

浙江大学

本科周报告

姓 名:	赖梓林
学 院:	计算机科学与技术
系:	计算机科学与技术
专 业:	计算机科学与技术
学 号:	3170104684

目录

第一章 摘要	3
第二章 正文	4
1.卷积神经网络	4
1.1 卷积神经网络定义	5
1.2 卷积神经网络历史和发展	7
1.3 卷积神经网络应用	7
1.4 卷积神经网络组成	7
1.4.1 卷积层	7
1.4.2 池化层	8
2.UNet	9
2.1 UNet 的简介	9
2.2 UNet 的结构	9
2.3 UNet 的优点	10
2.4 UNet 的策略	6
3.其他问题	11
3.1 数据增强相关	11
3.2 泛化能力相关	11
3.3 迁移学习	11
3.4 其他讨论	12

摘要

报告简要地总结了卷积神经网络(Convolution Neural Network)和一种着重于解决图像分割问题的卷积神经网络结构：U-Net 的相关基础知识。

二、正文：

1. 卷积神经网络(Convolution Neural Network)：

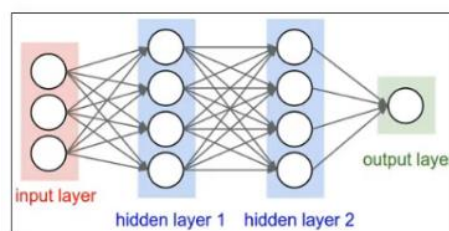
1.1. 卷积神经网络定义：

卷积神经网络是深度神经网络中的一种，通常由以下几种层结构组成：卷积层(Convolution Layer)，池化层(MaxPooling Layer / AveragePooling Layer)，标准化层(Normalization Layer)和全连接层(Fully-Connected Layer)。

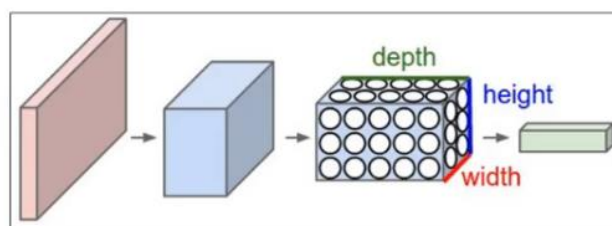
由于图像处理和影响处理等领域，输入层以图像中所有像素点作为输入，使得所需的神经元数量急剧上升，需要学习的参数数量也急剧上升，此时，前馈神经网络已经无法满足需求。转而采用卷积神经网络。

From NN to CNN

- A regular 3-layer Neural Network.



- A ConvNet arranges its neurons in three dimensions (3D output volume)



<http://cs231n.github.io/convolutional-networks/>

图 从神经网络到卷积神经网络

其中不包括全连接层的卷积神经网络网络称为全卷积网络。卷积神经网络通常用于图像识别问题等，输出层之前是全连接层，结果输出一个向量。而在图像分割等问题中，单单输出向量无法满足要求。

所以，在全卷积网络中，通过卷积层代替全连接层实现上采样操作，输出和输入图片等大的图片。由于舍弃了全连接层，需要学习的参数数量也会大幅减少。

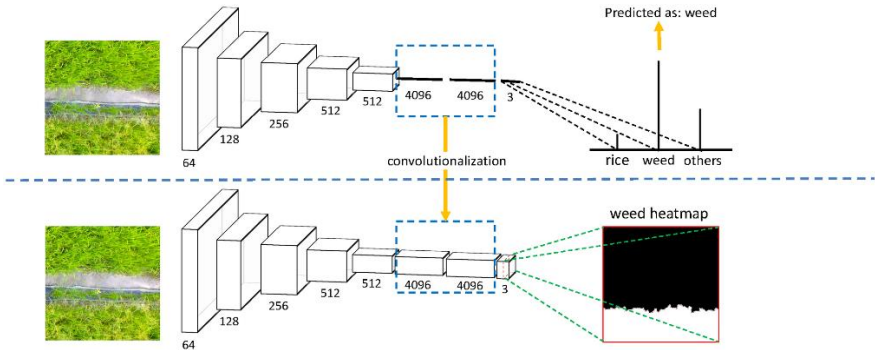


图 全卷积网络（下）

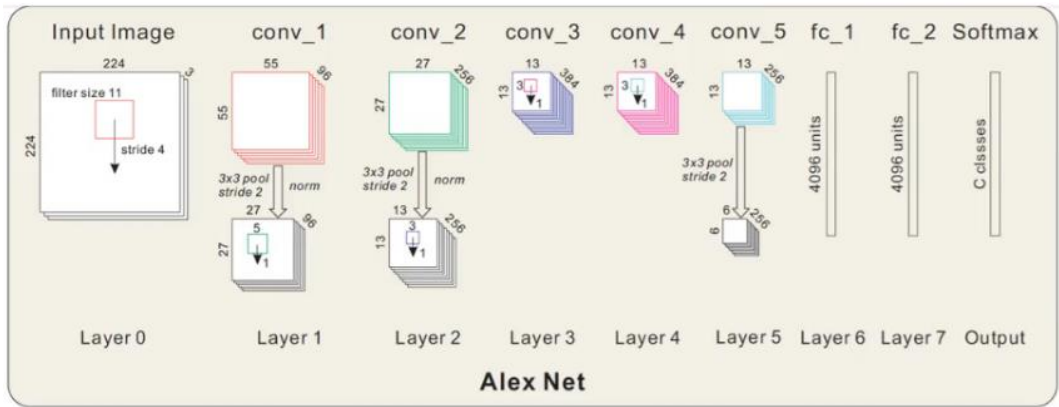


图 全卷积网络 ALEXNET 的结构

1.2. 卷积神经网络的历史和发展：

在 1994 年，由 Yann LeCun 提出的 LENETS，最早的神经网络之一，推动了深度学习领域的发展，经过了多次成功的迭代之后命名为 LeNet5。

2010 年，Dan Claudiu Ciresan 和 Jurgen Schmidhuber 提出了最早之一的 GPU 神经网络，他们使用一块 NVIDIA GTX 280 GPU 实现了九层结构的神经网络。

2012 年，Alex Krizhevsky 发表了具有更深更广的层结构版本的 LENETS，命名为 ALEXNET，这个网络以远超第二名的准确率赢得了 ImageNet 的比赛。

LENETS 和 ALEXNET 被称为卷积神经网络中两种突破性的结构。

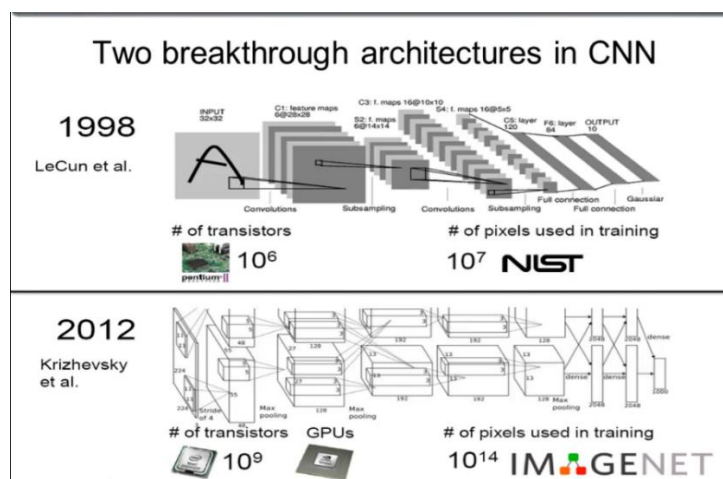


图 LENETS 和 ALEXNET 的对比

随着互联网的发展，越来越多的非互联网产业也开始涉及这一领域。许多在过去通过纸面、书面记录的数据转而通过网络储存，这使得经由计算机的数据量暴增。对于人工智能来说，这些数据都是潜在的学习样本。另外，计算机硬件的发展使得过去昂贵的 GPU 变得越来越亲民，计算机算力的限制逐渐变少。这两点都是促进人工智能相关产业发展的关键。

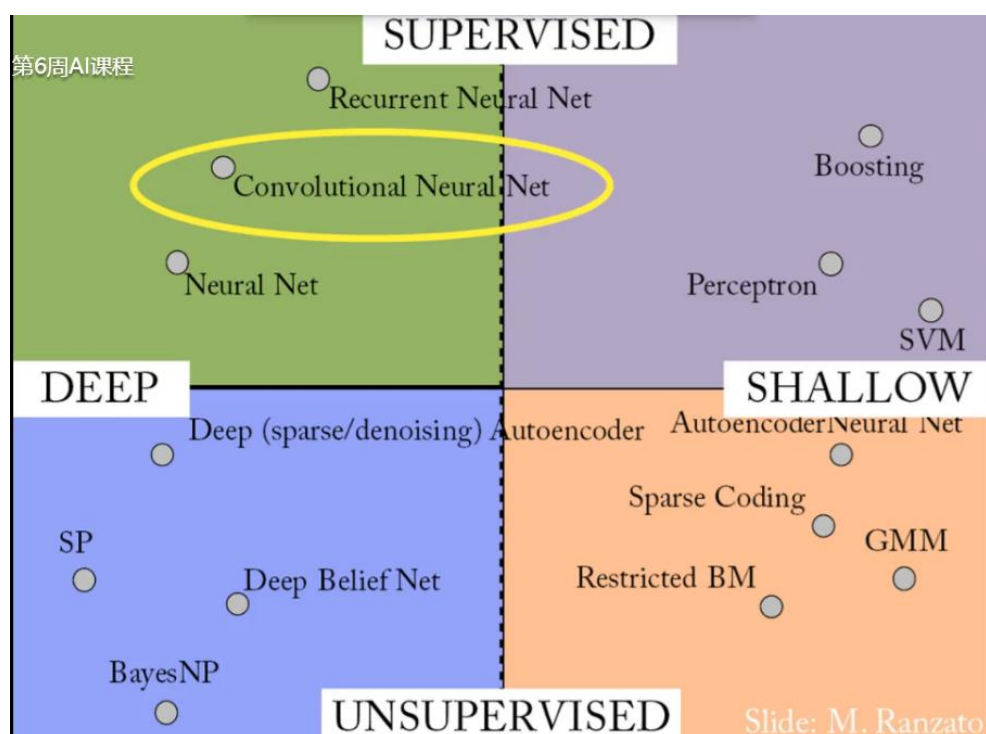


图 卷积神经网络的定位

1.3. 卷积神经网络的应用：

卷积神经网络弥补了前馈神经网络在图像领域的不足。图像和影像具有空间相关/依赖性，每个像素点都与邻近的像素点有关。通过卷积操作，可以从像素点中抽取主要特征，得到特征地图（Feature Map）。常用的操作是采用多个卷积核，将一张三通道的彩色图片变换为多张特征地图。

Convolutional Neural Network: convolution

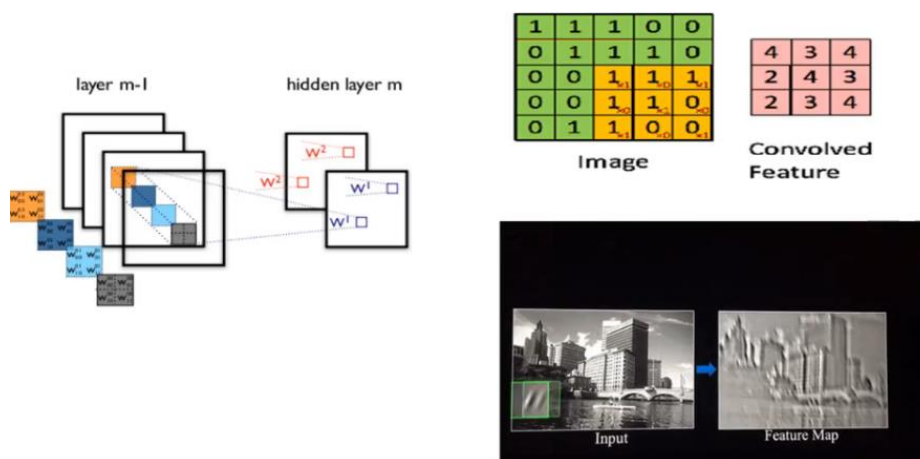


图 卷积操作范例

通过卷积和池化等操作后，和全连接层第一层相连的个数显著减少，达到了维度削减的目的。卷积后的结果进行非线性变换（Relu 等），降低了分辨率，实现下采样，减少信息冗余，得到更有区别意义的数据。

1.4. 卷积神经网络的组成：

卷积神经网络常包含的组件有：卷积层，池化层，标准化层（近期相对不太热门），全连接层。下面主要讨论较为重要的卷积层和池化层。

1.4.1 卷积层：

卷积层是卷积神经网络的核心组件，类似图像处理中的滑动窗口，卷积层主要工作通过一组卷积核完成。卷积核每次接收的只有图像的一部分，但整个图像都会被经过。卷积核在宽度和高度上卷积，计算卷积核和接收数据的点积，生成激活地图（Activation Map）。

卷积层具有三个超参数，分别是深度，步长和填充。

深度通常取输入图像的通道数，步长指卷积核在输入图像中滑动时，每次移动的像素数量。步长为 1 时，像素点将会被多次重复计算，产生大量输出。步长设定为 3 或以上的情况较少见。填充参数的设定将会影响输出图像的大小。

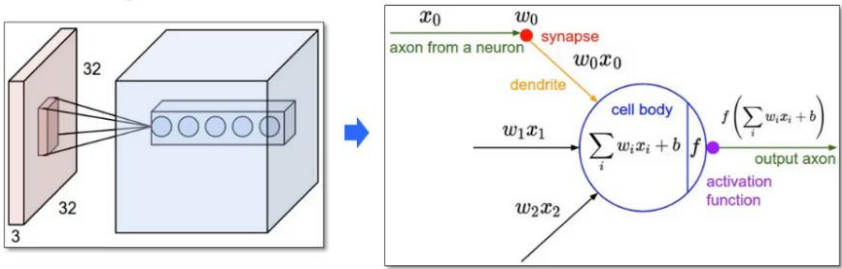


图 卷积层

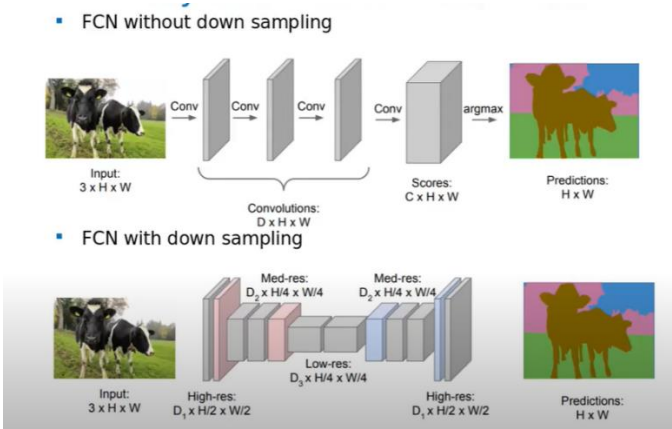


图 下采样

1.4.2 池化层：

卷积层是卷积神经网络另一个重要的组件。在图像中，比起一个特征的准确位置，更重要的是这个特征与邻近的其他特征的相对位置。池化层就是基于这个概念工作的。常在连续的卷积层中插入池化层。

池化层的工作同样由核 (Kernel) 来完成，但卷积层计算点积，而池化层用最大值点或点的平均值等代替中心点。

常见的池化层组件有最大化池化层和平均池化层，其中最大化池化层属于非线性函数池化层，平均池化层属于线性函数池化层。

2. UNet:

2.1. UNet 的简介:

UNet 结构是由德国 Freiburg 大学计算机部门开发的一种全连接神经网络结构。开发时主要着重于解决生物医学领域图像分割问题。但近期逐渐被广泛用于各种图像分割相关领域。

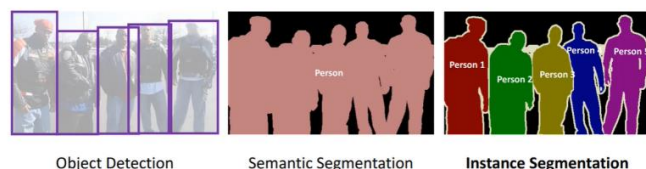


图 不同的图像处理任务

UNet 比较特别的点是包含了一条收缩路径(contracting path)和一条扩展路径(expanding path)，且通过数据增强，UNet 可以更有效率地使用标注样本。在较小的训练数据集的训练下，UNet 仍有良好的表现，并且耗时极低。

UNet 在比赛中的表现非常好，准确率远超历届最高水准和同届的第二名，并且对于不同的分割应用都有良好表现。

2.2. UNet 的结构:

UNet 首先属于全连接神经网络，不包含全连接层。

UNet 具有一条收缩路径，和与之相对的扩展路径。整体的结构形似“U”，所以称为 UNet。在收缩路径上实施下采样，图像的分辨率不断下降，而在扩展路径上实施上采样，分辨率上升。

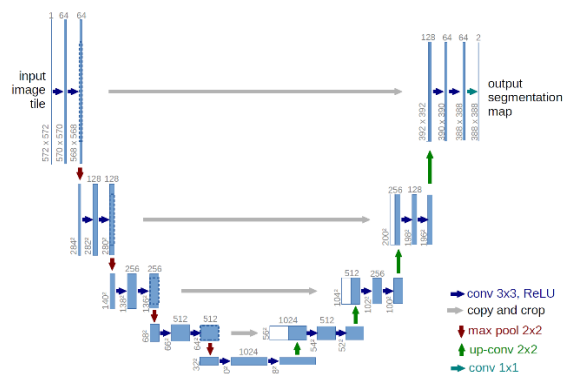


图 UNet 的具体结构

Unet 结构包括了卷积神经网络的核心组件卷积层和池化层，除此之外，添加了分割层，拼接层和上采样层（也属于卷积层）。

其中细节部分，池化层部分使用最大值池化层，在原论文中未采用填充策略，但在有些讨论中指出启用填充策略后能取得更好的表现。

2.3. UNet 的优点：

因为并不存在一种对于所有问题都有良好表现的结构，需要根据问题决定特定的结构。在计算机视觉领域，卷积神经网络在图像识别、图像分类的问题中表现良好，而 UNet 则注重于图像分割。

在普通的卷积神经网络中，数据不断通过卷积层、池化层，分辨率的下降使得偏底层的数据，如边界和边缘等逐渐模糊、丢失。最后只剩下特征信息。这样无法达到图像分割的目的。

随着卷积层的深入，我们获得特征信息，失去语义信息。而在这种情况下，因为忽视了局部和整体关系，只关注于局部的特征信息，会出现歧义。

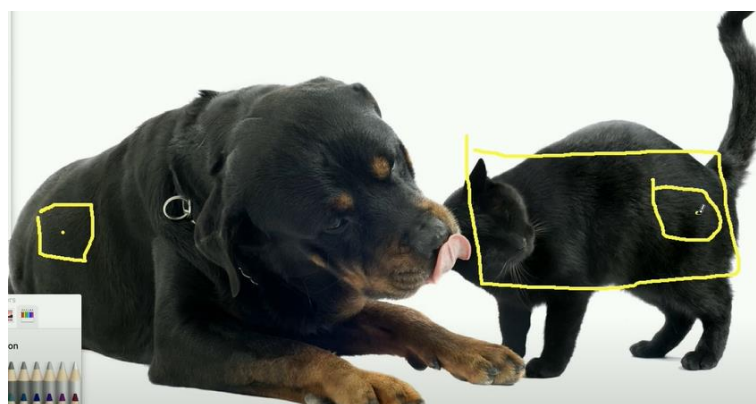


图 歧义问题

但在 UNet 中，每个进一步的下采样操作前都会将图片裁剪并和扩展路径上采样的图片拼接在一起，这样保证了既有深层的特征信息，又保留了浅层的语义信息，特征信息和语义信息的结合完善了整个网络的表现。

同时扩展路径取代了直接的上采样，在结果上来说相当于网络最后会学习到最好的上采样方式。

2.4. UNet 的策略:

文章中提到 UNet 的良好表现也得益于它充分使用了数据增强策略。

对于显微图像来说，主要需要的是平移和旋转不变性，以及对灰度变化和图像变形的鲁棒性。文章提到，在数据增强步骤中，对训练样本实施随机弹性变化，可能是以较少训练样本训练分割网络的关键。

他们在 3×3 的网格上施加平滑化变形，以 10 个像素为标准的高斯分布进行采样位移。之后通过双三次插值算法计算每个像素的位移，随后在收缩路径结尾的丢弃层(Dropout Layer)进一步执行隐式的数据增强。

3. 其他问题:

在学习过程中，也遇到了一些主题之外的值得关注的点，在以下的小节中简要讨论。

3.1. 数据增强相关:

数据增强是训练网络中关键的一步。主要的实现方式是人为的改变图像的像素分布，得以学习到鲁棒性更强的特征信息。对于不同的应用，常用的数据增强手段包括：旋转、放大、亮度调整、添加噪声、添加丢弃层等。

3.2. 泛化能力相关:

当网络对于训练数据表现良好，但训练数据以外表现不好时（过拟合问题），我们需要提高网络的泛化能力，常用的手段包括：添加噪声、预训练、数据增强、添加丢弃层、丢弃连接层等。

3.3. 迁移学习:

迁移学习指对于同一个网络，假设在某个领域中训练、学习后表现良好，现要迁移至另一个领域。迁移学习主要调整靠近输出端的层，保留已完善的部

分，让其他部分训练、学习。常用于大数据驱动下学习到的神经网络应用至小数据的任务当中。

- Generalization:
– Transfer learning

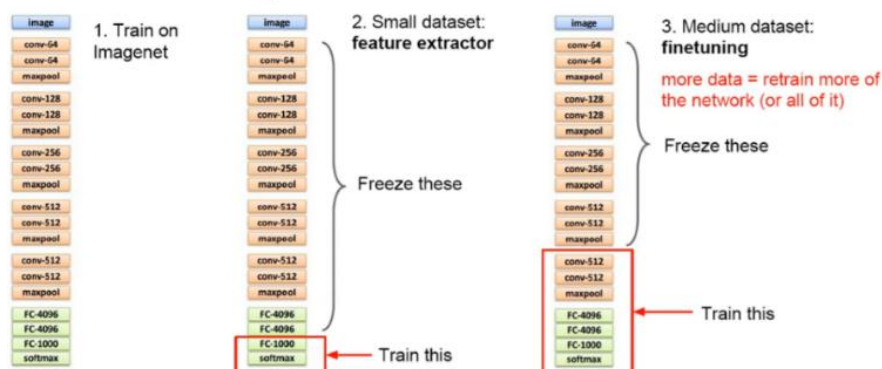


图 迁移学习

3.4. 其他讨论:

网络的训练策略很大一部分取决于我们对性能和算力/时间开销的平衡。

对于 ALEXNET 来说，ALEXNET 邻近输出层的两个全连接层使得整个网络需要学习的参数骤增。

有研究指出：ALEXNET 去除第二个全连接层后，性能仅下降 1.1%，需要学习的参数下降了一千六百万个。去掉两个全连接层后，性能下降 5.7%，需要学习的参数下降了五千万个。去掉第三、四层（卷积层）时，降低了 3.0%的性能，但仅减少了一百万个参数。而去掉第三、四、六、七层时，整体性能直接下降了 33.5%。

从这个研究中我们可以看出，网络的深度是性能的核心指标。这一点从最近热门的残差神经网络也可以看出。层数越多，对原始数据特征区分的表达能力也越强，非线性变换更多，处理复杂数据的能力就更强。