

基于卷积神经网络的皮肤病诊断多二分类器研究^①

代闯闯^{②***} 栾海晶^{***} 杨雪莹^{***} 过晓冰^{***} 牛北方^{***} 陆忠华^{③***}

(* 中国科学院计算机网络信息中心 北京 100190)

(** 中国科学院大学 北京 100049)

(*** 联想研究院 北京 100094)

摘要 近年来,随着深度学习技术的日益普及与发展,卷积神经网络(CNN)被广泛应用于辅助医学诊断,并在医学影像诊断领域取得了重要的研究成果。本研究基于皮肤病数据种类繁多、特征不显著等特点,引入多二分类的研究方法搭建了从医学影像到计算机辅助诊断的框架,解决了目前皮肤病难以区分的问题,并在常见的皮肤病分类识别问题中得到具体检验。首先,本研究以 3 类常见的皮肤病数据集(白癜风、痤疮和银屑病)为例,实现了图像数据的增强、分割、多二分类器的构建、图像块的分类、皮肤病的判别等完整的工作流程。其次,在分组数据交叉验证下,三分类判别准确率为 0.8320,四分类判别的准确率达到 0.9125。最后,为了获得更高的准确率,在随机森林方法结果不理想的情况下,本研究引入了多二分类器网络架构,准确率达到了 0.9377。

关键词 辅助医学诊断;卷积神经网络(CNN);皮肤病;多二分类器

0 引言

卷积神经网络是受生物学上感受野机制启发而提出的^[1-2]。在 20 世纪末,卷积神经网络 LeNet-5^[3-5]应用到手写数字识别领域,除此之外,基于 LeNet-5 的手写数字识别系统被美国很多银行所使用,用于识别支票上的手写数字。目前,随着计算能力的提高,在基于图像处理的深度学习领域中,新的网络模型不断被提出并得到应用。2012 年,AlexNet^[6-8]模型的提出证明了卷积神经网络(convolutional neural network,CNN)在复杂模型下的有效性,该模型结合图形处理器(graphics processing unit,GPU)进行并行训练可在有限时间内得到预期的效果^[9],并赢得了 2012 年 ImageNet 图像分类竞赛的冠军^[10]。随着 AlexNet 结构的进一步发展,2014 年,ILSVRC 竞赛中 VGG^[11-13]模型应运而生,仅次于本次竞赛中的冠

军 GoogLeNet^[14-15]。然而,VGG 模型在多个迁移学习任务中的表现要优于 GoogLeNet,并且从图像中提取 CNN 特征,VGG 模型是首选算法。GoogLeNet 模型的提出则证明随着卷积层数的增加和网络结构的深入,模型的学习效果会更佳^[16],因此获得了 2014 年 ImageNet 图像分类竞赛的冠军。2015 年,深度残差网络的提出使得更加简化的卷积结构和更深的网络得到应用^[17]。

近年来,随着人工智能和大数据科学的发展,新的深度学习模型尤其是卷积神经网络模型得到不断发展^[10,16-17],深度学习在图像识别问题中得到越来越多的应用。深度学习方法受数据驱动在医疗领域得到应用,因为基于不同的医疗领域,各数据来源和疾病特征不同,需要对现有的模型和分析框架做相应的调整,一些像素过大的图片数据如 CT 扫描作为输入时需要进行信息压缩或者特征提取再进一步

① 中国科学院战略性先导科技专项(XDC01040100)和广西科技重大专项(桂科 AA18118054)资助。

② 男,1993 年生,博士生;研究方向:高性能计算,区块链技术;E-mail: dcc@sccas.cn。

③ 通信作者,E-mail: zhlu@sccas.cn。

(收稿日期:2021-08-07)

使用网络进行学习^[18],有关分类诊断如黑色素瘤的良性或恶性的判别需要借助皮肤癌领域对于疾病的细分知识^[19-20]。

1 神经网络模型介绍

1.1 AlexNet

AlexNet 是 Hinton 和他的学生 Alex Krizhevsky

在图像分类挑战上提出的一种卷积神经网络结构模型,用来对 ImageNet 上 120 万张高清图像中的 1000 类进行分类,在测试集上实现了 37.5% 的 top1 错误率,17.0% 的 top5 错误率。AlexNet 是卷积神经网络在图像分类上的一个经典的网络结构,具体的网络结构如图 1 所示。

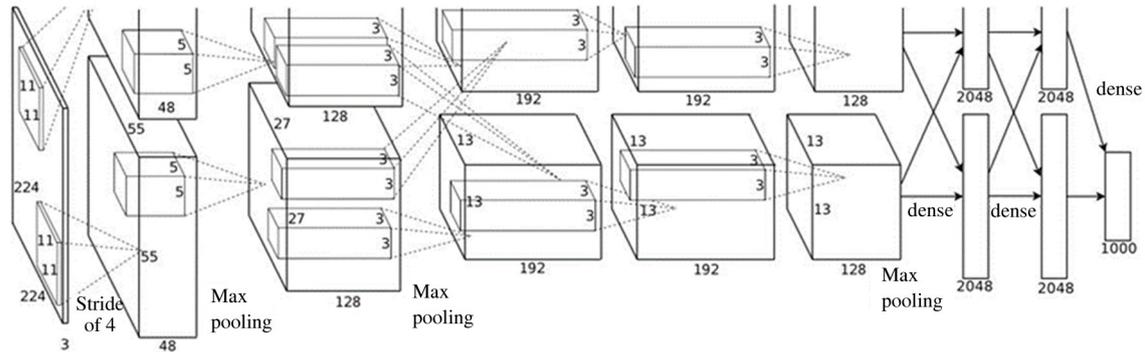


图 1 AlexNet 网络结构

AlexNet 模型中使用了修正线性单元 (rectified linear unit, ReLU), ReLU 函数的公式如下:

$$ReLU(z) = \max(0, z) \quad (1)$$

ReLU 函数曲线如图 2 所示。

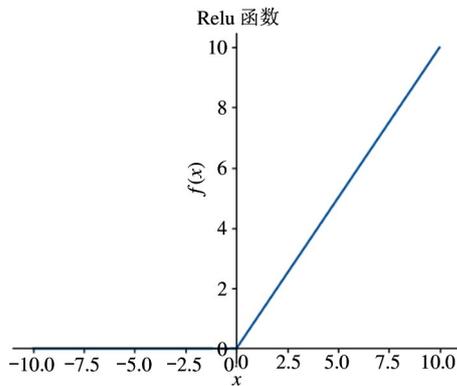


图 2 ReLU 函数曲线

通过式(1)和图 2 可以看出,ReLU 函数非常简单,在 $x < 0$ 处取值为 0,在其他地方取值为 x 。ReLU 代替了传统的 Tanh 和 Logistic 等激活函数,运算速度快,能够很好地解决梯度消失的问题^[21-22]。ReLU 函数最先在 Hinton 的 ImageNet 的论文中^[10]提出,截至目前,已得到了广泛的应用。然而,ReLU 函数在网络模型训练过程中最大的缺点是容易出现“死

亡现象”,即如果反向传播过程中有一个参数为零,则会导致其后的参数均无法进行更新,随后提出的 Leak ReLU、PreLU、Random ReLU 等改进版本的激活函数则可很好地解决这个问题。

1.2 GoogLeNet

随着当前深度学习的网络规模的扩大以及训练参数的增多,在训练过程中容易出现过拟合的现象。除此之外,深度学习涉及的数据计算量巨大,且随着卷积层数的增加使得数据计算量上升至指数级别,因此仅单纯地追求模型识别准确率的提高而盲目增加网络规模是不可取的,需要着眼于特征提取模块的设计,在基本的特征提取单元上进行一些优化操作,然后运用优化后的特征提取模块去构建网络,可能会有利于模型准确率的提高。GoogLeNet 提出了 Inception module 的概念, Inception 模块同时使用 1×1 、 3×3 、 5×5 等大小不同的卷积核,并将得到的特征映射在深度上拼接起来作为输出特征映射,一方面使得网络层具有更加强大的特征提取功能,另一方面使得网络层更深、更宽,结构变得更加稀疏,但参数量没有增加,计算量维持在一定范围内^[23-26]。Inception 是一种网中网的结构,即原来的结点也是一个网络。Inception 一直在不断发展,目前已经提

出了 V1^[16]、V2、V3^[27]、V4。Inception V1 的结构如图 3 所示,模块中的前置 1×1 卷积用来减少特征映射的深度,起到了降维的作用。

GoogLeNet 是由上图中的 Inception 模块所构建的模型,包含 9 个 Inception 模块、5 个池化层、一些卷积层和全连接层构成,一共 22 层网络,如图 4 所示。GoogLeNet 由于网络的卷积核可共享,计算梯度可累加,因此不会出现梯度消失的现象,其典型代

表应为 ResNet^[17,28]。

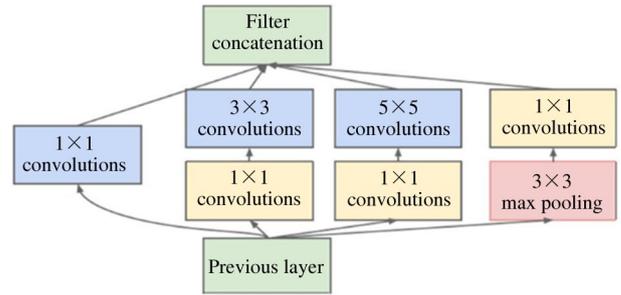


图 3 Inception V1 的模块结构

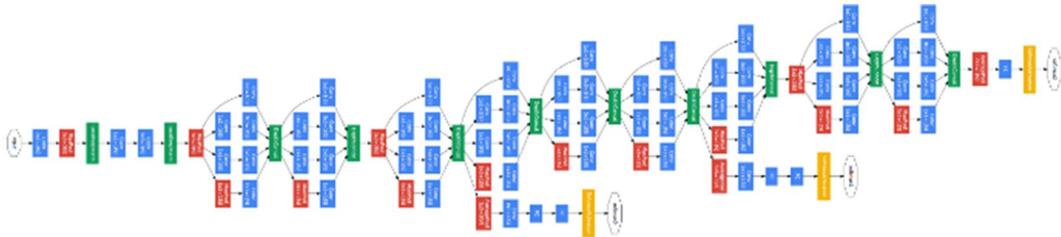


图 4 GoogLeNet 网络结构^[27]

2 数据预处理及预分类

本研究将基于皮肤病图像数据集的特点以及皮肤病的病理特征,提出分割、筛选、分类和分析的框架进行皮肤病的分类。首先对原始数据集进行数据预处理及预分类,预处理包括对原始图像的光强调节、亮度归一和切片,预分类是指通过一个 CNN 分类器将图像块分成皮肤、毛发、背景、器官 4 类;再对

皮肤类图像块进行疾病的细分类,分别通过基于 AlexNet、GoogLeNet、多二分类器和随机森林^[29]的方法得到相应图像块的概率向量;最后根据图像块的概率向量从不同的角度进行信息的整合及预测,最终对相应的图像做确切诊断。整体研究流程如图 5 所示。

由于医学影像数据具有来源复杂以及数据采集不够规范的特点,将图像预处理划分为如下 3 步:

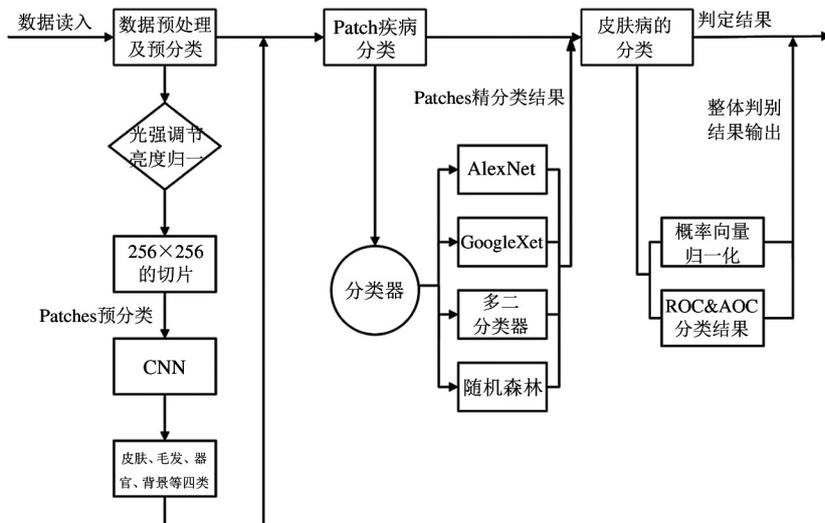


图 5 研究整体流程

(1)亮度归一化调节;(2)图像块分割;(3)图像块分类筛选。非线性变换的亮度调整减轻了图像采集过程中外部光线因素的不利影响,可提高对图像识别的准确率。图像块分割是指对原始影像数据进行切片处理,根据一致性的规则从每张图片中切割出若干小的分块(256×256 像素,RGB 三通道),完成数据的规范化,使得不同像素的医学影像能够适用于同一个模型。图像块分类筛选指的是使用卷积神经网络训练识别医学影像分割的有效部分,排除干扰信息。本研究使用 CNN 对图像块进行训练,筛选出皮肤面积占比大于 3/4 的有效图像块用于进一步的疾病分类。

对经预处理后筛选出的包含皮肤信息的图像块,使用 CNN 进行分类,可将其分为白癜风、痤疮、银屑病、健康 4 类。使用经归一化的四维概率向量表示每一个图像块经过 CNN 预测后的结果。研究中使用 Caffe^[30] 深度学习框架,并在多个经典的 CNN 和自定义的网络结构进行了测试,最终选取 GoogLeNet 和 AlexNet 作为本研究的核心网络,调整了网络训练中的 softmax-loss 层,修改参数权重提升整个网络对皮肤病诊断的敏感度。针对某一病例,均对应多个图像信息,因此需对疾病分类 CNN 得到的结果进行信息的整合。

2.1 数据预处理

2.1.1 数据集简介

数据集由 128 张 2000~4000 像素不等的病人皮肤照片组成。病人的皮肤照片包括了银屑病、痤疮和白癜风等 3 种皮肤病。图 6 是病人的侧面照片的示例图。



图 6 示例图

2.1.2 预处理步骤

(1) 亮度调整

在进行图像采集时,不同的光线采集条件会对所采集图像的颜色、亮度等特征产生干扰,因此需对图像亮度进行归一化处理。经过实验验证,当亮度归一化比例取 0.62 时效果最好。亮度调整的示意图如图 7 所示。



图 7 亮度归一化

(2) 数据集的扩充及实现

由于深度学习需要海量数据进行模型训练,因此本实验利用常规图像扩充技术进行数据扩增。数据的扩增主要包括图像的旋转、镜像、缩放和平移等,本研究中采用了镜像和旋转两种方法对常规图像进行集合扩增。

(3) 图片剪裁

由于初始数据集中图片的大小不等,且某些网络的结构对输入图片的大小也有严格的要求,实验中会将初始宽高在 2000~4000 像素不等的图片剪裁成 256×256 的分块。图片剪裁操作的示意图如图 8 所示。



图 8 图片分割

在图像块中,如图 9 中包含背景图像块、图 10 中健康皮肤的图像块和图 11 中包含其他无关信息的图像块,会对判别结果产生误导,实验中需要将它们从训练集中剔除。

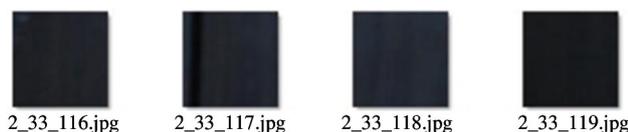


图9 包含背景的图片块



图10 健康皮肤的图像块

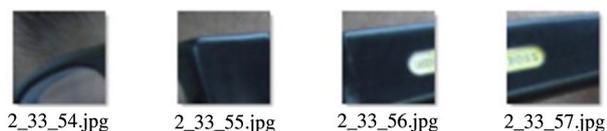


图11 其他无关信息的图像块

2.2 数据预分类

对经预处理后的图像块进行疾病的细分类,本研究采用随机森林和普通的卷积神经网络等方法并交叉验证其稳定性,最终确定 GoogLeNet 作为预分类的分类模型。经多次实验与交叉验证,测试集准确率稳定在 94% 左右,实验的混淆矩阵和准确率如表 1 所示。

经图像块预分类后,可得到包含患者皮肤信息的有效图像块样本。经人工筛选标记得到健康、银屑病、痤疮、白癜风 4 类皮肤图像块样本。然后将数据集划分为训练集和测试集,训练集每类包含 682 个样本,测试集每类包含 76 个样本。

表 1 预分类实验结果

TRAIN	Sk (P)	Ot (P)	Or (P)	Ha (P)	TEST	Sk (P)	Ot (P)	Or (P)	Ha (P)
Sk (T)	953	0	1	0	Sk (T)	94	1	9	2
Ot (T)	0	952	0	2	Ot (T)	1	98	4	3
Or (T)	0	0	954	0	Or (T)	2	0	104	0
Ha (T)	0	0	0	954	Ha (T)	0	2	0	104
Train acc	0.999				Test acc	0.943			

3 皮肤病分类的多二分类器设计

针对皮肤病的多分类问题,若要通过神经网络训练获得一个好的结果,则需设计一个合适的神经网络。随着深度学习的快速发展,网络结构也不断更迭,本文旨在不改变网络结构的前提下,尽可能提高分类器的准确率。卷积神经网络相对于其他的分类器来说,可看作是一个特征提取的过程,而一般的分类器往往是通过某些已经发现的特征来进行分类。如果将具有某些共同特征的元素归到同一类,再进行训练,这些共同的特征可能会更容易地被提取出来,因此拟通过这个网络来训练多个二分类器从而实现对初始多分类器的优化。

多个二分类器类似于二叉树的模型,先将数据分成两类,再对每一类构建二分类器,得到的一个“二叉树”如图 12 所示。将皮肤病分类(白癜风、痤疮、银屑病、健康)的四分类的问题,标签记为 1(白癜风)、2(痤疮)、3(银屑病)、4(健康),第一个分类

器先区分 1(白癜风)、2(痤疮)和 3(银屑病)、4(健康),再构建 2 个分类器分别区分 1(白癜风)和 2(痤疮)、3(银屑病)和 4(健康),一个数据经过 2 个分类器即可获得最终结果。然而该方法却存在很多不足之处,由于图像种类繁多,为每个二分类器单独设计网络不太现实,数据量的变化对网络的训练有很大的影响,在数据量不足的时候很容易过拟合。基于神经网络的多二分类器训练中若某个网络中的二分类器出现分类错误,该分类器就无法获得正确的分类结果。

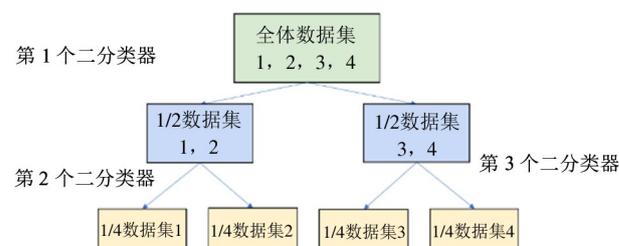


图 12 多个二分类器

为降低分类时的错误率,随后提出一种改进的方法。该方法的主要思想是首先获得一个二叉树,针对该二叉树的每一层只构造一个分类器,每个分类器都使用所有的数据集,使得每一个节点的孩子节点在其所在层对应的分类器都属于不同类别,具体如图 13 所示。对于最后一层若不包含所有标签类别,将剩余标签随意加入两类中的一类。以四分类为例,则只需要构造 2 个二分类器,第 1 个用于区分 1(白癜风)、2(痤疮)和 3(银屑病)、4(健康);第 2 个用于区分 1(白癜风)、3(银屑病)和 2(痤疮)、4(健康);经过 2 个分类器就可以得到结果。假如第 1 个分类器将 1(白癜风)、2(痤疮)标记为 0,3(银屑病)、4(健康)标记为 1;第 2 个分类器将 1(白癜风)、3(银屑病)标记为 0,2(痤疮)、4(健康)标记为 1。若一个数据,第 1 个分类器将其标记为 1,第 2 个分类器也将其标记为 1,则不对应任何的情况,对于此问题,本研究提出了如下 2 种解决方法。

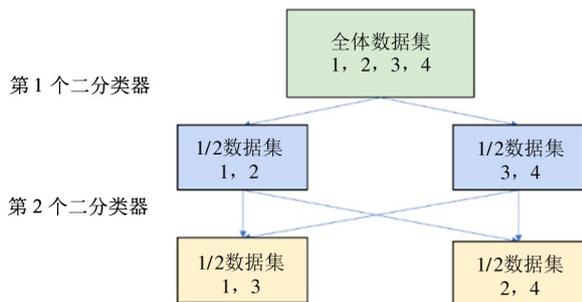


图 13 改进的多个二分类器

(1) 对无法分类得出结果的数据再进一步分类。既然是不对应任何结果的情况,那么至少有一个分类器在分类时出现了错误。那么由多个二分类器构成的分类器则对这个数据不太适用,可以将这个难以识别的数据再放入多分类器进行判别。

(2) 对无法分类得出结果的数据不做处理,由于这个数据在某个网络中出现了错误,则多分类网络在处理这个数据的时候很可能会出现错误,实验中可再定义出一类记为无法判断,不做任何处理。

这 2 种处理方法各有优劣,第 1 种不会出现无法判断的情况,每一个数据都能给出一个结果,而第 2 种方法简单高效,可以对复杂特征减少误判,后续的研究将对这 2 种方案分别进行具体实验分析。

4 实验结果分析

4.1 AlexNet—实验结果分析

在训练皮肤病四分类(健康、银屑病、痤疮、白癜风)的 CNN 时,首先尝试了卷积神经网络 AlexNet。数据集的训练集每类包含 682 个样本,测试集每类包含 76 个样本,该卷积神经网络的四分类结果如表 2 所示。综合实验结果计算得到网络的准确率是 0.8750。

表 2 AlexNet 网络四分类结果

分类\真实	白癜风	痤疮	银屑病	健康
白癜风	69	2	1	4
痤疮	2	67	4	3
银屑病	2	2	72	0
健康	6	7	5	58

为了提高准确率,将数据集随机分为三类(去掉健康皮肤的分类)和四类进行十折交叉验证,三分类实验结果为 0.8202、0.8465、0.8596、0.8070、0.8356、0.8465、0.8421、0.8246、0.8133、0.8202,平均准确率为 0.8320;四分类实验结果为 0.9049、0.8721、0.9049、0.9114、0.9102、0.9245、0.9213、0.9245、0.9278、0.9235,平均值为 0.9125。该实验表明实验结果过分依赖于数据集的选取,可能是数据集少,不完善所造成的结果。

为进一步提高准确率,实验随后使用了多二分类的方法,训练 3 个二分类网络后,针对无法判断分类结果的数据再使用上述的四分类器,最终的实验结果如表 3 所示,计算实验结果准确率为 0.9177;如果将无法判断的部分不做处理,得到的结果如表 4 所示,计算实验结果准确率为 0.9377。

表 3 AlexNet 多二分类和四分类实验结果

分类\真实	白癜风	痤疮	银屑病	健康
白癜风	72	1	0	3
痤疮	1	72	1	2
银屑病	2	2	71	1
健康	4	5	3	64

表4 AlexNet 网络多二分类实验结果

分类\真实	白癜风	痤疮	银屑病	健康
白癜风	62	0	0	1
痤疮	1	57	0	1
银屑病	1	1	65	1
健康	5	3	2	57

通过上面的结果可以发现,使用这个方法对结果的提升还是比较明显的,在舍弃数据的情况下也可将准确率提升至93%。

4.2 GoogLeNet-实验结果分析

在训练皮肤病四分类(健康、银屑病、痤疮、白癜风)的CNN时,实验又尝试了卷积神经网络GoogLeNet,当使用GoogLeNet网络时的四分类结果如表5所示,计算实验结果准确率为0.8849。

表5 GoogLeNet 网络四分类结果

分类\真实	白癜风	痤疮	银屑病	健康
白癜风	70	1	0	5
痤疮	2	66	4	4
银屑病	0	5	71	0
健康	2	12	0	62

为进一步提高准确率,实验随后又使用了多二分类的方法,在训练3个二分类网络后,针对无法判断分类结果的数据再运用四分类器,最终的实验结果如表6所示,计算实验结果准确率为0.88196;如果将无法判断的数据不做处理,得到的结果如表7所示,计算实验结果准确率为0.9242。在运用GoogLeNet进行测试后,发现在舍弃一部分数据后准确率仍能得到显著提高,即使舍弃数据仅占一小部分的比例。

表6 GoogLeNet 多二分类和四分类实验结果

分类\真实	白癜风	痤疮	银屑病	健康
白癜风	70	2	0	4
痤疮	2	67	4	3
银屑病	2	2	71	1
健康	8	7	3	58

表7 GoogLeNet 多二分类实验结果

分类\真实	白癜风	痤疮	银屑病	健康
白癜风	61	0	0	6
痤疮	0	64	1	2
银屑病	0	2	67	1
健康	1	7	0	52

4.3 实验分析

综合来看,多二分类器下 AlexNet 和 GoogLeNet 模型对于准确率的提升效果一致。与普通的分类器相比,多二分类器能够快速发现数据集的共同特征,在第1种策略数据出错时至少要有2个二分类器出错,降低了误判的概率,准确率会比直接运用四分类器时高。第2种策略舍弃的那部分数据,必定是某个二分类器错判的数据,也是模型不易判断的数据,该部分数据在判断的时候准确率会较低,舍弃数据后准确率会得到提升。与使用 AlexNet 的结果相比,GoogLeNet 的准确率反而较低,所以不是网络的结构越复杂,效果越好。GoogLeNet 虽然在结构上做了不少优化,但是其二十几层的深度要求更多的数据去训练,否则带来的过拟合现象会降低测试集的准确率,因为本实验数据相对较少,容易出现过拟合现象。

5 皮肤病诊断

经过对图像块的分类,得到了对于每个图像块的疾病概率,需将图像的所有图像块信息进行整合以获取最终的分类结果。通过对图像块的疾病分类预测得到每个图像块对应的概率向量如下:

$$\text{Patch 1: } (P_{11}, P_{12}, P_{13}, P_{14})$$

$$\text{Patch 2: } (P_{21}, P_{22}, P_{23}, P_{24})$$

$$\text{Patch 3: } (P_{31}, P_{32}, P_{33}, P_{34})$$

$$\text{Patch } n: (P_{n1}, P_{n2}, P_{n3}, P_{n4})$$

其中 P_{11} 、 P_{12} 、 P_{13} 、 P_{14} 分别是第 i 个图像块被判别为健康、银屑病、痤疮、白癜风的概率。然后用式(2)分别算出当前患者患有银屑病、痤疮、白癜风的归一化概率。

$$P_i = \frac{\sum_{j=1}^n P_{ij}}{\sum_{i=1}^3 \sum_{j=1}^n P_{ij}} \quad (2)$$

将最大概率所对应的皮肤病类别判定为患者所患的疾病类型。图 14 是运用混淆矩阵的方法来展示处理后的效果图。

其中 (i, j) 位置上的图是实际第 i 类样本的图

像块在第 j 类疾病上对应的概率累计值的统计直方图。通过观察 (x, x) 、 (y, y) 、 (z, z) 位置上的直方图,每个样本值都较高并均接近于 1,说明 patches 信息的整合是很有效的。

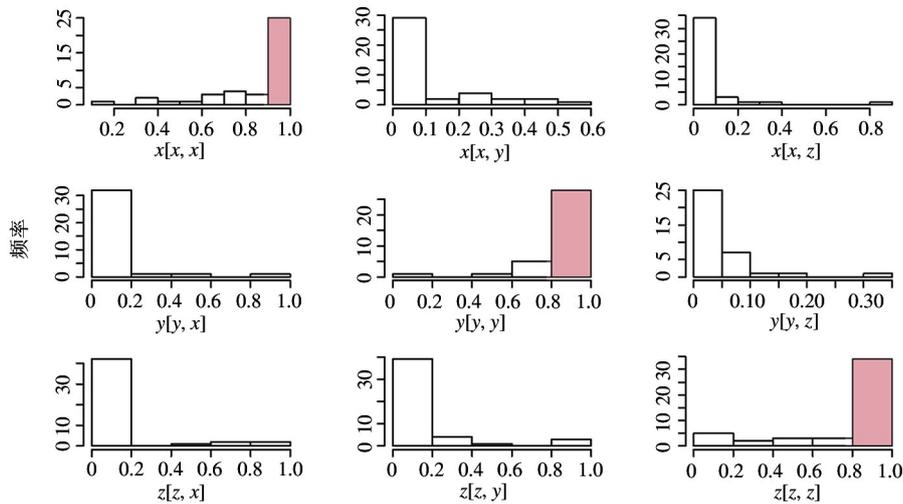


图 14 皮肤病分类

除了“混淆矩阵”的方法之外,本方法还采用了受试者工作特征 (receiver operating characteristic, ROC) 曲线来评价分类器的效果,如图 15 所示。在 ROC 曲线中代表白癜风的曲线与坐标轴围成的面积(即 AUC 统计量)是 0.9707;代表痤疮的曲线与坐标轴围成的面积(即 AUC 统计量)是 0.9668;表示银屑病的曲线与坐标轴围成的面积(即 AUC 统计量)是 0.9234。

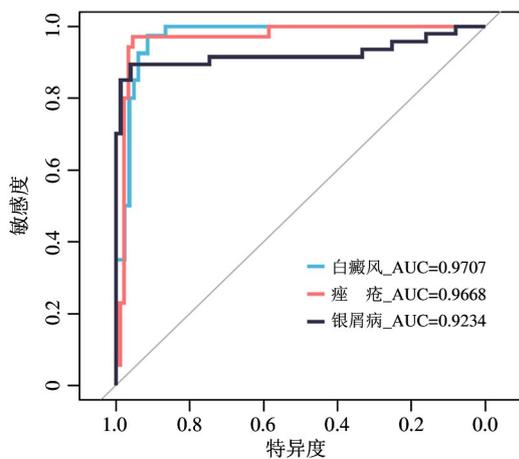


图 15 ROC 曲线

率可达到 90.9836%,现实情况下,阈值的选取会影响分类器的真阳率(true positive rate, TPR)、假阳率(false positive rate, FPR), TPR、FPR 计算公式分别如式(3)、式(4)所示。

准确率并不是唯一的分类器评价标准。某些情况下 TPR 和 FPR 的重要程度不同,由于某一种错误判断需要承担的风险和代价很大,所以不能将提高诊断准确率作为唯一标准,得到高的 TPR 才是至关重要。

$$TPR = \frac{TP}{(TP + FN)} \tag{3}$$

$$FPR = \frac{FP}{(FP + TN)} \tag{4}$$

下图是在皮肤病分类诊断问题中,考察归一化的概率累计得分在阈值 λ 变动下 TPR、FPR 和 TPR-FPR 的变化如图 16 所示(其中从上到下分别对应白癜风、痤疮、银屑病)。

在现实问题中,可通过构造一个关于 TPR、FPR 的函数用优化方法确定阈值。记阈值为 λ ,目标是选取合适的阈值使得 TPR 比较大并控制 FPR,可如下表示:

$$\max_{\lambda} f(TPR_{\lambda}, FPR_{\lambda})$$

对概率累计值取 softmax 的诊断分类方法准确

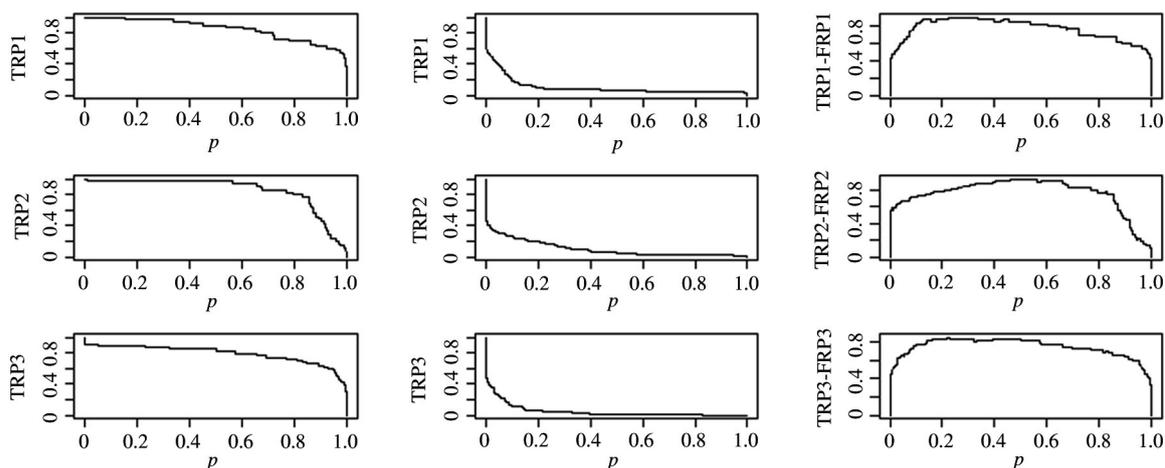


图 16 TPR、FPR、TPR-FPR 变化曲线图

其中 f 关于 TPR 单调递增, 关于 FPR 单调递减, 可以如式(5)所示。

$$f(TPR_{\lambda}, FPR_{\lambda}) = t \times TPR_{\lambda} - (1 - t) \times FPR_{\lambda} \quad (5)$$

最佳阈值 λ 的范围和参数 t 的函数关系如图 17 所示, 其中两条线段分别是最佳阈值的上下界。

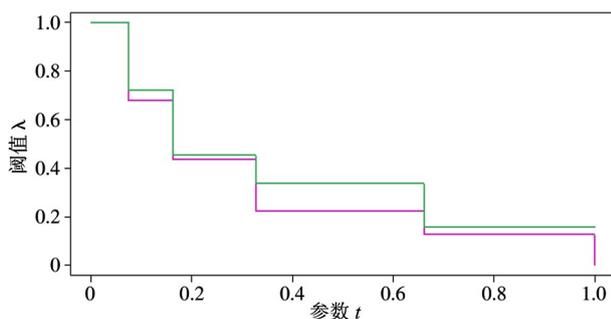


图 17 最佳阈值变化折线图

6 结论

本研究构建了从原始的医学图像数据预处理到疾病分类判断的完整流程, 并在三类常见皮肤病的诊断问题中得到检验。目前的工作仍存在很多问题与不足: (1) 分类器过分依赖于计算资源; (2) 使用的数据集过小, 覆盖的疾病种类有限, 过多的网络参数容易产生过拟合的问题; (3) 基于图像的疾病分类判断未充分利用其他辅助信息, 例如病人基本信息、化验信息等; (4) 未使用三类常见皮肤病的分类结果来研究这些疾病间的联系。

数据驱动的医疗辅助诊断方兴未艾, 相信在未

来, 随着联邦学习、集成学习等相关研究应用到医疗影像诊断领域中, 其在保护数据隐私的前提下会取得更大的进展和应用。

参考文献

- [1] BICANSKI A, BURGESS N. Neuronal vector coding in spatial cognition [J]. *Nature Reviews Neuroscience*, 2020, 21(9): 453-470
- [2] KELLER A J, ROTH M M, SCANZIANI M. Feedback generates a second receptive field in neurons of the visual cortex [J]. *Nature*, 2020, 582(7813): 545-549
- [3] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324
- [4] BOTTOU L, BENGIO Y, LE CUN Y. Global training of document processing systems using graph transformer networks [C] // *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, USA, 1997: 489-494
- [5] LE CUN Y, BOTTOU L, BENGIO Y. Reading checks with multilayer graph transformer networks [C] // *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Munich, Germany, 1997: 151-154
- [6] HAN X, ZHONG Y, CAO L, et al. Pre-trained AlexNet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification [J]. *Remote Sensing*, 2017, 9(8): 848
- [7] ZHANG Y D, GOVINDARAJ V V, TANG C, et al. High

- performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model[J]. *Journal of Medical Imaging Health Informatics*, 2019, 9(9): 2012-2021
- [8] SAMIR S, EMARY E, EL-SAYED K, et al. Optimization of a pre-trained AlexNet model for detecting and localizing image forgeries[J]. *Information*, 2020, 11(5): 275
- [9] PHILLIPS J C, HARDY D J, MAIA J D, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD[J]. *The Journal of Chemical Physics*, 2020, 153(4): 044130
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1097-1105
- [11] 赵文仓, 徐长凯, 王春鑫. 基于优化边界框回归的目标检测[J]. *高技术通讯*, 2021, 31(7): 747-753
- [12] MATEEN M, WEN J, SONG S, et al. Fundus image classification using VGG-19 architecture with PCA and SVD[J]. *Symmetry*, 2019, 11(1): 1
- [13] HA I, KIM H, PARK S, et al. Image retrieval using BIM and features from pretrained VGG network for indoor localization[J]. *Building Environment*, 2018, 140: 23-31
- [14] XIE S, ZHENG X, CHEN Y, et al. Artifact removal using improved GoogLeNet for sparse-view CT reconstruction[J]. *Scientific Reports*, 2018, 8(1): 1-9
- [15] LI C, ZHANG H, WU P, et al. A complex junction recognition method based on GoogLeNet model[J]. *Transactions in GIS*, 2020, 24(6): 1756-1778
- [16] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 1-9
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770-778
- [18] KURUVILLA J, GUNAVATHI K. Lung cancer classification using neural networks for CT images [J]. *Computer Methods Programs in Biomedicine*, 2014, 113(1): 202-209
- [19] HIRANO G, NEMOTO M, KIMURA Y, et al. Automatic diagnosis of melanoma using hyperspectral data and GoogLeNet [J]. *Skin Research Technology*, 2020, 26(6): 891-897
- [20] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks [J]. *Nature*, 2017, 542(7639): 115-118
- [21] WANG S H, MUHAMMAD K, HONG J, et al. Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization [J]. *Neural Computing*, 2020, 32(3): 665-680
- [22] OPSCHOOR J A, PETERSEN P C, SCHWAB C, et al. Deep ReLU networks and high-order finite element methods [J]. *Analysis*, 2020, 18(05): 715-770
- [23] KIM Y G, CHA E Y. Streamlined GoogLeNet algorithm based on CNN for Korean character recognition [J]. *Journal of the Korea Institute of Information Communication Engineering*, 2016, 20(9): 1657-1665
- [24] KIM H, CHUNG Y. Improved handwritten Hangeul recognition using deep learning based on GoogleNet [J]. *The Journal of the Korea Contents Association*, 2018, 18(7): 495-502
- [25] HUANG X, CHEN W, YANG W. Improved algorithm based on the deep integration of GoogleNet and residual neural network [J]. *Journal of Physics: Conference Series*, 2021, 1757(1): 012069
- [26] SALMA A, BUSTAMAM A, YUDANTHA A R, et al. Diabetic retinopathy detection and classification using GoogleNet and attention mechanism through Fundus images [J]. *Turkish Journal of Computer Mathematics Education*, 2021, 12(14): 590-597
- [27] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 2818-2826
- [28] ZHU J, YU F, LIU G, et al. Classroom roll-call system based on ResNet networks [J]. *Journal of Information Processing Systems*, 2020, 16(5): 1145-1157
- [29] KIM S J, KWAK S Y, CHUL K O B, et al. Fast pedestrian detection in surveillance video based on soft target training of shallow random forest [J]. *IEEE Access*, 2019, 13(1): 40-49

- [30] JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding[C] //Proceedings of the 22nd ACM International Conference on Multimedia, Orlando, USA, 2014: 675-678

Research on multi binary classifier for skin disease diagnosis based on convolutional neural network

DAI Chuangchuang^{* **}, LUAN Haijing^{* **}, YANG Xueying^{* **}, GUO Xiaobing^{***},
NIU Beifang^{* **}, LU Zhonghua^{* **}

(^{*} Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190)

(^{**} University of Chinese Academy of Sciences, Beijing 100049)

(^{***} The Lenovo Research, Beijing 100094)

Abstract

In recent years, with the development of deep learning technology, the convolutional neural network (CNN) has been applied to assist medical diagnosis in some medical fields and has made great progress in the field of medical imaging diagnosis. In this study, based on the characteristics of skin disease data, such as a wide variety of features are not significant, a multi binary classification method is introduced to build a framework from medical images to computer-aided diagnosis, which overcomes the difficulty of distinguishing skin diseases, and has been tested in common skin disease classification and recognition. First of all, this study takes three kinds of common skin disease data sets (vitiligo, acne, psoriasis) as examples to realize the complete workflow of image data enhancement, segmentation, construction of multi binary classifier, classification of patches, skin disease discrimination. Secondly, under the cross validation of grouped data, the accuracy of three classification is 0.8320, and the accuracy of four classification is 0.9125. Finally, in order to obtain higher accuracy, in the case that the results of random forest method are not ideal, this study introduces a multi two classifier network architecture, and its accuracy reaches 0.9377.

Key words: auxiliary medical diagnosis, convolutional neural network (CNN), dermatosis, multi-binary classier