

## 天津大学深度学习暑期研讨班会议总结

天津大学深度学习暑期研讨班主要目的是为从事深度学习研究的博硕士等提供一个高水平的集中研讨交流会。天津大学作为主办方也邀请了来自国内深度学习一线的博硕士和技术专家参会。本次会议主要是中科院计算所刘昕博士，上海大学研究生二年级赵凯同学针对深度学习相关知识做了详细，深入的报告。我的此次总结，首先对刘博士及赵凯同学讲解的内容进行简单的介绍，然后对个人的收获和体会进行总结。

刘昕博士第一个报告的题目是《深度学习基础》，顾名思义，刘昕博士的第一个报告是对深度学习的基础知识进行了系统的阐释和说明。刘博首先从计算机视觉说起，讲到了离散和连续两类 task，并用一副漫画解释了传统的两段式方法从特征确定，特征选择到特征分析等解决问题的过程及两段式方法在计算机视觉领域“包治百病”的地位。以此作为转折，刘博引出了深度学习的端到端式的解决问题的方式，进一步引出了深度学习这一概念。刘博首先解释了深度学习的定义：以不少于两个隐含层的神经网络对输入进行非线性变换或表示学习的技术，包括 RNN，CNN 等多种结构。而深度学习在图像分类，人脸识别及物体检测领域等领域带来的提升，则是让我们为之振奋的，刘博对此也列举了详细的数据。从这些数据中，我们也可看出深度学习网络逐渐变大变深，数据量不断增大的趋势。接下来，刘博讲到了深度学习给我们带来的新端到端的层级特征学习，并以 image，text，speech 为例，解释了深度学习的层级性特征。除了层级性外，刘博也简单介绍了感受野的概念。接下来，刘博介绍了人工神经网络研究的发展进程，讲述了各个时期神经网络的结构特点及存在的缺陷。

接下来，刘博介绍了深度学习的基础理论。刘博首先讲到的是基础结构单元，对基础结构单元中的全连接层，卷积层，pooling 层等进行了详细的讲解。其中，刘博重点讲解了卷积层。从多个 Feature Map 的计算，误差反向传导，卷积层梯度的更新讲起，用 caffe 为例，讲解了多通道卷积的快速实现。而后，刘博详细讲解了 RNN 的结构特点，形式及缺陷等问题，而 LSTM 则解决了 RNN 无法建模长时间依赖的问题。而后，刘博对 LSTM 的核心思想，结构组成，及应用等问题进行了详细的解释说明。在对基础结构单元进行了详细的解释之后，刘博进一步讲解了深度学习的损失函数和训练方法。在对深度学习进行充足解释之后，提出了深度学习存在的两个理论性问题：一、激活函数导致的梯度消失，并提出了解决梯度消失问题的策略。二、梯度溢出问题。在报告的最后，刘博推荐了一系列可供阅读的课本、博客、公众号等。

刘昕博士第二个报告的名字是：《深度学习一线实战暑期研讨班 CNN 基础与 Caffe 实践》。从报告的名称可以看出，本次报告的主题是 CNN 和 Caffe。

刘博首先讲到了 CNN 早期结构的演化，刘博从 1989 年 LeCun 结构讲起，到 1998 年的 LeNet 再到 2012 年的 AlexNet, 2014 年的 VGG-Net，一直讲到了 2015 年的 MSRA-Net。刘博详细地讲解了每种结构的组成，做出的改进，实现的功能，存在的问题等，带我们了 CNN 网络的发展历程。在 CNN 结构演化的增强卷积模块功能的部分，刘博重点讲解了 GoogLeNet，其中对多个 Inception 结构串联和两个辅助损失层进行了详细的讲解。并重点推荐了 Inception 结构。在对 Inception V3 的讲解中，刘博提到了将  $5 \times 5$  卷积层拆成 2 层的  $3 \times 3$  卷积层，这样则使得计算量变小，而深度加深，而进一步的加入  $1 \times n$  和  $n \times 1$  的卷积和，则进一步地减少了计算量，增加了深度。但是，在具体地实践操作中，到底要不要替换成更“瘦”更深地网络，要自己尝试后确定，实践中的结果往往无法和理论值完美匹配，所以要不断地试错。而后，R-CNN，SPP-Net 则实现了从分类到检测地跨越，Fast R-CNN，Faster-R-CNN 则大大加快了检测地速度。最后，刘博简单提到了 FCN、DeepLab、ST-Net、CNN+RNN/LSTM 等新的功能模块。

在讲到的 caffe 部分，刘博首先对 caffe 的安装，概念，数据准备等基本的概念进行了简单的讲解。而后，详细地讲解了 caffe 的使用中的定义网络结构部分，对 data 层，卷积层，pooling 层，全连接层及特殊功能层，在此之后，刘博推荐了 Python 定义网络结构，Python 定义网络结构是一个可视化网络结构地在线工具。可用 python 代码生成每个 layer，可构建任意深度地参差风格。除此之外，刘博还推荐了一些值得一看地 DL 开源平台和 caffe 分支。

在 CNN 实战方面，刘博也以 caffe 为例，进行了系统的讲解。在此不再细说。在开放问题讨论部分，刘博提出可以通过 Batch Normalization & Residual Net 来解决收敛困难的问题。而在“高瘦”网络是否一定优于“矮胖”网络的问题上，答案也是否定的，具体性能的优劣要视情况而定，是否使用 DL 也要视情况而定。在最后，刘博提出，要学会 DL 的思想，适应 DL 带来的新变化，但是不要只会 DL，数学和统计机器学习是至关重要的。

中科院计算所的邬书哲博士分享了目标检测的相关知识。邬博士从目标检测的任务，评价指标，基本检测思路，主要的检测方法各类检测器的比较几个方面简单介绍了目标检测的基础知识。而后，他简单介绍了新的检测框架，候选窗口生成方法，R-CNN、Fast R-CNN 的功能和性能分析，并列举了 FASTER-RCNN 实用案例，并对现在方法存在的问题，提出了一系列的解决方案。为了更快速地实现目标检测，邬博士提出了减小单个 CNN 的复杂度，采用多个级连的方式。他以人脸检测为例，展示了窗口分类网络，边框校准网络，MT-CNN 等。在最后的的部分，他展示了人脸检测的案例，展示了基于 Faster R-CNN 的人脸检测器。同时，他也提出了现实存在的各种困境。A

上海大学研究生二年级的赵凯同学今年在 cvpr2016 上发表论文一篇，他也

在此次会议上做了报告。他报告的题目是《全卷积网络的细节与应用》，他首先讲到了如何将非线性不可分模型通过二维、三维的映射转变成非线性可分模型，在此他提到，在低维空间中很难处理的问题，到高维空间中后，可能会变得更简单。接下来，赵凯同学详细地讲解了卷积与反卷积，在这一部分，他强烈推荐了用矩阵乘法做卷积地方式。在 FCN 应用地方面，他也给出了很多例子。在接近结尾部分，他展示了他们课题组做的电表数字识别地项目，强调了感受野的重要性。最后，他推荐了一个很好用的在线编辑器：Jupyter.

对本次研讨班内容进行简单的介绍之后，下面谈一下我参加本次研讨班的收获。

一、 带我认识了深度学习。以前总是听大家讲深度学习，对深度学习的具体内容却不是很了解。这次刘博士的深度学习基础讲的东西都很实在，很详细，对深度学习进行了深入浅出的介绍，尤其对深度学习为我们带来的新的机遇和挑战的介绍，让我对深度学习产生了浓厚的兴趣。这是我认为对我来说最重要的收获。让我从茫然无知阶段步入了开始对某个方面有了浓厚的兴趣，并想要开始去深入地了解，实践。

二、 刘博士，邬博士和赵凯同学都指出了他们在实践过程中出现的各种问题以及新手开始深度学习的学习时可能会遇到的问题，并给出了一系列的解决方案。例如刘博提到的鞍点的问题，梯度消失、梯度溢出问题，以及赵凯提出的反卷积后，与输入不对等的问题，需加大 padding 等，这些问题的提出及解答为我以后自己做东西都会有一定的帮助。

三、 会议报告者推荐了很多优秀的工具，很实用的结构及开发平台等。刘昕博士为我们推荐了很多适用的结构，像他推荐的很好用 python 定义网络结构，还有一些 DL 开源平台和 caffe 分支以及可供阅读的书、论文、微博等以及赵凯推荐的实用工具和计算方法等。我计划本次报告结束后，结合刘昕博士的《深度学习基础》，对深度学习做一个深入的了解，对我以后要做什么方向，具体做什么事情做一个思考。

四、 此次报告，刘昕博士最让我印象深刻。他对深度学习研究得很深入，很透彻，对各方面的知识了解的都很多，读的，做的东西也很多。刘昕博士是身边的榜样，要向他学习，多读，多写，多做。其次，刘昕博士非常的风趣幽默，把古板的也讲解得很有趣，这是在表达上要向他学习的地方。

总结：这次报告，这三位讲者都讲的很细致，在一些具体的问题上，也做了详细的讲解。虽然听懂的东西有限，但是我个人觉得是收获最大的一次，听了优秀的人的报告，也认识了很多深度学习领域的同学、老师，对我来说还是觉得收获很大的。