

报告总结

郑自强

这次参加天津大学的深度学习暑假研讨班，四位讲师给我们带来的精彩的报告，我也对深度学习有了更深刻的认识 and 思考，其中以刘昕博士和赵凯同学带来的报告对我震撼最大。

刘昕博士首先给我们描述了深度学习的历史和发展过程，经过他的梳理，我对深度学习有了更深刻的认识。同时他也分享了一些 caffe 的实战经验，其中怎么设置 solver 参数是当前我们参加 imagenet 比赛所迫切需要的。同时我也深切感受到深度学习是一个深刻和值得人思考的一个新兴领域，当前我所学习到的还只是皮毛，还有很多知识值得我不断学习和探索。经过这次长达两天的报告，我也有了我自己的一些认识。首先，深度学习虽然现在迎来了第三个发展的高峰期，但是它远远没有达到那么智能，换句话说，他还有很多的潜力和发展空间。现在的深度学习就是通过一个模型和网络去实现一个功能，通过监督学习和无监督学习等方法来调整模型的参数让其更具有泛化能力，具有更强的鲁棒性。

下面是我自己这两天来的总结和体会：

1. 在图像识别和语音检测方面，深度学习较以前的 SIFT, HOG 等方法有了更好的效果，因为深度学习依赖监督学习或无监督学习，让模型自己去调整自身参数更好的拟合。至于深度学习的具体过程我有了一个较为浅显的理解，以图像识别为例，对于一张图片，其包含的信息就是其每个格点的像素值，对于一张照片，我们能很轻松的辨认和“学习”它包含的信息，而计算机看到的只是一串储存在不同空间的数据，为了让计算机能看懂图片，我们可以类比我们学习辨认照片的过程，当我们接触一件新的事物时，我们首先记住的是它的外形特征，它的边缘形态，然后了解这个东西是什么？它是属于什么 label，再然后就是它所包括的语义和特征。以长颈鹿为例，当我们第一次看到长颈鹿是印象最深刻的就应该是它的长长的脖子，当我们逐渐熟悉长颈鹿这个类别后，当别人提起长颈鹿时，我们在脑海里就能演化出长颈鹿的形态，再然后就是想到长颈鹿是植食动物等信息。

类比人类认知事物的过程以及人类认知事物有不同的分层次的特点，深度学习首先要去认识事物的特征，第一步就是通过不同的方法去提取事物的特征，在 caffe 这个深度学习框架中，我们通过卷积去提取事物特征，将事物的特征更抽象，更具有代表意义，然后通过一个激活函数，达到阈值的为 1，其他的都为 0，然后为了我们要得到事物的更深层次的特点，我们可以可以通过卷积和激活函数提取特征，然后通过分类器给提取出的分类特征加上 label。那深度学习是通过什么学习到的事物特征呢？损失函数是解决这个问题的关键，为了解决单层感知器不能解决的异或问题，我们引进 hidden 层，通过多个 hidden 层和 input 层，output 层的连接实现拟合事物特征。万能逼近定理证明了只要 hidden 层足够多我们就能逼近这种非线性特征。那我们如何去拟合这些非线性特征呢？我们首先初始化 input 层和 output 层之间的权重（初始化不均匀会造成拟合速度很慢或者不能拟合）。然后通过不断的训练数据去更新权重，损失函数就起到了这样的作用，我们通过反向传播和梯度下降更新调整权重，通过不断的奖惩来改变参数和权重。起到更好的拟合，但是为了防止模型对训练数据过拟合，我们引进 dropout。加入正则化系数，让模型有更强的鲁棒性。在这里，正则化系数共有 $L1$ 和 $L2$ 两种。我们还可以通过 dropout 防止过拟合，每次训练只让部分隐含层与输入输出层相连。这样权值的更新不再依赖于有固定关系隐含节点的共同作用，阻止了某些特征仅仅在其它特定特征下才有效果的情况。总之我们加入这种种参数都是为了防止模型对特定的情况过分拟合而泛化性不强。当我们将模型训练好之后就是用它去提取一张图片的特征，一张图片的特征最底层、最基本的就是它的边缘特征，然后就是它的 object，最高层的特征就是它的 background。以西瓜为例，它最底层的特征就是它是圆形的，它的边缘特征是弧形的，更高层的特征提取就是如果我的计算机检测到一个西瓜放在不同的位置能提取出不同的信息，根据不同的背景特征做出场景分类。解决这一问题我感觉十分困难，因为模型中的算法是我们预先设计好了的，它是否能够满足所有的实际情况，当没

经过训练的特征出现时，我们的模型是否能学习到新的特征。还有一个问题就是物体的背景会对物体的特征提取造成干扰，怎么才能实现最小化这些干扰。

当模型对物体的特征提取实现后，接下来模型对物体进行分类，caffe 中的分类算法是 softmax，具体的实现过程我现在还不是很明白，我的理解是它将提取到的特征用 0，1 表示，然后特征进行矩阵计算得到一个分类器。

听完了刘博士的讲解，我有一些疑问，就是我们知道从不同的角度看一个物体，一个物体会有不同的特征，caffe 是对图形就行卷积操作，那么如何能够实现这种不同角度的提取特征呢？另外就是赵凯分享的他们进行的一个识别数字的一个项目，他们所做的数字识别是建立在很多局限的基础上的，如何才能在自然环境下对物体进行识别和检测。

2. 这次报告对我一个很好的帮助就是我有利于我捋清 caffe 的框架，知道它的实现过程，caffe 最底层数据是 blob，每一层之间的数据传输和新建数据都是用 blob 进行封装传递。经过刘昕博士的整体介绍，我大致了解了 caffe 的具体实现流程。也自己写了一个怎么看 caffe 源码的总结，经过看源码，里面的很多实现细节现在也更清晰了，我接下来要做的就是先将 caffe 中的一些数学公式是怎么实现的推导出来，然后再联系代码加深印象，进一步了解实现细节。另外我感觉 caffe 是一个实践学科，许多调参都是建立在不断的尝试和纠正改错的基础上的，并没有实际的理论支持，许多调试工作都是经验性的。经过这一个星期的源码分析，真的感觉 caffe 是一个很好的框架，它对函数的封装和实现以及它内部的一些联系都是做的十分好的。所以坚持看 caffe 源码和做注释能让我学到许多知识，还有就是俞老师介绍的 Qt 也是一个非常帮的调试工具，它能够帮助我们理清各个函数之间的关系以及互相之间的依赖关系。是一个很好的帮助我们阅读代码的工具。

3. 关于深度学习，我有一些疑问，就是当模型对物体特征不能很好的拟合时，我们的模型能不能自己做出调整，或者表现得更加智能，而不仅仅表现得像一堆代码和算法的合成。就是深度学习能否能自己从物体上提取特征，而不是我们先给它设定好卷积核的大小，中间各个层之间的连接。当遇到不能拟合的情况，机器又能否放弃使用这个模型去拟合，转而用其他方法去实现学习。类比我们人的思考过程，我们接受外界信息的刺激从而引起神经元兴奋，然后通过强化某些神经元之间的连接以实现学习和记忆，那么类比机器学习，我们能不能将机器学习的某些基础的很小的功能进行封装，然后使用遗传算法对机器学习进行用进废退的选择，让他拥有更好的学习能力，能根据具体情况实现自己的调整，而不是我们首先预设一大堆参数。

这次参加天津大学的暑期深度学习研讨班，我收货了很多知识，我明白了一些我在跑 caffe 程序时候遇见的一些错误的原因以及 caffe 的一些基本的调参操作，对怎么进行深度学习的深度学习有了更清晰的认识。听了这些讲者的报告，我也对深度学习的发展历史和近况有了更多的了解，我感觉深度学习现在还不是那么想象中的智能，首先它的实现是有很多的限制条件的，而且开发出一个新的网络模型现在看来也是很难的，所以深度学习还有很大的发展空间和潜力。我自己也制定了一些计划，我认为我最大的收货就是我知道以怎样的顺序和方法去学习 caffe，学习深度学习了。我现在要做的就是要多看博客和教程了解原理和实现方案，然后多跑程序和调节参数积累经验。

2016.7.20