硕士学位论文

基于深度学习的足踝关节骨损辅助诊断研 究

Research on Auxiliary Diagnosis of Ankle Joint Bone Damage Based on Deep Learning

学科专业	工程硕士
专业领域	控制工程
作者姓名	彭哲
指导教师	谢斌 副教授

中图分类号_TP391.4_

UDC <u>004.8</u>

学校代码<u>10533</u> 学位类别<u>专业学位</u>

硕士学位论文

基于深度学习的足踝关节骨损辅助诊断研 究

Research on Auxiliary Diagnosis of Ankle Joint Bone Damage Based on Deep Learning

作者姓名	彭哲
学科专业	工程硕士
专业领域	控制工程
研究方向	模式识别与智能系统
二级培养单位	自动化学院
指导教师	谢斌 副教授

论文答辩日期 2021.5.30

答辩委员会主席 2 4

中南大学 2021年5月

学位论文原创性声明

本人郑重声明,所呈交的学位论文是本人在导师指导下进行的研 究工作及取得的研究成果。尽我所知,除了论文中特别加以标注和致 谢的地方外,论文中不包含其他人已经发表或撰写过的研究成果,也 不包含为获得或其他教育机构的学位或证书而使用过的材料。与我共 同工作的同志对本研究所作的贡献均已在论文中作了明确的说明。

申请学位论文与资料若有不实之处,本人承担一切相关责任。



学位论文版权使用授权书

本学位论文作者和指导教师完全了解有关保留、使用学位论文的 规定:即学校有权保留并向国家有关部门或机构送交学位论文的复印 件和电子版;本人允许本学位论文被查阅和借阅;学校可以将本学位 论文的全部或部分内容编入有关数据库进行检索,可以采用复印、缩 印或其它手段保存和汇编本学位论文。

保密论文待解密后适应本声明。

2 zer 作者签名: 查 折 导师签名: 日期: 2021年_6月2日 日期: 2021年6月2日

基于深度学习的足踝关节骨损辅助诊断研究

摘要:骨骼是人体的重要支撑组织结构,随着人口年龄结构的改 变以及人们生活水平的提高,踝关节骨损作为骨骼损伤中一种高发 疾病,受到越来越多的关注,它们易于发生、隐蔽性高、治疗昂贵 复杂、恢复周期长。踝关节骨损有多种类型,在临床诊断中小尺寸 的骨损和肌肉韧带损伤容易被忽视。基于智能算法的医疗辅助诊断 系统能够提升医生工作效率,减轻临床工作负担。但该领域的研究 面临着相关数据量不足,小尺寸骨损容易发生漏检等问题,因此现 阶段针对踝关节骨损检测以及提取骨骼几何物理信息的相关研究比 较缺乏。针对以上这些问题,本文做出以下工作:

1.针对踝关节骨损数据不足的问题,本文构建了一个踝关节骨损数据库,其由湘雅三医院骨科和影像科合作提供原始数据,包含1852 张足踝关节 x 光片,其中 1246 张为非完全粉碎型骨损数据, 606 张为完全粉碎型骨损数据。非完全粉碎型骨损数据中有 511 张数据为健康无骨损数据,735 张为骨损数据。据统计,数据的分辨率均大于 2000×2500,骨损区域的大小不一,绝大多数都在 300×300 以下,最小的骨损区域尺寸只有 50×50,且骨损主要集中在踝关节 周围。通过整理和标注,建立了一个面向足踝关节骨损检测的数据 集,以及一个面向踝关节缝隙语义分割任务的数据集。

2.通过分析以上骨骼原始数据集的特点,本文面临的问题如下: 由于完全粉碎性骨损数据的骨骼物理结构呈现出较大的离散性,该 类数据增加了辅助系统的训练拟合难度,对系统的稳定性干扰较大; 骨损区域在数据中的相对尺寸较小,容易出现漏检问题;在临床诊 断中,缺少经验的普通医生,很难仅通过 x 光片确诊肌肉韧带损伤 的问题。本文针对以上问题,提出一种基于神经网络的多阶段踝关 节骨损辅助诊断方案。该方案包含三个阶段,分别为:过滤干扰数 据阶段、识别踝关节阶段、骨损识别以及关节缝隙分割阶段。在过 滤干扰数据阶段,本文通过二分类网络过滤掉数据集中的完全粉碎 性骨损数据,只留下非完全粉碎型骨损数据,保证了系统后阶段不 受干扰数据的影响,在提升系统的稳定性的同时,减少系统网络的 训练难度;在识别踝关节阶段,本文将区域预测网络 RPN 融入到 Resnet 网络中,实现了对踝关节的识别,将识别出的踝关节作为下 一阶段的输入,从而增加了小尺寸目标在输入数据中的相对比例, 提升系统对相对小目标的检测能力;在骨损识别以及关节缝隙分割 阶段,本文提出在目标识别算法中添加注意力机制优化特征网络结 构,提升系统对小尺寸骨损地识别性能。为了辅助临床医生更好的 识别肌肉韧带损伤,本文提出利用语义分割算法提取踝关节间隙。

3.为了解决不同区域之间医疗资源不平衡的问题,让更多患者可 以享受同样优质的医疗资源,缓解医生工作压力,提升工作效率, 本文基于安卓系统开发了一套踝关节骨损自动检测系统。用户可以 在 App 上快捷方便的上传数据以及查看检测结果,也能邀请专家进 行联合会诊。

相比只进行骨损分类的辅助系统,本文辅助诊断系统除了可以对 骨损进行精确定位,还提升了对踝关节相对小尺寸骨损目标的识别 能力,同时可以提取出踝关节间隙的几何信息,为骨科医生诊断肌 肉韧带损伤提供重要的物理参考依据。

图 46 篇,表 14 个,参考文献 66 篇

关键字: 足踝关节疾病; 目标识别; 语义分割; 辅助诊断系统; 注 意力机制; 肌肉韧带损伤

Research on Auxiliary Diagnosis of Ankle Joint Bone

Damage Based on Deep Learning

Abstract: Bone is an important supporting structure of the human body. With the changes in the age structure of the population and the improvement of people's living standards, ankle bone damage, as a high-incidence disease in bone injuries, has received more and more attention. They are prone to occur, highly concealed, expensive and complex to treat, and have a long recovery period. There are many types of ankle bone damage, where small-sized bone damage and muscle or ligament injury are easily overlooked in clinical diagnosis. The medical assistant diagnosis system based on intelligent algorithms can improve the work efficiency of doctors and reduce the burden of clinical work. However, the research in this field faces the problems of insufficient data and missed detection in small-size bone damage. Therefore, at this stage, there is a lack of related research on ankle bone damage detection and extraction of bone geometry and physical information. In response to the above problems, this article makes the following work:

1. Aiming at the problem of insufficient data on ankle joint bone damage, this paper constructs an ankle joint bone damage database. The original data is provided by the Department of Orthopedics and Imaging of the Third Xiangya Hospital. It contains 1852 foot and ankle x-rays, where 1246 are Non-complete comminuted bone damage and 606 sheets are completely comminuted bone lesion. Among the Non-complete comminuted bone damage data, 511 pictures are healthy and non-damaged data, and 735 pieces are bone damage data. According to statistics, the resolution of the data is greater than 2000×2500 . The size of the bone damage area is different. Most of them are below 300×300 , but the smallest bone damage area is only 50×50 . And the bone damage is mainly concentrated in the ankle joint around. Through sorting and labeling, a data set for the detection of bone damage in the foot and ankle joint

and a data set for the task of semantic segmentation of the ankle joint gap were established.

2. By analyzing the characteristics of the original bone data set, the problems faced by this article are as follows: Since the physical structure of the bones of the completely comminuted bone damage presents a large discreteness, this type of data increases the difficulty of training and fitting the auxiliary system. It has a large interference to the stability of the system. The relative size of the bone damage area in the data is small, and it is easy to miss detection .In clinical diagnosis, it is difficult for ordinary doctors who lack experience to diagnose muscle and ligament damage only through x-rays. Aiming at the above problems, this paper proposes a multi-stage ankle joint bone damage auxiliary diagnosis scheme based on neural network. This program consists of three stages: filtering interference data, identifying ankle joints, identifying bone damage, and segmenting joint gaps. In the stage of filtering interference data, this paper filters out the completely crushed bone loss data in the data set through the two-class network, leaving only the incomplete crushed bone loss data, to ensure that the system is not affected by the interference data in the later stage of the system. It not only improves the stability of the system but also reduces the training difficulty of the system network. In the stage of identifying the ankle joint, this article integrates the regional prediction network RPN into the Resnet network to realize the recognition of the ankle joint, and use the identified ankle joint as the input of the next stage. Thus this work increase the relative proportion of small-sized targets in the input data and improve the ability of system to detect relatively small targets. In the stage of bone damage recognition and joint gap segmentation, this paper proposes to add an attention mechanism to the target recognition algorithm to optimize the feature network structure .This work can improve the ability of system to recognize small-size bone damage. To assist clinicians to better identify muscle and ligament injuries, this paper proposes to use semantic segmentation algorithm to extract the ankle joint space.

3. In order to solve the problem of the imbalance of medical resources between different regions, meanwhile let more patients can enjoy the same high-quality medical resources, and doctors work easy, this article developed a set of automatic ankle bone damage detection based on the android system. Users can quickly and conveniently upload data and view test results on the App. Besides, users can invite experts for joint consultations.

Compared with the auxiliary system that only performs bone damage classification, the auxiliary diagnosis system in this paper can not only accurately locate the bone damage, but also improve the recognition ability of the relatively small-sized bone damage target of the ankle joint. At the same time, it can extract the geometric information of the ankle joint space, which provides an important physical reference basis for orthopedic surgeons to diagnose muscle and ligament injuries.

Figure 46, Table 14, Reference 66.

Keywords: foot and ankle joint disease; target recognition; semantic segmentation; auxiliary diagnosis system; attention modeling, muscle and ligament injury

摘要	I
Abstract	III
第1章绪论	1
1.1. 研究背景及意义	1
1.2. 国内外研究现状	2
1.2.1. 骨科疾病相关数据集	4
1.2.2. 相关技术分析	4
1.3. 本文研究难点和主要内容	7
1.3.1. 数据特性以及难点	7
1.3.2. 本文主要研究内容	8
1.4. 文章结构安排	10
第2章基于卷积神经网络的踝关节骨损分类	12
2.1. 踝关节骨损分类背景及意义	12
2.2. 主流卷积神经网络分类算法介绍	13
2.3. 实验方法与设计	15
2.3.1. 实验设计及过程	15
2.3.2. 评价指标	16
2.4. 实验结果与分析	17
2.5. 本章小结	
第3章 踝关节小目标骨损识别研究	19
3.1. 问题分析	19
3.2. 基于目标识别算法的踝关节检测	20
3.2.1. Resnet-RPN 网络结构	20
3.2.2. 实验设计与评估指标	23
3.2.3. 实验结果与分析	24
3.3. 基于注意力机制的小尺寸骨损检测	26
3.3.1. 注意力机制	26
3.3.2. 融合注意力机制的目标识别网络	
3.3.3. 实验方法与设计	35
3.3.4. 实验结果与分析	
3.4. 本章小结	40
第4章 基于语义分割网络的踝关节缝隙提取	41
4.1. 踝关节缝隙提取背景与意义	41

目录

4.2. 踝关节缝隙语义分割数据集的构建	42
4.3. 主流医疗语义分割算法	43
4.4. 实验方案与评估指标	44
4.4.1. 实验方案	44
4.4.2. 训练过程	45
4.4.3. 评估指标	46
4.5. 实验结果分析	47
4.6. 本章小结	52
第5章 足踝关节骨损辅助检测系统设计与实现	53
5.1. 系统功能	53
5.1.1. 核心功能	53
5.1.2. 全自动骨损伤检测定位功能	54
5.1.3. 韧带肌肉损伤检测功能	55
5.1.4. 专家联合会诊功能	55
5.2. 系统设计与实现	56
5.2.1. 客户端实现方法以及主要工作流程	56
5.2.2. 服务端实现方法和工作流程	57
5.3. 本章小结	59
第6章总结与展望	60
6.1. 论文总结	60
6.2. 研究展望	61
参考文献	62
攻读学位期间主要研究成果	69
致谢	70

第1章 绪论

1.1. 研究背景及意义

近年来骨质疏松性骨折,逐渐成为影响老人健康的主要问题,在生活中对 于老年人它们易于发生,然而治疗变得越来越昂贵和复杂,并发症也难以预防。 根据 Charles M. Court-Brown[1]的研究可知成人骨折的流行病学变化迅速,该团 队在 2000 年对 5953 处骨折病例进行的分析表明老年患者骨折的发生率逐年增 加。通常普遍认为骨质疏松性骨折主要见于胸腰椎,股骨近端,肱骨近端和骨 远端,但数据分析表明,现在应将 14 种不同的骨折视为潜在的骨质疏松。男性 中约 30%的骨折,女性中 66%的骨折和 70%的住院患者的骨折可能患有骨质 疏松症。根据 Gele B.Moloney[2]的研究可知,低能股骨远端骨折常发生在脆弱 的老年人群中,在对四肢骨折进行手术治疗后,容易发生局部和全身性并发症。 根据 xU Hal-lin[3]的调查可知,每年踝关节骨折的发生率为 1.8%,占所有骨折 的 3.92%, 而踝关节骨折患者通过治疗大约 3 天后, 疼痛减轻, 可以开始进行 膝部和踝部的功能锻炼。6-8 周后, x 射线检查显示骨折开始愈合, 可以开始负 重运动,大约 3 个月才能正常负重。根据王智勇[4]等人的研究可知,临床骨科 中最多见的疾病是骨折,基于 x 光片的检测方法是临床诊断中检测骨折的最重 要的手段。在实际生活中,踝关节对人体的活动具有非常重要的作用,由于其 活动较为频繁,所以很容易出现扭伤或者骨折的情况。根据[5]可知,在临床诊 断中,踝关节损伤通常需要骨科医生借助 x 线影像进行临床症状观察以判断损 伤程度,但是这种方式不能很好的全面反映踝关节创伤的真实情况,容易延误 患者病情。

医疗行业长期存在优质资源分配不均,医疗费用成本过高,诊断误诊漏诊 率较高,医生培养周期长,医疗资源供需缺口较大等多种问题。在实际的医疗 诊断过程,由于一些不可抗拒的因素如临床医生的疲劳、分心和时间压力[6], 造成一线医生在诊断骨损伤时极易出现误诊漏诊的现象,进而影响最终的治疗 效果。近些年随着深度学习网络的不断进步,使得自动检测骨损成为可能,其 可以检测出 x 光片中病灶部分,供临床医生作为诊断参考,缓解骨科医生的工 作压力,帮助医生提高工作效率,从而减少医疗失误,在减少误诊漏诊上具有 临床意义[7]。基于神经网络的医疗辅助诊断系统也可以助力缓解医疗资源的紧 缺问题以及缓解区域医疗资源不平衡的现状,对于医疗资源相对落后的地区, 自动检测骨折系统可以结合互联网来弥补医疗资源短板,对于医疗资源密集区 域则可以起到患者分流导流的作用,减轻大医院的负荷。基于以上背景,本课

题立足于研究效率更高,性能更好的踝关节骨损疾病自动检测系统。

具有针对性的高质量数据集是研发辅助诊断系统的关键,就骨科数据而言, 常见的数据为骨龄数据集,骨骼肌肉损伤分类数据集,骨损分类数据集等等。 踝关节骨损的数据比较少,带有小尺寸骨损的踝关节数据集较为少见,踝关节 缝隙语义分割的数据集几乎没有。然而小尺寸骨损的识别和关节缝隙的提取, 恰恰能够体现出辅助诊断系统在临床应用中给医生带来的价值。此外,就骨损 检测而言,目前骨骼智能辅助诊断领域的研究大多数是对骨损进行分类,缺少 对骨损定位的研究,特别是如何提高对小尺寸骨损目标定位方面的研究,对骨 骼几何物理信息提取方面的研究同样也极为稀少。

基于以上问题,本文介绍了踝关节骨损数据集的构建过程,其覆盖了小目标骨损类型数据集和一个高质量的语义分割数据集。基于完全粉碎性骨损的检测意义不足,但对算法干扰性大的问题,本文通过训练二分类模型,将数据集分为完全粉碎型骨损和非完全粉碎型骨损,只有非完全粉碎型骨损才进入骨损检测和踝关节缝隙提取部分,减少了完全粉碎型骨损给系统带来的干扰,提升了系统在完全粉碎型骨损伤的检测能力。

1.2. 国内外研究现状

近年来,随着 GPU 计算能力的提升以及高质量有效数据的增加,基于大数 据的神经网络得到了极大地发展,先后出现了以 AlexNet、VGGNet、ResNet 等 为代表的一系列优秀神经网络。同时在智慧医疗领域,越来越多的研究利用深 度神经网络来进行深度特征的提取。在 Litjens G[8]调查的综述中,详细概述了 深度学习在每个医疗领域的运用,如神经,视网膜,肺,数字病理学,乳腺, 心脏,腹部,肌肉骨骼等等。Rajpurkar P 在[9]中提出 CheXNet 网络来学习胸部 X 光片特征,辅助诊断肺炎,并且识别精度优于人类医生。在脑肿瘤领域中, Dong H 等人[10]提出利用 U-Net 网络在脑部 MRI 影像中识别提取并分割出病患 区域,识别分割结果可以和医生相媲美。Esteva[11]、Haenssle[12]、Gutman[13]、 Jianpeng Zhan 等人[14]将人工智能算法用于皮肤疾病诊断领域,为开创皮肤疾 病智能诊断领域做出了巨大的贡献。

结合深度学习的骨科辅助诊断的研究日趋完善,国内的汇医慧影公司通过 深度学习算法实现针对手骨、足骨等骨折和隐匿性骨折的识别检测;胡颜民等 人[15]通过图像处理计算从医学图像中分割出软骨,进而区分骨质与尿酸盐结 晶区域;贾文慧等人[16]基于 XGBoost 算法完成了骨科辅助诊断分类预测模型 的建立。在国外,Lindsey 等人[17]证明了深度学习对于辅助临床医生在骨折诊 断上有实际的帮助与深远的意义。Chung S W[18]团队利用卷积网络识别肱骨近

端骨折,表现出较好的分类结果,top-1 准确性为 65-86%,AUC 为 0.90-0.98, 敏感性/特异性为 0.88 / 0.83-0.97 / 0.94。与人类人群相比,CNN 的表现优于普 通医师和骨科医生,与专门从事肩部的骨科医生的表现相似。Olczak[19]团队将 VGG 等五种网络用于腕部,手部和踝部 x 光片骨折的识别上,在识别侧面,身 体部位和检查视图时,所有网络均显示至少 90%的准确性。对于表现最佳的网 络,骨折的最终准确性估计为 83%。当以相同的分辨率显示图像时,该团队的 网络模型的性能表现类似于高级骨科医生。

为了使研究更具针对性,本文收集了特定部位骨折诊断的相关文献。在踝 部骨折的诊断分析上, Jakub Olczak 等人[20]根据 AO / OTA 2018 分类协议将踝 部骨折进行分类,基于 ResNet 架构的神经网络训练出分类器,分类效果可以和 医生媲美; Kitamura[21]等人结合小样本与多视图的技术, 实现基于深度网络的 踝关节骨折的分类识别。对于髋关节, Yamada 等人[22]使用深层卷积神经网络 对髋部骨折进行自动分类,实现了具有骨科医师水平准确性的分类性能。 腕关 节作为骨折最易发生的部位,也有不错的研究,如 Gan 等人[23]利用 CNN 实现 了对腕部 X 光片上骨折的检测, Ebsim R 等人[24]结合正侧位的腕部图像实现对 于腕部骨折的识别。对于膝关节, Lefevre[25]等人基于 Mask RCNN 实现了对于 膝关节骨折的自动检测与定位; Prasoon A[26]等人则对膝关节软骨进行特征的 学习以进一步的辅助诊断。针对股骨与胫骨骨折, Matthew 等人[27]基于深度学 习对股骨胫骨骨折检测并且在诊断效果上与人类进行对比; Urakawa 等人[28]则 研究股骨转子的自动诊断并且达到专业骨科医生的标准。对于脊椎部位,Brett 等人[29]基于传统数字图像处理方法设计了一套增强脊椎骨折检测的工具大量 研究; Sa, Ruhan 等人[30]使用 Faster RCNN 对椎间盘进行定位以辅助骨科医生 的临床诊断。Heimer 等人[31]证明了深度学习在颅骨骨折识别分类中应用的可 行性。

除了对于骨折的诊断,深度学习在骨科领域也有一些其他应用:如在骨质 疏松方面,Tomita 等人[32]设计了对骨质疏松进行识别的深度网络;而 Stern D 等人[33]则基于手部 MRI 图像,完成了对于骨龄的自动估计。

虽然骨骼辅助诊断研究在多个部位的骨损识别上取得了一定的进步,但是 还有一些细分的研究点需要去完善和实现,比如使用迁移学习对网络效果进行 优化与改进[34],再比如小目标骨损对象的识别问题,骨骼几何信息的提取和 利用方面的研究,本文针对这些细分问题以及难点进行了相关的研究。

1.2.1. 骨科疾病相关数据集

现有公开可获取的骨科数据集主要是以 X 光片图像为主, 整理如表 1-1 所示:

表 1-1 现有公开可获取的骨科数据集

数据集	部位	标签	图像数目
MURA[35]	上肢骨骼	是否正常	40561
Pediatric Bone Age (AIMI) [36]	手部骨骼	骨龄	14236
0.E.1 (OAI) [37]	膝部骨骼	K&L 等级	8892
ChestX-ray8[38]	胸部	多种病理分类	112120
OpenI[39]	胸部	多种病理分类	7470
MC[40]	胸部	是否正常	138

其中最著名的是 Rajpurkar P 等人[35]收集的一个包含有 40561 张数据的骨 骼数据集,其包括手指、前臂、肩部,手肘,肱骨,手掌等上肢部位的骨骼 x 光片。但是该数据集的标签仅仅只能用来训练分类网络,并且数据成像质量一般。从表中可以看出,骨科数据以分类标签居多,比如预测是否有肌肉损伤,预测骨折的 K&L 等级,预测病理类别等等。用于骨损检测的数据集较为少见。

1.2.2. 相关技术分析

本文重点运用了目标识别算法以及语义分割算法。其目标识别是计算机视觉领域中一个重要的研究方向,在无人驾驶、人脸检测等领域有着非常广泛而 丰富的应用场景。其算法主要分为传统目标识别算法以及基于深度学习的目标 识别算法,经典的 Faster R-CNN 和 YOLO V3 目标识别算法均属于后者。目标 识别算法主要解决对图片中前景目标进行定位和分类的问题,目标检测通过预 测目标位置坐标来实现目标定位,获得预测边界框后,目标检测利用分类器, 预测边界框内前景目标的类别,从而实现目标的定位和分类。语义分割则主要 研究的是像素级分类问题,也是图像处理和计算机视觉领域中一个重要研究方 向,在医学图像分析、机器人感知、视频监控等领域发挥着极其重要的作用, 经典的算法有 FCN[41]、U-net[42]等等。

1.2.2.1. 传统目标识别算法

传统目标识别算法一般分为区域选择、特征提取、分类器分类三个步骤。 对于输入图片,通常先用滑动窗口获取许多不同尺寸的窗口,用来选取特征候 选区域。然后依赖人的先验知识设计有效的特征,方便从数据中提取特征。最后用提取的特征训练分类器,不同的任务适用不同的分类器,其中支持向量机(Support Vector Machine,SVM)[43]是一种常见的二分类器器。

传统的目标识别算法虽然简单,针对特定问题可以得到不错的识别性能, 但是由于采用了滑动窗口获取特征区域,然后再进行遍历搜索,因此实际运行 中该类算法实时性不高,无法满足现在大数据低时延的需求。并且传统目标识 别算法非常依赖人的先验知识,往往在相同的任务上,不同的特征设计者设计 出的算法性能差异非常大,导致该类可移植性差、试错成本高。

1.2.2.2. 基于神经网络的目标识别

深度神经网络(DNN)的出现[44]-[45],促进了目标识别算法的发展,比如 R-CNN[46]通过引入卷积神经网络 CNN 来自动提取特征,发展出了可以实现分 类和边界框回归任务共同优化的 Fast-RCNN [47]和需要区域预测网络(Region Proposal Network, RPN)生成 ROI 的 Faster R-CNN[48]。还有可通过深度神经 网络一次性生成对象坐标和类别的 YOLO 算法[49]-[51]。其中以 Faster R-CNN 为代表的 R-CNN 系列目标识别算法基于分类思想,以 YOLO 为代表的目标识 别算法则基于回归思想。

基于分类的目标识别算法继承了传统识别算法的思想,即分类和定位是分 开进行的。首先通过提取图像的特征来产生建议区域(Region proposal, RP), 然后对建议区域进行进一步的分类和边框回归。R-CNN[46]、Faster R-CNN [48]、 空间金字塔池 spp[52]、基于区域的全卷积网络 R-FCN[53]、特征金字塔网络 FPN[54]等模型都循这样的思想。

R-CNN[46]算法首先依赖 Selective Search[55]方法在输入图像上产生 N 个建 议区域,然后用 AlexNet[56]神经网络从每个建议区域自动提取抽象特征,生成 一个 Nx4096 的特征向量。特征向量经 SVM 分类处理后,用非极大值抑制 (Non-maximal Suppression, NMS)算法剔除非目标的区域留下最好的建议区域, 最后对剩下的建议区域进行回归操作,得到最终目标预测边界框。

为了解决 R-CNN 计算量大的问题, Fast R-CNN [47]算法利用 CNN 提取出 图像的高层特征,减少候选框的数量来进行优化。具体来说深度卷积神经网络 会对 Selective Search 算法产生的建议区域的特征,进行分类和边框回归。Fast R-CNN 算法相比于 R-CNN 算法,检测过程变得简单,计算量也得到了很大的 缩减。

Faster R-CNN[48]用区域建议网络(RPN)代替 Fast R-CNN[46]中的 Selective Search 方法来产生感兴趣区域,其会输出一个前景的预测边界框和一

个前景置信度,可以实现更快速高效的定位。同时为了减少边界框的回归难度 提出了 Anchor 机制,即在区域建议网络(RPN)中预设三种尺寸三种比例的参 考框以适配不同尺寸的目标。Faste R-CNN 在 R-CNN 的基础上,减少了重复计 算,交替训练,实现了特征共享。

基于回归目标识别算法不需要把边界框预测和分类分开,其使用深度神经 网络直接从图像中预测出边界框的坐标、类别及其边界框前景置信度。因此目 标检测算法的检测速度一般相对基于分类的目标识别算法速度较快,如: YOLO[49]模型的训练以及检测速度比 Faster R-CNN 快。

YOLO[49]算法广泛应用于对于实时性较高的任务中,如视频中追踪目标。 其将图像分成独立的单元格,每个单元格会生成若干个边界框坐标,用于预测 中心落在该单元格内的目标物体边界,同时还会预测出每个边界框的置信度和 类别。YOLO 以每个单元格为中心进行多尺度融合检测,以此取代建议区域网 络的兴趣区域预测,从而极大地提升目标检测速度。

1.2.2.3. 基于深度学习的语义分割

随着深度学习的发展,语义分割算法也得到了长远的发展,不同于早期的 语义分割算法基于灰度、条件随机场,基于深度学习的语义分割融入了语义和 实例级分割的大量开创性工作,包括端对端的卷积网络、编码解码器架构、多 尺度、金字塔方法等等,这些开创性工作极大地促进了基于深度学习的语义分 割算法的发展。

FCN[41]网络的提出是图像分割领域的一个里程碑,它端到端的训练方式证 明了神经网络可以在不同尺寸的图像上进行分割。FCN 原于分类网络,对于一 般的基于深度学习的分类网络,因为最后的分类依赖于全连接层,所以只能产 生一维的概率信息,不能产生标识每个像素点类别的二维矩阵信息。FCN 在分 类网络的基础上把后面几个全连接都换成了卷积,获得了二维的特征图,然后 由 softmax 层对每个像素点进行分类,从而实现了图像分割。FCN 已经应用于 多种分割问题,如脑肿瘤胶质瘤的分割、皮肤损伤分割等等。FCN 虽然开创了 深度学习语义分割的先河,有效性也得到了验证,但是固有的局限性仍然不可 忽视。比如由于本身的网络结构简单,FCN 不能有效地考虑全局上下文,也不 能快速地进行实时推理。

基于编码器解码器体系结构的语义分割算法是另一个流行的图像分割的深度模型,该体系结构在医学领域应用较多。Ronneberger 等人[42]最早提出的 U-net 结构就是一种用于在显微镜下分割细胞的卷积编码解码架构。该结构的编码器将输入图像进行一个卷积加池化的下采样操作,解码器通过反卷积进行上采

样操作,同时进行通道维度的特征拼接实现多尺度的特征融合。随着行业的发展,为了实现更多的功能,U-net 加入了新的模块融入了其他的设计理念。比如为了适应三维的数据的分割,U-Net 扩展为 3D U-Net[57],该结构在 2018 年 MICCAI 脑肿瘤分割挑战赛中取得了第二名的成绩。Ibtehaz[58]在 2019 年提出 将 MutiRes 模块与 U-Net 结合,该模块是一个残差连接的扩展,它将三个 3×3 的卷积结果拼接起来作为一个特征图,然后与输入特征相加。Alom[59]在 2018 年提出 R2U-Net,该方法将残差连接和循环卷积结合起来替换 U-Net 中的卷积 子模块,该方法在皮肤病图像、视网膜图像、肺部图像等几个公共数据集上获 得了不错的性能。Oktay[60]在 U-Net 中引入了注意力机制,提出了 Attation U-Net,在背景信息复杂目标较小的情况下提高了图像分割的精度。

1.3. 本文研究难点和主要内容

本文聚焦于踝关节骨科疾病的辅助诊断研究,不同于一般的骨折等级分类 方案,本文的辅助诊断系统可以实现骨损的精确定位以及对踝关节缝隙进行提 取。

本文原始数据分辨率高,携带丰富的骨骼信息,但骨损目标占全图比例小。如 果直接进行检测容易发生误检漏检。并且本文的检测目标都集中在踝关节部位, 而原始图像包含了整个小腿以及踝关节,因此为了能更好的识别出踝关节骨损, 本文提出通过预先训练踝关节识别模型,截取高价值踝关节的预处理方式,来 提升系统对相对小尺寸骨损的检测能力和踝关节缝隙的识别能力。该方案的好 处是通过第一阶段的目标识别缩小了第二阶段的检测范围,提升了第二阶段目 标的检测精度,减小了小目标的漏检误检概率,同时也可以提升语义分割的识 别精度。

1.3.1. 数据特性以及难点

本课题数据由湘雅三医院放射科以及骨科合作提供,本文团队进行整理标注,共同收集了1852 张数据,包括1246 张非完全粉碎的踝关节以及 606 张完全粉碎的踝关节。原始数据为高分辨率的小腿踝关节 x 光片,携带丰富信息,成像清晰,拍摄角度主要为正视位和侧视位。该数据集包含了临床上出现的大部分足踝关节骨折的类型以及各种尺度的骨折程度。有已经完全粉碎脱落的踝关节如图 1-1 中 b、c、d 所示。在非完全粉碎的踝关节中骨损尺寸大小不一,有的相对比例适中有的则比例较小如图 1 中 b、c、d 中黄框所示,特别是图 1-1d 中骨折是一条细缝。数据原始

尺寸太大,基本在 2000×2500 分辨率以上,如果直接输入到网络中,将会导致 GPU 内存不足,如果将尺寸进行压缩,网络容易误检漏检。



图 1-1 非完全粉碎型骨骼和完全粉碎型骨骼数据示例, a 为完全粉碎型骨骼, b、c、d 为非 完全粉碎型骨骼

1.3.2. 本文主要研究内容

针对本文数据集的图像特点,本文提出了一种多阶段的辅助诊断方案有效 地解决了相对小目标骨损的识别问题。如图 1-2 所示为拟采取的技术路线整体 框架,本文系统的整体框架分为多个阶段。



图 1-2 拟采取的技术路线整体框架

首先鉴于完全粉碎型骨损的踝关节已经没有必要再进行伤损检测了,因此 第一阶段本文训练了一个二分类模型,将输入图片分为非完全粉碎型骨损和完 硕士学位论文

全粉碎型骨损,只有非完全错位的数据进入下个阶段。在该阶段本文训练了 VGG-16[61]、GoogleNet[62]、ResNet[63]三种分类网络的二分类模型,用于识 别数据中的非完全粉碎型骨损和完全粉碎型骨损,其中正常健康的骨骼被归类 为非完全粉碎型。然后由于本数据集数据尺寸大、分辨率高、踝关节骨损目标 比例小,一般的目标识别算法容易漏检误检,小目标检测性能不佳,因此研究 探讨了缩放和截图两种不同的输入数据预处理方式给网络识别带来的性能差异。 通过实验可知在本数据上预先训练一个识别踝关节的模型的预处理方式在提升 小目标检测性能方面具有很好的正向作用,如图 1-3 所示。



图 1-3 输入预处理从整个图片中提取踝关节

最后本文在不同目标识别算法中添加注意力机制,研究注意力机制模块是 否能提升小目标的识别性能,部分识别结果如图 1-4 所示,其中性能最好的模 型是添加了注意力机制 CBAM 的 YOLO V3,在预测方框的实验中,精确率为 0.9324, 召回率为 0.8415, F1 为 0.87。在判断是否有骨损的结果中,准确度为 0.9478,精确率为 0.9787,召回率为 0.9324, F1 为 0.9550。



图 1-4 小尺寸骨损识别模型预测结果样例

提取关节间隙的几何特性对辅助临床医生诊断同样具有重要意义,正常的 踝关节 x 光片中腓骨、胫骨、踝骨的边界清晰关节间隙规则,关节间隙上下边 界大体平行如图 1-5 所示,左边为正常的踝关节间隙图,右边为关节间隙提取 图,而肌肉损伤或者韧带损伤会导致关节间隙出现变形。本文将使用语义分割 算法来提取该部分,医疗的语义分割相对自然物体的语义分割对精度和边界清晰度有更高的要求,因此本文对比实验了医疗领域常用的语义分割算法,U-net[41]和 U-net++[71]两种网络。



图 1-5 正常的踝关节间隙(左)以及关节间隙提取图(右)

1.4. 文章结构安排

本文将分 6 个章节对足踝关节骨损疾病辅助系统进行详细介绍,各章节的 主要内容如下:

第1章,绪论。本章节首先阐述了骨科辅助诊断系统研究的背景和意义,然 后通过深入调查国内外骨科相关的研究归纳总结了一些骨科可以研究的方向, 介绍了本文的数据集的特点以及介绍了本文的研究难点和相关技术分析,最后 说明了本文的组织结构。

第2章,骨损二分类问题的研究。首先介绍了本文数据集的数据组成成分,介绍了骨损二分类的意义,然后介绍了VGG[61]、GoogleNet[62]、ResNet[63] 三个经典分类网络的原理,比较了不同网络在 ImageNet 分类数据集上的性能差异。接着设计了基于 VGG、GoogleNet、ResNet 的二分类实验,最后探讨了三个网络在本文数据集上的实验性能差异。

第3章,基于目标识别算法的踝关节骨损特征提取研究。首先分析了踝关节 中骨损的特点,提出一般的目标识别算法直接识别小目标骨损的局限性。然后 介绍了注意力机制的原理和特点,提出利用注意力机制提升目标识别算法对小 目标的检测性能,并详细介绍了如何优化 YOLO V3 和 Faster R-CNN 两种网络,

同时探讨了不同输入预处理带来的差异。最后实验验证了不同预处理带来的性能差异以及注意力机制在目标识别网络中的有效性。

第4章,基于语义分割算法的踝关节缝隙特征提取研究。首先探讨了踝关节 缝隙的物理几何信息在肌肉韧带损伤诊断中的临床意义,提出用语义分割算法 提取该信息,介绍了像素级语义分割标签的情况。然后介绍了 U-net 网络和 Unet++网络的原理和差异,最后探讨了这两个网络在本文数据集上的分割性能差 异。

第5章,踝关节骨损自动检测平台系统设计及实现。首先阐述了自动检测平台的框架,然后介绍了不同部分的主要功能,最后着重介绍了终端 app 的界面和操作流程以及后端系统。

第6章,总结与展望。总结了本文的研究工作,对所提出的改进方法进行了 归纳,并展望了未来的研究方向。

第2章基于卷积神经网络的踝关节骨损分类

医疗实际环境中,骨损患者产生的数据特征分布具有一定的差异性。在利 用大量数据训练网络模型时,差异较大的数据往往会增加模型的复杂度,导致 拟合难度增加。为了让辅助诊断系统中的目标识别算法和语义分割算法更容易 拟合,本章采用基于深度卷积神经网络的图像分类算法对骨损特征进行深度学 习,为骨损辅助系统构建一个骨损分类模型,用于滤除自然数据中的干扰数据。

2.1. 踝关节骨损分类背景及意义

在本文 1852 张数据中,1246 张为非完全粉碎型骨损,606 张为完全粉碎型 骨损,完全粉碎型骨损占总数据的三分之一。其中完全粉碎型骨损是指踝关节 受不可抗的外力,导致胫骨、腓骨或踝骨出现面积超 30%破损的病例,其失去 了原有的物理几何形态。如图 2-1 所示为完全粉碎型骨损病例样本,图 2-1 中 a、 b、d 腓骨严重折断,胫骨腓骨踝骨三骨发生严重的位置偏移,c、d、e 的踝关 节缝隙完全被遮蔽。在临床治疗上,完全粉碎型骨损和非完全粉碎型骨损遵循 不同地检测治疗方案。如果在定位骨损阶段和语义分割踝关节阶段,训练网络 模型时不做区别地利用两种数据,其模型的复杂度和训练难度都会增加,但模 型的准确度会降低。因此在本文中,辅助诊断系统的第一阶段将训练一个踝关 节骨损分类模型,用于区分非完全粉碎型骨损和完全粉碎型骨损,只有非完全 粉碎型的数据才能进入下个辅助诊断阶段。该阶段为整个辅助诊断系统滤除掉 数据集中的完全粉碎型骨损数据,降低了任务复杂度,减少数据干扰项,使得 系统的模型复杂度减少,提升了整个系统的鲁棒性。



图 2-1 完全粉碎性骨损病例样本图

2.2. 主流卷积神经网络分类算法介绍

图像分类问题并不是一个新课题,相应的算法一直在发展,在深度神经网 络算法还没有得到充足发展的早期,就有科学家们通过设计图像预处理、特征 提取和分类器来实现图像的分类。图像预处理一般为中值滤波、均值滤波、高 斯滤波等等,这些滤波器过滤掉图片中无关的信息保留有用的信息,使图片特 征具有低纬度、低噪声等特点。早期图像分类算法的特征提取主要依赖于人工 设计特征,比如纹理、颜色、形状等等,这种特征提取方式的好坏取决于特征 设计者的先验知识,因此普适性差,并且随着数据量的增加以及维度的增加人 工设计特征的难度呈指数级增加。传统的分类器包括 BP 神经网络、决策树、 支持向量分类器等等,虽然这些分类器提升了图像分类的性能,但是对于处理 复杂量大的图像分类还是无法满足需求。

随着算力和算法的进步,深度学习在提取图像特征方面显露出的优良性能 引起了研究者们的关注。比如 Alex Krizhevsky 等人[56]提出来的 AlexNet 网络 夺得了 2012 年 ImageNet ILSVRC 比赛的冠军,其 top-5 错误率为 15.3%远低于 当时的第二名 26.2%,该网络的提出在深度学习的发展史中是个重大的转折点, 使得卷积神经网络成为了计算机视觉的重大核心突破领域。深度神经网络相较 于传统的图像分类算法,其不需要人工设计特征和提取,而是通过神经网络在 大量的训练样本中自动学习样本特征,抽取出更有效的特征,减少了对算法人 员先验知识的依赖,提高了算法的普适性。同时这些特征提取网络与分类器联 系紧密,解决了分类器与特征不匹配的问题。

近年来深度神经网络发展迅速,在 AlexNet 提出后,研究者们针对深度神经 网络的许多问题做了很多的优化改进,使得基于深度卷积神经网络的分类算法 性能得到了非常大的提升,如下表 2-1 所示为多种经典网络在 ImageNet 数据集 上的性能对比以及相关的改进特点。

AlexNet 网络结构包含了 5 个卷积层和 3 个全连接层,总共有 8 层,因为 Image Net 有 1000 个类别,所以 AlexNet 在最后一层是一个 1000 个节点的 Softmax。当 AlexNet 在 ImageNet ILSVRC 比赛中获得分类方面第一名以后,研 究者们注意到增加网络的深度可以让模型能获取更多的特征且更加稳定。因此 在 2014 年 Simonyan 等人[61]提出了 VGG 网络,该网络在 ImageNet ILSVRC 分 类比赛中获得第二名,VGG-16 的错误率为 8.43%,与 AlexNet 的 8 层结构不同, VGG 增加了网络深度,常用的结构有 16 层卷积层加 3 层全连接层的 VGG-16 和 19 层卷积层加 3 层全连接层的 VGG-19,并且 VGG 所有卷积层均使用两个 3 ×3 的卷积核,这样在未改变视野域的前提下,两层的卷积核比一层的卷积核

多一次非线性变化,当网络深度在一定范围内增加时可以保证学习更复杂的模型,并且采用小卷积核的卷积堆可以减少网络参数数目。但由于全连接层的大量使用,VGG 耗费更多的计算资源并且使用了过度的参数,其中绝大多数的参数来自第一个全连接层,比如 VGGNet-16 的参数数量为 1.38×10⁸ 个,AlexNet 的参数数量为 6×10⁷ 个,GoogleNe 参数数量为 5×10⁶ 个。在同等的任务和计算资源下 VGG-16 每秒浮点运算次数为 1.5×10¹¹ 次,AlexNet 每秒浮点运算次数为 7×10⁹ 次。对于 VGG 框架随着卷积网络的加深,会发生梯度消失的现象导致模型难以优化。

时间	卷积神经网络	Top5 错误 率(val)	FLOPS(每秒浮 点运算次数)	特点	
2012	AlexNet[56]	15.4%	7×10 ⁹	首次将深度卷积神经网络运用于 图像分类问题上,开创了深度卷 积神经网络新进程	
2014	VggNet-16[61]	8.43%	1.5×10^{11}	探索了不同网络深度的性能差异	
2014	GoogleNet[62]	7.89%	1.5×10^{10}	增加特征网络的深度宽度减少了 网络参数,提升了网络内部计算 资源的利用	
2015	ResNet- 152[63]	4.49%	1.13×10 ¹¹	引入深度残差框架解决网络深度 增加带来的网络退化问题	
2017	SENet-152[64]	4,47%	1.13×10^{11}	根据每个特征通道的重要程度决 定特征通道的权重	

表 2-1 各个 CNN 图像分类模型在 ImageNet 数据集上的性能比较

由 Szegedy 等人于 2014 年提出的 GoogleNet[58] 在 2014 年的 ImageNetILSVRC 比赛中获得分类方面第一名,其 top5 的验证集错误率为 7.89%,超过了同年一起提出的 VGG。网络采用了模块化的结构,并且采用平 均地化层代替全连接层,减少了全连接层参数多的问题,减轻了模型训练难度, GoogleNet 模型参数数量为 5×10⁶ 个比 VGGNet-16 的 1.38×10⁸ 个参数数量少 了两个数量等级,其每秒浮点运算次数为 1.5×10¹⁰ 次,模型参数比 2012 年提 出的 AlexNet 网络少 12 倍,但准确率提高了两倍。为了提升特征提取能力 GoogleNet 加深了卷积层的深度,GoogleNet V1 总共有 22 层,包括 21 个卷积层 和一个 Softmax 层,对于一般的深度卷积神经网络,提升卷积深度容易出现梯 度弥散问题,

也就是梯度越往后越容易消失,GoogleNet为了避免梯度消失,网络额外增加了两个辅助用的Softmax 层用于前向传导梯度。GoogleNet 使用 Dropout 解决 过拟合问题,并引入 Inception 结构融合不同尺寸的特征信息,提升了网络对内 部资源的利用率,可以做到用更少的参数实现更好的性能。

在普通的网络中,对浅层网络叠加卷积层,可以使模型复杂度更高,可拟 合的映射关系更多,所以同等的训练集和测试集上,模型的会有更好的性能表 现。但当卷积层数加到一定程度后,会出现网络模型性能退化的现象,即随着 网络深度的增加训练集的损失函数不降反增。为了解决该问题,Kaiming He 等 人[63]于 2016 年提出了一种 ResNet 网络,该网络解决退化问题的思路是调整模 型结构,让模型更易于优化。具体的做法是提出了一种残差块,然后由残差块 搭建网络。该方案训练模型要比通过一堆非线性层来匹配一个恒等映射要容易 很多。Kaiming He 使用 ResNet 在 ImageNetILSVRC2015 比赛中成功训练出 152 层的神经网络,取得当年的分类组冠军,其模型在验证数据集上 top5 的错误率 为 4.49%,验证集上的 top5 错误率为 3.57%,同时因为减少了全连接的使用, 模型参数比 VGGNet 少。得益于残差块,在保证网络不退化的前提下,现可以 搭建突破 1000 层的超深网络结构。在 ResNet 中使用 Batch Normalization 代替 Dropout 层,加快了模型的训练。

2.3. 实验方法与设计

2.3.1. 实验设计及过程

本文收集的踝关节数据集共有 1852 张有效数据,其中包含 1246 张非完全 粉碎型骨损和 606 张完全粉碎型骨损数据,本实验的训练集、验证集、测试集 的数据量以近似 3:1:1 的比例划分。1246 张非完全粉碎型数据选取 747 张为训 练集,249 张为验证集,250 张为测试集。606 张完全粉碎型数据选取 363 张为 训练集,121 张为验证集,122 张为测试集。本实验中的训练集总共有 1110 张 数据,验证集有 370 张数据,测试集有 372 张数据。数据原尺寸过大且不统一,为了避免出现 GPU 内存溢出,训练时将图像尺寸统一为 1000×1000。在该数 据集上本文对比了 VGG-16、GoogleNet、ResNet-50 三种网络的性能,实验中 损失函数统一设定为二元交叉熵,公式如式 2-1 所示,p 为预测正类的概率,1-p 为预测负类的概率。y 为标签值,在本文数据中非完全粉碎骨骼标为 1,完全 粉碎骨骼标为 0。

$$L = -(y * \log(p) + (1 - y) * \log(1 - p))$$
(2-1)

本节所有分类网络代码均基于 Facebook 开原深度学习框架 Pytorch 实现, 在训练阶段,使用自适应矩估计优化算法 Adam (Adaptive Moment Estimation) 更新网络参数,设置学习率为 lr=0.0001。每个 batch 随机选取 5 个训练样本, 在一块 Nvidia Geforce RTx 2060 SUPER 显卡上训练 200 个 epoch 后停止。 2.3.2. 评价指标

本章使用固定的置信度阈值计算精确率(Precision)、召回率(Recall)、准确度(Accuracy)、F1 Score 用来评估模型的分类能力,设置不同的置信度绘制 ROC 曲线,更直观地评估模型的分类效果。

计算以上评估指标之前,需要先根据模型预测结果构建一个混淆矩阵。如下表 2-2 所示,其中包含真正例(True Positive, TP),真反例(True Negative, TN),假正例(False Positive, FP)和假反例(False Negative, FN)。

	·····	
Confusion Matrix	预测正例(Positive)	预测反例(Negative)
真实正例(Positive)	真正例 (TP)	假反例 (FN)
真实反例(Negative)	假正例(FP)	真反例 (TN)

表 2-2 混淆矩阵

准确率(Accuracy)表示预测正确的样本占所有样本的比例,其公式如式 2-2 所示:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$
(2-2)

准确率在数据平衡时,可以表示出模型的分类性能。但当数据量不平衡时,准确率不能很好地代表模型好坏,比如模型把少样本类全分错,但是准确率仍然 可以很高。

精确率(Precision)表示所有预测为正例的样本中真实正例样本所占比例, 其公式如式 2-3 所示:

$$Precision = \frac{TP}{TP+FP}$$
(2-3)

召回率(Recall)表示在所有真实正例中预测正例所占比重,其公式如式 2-4 所示:

$$\operatorname{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$
(2-4)

精确率(Precision)和召回率(Recall)这两个指标是此消彼长的关系,当置信度阈值较高时,计算出的精确值要高,召回率会变低,反之亦然。因此,为了综合对比两种网络的性能,可以通过 F1 值和 P-R 曲线进行更好的评估。F1 值 是精确率和召回率的调和平均值,其公式如式 2-5 所示:

$$F_{1} = \frac{2}{\frac{1}{\text{Precision}^{+} + \frac{1}{\text{Recall}}}} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}}$$
(2-5)

ROC 曲线以不同置信度阈值下的 FPR 为横轴, TPR 为纵轴, 曲线包围的面积所占比重为 AUC 值, AUC 值越大曲线越趋向左上角, 代表模型性能越好。

2.4. 实验结果与分析

本章节在收集的骨科数据集上测试对比了 VGG-16、GoogleNet、ResNet-50 三种网络的性能,结果如表 2-3 所示:

表 2-3 VGG-16、GoogleNet、ResNet-50 三个模型在测试集上预测结果的混淆矩阵

	VGG-16		GoogleNet		ResNet-50	
	非完全粉碎	完全粉碎	非完全粉碎	完全粉碎	非完全粉碎	完全粉碎
非完全粉碎	242	8	245	5	244	6
完全粉碎	9	113	6	116	5	117

表 2-3 中统计了 VGG-16、GoogleNet、ResNet-50 三个模型在测试集上预测 结果的混淆矩阵。对于 122 个完全粉碎型骨损数据,GoogleNet 预测出 116 个正 确的结果,6 个被预测成非完全骨损。而 ResNet-50 预测出 117 个正确结果,5 个非完全粉碎型骨损。可以看出 GoogleNet 更容易把完全粉碎型骨损分为非完 全粉碎型骨损。对于 250 个非完全粉碎型骨损,GoogleNet 正确预测出 245 个数 据,有5 个数据错误地被预测为完全粉碎型骨损,而 ResNet-50 预测出 244 个 正确结果,6 个非完全粉碎型骨损。可以看出 ResNet-50 更容易把非完全粉碎型 骨损分类为完全粉碎型骨损。在本文的辅助诊断系统中,二分类网络只是在数 据输入的时候起到过滤完全粉碎型骨损干扰的作用。如果有个别完全粉碎型骨 损被判为非完全,在后阶段的目标识别任务中网络仍会判断该数据为有损数据, 因此不会影响到临床的诊断。但如果是非完全粉碎型骨损被判为完全粉碎型骨 损,那么该数据将不会进入后阶段的诊断,因此会影响到临床诊断。

VGG-16、GoogleNet、ResNet-50 二分类结果混淆矩阵评估指标的计算值如 表 2-4 所示:

表 2-4 VGG-16、GoogleNet、ResNet-50 三个模型在测试机上的评估指标

	Accuracy	Precision	Recall	F1 Score	AUC
VGG-16	0.9543	0.9642	0.968	0.9666	0.95
GoogleNet	0.9704	0.9760	0.98	0.9780	0.97
ResNet-50	0.9704	0.9800	0.976	0.9780	0.96

GoogleNet 和 ResNet-50 具有相同的准确率,均为 0.9704,而 VGG-16 的准确率为 0.9543。这是因为相比于 VGG-16,其余两个网络具有更深的特征提取 网络,因此可以拟合更有效的模型。从表中可知,GoogleNet 的精确率为 0.9760,

比 ResNet-50 的精确率 0.98 低了 0.004; 而 GoogleNet 的召回率为 0.98, 比 ResNet-50 的精确率 0.976 高了 0.004。从图 2-2 ROC 曲线图中可以看出, VGG-16 性能要要比其他两个网络差些, 其 AUC 值为 0.95, GoogleNet 的 AUC 值为 0.97, ResNet-50 的 AUC 值为 0.96。



图 2-2 VGG-16、GoogleNet、ResNet-50 在测试数据集上的 ROC

通过以上分析可知,对于骨损二分类问题,GoogleNet和 ResNet-50两个网络的性能要优于 VGG-16。GoogleNet和 ResNet-50相比两者性能接近,但是考虑到实际的临床诊断中,召回率高的网络可用性更高,因此 GoogleNet 在本实验中要优于 ResNet-50 网络。

2.5. 本章小结

本章主要围绕本文骨科数据集实现了一个二分类的模型,该分类模型可以 准确地区分非完全粉碎型骨损和完全粉碎型骨损。首先介绍了数据集的由来、 构成和数据特点,明确了骨损的类别,分析了完全粉碎型骨损对整个系统的影 响。然后介绍了分类网络的发展,分别介绍了 AlexNet 网络的特点以及历史意 义,VGG 网络和 AlexNet 的差别和缺点,以及 GoogleNet 网络和 ResNet 网络在 解决网络退化和梯度消失方面做出的改进优化,对比了多种网络在 ImageNet 数 据集上 1000 分类的性能差异。本章重点做了一个基于 VGG-16、GoogleNet、 Resnet-50 三个网络的骨损二分类实验,对比了三个网络的多个性能指标差异, 分析了三个网络的骨损二分类实验,对比了三个网络的多个性能指标差异, 分析了三个网络的性能差异,得出了 GoogleNet、Resnet-50 在骨损分类任务中 要优于 VGG-16 的结论。虽然 GoogleNet、Resnet-50 各项指标都很接近,但是 考虑到临床诊断中召回率高的网络可用性更高,因此本文辅助系统选用 GoogleNet 作为分类网络。

第3章 踝关节小目标骨损识别研究

本文数据库由高质量的 x 光片组成,每张图片内容包括整个小腿和脚踝。 图像分辨率一般为 2000×2500 以上,骨损主要分别在踝关节周围,骨损区域的 大小不一,绝大多数都在 300×300 以下,最小的骨损区域尺寸只有 50×50。 骨损区域相对原始图像比例小,临床诊断容易被忽视,在深度卷积神经网络的 不断下采样下,很容易造成这些小目标的细微特征在模型中被覆盖,从而出现 小目标漏检现象,导致模型整体的性能较差。

本章节旨在研究如何提升目标识别算法在足踝骨数据集上的骨损识别能力, 特别是优化解决相对原图片尺寸较小的骨损目标容易发生漏检的问题。基于骨 损位置较为固定的特点,本章首先采用目标识别算法实现对踝关节地识别,然 后将注意力机制融入到目标识别算法中用以优化目标识别网络的特征提取网络, 提升模型对小目标骨损的检测能力。

3.1. 问题分析

近年随着深度神经网络的发展,目标识别算法从传统的人工设计特征的目标检测算法发展出了以 Fast-RCNN 为代表的基于深度卷积神经网络的目标识别算法。虽然各种卷积神经网络融合到目标识别算法后带来了许多性能方面的改进,比如在 Fast-RCNN 中用 VGG16 代替 AlexNet 作为特征提取网络后,平均准确率达到了 70%,训练速度比 R-CNN 快了 9 倍,但是现有常见目标识别算法仍然不能很好地处理小目标识别任务。而在本文的骨骼数据集中,骨损大小相对整个图像尺寸比例较小,直接用目标识别算法检测容易漏检。

常见目标检测的研究使用的数据集主要有 ImageNet[64]、PASCAL VOC[66]、 MS COCO[67]等,这些都是比较权威的公开数据集。本文目标识别数据集由踝 关节骨骼数据构成,其骨损目标存在相对尺寸小的问题,即骨损小目标的尺寸 仅相对整个照片的尺寸较小,在深度神经网络中随着进行越来越深的下采样操 作,这些小目标的特征在整个网络中的比例会越来越少直至消失,最终导致漏 检问题。如下图 3-1 所示,为本数据库中常见的小尺寸骨损样例,黄色边界框 内前景目标是踝关节,绿色边界框内的前景目标是小尺寸骨损。

为了提升目标识别算法在小尺寸骨损目标上的检测性能,本文从不同的优 化角度出发提出了两种优化方式。一种是优化输入预处理方式,让小尺寸目标 的特征放大,即训练一个踝关节识别模型预先截取高价值踝关节,使得尺寸相 对较小的骨损目标比例放大,从而提升系统对相对小目标的检测能力;另一种 是优化特征提取网络结构,即在目标识别网络中添加注意力机制,增强小尺寸 骨损特征的空间和通道特征,衰减其他干扰特征,从而提升检测能力。



图 3-1 数据库中常见骨损样例

3.2. 基于目标识别算法的踝关节检测

3.2.1. Resnet-RPN 网络结构

本文 Resnet-RPN 网络由 Faster R-CNN[48]根据本文任务特点优化得到, Faster R-CNN 继承了 R-CNN 的设计思路,优化了 Fast R-CNN 用 Selective Search 生成候选目标区域的做法,直接利用区域预测网络(Region Proposal Network, RPN)生成 ROI,让整个网络可以完成端到端的训练,显著地提升了 算法的实时性。Faster R-CNN 主要由四个部分组成,分别是深度卷积神经网络、 区域预测网络 RPN、ROI 池化层、分类和回归网络。其结构图如下图 3-2 所示:



图 3-2 Faster R-CNN 整体结构图

深度卷积神经网络由一系列的卷积层池化层等等结构组成,其主要用于自动高效地提取数据中高层次高维度的特征,一般会用到现有的经典网络模型,比如 VGG、ResNet、Inception 等等,在本实验中本文使用 ResNet-50 作为特征提取网络。

区域预测网络 RPN 是 Faster R-CNN 最大的改进点,其作用是生成区域候选框。该网络输出两类信息,一类是目标的分类判决信息,即通过 Softmax 判断参考框(anchors)属于前景还是背景。前景表示的是该目标与真实目标重叠并且重叠比值大于 0.7,背景表示的是该目标不与任何真实目标重叠或者重叠比值小于 0.3。另一类是边框回归调整信息,即利用边界回归修正参考框位置获得目标的准确位置。特征图中参考框(anchors)的设定可以让 Faster R-CNN 有能力预测不同尺寸不同长宽比例的目标边界框,特征图中每个点都会预设 3 种尺寸 3 种比例共 9 种类型的不同大小长宽比例的参考框,每个参考框输入 RPN 网络中后都会预测出 6 个参数,两个参数表示前景和背景的预测概率,四个参数表示参考框与前景真实框的位置偏移量以及大小的缩放量。如下图 3-3 所示:



图 3-3 参考框计算示意图

RPN 的损失函数包括是二分类损失和回归损失,其公式如式 3-1 所示:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)$$
(3-1)

公式中 i 表示预选框的下标, P_i表示第 i 个预选框是前景目标的概率, P_i*表示第 i 个预选框的真实类别标签, 前景为 1 背景为 0, t_i表示预测边界框坐标的向量, t_i*表示前景真实边界框的坐标向量, λ 表示分类损失和回归损失的平衡权重系数, 在原文中设置为 10。L_{cls}在 RPN 中是一个经典的二分类交叉熵损失, 公式如式 3-2 所示:

$$L_{cls}(p_i, p_i^*) = -(p_i^* \cdot \log(p)_i + (1 - p_i^*) \cdot \log(1 - p_i))$$
(3-2)

Lreg 表示的是 smooth L1 函数,其公式如式 3-3 所示:

$$smooth_{L_1}(x) = \begin{cases} \frac{0.5 \cdot x^2}{\sigma^2} & \text{if} |x| < 1/\sigma^2 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$
(3-3)

区域预测网络最后会将分类结果、边界框预测结果以及原图信息输入到 proposal 层,最终得到建议框。

ROI 池化层介于区域预测网络和 Faster R-CNN 的分类回归网络中间,其主要作用是综合卷积层特征和候选框特征的信息,并把该信息池化成 7×7 大小的特征,满足后阶段分类回归网络固定输入特征尺寸大小的要求。

分类和回归网络是 Faster R-CNN 最后的一个结构,其包含分类层和边界框回归层,分类层通过全连接层和 SoftMax 激活函数预测每个由 ROI 池化层传来的建议框的类别,边界框回归层则是进一步微调建议框的边界信息,分类层和边界回归层共享一个全连接层。

在本文中目标识别任务只需要定位骨损位置而不需要进行多分类任务,因此 Faster R-CNN 网络最后一层的分类器有优化的余地,而在 Faster-ECNN 区域预测网络 RPN 已经完全可以满足边界框预测以及前景概率预测的需求。因此基于 Faster R-CNN 原理,我们提出了 Resnet-RPN 网络,其网络结构如图 3-4 所示,其主要结构由 Resnet 网络和 RPN 两部分组成,去除掉了 Faster R-CNN 中 ROI 池化层、分类和回归网络,简化了网络结构也简化了模型训练难度。



图 3-4 Resnet-RPN 网络结构

本文 Resnet-RPN 选用 ResNet-50 作为主干网络提取特征,可将输入图片压

缩为原尺寸的 32 倍的大小,即在 ResNet-50 的 Layer4 层中每个单元格对应原图中的 32×32 的小方框。在本次实验中,每个单元格会选定 3 种尺寸 3 种比例共9 种 anchor,每个方框在 RPN 中会预测其包含的前景和背景的概率以及预测边界框的中心点位置偏移量以及长宽缩放比例,因此每个单元格最终会预测 18 个概率参数和 36 个位置参数。当输入图片尺寸为 512×512,则 RPN 的 cls_prob输出(18,16,16),bbox_pred 输出(36,16,16)。最后 RPN 中的 proposal部分会将预测边界框映射到原尺寸输入数据中并对预测边界框做回归,根据非极大值抑制(non maximum suppression, NMS)以及是否超出边界去除不合格的预测框,然后通过对前景概率由大到小的排序选取出前 128 个边界框。本文中 Resnet-RPN 损失函数包括是前景背景分类的二分类损失和边界框预测的回归损失,其公式如下式 3-4 所示:

$$L(p_i, t_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)$$
(3-4)

公式中 i 表示预选框的下标, P_i表示第 i 个预选框是前景目标的概率, P_i*表 示第 i 个预选框的真实类别标签, 前景为 1 背景为 0, t_i表示预测边界框坐标的 向量, t_i*表示前景真实边界框的坐标向量, λ 表示分类损失和回归损失的平衡 权重系数, 在原文中设置为 10。L_{cls}在 RPN 中是一个经典的二分类交叉熵损失, L_{reg}表示的是 smooth L1 函数。

3.2.2. 实验设计与评估指标

本小节数据均为非完全粉碎骨损数据,总共包含 1246 张数据。为让实验具 有说服力,本实验的训练集、验证集以及测试集的数据量以近似 3:1:1 的比例划 分,训练集数据共有 747 张,验证集有 247 张,测试集 249 张。数据标签包含 踝关节边界框位置信息,如图 3-5 所示。



图 3-5 踝关节标注

所有的数据标签均由湘雅三医院骨科以及影像科室的临床医生监督本文团 队标定所得,标注工具为标注精灵,输出格式为 xml 文件。本实验对比了 Faster R-CNN 和 Resnet-RPN 两种网络在踝关节数据上的检测差异,网络代码均 基于开源深度学习框架 Pytorch 实现。网络模型的训练及预测过程如下:训练阶 段,使用自适应矩估计优化算法 Adam (Adaptive Moment Estimation),设置学 习率 lr=0.0001。输入图像尺寸统一化为 512×512。每个 batch 随机在训练集中 选取 10 个训练样本,在验证集中选取 3 个验证样本,在 Nvidia Geforce RTx 2060 SUPER 显卡上训练 100个 epoch 后停止; 预测阶段, 输入数据尺寸同样统 一化为 512×512, 然后将该测试样本输入到训练好的踝关节识别模型中, 最后 经非极大值抑制算法(NMS)后输出最终边界预测框。

本章节使用的模型评价指标为精确率(Precision)、召回率(Recall)以及两 者的综合指标(F1-score)。精确率表示所有预测为正例的样本中真实正例样本 所占比例。在本任务中由于网络只会预测前景目标的边界框,所以精确率表示 在所有的预测边界框中正确的边界框所占数量比例。召回率表示在所有真实正 例中预测正例所占比重,本任务中则为正确预测边界框数量占所有真实边界框 数量的比例。本实验中,当预测框与标签边界框的 IOU 大于 0.7 时,则判为阳, 反之为阴。

3.2.3. 实验结果与分析

经过实验,本文得到了 Faster R-CNN 与 Resnet-RPN 在测试数据集上的对比 实验结果,在实验中测试集有 249 张图片,每个图片都有一个踝关节的边界框 标签,因此测试集中真实目标边界框有249个,如表3-1所示:

 表 3-1A Resnet-RPN 和 Faster R-CNN 在测试桌上预测跳天卫边芥框头验结果						
网络	真实目标的 框数	检测所得 的框数	检测中正确 的框数	Recall	Precision	F1-score
Faster R-CNN	249	249	249	100%	100%	1
Resnet-RPN	249	249	249	100%	100%	1

D のDIを明い住して明四ノナル田にウみ4日

表 3-1B Resnet-RPN 和 Faste-RCN	的参数规模和在测设	式集上的推理计算速度
-------------------------------	-----------	------------

网络	Parameters	FPS
Faster R-CNN	53M	11
Resnet-RPN	42M	26

由于该任务难度不大,数据复杂度不高,踝关节尺寸相对比例合适,干扰 很少,所以从表 3-1A 中可以看出,两个网络均可以把所有的踝关节识别出来,

并且框出来的所有的踝关节均是正确的踝关节。表 3-1B 中记录了两种网络的参数规模以及在测试数据集上计算速度,由于 Resnet-RPN 减少了 ROIPooling 以及分类器等网络结构,其网络结构更加紧凑,因此其参参数规模要小于 Faster R-CNN,而推理速度则更快。其中部分踝关节边界框预测结果如图 3-6 所示,黄色框为真实标签,红色框为 Faster R-CNN 网络预测边界框,蓝色框为 Resnet-RPN 预测边界框。



图 3-6 Faster R-CNN 和 Resnet-RPN 预测踝关节边界框部分结果图, 黄色框为真实标签, 红色框为 Faster R-CNN 网络预测边界框框, 蓝色框为 Resnet-RPN 预测边界框

从图 3-6 中红框和蓝框和黄色框的重合度可以看出, Resnet-RPN 和 Faster R-CNN 两个网络可以很好地识别出数据中的踝关节。



Resnet-RPN 和 Faster R-CNN 两个网络在测试集上预测边界框与真实边界框 IOU 结果如图 3-7 所示:

图 3-7 Resnet-RPN 和 Faster-RCNN 在测试数据集上预测结果的 IOU 值统计图, Resnet-RPN 的 IOU 平均值为 0.8492, Faster R-CNN 的 IOU 平均值为 0.8564

可以看到 Resnet-RPN 和 Faster R-CNN 的 IOU 结果均集中在 0.7~0.99,当预 测框与标签边界框的 IOU 大于 0.7 时,则判为阳,反之为阴,因此两个目标识 别网络都可以很好的识别出踝关节。Resnet-RPN 的 IOU 平均值为 0.8492, Faster R-CNN 的 IOU 平均值为 0.8564。这说明两种网络模型都可以有效地学习 到踝关节的特征, Faster R-CNN 的性能略好于 Resnet-RPN。

通过统计分析 Resnet-RPN 网络在预测踝关节边界框任务中的预测结果,可 以证明 Resnet-RPN 在本任务上具有有效性。在和 Faster-RCNN 进行对比实验时, 注意到 Resnet-RPN 在测试数据上平均 IOU 值仅略低于 Faster-RCNN。但 Resnet-RPN 的网络参数规模小于 Faster R-CNN,并且推理速度比 Faster R-CNN 快。

基于踝关节检测实验,本文原始数据在进行骨损目标识别以及语义分割提 取踝关节缝隙之前进行了裁剪,只留下了具有高价值目标的踝关节部分,如图 3-8 所示:



图 3-8 截取踝关节构成数据集的部分样例

3.3. 基于注意力机制的小尺寸骨损检测

3.3.1. 注意力机制

注意力机制(Attention Mechanism)最早运用于图像领域,在2014年, Google Mind 团队发表了[68],论文中该团队在 RNN 模型上使用注意力机制来 进行图像分类,取得了很好的效果。此后融合了注意力机制的深度学习网络被 广泛地应用在自然语言处理(Nature Language Processing, NLP)、图像分类、 语言识别等各种不同类型的任务中,取得了不错的性能提升。深度学习领域中 的注意力机制有其生理学领域的依据,其借鉴了生物观察思维方式(生物视觉
注意力机制是生物视觉所特有的大脑信号处理机制)。通过大量对人类的视觉研 究可知,人类为了能从大量的视觉信息中快速地筛选出高价值的信息,人类的 视觉系统进化出了注意力焦点的观察模式,即人类视觉每次处理视觉信息时只 会把大量的视觉注意力资源投入到重点关注的目标区域,抑制其他无用信息的 处理,从而获取更多的目标对象的细节信息。人类视觉注意力机制极大地提升 了信息处理速度和准确度,缓解大脑的处理压力。

深度神经网络中的注意力机制其本质上是通过深度学习网络学习到对目标 特征敏感的权重,使得有用的目标特征被增益,无用的特征被抑制,从而使得 添加了注意力机制的深度卷积神经网络的参数减少,处理速度增加,性能变好。 深度神经网络注意力机制根据注意力的计算区域可以分为软注意力(Soft Attention)、硬注意力(Hard Attention)以及局部注意力(Local Attention)三大 类。软注意力是一种全局计算的注意力学习方法,该方法会对所有的特征求权 重概率,其概率范围为 0~1,权重概率越大表示该特征越重要。软注意力完全 可微分,因此在添加软注意力机制的深度卷积网络中梯度能够正常地进行反向 传播。该方法虽然能够扫视全局使得性能更好,但是也增加了计算量,导致运 行速度不快。硬注意力对感兴趣的特征权重赋值 1,其余不感兴趣的特征权重 赋值 0。因此这种方法需要一步到位地对齐,并且由于该方法不可导,一般需 要用强化学习的方法进行训练。局部注意力方法是软注意力机制和硬注意力机 制的一个折中方案,该方法先用硬注意力机制定位到感兴趣区域的大致方位, 然后在该感兴趣区域用软注意力机制计算感兴趣区域全局的特征权重概率。把 注意力机制融入深度卷积神经网络中,可以增益兴趣区域抑制干扰区域,让深 度卷积神经网络更聚焦于目标特征,在减少网络参数增加运行速度的同时提升 了网络性能。比如 Woo, Sanghyun 等人在论文[69]中基于 Resnet 网络做过与软 注意力机制相关的对比实验, 其基于 ImageNet 数据集在 ResNet 网络中先后实 验对比了软注意力机制的空间注意力模型(Spatial Attention Module)和通道注 意力模型(Channel Attention Module)给网络带来的性能增益,其结果如下表 3-2 所示。

表 3-2 空间注意力模型和通道注意力模型基于	ImageNet-1K 数据集在 ResNet 网络上的性能
-------------------------	--------------------------------

模型	Parameters	GFLOPs	TOP-1 Error	TOP-5 Error
ResNet50 (baseline)	25.56M	3.86	24.56 %	7.50 %
ResNet50 + AvgPool & MaxPool	25.92M	4.02	22.80%	6.52 %
ResNet50 + channel + spatial	28.09M	3.864	22.66 %	6.31 %

本文将采用软注意力机制中的空间注意力模型(Spatial Attention Module) 以及通道注意力模型(Channel Attention Module),从结构上改善目标识别算法 在小目标任务中的识别性能。

深度神经网络中的注意力机制一般模型包括三个部分,分别是卷积单元 Convolutional Bottleneck、有隐藏层的多层感知机 MLP(Multi Layer Perceptron) 以及一个 Sigmoid 函数,该三部分共同构成了基础注意力模型(Basic Attention Module),该注意力模型可以嵌入到任何卷积神经网络中,如图 3-9 所示。



图 3-9 注意力机制结构图

输入 F 为 C×H×W 的特征图,通过注意力模型后获得 1×H×W 的目标注意力 特征图 F', 然后将目标注意力图 F'和输入特征 F 逐元素相乘, 最后得到目标特 征突出而无用信息被抑制的输出特征图 F"。注意力机制一般需要利用空间和通 道两个维度的特征信息,基于空间和通道特征的注意力模型分别称为空间注意 力模型(Spatial Attention Module)和通道注意力模型(Channel Attention Module),通道注意力模型利用深度卷积神经网络中通道所携带的特征来训练 注意力映射,其通过把输入特征空间的维度进行压缩,可以有效地计算通道注 意力映射,比如将 C×H×W 的输入特征压缩为 C×1×1,然后传递给多层感 知机 MLP, 最后由 Sigmoid 激活函数输出新的特征图。空间注意力模型 (Spatial Attention Module) 会生成一个空间注意力映射,该映射对特征图中目 标之间的位置关系敏感,其会将输入的特征图的通道特征压缩,最终得到一个 1×H×W 的新特征图。空间注意力模型和通道注意力模型高效地利用了深度卷 积神经网络的空间特征和通道特征,使得注意力机制可以从多维度抑制无用的 背景信息和提升有用信息的特征权重。一般注意力机制会同时包含这两种模型, 比如由 Park, Jongchan 等人提出的[70]BAM(Bottleneck Attention Modules) 就将 空间注意力模型和通道注意力模型并行地接在输入特征图后位,然后将两种模 型输出的新特征进行加运算得到一个融合两个空间特征和通道特征的新特征。 而[69]CBAM(Convolutional Block Attention Module)则是将两种注意力模型串 行地置于输入特征图的后方,网络的特征先输入到通道注意力模型得到空间压 缩的特征,然后交由空间注意力模型得到新的特征。

3.3.1.1 BAM 结构及原理

BAM (Bottleneck Attention Modules) [70]的总体结构由通道注意力模块和 空间注意力模块并行组成,其结构图如图 3-10 所示:



图 3-10 BAM 结构图

由于 BAM 的模块化,其可以很方便地嵌入到任何深度神经网络中,如图所 示 BAM 在保留原有网络的结构下并行地嵌入到一个卷积神经网络中。令输入 特征为 F∈R^{C×H×W},输出为 F',通道注意力特征映射为 M_c,空间注意力特征映 射为 M_s, sigmoid 激活函数为 σ,则 BAM 整体结构的数学表达式如式 3-5 所示:

 $F' = F + F \otimes \sigma(M_c(F) + M_s(F))$ (3-5) 其中 \otimes 为矩阵点积运算,输出 F'的尺寸为 C×H×W。BAM 在结构的最后部分 采用了残差结构促进梯度的传播,该残差运算表现为输入特征 F 加上 F 和注意 力映射权重的点乘结果。

BAM 的通道注意力模块会将输入特征图 F 压缩成 C×1×1 的通道特征向量, 其数学公式如式 3-6 所示:

$$M_{c}(F) = BN(MLP(AvgPool(F)))$$
(3-6)

通道注意力模块首先会对输入的特征图 F 做一个全局平均池化(Global Average Pooling)得到 F_c,然后将 F_c输入到一个有隐藏层的小型多层感知机中(MLP),最后对 MLP 输出的特征进行批量归一化。BAM 中的空间注意力模块会将输入特征图 F 的通道压缩,最终得到的特征向量尺寸为 1×H×W,其数学公式如式 3-7:

$$M_{s}(F) = BN(f_{3}^{1\times1}(f_{2}^{3\times3}(f_{1}^{3\times3}(f_{0}^{1\times1}(F)))))$$
(3-7)

空间注意力模块由一串卷积层组成,第一层的卷积层的卷积核为 1×1,该层可 以压缩特征图的通道维度,第二层和第三层卷积层的卷积核为 3×3,为了增加 网络的有效感受视野,该层的膨胀系数(dilation value)d=4,最后一层卷积层 的卷积核为 1×1,该层使得空间注意力模块地输出特征向量尺寸为 1×H×W, 最后和通道注意力模块一样添加了一个批量归一化层。为了使梯度反向传播时 更加顺畅,空间注意力模块和通道注意力模块的结果最后会进行加运算而不是 乘法运算,然后交由 Sigmoid 激活函数处理。

3.3.1.2 CBAM 结构及原理

CBAM (Convolutional Block Attention Module) [66]的整体结构和 BAM 并行结构构不同, CBAM 的通道注意力模块和空间注意力模块是串型结构,从深度卷积神经网络中传出的特征图先经通道注意力模块处理,然后转交由空间注意力模块处理,整体结构如下图 3-11 所示:



图 3-1 1CBAM 整体框架图

注意力计算数学表达总结如式 3-8 所示:

$$\mathbf{F}' = \mathbf{M}_{\mathbf{c}}(\mathbf{F}) \otimes \mathbf{F}; = \mathbf{M}_{\mathbf{s}}(\mathbf{F}') \otimes \mathbf{F}'$$
(3-8)

设定从卷积神经网络传入 CBAM 的特征图为 F,其向量尺寸为 C×H×W。 Mc 表示通道注意力模块的映射权重,Mc (F)为特征 F 经由通道注意力模块后 的运算结果,其特征向量尺寸为 C×1×1,F'为 Mc (F)和输入特征 F 的点乘 结果。Ms 为二维空间注意力模块的映射权重,其输入 F'经由 Ms 权重运算后得 到输出结果为 Ms (F'),该结果的向量尺寸为 1×H×W,Ms (F')和 F'进行点 乘运算得到最终的结果 F"。

CBAM 除了结构与 BAM 的结构不一样,其空间注意力模块以及通道注意 力模块也相应地进行了一些优化。CBAM 的通道注意力模块虽然遵循 BAM 通 道注意力模块一样的运算过程,但是 CBAM 的通道注意力模块添加了平均池化 层(Average Pooling)以及最大池化层(Max Pooling)用来获得更多的通道特 征,平均池化层可以聚合特征的空间信息,最大池化层则可以捕获更精细的通 道特征,两个池化层的同时使用可以使网络有更强的表达能力。结构如图 3-12 所示:



图 3-12 CBAM 通道注意力模块图

通道注意力模块表达式 M。的数学表达式如式 3-9 所示:

$$M_{c}(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F)))$$
(3-9)

由公式可知,输入特征 F 首先同时利用平均池化层和最大池化层来聚合其空间 信息,分别生成平均池化特征 AvgPool (F)和最大池化特征 MaxPool (F),然 后两个特征分别输入到一个由有隐藏层的多层感知机 (MLP)组成的共享网络 中,然后将两个特征逐元素相加,最后输入到 sigmoid 函数中。

CBAM 的空间注意力模块串接在通道注意力模块后,与 BAM 的空间注意 力模块相比其减少了卷积层的数量,同 CBAM 的通道注意力模块一样新增了平 均池化层和最大池化层,使得该模块的表达能力增强。其整体结构如下图 3-13 所示:



图 3-13 CBAM 中空间注意力模块的结构示意图

输入特征 F 先后通过最大池化层和平均池化层, 然后将结果输入一个卷积层, 该卷积层的卷积核大小为 7×7, 最后交由激活函数处理, 其数学表达式如式 3-10 所示:

$$M_{s}(F) = \sigma(f^{7\times7}([AvgPool(F); MaxPool(F)]))$$
(3-10)

空间注意力模块和通道注意力模块为卷积神经网络提供了多维度的注意力 特征,这两个模块分别聚焦于特征是什么以及特征在哪里。两个模块可以组成 串行结构也可以组成并行结构,通过 CBAM 论文中的实验可知,串行结构比并 行结构带来更好的性能优化。 3.3.2. 融合注意力机制的目标识别网络

3.3.2.1. 融合注意力机制的 YOLO V3 网络结构

Yolo 算法和 R-CNN 系列的目标识别算法的先检测出预选边界框后分类的思路不同,其将目标识别的定位任务和分类任务看作是一个回归任务,希望实现只通过一次检测就能得到分类预测信息和边界框预测信息。Yolo 系列最早由Redmon, Joseph 等人[49]在 2016 年的 CVPR 中提出,后又陆续发展出 Yolo v2[50]以及 Yolo v3[51]等等。Yolo 算法的总体思路是通过深度卷积神经主干网络把输入图像压缩至尺寸大小为 N×N 的特征单元格,每一个特征单元格都代表图片该区域经过一系列的特征提取后得到的高层次高抽象的特征向量,在Yolo 系列中该特征向量包含了前景边界框的信息、前景的类别信息以及置信度信息。在训练过程中,如果该特征单元格中包含有前景目标,那么该单元格提取的特征对前景目标的边界框信息、类别信息、置信度信息敏感;如果单元格中只有背景目标的信息,那么该单元格的特征只对置信度敏感。如图 3-14 所示,在踝关节 x 光片数据的第五行第四列的红色单元格中包含有前景骨损,而第七行第五列的绿色单元格包含的是背景脚趾关节,因此红的单元格最终会输出前景边界框信息等等,而绿色的单元格最终只会输出置信度的概率。



图 3-14 Yolo 算法检测骨骼原理示意图

在 Yolo V3 中,为了能获取更复杂更深层次的特征,网络的主干网络换成了 具有残差结构的 Darknet-53,Yolo V3 的网络结构如图 3-15 所示。为了加强对 小目标的检测精度,Yolo V3 中采用了多尺度融合的预测方式,在原文中 Yolo V3 有三种不同感受野的预测分支,分别输出把输入图像压缩 32 倍、16 倍和 8 倍后的特征图,特征图越小其拥有的感受视野就越大。因此小特征图检测原图 像中大尺寸的前景目标,大特征图检测精细的小尺寸前景目标。 原论文中 Yolo V3 每个预测任务都会得到一个 N×N×[3*(4+1+C)]的结果矩阵, N×N 代表特征单元格的数量, 3 表示每个单元格会有三种不同尺寸的预测输出, 4 表示每个尺寸的输出都会包含 4 个边界框的预测结果, 其中包括边界框中心相对单元格中心的偏移量以及边界框大小相对单元格大小的缩放比例, 1 表示置信度, 代表单元格中目标是前景的概率, C 代表前景类别的数量, 比如对于 COCO 的 80 种类别而言这里的 C 就等于 80。



图 3-15 YOLO V3 网络结构

Yolo 算法的损失函数包括三部分,分别是置信度损失函数 L_{ij}^{conf}、分类损失 函数 L_{ij}^{cls} 以及定位损失函数 L_{ij}^{reg},公式如式 3-11 所示:

$$Loss = \sum_{i=0}^{N^{2}} \sum_{j=0}^{B} (L_{ij}^{reg} + L_{ij}^{conf} + L_{ij}^{cls}),$$

$$L_{ij}^{reg} = \lambda_{coord} I_{ij}^{obj} [(x_{ij} - x_{ij}^{*})^{2} + (y_{ij} - y_{ij}^{*})^{2} + (\sqrt{w_{ij}} - \sqrt{w_{ij}^{*}})^{2} + (\sqrt{h_{ij}} - \sqrt{h_{ij}^{*}})^{2}], \quad (3-11)$$

$$L_{ij}^{conf} = -I_{ij}^{obj} \log (cof_{ij}) - \lambda_{noobj} I_{ij}^{noobj} \log (1 - cof_{ij}),$$

$$L_{ij}^{cls} = -I_{ij}^{obj} \sum_{c \in cls} (p_{ij}^{c^{*}} \log (p_{ij}^{c}) + (1 - p_{ij}^{c^{*}}) \log (1 - p_{ij}^{c}))$$

N 为特征图的大小, B 为每个特征单元格中包含的预测边界框数量,在 Yolo V3 中 B 取值为 3。I_{ij}^{obj}表示第 i 个单元格中的第 j 个预测边界框是否是前景,如果 是前景则 I_{ij}^{obj}取值为 1 否则为 0, I_{ij}^{noobj}表示第 i 个单元格中的第 j 个 anchor 是 否是背景,取值和 I_{ij}^{obj}相反。λ_{coord}和 λ_{noobj}分别为定位损失和置信度损失的平衡 系数。(x_{ij}, y_{ij})为网络预测输出的前景边界框的中心的坐标,(x_{ij}*, y_{ij}*)为真实框 的中心坐标,(w_{ij}, h_{ij})为网络预测的前景边界框的宽和高的信息(w_{ij}*, h_{ij}*)为真 实框的宽和高。cof_{ij}、p_{ij}°分别表示网络预测的第 i 个单元格中第 j 个预测框的置 信度以及前景目标是第 c 类的概率,p_{ij}°*表示前景目标实际的类别标签。 YOLO V3 的特征提取主干网络为具有残差结构的 Darknet-53。Darknet-53 由 28 个 Convolutional 组成, CBAM 和 BAM 可以很方便地添加到 Darknet-53 的 每个 Convolutional 中, 如图 3-16 所示。



图 3-16 CBAM 和 BAM 注意力机制添加到 Darknet-53 网络示意图

因此注意力模块可以很方便地融合到 YOLO V3 的结构中,其总体网络结构 示意图如图 3-17 所示:



图 3-17 YOLO V3 中添加注意力机制示意图

3.3.2.2 融合注意力机制的 Resnet-RPN 网络

本文中 Resnet-RPN 由 Resnet-50 和 RPN 组成, Resnet-50 由 16 个 Bottleneck 组成,本章节将 CBAM 和 BAM 添加都每个 Bottleneck 中,从而搭建出融合注 意力机制的 Resnet-RPN 网络,如图 3-18 所示,



图 3-18 Resnet-RPN 中添加注意力机制示意图

3.3.3. 实验方法与设计

为了评估不同预处理方式对骨损目标识别任务的影响,本章节对原始尺寸 数据进行了两种不同的输入预处理,一种是将原图像缩小为 512×512,另一种 是用踝关节目标识别模型识别踝关节,然后通过截取目标缩小输入图片尺寸。 两种预处理方式得到的数据集均用来训练 YOLO V3 网络和 Resnet-RPN 网络, 将得到面向不同预处理方式数据的骨损目标识别模型。此外本章将在以上实验 基础上,通过在目标识别算法的特征提取网络中添加注意力机制进行对比试验, 来评估注意力机制在本任务中的有效性。

本章数据均为非完全粉碎骨损数据,其中包括 511 张健康骨骼数据和 735 张含小目标骨损的数据,总共 1246 张数据。为让实验具有说服力,本实验的训 练集、验证集以及测试集的数据量以近似 3:1:1 的比例划分,在 511 张健康骨骼 数据集中分别随机取 307 张为训练集、103 张为验证集、101 张为测试集,在 735 张含小目标骨损数据集中分别随机取 440 张为训练集、147 张为验证集、 148 张为测试集。训练集数据共有 747 张,验证集有 247 张,测试集 249 张。数 据标签标注出了数据中的小目标骨损边界框位置信息,如果是健康骨骼则不会 有边界框位置信息。所有的数据标签均由湘雅三医院骨科以及影像科室的临床 医生监督本文团队标定所得,标注工具为标注精灵,标签输出为 xml 文件。

两种目标识别网络代码均基于开源深度学习框架 Pytorch 实现,网络模型的 训练及预测过程如下:小目标骨损识别阶段的训练阶段,使用自适应矩估计优 化算法 Adam (Adaptive Moment Estimation),设置学习率 lr=0.0001。输入图像 均尺寸统一化为 512×512。每个 batch 随机在训练集中选取 10 个训练样本,随 机在验证集中选取 3 个验证样本,在 Nvidia Geforce RTx 2060 SUPER 显卡上训 练 100 个 epoch 后停止。测试阶段,输入图像尺寸同样统一化为 512×512,将 测试样本输入到小目标骨损识别模型中进行预测,经非极大值抑制算法 (NMS) 后输出最终边界预测框。当预测框与标签边界框的 IOU 大于 0.7 时,则判为阳,反之为阴。

3.3.4. 实验结果与分析

3.3.4.1 不同预处理方式下骨损识别的结果与分析

本章完成了两种不同预处理方式下骨损识别对比实验,一种是把原始图片 尺寸统一缩小为 512×512,简称缩放方式;另一种是通过踝关节识别模型从原 始尺寸数据中识别并截取出踝关节,简称截取方式。这两种方式的数据包含有 相同数量的骨损目标。测试数据包含 101 例健康骨骼数据和 148 例有小尺寸骨 损的数据总共 249 例数据,148 例骨损数据包含 370 处独立骨损,平均每例数据 有 2.5 个独立的骨损,其尺寸大小不一。Resnet-RPN 和 YOLO V3 网络在两种输 入数据上的骨损目标识别结果如表 3-3 所示:

		•					
预处理 方式	网络	真实框 个数	预测框 个数	预测正确 的框数	Recall	Precision	F1- score
佐 苗	YOLO V3	370	274	196	0.5297	0.7153	0.6087
细瓜	Resnet-RPN	370	246	174	0.4703	0.7073	0.5649
北雨	YOLO V3	370	412	340	0.9189	0.8252	0.8696
截収	Resnet-RPN	370	418	335	0.9054	0.8014	0.8503

表 3-3 不同输入预处理的边界框预测结果统计

表 3-3 中可知,在 Resnet-RPN 和 YOLO V3 两个网络中,其截取方式的结果均 优于缩放方式的结果,其 Recall 有 0.3~0.4 的差距, Precision 有约 0.1 的差距。 表 3-3 中也可以看到 YOLO V3 的结果要优于 Resnet-RPN 的结果,其 Recall 相 差 0.013, Precision 相差 0.023, F1 值相差 0.02。

对于每例数据而言, Resnet-RPN 和 YOLO V3 网络预测结果只要有骨损边 界框输出, 就可判定该例数据为有骨损。表 3-4 统计了两个网络在不同预处理 方式下, 判别有无骨损的混淆矩阵, 表 3-5 中计算出了其对应的性能指标数值。

衣 3-4 两种目标识别网络识别宜船走否有预伤结末讹闲矩阵								
	预处理方式	缩	放	截	取			
网络	混淆矩阵	预测无骨损	预测有骨损	预测无骨损	预测有骨损			
$VOI \cap V^2$	真实无骨损	86	15	97	4			
TOLO VS	真实有骨损	43	105	13	135			
Resnet-RPN	真实无骨损	84	17	95	6			
	真实有骨损	46	102	14	134			

责 3-4 两种目标识别网络识别骨骼是否有损伤结果混淆矩阵

表	3-:	;两	种输	入预	处理:	方式在	两种	网络	上的	实验	数据
---	-----	----	----	----	-----	-----	----	----	----	----	----

	YOL	O V3	Resne	Resnet-RPN		
	缩放	截取	缩放	截取		
Acc	0.7671	0.9317	0.7470	0.9197		
Precision	0.8750	0.9712	0.8571	0.9571		
Recall	0.7095	0.9122	0.6892	0.9054		
F1	0.7836	0.9408	0.7640	0.9306		

从混淆矩阵表 3-4 中可知,同一网络下截取的方式比缩放方式能在骨损数据 中检测出更多真实骨损病例。比如在 148 例有损数据中,YOLO V3 截取方式可 以检测出 135 例真实有损数据,而缩放的方式只能检测出 105 例有损数据。两 种网络在截取的方式下其准确率比缩放的方式的准确率高 0.1~0.2,如表 3-5 中 所示。由以上实验结果可知,在 Resnet-RPN 和 YOLO V3 两种网络上,截取的 输入预处理方式在各个评估参数上都优于缩放的预处理方式。这是由于缩放的 预处理方式压缩了小目标骨损尺寸,使得相对小目标任务转变为绝对小目标任 务,小目标的特征细节严重损失,故而导致许多相对小目标骨损无法识别。如 图 3-19 所示,用截取的预处理方式可以识别出踝关节中细小的骨损,如图右边 绿色框所示,而缩放预处理方式因为小目标尺寸被压缩,特征消失,因此无法 识别出细小目标骨损。



图 3-19 缩放预处理和截取预处理方式的差异

从表 3-4 和表 3-5 中结果数据可以看出,YOLO V3 的准确率为 0.9317,F1 值为 0.9408,Resnet-RPN 的准确率为 0.9197,F1 值为 0.9306,YOLO V3 的性 能略优于 Resnet-RPN。在 148 个真实有骨损的数据中,在截取方式下 YOLO V3 检测出了 135 个有骨损的数据,其召回率为 0.9122,而 Resnet-RPN 检测出 了 134 个有骨损的数据,召回率为 0.9054。在所有检测出有骨损的数据中 YOLO V3 有 4 个数据误检,精确率为 0.9712,Resnet-RPN 有 6 个数据误检,精 确率为 0.9571。YOLO V3 的效果好是因为有多尺寸融合机制,可以在 8×8、 16×16、32×32 三种不同尺寸的特征图中检测对象,所以对小目标的检测更准, 而 Resnet-RPN 只能在最后一层 16×16 的特征图中检测对象,所以对小目标任 务相对没那么敏感。具有代表性的案例如图 3-20 所示,左边 Resnet-RPN 没有 检测出小尺寸骨损,而右边的 YOLO V3 检测出了骨损。



图 3-20 Resnet-RPN 和 YOLO V3 的检测结果对比样例, 左边 Resnet-RPN 没有检测出小尺 寸骨损, 而右边的 YOLO V3 检测出了骨损

3.3.4.2 注意力机制有效性实验

本小节通过在 YOLO V3 和 Resnet-RPN 上添加 CBAM 和 BAM 模块,对比 实验验证了注意力机制模块的有效性,并比较了 CBAM 和 BAM 给网络带来的 性能差异。实验中有 249 个数据作为测试数据,共计 370 个边界框,输入图片 都通过截取的方式进行预处理,只留下踝关节,其边界框预测结果统计数据如 表 3-6 所示。在两个网络中添加注意力机制后,其预测的总框数有降低的趋势, 而预测正确的框数则增加了,如在 YOLO V3 中预测框数从 412 个降到 410 个, 预测正确框数从 340 个上升到 345 个。两个网络的召回率和准确率有约 0.01 以 上的提升,其中 YOLO V3 添加 CBAM 模块的网络的提升最多,召回率从 0.9054 提升至 0.9324,精确率从 0.8014 提升至 0.8415。

	网络	真实框数	预测框数	正确预测 框数	Recall	Precision	F1-score
有	YOLO V3	370	410	345	0.9324	0.8415	0.8846
CBAM	Resnet-RPN	370	415	341	0.9216	0.8217	0.8688
有	YOLO V3	370	412	343	0.9270	0.8325	0.8772
BAM	Resnet-RPN	370	416	341	0.9216	0.8197	0.8677
无注意	YOLO V3	370	412	340	0.9189	0.8252	0.8696
力机制	Resnet-RPN	370	418	335	0.9054	0.8014	0.8503

表 3-6 YOLO V3 和 Resnet-RPN 添加注意力机制后边界框数预测结果

表 3-7 统计了两个网络在不同注意力机制模块下,判别有无骨损的混淆矩阵,表 3-8 中计算出了其对应的性能指标数值。

表 3-7 YOLO V3 和 Resnet-RPN 添加注意力机制后检测出有骨损的结果统计

		有 C	BAM	有B	BAM	无注意	力模块
网络	混淆矩阵	预测无 损	预测有 损	预测无 损	预测有 损	预测无 损	预测有 损
VOLOV2	真实无损	98	3	98	3	97	4
YOLU V3	真实有损	10	138	11	137	13	135
Desnet DDN	真实无损	97	4	96	5	95	6
Keshet-KF IN	真实有损	12	136	12	136	14	134

表 3-8 YOLO V3 和 Resnet-RPN 添加注意力机制后检测出有骨损的评估指标

	YOLO V3			Resnet-RPN		
性能指标	有 CBAM	有 BAM	无注意力模块	有 CBAM	有 BAM	无注意力模块
Acc	0.9478	0.9438	0.9317	0.9357	0.9317	0.9197
Precision	0.9787	0.9786	0.9712	0.9714	0.9645	0.9571
Recall	0.9324	0.9257	0.9122	0.9189	0.9189	0.9054
F1	0.9550	0.9514	0.9408	0.9444	0.9412	0.9306

从表 3-8 中可知,添加了 CBAM 的 YOLO V3 网络,在 148 例有损数据中正确 检测出了 138 例,其总共预测出了 141 例有损数据,其召回率为 0.9324,精确 率为 0.9787。而无注意力机制的 YOLO V3 的召回率只有 0.9122,精确率为 0.9712。可以看出添加 CBAM 后大幅度地降低了网络的漏检情况,小幅度地降 低了误检的情况。由表中可以看出在 Resnet-RPN 网络中,添加注意力机制具有 同样骨损检测性能的提升。从表中还可知 BAM 虽然具有同样提升性能的能力, 但是 CBAM 的提升更高。

100								
网络	YOLO V3+CBAM	Resnet-RPN+CBAM	YOLO V3	Resnet-RPN				
Parameters	65.3M	44.8M	61.9M	42M				
FPS	77	30	75	26				

表 3-9 YOLO V3 和 Resnet-RPN 有无注意力机制的参数规模和推理计算效率

从表 3-9 中可知, YOLO V3 的参数规模要比 Resnet-RPN 大, 但 YOLO V3 的推理速度更快,几乎比 Resnet-RPN 快三倍。在添加注意力机制后,YOLO V3 的参数规模稍微增加,但其推理速度也同时得到了提升。

由上述两个目标识别网络的所有实验结果分析可以证明,注意力机制添加 到目标识别网络中后,的确可以提升网络对骨损目标的检测能力。通过对比 CBAM 和 BAM 的结果分析可知,CBAM 在检测骨损目标任务上比 BAM 能给 目标识别网络带来更好的性能提升。

3.4. 本章小结

本章首先介绍了骨损目标识别数据集的组成以及数据特点,其中点出小尺 骨损检测是一个研究难点。然后提出本文在优化小尺寸骨损目标识别方面有两 种优化方案,第一种是优化其数据的预处理方式,让输入数据在保留最多的骨 损目标特征的前提下尺寸尽可能适合网络训练。第二种就是优化目标识别网络 的特征网络,即在网络中添加注意力机制模块。围绕注意力机制,本章介绍了 CBAM 和 BAM 两种注意力机制模块的原理和特点,介绍了如何将这两种模块 融合到 YOLO V3 以及 Resnet-RPN 的网络中。然后介绍了 Faster R-CNN 网络的 结构,以及针对本文数据二分类的特点,基于 Faster R-CNN 提出了 Resnet-RPN 网络,同时也介绍优化了 YOLO V3 网络。最后就是进行了相关的对比试验, 包括了三个实验,第一个是对比 Resnet-RPN 和 Faster R-CNN 在识别踝关节任 务上的性能差异,验证了 Resnet-RPN 的有效性;第二个对比了两种输入预处理 方式给骨损检测带来的性能差异,验证了截取的预处理方式可以很好地提升目 标识别网络对小尺寸骨损的识别,从而提升整个网络对骨损的识别能力;第三 个实验通过在两个目标识别网络上添加 CBAM 和 BAM 模块,验证了注意力机 制模块在提升骨损检测能力方面具有可行性。

第4章 基于语义分割网络的踝关节缝隙提取

踝关节缝隙物理几何信息对关节肌肉韧带损伤辅助诊断具有非常高的临床价值。临床上,肌肉韧带在 x 光片上呈现透明状态,缺少丰富临床经验的医生 很难准确的仅通过 x 光片诊断出肌肉韧带是否有损伤,经验丰富的专家和医生 可以通过观察踝关节缝隙的物理几何信息进行诊断。本文针对肌肉韧带损伤在 临床 x 光片上难检测的问题,提出利用 U-net 和 U-net++语义分割网络识别提取 踝关节缝隙的物理几何信息,辅助医生进行踝关节肌肉韧带损伤的诊断。

4.1. 踝关节缝隙提取背景与意义

以往的肌肉韧带智能检测辅助方法基于分类网络直接判断骨骼是否有肌肉 韧带损伤,比如 Pranav Rajpurkar1 等人[35]基于 MURA 肌肉骨骼数据库训练了 一个鉴别骨骼肌异常的分类模型。该类方案数据和实现相对简单,无法给医生 提供一个准确的几何物理信息作为参考,并且精度不高,实际使用效果一般。 Olczak 等人[19]提到,骨骼损伤的分类辅助系统对临床医生辅助诊断的帮助十 分有限,而骨骼几何信息的提取对外科医生具有十分重要的临床辅助诊断意义, 比如 x 光片中骨骼的相对位置、缝隙的物理几何信息、骨骼的几何形状等等都 是临床医生进行临床诊断工作时重要的参考依据。

本文和湘雅三医院外科骨科医生以及影像科医生讨论得知,踝关节间隙的 几何特征对诊断韧带以及肌肉损失有重要参考价值,因为在 x 光片数据中韧带 和肌肉均为透明,无法直接观察是否有损伤,临床上只能通过观察关节间隙等 间接的信息来诊断。无肌肉韧带损伤的正常踝关节,其关节缝隙上下边界平行, 整体结构规整无异样;有肌肉韧带损伤的踝关节,关节缝隙物理几何信息将发 生形变。在 x 光片中因为存在遮挡叠影,这往往导致临床医生诊断韧带肌肉损 伤难度较大,经验缺少的医生容易忽略这些信息。

为了提供更有效的几何物理信息,本文提出用语义分割算法提取踝关节间 隙几何特征。如图 4-1 所示为踝关节语义分割标签示意图,其中蓝色部分为踝 关节间隙,绿色为腓骨,红色为胫骨,黄色为踝骨。本章节目标便是从腓骨、 胫骨、踝骨中语义分割出蓝色的踝关节间隙。



图 4-1 胫骨为红色, 腓骨为绿色, 踝骨为黄色, 关节间隙为蓝色

4.2. 踝关节缝隙语义分割数据集的构建

经过多年的临床积累, 湘雅三医院骨科以及放射科具有丰富的临床指导经 验以及完备的病人历史数据库。通过合作现收集了 1246 张符合要求的高质量踝 关节 x 光片, 病人年龄均大于 25 周岁, 男女性别比例各占一半, 数据方位包括 正视方位和侧视方位两种, 不同视角的数据量所占比例均等。每张图像的尺寸 大小不一, 分辨率高, 基本大于 2000×2500。本文为每张图片都标注了高精度 的语义分割的标签, 如图 4-2 所示, 每张图片包含 4 种类别, 分别为胫骨、腓 骨、踝骨、关节间隙, 在本文中只用到了关节缝隙的语义标签。



图 4-2 图的第一行为原始数据, 第二行为标签图, 其中腓骨为绿色, 胫骨为红色, 踝骨为 黄色, 关节间隙为蓝色。a和d为侧视方位, b和c为正视方位。

4.3. 主流医疗语义分割算法

在计算机视觉领域,图像的语义分割具有非常重要的作用,在许多运用领 域,例如无人驾驶、智慧医疗、机器人技术等等,语义分割算法至关重要。近 年随着深度卷积神经网络的发展,语义分割算法更是大放异彩,特别是在医学 图像分析方面。在医学图像分析领域,基于语义分割的医学图像分割可用于影 像引导、定向放疗等过程,也可以用于提取显微成像图中的人体组织等等。

本项目将基于 U-net[42]和 U-net++[71]两种医疗领域常见的语义分割网络提 取踝关节间隙。U-net 网络本身是针对医疗图像分割而提出的,其由 FCN 优化 而来。U-net 网络是一个深度监督的编码-解码网络,编码器用于特征提取,解 码器通过上采样对图像中每个像素进行分类。U-net++在 U-net 的基础上优化了 网络结构,其编码和解码的子网络通过一个系列稠密的嵌套跳跃路径来相互连 接。这种重新设计的跳跃连接主要是用来降低编码解码子网络中特征图的语义 缺失,当编码解码网络的特征图语义类似时,这种优化器可以解决一个更简单 的学习任务。在编码解码网络中跳跃连接是一个恢复分割细节的关键点,跳跃 连接将来自解码子网络中深度语义粗略的特征图和来自编码子网络中浅层低级 详细的特征图进行融合,可以更高效地捕获目标对象的深层细节,在最终的复 杂背景中产生更加细节的分割图。Zongwei Zhou[71]论文中数据显示 U-net++在 肺结节、结肠息肉、肝脏和细胞核的数量参数和分割准确性上比 U-net 和 wide U-net 有更好的结果。U-net 和 U-net++网络结构如图 4-3 所示:



图 4-3 A 为 U-net 的整体结构, B 为 U-net++网络结构

U-net 解码层的输入特征图直接来自编码层,而 U-net++在编码层和解码层之间 有一个稠密卷积块,这个块的卷积层数量取决于"金字塔"的级别。 在图 4-3 中 X_{i,j}代表一个卷积操作节点 X_{i,j}的输出,i指的是沿着编码方向的 第 i 个下采样层,j指的是沿着跳跃路径方向的第 j 个卷积层,X_{i,j}计算如式 4-1:

$$X_{i,j} = \begin{cases} H(X_{i-1,j}) & j = 0\\ H([[X_{i,k}]_{k=0}^{j-1}, U(X_{i+1,j-1})]) & j > 0 \end{cases}$$
(4-1)

函数 H 表示一个卷积操作, U 表示一个上采样, []表示连接层。拿 X_{0,0} 到 X_{0,4} 的过程举例可知以下方程 4-2 所示:

$$X_{0,4} = H[X_{0,0}, X_{0,1}, X_{0,2}, X_{0,3}, U(X_{1,-3})]$$
(4-2)

U-net++的损失函数由二元交叉熵和 Dice 系数组合,公试如 4-3 所示:

$$L(Y, \bar{Y}) = -\frac{1}{N} \sum_{b=1}^{N} (\frac{1}{2} \cdot Y_{b} \cdot \log \bar{Y}_{b} + \frac{2 \cdot Y_{b} \cdot \bar{Y}_{b}}{Y_{b} + \bar{Y}_{b}})$$
(4-3)

Y为原始标签, \bar{Y} 为网络预测结果, N为批次大小, Y_b表示第b批的标签, 相应的 \bar{Y}_b 表示第b批数据的预测结果,

4.4. 实验方案与评估指标

4.4.1. 实验方案

本文在 U-net[42]网络和 U-net++[71]上进行对比实验,所有的实验均在 Intel Core i5-9400F 2.90GHz x6, Nvidia Geforce RTx 2060 SUPER 配置的电脑上实现, 本章代码基于深度学习框架 Pytorch。数据集有 1246 张图像,为让实验具有说 服力,本实验的训练集、验证集以及测试集的数据量以近似 3:1:1 的比例划分, 训练集数据共有 747 张,验证集有 247 张,测试集 249 张。所有的数据标签均 由湘雅三医院骨科以及影像科室的临床医生监督本文团队标定所得,标注工具 为标注精灵,输出为 Json 文件。每张数据均标有 4 类标签,分别为胫骨、腓骨、 踝骨、关节间隙,在本章实验中只用到关节间隙标签。本章实验中所有的数据 均为踝关节目标识别算法识别后裁剪出来的数据,尺寸统一为 256×256,以 0.5 的概率随机水平翻转。在训练阶段初始的学习率为 0.001, epoch 设置为 250。 每个 epoch 都记录一次模型的损失值。

4.4.2. 训练过程



如图 4-4 所示为 U-net 和 U-net++网络训练过程中损失函数的趋势。

可以看到在第 50 个 epoch 时,两个网络在验证数据集上的损失分数下降非常明显,在 A 点即第 68 个 epoch 时,U-net++验证数据集的损失分数已经达到最小值 0.102105。由图 4-4 可以看出,在训练过程中 U-net++在验证数据集上的损失值比 U-net 波动更稳定,U-net++更容易拟合本任务目标。

为了更直观地观察训练效果,在每一个 epoch 训练后,用一个非训练数据查 看模型的拟合程度,结果如图 4-5 所示:



图 4-5 不同的训练批次下的 U-net++预测结果图以及损失函数值,在第 68 次 epoch 时,损失值最小,分割效果也最好

由每幅图右上角的 Loss 值的变化过程可以看出在第 68 个 epoch 之前训练损失值下降明显,预测结果也出现较大的转变,获得的预测结果与标签最接近。

图 4-4 U-net 和 U-net++网络训练过程损失函数趋势,50次循环前,验证数据的损失分数下降非常明显,U-net++网络在 A 点即第68次循环时,验证数据集的损失值达到最小值。

在训练后期损失分数变动较小,预测结果变动较小。

4.4.3. 评估指标

分割模型常用 IOU (intersection over union)和 Dice (dice similarity coefficient)系数作为评价指标,这两个指标都可以用来度量两个集合之间的相 似程度,在评估图像分割模型的性能上则是反映出模型预测结果和真实标签之 间的相似度。另 P 为语义分割模型的预测分割结果,G 为真实标签 mask,则 IOU 和 Dice 的计算公式如式 4-4 所示:

$$Dice(\mathbf{P},\mathbf{G}) = \frac{2|\mathbf{P} \cap \mathbf{G}|}{|\mathbf{P}| + |\mathbf{G}|},$$

$$IOU(\mathbf{P},\mathbf{G}) = \frac{|\mathbf{P} \cap \mathbf{G}|}{|\mathbf{P} \cup \mathbf{G}|}$$
 (4-4)

其中|P∩G|表示真实标签 G 和预测结果 P 的交集,数学实现为矩阵点乘运算,然 后将点乘的元素结果相加。

为了更直观地理解,下面将通过几何方法给出解释。如图 4-6 所示,令蓝色 方框加黄色方框是真实 mask 标签,红色加黄色方框是模型预测结果,那么黄色 方框为 TP (True Positive),蓝色为 FP (False Positive),红色为 FN (False Positive)。



图 4-6 分割模型结果几何解释,蓝色为 FP (False Positive),黄色为 TP (True Positive),红色为 FN (False Positive)

因此 IOU 几何描述公式如式 4-5 所示:

$$IOU = \frac{TP}{TP + FN + FP}$$
(4-5)

而 Dice 的几何描述公式如式 4-6 所示:

$$Dice = \frac{2*TP}{TP + FN + TP + FP}$$
(4-6)

从几何描述可知, Dice 比 IOU 更接近分类网络中准确率(Accuracy)的概念, 而语义分割网络原理上就是一个像素级分类网络,因此本文用 Dice 评估模型的分割准确率。同理评估语义分割模型的精确率(Precision),其公式如式 4-7 所示:

$$Precision = \frac{TP}{TP + FP}$$
(4-7)

召回率(Recall)的公式如式 4-8 所示:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$
(4-8)

4.5. 实验结果分析

本文踝关节实验测试集包括 249 张数据,所有的测试数据经过 U-net 和 U-net++网络后都会生成关节缝隙预测图,通过预测图和标签图可以计算每个预测 图的 Dice、Precision、Recall 数值,由此可以评估比较 U-net 和 U-net++两种模 型在识别踝关节缝隙任务中的性能。为了能更好地分析数据,由医生根据其工 作经验,从测试集中选取了 87 个具有代表性的数据,进行详细讨论。

U-net 网络和 U-net++网络的 Dice 分数对比结果如图 4-7 所示:



图 4-7 U-net 和 U-net++网络在测试数据集上的部分 Dice 结果比较, 横轴是数据索引号, 纵轴是 Dice 分数, 图中 a、b、c、d 四点为最差分割结果

在测试数据集上,U-net 的平均 Dice 为 0.7656,U-net++的平均 Dice 为 0.7465。从平均 Dice 分数来看,U-net 的分割准确率比 U-net++高 0.02。通过对 比图 4-7 中的两个网络每例数据的 Dice 可知,大多数 U-net 预测结果的 Dice 值 要高于 U-net++。图 4-7 中 a、b、c、d 四个数据的 Dice 分数很低,其对应的数 据以及预测结果如图 4-8 中 a、b、c、d 所示。这四个数据均为有肌肉韧带损伤 的踝关节侧视方位图,在 a、c、d 三个图上 U-net 的 Dice 分数均高于 U-net++, 特别是在 d 图上,U-net++几乎没有预测到任何结果,U-net 的 Dice 值为 0.2278。从图 a、b、c 图中可以看出,由于踝关节中肌肉韧带的损伤,其胫骨和踝骨的 相对位置发生严重的移位,关节缝隙已经完全形变,失去了原有的结构,故而 导致 U-net 和 U-net++的预测结果极差。如图 4-8 中第三行和第四行所示,蓝色 预测掩码与真实标签重合面积较小,因此 Dice 分数非常低。



图 4-8 测试数据集中具有代表性的四个低 Dice 分值的数据样例,最上面一行为原始图像, 第二行是真实标签,第三行蓝色掩膜为 U-net++网络的预测值,第四行为 U-net 的预测值

U-net 和 U-net++在 87 个代表性数据上的 Precision 值如图 4-9 所示:



图 4-9 87 个代表性数据在两个网络中的 Precision 值对比,图中 a、b、c、d 点对应 4-7 中的 a、b、c、d

U-net 的平均 Precision 为 0.8135, U-net++平均 Precision 为 0.7784。U-net 的 平均 Precision 比 U-net++约高 0.04,这意味着相比于 U-net++而言,U-net 的预 测结果有更少的错误预测,因此其预测出来的结果可靠性更高。如图中 c 点对 应的数据,U-net 的 Precision 为 0.543, U-net++的 Precision 为 0.3465,结合 c 在图 4-8 中对应的预测掩膜结果可以看出,U-net 模型对于 c 数据的预测结果有 更少的错误,因此其预测结果更具有临床辅助参考价值。

U-net 和 U-net++在 87 个代表性数据上的 Recall 值如图 4-10 所示:



图 4-10 87 个代表性数据在两个网络中的 Recall 值对比, 图中 a、b、c、d 点对应图 4-8 中的 a、b、c、d

U-net 的平均 Recall 为 0.7416, U-net++平均 Recall 为 0.7419, 两个网络的召回

率差距不大。Recall 在评估语义分割结果时,表示的是预测的掩码与真实标签的重合面积占真实标签面积的比值,Recall 越高,则网络漏检的可能性越少。 踝关节缝隙语义分割实验较好的预测结果如图 4-11 所示:



图 4-11 U-net 网络和 U-net++网络测试数据预测结果对比,从整体来看,两个网络可以很好 地的分割出踝关节缝隙,从 e 项和 d 项可以看出 U-net 网络预测结果连续性更好

图 4-11 中, a、c、e 都是踝关节正视图,在这三副图中两个网络的 Dice 值均大 于 0.6。其中 a 小图在两网络上的 Dice 值分别为 0.832 和 0.8319,说明预测图和 真实 mask 重合度很高,从 Dice 值和预测图像上来看 a 小图上 U-net++的预测结 果略好于 U-net。在 e 小图上,两网络的 Dice 分数差距比较大,U-net++为 0.6373,而 U-net 为 0.8159,可以很明显地看到 U-net++预测图损失了很多胫骨 和腓骨缝隙的细节,U-net 则能很好地预测结果。在 c 小图上,U-net 的 Dice 值 略好于 U-net++。b、d、f 三个小图都是踝关节侧视图,U-net++的 Dice 值均优 于 U-net 网络,从 d、f 两个小图可以很明显地看出 U-net++的预测图比 U-net 的 预测图具有更完整的结构,但是边界更加模糊。由上分析可知,U-net 在踝关节 正视图这种干扰较少的情况下有更好的识别性能,能识别出更完整的踝关节缝 隙结构。U-net++在踝关节侧视图这种干扰较多的环境中能识别出更完整的踝关 节缝隙结构,但是边界更模糊。b、f 关节两网络 Dice 分数不高的主要原因是 x 光片中叠影造成了干扰,如图 4-12 所示红色和黄色线是踝关节缝隙的边界,由 于 x 光片拍摄视角原因,缝隙边界相互叠加导致叠影,故而影响了机器的预测 结果。在实际临床中,侧视位拍摄无法避免叠影的产生,因此在数据量不大的

b

条件下还无法很好地解决侧视位预测分割性能低的问题。

图 4-12 骨骼 x 光片中叠影的产生和干扰导致 Dice 分数不高

通过以上的对比实验证明本文收集整理的踝关节缝隙语义分割数据集,在 U-net 和 U-net++上均可以训练出收敛的语义分割模型,因此具有有效性。另外 U-net 网络训练出来的模型比 U-net++的模型可靠性更高,准确率的更好。

表 4-1 中列出了 U-net 和 U-net++两个网络在本任务中的平均 Dice、平均 Precision 以及平均 Recall。

表 4-1 U-net 和 U-net++在测试数据集上的平均 Dice、平均 Precision 以及平均 Recall

网络	平均 Dice	平均 Precision	平均 Recall
U-net	0.7656	0.8135	0.7416
U-net++	0.7465	0.7784	0.7419

从表中可以看出在验证数据集上, U-net 的平均 Dice 和平均 Precision 都要 比 U-net++更高。

表 4-2 中列出了 U-net 和 U-net++两个网络的参数规模和推理计算效率。

网络	Parameters	FPS				
U-net	45.9M	25				
U-net++	67.2M	30				

. イ .. . 儿 台 如 旧 世 イ 门 旋 江 南

从表中可以看出由于 U-net++包含许多跳连接,因此其参数规模比 U-net 大, U-net++为 67.2M 而 U-net 为 45.9M。在推理阶段, U-net 为每秒 25 帧, U-net++ 为每秒 30 帧, U-net++略快于 U-net。

4.6. 本章小结

本章首先探讨了踝关节缝隙特征提取的应用背景及临床辅助意义,明确了 本章节的目标是训练一个语义分割模型,其可以从 x 光片中提取出关节缝隙, 使得临床医生可以获得更好辅助诊断信息,即踝关节缝隙的物理几何信息。然 后介绍了本文针对踝关节缝隙语义分割任务构建的数据集,其原数据由湘雅三 医院骨科和影像科提供,共包含 1246 张符合要求的高质量踝关节 x 光片,经整 理和标注得到了一个有高质量像素级语义分割标签的数据集。接着描述了相关 的实验,本章节所有实验均基于 U-net 和 U-net++两种语义分割算法,通过观察 这两种网络训练过程中损失函数值的变化趋势以及对验证集预测结果的输出, 可知网络模型能够收敛。最后通过分析实验预测结果和 Dice、Precision 等评价 指标,证明了本文踝关节缝隙语义分割数据集的有效性,以及提取踝关节缝隙 的可行性。从训练过程和实验结果可以看出,在本任务中 U-net 和 U-et++网络 各具特点,从训练过程可以看出,U-net++具有更好的收敛性。而在验证阶段 U-net 网络比 U-net++网络具有更好的性能。

第5章 足踝关节骨损辅助检测系统设计与实现

骨科疾病中,骨折是一种老年常见疾病,其隐蔽性高,并发症严重,恢复 时间长,严重影响了老年人生活质量。踝关节肌肉韧带损伤常发生于青少年和 运动员等运动量大的人群,其受伤初期难以诊断,容易被忽视,日积月累易留 下后遗症。目前骨科临床医生诊断上述疾病主要依靠多年积攒的临床经验,工 作经验不足容易对小尺寸骨损以及韧带损伤地做出错误的诊断,从而影响治疗 质量。在中国许多基层地区,骨科医生少检测设备相对较差,工作经验和诊断 能力参差不齐,该类疾病容易被忽视。为了解决这一问题,本文针对足踝关节 开发了一套足踝骨科疾病自动检测系统。本系统能辅助医生识别足踝关节小尺 寸骨损以及辅助鉴别韧带肌肉损伤,患者亦可使用本套系统智能自诊,同时为 算法的更新和迭代提供一个数据收集平台。

5.1. 系统功能

基于本文足踝关节骨损检测算法,本系统以骨骼 x 光片临床图像为输入, 涉及骨骼疾种包括足踝关节骨折、关节肌肉韧带损伤,系统功能强大且操作容 易上手,具体系统功能如下:

5.1.1. 核心功能

该套系统主要面向患者以及相关医生,所有功能都以患者和医生的操作使 用体验为核心,为了能让用户有更加直观的感受,在智能诊断首页选患病部位 界面我们除了在醒目的位置添加了常见病患处,还添加了全身骨骼三维导航图, 如图 5-1 中 A 所示。选好部位,我们可以通过拍照上传的方式和文件上传的方 式上传数据,该功能满足了医生上传高质量 x 光片的需求也满足了患者操作快 捷的需求,在上传符号下面显示有患者的基本信息供用户核验,如图 5-1 中 B 所示。点击上传按钮后终端就会将患者信息通过后端系统发给算法,最终终端 得到算法返回的结果如图 5-1 中 C 所示。点击查看报告按钮可获得详细信息, 如图 5-1 中 D 所示,其中最上面会显示传入数据缩略图以及患者基本信息,分 享报告按钮可供用户之间分享结果报告,中间的黄色页面显示了报告粗略信息, 其中包括是否有骨折,是否可能有肌肉韧带损伤,骨龄的预测值在全国的排名 情况,下面的小页面分别是骨损检测结果详情、缝隙检测详情、专家联合诊断 以及康复计划,点击可以查看详情。



图 5-1 终端上传数据以及查看结果流程图

5.1.2. 全自动骨损伤检测定位功能

基于本文性能稳定、识别准确率高的骨损检测算法,不管是 300×300 以上的大尺寸骨损还是 50×50 的小尺寸骨损,该辅助系统都可以快速地识别踝关节骨损并给出定位边界框。如图 5-2 所示,在智能诊断结果报告中点击骨损检测结果后便进入了骨损分析结果界面,其详解地展现了骨损的信息,如图 5-2 中B 所示,其中包括骨损定位边界框以及骨损置信度,用图片显示和文字描述的形式展现给用户。



图 5-2 骨损结果详情查看界面

5.1.3. 韧带肌肉损伤检测功能

该辅助诊断系统依靠本文中的关节缝隙提取算法,可以很精确地提取出踝 关节缝隙的物理几何信息,如图 5-3 中结果所示,算法将识别出踝关节缝隙并 用掩膜显著标注。然后辅助系统将该结果发送给相应的骨科专家,结合患者的 基础信息以及骨损位置信息,骨科专家医生可以做出高价值的临床诊断报告, 辅助系统结合用户信息将检测诊断结果反馈给用户,如图 5-3 中 B 所示。



图 5-3 韧带肌肉损伤检测流程图

5.1.4. 专家联合会诊功能

本算法暂处于试验阶段,相应的算法能力和疾病识别种类还有待提高和扩展,面对疑难杂症仍需要专业医生进行合作诊断。因此本系统设置了专家联合会诊功能,为用户提供多维度的诊断通道。如图 5-4 所示,在图 A 中点击专家联合会诊按钮后,界面跳到图 B 专家联合会诊界面,可以从中选择会诊日期,并罗列了会诊日期当日可选择的专家,每个专家都会给出职称、所属医院科室、擅长领域供用户参考,用户点击专家简介后方邀请按钮,便可以邀请该专家进入联合会诊,每名病人可以同时要求多名专家经行会诊。



图 5-4 专家联合会诊界面示意图

5.2. 系统设计与实现

本文足踝骨科疾病自动检测系统基于 C/S 网络架构开发,系统由客户端(C) 和服务器(S)两个组成部分。如图 5-5 所示,客户端 App 可以上传用户医疗信 息并接受来自服务端处理的结果。服务端由后端分发平台、文件存储集群、数 据库以及核心算法集群四大部分组成,共同为客户端的请求数据服务。客户端 和服务端通过 HTTP 协议实现通信。下面将分别介绍客户端与服务端的工作流 程和实现方法。



图 5-5 基于 C/S 网络框架的系统框架图, 通过 HTPP 进行通信

5.2.1. 客户端实现方法以及主要工作流程

客户端 App 基于 Android 操作系统 10 版本开发,允许运行最低硬件环境为: CPU: 骁龙 801;内存: 2GB 以上;分辨率: 1920×1080;内存: 64MB,推荐 128MB 以上,颜色 256。App 菜单界面主要包括注册登录界面、骨科学院界面、 智能诊断界面、用户信息界面,如图 5-6 所示。客户端 APP 的核心功能是智能 检测,该功能可以实现踝关节骨损自动检测以及踝关节缝隙自动分割,检测结 果可以辅助医生进行临床诊断。智能检测界面主要操作逻辑介绍如下:

 1)菜单选择界面:用户输入账号密码成功后会默认显示首页内容,最下面 菜单条会显示"首页"、"骨科学院"、"智能诊断"、"我的"以及"地图导航", 用户可根据需要选择菜单,需要智能诊断的用户点击智能诊断菜单选项;

 2)患者信息输入:智能诊断界面首先需要获取患者的基本信息,包括年龄、 性别、数据来源以及病患部位。App 中提供多种病患部位的选择方式,分别为 关键字搜索、图标选择以及三维骨架选择,患者根据需要选择病患部位; 3) 文件上传:输入患者信息后,进入文件上传界面。App提供了两种提交方式,一种是手机拍照方式,另一种是文件上传方式,其中可以接受JPG、 PNG等多种格式。确认信息无误后,用户点击上传按钮,完成患者信息输入。

4) 等待结果页面: 客户端端把数据发送给服务端后,客户端保留一结果查 询个进程,每隔 5s 向服务端发送一个结果查询请求,直到服务端将计算结果返 回给客户端。客户端其他空闲进程可以继续服务用户的操作,实现异步请求的 效果。

5)结果显示页面:客户端得到计算结果后,首先将结果数据和图片保留在 客户端本地,同时会提醒用户结果已收到请尽快查收。用户可以点击界面骨损 检测报告和缝隙检测报告按钮查看检测详情,也可以点击专家联合会诊按钮邀 请相关专家给出进一步的治疗建议。



图 5-6 终 App 注册登录以及菜单界面

5.2.2. 服务端实现方法和工作流程

服务端由后端分发平台、文件存储集群、数据库以及核心算法集群四大部 分组成,如图 5-7 所示:



图 5-7 后端分发平台、文件存储集群、数据库以及核心算法集群共同构成服务端

整个服务端基于 SpingMVC 框架搭建而成,该框架是一个基于 Java 语言开

发实现的轻量级 Web 框架,其中包含了表现层(Controller)、业务层(Service)、 持久层(Dao)三层,是一个典型的 MVC 服务框架,每一层分工合作完成不同 的任务,实现了业务代码的高内聚低耦合的思想,该框架配置简单又不失灵活 性,代码的可重用性很高,可扩展性好。WEB 容器为 Tomcat,数据库采用 MySQL,,文件存储集群采用 NAS (Network Attached Storage: 网络附属存储)。 服务端的逻辑如图 5-8 所示,简要的工作流程介绍如下:

1)WEB 容器 Tomcat 一直等待客户端的请求,当请求为调度算法进行智能 诊断时,后台分发平台将获取将请求信息转发给 MySQL,同时把图像存储到 NAS,并把图像地址更新到 MySQL 中,然后调度计算资源分析疾病。当请求 是查询结果时,分发平台会根据请求中的信息去 RabbitMQ 结果消息队列查询, 如果查到有相关消息,则把结果反馈给客户端,如果没有则告知客户端还没有 结果。

2)图片数据存储:后端分发平台会将图像存储到 NAS 中,存储成功后 NAS 会返回一个存储地址。

3)信息存储:后端分发平台将获得的图片存储信息和患病部位信息存储到 MySQL数据库中。

4)调用计算资源:后端分发平台将图片存储地址发送给算法集群,经过核心算法的检测获得病患相关检测结果信息,最后把结果信息发送至基于 RabbitMQ的结果消息队列。



5.3. 本章小结

本章首先介绍了足踝关节骨损辅助诊断系统的主要功能,包括全自动骨损检测 定位以及肌肉韧带损伤检测功能以及专家联合会诊功能。接着,本章对系统做 了一个简单概要,然后分别介绍了客户端和服务端的实现方法和工作流程。客 户端基于安卓操作系统开发,主要面向患者和医生,该 App 为用户提供了图片 数据上传以及检测结果显示功能。服务端基于 Spring MVC 框架开发,其中主 要包括后端分发平台、文件存储集群、数据库以及核心算法集群四大部分,集 数据存储、智能诊断、信息分发于一体,为客户端提供强大的服务。

第6章 总结与展望

6.1. 论文总结

随着人民生活水平质量的提高,人们对高质量的医疗服务需求越来越旺盛, 短缺的医疗人才难以支撑庞大的市场需求。智能辅助诊断系统的出现可以辅助 医生处理许多枯燥重复的工作,提高医疗工作者的工作效率,并且稳定可靠的 辅助诊断系统可以平衡不同区域之间的医疗资源,具有很好的社会意义以及研 究价值。随着深度神经网络的迭代发展,图像识别技术得到了长足的进步,基 于图像识别的医疗辅助诊断如雨后春笋般发展迅速,骨骼方面的辅助诊断研究 也日新月盛,但是对于踝关节骨损的检测以及踝关节缝隙的检测却寥寥无几, 针对骨损小目标识别的研究就更加风毛麟角。在本文中,本文和湘雅三医院骨 科以及影像科合作开发了一套踝关节骨损检测系统,主要包含三大块,一是把 数据分为非完全粉碎型骨损和完全粉碎型骨损的分类网络,二是在非完全粉碎 型骨损数据中识别踝关节骨损的目标识别网络,三是在非完全粉碎型骨损数据 中分割踝关节间隙的语义分割网络。在每个部分有不同的研究侧重点,本文具 体工作可以概述为以下几点:

1. 本文中所有研究数据均由本文团队和湘雅三医院的骨科以及影像科合作整理所得,在实际临床数据中踝关节骨损存在完全粉碎型骨损类型,该类骨损损伤严重,临床处理与非完全粉碎型骨损大相径庭,已不在本文研究范围内,所以本文训练分类网络识别数据中的完全粉碎型骨损和非完全粉碎型骨损。在文中本文训练对比了 VGG-16、GoogleNet、ResNet-50 三种分类网络,其中GoogleNet 和 ResNet-50 的分类效果最好,其准确为 0.9704,但是 GoogleNet 的 召回率高,更适合本文辅助系统。

2. 在非完全粉碎型骨损数据上本文研究了相关的小目标识别问题,因为在 数据中原始图像尺寸很大,而骨损目标尺寸差异性大,存在相对尺寸较小的骨 损目标。本文实验对比了缩放和截图两种不同的输入数据预处理方式给网络识 别带来的性能差异,结果显示训练踝关节识别模型预先截取高价值踝关节的预 处理方式比直接把原图压缩为 1000×1000 的预处理方式效果更好。其原因是骨 损主要集中分布在踝关节周围,因此踝关节以外的信息便成了无效干扰信息, 截图的方式就是排除这些干扰信息从而提升骨损有效信息的比例,而压缩则会 减少骨损有效信息的比例。本研究从优化网络结构提升小目标识别性能角度出 发,将注意力机制模块 CBAM 融合到目标识别算法的主干网络中,通过对比实

验 YOLO V3 等目标识别网络添加 CBAM 和不添加 CBAM 的性能差异,验证了 注意力机制在本研究任务中的有效性。

3. 经本文和湘雅三医院骨科专家的讨论, 踝关节缝隙的物理几何信息在临床上对辅助诊断踝关节的肌肉韧带损伤具有很好的指导意义,因此针对踝关节缝隙的提取本文做了语义分割算法的相关研究。首先本文在医生专家的指导下对数据进行了高质量的像素级语义标注, 然后对比了 U-net++和 U-net 两种语义分割网络的分割性能差异。实验结果显示 U-net 的平均 Dice 为 0.7656, U-net++的平均 Dice 为 0.7465,因此 U-net 的分割效果比 U-net++效果更好。

4. 本文基于踝关节骨损检测算法开发了一套踝关节辅助诊断系统,其包括 了终端 App、后端分发平台、算法集群、数据库以及文件存储集群。患者和医 生只需在 App 中提供必要的基本信息和骨骼数据,就能很方便地利用本文的算 法资源进行一个快速的踝关节疾病的辅助检测,大大节省了患者和医生的时间, 后期经多地推广可以缓解地区性的医疗资源不平衡问题。

6.2. 研究展望

虽然本文的踝关节疾病辅助诊断研究取得了不错的结果,在本文的数据集 上有不错的性能指标,并且基于该算法开发的踝关节骨损自动检测系统已经在 小范围内开始测试,达到了医生和本文都满意的表现。但是在真实使用过程中, 暴露出了许多还需要进一步研究的问题,这些问题如下所示:

1. 覆盖更加多样全面的骨科病例。在本研究中本文只收集了踝关节骨损相 关的数据,也只做了这一种疾病的研究,但在实际的生活中,还有许多其他部 位的骨骼也会发生类似的骨折骨损以及肌肉韧带损伤,比如手腕、肩关节、髋 关节、颈椎等等。因此本文希望在该项目的基础上,进一步研究更多部位的相 关疾病,打造一个更完善且多样化的骨骼辅助诊断平台。

2. 融合多种类型的数据。本文中的研究内容都基于图像数据,不管是分类、 目标识别还是语义分割都是基于骨骼的 x 光线图像。但是在实际的临床诊断中, 医生会综合多维度的信息进行综合分析,比如病人的身体基础信息、过往病史、 先天性疾病等等。因此以后本文还需要进行多数据融合方面的研究,提升有效 信息的利用效率,综合多种数据给出可靠性更高的诊断结果。

3. 提升踝关节缝隙分割的性能。本文中运用了 U-net 和 U-net++两种语义 分割算法验证了分割踝关节缝隙的有效性,但是由于骨骼叠影的干扰,制约了 分割性能的提升,其两个网络的 Dice 平均值都为 0.75-0.8 之间。因此该方向还 需要更深入的研究,以优化相关算法和数据,提升踝关节提取的性能。

参考文献

- Charles M. Court-Brown, Ben Caesarn: Epidemiology of adult fractures: A review. Injury 37: 691-697, 2006.
- [2] Moloney G B, Pan T, Van Eck C F, et al. Geriatric distal femur fracture: are we underestimating the rate of local and systemic complications?[J]. Injury, 2016, 47(8): 1732-1736.
- [3] xU Hal-lin, LIU Li-min, LI xuan, etc: Multicenter follow-up study of ankle fracture surgery. Chinese Medical Journal 125(4): 574-578, 2012.
- [4] 王智勇,钟秀,张云凤.医学影像学在临床骨科中的应用进展[J].中国卫生产业,2016,13(27):196-198.
- [5] 张程.医学影像学在不同骨科疾病诊断中运用[J].影像研究与医学应用,2020,4(19):3-5.
- [6] Yee Liang Thian, Yiting Li, Pooja Jagmohan, David Sia, Vincent Ern Yao Chan, Robby T. Tan: Convolutional Neural Networks for Automated Fracture Detection and Localization on Wrist Radiographs. Radiology: Artificial Intelligence, DOI: 10.1148/ryai.2019180001, Jan 30, 2019.
- [7] 翟禹樵,李开南.骨科人工智能诊断的研究进展[J].中国临床研 究,2021,34(04):542-545.
- [8] Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis, 2017, 42:60-88.
- [9] Rajpurkar P, Irvin J, Zhu K, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning[J]. arxiv preprint arxiv:1711.05225, 2017.
- [10]Dong H, Yang G, Liu F, et al. Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks[C]//annual conference on medical image understanding and analysis. Springer, Cham, 2017: 506-517.
- [11]Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639):115-118.
- [12]Haenssle, Holger A., et al. "Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists." Annals of Oncology 29.8 (2018): 1836-1842.
- [13]Gutman D, Codella N C F, Celebi E, et al. Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)[J]. 2016.
- [14]Zhang J, xie Y, Wu Q, et al. Skin Lesion Classification in Dermoscopy Images Using Synergic Deep Learning: 21st International Conference, Granada, Spain, September 16–20, 2018, Proceedings, Part II[M]// Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. 2018.
- [15] 胡颜民. 基于医学图像处理关键技术的骨科疾病计算机辅助诊断系统[D]. 华南理工大学,2020.贾文慧.
- [16]贾文慧. 基于 XGBoost 算法的骨科辅助诊断模型研究[D].太原理工大学,2018.
- [17]Lindsey, Robert, et al. "Deep neural network improves fracture detection by clinicians." Proceedings of the National Academy of Sciences 115.45 (2018): 11591-11596.
- [18]Chung S W, Han S S, Lee J W. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm.[J]. Acta Orthopaedica, 2018, 89(4):468-473.
- [19]Olczak, Jakub, et al. "Artificial intelligence for analyzing orthopedic trauma radiographs: deep learning algorithms—are they on par with humans for diagnosing fractures?." Acta orthopaedica 88.6 (2017): 581-586.
- [20] Jakub Olczak, Filip Emilson, Ali Razavian, Tone Antonsson, Andreas Stark, Max Gordon. (2020) Ankle fracture classification using deep learning: automating detailed AO Foundation/Orthopedic Trauma Association (AO/OTA) 2018 malleolar fracture identification reaches a high degree of correct classification. Acta Orthopaedica 0:0.
- [21]Kitamura, Gene, Chul Y. Chung, and Barry E. Moore. "Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation." Journal of digital imaging 32.4 (2019): 672-677.

- [22] Yamada, Yutoku, et al. "Automated classification of hip fractures using deep convolutional neural networks with orthopedic surgeon-level accuracy: ensemble decision-making with antero-posterior and lateral radiographs." Acta Orthopaedica (2020): 1-6.
- [23]Gan, Kaifeng, et al. "Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments." Acta orthopaedica 90.4 (2019): 394-400.
- [24]Ebsim R., Naqvi J., Cootes T.F.: Automatic Detection of Wrist Fractures From Posteroanterior and Lateral Radiographs: A Deep Learning-Based Approach. Computational Methods and Clinical Applications in Musculoskeletal Imaging 11404: 114-125, 2018.
- [25]Lefevre, T., et al. "Automatic knee meniscus tear detection and orientation classification with Mask-RCNN." (2019).
- [26]Prasoon A , Petersen K , Igel C , et al. Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network[M]. Medical Image Computing and Computer-Assisted Intervention. MICCAI 2013. Springer Berlin Heidelberg, 2013.
- [27] Matthew Adams, Weijia Chen, David Holcdorf, Mark W McCusker, Piers DL Howe, Frank Gaillard, Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures, Journal of Medical Imaging and Radiation Oncology, 63 (2019) 27–32.
- [28]Urakawa, Takaaki, et al. "Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network." Skeletal radiology 48.2 (2019): 239-244.
- [29]Brett, Alan, et al. "Development of a clinical workflow tool to enhance the detection of vertebral fractures: accuracy and precision evaluation." Spine 34.22 (2009): 2437-2443.
- [30]Sa, Ruhan, et al. "Intervertebral disc detection in x-ray images using faster R-CNN." 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017.
- [31]Heimer, Jakob, Michael J. Thali, and Lars Ebert. "Classification based on the presence of skull fractures on curved maximum intensity skull projections by means of deep learning." Journal of Forensic Radiology and Imaging 14 (2018): 16-20.

- [32] Tomita, Naofumi, Yvonne Y. Cheung, and Saeed Hassanpour. "Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans." Computers in biology and medicine 98 (2018): 8-15.
- [33]Stern D, Payer C, Lepetit V, et al. Automated Age Estimation from Hand MRI Volumes using Deep Learning[C]. Medical Image Computing and Computer Assisted Intervention. MICCAI 2016. Springer International Publishing, 2016.
- [34]Kim, D. H., and T. MacKinnon. "Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks." Clinical radiology 73.5 (2018): 439-445.
- [35]Rajpurkar P, Irvin J, Bagul A, et al. MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs[J]. 2017.
- [36] AIMI. Artificial intelligence in medicine & imaging: Available labeled medical datasets.https://aimi.stanford.edu/available-labeled-medicaldatasets.[Online;acce-ssed 2-December-2017].
- [37]OAI. Osteoarthritis initiative: a multi-center observational study of men and women.https://oai.epi-ucsf.org/datarelease/.[Online; accessed 2-December-2017].
- [38] Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong, Bagheri, Mohammadhadi, and Summers, Ronald M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. arXiv preprint arXiv:1705.02315, 2017.
- [39]Demner-Fushman, Dina, Kohli, Marc D, Rosenman, Marc B, Shooshan, Sonya E, Rodriguez, Laritza, Antani, Sameer, Thoma, George R, and McDonald, Clement J. Preparing a collection of radiology examinations for distribution and retrieval. Journal of the American Medical Informatics Association, 23(2):304–310, 2015.
- [40] Jaeger, Stefan, Candemir, Sema, Antani, Sameer, Wang, Yi-Xiang J, Lu, Pu-Xuan, and Thoma, George. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery, 4(6):475, 2014.
- [41]Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.

- [42]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. Springer, Cham, 2015: 234-241.
- [43]Borji A, Sihite D N, Itti L. Salient object detection: A benchmark[C]. Proc of European Conf on Computer Vision. Florence: Springer Berlin Heidelberg, 2012: 414- 429.
- [44] A. Krizhevsky et al., "ImageNet classification with deep convolutional neural networks", Proc. NIPS, pp. 1097-1105, 2012.
- [45]G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks", Science, vol. 313, no. 5786, pp. 504-507, 2006.
- [46]R. Girshick et al., "Rich feature hierarchies for accurate object detection and semantic segmentation", Proc. CVPR, pp. 580-587, 2014.
- [47] R. Girshick, "Fast R-CNN", Proc. ICCV, pp. 1440-1448, 2015.
- [48]S. Ren et al., "Faster R-CNN: Towards real-time object detection with region proposal networks", Proc. NIPS, pp. 91-99, 2015.
- [49]Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [50]Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [51]Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arxiv preprint arxiv:1804.02767 (2018).
- [52]K. He et al., "Spatial pyramid pooling in deep convolutional networks for visual recognition", IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904-1916, Sep. 2015.
- [53]J. Dai et al., "R-FCN: Object detection via region-based fully convolutional networks", Proc. NIPS, pp. 379-387, 2016.
- [54]T.-Y. Lin et al., "Feature pyramid networks for object detection", Proc. CVPR, pp. 936-944, 2017.
- [55]Uijlings J R R , K. E. A. van de Sande, et al. Selective Search for Object Recognition[J]. International Journal of Computer Vision, 2013, 104(2):154-171.

- [56]Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012) : 1097-1105.
- [57]Çiçek, Özgün, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. "3D U-Net: learning dense volumetric segmentation from sparse annotation." In International conference on medical image computing and computer-assisted intervention, pp. 424-432. Springer, Cham, 2016.
- [58] Ibtehaz N, Rahman M S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation[J]. Neural Networks, 2020, 121: 74-87.
- [59] Alom, Md Zahangir, Mahmudul Hasan, Chris Yakopcic, Tarek M. Taha, and Vijayan K. Asari. "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation." arxiv preprint arxiv:1802.06955 (2018).
- [60]Oktay, Ozan, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori et al. "Attention U-Net: learning where to look for the pancreas." arxiv preprint arxiv:1804.03999 (2018).
- [61]Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arxiv preprint arxiv:1409.1556 (2014).
- [62] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
- [63]He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [64]Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
- [65]Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]. IEEE Conference on Computer Vision & Pattern Recognition. 2009.
- [66] Everingham M, Eslami S M A, Van Gool L, et al. ThePascalVisual Object Classes Challenge: A Retrospective[J]. International Journal of Computer Vision, 2015, 111(1):98-136.

- [67]Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context[J]. 2014(04):740-751.
- [68]Mnih V, Heess N, Graves A. Recurrent models of visual attention[C]//Advances in neural information processing systems. 2014: 2204-2212.
- [69] Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." Proceedings of the European conference on computer vision (ECCV). 2018.
- [70]Park, Jongchan, et al. "Bam: Bottleneck attention module." arxiv preprint arxiv:1807.06514 (2018).
- [71]Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation[M]//Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, Cham, 2018: 3-11.

攻读学位期间主要研究成果

一、学术论文和专利成果

[1] 骨科疾病智能诊断软件 [简称: AI 骨科软件] V1.0, 软著申请登记号: 2021SR0179693, 证书号: 软著登字第 6904010 号, 开发完成时间: 2020 年 6 月 01 日, 申请日期: 2021 年 02

二、硕士期间获奖情况

- [1] 国家级,中国"互联网+"大学生创新创业比赛,国家银奖,2018年。
- [2] 国家级,全国研究生数学建模竞赛,国家三等奖,2019年。

[3] 省级,湖南省数学建模竞赛,省二等奖,2019年。

[4] 校级,格林美团队奖,2019

三、参加科研项目

[1] 基于临床图像皮肤疾病辅助诊断系统研究,2018.9~至今,项目组主要成员。

- [2] 多模态脑部图像的肿瘤特征分析, 2018.9-2019.5, 项目主要成员。
- [3] 踝关节骨损辅助诊断系统的研究, 2018.9~至今, 项目主要成员。
- [4] 书法评测系统及算法的研究, 2019.6~至今,项目主要开发人员。

致谢

时光荏苒,在中南大学读研的日子马上就要接近尾声,开启人生的下一个 阶段。在这三年里,我受益良多,在此感谢在中南大学遇到的所有人。

首先感谢我的导师,谢斌老师。谢斌老师对科研敏锐的洞察力和前瞻性开 阔了我的眼界,提高了我的格局,经常在科研的岔路口指引我前进的方向,而 且您对人亲切,关心学生,做事坚定,不仅传授给我专业知识,还教会了我许 多人生哲理。在此,我衷心地感谢我的老师!除此之外,还要感谢智能所的王 勇老师、陈白帆老师、刘丽珏老师、王勇老师和王冰川老师等,感谢你们的精 心授课和在学业上给予的指导!

感谢李沁师姐、何小宇师兄、宋迪师兄、秦觅师兄、林学斌师兄、张文婷 师姐、严孝鑫师兄、何志远师兄,你们奋发努力的精神激励着我,在我刚到中 南大学的时候,给我树立很好的榜样,指引了我前进的方向。在我对未来迷茫 不知的时候,你们无私地分享了宝贵经验,指导我新的努力方向,让我少走了 很多弯路。非常幸运能遇到这么优秀的师兄师姐们!

感谢赵新宇、丘文杰、徐勇、程清、孙铭蔚、吴迪、万思远、李旭晖、后 国炜、邝逸零、懈振南、聂海涛等实验室的小伙伴,我们度过了充满欢声笑语 和火锅的三年,感谢有你们的陪伴和帮助,让我的读研生活变的温馨且多彩。 感谢刘飞、魏亚东、宋宝军、汤松林、何诗惠、袁迪等同届小伙伴,每一位好 友都将是我未来人生路上最宝贵的财富。

最后,我还要特别感谢我亲爱的家人,正是有你们永远都我身后默默地支持和付出,我才能有信心一步一步获得研究生学位。二十多年的求学生涯,你 们辛苦了!在此,也对这一路来曾经给予过我帮助的人们致以最真挚的感谢!

> 2021年4月19日 彭哲

70