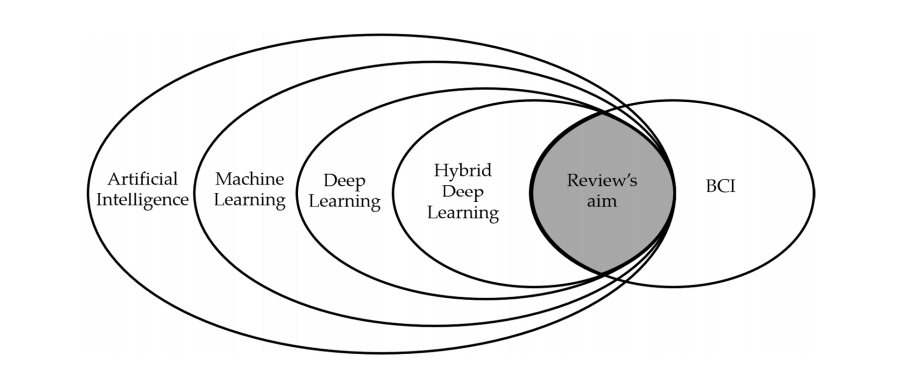
一个实例：HDL运用于脑机接口——概述与CNN技术简介

撰写：张杭磊

一、概述：

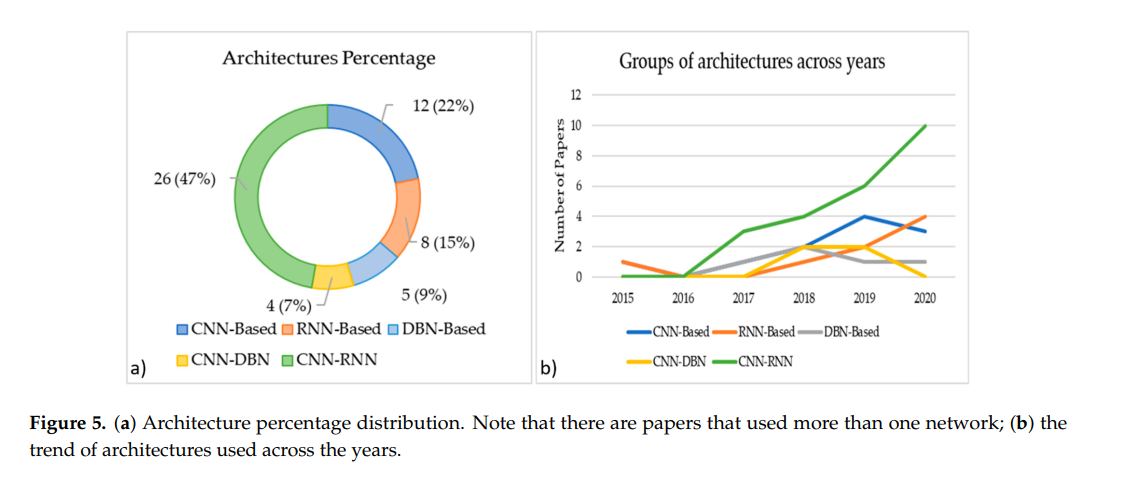
脑机接口（BCI）技术作为一项连接交互人脑与机器的技术，势必要求人工智能技术领域所能提供的提取、解析能力来开发其应用领域、改善其使用效果。而相比上面提及的支持向量机（SVM）技术外，人工智能领域中的深度学习（D相关技术由于其在简化计算量、筛选特征、分类效果等方面上的优势，能够在实际应用中发挥更好的效果，故而近年来成为了脑机接口领域中更为常用的人工智能技术。



人工智能及其下属（混合深度学习）领域与BCI的交互

（图源Nibras, A. A. Hybrid Deep Learning (hDL)-Based Brain-Computer Interface (BCI) Systems: A Systematic Review）

深度学习是“是一种以人工神经网络为架构，对资料进行表征学习的算法。”（Wikipedia），而人工神经网络种类众多，如深度神经网络（Deep Neural Networks, DNN）、卷积神经网络（convolutional neural networks，CNN）、循环神经网络（recurrent neural network, RNN）等，深度学习故而可以根据具体的网络而区分成不同种类。报告主要涉及的是深度学习中以卷积-循环神经网络（CNN-RNN）为构架的混合深度学习（hybrid deep learning，HDL）。作为一种混合了多种神经网络的深度学习，混合深度学习能够将不同神经网络的优势相结合，取得整体大于部分的优化效果。

混合深度学习实际上是一个统称，CNN-RNN只是其中一类，还有如CNN-DBN等也可称为混合深度学习。而之所以介绍CNN-RNN是因为它在脑机接口领域有极高的应用价值和较为优秀的成果，在许多论文中被加以采用，如下图所示：

（图源Nibras, A. A. Hybrid Deep Learning (hDL)-Based Brain-Computer Interface (BCI) Systems: A Systematic Review）

左图中展示了近年脑机接口-深度学习领域有关论文中所采用的神经网络架构的比例分布，其中CNN-RNN占了将近一半，而右图也表明CNN-RNN在近年愈来愈受到研究关注。故而CNN-RNN可以认为是深度学习在脑机接口领域最具有代表性的一种方法。通过了解CNN-RNN，我们可以对深度学习是如何应用于脑机接口的过程有一个较为初步但是全面的认识。

以下将分为三个部分对CNN-RNN在脑机接口领域的应用进行介绍，首先是CNN与RNN技术内容的分别介绍，两个部分之后，再对它们的混合网络整体如何作用于脑机接口领域进行介绍。

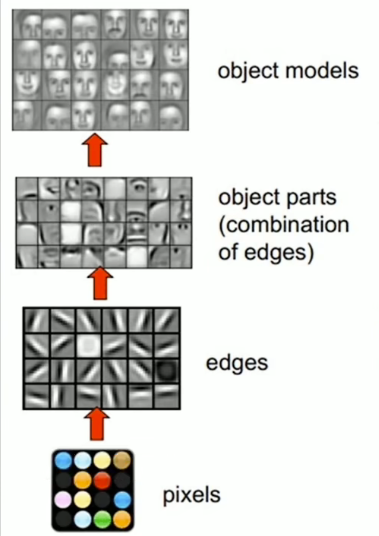
二、卷积神经网络（CNN）

1.概括与原理

图示

描述已自动生成

（图源：LeCun et al. Gradient-Based Learning Applied to Document Recognition）

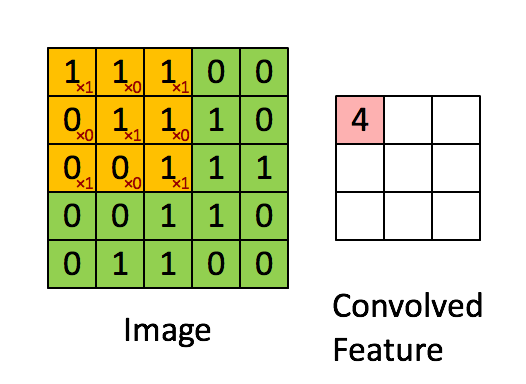
卷积神经网络的作用可以形容为“认照片”，是一种常用于对静态图像进行特征提取与分析的人工智能深度学习网络。其主要构成有卷积层（convolutional layer）、池化层（pooling layer）、全连接层（fully Connected layer）。其设计思想主要来自于对生物视觉处理的模仿：

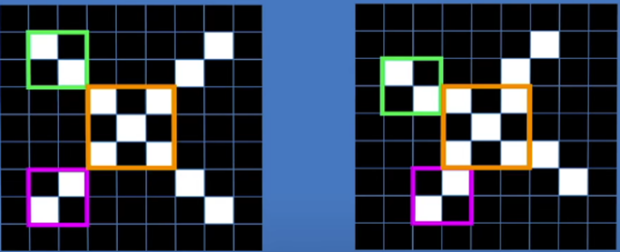
(图源：Lee H, Roger Grosse R, Ranganath R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations)

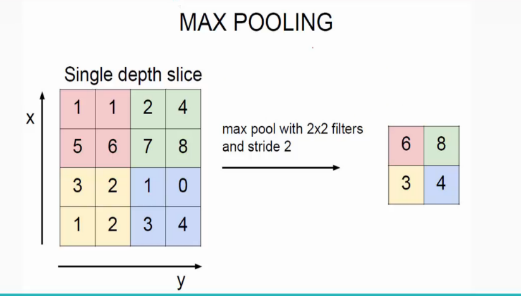
以人脸识别为例，一张照片由许多像素（pixels）组成，眼睛接受像素点的反射光，经过视网膜视神经初步的处理之后，零散的像素被组成许多零碎的点、线元素，或称为边缘（edges）；再之后，更高级的神经中枢将零散的edges进一步综合，得到了五官等视觉成分（object parts）；最后在皮层等最高级的中枢，所有的视觉成分被综合成人脸的视觉，我们由此识别出容貌。当我们希望计算机也能够拥有“视觉”时，我们可以模拟这一过程来设计电脑的思维机制，即把以上的过程抽象为“提取轮廓—突出特征—呈现形象”的反复操作，而这种反复正是卷积神经网络所进行的操作，在架构的物理层面上，由“卷积层-池化层-全连接层”实现这些操作。

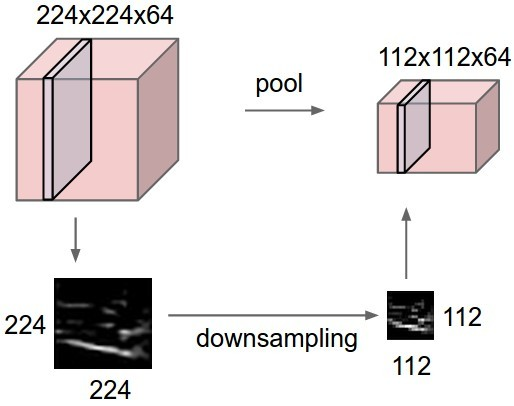
2.具体组成介绍

（1）卷积层

图像在计算机中的存储在一般表示为二进制的数组，如0表示黑1表示白的像素点（色彩的情况较为复杂暂不提及）。如动图所示卷积层的主要过程，卷积操作相当于将原来的image通过一个固定大小、数值特定（经过恰当的参数设置）的filter，格子对格子相乘，让image变成更小的convolved feature。这一过程实际上模拟的是视网膜对像素的初步处理，即将局部的特征（比像素更大一些的轮廓特征）分块提取，图片的细节变得模糊而平滑，而特征得以识别。

以具体例子而言，人眼识别右图的扭曲x，实际上是将右图的部分局部特征与左图的标准x进行对比（如都有四角的末端、中心的x结构），而不是按照像素的一一对应进行对比（因为二者的像素层次上并非黑白完全相同）。而卷积层就是在仿照这样的流程，对图片进行分块特征提取，起到了挑选局部特征的作用。（以上图片来自网络）

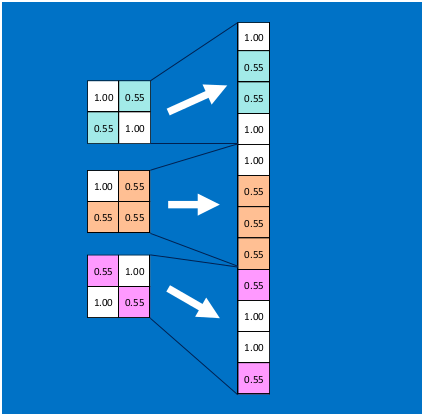
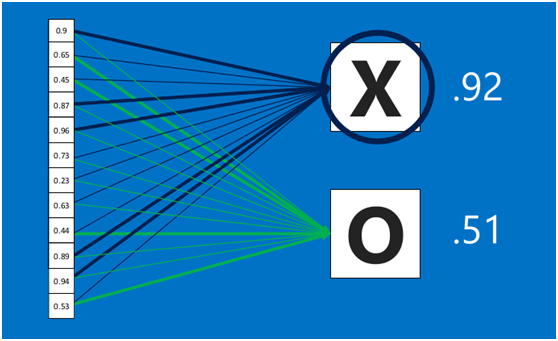
（2）池化层

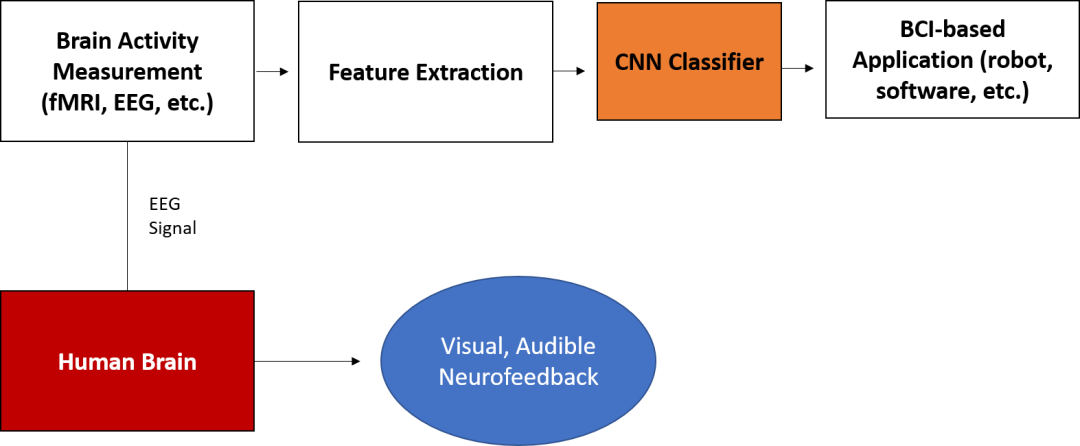
池化的作用让人联想到压缩图片尺寸：像素降低了，但图片的特征轮廓保留了。在数据上，常用的池化有Max Pooling，亦即打捞一个区域的最大值放入小池中。形象的例子是一张尺寸较大的图片，我们通过减小它的清晰度，使它的尺寸下降，但是图片中的线条依旧能够辨认（黑色的部分相对不重要而被缩小），这就是相当于池化中的选取最显著的特征，集中于更小的池中。池化具有的作用由此可见是减小数据的维度、排除冗余影响，进而让整体的计算量减小，提升了计算的速度。

其实这一过程也有常识的参照，我们可以考虑人类对形象的认识，我们认识一种轮廓往往也是抓大处放小处，最终轮廓留在记忆之中，之后就能更有效率的识别物体。

（3）全连接层

经过多次的卷积和池化处理后，我们得到了相对而言简单、尺寸较小的数据组，它包含着一张图片的特征信息，而为了让电脑读取和识别这样的信息，我们有必要将其转化成更为底层的形式，而计算机的这种形式往往是一个一维数组。全连接层的作用就是将特征信息的数据组转化为这样的一维数组。经过学习，计算机可以将“标准图案”经转化为标准的一维数组储存，在之后传入训练的图片时，也可以将训练图转化为“训练数组”，与“标准数组”进行比较，按照其中数据的相似程度，挑选最可能的“标准图案”作为辨认的结果。如下图，一张训练图转为数组之后，假设要辨认它为“X”还是“O”则可以将这个数组与“X”和“O”对应的标准数组进行相似度的对比，结果上与“X”的相似度接近100%而“O”只有50%左右，故可以判断这一训练图一般为“X”。





3.BCI实践

我们已经大致上理解了CNN网络的运作过程，当我们联系到BCI领域时，我们不难发现这项技术对BCI中的许多内容可以发挥作用。经由BCI我们能获取倒EEG等脑电的图像化信息，通过分析和学习，神经网络能寻找其中模式化的特征并进行分拣，这就是CNN这样对静态图像具有较高特征信息分拣辨认能力的神经网络的用武之地。

然而RNN仅仅作为一种静态的分析手段，具有较大局限性。由于脑部活动在时间上具有先后次序以及连续性，我们必须在动态的时序上分析考虑脑电的特征，故而我们引入了RNN弥补缺陷。