

ccf-cv 报告总结

郑自强

今天的报告主要是关于深度学习的一些介绍和最新技术.今天在场的三维嘉宾都针对他们各自所进行的科研活动和取得的成绩进行了详细的讲解.

(1)第一位嘉宾主要讲解了三个方面的内容,分别是特征选择,模型学习,知识迁移.首先特征选择必须要具有泛化性和不变性,要具有更好的鲁棒性.我们选择的特征不能过度拟合我们的训练数据,必须要具有泛化性,要更好地表达好体现我们的样本数据.第二就是模型学习,建立一个端到端的学习体制,使网络具有学习到一种模型的能力.第三知识迁移,我们的网络在已经学习到某一种特征和模型的基础上,再去学习另一种新的模型和特征,能够利用现有的已经学习到的知识,将学习上一次特征和模型的方法运用迁移到现在来,通俗的讲,就是已学的知识对现在起到一个指导的作用.

接下来该嘉宾谈到了深度学习下对于大数据进行分析和获取信息是必要和必须的,正是因为大数据时代的到来才推动了深度学习的发展,让深度学习能够学习到更多的特征和信息,保证了我们学习到的模型具有更好的泛化性和鲁棒性.但由于深度学习计算复杂度高,包含大量参数,所以上世纪八十年代深度学习由于当时硬件设施的落后而陷入寒冬也不足为奇了.另外我们可以模拟大脑认知的机理,可以设计我们的网络不同层之间的连接,如跨层连接,全连接.总之我们可以借鉴生物工作机理来设计我们的网络和优化算法.另外他还提到了分层学习的概念,将具有相似特征多的几个样本放在一个大类中,然后在这个大类中再不断细分,然后再不断检测在这个大类中的不同点,实现进一步分类学习.这样可以减轻我们的计算量.提高我们的分类速度.

最后该嘉宾提到了 rnn 对于时序的依赖问题,由于最近我正在编写 rnn 和 lstm,所以我对这方面的内容比较注意,rnn 主要是为了处理序列数据中的时间相依问题,在序列信号中,某时刻的信号对以后的信号造成影响.怎么让 rnn 实现这些功能?这些还需要我不断编程和调试积累经验.

(2)第二位嘉宾主要是带来了关于图片检索和语义检索,并阐述了当今社会记录信息的主要方式是视频,所以我们要针对视频进行分类和提取信息,将各个视频的类别通过深度学习的方法分出来并加上标签,它的另一个实际用途还包括鉴定是否与其他视频雷同,在保护个人视频版权的方面也会做出贡献.另外这位嘉宾提出了一个概念我觉得十分赞同,就是我们在处理数据的时候可以考虑数据之间的关系,可以针对数据之间的相似性关系来进行进一步的分类和检索.传统的提取特征的方法有 HOG,Harr 等方式,但是在数据量很大的时候这些方式提取特征所取得的效果并不好,所以为了提高提取特征的效果,我们要使用深度聚类的方法,通过无监督学习和高度非线性方式去拟合图片特征,从而取得较好的效果.接下来姜老师谈到了他们最新的工作主要是人脸校验,并阐述了他们的主要算法是条件玻尔兹曼机,提到了一些图像变换和流形可视化的一些基本概念.最后他对自己所从事的实际工作进行了介绍,从他的介绍中可以看出,深度学习已经较好的运用于实际生活中,而不仅仅只有理论价值.

(3)第三位嘉宾主要讲述了一些调参和具体操作实现方面的经验,他同时也提到了一些比较好的深度学习方式,如 rnn 和 cnn 进行融合,我的初步想法是当我把 rnn 和 cnn 的代码全都写完和实现后,我也要尝试一下将 rnn 和 cnn 进行融合,看是否取得的效果比只用其中一种所取得的效果更好.还有一个比较吸引我的概念就是集成学习,所谓的集成学习,就是用多重或多个弱分类器结合为一个强分类器,从而达到提升分类方法效果。严格来说,集成学习并不算是一种分类器,而是一种分类器结合的方法。从上面三位嘉宾所谈到的内容可以看出,深度学习中一个很重要的概念就是融合,将不同的算法和模型进行融合以求取得更好的结果,将不同算法的优势拼接在一起来实现更快,更准,更有效.最后该嘉宾提出了现在深度学习所处理的数据量还是不能达到实际需求所要求的数据量,只能在小范围的使用,如果在更大范围和更大量级的数据上进行操作,那么所取的效果并不好,所以深度学习发展的空间和潜力还非常大.

个人感悟:

听了上面三位嘉宾带来的报告,我主要有以下想法,首先深度学习绝不停留在理论阶段,它和实际运用结合得十分好.其次我这次所关注的点主要有这些

1.rnn 和其变种 lstm 关于帧与帧之间的时空依赖性,怎么样才能取得更好的结果,除了增加时序间的信息交流和储存以往时序的信息之外还有没有更好的方法来取得更好的记忆效果?我个人的想法是能不能借鉴光流的实现机制,在时序中像素变化最剧烈的地方勾勒和提取物体运动轨迹.

2.rnn 和 cnn 之间的模型融合,具体的实现机制还要以后慢慢写程序不断摸索.

3.集成学习和以及它的实现原理运用.