卷积神经网络（Convolutional Neural Network, CNN）是一种前馈神经网络，它的人工神经元可以响应一部分覆盖范围内的周围单元，对于大型图像处理有出色表现。

卷积神经网络利用输入是图片的特点，把神经元设计成三个维度：width, height, depth(注意这个depth不是神经网络的深度，而是用来描述神经元的) 。比如输入的图片大小是32×32×3(rgb)，那么输入神经元就也具有32×32×3 的维度。

卷积神经网络（CNN）由输入层、卷积层、线性整流层（激活函数）、池化层、全连接层组成，即INPUT-CONV-RELU-POOL-FC。

**卷积层（Convolutional layer）**，卷积神经网路中每层卷积层由若干卷积单元组成，每个卷积单元的参数都是通过反向传播算法优化得到的。卷积运算的目的是提取输入的不同特征，第一层卷积层可能只能提取一些低级的特征如边缘、线条和角等层级，更多层的网络能从低级特征中迭代提取更复杂的特征。

**线性整流层（Rectified Linear Units layer, ReLU layer）**，这一层的激活函数（Activation function）使用线性整流（Rectified Linear Units, ReLU）。

**池化层（Pooling layer）**，通常在卷积层之后会得到维度很大的特征，将特征切成几个区域，取其最大值或平均值，得到新的、维度较小的特征。

**全连接层（Fully-Connected layer）**, 把所有局部特征结合变成全局特征，用来计算最后每一类的得分。

普通神经网络把输入层和隐含层进行“全连接(Full Connected)“的设计。从计算的角度来讲，相对较小的图像从整幅图像中计算特征是可行的。但是，如果是更大的图像（如96x96 的图像），要通过这种全联通网络的这种方法来学习整幅图像上的特征，从计算角度而言，将变得非常耗时。

卷积层解决这类问题的一种简单方法是对隐含单元和输入单元间的连接加以限制：每个隐含单元仅仅只能连接输入单元的一部分。每个隐含单元连接的输入区域大小叫r神经元的感受野(receptive field)。