摘 要

糖尿病性视网膜病变（DR）是糖尿病性微血管病变中最重要的表现，是一种具有特异性改变的眼底病变，是糖尿病的严重并发证之一。DR是一种隐性疾病,但由于视网膜病变的患病率与患糖尿病的年龄和疾病的持续时间有，而且它只有当视网膜的变化发展到一个水平才会被患者注意到，这时治疗已经相当复杂，甚至是不可能的。

不同分期的DR患者激光光凝治疗后其视力预后不同。因此，尽早通过定期筛查发现症状至关重要的。数字图像处理技术可以十分有效地降低筛查及分期成本，而且近年来，深度学习在计算机视觉方面取得了巨大的进步，并在利用计算机视觉完成医学影像的阅片工作方面展现出了良好的应用前景。

本次毕业设计针对DR对眼底图像进行分类工作，通过构建深度卷积神经网络，完成了原始照片的特征提取、特征组合和结果分类，最终得出DR分期结果。

具体方法分为下列个步骤（1）通过均衡图片质量和调整图片尺寸对眼底图像进行预处理，达到压缩数据大小的目的。（2）通过对容量较少的类别进行过采样解决类别不平衡的问题。（3）使用较小的图片训练较小的网络，以初始化网络的参数。（4）分别训练两个结构不同的卷积神经网络（Convolutional Neural Network, CNN）（5）最终的结果由两个CNN所集成学习的结果得到。

实验结果在kaggle竞赛的数据集中进行测试，二次加权Kappa系数达到0.844，表明了输出结果与正确结果的高度一致。

关键词：糖尿病视网膜病变，图像处理，卷积神经网络

ABSTRACT

Diabetic retinopathy (DR) is the most important manifestation of of diabetic capillary blood vesse, and it is a kind of retinopathy with specific changes. It is a severe complication of diabetes. DR is a recessive disease. However, the prevalence of retinopathy is related to the age of diabetes ,as well as the duration of the disease, it will not be noticed by the patient before the retinal changes reach to a level, when the treatment already becomes rather complicated , even impossible.

Patients with different stages of DR have different prognosis of eyesight after laser photocoagulation treatment. Therefore, early detection through regular screening is crucial. Digital image processing technology can effectively reduce the cost of screening and staging, and in recent years, deep learning has made great progress in computer vision, showing good application and prospect in the medical image reading..

This graduation project aims at constructing a deep convolutional neural network to classify the fundus images by DR. The projesct dealt with the topics about feature extraction, feature combination, and result classification of the original photos, and finally obtained the DR staging results.

The specific method is divided into the following steps: (1) To compress the size of the fundus image, I crop away all background and resize the images to squares of 128, 256 and 512 pixel. (2) because the classes in the dataset are highly imbalanced ,oversampling the rare classes is used. (3) Use a smaller picture to train a smaller network to initialize the network parameters. (4) Train the two convolutional neural networks (CNNs) with different structures separately (5) The ﬁnal solution is a simple average of blends with features from two deep convolutional networks and three sets of weights for each network.

The experimental results were tested in the data set of the Kaggle challenge. The quadratic weighted Kappa reached 0.844, indicating that the output was highly consistent with the real classes.

**KEY WORDS**：Diabetic retinopathy, image processing, Convolutional Neural Network

目 录

[第一章 绪论 1](#_Toc516447871)

[1.1研究背景和意义 1](#_Toc516447872)

[1.1.1 糖尿病性视网膜病变（DR） 1](#_Toc516447873)

[1.1.2眼底图像 1](#_Toc516447874)

[1.2研究现状 2](#_Toc516447875)

[1.2.1眼底图像预处理 2](#_Toc516447876)

[1.2.2病变区域检测 2](#_Toc516447877)

[1.2.3 DR筛查 4](#_Toc516447878)

[1.3 论文内容及结构安排 5](#_Toc516447879)

[第二章 理论基础 7](#_Toc516447880)

[2.1引言 7](#_Toc516447881)

[2.2预处理 7](#_Toc516447882)

[2.2.1 眼底图像中分量图像的提取与分析 7](#_Toc516447883)

[2.2.2 直方图均衡化 7](#_Toc516447884)

[2.3神经网络理论 9](#_Toc516447885)

[2.3.1机器学习与神经网络 9](#_Toc516447886)

[2.3.2神经元模型 9](#_Toc516447887)

[2.3.3感知机与多层网络模型 11](#_Toc516447888)

[2.3.4误差逆传播算法 12](#_Toc516447889)

[2.4卷积神经网络理论 15](#_Toc516447890)

[2.4.1概述 15](#_Toc516447891)

[2.4.2卷积神经网络的基本结构 15](#_Toc516447892)

[2.4.3 卷积 16](#_Toc516447893)

[2.4.4 池化 17](#_Toc516447894)

[2.4.5 卷积神经网络常用的训练方法 18](#_Toc516447895)

[第三章 糖尿病视网膜病变分期算法 21](#_Toc516447896)

[3.1引言 21](#_Toc516447897)

[3.2预处理 21](#_Toc516447898)

[3.2.1 尺寸调整 21](#_Toc516447899)

[3.2.2 数据集扩充 21](#_Toc516447900)

[3.2.3 重采样 21](#_Toc516447901)

[3.3网络模型 22](#_Toc516447902)

[3.4网络初始化和预训练 22](#_Toc516447903)

[3.5特征抽取 23](#_Toc516447904)

[3.6左右眼混合 23](#_Toc516447905)

[第四章 实验与讨论 25](#_Toc516447906)

[4.1数据集 25](#_Toc516447907)

[4.1.1 Messidor 数据集 25](#_Toc516447908)

[4.1.2 ROC数据集 25](#_Toc516447909)

[4.1.3 DIARETDB1 数据集 25](#_Toc516447910)

[4.1.4 DR1 数据集 25](#_Toc516447911)

[4.1.5 Kaggle-DR数据集 25](#_Toc516447912)

[4.2衡量指标 26](#_Toc516447913)

[4.2.1 Kappa系数 26](#_Toc516447914)

[4.2.2 加权Kappa系数 26](#_Toc516447915)

[4.3实验设计 27](#_Toc516447916)

[4.4时间复杂度 29](#_Toc516447917)

[4.5结果比较 29](#_Toc516447918)

[4.6本章小结 30](#_Toc516447919)

[第五章 总结与展望 33](#_Toc516447920)

[5.1总结 33](#_Toc516447921)

[5.2展望 33](#_Toc516447922)

[参考文献 35](#_Toc516447923)

[致 谢 39](#_Toc516447924)

[毕业设计小结 41](#_Toc516447925)

第一章 绪论

1.1研究背景和意义

1.1.1 糖尿病性视网膜病变（DR）

糖尿病性[视网膜病变](https://baike.baidu.com/item/视网膜病变#_blank)（DR）是糖尿病性微血管病变中最重要的表现，是一种具有特异性改变的眼底病变，是糖尿病的严重并发证之一。DR是一重隐性疾病,它只有当视网膜的变化发展到一个水平才会被患者注意到，这时治疗已经相当复杂，甚至是不可能的。而且视网膜病变的患病率也与患糖尿病的年龄和疾病的持续时间有关[6]。

根据患者眼底视网膜病变程度的不同，临床将糖尿病视网膜病变分为5期[15]。本次毕业设计针对DR对眼底图像进行分类工作，通过构建深度卷积神经网络，完成了原始照片的特征提取、特征组合和结果分类，最终得出DR分期结果

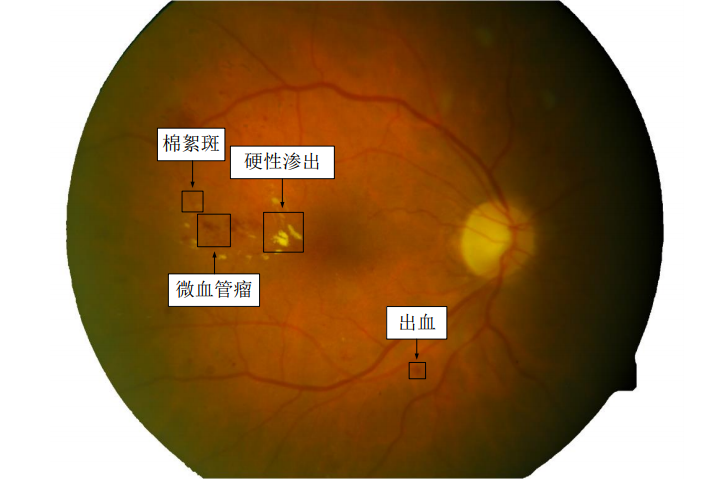


图1-1 几种常见糖尿病视网膜病变病理特征

1.1.2眼底图像

不同分期的DR患者激光光凝治疗后其视力预后不同[14]。因此，尽早通过定期筛查发现症状至关重要的。数字图像处理技术可以十分有效地降低筛查及分期成本，而且近年来，深度学习在计算机视觉方面取得了巨大的进步，并在利用计算机视觉完成医学影像的阅片工作方面展现出了良好的应用前景。

医学图像分析这一研究领域如今吸引了来自科学家和医生的广泛关注。这个领域的目标是开发用于辅助诊断的计算工具，这将有助于将病理学和解剖结构的关键区域的量化和可视化。DR分期算法在眼底图像上进行。眼底图像拍摄首先需要患者使用用于放大瞳孔制药眼药水，之后，患者盯住一个固定设备保持眼球的静止。拍照时,患者会看到一系列的明亮的闪光。整个过程大约需要5到10分钟。同时，为了确保及时的DR治疗,糖尿病患者必须每年至少进行一次的眼底照片拍摄[6]。

1.2研究现状

1.2.1眼底图像预处理

眼底图像预处理的主要目的是为了从眼底图像的候选区域中提取感兴趣区域（Rigion of Interest，ROI），并平滑图像，使图像利于对眼底图像的分割。

Datta[15]等采用一种模糊对比增强预处理操作方法，先从彩色眼底图像中选择绿色分量图像，进行模糊直方图计算，并对直方图信息进行分块，再均衡化处理直方图得到新的图像。去除血管和视盘等干扰后，设置图像绿色分量阈值和病灶块大小阈值，对图像微血管瘤进行检测，根据微血管瘤个数判断DR病变等级。Pratt[16]等运用卷积神经网络处理 Kaggle 竞赛公开眼底图像数据，先对图像进行颜色规范化处理以去除光照不均等影响，将图像大小设置为512×512像素。

1.2.2病变区域检测

糖尿病是一组以高血糖为特征的代谢性疾病。高血糖则是由于胰岛素分泌缺陷或其生物作用受损，或两者兼有引起。糖尿病时长期存在的高血糖将会影响包括视网膜在内的循环系统。DR起因为视网膜血管受损，液体渗出到视网膜。DR表现为动脉瘤、出血斑点、硬性渗出、棉绒斑、静脉串珠状、视网膜内微血管异常（IRMA），以及黄斑水肿等。目前针对DR图像处理的算法通常就从数字眼底图像提取这些特征再进行分类。

基于全局图像的DR分类方法是针对全图的整体特征信息来设计不同的特征提取方法进行分类，并不需要针对单一的病灶特征信息进行特定的检测方法。基于局部病灶的分类方法是对眼底图像中不同类型的病灶进行分析，综合病灶的数量和类型进行分类。主要检测的病灶包括硬性渗出物、出血点和微血管瘤 3 大类。病灶检测方法可分为两种: 一种是基于去除眼底图像正常结构的方法，先分割眼底图像中正常结构，再提取病灶区域。进而根据所得病灶区域对眼底图像进行分类；另一种是基于直接提取病灶的方法，对眼底图像提取病灶区域进行分类，不需要检测眼底图像背景结构信息。

Rocha[16]等把视觉描述子词袋的概念引入到硬性渗出物的检测过程，首先通过特定的算法在训练集的每幅图像上检测出一定数量的兴趣点，并用向量描述其邻域，将每幅图像表示为它所包含兴趣点的集合．然后，分别用表示正常区域和渗出区域的向量建立有代表性的视觉字典，再将每幅图像的兴趣点映射到该字典中生成直方图来表示原始图像．最后，向支持向量机输入生成的直方图信号来进行训练和分类，判断测试图像是否含有硬性渗出物．该方法不用对图像做任何预处理和后处理，在RetiDB 和 Messidor 数据集上该方法的灵敏度为90%，特异度为87%。该方法使用特征描述子对图像进行感兴趣点检测，用感兴趣点编码来组合特征进行分类，不用检测眼底图像正常结构，因此方法简单．但是对微血管瘤等小型病灶检测能力不足，同时加速稳健特征( speeded up robust features，SURF) 和尺度不变特性转换(scale-invariant featuretransform，SIFT)不能检测棉絮斑和新生血管等病变结构。

Pires[17]等在文献的基础上应用描述子词袋方法对眼底图像的多种病灶进行识别。以DR1数据集作为训练集，在已标注病灶区域的图像上对感兴趣点进行 SIFT和SURF 描述子表示，聚类生成字典集，再对图像进行编码生成特征向量，运用支持向量机分类器进行正常和异常分类，最后根据所得模型对自己的数据集进行多种病灶分类。在DR1 数据集中，模型识别硬性渗出物区域的特异度为90.0%，灵敏度为86.7%; 在识别出血区域的特异度为90. 0%，灵敏度为83.3%。在Pires数据集上检测硬性渗出物的特异度和灵敏度分别为88。9% 和100%，检测出血的特异度和灵敏度分别为95%和67%．该方法可实现跨数据集检测多种类型的病灶，可以检测不同分辨率的图像病灶，但描述子不能对微血管瘤等其他病灶进行检测。

Antal [18]等综合眼底图像水平信息、损伤区域检测信息和解剖结构信息来对图像进行分类。其中，图像水平信息有血管分割、纹理和频率信息，损伤检测信息有微血管瘤和硬性渗出物检测，解剖结构信息有视盘和黄斑结构信息．综合这3类信息，选择其中18个特征，设计分类器进行图像分类。 图像所用数据是Messidor数据集，在正常图像和轻度DR图像两类数据分类的正确率为 90%，灵敏度和正确率分别为94%和90%; 在正常数据和异常数据 (包括轻度、中度和重度) 两分类实验中正确率为90%，灵敏度和正确率分别为90% 和91%。该方法综合了3种眼底图像信息，组合多种分类器，对DR病变程度分类有很好的效果。

Pratt[19]等运用卷积神经网络处理Kaggle竞赛公开眼底图像数据，先对图像进行颜色规范化处理以去除光照不均等影响，将图像大小设置为 512×512像素。使用随机角度旋转和翻转等数据扩增技术来增加训练数据量。设计一个15层网络结构，使用随机梯度下降法训练网络.测试5000 张图像，包含5类病变程度图像，分类正确率为75%，分类灵敏度和特异度分别为 30%和95%。该方法没有运用眼底图像基本结构信息。基于全局图像的分类方法较基于局部病灶的分类方法有其独特优势，通过对全局图像提取特征，检测所有异常区域，实现关于眼底图像病灶的规则和特征的综合表示，不考虑各类病灶的结构特点，避免复杂的特征提取工作。

1.2.3 DR筛查

由于筛查 DR 所需要的基础设施的本来就少，糖尿病患者数量却又持续增加，将会变的更加不足，这一矛盾短期内无法调和。对一个全面的、自动化的 DR 筛查方法的需要早已成为共识。DR 筛查是一个热门的研究领域，大量的研究人员投身于这一领域并致力于其研究发展。将糖尿病视网膜病变眼底图像作为输入使用模式识别，图像分类，机器学习等方法的早期研究工作已经取得了不错的进展。便提出了自动化筛查的方法以解决人工筛查中诸如灵敏度低，成本高，耗时多和人工检测能力低等问题。自动筛选方法的目标是确定患者是否有进一步治疗的需要。Abramoff 和 Niemeijer 关注了自动化方法在大规模筛查中的高灵敏度和高生产力的问题[28]。Donsa 等人强调了决策支持系统和机器学习方法在糖尿病治疗中存在的若干问题和挑战[29]。部分研究人员在一些文章中提出了针对早期 DR 信号微血管瘤的自动诊查系统[30]。在其他一些研究文章中也提出了针对渗出区的自动检测系统[31]。渗出区是带有尖锐边缘的白色或淡黄色的区域[32]。针对其他特征比如出血，异常新生血管，在一些文献中提到了许多检查方法。Larsen 等人结合图像处理技术提出了一个自动分级系统，可以处理出血点和微血管瘤 [33]。Jelinek 等人开发了一个用于检测微血管瘤的有效工具[34]。Priya 和 Aruna 研究提出了一个可以识别正常，非增殖和增殖三个不同阶段的系统，该系统通过对原始彩色眼底图像使用各种图像处理技术从中抽取特征并用 SVM 进行分类[35]。之后又结合概率神经网络（PNN）和支持向量机（SVM）对该系统进行了加强[36]。文献[37]提出了一种利用径向基函数神经网络的分类方法。Permalsamy 等人提出并开发了一个可以处理并进行分级的 ADRES3.0系统，对于眼科医生它可以用做 DR 常规临床检查诊断的一个辅助工具[38]

以Google为例，Google通过与印度和美国的医生密切合作，创建了一个包含 12.8 万张眼底扫描图片的数据集，用于训练一个检测糖尿病性视网膜病变的深度神经网络。随后，通过把神经网络的诊断结果，与7个美国委员会认证的眼科医生的检查结果相对比，来验证其诊断灵敏度。结果表明，该算法的效果与眼科医生小组的表现不分伯仲。例如，下图描述的验证集中，该算法的F-Score 值（结合灵敏度和特异性的度量，最大值为1），为 0.95 分，稍微高于眼科医生的F-Score中位数（测得为0.91分）。这些成果令人倍感振奋，但接下来仍任重而道远。首先，尽管用来评估算法的传统质量衡量指标的效果十分鼓舞人心，Google仍在加深与视网膜研究领域的专科医生合作，以便确立更具参考价值的量化衡量标准。此外，眼底图像的判读只是糖尿病视网膜病变诊断的众多步骤中的一步，其余的步骤在某些情况下也是不可缺少的，比如用来检查视网膜各层的光学相干断层扫描（Ocular Coherence Tomography, OCT）。Google 正在探索如何将此研究成果与DeepMind的OCT研究进行结合，以便进一步协助医生们对于糖尿病视网膜病变和其他眼部疾病的诊断。

1.3 论文内容及结构安排

本文首先对深度学习理论，特别是神经网络进行了介绍，然后对所使用方法进行了详细地研究与介绍，并在相关数据集上对该方法进行了相应实验。

本文的组织结构如下：

第一章：绪论。对于本文的研究背景及意义，国内外研究现状，本文的研究内筒和全文的组织结构进行介绍。

第二章：理论基础。详细地介绍了深度学习中，神经网络的基本概念及模型。

第三章：糖尿病视网膜病变分期算法。对提出的方法进行了详细地介绍。

第四章：对比实验。对前面已经提到的方法进行对比实验。

第五章：总结与展望。对全文工作进行总结分析，并对今后的研究工作进行了展望。

第二章 理论基础

2.1引言

本章首先介绍对眼底图像进行预处理的图像处理方法和原理，接下来介绍神经网络的理论及模型，并且重点介绍卷积神经网络中的卷积、池化、全连接方法及其训练原理。

2.2预处理

2.2.1 眼底图像中分量图像的提取与分析

彩色图像的RGB各通道的图像如图2-1所示。 比较图2-1中分离的三个通道图像分别有不同的灰度特征。

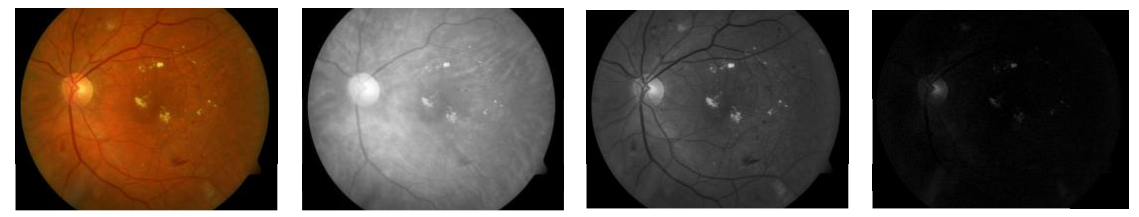


图2-1 由左到右依次为原彩色眼底图像， R 通道图像，G 通道图像，B 通道图像

在预处理阶段，从RGB颜色平面中选择绿色的通道，因为微动脉瘤在这个颜色通道与背景有最大的对比度

 (2-1)

其中，G是绿色通道（G）成分，R和B分别是红色通道和蓝色通道成分。

2.2.2 直方图均衡化

图像灰度直方图图均衡化操作，可以使图像的灰度分部更加均匀，图像所 占有的像素灰度之间距离拉开，进而加大图像反差，改善图像显示效果，达到图像增强的目的。若一幅图像直方图均匀分布，则其反馈到人眼的色调感觉就会比较协调。 假设对原图像灰度级在0~1间进行归一化，即。为原图像 的灰度分部概率密度函数，直方图均衡化操作的目的其实就是找到一个灰度变换函数，让转换灰度值能满足，这里，归一化为，使得与之间形成对应的映射关系，要求转换后的图像的灰度分部概率密度函数等于1（转换后概率密度函数在上均匀分布），期望所有灰度级出现概率相同。直方图均衡变换函数如图2-2所示。

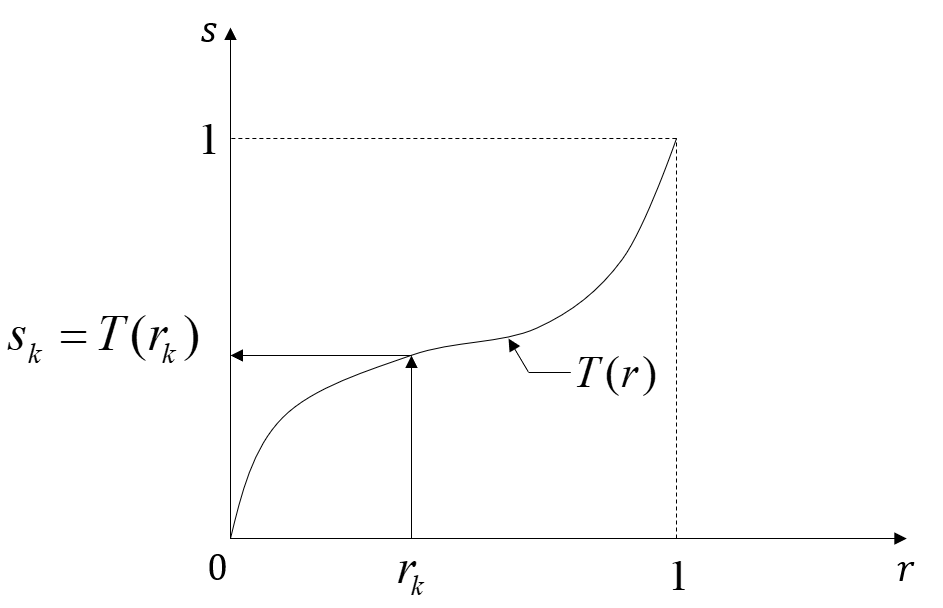


图2-2 直方图均衡变换函数

由图2-2可发现灰度变换在和区间内，像素点个数不变，所以有：

 (2-2)

当，，略去下标 ，有 。由于，，则 ，最终的到直方图均衡化的灰度变换函数为：

 (2-3)

它是原始图像灰度的累积分部函数（Cumulative Density Function，CDF）。 离散数字图像直方图均衡化操作算法步骤如下：

第一步，计算原图像直方图：

 (2-4)

其中为归一化输入图像灰度级。

第二步，统计直方图累积分布曲线

 (2-5)

第三步，以累积分布函数为变换函数，实施灰度变换。根据以上步骤获取的sk，将输入与输出图像灰度级的关系进行一一对应，亦即对累积分布函数sk重定位（与已归一化的灰度等级rk进行比较，寻找最接近的一个作为原灰度级k变换后的新灰度级）。

2.3神经网络理论

2.3.1机器学习与神经网络

学习行为是人类具有的重要智能行为，机器学习正是一门关于理解与研究学习的内在机制、试图建立能够通过学习自动提高自身水平的计算机程序的理论方法的学科。[20]机器学习是计算机在算法的指导下从已知数据中进行数据分析的一个过程，在不断地优化和调整中提升性能，最终具有一定的泛化能力，在遇到相似的问题或场景可以进行迁移。近年来，机器学习理论在诸多应用领域尤其工业界中得到成功的应用与发展，在智能领域研究中具有十分重要的地位。根据学习方式的不同，可以分为监督式学习，无监督式学习[21]，强化学习等几种类型。实践中应用较多的主要是前两种，在深度学习中，数据输入常常带有标签属于监督学习。神经网络是机器学习研究领域的一类重要算法，神经网络有悠久的发展历史，从20世纪40年代起，对神经网络的研究就已经展开。相关理论研究以M-P模型为里程碑，1957年，Rosenblatt在对M-P模型的长期研究基础上[22]，提出了感知器(Perceptron)模型，是第一个真正意义上具有现代神经网络的基本原则的神经网络。1982年，出现了一种离散型神经网络Hopfield[23]。 Hinton和Sejnowski于1986年首次提出了多层网络的学习算法，称为Boltzmann机模型[24]。D.E.Rumelhart等人同年提出了适用于多层神经网络中进行权值修正的反向传播学习算法，解决了多层前向神经网络中的学习问题。90年代初，Vapnik等人提出了支持向量机(SVM)和VC(Vapnik-Chervonenkis)维数的概念[25]。深度学习是神经网络2006年后重获新生的名称，是机器学习研究发展的新领域。

2.3.2神经元模型

在生物学上，神经元通常由细胞体，细胞核，树突和轴突构成。树突用来接收其他神经元传导过来的信号，其中一个神经元可以有多个树突；细胞核是神经元中的核心，用以处理所有的传入信号；轴突是输出信号的单元，它有很多个轴突末梢，可以给其它神经元的树突传递信号。

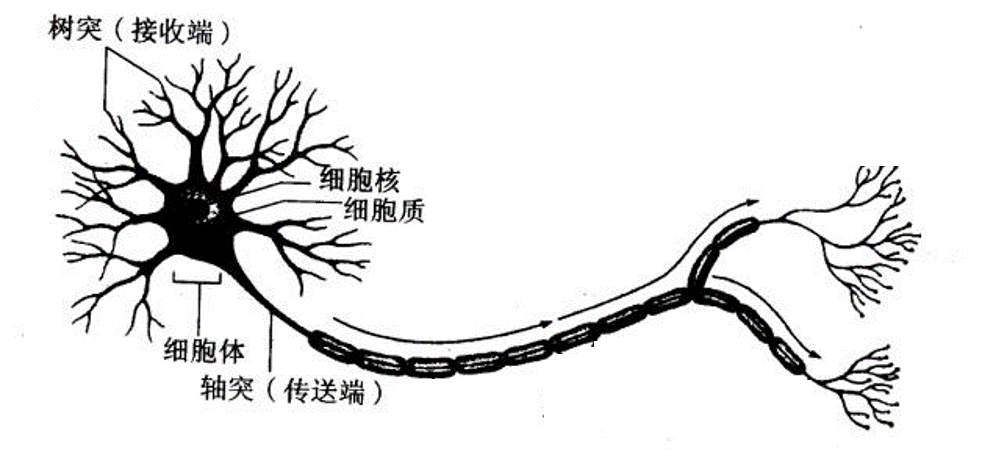


图2-3 生物神经元[13]

神经网络中最基本的成分是神经元模型，生物神经网络中，每个神经元与其他神经元连接，当其中一个神经元“兴奋”时，就会向相连接的神经元发送化学物质，这种化学物质称为神经递质，从而改变相连接神经元内的电位；如果某个神经元的电位超过某一个阈值，那么它就将会被激活，向其他神经元发送化学物质。

20世纪40年代初，由McCulloch和Pitts[40]提出的如图 2-4所示的“M-P神经元模型”，在该模型中，神经元接收到来自n个神经元传递过来的输入信号，传递过来的输入信号通过带权值的连接来传递，神经元收到的总输入将于神经元阈值进行比较，然后通过激活函数处理来产生神经元的输出。

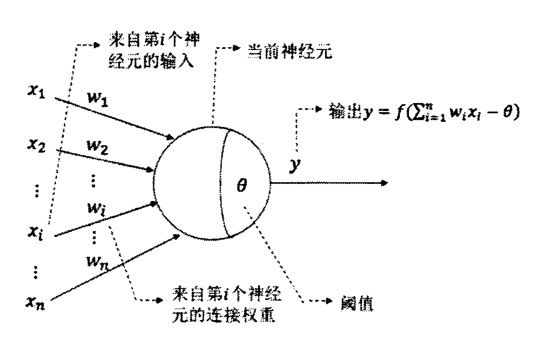


图 2-4 M-P神经元模型[20]

在理想情况下，激活函数是输出值为“0”或“1”的阶跃函数，它将输入值映射为输出值“0”和“1”，“1”对应于神经元兴奋，“0”对应于神经元抑制。由于阶跃函数具有不连续性，不光滑等特性，因此在实际操作中，常用Sigmoid函数作为其激活函数。典型的Sigmoid函数如图2-5中所示，它可以把在比较大范围内变化的输入值压缩到(0 , 1)输出值范围内。

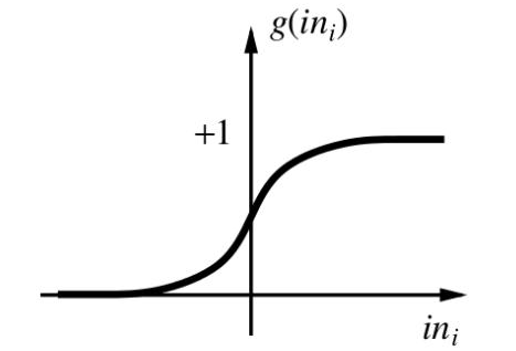


图 2-5 Sigmoid函数

Sigmoid函数的计算公式如下式(2-6)所示：

(2-6)



把多个如上述模型的神经元按一定的层次结构连接便得到了神经网络。在计算机科学的角度来说，可以先不考虑上述所提出的人工神经网络是否真的模拟了生物神经网络，只需将一个神经网络视为包含了多个参数的数学模型即可，而这个模型是由若干个函数相互嵌套得到的。

2.3.3感知机与多层网络模型

感知机由两层神经元组成，输入层接收到外界的输入信号之后，传递给输出层，输出层为M-P神经元，也称为“阈值逻辑单元”。

一般地来说，给定训练数据集，权重 以及阈值可以通过学习来得到，阈值可看做一个固定输入为 -1.0的“哑结点”，所对应的权重为，这样的权重和阈值的学习便可简化为对权重的学习。感知机学习规则非常地简单，对训练样例(x,y)，若当感知机的输出为，则感知机权重将做如下调整：

 (2-7)

 (2-8)

其中称为学习率。从式(2-8)中可以看出，如果感知机对训练样例(x,y)预测正确，即，则感知机不发生变化，否则将根据错误的成都进行权重调整。

一般来说，常见的神经网络是层级结构，形如图2-6所示，每层神经元和下一层神经元全互连，同层神经元之间不存在相互连接，跨层的神经元之间也不存在相互连接。如图2-6所示的神经网络成为多层前馈神经网络，其中输入层神经元接收外界的输入，隐含层神经元和输出层神经元对信号加工处理，输出层神经元输出最终结果。只要网络中存在隐含层，便可以称为多层网络，其中，隐含层可以有多层。我们深入地揭开神经网络的神秘面纱，根据训练数据来调整神经元之间的“权重”以及每个功能神经元的“阈值”，这就是神经网络的学习过程。神经网络的学习过程，即在学习中找到合适的权重和阈值的过程。

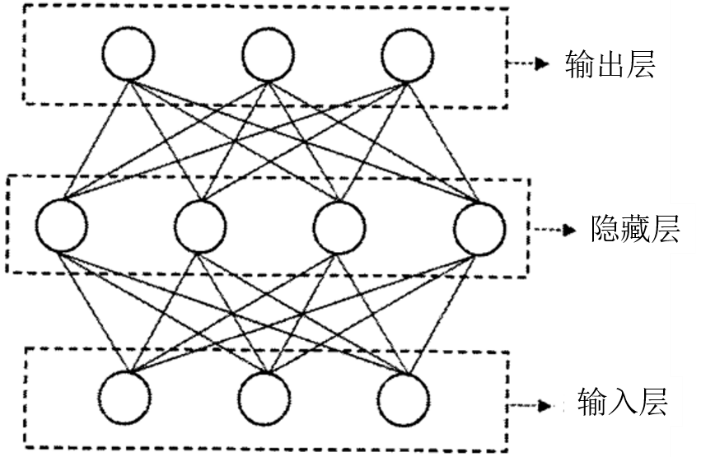


图 2-6 多层前馈神经网络

2.3.4误差逆传播算法

相对单层感知机来说，多层网络的学习能力要强得多。要训练多层网络，要比训练单层网络需要更强大的算法。误差逆传播算法（Back Propagation，BP）是多层网络中典型的代表。BP网络是1986年由Rumblhart 和McCelland[40]为首的科学家小组提出，是一种按误差逆传播算法训练的多层前馈网络，是目前应用最广泛的神经网络模型之一。BP算法除了应用于多层前馈神经网络，还可以用于其他神经网络，如递归神经网络。一般来说的“BP网络”指的是用BP算法训练的多层前馈神经网络。

给定训练集，，，输入实例由个属性描述，输出维的实值向量。如图 2-7所示，我们设计了一个有个输入神经元，个输出神经元，q个隐含层神经元的多层前馈网络结构，其中用表示输出层第j个神经元的阈值，用表示隐含层第个神经元的阈值。输入层第i个神经元之间的连接权值为。隐含层第个神经元接收到的输入为，输出层第个神经元接收到的输入为，其中为隐含层第个神经元的输出。假设隐含层和输出层神经元都用图 2-2中的Sigmoid函数。

对训练例，假定神经网络的输出为，即

 (2-9)

则网络在上的均方误差为

 (2-10)

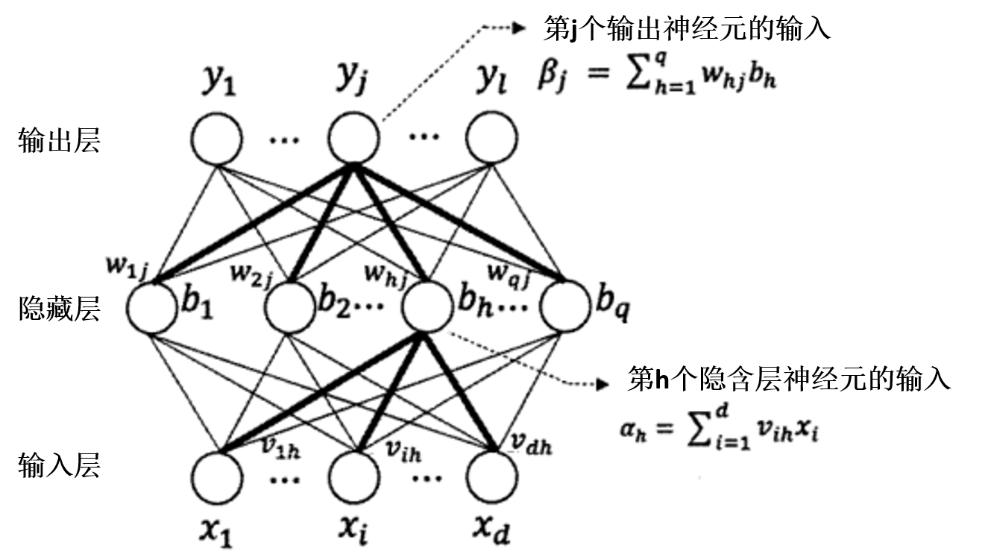


图2-7 BP网络及算法中的变量符号

在2-7所示的网络，有个参数需要确定：输入层到隐含层的个权值，隐含层到输出层的个权值，q个隐含层神经元阈值，l个输出层神经元阈值。BP算法是一种迭代学习算法，通过每一次的迭代，所使用的广义感知机学习规则会对参数进行更新估计，即与式(2-7)类似，任意参数v的更新估计式为

 (2-11)

以图2-7中隐含层到输出层的连接权值为例进行推倒：

基于梯度下降策略的BP算法，以目标函数的负梯度方向对参数进行调整。对式(2-10)的误差，给定学习率，有

 (2-12)

注意到先影响到第个输出层神经元的输入值，再影响到其输出值，然后影响到，有

 (2-13)

根据的定义，显然有

 (2-14)

Sigmoid函数有一个很好的性质，如下：

 (2-15)

根据式（2-9）和式（2-10），有





 (2-16)

将式（2-14）和式（2-16）带入式（2-13）就得到了BP算法那关于的更新公式：

 (2-17)

类似可以得到：

 (2-18)

 (2-19)

 (2-20)

式(2-19)和式(2-20)中







 (2-21)

我们通过学习率来控制着算法中每一次迭代中的更新步长，步长如果太大，会容易震荡，如果步长太小则会使得收敛速度太过缓慢，所以控制好算法中的更新迭代步长，能够更快更好地得到合适的结果。

2.4卷积神经网络理论

2.4.1概述

卷积神经网络（Convolutional Neural Network，CNN），是一种[前馈神经网络](https://zh.wikipedia.org/wiki/%E5%89%8D%E9%A6%88%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C)，它的人工神经元可以响应一部分覆盖范围内的周围单元，对于大型图像处理有出色表现。

卷积神经网络是近年发展起来，并引起广泛重视的一种高效识别方法。20世纪60年代，Hubel和Wiesel在研究猫脑皮层中用于局部敏感和方向选择的神经元时发现其独特的网络结构可以有效地降低反馈神经网络的复杂性，继而提出了卷积神经网络。现在，CNN已经成为众多科学领域的研究热点之一，特别是在模式分类领域，由于该网络避免了对图像的复杂前期预处理，可以直接输入原始图像，因而得到了更为广泛的应用。K. Fukushima[41]在1980年提出的新识别机是卷积神经网络的第一个实现网络。随后，更多的科研工作者对该网络进行了改进。其中，具有代表性的研究成果是Alexander和Taylor提出的“改进认知机”，该方法综合了各种改进方法的优点并避免了耗时的误差反向传播。

卷积神经网络由一个或多个卷积层（Convolutions）组成，同时也包括关联权重和[池化](https://zh.wikipedia.org/w/index.php?title=%E6%B1%A0%E5%8C%96&action=edit&redlink=1)层（pooling layer）。这一结构使得卷积神经网络能够利用输入数据的二维结构。层与层之间特征图的尺度递减，每层所含的特征图的数目或维度递增，因此，随着层数的加深，建立起了更高的空间，拥有了数据结构不变性，抽象出了更多的特征信息。卷积层和池化层，是实现卷积神经网络特征提取功能的核心模块。根据任务的不同，输出层的分类器可以采用Softmax回归，Logistic回归或者支持向量机等形式。卷积神经网络中的学习过程，更多地使用BP算法，如果最终的分类结果与预期结果不同，则将该误差以某种形式反向地，向输入层一层一层地传播。在传播过程中，将误差按照一定规则给每层的每个神经单元，每个神经单元会根据这些误差对各个单元的权值进行更新。重复进行这样的前向传播和反向传播的修正，直至最终的误差减少到预设的可接受范围或者重复的次数达到预定的次数。

2.4.2卷积神经网络的基本结构

卷积神经网络相对于传统的多层前馈网络，具有以下特点，使得其解决了网络参数过多问题，提升了学习效率以及对平移，缩放，倾斜等形变具有高度不变性。

（1）稀疏连接：CNN通过加强神经网络中相邻层之间节点的局部连接模式来挖掘自然图像的空间局部关联信息。第m层节点的输入只是第m−1层节点的一部分，这些节点具有空间相邻的视觉感受野。假设m−1层为视网膜输入层，它上面的m层每个神经元节点都具有宽度为3的感受野，即分别与m−1层的3个相邻的节点相连。m+1层与m层有着相似的连接属性，即m+1层的节点仍具有宽度为3的感受野，但是对于m−1层感受野会更大。每个神经元对其感受野外的变化没有响应，这种结构保证了训练好的每个滤波器只对空间局部输入模式产生最强烈的响应。多层神经元堆积形成了非线性滤波器，并逐层递增的具有了全局性。比如，第m+1层能够对宽度为5的非线性特征进行编码。如图2-8所示。

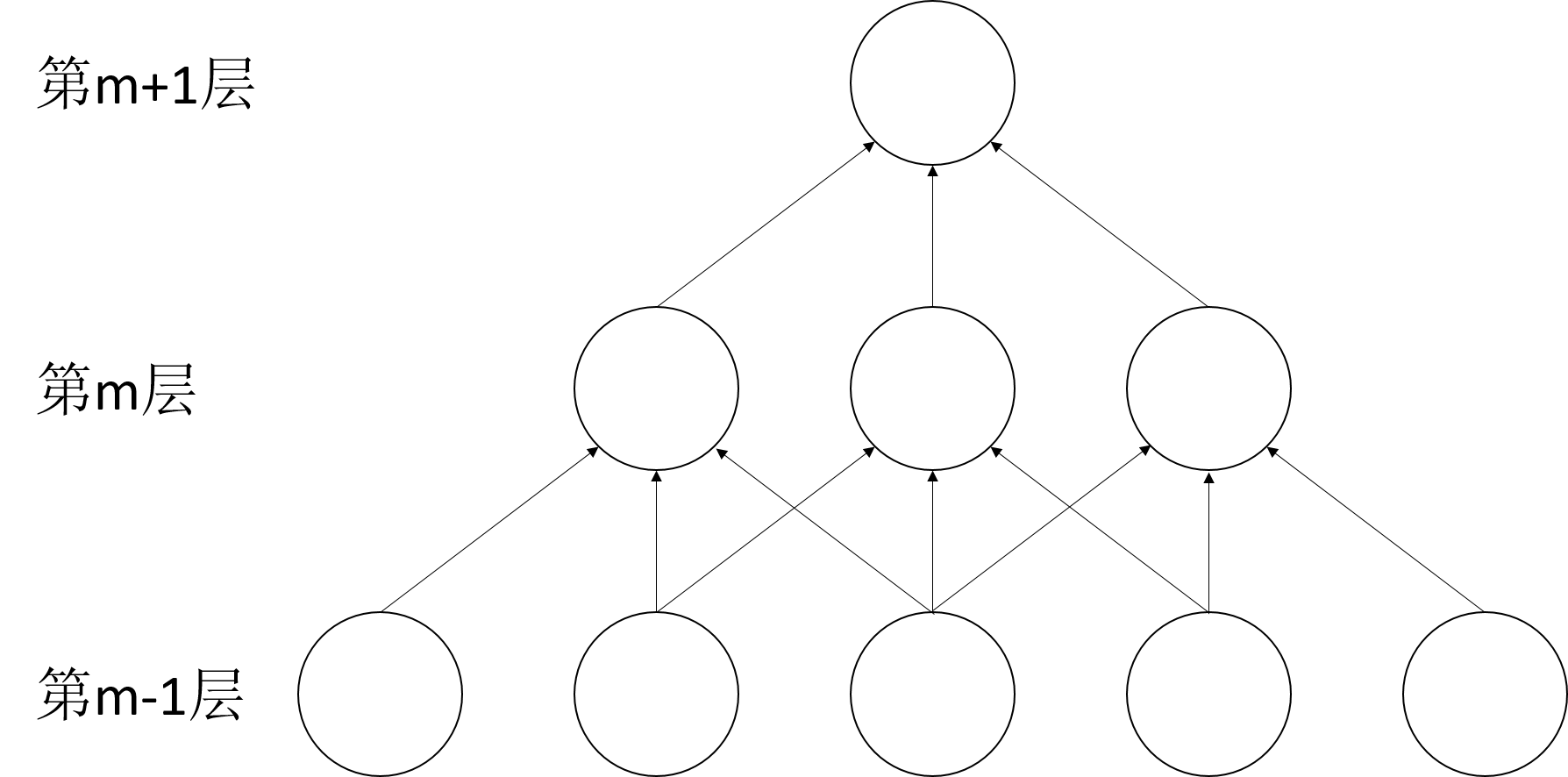


图 2-8 卷积神经网络的局部连接方式

（2）参数共享：在卷积神经网络中，卷积层的每一个滤波器重复地作用于整个感受野，对输入图像进行卷积操作，抽取出图像的局部特征，形成特征图。每一个卷积滤波器共享相同的参数，包括相同的权重矩阵和偏置项。如图2-9所示。

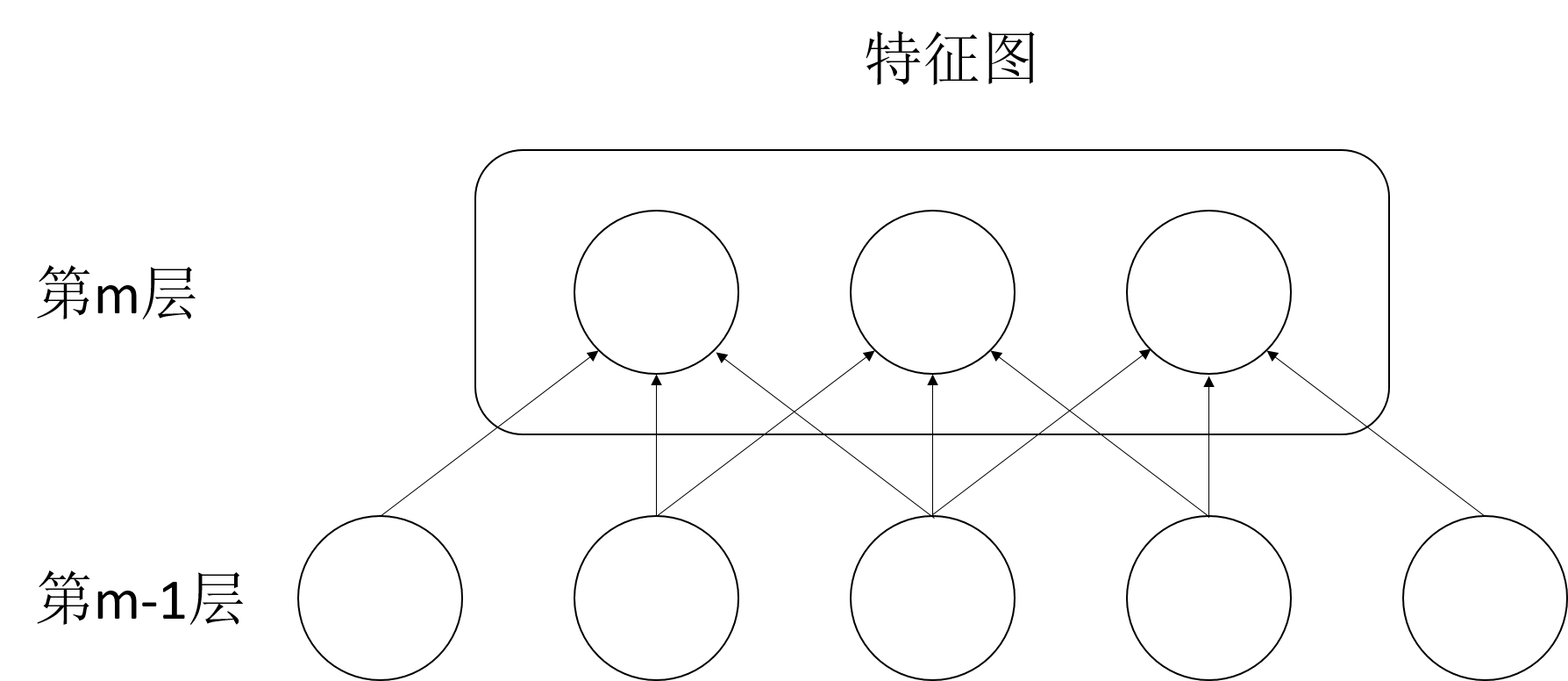


图 2-9 卷积神经网络的参数共享

2.4.3 卷积

卷积（Convolution）是在信号与线性系统的基础上出现的，卷积操作是人为定义的一种运算，作为工具使正在研究的问题变得简单，数学里很多地方都用到卷积，但不同地方卷积的具体定义也不尽相同。具体到图像处理，广泛应用于图像滤波。在卷积网络中用于特征学习。卷积核的设计主要考虑卷积核的大小、数目，滑动步长和边界处理等问题。

卷积核的数目就是对上层输入进行卷积操作后得到的特征映射的个数。一般认为，得到的特征图越多，该层对上一网络层传入的输入的刻画越准确，特征学习越精确，但卷积核过多时，网络参数增加，网络计算的复杂度也随之增加，同时也可能会导致过拟合现象的发生。针对具体问题，卷积核数目的选择主要还是具体分析。同理，卷积核的大小即为特征映射的大小，卷积核大小和滑动步长决定了特征图的个数。越小的窗口滑动步长，两次滑动取得的数据块重叠的部分就越多，相应的窗口停留次数就多，运算量大一些;窗口滑动步长越大，重叠部分越少，窗口停留次数就少运算量小。但在一定程度上，数据信息不如步长小时丰富。因此，通过某种方式将不同步长的卷积合并起来，可以节省卷积层的空间复杂度。图2-11中，黄色区域是一个大小的卷积核，以步长大小为1，在一张的图片上由左至右，从上往下进行卷积的一个示意。

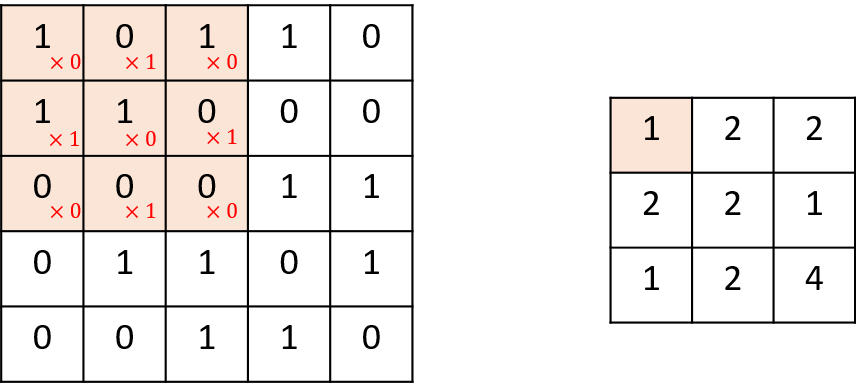


图2-10 卷积过程

当对一个彩色图像进行卷积时，卷积核需要分别在RGB三个通道上分别进行操作，然后对结果进行求和，经激活函数处理后得到特征映射。

2.4.4 池化

池化（pooling）这是一种非线性的降采样方法，通常将输入特征图中每一个不同的n×n区域的元素来做池化操作。最大值池化（max pooling），随机池化（stochastic pooling），平局池化（average pooling）等池化方法，是目前比较常用的几种方法。通过卷积在获取图像的卷积特征以后，通过池化操作，主要目的是使用下采样的方式对已经获取的卷及特征进行降维。如果选择特征图中的连续范围作为池化区域，那么经过池化后的特征就具有了一定的平移不变性。这意味着，即使输入图像经历了一个微小的平移，但经过卷积和池化后，依然能够提取出相同的特征图，这样来看的话，有利于特征分类。

通过卷积获得特征后，往往希望对得到的特征进行分类，理论上可以将得到的特征解析出来关联到一个分类器上，但计算量比较大。以最大池化为例，对一张图像进行池化操作如图2-11所示。

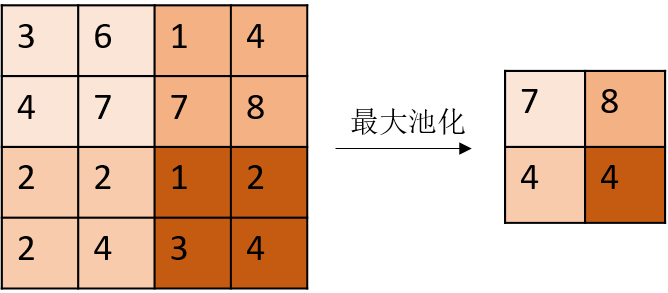


图2-11 池化操作中的最大池化操作

最大池化即在的窗口中选择最大的值作为输出，平均池化时取窗口覆盖所有值的平均值，随机池化根据概率矩阵选择输出值，概率值越大被选中的概率越大，输出不再始终是最大值。对图像卷积操作后获取的特征存在“不变性”，即这种特征在图像中是多区域分布，所以通过对局部的特征进行聚合统计来描述较大的特征，经验显示最大池化法效果可能教好。池化主要还是仿照人的视觉系统进行降维，缩小卷积特征维数，同时不容易过拟合，减小分类器的设计负担。

2.4.5 卷积神经网络常用的训练方法

训练深度网络的常规方法是BP算法指导下的随机梯度下降。该方法通常行之有效，但对于缺乏经验的使用者而言，失效的情况却更多一些。网络的一般训练过程是这样的：(1) 初始化参数空间W (2)输入数据，开始前向传播(3) 返回损失，反向传播各层的梯度并更新权值，重复(2)(3)直到结果收敛。但是通常情况下，如果初始参数稍微偏大或者偏小，在级联结构的作用下，总会出现两个问题，一个是梯度爆炸，一个是梯度消失。所以，深度网络需要更合理的初始化策略，更合适的非线性激活函数，以及选择学习率的更新策略。

同时当网络模型过于复杂，而训练数据不足时，往往会产生过拟合的问题。这时也需要一些策略手段来辅助模型的训练，比如说Dropout等正则化方法。

假设我们想训练一个网络，去判断一个物体是不是叶子，如果网络将一些叶子的锯齿状边缘也学习到，当该物体只有出现锯齿状边缘时，网络才会判定该物体是叶子。这种情况被称之为过拟合。为了解决模型过拟合的问题，Hinton[42]提出来了Dropout，Dropout是一个同正则化完全不同的技术，它不会修改代价函数而是修改深度网络本身。Dropout在训练模型时，每次随机的让模型中某些节点之间的连接断掉，也就是让这些权值不起作用。其操作示意图如下：



图2-12 Dropout操作示意图

仿照Dropout的思想，IJ Goodfellow [43]提出了Maxout。对于Dropout来说是使每个输入神经元以一定的概率不工作，而Maxout更加极端，是对上一层的多个特征映射跨通道去最大值作为输出。Maxout层在网络中，可以被看做是网络的激活函数，如图2-13所示。

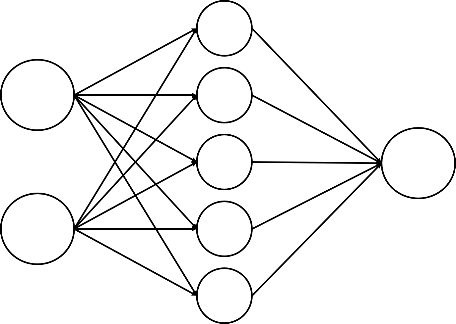


图2-13 Maxout操作示意图

如果使用Maxout作为激活函数，则其隐含层节点的输出表达式为：

 (2-22)

其中为三维向量，大小为，表示输入层节点的个数，表示隐含层节点的个数，表示每个隐含层节点对应了个“隐隐含层”节点，这个“隐隐含层”节点都是线性输出的，Maxout的每个节点作用就是去掉这个“隐隐含层”及保留输出值中的最大值。Maxout的拟合能力是非常强的，可以拟合任何的凸函数。

第三章 糖尿病视网膜病变分期算法

3.1引言

本次实验的目的是将眼底图像进行分类，根据病灶信息进行特征提取将其分为五期，通过卷积神经网络进行特征提取和分类。

具体方法分为下列个步骤（1）对眼底图像进行预处理，通过均衡图片质量和调整图片尺寸达到压缩数据大小的目的。（2）通过对容量较少的类别进行过采样解决类别不平衡的问题。（3）使用较小的图片训练较小的网络，以初始化网络的参数。（4）分别训练两个结构不同的卷积神经网络CNN（5）最终的结果由两个CNN集成学习的结果得到。

3.2预处理

3.2.1 尺寸调整

由于图像尺寸太大，使用卷积神经网络将带来巨大的计算量。算法选择去掉背景，以图像中的圆形眼底图片为边界进行裁剪，取中心部分。调整分辨率分别为128x128,256x256,512x512，用于之后的训练。除此以外没有采用任何其他图像(预)处理步骤。

3.2.2 数据集扩充

采用的数据集扩充方法包括镜面变换、随机比例拉伸、随机角度旋转、上下翻折以及色彩增强。同时还在每个图像通道(RGB)进行中心化和比例化处理，以保障其在训练集上具有零均值和单位方差。数据增强后对于输入大小为128/256/512像素的图像，其输出为112/224/448像素。

3.2.3 重采样

为了解决数据集中类别不平衡的问题。在实验中使用了以下的重采样策略：首先从所有类中抽取同样多的样本，然后逐渐减少对稀有类的过采样。设w0为初始重采样权值和wf终止重采样权值。在第t轮训练中，分期为0级到4级的重采样权值由式（3-1）给出

 (3-1)

其中，设 = 0.975, ≈ (1.36,14.4,6.64,1.36,14.4)， =(1,2,2,2,2) 。

3.3网络模型

本文中的卷积神经网络的网络结构如图3-1所示。我们训练了两个不同结构的卷积神经网络网络NetA和NetB。

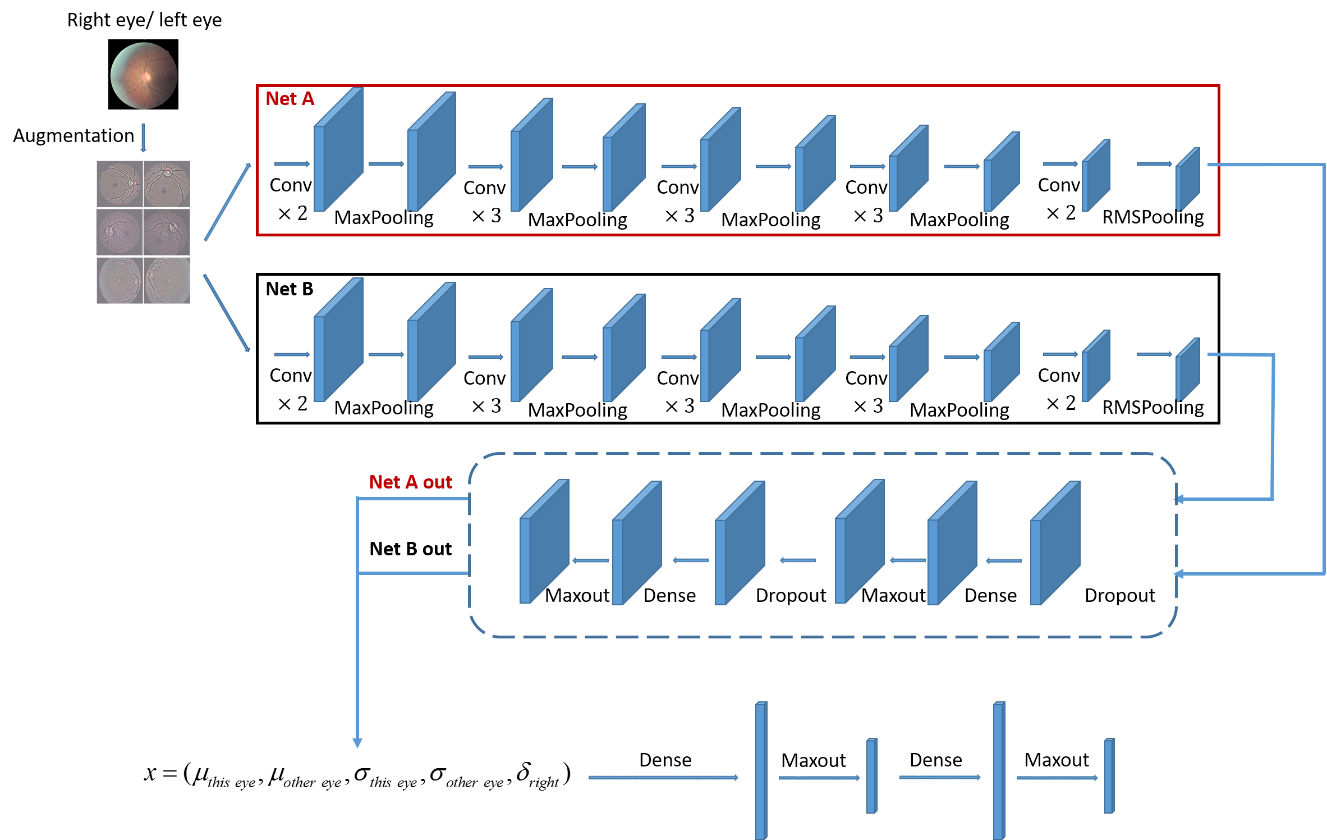


图3-1 全卷积网络结构示意

第一阶段：将一只眼的眼底图像预处理后的数据分别输入具有不同结构的卷积神经网络Net A和Net B。

第二阶段：Net A和Net B在前20-25层具有相同网络结构：将RMSPooling后的结果通过Dropout正则化。

第三阶段：将Net A和Net B集成学习得到的特征这只眼的特征与同样方法得到的另一只眼的特征作为输入，经过两个全连接层和最大值输出层得到分类结果。

3.4网络初始化和预训练

由于网络层数过深，本算法没有从头开始训练庞大的卷积网络。而是采用以下方法进行参数的初始化和网络预训练：

（1）首先使用10%的训练集作为验证集

（2）先使用正交法初始化权值和偏置，并用128x128像素图片训练1-11 和 20-25层。

（3）用256x256像素图片训练1-15 和 20-25层。初始时，使用上一次训练得到的权重作为1-11层的权重。

（4）用512x512像素图片训练全部层。初始时，使用上一次训练得到的权重作为1-15层的权重。

（5）最后得到三组权重（weights）：正确率最高的一组权值，Kappa系数最高的一组权值和最后得到的一组权值。

3.5特征抽取

为了提高抽取特征的质量，对于每个图像使用50种不同的数据集扩充方法得到一组输入，从第二阶段卷积网络的最后一个全连接层中提取特征，并将每个特征的均值和标准偏差作为之后第三阶段全连接层的输入。

3.6左右眼混合

对于每只眼睛（或每个病人），利用上面说到的从A网络和B网络的RMSPool layer中得到的输出（均值(µ)和标准偏差(σ)）作为混合网络的输入。由式(3-2)给出：

 (3-2)

其中为左眼的一个指示变量，表明是否为左眼。我们将所有的特性标准化，使其具有零均值和单位方差，并将它们作为第三阶段全连接层的输入。

第四章 实验与讨论

4.1数据集

4.1.1 Messidor 数据集

该数据集包含540幅正常图像和660幅异常图像，图像有3种分辨率：1 440×960、2 240×1 488和2 304×1 536像素，图像为 TIF 格式。根据微血管瘤、出血、硬性渗出物和新生血管等病灶来对每张图像标注从R0到R3的病变等级，R0表示无病变图像，R1和R2分别表示轻度和重度非增殖型视网膜病变，R3表示最严重的增殖型病变。其中，R0等级的图像有540张，占46%；R1等级的图像有153张，占12.75%；R2等级图像有247张，占20.58%；R3等级图像有260张，占21.67% 。

4.1.2 ROC数据集

该数据集是2009年爱荷华洲大学公开的一个用于DR筛查竞赛的眼底图库，包含 100张DR病变和微血管瘤的眼底图像。其中，训练组50张，测试组50张，使用拓普康 NW100 或者佳能CR5-45NM 相机在45°视场下采集的图像，分辨率有：768×576、1 058×1 061和1 389×1 383像素，图像为JPEG格式。

4.1.3 DIARETDB1 数据集

该数据集包含89张眼底图像，分辨率为1 500×1 152像素，包含27张正常图像，7张DR轻度图像，28张DR中度图像，27张DR重度图像。

4.1.4 DR1 数据集

该数据集包含687幅正常图像和327幅异常图像，其中245幅含有硬性渗出物，191幅含有微血管瘤或出血，两种异常病灶都有的图像有109幅．图像分辨率为 640×480像素，均为彩色图像，有TIF和JPG两种图像格式。

4.1.5 Kaggle-DR数据集

本文采用的次数据集进行训练和测试。该数据集是Kaggle在2015年的 DR竞赛中的数据集，数据集中眼底图像有多种分辨率，每一个被试患者有左右眼两张图像，且两张图像相同分辨率。测试数据35 126张图像，训练数据53 576张图像。图像按病变程度分为5类，数据集已标注图像病变程度等级。

4.2衡量指标

4.2.1 Kappa系数

Kappa系数是一种衡量分类精度的指标。它是通过把所有地表真实分类中的[像元](https://baike.baidu.com/item/%E5%83%8F%E5%85%83)总数（）乘以混淆矩阵对角线（）的和，再减去某一类地表真实像元总数与该类中被分类像元总数之积对所有类别求和的结果，再除以总像元数的[平方](https://baike.baidu.com/item/%E5%B9%B3%E6%96%B9)减去某一类地表真实像元总数与该类中被分类像元总数之积对所有类别求和的结果所得到的 。

用表示个被划分为类的类样本个数，为类样本个数，为被划分为类的样本个数。

 （4-1）

其中是每一类正确分类的样本数量之和除以总样本数，也就是总体分类精度。另外，假设每一类的真实样本个数分别为，而预测出来的每一类的样本个数分别为，总样本个数为n则有：

 （4-2）

Kappa计算结果为-1~1，但通常Kappa是落在 0~1 间，可分为五组来表示不同级别的一致性：0.0~0.20极低的一致性(slight)、0.21~0.40一般的一致性(fair)、0.41~0.60 中等的一致性(moderate)、0.61~0.80 高度的一致性(substantial)和0.81~1几乎完全一致(almost perfect)。

4.2.2 加权Kappa系数

本文研究的DR分期问题属于有序分类变量。有序分类变量是指其取值的各类别之间存在着程度上的差别，给人以“半定量”的感觉，因此也称为等级变量。是根据取值特征而分类的一种定性变量。对于这种变量，常使用加权Kappa系数。其计算公式为：

 (4-3)

其中

 (4-4)

 (4-5)

 (4-6)

可知传统Kappa统计量只从样本个数角度度量一致性,在有序问题中分类结果之间存在大小关系,只计算个数会对一致性判断结果造成偏差。加权Kappa统计量通过引入权重wst刻画序差距,序差距越小,一致性越高，权重wst就越大。特别地，二类问题中，加权Kappa与不加权Kappa是等价的。[27]

4.3实验设计

对于网络结构中各参数的设定如表4-1所示。

表4-1 网络结构参数

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 层数 | net A | | | | net B | | | |
|  |  | units | filter | stride | size | filter | stride | size |
| 1 | Input |  |  |  | 448 | 4 |  | 448 |
| 2 | Conv | 32 | 5 | 2 | 224 | 4 | 2 | 224 |
| 3 | Conv | 32 | 3 |  | 224 | 4 |  | 225 |
| 4 | MaxPool |  | 3 | 2 | 111 | 3 | 2 | 112 |
| 5 | Conv | 64 | 5 | 2 | 56 | 4 | 2 | 56 |
| 6 | Conv | 64 | 3 |  | 56 | 4 |  | 57 |
| 7 | Conv | 64 | 3 |  | 56 | 4 |  | 56 |
| 8 | MaxPool |  | 3 | 2 | 27 | 3 | 2 | 27 |
| 9 | Conv | 128 | 3 |  | 27 | 4 |  | 28 |
| 10 | Conv | 128 | 3 |  | 27 | 4 |  | 27 |
| 11 | Conv | 128 | 3 |  | 27 | 4 |  | 28 |
| 12 | MaxPool |  | 3 | 2 | 13 | 3 | 2 | 13 |
| 13 | Conv | 256 | 3 |  | 13 | 4 |  | 14 |
| 14 | Conv | 256 | 3 |  | 13 | 4 |  | 13 |
| 15 | Conv | 256 | 3 |  | 13 | 4 |  | 14 |
| 16 | MaxPool |  | 3 | 2 | 6 | 3 | 2 | 6 |
| 17 | Conv | 512 | 3 |  | 6 | 4 |  | 5 |
| 18 | Conv | 512 | 3 |  | 6 | n/a |  | n/a |
| 19 | RMSPool |  | 3 | 3 | 2 | 4 | 2 | 2 |
| 20 | Dropout |  |  |  |  |  |  |  |
| 21 | Dense | 1024 |  |  |  |  |  |  |
| 22 | Maxout | 512 |  |  |  |  |  |  |
| 23 | Droupout |  |  |  |  |  |  |  |
| 24 | Dense | 1024 |  |  |  |  |  |  |
| 25 | Maxout | 512 |  |  |  |  |  |  |

另外，在处理非线性的问题上，采取的方法是在每一个卷积和全连通层后使用漏极(0.01)整流单元。该网络采用牛顿动量优化方法训练超过250次的固定轮数（epoch）。

对于输入为256和128像素的网络，均进行200轮训练，其中每一层使用L2进行正则化，权值衰减因子为0.0005。

另外，将这个问题视为一个回归问题。对于256像素网络来说，目标均方误差和阈值分别为0.5和1.5；对于128像素网络，其目标均方误差和阈值为2.5,3.5.

我们将所有的特性标准化，使其具有零均值和单位方差，并使用Net A和Net B所示的体系结构完全连接的网络。混合网络的训练过程总结如下：

（1）每个全连接层后连接一个整流非线性单元。

（2）第1层采用因子为2e-5的L1正则化，每一层均采用权值衰减因子为0.005的L2正则化。

（3）以固定100轮以上训练次数用Adam优化算法更新。

使用训练集中10%作为验证集，得到如表4-3所示的混淆矩阵。

表4-3所得混淆矩阵

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2355 | 193 | 25 | 5 | 0 |
| 119 | 105 | 34 | 0 | 0 |
| 62 | 125 | 256 | 69 | 2 |
| 2 | 1 | 24 | 66 | 0 |
| 0 | 2 | 10 | 30 | 29 |

正确率为92%，加权Kappa系数为0.844。

4.4时间复杂度

网络训练时间为7天(英特尔Xeon E5-2678V3，4个NVIDIA Titan Xp GPU，128 GB内存，120 GB 固态硬盘)。

实际上，经实验发现，如果第一阶段只使用B网络进行训练，只需要3天时间，且加权Kappa系数可以达到0.842。

4.5结果比较

表4-4 几种公开眼底图像数据集的DR分类方法比较

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集名称 | 数据描述 | 主要方法 | 正确率/% | 加权Kappa |
| Messidor | 正常数据(R0)和  异常数据（R1、R2和R3）进行2分类，10倍交叉验证测试数据 | 联合图像水平、损伤区域和解剖结构特征进行特征提取 | 90.00 | 0.910 |
| R0和R1数据共693 张图像进行2 分类，10 倍交叉验证测试数据 | 联合图像水平、损伤区域和解剖结构特征进行特征提取 | 90.00 | 0.900 |
| 正常数据(R0)和  异常数据（R1、R2和R3）进行2分类，10倍交叉验证测试数据 | 卷积神经网络，5层网络结构，使用dropout  策略和maxout激活函数等参数 | 96.00 | 0.960 |
| R0和R1数据共693 张图像进行2 分类，10 倍交叉验证测试数据 | 卷积神经网络，5 层网络，使用dropout 和 maxout 激活函数等 | 95.40 | 0.950 |
| 出血点检测，测试数据包含348张图像 | 均衡化对比增强，区域生长 | 98.56 | — |
| Diaretdb1 | 检测硬性渗出物，基于像素点水平测试实验结果 | 综合强度、几何、纹理、病灶区域特征和空间关系5大类特征，使用随机森林检测 | — | 0.890 |
| 使用DIARETDB1中89张图像，检测硬性渗出物，基于像素点计算分类效果 | 数学形态学的高斯尺度空间方法，使用支持向量机分类 | 98.35 | 0.931 |
| 检测微血管瘤，使用 89 张眼底图像训练，正常图像 5 张，异常图像 84 张，基于像素点检测分类结果 | 模糊对比增强预处理，去除血管等结构影响，计算微血管瘤个数 | — | 0.985 |
| 出血点检测，测试图像39张 | 均衡化对比增强，区域生长 | 94.87 | — |
| DR1 | 训练数据包括687张正常图像，245张含硬性渗出物图像，5倍交叉验证 | SIFT 和 SURF 描述子，使用词袋模型对图像进行编码，使用支持向量机分类器 | — | 0.900 |
| 训练数据包括687张正常图像，245张含硬性渗出物图像，5倍交叉验证 | SIFT 和 SURF 描述子，使用词袋模型对图像进行编码，使用支持向量机分类器 | — | 0.900 |
| Kaggle | 选择Kaggle 数据中500张5类图像 | Canny 边缘检测和阈值提取特征 | 85.00 | — |
| 选择Kaggle 数据中500张5类图像 | 卷积神经网络，输入图像大小为41 ×41 像素，9层网络结  构 | — | 0.821 |
| **Kaggle** | **选择Kaggle 数据中500张5类图像** | **卷积神经网络，22层**  **（本文方法）** | **92.00** | **0.844** |

4.6本章小结

这一竞赛的数据集中的图像包含了一些重要的特征，如微型动脉瘤，它在眼底图像中非常小。但是使用相当大的输入图像对于特征提取十分重要。在10%验证集上，本文的卷积神经网络输入为256像素图像时竞赛中评分达到0.70, 输入为256像素图像时竞赛中评分达到0.80。当使用768像素的图像，这进一步增加了kappa到0.81。但是计算量将会增大许多。

同时，当整流器的漏率增大时，计算量将会减小。但是当使用漏率很高的整流器时，从头开始训练大网络时，分数会下降。因为性能下降只会在混合后出现，而不是在训练中，即整流单元的变化或训练过程的变化(从较小的网络到训练的初始化)导致了性能的下降。解决方案是使用两个卷积网络结构和三组不同的训练权重进行平均。

可以看出，本文使用的方法具有较高的正确率和一致性。

第五章 总结与展望

5.1总结

本项目的研究内容，实现了DR自动分期，并解决了以下关键科学问题：

（1）数据集偏小是目前围绕医学影像数据集的一个普遍问题，目前的模型很少出现过拟合问题，一旦能够获得更多的数据，必定能够得到更好的结果．事实上，受制于成本、时间等因素，完全依赖有标注数据是不现实的。且Kaggle竞赛还有5万多张未标记数据，本文构建了一个半监督学习模型、充分利用未标记数据。

（2）DR检测任务属于细粒度图像分类(fine．grained image classification)任务，因为相邻DR级别的眼底图像差异很小，需要经过专业培训的医生才能进行诊断。使用解决通用图像任务的方法，模型的能力会受到很大的限制。针对细粒度图像分类，有一些算法使用了部位标注点等额外的人工标注信息，以便定位关键特征来源。而在DR检测场景下，医生只会给出诊断结果，不会在眼底照片中标记特征的区域和特征的性质。因此，本文仅使用图像级别标注信息进行DR检测，综合使用弱监督学习方法定位特征区域。

（3）本文对糖尿病视网膜病变眼底图像分类进行了实验，将计算机辅助诊断应用到大规模糖尿病视网膜病变筛查中，理论上可以提供一个眼底图像自动分析系统给眼科医生，帮助提高诊断效率和诊断准确率。

5.2展望

本文虽对糖尿病视网膜病变检测相关问题展开了研究， 且取得了一些研究成果，但由于实验环境和研究时间以及个人能力等条件限制，研究还可以更加深入和完善。后期的研究可以从以下几个方面着手：

（1）本文使用基于全局图像的分类方法，较基于局部病灶的分类方法有其独特优势，通过对全局图像提取特征，检测所有异常区域，实现关于眼底图像病灶的规则和特征的综合表示，不考虑各类病灶的结构特点，避免复杂的特征提取工作。但是该方法的可解释性较弱，无法确切说明检测出的病灶区域，而且单一的特征提取方法对多种病灶的描述性不足 。

（2）对整个 DR 检测系统来说，对于眼底图像的预处理，即如何将眼底图像质量评价算法融合到采集设备，形成双向的反馈机制以利于下一步的分割和特征提取将是下一步的研究目标。

（3）由于图像质量的不均衡，在预处理方面应该对图像进行颜色，亮度的均衡化，考虑用此方法提高算法的鲁棒性。

参考文献

1. Diabetic Retinopathy Detection — Kaggle: https://www.kaggle.com/c/diabetic -retinopathy-detection, 2015.
2. Ben Graham∗b. Kaggle Diabetic Retinopathy Detection competition report. Source: b.graham@warwick.ac.uk,2015.
3. M. Antony and S. Brggemann. Kaggle Diabetic Retinopathy Detection: Team o\_O solution, 2015.
4. Alexander Rakhlin. Diabetic Retinopathy detection through integration of Deep Learning classification framework, bioRxiv 225508.
5. DLT Jord, A Valls, D PuigA. Deep Learning Interpretable Classifier for Diabetic Retinopathy Disease Grading, arXiv:1712.08107[cs.LG]

URL:http://arxiv.org/abs/1712.08107.

1. Faust O, Acharya U R, Ng E Y, et al. Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review[J]. Journal of Medical Systems, 2012, 36(1):145.
2. Fong D S, Aiello L, Gardner T W, et al. Diabetic retinopathy.[J]. Diabetes Care, 2003, 26 Suppl 1(1):S99.
3. Wang Z, Yang J. Diabetic Retinopathy Detection via Deep Convolutional Networks for Discriminative Localization and Visual Explanation[J]. 2017.
4. Acharya, U. R., Lim, C. M., Ng, E. Y. K., Chee, C., and Tamura, T., Computer based detection of diabetes retinopathy stages using digital fundus images. J. Eng. Med. 223(H5):545–553, 2009.
5. Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., and Goldbaum, M., Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Trans. Med. Imag. 8(3):263–269, 1989.
6. Hoover, A. D., Kouzanetsova, V., and Goldbaum, M., Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans. Med. Imag. 19(3):203–210, 2000.
7. Hayashi, J., Kunieda, T., Cole, J., Soga, R., Hatanaka, Y., Lu, M.,Hara, T., and Fujita, F., A development of computer-aided diagnosis system using fundus images, Proceeding of the 7th International Conference on Virtual Systems and MultiMedia (VSMM 2001), pp. 429-438, 2001.
8. 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(11).
9. 齐慧君, 黎晓新. 不同分期糖尿病视网膜病变激光治疗预后分析[J]. 中国实用眼科杂志, 2004, 22(10):801-803.
10. 杨玺. 糖尿病视网膜病变的分期与防治[J]. 中国社区医师, 2012(11):8-8.
11. Rocha A, Carvalho T, Jelinek H F, et al. Points of interest and visual dictionaries for automatic retinal lesion detection[J]. IEEE transactions on bio-medical engineering, 2012, 59(8):2244.
12. Pires R, Carvalho T, Spurling G, et al. Automated Multi-Lesion Detection for Referable Diabetic Retinopathy in Indigenous Health Care[J]. Plos One, 2015, 10(6):e0127664.
13. Antal B, Hajdu A. An ensemble-based system for automatic screening of diabetic retinopathy[J]. Knowledge-Based Systems, 2014, 60(2):20-27.
14. Pratt H, Coenen F, Broadbent D M, et al. Convolutional Neural Networks for Diabetic Retinopathy [J]. Procedia Computer Science, 2016, 90:200-205.
15. 周志华. 机器学习[M].清华大学出版社, 2016: 10-16页.
16. Hinton G E. Unsupervised learning: foundations of neural computation[M]. MIT Press 1999:101-102P.
17. 刘建伟，申芳林，罗雄麟.感知器学习算法研究[J].计算机工程，2010,07:190-192页.
18. Karapataki M, Wilde P D. Hopfield network applied to blood vessel detection in Angiograms [J].Medical&Biological Engineering&Computing, 1997, 35(4):428-430P.
19. Hinton, G. E, Sejnowski,T.J.Learning and relearning in Boltzmann machines[M] Parallel distributed processing: explorations in the microstructure of cognition, vol. 1.MIT Press, 1986:45-76P.
20. Vapnik V.SVM method of estimating density, conditional probability,and conditional density[C] IEEE International Symposium on Circuits and Systems, 2000. Proceedings. ISCAS. IEEE, 2000,2(1):749-752P.
21. Datta N S, Dutta H S, Majumder K． Brightness-preservingfuzzy contrast enhancement scheme for the detection andclassification of diabetic retinopathy disease[J].Journalof Medical Imaging, 2016, 3(1): 014502.
22. 张晓琴, 王婕婷, 钱宇华. 一种基于加权kappa增益的有序分类的选择性集成算法[J]. 山西大学学报（自然科学版）, 2015,(1):58-66.
23. Abramoff M D, Neumeier M.Mass screening of diabetic retinopathy using automated Methods[M].Teleophthalmology in Preventive Medicine. Springer Berlin Heidelberg, 2015: 41-50P.
24. Donsa K, Spat S, Beck P, et al. Towards personalization of diabetes therapy using computerized decision support and machine learning: some open problems and challenges[M].Smart Health.Springer International Publishing,2015: 237-260P.
25. Akram M U, Khalid S, Khan S A. Identification and classification of microaneurysms for early detection of diabetic retinopathy[J]. Pattern Recognition, 2013, 46(1).
26. Wisaeng K, Hiransakolwong N, Pothiruk E. Automatic Detection of Exudates in Digital Retinal Images[J]. International Journal of Computer Applications, 2013, 64(4):19-26P.
27. Harangi B, Antal B, Hajdu A. Automatic exudate detection with improved Naive-bayes classifier[C]. International Symposium on Computer-Based Medical Systems. IEEE, 2012:1-4P.
28. Larsen N, Godt J, Grunkin M, et al. Automated detection of diabetic retinopathy in a fundus photographic screening population[J]. Investigative Ophthalmology & Visual Science, 2003, 44(2):767-771P.
29. Jelinek H J, Cree M J, Worsley D, et al. An automated microaneurysm detector as a tool for identification of diabetic retinopathy in rural optometric practice[J]. Clinical and Experimental Optometry, 2006, 89(5):299-305P.
30. Priya. R et al. Review of automated diagnosis of diabetic retinopathy using the support vector machine[J] International Journal of Applied Engineering Research, 2011, 1(4):1-4P.
31. Priya R, Aruna P. SVM and Neural Network based Diagnosis of Diabetic Retinopathy[J]. International Journal of Computer Applications, 2012, 41(1):6-12P.
32. Aravind C, Ponnibala M, Vijayachitra S. Automatic detection of microaneurysms and classification of diabetic retinopathy images using SVM technique[C]. IJCA Proceedings on international conference on innovations in intelligent instrumentation, optimization and electrical sciences, 2013, (11):18-22P.
33. Perumalsamy N, Prasad N M, Sathya S, et al. Software for reading and grading diabetic retinopathy: Aravind Diabetic Retinopathy Screening 3.0[J]. Diabetes Care, 2007, 30(9):2302-2306P.
34. McCuIIoch WS，PITTS W. A IogicaI caIcuIus of the ideas immanent in nervous activity[J]. BuII Math Biophys,1943,( 5):115 - 133.
35. Rumelhart D, Mcclelland J. Learning Internal Representations by Error Propagation[M]// Neurocomputing: foundations of research. MIT Press, 1988:318-362.
36. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics, 1980, 36(4):193-202.
37. J. Nowlan and G. E. Hinton. Simplifying neural networks by soft weight-sharing. Neural Computation, 4(4), 1992.
38. Goodfellow I J, Warde-Farley D, Mirza M, et al. Maxout networks[J]. arXiv preprint arXiv:1302.4389, 2013.

致 谢

整个毕业设计过程中十分感谢夏勇老师对我的帮助。

毕设期间夏老师给了我许多的建议和指导，从一开始的选题，指导修改开题报告，提供参考文献，到后面孜孜不倦的解疑答惑，一直都十分负责和严谨。其实夏老师有很多工作要做，一直都很忙碌，但是也总是会在看到对我们毕设有指导意义的文章时推荐给我们，他能清楚地记着大家的题目是什么，用的方法是什么，进展到哪一步了，又遇到了怎样的问题。如果不是极其负责的老师，是断然做不到这一步。我常常都是凌晨一两点的时候还能收到老师的消息，督促和指导贯穿始终。

除此以外，也感谢夏老师身体力行，为我以后工作学习树立了标杆。从每次组会的安排到对待论文的细致，我学习到了如何是做科研的正确态度：要有梦想，保持热情，对待自己的心工作是信任的，也是喜爱的；文章要多读，不论是不是准确地与所做工作相关的，都要去关注，这样才有可能有新的想法；基础要扎实，很多人做研究可能做着做着会忽略一些很基础的东西，但是夏老师总是强调理论基础，并且安排集体学习……真的从很多很多方面教育着我。

也感谢教研室的师兄师姐，给我提供了各种各样的帮助，感谢陈媛媛师姐对我论文的指导，帮我指出了许多问题；感谢王勤芬师姐帮我解决了一些运行代码的问题；感谢董晓云师姐的鼓励以及生活上的帮助；感谢秦子博师兄提供服务器技术支持；还要感谢张静宜，王妍，路梦康等一同做毕设的同学的帮助和提醒。

还要我的感谢父母和爷爷奶奶，他们常常给我打电话，询问我，督促我。感谢他们一直作为我最大的力量支持和鼓励着我。虽然离家千里，他们送我来此求学，但是能够感觉到他们的目光，一直在我的身后，欣慰又关切的望着我。

在西北工业大学计算机学院进行毕设真的是一件幸运的事情，无论是从硬件方面，还是数据方面学校都能给我们提供最好的支持，让我有机会学习到各种算法去完成这个数据量很大的工作。

毕业设计小结

准确来讲，这次的毕业设计还有很多不足的地方，我投入的时间和精力可能不及其他同学，从考研到找工作，考雅思，申请留学，感觉好像是经历了人生最难的阶段。我常常感到焦虑和不安，压力也一直都有，真的有一段时间每天晚上睡不着觉，不想去面对又一个崭新的太阳。

我很遗憾自己的毕业设计没有像计划般那样进行，很多想要试的算法和改变都没有去尝试，希望答辩以后还能趁着自己有这么好的环境和这么多的良师益友，在毕业以前还能多学习学习，做一些东西。也希望能够在接下来做好以后的每一份工作，不让努力过的东西变成遗憾。

这个毕业设计意义重大，收到了太多的帮助和支持。同时，这个毕设中用到的机器学习的方法我相信对我以后的工作学习也会有很大的帮助。毕设的题目也是十分有意义有前景的，我愿意在这条路上走得更远一些，感谢老师提供方向。我也想要继续，带着梦想前行。