

The essential advantage of the pooling layer is its ability to inject location invariance into the network. Location invariance means that features can be detected by the network no matter where they are on the image.

The pooling layer applies its aggregation function to all the channels of the input image. For example, in an R, G, B image (i.e., an image with three channels, red, green, and blue), the MaxPool will be applied independently to all the three channels. Similarly, for feature maps with a particular depth, the pooling aggregation will be applied separately to each feature map. See Figure 35-14 as an example of applying pooling to the channel depth of its inputs.

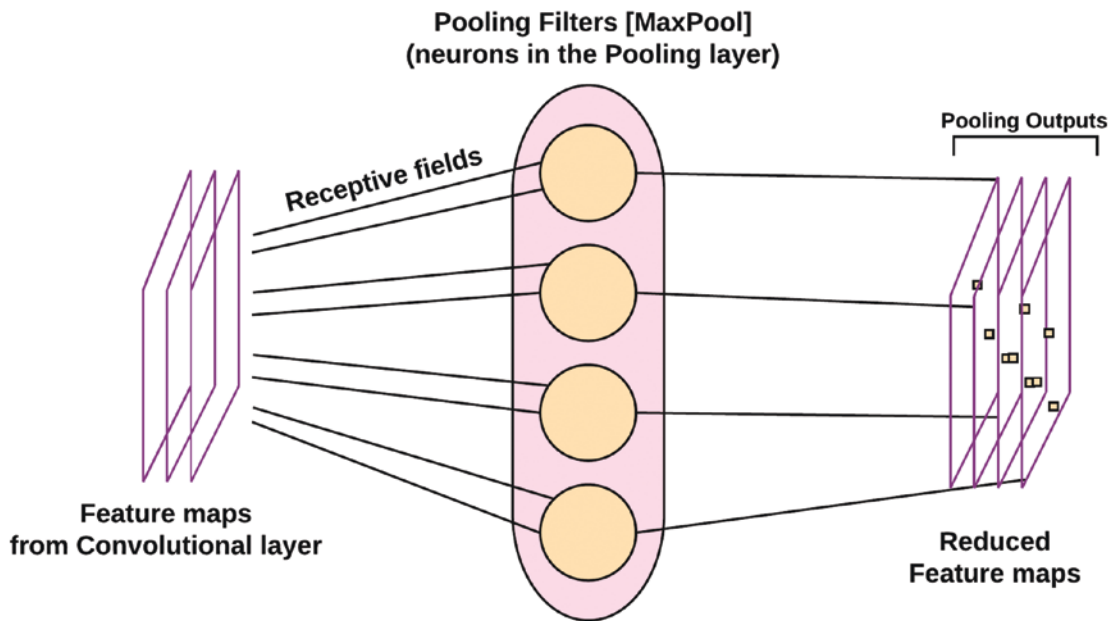


Figure 35-14. Example of applying pooling to input with depth. Note that the filters in the pooling layer have no weights or biases

The Fully Connected Network Layer

The fully connected network (FCN) layer is our regular feedforward neural network or multilayer perceptron. These layers typically have a non-linear activation function. In any case, the FCN is the final layer of the convolutional neural network. In this case, a softmax activation is used to output the probabilities that an input belongs to a particular class.

Before passing an input into the FCN, the image matrix will have to be flattened. For example, a $28 \times 28 \times 3$ image matrix will become 2352 input weights plus a bias of 1 into the fully connected network.

In the case of our convolutional network, the feature maps of either the convolutional or pooling layer are flattened before passing into the FCN to compute the final network probabilities using the softmax function.

An Example CNN Architecture

We have discussed the building blocks of a convolutional neural network system. As you've seen, a CNN system is principally composed of convolution layers, pooling layers, and the fully connected layer. However, the way these layers are arranged and in what number are down to the preferred heuristics of the particular use case that a CNN is employed in solving.

An example CNN modeling pipeline is shown here:

1. The first layer following the input layer of images must be a convolutional layer for extracting image features. A 3×3 image filter is commonly used depending on the size of the input image.
2. Pooling layers typically follow a set of one or more convolutional layers. Typically, a 2×2 filter size is used in the pooling layer.
3. The fully connected layer must be the final layer of the CNN. It is also called the dense layer. It contains the softmax activation function to give the probabilities of class membership.
4. CNN may include one or more Dropout layers to prevent the network from overfitting.