

卷积神经网络(CNN)

目录：

深度学习与卷积神经网络

卷积神经网络在图像处理中的应用

神经元(感知器)

构成神经网络

卷积神经网络

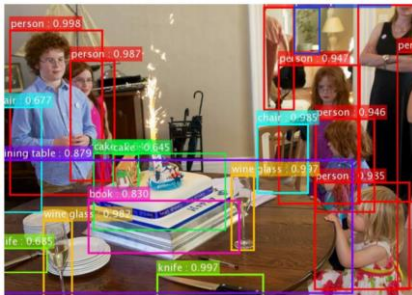
1. 深度学习与卷积神经网络

深度学习(Deep Learning, DL) 是机器学习(ML, Machine Learning)领域中一个新的研究方向。是学习样本数据的内在规律和表示层次, 这些学习过程中获得的信息对诸如文字, 图像和声音等数据的解释有很大的帮助。在搜索技术, 数据挖掘, 机器学习, 机器翻译, 自然语言处理, 多媒体学习, 语音, 推荐和个性化技术, 以及其他相关领域都取得了很多成果。深度学习又分为卷积神经网络 (CNN), 深度置信网 (DBN), 对抗神经网络 (GAN) 等。

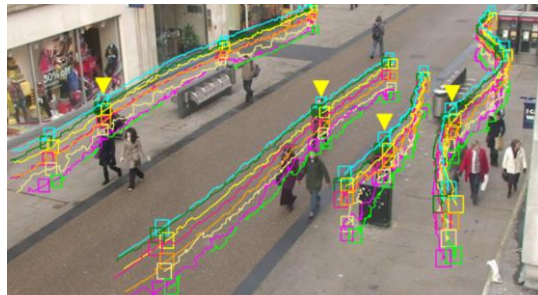
卷积神经网络 (Convolutional Neural Networks, CNN) 是在多层神经网络的基础上发展起来的针对图像分类和识别而特别设计的一种深度学习方法。与其他深度学习结构相比, 卷积神经网络在图像和语音识别方面能够给出更好的结果。这一模型也可以使用反向传播算法进行训练。相比较其他深度、前馈神经网络, 卷积神经网络需要考量的参数更少, 使之成为一种颇具吸引力的深度学习结构。

2. 卷积神经网络在图像处理中的应用

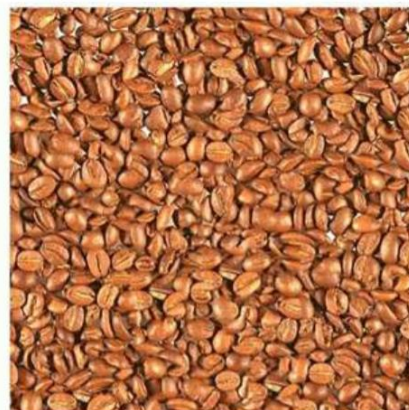
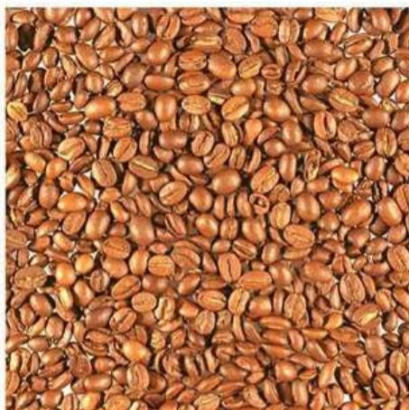
影像识别, 视频分析:



物体识别



物体追踪



纹理合成



风格转移

自然语言处理：

CNN 的模型被证明可以有效的处理各种自然语言处理的问题，如语义分析、搜索结果提取、句子建模、分类、预测、和其他传统的 NLP 任务等。



手写数字识别

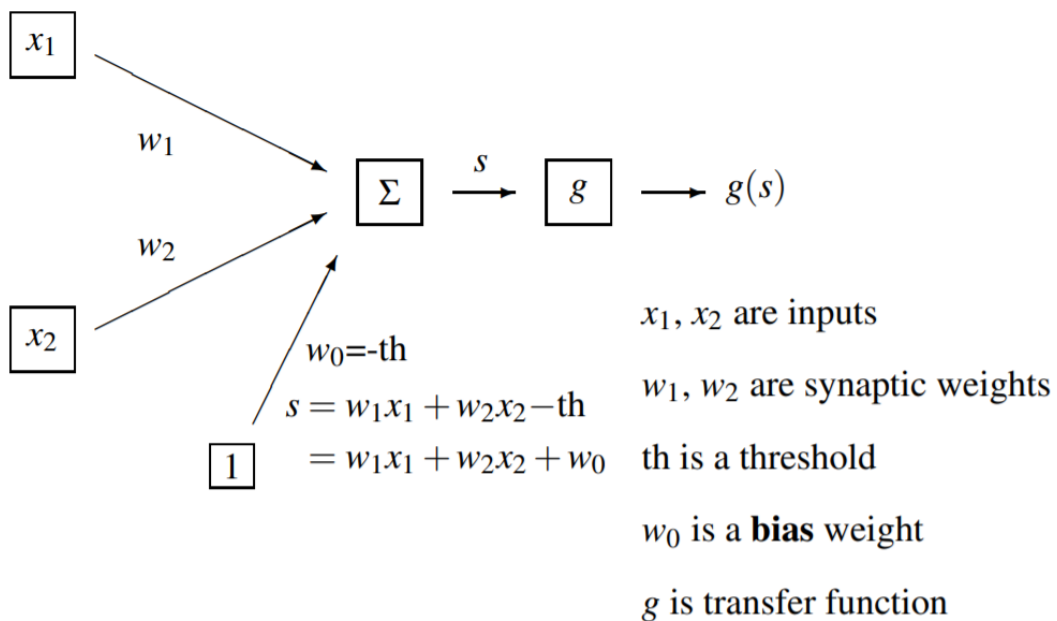
药物发现，医学造影：

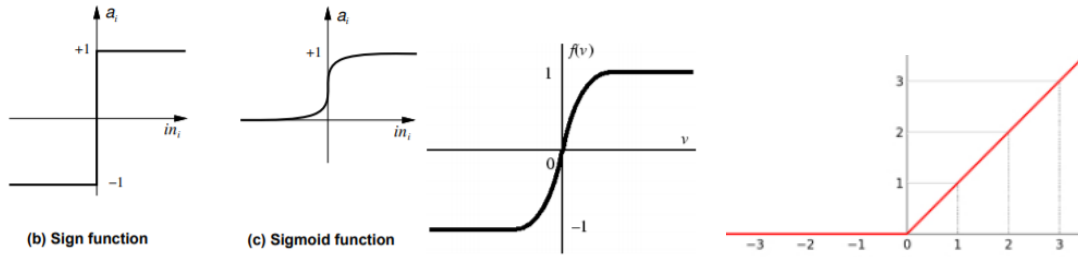
卷积神经网络已在药物发现中使用。卷积神经网络被用来预测的分子与蛋白质之间的相互作用，以此来寻找靶向位点，寻找出更安全有效的潜在治疗方法。

竞技：

AlphaGo, AlphaStar, OpenAI Five

3. 神经元(neurons) -> 感知器





if $g(s) = 0$ but should be 1,

if $g(s) = 1$ but should be 0,

$$w_k \leftarrow w_k + \eta x_k$$

$$w_k \leftarrow w_k - \eta x_k$$

$$w_0 \leftarrow w_0 + \eta$$

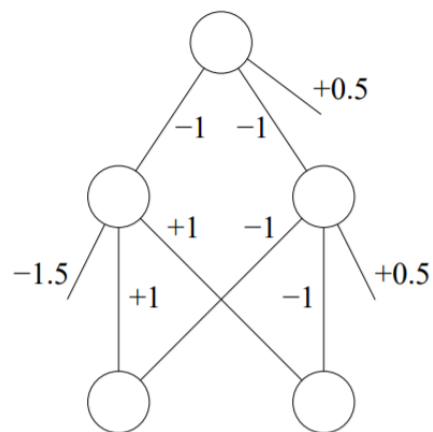
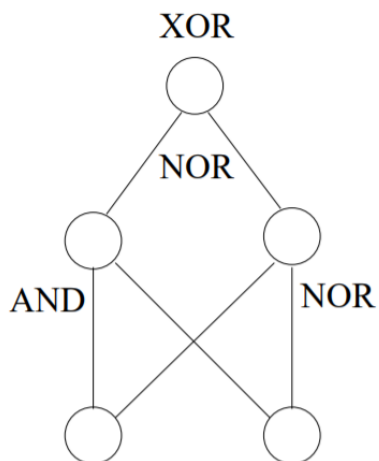
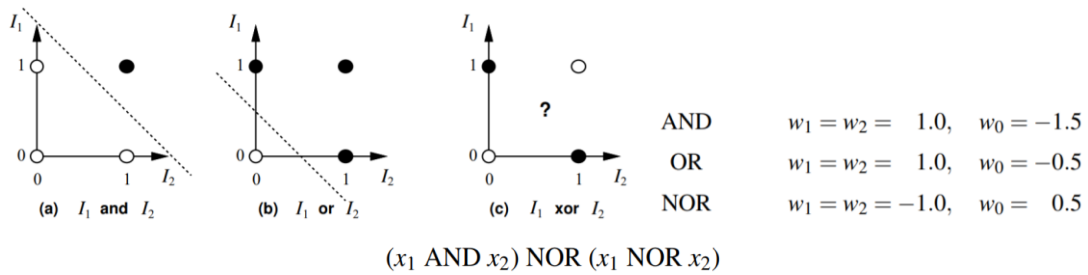
$$w_0 \leftarrow w_0 - \eta$$

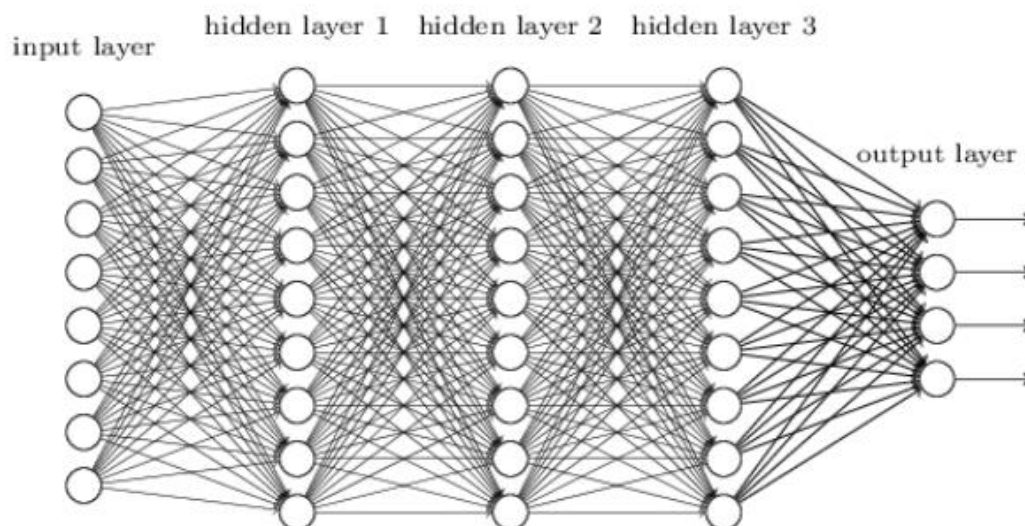
$$\text{so } s \leftarrow s + \eta (1 + \sum_k x_k^2)$$

$$\text{so } s \leftarrow s - \eta (1 + \sum_k x_k^2)$$

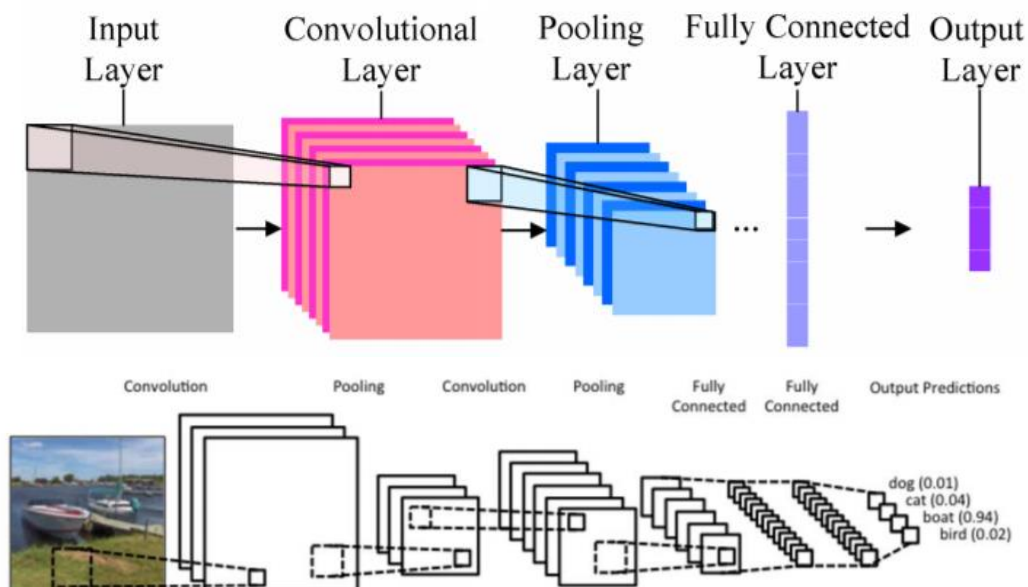
otherwise, weights are unchanged. ($\eta > 0$ is called the **learning rate**)

4. 神经网络(neural networks, NNs)





5. 卷积神经网络 (Constitutional Neural Networks, CNN)



First Layer



Second Layer

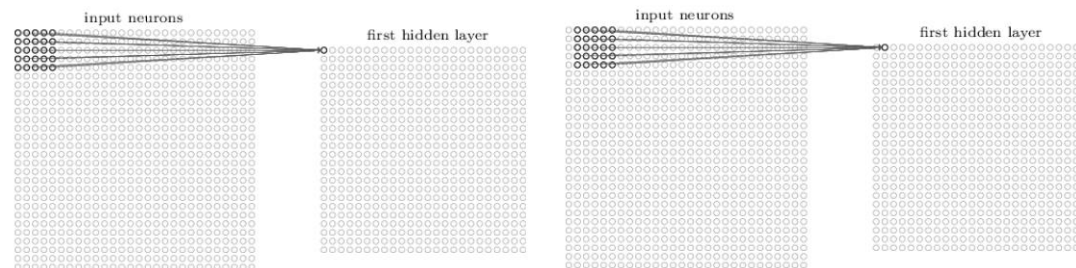


Third Layer

卷积神经网络有三个基本概念：局部感知域 (local receptive fields), 共享权重(shared weights) 和池化 (pooling)。

卷积层：

局部感知域： 在神经网络中输入层是用一列的神经元来表示的，在 CNN 中，将输入层作为二维矩阵排列的神经元（卷积核）。

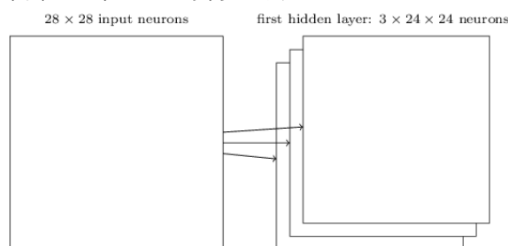


共享权重： 上面得到的第一隐藏层中的 24×24 个神经元都使用同样的 5×5 个权重。第 j 个隐藏层中第 k 个神经元的输出为：

$$\sigma(b + \sum_{l=0}^4 \sum_{m=0}^4 w_{l,m} a_{j+l,k+m})$$

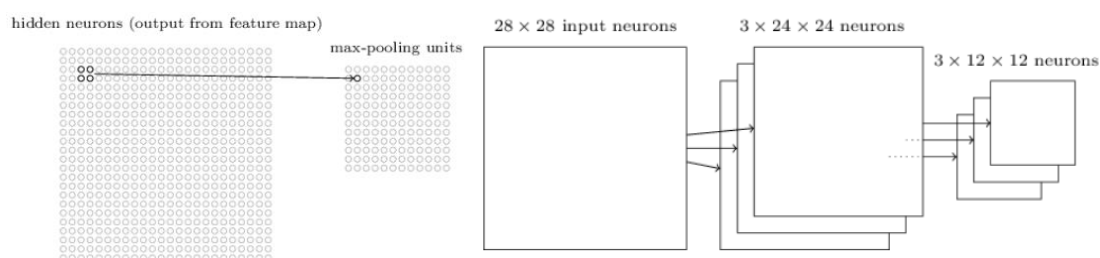
σ 是神经元的激励函数（可以是 sigmoid 函数、tanh 函数或者 rectified linear unit 函数等）。 b 是该感知域连接的共享偏差。 $w_{l,m}$ 是个 5×5 共享权重矩阵。 $a_{x,y}$ 代表在输入层的 x,y 处的输入激励。

为了做图像识别，通常需要不止一个的**特征映射**(filters, kernels)，因此一个完整的卷积层包含若干个不同的特征映射。

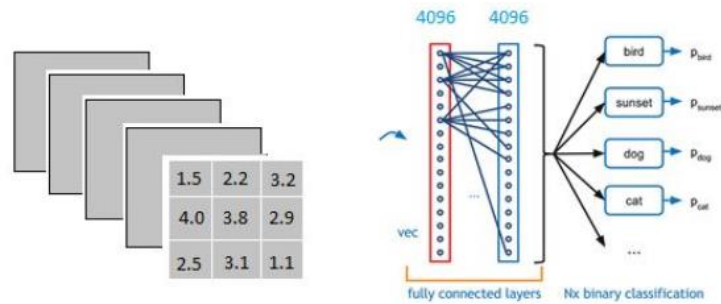


***线性整流层：** 使用线性整流 (Rectified Linear Units, ReLU) 作为这一层神经的激励函数。它可以增强判定函数和整个神经网络的非线性特性，而本身并不会改变卷积层。其他的一些函数也可以用于增强网络的非线性特性，如双曲正切函数 tanh 或者 Sigmoid 函数。相比其它函数来说，ReLU 函数更受青睐，这是因为它可以将神经网络的训练速度提升数倍，而并不会对模型的泛化准确度造成显著影响。

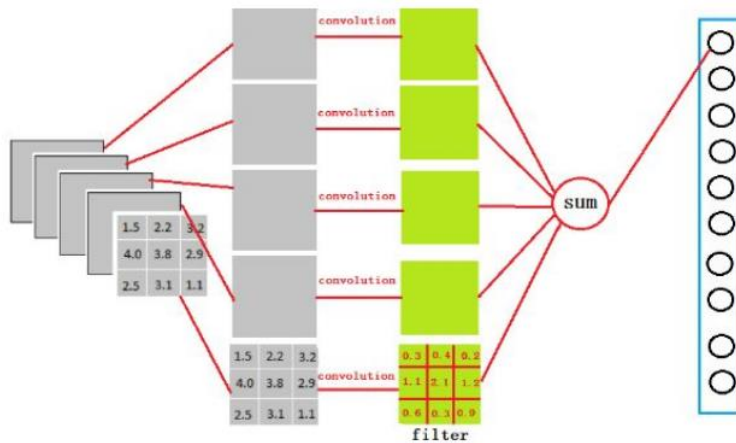
池化层： 通常紧随卷积层之后使用，其作用是简化卷积层的输出。经常使用的方法有 max-pooling, average-pooling, mean-pooling



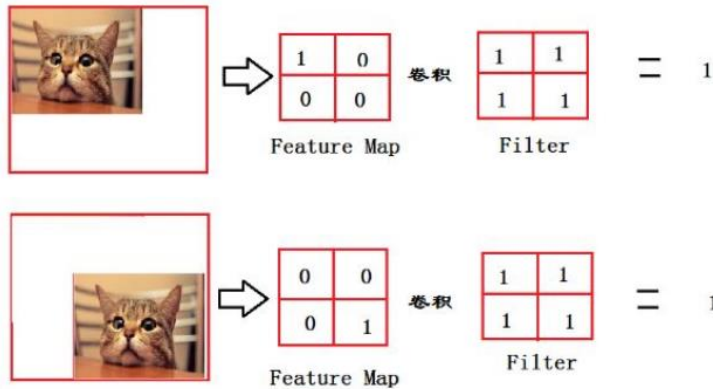
全连接层：该层的每个神经元都与最后一个 pooling 层的每个神经元连接



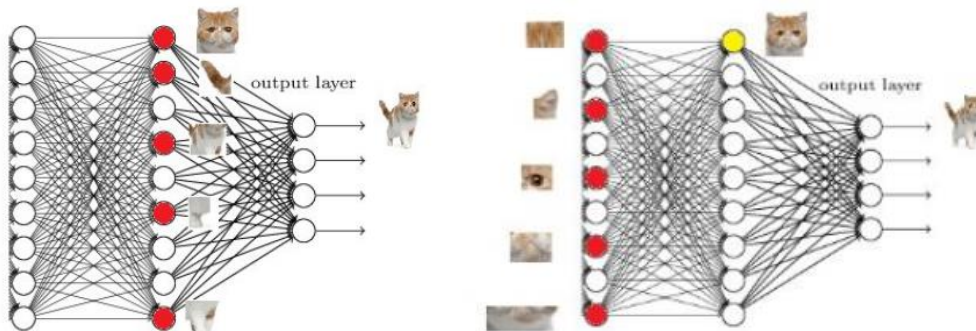
3x3x5 的输出转化为 1x 4096，进行卷积



全连接层卷积的作用：把分布式特征 representation 映射到样本标记空间（整合到一起，输出为一个值，减少特征位置对分类带来的影响）



多个全连接层的作用：全连接层中一层的一个神经元就可以看成一个多项式，我们用许多神经元去拟合数据分布，但是只用一层 fully-connected layer 有时候没法解决非线性问题，而如果有两层或以上 fully-connected layer 就可以很好地解决非线性问题了。



***损失函数层：**用于决定训练过程如何来“惩罚”网络的预测结果和真实结果之间的差异，它通常是网络的最后一层。各种不同的损失函数适用于不同类型的任务。例如，SoftMax 交叉熵损失函数常常被用于在 K 个类别中选出一个，而 Sigmoid 交叉熵损失函数常用于多个独立的二分类问题。欧几里德损失函数常用于标签取值范围为任意实数的问题。