分类号:

密级:

单位代码: 10441

学 号: 2019111336

① 少东中区药大学 硕士学位论文

中文题目:糖尿病视网膜病变智能诊断技术研究

英文题目: Research on Intelligent Diagnosis technology of Diabetic Retinopath

2022年6月20日

原创性声明

本人郑重声明: 所呈交的学位论文,是在导师的指导下独立完成的,文中除注明引用的内容外,不包含任何其他已经发表的科研成果。对本文研究做出重要贡献者,均已在文中以明确方式表明。本声明的法律责任完全由自己承担。

论文作者签名。一号师签名。一号师签名。一号师签名。一号师签名。

关于学位论文使用授权的声明

本人完全了解山东中医药大学有关保留使用学位论文的规定,同意学校保留或向国家有关部门机构送交论文的复印件和电子版,允许被查阅和借阅。本人授权山东中医药大学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,可以采用影印、缩印或其他复印手段保存和汇编本学位论文。

(保密论文在解密后应遵守此规定)

论文作者签名 导师签名 与期 2011. 6.20

糖尿病性视网膜病变 (DR) 是引起人类视觉损伤和失明的主要因素。临床研究显示,只有对糖尿病人进行早期 DR 筛查和及时诊断,才能预防其视力衰退。目前病变主要通过人工诊断的方式进行治疗,然而随着筛查数量的增加,医生在诊断时非常容易出现漏诊、误诊、反馈不及时等问题,从而错过病情的最佳治疗时机。为此,亟需开发有效的智能技术来进行病变诊断。近年来,机器学习和深度学习等智能技术被广泛应用于糖尿病视网膜病变诊断领域,对于辅助医生进行大规模筛查,提高分类精度和诊断效率发挥了重要作用。本文结合经典图像处理技术,基于机器学习和深度学习方法研究 DR 等级分类问题,具体研究内容如下:

- (1)为了解决图像光照不均、数据不平衡等问题的干扰,采用裁剪、灰度化、 颜色与亮度归一化、图像增强、伽马校正以及图像扩充等预处理操作。其中图像扩充 采用翻转、平移、旋转等。
- (2)基于已有模型仅针对深度特征进行融合的问题,提出一种多类别特征融合分类模型。根据眼底病灶的特点,在提取 AlexNet、VGG-16 和 ResNet-50 三种深度特征的基础上,新添加了 Gabor、LBP、HOG 以及 Haralick 等四种影像组学特征,并采用双层特征融合网络对四种影像组学特征和三种深度特征进行筛选,最后通过 SVM、RF、KNN、ELM 对病变进行分级,其准确率分别达到 88.64%、86.63%、86.27%、85.61%。有效避免了深度网络底层特征提取不足以及因影像组学特征维度较少造成的特征丢失问题。
- (3)为提高模型对小病灶的关注程度,提出一种基于深度可分离卷积和注意力机制的分类模型。其中,注意力机制基于 1×1 卷积和激活函数对关键特征进行识别,同时采用轻量化模块对模型的标准卷积进行替换,在保证精度的同时,实现病变的快速分类。经过在数据集上验证,其分类准确率达到 91.5%,召回率达到 90.3%,参数量为 12.24M,FLOPs 为 9.74G,运行时间为 158.7ms,Kappa 值为 0.893。在解决多类别特征网络提取背景、结构等无关信息过多,增强对小病灶特征关注程度的同时,有效降低了模型参数,减少运行时间。

关键词 糖尿病视网膜病变; 医学图像分类; 特征融合; 轻量化模型; 注意力机制

Research on Intelligent Diagnosis technology of Diabetic Retinopathy

Speciality: Biomedical Engineering

Author: Nie Yongqi

Tutor: Cao Hui

ABSTRACT

Diabetic retinopathy (DR) is a major cause of visual impairment and blindness in humans. Clinical studies show that only early DR screening and timely diagnosis of diabetic patients can prevent their vision decline. At present, the disease is mainly treated by manual diagnosis. However, with the increase of the number of screening, it is very easy for doctors to miss diagnosis, misdiagnosis, feedback is not timely and other problems, so as to miss the best treatment opportunity of the disease. Therefore, it is urgent to develop effective intelligent technology for disease diagnosis. In recent years, intelligent technologies such as machine learning and deep learning have been widely applied in the diagnosis of diabetic retinopathy, playing an important role in assisting doctors in large-scale screening and improving classification accuracy and diagnostic efficiency. Combined with classical image processing technology, this paper studies DR classification based on machine learning and deep learning methods. The specific research contents are as follows:

- (1) In order to solve the interference caused by uneven image illumination and unbalanced data, pre-processing operations such as clipping, grayscale, color and brightness normalization, image enhancement, gamma correction and image expansion are adopted. The image expansion adopts flip, translation, rotation and so on.
 - (2) Based on the problem that existing models only fuse depth features,

a multi-category feature fusion classification model is proposed. According to the characteristics of fundus lesions, on the basis of extracting AlexNet, VGG-16 and RESNET-50 depth features, four new image omics features, including Gabor, LBP, HOG and Haralick, are added, and the four image omics features and three depth features are screened by the two-layer feature fusion network. Finally, SVM, RF, kNN and ELM are used to classify the lesions, and the accuracy reaches 88.64%, 86.63%, 86.27% and 85.61%, respectively. It effectively avoids the problem of feature loss caused by insufficient feature extraction at the bottom of deep network and less feature dimension of image omics.

(3) In order to improve the model's attention to small lesions, a classification model based on depth separable convolution and attention mechanism is proposed. The attention mechanism is based on 1×1 convolution and activation function to identify key features, and the lightweight module is used to replace the standard convolution of the model, so as to achieve rapid classification of lesions while ensuring accuracy. After verification on the data set, its classification accuracy reaches 91.5%, recall rate reaches 90.3%, parameter number is 12.24m, FLOPs is 9.74g, running time is 158.7ms, Kappa value is 0.893. In addition to solving the problem of excessive irrelevant information such as background and structure of multi-category feature network extraction, the model parameters are effectively reduced and the running time is reduced while the attention to the features of small lesions is enhanced.

Keywords diabetic retinopathy; medical image classification; feature fusion; lightweight model; attention mechanism

目 录

第	一章 绪 论1
	1.1 研究背景及意义 1
	1.2 糖尿病视网膜病变介绍 2
	1.3 国内外研究现状 5
	1.3.1 医学图像诊断的研究现状5
	1.3.2 糖尿病视网膜病变诊断的研究现状6
	1.4 研究内容与创新点 8
	1.5 章节安排 9
第	二章 论文相关技术10
	2.1 传统机器学习分类方法 10
	2.1.1 支持向量机10
	2.1.2 极限学习机11
	2.1.3 随机森林12
	2.1.4 K 近邻算法12
	2.2 深度学习方法
	2.2.1 深度学习概论13
	2.2.2 AlexNet
	2.2.3 VGGNet
	2.2.4 GoogleNet
	2.2.5 ResNet
	2.2.6 DenseNet
	2.3 本章小结 19
第	三章 基于多类别特征融合的 DR 分级模型20
	3.1 实验数据介绍及预处理 20
	3.2 影像组学特征提取 24
	3.2.1 Gabor 特征24
	3. 2. 2 LBP 特征25
	3.2.3 HOG 特征25

3.2.4 Haralick 纹理特征	25
3.3 深度特征提取	26
3.4 特征融合模型	28
3.4.1 模型结构	28
3.4.2 性能评价指标	30
3.4.3 融合实验结果对比分析	31
3.5 本章小结	33
第四章 基于深度可分离卷积和注意力机制的 DR 分级模型	34
4.1 深度可分离卷积	34
4.2注意力机制	35
4.3 ResNet 算法改进	37
4.4 实验设置	39
4.4.1 数据集准备	39
4.4.2 模型参数	39
4.4.3 实验步骤	40
4.5 性能评价指标	40
4.6 实验结果与分析	41
4.7 本章小结 4	43
第五章 总结与展望	45
5.1总结	45
5.2展望	46
参考文献	47
附 录	54
致 谢	55
攻读硕士期间取得的研究成果	56

第一章 绪 论

1.1 研究背景及意义

近年来,在经济飞速发展的同时,人们的饮食习惯也在不断地改变,由于粮食产量的增加,现代人的主要食物是含糖量较高的碳水化合物,同时人们在出行方面多采用代步工具,人们的运动量相比上个世纪要小得多。上述两个因素导致了糖尿病人数量的激增,这已经成为了一种高发病率的疾病。根据数据显示,截止到 2021 年,全球糖尿病人的数量已经增加至 5.37 亿例,而我国糖尿病人数目也高达 1.4 亿例,这一数字不容乐观,甚至还在不断上升,预计到 2025 年全球患病人数将增加至 7.83 亿例,而到 2045 年,患病人数将高达 7.002 亿[1]。

糖尿病是一种高风险疾病,会导致糖尿病性肾病、心血管疾病、视网膜病变、下肢血管病变等多种并发症,其中最严重的并发症当属 DR。DR 是造成人眼失明的三大病症之一,它是由糖尿病引起的,当患者身体中的血糖无法控制时,血液中的高浓度血糖会在血管与脏器中流动,过量的糖分流进视网膜后,极易在眼底毛细血管中形成多糖粒子,从而导致血管堵塞,血液无法流通,而长期堵塞会造成眼底局部组织缺氧,血红蛋白减少,血管破裂等问题,最终导致患者视力下降甚至失明^[2-3]。据统计,我国有 0.42 亿 DR 患者,如果这些患者无法在 DR 早期得到诊断和有效治疗,极大可能造成视力障碍。DR 造成的视觉损害是不可恢复的,应及时通过筛查发现病变,并及早就诊,以达到早期诊断的目的^[4]。

目前,DR 在临床上的诊断具有很大的局限性,主要以患者视网膜图像上的病理特征为依据,由经验丰富的医生进行病情评估并得出诊断结果^[5]。据《中国眼健康白皮书》可知,全国眼科医师总数只有 4.48 万人,根本无法满足广大 DR 患者的需求,这就意味着很多糖尿病人根本无从知晓自己的病情,从而导致病情的恶化,甚至出现失明的严重情况,仅仅依靠医生的临床经验来分析视网膜图像的 DR 诊断方法不能填补当前病患的巨大需求。因此亟需寻找一种高效率、高准确率且具有客观性的诊断方法。

随着机器学习算法在人工智能等领域的广泛应用,智能诊断技术在图像分析、图像处理等方面也取得了进步。通过智能诊断技术提前判断 DR 等级,为眼科医生的后

续诊断提供参考,既可避免病人再次做第二次检查,又可减少财力物力的浪费,同时也积极支持地方医院实行的分级诊疗政策,缓解了大城市的医疗压力。此外,智能诊断技术通过筛查和分析海量图像数据提高病变诊断的准确性,极大程度减少了对昂贵专业知识的依赖度,也减轻了眼科医生的负担,还可为患者提供及时的反馈。因此使用智能诊断技术对 DR 进行分级诊断是非常必要的,它可以为眼科医生的后期诊断提供参考性意见,可以有效提高眼科医生的诊断速度和诊断准确率。目前智能诊断技术包括很多操作步骤,其灵活性较差,无法实现端到端检测,所以分类的自动化程度较低,因此基于智能诊断技术的 DR 分类诊断模型仍有较大提升空间。

基于此,本文深入研究 DR 的等级分类算法,利用相关算法提取 DR 图像的影像组学特征和深度特征,并探索多种特征融合算法以设定特殊的网络结构模型。此外,对经典分类网络 ResNet-50 的特征提取步骤进行改进,通过引入注意力机制和轻量化模型来构造新的诊断网络,以此提高诊断准确率,减少医生重复工作,为患者的健康提供保障。

1.2糖尿病视网膜病变介绍

视网膜是一种软质的透明薄膜,存在于眼球的内壁后方,下面是大量的血管,是身体内仅存的可见血管。在临床应用中,医生根据眼底图像来判断患者身体各处的血管改变情况。此外,视网膜还含有黄斑、视盘等眼部组织结构。黄斑是视网膜中最薄,也是视觉最敏锐的区域,位于视网膜后方,因为它含有叶黄素,所以颜色深于周围区域,黄斑中心的凹陷处称作中央凹,视网膜图像中的血管表现为红色细长线条。视盘又称之为视神经盘,位于视神经的前端,颜色呈淡红色。健康的视网膜具有轻微的红晕,充盈的血管脉络,反光的黄斑中心与清晰可见的视盘,其结构如图 1.1 所示[6]。

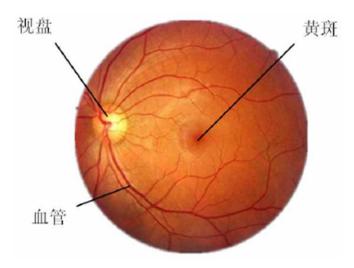


图 1.1 健康视网膜结构图

随着糖尿病患者出现糖尿病性视网膜病变的并发症,视网膜会发生某些病理变化,不同病变程度对应不同的病变特征。一般情况下,DR 的特征包含微动脉瘤(MA),硬性渗出物(EX),软性渗出物(SE),出血(HEs),增生血管等。MA 通常于病变早期出现,表现为深红色的点状物,这是因为眼底血管缺氧造成血管壁变薄引起的。HE 通常位于血管附近,表现为暗斑状,这是因为眼底血管阻塞造成血液渗出引起的。EX 与 SE 的形成表示病变已经发展为较严重的状态,血管出现长时间的血液供应不足极易造成毛细血管末端坏死,而这部分坏死血管就以絮状物或亮点状物出现在视网膜中。当眼睛长期缺氧的时候,它就会产生新的血管,以此获得更多的氧气和血液,而该过程极易造成糖尿病人失明[7-9]。视网膜的主要病变特征如图 1.2 所示。

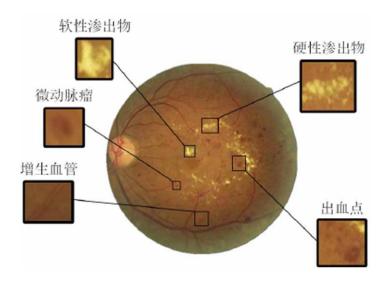


图 1.2 视网膜病变特征

国际眼科协会将 DR 的程度分为五个等级,通过标签 0、1、2、3、4 来表示。其

中,标签 0 表示无病变(No Diabetic Retinopathy),标签 1 代表轻度非增殖糖尿病视网膜病变(Mild Diabetic Retinopathy),标签 2 代表中度非增殖糖尿病视网膜病变(Moderate Diabetic Retionpathy),标签 3 代表重度非增殖糖尿病视网膜病变(Severe Diabetic Retionpathy),类别 4 代表增殖糖尿病视网膜病变(Proliferative Diabetic Retionpathy)。不同病变等级的视网膜存在不同病理特征,以此判断 DR 眼底图像的所处类别。具体分类标准如表 1.1 所示。

标签 0 代表健康的眼底视网膜图像,表现为淡淡的橘红色,视网膜动脉血管颜色鲜艳,静脉血管颜色暗红,视盘边界清晰,无病变特征。标签 1 具有的临床症状: 视网膜出现以灰色、黑色斑点或斑块样式的 MA。标签 2 具有临床症状: 除 MA 外,视网膜还有少量出血,即亮色斑状的 EX,离黄斑存在一段距离。标签 3 具有临床症状: 出现异常的微血管,四个象限均出现至少 20 个 HE,大片的明亮絮状物,已经涉及到黄斑区域,表示 SE 已经出现。标签 4 具有的临床症状: 除标签 1、2、3 的症状外,出现了占据视盘面积三分之一以上的增生血管。

在不同级别的图像中,一般不会出现单个的病灶,其反复率和重复性较高,而且相邻 DR 等级间的病灶特征也很容易发生重叠,这就给分类带来了很大的难度,但是可以根据不同级别病变的特有病灶特征对 DR 图像进行等级分类^[10-12]。表 1.1 为五个类别的视网膜图像。

类别标签 视网膜病变等级 病变特征 0 无病变特征 无病变 轻度非增殖病变 1 MA 2 中度非增殖病变 MA、EX、絮状静脉串珠 大于两个象限出现静脉串珠, 四象限内每个象限出现 3 重度非增殖病变 20 个以上 IE, 至少一个象限出现微血管异常 4 增殖病变 MA、EX、絮状静脉串珠、HE、增生新血管

表 1.1 视网膜病变等级及特征

1.3 国内外研究现状

1.3.1 医学图像诊断的研究现状

随着科技的发展,作为影响诊断准确率的重要环节,医学图像的成像方式也从二维图像发展到三维图像。所以研究出与高精度成像技术相匹配的智能诊断技术,已成为医疗界亟需解决的问题。早在二十世纪七十年代,学者利用边缘检测、区域生长等技术,建立了用于医学诊断的数学模型。直到本世纪初,机器学习算法迎来了新的研究高潮,因此相当一部分学者将方向转至与科技紧密联系的医学领域,将有监督学习运用至图像分析领域,通过机器学习利用手工设计的特征在医学图像处理中获得了更高准确率。然而随着科学技术的不断发展,通过手工设计特征的方式无法满足人们对于高效率与低成本的需求,同时该方式也无法取得解放人力的作用,所以深度学习再度兴起。深度学习可以自主迭代地提取特征,所以被广泛应用于医学图像分析,其中CNN的时间最为长远,其应用也最为广泛。各学者提出一系列具有不同结构的网络模型用以处理不同部位、不同种类的图像。

在肺结节诊断领域,深度学习取得了显著成就。Setio 等人[13]提出一种多视角 CNN 诊断模型。该模型通过两步算法进行训练学习,首先通过集成实性、亚实性与大结节三种模型来确定结节位置,然后划分分类网络框架为 9 个通道,并将输出的结节位置作为每个通道的输出,最后采用全连接层进行级联输出,有效减少了分割过程存在的假阳性问题。该二维 CNN 具有结构简单、运行速度快等优势,但是当输入数据为三维图像时,该模型极易丢失大量的空间位置信息,因此还应提出相应的三维CNN 进行研究。Xie 等人[14]将二维 FCN 与三维 CNN 相结合,建立一种用于处理假阳性问题的端到端融合诊断网络。首先通过 2D-FCN 输出与原始尺寸大小相同的肺结节区域标注图像,然后根据标注区域的位置信息,对结节的三维图像进行提取,并将其输入至 3D-CNN 中进行分类训练,从而消除假阳性。通过实验验证,该模型在 2 个假阳性时,分割召回率可以达到 98.2%,该模型还具有优良的泛化能力,可以拓展至其他医学图像分割任务中。

在乳腺癌诊断领域,神经网络也取得了广泛的应用。为实现补丁级预测,Vang 等人[15]提出一种集成多数投票、梯度提升机、Logistic 回归的分类诊断模型。该模型采用 Inception 结构取代了传统的卷积和激活函数,并引入了 1x1 卷积来降低参数过多

的问题,相较于 AlexNet,模型参数量仅为其 1/12 倍。Brancati 等人^[16]针对模型训练中的梯度消失问题,提出一种结合多种网络深度的集成诊断模型。该模型基于 ResNet-34、ResNet-50 与 ResNet-101 的不同网络层进行特征提取与融合,该方法解决了特征提取单一的问题,并获得较好的分类结果。Kohl 等人^[17]提出一种基于 DenseNet-101 的特征加强分类模型,通过稠密块增强特征传播,减少计算参数。将该模型与 VGG19、Inception-v3 进行实验对比,其模型精度高于其他两种网络 4%。为解决样本量偏少的问题,Shen 等人^[18]提出一种自动提取病理组织高级图像特征的编码器分类诊断模型,同时利用标签信息与高层次特征对模型进行端到端的微调,从而达到最优。

1.3.2 糖尿病视网膜病变诊断的研究现状

早期 DR 等级分类研究多为二分类与三分类,多数方法都是从图像中提取对比度、形状等低级视觉特征,然后再采用机器学习的方法对其进行分类,从而实现 DR 的病变等级诊断。同时,受图像数据集规模的影响,算法的可靠性在一定程度受到了限制。

对于二分类问题,学者将视网膜图像划分为两种情况,即正常与病变图像、增殖性病变与非增殖性病变图像,其分类标准相对粗略,属于前期 DR 筛查的范畴。 EITanboly 等人^[19]基于 OCT 图像所具有的视网膜厚度、反射率以及曲度三个特征,通过 SNCAE 实现病变等级的二分类,其准确率达到 96%。Abramoff 等人^[20]将 SVM 与 VGGNet、GoogLeNet 等深度模型相结合,以此对 DR 图像进行分级,其敏感性与特异性分别达到 96.8%和 87.0%。对于三分类问题,学者将视网膜图像划分为正常、非增殖性和增殖性病变,对于这一任务,Nayak 等人^[21]同样提取不同病灶特征,然后在包含 149 张眼底图像的测试集上训练神经网络,其病变分级精度达到 93%。

为进一步精确 DR 的发展阶段,ICO 和 ADA 制定了关于筛查糖尿病视网膜的最新指南,将 DR 划分为五个等级^[22]。在此基础上,国内外学者进行了大量研究,并取得了良好的分类结果^[23-24]。在传统的机器学习算法中,为避免假阳性问题的出现,Srivastava 等人^[25]在病灶特征提取时,通过不同网格尺寸的滤波器对血管和红色病灶进行区分,但网格尺寸过大极易造成计算量的增加。孟琳等人^[26]采用 Frangi 滤波器进行视网膜血管分割,并将分割后的图像采用 Otsu 进行增强以获得清晰的血管分割图。Raman 等人^[27]在对图像进行对比度增强和噪音去除后,通过提取病灶特征实现

DR 分类。

除了以上所提到的机器学习算法,深度学习的兴起也为医学图像处理的发展提供了新的思路。2017年,庞浩等人^[28]基于 ResNet 架构引入多个 Inception 模块,以此构成多分支网络模型,从而实现 DR 病变等级诊断,其 Kappa 值达到 0.85。Gargeya 等人^[29]基于 ResNet 模型的池化层提取图像 1024 维特征,然后通过训练树二分类模型实现 DR 与正常图像分类。Li 等人^[30]在结合深度学习和机器学习的基础上,采用迁移学习对预训练模型微调,以此提高 DR 诊断准确率。

2018 年,Suvajit 等人^[31]将经过中值滤波处理后的样本数据输入网络进行病变等级分类,其分类准确率达到 86.3%。Wang 等人^[32]从 Kaggle 数据集挑选了 166 张高质量视网膜图像作为训练集以此保证训练数据的质量,但是由于训练样本过少,Inception V3 模型无法提取足量的有效特征,因此分类效果不佳。为了避免以上问题,Ghosh 等人^[33]采用超过 30000 张图像的样本数据集进行网络训练,并结合去噪方法识别 MA 和出血等特征,其诊断精度达到 85%。

2019年,为了获取多层级的病灶特征信息,提高病变诊断精度,Jiang 等人^[34]采用扩张卷积替换深度卷积神经网络(DCNN)中原来的标准卷积,能够在不损失图像空间分辨率的情况下获得更大的接收场,减少特征信息的丢失。为提高病灶检测效率,Chen 等人^[35]基于 ResNet 网络搭建外接金字塔(FPN)模型。该模型将 ResNet 的卷积模块数目增加至六个,并采用自上而下的路径构造七个不同尺度大小的特征图像,以此提高特征图像分辨率,保留更丰富的病灶特征信息,从而加快模型检测进度。Jiang等人^[36]通过 Adaboost 算法集成 Inception-V3、ResNet-152、Inception-ResNet-V2 三个并行网络的输出,以获得检测结果。与其他方法不同的是,该模型引入一个图像增强模块,以此提高输入图像的多样性。该方法将病灶检测问题转换为多标签分类问题,降低了单一偏见对模型的影响,有效提高了模型诊断的鲁棒性。

2020年,张子振等人^[37]对高效网络 EfficientNet 架构进行改进,在其末端添加注意力机制,以此提高模型对小病灶的识别能力,其准确率和 Kappa 值分别达到 97.2%、0.84,解决了相邻 DR 等级难以区分的问题。Xie 等人^[38]提出在线困难样本挖掘机制(OHEM),用以改进 Faster R-CNN 模型。OHEM 将原始 ROI 网络拓展为两个,其中,原始的 ROI 网络采用前向传播算法筛选出困难样本,并将这些样本输入第二个ROI 网络,用于输出识别的病灶边框,有效克服了样本数据正负失衡的问题,从而获

得更高的诊断精度。Zhang 等人^[39]对 U-Net 检测模型进行改进,以此提高病灶检测效率。该模型通过融合 Inceptin 模块、残差模块和稠密模块来增加网络的宽度和深度,控制参数的增长,加快网络的收敛。将改进的 U-Net 模型在多个数据集上进行验证,实验结果表明上述改进可以有效降低冗余计算量,快速定位疑似病灶区域,提高诊断模型检测效率。

2021年,孙雨琛等人^[40]为提高 Inception 模型的诊断效率,对其网络结构进行改进。基于原始 Inception 模块,添加 1*1 的瓶颈层,采用 2 个 3*3 卷积替换 5*5 卷积,在增加网络深度的同时,减少模型计算量。此外,还加入了残差模块来解决模型深度增加造成的梯度消失问题。将改进的分级模型在 EyePACS 数据库上进行实验,模型召回率较 AlexNet 等经典网络有显著提升,证明了上述改进方法的有效性。Erciyas 等人^[41]提出一种级联多个深度网络的分类模型。首先将不同数据集进行融合以此扩充模型训练的样本集,其次通过 Faster RCNN 输出标记后的病灶区域,最后采用迁移学习与注意力机制实现病变等级诊断。Ai 等人^[42]提出了 DR 检测模型 DR-IIXRN。该模型首先基于 Inception-V3、InceptionResNet-V2、Xception、ResNet101 和 NASNetLarge添加注意力机制,并分别对上述模型进行微调训练,最后采用加权投票算法对不同模型算法进行集成,以此实现 DR 等级分类。Gayathri 等人^[43]基于多路径 CNN 提取图像的全局和局部特征,然后分别采用 SVM 和随机森林对图像进行分类,通过传统学习和深度学习相结合的方法实现了 DR 分级诊断。

1.4 研究内容与创新点

本论文主要对 DR 图像的等级分类方法进行研究,以提高病灶区域和图像分类精度,改善网络性能,主要在数据处理、特征融合、网络结构、轻量化模型和注意力机制方面进行研究和改进,具体内容如下:

- (1) 针对 DR 图像存在的尺寸不一、光照不均、过度曝光、对比度低、焦点模糊等问题,本文通过不同功能的预处理操作和增强算法进行尺寸标准化、提升图像亮度、增加对比度,以此突出 DR 图像的内部病灶细节,同时为解决数据集类别不均衡和数据量偏小的问题,采用随机方法对数据集进行扩充,为后续病变等级分类打好基础。
 - (2) 结合 DR 图像本身所具有的特点,设计并提取了 DR 病灶区域的影像组学

特征和深度特征,为充分发挥特征的有效性,提高模型的分类精度,提出了一个新的双层特征融合网络,有效避免影像组学特征的丢失,更好地实现图像特征表达,将融合后的特征输入 SVM、RF、kNN、ELM 等学习分类器,从而更好的完成图像分类任务。

(3) 针对 DR 图像中病灶占比小、特征提取时容易遗漏以及模型参数量过大,运行时间长的问题,本文引入了注意力机制和轻量化模块,通过改进经典图像分类算法 ResNet-50 的特征提取步骤来提高模型的分类精度和运算速度。其中注意力机制通过 1×1 和各种激活函数组成,在保证参数量的同时使模型重点关注感兴趣区域。

1.5 章节安排

本文共有五个章节,每章结构设置如下:

第一章:绪论。对 DR 等级分类研究的背景与意义进行了阐述、介绍 DR 的病理基础,梳理了国内外研究现状,主要分为医学图像诊断和 DR 图像诊断两方面,对论文的研究内容进行总结。

第二章:相关理论介绍。主要介绍本文所使用的数据集以及数据处理过程所用到的预处理方法与图像扩充方法,列举了用于评价分类模型的指标,最后对相关机器学习和深度学习的理论知识进行介绍。

第三章: 多类别特征融合。将 DR 图像的分类工作分为两个方面: 特征抽取和特征融合。在特征提取部分,根据眼底病灶特点,提取了 3 种深度特征和 4 种影像组学特征,基于 7 种不同的图像特征,提出双层特征融合网络将其进行有效融合,以获取充分发挥各类别特征优势的特征表达。

第四章: 注意力机制与轻量化模块。首先提出一种由 1×1 卷积和激活函数组成的新型注意力机制,并将该机制嵌入至 ResNet-50 模型中每一个 Stage 的后端,然后采用轻量化模块对模型的标准卷积进行代替,以此快速、准确地完成病变的等级分类任务。

第五章: 总结。对本文的工作进行了总结,并对今后的工作进行了展望。

第二章 论文相关技术

2.1 传统机器学习分类方法

基于传统机器学习方法的病变诊断流程主要通过提取 DR 图像的传统特征,并将提取的特征进行归一化处理,采用机器学习分类器以此完成病变的分级任务,下文对几个典型的机器学习分类器进行处理。

2.1.1 支持向量机

目前,SVM 是机器学习中最流行的分类算法之一,该模型通过实现结构风险最小化来提高模型泛化能力。其原理主要分为两步,首先通过构造核函数将待分类的实验数据映射至高维空间。在高维空间映射之后,为实现对样本的有效划分,首先要找到样本数据间的最大超平面,该平面即为所求的支持向量。目前,多采用以线性、多项式和径向基为核函数的 SVM。

SVM 具体分类步骤:

对于样本集 $D=\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$,其中 $y_i \in \{-1,1\}$ 表示样本类别。向量 x_i 代表映射至高维空间后的特征向量,假设存在能够将 D 线性分开的超平面(w,b),则可以得到:

$$\mathbf{w} \cdot \mathbf{x} + \mathbf{b} = 0 \tag{2.1}$$

公式 2.1 中,法向量 $w = (w_1; w_2; \dots; w_d)$ 代表平面方向,b 代表位移。

那么可以通过下列公式来表达任意点 x 至超平面的距离:

$$r = \frac{|w^T x_i + b|}{\|w\|} \tag{2.2}$$

为了对数据进行正确的分类,超平面须符合下列条件:

$$f(x) = \begin{cases} w^T x_i + b \ge +1, \ y_i < +1 \\ w^T x_i + b \ge -1, \ y_i \ge -1 \end{cases}$$
 (2.3)

SVM 的原理示意图由图 2.1 所示。其中,平面 l_1 、 l_2 的支持向量分别与超平面的距离之和称之为

$$r = \frac{2}{\|w\|} \tag{2.4}$$

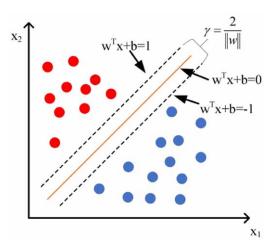


图 2.1 SVM 结构示意图

使用最大间隔划分超平面时,必须符合下列条件:

$$min\frac{2}{\|w\|}^2$$
, $y_i(w^Tx_i+b) \ge 1$, $i = 1,2$, ..., m (2.5)

在此,用拉格朗日乘数法求解 SVM 超平面的最优解,首先将式(2.5)的每条约束条件添加拉格朗日乘数 a_i ,即当 $\sum_{i=1}^m a_i y_i = 0$ 时,求出公式(2.6)的最大值。

$$\hbar = \max \sum_{i=1}^{m} a_i - \frac{1}{2} \sum_{i=1}^{m} \cdot \sum_{i=1}^{m} a_i a_j y_i y_j x_i^T x_j$$
 (2.6)

求解之后的最优函数如公式(2.7)表示。

$$f(x) = sgn\{(w \cdot x) + b\} = sgn\{\sum_{i=1}^{N} a_i y_i(x_i, x) + b\}$$
 (2.7)

SVM 是典型的二分类器,而本文所进行的 DR 分级为五分类任务,因此需要通过多个分类器实现不同等级图像数据的分类任务。

2.1.2 极限学习机

传统的前馈神经网络通常不能满足模型训练的要求,ELM 是一种能设置输入层连接权重、隐含层阈值的单层前向神经网络。该网络无需再进行二次调整,从而大大减少了模型参数,提升了运行效率^[44]。相比于 BP 网络,ELM 算法无需更迭连接权值β,可有效解决训练过程通过不断逆向传播调整隐含层阈值与权值的问题^[45]。通过上述方法确定连接权值β,不仅可以加快模型训练速度,还可优化网络的泛化性能。ELM的分类原理图如图 2.2 所示。

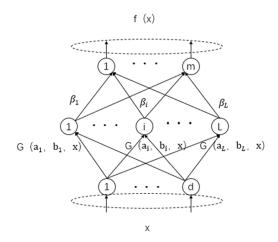


图 2.2 ELM 结构示意图

2.1.3 随机森林

随机森林^[46]依赖于树对样本进行训练,并由此完成对其所属类别的准确分类。RF的操作原理如下:以RF的所有实验决策间不存在关联为前提,在分类模型中输入样本数据,完成操作并呈现森林后,根据决策树的判断确定结果,基于判断结果进行投票,最终取得分类结果,RF原理示意图如图 2.3 所示。

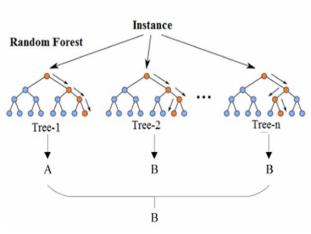
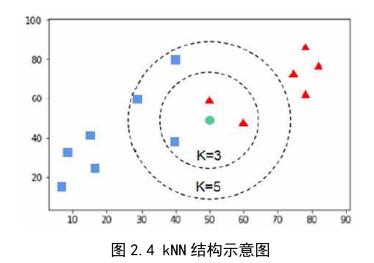


图 2.3 RF 结构示意图

2.1.4 K 近邻算法

kNN 实现样本分类的具体过程可理解为基于数据样本和标签,输入待分类的样本数据,实验过程中对待分类样本、训练组样本间距进行计算,并按照由大到小的顺序对计算间距进行排序,找到最为接近之前训练样本的前 K 个数据,将距离最短、出现次数最多的类别划分为待分类的样本集,其原理如图 2.4 所示。



2.2 深度学习方法

2.2.1 深度学习概论

传统机器学习算法虽然取得了良好的分类效果,但其提取样本特征所消耗的时间 较长,操作步骤较为繁琐,其功能具有较大狭隘性。深度学习算法在其基础上进行拓 展延伸,基于海量数据进行训练,构建包含多个隐藏层的网络结构,进而提取样本数 据深层次的抽象特征,最终取得精准的预测结果,现已成为现代机器学习的重要内容。

基于模型不同的应用领域,深度学习模型可划分为以下几个类别,如 CNN、递 归自动编码器(Recursive Auto Encoders)^[47]、受限玻尔兹曼机(Restricted Boltzmann Machines,RBM)^[48]、循环神经网络(Recurrent Neural Network, RNN)^[49]深度置信网络(Deep Belief Nets,DBN)^[50]等。本文主要以 CNN 模型为基础^[51-52],对 DR 进行了研究。 CNN 主要包括以下三部分:输入层、隐含层、输出层,CNN 基本结构示意图由图 2.5 所示。

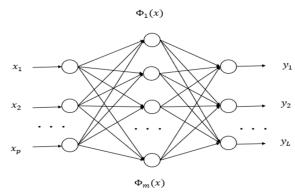


图 2.5 CNN 结构示意图

1、输入层

在输入层中,可以实现不同种类、不同维度的样本训练。如 1D CNN 的输入层可以接受 1D 和 2D 数组,1D 数组指的是时间、频域采样,2D 数组则涵盖多个通道。 2D CNN 的输入层可接受 2D 和 3D 数组,3D CNN 的输入层则可以接收 4D 数组。

2、隐含层

一般情况下,隐含层包含卷积层、池化层、全连接层,其中池化层与卷积层为 CNN 专属结构。此外,部分深层次模型还包含其他复杂的模块结构。

(1) 卷积层

通常情况下,CNN 卷积层采用两种方法来提取实验对象的特征,并且可减少网络参数冗余,降低模型参数量。第一种方法采用局部感知,一般情况下人们对事物的认知是按照从局部到整体的顺序,对于样本数据特征的训练也可采用这一原理进行分析,同时也可根据该原理分析信息相关性与像素间距的关系。通俗来讲,局部感知原理表示仅仅依靠高层整合局部信息即可实现全局信息的提取。

局部感知虽然可以优化网络训练的参数体积,但该优势仍然无法满足样本数据训练对 CNN 的功能需求。为解决上述问题,提出了权值共享策略这一种方法。在实际的模型训练中,当进行局部连接时,假定神经元数目为 m,参数数目为 n,在 m 与 n 对应时,m 与 n 实现共享,由此得到总参数量的数值为 n,这个过程就是通过卷积操作得到的。另外,在模型训练中,不能只用单一的卷积核进行特征抽取,如有必要,还需要使用多个卷积核,而从不同的卷积核提取的特征图可以用作不同通道的输入^[53]。

(2) 池化层

池化层(pooling layer)处于卷积层之间,它可以通过卷积层传送产生的图像,并通过向下取样来缩减图像的大小,从而降低过拟合现象发生。常用的池化操作有两种,即最大池化与平均池化,这两种方法都是根据相应滤波器的最大值和平均值来区分。

(3) 全连接层

全连接层(fully-connected layer)位于隐含层的末尾,它主要是对经过多次卷积与池化操作后的特征图像进行传送,并使之展开为特征向量,传递至下一层。

3、输出层

在图像分类中,输出层位于整个网络模型的末端,通常使用归一化函数或逻辑功能来计算分类结果的概率,并输出最终的图像类别。

2.2.2 AlexNet

2012 年,Krizhevsky 等人提出 AlexNet^[54],相较于 LeNet,AlexNet 具有更加泛量的深度和广度,该模型由 5 个卷积层和 3 个全连接层组成。该模型具有以下特点:

- 1、在实验训练中采用并行加速,其训练周期大幅度降低。
- 2、采用 ReLU 作为激活函数,很好地抑制了深层网络的梯度分化,同时也加快了最速下降法(SGD)的收敛速度。
- 3、为了避免过拟合现象发生,引入了 LPN 层、dropout 和数据推广技术。其结构如图 2.6 所示。

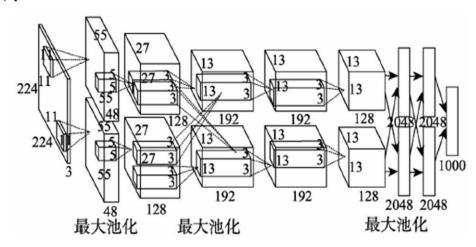
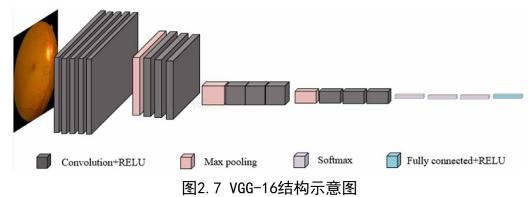


图 2.6 AlexNet 结构示意图

2. 2. 3 VGGNet

VGGNet [55]是谷歌公司和牛津大学共同开发的、具有代表性的CNN模型,该模型充分说明了增加网络层深度对于改善模型性能的重要性。基于VGGNet框架已经设计出多种CNN模型,如FCN(Fully convolutional networks)、U-Net、SegNet 等。目前,VGG-16和VGG-19是VGGNet的两种常用结构,其中VGG-16网络结构如图2.7所示。



VGG-16虽与AlexNet模型的结构基本相同,但存在以下区别:一是相比于AlexNet,其网络层数包含16-19层,以此呈现出更深的网络结构。二是VGG16将卷积层的概念扩展到由两个或三个卷积层组成的卷积块。这种扩展不但扩大了感受野范围、减小了网络的参数,而且还可以多次利用 ReLU增强模型的深度学习性能,并在实验中获得更多的线性交换。

从其他意义讨论VGG-16可以发现,其良好的泛化能力可以便捷转移至其他图像识别任务中,而且还可以通过下载VGGNet中已经分类完成的实验参数来进行权重的初始化操作。

2.2.4 GoogleNet

GoogLeNet^[56]作为2014 ILSVRC竞赛的年度冠军,将TOP-5的错误率降低到67%。从网络结构出发,其网络结构与VGGNet大体相似,同时,为了有效减少网络输出,在主要的卷积环节中,使用5个模块且模块间采用步长为2的3×3卷积进行连接。不同于VGGNet在不同卷积层增加loss函数和Inception结构,即使GoogleNet修改通道难度较大且模型计算复杂度高,但该模型在较大幅度拓展网络结构和减少参数量方面具有明显优势。

图2.8为Inception模块结构示意图,由于Inception模块中不存在全连接层,通过4条并行路径可以有效提升模型运算效率,减少训练时间。为降低模型训练的复杂度,将GoogLeNet的参数量减少到AlexNet的1/12,并在后面3条路径中添加1x1卷积层,以此简化VGGNet的结构。实验过程中,为提取图像多尺度空间信息并便于对齐,在前3条路径分别采用卷积大小为1×1、3×3、5×5的卷积层;为实现输入输出图像的尺寸一致性,4条并行路径均采用上述填充方法,在通道维度上衔接所有路径的输出,最终输入下一层级。

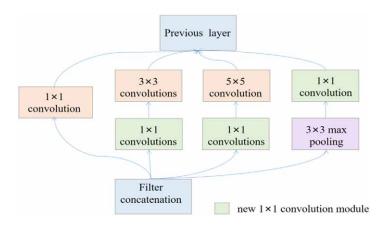


图2.8 GoogleNet结构示意图

2. 2. 5 ResNet

在不存在过拟合现象时,网络的分类精度通常会随着网络的加深而提高。虽然准确率得到显著提升,但无法避免的是网络加深过程所带来的梯度消失问题。这一现象出现的原因是由于新增加的网络层作为参数在更新信号并伴随梯度传播时,传播的方向由后向前进行,随着网络层数的更迭,靠前部分的网络层梯度减小,从而影响前面部分网络层的学习,导致学习停滞。

ResNet^[57]的提出是 CNN 图像史上的一个里程碑,它通过残差模块有效解决了深度网络的退化问题,因此具有非常深的网络结构而不存在梯度消失或梯度爆炸的问题,降低了网络深度对分类性能的限制度。根据计算表明,ResNet-50 的层数虽然远大于VGG-16 的层数,但其参数量却远远小于 VGG-16,因此,与其它深度神经网络相比,使用残差模块可以在加深网络的同时控制参数量的增长。此外,ResNet 采用 Batch normalization 代替 drop out 操作,用以加速模型的训练速度,上述特点在提高模型训练准确度的同时减少网络参数量和计算量,因此该网络被广泛用于各图像处理领域。

如图 2.9 所示的 ResNet 中的残差模块,图中曲线表示为 short cut,该曲线表示在主分支的特征矩阵经过多次卷积池化操作后与输入特征矩阵相加,ResNet 分为浅层和深层两种类型,图 2.9 (a) 为 ResNet-34 等浅层残差网络的核心模块,图 2.9 (b) 为 ResNet-50/101/152 等深层残差网络的核心模块。右图称为 bottleneck design,其卷积核的设计发挥了降维的作用,并且采用1×1+3×3+1×1的卷积形式替换两个3×3卷积层,在保证精度的同时降低计算量。

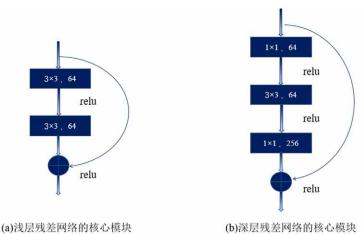


图 2.9 残差模块结构示意图

ResNet 有多种不同网络层数的结构,具体结构如表 2.2 所示。经过试验验证, ResNet-50 在眼底图像进行等级分类时既能保证分级准确性又能控制参数量爆发性的 增加,所以本文选用 ResNet-50 作为深度学习模型框架。

layer output 18-layer 34-layer 50-layer 101-layer size name 112×112 $7 \times 7,64$, stride 2 conv1 3×3 max pool, stride 2 $\begin{bmatrix} 1 \times 1,64 \\ 3 \times 3,64 \end{bmatrix} \times 3$ 56×56 conv2 x $\begin{bmatrix} 1 \times 1,64 \\ 3 \times 3,64 \end{bmatrix} \times 3$ $\begin{bmatrix} 3 \times 3,64 \\ 3 \times 3,64 \end{bmatrix} \times 2 \qquad \begin{bmatrix} 3 \times 3,64 \\ 3 \times 3,64 \end{bmatrix} \times 3$ $\begin{bmatrix} 1 \times 1,128 \\ 3 \times 3,128 \end{bmatrix} \times 4$ $\begin{bmatrix} 1 \times 1,128 \\ 3 \times 3,128 \end{bmatrix} \times 4$ $\begin{bmatrix} 3 \times 3,128 \\ 3 \times 3,128 \end{bmatrix} \times 2 \quad \begin{bmatrix} 3 \times 3,128 \\ 3 \times 3,128 \end{bmatrix} \times 4$ 28×28 conv3 x $\begin{bmatrix} 1 \times 1,256 \\ 3 \times 3,256 \\ 1 \times 1,1024 \end{bmatrix} \times 6$ $1 \times 1,256$ $\begin{bmatrix} 3 \times 3,256 \\ 3 \times 3,256 \end{bmatrix} \times 2 \quad \begin{bmatrix} 3 \times 3,256 \\ 3 \times 3,256 \end{bmatrix} \times 6$ $3 \times 3,256 \times 23$ 14×14 conv4 x $[1 \times 1,1024]$ $1 \times 1,512^{-1}$ $1 \times 1,512$ $\begin{bmatrix} 3 \times 3,512 \\ 3 \times 3,512 \end{bmatrix} \times 2 \quad \begin{bmatrix} 3 \times 3,512 \\ 3 \times 3,512 \end{bmatrix} \times 3$ $\begin{bmatrix} 3 \times 3,512 \\ 1 \times 1,2048 \end{bmatrix} \times 3$ 7×7 $3 \times 3,512$ conv5 x 1×1 average pool, 1000-d fc, softmax **FLOPs** 1.8×10^{9} 3.6×10^{9} 3.8×10^{9} 7.6×10^{9}

表 2.2 ResNet 网络结构示意图

2. 2. 6 DenseNet

2017 年,Huang 等人提出 DenseNet^[58],该模型通过密集连接将前一层的特征图

像输出当作其后所有层的输入,并在最后一层拼接特征图像。DenseNet 所特有的结构在一定程度可以有效解决梯度消失问题,所有卷积层实现从损失函数到梯度的径直访问,加强特征传播性,提倡特征复用,大幅降低参数量,提高参数利用率。这种密集连接相当于将各层与输入和损失函数直接相连,从而有效减少了梯度消失问题的出现。但是只有一个 dense block 中存在 dense connectivity,不同 block 间不存在 dense connectivity,其网络结构如图 2.10 所示。

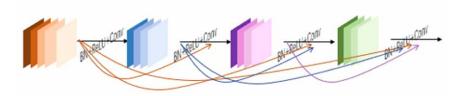


图2.10 DenseNet结构示意图

2.3 本章小结

本章主要介绍了 DR 病变等级辅助诊断的相关理论,共包括四部分,首先对实验数据及预处理过程进行介绍,其次列举了用于评估模型的指标,然后介绍了机器学习相关的理论,最后重点介绍深度学习中 CNN 模型的基本结构。

第三章 基于多类别特征融合的 DR 分级模型

本文所研究的 DR 图像病变等级分类,属于细粒度分类范畴,丰富、广泛的病灶特征信息对于模型的分类精度至关重要,因此本文将多类别特征融合策略引入 DR 图像的病变等级分类任务中,模型结构如图 3.1 所示。

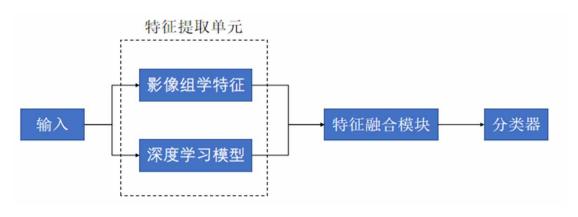


图 3.1 多类别特征融合模型框架图

3.1 实验数据介绍及预处理

本文采用公开的数据集 diabetic retinopathy, 该数据集于 2019 年发布, 其下载链接是 https://www.kaggle.com/himanshuagarwal1998/diabetic-retinopathy。数据集大小为8.09G, 视网膜图像均为 png 格式, 共包含 3662 张, 全部按照 DR 的严重程度将图像标注为0-4级。该数据集来源于不同的医院, 受采集设备的影响, 原始图片大小不等, 其数值主要分布在819*614、1050*1050、2048*1536、2088*1958等范围内, 除此之外受摄影环境、成像装备以及摄影师拍摄习惯的影响, 图片的明暗、质量等均有差别。为了提高图像亮度、增加眼底结构与背景的对比度, 从而更好地显示病灶信息, 对数据集进行预处理。预处理的操作如下:

1、图像裁剪

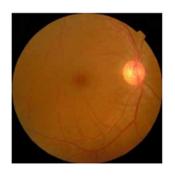
视网膜图像通常为不同大小的矩形,但有效的图像区域仅位于圆形眼球内部,因此 DR 图像的周围存在大量黑色边框,为了消除背景信息干扰,减少无关信息,缩减图片大小,以此方便后续学习,以各眼底图像中眼球直径作为边长,眼球中心定位图片中心来裁剪图片。

2、图像灰度化

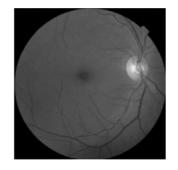
我们在手机、电脑、电视等电子设备上看到的彩色图像均以 RGB 模型形式来显示,即利用红、绿、蓝三色光相叠加而产生混叠效应,然后产生最柔和的视觉效果。但在医学图像的处理中,为了加快运算速度往往需要对图像进行灰度化处理。灰度化就是在彩色图像中使 R、G、B 三分量相等,即 R=G=B,这样就使得原 R、G、B 分别具有不同彩色图像,仅通过灰度值变化就可表示。灰度化图像方法有以下四种:

- (1)分量法:将彩色图像中 R、G、B 三分量的亮度取作三个图像的灰度值,处理成为灰度图像。
- (2) 最大值法: 灰度图的灰度值采用彩色图像中三分量的亮度最大值来表示。
- (3) 平均值法: 灰度图的灰度值采用彩色图像中 R、G、B 的亮度平均值来表示。
- (4) 加权平均法: 依据权重和其他指标,对R、G、B进行加权平均。

但是由于视网膜图像的特殊性,经过实验验证,上述灰度化方法获取的图像难以达到理想效果,同时发现眼底图像去绿色通道时其眼底图像效果更好,也更利于后期的处理,所以通过绿色通道来提取视网膜图像,如图 3.2 所示。



(a) 原始图像

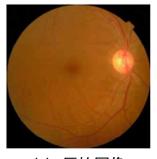


(b) 处理后的图像

图 3.2 图像灰度化处理

3、图像颜色与亮度归一化

在深度网络模型中,图像的局部特征对分类精度的影响很大,尤其是图像亮度与色彩状况。为减小类内方差带来的影响,本文利用高斯滤波算法对 DR 图像进行处理,以获得图像的局部均值,然后将原始图像和局部平均进行相减,从而获得最后的归一化图像,其处理结果见图 3.3。



(a) 原始图像



(b) 处理后的图像

图 3.3 图像归一化处理

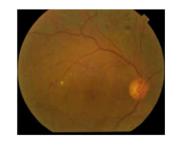
其运算过程如公式 3.1 所示。

$$I = \alpha I_0 + \beta G(\theta)^{\circ} I_0 + \gamma \tag{3.1}$$

 I_0 代表原始图像,P 代表归一化后的图像, $G(\theta)$ 代表高斯和, θ 表示滤波尺寸,。 代表滤波运算, α 、 β 、 γ 代表图像融合的权重,本文中将 θ 、 α 、 β 、 γ 设置为 30、4.14、-4.14、128.3。

4、图像增强

为增加图像中细节信息的对比度,提高后期分类准确性,本实验采取图像增强处理。传统的增强方法分为直接增强与间接增强两种,而间接增强又可分为两种方式,即直方图均衡化^[59]与直方图拉伸,其中直方图均衡化通过对图像灰度值的调整来实现对比度增强^[60]。自适应直方图均衡是一种直方图均衡方法,它是通过计算局部直方图,再重新排列原始图像,来强化局部信息处理,提高局部信息的对比性,有效地克服了直方图均衡算法由于计算整体直方图而造成的细节损失。但是这导致异常数据中的细节信息被放大^[61]。因此,基于上述方法提出了限制对比度直方图均衡算法(Contrast Limited Adaptive Histogram Equalization,CLAHE),该方法有效地克服了异常数据信息被放大的缺点。因此,本文采用 CLAHE 对图像进行强化,其效果见图3.4。



(a) 原始图像



(b) 处理后的图像

图 3.4 图像增强

具体步骤如下所示:

- (1) 将视网膜图像分为 M*N个非交叠的区块。
- (2)计算分割后图像块的灰度直方图,将每个图像块的像素平均至各个灰度级,灰度级平均像素数 $\overline{N_a}$ 可由公式3.2求出

$$\overline{N_a} = \frac{N_{xp} - N_{yp}}{N_a} \tag{3.2}$$

其中 N_g 为图像块中具体灰度级的数量, N_{xp} 为x轴方向图像块中的像素数目, N_{yp} 为y轴方向图像块中的像素数。

(3) 截掉灰度直方图中大于 N_l 的数值, N_l 由公式3.3求出。

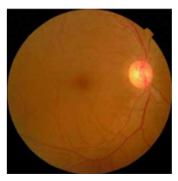
$$N_l = S\overline{N_a} \tag{3.3}$$

其中S为截取系数。

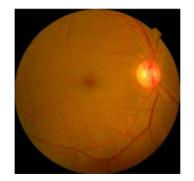
(4) 直方图均衡化。将所有图像块的直方图进行均衡化,利用变换函数计算出变换后的灰度值。

5、图像伽马校正

伽马校正^[62]通过作用于伽马曲线对原始 DR 图像进行色调处理,检测出图像中深色与浅色信号,并对其比例进行调衡,然后基于原始图像对其对比度进分行加强。其中伽马曲线由函数 $Y=(X+e)^{\gamma}$ 。其中 Y、X、e、 γ 分别代表亮度、输出电压、补偿系数、伽马值。伽马曲线的改变伴随伽马值大小的变化。本文设置的伽马值为 0.45,其结果如图 3.5 所示。



(a) 原始图像



(b) 处理后的图像

图 3.5 伽马校正

6、图像扩充

本数据集中,不同病变等级的视网膜图像数目差异较大,其中,0级样本量为1804,接近总样本量的50%,3级样本量所占比例最小,约为8.1%。由于图像数据不平衡极

易导致模型训练出现偏差,为保证良好的学习效果,需要扩充数据集,使其比例接近1:1:1:1,选择随机的方法对视网膜图像进行扩充,使得 0~4 级样本的数量分别为 6017:5986:6325:6480:6048,图像总量达到 30856,本文采用以下方法进行扩充。

- (1) 将图像在[-180°,180°]范围内随机旋转。
- (2) 将图像沿水平和竖直方向翻转。
- (3) 将图像沿水平和竖直方向随机移动。
- (4) 将图像按照[0.8.1.2]的随机因子进行缩放或扩大。

本文实验中,在经扩展后的数据集上按照训练集:测试集=7:3 的比例对数据集进行划分,数据集的实际情况数量如表 3.1 所示。

	训练集	测试集	合计
0 级	4217	1804	6017
1级	4190	1796	5986
2级	4427	1898	6325
3级	4536	1944	6480
4级	4234	1814	6048

表 3.1 数据集数量分布

3.2 影像组学特征提取

DR 图像因其内部构造复杂,各等级图像间差异小,病灶占比低,导致单纯意义上的传统特征难以适用于 DR 等级分类这一复杂任务,因此需要选择合适的特征表达实现 DR 图像分级。本节选用四种常见的影像组学特征作为 DR 图像的特征表示。目前,已被广泛应用的影像组学特征有 Gabor 特征^[63]、局部二值模式特征(Local Binary Patterns,LBP) ^[64]、方向梯度直方图特征(Histogram of Oriented Gradient,HOG) ^[65]以及 Haralick 纹理特征^[66]等。

3. 2. 1 Gabor 特征

Gabor 特征表示频域中图像方向、尺度上的相关特征。其特征提取步骤为: 首先

设计四个不同尺度(v=0,1…,3)、六个不同方向(θ =0, π /6, π /3, π /2,2 π /3,5 π /6)的 Gabor 滤波器组,其次将待处理图像 I(x, y),x, y \in Ω (Ω 代表该图像的像素点集合)分别与 Gabor 滤波器 G(x, y) 的实部和虚部进行卷积和运算,以此生成特征图像 R(x, y),最后依次计算特征图像 R(x, y) 中所有像素灰度值的均值与方差,该均值与方差即 为所求取的 Gabor 特征。

3. 2. 2 LBP 特征

LBP 表示图像中的局部纹理特征。该特征的提取步骤为:对待处理图像进行随机划分(假设大小为16×16),依次对比所有图像块中每一个像素的具体灰度值和其相邻像素(8个)数值的大小,按顺序将每个方块中各个像素点的灰度值与其邻近像素(8个)的灰度值进行比较,如果中间位置像素值较小,则将邻域内的像素数值转化为1,反之转换为0.8,各位置的像素值替换后,在该区域将会生成8位二进制数字,然后通过十进制变换得到该位置像素点的LBP特征值。最后依据各数字出现的频率绘制所有图像块的直方图,并将直方图进行归一化以生成特征向量,由此得到图像的LBP特征值。

3.2.3 HOG 特征

HOG 特征具有多种良好的特性,如:光学不变性、几何不变性,该特征的提取步骤可总结为计算输入图像中局部区域的方向梯度直方图,其主要提取步骤可总结如下:利用特定大小、形状的窗口划分原始图像,以此生成多个小连通区域,称为 cell,然后计算出所有 cell 的梯度直方图。其具体计算方法是将 DR 图像分为多个 cell,将 0~360°的梯度方向平均分成 9 个方向,4 个相邻 cell 构成一个 2×2 分区,并对各分区的梯度直方图进行统计,然后串联各分区的特征向量以此获取多维特征,上述特征向量即为所求取的 HOG 特征。

3. 2. 4 Haralick 纹理特征

Haralick 特征代表图像中像素的排列规律与局部模式,通过借助像素方向与距离间的空间位置关系来构建 GLCM。但 GLCM 不能直接用以表示特征,需要建立多个特征表达的统计量来表达图像的特征信息。因此,Haralick 纹理特征的提取步骤为:

首先构建原始图像对应的 GLCM,其次计算通过特征统计量构建的特征向量。其中 GLCM 用于计算像素间的相对距离(d=1)与不同方向($\theta=0,\pi/4,\pi/2,3\pi/4$),并对图 像灰度级进行量化处理,最后得出上述方向下特征向量的均值和标准差,由此取得 Haralick 纹理特征表示。表 3.2 列举了上述影像组学特征向量的维度。

影像组学特征 特征向量维度

Gabor 特征 48

LBP 特征 59

HOG 特征 324

Haralick 纹理特征 26

表 3.2 影像组学特征及其向量维度

采用 Min-max normalization 算法^[67]对提取出来的上述四种特征数据进行归一化,用以模型后续学习与训练。最后利用 SVM、RF、kNN、ELM 等多个学习分类器在训练集上建立不同的 DR 分级模型,并在测试集中验证不同分类器的分级性能。

3.3 深度特征提取

近年来,随着数据量爆炸式的增长,深度学习作为新兴的机器学习算法实现了飞速发展。通过深度学习方法实现图像识别,其优势在于图像的识别精度较高,适应性强更易于转换,可利用数据进行有效的缩放,方便实际部署和应用。现有深度学习网络的浅层次卷积主要通过对原始图像进行纹理、颜色以及边缘等特征的提取,然后降采样以获得低阶特征,深层次卷积则用来获取图像的语义信息和高分辨率特征图。

实验过程中,利用梯度下降与交叉熵损失函数对模型进行微调。本实验选择 AlexNet 中 FC7 层输出作为 DR 图像的深度特征表达,深度特征提取的具体过程如图 3.6 所示。

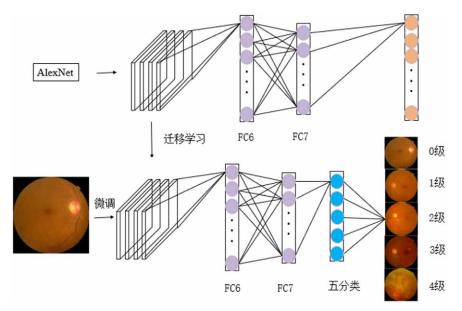


图3.6 深度特征提取过程示意图

DR 图像深度特征的提取步骤如下:

输入: 训练集的DR图像 $\{x_1, x_2 \cdots, x_m\}$, 病变的具体分级由表示 $y^{(i)} \in \{0,1,2\}$ 。

1、DR图像的深度特征提取

将DR图像数据输入AlexNet进行卷积操作,逐层提取图像信息,将其FC7层的输出作为深度特征表达,记为 $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \cdots, (x^{(m)}, y^{(m)})\}$ 。

2、计算代价函数 $I(\theta)$

对于输出图像特征 $x^{(i)}$,利用公式3.4计算其代价函数:

$$J(\theta) = \frac{-1}{m} \sum_{i=1}^{m} y^{(i)} lg\left(h_{\theta}(x^{(i)})\right) + \left(1 - y^{(i)}\right) lg\left(1 - h_{\theta}(x^{(i)})\right)$$
(3.4)

在公式3.1中, $y^{(i)}$ 表示图像类别, $h_{\theta}(x^{(i)})$ 表示Softmax层输出。

3、更新深度网络的参数 θ

利用梯度下降算法优化代价函数 $J(\theta)$,根据梯度方向优化深度学习网络中的模型参数 θ 。最后使用训练好的AlexNet提取DR图像的深度特征。

输出: DR图像的深度特征 $x_j = [g_1, g_2 \cdots, g_{4096}]$

此外,本文也提取了VGG-16、ResNet-50模型的深度特征。3种深度特征向量的具体维度如表3.3所示。

深度特征	向量维度
AlexNet	4096
VGG-16	512

表3.3 深度特征及向量维度

2048

3.4 特征融合模型

ResNet-50

3.4.1 模型结构

由于不同深度学习网络在进行特征提取时,其图像关注点不一致,因此本文尝试串接不同深度学习网络提取的特征向量,以此融合DR图像中多个兴趣位置的特征信息^[68]。本方案借鉴DenseNet模型的思想,通过引入网络结构多尺度特性拼接模型所提取的不同尺度特征向量。特征向量串接结构如图3.7所示。

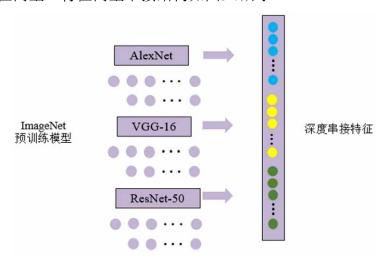


图3.7 串接融合示意图

在不进行模型参数更新的情况下,从AlexNet、VGGNet以及ResNet中抽取模型的特征向量。采用串行融合方法,将三个不同的深度特征合成一个具有6656维的深度串接特征。同时,为充分发挥四种传统特征在图像分级任务中的作用,同样采用特征串接方式将不同特征进行合并,形成维数大小为457维的影像组学特征和7113维的影像组学+深度串接特征。

假如仅对上述过程提取的特征向量进行简单的串联融合,通常会导致特征向量维度大幅度增加,由于影像组学特征维度相较于深度特征较小,在模型分类阶段容易遗

漏病灶特征,从而降低DR分级精度。为了将提取的多尺度特征向量进行有效融合,且使不同类型特征的融合结果在病变分级中发挥最大作用。本章提出了一个双层特征融合网络,并通过最小化损失函数来训练网络模型。图3.8为本章设计的特征融合网络模型。

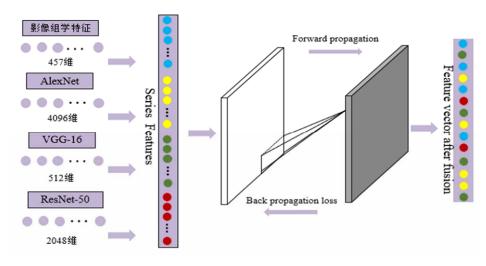


图3.8 融合网络示意图

在特征融合网络中,采用四种不同类型的特征向量作为输入,将串联融合后形成的特征向量记为f。该特征向量通过学习网络隐含层的权值以获取隐含层的节点数,获取的节点数即为融合向量的最终维数。特征融合算法的具体步骤如下所示:

将4种不同类型的特征进行串接以得到融合特征的串接向量

1、进行前向传播;

将步骤1获得的串接向量输入特征融合模型进行前向传播,目标函数如公式3.5所示。

$$o(\theta, \gamma) = softmax(\theta_2 \cdot tanh(\theta_2 \cdot f + \gamma_1) + \gamma_2)$$
(3.5)

2、进行反向传播;

使用交叉熵损失函数进行反向传播,通过梯度下降算法优化代价函数,并根据梯度变化方向对参数θ进行调整。

3、获得DR图像的融合特征:

利用特征向量融合模型,可以获得代表DR图像的融合特征向量。

输出: DR图像的融合特征 $x = [u_1, u_2 \cdots, u_n]$

在特征融合模型中,选取了tanh和Softmax分别作为隐含层和输出层,通过最小化 损失函数来反向传播隐藏层的权值,并利用dropout方法来避免过拟合,以此获得融合 后的图像特征。然后将图像特征输入分类器中,以此获得最终分类结果,具体分类模型如图3.9所示。

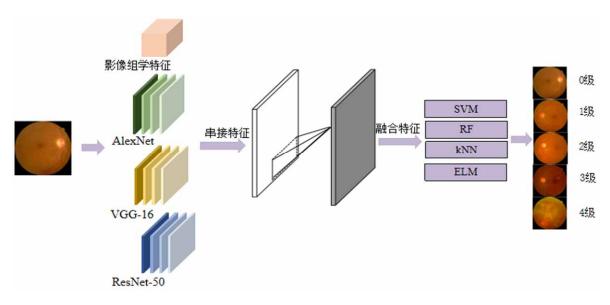


图3.9 分类模型示意图

3.4.2 性能评价指标

为更好地验证本研究所提出的 DR 辅助诊断算法的有效性与可行性,本文采用准确率、查准率和查全率对诊断算法进行评估,一般情况下,样本的实际分类和模型的预测分类,有下列四种可能的组合,如表 3.4 所示。

表 3.4 模型预测类别

组合	含义	
TP(True Positive)	预测和真实类别均为正样本	
FP(False Positive)	预测为正样本, 真实为负样本	
FN(False Negative)	预测为负样本,真实为正样本	
TN(True Negative)	预测和真实类别均为负样本	

通过 TP、TN、FP、FN 表示的具体指标如下:

准确率:
$$ACC = \frac{TP + TN}{TP + FP + FN + TN}$$
 (3.6)

查准率:
$$P = \frac{TN}{TN + FP}$$
 (3.7)

查全率:
$$R = \frac{TP}{TP + FN}$$
 (3.8)

3.4.3 融合实验结果对比分析

利用上述特征融合方法获取DR图像的特征信息,分别用于后续的训练与测试。最后使用SVM、RF、KNN、ELM等多个学习分类器建立不同的DR等级分类模型,并在测试集进行分类性能评估。分类性能结果如表3.5、3.6所示。其中,表3.5列举了DR图像中不同影像组学特征对应的分级实验结果,表3.6列举了DR图像中各种融合特征对应的分级实验结果。

特征 评价指标(%) SVM RF KNN ELM Р 66.79 59.52 63.63 63.93 70.70 69.15 71.70 67.79 Gabor R ACC 71.55 67.01 70.64 70.91 Р 68.36 61.86 70.77 59.76 LBP 70.43 66.33 R 71.16 73.80 ACC 73.60 70.49 71.89 69.36 Р 66.57 53.12 58.60 56.44 HOG 69.88 65.61 64.06 61.78 R ACC 71.10 69.69 63.15 65.37 Р 54.85 64.92 58.74 60.58 Haralick R 66.42 67.06 65.70 67.06 ACC 69.42 66.52 64.88 68.76

表 3.5 各影像组学特征分级实验结果

由表 3.5 的实验数据可知: 在相同分类器的条件下,四种影像组学特征中 LBP 特征和 Gabor 特征相较于另外两种表现较好,其中 LBP 特征的四个评价指标最高,由于该特征侧重于提取图像的局部纹理信息,因此针对图像细节的分级效果要优于其他 3种特征。经过横向对比,在同一特征表达下,SVM 的分级效果最好,其中 HOG 特征在SVM 分类器下的分类效果优势最明显,平均分类准确率高于其他分类器 5%左右。

表3 6	特征融合模型分类结果
12CO. O	

特征	评价指标(%)	SVM	RF	KNN	ELM
影像组学特征	Р	74.95	70.08	71.62	69. 02
	R	74.69	74. 16	75.67	73. 07
	ACC	75. 48	75. 13	74. 70	74. 93
深度串接特征	Р	78. 12	77.77	77. 90	75. 51
	R	79.89	79.67	79. 98	78. 62
	ACC	81.78	80.12	79. 16	78. 52
影像组学+	Р	81.45	79.74	80.07	78. 28
深度串接特征	R	83.62	82. 16	81.61	82. 44
体及中女村仙	ACC	84. 53	83. 18	83. 32	82. 37
基于特征融合的诊断方法	P	85. 42	83.85	84. 57	81. 43
	R	86.88	85. 53	85.02	84. 44
	ACC	88.64	86.63	86. 27	85. 61

由表3.6中实验结果对比可知,在三种串接特征中,影像组学+深度特征融合方式 所提取的图像特征对模型的分类效果普遍优于另外两种,在深度串接特征基础上,添加图像的影像组学特征有利于模型分类精度的大幅度提升,弥补了深度网络浅层次特征提取的不足,丰富了特征的种类,使之更符合DR特点。但是简单的串接融合模式会造成影像组学特征的遗漏,以及特征冗余,因此采用有效的特征融合网络来获取病灶特征,最终取得较好的分类结果,其分类准确率高于简单串接特征4%左右。而对于同一特征,SVM的分类性能最好。

表3.7为不同分类算法在数据集上获取的分类准确率。基于多类别特征融合的诊断方法准确率最高。证明了特征融合网络模型设计合理,优于之前的研究结果。

表3.7 不同算法实验结果对比

方法	准确率(%)
AlexNet	62
DCNN+SVM ^[69]	86. 17
WSDIM ^[70]	80. 04
基于多类别特征融合的诊断方法	88. 64

3.5 本章小结

本章针对 DR 图像的等级分类问题进行了实验,为了提高模型的特征提取能力,设计了特征融合网络来融合多种特征,主要包括影像组学特征和深度特征,其中影像组学特征分为 Gabor 特征、LBP 特征、HOG 特征以及 Haralick 纹理特征,并通过SVM、RF、KNN、ELM 进行病变等级分类。最终结果表明了多类别特征融合模型的有效性。

第四章 基于深度可分离卷积和注意力机制的 DR 分级模型

前一章的图像分类过程中,虽然融合了不同模型的特点,提高了特征多样性,但由于病灶在图像中占比小,模型训练过程中提取了背景、视盘等无关特征,因此提出一种全新的注意力机制,对关键特征予以更多的注意,以此提高小目标病灶特征的提取能力。由于注意力机制的引入会造成模型的加深,为了防止梯度消失问题,选用ResNet-50 作为基础网络结构。此外,在增加网络深度,提高网络复杂度的同时,采用深度可分离卷积代替标准卷积,以降低网络参数量,提高运行速度。最终,本章提出一种基于深度可分离卷积和注意力机制的等级分类模型,经过实验验证改进的模型更有利于病变等级分类。

4.1 深度可分离卷积

目前,多数网络通过加深网络层数,引入卷积模块等方法提高模型识别的准确率。该方法使模型的性能有了大幅提升,但相较于之前的网络结构更为复杂,大幅提升了内存和所需计算量。为解决上述问题,多数学者在标准卷积功能模块的基础上,进行轻量化处理。其中,深度可分离卷积就是通过调整卷积的通道和空间维度,降低卷积冗余、提高计算速度,以此实现了轻量化目的[71-73]。

深度可分离卷积由逐通道卷积(Depthwise separable convolution, DW)和逐点卷积(Pointwise convolution, PW)两部分组成,其结构分别如图 4.1、4.2 所示。深度可分离卷积把标准卷积分解为 2D 空间的 DW 与修改通道数的 PW,将输入的 3D 特征变为独立的 2D 特征与通道维度。卷积过程中为各通道分配一个卷积核,并将输出的特征图作为输入的通道数,利用一个卷积滤波器进行实施;然后,采用 contact 方式串联特征图,以此进行 PW,通过 PW 计算输入通道的线性组合以构建新特征。卷积部分的参数个数由公式 4.1 表示。

$$N_{-}dept\hbar wise = M \times M \times C$$
 (4.1)

其中,M×M为卷积核大小,C为输入通道数。PW的卷积核大小为1×1×C,其运算与常规卷积类似,将上一层的输出特征图在深度方向上进行加权组合,形成与卷积核个数相同的特征图。卷积部分的参数个数由公式4.2表示。

$$N_{pointwise} = 1 \times 1 \times C \times N$$
 (4.2)

其中,C表示输入层的通道数,N表示卷积核的个数。

与标准3×3卷积相比,3×3的深度可分离卷积在准确度略有下降的基础上,将 计算量减少了8到9倍,达到了轻量化目的。

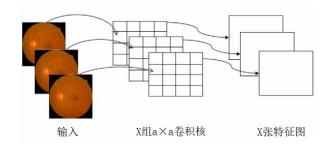


图 4.1 逐通道卷积

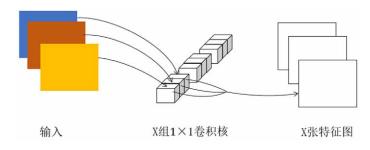


图 4.2 逐点卷积

4.2 注意力机制

近年来,注意力机制作为一种重要的深度学习手段,已经在模式识别^[74]、语音处理^[75]等领域得到广泛应用。添加注意力机制,能有效提高模型对于图像的特征提取能力。常用的注意力机制有通道注意力机制、空间注意力机制和混合注意力机制。通道注意力机制,如 Squeeze and Excitation(SE)block^[76],需要在不同的通道中引入注意力机制,以此进行特征图像压缩、加权。空间注意力机制,如 Non-local Neural Networks^[77],则是在空间域上加入注意力机制。混合注意力机制,如 Convolutional Block Attention Module(CBAM)^[78],是在 CNN 中加入通道注意力机制与空间注意力机制,通过池化操作计算通道与空间权重,并将两种权重结合起来。Google DeepMind 团队在 2014 年,将注意力机制引入循环卷积神经网络,实现多物体识别,在 MNIS 分类识别中,将错误率减少了 4%,这充分证明了在卷积网络中引入注意力机制的有效性。

在观察一幅图像时,人的眼睛会迅速扫描整幅图像来获得目标区域,然后更关注

某一区域,以此获得有用的信息,这是人脑中所特有的注意力机制。根据这一原理,深度学习算法中的注意力机制应运而生。目前,注意力机制在计算机视觉领域产生了重大影响,用以指导神经网络关注图像中最具辨识度的信息。注意力机制 $[^{79}]$ 是一种通过计算概率分布来强调某个输入对输出影响的组合函数。基于 Encoder-Decode 框架的注意力机制结构图如 4.3 所示。该模型将输入序列 $\mathbf{X}=(x_1, x_2, \cdots, x_n)$,通过一系列非线性变化映射至输出序列 $\mathbf{Y}=(y_1, y_2, \cdots, y_m)$ 。

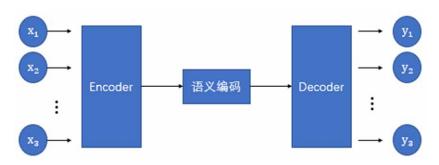


图 4.3 基于 Encoder-Decode 框架的注意力机制结构图

该模型将输入序列 $X=(x_1, x_2, ..., x_n)$,通过一系列非线性变化映射至输出序列 $Y=(y_1, y_2, ..., y_m)$ 。但 Encoder-Decode 框架无法对输入的序列进行区分,为解决上述问题,Bahdanau 等人[80]提出注意力机制,其结构如图 4.4 所示。

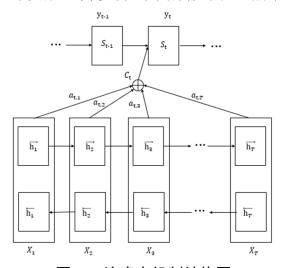


图 4.4 注意力机制结构图

在图中, y_i 表示目标词, C_t 表示上下文向量, S_t 表示解码器在 t 时刻的隐状态,且第 t 时刻隐状态由以下公式 4.3 来表示。

$$s_t = f(s_{t-1}, y_{t-1}, C_t)$$
 (4.3)

其中,上下文向量 C_t 依赖输入序列 X 的隐藏层,加权后由下列公式 4.4 所示。

$$C_t = \sum_{j=1}^{T} \alpha_{t, j} \hbar_j \tag{4.4}$$

在上述公式中,T 表示输入序列 X 的长度。 \hbar_j 是编码器第 j 个词的隐向量,包括了输入序列 X 的全部信息,且重点关注第 j 个词周围的信息。 $\alpha_{t,j}$ 是编码器部分第 j 个词对解码器第 t 个词的注意力分配系数,概率和为 1。 $\alpha_{t,j}$ 可由公式 4.5 来计算。

$$\alpha_{t, j} = \frac{\exp(\alpha_{t, j})}{\sum_{j=1}^{T} \exp(\alpha_{t, j})}$$

$$\alpha_{t, j} = \alpha(s_{t-1}, \alpha_{j})$$
(4.5)

公式 4.5 中, $\alpha_{t,j}$ 用以计算编码部分位置 j 的词相较于解码器部分位置 t 的词影响程度。一般将 a 参数化为前馈神经网络和系统的其余部分并且共同训练。

4.3 ResNet 算法改进

本章节改进 ResNet-50 模型进行 DR 等级分类任务的目的在于准确定位病灶区域,缩小特征提取范围,减少网络参数,从而提高后续模型的分类精度,加快运行速度。本文通过微调 ResNet-50 基础网络模型实现糖尿病视网膜病变的精准分级。改进后的网络模型如图 4.5 所示。

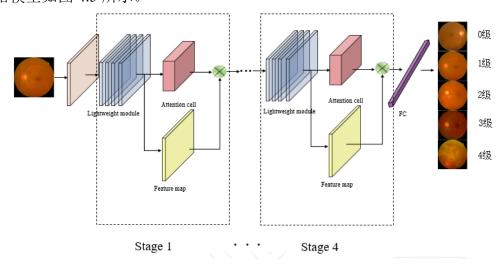


图 4.5 网络结构改进示意图

在 DR 的分级诊断中, 微小的病理特征如 MA 等类别差异小的细粒度图像在医生的诊断中发挥重要作用, 但是眼底图像中包含视盘、黄斑等组织结构, 这些无关信息会对 DR 等级判断造成干扰。因此, 将注意力机制引入 ResNet-50 的残差模块中, 以

此构建 AResNet-50 模型,使网络能够在提取特征时,选择更需要关注的病灶区域,更好地模仿临床医生的诊断过程,不断提高准确率。全局平均池化层(Global Average Pool, GAP)可以压缩空间信息,从而对输入的信息空间变化具有较好的鲁棒性,但相比于其他区域,特征相关性要低于同一区域,GAP 太过简单,无法很好地提取相关性信息。

基于此,本文提出一种基于 ResNet-50 模型输出特征 F 作为其输入的新注意力机制。放入包含 3 个 1×1 的卷积层中,并且加入 ReLU 和 Sigmoid 激活函数以获得注意力映射,然后将特征图像的原始信息和注意力单元所获取的通道信息相结合,以增加微小病灶特征的网络权重。注意力机制的结构如图 4.6 所示。

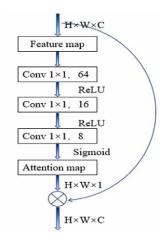


图 4.6 注意力单元结构图

其中 Conv1×1 表示核为 1×1 的卷积, Feature Map 表示 ResNet 提取的特征图。通过 ResNet 模型的卷积操作得到特征向量 z。

$$Z = \Phi\left(\sum_{i, j}^{N_p} X_{i, j} W_{i, j} + b\right)$$
 (4.6)

 Φ 为激活函数, N_p 为图像尺寸大小 $M \times N$,X 为输入,W 为权重,b 为偏置。将 Z 输入 3 个 1×1 卷积,以提取通道信息,并学习得到权重 W_k ,通过激活函数 ReLU 和 Sigmoid,将输入转为非线性特征,并与输入特征进行融合,生成注意力图像 T,作为下一 Block 层的输入。

$$T_{i, j} = z_{i, j} \sum_{j=1}^{N_p} \frac{1}{1 + exp(-W_k z_j)}$$
 (4.7)

AResNet-50 模型通过添加注意力机制提高了模型对小病灶的识别能力,但该过程也增加了模型计算复杂度。为提高后续病变分级的速度,本实验继续对 AResNet-50 的残差模块进行改进,将标准的 3×3卷积替换为更轻量化的深度可分离卷积,既

解决了网络层数过深导致的梯度消失与退化问题,又降低了网络的参数量,提高了运算速度。改进后的卷积结构如图 4.7 所示。

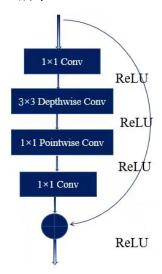


图 4.7 改进后的残差卷积块

4.4 实验设置

本实验采用了 TensorFlow 深度学习框架,使用 python 编程语言,在 Anaconda 平台上搭建模型,在 GPU 平台上运行。操作系统选用 Windows10,处理器为 i9-9820X,显卡为 NVIDIA GeForce RTX 2080Ti,内存为 128G。

4.4.1 数据集准备

在对数据集进行预处理和扩充后,为更好的测试学习好的病变检测模型,将样本集按 7:3 的比例随机划分为两组,即训练集和测试集。其中,训练集主要用以神经网络参数的训练,验证集主要在模型训练完成后,用以评估网络的性能,但其无法对超参数或参数提出修改意见,只能将其作为一个客观的衡量标准。

4. 4. 2 模型参数

模型使用交叉熵损失函数 categorical_crossentropy, 该函数可以衡量模型训练所获得的概率分布与实际分布间的差距,其函数值越小,这两个概率分布就越接近。由于模型预加载了权重,训练是一个不断微调的过程,实验中模型训练的参数设定如表4.1 所示。

参数	设定值
学习率	0.001
学习轮次数	100
批次大小	32
损失函数	categorical_crossentropy
优化器	Adam

表 4.1 网络参数设定

4.4.3 实验步骤

模型训练的具体步骤如下。

- 1、对原始样本集进行预处理,并将处理后的结果按照 7:3 的比例分成训练集和 测试集
- 2、基于 Tensorflow 框架搭建网络模型,采用 tf.global_variables_initializer ()对模型进行初始化,drop out 设置为 0.5。
- 3、将训练集中的图像样本每次随机抽取 32 张进行训练,通过 Adam 优化器来更新权重参数,计算公式如下:

$$m_t = \beta_1 v_{t-1} + (1 - \beta_1) g_t$$
 (4.8)

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$
 (4.9)

其中, \mathbf{m} 与 \mathbf{v} 的指数衰减率分别由 β_1 与 β_2 表示,时间步由 \mathbf{t} 表示, g_t 为 \mathbf{t} 的梯度。设置 η 为学习率的初始值,则梯度更新公式由 4.10 所示。

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\widehat{v_t}}} \widehat{m_t}$$
 (4.10)

4、将测试集上的图像数据随机输入至训练好的网络中,计算其输出值,并与标签值进行对比,以此计算准确率等指标,从而评价模型的学习性能。

4.5 性能评价指标

本研究在准确率、查全率的基础上,新引入参数量、FLOPs、运行时间以及 Kappa 一致性作为模型的评价指标,具体表示方式如下:

Parameters=
$$(K_h * K_w * C_{in}) * C_{out} + C_{out}$$
 (4.13)

其中, K_h 为卷积核(filter 或 kernel)的高度, K_w 表示卷积核宽度, C_{in}表示输入通道数(Channels)。

$$K = \frac{p_0 - p_e}{1 - p_e} \tag{4.14}$$

其中, p_0 表示总体分类精度, $p_e = \frac{a1 \times b1 + a2 \times b2 + \cdots + aC \times bC}{n \times n}$, $a1,a2,\cdots,aC$ 为每一类真实样本的个数, $b1,b2,\cdots,bC$ 为预测出的每一类样本个数,n 为总样本量。

4.6 实验结果与分析

本文所提出的用以糖尿病视网膜病变等级分类的深度学习模型,在训练集、测试集上的训练过程如图 4.8 所示,由图 b 可以看到,本文模型在迭代至 53 次时达到 91.5%,然后上下浮动,趋于稳定。

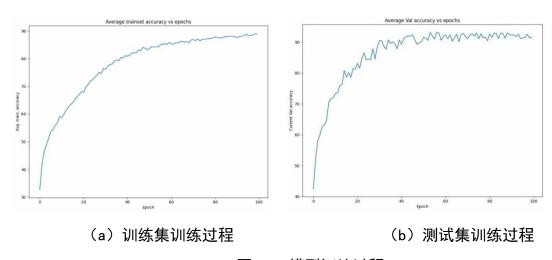


图 4.8 模型训练过程

为验证模型有效性,在本节将网络结构 ResNet-50、AResNet-50 与改进后的网络结构在同一个数据集上进行对比实验,衡量标准采用分类准确率、召回率、参数量、FLOPs、运行时间以及 Kappa 一致性,实验结果对比如表 4.2 所示。

网络	ACC (%)	R (%)	参数量(M)	FLOPs(G)	运行时间(ms)	Kappa
ResNet-50	90.2	88. 7	23. 52	17.61	181.07	0.877
AResNet-50	92.9	91.7	24.04	17.62	203.1	0.911
庞浩等人[28]	83.8	_	_	_	-	_
徐宏等人[81]	_	89. 5	=	=	-	=
基干深度可分离卷积						

12.24

9.74

158.7

0.893

91.5

和注意力机制的方法

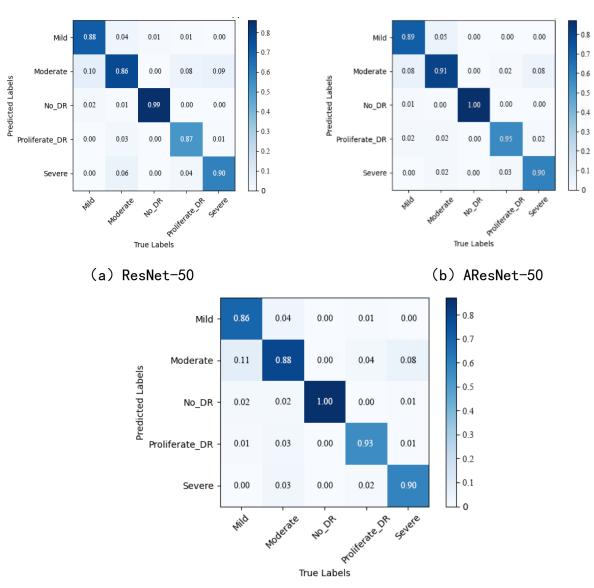
90.3

表 4.2 算法性能对比分析

由表 4.3 可知,在数据集中,通过本文方法获取的模型准确率、召回率以及 Kappa 值分别为 91.5%、90.3%以及 0.893。相比于 ResNet-50 模型,AResNet-50 和本文方法通过添加注意力机制获得了更高的分类精度,说明提出的注意力机制可关注更为有效的病灶特征,降低了背景、视盘等无关信息的干扰。相比于 AResNet-50 模型,基于深度可分离卷积和注意力机制的方法在其基础上又增添了轻量化模块,虽然分类准确率降低了 1.4%,但网络复杂度得到了改善,其中模型参数量减少了 49.08%,计算量减少了 7.88G,运行时间减少了 44.4ms。此外,为了评价本文模型在糖尿病视网膜病变分级领域的效用,还引入了另外两篇模型在本文数据集中进行比较,本文设计的病变等级分类模型,在准确率方面高于庞浩等人的方法,说明网络结构设计合理,结果优于前人的研究方法,在查全率方面优于徐宏等人的方法,说明了模型的漏诊率较低,具有一定优越性。

为证明模型的合理性,绘制了 ResNet-50、AResNet-50 以及基于深度可分离卷积和注意力机制方法的混淆矩阵图,结果如图 4.10 所示,其主对角线上的每一个元素表示每一类被分类正确的概率,其他位置则代表分类错误的概率,例如[1,0]位置上的数值表示原来应当是 1 级却被错误分类成 0 级的概率,[3,2]表示原来是 3 级却被错误分类成 2 级的概率。通过对比三种模型的混淆矩阵可知,本文方法对 1 级图像的分类准确率低于另外两种,但对于 0 级图像的分类可以达到完全正确。由图 4.9(e)可知,实际标签为 1 级、3 级和 4 级的 DR 图像更容易被误分类为 2 级病变图像,而实际标签为 2 级的 DR 图像更易被误分为 3 级和 4 级病变图像,该混淆问题的出现是由于 2 级、3 级和 4 级图像间存在 MA、EX、絮状静脉串珠等相同病灶。通过引入注意力机制,模型可以有效提取微小病灶特征,但对于异常血管、HE 等特殊病变特征的提取

能力还有待加强。相比于 AResNet-50,本文方法由于引入了深度可分离卷积,导致模型在进行分类时将 1 级和 4 级误分类为 2 级的概率较高。



(c) 基于深度可分离卷积和注意力机制的诊断方法 图 4.9 混淆矩阵结果图

4.7 本章小结

在本章节中,采用本文提出的融合了 ResNet-50、深度可分离卷积、注意力机制的集成网络模型在数据集中进行 DR 分级的实验。首先介绍 ResNet-50 分类算法、深度可分离卷积、注意力机制的基本原理,并提出一种改进的注意力机制,对 ResNet-50 的特征提取步骤进行改进,采用深度可分离卷积来替换模型中残差模块的标准卷

积,以此完成 DR 的分类任务。实验结果表明,改进后的 ResNet 方法在 DR 分级过程中具有不错的分类效果。

第五章 总结与展望

5.1 总结

随着人工智能以及大数据时代的发展,医工结合已成为热门方向,利用人工智能技术来处理与分析医学图像大幅度减少了人力资源方面的投资成本,且通过人工方式进行医学图像处理极易受到主观因素的干扰,如果眼科医生缺乏经验或长时间工作比较疲劳,极易发生漏诊、误诊等问题,如果想要提高医生的诊断精度,需要寻求具有丰富经验的医生进行协助,由此会造成医疗资源的浪费。随着经济的飞速发展,人的生活习惯也在随之改变,不合理的饮食结构以及便捷的交通方式导致高血糖、高血压、高血脂等疾病,作为现代流行病之一的高血糖极易诱发糖尿病,而糖尿病则会进一步引发眼、肾、血管等器官的病变,其中,眼作为能够直观感受到其变化的器官,及时发现病变情况就可以有针对性的治疗。此外,眼底图像包含丰富的信息,通过这些细节信息还可发现高血压、阿尔兹海默症、动脉硬化等慢性疾病,因此加大对视网膜图像的研究,对医疗界的发展具有重要意义。本文基于上述出发点,对 DR 等级的分类模型展开研究,希望尽自己最大的努力为医疗界奉献绵薄之力。本文主要完成了以下内容:

- (1) 数据预处理。针对 Kaggle 数据集中图像尺寸不一致、光照不均匀、过度曝光、对比度低、焦点模糊等问题,本文采用图像剪裁、图像灰度化、颜色与亮度归一化、CLAHE、伽马校正以及图像扩充等方式进行数据预处理,以此提高模型的分类准确率。
- (2) 多类别特征融合。针对 DR 图像的特点,首先提取 Gabor、LBP、HOG、Haralick 等影像组学特征,其次通过 AlexNet、VGG-16、ResNet-50 三种深度网络提取图像深层级特征,提出两种特征融合方式,并将融合后的特征输入 SVM、RF、kNN、ELM 等经典分类器,从而完成图像分类任务。其中,采用融合特征训练的网络模型优于单一特征的分类模型,通过融合网络提取的模型分类结果优于串行连接融合方式。
- (3) 模型结构改进。针对 DR 图像中病灶占比小、特征提取时容易遗漏以及模型参数量过大,运行时间长的问题,本文引入了注意力机制和轻量化模块,通过改进经典图像分类算法 ResNet-50 的特征提取步骤来提高模型的分类精度和运算速度。其

中注意力机制通过 1×1 和各种激活函数组成,在保证参数量的同时使模型重点关注感兴趣区域。

5.2 展望

本文提出的 DR 等级分类模型已经通过相关实验证明了其有效性,但是距离实际投入临床应用还有很大发展空间。基于研究中存在的问题,未来可以从以下几点进行开展:

- (1)数据集方面。虽然现有公开数据集较多,但大都面临数据量不足、标注粗略的问题,且由于不同数据集之间的标记异构性,导致无法在实际研究中直接使用不同数据集。因此在未来的研究工作中,可以将 GAN 与现有网络相结合,以满足现实的需要。
- (2)训练方式。现阶段 DR 等级分类模型多采用有监督的训练方式,但由于数据的标注需要耗费大量的人力、物力与资金,因此带有像素语义标注的视网膜较少。寻找数据本质特点,弱化对人工标签的依赖性,设计半监督、无监督网络模型是未来研究方向。
- (3) 网络结构。视网膜图像中,不同病灶的形状与大小差异较大。但大部分深度网络及其变体结构采用固定大小尺寸的卷积核进行处理,这就极易导致病灶特征的遗漏,因此,为了灵活地学习感受野,采用自适应卷积取代标准卷积,但这种方法的运算量很大,还应及时解决模型参数量的问题。

参考文献

- [1] 赵刚. 基于深度学习的糖尿病视网膜病变分类研究[D]. 兰州理工大学,2021.DOI:10.27206/d.cnki.ggsgu.2021.000053.
- [2] 张进松. 糖尿病视网膜病变治疗进展[J]. 湖北民族大学学报(医学版), 2021,38(01): 88-90.
- [3] 梁静.探究视网膜静脉阻塞患者视力预后的影响因素[J].中国现代药物应用,2020,14(10):27-29.
- [4] 陈 艳 燕 . 糖 尿 病 性 视 网 膜 病 变 的 发 病 及 致 盲 白 皮 书 [J]. 糖 尿 病 新 世 界 ,2000(04): 30-3.
- [5] 褚海琳,任永霞.糖尿病视网膜病变患者就诊延迟情况及就诊时间的影响因素分析[J].中国全科 医学 2021:1-6.
- [6] 郑二来, 张海玲, 黄剑真, 等. 眼球在解剖学实验中的应用探索[J]. 医学理论与实践, 2019, 32(15): 2490-2491.
- [8] Kang Q, Yang C.Oxidative stress and diabetic retinopathy: Molecular mechanisms, pathogenetic role and therapeutic implications. Redox Biol. 2020 Oct;37:101799.
- [9] Eid S,Sas KM,Abcouwer SF,et al.New insights into the mechanisms of diabetic complications:role of lipids and lipid metabolism. Diabetologia.2019 Sep;62(9):1539-1549.
- [10] 王光璐. 糖尿病视网膜病变的分级标准初探[J]. 眼科, 2005(04): 218-220.
- [11] 杨英,向彦锋.糖尿病视网膜病变患者血糖指标与视力严重程度的相关性[J].血栓与止血 学,2022,28(04):553-554+557.
- [12] Lakshminarayanan V, Kheradfallah H, Sarkar A, et al. Automated Detection and Diagnosis of Diabetic Retinopathy: A Comprehensive Survey. J Imaging. 2021 Aug 27;7(9):165.
- [13] Setio A A A, Traverso A, Bel T, et al. Validation comparison and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images the LUNA16challenge J MedicalImage Analysis 2017:42 1-13.
- [14] Xie H T, Yang DB, Sun NN, et al. Automated pulmonary nodule detection in CT images using deep convolutional neural networks J Pattern Recognition 2019:85 109-119.
- [15] Vang Y S, Chen Z, Xie X.Deep learning framework for multi-class breast cancer histology image

- classification[C]//Proceedings of International Conference on Image Analysis & Recognition, 2021:914-922.
- [16] Brancati N, Frucci M, Riccio D.Multi-classification of Breast Cancer Histology Images by Using a Fine-Tuning Strategy[C]//Proceedings of International Conference on Image Analysis & Recognition, 2019:771-778.
- [17] Kohl M,Walz C,Ludwig F,et al.Assessment of breast cancer histology using densely connected convolutional networks[C]//Proceedings of International Conference on Image Analysis & Recognition,2020:903-913.
- [18] Shen W, Zhou M, Yang F, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification J Pattern recognition 2019:61 663-673.
- [19] Eltanboly A, Ghazaf M, Khalil A, et al. An integrated framework for automatic clinical assessment of diabetic retinopathy grade using spectral domain OCT images[C]// 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018:1431-1435.
- [20] Abramoff M D, Lavin P T, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices[J]. Yearbook of Paediatric Endocrinology, 2019,1:39.
- [21] Nayak J, Bhat P S, Acharya R, et al. Automated identification of diabetic retinopathy stages using diagital fundus images Journal of medical systems,2008,32(2):107-115.
- [22] 范家伟,张如如,陆萌,等.深度学习方法在糖尿病视网膜病变诊断中的应用[J].自动化学报,2021,47(05):985-1004.
- [23] Lee AY, Yanagihara RT, Lee CS, et al. Multicenter, Head-to-Head, Real-World Validation Study of Seven Automated Artificial Intelligence Diabetic Retinopathy Screening Systems. Diabetes Care. 2021 May;44(5):1168-1175.
- [24] Wang X N, Dai L,Li ST,et al. Automatic Grading System for Diabetic Retinopathy Diagnosis Using Deep Learning Artificial Intelligence Software. Curr Eye Res. 2020 Dec;45(12):1550-1555.
- [25] Srivastava R, Duan L, Wong D W, et al. Detecting retinal microaneurysms and hemorrhages with robustness to the presence of blood vessels[J]. Computer Methods and Programs in Biomedicine, 2017: 83-91.
- [26] 孟琳, 刘静, 曹慧,等. 基于Frangi滤波器和Otsu视网膜血管分割[J]. 激光与光电子学进展,

- 2019, 56(18):7.
- [27] Raman V, Then P, Sumari P. Proposed retinal abnormality detection and classification approach: Computer aided detection for diabetic retinopathy by machine learning approaches. 8th IEEE International Conference on Communication Software and Networks (ICCSN). June 04-06, 2016[C]. Beijing, PEOPLES R CHINA: IEEE, 2016: 636-641.
- [28] 庞浩, 王枞. 用于糖尿病视网膜病变检测的深度学习模型[J]. 软件学报, 2017,28(11): 3018-3029.
- [29] Gargey A R, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning[J].Ophthalmology, 2017:962-969.
- [30] Li X, Pang T, Xong B, et al. Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification[C]// 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE, 2017.
- [31] Dutta S, Manideep B C, Basha S M, et al. Classification of Diabetic Retinopathy Images by Using Deep Learning Models[J]. International Journal of Grid & Distributed Computing, 2018, 11(1): 99-106.
- [32] Wang X, Lu Y, Wang Y. Diabetic retinopathy stage classification using convolutional neural networks[C]//2018 IEEE International Conference on Information Reuse and Integration (IRI). IEEE, 2018: 465-471.
- [33] Ghosh R, Ghosh K, Maitra S, Automatic Detection and Classification of Diabetic Retinopathy Stages Using CNN[C]. International Conference on Signal Processing & Integrated Networks. IEEE, 2018.
- [34] Jiang Y , Tan N , Peng T , et al. Retinal Vessels Segmentation Based on Dilated Multi-Scale Convolutional Neural Network[J]. IEEE Access, 2019.
- [35] Chen Q, Sun X, Zhang N, et al. Mini Lesions Detection on Diabetic Retinopathy Images via Large Scale CNN Features[J]. 2019.
- [36] Jiang H, Yang K, Gao M, et al. An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification[C]// 2019 41st Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2019.
- [37] 张子振, 刘明, 朱德江. 融合注意力机制和高效网络的糖尿病视网膜病变识别与分类[J]. 中国图象图形学报, 2020, 25(08): 1708-1718.

- [38] 谢云霞,黄海于,胡建斌.基于深度卷积神经网络的糖尿病视网膜病变分期及病灶检测[J].计算机应用,2020,40(08):2460-2464.
- [39] Zhang Z , Wu C , Coleman S , et al. DENSE-INception U-net for medical image segmentation[J]. Computer Methods and Programs in Biomedicine, 2020, 192:105395.
- [40] 孙雨琛,刘宇红,张达峰,等.基于深度学习的糖尿病视网膜病变诊断方法[J].激光与光电子学进展,2020,57(24):359-366.
- [41] Erciyas A, Barışçı N. An Effective Method for Detecting and Classifying Diabetic Retinopathy Lesions Based on Deep Learning. Comput Math Methods Med. 2021 May 31;2021:9928899.
- [42] Ai Z, Huang X, Fan Y, et al. DR-IIXRN: Detection Algorithm of Diabetic Retinopathy Based on Deep Ensemble Learning and Attention Mechanism. Front Neuroinform. 2021 Dec 24;15:778552.
- [43] Gayathri S, Gopi V P, Palanisamy P. Diabetic retinopathy classification based on multipath CNN and machine learning classifiers. Phys Eng Sci Med. 2021 Sep;44(3):639-653.
- [44] Kajan S, Goga J, Lacko K, et al. Detection of Diabetic Retinopathy Using Pretrained Deep Neural Networks[C]// 2020 Cybernetics & Informatics (K&I). 2020.
- [45] Zeng Y, Li Y, Chen J, et al. ELM embedded discriminative dictionary learning for image classification. Neural Netw. 2020 Mar;123:331-342.
- [46] Tong Y, Lu K, Yang Y, et al. Can natural language processing help differentiate inflammatory intestinal diseases in China? Models applying random forest and convolutional neural network approaches. BMC Med Inform Decis Mak. 2020 Sep 29;20(1):248.
- [47] 孙琦,梁永全.基于PMI的递归自编码器的情感分析方法[J].软件导刊,2021,20(06):59-62.
- [48] Bulso N, Roudi Y. Restricted Boltzmann Machines as Models of Interacting Variables. Neural Comput. 2021 Sep 16;33(10):2646-2681.
- [49] Liu D, Zhang Y, Luo L, et al. PDC-Net: Robust point cloud registration using deep cyclic neural network combined with PCA. Appl Opt. 2021 Apr 10;60(11):2990-2997.
- [50] Arrizza AM, Caimo A. Bayesian dynamic network actor models with application to South Korean COVID-19 patient movement data. Stat Methods Appt. 2021;30(5):1465-1483.
- [51] Apostolopoulos ID, Pintelas EG, Livieris IE, et al. Automatic classification of solitary pulmonary nodules in PET/CT imaging employing transfer learning techniques. Med Biol Eng Comput. 2021 Jun;59(6):1299-1310.

- [52] Wei G, Qiu M, Zhang K, et al. A multi-feature image retrieval scheme for pulmonary nodule diagnosis[J]. Medicine, 2020, 99(4):e18724.
- [53] Sharan RV, Xiong H, Berkovsky S. Benchmarking Audio Signal Representation Techniques for Classification with Convolutional Neural Networks. Sensors (Basel). 2021 May 14;21(10):3434.
- [54] Krizhevsky A, Sutskever I, Hinton G E. Imagenet Classification with Deep Convolutional Neural Networks [J]. Advances in neural information processing systems, 2012, 25: 1097-1105.
- [55] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. arXiv preprint arXiv:14091556, 2014.
- [56] Christian S, Sergey I, Vincent V,et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.arXiv:1602.07261, 2016.
- [57] Bartlett PL, Helmbold DP, Long PM. Gradient Descent with Identity Initialization Efficiently Learns Positive-Definite Linear Transformations by Deep Residual Networks. Neural Comput. 2019 Mar;31(3):477-502.
- [58] Huang G, Liu Z, Van Der Maaten L, et al. Densely Connected Convolutional Networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, Hawaii: IEEE, 2017: 4700–4708.
- [59] 黄卫. 医学 CT 图像对比度增强算法研充综述[J]. 电脑知识与技术, 2018, 14(20):175-178.
- [60] Jadhav AS, Patil PB, Biradar S. Analysis on diagnosing diabetic retinopathy by segmenting blood vessels, optic disc and retinal abnormalities. J Med Eng Technol. 2020 Aug;44(6):299-316.
- [61] 纪平, 胡学友, 张瑞琦. 基于直方图均衡算法的图像增强技术研究[J]. 蚌埠学院学报,2021,10(02):40-43.
- [62] 龙鑫, 何国田. 基于多层融合和细节恢复的图像增强方法[J]. 计算机应用研究, 2020(2):4.
- [63] Barshooi AH, Amirkhani A. A novel data augmentation based on Gabor filter and convolutional deep learning for improving the classification of COVID-19 chest X-Ray images. Biomed Signal Process Control. 2022 Feb;72:103326.
- [64] Maheshwari S, Sharma RR, Kumar M. LBP-based information assisted intelligent system for COVID-19 identification. Comput Biol Med. 2021 Jul;134:104453.
- [65] Zhang C, Wu S, Lu Z, et al. Hybrid adversarial-discriminative network for leukocyte classification in leukemia. Med Phys. 2020 Aug;47(8):3732-3744.

- [66] Perumal V, Narayanan V, Rajasekar S J S. Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. Appl Intell (Dordr). 2021;51(1):341-358.
- [67] Jacques F,Gregor V L,Cailos T,et al.Time Series Analysis of Blockchain-Based Cryptocurrency Price Changes.arXiv:2202.13874,2022.
- [68] 闫涵, 张旭秀, 张净丹. 多感知兴趣区域特征融合的图像识别方法[J].智能系统学报, 2020, 22(3):1-7.
- [69] Li Y H, Yeh N N, Chen S J, et al. Computer-Assisted Diagnosis for Diabetic Retinopathy Based on Fundus Images Using Deep Convolutional Neural Network[J]. Mobile Information Systems, 2019, 2019:1-14.
- [70] 张慧婷. 用于糖尿病视网膜病变分级的深度学习模型研究[D].太原理工大学,2021.
- [71] Dhanasri V, Dheekshitha S, Soundarya M. Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data[C]//Proceedings of the 2021 International Conference on Computing, Communication, Electrical and Biomedical Systems (ICCCEBS) Coimbatore, India, Mar 25-26,2021. United Kingdom: IOP Publishing,2021, 1916(1): 012168.
- [72] Chen B, Li P, Li B, et al. Psvit: Better Vision Transformer Via Token Pooling and Attention Sharing [J]. arXiv:210803428, 2021.
- [73] Fan D-P, Zhou T, Ji G-P, et al. Inf-Net: Automatic Covid-19 Lung Infection Segmentation from Ct Images[J]. IEEE Transactions on Medical Imaging, 2020, 39(8): 2626-2637.
- [74] Huang Z L, Wang X G, Huang L C, et al. CCNet: criss-cross attention for semantic segmentation[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2019: 603-612.
- [75] Tóth, F Honbolygó, O Szalárdy, et al. The effects of speech processing units on auditory stream segregation and selective attention in a multi-talker (cocktail party) situation[J]. Cortex, 2020.
- [76] Li X, Wei Y, Wang L, et al. MSGSE-Net: Multi-scale guided squeeze-and-excitation network for subcortical brain structure segmentation[J]. Neurocomputing, 2021, 461(3).
- [77] Tan T, Wang Z, Du H, et al. Lightweight pyramid network with spatial attention mechanism for accurate retinal vessel segmentation[J]. International Journal of Computer Assisted Radiology and Surgery, 2021, 16(1).
- [78] Xiao Y, Yin H, Wang SH, et al. TReC: Transferred ResNet and CBAM for Detecting Brain Diseases.

- Front Neuroinform. 2021 Dec 23;15:781551.
- [79] Lee SH, Goëau H, Bonnet P, et al. Attention-Based Recurrent Neural Network for Plant Disease Classification. Front Plant Sci. 2020 Dec 14;11:601250.
- [80] Bahdanau D,Cho K,Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv peprint arXiv: 1409. 0473, 2014.
- [81] 徐宏. 基于眼底图像的糖尿病视网膜病变智能诊断[D]. 电子科技大学, 2019.

附 录

英文缩略词

缩写	英文全称	中文全称
DR	Diabetic retinopathy	糖尿病视网膜病变
CAD	Computer Aided Diagnosis	计算机辅助诊断技术
CLAHE	Contrast Limited Adaptive Histogram Equalization	限制对比度自适应直方图均衡算法
SVM	Support Vector Machine	支持向量机
RF	Random Forest	随机森林
KNN	k-Nearest Neighbor	K 最近邻算法
CNN	convolutional neural network	卷积神经网络
RAE	Recursive Auto Encoders	递归自动编码器
RBM	Restricted Boltzmann Machines	受限玻尔兹曼机
RNN	Recurrent Neural Network	循环神经网络
DBN	Deep Belief Nets	深度置信网络
LBP	Local Binary Pattern	局部二值模式
HOG	Histogram of Oriented Gradient	方向梯度直方图特征
FCN	Fully convolutional networks	全卷积神经网络
DW	Depthwise separable convolution	逐通道卷积
PW	Pointwise convolution	逐点卷积
SE	Sequeze and Excitationblock	通道注意力机制
CBAM	Convolutional Block Attention Module	混合注意力机制
GAP	Global Average Pool	全局平均池化层

致 谢

转眼三年的研究生学习生涯马上要划上句号,回忆三年的研究生涯令我感慨不已。 值得欣慰的是,三年的研究生涯没有白费,我在无数个学习和拼搏的日夜度过, 接触的老师教会我做研究和做事的态度让我受益匪浅。而我本人的学习科研能力以及 生活处事方面与本科阶段将较已有大幅提升。

在此首先感谢感谢我的母校山东中医药大学为我创造的良好温馨的科研和生活 环境,这三年里我学到了很多有用的知识,使我获益匪浅,希望今后我可以将这些知识更多地应用到实践中去。

在论文撰写过程中,要特别感谢曹慧老师对我的耐心细致的指导与帮助,从课题选择、资料查阅、一次次地实验,到论文的最终完成,各个环节都是在曹老师的耐心指导下完成的,倾注了导师大量的心血。老师对我的严格要求,使我没有在研究过程中放弃和懈怠,对于本课题所涉及的重点难点,老师的耐心讲解总能为我的实验指明正确的方向。导师一丝不苟的工作作风、严谨认真的治学态度、平易近人的人格魅力对我产生深远的影响。

衷心感谢研究生阶段无私帮助我的邱鹏老师、刘静老师、魏德建老师、马金刚老师、李明老师;感谢 508 实验室帮助过我的同学、师哥师姐、师弟师妹们,你们给予了我莫大的鼓励与支持,和你们度过了许多欢乐时光。

最后感谢我的父母,感谢在我学习生涯从未出现过的男朋友,使我能够专注学业,感谢一路走来所有默默关心和支持我的大家。

攻读硕士期间取得的研究成果

一、论文发表

- [1] 聂永琦,曹慧,杨锋,等.深度学习在糖尿病视网膜病灶检测中的应用综述[J].计算机工程与应用,2021,57(20):25-41.
- [2] Nie Yongqi, Ma Chong, Qiu Peng, et al. Computer Assisted Grade Diagnosis Using Feature Fusion and Network Algorithm Flow[J]. Journal of Physics: Conference Series, 2021, 2033(1).
- [3] 刘书轩,邱鹏,聂永琦,刘静,曹慧.虚拟现实在脑卒中患者运动功能康复中的应用[J]. 计算机工程与应用,2021,57(21):68-79.
- 二、获得专利授权
- [1] 一种基于深度可分离卷积和注意力机制的糖尿病视网膜病变图像处理方法。(已 受理)