

西安邮电大学

毕业设计（论文）

题目： 基于深度学习的移动超声影像识别
与分析

学院： 计算机学院

专业： 计算机科学与技术

班级： 计科 1303

学生姓名： 陈迪

学号： 04132070

导师姓名： 孟伟君 职称： 讲师

起止时间： 2016 年 12 月 5 日 至 2017 年 6 月 2 日

毕业设计（论文）声明书

本人所提交的毕业论文《基于深度学习的移动超声影像识别与分析》是本人在指导教师指导下独立研究、写作的成果，论文中所引用他人的文献、数据、图件、资料均已明确标注；对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式注明并表示感谢。

本人完全理解《西安邮电大学本科毕业设计（论文）管理办法》的各项规定并自愿遵守。

本人深知本声明书的法律责任，违规后果由本人承担。

论文作者签名：

日期： 年 月 日

西安邮电大学本科毕业设计(论文)选题审批表

申报人	孟伟君	职 称	讲师	学 院	计算机学院		
题目名称	基于深度学习的移动超声影像识别与分析						
题目来源	科研	√			教学		其它
题目类型	硬件设计		软件设计	√	论文		艺术作品
题目性质	应用研究		√		理论研究		
题目简述	<p>对疾病能够进行及时准确的诊断是临床医学的基础。中医可以依靠望闻问切，而对西医来说借助医学影像做出更准确诊断已经是常用手段，而临床诊断的 70%依靠医学影像。医学影像产业整体市场有着相当客观的需求规模。传统的单纯人工诊断受医生水平的制约，误诊率较高。随着计算机软件新技术发展，采用深度学习等人工智能的方法，研究对影像的识别分析有着广阔的应用前景和巨大的市场潜力和可预期的良好的经济效益和社会效益。</p>						
对学生知识与能力要求	了解超声图像的背景知识，掌握大数据处理相关技术。						
具体任务以及预期目标	移动超声图像大数据集的构建；大数据分类器的开发；临床诊断决策支持模型的构建。 成果形式是软件系统和毕设论文						
时间进度	2016.12.5-2016.12.10 查阅资料、理解任务、进行调研，完成开题报告 2016.12.10-2016.12.20 系统需求分析 2016.12.21-2017.1.3 系统概要设计 2017.1.4-2017.1.24 系 统 详 细 设 计 2017.1.25-2017.3.27 系统实现 2017.3.28-2017.4.10 系统测试与完善 2017.4.11-2017.5.22 撰写论文 2017.5.23-2017.6.2 准备答辩						
系（教研室）主任 签字	年 月 日			主管院长 签字		年 月 日	

西安邮电大学本科毕业设计（论文）开题报告

学生姓名	陈迪	学号	04132070	专业班级	计科 1303
指导教师	孟伟君	题目	基于深度学习的移动超声影像识别与分析		
<p>选题目的（为什么选该课题）</p> <p>问题现状及需求：近年来，深度学习方法得到广泛应用，并且已经在图像识别、语音识别等领域取得了重要的成果。深度学习方法是从人工神经网络发展而来的。在医学中会存在有大量医疗影像数据，如核磁共振影像、X 光等。人工分析这部分数据目前困难较多，易受干扰且效率不高。深度学习技术，尤其是 CNN 模型，在图片分类、关键部位发现等方面表现优异，能够更加快速准确地进行疾病诊断。</p> <p>自身条件：我是计算机科学与技术专业的学生，有良好的专业基础，熟悉 Python 语言及机器学习算法；自己对深度学习很感兴趣，升学后从事的研究方向也是深度学习方面的研究；选择这个课题，一方面能提升自己的科学研究能力，另一方面自己也能提高自己的算法运用能力，为今后的深入研究打下基础。</p>					
<p>前期基础（已学课程、掌握的工具，资料积累、软硬件条件等）</p> <p>1.已学课程：机器学习课程（Andrew Ng）,数据库原理及应用、软件工程导论。</p> <p>2.掌握的工具：Python、MySQL、sklearn。</p> <p>3.通过上网查找资料，对研究所需要用的技术有了大致的了解。</p> <p>4.软硬件条件：</p> <p> 软件：Linux 操作系统、vim 文本编辑器、MySQL 数据库。</p> <p> 硬件：PC 机一台。</p>					
<p>要研究和解决的问题（做什么）</p> <p>项目的研究内容是基于深度学习的移动超声影像识别与分析。移动超声影像识别与分析主要由以下部分构成：疾病超声图像大数据集，基于 GPU 集群的大数据分析器和临床诊断决策支持模型，其中大数据分析器由多个相互独立的自适应深度学习分类器构成。基于深度学习的移动超声影像识别与分析的输入为疾病超声数据。</p>					

工作思路和方案（怎么做）

思路：

- 1.移动超声图像大数据集的构建；
- 2.大数据分类器的开发和 GPU 集群的部署；
- 3.临床诊断决策支持模型的构建；
- 4.决策支持算法的准确率、灵敏度和速度的检验。

方案：

1.移动超声图像大数据集的构建。第一步是将不同来源的超声图像数据：超声图像数据 1，超声图像数据 2，…，超声图像数据 n，构成大数据集。

2.每一个超声图像数据将产生一个相应的诊断结果：超声图像数据 1 产生诊断结果 r_1 ，超声图像数据 2 产生诊断结果 r_2 ，…，超声图像数据 n 产生诊断结果 r_n ，所有的诊断结果构成早期预警集合 $R = (r_1, r_2, \dots, r_n)$ 。主要包括基于自适应深度学习的超声图像分类器及其在单 GPU 上实现，高效快速分类器的异构并行算法及 GPU 集群的实现。

3.临床诊断决策支持模型的构建。

来自于各个移动超声图像数据的分类结果由临床诊断决策支持模型汇总，得出临床诊断决策支持的结论。因此，临床诊断决策支持模型的设计是早期诊断的决策支持算法的重要组成部分。

4.早期诊断的决策支持算法的灵敏度、准确率和速度的检验

基于移动超声图像数据疾病早期诊断的决策支持的 GPU 算法开发完成以后，其准确率、灵敏度和速度必须经过严格检验。可以按照机器学习算法检验的方法，采用医院的超声图像作为测试集，测试基于 MRI 图像大数据分析的老年神经退行性疾病早期诊断的决策支持的 GPU 算法的准确率(accuracy)、接收器工作特性曲线(Receiver Operating Characteristic, ROC)和灵敏度(sensitivity)。

5.撰写论文,准备答辩。

指导教师意见

签字：

年 月 日

西安邮电大学毕业设计 (论文)成绩评定表

学生姓名	陈迪	性别	男	学号	04132070	专业 班级	计科 1303
课题名称	基于深度学习的移动超声影像识别与分析						
指导教师 意见	<div>(从开题论证、论文内容、撰写规范性、学习态度、创新等方面进行考核)</div> <div>评分 (百分制): 指导教师(签字): _____ 年 月 日</div>						
评阅 教师 意见	<div>(从选题、开题论证、论文内容、撰写规范性、创新和预期成果等方面进行考核)</div> <div>评分 (百分制): 评阅教师(签字): _____ 年 月 日</div>						
验收 小组 意见	<div>(从毕业设计质量、准备、操作情况等方面进行考核)</div> <div>评分 (百分制): 验收教师(签字): _____ 年 月 日</div>						
答辩 小组 意见	<div>(从准备、陈述、回答、仪表等方面进行考核)</div> <div>评分 (百分制): 答辩小组组长(签字): _____ 年 月 日</div>						
评分比例	指导教师评分 (30%) 评阅教师评分 (20%) 验收小组评分 (20%) 答辩小组评分 (30%)						
学生总评 成绩	百分制成绩				等级制成绩		
答辩委员 会意见	<div>毕业论文(设计)最终成绩(等级):</div> <div>学院答辩委员会主任(签字、学院盖章): _____ 年 月 日</div>						

摘 要

在这个深度学习技术高速发展的时代,这些领先的技术被应用在了各个行业中,比如金融、医疗、基础服务等行业。并且,随着技术的不断更新与突破,在一些用传统的机器学习算法不易处理或者处理效果不理想的领域也取得了重要成就,比如自然语言处理、图像识别、声音识别、个性化推荐等。本文主要是研究深度学习技术在医疗行业的应用,而在中国,越来越多的老人忍受着阿尔兹海默病(Alzheimer disease, AD)的痛苦。阿尔兹海默病俗称老年痴呆症,这种病症的早期诊断尤为重要,所以本文选取了阿尔兹海默症作为研究数据源,希望能提高该病的诊断准确率。

本文基于深度学习技术对阿尔兹海默症做了以下研究与工作:

首先,介绍深度学习技术在医疗相关领域的发展现状与研究进展,回顾了传统浅层神经网络的相关理论,引出了卷积神经网络(CNN)的优势与局限性,再谈到了迁移学习技术针对该规模的数据源的优越性。

其次,介绍了本文使用的深度学习框架 Caffe 的搭建与使用,并对 Caffe 的几大特性进行了详尽的介绍。

然后,详细地说明了本文所使用的迁移学习技术的模型设计与调参方法,以及模型训练的成果。

最后,开发了一个可视化的在线智能诊断阿尔兹海默病的 web 系统,方便用户体验到深度学习技术在医疗领域实际效果,使读者感受到人工智能的切身美妙体验。

关键字: 深度学习; 机器学习; 阿尔兹海默症; 卷积神经网络

ABSTRACT

At the depth of this depth of learning technology, these leading technologies have been used in various industries, such as finance, healthcare, and basic services. In addition, with the continuous updating and breakthrough of technology, in some with the traditional machine learning algorithm is not easy to deal with or processing ineffective areas have also made important achievements, such as natural language processing, image recognition, voice recognition, personalized recommendation. This paper focuses on the application of deep learning techniques in the medical industry, while in China, more and more elderly people suffer from the pain of Alzheimer disease (AD). Alzheimer's disease is known as Alzheimer's disease, the early diagnosis of this disease is particularly important, so this article selected Alzheimer's disease as a research data source, hoping to improve the diagnostic accuracy of the disease.

This article based on the depth of learning techniques for Alzheimer's disease made the following research and work:

Firstly, the present situation and research progress of depth learning technology in medical related field are introduced. The theory of traditional shallow neural network is reviewed, and the advantages and limitations of convolution neural network (CNN) are introduced. The superiority of the data source for this scale.

Secondly, it introduces the construction and use of Caffe, the depth learning framework used in this paper, and introduces the several characteristics of Caffe in detail.

Then, the model design and tuning method of migration learning technology and the result of model training are described in detail.

Finally, a visualized online intelligent diagnosis of Alzheimer's disease web system was developed to facilitate the user to experience the actual effect of the depth learning technique in the medical field and to enable the reader to feel the wonderful experience of artificial intelligence.

Key Words:Deep Learning;Machine Learning;Alzheimer's disease;CNN

目 录

第一章 绪论	1
1.1 课题背景与研究意义.....	1
1.2 国内外研究现状.....	1
1.3 本文组织安排.....	2
第二章 神经网络概述与结构	3
2.1 神经网络概述.....	3
2.2 卷积神经网络结构.....	5
2.3 迁移学习.....	6
第三章 Caffe 深度学习框架的介绍与使用	8
3.1 Caffe 深度学习框架介绍.....	8
3.2 Caffe 深度学习框架使用.....	8
第四章 迁移学习模型训练	11
4.1 数据集描述.....	11
4.2 数据集的预处理方法.....	11
4.3 训练过程的参数设置方法.....	12
4.4 模型的成果与评价.....	13
第五章 智能诊断可视化	15
5.1 智能诊断可视化的介绍.....	15
5.2 可视化系统相关开发技术介绍.....	15
5.3 系统设计.....	16
5.4 界面展示.....	19
第六章 结论与展望	21
6.1 结论.....	21
6.2 展望.....	21
致谢.....	22
参考文献.....	23

第一章 绪论

1.1 课题背景与研究意义

近年来，阿尔法狗横空出世，在围棋大战中战胜了韩国围棋大师李世石，瞬间将人工智能和深度学习这些“名词”推向了风口浪尖。我们生在这个时代是幸运的，因为真正的感受到了人工智能给生活带来的便利性。

深度学习在近年来的发展中，不仅成就了阿尔法狗，还在各行各业的发展进程中起到了不可或缺的作用。比如，金融、医疗、服务业等行业。由于医疗是与人们生活密切相关的，所以深度学习在医疗领域的研究尤为重要。

深度学习在医疗领域的研究发展有四大方向：一是提供早期诊断与风险预测；二是医学影像的识别，协助医生对疾病的争端；三是在医药研发领域，缩短制药周期，降低制药成本。四是用在基因测序，促进精准医疗。

2012 年，大规模视觉识别竞赛（ILSVRC）中，深度神经网络在图像识别与分类的领域中表现效果相较于传统的机器学习识别更加优秀，所以文本希望把深度学习技术在图像处理的优秀应用到医学影像处理中。在 ILSVRC2014 中，Googlenet 卷积神经网络获得了冠军，但其在医学影像的识别效果有待分析，所以本文将该神经网络迁移学习至医学影像中，并通过调参的方法以求得一个优秀的针对医学影像的预测模型。并且，作者开发了一个在线可视化诊断系统，以此给读者清晰的展示深度学习技术在医学影像的实际应用。

1.2 国内外研究现状

医学影像在医疗诊断的作用非常重要，医学影像是反映患者身体状况的重要手段。在神经外科、普通外科、手足外科等领域都需要医学影像的帮助。而深度学习技术对图像识别的高准确率就可以应用在医学影像处理中，可以协助医疗工作人员对患者进行诊断。

但是，在当下，深度学习在医学影像的应用存在一些问题：一是数据质量问题，对于深度学习来说，数据意味着教材，其质量决定着学习的效果与准确率；二是数据数量问题，要获取到大量的同一类型的数据实属不易；三是患者的隐私问题，许多患者不愿意将自己的病历提供于研究机构，这增加了我们获取数据的难度。四是观念问题，人们对人工智能在医疗诊断方面的认可度不够，从“人医”到“机医”的观念转变困难。

1.3 本文组织安排

第一章先概括介绍了整个论文的研究背景与国内外研究现状。第二章具体介绍了神经网络的结构与思想以及每一种神经网络的优缺点。第三章介绍了本文所使用的深度学习框架的优势以及使用方法。第四章讲解了本文所使用的迁移学习的参数设置方法以及训练的要点。第五章介绍了一款作者开发的疾病诊断系统的设计开发流程，以表现深度学习技术在实际生活中的应用。第六章是总结章节，总结了本文的研究与设计成果，并表达了作者对未来深度学以及人工智能发展的期望与研究方向。

第二章 神经网络概述与结构

2.1 神经网络概述

2.1.1 传统神经网络

深度学习的思想是构建多层神经网络，输入数据源通过一层一层的神经网络进行感知下去，每一层都会提取出当前层的特征，然后在最后一层中，提取出的特征与已知特征进行比对，利用该误差来调整前几层的参数，使最终的误差达到最小值。图 2.1 是个传统神经网络的层次结构^[1]。

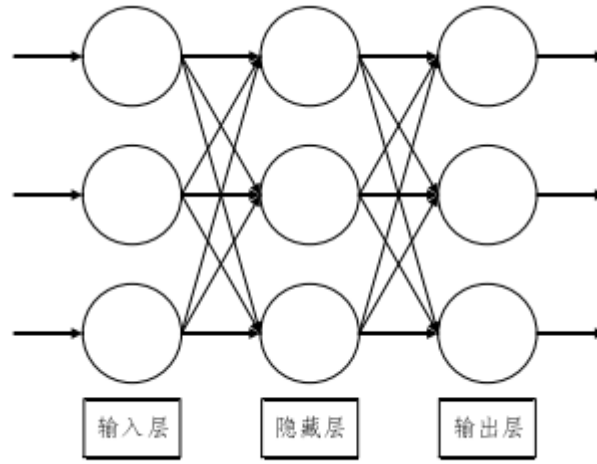


图 2.1 传统神经网络层次结构

由图可见，传统神经网络中间只有一层隐藏层，也就是说数据输入后，只进行一次中间层的感知计算就要进行输出层与已知特征进行误差比对。传统神经网络应用了 BP 算法，它是基于统计模型的机器学习算法，在早些年受到热捧。因为其可以从大量数据样本中基于统计原理得到特征并对其分类，解放了手工分类。BP 算法如公式 2.1 所示。

$$Z_k^i = \sum_j w_{ij} x_j^{k-1} \quad (2.1)$$

其中， k 表示第 k 层， i 表示第 k 层的神经元 i ，输入结果表示为 Z_k^i ， x_j^{k-1} 表示第 $k-1$ 层的第 j 个神经元， w 表示神经元的权重系数。BP 算法利用输出特征与实际特征之间的误差来调整权重系数。在调整参数的过程中，主要是应用梯度下降法^[2]。

熟悉了传统神经网络，可以更好地理解深度神经网络，深度神经网络模型是指隐藏层不止一层，甚至有很多层，比如本文所使用的 Googlenet 卷积神经网络

就达 22 层。层数变多后好处是能够使整个模型更加贴近实际数据样本，但缺点是参数难以调整，且容易出现过拟合现象。

2.1.2 卷积神经网络

以上谈到了深度神经网络的一个重要问题——参数过多，利用了自然图像的固有特性，可以运用卷积神经网络（CNN）来大量降低参数量，介绍一个卷积神经网络的一些重要概念。

（1）局部感受野：在全连接中，每个神经元与前一层的所有神经元均连接起来，这样的话训练起来就非常的费力，参数也太多。而局部感受野就很好的解决了这个问题，局部感受野是指每个隐层神经元仅仅前一层输入的一部分，不是所有都连接，这样的话就大大减少了参数个数，使训练简化。这一理论的是受神经系统的启发，视觉神经元也是局部感知信息的。况且，图像是具有局部特征的特性的，如图 2.2 所示，左边是全连接网络，右边是部分连接^[3]。

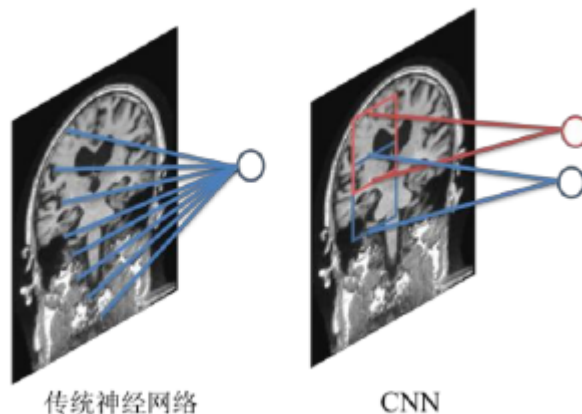


图 2.2 神经网络对比图

（2）权值共享：虽然部分连接降低了很多参数，但参数依然不少。这个时候就需要权值共享了，权值共享的意思是，当前层在卷积的时候参数一致的，不需要修改，这样的话参数会以指数级下降。这一理论是利用图像的固有特性，意思是某一部分学习到的特征与其他部分学习到的特征是一致的，可以用一个相同的权重来表示。卷积神经网络的卷积意思就是过滤器（filter）与前一层神经元特征的局部求和过程，只用一个过滤器提取到的特征不完整充分，所以需要多个过滤器来提取。卷积的过程又叫特征提取^[4]。

（3）池化：我们在通过卷积得到了特征后，理论上我们根据这些特征向量使用业界的一些分类器，比如 softmax、svm 等进行分类了。但是由于是图片数据，所以特征向量的维数巨大，直接上分类器会导致运算繁重，甚至出现过拟合

(over-fitting) 的现象。为了解决这个问题，首先上述提到图像具有一种固有特性，某一部分学习到的特征对其他区域是适用的。所以，我们可以对不同位置的特征进行一系列的运算统计，比如求平均值或者求最大值。这样的话可以将特征的维数大幅度下降，这样还不容易过拟合。这种运算统计的操作就叫做池化。池化的过程通常也叫特征映射^[5]。

2.2 卷积神经网络结构

本文所使用的卷积神经网络为 GoogLeNet 卷积神经网络，GoogLeNet 是在 ILSVRC2014 比赛中排名第一的神经网络结构，所以我使用该网络迁移学习到医学影像中。GoogLeNet 的网络结构图如图 2.3。

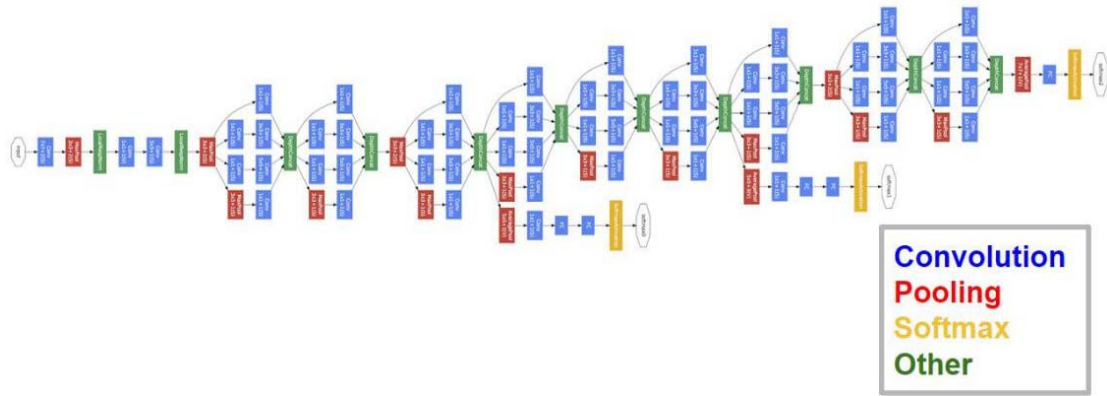


图 2.3 GoogLeNet 网络结构

GoogLeNet 提出了 Inception module 的概念，旨在强化基本特征提取模块的功能，一般的卷积层只是一味增加卷积层的深度，但是在单层上卷积核却只有一种，比如对于 VGG，单层卷积核只有 3x3 大小的，这样特征提取的功能可能就比较弱。GoogLeNet 想的就是能不能增加单层卷积层的宽度，即在单层卷积层上使用不同尺度的卷积核，GoogLeNet 构建了 Inception module 这个基本单元^[6]。基本的 Inception module 中有 1x1 卷积核，3x3 卷积核，5x5 卷积核还有一个 3x3 下采样，如图 2.4 所示。

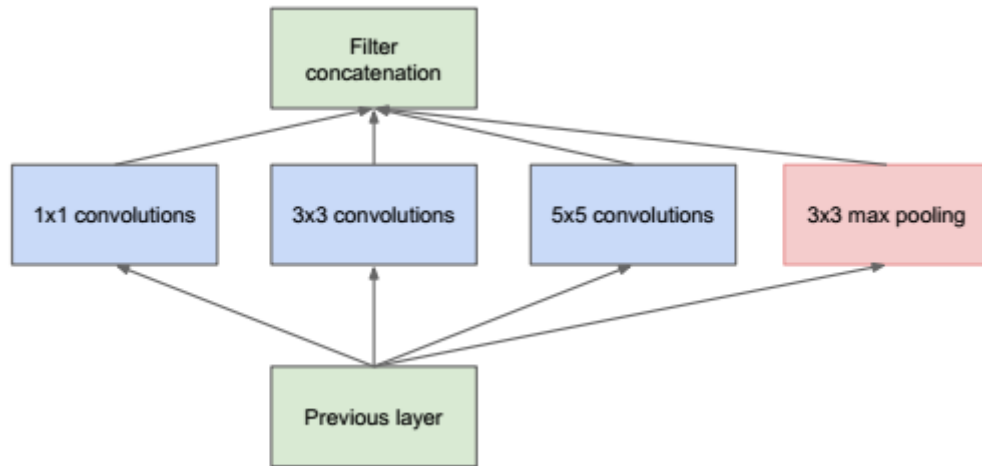


图 2.4 Inception Module 基本单元

这样尺寸不同的卷积核就可以提取不同尺寸的特征，单层的特征提取能力增强了,上面 Inception module 结构会存在一个问题，前一层的输出不经任何处理直接输入到下一层的卷积层中，这样如果前一层的特征图数量很多，又经过 5x5 这种大尺寸的卷积，带来的计算量也是非常大，所以在修正过后的 Inception module 在输出到下一层卷积之前，会对数据进行一次降维，采用 1x1 卷积核实现^[7]。比如原来有 256 个输入特征图，先用 1x1 的卷积核对特征图进行线性组合，使得输出只有 96 的特征图，这样不会带来信息损失，又减小了下一层卷积的计算量，修正后的 Inception module 如图 2.5 所示。

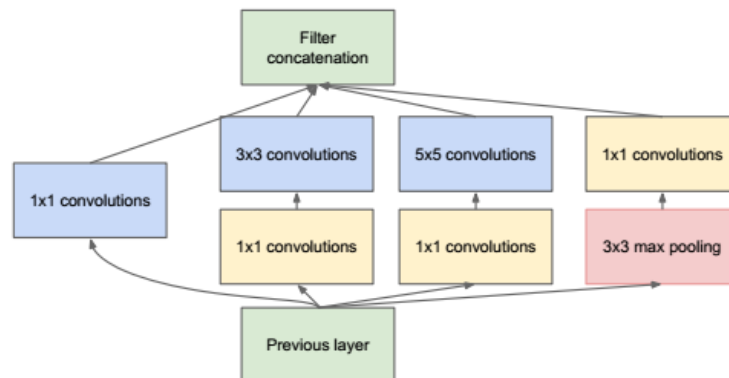


图 2.5 修正后的 Inception Module

这种 Inception module 既能大大提升卷积提取特征的能力，又不会使计算量提升很多。

2.3 迁移学习

迁移学习（Transfer Learning）最早出现在人类学习中，利用现有知识的基本

架构运用到新的知识中。因此迁移学习关心的问题是：什么是“知识”以及如何更好地运用之前得到的“知识”。这种方式在语言的学习中尤为明显，比如一个人学习了英语后，他再学习法语，由于两者有诸多的相似之处，因此往往在学习法语的过程中会利用到现有英语的相关知识^[8]。

因为现阶段的模型由于其大量参数，因此在训练中需要海量的数据。在面对某一领域的具体问题时，通常可能无法得到构建模型所需规模的数据。例如最常见的障碍在于训练模型过程中所需的海量数据，收集这些标定好的训练数据会耗费大量的人力与物力^[9]。另外，机器学习假设训练数据与测试数据服从相同的数据分布。然而许多情况下，这种同分布假设并不满足。通常可能发生的情况如训练数据过期，也就是好不容易标定的数据要被丢弃，而另外有一大堆新的数据要重新标定^[9]。

正是人们使用深度学习遇到了这些用传统手段无法克服的困难，人们就想到自己在学习中遇到这种困难如何克服，即利用现有知识来加快速度并且减少对客观条件的需要，所以迁移学习这种源自于人类的一项高级技能很自然的被科学家运用到人工智能中，例如我们可以通过一项涉及两套机器学习系统的实验来思考这个问题，为了简单起见，我们将它们称为机器 A 和机器 B。机器 A 使用全新的 DNN(深度神经网络)，机器 B 则使用已经接受训练并能理解英语的 DNN^[10]。现在，假设我们用一组完全相同的法语录音及对应文本来对机器 A 和 B 进行训练，大家觉得结果会怎样？令人称奇的是，机器 B（曾接受英语训练的机器）展现出比机器 A 更好的法语技能，因为它之前接受的英语训练将相关能力迁移到了法语理解任务中。

而本文的数据源——阿尔兹海默症的医学影像的样本量很少，故作者想要运用迁移学习的技术，以求得在少量的样本量的情况下，达到一个很高的正确识别率。

第三章 Caffe 深度学习框架的介绍与使用

3.1 Caffe 深度学习框架介绍

Caffe 是一个高效实用的深度学习框架,它是用 C++语言编写,提供了 Python 接口和 Matlab 接口,也就是说你可以不了解 C++也可以熟练地上手使用它。Caffe 有三种模式:命令行模式、Python 模式、Matlab 模式。使用任何一种模式都可以快速进行深度学习训练。Caffe 是基于 OpenCV 和 Cuda 开发的,在 CPU 和 GPU 的切换也只需要一行代码即可完成^[1]。

Caffe 框架有几大特性:

- 速度快:支持 CuDnn 加速,使你的训练速度更上一层楼。
- 结构清晰:每个模块编写规范,功能明确。
- 模型易修改:修改模型只需要类似于拼积木的方式拼凑完成。
- 参数易设置:所有参数均可以修改参数设置文件修改。
- 新增层方便:如果用户想要自己开发一层崭新的层,只需要在相应位置替换即可。
- 社区化:可以通过 Github 于业界讨论交流以及提交代码。

3.2 Caffe 深度学习框架使用

Caffe 的使用关键在编写两个 protxt 文件,首先是 train.protxt,该文件定义了神经网络结构,该文件的样例如图 3.1,有两种方式生成该文件。第一,可以直接根据 train.protxt 的结构规范,手写出神经网络的结构。第二,可以根据 caffe 提供的 Python 接口,以 Python 的方式编写神经网络结构,然后 Caffe 会自动生成 train.protxt 文件。其次是 solver.protxt,该文件定义了训练参数的设置,格式见图 3.2。

```
layer {
  name: "norm1"
  type: "LRN"
  bottom: "conv1"
  top: "norm1"
  lrn_param {
    local_size: 5
    alpha: 0.0001
    beta: 0.75
  }
}
```

图 3.1 train.protxt 格式图

```

net: "models/bvlc_alexnet/train_val.prototxt"
test_iter: 1000
test_interval: 1000
base_lr: 0.01
lr_policy: "step"
gamma: 0.1
stepsize: 100000
display: 20
max_iter: 450000
momentum: 0.9
weight_decay: 0.0005
snapshot: 10000
snapshot_prefix: "models/bvlc_alexnet/caffe_alexnet_train"
solver_mode: GPU
~
~

```

图 3.2 solver.protxt 参数设置

编写好这两个文件后，下面就是需要指定数据集的来源。指定数据集来源需要编写脚本生成一个 filelist，样例如图 3.3。前面指定的是文件相对目录位置，后面指定的是数据样本的特征。

```

TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_1.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_10.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_100.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_101.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_102.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_103.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_104.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_105.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_106.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_107.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_108.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_109.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_11.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_110.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_111.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_112.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_113.png 0
TRAIN/ADHC/AD/AD_002_S_1018/AD_002_S_1018_114.png 0

```

图 3.3 filelist 样例图

完成 filelist 的编写后，可以调用 caffe 里面自带的工具 convert_imageset 将图像数据转化为 lmdb 格式的数据集，方便后续处理。当然，也可以不转化，但是 Caffe 处理 lmdb 格式的数据集速度更快。

前面的所有工作准备就绪后，就可以使用 caffe 的 train 指令进行训练了。训练过程如图 3.4

```

I0521 23:16:50.948426 16613 caffe.cpp:248] Starting Optimization
I0521 23:16:50.948438 16613 solver.cpp:272] Solving LeNet
I0521 23:16:50.948444 16613 solver.cpp:273] Learning Rate Policy: inv
I0521 23:16:50.949869 16613 solver.cpp:330] Iteration 0, Testing net (#0)
I0521 23:16:52.311308 16621 data_layer.cpp:73] Restarting data prefetching from start.
I0521 23:16:52.361565 16613 solver.cpp:397] Test net output #0: accuracy = 0.1132
I0521 23:16:52.361588 16613 solver.cpp:397] Test net output #1: loss = 2.40542 (* 1 = 2.40542 loss)
I0521 23:16:52.385943 16613 solver.cpp:218] Iteration 0 (1.03014e+33 iter/s, 1.4375s/100 iters), loss = 2.38277
I0521 23:16:52.385969 16613 solver.cpp:237] Train net output #0: loss = 2.38277 (* 1 = 2.38277 loss)
I0521 23:16:52.385988 16613 sgd_solver.cpp:105] Iteration 0, lr = 0.01
I0521 23:16:54.665733 16613 solver.cpp:218] Iteration 100 (43.8634 iter/s, 2.2798s/100 iters), loss = 0.196593
I0521 23:16:54.665769 16613 solver.cpp:237] Train net output #0: loss = 0.196593 (* 1 = 0.196593 loss)
I0521 23:16:54.665778 16613 sgd_solver.cpp:105] Iteration 100, lr = 0.00992565
I0521 23:16:56.945461 16613 solver.cpp:218] Iteration 200 (43.8649 iter/s, 2.27973s/100 iters), loss = 0.159388
I0521 23:16:56.945488 16613 solver.cpp:237] Train net output #0: loss = 0.159389 (* 1 = 0.159389 loss)
I0521 23:16:56.945497 16613 sgd_solver.cpp:105] Iteration 200, lr = 0.00985258
I0521 23:16:59.224871 16613 solver.cpp:218] Iteration 300 (43.8707 iter/s, 2.27943s/100 iters), loss = 0.167608
I0521 23:16:59.224902 16613 solver.cpp:237] Train net output #0: loss = 0.167608 (* 1 = 0.167608 loss)
I0521 23:16:59.224911 16613 sgd_solver.cpp:105] Iteration 300, lr = 0.00978075
I0521 23:17:01.506147 16613 solver.cpp:218] Iteration 400 (43.8349 iter/s, 2.28129s/100 iters), loss = 0.0732997
I0521 23:17:01.506180 16613 solver.cpp:237] Train net output #0: loss = 0.0732998 (* 1 = 0.0732998 loss)

```

图 3.4 训练过程流程图

第四章 迁移学习模型训练

4.1 数据集描述

本文的数据集是从阿尔兹海默症影像学网站(ADNI)上面下载的。这是一个由国家生物医学成像和生物工程研究所(NIBIB)，美国食品和药物管理局(FDA)，民营医药企业与非营利组织组成的研究协会，协会创建在 2014 年。虽然看上去很年轻，但网站上的数据集一分庞大。ADNI 收集了超过 800 名来自全球 59 个地方的测试者。数据集的具体规模见表 4.1：

表 4.1 数据集介绍表

类别	人数	男/女	年龄（平均值±偏差）	样本数量
AD（患病）	43	24/19	76.85±6.36	6768
HC（健康）	148	69/79	75.74±7.34	10164

4.2 数据集的预处理方法

如果想用深度学习来诊断疾病，数据样本量必须足够大。实验没有借鉴国外的标准化处理过程，因为当样本量足够大的时候，各种各样的情况都会包括进去，实验也力求降低信息的缺损^[12]。但是，不做标准化处理小代表对数据不做处理，实验需要的是没有信息缺损，脑动微小的医学影像。在这点上，实验得到了北京天坛医院医生的医学帮助，使用医学软件 MRlcron 进行了数据集的甄选。然后，将经过分解的 MRI 图像的大小调整为 256*256。实验选用了大脑的冠状位，这是由于医生在确诊 AD 时更多的根据冠状位图，实验力求可以和医生临床诊断过程一致。图 4.1 为 AD，HC 的冠状位图。

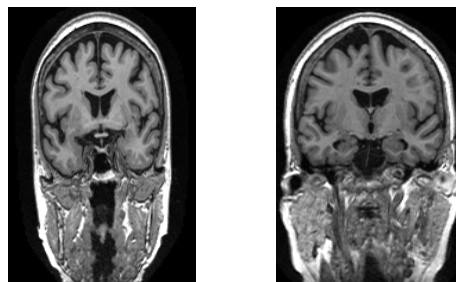


图 4.1 AD、HC 冠状位图

4.3 训练过程的参数设置方法

首先，solver.protxt 的设置如图 4.2。

```
net: "/home/chendi/ad_cure/train_val.prototxt"
test_iter: 40
test_interval: 500
test_initialization: false
display: 40
average_loss: 40
base_lr: 0.0001
lr_policy: "step"
stepsize: 20000
gamma: 0.96
max_iter: 100000
momentum: 0.9
weight_decay: 0.00002
snapshot: 5000
snapshot_prefix: "/home/chendi/ad_cure/ttt/bvlc_googlenet_snapshot"
solver_mode: GPU
```

图 4.2 训练参数设置图

对各项参数的解释如下：

net: 训练或测试的神经网络结构文件，就是前文提到的以特定格式表述神经网络的 protxt 文件。

test_iter: 完成一次测试所需要的迭代次数。本文测试样例有 2000 个样本而测试的 batch_size 是 50，故完成一次测试需要迭代 $2000/50=40$ 次。

test_interval: 测试间隔。因为训练样本一共有 15838 个，batch 是 32，将所有样本处理完一次需要 $15838/32 \approx 495$ ，所以这里将 test_interval 设置为 500，就是将所有数据处理完一遍后，才去进行测试。故这个数要大于等于 495。

Base_lr: 基础学习率，神经网络里面的学习率都是以这个为倍数制定的，因为是迁移学习，所以要将此参数设置得足够小。

lr_policy: 学习率的变化规律。

gamma: 学习率变化指数。

max_iter: 最大迭代次数。

momentum: 动量。

weight_decay: 权重衰减。

snapshot: 保存模型的间隔。因为 caffe 支持中间中断训练，调整参数，再继续训练，所以可以在训练的时候将过程记录下来。

Snapshot_prefix: 模型的保存路径。

Solver_mode: 使用 GPU 还是 CPU。

下面介绍神经网络模型的参数设置与调节方法，由于是将 GoogleNet 神经网络模型迁移学习至阿尔兹海默症的医学影像中，所以网络模型中的层间参数设置尤为重要。首先由于是二分类问题，需要将最后的全连接层的 `num_output` 设置为 2，注意由于 GoogleNet 是三输出模型，所以三个全连接层的 `num_put` 均要设置^[13]。并且更改一个名字，保证在迁移学习训练的时候能够重新训练该层的参数，该层的设置如图 4.3。

```
layer {
  name: "loss3/classifier_new"
  type: "InnerProduct"
  bottom: "pool5/7x7_s1"
  top: "loss3/classifier_new"
  param {
    lr_mult: 10
    decay_mult: 1
  }
  param {
    lr_mult: 20
    decay_mult: 0
  }
  inner_product_param {
    num_output: 2
    weight_filler {
      type: "xavier"
    }
    bias_filler {
      type: "constant"
      value: 0
    }
  }
}
```

图 4.3 模型层中参数设置

其次，需要将前面的所有层的学习率调为极低或者 0，而将最后一层全连接层的学习率提为前面的 20 到 30 倍，这样，除了全连接层以外预训练的模型参数就会不做很大的改动，而变化最大的仅仅是最后一层全连接层参数。这样，可以把在 ILSVC2014 比赛使用的预训练模型迁移到我们的阿尔兹海默症中，训练准确度较高，且不容易出现过拟合现象。

4.4 模型的成果与评价

本文将 GoogLeNet 卷积神经网络迁移学习至阿尔兹海默症的医学图像中，通过上述的调参方法，效果显著。模型的评价效果图如图 4.4、图 4.5。

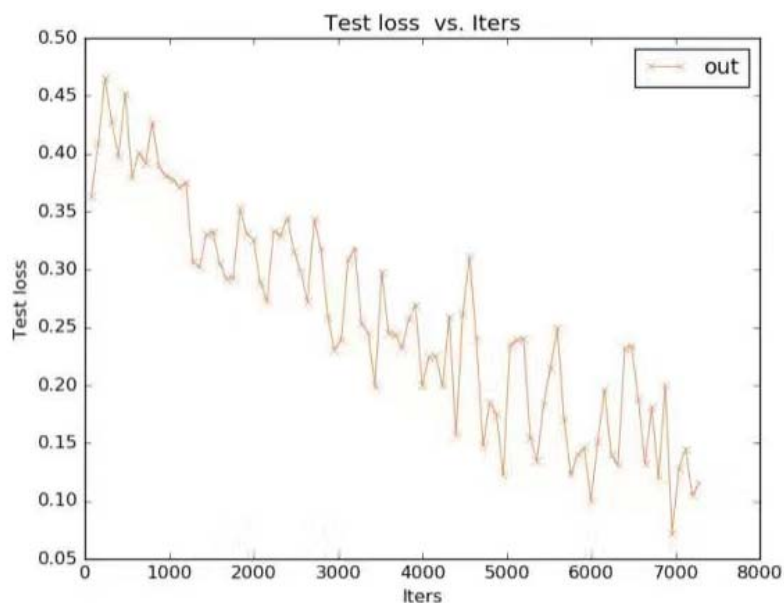


图 4.4 loss 图像

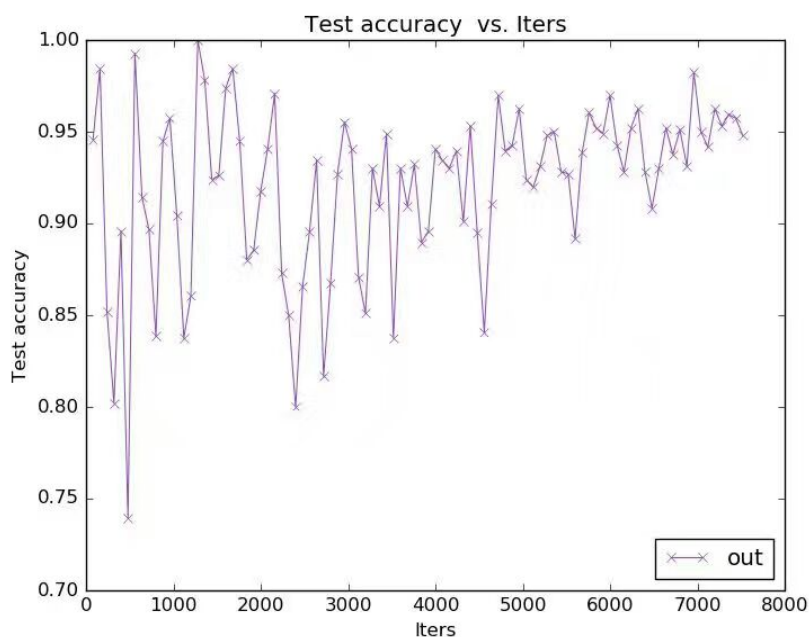


图 4.5 accuracy 图像

图 13 展示了 Caffe 在迭代了 8000 次以后代价函数返回值已经降低至 0.1 左右,已经很小了,本文前面说到过 BP 算法的目的就是调整参数使得误差最小化,可见此模型的针对阿尔兹海默症的适用性很高。图 14 展示了 Caffe 在迭代了 8000 次后,与测试样本匹配的准确率已经达到了 95%左右,也就是说,该模型确实在阿尔兹海默症的诊断中识别准确率很高。

第五章 智能诊断可视化

5.1 智能诊断可视化的介绍

现在业界在阿尔兹海默症的诊断上面还主要是医疗从业人员根据以往的诊断经验观察判断患者的脑部医学扫描图像。这种方式存在两个问题：第一，诊断速度缓慢；第二，诊断准确率不高。而本文迁移训练出的卷积神经网络模型能部分弥补上述两个缺点。分类识别的速度快，平均一张医学影像 2 秒可以识别完成；准确率高，前文已经提到了该模型针对阿尔兹海默症的识别已经达到了 95% 的准确率。

为了将前文的研究成果直观地体现在实际应用中，作者开发了一款阿尔兹海默症的在线智能诊断系统。整个可视化系统分为两个模式——训练模式和诊断模式。在训练模式中，用户可以在线上传阿尔兹海默症的医学影像数据集、在线设置调节训练参数、在线训练模型、查看训练模型成果。在诊断模式中，用户可以上传一张自己的脑部扫描图片，提交至系统中系统会智能诊断出该用户是否患有阿尔兹海默症，以及显示出发病的概率。

5.2 可视化系统相关开发技术介绍

5.2.1 web 框架 django 介绍

Django 是一个基于 Python 编写的 Web 框架，功能强大，各个模块层次分明功能清晰。Django 框架的核心组件有：

- (1) 用于创建模型的对象关系映射
- (2) 为最终用户设计的完美管理界面
- (3) 一流的 URL 设计
- (4) 设计者友好的模板语言
- (5) 缓存系统。

Django 是一个基于 MVC 构造的框架。有三个模块，模型层、存储层、模板层，它们各自的职责如表 5.1

表 5.1 Django 模块分工表

层次	职责
模型（Model），即数据存取层	处理与数据相关的所有事务：如何存取、如何验证有效性、包含哪些行为以及数据之间的关系等。
视图（View），即表现层	处理与表现相关的决定：如何在页面或其他类型文档中进行显示。
模板(Template)，即业务逻辑层	存取模型及调取恰当模板的相关逻辑。模型与模板的桥梁。

视图层仅仅决定了要把哪些内容表现给 Web 页面，其中模板决定哪些内容是不可变的，固定展现。

模式层将数据库的操作封装起来了，不需要直接使用 SQL 语句，更加的安全渐变，所有操作均可以调用 python 接口完成。

5.2.2ajax 介绍

AJAX 是“Asynchronous Javascript And XML”（异步 JavaScript 和 XML）的缩写，是一种异步处理前后端数据交互的技术。没有 AJAX 的话，WEB 页面要想更新整个页面内容的话必须刷新整个页面，用户体验很不友好。AJAX 的核心对象是 XMLHttpRequest，通过这个对象 JavaScript 可在重新刷新页面的情况给 web 服务器传输数据，通过这个技术可以在不刷新页面的情况下进行前后端交互，提高了用户体验。

5.3 系统设计

5.3.1 系统总体设计

系统总体架构在两台服务器上面，一台 Web 服务器与一台超算服务器。Web 服务器搭着 DjangoWeb 开发框架，专门负责 Web 页面的展示与业务逻辑的控制。超算服务器搭建 Caffe 深度学习框架，负责模型训练与样本分类等运算。具体分工流程图设计如图 5.1。

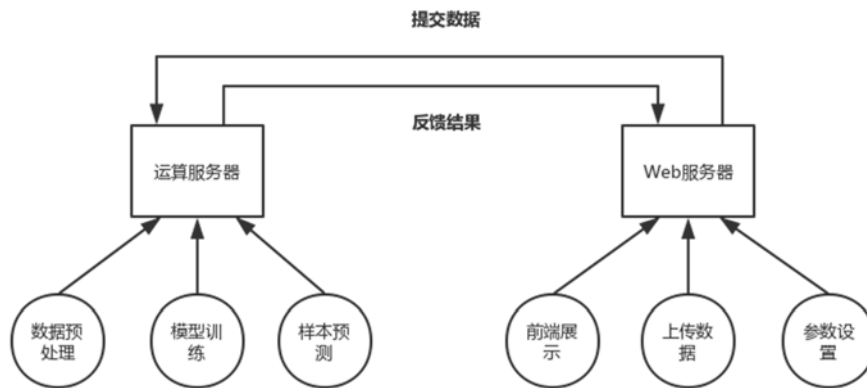


图 5.1 分工流程图

5.3.2 训练模块设计

训练模块的功能分为 4 个，上传数据集、参数设置、模型训练、查看训练模型成果。

a.事件流：

- 用户点击上传数据集，上传数据集（ZIP 格式）至 Web 服务器。
- Web 服务器将数据集包发送给超算服务器指定位置。
- 超算服务器执行循环检测脚本，发现指定位置有该数据集包，则解压数据集包，并调用数据预处理脚本，生成 filelist。
- 用户点击设置参数，用户填写参数表单，上传 Web 服务器，Web 服务器把表单内容生成 solver.protxt 的文本文件发送至超算服务器指定位置。
- 用户点击开始训练，Web 服务器向超算服务器发送运算指令，超算服务器根据 filelist 表以及 solver.protxt 开始训练模型，并将训练过程保存日志。
- 训练完成后，超算服务器将训练日志绘制成 accuracy 和 loss 的图像，并将图像发送至 Web 服务器指定位置。
- 用户点击查看图像，从指定位置查看图像。

该模块的设计流程图如图 5.2。

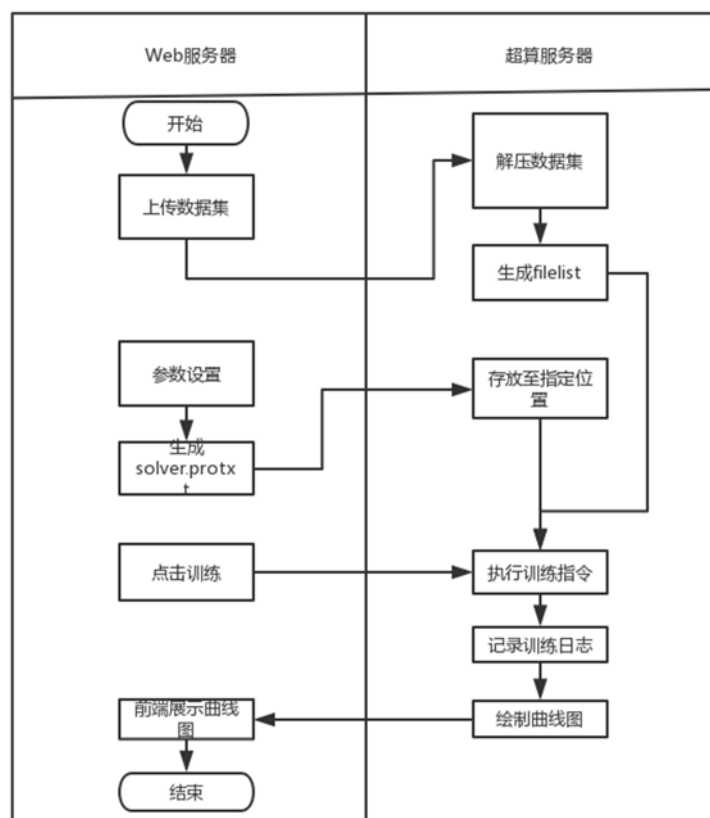


图 5.2 训练模块设计流程图

5.3.3 在线诊断模块设计

诊断模块功能有两个：上传医学图像、获得诊断结果。

b. 事件流

- 用户上传阿尔兹海默症的脑部扫描图像至 Web 服务器。
- Web 服务器将图像发送至超算服务器。
- 超算服务器执行分类预测程序。
- 返回结果给前端页面展示。

该模块的设计流程如图 5.3。

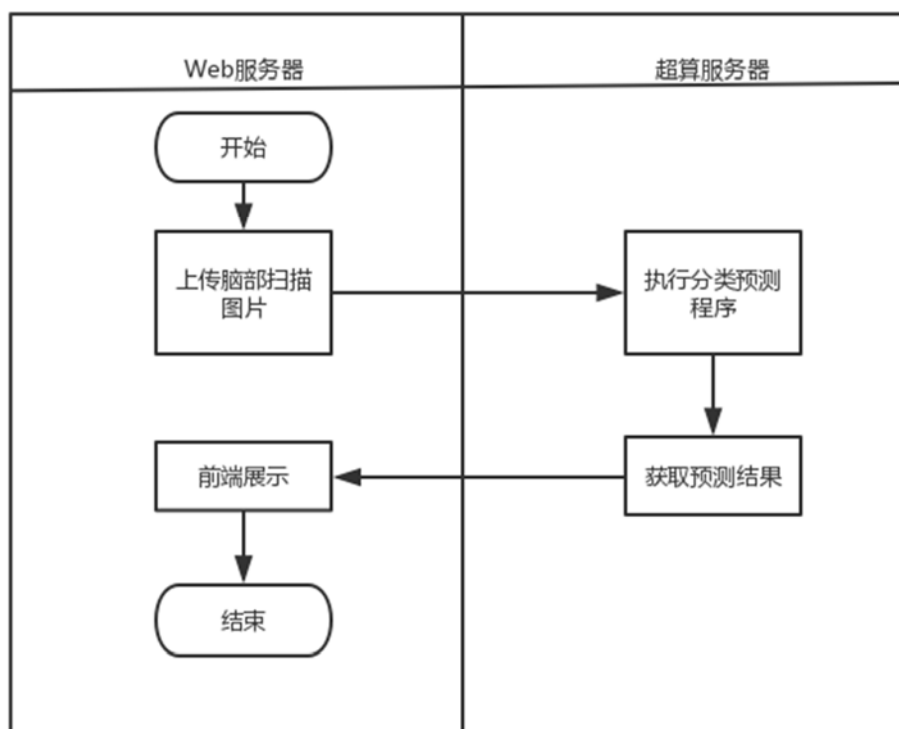


图 5.3 在线诊断模块设计流程图

5.4 界面展示

5.4.1 主界面

主界面如图 5.4，可通过主界面进入训练模式和诊断模式。



图 5.4 主界面图

5.4.2 训练模式界面

训练模式界面如图 5.5，有上传数据集、参数设置、模型训练、查看训练结果四个功能可供选择。

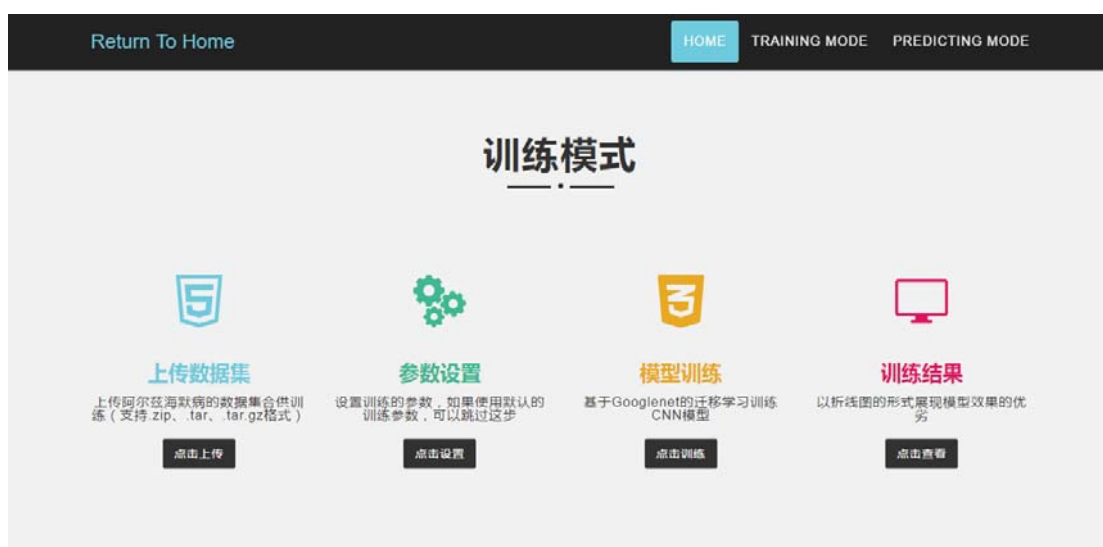


图 5.5 训练模式界面图

5.4.3 诊断模式界面

诊断模式如图 5.6，有上传图片，查看诊断结果两个功能。



图 5.6 诊断模式界面图

第六章 结论与展望

6.1 结论

在这个人工智能越发火热的时代，人工智能在各行各业的应用层出不穷，取得了各个领域的不同阶段的成就。深度学习作为人工智能的分支也在医疗领域，尤其是医学影像识别中效果突出。

本文迁移学习训练出的针对阿尔兹海默症的卷积神经网络模型在该病症的医学影像分析中表现突出，识别准确率 95% 左右，超越了业界诊断的准确率。

第五章开发的智能诊断系统采用 MVC 设计模式，层次清晰。该系统让读者不再认为深度学习技术遥不可及，切身体验到了深度学习的实际应用，感受到了人工智能给我们生活带来的变化以及给我们生活带来的便利。

6.2 展望

在这次毕业设计之前，我对深度学习这个领域不是很了解，之前只接触过机器学习相关的内容，有人说深度学习比机器学习更加复杂，在我深入接触过后，发现深度学习的难度都在调参和建模上面，跟机器学习的处理方法不太相同。用一种不恰当的形容词来形容它的话，就是深度学习更加“暴力”，这导致对机器的要求很高以及对数据集的要求很高。

迁移学习的方法有很多，比如将两个模型相结合或者提取最后一层的特征然后添加一个分类器，也许这一系列的方法都可以提高分类的准确率。所以，我希望以后在这方面的研究中，都可以逐一尝试。深度学习这个领域目前正在高速发展中，还有数不胜数的研究内容等待着我去追寻。我希望以后将深度学习应用在更多的领域中，比如，视频识别、自然语言处理、声音识别、软件自动化中。

我对人工智能这个领域充满着极大的兴趣，每天都有新的技术与新的知识等待着我去追寻。

致 谢

在本次的论文中，首先要感谢我的论文指导老师，孟伟君老师，对我的整个设计提供了很大的帮助，尤其是提供了一台带有两块 K80 GPU 的服务器。其次感谢我在中国科学院实习时指导我的赵地教授以及众多博士师兄师姐，给我提供了很多理论与实践方面的帮助。还有，要感谢北京天坛医院提供对数据预处理方法的支持。最后，再一次感谢所有给我这次论文和设计提供帮助的人。

参考文献

- [1]孙志远,鲁成祥,史忠植,马刚. 深度学习研究与进展[J]. 计算机科学,2016,(02):1-8.
- [2]曾明星,李桂平,周清平,徐洪智,董坚峰,覃遵跃,郭鑫. MOOC 与翻转课堂融合的深度学习场域建构[J]. 现代远程教育研究,2016,(01):41-49.
- [3]曾明星,李桂平,周清平,覃遵跃,徐洪智,张彬连,黄云,郭鑫. 从 MOOC 到 SPOC:一种深度学习模式建构[J]. 中国电化教育,2015,(11):28-34+53.
- [4]樊雅琴,王炳皓,王伟,唐烨伟. 深度学习国内研究综述[J]. 中国远程教育,2015,(06):27-33+79.
- [5]孙志军,薛磊,许阳明,王正. 深度学习研究综述[J]. 计算机应用研究,2012,(08):2806-2810.
- [6]张琪. e-Learning 环境中大学生自我效能感与深度学习的相关性研究[J]. 电化教育研究,2015,(04):55-61.
- [7]顾小清,冯园园,胡思畅. 超越碎片化学习:语义图示与深度学习[J]. 中国电化教育,2015,(03):39-48.
- [8]张建明,詹智财,成科扬,詹永照. 深度学习的研究与发展[J]. 江苏大学学报(自然科学版),2015,(02):191-200.
- [9]张春兰,李子运. 创客空间支持的深度学习设计[J]. 现代教育技术,2015,(01):25-31.
- [10]尹宝才,王文通,王立春. 深度学习研究综述[J]. 北京工业大学学报,2015,(01):48-59.
- [11]SONG Xiaodong,ZHANG Ganlin,LIU Feng,LI Decheng,ZHAO Yuguo,YANG Jinling. Modeling spatio-temporal distribution of soil moisture by deep learning-based cellular automata model[J]. Journal of Arid Land,2016,(05):734-748.
- [12]Zhenlong Yuan,Yongqiang Lu,Yibo Xue. Droid Detector:Android Malware Characterization and Detection Using Deep Learning[J]. Tsinghua Science and Technology,2016,(01):114-123.
- [13]Fei-wei QIN,Lu-ye LI,Shu-ming GAO,Xiao-ling YANG,Xiang CHEN. A deep learning approach to the classification of 3D CAD models[J]. Journal of Zhejiang University-Science C(Computers & Electronics),2014,(02):91-106.