

哈爾濱工業大學

畢業設計（論文）

題 目 基于深度学习的舌体与舌下络脉  
分割及其颜色特征分类研究

专 业 计算机科学与技术

学 号 160400105

学 生 侯英妍

指 导 教 师 张盛平

答 辩 日 期 2020 年 6 月 8 日

## 摘 要

本文基于国家重点研发计划，目的是推动中医舌诊的客观化发展，借助先进的计算机技术，用于传统的中医诊断领域，将人工智能与中医舌诊相结合，为中医舌诊的变革提供新的驱动力。舌象客观化研究中，舌体与舌下络脉的分割提取的结果将会影响到之后的特征分析。舌下络脉的颜色对中医诊断有重要作用，中医可以通过观察舌下络脉的颜色，来判断病人的血液循环情况。目前对舌下络脉分割的研究方法，对部分图片无法分割，且效果不佳。对舌下络脉分类的方法基于单一颜色块，无法代表全部颜色信息。基于此，本文对舌体与舌下络脉语义分割进行研究，并对舌下络脉进行颜色分类。研究内容如下：

（1）舌体与舌下络脉的语义分割。本文基于深度学习的全卷积网络语义分割，设计无下采样语义分割模型，加入膨胀卷积扩大网络的感受野，同时对膨胀卷积率进行设计，避免了栅格效应。针对舌下络脉小目标特性，设计先定位再分割的小样本分割方法。并采用去除反光点、数据增强等数据预处理方法和边缘腐蚀等数据后处理方法，进一步提升了语义分割的效果。经实验，模型效果优于经典医学图像语义分割 U-net 模型；

（2）基于上一步语义分割后的舌下络脉，对其颜色进行颜色三分类，分为淡紫、青紫、紫黑三类。基于语义分割后的络脉，设计不同的输入到分类模型的络脉提取特征方案，经实验选取效果最佳的双络脉矩形输入。并实验不同的深度学习模型，如 ResNet、VGG Net、DenseNet 等，通过实验选择分类效果较好的模型，对其进行调整，以得到最佳的分类效果。

基于上述所提出的中医舌体与舌下络脉分割提取方法、数据的预处理和后处理算法及颜色分类算法，设计系统对算法进行展示。系统整体基于 PyQt5 实现，并采用模块化的设计，完成了算法与实际项目的结合，较好地将舌体与舌下络脉语义分割和颜色分类算法效果进行了演示。

**关键词：**舌下络脉；医学图像语义分割；络脉颜色分类；中医客观化

## Abstract

This article is based on the national key research and development plan, the purpose is to promote the objective development of Chinese medicine tongue diagnosis. In the objective study of tongue image, the results of segmentation and extraction of tongue body and sublingual choroid will affect the subsequent feature analysis. The color of the sublingual collaterals plays an important role in the diagnosis of traditional Chinese medicine. The current research method of sublingual collateral segmentation cannot segment some pictures, and the effect is not good. The method for classifying the sublingual collaterals is based on a single color block and cannot represent all color information. The research content is as follows:

(1) Semantic segmentation of tongue and sublingual veins. In this paper, based on the deep learning of full convolutional network semantic segmentation, we design a semantic segmentation model without downsampling, add dilated convolution to expand the receptive field of the network, and at the same time design the dilated convolution rate to avoid grid effects. Aiming at the characteristics of the small target of the sublingual collaterals, a small sample segmentation method is designed first, and then segmented. And the use of data pre-processing methods such as removing reflective points, data enhancement, and post-processing methods such as edge corrosion further improves the effect of semantic segmentation. Through experiments, the model effect is better than the classical medical image semantic segmentation U-net model;

(2) The colors are classified into three types, which are divided into three types: light purple, blue purple, and purple black. Based on the collaterals after semantic segmentation, a collateral extraction feature scheme with different inputs to the classification model is designed, and the double-column rectangular input with the best effect is selected through experiments. And experiment with different deep learning models, choose the model with better classification effect through experiment, and adjust it to get the best classification effect.

Based on the above mentioned extraction method of tongue and sublingual choroids of TCM, pre-processing and post-processing algorithms of data and color classification algorithm, the design system displays the algorithm. The whole system is implemented based on PyQt5, and adopts a modular design, and demonstrates the effect of semantic segmentation and color classification algorithm of tongue and sublingual veins

**Keywords:** sublingual collaterals, semantic segmentation of medical images, collateral color classification, objectification of traditional Chinese medicine

# 目 录

摘 要.....	I
Abstract .....	II
第 1 章 绪 论 .....	1
1.1 课题背景及研究的目的和意义 .....	1
1.1.1 课题背景 .....	1
1.1.2 研究的目的和意义 .....	1
1.2 国内外在该方向的研究现状及分析 .....	2
1.3 本文研究内容与研究流程 .....	5
1.4 本文结构安排 .....	6
第 2 章 相关技术 .....	8
2.1 引言 .....	8
2.2 语义分割神经网络介绍 .....	8
2.2.1 全卷积神经网络 .....	8
2.2.2 膨胀卷积 .....	9
2.2.3 无下采样语义分割神经网络 .....	10
2.3 去除镜面反射算法 .....	11
2.4 深度学习图像分类模型 .....	11
2.5 本章小结 .....	14
第 3 章 舌体及舌下络脉语义分割 .....	15
3.1 引言 .....	15
3.2 数据预处理方法 .....	15
3.2.1 数据集生成 .....	15
3.2.2 反光点去除算法 .....	16
3.2.3 数据扩充 .....	17
3.3 基于全卷积网络的语义分割算法模型 .....	18
3.3.1 语义分割模型 .....	18
3.3.1 小目标语义分割策略 .....	19
3.3.2 损失函数 .....	20
3.4 数据后处理 .....	21

3.4.1 去除孔洞和杂散点 .....	21
3.4.2 分割边缘的膨胀与腐蚀 .....	23
3.5 实验结果.....	23
3.5.1 数据集描述 .....	23
3.5.2 语义分割评价指标 .....	24
3.5.3 评价指标结果 .....	25
3.5.4 与其他模型实验对比 .....	26
3.5.5 语义分割图像效果展示 .....	27
3.6 本章小结.....	29
<b>第 4 章 舌下络脉颜色分类 .....</b>	<b>30</b>
4.1 引言.....	30
4.2 舌下络脉输入分类模型方法.....	30
4.3 基于深度学习的分类模型.....	32
4.4 实验结果.....	35
4.4.1 数据集描述 .....	35
4.4.2 颜色分类评价指标 .....	35
4.4.3 实验输入络脉方案及结果 .....	35
4.4.4 图像分类模型实验 .....	35
4.5 本章小结.....	36
<b>第 5 章 舌体与舌下络脉分割及颜色分类系统设计实现.....</b>	<b>38</b>
5.1 引言.....	38
5.2 系统总体实现流程.....	38
5.3 系统功能架构设计.....	39
5.4 系统功能实现.....	41
5.5 系统功能展示.....	42
5.6 本章小结.....	46
<b>结 论 .....</b>	<b>47</b>
<b>参考文献 .....</b>	<b>48</b>
<b>致 谢 .....</b>	<b>51</b>

# 第1章 绪 论

## 1.1 课题背景及研究的目的和意义

该毕业设计课题来源于国家重点研发计划——“中医健康数据融合标准与典型疾病诊断建模研究”项目。

### 1.1.1 课题背景

中医从人体状态的变化规律入手，建立了“四诊合参，辨证论治”的独特理论体系。中医四诊包括“望、闻、问、切”四个部分，每种检测都是从不同的角度观察病情，具有不同的方法和意义。

中医学是中国悠久的文化的重要组成部分。中医从人体的变化当中寻找规律，建立了自己的理论体系：“四诊合参，辨证论治”。“望、闻、问、切”是传统中医诊断的四个部分，均有着重要的意义，在每一个部分中医都会从不同的角度来观察疾病的变化。

舌诊，又称“望舌”，是通过观察舌象（舌质、舌苔等）以了解人体生理功能和病理变化的一种诊察方法。中医望诊中，舌诊是其中的主要内容，历代以来，都受到医家的重视。可以说，在临床诊疗实践和中医理论体系，舌诊都有着举足轻重的地位。传统的舌象诊断方法，比起现代医学诊断方法，具有先天独特的优势。例如，作为非侵入的诊断检测方法，舌诊不会给患者带来辐射照射伤害或是身体的损伤。然而，缺乏定量的描述，且舌诊受到医生和受治疗者的各种因素影响，再加上中医诊病有一定的主观性，通常取决于中医医生的临床经验和实践经验，重复性较差，主观性较大。如上述原因所述，这给中医的研究造成很多负面影响，更是影响了舌诊的继承和发扬。

### 1.1.2 研究的目的和意义

现如今，国家对中医药的发展十分重视，颁布并实施了《“健康中国 2030”规划纲要》、《中医药发展“十三五”规划》以及《中医药法》等等，把发展中医药这一事宜提升到了国家战略的层面。同时，中医临床现代化的进一步发展也对中医舌诊客观化、量化提出新的挑战和要求。舌诊研究的大方向，就是以前沿的计算机技术手段来研究舌诊，提高中医舌诊的临床应用价值，加快舌诊的科学化、客观化、具体化、实时化、智能化<sup>[1]</sup>。

结合中医学理论、计算机相关的图像处理和深度学习技术，用客观的检测指标，对中医学的舌象进行定性、定量、定位的研究，以推动舌诊客观化的发展。舌体分割的目的，是为了减少图像中其他因素的干扰，从背景中把舌体分离出来。中医对舌下络脉的观察，同样是舌诊的重要部分，从原始图像中将舌下络脉进行提取，便于之后的疾病诊断、特征分析（如颜色特征分析），有着重要的意义。舌象客观化研究中，舌体与舌下络脉的分割提取是十分重要的，分割结果将会影响到之后的特征分析。舌下络脉的颜色对中医诊断有着很重要的作用。中医可以通过观察舌下络脉的颜色，来判断病人的血液循环情况。对于健康人群来说，舌下络脉的颜色一般为淡紫，如果舌下络脉太突出、发青紫或者紫黑色，有可能代表着身体亚健康或有一定疾病。基于舌下络脉的颜色特征分析进行颜色三分类，能够初步进行颜色分类，以便于后续的诊断。

该项目旨在推动中医舌诊的客观化发展，借助先进的计算机技术，用于传统的中医诊断领域，将人工智能与中医舌诊相结合，形成人工智能领域的单点突破，为中医舌诊的变革提供新的驱动力，带来社会、经济、生态效益。

## 1.2 国内外在该方向的研究现状及分析

近年来，随着计算机视觉和深度学习等学科的不断进步和发展，语义分割问题成为计算机视觉的关键问题之一。语义分割，与普通分割不同的是，其是对图像的每一个像素都进行分类。如今，在越来越多的应用场景中，都需要精确的分割技术，如医学图像分割、虚拟现实、甚至在自动驾驶等等场景中，都会涉及到语义分割。语义分割一般分为两条路线，一条是基于传统的机器学习方法，例如，Li 等人提出了一个完整的舌下络脉提取系统<sup>[2]</sup>，获取舌脉重要特征。另外一条路线是基于深度学习卷积神经网络的语义分割方法。随着深度学习的不断进步和前人的不断探索，卷积神经网络（Convolutional Neural Networks, CNN）使得语义分割领域有了显著的发展。如图 1-1 所示，为利用深度学习解决语义分割问题的主要模型。

2014 年，Long 等人提出了完全卷积神经网络<sup>[3]</sup>，该网络去掉了原先末端的全连接层，这个模型是语义分割领域的“鼻祖”，随后的语义分割几乎都采用了这种架构。2015 年，Ron 等人提出 U-net 模型<sup>[4]</sup>，尝试把池化层的结果应用到译码的过程中，为网络引入了更多的编码信息。在生物医学图像处理中，得到图像中的每一个细胞的类别标签是非常关键的。如图 1-2 所示，U-Net 是非常著名的解决方案，它在全连接卷积层上构建模型，对其做了修改使得它能够在少量的训练图像数据上运行，得到了更加精确的分割。





Pyramid Pooling, ASPP)<sup>[6]</sup>, 这种架构可以对不同尺度的信息进行结合, 并使用了全连接条件随机场 (Conditional Random Field, CRF) 来进行后处理。Zhao 等人提出了一种金字塔场景解析网络<sup>[7]</sup>, 其通过应用大的核心池化层来提高模型的感受野, 并加入了膨胀卷积来修改 ResNet 神经网络, 还加入了金字塔池化模块。

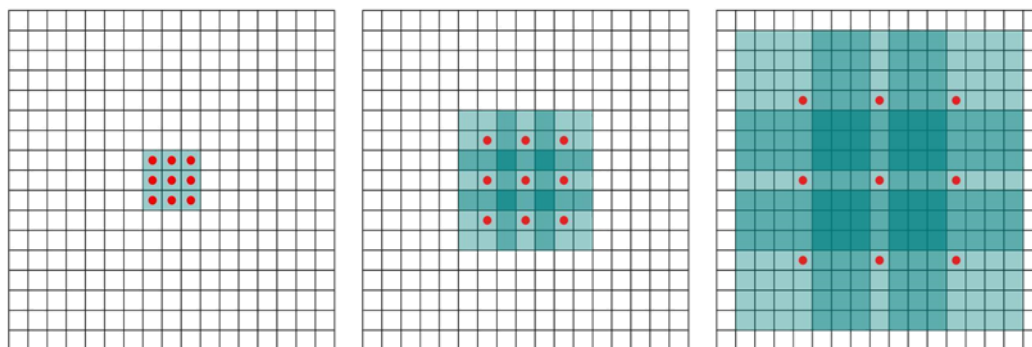


图 1-3 膨胀卷积核与感知域

2017 年, Peng 等人提出了使用具有非常大的内核卷积的编码器-解码器体系结构<sup>[7]</sup>, 该模型编码器采用 ResNet, 译码器采用反卷积和图卷积网络 (Graph Convolutional Network, GCN), 还使用了名为 Boundary Refinement 的残余模块。2019 年, KaiMing He 等人巧妙地考虑把图像的分割问题当作一种渲染问题, 提出了 PointRend<sup>[8]</sup>, 其可以被合并到流行的元架构中, 用于语义分割, 如全卷积网络 (Fully Convolutional Networks, FCN)。PointRend 显著改进了 DeepLabV3 模型。

针对具体的舌体与舌下络脉问题的研究, 国外的相关文献极少, 对于舌下图像分割问题, 2018 年, 日本的 Yang 等人用相同训练图像的有限数量标签的不同集合重新训练 U-net 模型<sup>[9]</sup>, 提出了自动分割舌下络脉的方法。2019 年, 同一团队改进了原来的模型, 分解整个编码器-解码器框架并协同更新参数, 提出了一个协同注意力网络<sup>[11]</sup>。最后对得到的舌下络脉的特征向量进行分析, 本课题进行的是颜色特征的三分类研究。前人进行过分类算法的一些相关研究。对于机器学习的分类器算法来说, 包括稀疏表示分类器 (Sparse Representation-based Classifier, SRC)<sup>[12]</sup>, 协同表示分类器 (Collaborative Representation Classifier, CRC)<sup>[13]</sup>, 线性分类器 (Linear Regression Classifier, LRC)<sup>[14]</sup>, 最近邻算法 (Nearest Neighbor, NN)<sup>[15]</sup>, 支持向量机 (Support Vector Machines, SVM)<sup>[16]</sup>等多种分类算法。对于舌下络脉的颜色特征分析, Takeichi 等人<sup>[17]</sup>利用计算机量化舌下静脉的颜色信息, 实现了血瘀诊断的客观化。

在舌诊现代化早期, 国内已经有一些研究机构利用图像处理及模式识别等技术对舌象进行研究, 在舌体和舌色等特征提取方面积累了大量经验, 取得了一定的

成果。在 1980 年代初期,我国的一些专家<sup>[20]</sup>就对舌诊进行定量研究,主要研究关注在色彩方面。当时的技术有限,舌头图像的采集主要使用光纤,或者使用相机,来制作幻灯片<sup>[19]</sup>。在 1990 年代中期,北京中医学院对中医舌象分析的主要方法做了一些研究<sup>[21]</sup>,对自动分割舌体,以及系统地对常见舌象特征(包括舌色、苔色、湿度等)等问题开始研究。

舌体分割将舌体从背景中分离,分割效果影响到后期特征分析的准确性。近年来,前人在这方面做了诸多研究,获得了一定的成果。Wu 等人<sup>[23]</sup>利用分水岭算法,获得一定的舌体轮廓,并基于这一步地轮廓,再利用 Snake 模型获得最终的舌体轮廓。可看出,在舌体分割中,活动轮廓模型<sup>[22]</sup>及一些变体得到了研究。张宏志等人<sup>[24]</sup>提出了 PolarSnake 模型,考虑到舌体轮廓为封闭曲线,他们将舌体图像转换到极坐标系下,在极坐标系下分割出舌体轮廓,并利用这个模型分割舌体边缘,再转换回到直角坐标系下,得到最终地舌体分割效果。

除了上述的 PolarSnake 模型,水平集方法也常用于舌体分割中。Zhu 等人<sup>[25]</sup>利用水平集方法,开发了一种快速地舌体轮廓分割方法。Li 等人<sup>[26]</sup>利用图的方法来分割舌体,并将舌体的几何信息加入进去。Zhang 等人<sup>[27]</sup>提出了一种动态阈值和修正模型的舌体分割算法,即对动态阈值法提取的舌象轮廓再做集合修正得到最终轮廓,他们提出的算法改进了自适应阈值分割法的问题,如用于舌体部分轮廓丢失、舌体边缘轮廓不连续等等。

舌下络脉的语义分割,也属于本文分割算法的部分。之前对于舌下络脉分割地传统图像处理算法大致分为三种:基于掩膜的方法、基于分类器的方法以及基于搜索<sup>[28]</sup>的方法。基于分类器的方法例如闫子飞等人<sup>[29]</sup>中对舌下络脉分割的研究。Zhang 等人<sup>[30]</sup>也采用同类方法,并在舌下络脉分割后又对病例开展研究。Zheng 等人在中医舌诊计算机舌下络脉特征攫取及分析中提出了利用计算机提供舌下静脉诊断的客观数据<sup>[31]</sup>。Sun 等人在基于特征聚类的舌下络脉自动提取方法中实现了对舌下静脉样本图像分割<sup>[32]</sup>,并计算出一些络脉参数,包括络脉长宽比、络脉颜色信息等等。目前国内相关文献还没有采用深度学习方法进行舌体或舌下络脉语义分割的相关研究。

目前对于舌体及舌下络脉语义分割的问题上,绝大多数文献采用的方法仍为机器学习方法,存在着费时费力、算法复杂复用性差、效率不尽如人意等缺点。相比之下,卷积神经网络等深度学习模型可以更好地学习特征,可以在较少人工参与的情况下获得更好的结果。因此,针对该课题,拟采用深度学习的方法,用“新的方法”解决“旧的问题”。

### 1.3 本文研究内容与研究流程

该课题的研究主要是针对舌部图像进行舌体分割和舌下络脉分割，并对提取出的舌下络脉进行颜色特征分析，其研究内容包括一个完整的数据分析的过程。数据的预处理属于研究的基础部分，是很关键的一步，高质量的数据集对于后续的语义分割和特征分析都有很大的帮助。为了得到可用的数据集，对采集的数据需要进行语义标注、标签选择、批量处理等一些预处理操作。

如图 1-4 所示，本研究课题的主要研究内容有三个方面的：首先，搭建模型提取舌体部分（如图 1-4 蓝色梯形部分）；其次，设计舌下络脉的语义分割模型提取舌下络脉（如图 1-4 黑色标线部分）；最后，对上一步提取出的舌下络脉进行颜色特征分类研究。

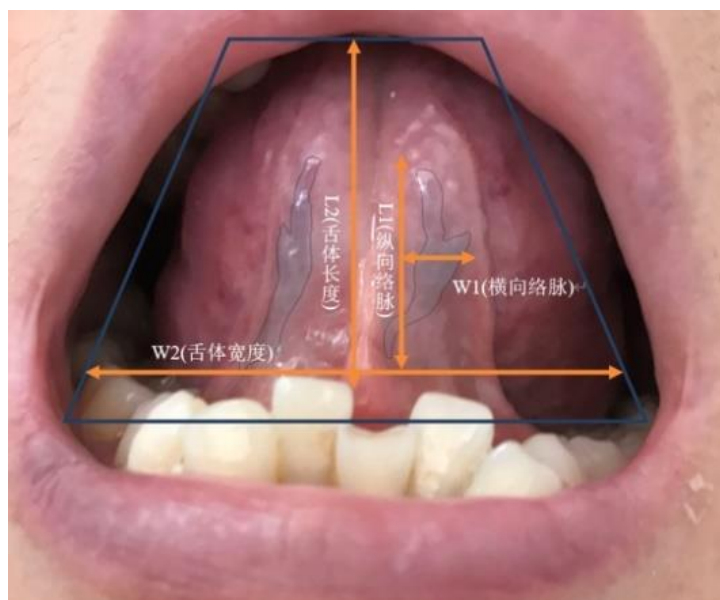


图 1-4 舌下络脉特征示意图

研究内容如图 1-5 所示，对于原始图片语义标注制作数据集，将两类数据集输入到舌体提取模型与舌下络脉提取模型中，输出模型预测结果，并对提取出的舌下络脉进行颜色特征分类研究。

## 1.4 本文结构安排

本论文总共安排五章内容，分别第一章是绪论，第二章相关研究综述，第三章基于深度学习的舌体与舌下络脉语义分割研究，第四章为舌下络脉的颜色分类，第五章为系统设计与实现，最后对本文的研究进行总结与展望。

第 1 章介绍相关的课题背景和研究的的内容、目的和意义，对研究的三个方面的研究现状进行了介绍，并进行了总结与分析。

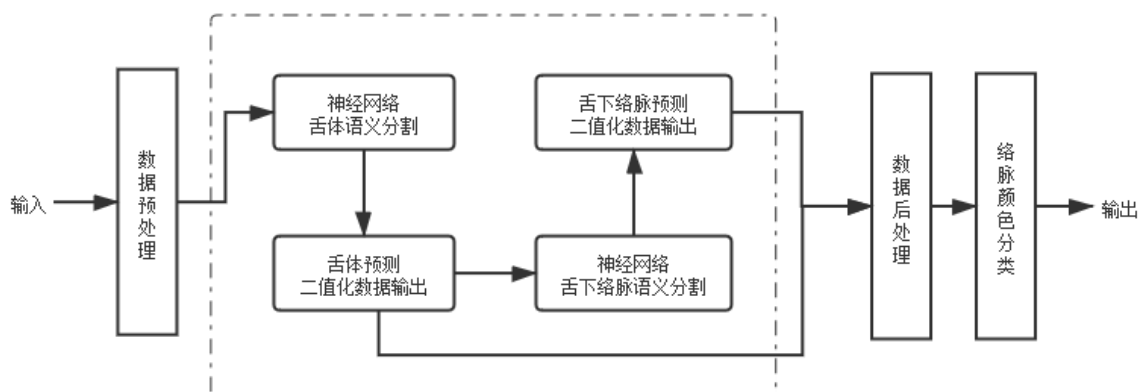


图 1-5 研究内容示意图

第 2 章为相关技术，主要介绍了与本文密切相关的相关研究，首先是语义分割领域，全卷积神经网络、膨胀卷积及无下采样模型设计，其次，对去反光点的相关研究，以及关于基于深度学习的分类模型。

第 3 章为基于深度学习的舌体与舌下络脉语义分割研究。针对舌下络脉小目标特征进行模型设计。最后对本章方法进行实验，介绍了实验使用的数据集、评估指标、实验参数和可视化结果，并与其他模型进行了对比。

第 4 章基于分割出的舌下络脉，设计不同方案提取颜色信息，并通过尝试不同的深度学习分类模型对舌下络脉进行颜色分类。最后对本章方法进行实验，介绍了实验评估指标和实验结果。

第 5 章介绍可应用系统的设计与实现。系统采用 PyQt5 实现，是第 3、4 章算法的应用展示。

最后为总结与展望，对本文研究工作进行总结，并对未来研究工作进行了展望。

## 第 2 章 相关技术

### 2.1 引言

本文所研究的对于舌下络脉的语义分割及颜色分类是计算机科学与中医科学所交叉的研究领域，是针对具体问题的研究，要根据舌下络脉的特点进行模型的设计，因此需要进行大量相关的文献调研。本章主要对该毕业设计涉及的相关研究及技术进行了介绍，其中包括语义分割领域常用的全卷积神经网络，以及常用的深度学习分类网络。

本章节接下来的内容如下：第 2.2 节介绍了语义分割神经网络，包括全卷积神经网络在语义分割领域的相关研究，无下采样的神经网络模型，膨胀卷积的引入及栅格效应的解决方法等相关研究；第 2.3 节介绍了去反光点的相关研究；第 2.4 节介绍了深度学习分类模型；最后 2.5 节对本章节进行了总结。

### 2.2 语义分割神经网络介绍

深度学习语义分割在前人的研究中不断发展，接下来将对本文中语义分割模型设计涉及到的相关研究进行介绍。在第 2.2.1 小节中，介绍全卷积神经网络模型；在第 2.2.2 小节中，对膨胀卷积及其栅格效应进行介绍；在第 2.2.3 小节中，将介绍无下采样的语义分割模型。

#### 2.2.1 全卷积神经网络

深度学习的第一种比较流行的分割方法是补丁分类。这种方法对中心像素进行分类，通过逐像素提取周围像素的方式。当时在卷积神经网络的末端，仅使用了完全连接的层，并且输入必须是图像块，大小有限制，因此我们只能使用这种按像素分割的方法。

全卷积网络（Fully Convolutional Networks, FCN）<sup>[3]</sup>，最后删除全连接层。后续语义分割模型基本上采用此结构，这是当前深度学习语义分割技术的发起者。FCN 将传统卷积神经网络（Convolutional Neural Networks, CNN）的完全连接层转换为卷积层。因为所有层都是卷积层，称为完全卷积网络。全卷积网络在像素级别对图像进行分类，并解决了图像语义分割问题。与传统的 CNN 不同，传统的 CNN 使用完全连接的层来获取固定长度的特征向量以在卷积层中进行分类，而 FCN 则接受任意大小的输入图像，并使用反卷积层，可以运行最终的卷积层要素图，上采

样以恢复与输入图像相同的大小。这样，就可以在保留原始输入图像的空间信息的同时为每个像素生成预测，并最终对像素的特征图执行奇/偶采样，如图 2-1 所示，FCN 从图像级别的分类进一步延伸到像素级别的分类。

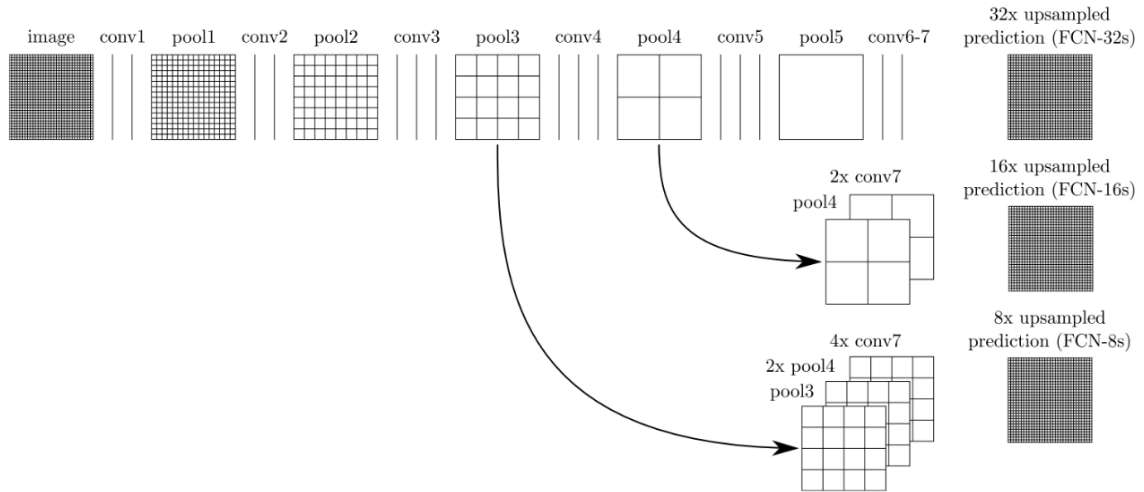


图 2-1 全卷积神经网络卷积和反卷积上采样过程

但完全卷积网络中，图像信息丢失是不可逆的。通过池化进行的下采样，会丢失一部分图像信息。深度学习分类的模型，只需要预测每一类的概率，不需要考虑池化导致损失图像细节信息的问题，但是做语义分割时，就要考虑到这个问题。池化虽然能够扩大感受野，但降低了分辨率。

### 2.2.2 膨胀卷积

膨胀卷积在标准的卷积中注入空洞，以此来增加感受野。膨胀卷积层可以聚合多尺度的信息，在特征图缩小到同样倍数的情况下可以获得更多关于图像的信息。相比原来的正常卷积，膨胀卷积多了一个超参数称之为膨胀卷积率，指的是卷积核的间隔数量。通过在卷积内核的像素之间插入“孔”的方式，它可以比普通的卷积获得更大的感受野。在不同的卷积层设置不同的膨胀卷积率，可以成倍的增加不同的感受野。

膨胀卷积的优点是可以以高空间分辨率计算特征，并且可以任意扩大感受野的大小。Yu 等人<sup>[5]</sup>使用序列化的层并以增加的扩张速率来实现上下文聚合。Chen 等<sup>[6]</sup>使用带有超采样内核的膨胀卷积进行语义图像分割。沿边界生成较为精准的密集预测和分割图，可以运用膨胀卷积，并且它已经被应用到更广泛的任务中，例如光流和音频生成。

栅格效应，是当前的膨胀卷积框架中存在一个固有的问题。Cottrell 等<sup>[33]</sup>减轻

了该问题，该问题是由标准的膨胀卷积运算引起的，膨胀卷积率间无公因子时，栅格效应得到减轻，如图 2-2 所示。

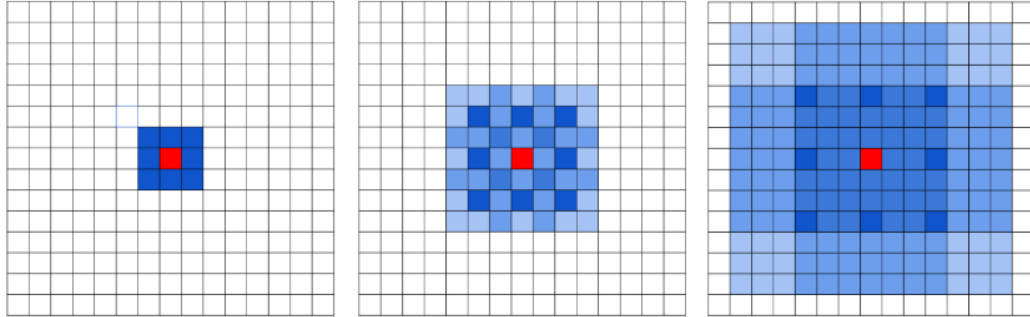


图 2-2 膨胀卷积去除栅格效应

### 2.2.3 无下采样语义分割神经网络

在逐像素语义分割任务中，大多数现有网络体系结构遵循完全卷积的下采样-上采样路径。为了进一步扩大感受野以捕获图像上下文信息，大多数现有的分割网络对中间特征图进行了下采样，这大大降低了分辨率。随后经常使用上采样层（例如反卷积）来部分恢复输入的高分辨率。但是，它无法完全恢复，并且添加反卷积层还引入了额外的计算成本。Li 等人<sup>[34]</sup>提出了一种架构，在语义分割领域是比较新颖的，如图 2-3 所示，他们设计了一种高分辨率且紧凑的网络结构，该架构在整个图层中都保持了高空间分辨率的特征图，没有进行任何下采样操作，用于在图像中语义分割精细的结构。

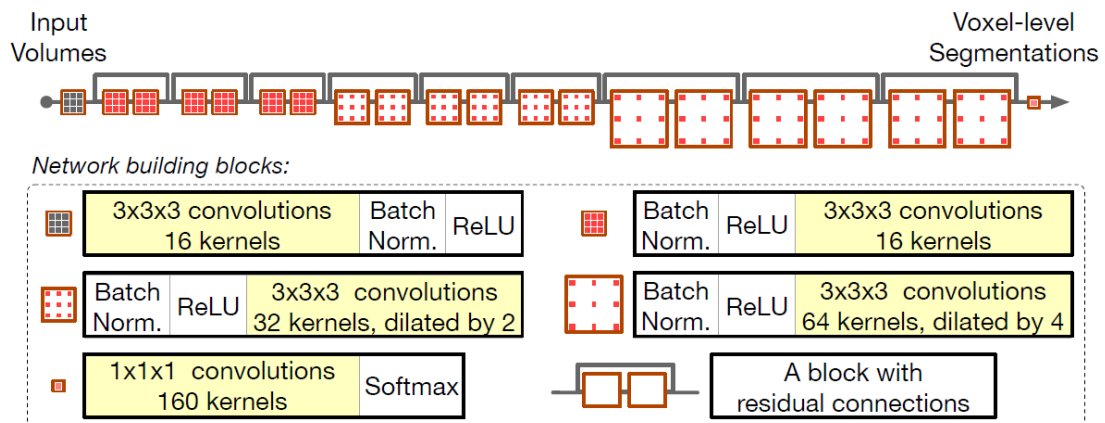


图 2-3 无下采样网络架构

## 2.3 去除镜面反射算法

为了对图片进行去反光（去除镜面反射）操作，Tan 等人<sup>[35]</sup>通过每个镜面反射像素的漫反射分量去除镜面反射，从而求得最大漫反射色度和像素颜色之间的非线性函数关系。

由 Dichromatic 反射模型<sup>错误!未找到引用源。</sup>可知，反射光线分为镜面反射和漫反射两个分量。

$$I(X) = w_d(X)D(X) + w_g(X)S(X) \quad (2-1)$$

式中  $I(X)$ ——像素的颜色向量；

$X$ ——像素坐标；

$w_d(X)$ ——像素的漫反射权重因子；

$w_g(X)$ ——像素的镜面反射权重因子；

$D(X)$ ——像素的漫反射分量；

$S(X)$ ——像素的镜面反射分量。

接下来对色度进行定义：

$$\begin{aligned} \sigma(X) &= \frac{I(X)}{I_r(X) + I_g(X) + I_b(X)} \\ \lambda(X) &= \frac{D(X)}{D_r(X) + D_g(X) + D_b(X)} \\ \gamma(X) &= \frac{S(X)}{S_r(X) + S_g(X) + S_b(X)} \end{aligned} \quad (2-2)$$

式中  $\sigma(X)$ ——与像素颜色相关的色度；

$\gamma(X)$ ——镜面反射色度；

$\lambda(X)$ ——漫反射色度。

接下来对最大色度进行定义：

$$\tilde{\sigma}(X) = \max\{\sigma_r(X), \sigma_g(X), \sigma_b(X)\} = \frac{\tilde{I}(X)}{I_r(X) + I_g(X) + I_b(X)} \quad (2-3)$$

式中  $\tilde{I}(X)$ ——为像素的颜色向量的最大值。

接下来，像素的漫反射分量即可通过下式求得：

$$m_d(X)\lambda(X) = I(X) - \frac{\tilde{\lambda}(X)(I_r(X) + I_g(X) + I_b(X)) - \tilde{I}(X)}{3\tilde{\lambda}(X) - 1} \quad (2-4)$$

式中  $\tilde{\lambda}(X)$ ——漫反射色度的最大元素。由上式可知，根据最大漫反射色度即可得到去除镜面反射后的图像，及去除反光点后的图像。

## 2.4 深度学习图像分类模型



2014 年 Karen 和 Andrew 创建 VGG Net，在 ISL VRC 上获得了第二好的成绩。VGG Net 包含 16 到 19 层的加权网络，比以前的网络体系结构更深，VGG Net 已经表明，它可以通过基于以前的网络体系结构增加网络层的数量和深度来提高网络性能。VGG Net 架构如图 2-4 所示。

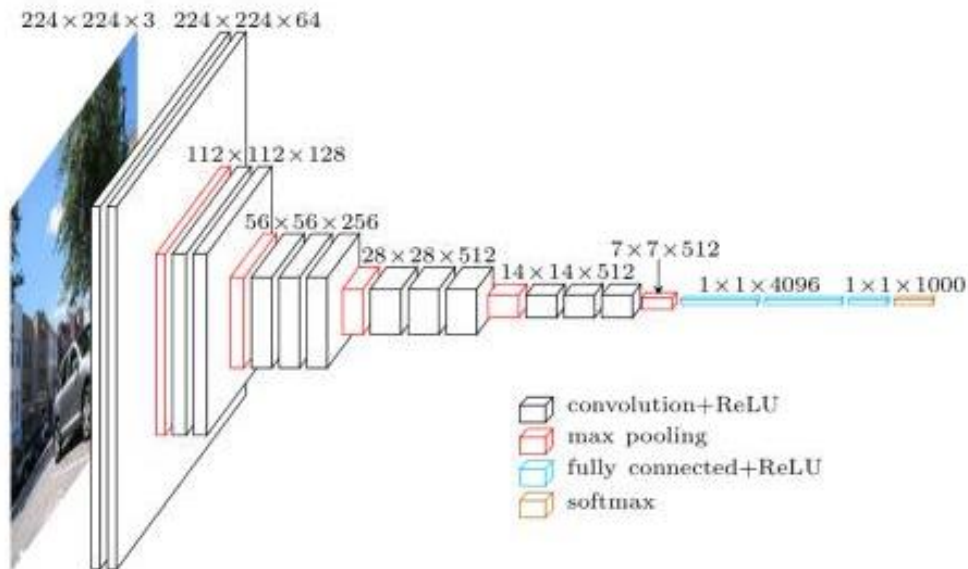


图 2-4 VGG Net 架构

之所以广受赞赏，是因为其结构简单。这很容易理解，但是有优化的空间。现在，其模型已广泛用于转移学习和其他需要训练的网络结构，例如大多数生成式对抗网络（Generative Adversarial Networks，GAN）。

何凯明等人提出 ResNet<sup>[37]</sup>，其残差单元如图 2-5 所示，该网络的优点是易于优化，并且通过增加网络层数大大提高了准确性。与以前的网络相比，这是一个更深入的残差网络学习框架。ResNet 以 3.57% 的性能超越了人类的感知水平，并创建了具有 152 层网络架构的新模型记录。

SqueezeNet 网络结构如图 2-6 所示，其核心结构为 FireModule 的组合形式。SqueezeNet 是 Han 等<sup>错误!未找到引用源。</sup>提出的一种轻量且高效的 CNN 模型，它参数比 AlexNet 少 50x，但模型性能与 AlexNet 接近。小模型相比大模型，具有很多优势：

（1）更高效的分布式训练，小模型参数小，网络通信量减少；（2）便于模型更新，模型小，客户端程序容易更新；（3）利于部署在特定硬件如 Field Programmable Gate Array（FPGA）。该网络能在保证不损失精度的同时，将原始 AlexNet 模型压缩至原来的 510 倍左右。

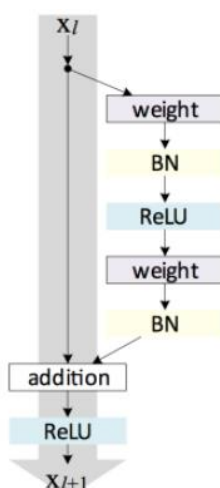


图 2-5 ResNet 架构中的残差单元

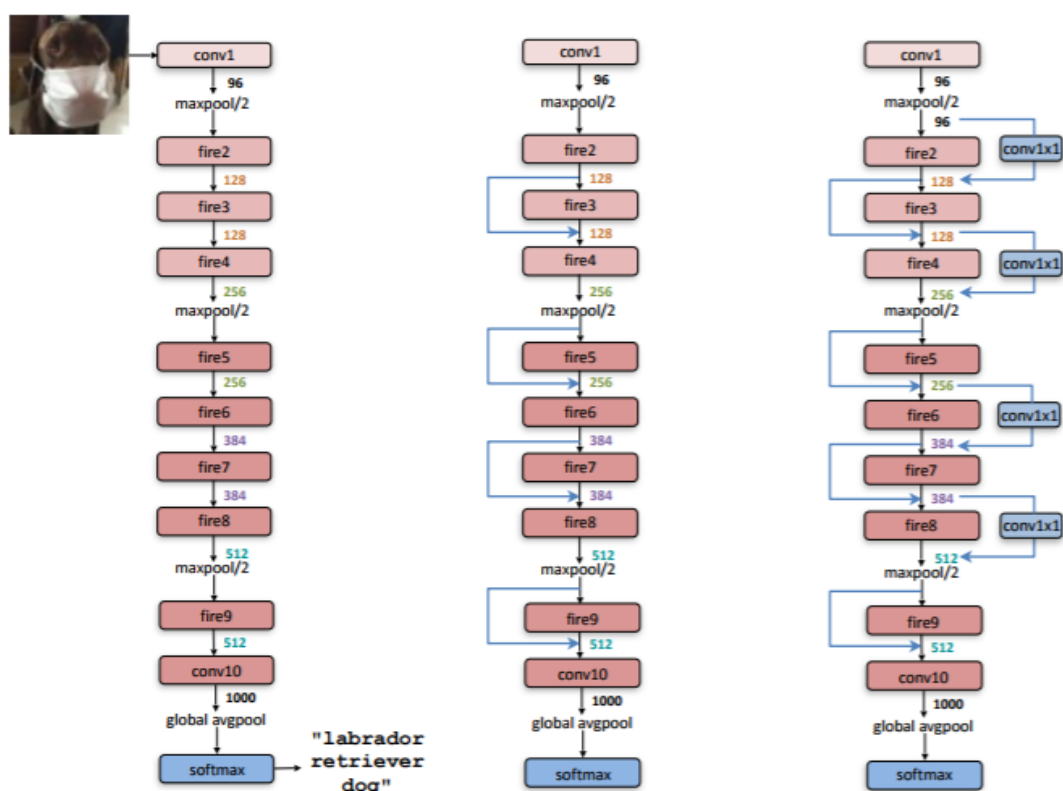


图 2-6 SqueezeNet 网络结构

2016 年，Gao 等人提出了密集卷积神经网络（DenseCNN）<sup>[39]</sup>概念，，如图 2-7 所示。该网络在前馈过程中将每一层与其他的层都链接起来。在网络的每一层，所有先前网络的功能图都用作输入，而特征图在其他网络层上也用作输入。

DenseCNN 具有许多优势，例如，缓解了梯度消失的问题，增强了特征传播和复用以及减少了参数数量。与 ResNet 相比，DenseNet 提供更好的性能，并只需要更少的内存和计算资源。

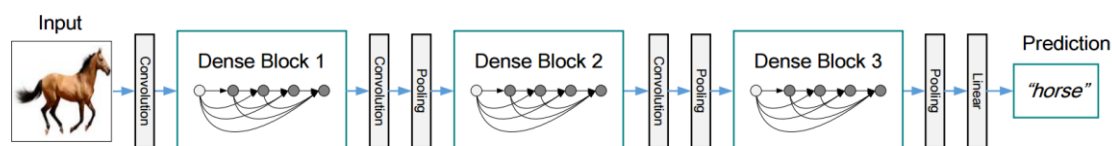


图 2-7 DenseNet 模型

## 2.5 本章小结

本章节对本文所采用的研究方法所涉及的相关研究进行了文献总结。在第 2.2 节描述了语义分割全卷积神经网络，并引入了无下采样的思路来构建模型，同时在扩大感受野的研究中，也考虑到了去除膨胀卷积栅格效应的问题，对模型构建中膨胀卷积率的设置做好了理论准备。在第 2.3 节中，详细介绍了去除镜面反射的算法；在第 2.4 节中，介绍了一些经典的深度学习分类网络。

基于舌下络脉小目标的特点进行医学图像语义分割模型的设计，如何对数据进行预处理和后处理，以及对分割出的舌下络脉进行颜色分类，就是接下来要介绍的本文的重点工作内容。

## 第3章 舌体及舌下络脉语义分割

### 3.1 引言

结合中医学理论、计算机相关的图像处理和深度学习技术，用客观的检测指标，对中医学的舌象进行定性、定量、定位的研究，以推动舌诊客观化的发展。舌体分割的目的，是为了减少图像中其他因素的干扰，从背景中把舌体分离出来。中医对舌下络脉的观察，同样是舌诊的重要部分，从原始图像中对舌下络脉进行提取，便于之后的疾病诊断、特征分析（如颜色特征分析），有着重要的意义。

为了实现舌下络脉的语义分割，需要解决三个方面的主要难题：（1）对现有数据进行处理，生成可以由深度学习模型训练的数据集；（2）由于舌下络脉占图像比例极小，属于小目标语义分割，因此基于上述数据集，实行感兴趣区域（Region of Interest, ROI）策略，先行提取舌体作为 ROI 区域；（3）进一步基于上述 ROI 区域，构建舌下络脉自动分割算法模型。针对上述问题，本文主要采用图像处理和数据增强相关技术进行数据预处理和数据后处理，设计基于深度学习的无下采样语义分割模型来实现舌下络脉的语义分割，最后对实验结果进行描述。

### 3.2 数据预处理方法

一般来说，由于深度学习数据集图像容易受到分辨率过高、数据量太少等问题的侵扰，因此在进行深度学习相关的研发工作之前，在进行语义分割工作前，还必须对其作进一步的数据预处理操作。数据的预处理属于课题研究的基础部分，高质量的数据集对于后续的语义分割有很大的帮助。第一步是数据集的生成，这里是指对图像舌体与舌下络脉标签的标注，并从标注文件中提取出相应信息。正如深度学习方法应用在生物学、地质学等领域一样，这个工作往往需要相关的专业知识。本研究通过与中医领域专家沟通合作，对原始图像数据进行了标注工作。

得到相应标签的标注信息后，由于原图像分辨率过高，模型难以训练。因此输入模型前再对数据集进行 Edge-padding 操作。不直接调整大小原因是避免舌体图片被拉伸，导致结构被破坏。

#### 3.2.1 数据集生成

为了得到深度学习语义分割可用的数据集，对采集的原始图像需要进行二重语义标注，即舌体与舌下络脉的语义标注。标注后，设计标签选择算法，根据舌体

和络脉的不同标签位置，对标注文件进行批量处理，分成舌体标注二值图像数据和舌下络脉标注二值图像数据，如图 3-1 所示。

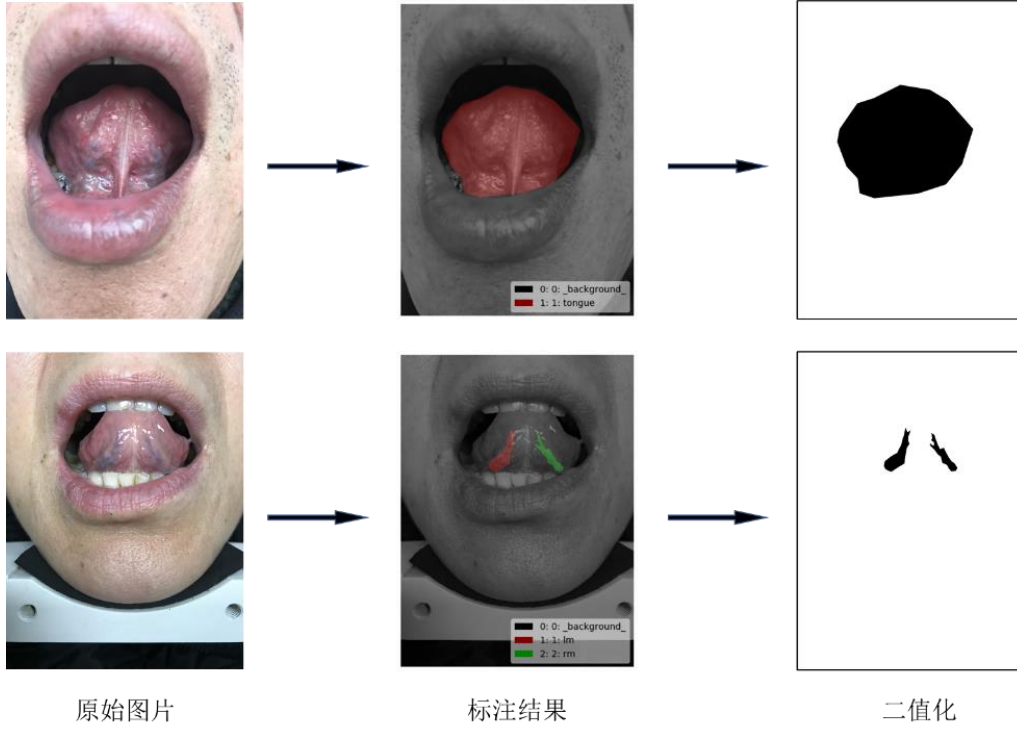


图 3-1 标注二值化数据示意图

得到相应 Label 的标注信息后，由于原图像较大，输入模型前再对数据集进行 Edge-padding 操作。

### 3.2.2 反光点去除算法

由于所给图像数据几乎都存在不同程度的反光点问题，影响舌体分割和舌下络脉分割，影响语义分割效果，如图 3-2 所示。

本文基于第二章中提到的去除镜面反射方法，进行了改进和简化，对数据集进行去除反光点处理。

去除镜面反射算法描述如下：求得图像  $RGB$  三通道值  $R, G, B$  与三通道中的最大值  $RGB_{max}$ ，再将图像  $RGB$  三通道转换为  $Alpha$  通道，并求得  $\alpha_r, \alpha_g, \alpha_b$  三通道值与三通道中的最小值  $\alpha_{min}$ 。根据  $\alpha_r, \alpha_g, \alpha_b$  三通道值，计算漫反射系数  $\beta_k$ ：

$$\beta_k = \frac{1 - (\alpha_{min} - \alpha_k)}{3 \times \alpha_{min} - 1} \quad (k = r, g, b) \quad (3-1)$$

再计算最大漫反射系数  $\beta_{max}$ ：

$$\beta_{max} = \max(\max(\beta_r, \beta_g), \beta_b) \quad (3-2)$$

根据 $\alpha_r, \alpha_g, \alpha_b$ 三通道值，计算 $\gamma_k$ ：

$$\gamma_k = \frac{\alpha_k - \alpha_{\min}}{1 - 3 \times \alpha_{\min}} \quad (k = r, g, b) \quad (3-3)$$



图 3-2 反光点影响语义分割效果

再计算最大系数 $\gamma_{\max}$ ：

$$\gamma_{\max} = \max(\max(\gamma_r, \gamma_g), \gamma_b) \quad (3-4)$$

根据 $\gamma_{\max}$ 计算镜面反射因子 $\xi$ ：

$$\xi = \frac{\gamma_{\max} \times (R + G + B) - RGB_{\max}}{3 \times \gamma_{\max} - 1} \quad (k = r, g, b) \quad (3-5)$$

最终可求得去除反光点后的三通道值 $R_{\text{new}}, G_{\text{new}}, B_{\text{new}}$ 为：

$$\begin{aligned} R_{\text{new}} &= R - (\xi + 0.5) \\ G_{\text{new}} &= G - (\xi + 0.5) \\ B_{\text{new}} &= B - (\xi + 0.5) \end{aligned} \quad (3-6)$$

去除反光点图像效果前后对比如图 3-3 所示。

### 3.2.3 数据扩充

由于深度学习模型需要充足的数据，可能才会有更好的表现，因此还需要对图像进行数据增强，来扩大数据集的规模。在去除反光点后的图像上，我们使用了数据扩充的一些方法，将数据集扩充了五倍。当前数据增强技术有很多，由于考虑到数据集涉及到语义分割的问题，因此不能采用随机裁剪、镜头拉伸等方法；还涉及到颜色分类的问题，因此不能采用对比度变换、彩色转换、颜色失真、像素色彩抖动等方法。



图 3-3 去除反光点效果示意

最终数据扩充采用了垂直镜像对称，水平、垂直翻转的方法，将训练集进行扩充。特别地，要将图像数据和标记数据进行同步扩充。

### 3.3 基于全卷积网络的语义分割算法模型

对于该问题，基于仅由卷积层构成的神经网络（FCN）进行设计语义分割网络。为了利用图像的全局信息，我们需要让神经网络在空间上进行更大范围的“交流”。池化层和全连接层可以达到这一目的，但是，池化层会降低空间精度。图像分割预测是逐像素输出，因此应将合并后的小图像尺寸上采样到原始图像尺寸以进行预测。

因此，图像分割 FCN 具有两个重要部分。一是池化以减小图像大小以扩展感知域，二是向上采样以增加图像大小。先减小然后增大大小的过程会丢失内部结构和空间层次结构信息，并且无法为小型对象重建信息（假设有 4 个池化层，物体信息为 16 个像素，如果小于 16 像素则理论上无法复原）。

舌下络脉宽度仅占图片的几个像素，因此需要避免对图片再在空间上降维。最终选择不采用池化层，使用膨胀卷积，高效地扩张感受野。通过膨胀卷积聚合多尺度的信息，扩大感知域的同时不会降低分辨率。

#### 3.3.1 语义分割模型

采用基于膨胀卷积的深度学习网络构建了语义分割模型，这也是整个算法的核心部件。模型设计如下图 3-4 所示，具体参数如表 3-1 所示。

通过采用 Batch Normalization 的方法，可以降低模型对参数的敏感度，使网络学习过程更加稳定。

为了规范数据，让输入数据尽可能具有独立同分布的特征，即输入特征分布具有相同的方差和均值，且特征之间彼此独立，没有相关性，机器学习领域往往采用白化的处理方法。但由于白化的计算成本太大，而且会丢失底层网络学习的参数信



息，模型加入 Batch Normalization 来解决这一问题。与 Layer Normalization 不同，Batch Normalization 是一种纵向的规范化方法，它与训练批次的大小有关，更适用于 CNN 网络。

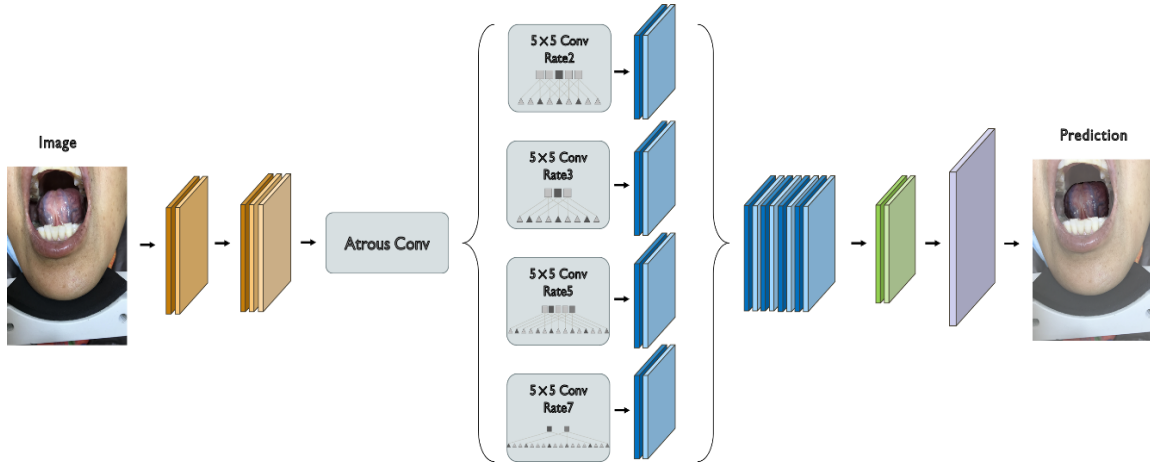


图 3-4 语义分割模型

表 3-1 语义分割模型参数设置

Layer	Num of Kernel	Kernel Size	Activation	Dilation Rate_ veins	Dilation Rate_ tongue
1	32	5*5	ReLU	/	/
2	32	5*5	ReLU	/	/
3	32	5*5	ReLU	(2,2)	(2,2)
4	32	5*5	ReLU	(3,3)	(5,5)
5	32	5*5	ReLU	(5,5)	(7,7)
6	32	5*5	ReLU	(7,7)	(11,11)
7	32	1*1	ReLU	/	/
8	1	1*1	/	/	/

### 3.3.1 小目标语义分割策略

舌下络脉属于典型的小样本目标，对于小目标语义分割，可以选取先定位缩小目标区域再分割的策略。络脉的语义分割问题属于典型的小目标分割，样本极不平衡。因此，将舌体的分割结果作为模型的输入,再进行舌下络脉的语义分割，如下图 3-5 所示。

我们先进行舌体的语义分割。舌体的语义分割模型结构图见图 3-6。舌体的语义分割采用全卷积神经网络，并加入膨胀卷积扩大感受野。舌体语义分割采用由八层卷积层构成的桶形网络，卷积核大小为 5\*5，激活函数采用 ReLU 函数。



其中，膨胀卷积率组合采用（2，2）（5，5）（7，7）（11，11）这一组合，基于舌体这一目标分割物体所占图片整体像素较多，因此采用较大的膨胀卷积率组合，同时这一设置也避免了膨胀卷积的栅格效应。

舌体的语义分割结果可作为络脉的输入。在训练过程中，模型以原始图像和语义标注图像作为输入，输出预测结果。测试时，模型仅需要原始图像，可输出对应的预测结果。为形象化地展示，将模型预测的二值化图像与原图进行叠加，从而得到最终的效果图。

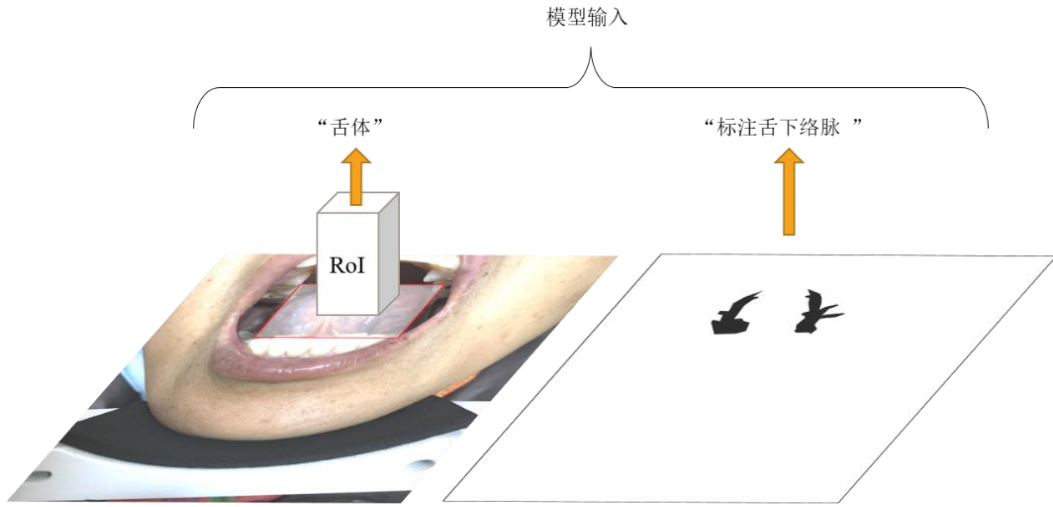


图 3-5 小样本语义分割策略

根据深度学习语义分割模型对舌体的语义分割输出结果，裁剪目标图像块，再进行舌下络脉的语义分割。舌体分割结果作为定位，缩小了舌下络脉的分割目标区域。基于该小样本语义分割策略，我们将舌体和舌下络脉标注作为舌下络脉语义分割模型的输入。舌下络脉语义分割算法的总体流程如图 3-7 所示。将提取出的舌体，作为络脉语义分割的数据输入，再进行络脉语义分割模型训练，比原图直接输入模型训练络脉语义分割，效果有了进一步的提升。

### 3.3.2 损失函数

默认的损失函数易将所有像素预测为非络脉。因此应该让网络对于正样本部分给予更多的“注意”，一定程度上提高模型的表现。对 Hinge 损失函数进行修改：

$$L_{hinge}^* = \begin{cases} \alpha \cdot \max\{0, 1 - \hat{y}\}, & y = 1 \\ \max\{0, \hat{y} + 1\}, & y = -1 \end{cases} \quad (3-7)$$

惩罚权重 $\alpha$ 大于 1，有利于提高模型的表现，提升 F1-score，经过实验，如图 3-8 可看出 Hinge 损失函数变化， $\alpha = 2$  时训练效果最好，且有最高的 F1-score，比原有

的模型 F1 值提升了 4-5 个点。

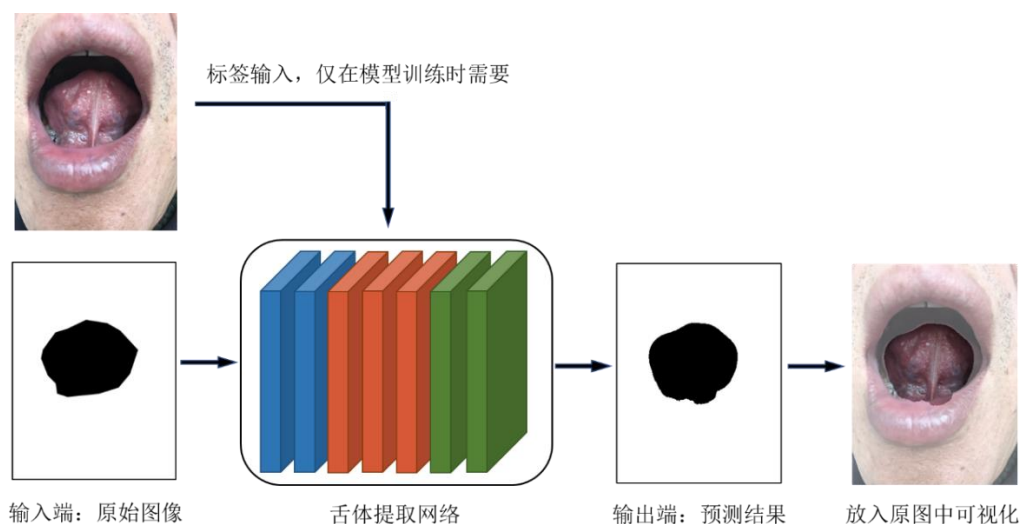


图 3-6 舌体语义分割算法总体流程

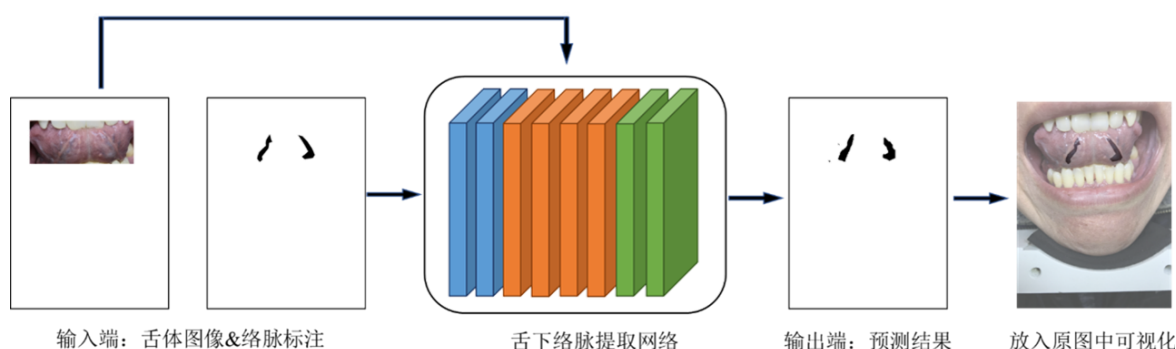


图 3-7 舌下络脉语义分割算法总体流程

### 3.4 数据后处理

模型预测输出的图像数据与原 Edge-padding 后输入模型的数据尺寸一致，需将模型输出恢复至原图数据的分辨率。因此对模型输出结果进行 Edge-clipping 操作。在得到语义分割的二值化图像后，需要进行一些后处理来提升分割效果。

#### 3.4.1 去除孔洞和杂散点

由于模型预测输出的舌体二值化图像中，个别图像存在一些问题，如外部有杂散点，内部有孔洞等问题。这将影响舌下络脉的语义分割效果、甚至医生的诊断。因此我们对模型预测出的舌体图像进行数据后处理，去除二值化分割图像中

的杂散点和孔洞，处理如图 3-9 所示。

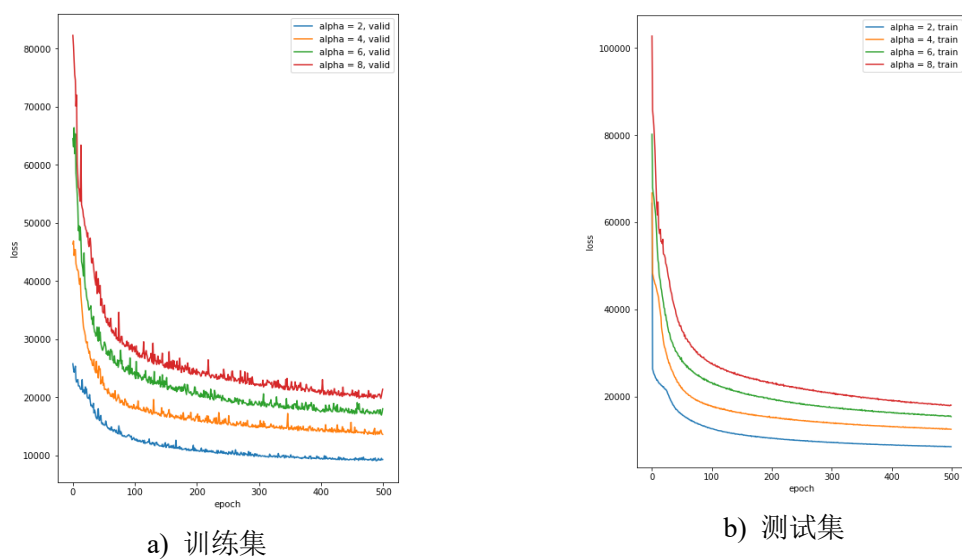


图 3-8 损失函数不同惩罚权重实验结果

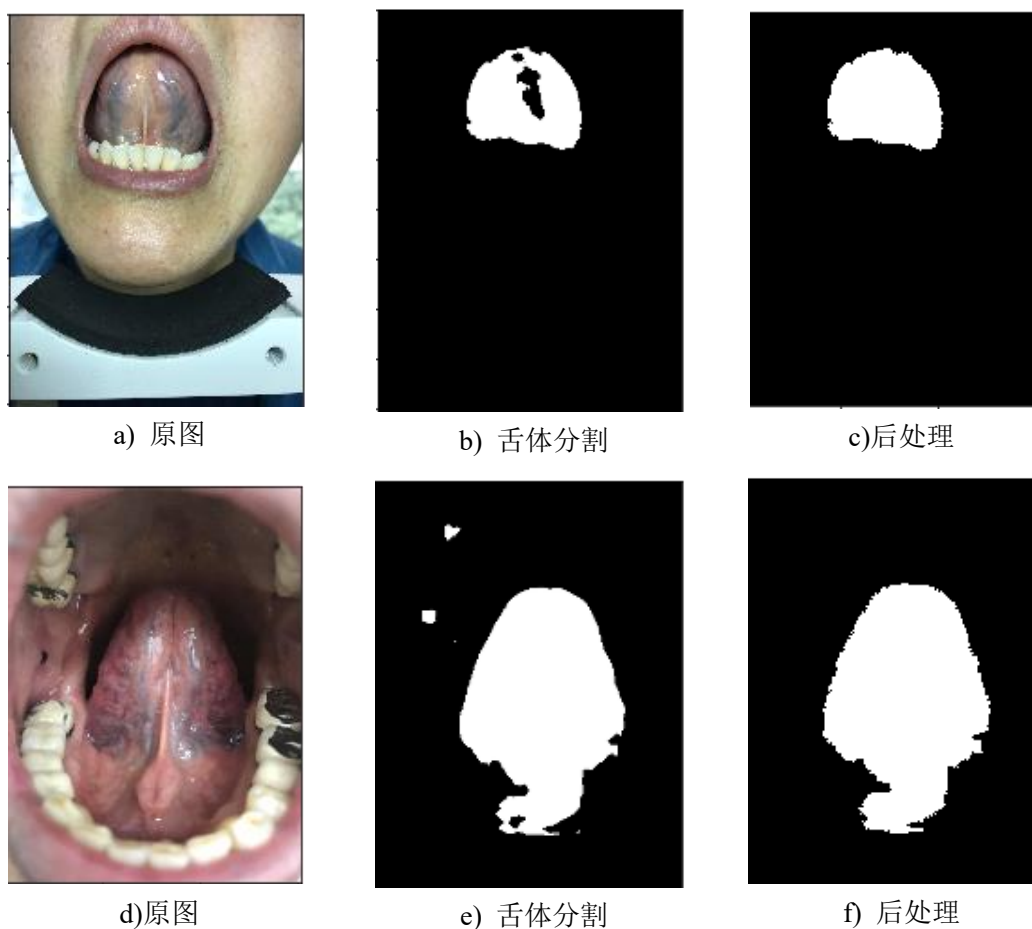


图 3-9 部分图像舌体后处理

### 3.4.2 分割边缘的膨胀与腐蚀

由于模型预测输出的二值化图像边缘存在锯齿现象，因此本文设计算法对预测图像进行后处理。对预测输出的二值化图像进行腐蚀和膨胀，去除边缘的毛刺锯齿现象。

经处理后，如图 3-10 所示，语义分割目标的边缘消除了锯齿现象，语义分割对应的几个评价指标也有了一定的增长，效果进一步提升。

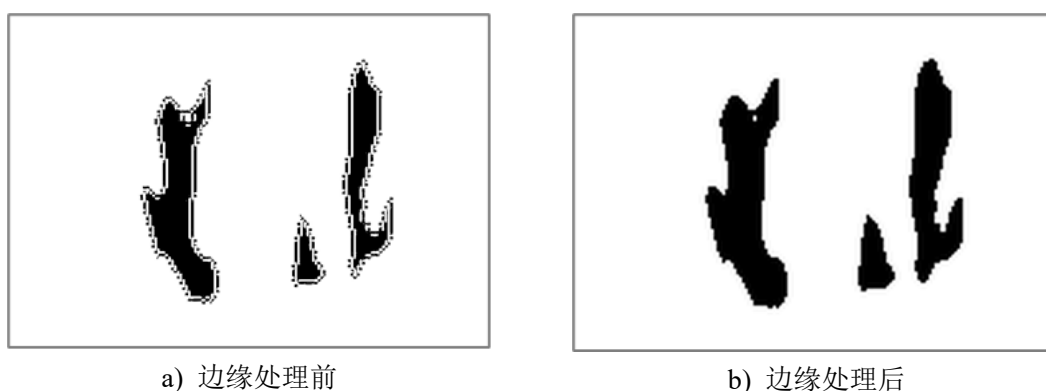


图 3-10 舌下络脉边缘图像后处理

## 3.5 实验结果

在这一部分将对舌体及舌下络脉的语义分割进行实验。在第 3.5.1 节中，对实验所用到的数据集进行描述；在第 3.5.2 节中，对实验中用到的语义分割评价指标进行介绍；在后续的第 3.5.3 节、第 3.5.4 节和第 3.5.5 节中，我们将从评价指标、与其他模型效果对比、分割效果图方面对舌体语义分割和舌下络脉语义分割效果进行评估。

### 3.5.1 数据集描述

数据集来源与上海中医药大学。语义分割原始数据为 322 张 24 位舌部图像，包括一些背景信息，分辨率为 3024\*4032，数据采集标准由上海中医药大学制定，并根据标准进行采集。我们与中医领域专家沟通合作，对原始图像数据进行了标注工作。最终数据集为原图像、舌体标注二值化图像、舌下络脉标注二值化图像各 322 张。我们将数据集划分训练集和测试集，比例为 8:2。特别注意的是，进行数据增强时，要将训练集的图像数据和标记数据进行同步扩充。

我们利用以 Tensorflow 实现的语义分割模型作为神经网络的主体框架。训练迭代次数为 350 次，学习率 learning\_rate 为 2。

### 3.5.2 语义分割评价指标

对于语义分割模型的评价指标,我们利用以下性能评价指标:F1 值(F1-score)、受试者工作特征曲线(Receiver Operating Characteristic, ROC)、ROC 曲线下的面积(Area Under Curve, AUC)、平均交并比(Mean Intersection over Union, mIoU)对结果进行评定。

我们利用常见评估指标:F1-score 评估模型的性能。F1-score 的计算方式略为复杂,涉及到其他的评估指标,在该实验中是对每一个像素点进行分类,我们计算方式如下:

$$F_1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3-8)$$

式中 *Precision*——精确率;

*Recall*——召回率。

F1-score 通过计算正类(即语义分割目标)的精确率和召回率的调和平均值得到。F1-score 能够综合考虑精确率和召回率两个评估指标,反应整体性能。精确率和召回率的计算方式如下:

$$Precision = \frac{TP}{TP + FP} \quad (3-9)$$

$$Recall = \frac{TP}{TP + FN} \quad (3-10)$$

式中 *TP*——预测值为正类(语义分割目标)、真实值为正类的样本数量;

*FP*——预测值为正类、真实值为负类(其他背景像素)的样本数量;

*FN*——预测值为负类、真实值为正类的样本数量。

精确率表示预测为正的样本中有多少是真正的正样本,即图像中每一像素模型判断是否是舌体/舌下络脉,等于提取出的准确像素数/提取出的像素总数。召回率表示样本中的正例有多少被预测正确,即标注中的舌体/舌下络脉部分像素有多少被预测正确。

采用曲线对模型语义分割效果进行判断。曲线是判断二分类模型的主要评价指标,能够不受样本不均衡影响。由于该语义分割任务存在明显的样本不均衡问题,因此引入曲线和值进行评估是必要的。该任务可视作对图像每个像素进行二分类(像素是否为舌体/舌下络脉),曲线越靠拢(0,1)点,越偏离 45 度对角线越好。AUC 值为曲线下的面积,越大越好,1 为理想状态。

我们还采用语义分割领域常用指标值,mIoU 即平均交并比来对模型语义分割进行评估。mIoU 值计算真实值和预测值两个集合的交集和并集之比。

$$mIoU = \frac{TP}{TP + FP + FN} \quad (3-11)$$

式中  $TP$ ——预测值为正类（语义分割目标）、真实值为正类的样本数量；  
 $FP$ ——预测值为正类、真实值为负类（其他背景像素）的样本数量；  
 $FN$ ——预测值为负类、真实值为正类的样本数量。

### 3.5.3 评价指标结果

我们在这一节中将从评价指标方面对舌体语义分割和舌下络脉语义分割效果进行评估。

#### 3.5.3.1 舌体语义分割

我们设置舌体语义分割实验为未去除反光点，去除反光点，去除反光点+数据扩充，去除反光点+数据扩充+后处理去除杂散点和内部孔洞四项实验，并保存模型进行评价指标对比。

首先未去除反光点，将数据集输入网络中进行训练。再将原始图像进行去除反光点处理后，将图像输入网络中进行训练得到模型。再对去除反光点后的训练集图像进行数据扩充，同时也对对应的舌体二值化图像进行同步扩充，再输入网络中进行训练。最终，在前面实验的基础上，对输出的图像进行去除杂散点和内部孔洞的图像后处理，并计算评价指标结果。如下表 3-2 所示，为舌体语义分割在测试集上的性能指标结果，如图 3-11 为对应的 ROC 曲线。

表 3-2 性能指标评价结果对比（舌体语义分割\_测试集）

任务	F1 值	精确率	召回率	AUC 值	mIoU
未去除反光点	0.9055	0.9324	0.8804	0.9853	0.8591
去除反光点	0.9190	0.9426	0.8965	0.9896	0.8501
去除反光点+数据扩充	0.9387	0.9608	0.9174	0.9979	0.8843
除反光+扩充+后处理	<b>0.9473</b>	<b>0.9559</b>	<b>0.9390</b>	<b>0.9984</b>	<b>0.9000</b>

#### 3.5.3.2 舌下络脉语义分割

我们设置舌下络脉语义分割实验为未去除反光点，去除反光点，去除反光点+语义分割舌体结果作为输入，去除反光点+语义分割舌体结果作为输入+数据扩充，去除反光点+语义分割舌体结果作为输入+数据扩充+后处理边缘膨胀腐蚀五项实验，并保存模型进行评价指标对比。

首先未去除反光点，将数据集输入网络中进行训练。再将原始图像进行去除反光点处理后，将图像输入网络中进行训练得到模型。其次，在如上基础下，将原图像替换为提取出的舌体图像作为网络的输入，训练模型。再引入数据扩充，对提取出的舌体图像进行数据扩充，同时也对对应的舌下络脉二值化图像进行同步扩充，

输入网络中进行训练。最终，在前面实验的基础上，对输出的图像进行边缘腐蚀膨胀的图像后处理，并计算评价指标结果。如下表 3-3 所示，为舌体语义分割在测试集上的性能指标结果，如图 3-12 为对应的 ROC 曲线。

### 3.5.4 与其他模型实验对比

对比的研究方法为经典 U-net 模型，及 Yang 等人对 U-net 模型进行修改后的对于舌下络脉语义分割的模型。我们对这两种模型进行了复现，并在同一数据集上进行了实验。为了对比模型本身的语义分割能力，我们除了上文提到的最优效果外，设置无数据预处理和后处理、原图像模型直接输入的实验，与经典和改进 U-net 模型进行对比。

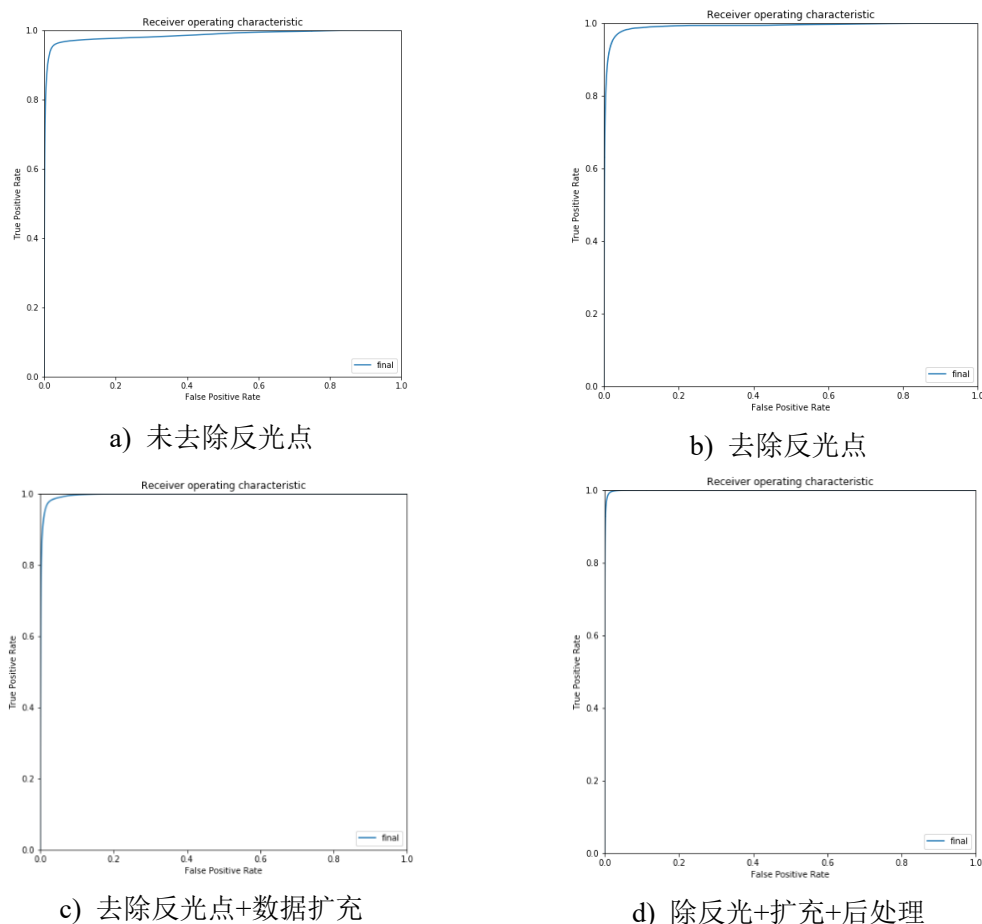


图 3-11 ROC 曲线

如表 3-4 所示，可看出模型本身效果对于该问题就优于经典 U-net 模型及其改进后的模型，更适用于该医学语义分割问题。经图像预处理、后处理及小目标分割策略处理后，获得模型最优效果。

表 3-3 性能指标评价结果对比（舌下络脉语义分割\_测试集）

任务	F1 值	精确率	召回率	AUC 值	mIoU
未去除反光点	0.6941	0.7628	0.6368	0.9811	0.5315
去除反光点	0.7013	0.7665	0.6463	0.9802	0.5400
去除反光点+舌体输入	0.7131	0.7693	0.6646	0.9872	0.5541
去除反光点+舌体输入+数据扩充	0.7283	0.7708	0.6908	0.9893	0.5731
除反光+舌体输入+扩充+膨胀腐蚀	<b>0.7383</b>	<b>0.7784</b>	<b>0.7021</b>	<b>0.9901</b>	<b>0.5851</b>

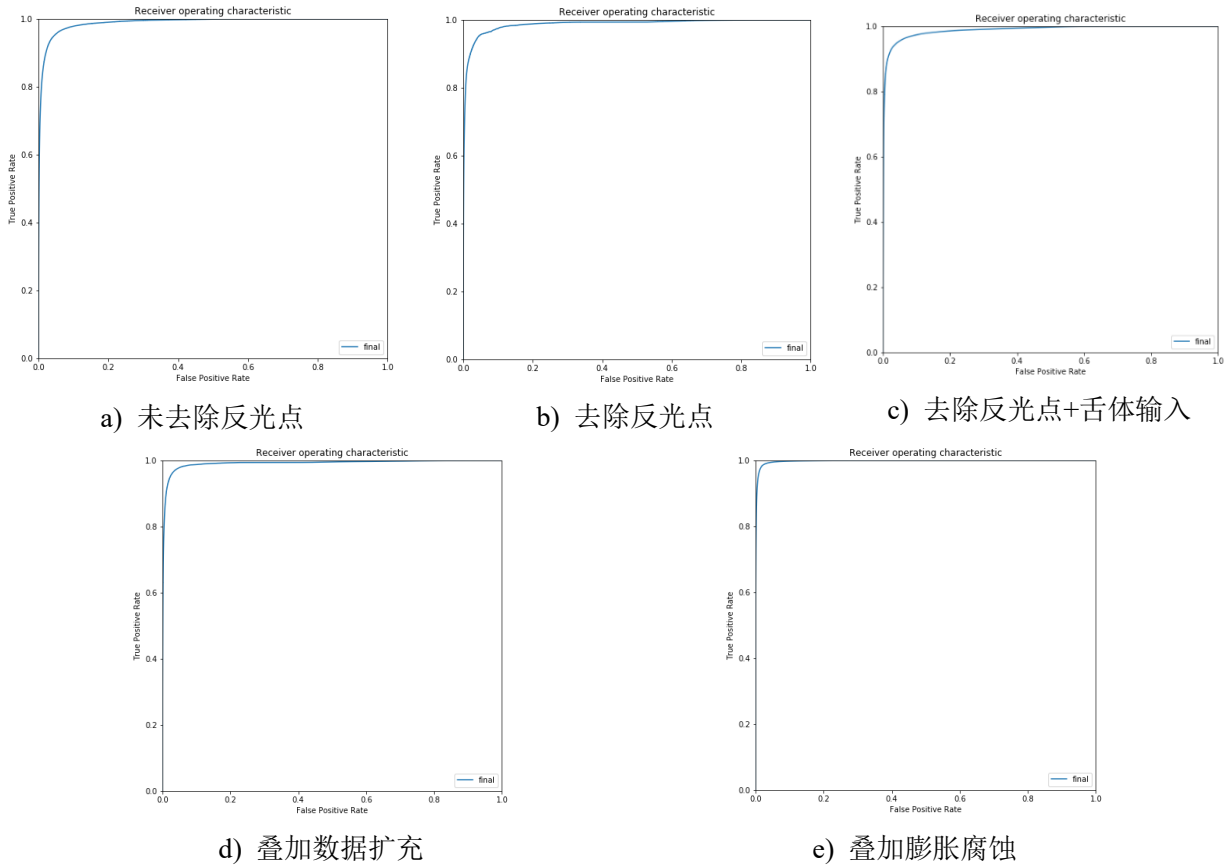


图 3-12 ROC 曲线

表 3-4 性能指标评价结果对比（舌体语义分割\_测试集）

模型	舌体分割		络脉分割	
	mIoU	F1-score	mIoU	F1-score
经典 U-net	0.8425	0.8841	0.5068	0.6673
改进 U-net	0.8390	0.8703	0.5112	0.6754
本研究模型无处理	0.8594	0.9032	0.5323	0.6931
本研究最佳效果	<b>0.9000</b>	<b>0.9473</b>	<b>0.5851</b>	<b>0.7383</b>

### 3.5.5 语义分割图像效果展示

我们在这一节中将从语义分割图像效果方面对舌体语义分割和舌下络脉语义



分割效果进行评估。

### 3.5.5.1 舌体语义分割

如图 3-13 所示，随机选取两张测试集图像进行舌体语义分割效果展示。上图为标注结果，下图为预测结果，可看出舌体语义分割效果较为精准，模型可较好的完成舌体语义分割任务。

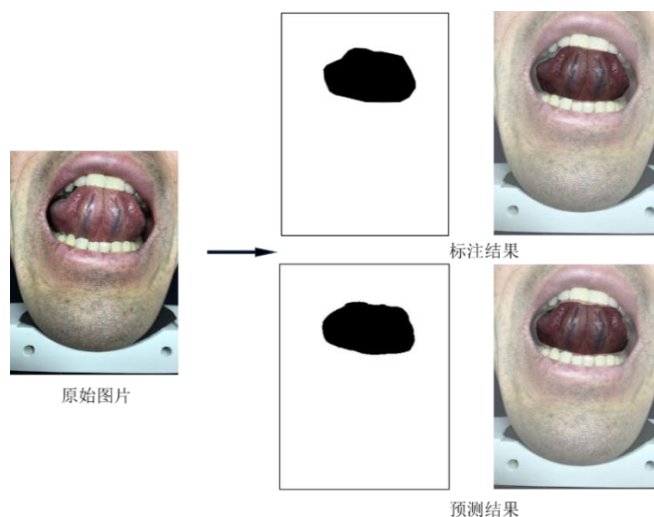


图 3-13 舌体语义分割效果示意图

### 3.5.5.2 舌下络脉语义分割

如图 3-14 所示，随机选取两张测试集图像进行舌下络脉语义分割效果展示。上图为标注结果，下图为预测结果，可看出舌下络脉语义分割效果基本较为精准，可分割出一些络脉的细节处，模型可较好的完成舌下络脉语义分割任务。

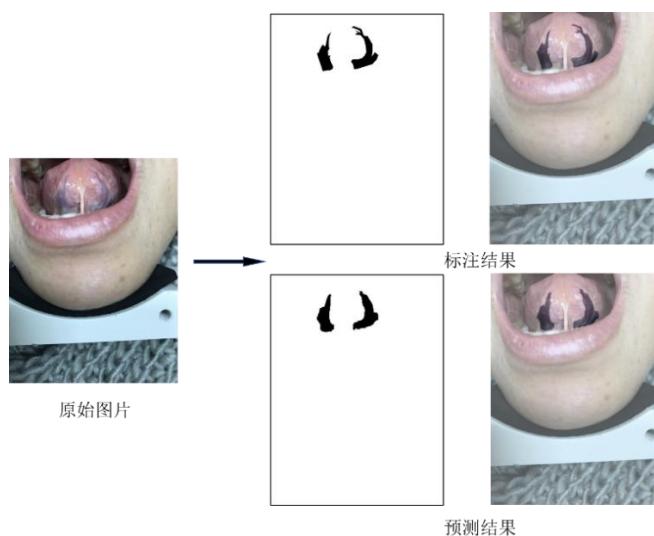


图 3-14 舌下络脉语义分割效果示意图

### 3.6 本章小结

本章节详细叙述了对舌体及舌下络脉语义分割的研究过程。全面地描述了数据集的处理过程，并且结合数据集的特征主要采用去除镜面反射和数据增强的方法进行了预处理工作。在处理后的数据上，设计无下采样语义分割模型，采用膨胀卷积。基于舌下络脉小目标特性，设计了基于小目标的语义分割策略，效果得到进一步提升。对数据进行后处理，得到了较好的语义分割效果。最后对实验结果进行了描述，实验结果表明，我们的模型在舌体及舌下络脉语义分割问题上表现优于经典的 U-net 模型，进行优化后得到了更好的效果。

## 第 4 章 舌下络脉颜色分类

### 4.1 引言

如上一章所述，提取出舌下络脉后，需对络脉进行颜色分类。本章节将基于语义分割后的络脉，设计不同的输入分类模型的络脉提取特征方案。在第 2.4 节中已经介绍了基于深度学习的经典分类模型，目前常常通过 ResNet、VGG Net、DenseNet 等模型进行分类。本章节将实验不同的深度学习模型，并对其进行调整，以得到最佳的分类效果。

接下来，本章的结构如下：第 4.2 节将对输入分类模型的络脉提取方案进行描述；第 4.3 节介绍实验的深度学习模型，并对效果最优的深度学习模型进行调整；在第 4.4 节对实验结果进行描述；最后在第 4.5 节将会对本章节进行总结。

### 4.2 舌下络脉输入分类模型方法

模型颜色分类受到输入图像络脉占比、络脉颜色信息多少和络脉形状等因素的干扰。考虑到干扰因素，设计了不同舌下络脉输入分类模型的方法，进行实验来检测效果。

原图尺寸提取络脉，如图 4-1 所示，输入深度学习模型，这种方法没有刨除掉络脉形状的干扰，同时图像络脉占比太少，该方法实验分类效果不佳。

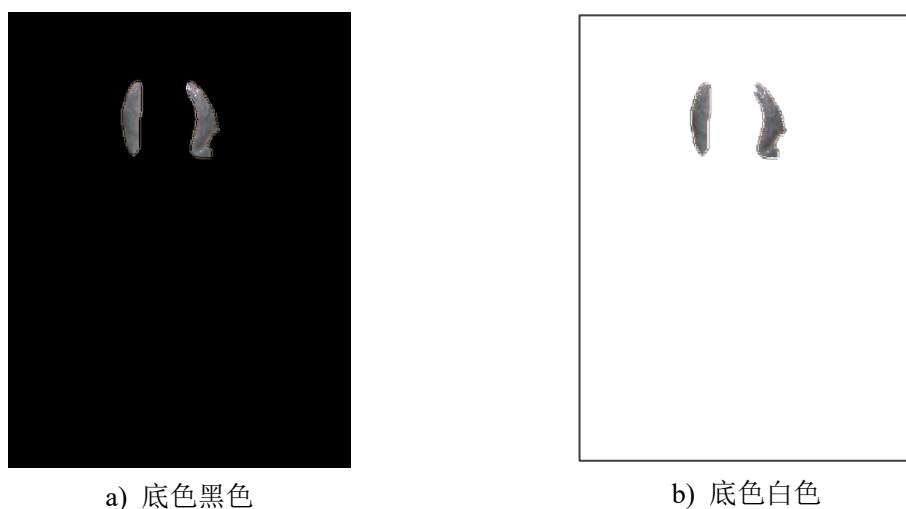


图 4-1 原图尺寸提取络脉

设计双络脉矩形提取络脉，如图 4-2 所示，输入深度学习模型，这种方法没有

刨除掉络脉形状的干扰，但颜色信息并未丢失，且图像络脉占比高。

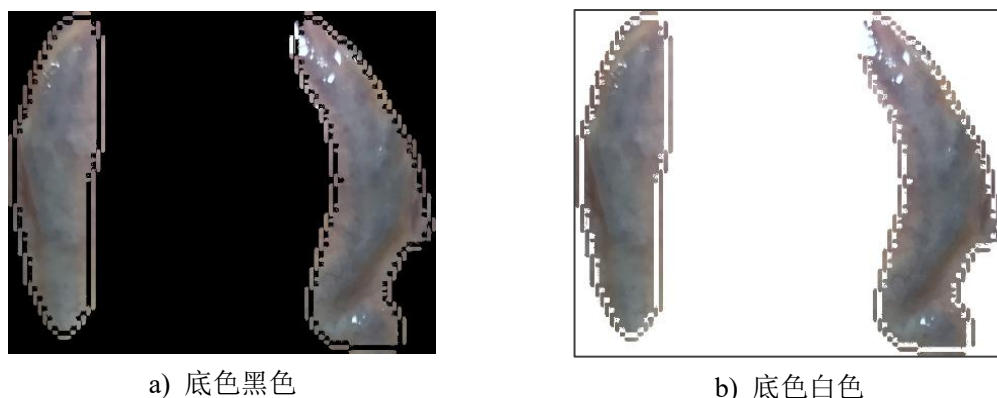


图 4-2 双络脉矩形提取络脉

设计单络脉矩形，如图 4-3 所示，输入深度学习模型，这种方法减少了络脉形状的干扰，但是丢失了较多颜色信息，实验效果不如双络脉，由此考虑影响模型颜色分类的因素中，络脉形状的干扰因素对颜色分类的影响弱于输入图像中的颜色信息量的因素。

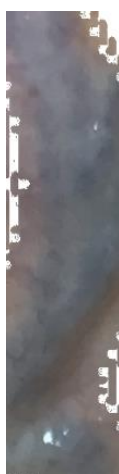


图 4-3 单络脉矩形提取络脉

设计提取 100\*100 颜色块方法，如图 4-4 所示，输入深度学习模型，选择颜色块代表不受形状干扰，但络脉颜色信息大大减少。颜色块实验有一个普遍的问题，即训练集效果不如测试集，说明模型无法充分学习到颜色信息，输入 100\*100 颜色块方法不适用于络脉颜色分类。



图 4-4 100\*100 颜色块提取络脉

从双络脉矩形对比单络脉矩形和颜色块的方法颜色分类效果看，颜色信息减少受影响更大些，而双络脉矩形这一方案效果最佳。

最后设计在原图上根据络脉坐标来切割出络脉存在的原图部分，如图 4-5 所示，作为实验的对照部分，输入深度学习模型。



图 4-5 原图尺寸提取络脉

### 4.3 基于深度学习的分类模型

由第 2.4 节中所介绍的基于深度学习的经典分类模型，我们实现了 19 种深度学习分类模型来进行络脉颜色分类。

VGGNet 包含 16 到 19 层的加权网络，比以前的网络体系结构更深，VGGNet 已经表明，它可以通过基于以前的网络体系结构增加网络层的数量和深度来提高网络性能。其模型已广泛用于转移学习和其他需要训练的网络结构。这一系列模型在舌下络脉颜色分类上效果不佳，弱于其他分类模型。

ResNet 与以前的网络相比，这是一个更深入的残差网络学习框架。该网络的优点是易于优化，并且通过增加网络层数大大提高了准确性。SqueezeNet 模型，具有更高效的分布式训练，小模型参数小，网络通信量减少。其属于 AlexNet 级的准确度，但模型尺寸小。ResNet 系列与 SqueezeNet 系列分类模型效果相仿。

如图 4-6 所示，DenseCNN 将每一层与其他的层都链接起来，在网络的每一层，所有先前网络的功能图都用作输入，与 ResNet 相比，DenseNet 提供更好的性能，并只需要更少的内存和计算资源。对于颜色分类，DenseNet 系列模型效果也普遍好于其余系列模型。Dense block 内部结构如图 4-7 所示，其密集连接的特点，使得网络脱离了加深网络层数和加宽网络结构来提升网络性能的定式思维，从特征的角度考虑，通过特征重用，既大幅度减少了网络的参数量，又在一定程度上缓解了梯度消失问题的产生。

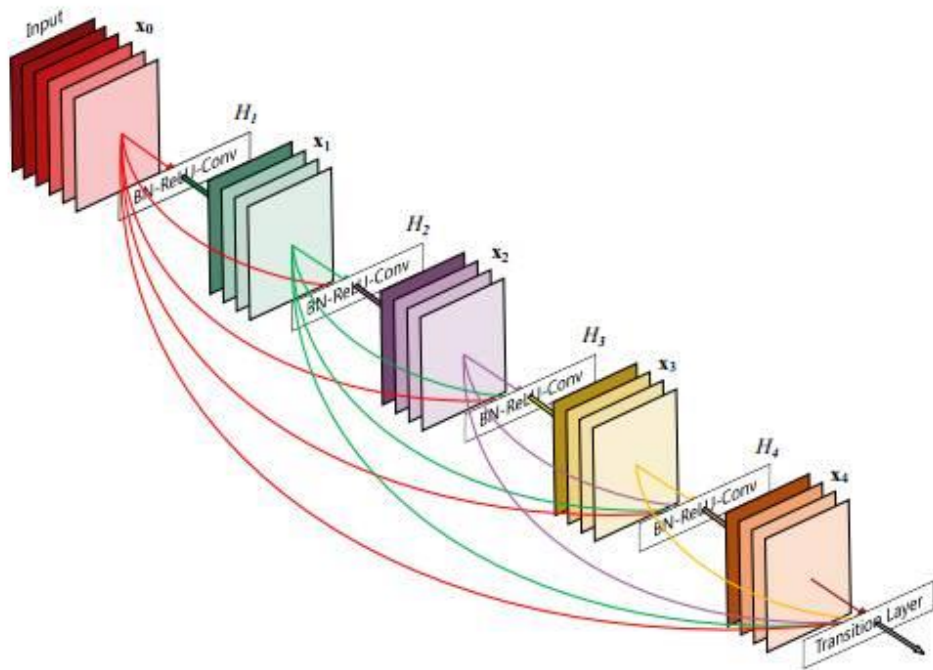


图 4-6 五层的致密网络，每一层将所有层都视为自己的输入

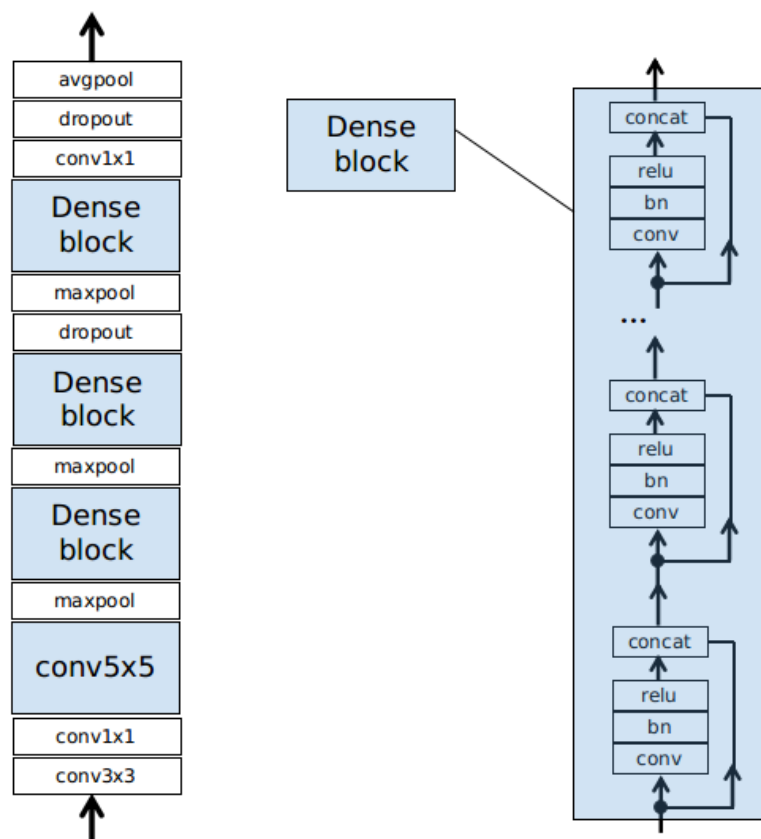


图 4-7 Dense block 内部结构

DenseNet 网络结构如表 4-1 所示，其中 conv 的结构是 BN-ReLU-Conv。在 DenseNet 系列模型中，Densenet-161 对络脉颜色分类效果最佳。

表 4-1 网络具体结构

Layers	Output Size	DenseNet-121	DenseNet-161	DenseNet-169	DenseNet-201
Convolution	112×112	7×7 conv, stride2			
Pooling	56×56	3×3 max pool, stride2			
Dense Block (1)	56×56	$6 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$6 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$6 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$6 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$
Transiton Layer (1)	56×56 28×28	1×1 conv 2×2 average pool, stride2			
Dense Block (2)	28×28	$12 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$12 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$12 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$12 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$
Transiton Layer (2)	28×28 14×14	1×1 conv 2×2 average pool, stride2			
Dense Block (3)	14×14	$24 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$32 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$32 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$48 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$
Transiton Layer (3)	14×14 7×7	1×1 conv 2×2 average pool, stride2			
Dense Block (4)	7×7	$24 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$32 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$32 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$	$48 * \begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix}$
Classification Layer	1×1	7×7 global average pool 1000D fully-connected, softmax			

在此基础上，对模型加入了瓶颈层和压缩。瓶颈层（Bottleneck layers）如图 4-8 所示，将卷积作为瓶颈层，对每个卷积缩小了输入的特征图，从而使模型计算效率得到提升。瓶颈层加入 DenseNet，记为 DenseNet-B。

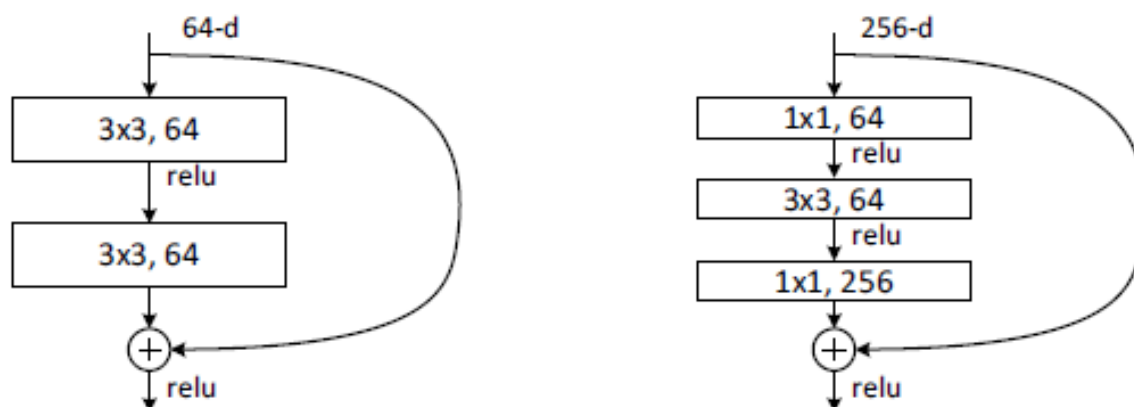


图 4-8 瓶颈层示意图

压缩（Compression）可以在过渡层减少特征图的数量，从而更好地提升模型的紧凑程度。在 DenseNet 加入压缩，记为 DenseNet-BC。

同时，对模型进行 early stopping，减轻过拟合。

## 4.4 实验结果

在这一部分将对舌下络脉颜色分类进行实验。在第 4.4.1 节中，对舌下络脉颜色分类实验所用到的数据集进行描述；在第 4.4.2 节中，对实验中用到的评价指标进行介绍；在后续的第 4.4.3 节和第 4.4.4 节中，对输入络脉方案、分类模型实验结果进行评估。

### 4.4.1 数据集描述

数据集来源与上海中医药大学。上海中医药大学提供了标定颜色的图像共 112 张，其中淡紫 53 张；青紫 48 张；紫黑 11 张。需要根据提取舌下络脉部分，分辨络脉颜色（淡紫、青紫、紫黑三类）。

### 4.4.2 颜色分类评价指标

我们利用准确率来评估络脉颜色分类的性能。其中准确率的计算方式如下：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4-1)$$

式中  $TP$ ——预测值为正类（语义分割目标）、真实值为正类的样本数量；  
 $TN$ ——预测值为负类、真实值为负类的样本数量；  
 $FP$ ——预测值为正类、真实值为负类（其他背景像素）的样本数量；  
 $FN$ ——预测值为负类、真实值为正类的样本数量。

### 4.4.3 实验输入络脉方案及结果

根据络脉的分割结果，对模型的输入图像的选择进行实验。设计原图/原图提取络脉/络脉坐标点切割原图/双络脉矩形/单络脉矩形/颜色块等方法，基于此作为输入进行实验，实验结果如表 4-2 所示。

最终选择双络脉矩形，如图 4-9 所示，作为分类模型的输入。

### 4.4.4 图像分类模型实验

基于双络脉矩形的输入，对 20 种图像分类模型进行实验。DenseNet 系列的模型在训练集上效果明显较好，测试集上效果与其他模型相比也不差。最终选择效果最佳的 DenseNet161-bc，并为了减少过拟合进行了 early\_stopping。实验结果如表 4-3 所示。



表 4-2 输入舌下络脉方案

任务	train_acc	val_acc
原图	0.7949	0.4706
原图提取络脉（除络脉其余部分白色）	0.8462	0.4412
原图提取络脉（除络脉其余部分黑色）	0.8077	0.4412
络脉坐标点切割原图	0.8333	0.7353
切割双络脉的矩形（除络脉其余部分白色）	<b>0.8846</b>	<b>0.7647</b>
切割双络脉的矩形（除络脉其余部分黑色）	0.8077	0.6471
切割单络脉的矩形（除络脉其余部分白色）	0.7436	0.7647
切割单络脉的矩形（除络脉其余部分黑色）	0.7308	0.6765
100*100 颜色块	0.6410	0.4412
切割单络脉的矩形（除络脉其余部分填充）	0.7564	0.6647

表 4-3 舌下络脉颜色分类效果

图像分类模型	切割双络脉的矩形	train_acc	val_acc
Resnet 残差网络, Highway Networks 网络较深	resnet18	0.8257	0.7447
	resnet34	0.7339	0.7021
	resnet50	0.6147	0.6383
	resnet101	0.7248	0.5532
	resnet151	0.7523	0.617
VGG Net 2*2 最大池化 使得层数更深	vgg11	0.5138	0.2979
	vgg11+bn	0.6055	0.4468
	vgg13	0.3028	0.3404
	vgg13+bn	0.6239	0.4043
	vgg16	0.3028	0.3404
	vgg16+bn	0.4862	0.4894
	vgg19	0.5138	0.2979
	vgg19+bn	0.4679	0.5319
SqueezeNet 参数减少 50 倍	SqueezeNet 1.0	0.7431	0.617
	SqueezeNet 1.1	0.5872	0.5532
DenseNet 密集连接特性	densenet121	0.9174	0.6809
	densenet161	0.9174	0.7021
	densenet169	0.8899	0.7021
	Densenet201	0.9266	0.6809
	densenet161- bc+early_stopping	<b>0.8363</b>	<b>0.8037</b>

## 4.5 本章小结

本章节基于第三章语义分割后的络脉，设计了不同的输入分类模型的络脉提取特征方案。并通过 ResNet、VGG Net、DenseNet 等深度学习分类模型对络脉颜

色进行分类。



图 4-9 双络脉矩形-分类模型输入

对效果最佳 DenseNet161 模型进行调整，加入瓶颈层和压缩，并提前中止减轻过拟合，以得到最佳的分类效果。最后对实验结果进行了描述，我们的颜色分类方法得到了较好的结果。

## 第5章 舌体与舌下络脉分割及颜色分类系统设计实现

### 5.1 引言

基于前文所提出的中医舌体与舌下络脉分割提取方法、数据的预处理和后处理算法及颜色分类算法，可以开发出一套中医舌体与舌下络脉分割提取及颜色分类系统，以展示算法。基于所给的数据，对于分割算法的相应模型进行训练；使用经过训练的模型，预测图像的舌体与舌下络脉位置，并进行分割提取。

### 5.2 系统总体实现流程

基于前文中的舌体与络脉分割方法，开发了一套中医舌体与舌下络脉分割提取及颜色分类系统。使用中医舌体与舌下络脉分割及颜色分类系统进行中医舌体与舌下络脉的预测提取操作主要可以分为两个阶段，第一是系统的准备阶段；第二是系统的使用阶段。

在系统的准备阶段，主要完成四个步骤的准备工作。首先，针对所给数据进行预处理，去除反光点，减少对深度学习模型学习的影响。其次，根据去除反光点后的数据，进行标注，并根据不同的标签形成舌体数据集和舌下络脉的数据集。第三，根据舌体数据集，对训练集进行数据增强操作，得到扩充的训练数据集。再将增强后的训练集输入深度学习模型进行训练，获得经过训练的舌体预测模型。第四，根据第三步提取出的舌体数据，将其进行数据增强操作，将其与舌下络脉的标注作为舌下络脉的训练集，输入深度学习模型进行训练，获得经过训练的舌下络脉预测模型。

在系统的使用阶段，测试者将使用已经训练完成的预测模型，实施所需的图像预测提取分割操作。第五步，在目录中选择所需处理图像，将原始图像显示在系统中。第六步，选择所要进行的操作。共给出六种操作：（1）图像去除反光点；（2）图像显示舌体标注信息；（3）图像预测舌体部位并画出轮廓；（4）图像显示舌下络脉标注信息；（5）图像预测舌下络脉部位并画出轮廓；（6）图像预测舌下络脉颜色；第七步，系统可对所选的操作进行删除。根据所需操作进行选择和删除。第八步，对于所需操作生成的图像，可以选择保存。保存图像的路径和文件名由使用者选择。

需要说明的是，在分割及颜色分类系统第一次使用前，需要完成系统准备阶段的四大准备步骤。在准备阶段完成后，就可以多次对系统进行使用操作，而不

必再次重复系统的准备阶段工作。整个系统的操作流程如图 5-1 所示。

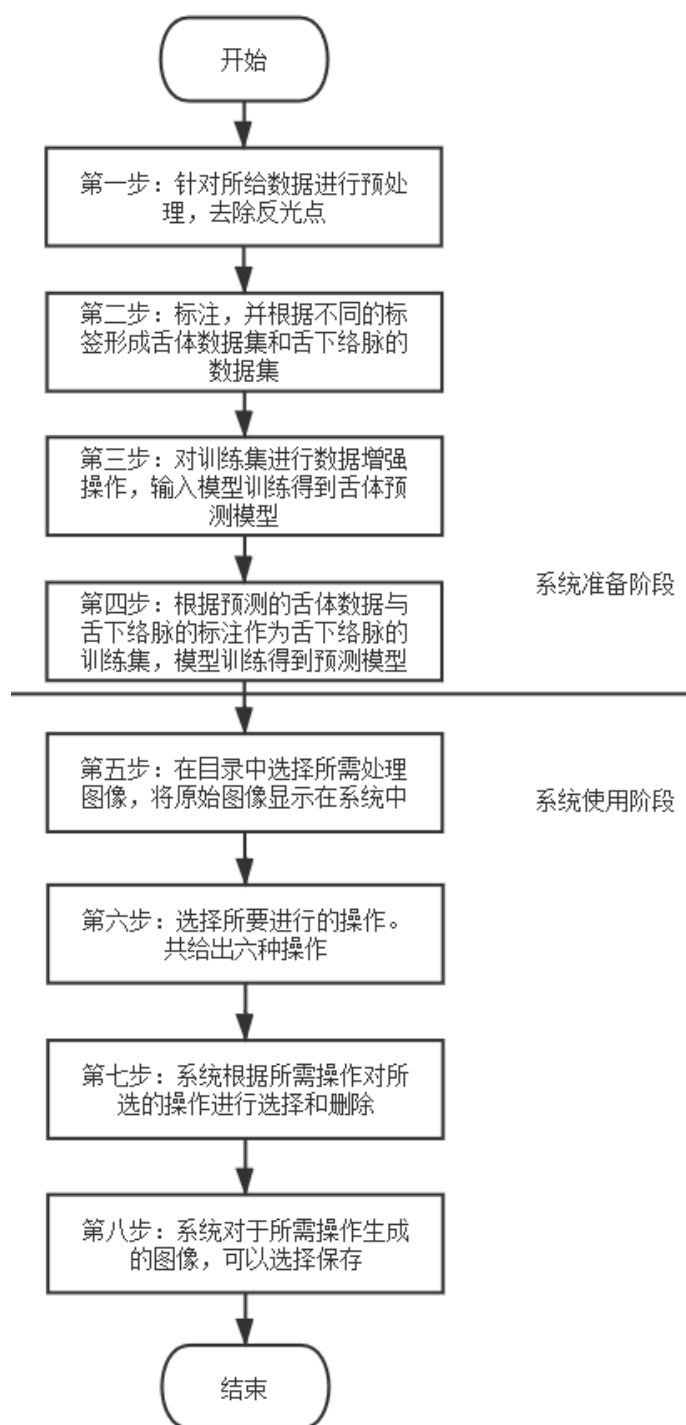


图 5-1 系统操作流程

### 5.3 系统功能架构设计

系统的总体架构见图 5-2。

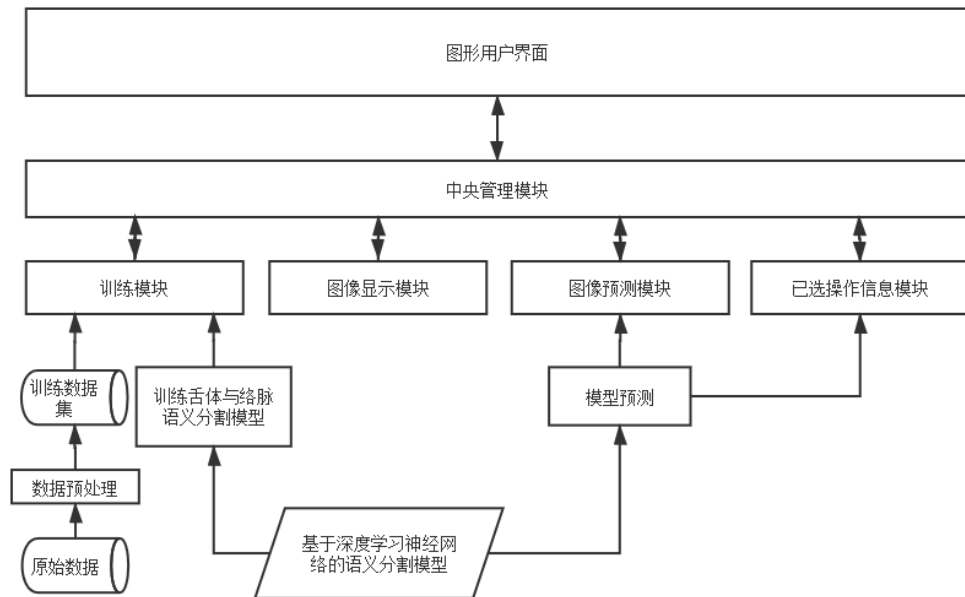


图 5-2 系统整体架构设计

如图 5-2 所示，整个中医舌体与舌下络脉分割及颜色分类系统自顶向下可以分成三大主要层次，顶层是用户界面模块；中间是中央控制模块；底层是各个功能模块。

用户界面模块主要是给中医舌体与舌下络脉分割及颜色分类系统的使用者提供一个图形化的友好的用户操作界面，以方便用户对于中医舌体与舌下络脉分割及颜色分类系统的使用。

中央控制模块，一方面负责协调控制各个底层功能模块之间的工作，另外一方面，负责将界面获得的用户操作指令转化为各个底层功能模块的协调动作执行起来，同时将底层功能模块的执行结果反馈给用户界面模块。

底层功能模块主要包括如下功能模块：（1）训练模块：主要包含三大算法功能。第一，将训练数据进行预处理，进行数据增强和标签分类，形成两类规范的舌体数据集和舌下络脉数据集。第二，针对舌体数据集对深度学习模型进行训练，形成稳定的预测模型，并保存相应的训练模型结果。第三，根据第二步预测的舌体数据，将其与络脉标注作为舌下络脉模型预测的输入，对深度学习模型进行训练，形成稳定的预测模型，并保存相应的训练模型结果；（2）图像操作显示模块：主要是基于原始数据，显示去除反光点、舌体标注、舌下络脉标注效果；（3）图像预测显示模块：主要是基于已经训练好的语义分割模型，针对使用者

的需求，选择舌体轮廓预测显示、舌下络脉轮廓预测显示及舌下络脉颜色分类；  
（4）已选操作信息模块：显示已选的操作，并显示相应属性。使用者可以根据需求对已选操作进行删除。

## 5.4 系统功能实现

系统功能设计如图 5-3 所示。

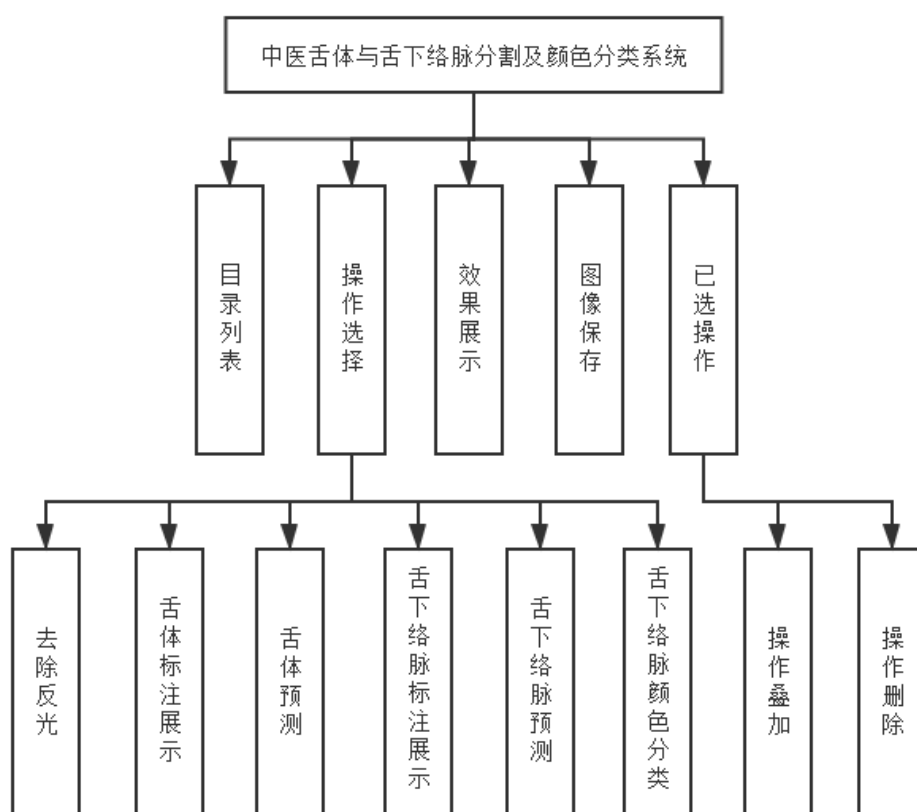


图 5-3 系统功能设计

系统的运行需要如下几个层次运行环境的支撑。首先系统的准备阶段需要在 Tensorflow 和 Keras 相适配环境下运行；系统的运行需要在 PyQt5 环境中运行。只有具备了上述支撑环境，中医舌体与舌下络脉分割及颜色分类系统才能正常地运行。系统的使用者只需要通过图形界面就能完成舌体与舌下络脉的预测轮廓提取功能。

数据处理的操作部分，如去除反光点操作、舌体标注展示操作、舌下络脉标注展示操作主要使用 Python 实现。舌体与舌下络脉语义分割的操作部分主要基于 Tensorflow 和 Keras 框架进行模型的编写。为了提高系统的速度，方便用户使

用，首先根据模型的训练集在 GPU 服务器上进行训练。在系统执行舌体与舌下络脉语义分割功能时，会调用已经训练好的模型文件进行分析，从而大大提高系统语义分割操作的速度。

为了方便用户进行本地化存储，系统中所有的图片支持导出。

## 5.5 系统功能展示

首先，作为中医舌体与舌下络脉分割提取及颜色分类系统，主界面如图 5-4 所示，用户可从左栏中选择本地数据集中需处理图片。

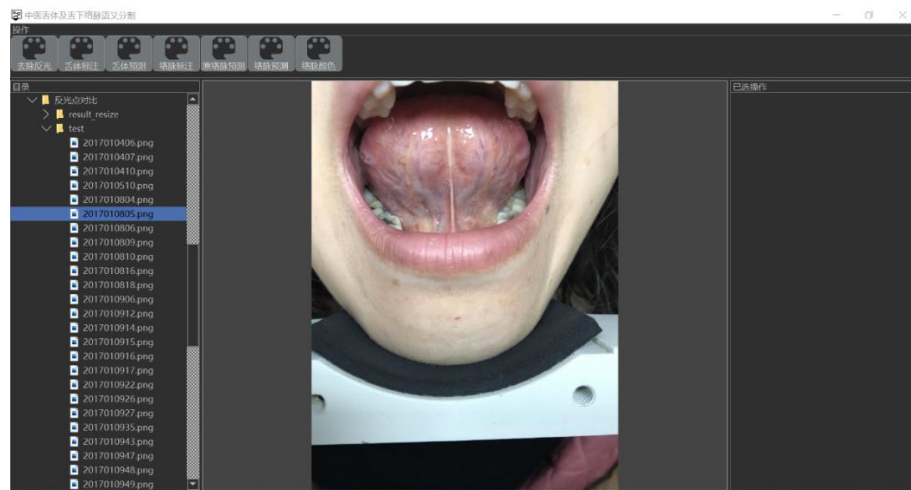


图 5-4 中医舌体与舌下络脉分割提取及颜色分类系统主界面

用户可以选择去除反光操作，去除反光后图片展示在中间区域，右侧已选操作栏显示选中的去除反光操作，如图 5-5 所示。

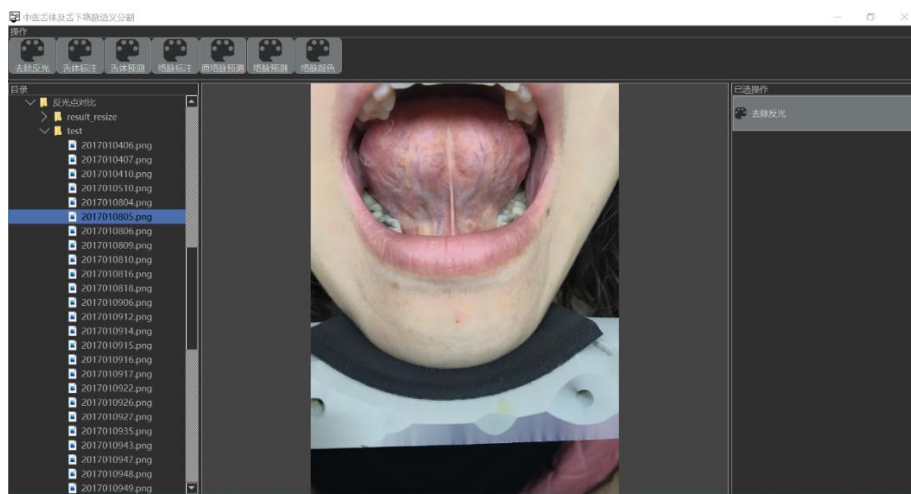


图 5-5 去除反光操作

在已有操作的基础上，中间显示叠加的操作后的图片。如图 5-6 所示，展示舌轮廓标注。

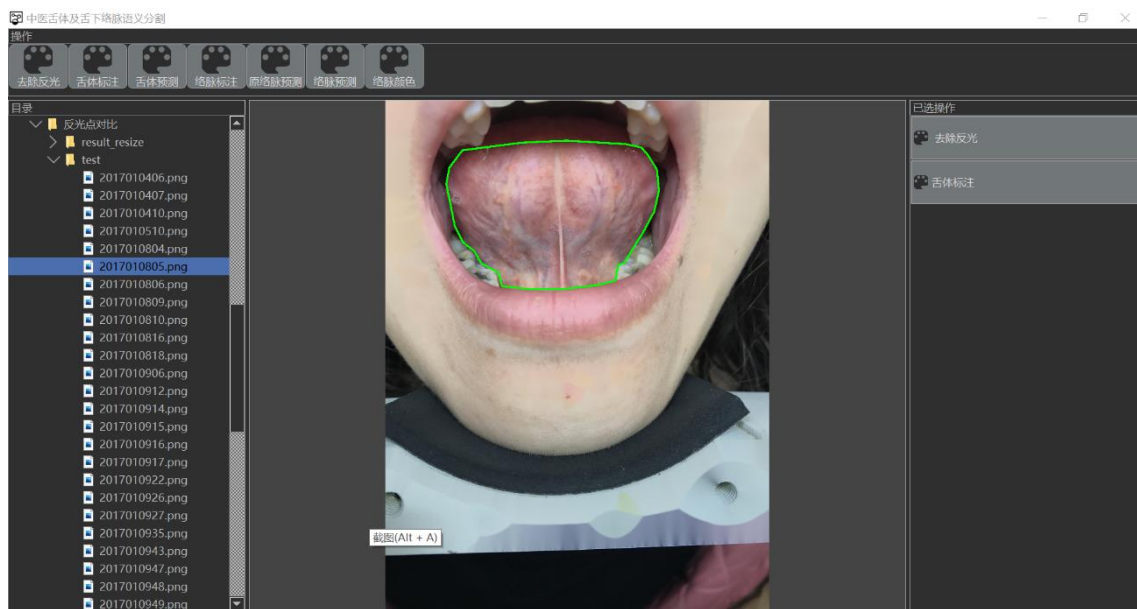


图 5-6 舌体轮廓标注显示操作

在已有操作的基础上，中间显示叠加的操作后的图片。如图 5-7 所示，展示舌下络脉轮廓标注。

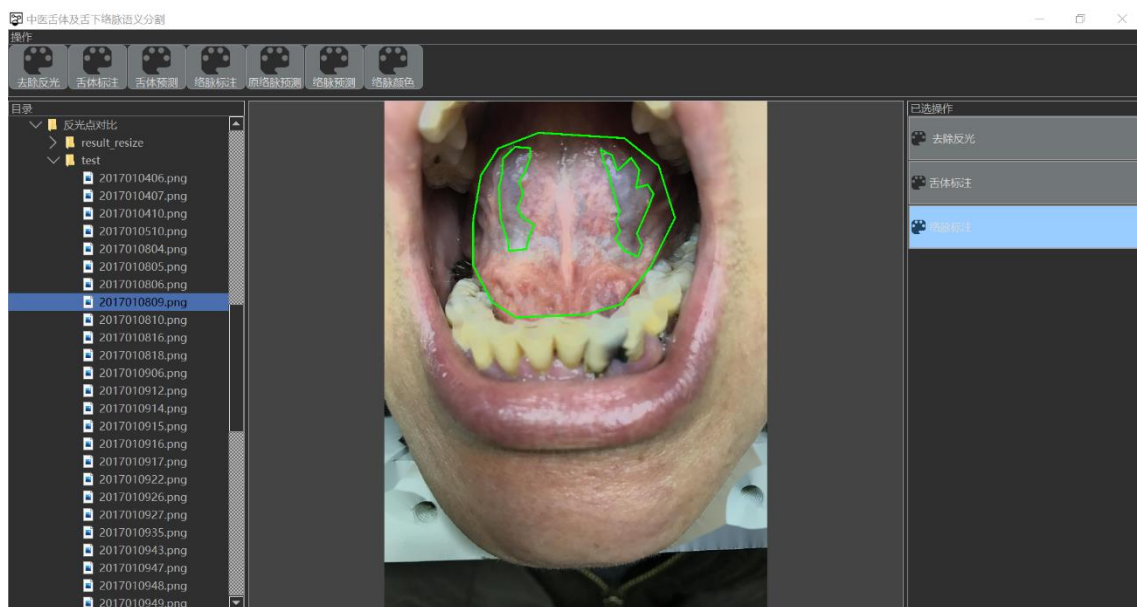


图 5-7 舌下络脉轮廓标注显示操作

模型对舌体进行预测，并显示出语义分割结果，为与舌体轮廓标注区别，预测



舌体轮廓结果展示采用红色绘制。如图 5-8 所示，模型预测舌体语义分割。

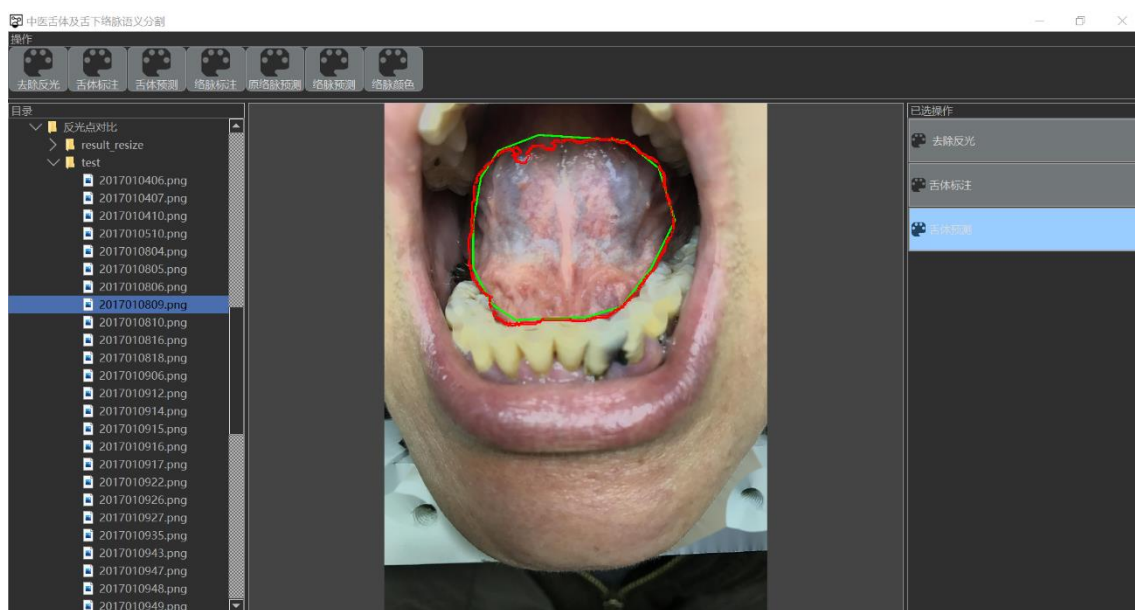


图 5-8 模型预测舌体语义分割

模型对舌下络脉进行预测，并显示出语义分割结果，为与舌下络脉轮廓标注区别，预测舌下络脉轮廓结果展示采用红色绘制。如图 5-9 所示，模型预测舌下络脉语义分割。

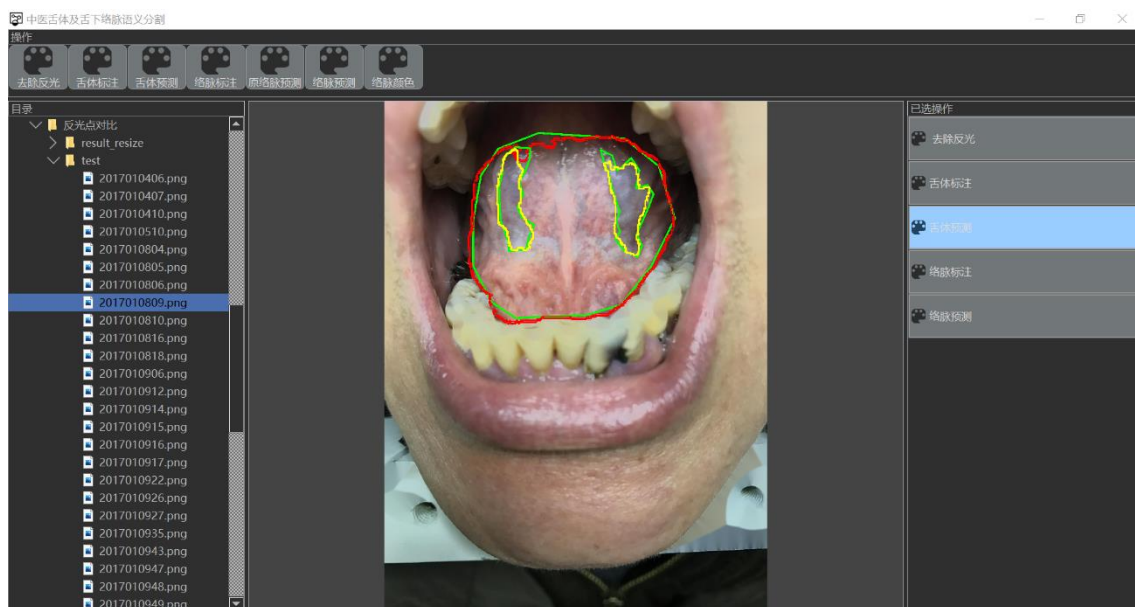


图 5-9 模型预测舌下络脉语义分割

模型对舌下络脉颜色进行预测，单击已选操作中的络脉颜色，显示出络脉的颜

色信息。如图 5-10 所示，模型预测舌下络脉语义分割。

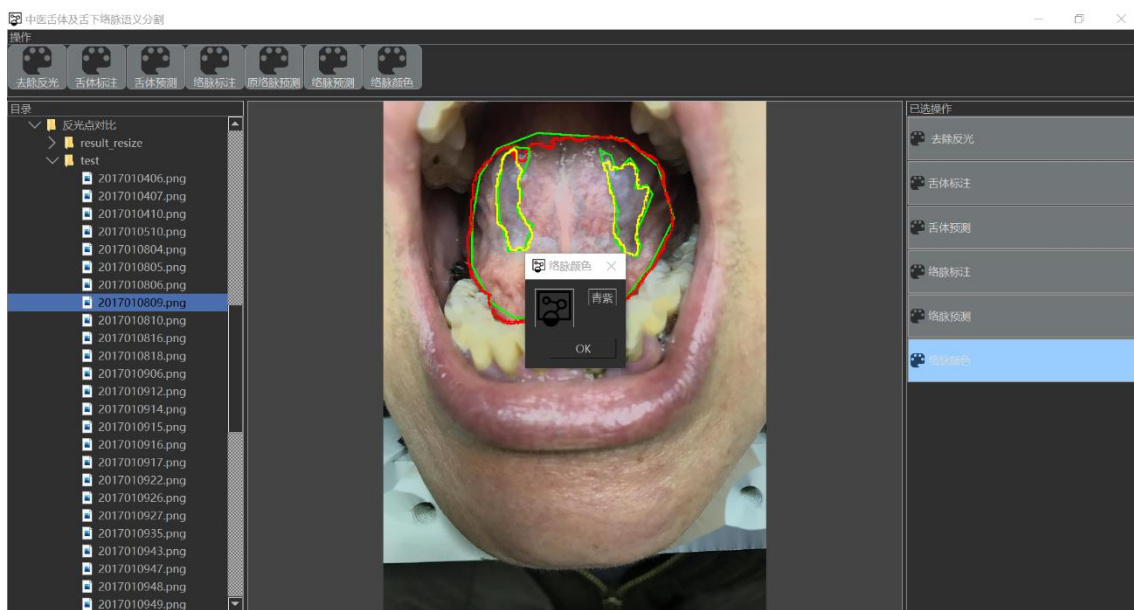


图 5-10 系统基本信息展示部分（卡号基本信息展示）

右侧操作栏的已选操作可以进行单击右键对其删除。如图 5-11 所示。

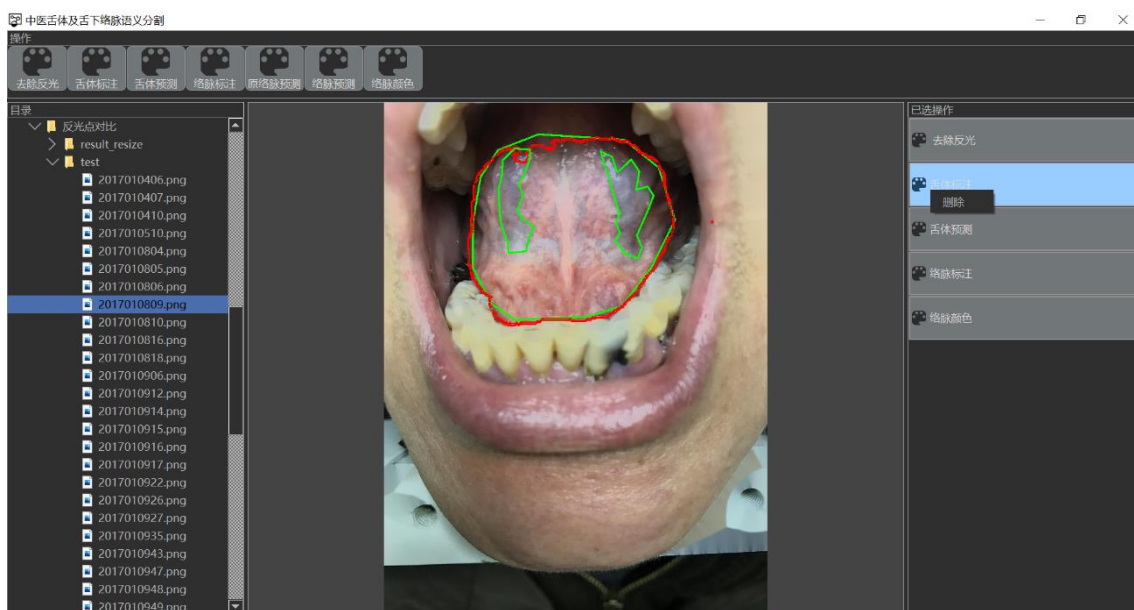


图 5-11 已选操作删除

全部操作过程中的图片，都可以根据用户的需求，进行保存。图片可另存为.jpg,.png,.jig 格式。图片另存为界面如图 5-12 所示。

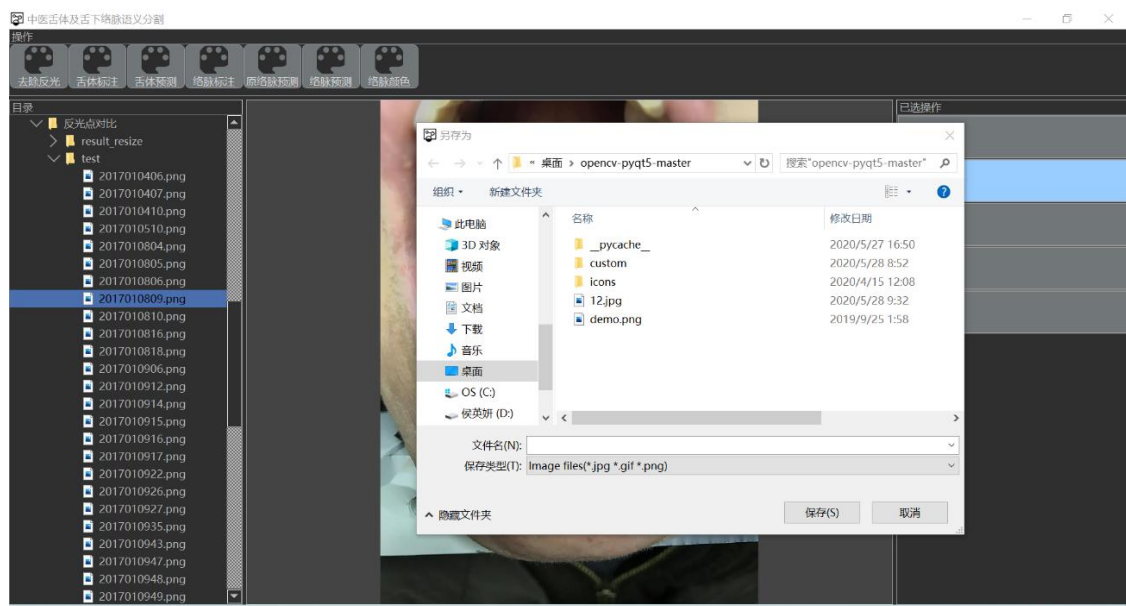


图 5-12 图片另存为

## 5.6 本章小结

本章在语义分割舌体与络脉相关算法研究以及络脉颜色分类研究的基础上，设计并实现了中医舌体与舌下络脉分割及颜色分类系统。本章从系统实现流程、系统架构设计、各功能模块设计及展示等方面对中医舌体与舌下络脉分割及颜色分类系统进行了介绍。总的来看，该系统完成了算法的展示功能，具有一定的实用性。

## 结 论

为推动中医舌诊的客观化发展，本文借助先进的计算机技术，用于传统的中医诊断领域，本文将人工智能与中医舌诊相结合，主要对中医舌下图像进行语义分割，提取舌体与舌下络脉，并对舌下络脉进行颜色分类。

本文的第一个研究内容是舌体与舌下络脉的语义分割。舌体分割的目的，是为了减少图像中其他因素的干扰，从背景中把舌体分离出来。中医对舌下络脉的观察是舌诊的重要部分，从原始图像中将舌下络脉进行提取，便于之后的疾病诊断。本文基于深度学习语义分割方法，对舌体及舌下络脉设计无下采样的语义分割模型，加入膨胀卷积扩大网络的感受野，同时对膨胀卷积率进行设计，避免了栅格效应。并基于络脉小样本目标特性，选取先定位缩小目标区域再分割的策略。并辅以数据增强、去除镜面反射等数据处理，对舌体与舌下络脉语义分割达到了不错的效果。经实验，模型效果优于经典医学图像语义分割 U-net 模型。

本文的第二个研究内容是舌下络脉的颜色分类进行颜色三分类，分为淡紫、青紫、紫黑三类。络脉的颜色分类对中医诊断有着很重要的作用。中医可以通过观察舌下络脉的颜色，来判断病人的血液循环情况。本文基于语义分割出的舌下络脉，进行了多种输入设计，并选取效果最佳的双络脉矩形，经实验选择 Densenet161 模型，并对模型加入瓶颈层和压缩，进行舌下络脉颜色分类。

本文的创新点主要在于（1）语义分割模型的设计。摒弃了常用的下采样-上采样架构，未池化而是加入消除栅格效应的膨胀卷积，获得了很好的效果；（2）去除镜面反射算法。根据数据特点，加入这一项图像预处理；（3）小目标分割策略。基于络脉宽度占比小的小目标特性，将前一步的舌体提取效果作为络脉模型训练的输入；（4）络脉颜色分类的模型输入，设计不同方案选取颜色信息未丢失的双络脉矩形方案。

本文对中医舌下图像进行语义分割，提取舌体与舌下络脉，并对舌下络脉进行颜色分类，均取得了优秀的实验结果。语义分割模型在该数据集上有较好效果，未来将尝试将本研究语义分割模型泛化，在其它类型的公开数据集上进行语义分割实验，尤其是对于小目标语义分割。

## 参考文献

- [1] 魏敏. 中医药发展“十三五”规划发布[J]. 中医药管理杂志, 2016, 17: 79.
- [2] Li Q, Wang Y, Liu H, et al. Sublingual vein extraction algorithm based on hyperspectral tongue imaging technology[J]. Computerized Medical Imaging & Graphics, 2011, 35(4): 179-185.
- [3] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [4] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-assisted Intervention, 2015: 234-241.
- [5] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[J]. arXiv preprint arXiv: 1511.07122, 2015.
- [6] Chen L, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834.
- [7] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.
- [8] Peng C, Zhang X, Yu G, et al. Large Kernel Matters--Improve Semantic Segmentation by Global Convolutional Network[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4353-4361.
- [9] Kirillov A, Wu Y, He K, et al. PointRend: Image Segmentation as Rendering[J]. arXiv preprint arXiv: 1912.08193, 2019.
- [10] Yang T, Yoshimura Y, Morita A, et al. Fully Automatic Segmentation of Sublingual Veins from Retrained U-Net Model for Few Near Infrared Images[J]. arXiv preprint arXiv: 1812.09477, 2018.
- [11] Yang T, Yoshimura Y, Morita A, et al. Synergistic attention U-Net for sublingual vein segmentation[J]. Artificial Life and Robotics, 2019, 24(4): 550-559.
- [12] Li Z, Weida Z, Peichann C, et al. Kernel sparse representation-based classifier[J]. IEEE Transactions on Signal Processing, 2011, 60(4): 1684-1695.
- [13] Zhang L, Yang M, Feng X. Sparse representation or collaborative representation: Which helps face recognition?[C]. International Journal of Computer Vis

- ion, 2011: 471-478.
- [14] Naseem I, Togneri R, Bennamoun M. Linear regression for face recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(11): 2106-2112.
- [15] Cover T, Hart P. Nearest neighbor pattern classification[J]. IEEE Transactions on Information Theory, 1967, 13(1): 21-27.
- [16] Xuegong Z. Introduction to statistical learning theory and support vector machines[J]. Acta Automatica Sinica, 2000, 26(1): 32-42.
- [17] Takeichi M, Sato T. Computerized color analysis of "xue yu" (blood stasis) in the sublingual vein using a new technology[J]. The American Journal of Chinese Medicine, 1997, 25(02): 213-219.
- [18] 余娅. 数字式舌诊仪的研制与应用[J]. 计算技术与自动化, 1990, 9(1): 68-72.
- [19] 赵荣莱, 许胜. 舌质舌苔的计算机定量描述和分类[J]. 中医杂志, 1989, 30(2): 47-48.
- [20] 孙立友, 程钊, 高逢生, 等. 利用计算机图像识别技术进行舌诊客观化研究的探讨[J]. 安徽中医学院学报, 1986, 5(4): 5-7.
- [21] Shen L, Wei B, Cai Y, et al. Image analysis for tongue characterization[J]. Chinese Journal of Electronics, 2003, 12(3): 317-323.
- [22] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models[J]. International Journal of Computer Vision, 1988, 1(4): 321-331.
- [23] 吴佳, 张永红, 白净, 等. 基于分水岭变换和主动轮廓模型的舌体轮廓图像分割[J]. 清华大学学报: 自然科学版, 2008, 48(6): 1040-1043.
- [24] 王宽全, 张宏志, 李乃民, 等. 计算机化中医舌诊[J]. 第三届世界中西医结合大会论文摘要集, 2007.
- [25] Zhu M, Du Q. A novel approach for color tongue image fast segmentation based on level sets[C]. Advanced Materials Research, 2012, 341: 714-719.
- [26] Li W, Yao J, Yuan L, et al. The segmentation of the body of tongue based on the improved level set in TCM[M]. Life System Modeling and Intelligent Computing, 2010: 220-229.
- [27] 张志顺, 刘勇. 基于动态阈值和修正模型的舌体提取算法[J]. 计算机与现代化, 2014(11): 49-52.
- [28] Hoover A, Kouznetsova V, Goldbaum M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response[J]. IEEE Transactions on Medical Imaging, 2000, 19(3): 203-210.
- [29] Yan Z, Ding H, Li N. Automatic segmentation of sublingual excrescences in color sublingual images[C]. International Conference on Medical Biometrics,

2014: 85-89.

- [30] 张纬博. 中医舌诊中舌下络脉特征提取与分析研究[J], 2011.
- [31] 郑舜仁, 蒋依吾, 陈建仲, 等. 中医舌诊计算机化舌下络脉特征提取及分析[J]. 苏州大学学报: 医学版, 2005, 25(3): 426-430.
- [32] 孙丹萍, 吴佳, 张永红, 等. 基于特征聚类的舌下络脉自动提取方法[J]. 中国生物医学工程学报, 2008, 27(2): 265-269.
- [33] Wang P, Chen P, Yuan Y, et al. Understanding convolution for semantic segmentation[C]. IEEE Winter Conference on Applications of Computer Vision, 2018: 1451-1460.
- [34] Li W, Wang G, Fidon L, et al. On the compactness, efficiency, and representation of 3D convolutional networks: brain parcellation as a pretext task[C]. International Conference on Information Processing in Medical Imaging, 2017: 348-360.
- [35] Tan R T, Ikeuchi K. Separating reflection components of textured surfaces using a single image[M]. Digitally Archiving Cultural Objects, 2008: 353-384.
- [36] Shafer S A. Using color to separate reflection components[J]. Color Research & Application, 1985, 10(4): 210-218.
- [37] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks[C]. European Conference on Computer Vision, 2016: 630-645..
- [38] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and<0.5MB model size[J]. arXiv preprint arXiv:1602.07360, 2016.
- [39] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700-4708.

## 致 谢

四年大学生活已进入了倒计时，回顾这四年，可以说很平凡，也可以说很特别。在此，特别感谢我的老师、同学们。

首先，感谢我的导师张盛平老师。张盛平老师治学严谨，为人和善。作为我的班主任老师和实验室导师，张老师在我大学四年期间给予了许多指导和关怀，更给予了我很多的帮助。当我遇到问题时，就会和老师进行讨论，老师向我讲述很多人生的哲理，为我的大学生活指引方向。当生活或者学习上有问题的时候，老师总是有求必应。大学四年张老师的治学态度让我敬佩，在老师的指导下我受益良多。衷心感谢张盛平老师的指导和帮助。

同时也真诚地感谢导师徐华老师，在毕业设计的半年以来，每一次组会上都耐心且尽心指导我，一直激励我前进。同时，也想感谢同组的余文梦师兄和伏天韵师姐，他们丰富的科研经历和知识给我的工作非常多的启发，许多问题也在讨论的过程中迎刃而解。

感谢王佳浚老师大学四年对我的关怀，感谢计算机学院的各位老师给我的支持和帮助，感谢辩论队学长学姐们大学给予我的关怀，感谢我大学期间认识的同学们，他们让我度过了充实而快乐的大学四年的生活。愿大家都前程似锦，有一个美好的未来。

感谢我的父母，谢谢他们对我的支持，他们多年来无私的付出是我如今能享有这一切的根源。

最后，特别感谢答辩组的各位老师，感谢你们的指导和帮助。