

◆ 生物医学信息计算 ◆

基于 ME-ANet 模型的糖尿病视网膜病变分级*

程小辉^{1,2**},李贺军^{1,2},邓昀^{1,2},陶小梅¹,黎辛晓^{1,2}

(1. 桂林理工大学信息科学与工程学院,广西桂林 541006;2. 广西嵌入式技术与智能系统重点实验室,广西桂林 541006)

摘要:糖尿病视网膜病变(Diabetic Retinopathy, DR)是一种致盲率很高的眼科疾病。不同病变等级的视网膜图像之间差异小且病灶点分布无规律。针对现有深度模型对 DR 中的相似病灶点识别率低,严重影响模型分类精度的问题,本研究以深度学习为基础,构建新的模型架构进行训练,提出一种集成 MobileNetV2 和 EfficientNetB0 深度模型的注意力网络:ME-ANet。模型集成分为头部和主干两部分,将深度模型的浅层部分融合构成网络的头部,训练时采用迁移学习的策略对网络模型参数进行初始化,减少训练中的过拟合问题。主干部分利用上述两种模型的核心结构,设计 3 个阶段集成模块进行特征提取。同时设计全局注意力机制(Global Attention Mechanism, GAM)并分别嵌入到 3 个阶段的集成模块中。模型的改进加速了网络的收敛速度,该网络模型实现了对图像浅层信息的特征融合提取,减少了微病灶特征信息在训练过程中的卷积丢失问题,模型分类精度进一步得到改善。通过模型集成构建特征提取主干网络,提高了模型对低级特征信息的学习,注意力机制抑制非病变特征信息,强化典型病灶特征学习,从而实现细粒度分类,进一步提升了模型分类性能。

关键词:糖尿病视网膜病变 迁移学习 集成模块 注意力机制 特征融合

中图分类号:TP3 文献标识码:A 文章编号:1005-9164(2022)02-0249-11

DOI:10.13656/j.cnki.gxkx.20220526.004

糖尿病视网膜病变(Diabetic Retinopathy, DR)是一种发病率和致盲率都很高的糖尿病并发症^[1],近年来成为造成视力模糊甚至失明的主要原因之一。糖尿病视网膜病变的主要病理特征有微动脉瘤、出血点、软性渗出物以及硬性渗出物^[2]。根据眼底图像病变的类型和程度,DR 可分为 5 个阶段:无(DR0)、轻度(DR1)、中度(DR2)、重度(DR3)和增生性(DR4)。

通常,DR 的早期阶段症状不明显,患者往往会忽视 DR 早期对自身的影响而错过最佳的治疗时间。目前,糖尿病视网膜病变的诊断需要专业的眼科医生根据自身的诊断经验来判断,由于糖尿病视网膜图像中不同等级之间的差异性很小,再加上专业眼科医生就诊经验的不同,可能会出现误诊、漏诊等情况^[3,4]。利用深度学习来提高 DR 模型分类准确性和诊断效

收稿日期:2022-02-13

* 国家自然科学基金项目(61906051),广西自然科学基金项目(2018GXNSFAA281235)和广西中青年骨干教师基础能力提升项目(2018KY0248, 2020KY06026)资助。

【作者简介】

程小辉(1961-),男,教授,博士生导师,主要从事嵌入式系统、物联网和人工智能研究,E-mail:cxiaohui@glut.edu.cn。

【**通信作者】

【引用本文】

程小辉,李贺军,邓昀,等.基于 ME-ANet 模型的糖尿病视网膜病变分级[J].广西科学,2022,29(2):249-259.

CHENG X H, LI H J, DENG Y, et al. Study on Grading of Diabetic Retinopathy Based on Me-ANet Model [J]. Guangxi Sciences, 2022, 29(2): 249-259.

率对预防糖尿病视网膜病变有重大意义^[5]。深度卷积神经网络(Convolutional Neural Network, CNN)对图像的特征提取和学习能力突出,被广泛应用在医学图像处理领域,特别是在视网膜图像分类、分割和目标检测任务中表现出色^[6-8]。Gulshan等^[9]利用Inception-v3模型框架并采用12多万张糖尿病视网膜眼底图像来检测病变,由于数据量比较充分以及眼科专家对眼底图像的筛选,该模型取得了良好的效果。李琼等^[10]在AlexNet网络的基础上,在每一个卷积层后紧跟一个批量归一化层来加深网络,用于视网膜图像的特征提取,训练时采用迁移学习的策略,该模型具有较好的鲁棒性和泛化性。Sharma等^[11]提出一种深度CNN架构来对非糖尿病视网膜病变和糖尿病视网膜病变眼底图像进行分类,具有较好的效果。丁蓬莉等^[12]提出一种密集型CompactNet网络模型,该模型使用AlexNet的浅层结构参数,由于缺少标记的视网膜眼底图像,模型在训练过程中提取特征信息不足,导致分类准确率不高。Zhou等^[13]利用多模块结构对高分辨率的眼底图片进行学习,训练时以分类和回归的方法对标签进行预测,最后利用EyePacs数据集对模型有效性进行测试,并取得了不错的效果。但是该模型仅适合像素较高的眼底图片,对于分辨率较低的图像无法保证其精度。郑婷月等^[14]采用编码器-解码器对称全卷积结构,该结构将高层语义信息与低层特征信息进行特征融合,并利用空洞卷积构建多尺度空间金字塔结构,同时引入残差结构块进一步提取图像特征信息,获得了较好的分类准确率。近年来,基于深度学习的方法对DR等级进行分类取得了一定的进步。然而,在糖尿病视网膜病变过程中,病理图像质量差异大且不同DR等级图像之间相似度较高,增加了DR分类模型在无标记病变区域条件下自动精确分类的难度^[15]。不同的眼底图片在颜色和纹理方面一致性较高,在模型训练过程中容易混淆,导致模型难以对类别特征作出精确判断。除此之外,眼底数据集的高度不平衡性也严重制约着模型的性能。如何设计合理高效的深度模型实现高精度的识别分类,进而推动深度学习在医学图像中的可靠应用成为关键的一环。为了更加充分地提取图像特征信息,增加模型对微病灶点的学习能力,本研究提出一种基于集成MobileNetV2和EfficientNetB0深度模型的注意力网络(ME-ANet)进行细粒度学习分类,旨在进一步强化深度模型对DR病变特征的学习,提高模型的准确率和鲁棒性。

1 模型设计

1.1 糖尿病视网膜病变分类流程

算法的工作流程主要分为数据处理、模型训练和模型性能评估3个阶段。在数据处理阶段,主要包括数据预处理和数据扩增两个步骤,解决原始数据集中类别分布不均衡的问题,对预处理后的眼底图像按照一定的比例进行划分,将扩增后数据集划分为训练集、验证集和测试集,比例为8:1:1。模型训练阶段采用ME-ANet集成网络,将眼底图像尺度大小固定为 $224 \times 224 \times 3$ 作为网络的输入层数据。为了提高模型的收敛速度,模型训练时采用迁移学习的方式对集成模型头部进行权重初始化,最后输出使用softmax分类器完成DR五分类预测任务;在模型性能评估阶段,利用训练好的模型进行测试集的分类,计算相应的评价指标并作对比分析,对模型进行评价总结。DR分类系统的工作流程如图1所示。

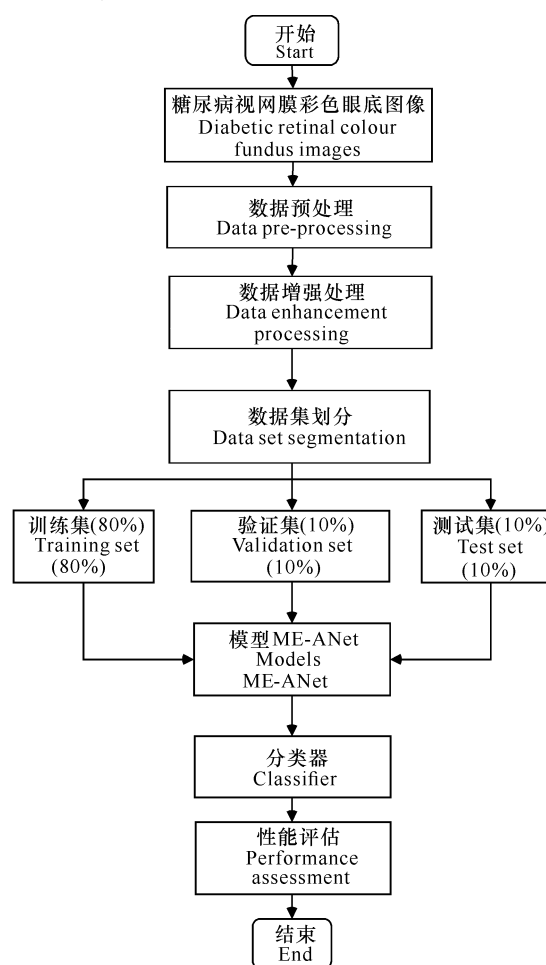


图1 DR分类系统流程图

Fig. 1 Flowchart of DR classification

1.2 特征提取子网络

为了强化学习眼底病变图像中的病灶特征,提高模型的分类精度,本研究采用集成的方式,将轻量化的深度模型架构 MobileNetV2 与高效神经网络 EfficientNetB0 部分结构集成组合为特征提取主干,并将其命名为 ME 特征提取主干。MobileNetV2 是由一系列具有线性瓶颈层的倒残差结构(Inverted residual)堆叠而成。模型利用倒残差结构代替传统卷积顺序叠加的方式,该结构采用先升维再降维的信息处理方式,另外,深度可分离卷积替换传统卷积,大大减少了模型的参数。捷径分支位于两端较窄的瓶颈层之间,中间部分采用轻量级的深度卷积来提取数据特征,卷积过程中均使用 ReLU6 激活函数。为了减少非线性激活函数对压缩通道特征信息的丢失,维持网络的特征表达能力,在最后的卷积操作后使用线性激活函数 Linear 进行激活。具有线性瓶颈层的残差结构 Bottleneck 如图 2 所示,带跳跃连接的残差结构为优化后的结构,均采用 1×1 卷积对通道进行升维和降维处理。在模型训练过程中,由于跳跃连接的存在,深层的梯度更容易传回至浅层,该结构避免了深度网络反向参数更新引起的梯度弥散问题。

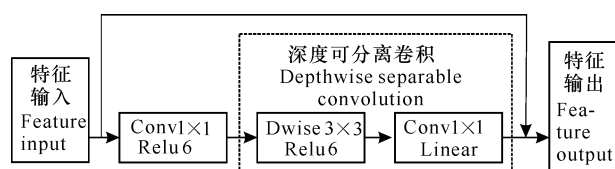


图 2 Bottleneck 残差块

Fig. 2 Bottleneck residual block

当利用深度模型对处理完成后的糖尿病视网膜图像进行训练时,增大模型的深度往往能够提取到更加丰富的特征,但网络深度过深容易出现梯度消失,

导致模型训练失败。增加网络的宽度、增加特征图的通道数虽然可以获得更高细粒度的特征信息,但宽度过大且深度较浅的模型很难学习到高层次的语义特征。高分辨率的图像输入虽然可以使模型学习到更加细粒度的信息,能够扩大网络的感受进而提升网络性能,但是会增大模型的计算量。EfficientNetB0 采用自适应算法平衡模型深度、宽度和分辨率三者之间的比例关系,通过比例系数对模型进行 3 个维度的自适应缩放,合理优化网络结构。模型核心结构为移动翻转瓶颈卷积(Mobile Inverted Bottleneck Convolution, MBConv)模块(图 3)。输入特征首先经过第一次 1×1 卷积和深度卷积(之后都进行 BN 和 Swish 激活操作),需要注意的是,第一次卷积是为了扩展通道维度,扩展比例设置为 n ,当 $n = 1$ 时,不进行通道扩展,直接进行 $k \times k$ 的深度卷积操作。接着将输出特征输入到 SE 注意力模块结构中,再通过卷积(第二次 1×1 卷积之后只进行 BN 操作)进行降维,调整输出通道维度使之与输入通道维度相等。最后进行随机失活处理,同时在输入和输出之间设置捷径分支,该结构有利于模型融合更多的特征信息,而只有当输入 MBConv 结构的特征矩阵与输出的特征矩阵形状相同时才有捷径分支的存在。

图像分类、图像分割和目标检测任务中,SE 注意力模块效果显著。该模块由一个全局平均池化(Global Average Pooling, GAP)、两个全连接层构成。第一个全连接层的输出通道数是输入该 MBConv 特征矩阵 channels 的 $1/4$,且使用 Swish 激活函数;第二个全连接层的节点个数等于深度卷积层输出的特征矩阵通道数,且使用 Sigmoid 激活函数。假设一个输入特征 $H \times W \times C$,首先通过全局平均池化

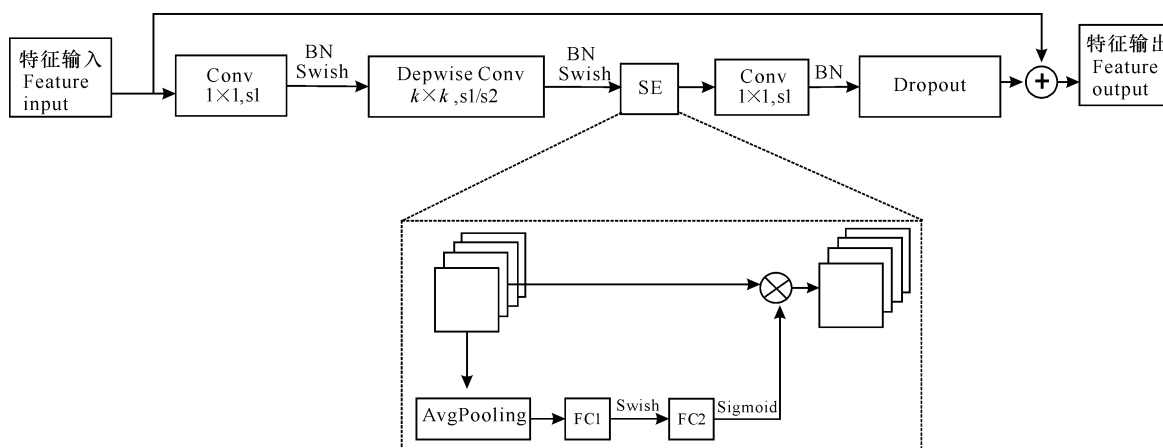


图 3 MBConv 模块

Fig. 3 MBConv block

和全连接层将其拉伸为一个 $1 \times 1 \times C$ 的特征向量, 然后与原始图像相乘, 对通道特征重新加权计算, 从而学习更多目标特征信息。

1.3 迁移学习

在深度学习医学图像领域, 缺乏大型公开的数据集成为医疗图像处理中的难题之一。在缺少训练样本的情况下, 容易导致模型在训练过程中产生过拟合甚至出现不收敛状态, 导致训练出的模型鲁棒性较差。

为了解决视网膜图像数据量少、模型训练困难等问题, 除了对数据进行扩增处理外, 采用迁移学习的方式加载模型相应层的预训练权重参数有利于模型性能的提升。由于浅层卷积能够学习到更加丰富的低层语义信息, 本研究将加载 MobileNetV2 和 EfficientNetB0 已经在 ImageNet 数据集上训练好的模型参数, 在训练过程中, 冻结浅层模型参数, 仅保留相关卷积层, 将输入图像的特征信息和相应的标签作为输入, 通过计算预测标签与真实标签之间的误差, 自动调整网络参数。

1.4 注意力机制设计

在糖尿病视网膜病变分类任务中, 微小的病灶点如渗出物和微动脉瘤等类别之间差异较小, 造成视网膜不同类别之间差异非常细微, 模型难以准确识别相

似度较高的病变图像, 为提高模型分类精度, 需要模型具备更加细粒度分类能力, 引入更加细粒度分类实现视网膜病变的分级。对于细粒度的分类来说, 如何使模型学习到有效的目标特征成为决定细粒度分类模型性能的关键。

在注意力机制中, 通道注意力机制通过学习通道注意权重, 判别每个特征通道的重要程度同时抑制信息量较小的通道, 而空间注意力机制通过学习空间注意权重说明每个空间位置的重要性, 与通道注意力机制形成互补机制。基于此, 本研究设计全局注意力机制(Global Attention Mechanism, GAM)模块来学习目标病灶特征。GAM 模块由通道注意力和空间注意力模块组成, 采用平均池化单分支结构。全局平均池化(GAP)可以实现图像通道特征的压缩, 对输入特征图的空间变化具有较强的鲁棒性。但是, 仅考虑建模评估特征图中每个通道之间的重要性过于简单, 无法准确学习特征之间的长依赖关系。由于空间位置信息决定生成空间选择性注意力图, 在计算机视觉任务中, 这些空间位置信息往往更有利于模型捕获学习目标特征。故重新设计注意力机制, 具体使用平均池化对图像通道和空间特征信息进行转换处理, 具体结构如图 4 所示。

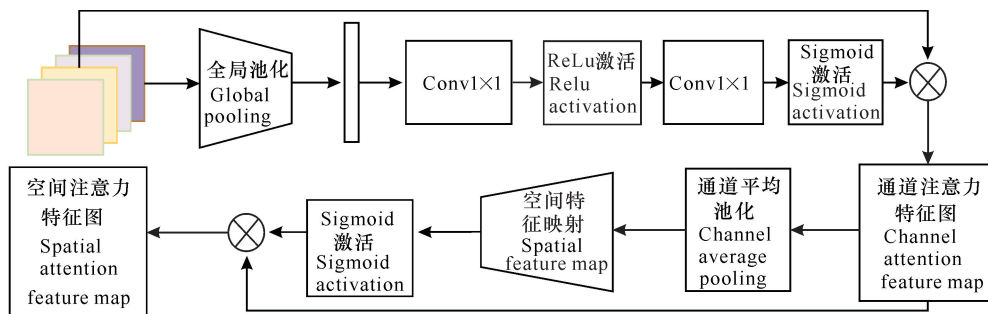


图 4 全局注意力机制

Fig. 4 Global attention mechanism

注意力网络将集成模型输出的融合特征 $F_{GAM-IN} \in R^{H \times W \times C}$ 作为输入, 利用公式(1)计算生成通道注意力特征图 $F_{CA-OUT} \in R^{H \times W \times C}$ 。

$$F_{CA-OUT} = (\sigma(\text{Conv2}(\text{GAP}(F_{GAM-IN})))) \otimes F_{GAM-IN}, \quad (1)$$

式中, $\sigma(\cdot)$ 表示 sigmoid 激活函数, 将变量值映射到 $[0, 1]$ 区间; Conv2 表示两个卷积核大小为 1×1 的卷积层; GAP(\cdot) 为全局平均池化操作, 即在通道维度上进行全局平均池化; \otimes 表示特征同位元素对应相乘。

利用公式(2)计算 GAM 的输出特征 $F_{SA-Feature} \in R^{H \times W \times 1}$, 即空间注意力特征图。

$$F_{SA-Feature} = F_{CA-OUT} \otimes \sigma(\text{C_GAP}(F_{CA-OUT})), \quad (2)$$

式中, C_GAP(\cdot) 为空间维度平均池化操作, $\sigma(\cdot)$ 为 sigmoid 激活操作。空间注意力结构简化为单链路, 针对模型的原始特征输入, 只采用平均池化操作, 更好地维护了图像的空间信息, 有利于对空间注意力权值的学习, 同时指出每个空间位置的重要性。

1.5 模型集成

在特征提取主干网由 MobileNetV2 的核心结构 Bottleneck 和 EfficientNetB0 的核心结合 MBConv 卷积结构进行相应的集成操作,其中的 MBConv6 表示其卷积扩展比例设置为 6。具体的头部集成模型 ME-Head 结构如图 5 所示,其中头部训练时采用迁移学习的方式加载与训练权重进行特征提取,提高模型的收敛时间。头部集成模型中,输入尺度为 $224 \times 224 \times 3$,首先将图像进行一步卷积操作,转化为 MB-

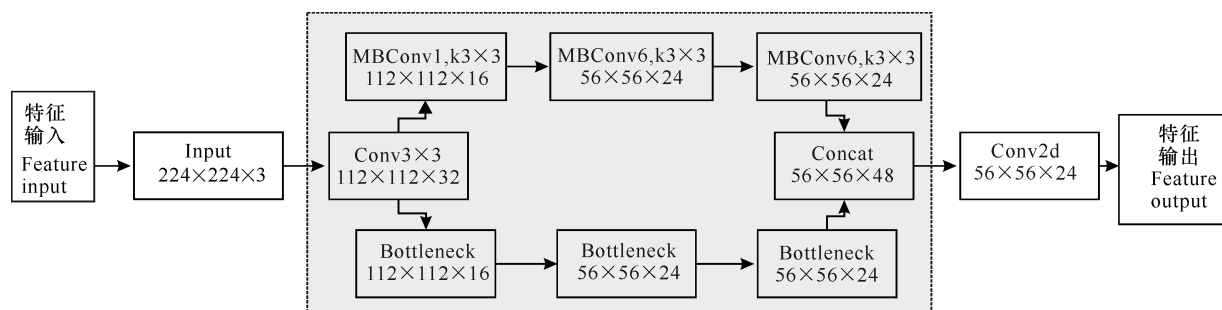


图 5 头部集成模型 ME-Head 结构

Fig. 5 Head integrated model ME-Head structure

在深度学习中,通常采用加深网络深度来获取医学图像中的深层次语义特征,而忽略模型对浅层特征的学习。为解决网络模型加深带来的浅层网络学习特征学习不充分问题,本研究设计 ME-ANet 模型(图 6)。除了特征提取头部集成外,设计模型主干 3 个阶段特征集成方案,实现充分学习不同尺度特征的效果,集成模块的 Bottleneck 和 MBConv6 结构的叠加层数分别设为 3,4,3 层。集成模型头部的输入尺度为 $56 \times 56 \times 24$,输入到模型集成的第一个阶段,经过下采样后分别得到两个 $28 \times 28 \times 40$ 的特征向量,然后使用通道融合技术进行 Concat 处理得到特征向量 $F_{Concat_1} \in R^{28 \times 28 \times 80}$,该特征输出至 GAM 模块,第二阶段的 Bottleneck 和 MBConv6 3 个模块中,特征图输入到 GAM 模块中,从通道和空间两个维度对网络权重更新进行训练优化,其中得到注意力特征图 $F_{GAM_1} \in R^{28 \times 28 \times 80}$ 。依次类推,可得到 $F_{Concat_2} \in R^{14 \times 14 \times 160}$ 、 $F_{GAM_2} \in R^{14 \times 14 \times 160}$ 、 $F_{Concat_3} \in R^{7 \times 7 \times 384}$ 、 $F_{GAM_3} \in R^{7 \times 7 \times 384}$,输出的 3 个阶段的注意力特征分别代表图像的浅层纹理特征、中间层过渡特征和深度语义特征。将 3 个阶段输出的注意力特征图进行多尺度特征融合操作,得到最后的特征输出 $F_{Final} \in R^{7 \times 7 \times 384}$ 。然后通过全局平均池化代替传统的全连接层,将特征提取网络学习到的深度特征映射到样本标记空间,不仅减少了参数的数量,还可以减轻模型过

Conv 和 Bottleneck 结构所需要的输入维度 $112 \times 112 \times 32$ 。然后分别选择两种基本模型核心结构的前两层,即对应 3 组 MBConv 和 Bottleneck 结构,进行 Concat 特征融合操作,将两部分的输出特征进行通道叠加,丰富通道特征信息,融合特征尺度为 $56 \times 56 \times 48$ 。最后再进行一次卷积运算,调整输出特征的通道维数,得到输入到下一个集成模块的特征图,尺度为 $56 \times 56 \times 24$ 。

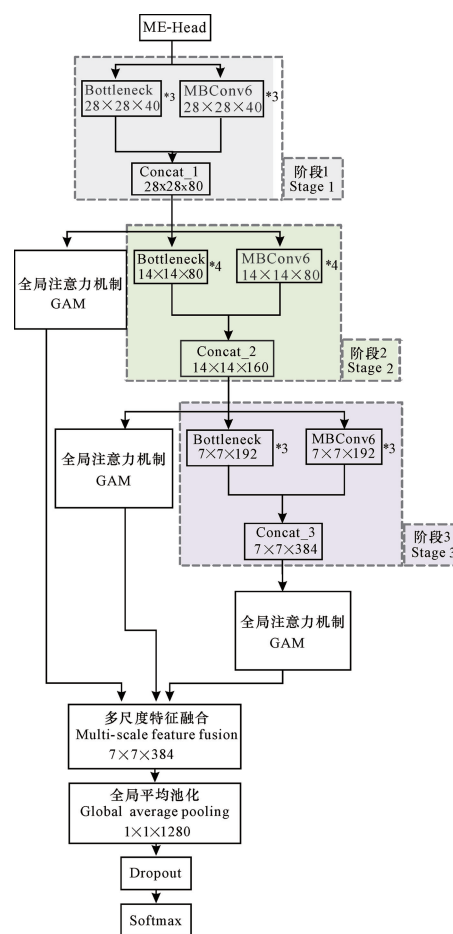


图 6 ME-ANet 模型

Fig. 6 ME-ANet model

拟合的问题。最后采用 softmax 函数完成多分类任务。

2 结果与分析

2.1 数据集

实验所用的数据集来源于数据建模和数据分析

竞赛平台 Kaggle 中的比赛 Diabetic Retinopathy Detection。数据集包含 35 126 张由专业眼科医生诊断为不同严重程度的高分辨率视网膜图片。眼科专家根据糖尿病引发的视网膜病变程度的不同,将其分为 5 个等级(图 7),分别为 DR0、DR1、DR2、DR3 和 DR4,分别对应正常、轻度、中度、重度和增殖性。

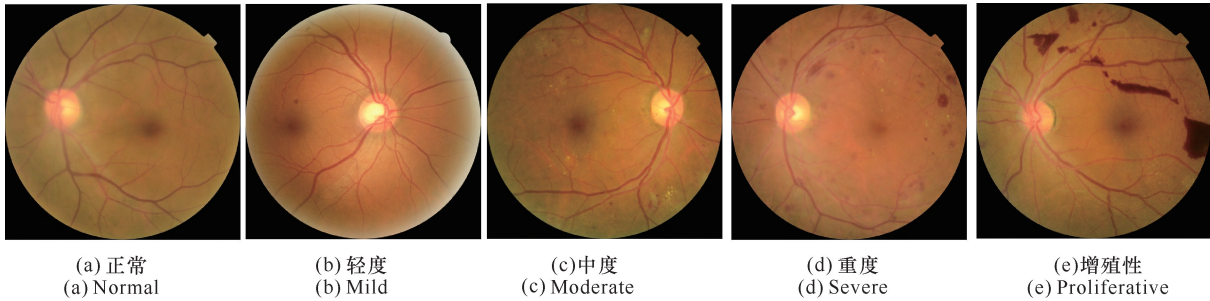


图 7 糖尿病视网膜病变图像

Fig. 7 Diabetic retinopathy images

2.2 数据预处理和扩增

由于数据集来源和成像设备不同,图像质量差别较大,存在图像尺寸及颜色差异、曝光过度、噪点较多等现象。为了提高眼底图片的质量,对原始图像进行降噪和归一化处理十分必要。同时为了适应模型训练,对分辨率大小进行统一操作,使用 OpenCV 对视网膜图像进行预处理,具体步骤如下:首先处理原始图像的边框信息,如图 8(a)所示;去除多余黑色背景区域,去除冗余特征噪声,如图 8(b)所示;使用 OpenCV 图像处理库中的 `resize()` 方法统一数据集分辨率为 224×224 像素,即进行尺寸归一化处理,具体如图 8(c)所示;缩放后的图像边缘被裁剪,为了得到与图像主体类似的轮廓,本研究以在图像中心点画圆的方式处理裁剪后的图像,处理结果如图 8(d)所示;同时为了减少亮度条件不一致因素的影响,利用高斯滤波对图像进行亮度、对比度等均衡化操作,具体公式如下:

$$I_r(x, y; \sigma) = \alpha(I(x, y) - G(x, y; \sigma) * I(x, y)) + \beta, \quad (3)$$

式中, $G(x, y; \sigma)$ 为高斯平滑函数; σ 代表标准差,用来提高图像的对比度; β 代表强度,像素值分布在 $[0 - 255]$ 。当 $\alpha = 4$, $\sigma = 256/30$, $\beta = 128$ 时,图像处理效果如图 8(e)所示。

糖尿病视网膜眼底图像 5 个等级对应的原始图像和预处理后生成的图像如图 9 所示。

视网膜数据量较小且各病变类别分布极不平衡,如果数据不进行均衡化处理,就容易导致模型出现过拟合现象。在深度学习中,通常采用 Dropout 随机失活和增加训练数据集来解决过拟合问题。数据集中每种类别数据数量如表 1 所示,健康的眼底图片占总数据集的 70% 以上。为了保证模型数据量的大致相等,提高模型的泛化能力。利用数据增强扩增训练数据是解决模型过拟合问题最直接有效的做法。

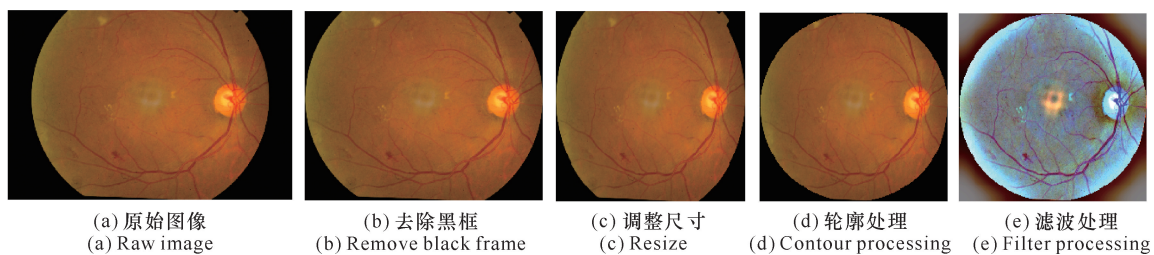


图 8 数据预处理对比图

Fig. 8 Comparison chart of data preprocessing

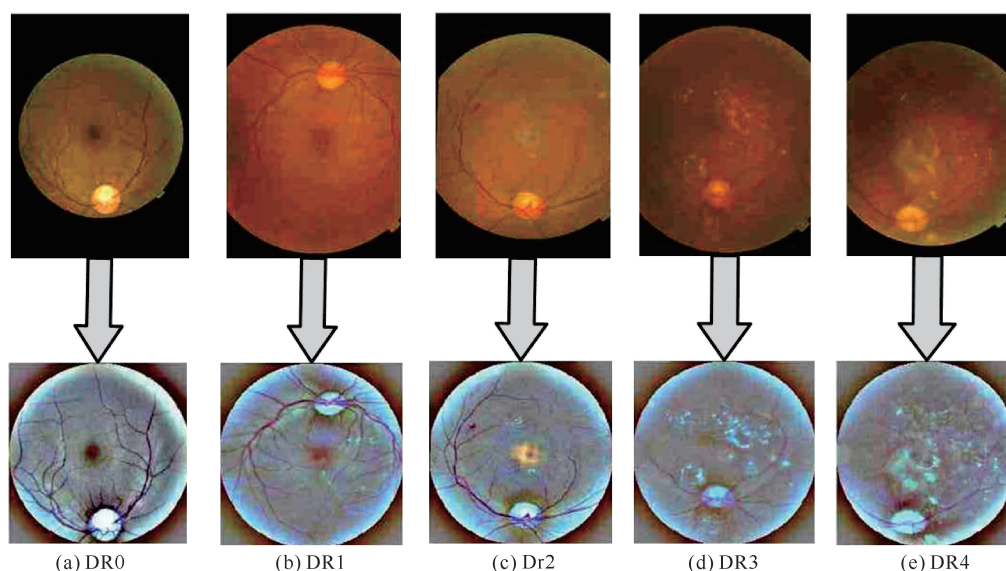


图9 5种眼底图片处理前后对比图

Fig.9 Comparison of 5 kinds of fundus images before and after processing

表1 Kaggle数据集分布情况

Table 1 Distribution situation of Kaggle dataset

病变等级 Grade of lesion	病变程度 Degree of lesion	数量 Number	比例(%) Proportion (%)
DR0	正常 Normal	25 810	73.40
DR1	轻度 Mild	2 443	6.96
DR2	中度 Moderate	5 292	15.07
DR3	重度 Severe	873	2.48
DR4	增殖性 Proliferative	708	2.01

在深度学习中,数据的增强可以分为离线增强和在线增强两种方式。离线增强就是事先处理好需要进行训练的数据,然后存储在本地,模型训练时自动加载预处理过的数据,但这种方式效率比较低,且不是实时增强。本研究采用实时数据增强的方式,通过keras框架提供的ImageDataGenerator()图片生成器进行实时数据增强,生成模型训练时所设定的batch_size大小的张量图像,且可以循环迭代,主要采用随机拉伸、旋转图像、随机水平和垂直翻转、比例缩放、随机水平和垂直移动等方法来增加样本数量。图10展示了本研究训练过程中图像随机增强效果。

2.3 实验设置与实验指标

本研究深度学习框架为Tensorflow2.0版本,采用Python3编程语言进行模型框架的搭建,在GPU型号为NVIDIA RTX 2080Ti、16 G显存的Ubuntu16.04操作系统,CUDA10.0,cudnn7.6.0的服務器上进行训练调试。将预处理后的各类数据以

8:1:1的比例进行划分,模型输入尺寸调整为 224×224 像素,训练过程中使用Adam优化器进行参数调优,初始学习率设为0.000 05,Loss计算过程中采用交叉熵(Cross Entropy,CE)损失函数,损失函数的权重衰减因子设为0.1,batch_size设置为32,一共训练30个epochs。

本研究通过分类准确率(Accuracy)、特异性(Specificity)、灵敏度(Sensitivity)和二次加权Kappa(Quadratic Weighted Kappa,QWK)系数对模型分类效果进行评估。评价指标定义如式(4)~(6)所示。

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}, \quad (4)$$

$$\text{Specificity} = \frac{TN}{TN + FP}, \quad (5)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \quad (6)$$

式中,统计定义真阳性(TP)、假阳性(FP)、真阴性(TN)和假阴性(FN)4个指标。其中,当输入正常眼底图像时,模型分类为正常则为真阳性,分类为非正常则为假阳性;当输入非正常眼底图像时,模型分类为非正常则为真阴性,分类为正常则为假阴性。

Kappa系数是统计学中度量一致性的指标,对于分类问题,一致性检验就是考察模型预测结果和实际分类结果是否一致^[13]。本研究采用二次加权Kappa(QWK)系数作为评估指标。QWK系数被设计用于衡量两个评分者在具有序列等级的标签上是否具有 consistency,并且已经被应用在糖尿病视网膜病变等级分

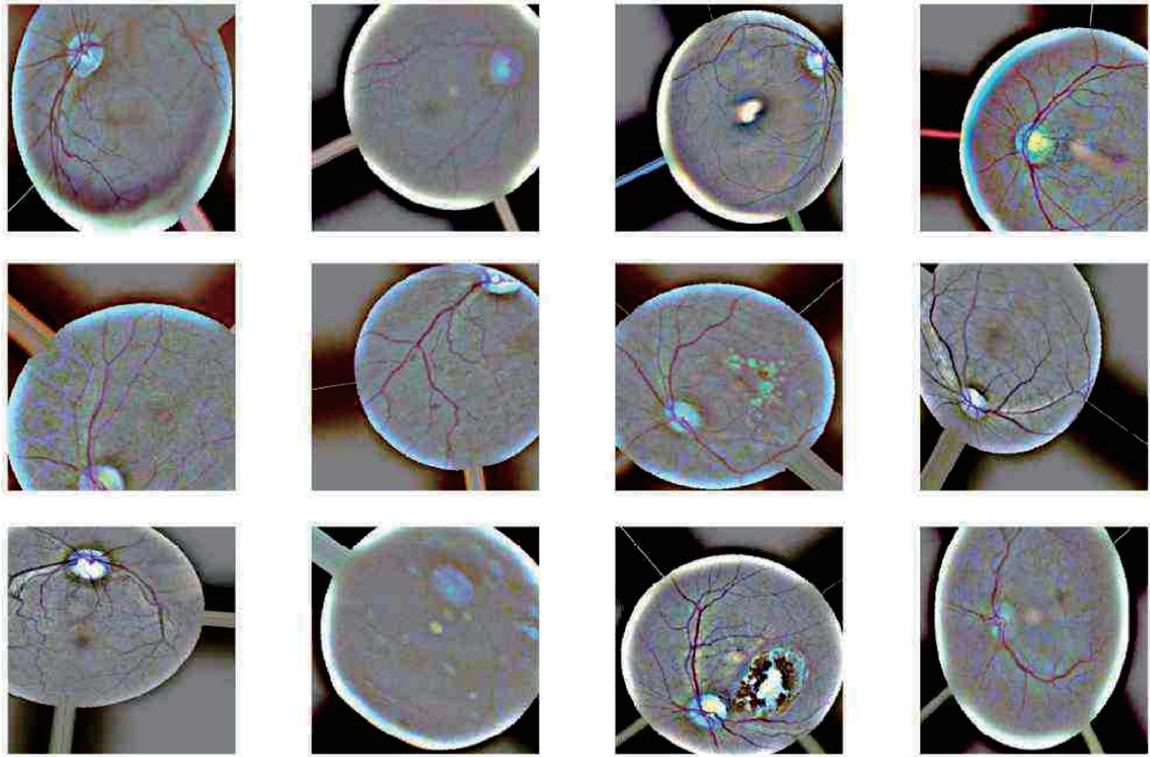


图 10 实时随机增强效果图

Fig. 10 Real-time random enhancement renderings

类中。QWK 系数的数学表达式如下:

$$k_w = 1 - \frac{\sum_{i,j} \omega_{i,j} O_{i,j}}{\sum_{i,j} \omega_{i,j} E_{i,j}}, \quad (7)$$

其中 $\omega_{i,j} = \frac{(i-j)^2}{(N-1)^2}$; $E_{i,j} = \sum_j O_{i,j} \times (\sum_j O_{i,j})^T$, $N=5$; $O_{i,j}$ 是一个 $N \times N$ 的矩阵, $O_{i,j}$ 表示将第 i 类图片分为第 j 类的样本数; ω 作为惩罚项。

2.4 实验设计与分析

为了验证所提出的模型集成的特征提取主干分类的有效性,实验采用相同的数据集对特征提取网络分别为 MobileNetV2、EfficientNetB0 以及两者之间的集成模型 ME-Net 进行训练。ME-Net 仅仅是将两模型按照图 6 的集成方式构建,只是缺少了 GAM 模块结构。实验采用迁移学习训练策略,模型加载预训练权重,结果如表 2 所示。将模型融合后作为主干网络,优于单个基础模型,评价指标均高于单个模型,表明提出的模型集成融合策略有助于提高模型的性能。

为了探究本研究所提出的 GAM 模块的分类效果,分别以 MobileNetV2、EfficientNetB0 和 ME-Net 作为基本的特征提取器,并加载初始化预训练权重,将 3 个模型的特征输出最终输入到一个 GAM 模块

中,分析模型效果。如表 3 所示,总体来看,相较于表 2 中的数据,加入该注意力模块后,除了特异性外,在分类准确率、灵敏度和 Kappa 值上均有所提升,特别是在 ME-Net 主干下, Kappa 值提升了近 3 个百分点,说明所提出的 GAM 模块更适用于直接与主干网络结合进行分类任务。

表 2 主干模型训练结果

Table 2 Results of backbone model training

模型 Models	准确率 Accuracy	特异性 Specificity	灵敏度 Sensitivity	Kappa 值 Kappa value
MobileNetV2	0.858	0.914	0.842	0.842
Efficient-NetB0	0.862	0.932	0.854	0.864
ME-Net	0.898	0.930	0.868	0.869

表 3 加入注意力机制模型性能对比

Table 3 Performance comparison of models adding attention mechanism

模型 Models	准确率 Accuracy	特异性 Specificity	灵敏度 Sensitivity	Kappa 值 Kappa value
MobileNetV2 + GAM	0.893	0.921	0.862	0.850
EfficientNetB0 + GAM	0.921	0.924	0.870	0.870
ME-Net + GAM	0.922	0.936	0.889	0.900

为进一步探究本研究所设计的集成 ME-ANet 模型分类效果以及头部迁移学习策略对模型性能的影响,再次设计一组对照实验,对 ME-ANet 模型进行训练时,将头部集成模型 ME-Head 结构加载预训练权重和未加载预训练权重作对比,实验结果如表 4 所示。

表 4 ME-ANet 模型训练策略性能对比

Table 4 Performance comparison of ME-ANet model training strategies

训练策略 Training strategies	准确率 Accuracy	特异性 Specificity	灵敏度 Sensitivity	Kappa 值 Kappa value
非迁移学习 Non-transferable learning	0.921	0.945	0.884	0.914
迁移学习 Transfer learning	0.945	0.956	0.905	0.925

由表 4 分析可知,无论是否采用迁移学习的方式,本研究设计的模型性能指标均达到不错的效果,头部集成的迁移训练策略更进一步提升了模型的性能。其中 Kappa 值超过 0.92,表明模型的训练和预测的一致性较高,模型预测稳定性较好。

从模型迁移训练过程中的准确率和损失变化曲线可以看出,模型并没有发生过拟合现象,模型在前 5 个 epoch 训练中准确率很快突破了 0.85,损失率降到了 0.35 左右。之后模型很快收敛,准确率稳定在 0.94 左右(图 11)。

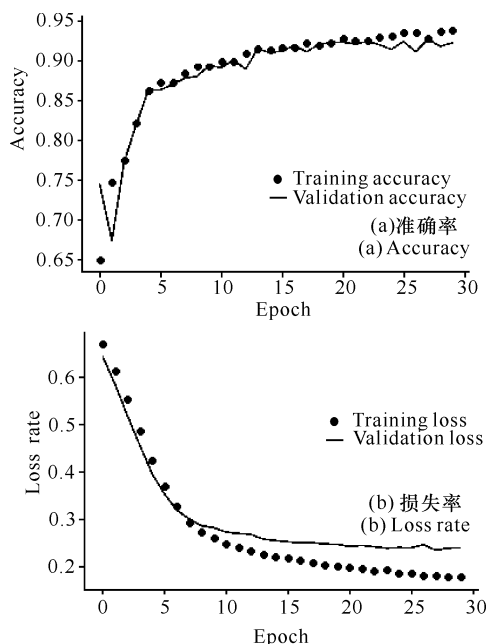


图 11 模型迁移训练和验证过程中准确率和损失率变化曲线

Fig. 11 Accuracy and loss rate curve in model transfer training and validation

最后,为了评估本研究分类模型 ME-ANet 的分类性能,使用相同的数据集和实验训练参数与近年来经典的方法进行实验对比,本研究模型的准确率和二次加权 Kappa 系数均达到了最大值,且特异性和灵敏度均超过 0.9(表 5),进一步表明,本研究的模型分类达到了不错的效果。

表 5 模型方法对比

Table 5 Comparison of model method

模型 Models	准确率 Accuracy	特异性 Specificity	灵敏度 Sensitivity	Kappa 值 Kappa value
AlexNet	0.897	0.940	0.812	0.864
GoogleNet	0.933	0.934	0.776	0.858
EfficientNet	0.925	0.975	0.918	0.800
InceptionV3	0.783	0.889	0.782	0.772
ME-ANet	0.945	0.956	0.905	0.925

3 结语

本研究提出一种基于模型集成的注意力网络 ME-ANet,用于糖尿病视网膜病变分类。利用 Kaggle 数据集进行模型训练,为解决数据不充分及分布不平衡的问题,本研究采用数据预处理及数据扩增等方式做均衡处理。采用模型融合构成主干特征提取器,并基于头部迁移学习的方式训练模型,以充分提取浅层特征信息。改进设计全局注意力机制,优化模型提取微病灶特征信息,强化眼底细粒度图像病变之间的差异性,抑制无关信息,使网络更好地学习到 DR 病变类型之间的细微差异。本研究提出的方法提高了 DR 模型分类的精确度。

在 DR 分类中,对于质量较差的原始图像没有进行剔除,导致在数据实时扩增阶段产生较多的无用特征数据,影响模型的效率和精度,后续的任务就是研究一套完整的视网膜图像质量评估方法,将着重注意提高数据集的质量。同时,探讨研究更适合图像差异小的样本(如视网膜图像)的注意力机制,进一步提高模型的性能。

参考文献

- [1] 何蓓蕾,何媛. 糖尿病与非视网膜眼部并发症相关性的研究进展[J]. 国际眼科杂志, 2021, 21(4): 623-627.
- [2] 徐宏. 基于眼底图像的糖尿病视网膜病变智能诊断[D]. 成都: 电子科技大学, 2019.
- [3] CHOE J, SHIM H. Attention-based dropout layer for weakly supervised object localization [J]. Proceedings of

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA; IEEE, 2019, 6:2219-2228.
- [4] ZHUANG J X, CAI J B, WANG R X, et al. CARE: Class attention to regions of lesion for classification on imbalanced data [C]//Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning. British: PMLR, 2019, 102:588-597.
- [5] BADAR M, HARISA M, FATIMA A. Application of deep learning for retinal image analysis: A review [J]. Computer Science Review, 2020, 35: 100203. DOI: 10.1016/j.cosrev.2019.100203.
- [6] LI X M, HU X W, YU L Q, et al. CANet: Cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading [J]. IEEE Transactions on Medical Imaging, 2019, 39(5):1483-1493.
- [7] LU L, JIANG Y, JAGANATHAN R, et al. Current advances in pharmacotherapy and technology for diabetic retinopathy: A systematic review [J]. Journal of Ophthalmology, 2018 (2): 1 - 13. DOI: 10.1155/2018/1694187.
- [8] DIAZ-PINTO A, COLOMER A, NARANJO V, et al. Retinal image synthesis and semi-supervised learning for glaucoma assessment [J]. IEEE Transactions on Medical Imaging, 2019, 38(9):2211-2218.
- [9] GULSHAN V, PENG L L, CORAM M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs [J]. Journal of the American Medical Association, 2016, 316(22):2402-2410.
- [10] 李琼, 柏正尧, 刘莹芳. 糖尿病性视网膜图像的深度学习分类方法[J]. 中国图象图形学报, 2018, 23(10): 1594-1603.
- [11] SHARMA S, MAHESHWARI S, SHUKLA A. An intelligible deep convolution neural network based approach for classification of diabetic retinopathy [J]. Bio-Algorithms and Med-Systems, 2018, 14(2): 20180011. DOI:10.1515/bams-2018-0011.
- [12] 丁蓬莉, 李清勇, 张振, 等. 糖尿病性视网膜图像的深度学习神经网络分类方法[J]. 计算机应用, 2017, 37(3): 699-704.
- [13] ZHOU K, GU Z W, LIU W, et al. Multi-cell multi-task convolutional neural networks for diabetic retinopathy grading [C]//Conference proceedings: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu, HI, USA; IEEE, 2018, 2724-2727.
- [14] 郑婷月, 唐晨, 雷振坤. 基于全卷积神经网络的多尺度视网膜血管分割[J]. 光学学报, 2019, 39(2): 119-126.
- [15] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [J]. International Journal of Computer Vision, 2020, 128(2): 336-359.

Study on Grading of Diabetic Retinopathy Based on Me-ANet Model

CHENG Xiaohui^{1,2}, LI Hejun^{1,2}, DENG Yun^{1,2}, TAO Xiaomei¹, LI Xinxiao^{1,2}

(1. College of Information Science and Engineering, Guilin University of Technology, Guilin, Guangxi, 541006, China; 2. Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guilin, Guangxi, 541006, China)

Abstract: Diabetic retinopathy (DR) is an ophthalmic disease with a high blindness rate. The difference between retinal images of different pathological grades was small and the distribution of lesion points was irregular. Aiming at the problem that the existing depth model has low recognition rate for similar lesions in DR, which seriously affects the classification accuracy of the model. This study builds a new model architecture for training based on deep learning, and proposes an attention network that integrates MobileNetV2 and EfficientNetB0 depth models: ME-ANet. The model integration is divided into two parts: head and backbone, and

the shallow part of the depth model is fused to form the head of the network. During the training, the transfer learning strategy is used to initialize the network model parameters to reduce the over-fitting problem in the training. The backbone part uses the core structure of the above two models to design a 3-stage integrated module for feature extraction. At the same time, the Global Attention Mechanism (GAM) is designed and embedded into the 3-stage integrated module respectively. The improvement of the model accelerates the convergence speed of the network. The network model achieves feature fusion extraction of shallow image information, reduces the problem of convolution loss of micro lesion feature information during the training, and further improves the classification accuracy of the model. The feature extraction backbone network is constructed through model integration, which improves the model's learning of low-level feature information, the attention mechanism suppresses non-lesion feature information, and strengthens the learning of typical lesion features, thereby realizing fine-grained classification and further improving the classification performance of the model.

Key words: diabetic retinopathy; transfer learning; integrated module; attention mechanism; feature fusion

责任编辑:梁 晓



微信公众号投稿更便捷

联系电话:0771-2503923

邮箱:gxxk@gxas.cn

投稿系统网址:<http://gxxk.ijournal.cn/gxxk/ch>