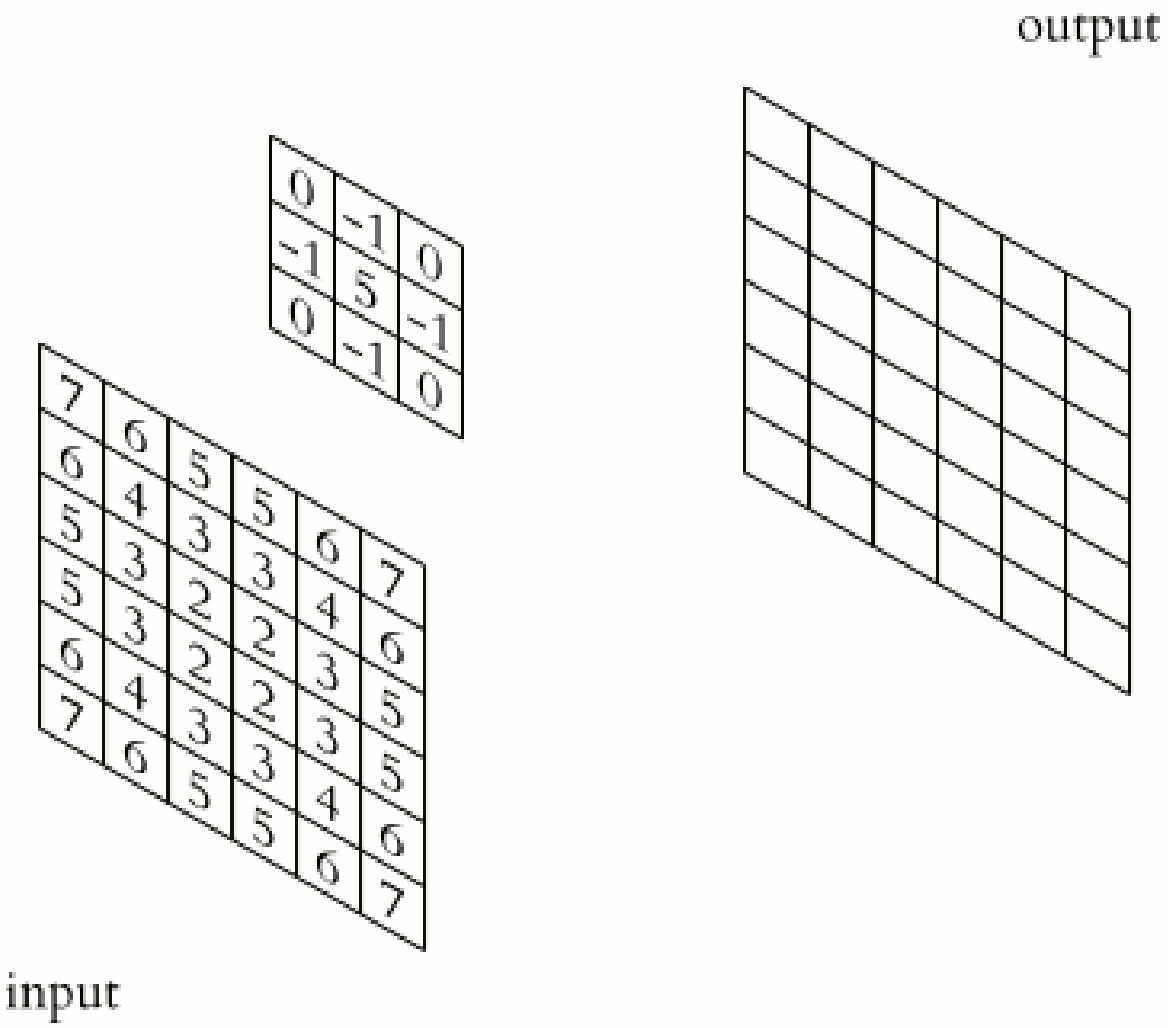
- Consider a network that has to classify images as “bird” or “not bird”.
- A fully connected network is undesirable due to its limitations.
- Some patterns, e.g. the wings or the tail, are much smaller than the whole image.
- These patterns can appear in different parts of the images.

CNN Structure

Convolutional Layer



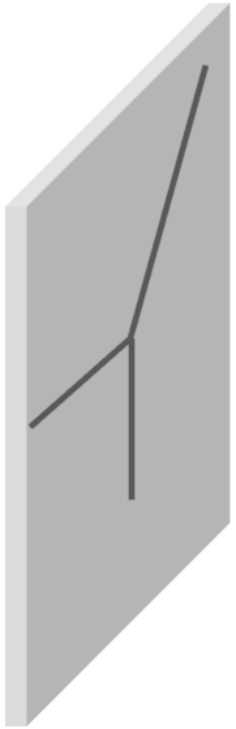
Convolution Operation



Filters

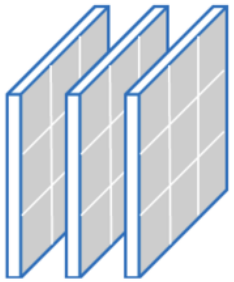
- Each filter gradually learns to detect a specific pattern.
- The number of channels in the output of a convolutional layer is equal to the number of its filters.

Filters



Input Image

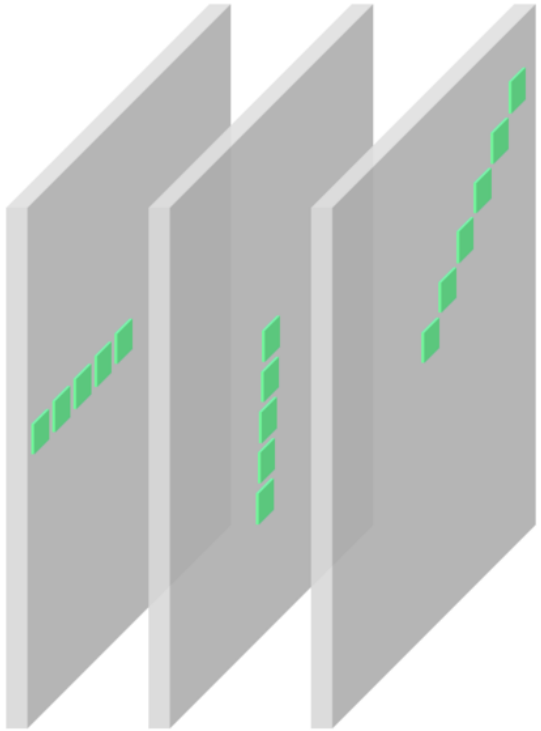
Filters learn to detect structural patterns



F_1 = Horizontal Lines

F_2 = Vertical Lines

F_3 = Diagonal Lines



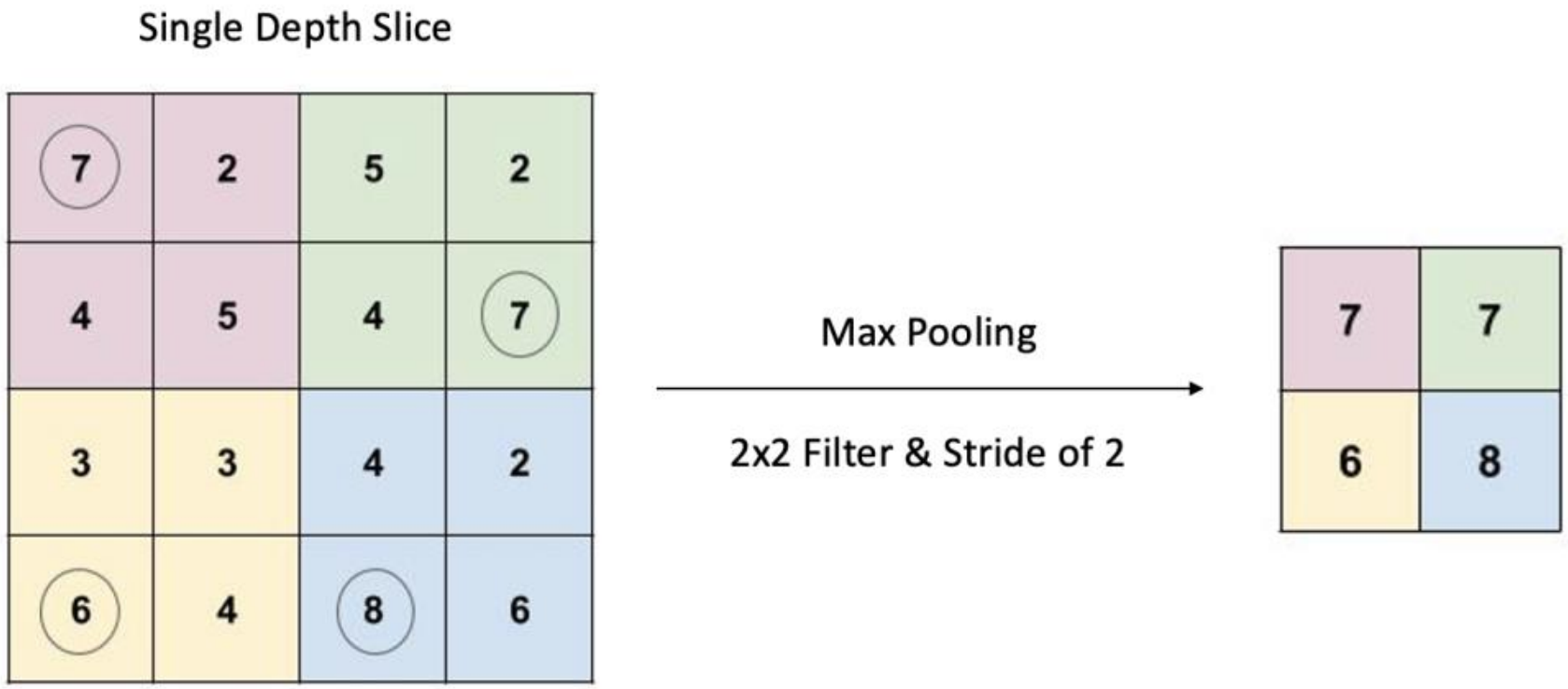
Activation Maps

= Highly activated neurons in activation maps

Padding and Stride

- Padding in convolution refers to adding extra border pixels around the input image to preserve spatial dimensions after convolution.
- Stride refers to the number of pixels the filter kernel moves at each step during the convolution operation.

Max Pooling Layer



Pooling Layers

- There are other types of pooling layers such as average pooling.
- However, max pooling is the most commonly used pooling layer in modern CNNs.

Convolutional Block

- A convolutional block typically contains one or more convolutional layers followed by a pooling layer.
- Other layers are also sometimes incorporated, but we will focus on these two layer types to keep things simple.

CNN Example: VGG-16