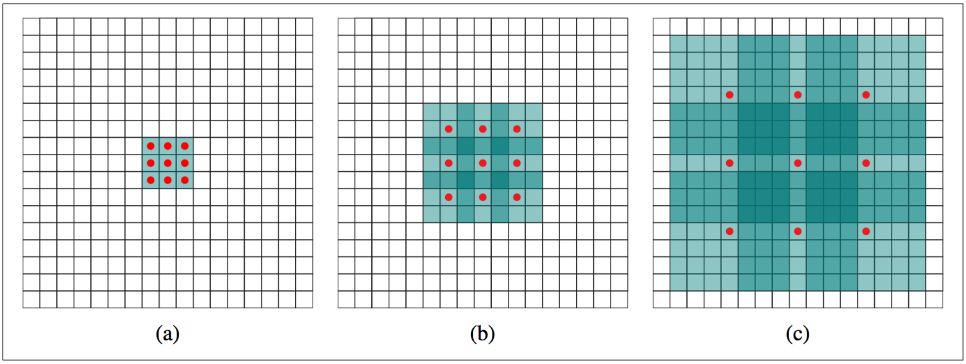

Figure 6-8. A dilated (or atrous) convolution. Gaps are left in the local receptive field for each neuron. Diagram (a) depicts a 1-dilated 3 × 3 convolution. Diagram (b) depicts the application of a 2-dilated 3 × 3 convolution to (a). Diagram (c) depicts the application of a 4-dilated 3 × 3 convolution to (b). Notice that the (a) layer has receptive field of width 3, the (b) layer has receptive field of width 7, and the (c) layer has receptive field of width 15.

Applications of Convolutional Networks

In the previous section, we covered the mechanics of convolutional networks and introduced you to many of the components that make up these networks. In this section, we describe some applications that convolutional architectures enable.

Object Detection and Localization

Object detection is the task of detecting the objects (or entities) present in a photograph. Object localization is the task of identifying where in the image the objects exist and drawing a "bounding box" around each occurrence. Figure 6-9 demonstrates what detection and localization on standard images looks like.

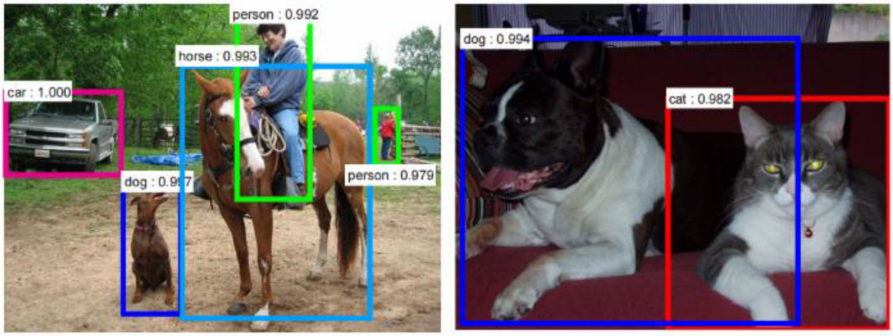

Figure 6-9. Objects detected and localized with bounding boxes in some example images.

Why is detection and localization important? One very useful localization task is detecting pedestrians in images taken from a self-driving car. Needless to say, it's extremely important that a self-driving car be able to identify all nearby pedestrians. Other applications of object detection could be used to find all instances of friends in photos uploaded to a social network. Yet another application could be to identify potential collision dangers from a drone.

This wealth of applications has made detection and localization the focus of tremendous amounts of research activity. The ILSVRC challenge mentioned multiple times in this book focused on detecting and localizing objects found in the ImagetNet collection.

Image Segmentation

Image segmentation is the task of labeling each pixel in an image with the object it belongs to. Segmentation is related to object localization, but is significantly harder since it requires precisely understanding the boundaries between objects in images. Until recently, image segmentation was often done with graphical models, an alternate form of machine learning (as opposed to deep networks), but recently convolutional segmentations have risen to prominence and allowed image segmentation algorithms to achieve new accuracy and speed records. Figure 6-10 displays an example of image segmentation applied to data for self-driving car imagery.

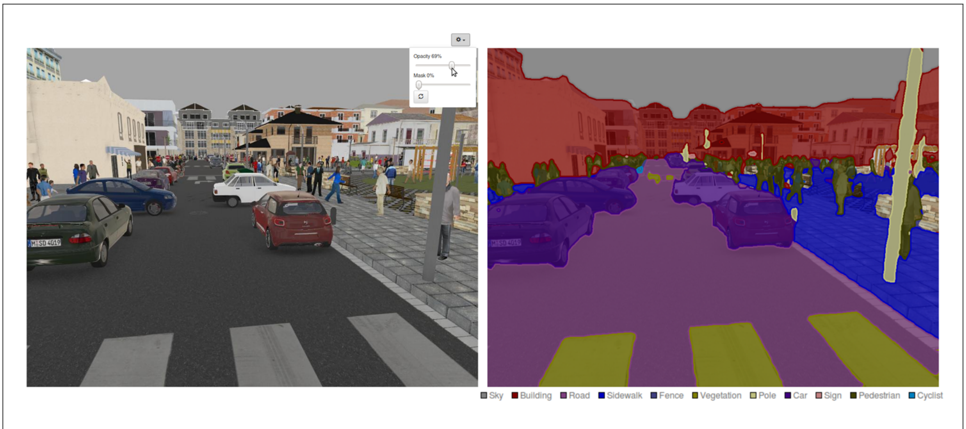

Figure 6-10. Objects in an image are "segmented" into various categories. Image segmentation is expected to prove very useful for applications such as self-driving cars and robotics since it will enable fine-grained scene understanding.