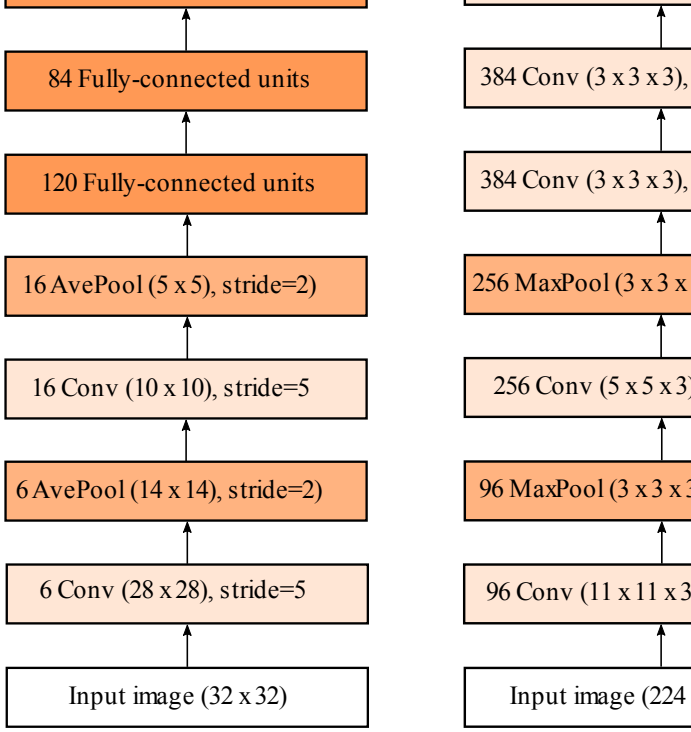

Notation:
<number of planes> **type of layer** *<width x hight x RGB>*, **stride/p**

Each layer in AlexNet is three dimensional because it was designed (LeNet-5 classified 10 grey-scale digits). The dimensions represent and blue) color values. This type of 3-D arrays of numbers is often tensors. The pooling operation is done by taking the maximum val the average of all units, which is known as **max pooling**. The patte convolutional and pooling layers is different from the one in LeNet utilizes the same building blocks and operations as LeNet-5.

Neural network models of vision and computer

In my experience, If you ask a random researcher in computer visio between the human visual/perceptual system and convolutional ne be something like: "*Well, CNN's are roughly inspired in the brain b brain. I care about solving the problem artificial vision by any mea biological correspondence to human vision, more or less in the sam without having to imitate birds flapping*". Or some version of that. as an independent area of research with its own goals, regardless o neural networks these days. Beyond the parallels with human visio AlexNet are designed to maximize object-recognition performance instance, ReLU units were introduced in AlexNet because they facil are more biologically realistic than sigmoids. For instance, the LeN in the context of the debate between traditional pattern recognition automated learning-based approach of neural nets. Nothing was s However, from our perspective, the issue of **whether convolution human perception and vision** is critical. This is an open debate. convolutional nets are useful models for human vision and percep