



(12) 发明专利



(10) 授权公告号 CN 114882995 B

(45) 授权公告日 2024. 12. 10

(21) 申请号 202210608744.6

(22) 申请日 2022.05.31

(65) 同一申请的已公布的文献号

申请公布号 CN 114882995 A

(43) 申请公布日 2022.08.09

(73) 专利权人 河北农业大学

地址 071001 河北省保定市灵雨寺街289号
河北农业大学

(72) 发明人 李聪聪 闫新晟 滕桂法 王芳

(74) 专利代理机构 北京卓岚智财知识产权代理
有限公司 11624

专利代理师 刘婷

(51) Int. Cl.

G16H 50/20 (2018.01)

G16H 50/70 (2018.01)

G06V 10/764 (2022.01)

G06V 10/774 (2022.01)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

G06N 3/045 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/084 (2023.01)

G06N 3/09 (2023.01)

(56) 对比文件

AU 2020103613 A4, 2021.02.04

CN 110189305 A, 2019.08.30

审查员 王文文

权利要求书1页 说明书12页 附图5页

(54) 发明名称

一种基于深度学习的舌象体质辨别方法

(57) 摘要

本发明所述一种基于深度学习的舌象体质辨别方法部署在移动终端,方便用户随时随地识别舌象图像。本发明方法采用卷积神经网络算法,对舌图像进行了同类增强和混类增强两种方法组合的方式对舌象数据集进行数据增强。使得模型能够适应各种环境下的舌象数据,提取的特征能够直接输出分类结果,计算量大大减少且泛化能力更强。

获取待识别的舌象图像,使用训练好的改进的卷积神经网络GoogleNet模型对待识别的舌象图像进行识别,得到舌象的识别结果

将舌象图像的识别结果在移动终端进行展示

1. 一种基于深度学习的舌象体质辨别方法,其特征在于,所述方法部署在移动终端,所述方法包括:

获取待识别的舌象图像,使用训练好的改进的卷积神经网络GoogleNet模型对待识别的舌象图像进行识别,得到舌象的识别结果;

将舌象图像的识别结果在移动终端进行展示;

其中,改进的卷积神经网络GoogleNet模型的训练过程如下:

获取舌象图像数据集,对舌象图像数据集进行增强处理得到舌象训练数据集,其中,增强处理包括同类增强和混类增强两种方式;

使用舌象训练数据集对改进的卷积神经网络GoogleNet模型进行训练,得到训练好的改进的卷积神经网络GoogleNet模型;

其中,改进的卷积神经网络GoogleNet模型改进部分包括:

模型的每一中间层做了标准化处理,有效避免了梯度消失和梯度爆炸,加快了网络收敛速度,优化了网络结构;

残差融入到Inception结构中,对Inception结构进行了改进优化,将输入特征与Inception结构输出的特征进行了特征融合,命名为Inception B;

将Inception模块中的3 x 3卷积分支分解为1 x 3和3 x 1卷积,将5 x 5卷积分解为3 x 3卷积和1 x 3,3 x 1卷积,并在每个分支后加入了Batch Normalization,有效减少了模型的参数量,命名为Inception C;

在GoogleNet原模型的基础上,保留了没有改进的Inception模块,将其命名为Inception A,把Inception A、Inception B和Inception C结构进行融合,应用到改进之后的GoogleNet网络模型中,浅层中加入Inception A模块,中层加入Inception B模块,深层加入Inception C模块,其中A和B模块分别用了两次,C模块对特征的提取和感知能力更强,所以用了五次。

2. 根据权利要求1所述的方法,其特征在于,同类增强是指对相同标签的图像使用数据归一化、水平或垂直翻转、按照比例放大或缩小图像、随机旋转、随机错切变换、亮度变化以及像素填充方法进行数据集扩充,颜色也是舌象体质辨识中的重要特征,故不对舌象数据集做颜色通道方面的变换。

3. 根据权利要求2所述的方法,其特征在于,混类增强的方式是为了解决样本数据集类间不平衡问题,混类增强方式对两类不同标签的样本数据进行混合,构建虚拟的训练样本丰富舌象数据集,其数学表达式如(1)所示:

$$\begin{aligned}\hat{x} &= \lambda x_i + (1 - \lambda)x_j \\ \hat{y} &= \lambda y_i + (1 - \lambda)y_j\end{aligned}\tag{1}$$

其中 (x_i, y_i) (x_j, y_j) 是随机选取的两个样本及对应标签, λ 为服从beta分布的随机数,且 $\lambda \in [0, 1]$ 。

一种基于深度学习的舌象体质辨别方法

技术领域

[0001] 本发明涉及图像领域,尤其涉及一种基于深度学习的舌象体质辨别方法。

背景技术

[0002] 现阶段利用机器学习、深度学习等人工智能技术进行图像识别已经趋于成熟。有些学者通过收集大量的带有体质标签的舌图像建立舌图像数据库,并设计算法将舌图像与其建立的数据库进行匹配,以此达到识别的目的。考虑到传统人工提取舌象特征的复杂性以及舌头区域分割的准确性问题,深度学习因其自动提取特征的优势被应用于体质辨识中,一些学者采用卷积神经网络的各种模型来进行舌象体质辨识。这些研究虽然推动了舌诊智能化体质辨识的发展,但是也存在一些普遍的问题,第一:深度学习需要大量的数据集作为支撑,舌图像的样本数据采集具有一定难度。第二:只是简单的使用模型进行分类并没有针对模型改进优化,导致用以上方法进行体质辨识的准确率并不高。

发明内容

[0003] 为解决上述问题,本申请提出了一种基于深度学习的舌象体质辨别方法,具体为一种基于改进卷积神经网络GoogleNet模型的舌象图像识别方法,在采集到的舌象图像数据集的基础上使用该方法进行体质辨识,本申请的主要贡献有:构建了带有体质标签的舌象数据集,通过添加批量归一化操作、改进优化Inception模块、添加注意力机制以及不同Inception结构组合使用等方法,提出一种适用于舌象体质辨识的GoogleNet网络模型(TCR-GoogleNet)。设计了消融实验验证各个改进模块对预测精度的贡献程度,并与5种图像分类的主流网络设计了对比实验,证明了本申请提出的模型在准确率和稳定性上都有较大的提升。

[0004] 本发明实施例提供了如下技术方案:

[0005] 基于深度学习的舌象体质辨别方法,所述方法部署在移动终端,所述方法包括:

[0006] 获取待识别的舌象图像,使用训练好的改进的卷积神经网络GoogleNet模型对待识别的舌象图像进行识别,得到舌象的识别结果;

[0007] 将舌象图像的识别结果在移动终端进行展示;

[0008] 其中,改进的卷积神经网络GoogleNet模型的训练过如下:

[0009] 获取舌象图像数据集,对舌象图像数据集进行增强处理得到舌象训练数据集,其中,增强处理包括同类增强和混类增强两种方式;

[0010] 使用舌象训练数据集对改进的卷积神经网络GoogleNet模型进行训练,得到训练好的改进的卷积神经网络GoogleNet模型;

[0011] 其中,改进的卷积神经网络GoogleNet模型改进部分包括:

[0012] 模型的每一中间层做了标准化处理,有效避免了梯度消失和梯度爆炸,加快了网络收敛速度,优化了网络结构;

[0013] 残差融入到Inception结构中,对Inception结构进行了改进优化,将输入特征与

Inception结构输出的特征进行了特征融合,命名为Inception B;

[0014] 将Inception模块中的3 x 3卷积分支分解为1 x 3和3 x 1卷积,将5 x 5卷积分解为3 x 3卷积和1 x 3,3 x 1卷积,并在每个分支后加入了Batch Normalization,有效减少了模型的参数量,命名为Inception C;

[0015] 在GoogleNet原模型的基础上,保留了没有改进的Inception模块,将其命名为InceptionA,把InceptionA、InceptionB和Inception C结构进行融合,应用到改进之后的GoogleNet网络模型中,浅层中加入InceptionA模块,中层加入InceptionB模块,深层加入Inception C模块,其中A和B模块分别用了两次,C模块对特征的提取和感知能力更强,所以用了五次。

[0016] 其中,同类增强是指对相同标签的图像使用数据归一化、水平或垂直翻转、按照比例放大或缩小图像、随机旋转、随机错切变换、亮度变化以及像素填充方法进行数据集扩充,颜色也是舌象体质辨识中的重要特征,故不对舌象数据集做颜色通道方面的变换。

[0017] 其中,混类增强的方式是了解决样本数据集类间不平衡问题,混类增强方式对两类不同标签的样本数据进行混合,构建虚拟的训练样本丰富舌象数据集,其数学表达式如(1)所示:

$$\begin{aligned}\hat{x} &= \lambda x_i + (1 - \lambda)x_j \\ \hat{y} &= \lambda y_i + (1 - \lambda)y_j\end{aligned}\quad (1)$$

[0019] 其中 (x_i, y_i) (x_j, y_j) 是随机选取的两个样本及对应标签, λ 为服从beta分布的随机数,且 $\lambda \in [0, 1]$ 。

[0020] 与现有技术相比,上述技术方案具有以下优点:

[0021] 本申请所述基于深度学习的舌象体质辨别方法部署在移动终端,方便用户随时随地识别舌象图像。本发明方法采用卷积神经网络算法,对舌象图像进行了同类增强和混类增强两种方法组合的方式对舌象数据集进行数据增强。使得模型能够适应各种环境下的舌象数据,提取的特征能够直接输出分类结果,计算量大大减少且泛化能力更强。

附图说明

[0022] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0023] 图1为本申请所述的基于深度学习的舌象体质辨别方法流程示意图;

[0024] 图2为提出的方法架构;

[0025] 图3为样本图像;

[0026] 图4为数据增强后的图像;

[0027] 图5为Mixup可视化;

[0028] 图6为Inception B结构;

[0029] 图7为Inception C结构;

[0030] 图8为SE模块;

- [0031] 图9为InceptionA结构;
- [0032] 图10为舌象体质辨识模型结构图;
- [0033] 图11为训练过程中的准确率曲线(1)和损失值曲线(2);
- [0034] 图12为四种模型测试集的准确率训练曲线。

具体实施方式

[0035] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0036] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本申请中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本申请所示的这些实施例,而是要符合与本申请所公开的原理和新颖特点相一致的最宽的范围。

[0037] 如图1所示,本申请提供了一种一种基于深度学习的舌象体质辨别方法,具体如下:

[0038] 获取待识别的舌象图像,使用训练好的改进的卷积神经网络GoogleNet模型对待识别的舌象图像进行识别,得到舌象的识别结果;

[0039] 将舌象图像的识别结果在移动终端进行展示;

[0040] 其中,改进的卷积神经网络GoogleNet模型的训练过如下:

[0041] 获取舌象图像数据集,对舌象图像数据集进行增强处理得到舌象训练数据集,其中,增强处理包括同类增强和混类增强两种方式;

[0042] 使用舌象训练数据集对改进的卷积神经网络GoogleNet模型进行训练,得到训练好的改进的卷积神经网络GoogleNet模型;

[0043] 其中,改进的卷积神经网络GoogleNet模型改进部分包括:

[0044] 模型的每一中间层做了标准化处理,有效避免了梯度消失和梯度爆炸,加快了网络收敛速度,优化了网络结构;

[0045] 残差融入到Inception结构中,对Inception结构进行了改进优化,将输入特征与Inception结构输出的特征进行了特征融合,命名为Inception B;

[0046] 将Inception模块中的 3×3 卷积分支分解为 1×3 和 3×1 卷积,将 5×5 卷积分解为 3×3 卷积和 1×3 , 3×1 卷积,并在每个分支后加入了Batch Normalization,有效减少了模型的参数量,命名为Inception C;

[0047] 在GoogleNet原模型的基础上,保留了没有改进的Inception模块,将其命名为InceptionA,把InceptionA、InceptionB和Inception C结构进行融合,应用到改进之后的GoogleNet网络模型中,浅层中加入InceptionA模块,中层加入InceptionB模块,深层加入Inception C模块,其中A和B模块分别用了两次,C模块对特征的提取和感知能力更强,所以用了五次。

[0048] 其中,同类增强是指对相同标签的图像使用数据归一化、水平或垂直翻转、按照比

例放大或缩小图像、随机旋转、随机错切变换、亮度变化以及像素填充方法进行数据集扩充,颜色也是舌象体质辨识中的重要特征,故不对舌象数据集做颜色通道方面的变换。

[0049] 其中,混类增强的方式是了解决样本数据集类间不平衡问题,混类增强方式对两类不同标签的样本数据进行混合,构建虚拟的训练样本丰富舌象数据集,其数学表达式如(1)所示:

$$\begin{aligned} \hat{x} &= \lambda x_i + (1 - \lambda) x_j \\ \hat{y} &= \lambda y_i + (1 - \lambda) y_j \end{aligned} \quad (1)$$

[0051] 其中 (x_i, y_i) (x_j, y_j) 是随机选取的两个样本及对应标签, λ 为服从beta分布的随机数,且 $\lambda \in [0, 1]$ 。

[0052] 本申请达到的技术效果如下,具体如下:

[0053] 1、在原GoogleNet模型中添加各种改进模块,例如添加同类增强和mixup混类增强两种方法组合的方式对舌象数据集进行数据增强,添加批量归一化操作,添加注意力机制模块。

[0054] 2、对GoogleNet的基础模块Inception结构进行改进和优化,例如在原始Inception结构中加入残差结构进行特征融合,在Inception结构中加入卷积替换和分解操作,将改进后的两种Inception结构与原始Inception结构组合共同构建模型。

[0055] 3、本发明方法采用卷积神经网络算法,对舌图像进行了同类增强和混类增强两种方法组合的方式对舌象数据集进行数据增强。使得模型能够适应各种环境下的舌象数据,提取的特征能够直接输出分类结果,计算量大大减少且泛化能力更强。

[0056] 4本发明方法基于现阶段主流的GoogleNet模型进行改进和优化,分类器采用softmax分类器,是一种监督学习方法适合用于辨识体质这类多分类问题,在计算方法和识别效果上得到更好的效果。

[0057] 5本发明方法与传统的中医体质辨识的判别时间相比,降低了诊断时间提升了诊断效率,与现阶段利用机器学习深度学习的方法相比,提升了识别精度。

[0058] 6本发明方法基于大量的舌象数据集,将深度学习技术应用到传统的中医体质识别领域,不仅能通过PC端进行体质判定,而且可以通过移植到手机端进行体质判别,十分方便,准确性高,节省时间。

[0059] 7本发明方法将深度学习方法与传统中医体质辨识相结合,在大数据的基础上进行识别,有效解决了传统体质辨识难的问题,该方法有一定的市场价值和推广价值。

[0060] 下面介绍基于深度学习的舌象体质辨别具体方法。

[0061] 1、引言

[0062] 中医(Traditional Chinese Medicine,TCM)是中华民族的瑰宝,有着数千年防病治病的历史,它不仅在中国的医疗保健以及“治未病”中发挥不可替代的作用,而且在国际上也被越来越多的人接受和频繁使用。同病异治和异病同治是中医了解疾病和治疗疾病的基本原则,基于这种辨证的治疗方法的前提和基础就取决于个体的体质。在中医理论中,认为体质是人在先天条件和后天获得的基础上表现出来的相对稳定的固有特性,包括了人的形态结构和心理气质。中医体质与人的健康息息相关,体质学也是中医体质研究的基础和核心内容,影响着中医整个临床学科的诊疗。结合《黄帝内经》提出的体质的定义以及

现代体质临床应用原则,王琦教授提出了体质九分法,即将体质类型归纳分为平和质、气虚质、阳虚质、阴虚质、痰湿质、湿热质、血瘀质、气郁质、特禀质,其中平和质为健康体质,其它八种体质为偏颇体质。

[0063] 舌诊是中医诊断中最为常用的一种手段,中医认为舌象与人体的脉络、脏腑有联系,并且反应了身体内在环境的寒暑虚实,阴阳盛衰。医生通过观察舌苔的颜色、质地、纹理以及形态的变化来判断病人的健康状况和对于疾病的易感性,并给出合理的建议。临床实践也证明了舌诊是一种简单有效并且无副作用的诊断方式,且舌象中蕴含着与人体体质直接相关的信息,因此将舌诊用于体质识别是可行的。2009年由中国中医药出版的《中医体质分类与判定》详细阐述了各个体质对应的舌象特征,例如书中规定“舌形胖嫩、边缘有齿痕,且舌头颜色淡或略带青暗,舌苔润滑”是为阳虚体质者的舌象特征,除此之外还介绍了体质的其他身体特征,这些身体特征可用于设计识别体质的调查问卷。但是调查问卷有很大的主观性,个体很难做出客观的选择,且量表的问题较多,需要耗费大量的时间,同时也会影响判断的准确性。随着机器学习、深度学习等人工智能方法的普及,如果能够拍摄个体的舌象图像,以数字图像形式存储,再结合人工智能技术就可以实现基于舌诊的智能化中医体质辨识,从而能够获得更加客观和准确的结果。

[0064] 现阶段利用机器学习、深度学习等人工智能技术进行图像识别已经趋于成熟。有些学者通过收集大量的带有体质标签的舌图像建立舌图像数据库,并设计算法将舌图像与其建立的数据库进行匹配,以此达到识别的目的。还有学者设计了定量化方法来提取客观化的舌象特征如舌质和舌苔的色度和饱和度、舌苔的纹理特征、平滑度以及齿痕特征等,建立基于支持向量机(SVM)、K-means、人工神经网络等机器学习方法的中医体质辨识模型。考虑到传统人工提取舌象特征的复杂性以及舌头区域分割的准确性问题,深度学习因其自动提取特征的优势被应用于体质辨识中,一些学者采用卷积神经网络的各种模型来进行舌象体质辨识。这些研究虽然推动了舌诊智能化体质辨识的发展,但是也存在一些普遍的问题,第一:深度学习需要大量的数据集作为支撑,舌图像的样本数据采集具有一定难度。第二:只是简单的使用模型进行分类并没有针对模型改进优化,导致用以上方法进行体质辨识的准确率并不高。为了应对这些问题,本申请提出了一种基于改进卷积神经网络GoogleNet模型的新方法,在采集到的舌图像数据集的基础上使用该方法进行体质辨识,本申请的主要贡献有:构建了带有体质标签的舌象数据集,通过添加批量归一化操作、改进优化Inception模块、添加注意力机制以及不同Inception结构组合使用等方法,提出一种适用于舌象体质辨识的GoogleNet网络模型(TCR-GoogleNet)。设计了消融实验验证各个改进模块对预测精度的贡献程度,并与5种图像分类的主流网络设计了对比实验,证明了本申请提出的模型在准确率和稳定性上都有较大的提升。

[0065] 2、方法

[0066] 本章节介绍了通过舌图像进行体质辨识的具体过程。第一部分简要说明了方法的具体步骤和架构。第二部分简要说明了舌图像集的构建,包括图像采集的基本流程和要求以及图像预处理的方法。第三部分介绍了舌头图像的预处理和对舌象数据集的数据增强操作。第四部分介绍了模块的改进方法以及模型整体的构建。

[0067] 2.1方法架构

[0068] 为了更好的识别舌象体质,本申请提出了TCR-Googlenet模型,通过对构建的舌象

数据集中的大量数据进行训练,以期得到较好的体质辨识结果。该方法的基础架构如图2所示,总共有五个步骤。第一步是舌头图像的采集。第二步第三步是图像预处理,包括从人脸图像中分割裁剪进而提取舌头区域,通过分割裁剪处理后可以大大提高舌象数据集的质量。第四步是将提取好的舌头图像输入到已经构建好的深度神经网络模型中进行训练和体质识别。最后分类识别出不同的体质类型。

[0069] 2.2数据集构建

[0070] (1) 常见九种体质类型特征描述

[0071] 在中医舌诊中不同的体质有不同的舌象特征,医生往往也是通过这些不同的舌象特征判断患者的体质和病理。本申请根据王琦教授提出的理论[6]将平和质、气虚质、阳虚质、阴虚质、痰湿质、湿热质、血瘀质、气郁质、特禀质九种体质的舌象特征描述作为研究对象,其特征如表1所示,这些特征描述可为深度学习网络模型对于中医体质识别的准确度提供主观判别的视觉依据。

[0072] 表1九种体质舌象特征描述

序号	中医体质类型	舌象特征描述
1	平和质	舌色淡红润泽, 苔薄白
2	气虚质	舌淡红, 舌体胖大, 边有齿痕
3	阳虚质	舌淡胖嫩, 边有齿痕, 苔润
4	阴虚质	舌质红, 瘦小, 少津
5	痰湿质	舌体胖大, 舌苔白腻, 口黏腻或甜
6	湿热质	舌质红, 舌苔黄腻
7	血瘀质	舌质紫暗, 有点、片状瘀斑, 舌下脉络发黑发紫
8	气郁质	舌质暗, 舌下静脉曲张
9	特禀质	舌质较嫩或舌体小, 舌淡胖或舌质红, 舌苔黄腻的气虚或阳虚体质为特禀质, 其舌象呈多样性, 舌象多变

[0075] (2) 舌象体质数据集构建

[0076] 深度学习网络模型模仿人脑从低级到高级抽象提取图像特征是基于大规模图像数据集,为了基于深度学习进行舌象体质识别,需要构建舌图像数据集,通过在中医院门诊采集患者的舌图像,共计1752张舌象数据,并且采集舌图像的同时通过让患者填写中医体质调查问卷确定其体质类型,进行舌图像标签标注,每张图像都经过人工处理准确的提取出舌区域图像,从而减少了原始采集的图像中其它部分对舌象体质识别的影响。舌头图像在光照明亮的自然条件下进行拍摄。舌象数据集的样本分布如表2所示,其中数值表示每个体质的舌图像数量。每个体质对应的样本如图3所示。

[0077] 表2数据集样本分布

体质类型	平和	气虚	阳虚	阴虚	痰湿	湿热	血瘀	气郁	特禀
数量	156	279	357	246	265	267	59	90	33

[0079] 2.3数据增强

[0080] 由于通过手机相机拍摄的图像,会受到光线、角度等外界因素的影响,且收集到的舌图像数量相对较少,为了防止过拟合,模型鲁棒性较低等现象,需要对收集到的图像进行数据增强。本申请主要采用了同类增强和混类增强两种方式组合使用对数据集进行了增强。

[0081] (1) 同类增强

[0082] 同类增强主要是指对相同标签的图像使用数据归一化、水平或垂直翻转、按照比

例放大或缩小图像、随机旋转、随机错切变换、亮度变化以及像素填充等方法进行数据集扩充,颜色也是舌象体质辨识中的重要特征,故不对舌象数据集做颜色通道方面的变换。图4是一组数据同类增强后的图片。

[0083] (2) 混类增强

[0084] 混类增强的方式是了解决样本数据集类间不平衡问题,例如特禀质是并不常见的体质类型,因此样本数据相对较少,本申请采用Mixup混类增强方式对两类不同标签的样本数据进行混合,构建虚拟的训练样本丰富舌象数据集。其数学表达式如(1)所示:

$$\hat{x} = \lambda x_i + (1 - \lambda) x_j$$

[0085]

$$\hat{y} = \lambda y_i + (1 - \lambda) y_j \quad (1)$$

[0086] 其中 (x_i, y_i) (x_j, y_j) 是随机选取的两个样本及对应标签, λ 为服从beta分布的随机数,且 $\lambda \in [0, 1]$ 。图5是Mixup随机生成两次的效果图。

[0087] 2.4舌图像体质识别模型构建

[0088] 本申请改进了GoogleNet模型用于舌象识别,GoogleNet是2014年ILSVRC竞赛的获胜者,属于深度卷积神经网络模型。自VGG模型发展以来,卷积神经网络朝着更深层次的架构发展,包含越来越多的网络层和参数,然而一味地增加网络必然会对计算资源造成负担。在GoogleNet中提出的Inception模块有效的缓解了加深网络造成的梯度消失和梯度爆炸问题,提高了模型的运算效率,本申请提出了一种基于GoogleNet模型改进之后的深度神经网络模型,更适合于本数据集且准确率更高。本章节主要介绍了舌图像识别模型的构建以及改进策略。

[0089] (1) 批量标准化操作 (BatchNormalization, BN)

[0090] 神经网络中随着参数的不断更新,中间每一层输入的数据分布往往会和参数更新前有较大差异,导致网络需要不断去适应新的数据分布,会加大模型训练的难度。2015年提出的BN操作有效的解决了这个问题。本申请在模型的每一中间层(激活函数前)做了标准化处理,有效避免了梯度消失和梯度爆炸等问题,加快了网络收敛速度,优化了网络结构。BN的计算流程如下:

[0091] 假如一批输入变量为 $x_1, x_2, \dots, x_k$, 分别求其均值和方差,其数学表达式如(2)(3)所示:

$$\mu = \frac{1}{k} \sum_{i=1}^k x_i \quad (2)$$

$$\sigma = \sqrt{\frac{1}{k} \sum_{i=1}^k (x_i - \mu)^2} \quad (3)$$

[0094] 然后对输入的一批x值归一化处理到均值为0,方差为1,其数学表达式如(4)所示:

$$\hat{x} = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}} \quad (4)$$

[0096] 式(4)把 x_i 从梯度消失的边缘矫正到梯度较大的区域,增加了训练速度,其中 ε 是一个很小的值,为了避免方差分母为0且不对计算造成影响。

[0097] 最后对归一化后的值进行重构,其数学表达式如(5)所示:

[0098]
$$y_i = \alpha \hat{x}_i + \beta \quad (5)$$

[0099] 其中 α 、 β 是两个可以学习的变量,目的是为了补偿模型的非线性表达能力,恢复出这层网络所要学到的分布。

[0100] (2) Inception结构残差设计

[0101] He在2015年提出了残差结构,通过理论和实践证明了特征加性合并的优点。在图像识别中,残差结构能够有效避免深度加深带来的网络退化问题,其同时也解决了梯度问题,使得网络的性能也得到提升。本申请借鉴此思想将残差融入到Inception结构中,对Inception结构进行了改进优化,将输入特征与Inception结构输出的特征进行了特征融合,图6为本申请加入残差的Inception模块,命名为Inception B。

[0102] (3) Inception结构卷积分解和替换

[0103] 在Inceptionv3中提出了卷积的分解和替换操作:将大卷积核分解为对称的小卷积核;分解为不对称的卷积核。第一种方法是用两个3 x 3的卷积核替换一个5 x 5的卷积核。第二种方法是将n*n的卷积核替换成1*n和n*1的卷积核然后进行堆叠。这两种方法都能在感受野不变的情况下,降低参数量和增加更多非线性能力。故本申请中采用这两种方法对Inception模块进行改进,将Inception模块中的3 x 3卷积分支分解为1 x 3和3 x 1卷积,将5 x 5卷积分解为3 x 3卷积和1 x 3,3 x 1卷积,并在每个分支后加入了Batch Normalization,有效减少了模型的参数量。图7为用该方法改进之后的Inception结构,命名为Inception C。

[0104] (4) SE注意力模块

[0105] SE模块是由Squeeze和Excitation两个操作组成,其中Squeeze是全局信息的低维嵌入,是一个压缩的过程,Excitation是一个映射变换的过程,SE模块主要探索通道之间的关系,对网络通道赋予权重来增强网络的学习能力,又称通道注意力。图8是SE模块的示意图。

[0106] (5) Inception结构融合与模型构建

[0107] 本申请在GoogleNet原模型的基础上,保留了没有改进的Inception模块,我们将其命名为Inception A如图9所示。我们把Inception A和上文提出的Inception B和Inception C结构进行融合,应用到改进之后的GoogleNet网络模型中,浅层中加入InceptionA模块,中层加入Inception B模块,深层加入Inception C模块,其中A和B模块分别用了两次,C模块对特征的提取和感知能力更强,所以用了五次。

[0108] 综合上述方法和改进方式,提出了改进之后的舌图像体质辨识模型整体结构图如图10所示。

[0109] 3、实验设计与结果分析

[0110] 本实验采用win10操作系统,CPU是Intel (R) Core (TM) i9-7900X3.30GHz,GPU是GeForce RTX 3090,Python版本为3.8.4,Cuda和Cudnn版本分别为11.2和8.1.0,使用tensorflow2.5.0框架为实验环境。

[0111] 3.1参数选择和优化

[0112] 深度学习模型训练过程中,参数的选择往往影响着模型的准确率,在参数的选择和优化过程中,本申请采用多次消融实验进行微调,在利用本模型进行训练时采用批量训练方式,批次大小(Batch size)为32,共进行了100轮Epoch,优化函数采用Adam算法

(Adaptive Moment Estimation自适应动量的适应性矩估计),动量为0.9,权重衰减值为 1×10^{-5} ,初始学习率为0.0001。为了进一步防止过拟合,将Dropout值设为0.3,即随机丢弃30%的结点,分类器采用SoftMax。

[0113] 3.2评价指标

[0114] 为了评价模型的优劣,选择模型达到收敛的轮数(E)即收敛速度、混淆矩阵、损失值(L)、Top-1准确率(Acc)作为模型训练中的评价指标,其中模型收敛需要的轮数意味着相同条件下对模型收敛速度的评价,体质识别是多分类问题,损失函数采用多分类的交叉熵损失函数(cross entropy loss),其损失值表达式为:

$$[0115] \quad L = -\frac{1}{K} \sum_{k=1}^K \sum_{i=1}^N y_i \cdot \log p_i \quad (6)$$

[0116] 式中 y_i 为第i个样本对应的真实标签; p_i 为模型训练该图像的预测值;N为体质类别总数;K为舌象样本总数。

[0117] Top-1准确率反映的是识别结果中识别正确的图像占全部识别图像的比例,反映了模型对数据集的训练效果,其数学表达式为:

$$[0118] \quad Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (7)$$

[0119] 其中TP(True positive)为模型预测为正的正样本;TN(True negative)为模型预测为负的负样本;FN(False negative)为模型预测为正的负样本;FP(False positive)为模型预测为负的正样本。

[0120] 3.3消融实验

[0121] 消融实验作为模型评估的常用方法,是指通过删除网络模型中的某一部分以便于我们更好的了解改进的网络模块对网络模型精度的影响和贡献程度,进一步说明改进算法的有效性,本申请采用消融实验对各个改进策略的有效性进行验证,分析各个改进的网络模块对模型精度的影响。

[0122] 本实验从添加批量标准化操作、分别使用和组合使用Inception A、InceptionB、Inception C结构以及添加注意力机制这几个方面组合设计实验,共设计了7组实验,编号为1-7,实验设计如表3所示。

[0123] 表3消融实验设计方案

实验 编号	批量标准化 操作 (BN)	Mixup 样本增强	Inception A	Inception B	Inception C	SE 模块
[0124]	1	x	x	√	x	x
	2	√	x	√	x	x
	3	√	√	√	x	x
	4	√	√	x	√	x
	5	√	√	x	x	√
	6	√	√	√	√	x
	7	√	√	√	√	√

[0125] 表4消融实验结果对比

	实验编号	Top-1 准确率/%	损失值
[0126]	1	75.9	0.882
	2	77.5	0.826
	3	78.9	0.785
	4	76.4	0.853
	5	83.8	0.674
	6	86.1	0.589
	7	87.4	0.501

[0127] 实验结果对比如表4所示,通过各组实验结果可得出以下结论:

[0128] (1) 批量标准化操作 (BN)

[0129] 实验1只用了Inception A模块且未对模型做任何改动,为GoogleNet原模型,在实验2中添加了BN操作,对比结果,实验2比实验1的Top-1准确率提升了1.6%,这说明BN操作优化了网络结构在一定程度上提升了模型准确率。

[0130] (2) Mixup

[0131] 实验3中使用了Mixup混类增强方法对数据集进行了扩充,对比实验2准确率和损失值都有不同程度的优化,与同类数据增强组合使用一定程度上弥补了舌象数据较少的缺陷,在几乎无额外计算开销的同时使模型更好的拟合了数据集。

[0132] (3) Inception结构

[0133] 实验4将模型中的Inception A模块全部替换为加入残差结构的Inception B模块,但是模型的准确率却下降了2.5%左右,证明一味的堆叠残差结构来加宽网络并不能很好的适应本申请的数据集。

[0134] 实验5全部使用Inception C模块,Top-1准确率提升显著,说明通过卷积分解和替换加深的网络结构,对舌象特征信息的提取有着更好的效果,减少网络模型参数的同时提高了识别精度。

[0135] 实验6综合了实验2、实验4和实验5的操作,采用了三种Inception结构融合的方式来改进模型,同时加宽和加深了网络模型,对舌象特征的提取更加全面,其准确率和损失值相比单独使用Inception结构效果更好。

[0136] (4) SE模块

[0137] 实验7是该研究改进之后的最终模型,从表中可以看出,相比之前的实验,加入SE模块Top-1准确率能够提升1.3%,从而说明引入注意力机制可以在不增加网络深度、宽度和不明显增加网络参数的同时提高了模型的Top-1准确率。

[0138] 3.4实验结果与分析

[0139] 基于上述模型的构建,我们把数据集的85%作为训练集,15%作为测试集,划分数据集时保证数据随机可复现,本申请从混淆矩阵、本模型训练集和测试集指标以及模型对比三个方面分析本申请模型的实验效果。

[0140] (1) 混淆矩阵

[0141] 将本模型应用到舌象数据集的测试集上,得出的混淆矩阵如表5所示。其包括了测试集中9种体质图像中预测正确和预测错误的样本数量。在混淆矩阵中,主对角线上的数值是预测正确的样本数量,其他位置的数值则为预测错误的样本数量。

[0142] 表5测试集的混淆矩阵

	体质类型	平和	气虚	阳虚	阴虚	痰湿	湿热	血瘀	气郁	特禀
[0143]	平和	15	2	0	1	0	0	0	0	0
	气虚	3	32	2	1	0	0	0	1	1
	阳虚	1	4	49	1	0	0	0	0	0
	阴虚	2	0	0	29	1	1	0	1	1
	痰湿	1	0	2	0	36	1	1	0	0
[0144]	湿热	0	1	0	2	1	37	0	0	0
	血瘀	1	0	0	0	1	0	9	0	0
	气郁	0	1	0	2	0	1	0	12	1
	特禀	0	2	0	0	0	0	0	0	2

[0145] 从表中可以得知,气虚质和阳虚质在分类过程中容易造成混淆,原因是这两种体质可能有较为接近的舌象特征,类间差异较小,所以模型易将这两种体质混淆。血瘀质、气郁质的舌象特征明显区别于其他体质,分类效果较好。其中特禀体质的识别率是最低的,因为特禀体质的样本很少,在现实中拥有特禀体质的人太少了,很难收集到大量的训练样本且特禀质的舌象特征多变。痰湿质和湿热质的识别准确率最高,得益于其训练样本充足,模型能够充分的提取舌象特征。

[0146] (2) 训练集和测试集识别结果

[0147] 图11为训练集和测试集训练过程中的准确率曲线和损失值曲线,训练准确率曲线(1)可以反映随着Epoch轮数的增加,模型预测的精度变化趋势。训练损失值曲线(2)表示随着Epoch轮数的增加,模型预测值与真实值之间的偏差和波动情况,损失值越低模型的预测精度越高,说明模型预测出错的概率越小。本申请模型的训练集准确率达到92.2%,损失值为0.375,测试集准确率达到87.4%,损失值为0.501。相比现阶段的其他同类型研究,本模型有着较高的识别准确率和稳定性。

[0148] (3) 模型对比结果

[0149] 为了验证本申请模型的有效性和改进之后的效果,本申请选取了图像识别的五种主流卷积神经网络模型来进行对比实验,图12是本申请模型与VGG-16、resnet-50、mobilenet-v3-large、mobilenet-v3-small、GoogleNet原模型五种模型针对测试集训练的准确率折线图对比,以示与其他模型的区别。表3直观列举了6种模型测试集的Top-1准确率、损失值和达到收敛的Epoch轮数(收敛速度)。我们把测试集的准确率当作模型对比的标准,因为测试集能更好的表征模型对数据集的拟合能力,对模型的训练精度有较为准确的评估。在模型评估中,模型预测精度的优先级大于收敛速度。

[0150] 在本实验和对比实验都采用从头训练神经网络模型的方式,所以初始准确率较低,但是本模型能够以较快的速度达到最高的Top-1准确率,为87.4%。虽然本申请模型对比GoogleNet、mobilenet-v3-large、mobilenet-v3-small这3种模型收敛速度较慢,但是本模型的Top-1准确率比GoogleNet模型高11.5%,比mobilenet-v3-large模型高17.1%,比mobilenet-v3-small模型高19.5%。mobilenet-v3模型的优势在于采用了深度可分离卷积,大幅度减少了运算量和参数量,所以mobilenet-v3系列在收敛速度这一方面非常有优

势,但是在本数据集上收敛之后震荡较大,模型的稳定性不佳,准确率低。加入残差的Resnet-50网络模型在Top-1准确率和损失值上都有较好的表现,说明融合残差结构对舌象特征有较好的表征能力。所以我们在本模型构建过程中也加入了残差结构,再结合其他的改进模块,识别效果比Resnet-50的Top-1准确率提升了5.6%,loss值下降了0.112,收敛速度也有小幅提升。VGG-16模型通过加深网络层数和减少卷积核大小来提升模型训练精度,但是在对比实验中效果却并没有达到预期,准确率仅为72.8%,与本申请模型实验结果有较大差距,证明单纯的加深网络结构并不能很好的提取舌象特征。在本申请网络模型中我们不仅从网络深度、网络宽度改进了模型,还通过消融实验证明了在此基础上添加的多个改进模块在一定程度上都提升了模型的预测精度。

[0151] 表6模型结果对比

	模型网络结构	测试集 Top-1 准确率	损失值	达到收敛的 Epoch(收敛速度)
[0152]	GoogleNet	75.9%	0.688	65
	VGG-16	72.8%	0.752	80
	Resnet-50	81.8%	0.613	85
	Mobilenet-v3-large	70.3%	0.722	50
	Mobilenet-v3-small	67.9%	0.813	65
	本申请模型	87.4%	0.501	75

[0153] 综合来看,本申请改进之后的模型在识别精度方面较其他模型提升较大,其收敛速度适中,但是模型预测精度的优先级大于收敛速度,故本申请提出改进的模型在舌象数据集上的识别效果是优于其他5种模型的。

[0154] 本申请基于GoogleNet提出了TCR-GoogleNet模型,构建了成熟的舌象数据集,同时我们也应继续采集图像,扩大舌图像数据集,以便为后续的研究工作提供更强大的数据支持。本模型在网络的宽度和深度优势以及各种改进的加持下,表现出对舌象数据集较强的特征提取能力和较好的识别精度。针对普通舌图像可以有效识别平和质、气虚质、阳虚质、阴虚质、痰湿质、湿热质、血瘀质、气郁质、特禀质这九种中医体质,在一定程度上缓解了体质分类需要丰富的专家经验这一困难,对比调查问卷的方式大大缩短了体质辨识的时间,为舌诊的客观化进程提供了借鉴和思路。

[0155] 基于舌诊的体质辨识未来还有更大的发展空间,舌象数据集可以有更丰富的扩充,在此基础上的识别准确率还可以得到更大的提升。由于每个人的体质具有多样性,一个人可能包含多种体质,因此后续研究可以围绕舌图像体质辨识的多标签学习展开。除此之外舌诊的研究不仅仅可以用来识别体质,现在已经有研究将舌诊和西医的病症结合起来防治疾病,但这方面的研究还有待进一步的发展。

[0156] 本说明书中各个部分采用递进的方式描述,每个部分重点说明的都是与其他部分的不同之处,各个部分之间相同相似部分互相参见即可。

[0157] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本申请中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本申请所示的实施例,而是要符合与本申请所公开的原理和和特点相一致的最宽的范围。

获取待识别的舌象图像，使用训练好的改进的卷积神经网络GoogleNet模型对待识别的舌象图像进行识别，得到舌象的识别结果

将舌象图像的识别结果在移动终端进行展示

图1

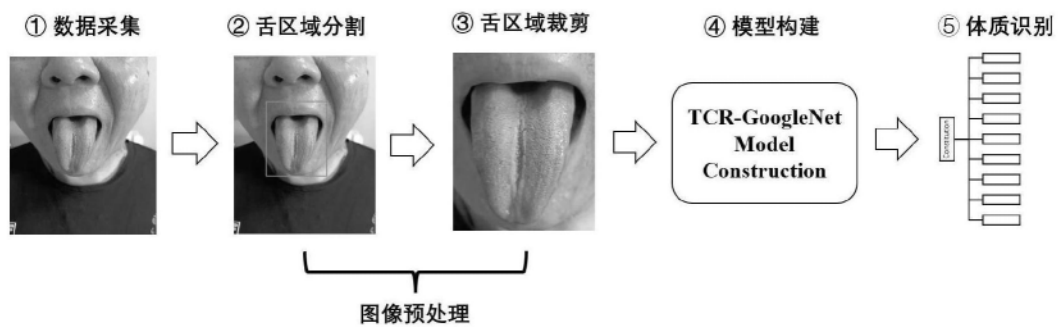


图2



图3

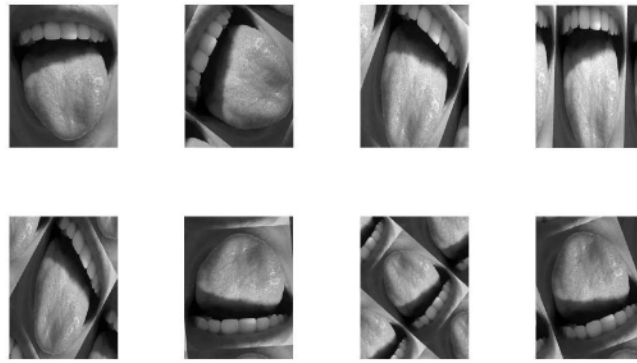


图4



图5

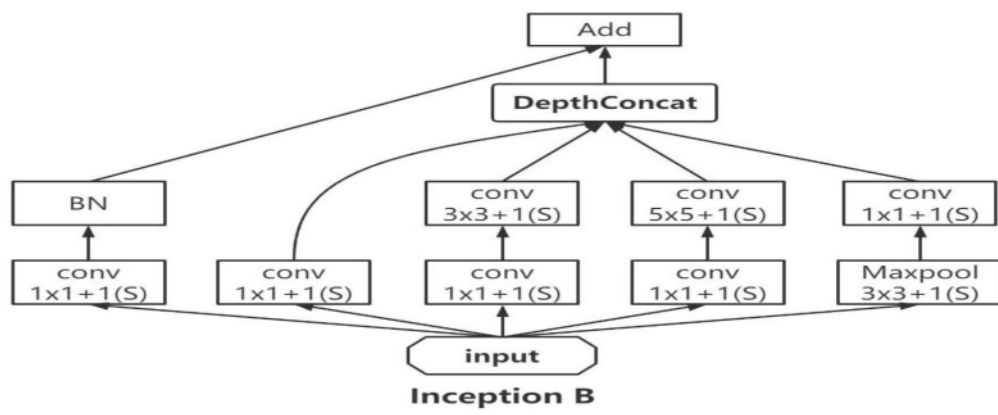


图6

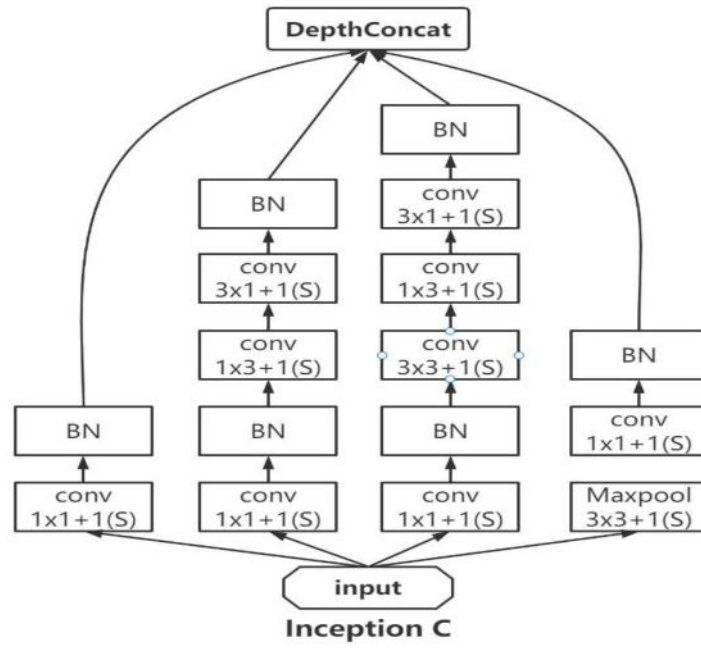


图7

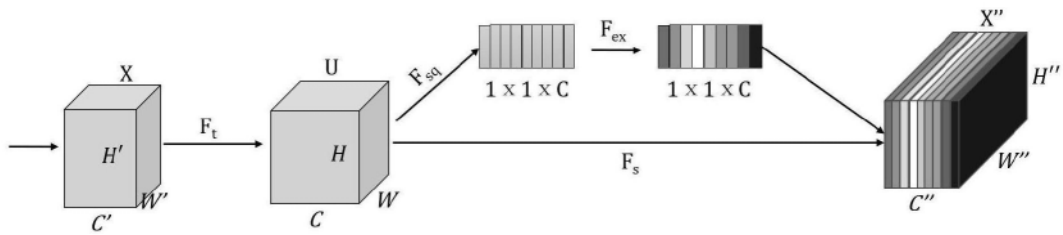


图8

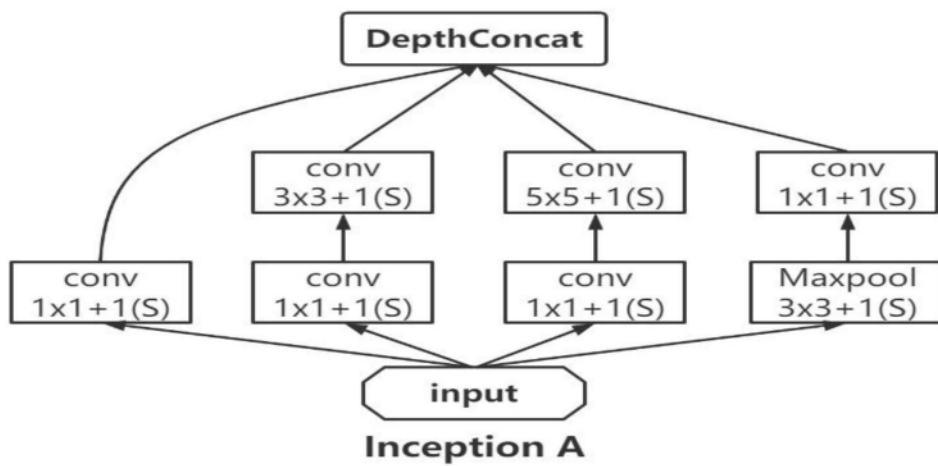


图9

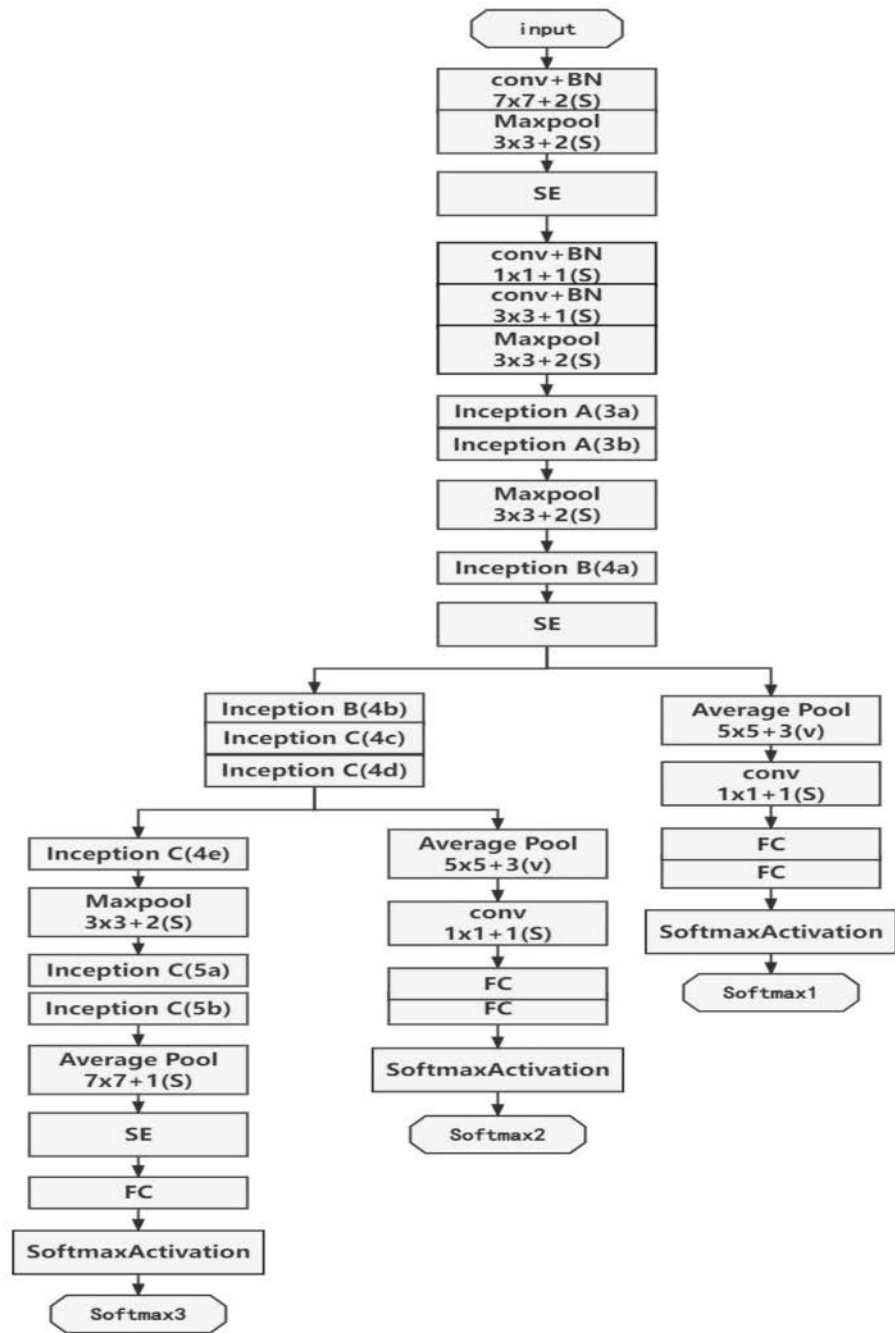


图10

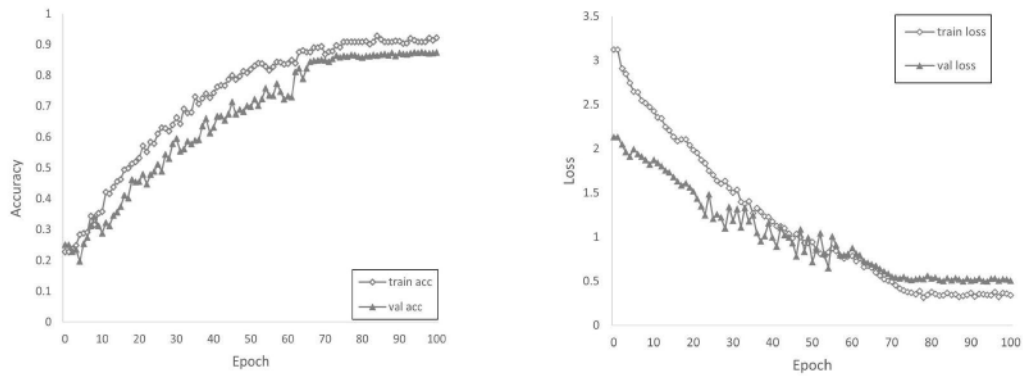


图11

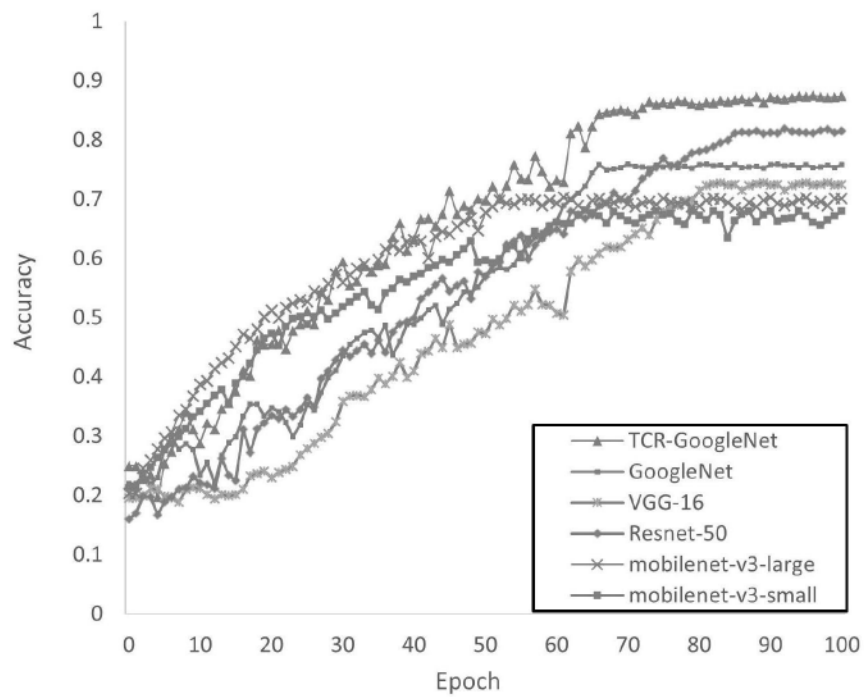


图12