

文本复制检测报告单(全文标明引文)

№:ADBD2018R_2018053015312720180530154844440174212949

检测时间:2018-05-30 15:48:44

检测文献: 53141116_左丰源_计算机科学与技术(网络与信息安全专业)_基于深度学习的脊柱骨骼识别方法的设计与实现

作者: 左丰源

检测范围: 中国学术期刊网络出版总库

中国博士学位论文全文数据库/中国优秀硕士学位论文全文数据库

中国重要会议论文全文数据库

中国重要报纸全文数据库

中国专利全文数据库

图书资源

优先出版文献库

大学生论文联合比对库

互联网资源(包含贴吧等论坛资源)

英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor&Francis 期刊数据库等)

港澳台学术文献库

互联网文档资源

CNKI大成编客-原创作品库

个人比对库

时间范围: 1900-01-01至2018-05-30

检测结果

总文字复制比: 14.6%

跨语言检测结果: 0%

去除引用文献复制比: 14.6%

去除本人已发表文献复制比: 14.6%

单篇最大文字复制比: 4.2%

重复字数: [3209]

总段落数: [6]

总字数: [21938]

疑似段落数: [3]

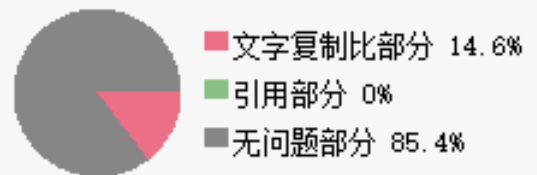
单篇最大重复字数: [916]

前部重合字数: [397]

疑似段落最大重合字数: [2455]

后部重合字数: [2812]

疑似段落最小重合字数: [98]



指标: ☐ 疑似剽窃观点 ☒ 疑似剽窃文字表述 ☐ 疑似自我剽窃 ☐ 疑似整体剽窃 ☐ 过度引用

表格: 0

公式: 0

疑似文字的图片: 0

脚注与尾注: 0

0% (0) 中英文摘要等 (总2168字)

16.5% (656) 第1章绪论 (总3965字)

4% (98) 第2章图像识别技术 (总2473字)

21.4% (2455) 第3章深度学习 (总11489字)

0% (0) 第4章实验分析与结果 (总1597字)

0% (0) 第5章结论与展望 (总246字)

(注释: 无问题部分 文字复制比部分 引用部分)

1. 中英文摘要等

总字数: 2168

相似文献列表 文字复制比: 0%(0) 疑似剽窃观点: (0)

原文内容 红色文字表示存在文字复制现象的内容; 绿色文字表示其中标明了引用的内容

吉林大学学士学位论文(设计)承诺书

本人郑重承诺: 所呈交的学士学位毕业论文(设计), 是本人在指导教师的指导下, 独立进行实验、设计、调研等工作基础上取得的成果。除文中已经注明引用的内容外, 本论文(设计)不包含任何其他个人或集体已经发表或撰写的作品成果。对本人实验或设计中做出重要贡献的个人或集体, 均已在文中以明确的方式注明。本人完全意识到本承诺书的法律结果由本人

承担。

学士学位论文(设计)作者签名:

2018年5月20日

摘要

基于深度学习的脊柱骨骼识别方法的设计与实现

作为当前计算机技术领域的热点话题之一,深度学习取得了高速的发展,被广泛应用于生活的各个领域,而图像识别作为其产生之初的主要目的,更是获得了巨大的进步,不仅在图像分类、目标检测、目标跟踪、语义分割、实例分割各个方面都表现出强大的性能优势,而且速度、精度和准确度有了飞跃的提升,深度学习逐渐在图像识别领域展现出了无与伦比的优势。骨骼X线图像黑白对比强,明暗差别大,界限分明,层次感强烈,是深度学习最优秀的素材之一,能在深度学习的图像识别领域发挥独特的优势,将其与深度学习图像识别相结合,能更好的诊断疾病,发挥计算机在图像识别的优势,减少一些人为因素带来的误差。本文综述了深度学习及其神经网络模型的基本概念,阐述了深度学习在图像识别领域的发展历程,重点介绍了图像识别目标检测的方法及其应用,详细的说明了Faster R-CNN算法,旨在设计一个程序对脊柱骨骼X线图像进行识别,对脊柱侧弯进行测定,衡量出Cobb角大小,从而对医生诊断给出一个较为准确的建议。

关键词:深度学习,图像识别,脊柱侧弯,目标检测

Abstract

Design and Implementation of Spine Bone Recognition based on Deep Learning Algorithm

As one of the current hot topics in the field of computer technology, deep learning has achieved high-speed development and has been widely used in various fields of life. Image recognition, as the main purpose of its creation, has made great progress, not only in Image classification, target detection, target tracking, semantic segmentation, and instance segmentation all show strong performance advantages, and speed, precision, and accuracy have improved dramatically. Deep learning has gradually exhibited unparalleled advantages in the field of image recognition. The X-ray images of bones have strong contrasts in black and white, big differences in light and shade, distinct boundaries and strong sense of depth. They are one of the best materials for deep learning and can play a unique advantage in the field of deep learning in image recognition. The combination can better diagnose the disease, exert the advantages of computer in image recognition, and reduce the errors caused by some human factors. This article reviews the basic concepts of deep learning and its neural network model, describes the development process of deep learning in the field of image recognition, highlights the methods and applications of image recognition target detection, and explains the Faster R-CNN algorithm in detail. A program was designed to identify the X-ray image of the spine, measure the scoliosis, and measure the Cobb angle, which gave a more accurate diagnosis for the doctor's diagnosis.

Keywords: Deep Learning, Image Recognition, Scoliosis, Target Detection

目录

第1章绪论	1
1.1 课题研究背景及意义	1
1.2 国内外研究现状	2
1.3 论文的结构安排	4
第2章图像识别技术	6
2.1 图像识别的基本结构	6
2.2 特征的提取和表达	7
2.3 分类器	7
第3章深度学习	9
3.1卷积神经网络	10
3.1.1 卷积神经网络的优势	10
3.1.2 卷积神经网络的结构	10
3.1.3 卷积神经网络的特性	11
3.1.4 卷积神经网络的训练过程	14
3.2 卷积神经网络实例ZF-Net	15
3.2.1 Alex-Net网络模型	15
3.2.2 ZF-Net网络模型	16
3.2.3 Alex-Net和ZF-Net网络图像处理效果的比较	16
3.3 Faster RCNN神经网络	17
3.3.1 Region Proposal (候选区域) 的概念	17

3.3.2 Faster R-CNN	18
第4章实验分析与结果	26
4.1 数据集的制作与处理	26
4.2 模型训练	27
4.3 训练结果及分析	27
4.4 曲线拟合及结果分析	27
4.5 实验结果及分析	28
第5章结论与展望	28

2. 第1章绪论		总字数：3965
相似文献列表 文字复制比：16.5%(656) 疑似剽窃观点：(0)		
1	深度学习及其算法的研究与发展 高永旺 - 《大学生论文联合比对库》 - 2017-05-20	9.3% (370) 是否引证：否
2	深度学习及其算法的研究与发展 高永旺 - 《大学生论文联合比对库》 - 2017-05-27	9.3% (370) 是否引证：否
3	基于深度卷积神经网络的图像检索系统 李猛 - 《大学生论文联合比对库》 - 2017-05-24	8.4% (335) 是否引证：否
4	降噪自编码及其在混合属性数据聚类分析中的应用 朱海洋 - 《大学生论文联合比对库》 - 2017-05-19	5.3% (209) 是否引证：否
5	朱海洋_降噪自编码及其在混合属性数据聚类分析中的应用 朱海洋 - 《大学生论文联合比对库》 - 2017-05-18	3.7% (148) 是否引证：否
6	基于视频监控的行人检测技术研究 于怀勇(导师：汪国强) - 《黑龙江大学博士论文》 - 2015-03-23	1.5% (61) 是否引证：否
7	软件学院_软件工程_20092552_高巍 软件工程 - 《大学生论文联合比对库》 - 2013-05-16	1.5% (61) 是否引证：否
8	冷藏集装箱远程故障诊断系统研究 张珍(导师：陈宝忠;韩厚德) - 《上海海事大学硕士论文》 - 2006-06-01	1.5% (59) 是否引证：否
9	嵌入式血液分析仪关键技术的研究 (可编辑) - 豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2016	0.8% (33) 是否引证：否
10	纺织纤维成分非破坏性快速鉴别技术研究现状 耿响;桂家祥;周丽萍; - 《纺织导报》 - 2016-01-08	0.8% (31) 是否引证：否

原文内容 红色文字表示存在文字复制现象的内容; 绿色文字表示其中标明了引用的内容

第1章绪论

1.1 课题研究背景及意义

随着人工智能时代的来临和计算机处理性能的显著增强，图像识别技术正向着更加智能化的方向发展，而基于深度学习的图像识别技术正是当下机器学习领域的研究热点。

作为人工智能的一个重要领域，图像识别技术是以图像的主要特征为基础，排除图像的冗余信息，抽取关键信息，通过比对图像之间的关键信息的相似性，如果相似性达到一定的程度，那么就可以认为两者一致，那么图像就识别了。图像识别技术的基本处理过程是首先获取信息，这里的信息可以值光电信号，图像文字，物理量或者逻辑值。然后需要对图像进行预处理，包括图像几何变换、平滑处理、滤波、边缘检测等，主要目的是增强相关信息，方便提取，消除部分无关信息，最大可能的提高图像识别的可靠性。接着需要特征提取选择，即对原始数据进行变换处理，获得最能反映出其本质的特征。最后是训练过程，训练过程分为两个部分，一个是分类器的设计，分类器的设计主要是通过训练过程来获得相应的决策信息，通过该决策信息来分类时能保证成功率最高。另一个是分类决策，即通过分类器设计获得的决策信息来设计的分类决策，应用于被识别图像的分类处理。在上述处理过程中，最核心的步骤是特征提取选择，一个好的特征的选择，可以使用简单的线性分类器就可以识别目标分类。然而，传统的特征提取都是人为选择的，识别不同目标，需要设计不同的特征选择方式，要求相关人员必须具有相应的专业技能，而且这些都是基于目标的底层视觉特征，只是代表了部分的目标属性。

与传统的图像识别不同的是，基于深度学习的图像识别获取的特征是通过训练大量的数据计算机学习到的，而非人工设计，能够有效地发挥大数据的优势，减少手工调整参数的困难，减少开发周期，提升工作效率。深度学习能够将图像识别中特征提取和分类器的设计相结合，联合优化，使图像识别的性能显著提升。不同于传统图像识别只能获取图像的底层信息，深度学习采用多层次神经网络结构，通过增加模型深度，从像素级数据到抽象的语义概念逐层提取特征信息，对图像的分层特征进

行分布式表达，能够有效地提取出图像的全局特征以及上下文信息，从而减少计算单元，所需的参数和训练样本的数目。随着图像识别问题复杂度的增加，这种优势会更加的明显，甚至能够达到几个数量级的差距，深度学习的出现，使图像识别的应用范围，实现方法，作用效果都有了显著的提升。

传统医学领域，医学成像是医生判断一些疾病的重要依据，然而人工辨识图像的效率与质量并不高，所以许多医疗机构急需能够熟练识图的医生。例如，本文要研究的脊柱侧弯的识别，对于脊柱侧弯的检测都是医生通过脊柱骨骼X线图像手工绘制，简单机械衡量出Cobb角的大小，可能由于医生的习惯，测量仪器的差别等原因导致不同医生的测量出现一定的误差，而且医生手工绘制效率比较低，很长时间才能绘制出一张图像，患者需要较长时间等待才能有一个较为准确的结果，这样很不利于患者的诊断和治疗。然而，基于深度学习的图像识别不仅能够减少人工识别所带来的误差，还可以加快诊断速度，提高医生的工作效率，也能够给患者一个治疗指导。另外，骨骼X线图像层次性高，对比度明显，是图像识别最好的样本之一。所以通过深度学习对脊柱图像进行识别是完全可行，并且拥有极高的应用前景和现实意义的。

1.2 国内外研究现状

1943年，神经科学家麦卡洛克 (W.S.McCulloch) 和数学家皮兹 (W.Pitts) 在发表的论文《神经活动中内在思想的逻辑演算》 (A Logical Calculus of the Ideas Immanent in Nervous Activity) 中提出的MCP人工神经元模型是最早的神经网络，该模型受到生物神经元工作方式的启发，将生物神经元抽象简化成输入信号线性加权，求和，阈值法三个过程，从而模型大脑的工作方式，人工神经网络就此诞生。

1958年，罗森布拉特 (Rosenblatt) 第一次将MCP用于机器学习分类问题。这个由两层神经元构成的感知器能够对多维的数据进行二分类，并且能够从样本中学习更新权值。1962年，该算法被证实能够收敛，推动了第一次神经网络的热潮。但是1969年，数学家Marvin Minsky证实感知器是一种线性模型，只能处理线性分类问题。感知机不能处理线性不可分问题，哪怕最简单的异或 (XOR) 问题都无法解决，巨大的打击，让神经网络的研究停滞了将近20年。

1986年，神经网络之父Geoffrey Hinton发明了适用于多层感知机的BP算法，采用Sigmoid函数作为神经网络的阈值函数进行变量的非线性映射，能够有效解决非线性分类的问题，引发了第二次神经网络的热潮。

1991年，BP算法被证实存在梯度消失的问题，误差梯度向后传递的过程中，后层梯度以乘性方式叠加到前层，但是由于Sigmoid函数的饱和性，后层梯度本来就小，误差梯度传到前层几乎为0，因此无法对前层进行有效的学习，也正是这个原因再次阻碍了神经网络的发展。

1986年，决策树方法被提出。1995年，线性SVM被提出，这种方法由非常完美的数学理论推导而来，并在线性分类问题上表现突出。1997年，AdaBoost被提出，通过将弱分类器进行集成达到强分类器的效果。2000年，KernelSVM被提出，这种方法能够将原空间线性不可分的问题通过Kernel映射成高维空间的线性可分问题，解决了非线性分类问题并且分类效果不错。2001年随机森林被提出，这种方法也是一种集成方法，但是比AdaBoost的效果更能一直过拟合问题。在这个阶段，统计学方法发展的相当迅速，各种前层模型相继提出，并且展现出相当不错的效果，但这个无疑令神经网络的发展雪上加霜。虽然在此期间，LeCun发明了卷积神经网络LeNet，在数字识别领域取得不错的效果。另外LSTM模型也在这个时期出现，在序列建模上表现优异。但是它们都没有引起足够的重视。神经网络的发展举步维艰，这种状态一直持续到了2006年。

2006年，加拿大多伦多大学教授Geoffrey Hinton和他的学生在发表的一篇论文中提出了通过无监督预训练对权值进行初始化和有监督训练微调的方式解决神经网络训练过程中梯度消失的问题。这篇论文提出了通过逐层训练的方式解决神经网络难以达到最优的问题的观点，并且指明多层神经网络具有很强的学习能力，其学习到的特征数据更能代表原始数据的本质。这篇论文不仅打破了限制神经网络的枷锁，更为神经网络的发展指明了方向，神经网络再次引发了热潮。

2011年，ReLU激活函数被提出，有效地抑制了梯度消失的问题。同年微软首次将DL用在语音识别上，表现出不错的效果。深度学习进入了快速发展的阶段。

2012年，Hinton课题小组通过构建的卷积神经网络AlexNet参加ImageNet图像识别比赛，碾压第二名的SVM分类方法，一举夺魁，正是这个原因，CNN吸引了广大研究人员的注意。AlexNet的创新之处在于：(1) 采用ReLU激活函数，增加了收敛速度并且解决了梯度消失问题；(2) 采用有完全的有监督学习；(3) 扩展了LeNet5结构，添加Dropout层减小过拟合，LRN层增强泛化能力；(4) 首次采用了GPU加速运算。

2006年以来国内也开始了深度学习方面的研究，但是由于起步较晚，虽然取得了一定的成果，但是目前还处于理论研究起步阶段。2012年国内互联网巨头百度公司将深度学习最新相关技术应用到人脸识别和自然语言处理的问题上。2013年百度成立研究院，并且成立了Institut of Deep Learning (IDL)，深度学习在国内受到了重视。

2013年，ZF神经网络被提出，基于AlexNet进行了调整，使用ReLU激活函数和交叉熵代价函数，通过较小的filter保留更多原始像素信息，取得了2013年ILSVRC冠军。2014年提出的GoogleNet神经网络取得了2014年ILSVRC冠军，其网络结构有了较大的创新，有效地改善了性能，其特点是：(1) 使用9个inception模块并行计算，改变了CNN的串行结构。(2) 不使用FC层，采用平均池化层代替，减少了大量的参数。(3) 检测模型采用R-CNN的概念。2015年提出的ResNet神经网络是一种超深网络，拥有152层，取得了2015年ILSVRC冠军，并且引入了残差网络的概念。2014年到2016年是神经网络的高速爆发期，2016年出现较大革新方法的脚步已经慢慢变缓了。

2016年，谷歌 (Google) 基于深度学习开发的AlphaGo与世界围棋高手李世石进行人机大战，4：1战胜。在2017年初与数十位围棋高手对决，连续60场胜利。2017年5月，它以3：0的成绩战胜世界排名第一的围棋高手柯洁。正因为AlphaGo，深度学习走进了人们的生活，开始为人们所熟知。

1.3 论文的结构安排

本论文的结构安排如下：

第1章绪论。介绍了本论文的研究背景和意义，然后介绍了国内外的研究进展。

第2章图像识别技术。本章我们较为详细的介绍了图像识别的研究过程，包括其实现方法，基本原理和相关概念的讲述

第3章深度学习。本章是全文的重点，首先我们引入深度学习的话题，并介绍了卷积神经网络的概念，重点讲述了卷积神经网络在分类和目标检测领域的应用，通过ZF卷积神经网络和Faster RCNN目标检测卷积神经网络举例说明了其实现方式和相关内容。

第4章实验过程与结论。在这个过程中我们重点展示了实验的结果，按照实验流程介绍了实验内容，并对结果进行了分析和总结。

第5章总结与展望。综述了全文的研究内容，并对未来工作做出展望。

指 标	
疑似剽窃文字表述	
1.	特征，只是代表了部分的目标属性。 与传统的图像识别不同的是，基于深度学习的图像识别获取的特征是通过训练大量的
2.	更新权值。1962年，该算法被证实能够收敛，推动了第一次神经网络的热潮。但是1969年，
3.	适用于多层感知机的BP算法，采用Sigmoid函数作为神经网络的阈值函数进行变量的非线性映射，能够有效解决非线性分类的问题，引发了第二次神经网络的热潮。 1991年，BP算法被证实存在梯度消失的问题，误差梯度向后传递的过程中，后层梯度以乘性方式叠加到前层，但是由于Sigmoid函数的饱和性，后层梯度本来就小，误差梯度传到前层几乎为0，因此无法对前层进行有效的学习，
4.	方法能够能够将原空间线性不可分的问题通过Kernel映射成高维空间的线性可分问题，
5.	2006年。 2006年，加拿大多伦多大学教授Geoffrey Hinton和他的学生
6.	监督预训练对权值进行初始化和有监督训练微调的方式解决神经网络训练过程中梯度消失的问题。
7.	1.3 论文的结构安排 本论文的结构安排如下： 第1章绪论。介绍了本论文的研究背景和意义，然后介绍了国内外的研究进展。

3. 第2章图像识别技术		总字数：2473
相似文献列表 文字复制比：4%(98) 疑似剽窃观点：(0)		
1	111180217-朱琳-通信工程 朱琳 - 《大学生论文联合比对库》 - 2015-05-21	1.4% (35) 是否引证：否
2	基于分数阶傅里叶变换的无刷直流电机逆变器故障诊断技术研究 吴鑫;崔江;陈则王; - 《机械制造与自动化》 - 2014-06-20	1.2% (30) 是否引证：否
3	医学图像数字处理及识别技术研究 叶晨(导师：刘镰斧) - 《电子科技大学博士论文》 - 2012-04-01	1.2% (30) 是否引证：否
原文内容 红色文字表示存在文字复制现象的内容; 绿色文字表示其中标明了引用的内容		

第2章图像识别技术

图像识别技术是人工智能的一个重要领域，主要通过计算机对图像进行处理、分析和理解，达到识别各种不同模式的目标的技术，其本质就是图像的模式识别，对输入的图像建立图像识别模型，分析选择图像的特征，建立分类器，达到图像分类的目的。图像识别的发展从简单到复杂依次经历了文字识别、数字图像处理和识别、物体识别三个阶段。随着计算机性能的提升以及数据量的扩大，图像识别技术得到了更加广泛的应用，如人脸识别、指纹识别等，并且图像识别技术在向着更加多元的方向发展，极大的丰富和便利了人们的生活。

2.1 图像识别的基本结构

研究表明人们在看图像时主要将视线集中在图像的主要特征上，然后将特征传到大脑，大脑负责整合信息，让后将信息整合成一个完整的映像。图像识别技术是模仿人们认识图像的方式，抽取图像的主要特征，排除多余的信息，从而根据主要特征建立一个信息分类器，达到对图像信息分类的目的。图像识别其实是一种图像的模式识别过程，在大量的数据和信息中根据已有的认知和经验，利用计算机和数学推理的方法自动完成和图像的识别、评价的过程。这个过程包括两个阶段，即学习阶段和实现阶段，前者是通过样本的特征选择，设计一个分类器，后者是根据前者设计的分类器对未知样本集进行分类从而识别图像，如下图所示：

图2-1 图像识别过程

这些信息在绪论中已经介绍过，我们知道这些步骤中最核心当是特征的选择，直接关系到识别效果和性能，一幅图像中

的特征有很多，但并不都是有用的信息，因此我们需要对特征进行判断，提取干扰能力强的特征。

2.2 特征的提取和表达

图像的特征提取分为底层特征和高层特征两个层次。底层特征提取包括颜色、形状、纹理，是图像分析的基础，计算简单较为稳定。高层特征一般是语义层次的，需要通过底层提取，然后根据机器学习获得。

(1) 颜色特征

颜色是图像重要的组成因素之一，颜色常常和具体的物体相互关联，例如，说到绿色人们常常想起植物，说起蓝色常常会想到蓝天大海，相似的事物颜色也相近。而且颜色有很强的鲁棒性，对图像进行旋转，形状表还，大小变化对图像的颜色影响不大，而且计算简单。常用于提取颜色特征的方法有颜色直方图、颜色集、颜色矩、颜色聚合向量、颜色相关图。

(2) 形状特征

形状特征不会随着环境变化而变化，是一种相对稳定的特征。物体形状特征的提取大致分为两种方法，分别是基于轮廓的提取方法和基于区域的提取方法。图像轮廓主要描述物体的边界，其特征维度比较低，相对而言符合人类区分物体的方法，因此使用也更广泛，但是易受形变因素的影响。基于轮廓的特征提取方法有傅里叶描述符、链码、小波描述符、曲率尺度空间等。图像区域特征是指整个形状区域的特征，不仅包括外边界，还有边界包围的内容。区域特征包含整个区域的像素，收到形变因素的影响相对较小，但是纬度高。基于区域的特征提取方法有离散度、几何不变矩、区域不变矩。

(3) 纹理特征

纹理特征表现为图像的表面特性，拥有局部不规则形和区域规律性，通过提取这个区域的像素进行统计计算来保证特征提取的准确性，它反映了物体表面的隐含特性，在整个区域表现出对噪声良好的鲁棒性。纹理特征的提取方法有：结构法、频谱法、统计法和模型法。

2.3 分类器

分类器主要是用机器学习来对图像特征进行计算分类。目前常用的分类算法有：决策树、贝叶斯分类法、支持向量机、人工神经网络和集成学习算法等。

(1) 决策树

决策树是基于监督学习的一种分类方法，是一种树形结构。在决策树中，它的每个节点都表示一种对象，每个路径表示一种属性，它的叶子节点表示对象的值，即某个节点通过一种路径所能达到的结果。决策树的方法实现简单，能够直接体现数据的特点，而且不需要准备数据，只要能够理解数据的类型和常规属性就能建立一个相对可靠的决策树。常用的决策树算法有C4.5、CART等。

(2) 贝叶斯分类法

贝叶斯分类是一种属于统计学范畴的非规则分类方法，贝叶斯理论上给出了最小化误差的最优解决方法，其主要是基于贝叶斯定理，通过计算对象的属于某个类别的概率来选择概率最大的类别对对象进行分类。贝叶斯网络假设对象的每个属性是相互独立的，一个属性不会影响其他属性在分类中的效果，这个假设大大简化分类所需要的计算量。但是很多时候对象的属性并不是相互独立的，所以贝叶斯分类会有一定的误差，影响分类效果。

(3) 支持向量机

支持向量机也是属于统计学范畴的分类方法，在解决小样本和非线性分类问题中表现突出，常在数据分析师模式识别中用于分类和回归分析。支持向量机一般针对线性可分问题进行分析，当遇到线性不可分问题，通过将非线性的低维特征空间转换成高维的线性可分特征空间。支持向量机建立在结构风险最小化理论的基础上，在特征空间建造最有超平面，从而使样本的非线性特征能够在高维的特征空间采用线性算法进行计算。支持向量机的优点是可以转换成凸优化问题，从而求得全局最优解，而其他方法一般只能获得局部最优解。支持向量机在二分类问题上表现突出，但是在多分类问题上表现却不尽人意。

(4) 人工神经网络

人工神经网络主要是通过模仿人们大脑神经元的工作方式来建立数学模型。在该模型中，每个节点代表一个神经元，每个节点采用一种特定的输出函数，相互连接的节点之间都有一个输出信号的加权值，相当于人类的记忆，节点之间相互连接就组成了神经网络。神经网络具有如下优点：(1) 具有自我学习能力，将某种类型样本和对应的样本结果传递给神经网络，神经网络就会通过训练慢慢识别这种类型的对象；(2) 具有联想存储功能；(3) 能够快速发现最优解。但是神经网络目前还存在计算量大，训练时间比较长等缺点。相信随着计算机性能的提升，神经网络展现的效果一定能够更加出色。

(5) 集成学习

集成学习不是一种独立的算法，它使用一系列的分类算法，并按照某种规则将它们结合在一起，这样取得的效果往往比一种分类算法取得效果要好。常用的集合学习算法有：bagging算法、Adaboost算法等。

4. 第3章深度学习		总字数：11489
相似文献列表 文字复制比：21.4%(2455) 疑似剽窃观点：(0)		
1	51127345629906399_一种基于深度学习的人民币红外对管鉴别方法 陈唐鑫 - 《大学生论文联合比对库》 - 2017-05-12	8.0% (916) 是否引证：否

2	CNN目标检测 (一) : Faster RCNN详解 - zy1034092330的专栏 - CSDN博客 - 《网络 (http://blog.csdn.net) 》 - 2017	7.3% (838) 是否引证 : 否
3	基于大数据处理的人体姿态识别研 陈丰铭 - 《大学生论文联合比对库》 - 2017-05-16	7.1% (813) 是否引证 : 否
4	Deep Learning学习 之 卷积神经网络 (CNN) - Mr.Phoebie的专栏 - CSDN博客 - 《网络 (http://blog.csdn.net) 》 - 2017	6.7% (775) 是否引证 : 否
5	[转载][转]Deep Learning (深度学习) 学习笔记整理系列_dubyy - 《网络 (http://blog.sina.com) 》 - 2015	6.2% (713) 是否引证 : 否
6	基于深度学习的气体识别算法研究 于洋 - 《大学生论文联合比对库》 - 2015-06-03	6.1% (705) 是否引证 : 否
7	Deep Learning (深度学习) 学习笔记整理系列之 (七) - zouxy09的专栏 - 博客频道 - CSDN.NET - 《网络 (http://blog.csdn.net) 》 - 2013	5.8% (666) 是否引证 : 否
8	小白学《神经网络与深度学习》笔记之四-深度学习的常用方法 (2) - yingwei13mei的博客 - CSDN博客 - 《网络 (http://blog.csdn.net) 》 - 2017	5.5% (637) 是否引证 : 否
9	基于深度学习的交友推荐系统 梁晓婷 - 《大学生论文联合比对库》 - 2017-04-27	5.3% (606) 是否引证 : 否
10	基于深度学习的交友推荐系统 林泽宇 - 《大学生论文联合比对库》 - 2017-04-28	5.3% (606) 是否引证 : 否
11	卷积神经网络 - 《互联网文档资源 (http://wenku.baidu.c) 》 - 2017	4.9% (568) 是否引证 : 否
12	卷积神经网络 - 豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2016	4.9% (564) 是否引证 : 否
13	基于卷积网络的人脸检测的研究与实现 肖柏旭(导师:张丽静) - 《华北电力大学 (河北) 硕士学位论文》 - 2006-12-26	4.4% (506) 是否引证 : 否
14	新型数字眼底荧光血管造影图像处理系统的研制 崔栋(导师:刘伯强) - 《山东大学硕士学位论文》 - 2010-03-10	4.0% (463) 是否引证 : 否
15	红外图像人脸识别研究 周羽(导师:华顺刚) - 《大连理工大学硕士学位论文》 - 2007-12-01	3.9% (447) 是否引证 : 否
16	基于2DPCA和多分类器融合的人脸识别 韩文静(导师:李晶;曹茂永) - 《山东科技大学硕士学位论文》 - 2009-05-01	3.8% (438) 是否引证 : 否
17	第4讲 深度学习应用中的常见模型 乔林;张雄伟;史海宁;贾冲; - 《军事通信技术》 - 2016-03-25	3.7% (423) 是否引证 : 否
18	基于深度学习的手写金融汉字识别 谢浩然 - 《大学生论文联合比对库》 - 2016-06-01	3.6% (419) 是否引证 : 否
19	测绘学院-2012301610182-涂丽(论文) 测绘学院 - 《大学生论文联合比对库》 - 2016-10-24	3.5% (407) 是否引证 : 否
20	焊接熔池动态过程及电弧能量分布的数值模拟 雷玉成(导师:程晓农) - 《江苏大学硕士学位论文》 - 2006-10-01	3.5% (406) 是否引证 : 否
21	2012301610182 涂丽 (论文) - 《大学生论文联合比对库》 - 2016-05-25	3.0% (345) 是否引证 : 否
22	基于神经网络PID的预测控制在危化品生产过程中的研究与应用 王振(导师:孟友新) - 《青岛科技大学硕士学位论文》 - 2008-06-12	2.7% (313) 是否引证 : 否
23	卷积神经网络在路牌识别中的应用 宋波;王启春; - 《公路交通技术》 - 2015-10-15 1	2.6% (293) 是否引证 : 否
24	201203010216+彭飞+基于深度学习的目标检测算法研究 - 《大学生论文联合比对库》 - 2017-06-16	1.2% (133) 是否引证 : 否
25	2013491965_彭飞_基于深度学习的目标检测算法研究 彭飞 - 《大学生论文联合比对库》 - 2017-06-12	1.2% (133) 是否引证 : 否
26	面向Android平台的智慧相册的设计与实现 孙家贺(导师:苗振江) - 《北京交通大学博士学位论文》 - 2017-05-02	0.5% (58) 是否引证 : 否
27	电子1203_1210910307_王利伟 王利伟 - 《大学生论文联合比对库》 - 2016-05-17	0.3% (38) 是否引证 : 否
28	天然气管道调度优化研究 刘奇(导师:郑云萍) - 《西南石油大学博士学位论文》 - 2014-06-01	0.3% (35) 是否引证 : 否

原文内容 **红色文字**表示存在文字复制现象的内容; **绿色文字**表示其中标明了引用的内容

第3章深度学习

随着科技的进步与发展, 计算机处理器性能显著增强, 硬件成本普遍降低, 大数据时代的来了, 机器学习研究的进步, 促成了深度学习的出生与发展。深度学习是机器学习下的一个新的研究领域, 通过模拟人类大脑学习认知的方式, 构建类似人类神经系统的多层人工神经网络, 逐层从输入信息提取数据特征, 获得数据的高层特征, 利用这样的多层非线性信息处理实现监督学习和无监督学习的特征提取, 从而令计算机能够通过对数据的自我学习解释数据的内容, 达到识别声音、图像、文本等信息的目的。自深度学习诞生以来, 以其优良的训练性能, 卓越的计算能力被广泛应用于计算机视觉、语音识别、自然语言处理等领域, 深受人工智能领域研究者的喜爱, 是当今计算机的研究热点之一。深度学习的基本思想是建立一个多层的网络, 当输入一个数据时, 它被传入下一层处理后, 再将本层的输出信息当作下一层的输入信息, 知道给出一个结果, 通过这种方式一层一层的提取数据的特征, 实现了对信息的分级表达。这种方式的流程是这样的, 假设我们有一个n层的神经网络, 它的第i层表示为 S_i , 并且把输入设为I, 输出设为O, 那么这个结构可以形象的表示为: $I \Rightarrow S_1 \Rightarrow S_2 \Rightarrow \dots \Rightarrow S_n \Rightarrow O$ 。理解这个系统首先要明白一个概念, 即当输出与输入一致时, 也就是说经过整个网络的处理输入没有发生变化, 则认为系统给没有信息损失。因为一般情况下当经过每一层的处理后, 信息会都是一部分, 在这个网络中就变成为O中所含的信息是少于I中所含信息的, 而这样的一个处理系统显然是不满足条件的, 只有经过每一层处理后信息虽然形式变化了, 但是信息却没有丢失的系统才是要求的。那么深度学习就是建立一个这样的系统, 然后通过自我学习获取适合的参数, 使系统满足数据经过处理后输出O和输入I保持一致, 那么经过每一层的时候系统就会自动获取数据的层次特征。当然我们不必假设输入和输出完全相等, 可以让给输入和输出的差别规定一个限度, 小于这个限度的就满足条件, 这就是深度学习的基本思想。其核心思想是将无监督学习算法用以每一层的训练提取特征, 每次只训练一层, 并将该层的结果传递给下一层, 作为下一层的输入, 另外还需要用监督学习算法调整所有的层。深度学习常用的方法有: 深度信念网络, 卷积神经网络、卷积深度信念网络、深度玻尔兹曼机、栈式自动编码器等。我们会详细介绍卷积神经网络的基本理论和相关方法, 用ZF和Faster RCNN两种**卷积神经网络说明卷积神经网络在图像分类和目标检测方面的应用。**

3.1 卷积神经网络

3.1.1 卷积神经网络的优势

卷积神经网络是一种用来识别二维图形的多层次前馈人工神经网络, 这种网络对于图形的位移、缩放和扭曲等操作具有高度的不变性。所以卷积神经网络在图像识别方面表现出巨大的优势: (1) 直接输入而且图像避免了复杂的预处理操作; (2) 特征提取和模式分类同时进行, 相互结合, 不断优化获得所需参数; (3) 采用局部感知和权值共享两种方法, 减少参数的数量, 性能逐步增强。接下俩我们会详细介绍CNN的相关内容。

3.1.2 卷积神经网络的结构

卷积神经网络是一个专门用来识别二维特征的神经网络, 它的结构如下: 输入的图像通过三个可训练滤波器进行卷积操作, 卷积后在C1层产生三个特征映射图, 然后特征图经过加权值和偏置后, 通过一个sigmoid函数得到三个s2层的特征映射图, 这些特征映射图经过滤波得到C3层, 然后再通过一个sigmoid函数得到三个s4层的特征映射图, 这些像素值被光栅化, 并连接成一个向量输入到传统的神经网络中, 得到输出。

图3-1 卷积神经网络结构

在卷积神经网络中, 一般是由输入层, 卷积层、下采样层、全连接层和分类器五部分组成。在上图中, C表示卷积层, 每个卷积层包含多个特征图, 每个特征图又含有多个独立的**神经元, 每个神经元与前一层的局部感受视野相连接, 并提取出局部特征, 一旦局部特征被提取, 那么它与其他特征层的位置关系也就确定下来了。**S层表示下采样层, 主要是对特征图进行简单的缩放, 降低特征图的维数, 减少网络训练中权值的数量, 经过卷积层和下采样层的操作, 在网络和输出之间连接一个全连接层将特征相互结合起来, 最终得到一个用于识别分类的神经网络。

3.1.3 卷积神经网络的特性

卷积神经网络最大的特点是局部感知和权值共享, 这使得卷积神经网络训练参数的数目大大降低, 泛化性能进一步增强。

(1) 局部感知。

图3-2 全连接和局部连接

人们对信息的**认知都是从局部到全局的, 而且在图像空间里也是相邻像素之间的关联相对比较紧密**, 距离较远的像素之间的关联程度就比较弱。因此每个神经元只需要对局部进行感知即可, 然后在高层对信息进行综合处理。在上图中, **假设我们有一张1000x1000像素的图像, 图左为全连接网络, 需要1百万个隐层神经元, 每个神经元连接一个像素点, 那么就需要1000x1000x1000000个连接, 也就是10^12个权值参数。而图右为局部连接网络, 每个节点与上层的10x10个窗口连接, 那么只需要10x10x1000x1000个参数, 即10^8个参数, 参数值比原来降低了四个数量级。**

(2) 权值共享。

在卷积神经网络中, 每个hi过滤器会通过权值共享覆盖整个可视域, 这些共享权值的单元构成了一个特征映射, 如图3-3所示。

图3-3 权值共享映射结构

在图3-3中，有三个隐藏单元，他们属于一个特征映射。同种颜色的链接权值是相同的，可以通过梯度下降的方式来学习这些权值。权值共享的优点是：重复单元能够对特征进行识别，不需要考虑其在可视域的位置。权值共享能够使我们更有效的进行特征的提取，减少需要学习的参数的个数。

(3) 卷积神经网络的基本部件

卷积层：卷积是一种局部操作，通过一定大小的卷积核作用于局部图象区域获得图像的局部特征。我们通过三种边缘卷积核来说明卷积操作的作用，我们在原图分别使用整体边缘滤波器，横向边缘滤波器和纵向边缘滤波器，分别将这三种滤波器定义为如下的3x3大小的卷积核 K_e 、 K_h 和 K_v ： $K_e = \begin{bmatrix} 0 & -4 & 0 \\ -4 & 16 & -4 \\ 0 & -4 & 0 \end{bmatrix}$ $K_h = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -4 & -1 & -2 \end{bmatrix}$ $K_v = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 0 & -1 & 0 \end{bmatrix}$

若像素 (x, y) 处于物体的边缘，那么其四周像素 $(x-1, y)$ ， $(x+1, y)$ ， $(x, y-1)$ ， $(x, y+1)$ 应该与 (x, y) 具有显著的差别，那么此时用整体滤波器 k_e 可以消除四周像素值差异小的区域而保留差异显著区域，而横向滤波器 k_h 和纵向滤波器 k_v 可以保留横向、纵向的边缘信息。

图3-4 三维场景卷积

卷积神经网络的卷积核参数是通过网络训练的得到的，不进入我们所举的三种滤波器，还可以得到任意角度的滤波器，也可以是用来检测颜色、形状、纹理等基本模式的滤波器，通过这些滤波器可以将一些底层特征抽象为具体的高层语义特征。

池化层：一方面可以把池化层当作是降采样层，池化操作显著降低了特征的维数。另一方面，池化层也可以看作是一个用p-范数作非线性映射的卷积层。池化层是仿照人的视觉系统对输入对象进行降维和抽象，普遍认为池化层具有以下三种功效：
(1) 特征不变性。池化操作关注的是特征是否存在而不是特征的位置，这样使特征具有一定的自由度，微小的位移不会影响特征的变化。
(2) 特征降维。池化操作可以用来降采样，池化操作后的每个元素对应输入数据中的一个子区域，其实相当于将某个区域的维度约减。不仅使模型能够获取的特征更加广泛，而且减少下一层输入的大小，进而减少了计算量和参数数目。
(3) 池化操作能够在一定程度上防止过拟合。当时池化操作并不是卷积神经网络所必需的部件。

激活函数：激活函数层又被称为非线性映射层，激活函数能够增强网络的表达能力。没有激活函数那么网络只能起到线性映射的作用，复杂函数没有办法进行表示。激活函数模拟了生物神经元，接收一组信号并产生输出，在生物神经学领域，生物神经元通常有一个限值，只有激活信号超过该限值时，神经元才会被激活，否则神经元就会处于抑制状态。人工神经网络中通过Sigmoid函数来模拟这一操作，经过Sigmoid函数作用后，输出的值会被限制在 $[0, 1]$ 之间，0相当于抑制状态，而1相当于激活状态。但是Sigmoid函数有一个严重的问题，当输入的值大于5（或者小于-5）的时候会被压缩在1（或0）之间，这样会导致梯度的“饱和效应”，在误差反向传播过程中该区域的值很难传播到前层，导致神经网络无法进行训练。为此专门在神经网络中引入修正线性单元的概念，修正线性单元实际上是一个分段函数，规定单梯度大于或等于0时为1，当梯度小于0时为0，这样就消除了大于或等于0部分Sigmoid的梯度饱和效应。修正线性单元能够有助于梯度下降方法的收敛，是目前深度网络模型常用的激活函数之一。

全连接层：全连接层在卷积神经网络中相当于分类器，如果卷积层、池化层和激活函数等操作是将数据映射到特征空间，全连接层的作用相当于将学到的特征映射到样本的标记空间。

目标函数：目标函数的作用是用来衡量预测值和真实值之间的差距，目前卷积神经网络分类问题常用的目标函数是交叉熵损失函数。

3.1.4 卷积神经网络的训练过程

卷积神经网络本质上是一种输入到输出的映射，它不需要任何已知的输入和输出之间的精确数学表达，而是通过对大量已知输入和输出之间映射关系的学习，用已知模式的网络进行训练，网络就能自动获得输入和输出之间的映射关系，从而能够对未知数据进行判别其类型。卷积神经网络执行的是有监督的学习，其样本形式如： $(\text{输入向量}, \text{输入向量对应结果})$ 的向量对构成。所有的向量对都来源于网络模拟系统的实际运行结果。在开始训练之前，所有权都应该用一些不同的小随机数进行初始化，这样可以保证网络不会因为权值过大而进入饱和状态，从而致使训练失败。小随机数不能相同，因为相同的数去初始化权矩阵，网络是没有学习的能力的，不同的随机数可以保证网络正常的学习。

卷积神经网络的训练包括4个步骤，这4个步骤可以分为两个阶段：

第一阶段，向前传播阶段：

a) 从样本集合中取出一个样本 (X, Y_p) ，将 X 输入网络；

b) 计算相应的输出 O_p 。

在这个阶段，信息从输入层经过逐级变换传递到了输出层。这个过程也是网络训练完之后能够正常运行时执行的过程。在这个过程中，网络执行的是计算（实际上就是输入与每层的权值矩阵相点乘，得到最后的结果）： $O_p = F_n(\dots F_2 F_1 X p W_1 W_2 \dots W(n))$

第二阶段，向后传播阶段：

a) 计算实际输出 O_p 与相应的理想输出 Y_p 的差；

b) 按极小化的误差方法反向传播调整权矩阵。

这两个阶段的工作一般受到精度要求的控制，用下面的式子计算 E_p ， E_p 计算网络关于第 p 个样本的误差测度，而将网络关于整个样本集的误差测度定义为 $E = E_p$ ： $E_p = \frac{1}{2} \sum_j (y_{pj} - o_{pj})^2$

综上所述，之所以将这个阶段称为向后传播阶段，是对于输入信号的正常传播方向而言的，因为在开始调整神经元的连接权时，只能求出输出层的误差，而其他层的误差都要通过次误差反向逐层后推得到。

3.2 卷积神经网络实例ZF-Net

本节我们会介绍卷积神经网络ZF-Net，ZF-Net是基于Alex-Net网络提出来的，二者之间的网络结构并没有发生太大的变化，所以本节将Alex-Net和ZF-Net比较说明ZF-Net的性能。

3.2.1 Alex-Net网络模型

Alex-Net网络模型掀起了深度学习在计算机视觉领域的研发热潮。Alex-Net在2012年的ImageNet竞赛中以高出第二名10.9个百分点的优异成绩向世人展现出了卷积神经网络的强大之处。Alex-Net的网络结构如下图所示，它含有五个卷积层和三个全连接层。Alex-Net分为上下两块连接，是为了方便两块CPU同时进行运算，在第三卷积层和全连接层上下两块可以进行信息交互。我们在下表中列出了Alex-Net网络的架构及具体参数。于前代卷积神经网络相比，Alex-Net在网络结构的改善十分微小，仅仅提升了网络深度和复杂度。

图3-5 Alex-Net结构示意图

但是Alex-Net的贡献是巨大的。Alex-Net在ImageNet竞赛上的突出表现，彰显了其强大的学习能力和优秀的数据处理能力，引发了深度学习的研究热潮，卷积神经网络的发展尤为迅速。另外Alex-Net借助GPU运算，强大的硬件设施大大缩短了大型网络模型的研发时间和成本，这也是各种网络模型层出不穷的硬件保障。Alex-Net还将ReLU修正线性单元、局部响应规范化操作、采用增加数据量的方式防止过拟合和随机失活等训练技巧引入神经网络，这些操作在今后的网络构建中依然被引做范本。

3.2.2 ZF-Net网络模型

ZF-Net的网络结构如下图所示，实际上ZF-Net与Alex-Net在网络结构上的差别并不大，差异是ZF-Net取消了Alex-Net将信息处理分为两层用GPU来操作，选用一块GPU的连接结构。而且通过可视化可以发现Alex-Net第一层中有大量的高低频信息混合，但是却没有中间的频率的信息。而且第二卷积层由于第一卷积层的步长太大导致了信息的重叠。因此ZF-Net将Alex-Net的第一层的滤波器大小由11x11变为了7x7，并将步长变为原来的一半为2。

图3-6 ZF-Net网络模型

3.2.3 Alex-Net和ZF-Net网络图像处理效果的比较

如图3-7，在图中（a）是没有经过处理的图片经过第一个卷积层后的特征可视化图，在该图中有一个特征块是纯白色的。（b）是（a）经过Alex-Net中的第一个卷积层后的特征可视化图。（c）是（a）经过ZF-Net中的第一个卷积层中的特征可视化图。（b）和（c）与（a）相比减少了很多无意义的特征从而保留了更多有意义的特征，但是我们是看不出来两者之间的差别，这个差别比较细微。（d）是（b）经过Alex-Net中第二个卷积层后得到的特征可视化图。（e）是（c）经过ZF-Net第二个卷积层后的特征可视化图。（e）相比于（d）表现的更加清晰，保留的信息更多。同一张图片分别由Alex-Net和ZF-Net经过两个卷积层处理得分别到（d）和（e），从（d）和（e）的差别我们可以看出ZF-Net虽然与Alex-Net网络结构相似，但是特征提取能力明显更强。

图3-7 两种神经网络的结果对比

3.3 Faster RCNN神经网络

Faster RCNN是一个实现目标检测的神经网络，它在2015年被提出，至今依然保持着旺盛的生命力，是很多目标检测算法的基础。本文就是通过Faster R-CNN目标检测神经网络来实现对脊柱骨骼中骨骼位置的定位，这是本文的核心内容。Faster R-CNN是在RCNN和Fast R-CNN的基础上提出来的，它们都是基于候选区域的卷积神经网络算法。本节我们会详细介绍Faster R-CNN的相关实现。

3.3.1 Region Proposal (候选区域) 的概念

上述介绍过Faster RCNN和Fast RCNN，RCNN都是属于基于候选区域的深度学习目标检测算法，候选区域指的是目标在图像中可能的位置，利用图像的颜色、纹理和边缘等特征，保证尽可能少的选择候选窗口的情况下能够维持一定的召回率（IOU，Intersection-over-Union）。物体检测需要找出需要检测物体的位置，并用bounding box将其包括。对于bounding box的定位精度，我们提出了IOU这个概念来衡量定位的准确度。IOU定义了两个bounding box的重叠度，如下图所示：

图3-8 IOU重叠方式示意图

矩形框A、B的重叠度IOU计算公式为： $IOU=(AnB)/(AuB)$

即指矩形框A、B重叠的面积占A、B所占区域的总面积的比例： $IOU= SI/(SA+SB-SI)$ 其中I表示A和B公共部分的面积

与传统的滑动窗口方法相比，Region Proposal方法更加高效，常用的Region Proposal方法有Selective Search（SS，选择性搜索）和Edge Boxes（EB）。

3.3.2 Faster R-CNN

Faster R-CNN目标检测卷积神经网络我们可以理解为区域生成网络RPN和Fast R-CNN的结合，Faster R-CNN采用区域生成网络RPN的方式来代替Fast R-CNN中的SS方法。Faster R-CNN着重解决了如何设计区域生成网络，如何训练区域生成网络以及如何让区域生成网络和Fast R-CNN共享特征提取网络三个问题。Faster R-CNN的网络结构如下图所示：

图3-9 Faster RCNN的网络结构

根据上图Faster RCNN的网络结构我们可以看出来Faster RCNN的主要由conv layers、Region Proposal Networks、Roi

Pooling和classifier四部分组成，我们会通过这四个部分来介绍Faster R-CNN网络。

(1) **conv layers** : conv layers包含了conv，pooling和relu三种层。我们的程序实现中采用的是ZF-Net卷积神经网络模型，在上节详细的说明了这种卷积神经网络，我们在这里不再进行详细的描述。在这个过程中图片经过ZF-Net网络后获得了图像对应的feature maps，feature maps会被传入后续网络再次进行处理。

(2) Region Proposal Networks (RPN)

传统的候选区域生成算法都比较耗时，比如adaboost用滑动窗口的方式生成检测区域，R-CNN用SS的方法生成候选区域。而Faster R-CNN抛弃了传统的候选区域生成方式，将Region Proposal交给CNN来做，为此提出了RPN区域间以网络获取检测区域，这种方式能够和检测网络共享图像的卷积特征，几乎不耗费时间，大大缩减了检测框的生成时间。RPN的结构如下图所示：

图3-10 RPN数据处理

RPN分为两部分执行，在图中上部分，通过softmax分类anchors获得检测目标和图像背景，下部分用来计算anchors的bounding box regression偏移量，以此获得精准的proposals，当程序到达Proposal部分，目标定位的功能其实已经被实现了。

(3) 多通道图像卷积

在本小节我们介绍一些关于RPN的基础知识，首先我们会将一些单通道图像用单卷积核做卷积的方式：

图3-11 单通道单卷积核

在上图中我们对图像进行了扩边处理，我们设置一个3×3的卷积核，并设置步长为1，那么经过运算之后大小为M×N的图像就变为大小为(M+2)×(N+2)。在精工一个3×3的卷积就变为了M×N大小。

然后我们介绍一下多通道图像用多卷积核做卷积的方式：

图3-12 多通道多卷积核

上图中输入图像layer m-1有四个通道，同时含有两个卷积核W1和W2，对于W1先在输入图像的四个通道分别做卷积，再将四个通道的结果相加作为W1的卷积输出，卷积核W2的处理方式和W1相同，所以对于一个卷积层，无论有多少个通道，其输出图像通道数总是等于卷积核数量。

(4) anchors

anchors是RPN网络中一个很重要的概念，它是RPN网络在运行过程中生成的一组矩形，每个矩形由左上角点和右下角点坐标来表示，其中共有9个矩形共有3个形状，长宽比分别为1:1，1:2，2:1，实际上这是一种多尺度检测方式，anchors示意图如下所示：

图3-13 anchors示意图

需要为conv layers获得的feature maps中的每个点都加9个这样的anchors作为初始检测区域，虽然这个位置并不准确，但是可以在后续过程中修正以获得准确的位置。

(5) softmax判定foreground与background

一个矩阵经过Faster R-CNN的处理到达RPN网络时其长宽变为原来的十六分之一，假设现在的长宽为w和h，那么我们经过一个1×1卷积才能达到reshape和softmax，在图中可以看到该卷积核的输出设置为18，即经过该卷积核输出图像变为的大小变为w×h×18。这是因为feature maps每个点都对应着9个anchors，而且每个anchors都有两种可能的情况，可能是foreground也可能是background，共有18种情况保存在w×h×(2×9)的矩阵中。我们认为foreground为目标所在区域而background为背景区域，经过卷积核后输出到softmax中分类获得foreground anchors，初步获得目标检测区域的检测框。

图3-14 RPN

对于softmax前后的Reshape层是为了方便softmax分类，因为softmax分类的时候需要进行foreground和background的二分类，并且这涉及到caffe中数据的存储结构问题，大致可以理解为变换数据格式特出一个维度便于softmax进行分类，之后再将其恢复原状。在这个过程中，RPN网络初步提取出了foreground anchors检测区域。

(6) bounding box regression原理

下图中绿色边框对应着目标的准确区域位置，而红色边框为上节所提取的foreground anchors检测区域，虽然红色边框确实包含了目标所在的区域，但是红色边框的定位并不准确，此时需要一种方法来使红色边框区域更加接近绿色边框区域，这种方法就是bounding box regression。

图3-15 bounding box regression

窗口也就是目标检测区域一般用其中心点坐标和窗口的长宽来表示。下图中，框A检测出来的foreground anchors，框G表示目标的准确位置区域，现在需要找到一种方法使框A经过变换后能够更加接近框G，将框A变换后的区域设为框G'。不妨设A=(Ax, Ay, Aw, Ah)，G=(Gx, Gy, Gw, Gh)，G'=(G'x, G'y, G'w, G'h)，此时需要一种变换F：使得FA=G'，并且要求G'≅G。

图3-16 变换F

对于F的实现，接下来给出了一个简单的方法来实现：

1. 平移操作实现公式 $G'x = Aw \cdot dx_A + Ax$ $G'y = Ah \cdot dy_A + Ay$
2. 缩放操作实现公式 $G'w = Aw \cdot \exp(dw(A))$ $G'h = Ah \cdot \exp(dh(A))$

在平移和缩放的过程中需要，需要学习得到 dx_A ， dy_A ， dw_A ， dh_A 这四个变换，当框A和G相差小的时候可以认为这是一种线性变换，只需要用线性回归的方式就可以对窗口进行调整。在Faster R-CNN的论文中，平移量 (tx, ty) 与尺度因子 (tw, th) 的定义如下： $tx=x-x_{awa}$ ， $ty=y-y_{ahatw}=\log w_{wa}$ ， $th=\log h_{ha}$

上述公式已经给出了调整的方法，现在的问题是获得 dx_A ， dy_A ， dw_A ， dh_A ，这里采用线性回归的方式。这个问题的输入是图像经过卷积后得到的feature map，将其表示为 Φ ，另外还有训练传入的框G，即 (tx, ty, tw, th) ，输出就是我们要求的 dx_A ， dy_A ， dw_A ， dh_A ，这个过程的函数表示如下： $d^*A=w^*T \cdot \Phi(A)$

其中 $\Phi(A)$ 是对应anchor的feature map组成的特征向量， w 是需要学习的参数， $d(A)$ 是得到的预测值（*表示 x, y, w, h ，也就是每一个变换对应一个上述目标函数）。

为了让预测值 (tx, ty, tw, th) 与真实值差距最小，设计损失函数： $Loss=iN(t^*i-w^*T \cdot \Phi(A_i))^2$

函数优化目标为： $w^*=\argmin w^*(t^*i-w^*T \cdot \Phi(A_i))^2+\lambda||w^*||^2$

上述过程即为bounding box regression的实现，这个过程对应RPN网络的下部分，即计算anchors的bounding box regression偏移量，如下图所示：

图3-17 RPN

在这条线路中经过一个 1×1 的卷积核，图像的大小变为 $w \times h \times 36$ ，因为feature maps每个点有9个anchors，而每个anchors又含有四个用于回归计算的 dx_A ， dy_A ， dw_A ， dh_A 变量。

(7) Proposal Layer

Proposal Layer有foreground/background anchors的分类结果，bbox reg的 dx_A ， dy_A ， dw_A ， dh_A 四个变换量以及im_info三个输入，它负责整合三个输入计算出准确的proposal传递给Roi Pooling Layer。其中im_info保存的图像的缩放信息，在计算anchors偏移量时候会用到。

Proposal Layer的处理顺序如下：

- 1.生成anchors，使用 dx_A ， dy_A ， dw_A ， dh_A 四个变换量对所有的anchors做 bounding box regression回归。
- 2.按照输入的foreground softmax分类结果由大到小排序anchors，提取排名靠前个anchors，即提取修正位置后的foreground anchors。
- 3.利用im_info将foreground anchors从缩放后的尺寸映射为原图的大小，判断foreground anchors是否大范围超过边界，删除超出边界foreground anchors。
- 4.进行nms (nonmaximum suppression，非极大值抑制)。
- 5.再次按照nms后的foreground softmax分类结果由大到小排序foreground anchors，提取排名靠前的结果作为proposal输出，按照实际需要可以设计输出的个数。

执行完这一步之后，RPN网络的任务就完成了，总的来说RPN网络的执行顺序为：先生成anchors，然后安装softmax分类器提取foreground anchors，再用bounding box regression来对foreground anchors进行回归，最后在Proposal层生成proposals。

(8) Roi pooling

Roi Pooling层通过获取原始的feature maps和RPN层输出的proposal boxes计算出proposal feature maps并送入后续的网络。之所以采用Roi Pooling是因为传统的卷积神经网络输入图像的尺寸必须固定，如果不固定的话，需要采用裁剪或者缩放的方式来使其大小固定。但是这两种方法的效果都不好，裁剪会破坏图像的完整结构，缩放会破坏图像的原始形状信息。而Roi Pooling就是为了解决这个问题。

Roi Pooling的原理如下，proposal是对应着图像的原尺寸的，首先我们应该先用im_info中的缩放信息可以将其映射为feature maps的大小。然后将proposal水平和数值方向等比例分为7份，对每一份进行max pooling处理，这样处理后，大小不同的proposal都会变成 7×7 大小，这样就会实现图像的大小固定，它的实现方式如下图所示：

图3-18 Proposals

(9) 分类器

分类器部分通过全连接层和softmax来计算获得的proposal feature maps来判断每个proposal属于什么类别，输出cls_prob概率向量，同时还利用bounding box regression来计算proposal的位置偏移量，获得更加精确的目标位置。Classification的结构如下图所示：

图3-19 分类方式

从Roi Pooling获得大小为 7×7 的proposal feature maps后，执行两部分计算，一部分用来获取更加精确的位置框，另一部分用来识别图像的类型。

指 标

疑似剽窃文字表述

1. 输入设为I，输出设为O，那么这个结构可以形象的表示为： $I \Rightarrow S1 \Rightarrow S2 \Rightarrow \dots \Rightarrow Sn \Rightarrow O$ 。

2. 卷积神经网络
 - 3.1.1 卷积神经网络的优势卷积神经网络是一种用来识别二维图形的多层次前馈人工神经网络，
3. 卷积神经网络的结构
- 卷积神经网络是一个专门用来识别二维特征的神经网络，它的结构如下：输入的图像通过三个可训练滤波器进行卷积操作，卷积后在C1层产生三个特征映射图，然后特征
4. 偏置后，通过一个sigmoid函数得到三个s2层的特征映射图，这些特征映射图经过滤波得到C3层，
5. 个s4层的特征映射图，这些像素值被光栅化，并连接成一个向量输入到传统的神经网络中，得到输出。
- 图3-1
6. 神经元，每个神经元与前一层的局部感受视野相连接，并提取出局部特征，一旦局部特征被提取，那么它与其他特征层的位置关系也就确定下来了。
7. 假设我们有一张1000x1000像素的图像，图左为全连接网络，需要1百万个隐层神经元，每个神经元连接一个像素点，那么就需要1000x1000x1000000个连接，也就是 10^{12} 个权值参数。而图右为局部连接网络，每个节点与上层
8. 在卷积神经网络中，每个hi过滤器会通过权值共享覆盖整个可视域，这些共享权值的单元构成了一个特征映射，如图3-3所示。
- 图3-3 权值共享映射结构
- 在图3-3中，有三个隐藏单元，他们属于一个特征映射。同种颜色的链接权值是相同的，可以通过梯度下降的方式来学习这些权值。
9. 共享的优点是：重复单元能够对特征进行识别，不需要考虑其在可视域的位置。权值共享能够使我们更有效的进行特征的提取，减少需要学习的参数的个数。
10. 3.1.4 卷积神经网络的训练过程
- 卷积神经网络本质上是一种输入到输出的映射，它不需要任何已知的输入和输出之间的精确数学表达，
11. 已知模式的网络进行训练，网络就能自动获得输入和输出之间的映射关系，从而能够对未知数据进行
12. 所有的向量对都来源于网络模拟系统的实际运行结果。在开始训练之前，所有权都应该用一些不同的小随机数进行初始化，这样可以保证网络不会因为权值过大而进入饱和状态，从而致使训练失败。小随机数不能相同，因为相同的数去初始化权矩阵，网络是没有学习的能力的，不同的随机数可以保证网络正常的学习。
13. 在这个阶段，信息从输入层经过逐级变换传递到了输出层。这个过程也是网络训练完之后能够正常运行时执行的过程。
14. 这两个阶段的工作一般受到精度要求的控制，用下面的式子计算 E_p ， E_p 计算网络关于第p个样本的误差测度，而将网络关于整个样本集的误差测度定义
15. 之所以将这个阶段称为向后传播阶段，是对于输入信号的正常传播方向而言的，因为在开始调整神经元的连接权时，只能求出输出层的误差，而其他层的误差都要通过次误差反向逐层后推得到。
- 3.2 卷积神经网络
16. 计算anchors的bounding box regression偏移量，以此获得精准的proposals，
17. 输入图像layer m-1有四个通道，同时含有两个卷积核W1和W2，对于W1先在输入图像的四个通道分别做卷积，再将四个通道的结果相加作为W1的卷积输出，卷积核W2的处理方式和W1相同，所以对于一个卷积层，无论有多少个通道，其输出图像通道数总是等于卷积核数量。
18. 框A检测出来的foreground anchors，框G表示目标的准确位置区域，

5. 第4章实验分析与结果

总字数：1597

相似文献列表 文字复制比：0%(0) 疑似剽窃观点：(0)

原文内容 红色文字表示存在文字复制现象的内容; 绿色文字表示其中标明了引用的内容

第4章实验分析与结果

本文采用脊柱骨骼X线图片数据集在Faster RCNN网络结构框架上进行脊柱关节的检测，然后根据脊柱关节的中心位置拟合曲线从而识别出脊柱侧弯Cobb角的大小。

4.1 数据集的制作与处理

本文中的脊柱骨骼数据集来自于医院，整个数据集共含有29张图片，图片的大小均为1024*1024，我们选择25张图片来做训练集，4张图片检测训练效果。

我们的数据集是均为未标注的脊柱骨骼X线图像，因此我们需要对图像进行预处理。首先采用图像标注工具对图像进行标注，生成XML信息。我们采用的标注工具为Labelme，这个标注工具实现简单，性能十分优异。

我们是仿照VOC 2007数据集的格式对图像进行标注，标注的详细方式为自最后一节脊柱骨骼向上依次标注16块脊柱骨骼的位置，并将它们归为同一类别‘1’，标注完成之后将每个边框的左上角点的横纵坐标和右上角点的横纵坐标保存在XML文件中，在这个处理过程中应尽量将标注框的中心与每块脊柱骨骼的中心重合。标注完数据之后需要对数据重命名，但是图像和对应的XML信息的命名应该一致。然后我们根据其名字生成随机数据保证数据传入神经网络不是有序的。

4.2 模型训练

Faster RCNN网络我们采用Caffe深度学习框架来实现，根据当前深度学习目标检测的主流实现方式，我们选择在

ImageNet分类问题上预训练好的模型初始化训练网络。我们算法的实现就是通过将ImageNet分类训练得到的ZF卷积网络模型初始化特征提取网络卷积层。我们将学习率设置为0.000001，共进行6000迭代。实验使用设备为Elementary OS操作系统，我们是在CPU上进行运算所以没有使用GPU资源，运算速度较慢。

4.3 训练结果及分析

我们进行6000次迭代之后，得到了最终的训练模型，通过该模型我们检测图片，其结果如下图所示：

图4-1 实验结果

从图中可以看出来，图中边框在精度设置较低的情况下，能够较为有效的识别出脊柱骨骼上的点，但是识别的精度并不高。据我分析，能够识别出脊柱骨骼大致位置是因为：1、我们选用的数据集十分适合进行深度学习目标检测，因为其具有优良的对比度和细致的层次性；2、Faster RCNN利用RPN生成脊柱骨骼关节候选区域，然后将其传入后续网络进行分类的结构效果十分优良；3、ZF卷积神经网络对于小数据集训练效果比较突出。至于实验结果精度不高具有如下几点原因：1、试验所得数据集所含图片太少，整个训练样本包含的脊柱骨骼结构不多；2、因为采用CPU来训练资源，导致训练时间过慢，所以减少了模型训练的迭代次数，模型迭代次数不多，模型并未达到完全收敛。

4.4 曲线拟合及结果分析

当经过Faster RCNN目标检测网络的训练之后生成了最终的模型，根据模型我们检测了几张图片，上述过程已经展示过了，我们将其中的蓝色边框去掉，并调细边框。我们根据图中边框的位置计算出边框的中心点，并将这些中心点保存下来，然后按照这些中心点调换其横纵坐标，因为在图像中其横纵坐标的排放位置不方便我们进行曲线拟合，交换横纵坐标之后，然后将其保存在文件中。读取文件中的信息，对文件中点进行拟合，然后按照拟合图像。

4.5 实验结果及分析

我们可以从上图看出来虽然我们输出了用于曲线拟合的图像目标检测区域的算法但是我们的图像识别结果还是有一定的误差的，这样的误差来自于两个方面：1、训练过程中的误差，我们的目标检测还是没有达到预想的精度，原因我们已经在上述内容分析过了；2、曲线拟合的时候存在拟合误差，这样的误差还没有办法避免，只能尽可能通过细致的操作来减少这样的误差。但是通过调整参数，增加数据量已经增加训练的次数，我认为这样的误差是可以避免的，然后通过曲线拟合来识别出来曲线的角度也可以更加精确，完全能够达到辅助诊断的效果。

6. 第5章结论与展望		总字数：246
相似文献列表 文字复制比：0%(0) 疑似剽窃观点：(0)		
原文内容 红色文字表示存在文字复制现象的内容; 绿色文字表示其中标明了引用的内容		
第5章结论与展望		

本文在Faster RCNN的基础上，提出一种基于Faster RCNN的脊柱骨骼目标检测算法，旨在帮助医生及患者辅助诊断脊柱侧弯及计算出脊柱侧弯的角度。通过对Faster RCNN进行参数调整，我们的实验结果有了不错的检测精度。通过曲线拟合，我们发现了Cobb角的大小。针对本文中出现的问題，我们还有很大的改进空间，下一步我们要研究的是：提高目标位置的检测精度，曲线拟合找到侧弯更精确的位置。希望我们的研究，能在医学领域得到应用，帮助医生和患者快速准确的诊断出结果。

参考文献

[1]周凯龙. 基于深度学习的图像识别应用研究[D].北京工业大学,2016.

[2]蒋树强,闵巍庆,王树徽.面向智能交互的图像识别技术综述与展望[J].计算机研究与发展,2016,53(01):113-122.

[3]丰晓霞. 基于深度学习的图像识别算法研究[D].太原理工大学,2015.

[4]孙亮.计算机智能化图像识别技术的理论性突破[J].数字技术与应用,2013(06):106-107.

[5]王宇新. 基于特征分布的图像识别方法研究与应用[D].大连理工大学,2012.

[6]胡婷鸿,万雷,刘太昂,汪茂文,陈腾,王亚辉.深度学习在图像识别及骨龄评估中的优势及应用前景[J].法医学杂志,2017,33(06):629-634+639.

[7]张永焕. 基于深度学习的肿瘤细胞图像识别[D].华东交通大学,2017.

[8]薛迪秀. 基于卷积神经网络的医学图像癌变识别研究[D].中国科学技术大学,2017.

[9]王前程. 基于深度学习的水果图像识别算法研究[D].河北大学,2016.

[10] Yong Jun Liu,Kun Qi Liu. The Fast Convolution Neural Networks Applied to License Plate Recognition[J]. Applied Mechanics and Materials,2014,3207(556).

[11] Yunzhi Wang,Yuchen Qiu,Theresa Thai,Kathleen Moore,Hong Liu,Bin Zheng. A two-step convolutional neural network based computer-aided detection scheme for automatically segmenting adipose tissue volume depicting on CT images[J]. Computer Methods and Programs in Biomedicine,2017,144.

[12] Ming Che Lee,Sheng Yu Chiu,Jia Wei Chang. A Deep Convolutional Neural Network based Chinese Menu Recognition App[J]. Information Processing Letters,2017,128.

说明：1.总文字复制比：被检测论文总重合字数在总字数中所占的比例

2.去除引用文献复制比：去除系统识别为引用的文献后，计算出来的重合字数在总字数中所占的比例

3.去除本人已发表文献复制比：去除作者本人已发表文献后，计算出来的重合字数在总字数中所占的比例

4.单篇最大文字复制比：被检测文献与所有相似文献比对后，重合字数占总字数的比例最大的那一篇文献的文字复制比

5.指标是由系统根据《学术论文不端行为的界定标准》自动生成的

6.红色文字表示文字复制部分;绿色文字表示引用部分

7.本报告单仅对您所选择比对资源范围内检测结果负责



 amlc@cnki.net

 <http://check.cnki.net/>

 <http://e.weibo.com/u/3194559873/>

中国知网“大学生论文检测系统”