



Figure 9.16: A comparison of locally connected layers, tiled convolution, and standard convolution. All three have the same sets of connections between units, when the same size of kernel is used. This diagram illustrates the use of a kernel that is two pixels wide. The differences between the methods lies in how they share parameters. (*Top*)A locally connected layer has no sharing at all. We indicate that each connection has its own weight by labeling each connection with a unique letter. (*Center*)Tiled convolution has a set of  $t$  different kernels. Here we illustrate the case of  $t = 2$ . One of these kernels has edges labeled “a” and “b,” while the other has edges labeled “c” and “d.” Each time we move one pixel to the right in the output, we move on to using a different kernel. This means that, like the locally connected layer, neighboring units in the output have different parameters. Unlike the locally connected layer, after we have gone through all  $t$  available kernels, we cycle back to the first kernel. If two output units are separated by a multiple of  $t$  steps, then they share parameters. (*Bottom*)Traditional convolution is equivalent to tiled convolution with  $t = 1$ . There is only one kernel and it is applied everywhere, as indicated in the diagram by using the kernel with weights labeled “a” and “b” everywhere.