

**本科毕业设计（论文）**

|  |  |
| --- | --- |
| **题目：** | **基于机器学习的舌苔检测** |
|  |  |

|  |  |  |
| --- | --- | --- |
| **学号** | ： |  |
| **姓名** | ： |  |
| **学院** | ： | 计算机与信息安全学院 |
| **专业** | ： | 信息安全 |
| **指导教师** | ： |  |
| **指导教师职称** | ： | 教授 |

年 月 日

**本科生毕业设计（论文）独创性声明**

本人声明所呈交的毕业设计说明书（论文）是我个人在指导老师指导下进行的工作及取得的成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，文中不包含其他人或其它机构已经发表或撰写过的成果；也不包含为获得桂林电子科技大学或其他教育机构的学位或证书而使用过的材料。其他同志对本工作的启发和所做的任何贡献均已在毕业设计说明书（论文）中做了明确的说明并表示了谢意。

毕业设计说明书（论文）若有不实之处，本人承担一切相关责任。

本人签名： 日期：

**本科生毕业设计（论文）使用授权说明**

桂林电子科技大学有权保留并向国家有关部门或机构送交毕业设计说明书（论文）的复印件，允许毕业设计说明书（论文）被查阅和借阅。本人授权桂林电子科技大学可以公布毕业设计说明书（论文）的全部或部分内容，可以采用影印、缩印或其它复制手段保存、汇编毕业设计说明书（论文）。

本人签名： 日期：

导师签名： 日期：

摘 要

在现代社会中,由于生活质量的日益改善，人们对自己的人体健康状况也日益关注，但是每个人都有机会忽视人体的微小改变，久而久之人体陷入亚健康容易产生各类病症。而中医学则讲究通过治疗"未病之病"，观察身体的微小改变从而确定人体的健康状况，把疾病拦截于腠理之间。

舌诊是中医中判断患者身体状况的主要方法之一，舌中蕴含着大量的人类健康数据，从而透过舌的颜色、形状的改变可以诊断出人的身体状况。但是由于过去的传统中医舌诊方法往往依赖的是医生的个人知识，而没有对舌诊原理的系统化归纳，使得舌诊对人力诊断的依赖性较高，病症识别准确度也受医师的医术及主观判断影响。如今随着现代计算机技术的不断发展，特别是人工智能技术的出现，使得利用计算机技术，结合传统中医的经验判断，对舌苔进行分析检测并判断人体健康状况成为可能。

本课题利用的是机器学习中比较热门的深度学习方法，结合中医理论对舌苔数据进行标注制作舌苔数据集，利用efficientNet网络实现对舌象图片的图像分类，能够较为精准的分辨多种舌象特征，对照中医经验，对体质进行诊断，从而能够判断人体健康状态，给出身体病症信息，实现了普通用户能够借助计算机对舌苔进行检测分析并结合中医理论和经验推断出用户自身的体质状况。

关键词： 舌苔检测；机器学习；神经网络；体质辨识

Abstract

In modern society, people pay more and more attention to their own human health because of the improvement of their quality of life. However, everyone has the opportunity to ignore small changes in the human body. Over time, people who have fallen into sub-health are prone to various diseases. Chinese medicine, on the other hand, pays attention to the treatment of "disease without disease", observing small changes in the body to determine the health of the human body, and to intercept the disease between muscles.

Tongue diagnosis is one of the main methods to judge the physical condition of patients in traditional Chinese medicine. The tongue contains a large amount of human health data, so the change of the color and shape of the tongue can be used to diagnose the physical condition of people. However, in the past, traditional methods of tongue diagnosis in traditional Chinese medicine often relied on the physician's personal knowledge, but did not systematically summarize the principles of tongue diagnosis, making tongue diagnosis highly dependent on human diagnosis, and the accuracy of disease identification is also affected by the physician's medical skills and subjective judgment. With the continuous development of modern computer technology, especially the emergence of artificial intelligence technology, it is possible to use computer technology, combined with the experience and judgment of traditional Chinese medicine, to analyze and detect tongue coating and judge human health status.

This topic makes use of the more popular method of deep learning in machine learning, labels the data of tongue coating with the theory of traditional Chinese medicine to make the data set of tongue coating, and uses efficientNet network to classify the image of tongue image. It can distinguish more accurate features of tongue image, diagnose the physique by comparing with the experience of traditional Chinese medicine, so as to judge the health status of human body and give the information of body diseases. The system enables ordinary users to detect and analyze tongue coating by computer and infer users'own physical condition with the theory and experience of traditional Chinese medicine.

**Key words**: Tongue coating detection; Machine learning; Neural network; Physical Identification

目 录

[引言 1](#_Toc103782701)

[1 课题背景 2](#_Toc103782702)

[1.1 研究背景及意义 2](#_Toc103782703)

[1.2 舌苔检测研究现状 2](#_Toc103782704)

[1.3 课题任务内容 2](#_Toc103782705)

[1.4 本章小结 3](#_Toc103782706)

[2 机器学习相关理论 4](#_Toc103782707)

[2.1 机器学习的现状与发展 4](#_Toc103782708)

[2.2 深度神经网络的结构和概念 4](#_Toc103782709)

[2.2.1 神经网络模型 5](#_Toc103782710)

[2.2.2 卷积神经网络 5](#_Toc103782711)

[2.3 神经网络的训练 7](#_Toc103782712)

[2.4 本章小结 7](#_Toc103782713)

[3 舌苔检测需求分析 8](#_Toc103782714)

[3.1 可行性分析 8](#_Toc103782715)

[3.1.1 技术可行性 8](#_Toc103782716)

[3.1.2 经济可行性 8](#_Toc103782717)

[3.1.3 文化可行性 8](#_Toc103782718)

[3.1.4 社会可行性 8](#_Toc103782719)

[3.2 功能性需求 8](#_Toc103782720)

[3.2.1 数据集构建 8](#_Toc103782721)

[3.2.2 舌苔检测 9](#_Toc103782722)

[3.2.3 体质辨识 9](#_Toc103782723)

[3.3 非功能性需求 9](#_Toc103782724)

[3.4 本章小结 9](#_Toc103782725)

[4 舌象数据集构建与扩充 10](#_Toc103782726)

[4.1 舌象图片数据的标注分类 10](#_Toc103782727)

[4.2 使用图像增强扩充数据集 10](#_Toc103782728)

[4.3 生成对抗网络 12](#_Toc103782729)

[4.3.1 生成对抗网络相关概念 12](#_Toc103782730)

[4.3.2 DCGAN生成舌象图片 13](#_Toc103782731)

[4.4 本章小结 15](#_Toc103782732)

[5 舌苔检测网络设计与实现 17](#_Toc103782733)

[5.1 网络模型介绍 17](#_Toc103782734)

[5.2 网络模型分析 17](#_Toc103782735)

[5.2.1 网络主要结构 18](#_Toc103782736)

[5.2.2 网络功能模块 18](#_Toc103782737)

[5.3 网络模型搭建及功能的实现 19](#_Toc103782738)

[5.3.1 网络模型模块 20](#_Toc103782739)

[5.3.2 数据模块 20](#_Toc103782740)

[5.3.3 训练模块 20](#_Toc103782741)

[5.3.4 检测模块 21](#_Toc103782742)

[5.3.5 体质辨识界面模块 21](#_Toc103782743)

[5.4 本章小结 22](#_Toc103782744)

[6 舌苔检测实验分析 23](#_Toc103782745)

[6.1 实验数据集 23](#_Toc103782746)

[6.2 数据图像预处理 23](#_Toc103782747)

[6.2.1 图像增强 23](#_Toc103782748)

[6.2.2 图像大小处理 23](#_Toc103782749)

[6.2.3 图像归一化 24](#_Toc103782750)

[6.3 实验参数 25](#_Toc103782751)

[6.3.1 学习率 25](#_Toc103782752)

[6.3.2 训练迭代次数 26](#_Toc103782753)

[6.3.3 训练批大小 26](#_Toc103782754)

[6.4 实验评估指标 26](#_Toc103782755)

[6.4.1 损失函数 27](#_Toc103782756)

[6.4.2 准确率 27](#_Toc103782757)

[6.5 对比实验 27](#_Toc103782758)

[6.5.1 预训练参数对比实验 28](#_Toc103782759)

[6.5.2 图像预处理对比实验 29](#_Toc103782760)

[6.5.3 学习率对比实验 32](#_Toc103782761)

[6.6 舌苔检测训练数据 34](#_Toc103782762)

[6.7 体质辨识功能展示 35](#_Toc103782763)

[6.8 本章小结 37](#_Toc103782764)

[7 结论 38](#_Toc103782765)

[7.1 全文总结 38](#_Toc103782766)

[7.2 不足之处 38](#_Toc103782767)

[谢 辞 39](#_Toc103782768)

[参考文献 40](#_Toc103782769)

# 引言

舌诊是传统中医中重要且独特的组成部分，起源西周时期，是中华民族的医学智慧结晶，通过识别舌苔的颜色和纹理等特征来获得病人的体质信息，目前成为了中医临床诊断中使用最广泛的诊断方式之一。随着现代计算机技术的不断发展，使得利用计算机技术辅助医学诊断成为了如今热门的研究方向。

传统的中医舌诊通常依靠的是医师的个人经验，使得舌诊对人力诊断的依赖性较高，而且体质的辨识通常也是依靠医师的主观判断，不同的医师因其医术和个人经验的差异，诊断的结果也有所不同。在机器学习飞速发展的今日，越来越多的相关医学项目也开始利用机器学习解决实际问题。对于舌苔检测的项目，通过采集舌象的图像信息，并利用机器学习的图像分类算法对舌象进行分类识别，再结合传统中医的经验和判断对舌象中所蕴含的人体体质信息，实现使用机器学习进行舌苔检测项目。

本次课题通过使用Python中的Pytorch 库等成熟完善的机器学习框架，搭建目前流行的神经网络模型efficient Net对舌象图片数据进行图像分类。训练网络模型时使用的在ImageNet上训练得到的训练参数efficientnet-b0作为网络模型的预训练参数，可以较快的实现收敛且能得到良好的效果，训练完成模型权重参数后，运行分类模型，对测试用的舌象图片进行识别分类，判断出舌头色泽、纹理等特征，最后结合传统中医舌诊经验对特征进行分析给出人体体质信息，完成舌苔检测流程。

对于没有多的舌象图片数据的问题，将使用数据增强（翻转、改变亮度、像素平移、增加噪声）和生成对抗网络来扩充舌象图片数据集的大小，以减少数据数量不足对舌苔检测准确率的影响。

论文作者将会对本课题的研发背景、技术背景等展开深入研究,并对本课题中所采用的机器学习的基本方案和关键技术问题作了一次简单的汇总和整理,并对使用到的深度学习算法和理论问题做出了简单的解析。

# 课题背景

## 研究背景及意义

目前随着人们生活水平的不断提高，对于中医主张的理念越来越认可，对中医的需求也越来越多。在诊断中，中医通过观察人的舌头的舌质、苔质等舌象特征，了解人体内的体质信息从而对症下药。

传统中医的舌诊主要依赖于医生的肉眼观察，仅仅通过这种人工诊断不但需要消耗大量人力，而且诊断的结果往往受医生经验和主观判断影响，甚至受到周围客观环境的影响（如：光照、温度等）[1]，通过10位中医专家对两百多例患者进行舌象诊断，发现仅仅有9例相同，为了减少主观判断和客观环境的影响，利用现代计算机技术结合传统中医的理论和中医专家的经验，使中医的舌诊客观化、数字化成为了目前十分热门研究方向。

利用机器学习实现对舌苔精确快捷的检测，结合中医经验智能化的给出体质信息判别不仅是对计算机技术应用领域的一大拓展，也对传统中医的传承、推广、创新和现代化具有重大意义。利用现代计算机技术使传统中医的应用更加广泛，诊断更加方便快捷，能一定程度上缓解医师人力不足的问题，减少医疗成本，辅助医师提高诊断的准确性，乃至推动中医走向全世界。

## 舌苔检测研究现状

目前中医通过舌诊辨识人体体质信息，主要依据中医千年发展累积的经验，使得舌象诊断需要十分丰富的专业知识才能准确无误，如何利用计算机技术实现这一复杂的诊断也是目前研究中的重点和难点。目前利用计算机技术实现的智能化舌诊项目虽然辨识较为精准，但是往往需要在特定的环境或者需要十分专业的设备进行辨识分析，使用专业的设备和环境使得检测门槛较高，检测的成本也随之提升，对舌诊的推广有一定的影响。

## 课题任务内容

本课题是基于机器学习的舌苔检测，目的是利用机器学习的方法实现对舌苔进行检测，对舌象信息进行识别，最后再结合传统中医的理论辨识出人体体质信息。

具体实现内容如下：

（1）因为 所使用的舌象数据集图片较少，在开始训练数据集前先使用图像增强和生成对抗网络扩充数据集。将数据集图片进行翻转、平移、加盐、亮度调整、模糊处理等，能将数据集扩充数十倍，以满足数据量的需求。除此之外还使用了生成对抗网络生成新的舌象图片扩充数据集使数据集多样化。

（2）本次课题主要使用的是机器学习中的深度学习方法实现，使用的是深度学习中的efficientNet网络实现对舌象图片的图像分类，同时还使用了其他网络进行了对比实验。训练网络efficientNet模型时使用的在其他图像分类数据集上训练得到的训练参数efficientnetb0作为网络模型的预训练参数，可以较快的实现收敛且能得到良好的效果。分类网络实现了舌象中的苔色、舌色、齿痕、裂纹信息分别进行分类，分类准确率能达到90%以上。

（3）上传舌象图片后，运行分类网络输出的舌象判断结果并展示，使用舌象结果对照传统中医的舌象诊断经验，对人体体质信息进行判断并展示给用户。

本次课题利用现代计算机的技术对舌苔进行检测，并辨识人体体质信息，实现了普通用户借助个人计算机就能对舌苔进行分析以及体质辨识，对计算机技术的发展、创新以及推广有重大意义。

## 本章小结

本章内容介绍了本课题的研究背景，同时简单阐述目前使用机器学习进行舌苔检测的研究现状、本课题的任务要点和本篇论文的大体结构。

# 机器学习相关理论

本章先介绍了机器学习相关的现状，并简单梳理了机器学习的发展状况。之后介绍本课题所使用的机器学习方法：深度学习的理论概念，并提及神经网络模型的训练方式。

## 机器学习的现状与发展

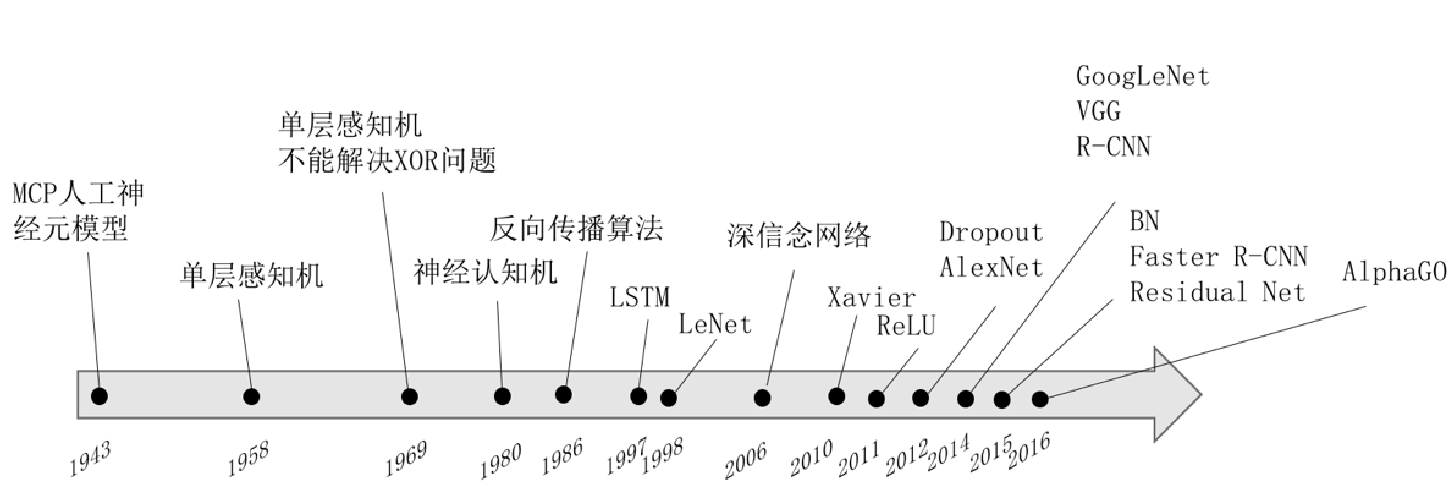


图2.1 机器学习的发展历史

机器学习的历史可以追溯到第一个神经网络数学模型——MCP人工神经元模型的提出开始[2]。早在20世纪的40年代人工智能开始起步发展，机器学习的理念就已经被提出，1952年Samuel设计了基于人工智能的跳棋程序，这是最早成功第一个基于人工智能的程序。在1958年，Rosenblatt提出了单层感知机（Perceptron），感知机的梯度下降算法被认为是深度神经网络的基础。到了1980年，Kunihiko Fukushima在猫视觉实验研究启发下，设计了一个网络——Neocognitrons，由多个卷积层和池化层构成，成为了现代卷积神经网络的最初的范例。2010年后，经历了推理期和知识期的机器学习开始蓬勃发展，机器学习中的深度学习方法兴起，各种深度神经网络模型相继出现，到近几年出现的efficient Net，使得深度学习如鱼得水。目前在机器学习的各种方法中，使用的最多的，同时也是火的便是深度学习。

## 深度神经网络的结构和概念

深度学习是目前机器学习中最热门的方法之一，而深度学习离不开深度神经网络，随着目前深度学习和深度神经网络的理论知识越来越丰富，针对不同的任务需求情况，使用不用同的网络结构。本小节主神经网络模型、卷积神经网络的相关理论和常用的优化算法。

### 神经网络模型

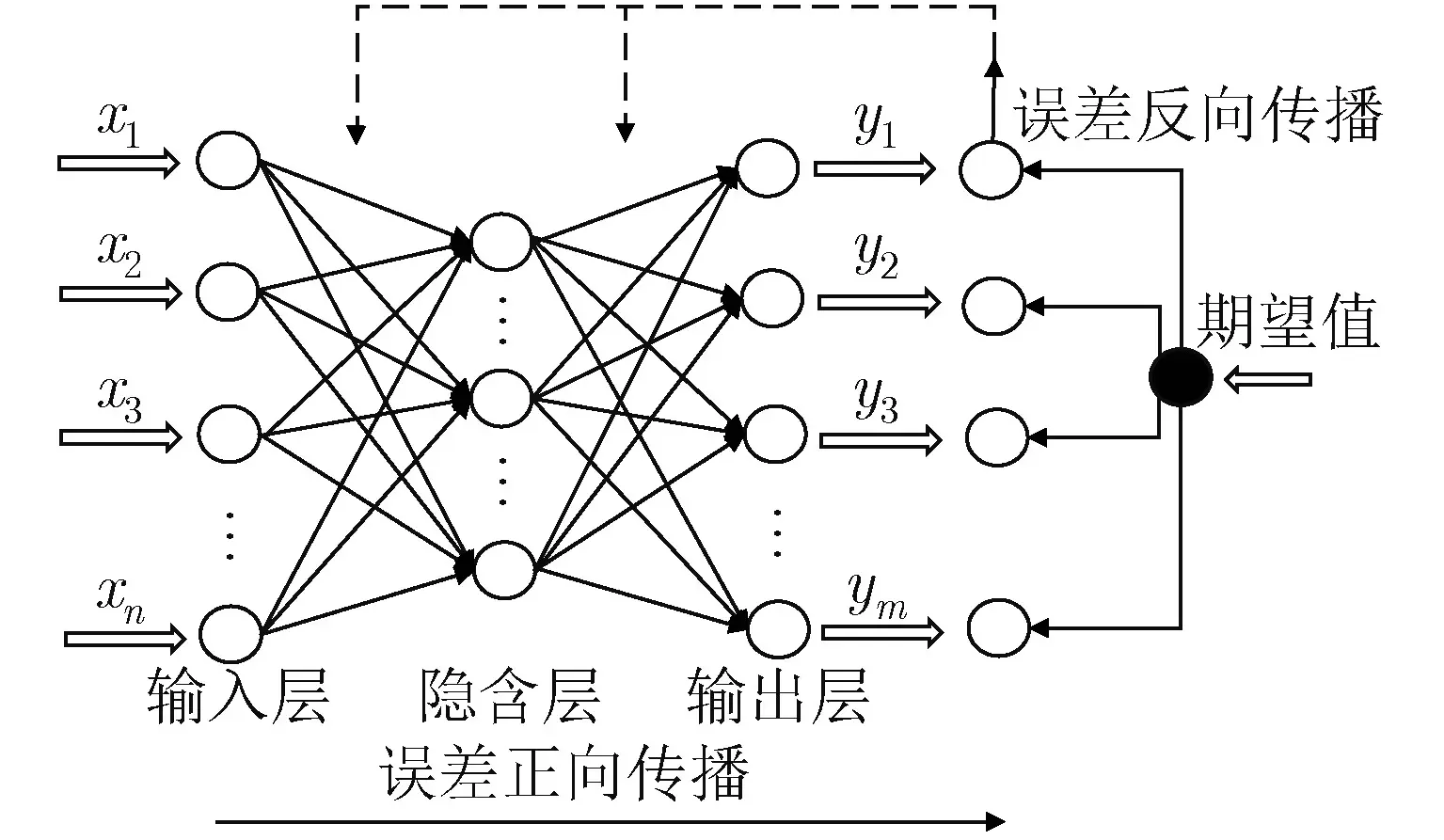


图2.2 神经网络结构图

神经网络是一种模拟人脑神经元的网络。它是由几个“神经元”组成的拓扑结构。它主要包括输入层、隐藏层和输出层。

如图2.2所示的神经网络结构表明，神经网络的原始输入数据首先进入网络的输入层，然后数据传播到隐藏层。隐层中的每个神经元可以对不同神经网络的输出层神经元设置不同的权重。该权重影响隐层神经元对输入信息的敏感性，而输出层根据隐层的组合权重输出结果。

当一份原始输入数据传入到输入层，经过正面传播，经过隐含层后从输出层输出。在网络首次使用时，隐含层中的初始权重可能并不合适原始输入数据，因此输出层得出的值可能和期望值不一。要想使输出值更加接近我们的期望，就得对网络模型进行训练，神经网络的负面传播即是网络模型学习调整的关键，负面传播过程为：首先向神经网络的输入原始数据，通过中间隐含层处理数据后，通过输出层获得输出值，计算出输出值和期望误差值，并将此误差从神经网络后端作为前端方向传播，更新了网络隐藏层的权重系数和阈值，将输出误差不断缩小，经过反复训练网络，不断对隐含层中的权重系数进行调整，最终得到和期望值相似输出值。

### 卷积神经网络

卷积神经网络是目前使用得最多、表现最好的一种深度神经网络，在其他领域上也崭露头角。卷积神经网络的隐含层由多次卷积、激活和池化完成其功能，在部分分类任务中，还会在隐含层的末端加入全连接层以完成分类任务的需求。

1. 卷积层

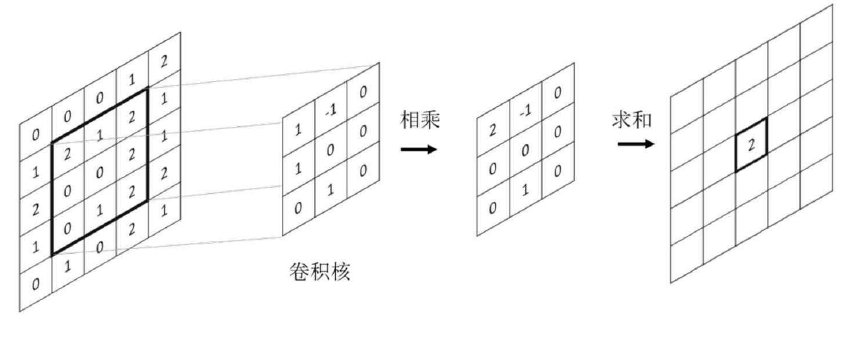


图2.3 卷积运算

卷积层是卷积神经网络的中心元件之一，它包括多组卷积元件，其中卷积层的参数由一组可改变的储核组成，而图像卷积是在乘积产生后，利用卷积像素点和输入图像的值进行相加，如图2.3 所示。因为卷积运算的特性，使得原始图像中的具有相关性的局部像素，能被更加有效的提取出来，使得图像的特征更加明显。通过卷积的运算使得图像大小发生改变，减少了神经网络模型的参数，提升了网络模型的效率和准确率。

1. 激活层

即激活函数层。激活层通常跟在卷积层之后，激活层通常不带有可以进行学习的参数。在深层神经网络中，激活函数一般采用又简单有快的线性函数，这更有利于学习图像的各种特征值。目前常见的激活函数主要是Sihmoid和ReLU激活函数，如图2.4所示：

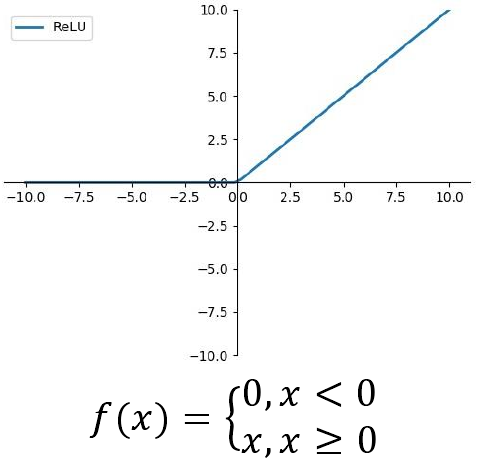
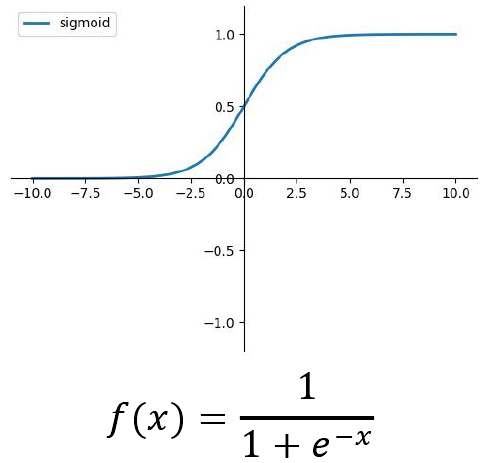


图2.4 Sihmoid和ReLU激活函数示意图

目前使用的最广的是ReLU激活函数，ReLU让小于0的神经元输出为0,为神经网络提供了稀疏性,能够缩短其学习周期[3]。

当卷积操作完成后，特征图像的维度发生变化，特征数量也会大幅增加，这一变化肯定会给之后网络的计算带来一定的困难。基于这些原因，所以需要将池化操作放至两个卷积层之间。池化操作如图2.5，在每一个像素区域内计算平均值（平均池化）或者最大值（最大池化）。

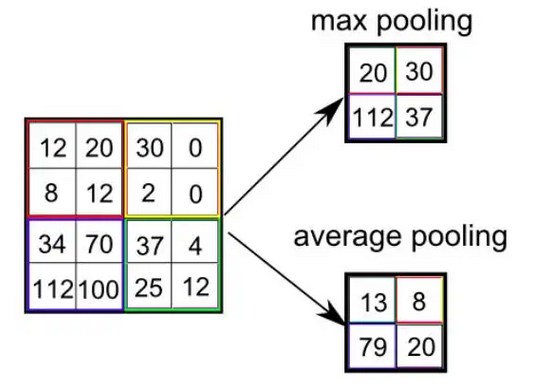


图2.5 池化运算

1. 全连接层

全连接层将前一层中的每一个神经元都与当前层中的所有神经元进行连接，其连接的方式和神经网络模型的链接方式一致的[4]。一般将全连接层放到卷积神经网络的最后，可以用来生成特征向量。

## 神经网络的训练

使用神经网络可以对数据进行训练，通过不断的迭代训练，网络的权重参数系数也会由输入数据的特征而调整，从而使得网络不断的拟合。在进行训练时将把数据集分为训练集和测试集。在利用神经网络对训练模块进行反复训练的基础上，对权重系数进行连续调整，得到最佳参数并将其保存，然后对测试集进行验证，确定网络模型的准确性，并对其质量进行评价.

## 本章小结

本章先介绍了机器学习相关的现状，并简单梳理了机器学习的发展状况。之后介绍本课题所使用的机器学习方法：深度学习的理论概念，并提及神经网络模型的训练方式，是本文的理论概括部分。

# 舌苔检测需求分析

## 可行性分析

### 技术可行性

如今，机器学习经过了半个世纪的发展，利用机器学习解决现实中各行各业的需求已经变得十分的成熟和可靠，而传统中医的数千年发展积累下来的舌诊经验在观察舌象辨识人体体质的方面也已经十分准确。基于成熟的中医技术和新兴饿计算机技术，利用机器学习完成舌苔检测也变得较为容易的实现，所以在技术上是可行的。

### 经济可行性

市面上利用计算机技术实现的智能化舌诊项目虽然辨识较为精准，但是往往需要在特定的环境或者需要十分专业的设备进行辨识分析，使用专业的设备和环境使得检测门槛较高，专业性很强，检测的成本也随之提升。本课题的是利用机器学习开源库开发的，没有硬件资源的花费开销的舌苔检测项目，在普通的计算机上就可以运行，所以在经济上是可行的。

### 文化可行性

中医药学包含着中华民族几千年的健康养生理念及其实践经验，是中华文明的一个瑰宝，凝聚着中国人民和中华民族的博大智慧，向世界推广中医药学也是传播中华文化的一环[5]。利用机器学习实现对舌苔精确快捷的检测，结合中医经验智能化的给出体质信息判别。这不仅仅是对计算机技术应用领域的一大拓展，也对传统中医的传承、推广、创新和现代化具有重大意义。利用现代计算机技术使传统中医的应用更加广泛，推动中医文化走向全世界。所以在文化上是可行的。

### 社会可行性

本次课题利用现代计算机的技术对舌苔进行检测，并辨识人体体质信息，实现了普通无需掌握专业的中医舌诊知识也能对自己舌苔进行分析，了解自身体质信息，有利于普通人更多的享受到中医的作用。在配合医生进行诊断时，也能减少医生的诊病人工成本，对社会是有益的，具备社会可行性。

## 功能性需求

### 数据集构建

数据集是深度学习的灵魂部分，在深度学习中，数据集的质量越高、数量越多，训练出来的模型就能拥有更高的准确率和泛化的能力，本课题需要对高质量的舌象进行精确的分类，制作成舌象数据集。由于舌象的一些特征类型是不冲突的，一张舌象中包含多种特征类型（如舌色、苔色、齿痕和裂纹），所以数据集的制作应当按照不同舌象特征分别制作数据集，例如按舌色特征制作数据集中可以分成淡红舌，淡白舌和深红舌三类；而按照苔色特征制作数据集中可以分成白苔，黄苔和无苔三类。原始的舌象数据图片不足，需要利用图像增强和生成对抗网络进行扩充。

### 舌苔检测

舌苔检测模块能利用舌象数据集训练网络模型，并能够用训练完成的网络对输入图像的舌象特征做图像分类，准确的识别出舌象中所含有的舌象特征信息，并将预测的结果输出。由于舌象的一些特征类型不冲突，一张舌象中包含多种特征类型，所以需要对几类舌象特征数据集进行分别训练，分别保存不同的训练权重参数，用于检测不同的舌象特征。

### 体质辨识

当普通用户上传舌象照片时，通过对舌苔检测模块给出的舌象特征信息结合汇总，对照舌象特征结构和体质对应图，辨识该舌象所反映出的人体体质信息，并能够用用户交互页面将其展示出来，界面能够简单易操作，尽可能满足更大的用户群体的需要。

## 非功能性需求

本课题非功能性需求主要是对计算机性能的需求，性能越高模型训练时的速度越快。硬件要求：计算机要有1.0 GHz以上CPU，512M内存，1G硬盘，独立显卡；软件方面：对于系统软件方面，操作系统最优为Windows 10 操作系统。开发语言用的是Python。

## 3.4 本章小结

本章主要针对舌苔检测的技术可行性、经济可行性、文化可行性、社会可行性、功能性需求和非功能性需求进行分析。

# 舌象数据集构建与扩充

数据集是机器学习的灵魂，正确的对数据进行标注，机器才能更好的发现图像中的特征，提升对图像样本的识别精确度。由于本课题所获取的数据集较少，所以本次课题还使用了其他方式来扩充数据集。

## 舌象图片数据的标注分类

本课题所使用的舌象图片由我的毕设导师所提供，共有高清舌象图片509张，图片像素2136 \* 1424，由于所使用的舌象图片均为未进行标注的图片，所以在分类时，使用人工检测的方式，针对舌象图片中所含有的特征对舌象图片进行分类处理。

在传统中医的舌诊中，舌象不是单纯的多分类问题，而是要结合多个特征综合判断后，才能准确的辨识体质。因此，本课题在分类时按照舌象诊断时的各个舌象特征：齿痕、裂纹、舌色和苔色分别构建4个舌象数据集：齿痕舌象数据集、裂纹舌象数据集、舌色舌象数据集、苔色舌象数据集。并对所获得的509张舌象图片采用人工的方式进行分类。

数据集具体分类如下：齿痕舌象数据集分为舌象有齿痕（64张）和舌象无齿痕（445张）；裂纹舌象数据集分为舌象有裂纹（258张）和舌象无裂纹（251张）；舌色舌象数据集分为淡红舌象（218张）、红绛舌象（118张）和淡白舌象（173张）；苔色舌象数据集分为白苔舌象（335张）、无苔舌象（52张）和黄苔舌象（122张）。每个数据集均使用的是导师所提供的509张舌象图片。

因为所获得的舌象图片数据较少，而且各分类的图片数量并不平均，训练一个准确率较高的模型需要大量的数据，所以必须扩充本课题所使用的数据集的大小规模。

## 使用图像增强扩充数据集

数据集是深度学习的灵魂部分，在深度学习中，数据集的质量越高、数量越多，训练出来的模型就能拥有更高的准确率和泛化的能力，本课题由于数据集的数量较小，扩充数据集十分有必要。通过网上搜索舌苔检测的数据信息后发现，研究使用的舌象数据集几乎没有在网上公开的，无法通过网上获取到大批量的舌象图片数据，故通过增加舌象图片数据来扩充本课题的舌象数据集是不可行的，所以本课题采用图像增强的方式来扩充数据集。

本课题中采用事先执行图像增强，通过实质的扩充数据集的大小，即线下增强，达到数据集扩充的目的，该方法适用于数据集过小的情况。本文主要通过使用opencv对原始舌象图片进行翻转、平移、对比度调整、噪声、模糊等等操作得到新的舌象图片数据，使用图像增强不但能扩充舌象数据集，还能有利于模型的训练。

（1）水平翻转：对原始图片进行水平镜像翻转，翻转后的图像能够和原图保持同一纬度，对于图像分类的数据集，水平翻转使用的较多。

（2）像素平移：将原始图片向左右平移，对图片像素进行位移操作，有利于深度学习模型在训练时不仅仅是针对于中心区域的特征学习，使计算机在执行图像分类时会将关注点转移至任意区域。这种操作有利于提升模型的稳定性。

（3）对比度调整：通常情况下图片采集时的亮度是几乎一致的，这时可能会造成采集的图像亮度覆盖范围不足，如此训练出的模型在面对不同亮度的检测样本时可能会导致模型准确率不稳定，所以为了提升模型的稳定性，将原始图片增加或减少对比度，提升数据集的亮度覆盖范围。

（4）噪声：本文采用的是0.05%的椒盐噪声，在原始图像上随机加入一些像素点，用来扰乱原图像的特征信息，提升网络模型的泛化力。

（5）模糊：对原图进行褶积操作，这种方法减少了图像中像素值的差异，从而提升了图像的模糊度，使得图片的特征信息较原图难以辨认，有利于训练网络的能力。

图4.1数据增强效果图为舌象数据集中的某一张舌象图片经过上述的图像增强方法后输出的效果图，图像增强后的图片虽然在人的眼中仍能清楚的看出舌象特征，且能发现其和原图十分相似能够，一眼认出是同一张图片，但在计算机视觉的眼中，图像的像素点的值以及发生了明显的改变，在机器眼里就是一张全新的图片，这便是利用图像增强扩充数据集的原理。

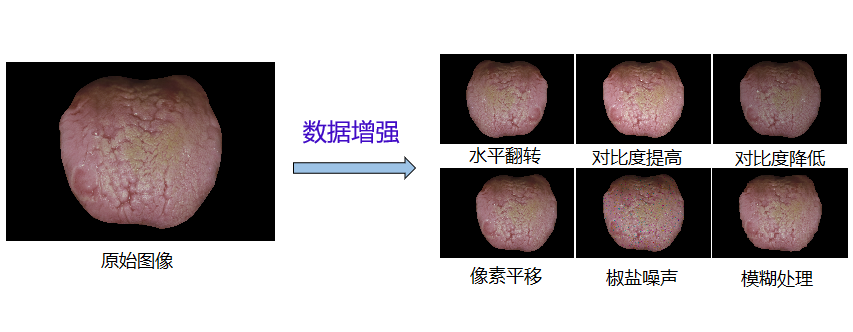


图4.1 数据增强效果图

## 生成对抗网络

图像增强的方法扩充数据集后扩充的图片都是基于原始图片处理得来的，与原始图片由一定的相同之处，为了使舌象图片的多样化，进一步的丰富数据集，本课题除了使用上一节所述的图像增强方法扩充数据集外，还使用了生成对抗网络中的DCGAN网络来生成新的舌象图片。

### 生成对抗网络相关概念

生成对抗网络(GAN)是在二零一四年由Ian Goodfellow教授开创性的提出的一个簇新的机器学习模型结果[6]。生成对抗网络是一个无监督模式机器学习的网络模式，极大的降低了深度神经网络中对数据集标记的依赖性，目前成了人工智能领域中的一个相当火热的科研方向。

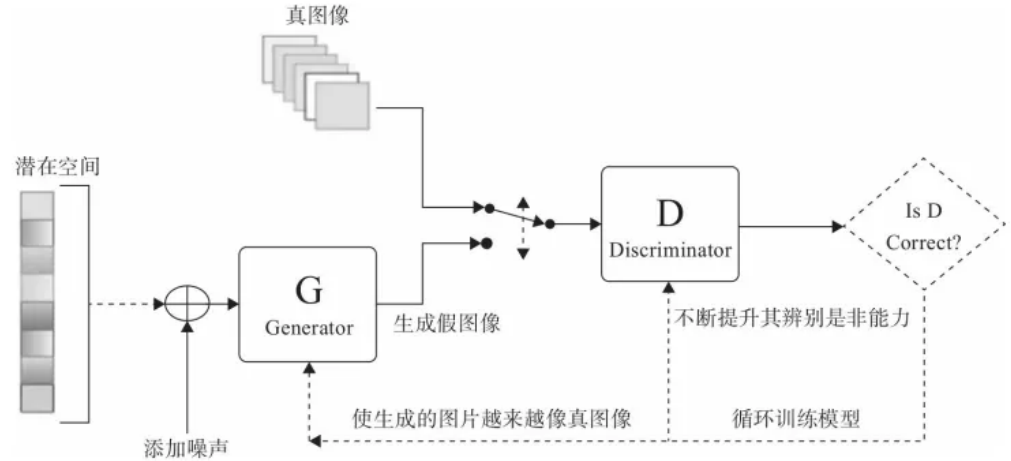


图4.2 生成对抗网络结构

生成对抗网络结构如图4.2所示，主要包含着两个部分，分别是生成器（Generator）和判别器（Discriminator），通过互相对抗的方式来对自身进行不断地优化调整参数，这是和其他一般的神经网络所不同的。生成网络的主要作用生成一幅尽可能与原始真实图片相近的随机图像样本；判别网络则是尽可能的分辨出数据的真伪。

在网络的训练过程中，首先把实际图像上传到判断网络，并告诉判断网络中实际图片是什么样的。然后生成网生成了一个随机噪声样本。传给判断网络作出真伪性的判别,当判断器的判别准确时，再通过实际的图片不断调整生成网络的参数。当判别器判断出错时，可以证明所产生的图像就已能迷惑判定网络了，这时便必须更改判定网络的参数，使之对图像能进行较为精确的判断。最具体的优化参数方法主要是通过使用损失函数，公式如下:

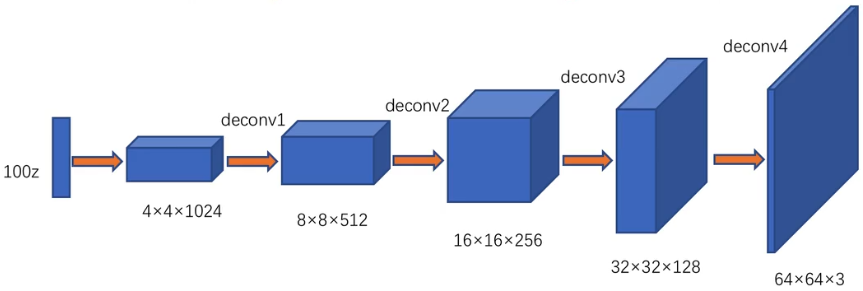
通过不断反复的训练，这时的生成器便可以通过机器创造用来生成大量的新图片用于扩充数据集。

### DCGAN生成舌象图片

本课题中使用的生成对抗网络是DCGAN，即通过深度卷积生成对抗网络，是生成对抗网络的一个改进的模式。和之前的GAN网络一样，DCGAN把卷积操作的思路带入到生成对抗网络当中做无监督的学习训练，通过使用卷积计算网络对图像特征的强大获取能力，使得网络的学生对图像特征的学习能力有了更进一步的提高。

本课题所使用的DCGAN网络同样包含生成网络和判别网络，但不同的是生成网络中使用反褶积方法扩大噪声数据维度，判别网络中的空间池化被褶积步长所取代，网络整体通过褶积层连接输入和输出层。

生成网络使用了四层反卷积算法的结构,输入初始噪声的通道维度为100，在经过四层反卷积算法上的采样扩大后可以输出64\*64,3通道的图像。除输出层采用tanh激活函数外，其他层次都使用了ReLU激活函数[7]。网络的具体构造如图4.3和表4.1所给出。



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **生成网络卷积层** | **输入分辨率** | **输出分辨率** | **输入通道数** | **输出通道数** |
| deconv1 | 4\*4 | 8\*8 | 1024 | 512 |
| deconv2 | 8\*8 | 16\*16 | 512 | 256 |
| deconv3 | 16\*16 | 32\*32 | 256 | 128 |
| deconv4 | 32\*32 | 64\*64 | 128 | 3 |

图4.3 DCGAN生成网络结构

表4.1 生成网络输入输出大小

判别网络使用四个strided=2的卷积，对64\*64,3通道的输入图像进行4次卷积操作最终输出一个一维度向量。在判别网络中使用ReLU激活函数。网络具体结构如图4.4和表4.2所示。

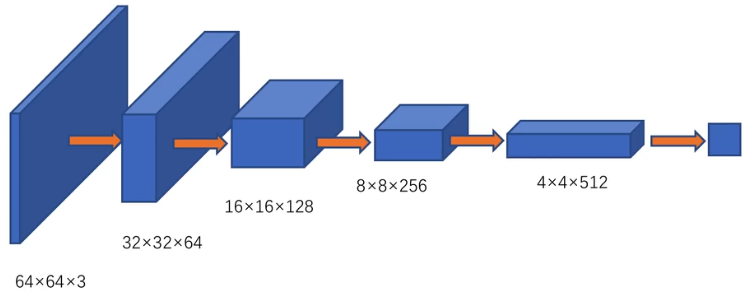


图4.4 判别网络结构

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **判别网络卷积层** | **输入分辨率** | **输出分辨率** | **输入通道数** | **输出通道数** |
| conv1 | 64\*64 | 32\*32 | 3 | 64 |
| conv2 | 32\*32 | 16\*16 | 64 | 128 |
| conv3 | 16\*16 | 8\*8 | 128 | 256 |
| deconv4 | 8\*8 | 4\*4 | 256 | 512 |

表4.1 判别网络输入输出大小

训练使用之前使用数据增强方法扩充后的舌象数据作为训练集，训练采用的优化算法为Adam优化算法。Adam优化算法在本网络中的效果较好，收敛速度较其他算法快，对内存的需求也比较少。训练模型时的初始学习率使用的是2e-4。

|  |  |
| --- | --- |
| **参数** | **值** |
| img\_size（图片大小） | 64\*64 |
| batch\_size（批大小） | 128 |
| optim（优化器） | Adam |
| Lr（学习率） |  |
| Batas | 0.5，0.99 |
| epoch（迭代次数） | 300 |
| noise\_dim（噪声通道维度） | 100 |

表4.2 网络参数

如图4.5，第一轮训练完成后生成网络生成的图片是一团随机的无意义的噪声，无法看出图像的任何特征。在网络模型训练到第五轮后，可以发现虽然图片还是一团模糊的噪声，但是已经能隐约看出是一个舌头状的物体。第十轮训练后，已经能够看出较为明显的舌头形状，且颜色与淡红舌十分相似。到了第五十轮训练后，舌头的形状已经十分明显，但是舌体的边缘仍有棱角等瑕疵，效果仍无法满足我们的需求。

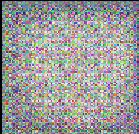
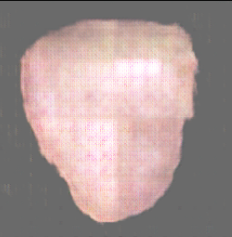
   

图4.5 第1、5、10、50轮训练效果示例图

在经过数百轮的训练后得到的舌象图片效果已经达到较为不错的效果，舌象的形状较为丰富，能够看到舌象的舌色大致颜色。如图4.6是训练300轮后生成网络随机输出的64张生成样例图片，可以看出下图中有舌色有淡红舌、淡白舌和红绛舌，个别生成舌象图片甚至能隐隐约约看到齿痕的形状。



图4.6 训练300轮后的生成网络生成样例图

在训练效果到达后，使用训练完成后的生成网络模型权重，单独通过导入权重的生成网络生成大量舌象图片，通过人工筛选，将生成效果不理想的舌象图片删去，由于生成的舌象图片分辨率较低，一些舌象特征无法展现，只能较为清楚的分辨舌色，所以在本课题中通过生成网络生成的舌象图片仅仅能用于扩充舌色数据集。

## 本章小结

本章为之后的模型训练准备了一份较为完善的舌象数据集，利用图像增强方法和生成对抗网络将原本仅有509张的舌象图片数据扩充到上万张舌象图片，原本标注分类不平衡，数据量过少的问题得到解决。对于本文提到的将舌象数据集按特征种类划分成四个不同的数据集是因为在舌象诊断中，舌象的一些特征类型是不冲突的，一张舌象中包含多个特征类型，所以数据集制作时和其他图像分类项目的单一数据集有所不同，采用了同网络多数据集训练多模型的办法，能明显提升各个特征分类的正确率。

# 舌苔检测网络设计与实现

## 网络模型介绍

EffcientNet是一个由谷歌团队研究的卷积神经网络，适用于图像分类领域。在2019年发布时，在ImageNet top-1上达到了当时的最高精确度84.3%，在参数数量方面也比当时其他的模型要少很多。是一个比较火的卷积神经网络模型，性能也十分优秀。

## 网络模型分析

网络通过混合参数对传统的卷积网络进行优化，同时改变网络的宽度、深度和输入图像大小提升网络性能，相比于传统的卷积神经网络，EffcientNet在效率上有一定优势。扩展网络深度今日也被广泛应用于神经网络的训练中，深度网络可以提取更复杂的特征值，以帮助提高对目标数据集的学习能力。然而，深层网络往往会出现消失梯度的问题。增加网络的宽能够使特征图获得更多通道数，同样得到广泛应用。通过在每个通道上进行卷积运算，这使得模型更具表现力。宽度大能让网络更好的学习各种特性，这使得训练网络变得容易。然而，也正是因为网络宽度足够导致网络深度是不够的。即使提取的特征丰富，但层次却不高，可以通过网络分辨率，丰富网络的感知区域来学习更加多的信息，网络的性能也可以得到提高。

综上所述，对网络的宽度、深度和输入图像分辨率的缩放需要找到一个合适的度，对其中的一方面进行盲目的缩放并不能达到最优的效果，EffcientNet则是非常成功的协调了上述三个方面的缩放，优化了网络性能。

### 网络主要结构

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stage** | **Operator** | **Resolution** | **#Channels** | **#Layers** |
| 1 | Conv 3\*3 | 224\*224 | 32 | 1 |
| 2 | MBConv1，k3\*3 | 112\*112 | 16 | 1 |
| 3 | MBConv6，k3\*3 | 112\*112 | 24 | 2 |
| 4 | MBConv6，k5\*5 | 56\*56 | 40 | 2 |
| 5 | MBConv6，k3\*3 | 28\*28 | 80 | 3 |
| 6 | MBConv6，k5\*5 | 14\*14 | 112 | 3 |
| 7 | MBConv6，k5\*5 | 14\*14 | 192 | 4 |
| 8 | MBConv6，k3\*3 | 7\*7 | 320 | 1 |
| 9 | Conv 1\*1 & Pooling & FC | 7\*7 | 1280 | 1 |

表5.1 EffcientNet-B0结构参数

本次实验所用到的是EffcientNet中的EffcientNet-B0模型，该模型由一个前处理层、十六个MBConv模块、池化层和全连接层组成，结构和主要参数如上表5.1。

### 网络功能模块

前处理层模块含有一个卷积层，主要的作用时使输入图像大小变成原先的一半，同时可以扩张图像通道。EffcientNet网络中的核心部分是中间的MBConv模块，主要起的是特征提取的作用，能够大大减少了网络参数和计算量，模块中含有的通道注意机制也可以增强各个特征映射的通道信息提取性能。MBConv模块主结构如图5.1，数据输入后先经过一个的升维卷积层，通过两个激活函数，再到达核大小为3或5的DW卷积，再通过两个激活函数，到达注意力机制模块SE，通过SE后到达一个的降维卷积层，通过一个激活函数后输出。

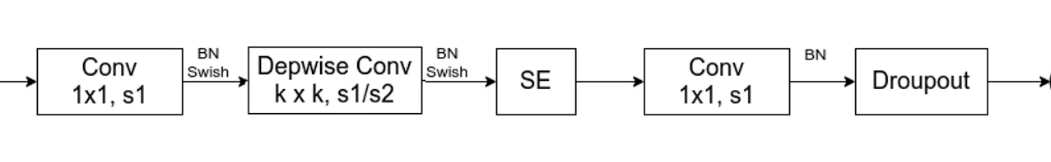


图5.1 MBConv模块主结构图

在MBConv模块引入的注意力机制模块SE结构如图5.2所示，当输入特征矩阵输入到注意力机制模块时，对输入特征矩阵中的每一个通道进行一次平均池化操作（AvgPooling）之后将分别通过两个全连接层FC1和FC2，在FC1中节点个数为输入MBConv模块的输入特征矩阵通道的1/4，而FC2中的节点个数为输入到注意力机制模块的特征矩阵的通道数。

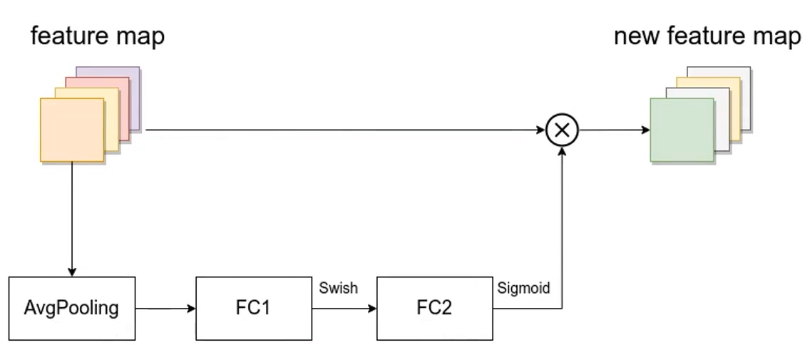


图5.2 注意力机制模块结构图

## 网络模型搭建及功能的实现

本课题是基于机器学习的舌苔检测,其实现主要包括以下几大模块：网络模型模块、数据模块、训练模块、检测模块、体质辨识模块。开发环境主要采用了PyCharm+Python3.6，工作环境主要为Windows10操作系统。舌苔检测程序设计的简要结构如图5.3所示：

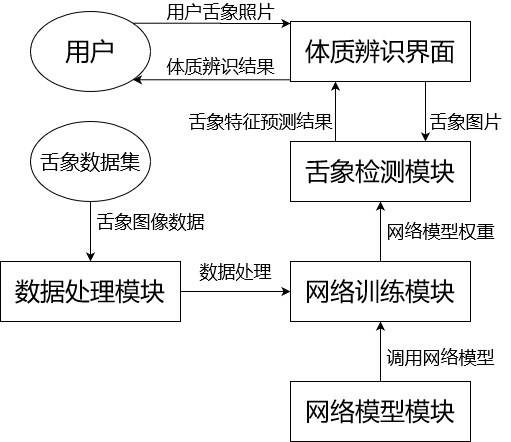


图5.3 舌苔检测程序简要结构图

接下来将简单说明各模块的基本实现以及功能。

### 网络模型模块

本课题使用的是EffcientNet网络中的EffcientNet-B0网络模型，使用Python中的Pytorch库中的torch.nn方法实现神经网络的搭建，网络模型主要分为四个小模块：网络主模块、卷积+BN模块、注意力机制模块、MBConv模块。三个功能模块分别实现了上一节介绍的EffcientNet网络模型中的主要功能，主模块整合各个功能部分，实现EffcientNet-B0卷积神经网络的功能。

### 数据模块

数据处理的主要功能是读取并处理舌象数据集，利用python中的os库中的path方法，遍历数据集目录，将舌象数据集中的80%划分为训练集输入到网络中进行训练，20%划分为验证集验证网络训练后的准确率，并在数据集图像训练前对输入图像进行预处理：使用torchvision中transforms的RandomResizedCrop方法，随机对数据图像裁剪为不同的大小和纵横比，然后将裁剪后的图像缩放到EffcientNet-B0网络输入图像的指定大小。使用transforms中Normalize方法对输入舌象数据集的数据集进行归一化操作。

### 训练模块

训练模块的主要功能是使用数据处理模块读取舌象数据集，再调用网络模型模块中进行训练。将每一个batch的数据作为一个step，在训练模块的每一个step训练流程如下图5.4。



图5.4 训练模块step流程图

训练模块中预留有多个参数设置：学习率（lr）、批大小（batch\_size）、迭代次数（epochs）、是否导入预训练权重、normalize参数等等，从而能在对比实验中不断调整参数，优化完拿过来模型的效果。

每一轮训练后都使用验证集对模型进行验证，并将训练和验证中得到的训练loss曲线以及验证准确率利用可视化工具Tensorboard在网页中展示出来，更直观的观察模型的训练情况，示例如图5.5，利用实验中对比模型优劣采用更加合适的训练参数设置。在完成训练后的将参数权重进行保存。

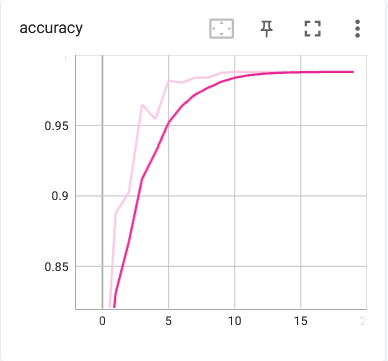
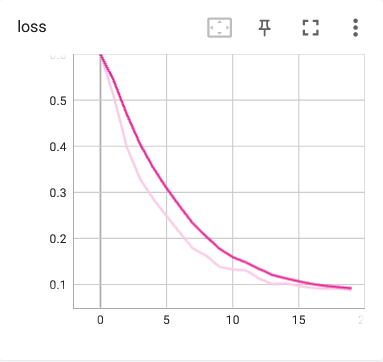


图5.5 训练loss和准确率曲线示例

### 检测模块

检测模块的主要功能是将待检测舌象图片输入后，将图像大小调整至224\*224，利用训练模块当中保存好的训练权重参数，使用torch中的no\_grad()方法对舌象图片进行分类并能预测其结果，给出各项分类预测的概率，如图5.6，并将预测概率最高的分类输出给体质辨识模块。

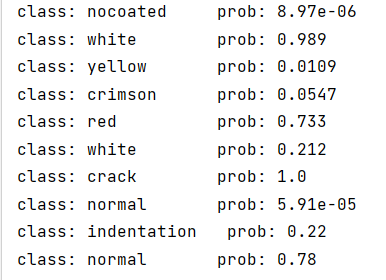


图5.6 多分类预测概率

### 体质辨识界面模块

体质辨识界面模块的主要功能是利用用户交互界面，获取用户上传的个人舌象图片，并将其送到检测模块中检测，得到舌象特征信息，对照舌象特征结构和体质对应图，辨识该舌象所反映出的人体体质信息，并将其展示出来，体质辨识界面如图5.7所示，该模块采用pyqt5库实现了用户的交互功能。简单易操作，方便普通用户能够更方便的对舌象进行检测。

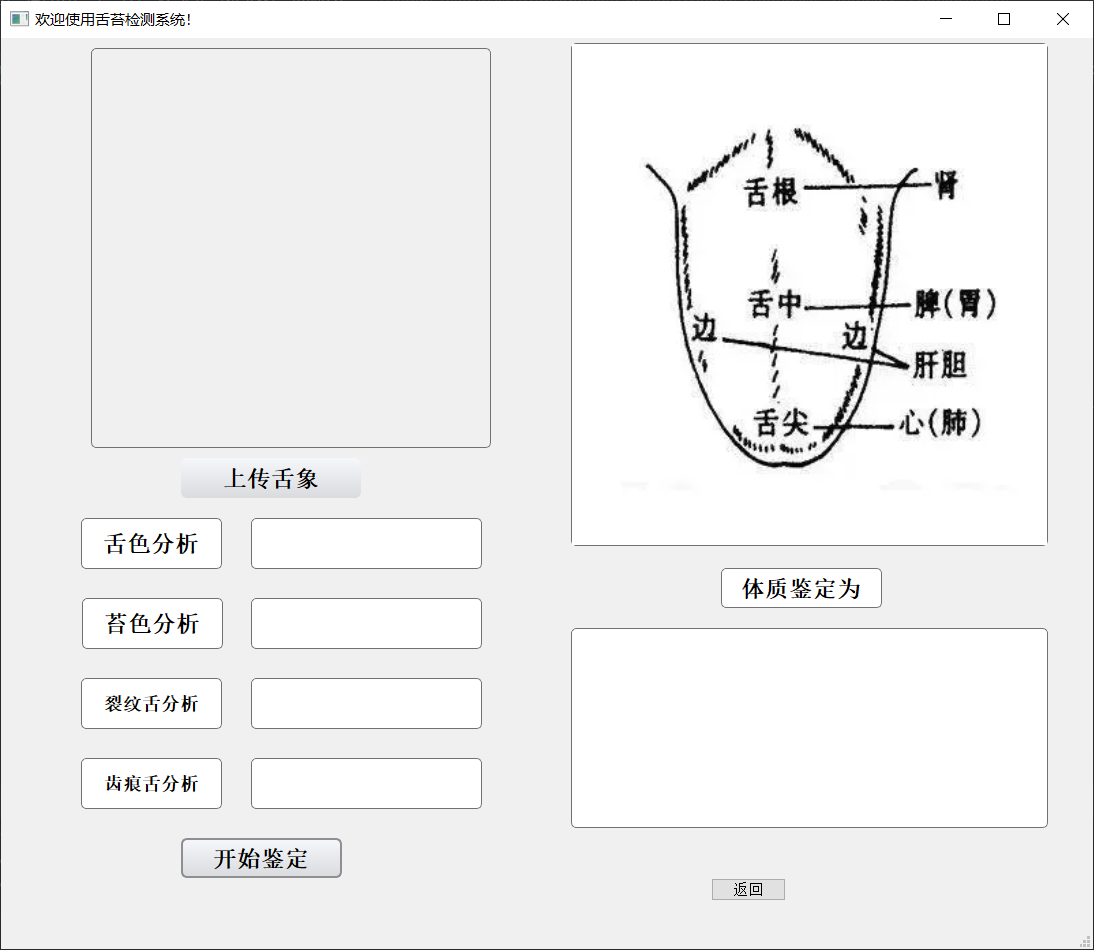


图5.7体质辨识界面

## 本章小结

本章主要介绍本课题中舌苔特征分类检测部分的网络搭建。简要介绍本课题所使用的EffcientNet神经网络，使用该网络对舌象数据训练网络模型，实现了整体舌苔检测网络模型的程序设计部分，为接下来的对比实验提供了可靠的网络模型。

# 舌苔检测实验分析

## 实验数据集

舌苔检测实验中所使用的数据集为第四章中的利用图像增强和生成对抗网络进行扩充后的数据集，扩充后的数据集图片数据大约有12000张，其中舌象数据集中的80%划分为训练集输入到网络中进行训练，20%划分为验证集验证网络训练后的准确率。数据集存放目录按照分类类别名称存放该分类的图像。训练时将依照舌像特征种类制作的四个舌象数据集：齿痕舌象数据集、裂纹舌象数据集、舌色舌象数据集、苔色舌象数据集分别对网络进行训练，获得对应的四个参数权重，这四个参数权重将分别用于对应舌象特征的检测。

## 数据图像预处理

在开始实验之前，在对数据集中采集的图像进行预处理之前，图像预处理的主要目的是去除图像中的外来信息，提高信息的真实性，提升可验证性，最大限度地简化图像数据，从而提高了目标图像提取、匹配和识别的可靠性，从而加快了实验中模型的收敛过程，同时也增强了网络模型的稳定性，提高了精度。

### 图像增强

在第四章中，对原始舌象数据图片进行图像增强，通过实质的扩充数据集的大小，即线下增强，达到数据集扩充的目的。通过使用opencv对原始舌象图片进行翻转、平移、对比度调整、噪声、模糊等等操作得到新的舌象图片数据，使用图像增强有利于模型的训练和对训练分布预测，解决样本不足的缺点。

### 图像大小处理

舌苔检测实验中所使用的舌象数据集中的舌象图像像素大小为图片像素2136 \* 1424，而EffcientNet-B0网络对输入图像的像素大小要求为224 \* 224，数据集中的图像像素并不符合EffcientNet-B0网络对输入图像的要求，故需要对图像大小进行变换。实验中图像大小处理使用torchvision中transforms的RandomResizedCrop方法，随机对数据图像裁剪为不同的大小和纵横比，然后将裁剪后的图像缩放到EffcientNet-B0网络输入图像的指定大小，即首先对数据图像随机采集，然后将剪裁图像缩放到224 \* 224。该操作的作用在于即使是物体的某一部分，也能让计算机认为是该类物体，能够提升网络对局部特征的识别能力。

如图6.1为齿痕舌象数据集中的某一张齿痕舌象图片使用该方法对输入图像进行转换后得出的示例图，可以看出经过随机裁剪后，舌象的齿痕特征得到了显著放大，更利于机器对齿痕特征的识别。

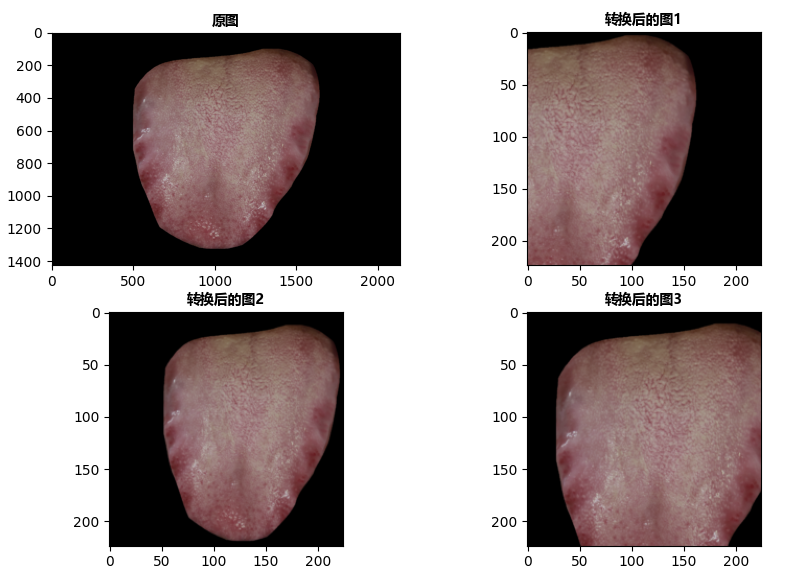


图6.1 RandomResizedCrop转换后示例图

对输入图像完成RandomResizedCrop操作后，接着对图像进行随机水平旋转，使用transforms中的RandomHorizontalFlip方法，示例如6.2所示。

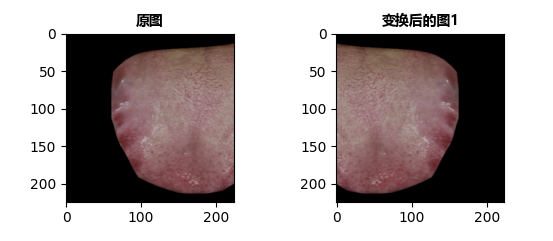


图6.2 RandomHorizontalFlip转换示例图

### 图像归一化

数据图像归一化的目的是在一定程度上限制预处理数据，从而去掉了复杂样本数据带来的负面影响。进行归一化处理后，可以让网络模型训练时加速损失梯度下降，提高网络的识别精度。

使用transforms中Normalize方法对输入舌象数据集的数据集进行归一化操作，该方法中所包含的参数mean、std由对应数据集图像数据的均值和方差求出，先将数据集图像全部转换为Tensor格式后，分别计算RGB三分量的均值和方差

均值计算公式：

方差计算公式：

计算各舌象数据集中图像数据的均值和方差结果如图6.3，将求得的小数保留3位，求得舌象数据集图像数据归一化所使用的Normalize方法的两个参数的最优值为：mean=[0.226，0.173，0.170]，std=[0.287，0.231，0.233]。

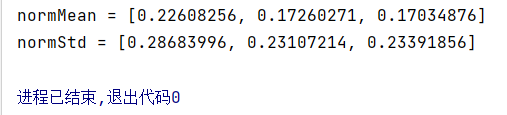


图6.3 数据集均值和方差计算结果

## 实验参数

卷积神经网络模型的训练过程是在不断地进行优化和调整的，而优化调整的实际效果和网络模型的训练参数设置是密切相关的，在网络模型训练开始前需要对网络的参数进行设置。本次实验使用的EffcientNet-B0网络的主要参数包括学习率(lr)、训练迭代次数(epochs)、训练批大小(batch-size)、优化函数等的。

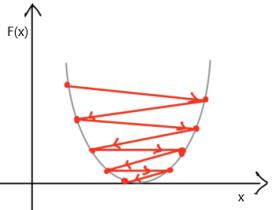
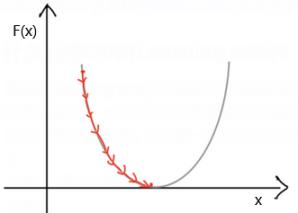
### 学习率

学习率代表着是深度学习中一个非常关键的参数，它决定了网络能否收敛到局部极小值，以及收敛效率。寻找到一个最优的学习率能够使得网络模型的既好又快，同时寻找最优学习率也是本次实验的重中之重。

网络模型梯度下降的公式：

学习率并不是越大越好，当学习率过大时，梯度可能会在全局最小值附件来回的进行抖动，在损失函数上带来不理想的结果，导致很难收敛，学习率过大反而适得其反。当学习率过小时，训练模型所需要的时间会变得非常非常的长，收敛到全局最小值会变得十分缓慢，对于实验者来说，收敛过慢导致浪费大量的时间是不可行的。

良好的模型训练离不开以后合适的学习率。



（a）学习率过小 （b）学习率过大

图6.4 过小或过大学习率下模型梯度下降曲线

### 训练迭代次数

在网络模型的训练过程中完成一个全部数据的前向和反向传播的完整过程叫做一次迭代（epochs），一般在训练过程中，模型往往需要多次迭代才能收敛至全局最小值，使模型从不拟合变成拟合状态。一般情况下，迭代的次数越多模型越收敛。由于本次课题中所使用的计算机显卡性能不够高，训练所用时间较长，所以不能盲目设置过大的epochs，默认设置为30轮，但在训练中训练到模型的最佳收敛状态即可停止训练。

### 训练批大小

训练批大小为每次训练时投喂一批图像数据的多少，训练批大小同样也是网络的一个重要的超参数。批大小过大时，会消耗更多的性能，过小时训练的收敛效果波动将会变大。本次使用使用的计算机GPU显存较小，批大小不能设置太大，所以将批大小设置为10。

## 实验评估指标

在实验中卷积神经网络不断的迭代训练，对于评定神经网络的训练效果，本课题主要使用损失函数（loss）和准确率（Acc）来评估模型的具体训练效果。实验中对比不同参数下的损失函数和准确率情况，从多次训练中找到最优的模型权重，从而找到上一节中所需要的最优的实验参数。

### 损失函数

损失函数就是用来表现预测的数据与真实的数据的差距程度，通常用于体现模型的训练效果，在训练中真实数据的实际值和神经网络模型的预测值一定会有误差，通常会使用一个函数表达这个误差就是损失函数，常用的损失函数有铰链损失（Hinge Loss）、平方损失（Square Loss）、交叉熵损失（Cross Entropy Loss）等等，本次实验中选择了交叉熵损失网络模型的损失函数。

交叉熵损失函数计算公式如下:

其中n为图像中像素数量；是真实值，取决于真实的样本；对输入样本的预测值。利用交叉熵损失公式6.3求出每一次批次输入图像训练后的损失，在计算出该轮迭代中的平均损失，利用可视化工具Tensorboard将每一轮迭代的平均损失记录在网页上，绘制loss曲线，方便在实验中更直观的判断模型的训练情况。

### 准确率

本实验主要是训练舌苔检测深度神经网络，舌苔检测属于图像分类的范畴，而在模型验证的时后，验证的精确率（Acc）能够最直观的反映模型训练的情况，精确率越高证明舌苔检测神经网络模型在该验证集下，对舌象特征的分类性能越优秀。准确率的计算公式如下：

在完成每一轮的迭代后，使用验证集对网络训练后保存的权重参数进行准确率测试，同时用可视化工具Tensorboard将每一轮迭代的平均损失记录在网页上，绘制loss曲线，方便在实验中更直观的判断模型的每一轮迭代后的性能。

## 对比实验

本小节为本次实验的具体内容，针对不同的实验参数进行对比实验，尽可能的找出最优的参数，使网络模型能够既好又快的收敛。

训练过程中编译语言使用的是Python3.6版本，实验环境是基于机器学习开源库Pytorch-1.10上进行的本次实验。操作系统为Windows10，配备了2.30GHz处理器、16GB内存和NVIDIA GeForce GTX 1050显卡。

### 预训练参数对比实验

模型训练时可以选择加载预训练模型，如果预训练模型对当前的项目的项目较为适合的话，能够极大的提升模型训练的效果。本次实验使用的EffcientNet-B0网络模型预训练参数来源于ImageNet数据集，该参数通过在ImageNet数据集上训练得出。本小节将对预训练参数进行对比实验，通过评估使用预训练权重的网络模型和未使用预训练权重的网络模型，比较网络性能，从而得到预训练参数是否有利于本课题舌苔检测网络模型的训练。

网络模型的超参数设置如表6.1，预训练参数在迭代开始初期的效果十分明显，迭代次数选择15是为了能够快速且显著的分辨出预训练参数的好坏。

|  |  |
| --- | --- |
| **参数** | **值** |
| img\_size（图片大小） | 224 \* 224 |
| batch\_size（批大小） | 10 |
| Lr（学习率） |  |
| optim（优化器） | SGD |
| weight\_decay （权值衰减） |  |
| momentum（动量） | 0.9 |
| epoch（迭代次数） | 15 |

表6.1 预训练参数对比实验网络训练参数

实验使用使用裂纹舌象数据集作为对比实验用例，神经网络模型训练好的参数权重使用上文提到的评估指标：loss和Acc进行优劣性能评估，先对使用预训练权重的网络对裂纹舌象数据集进行训练，训练过程中loss和Acc如图6.5所示。使用了预训练参数后，网络acc趋于上升，loss趋于下降，但有小幅度震荡，且未达到局部最小值。

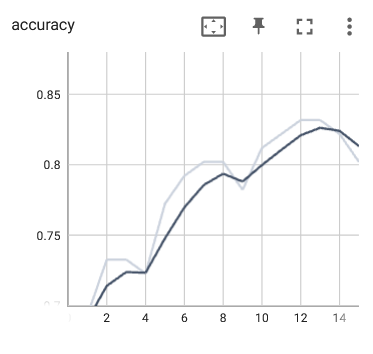


图6.5使用预训练参数网络训练loss、Acc曲线

接着对并未进行数据预处理的裂纹舌象原始数据集同样入同一网络训练迭代15轮，训练模型的loss和Acc如图6.6。从下图的loss和Acc曲线可以看出震荡非常大，而且loss十分的高，准确率基本在0.5上下震荡。

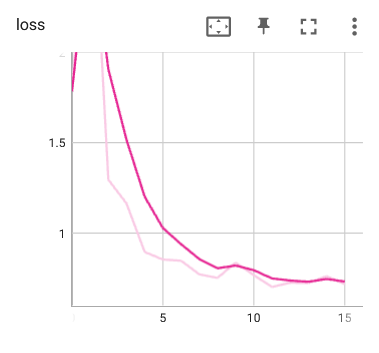


图6.6 未使用预训练参数网络训练loss、Acc曲线

通过对比使用预训练参数网络和未使用预训练参数网络的训练loss、Acc曲线，可以明显的看出使用预训练参数能让网络的前期迭代能够高效学习到舌象的特征，类似迁移学习，这样既节省了训练所用时间和计算资源，同时也能够很快的达到较好的效果，证明使用该预训练参数对模型训练的精度和效果都有较好提升，在接下来的实验中将会继续使用该预训练参数。

### 图像预处理对比实验

在上文中提到了实验前对图像进行预处理的重要性，本次对比实验的主要内容为对比使用了完整的图像预处理的图像数据与未经过任图像预处理的原始数据集图像分别放人相同的网络中训练。实验用的网络超参数如表6.2，所有参数均使用目前网络的默认设置，并使用上一小节所用到的预训练权重作为本实验的预训练权重参数。

|  |  |
| --- | --- |
| **参数** | **值** |
| img\_size（图片大小） | 224 \* 224 |
| batch\_size（批大小） | 10 |
| Lr（学习率） |  |
| optim（优化器） | SGD |
| weight\_decay （权值衰减） |  |
| momentum（动量） | 0.9 |
| epoch（迭代次数） | 30 |
| 预训练参数 | 有 |

表6.2 预处理对比实验网络训练参数

本次实验使用使用裂纹舌象数据集作为对比实验用例，神经网络模型训练好的参数权重使用上文提到的评估指标：loss和Acc进行优劣性能评估。首先是对进行了数据预处理（图像增强、输入处理、归一化处理）的裂纹舌象数据集进行训练迭代30轮，训练模型的loss和Acc如图6.5。可以发现迭代次数在15次左右时准确率基本已经逐步稳定，对验证集当中的裂纹舌象图片数据有着较好的辨识度，证明模型已经能比较稳定的对舌象的裂纹特征进行分类。在25轮迭代后，模型的损失函数已经开始有收敛的迹象。Loss曲线和Acc曲线都比较平滑，没有太大的抖动，训练的效果较为明显。

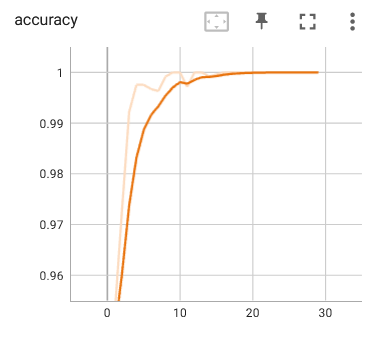
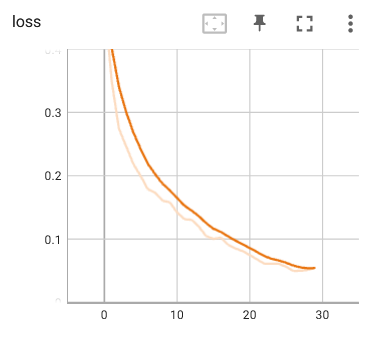
 

图6.7预处理后网络训练loss、Acc曲线

接着对并未进行数据预处理的裂纹舌象原始数据集同样入同一网络训练迭代30轮，训练模型的loss和Acc如图6.7。从下图的曲线可以看到，曲线震荡的比较厉害，准确率在10轮迭代前趋于上升，但在之后上下震荡十分不稳定，证明网络对舌象的裂纹特征进行分类稳定性不足。而loss曲线是趋于下降的，但是loss曲线同样震荡过大，而且曲线没有收敛的迹象，证明训练中模型对舌象的裂纹特征的预测值并不能很好的接近真实值，而且模型经过30轮迭代并没有到达局部最小值的迹象。

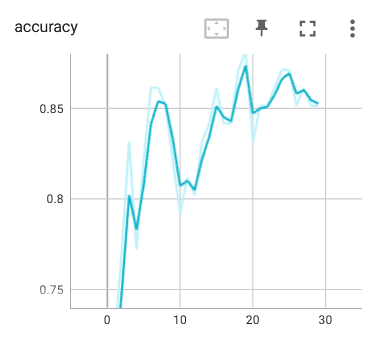
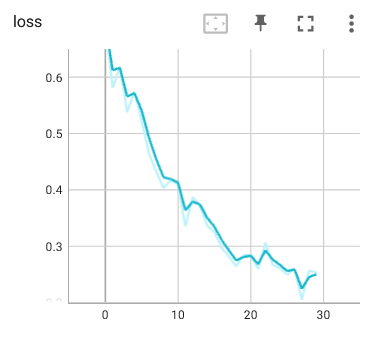
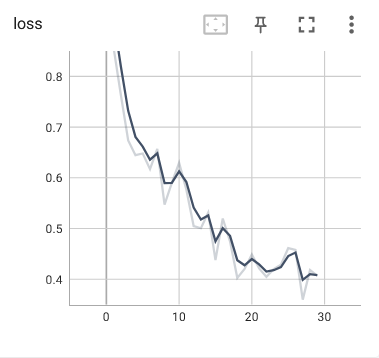
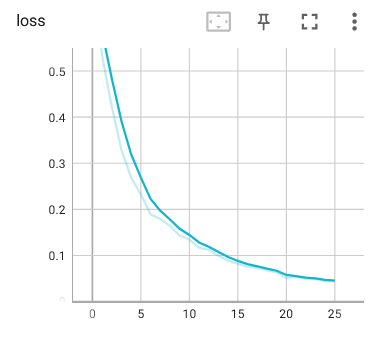
 

图6.8 未经预处理的网络训练loss、Acc曲线

同理也使用舌色舌象数据集进行了同样的对比实验，如图6.8。经过将进行过图像预处理的数据和未经过图像预处理的数据集放到同样的网络中对比实验可以明显的看出经过图像预处理后，模型的训练效果会更加优异一些，未经过预处理训练出来的loss曲线和Acc曲线震荡都十分明显，其主要原因图像的无关信息对训练时的干扰效果导致的，使得网络模型对裂纹舌象的特征分辨不够稳定。证明图像增强消除图像中的无关信息，提升有用的真实信息，增强信息的可检测性，最大程度的使图像数据简化，明显的提升了网络模型训练的效果，在接下来的实验中将会在图像输入网络之前对图像进行预处理，以提升训练效果达到舌苔检测的目的。



（a）经过预处理 （b）未经预处理

图6.9 舌色数据集训练loss曲线

### 学习率对比实验

学习率代表着是深度学习中一个非常关键的参数，它决定了网络能否收敛到局部极小值，以及收敛效率。寻找到一个最优的学习率能够使得网络模型的既好又快，本小节实验的目的就是通过对比在多个常用的学习率参数在本网络的表现，选择一个最为适合舌苔检测神经网络的学习率。

本次实验使用使用舌色舌象数据集作为对比实验用例，神经网络模型训练好的参数权重使用上文提到的评估指标：loss和Acc进行优劣性能评估。学习率使用目前常用的1，，，四个参数进行对比实验，其他参数不变，使用上述不同学习率进行网络训练，寻找最适合舌苔检测神经网络的学习率。实验用的网络超参数如表6.3

|  |  |
| --- | --- |
| **参数** | **值** |
| img\_size（图片大小） | 224 \* 224 |
| batch\_size（批大小） | 10 |
| Lr（学习率） | 1或或或 |
| optim（优化器） | SGD |
| weight\_decay （权值衰减） |  |
| momentum（动量） | 0.9 |
| epoch（迭代次数） | 30 |
| 预训练参数 | 有 |
| 图像预处理 | 有 |

表6.3 预处理对比实验网络训练参数

首先是使用学习率参数为1对舌象数据集进行训练迭代30轮，训练模型的loss和Acc如图6.10。loss非常之高，准确率低且震荡，故学习率为1不也行。

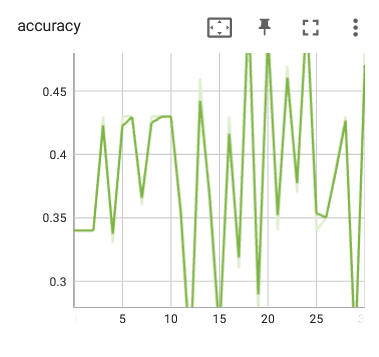
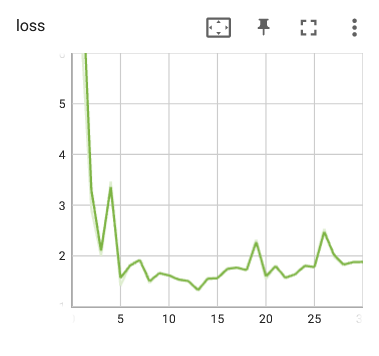
 

图6.10 学习率参数为1的Acc和loss

再使用学习率参数为对舌象数据集进行训练迭代30轮，训练模型的loss和Acc如图6.11。可以发现准确率相较学习率为1时提升不少，但是loss和Acc曲线波动较大，loss仍然较高收敛效果很差，证明学习率仍然过高，导致模型震荡无法收敛至局部最小值。

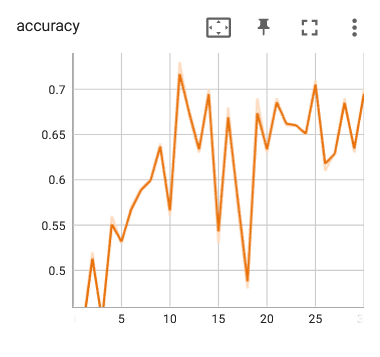


图6.11学习率参数为的Acc和loss

使用学习率参数为对舌象数据集进行训练迭代30轮，训练模型的loss和Acc如图6.12。从下图可以看出可以发现迭代次数在15次左右时准确率基本已经逐步稳定，对验证集当中的裂纹舌象图片数据有着较好的辨识度，证明模型已经能比较稳定的对舌象的裂纹特征进行分类。在25轮迭代后，模型的损失函数已经开始收敛的迹象，loss也较低。说明学习率对训练的模型的效果还不错。

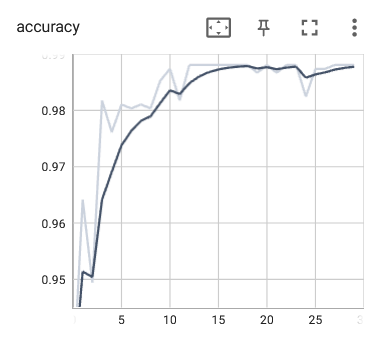
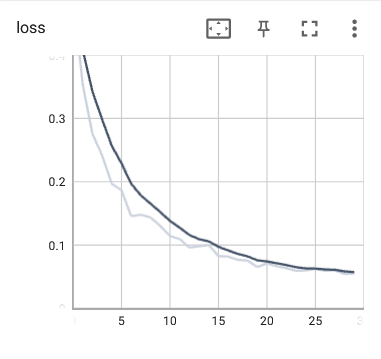
 

图6.12 学习率参数为的Acc和loss

使用学习率参数为对舌象数据集进行训练迭代30轮，训练模型的loss和Acc如图6.13。从下图可以发现，在的学习率下，模型的loss和Acc收敛效果也比较好。相较于学习率的情况，loss和Acc曲线比学习率的曲线要稳定一些，但是loss的收敛效果不如学习率的。该学习率的训练模型的效果也是同样不错的。

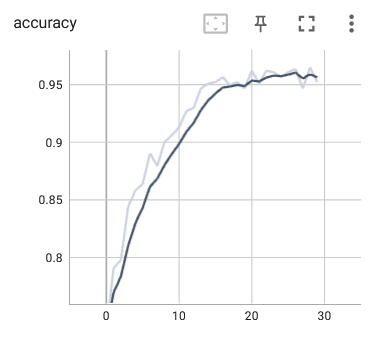
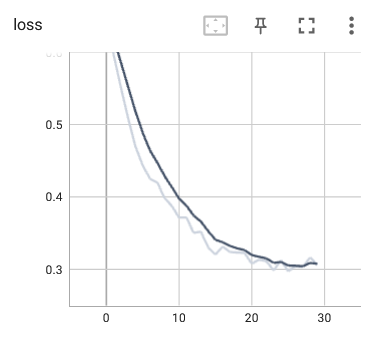
 

图6.13 学习率参数为的Acc和loss

通过4种不同学习率下训练完成的模型对比，和1的学习率对于舌苔检测的网络训练过高，训练震荡过大，无法收敛，所以和1的学习率并不适用于本实验。和在本次对比实验中的效果还不错，但是学习率的loss收敛效果较好，所以拟采用学习率作为舌苔检测训练的学习率参数。

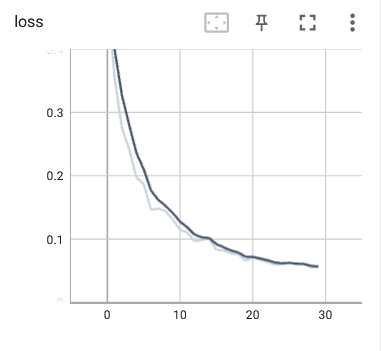
## 舌苔检测训练数据

经过上一节的对比实验找出了比较适合舌苔检测神经网络参数设置，本节使用通过对比实验找出的最优参数对第四章中构建四个舌象数据集分别训练，对每个数据集训练出来的网络权重分别保存，采用了这种同网络多数据集训练多模型的办法，能明显提升各个特征分类的正确率。对训练完成的神经网络模型训练好的参数权重使用上文提到的评估指标：loss和Acc进行优劣性能评估。

|  |  |
| --- | --- |
| **参数** | **值** |
| img\_size（图片大小） | 224 \* 224 |
| batch\_size（批大小） | 10 |
| Lr（学习率） |  |
| optim（优化器） | SGD |
| weight\_decay （权值衰减） |  |
| momentum（动量） | 0.9 |
| epoch（迭代次数） | 30 |

表6.4 舌苔检测神经网络参数

利用齿痕舌象数据集、裂纹舌象数据集、舌色舌象数据集、苔色舌象数据集分别训练完成的齿痕舌检测模型、裂纹舌检测模型、舌色检测模型、苔色检测模型，部分实验的损失函数loss曲线如图6.14。



6.14训练完成loss曲线

对完成训练的四个不同舌象特征的分类模型进行分类准确度测试，四类不同舌象特征的识别准确度都较高,由于苔色的分类中白苔的舌象图片占大多数导致苔色舌象的种类辨识偏低一些。

|  |  |
| --- | --- |
| **舌象特征** | **分类准确率** |
| 齿痕舌 | 0.93 |
| 裂纹舌 | 0.95 |
| 舌色 | 0.93 |
| 苔色 | 0.89 |

表6.5 各舌象特征识别准确率

## 体质辨识功能展示

利用上一章训练完成的舌苔检测网络模型，对用户输入的舌象图片进行检测，识别舌象图片中的各项舌诊特征（如齿痕、裂纹、舌色、苔色），分辨出各项特征后将其汇总，体质辨识模块将对照传统中医的舌象-体质图，对应图片信息解析出输入舌象图片中蕴含的体质信息。

当普通用户输入自己的舌象照片后，舌苔检测模块会将输入的图像使用训练好的网络模型权重进行分类并预测该舌象图片所含有的舌象特征，给出各项舌象特征分类预测的概率，如图7.1。

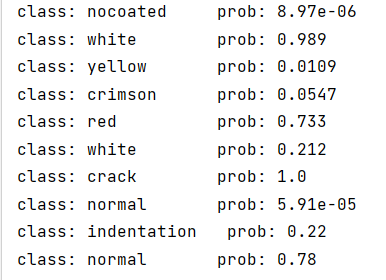


图7.1 各项舌象特征分类预测的概率

体质辨识模块统计舌苔检测模块对输入舌象图片的各舌象特征的预测值，并取最高的一组作为最终舌象特征预测，并通过对照传统中医的舌象-体质对应图如图7.2，解析出输入体质辨识结果。

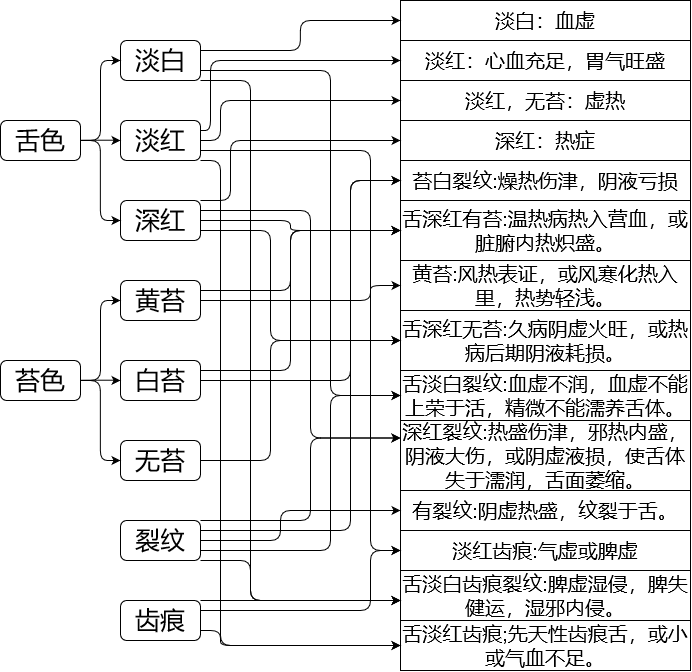


图7.2 传统中医的舌象-体质图

如图7.3，输入的图像为白舌苔，淡白舌色、舌面上有明显齿痕，根据传统中医的舌象-体质图可以辨识为：表证、寒证，主脾虚、血虚，水湿内盛证，舌胖大而多齿痕多属脾虚或湿困。

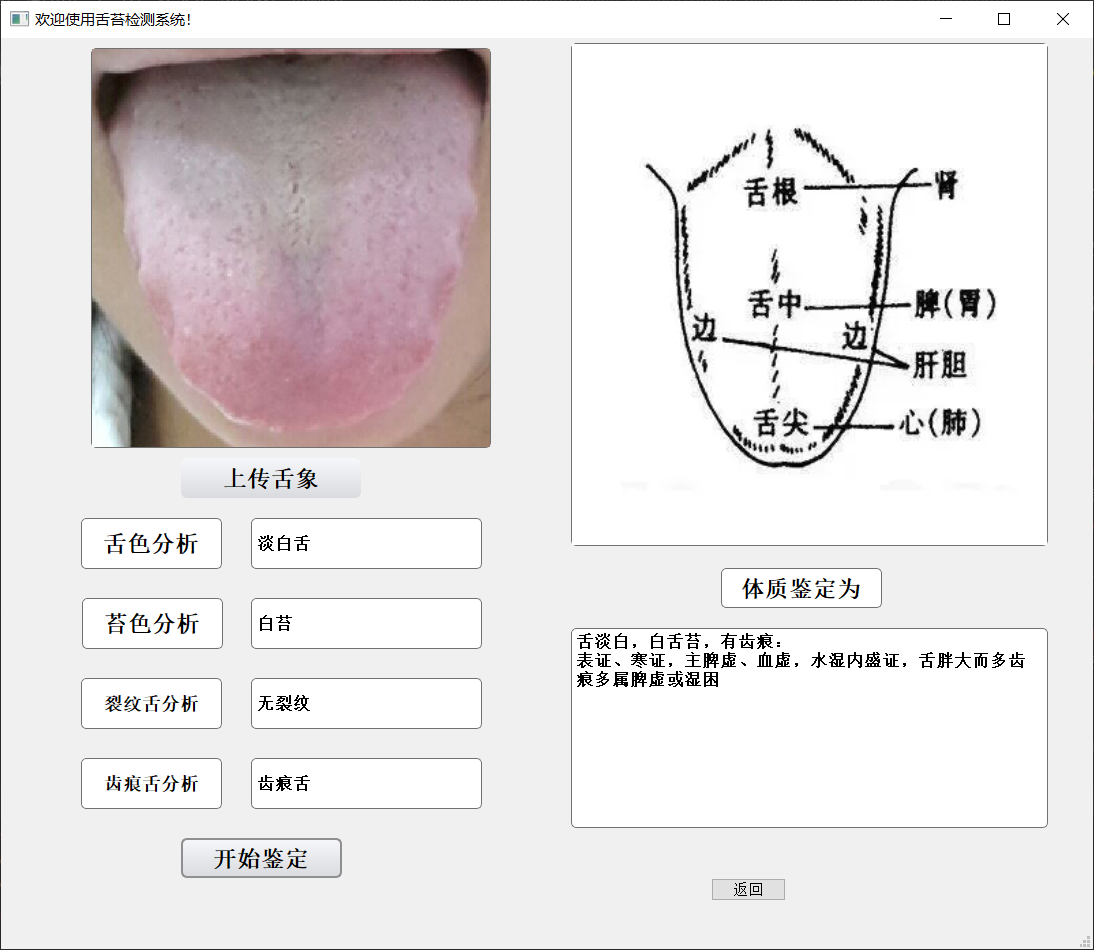


图7.3 舌苔检测体质辨识结果

## 本章小结

本章主要内容为利用上一章中所实现的舌苔检测网络模块进行网络训练、舌象检测、体质辨识等，并使用不同参数进行网络模型训练用于对比分析，对比得出较优的参数，用于后续的舌象检测，将舌苔检测后汇总舌象特征信息，对应舌象-体质对应图得出体质信息，实现对体质的辨识。

# 结论

## 全文总结

目前中医通过舌诊辨识人体体质信息，主要依据中医千年发展累积的经验，使得舌象诊断需要十分丰富的专业知识才能准确无误，所以传统中医面临着创新和传承的压力，对于传承和创新来说利用计算机技术实现这一复杂的诊断是一个具有可行性的手段。目前利用计算机技术实现的智能化舌诊项目，往往需要在特定的环境或者需要十分专业的设备进行辨识分析，使用专业的设备和环境使得检测门槛较高，对舌诊的推广有一定的影响。本文通过使用机器学习开源框架，深度学习应用到了中医舌诊和体质判断当中，实现了使用普通的计算机设备也实现对舌苔的检测和体质的判别。论文的主要工作如下：

（1）在没有足够多的舌象图片数据的情况下，利用图像增强和生成对抗网络DCGAN扩充数据集，数据集制作时和其他图像分类项目的单一数据集有所不同，采用了同网络多数据集训练多模型的办法，明显提升各个特征分类的正确率。

（2）利用EffcientNet-B0神经网络实现了对舌苔的检测。利用多次对比实验找出最适合舌苔检测神经网络的超参数，后对四个不同的舌象特征数据集进行训练，完成训练的四个不同舌象特征的分类模型进行分类准确度测试，四类不同舌象特征的识别准确度都较高，对所分类的舌象特征识别率能达到90%以上。

（3）实现了对舌象中蕴含的体质信息进行判别的功能，上传舌象图片后，运行分类网络输出的舌象判断结果并展示，使用舌象结果对照传统中医的舌象诊断经验，对人体体质信息进行判断并展示给用户。

## 不足之处

本课题利用深度学习实现了舌苔检测和人体体质识别，在一定程度上能够辅助中医师的舌诊以及方便普通用户对舌苔进行自我检测识别自身体质信息，但本课题仍然存在许多缺陷和不足之处。

首先是舌象图片数据样本不足，样本不足一定程度上制约了所使用的神经网络的学习能力和对舌象特征的分辨能力。但是我相信在进一步加大舌象样本的情况下，网络应该可以达到更好的效果。

其次，对舌象中蕴藏的人体体质信息判别还不够全面，人体是十分复杂的结构，本课题所使用的图像分类神经网络，也只能片面的预测人体体质，想要精确的进行舌苔检测判断人体体质信息，需要更多方面的数据分析。

最后则是目前本项目的功能还不够完善，在各方面仍有许多的提升空间，特别是体质判别的精度还要进一步提升。

# 谢 辞

时光飞逝。转眼间，大学生涯即将结束。回首往事，大学时光过得如此匆忙。在这个充满奋斗的过程中，给我的学生生涯带来了无限的激情和收获。

首先，我要诚挚的感谢我的导师刘忆宁教授对我的耐心指导和关心。在完成毕设的过程中，导师给我提供了很多帮助和有用的建议，在百忙之中的对我的论文提出许多宝贵的修改意见。

其次，我要感谢赵志强和罗力两位研究生学长，他们为我的毕设工作提供了很大的帮助，对我提出的问题都能不厌其烦的为我解答，和我一起仔细分析设计思路和方案。

而后，要感谢学院领导和辅导员还有我的舍友跟同学，感谢你们在过去四中的陪伴和关怀。我也感谢家人和朋友的支持和理解。他们总是我前进的动力。

最后，我要感谢所有参与本文评审和论文答辩的老师们。谢谢你的帮助和支持。我衷心祝愿所有帮助和支持我的人。我将继续努力工作，保持勤奋积极的生活态度，在未来的工作中取得更好成绩。

# 参考文献

1. 王爱民，赵忠旭，沈兰荪.基于矢量Prewitt算子的多尺度彩色图像边缘检测方法[J].中国图像图形学报,1999, 4(12):12~16.
2. 郑媛媛.基于深度学习和注意力机制的文本情感分类研究[D].焦作：河南理工大学,2019.
3. 曾海彬.基于深度神经网络的舌苔体质辨识[D].广州：华南理工大学,2017.
4. 孟婷婷.基于确定学习与深度学习的人体体表心电信号建模与分类[D].杭州：杭州电子科技大学，2020.
5. 孟立联."十四五"卫生健康发展新要求初探[J].中国农村卫生事业管理，2020，40(12): 842-847.
6. 俞彬.基于生成对抗网络的图像类别不平衡问题数据扩充方法[D]. 广州：华南理工大学,2018.
7. 於建.边云协同智慧医疗架构下的心电图智能诊断研究[D]. 长沙：湖南大学，2020.
8. 马佳炯.基于零样本学习的舌苔体质识别[D]. 广州：华南理工大学，2019.
9. 涂帅.基于舌象分析的中医体质检测研究与实现[D]. 湘潭：湘潭大学，2021.
10. 洪乐，叶双林，刘婷婷等.基于舌象分析的中医体质辨识系统的研究思路[J].中国中医药现代远程教育，2016，14(02)：34~36.
11. 陈松鹤，梁嵘，王召平.6种舌苔颜色数据的三维分布特征的描述[J].时珍国医国药，2009，20(11)：2852~2854.
12. 吴加林.基于增强深度学习的皮肤疾病图像分类方法研究[D].西安：西京学院，2021.
13. 巫毅，吴钢华.基于Efficientnet的宫颈图像的识别分类[J].计量与测试技术，2021，48(12): 62~65.
14. Tang Y, Sun Y. Research on Multiple-Instance Learning for Tongue Coating Classification [J]. IEEE ACCESS, 2021, 9: 66361-66370.
15. Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural computation, 2006, 18(7): 1527-1554.
16. Chiu C C. A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue[J]. Computer methods and programs in biomedicine, 2000, 61(2): 77-89.