

毕业设计(论文)开题报告

**学生姓名：** 匡乐 **学 号：** 20164484

**学　　院：**计算机信息与工程学院　

**专 业：** 计算机科学与技术

**设计(论文)题目：** 基于深度学习的图像分类研究

**指导教师：** 周惠斌

　　2020年 03月 29日

|  |
| --- |
| **文 献 综 述** 课题背景 人工智能是近几年的火热领域，在社会中受到很大的关注，因此发展速度迅速。借助人工智能，人类已经能够自动化的处理主观化、非规范化的任务，例如图像识别，这在以前的计算机是难以做到的。人工智能出现时间在上世纪，但真正受到人们广泛关注还是在近十年，这主要源于深度学习在许多人们以为计算机无法做到的领域做出了重大突破，例如AlphaGo战胜人类，生物识别领域在日常生活的广泛应用。吸引到了更多的组织和个人投入到这个蓬勃发展的领域。  计算机擅长抽象化和形式化的任务，远远超过人类能力，在语音和图像识别领域近十几年才达到了人类的水平。一个人在成长过程中要获得海量数据，才能在生活行动中体现出知识的存在。借鉴于人类的神经系统，计算机学习也需要海量数据才能表现出智能化，因此学者们的挑战就是将主观的、非形式化的知识教会给计算机。  通过海量数据学习，构建层次化结构拟合事物的规律，层次之间通过较为简单的方式连接，这样的方式避免了人类给计算机指定学习内容，这往往是无效的指定。借助类似人脑神经元的网络结构，加上反向传播的学习方法，计算机能够利用结构简单的模型学习到复杂的特征，这就是人工智能的一个分支——深度学习。 课题意义 Kaggle社区是面向全球的一个机器学习竞赛平台，提供有奖金竞赛，编写和分享代码，在平台中可以获得丰富的前沿知识分享和有挑战性的技能训练。在Kaggle平台竞赛中取得良好的成绩意味着自己架构的人工智能系统在全球学习者之间都是优秀的，能够客观的评价一个人工智能系统的准确度。  如今借助互联网的便利，在网络上学习是一种除学校之外重要的学习途径，有开放的交流平台，多种的学习和练习途径，原本需要耗费大量精力的数据集采集工作，可以在网络上分享得到，CIFAR-10数据集就是一个经典的图片数据集，常常被用作十分类学习任务的训练和测试任务。在该数据集上需要对网络结构和学习方法精心的设计才能达到满意的测试结果。是对技术锻炼很好的方法。 理论依据 训练神经网络需要有海量的数据喂入，该数据集有6万张图片，是10类别动物图像，基本满足数据量的需求。  检测图像类别，不是由整体决定的，而是一些局部区域特征决定；对于不同的图像，如果有相同的特征，那么可以用相同的检测模式去检测不同图像的相同特征；对于一个大图像，对其下采样，图像的基本特征不变。  卷积神经网络是深度学习在计算机视觉使用的最为广泛的一种技术，相较于手工设计需要的图像专业性相比，其自动学习参数更加智能，充分利用上述图类别像检测的特点。卷积层和池化层对数据特征进行提取，加快对高对比度图像的学习时间成本。在层与层之间非线性激活函数可以拟合出复杂的特征映射。  学者通过大量实践发现，网络深度的增加，能够显著提高模型精度，深层的网络需要更加精细的学习方法来调整。因此训练更深并且能够学习训练的网络成为当前学术的一个难点。 研究方法和研究内容 课题的研究方式是通过参加Kaggle竞赛的方式，评估模型，不断迭代优化模型，从而获得最佳的网络结构。  利用CIFAR-10数据集训练一个卷积神经网络，然后在Kaggle平台中提交测试，获得评分，不断的优化神经网络的结构和学习方法，模型的选择很难一步做到最适合状态，迭代升级不断调试超参数的选择，超参数正交化手段有利于对最优参数的调整。  对得到的数据集划分为三个部分，train\_set、dev\_set、test\_set，train\_set用来训练学习，模型在train\_set上获得很高的准确度后在dev\_set上验证其拟合能力，如果发生过拟合很严重，需要返回上一步，调整模型结构，降低过拟合，提高在dev\_set上的准确度，最后评估模型好坏的需要在test\_set上测试，其准确度标志着模型的泛化能力。  采用对比的方法，选择最适合的网络和超参数。设置训练多种网络，例如VGG16、VGG19、ResNet等等，比较其训练结果在相同测试集上的误差。对于同一种网络，通过调整超参数（学习率、学习衰减率、卷积核尺寸和数值等等），尝试出最适合的网络结构。 可能存在的问题 在深度学习中网络的规模很大程度上决定了模型的拟合能力，训练规模越大越深的网络是学者们追求的目标，越大的网络其需要学习的参数越多，例如GoogLeNet网络深度达到22层，参数个数6百多万，ResNet层数可达上百层，参数上亿，训练这样的模型需要有性能优秀的计算机和计算能力超强的GPU，这样的实验条件不容易达到，对此的解决的办法是采用迁移学习，获得训练效果良好的网络，对其局部冻结，训练部分。这在计算机视觉中是一种常见的学习方法，往往有很好的效果，也能节省精力，关注于对模型的调节。 预期结果 其预期目标，利用迁移学习训练出针对CIFAR-10十分类任务的多层卷积神经网络模型，需要在Kaggle平台CIFAR-10竞赛中获得较高的测试准确度。对训练过程中的结构调整能够总结出调整过程，流程数据可视化，便于时刻对模型能力的变化进行观察。 工作进度安排 设计时间安排前期先做专业知识储备学习，包括深度学习知识，编程知识，数据分析知识几个部分，然后是初步建立模型，对模型评估分析，迭代调整，提高模型能力，对过程中的数据进行可视化分析，并且记录。得到合适的模型后，汇总模型学习过程，拟写论文。 参考文献 [1] Matthew D. Zeiler and Rob Fergus（2014）: Visualizing and   * Understanding Convolutional Networks. In David Fleet, Tomas Pajdla, * Bernt Schiele, & Tinne Tuytelaars, eds. Computer Vision – ECCV 2014. * Lecture Notes in Computer Science. Springer International Publishing, * 818 – 833. * [2] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner（1998）: Gradient-based * learning applied to document recognition. Proceedings of the IEEE 86, * 11 (November 1998), 2278 – 2324. * [3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton（2012） : * ImageNet Classification with Deep Convolutional Neural Networks. * In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger, eds. * Advances in Neural Information Processing Systems 25. Curran * Associates, Inc., 1097 – 1105. * [4] Visual Object Classes Challenge 2012 VO(2012). * [5] Yann LeCun. Leon Bottou, Yoshua Bengio, and Patrick Haffner. * Gradient-Based Learning Applied to Document Recognition [EB/OL].http://vision.stanford.edu/cs598\_spring07/papers/Lecun98.pdf,1998. |
| **指导教师意见**： |
| 1．对“文献综述”的评语：  该学生的选题具有前瞻性，是在专业方向上的进一步的延申，紧扣专业领域的热点问题，做到理论与实际很好的结合。该生具有完成选题的能力和条件，对于本课题进行了较为深入的调研，参考了大量文献。最后确定的选题具有一定的现实意义和理论价值。  2．对本课题的深度、广度及工作量的意见和对设计（论文）结果的预测：  本课题是学生所学专业知识的延拓，符合学生专业发展方向，对于提高学生的科研能力大有裨益。难度适中，学生能够在预定时间内完成该课题。  指导教师：  年 月 日 |
|  |