**深度学习导论课程第六小组期末报告**

71117201姜子玥 71117206刘桂伶 71117204刘雪珂

71117223周威屹 71117122刘欣鹏

**摘要：**本学期的深度学习课程中，小组完成了数据集视频描述以及两次上机实验，实验基于卷积神经网络的相关算法，本篇报告中将介绍本学期完成实验包括实验相关原理、实验算法、实验结果分析、实验感想以及小组各位成员在本学期中各项工作的贡献量。

**1 引言**

深度学习是加拿大多伦多大学Hinton教授等人[1, 2]于2006年提出的一种新的机器学习方式，其将无监督的逐层初始化(Layer-wise Pretraining)结构和深度神经网络(Deep Neural Networks: DNN)结构进行了有效的结合。深度学习被《麻省理工技术评论》杂志评选为2013年世界十大技术突破之一，吸引了学术界和工业界的广泛关注，在语音识别和图像识别等众多领域[3-5]取得了突破性进展。卷积神经网络[6-8]因为建模难度适中且性能优异，逐渐成为深度学习中众多学习结构中被广泛采纳的一种结构。限制卷积神经网络大规模应用的很重要的一个因素就是其计算复杂度，尤其是其中线性卷积的计算[9-11]。Jaderberg等人[9]提出用低秩矩阵分解的方法来加速线性卷积的计算。Liu等人[10]提出同时结合低秩矩阵分解和稀疏约束的方法来加速线性卷积的计算。Vasilache等人[11]则提出在GPU环境下运用最简单的基-2 FFT[12]算法来加速线性卷积的实现。但是正如Liu等人[10]指出的虽然目前大型的卷积网络系统大部分是在GPU环境下实现，但是在普通的CPU环境中仍然有其自身的通用性的优势：基于CPU的系统能够方便地部署在目前通用的商品化集群系统（commodity clusters）中，而不需要任何特殊的GPU结点。因此研究CPU环境下卷积神经网络的构造也是一件十分有意义的研究工作。

本学期深度学习导论的课程中，我们小组在老师的教导下了解到了深度学习的基本知识，包括卷积神经网络、多层神经网络、线性神经网络等的基础知识。在动手实际过程中，我们小组完成了制作数据集、两次上机实验、制作课程PPT的任务。

本文将描述数据集以及两次实验的实验内容，包括实验原理、实验算法、实验结果分析、实验感想、参考文献。本文剩下结构组织如下：第2节描述三次作业的实验内容，第3节解释两次上机实验的实验原理以及实验算法；第4节对实验的结果进行分析；第5节中表达本学期参与实验后的实验感想。第6节将介绍小组成员对本学期所有工作的贡献以及占比。

**2 实验内容描述**

**2.1 作业1数据集的制作**

小组中每人观看600个视频，描述视频中画面内容，并对视频进行分类，得到相关数据集。

**2.2 作业2**

通过预处理数据、添加标签和提取c3d特征，调整c3d的维度训练模型网络，最后再利用代码中的测试代码来测试训练得到的网络模型。最后将生成测试语句提交到平台上得到本次实验的分数。

**2.3 作业3**

StarGAN 网络可以在一个数据集中进行多领域的图像转换任务，比如更换头发颜色、更换表情、更换年龄等。对数据集进行简单的预处理后，使用数据集对StarGan网络进行训练，再用专门的测试数据对模型进行测试，将结果压缩后以组为单位提交。

**3实验原理及算法**

**3.1作业2原理以及算法**

3D CNN主要运用在视频分类、动作识别等领域，它是在2D CNN的基础上改变而来。由于2D CNN不能很好的捕获时序上的信息，因此我们采用3D CNN，这样就能将视频中时序信息进行很好的利用。3D神经卷积网络相比于2D神经卷积网络，将时间维度看成是第三维，这里是对连续的四帧图像进行卷积操作，3D卷积是通过堆叠多个连续的帧组成一个立方体，然后在立方体中运用3D卷积核。在这个结构中，卷积层中每一个特征map都会与上一层中多个邻近的连续帧相连，因此捕捉运动信息。而3D卷积核只能从cube中提取一种类型的特征，因为在整个cube中卷积核的权值都是一样的，也就是共享权值，都是同一个卷积核（图中同一个颜色的连接线表示相同的权值）。我们可以采用多种卷积核，以提取多种特征 。

作业2中提取数据集的标签并利用数据集视频得到c3d特征，利用该特征训练卷积神经网络。

实验使用3D Convolutional Neural Networks for Human Action Recognition框架进行实验，涉及到了卷积神经网络算法、3D CNN。

**3.2作业3原理以及算法**

StarGAN原理：判别器D对真假图片进行判别，真图片判真，假图片判假，真图片被分类到相应域。生成器G接受真图片和目标域标签并生成假图片。生成器G在给定原始域标签的情况下将假图片重建为原始图片（重构损失）。生成器G尽可能生成与真实图像无法区分的图像，并且通过判别器D分类到目标域。

**4 实验结果分析**

**4.1作业2实验结果分析**

本次实验经过7次提交，得到评分0.0249。与其他小组对比，结果并不算好，进过分析，小组成员认为原因是由于实验有时间限制，而在此时间限制下设置的训练模型的维度只有200，而这200层也训练了整整一天，也是我们小组成员不够了解相关知识，没有尝试在配置更高的机器运行模型训练，导致模型训练不够，从而结果不够理想。

**4.2作业3实验结果分析**



实验二对模型的测试使用了1999个图像进行了五个不同的变换，分别是发色变黑、发色变金、发色变棕、改变性别、变老。上图为部分测试结果，可以看出最终学习出的模型是很成功的。

**5 实验感想**

**5.1作业2实验感想**

通过本次实验，小组成员对相关代码的调试更加熟练，同时对于相关模型训练更加理解。但通过本次实验也暴露了小组成员们的缺陷，我们对于相关模型训练不够熟练，在之前的步骤浪费了一些时间，导致后续没有足够的时间训练足够多层数的模型，导致结果不够理想；同时，我们不会善于使用工具，没有利用老师提供的算力卡提高训练的效率或者在配置更高的机器上运行训练模型。这些问题会给我们在之后深度学习相关内容的动手操作过程提供宝贵的经验。

**5.2作业3实验感想**

本次实验需要较高的计算机配置，小组成员个人的计算机无法达到配置要求，所以我们借用学校实验室的计算机，通过远程连接实现实验二的所需操作。由于我们对训练网络所用时间估计错误，导致尽管对时间充分利用，我们也只是刚好完成了实验二的必做任务，没有多余的时间去完成实验二的选做任务。我们推测如果一开始选择使用算力卡就不会有这方面的问题。由此可以看出小组成员对相关知识的掌握略有欠缺，我们会在之后相关内容的学习中对这方面更加注意。

**6成员工作**

**6.1作业1数据集的制作**

由每位成员共同参与完成，所有成员的工作量相同

**6.2作业2：实验一3D CNN实验**

由组长姜子玥单独完成该项作业

**6.3 作业3：实验二**

由组员刘雪珂和刘桂伶共同完成

**6.3课程答辩PPT制作以及演讲**

由组员刘欣鹏和周威屹共同合作完成，刘欣鹏负责了课程的答辩演讲

**6.4团队报告的撰写**

由组长姜子玥、组员刘雪珂、刘桂伶共同合作完成

**6.5实验一二程序说明文档**

由完成该实验的三位同学姜子玥、刘桂伶、刘雪珂各自完成

**6.6团队贡献占比**

组长姜子玥：33%

组员刘雪珂以及刘桂伶均占33%

组员周威屹、刘欣鹏均占33%

参考文献

1. HINTON G E and SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]*. Science*, 2006, 313: 504–507. doi: 10.1126/science.1127647.
2. HINTON G E, OSINDERO S, and TEH Y W. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527–1554. doi: 10.1162/neco.2006.18.7.1527.
3. BENGIO Y, COURVILLE A, and VINCENT P. Representation learning: a review and new perspectives [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2013, 35(8): 1798–1828. doi: 10.1109/TPAMI.2013.50.
4. LECUN Y, BENGIO Y, and HINTON G E. Deep learning [J]. *Nature*, 2015, 521(7553): 436–444. doi: 10.1038/nature14539.
5. DENG L and YU D. Deep Learning: Methods and Applications[J]. *Foundations and Trends in Signal Processing*, 2014, 7(3):197–387.
6. LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278–2324. doi: 10.1109/5.726791.
7. KRIZHEVSKY A, SUTSKEVER I, and HINTON G E. Imagenet classification with deep convolutional neural networks [C]. Advances in neural information processing systems*,* Lake Tahoe, NV, United States, 2012: 1097–1105.
8. SZEGEDY C, LIU W, JIA Y Q, *et al*.. Going deeper with convolutions [C]. IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, United states, 2015: 1–9.
9. JADERBERG M, VEDALDI A, and ZISSERMAN A. Speeding up convolutional neural networks with low rank expansions [J].*Computer Science*, 2014,4(4):XIII.
10. LIU B, WANG M, FOROOSH H, *et al*. Sparse Convolutional neural networks [C]. IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, United States, 2015: 806–814.