为什么要采用stride？如何选择？

取决于图片的分辨率。当stride较大时，对图片的处理就相当于down sampling，就有点类似pooling，有时候效果比pooling更好。同时，在应用stride的同时也是在减小feature map的输出尺寸，也就相当于减少了模型参数。因为在最终FCN时是要连接全部参数的，如果feature map的参数比较少，training cost也更低。