

# 基于改进 MobileNetV2 神经网络的视网膜 OCT 图像多分类<sup>①</sup>



姚娟, 乔焕, 方玲玲

(辽宁师范大学 计算机与人工智能学院, 大连 116081)

通信作者: 方玲玲, E-mail: [fanglingling@lnnu.edu.cn](mailto:fanglingling@lnnu.edu.cn)

**摘要:** 光学相干断层成像 (optical coherence tomography, OCT) 是一种具有无接触、高分辨率等特点的新型眼科医学诊断方法, 现在已经作为医生临床诊断眼科疾病的重要参考物, 但人工分类疾病费时费力, 视网膜病变的早期发现和临床诊断至关重要. 为了解决该类问题, 本文提出了一种基于改进 MobileNetV2 神经网络对视网膜 OCT 图像多分类识别方法. 此方法利用特征融合技术处理图像并设计增加注意力机制改进网络模型, 二者在极大程度上提高 OCT 图像的分类准确率. 与原有算法相比, 分类效果具有明显提升, 本文模型的分类准确率、召回值、精确度、*F1* 值分别达到 98.3%、98.44%、98.94%、98.69%, 已经超越人工分类的准确率. 此类方法不仅在实际诊断中加快诊断流程、降低医生负担、提高诊断质量, 同时也为眼科医疗研究提供新的方向.

**关键词:** 视网膜; 光学相干断层扫描; 注意力机制; 特征融合; 图像分类

引用格式: 姚娟, 乔焕, 方玲玲. 基于改进 MobileNetV2 神经网络的视网膜 OCT 图像多分类. 计算机系统应用, 2024, 33(5): 37-46. <http://www.c-s-a.org.cn/1003-3254/9482.html>

## Multi-classification of Retinal OCT Images Based on Improved MobileNetV2 Neural Network

YAO Juan, QIAO Huan, FANG Ling-Ling

(School of Computer and Artificial Intelligence, Liaoning Normal University, Dalian 116081, China)

**Abstract:** Optical coherence tomography (OCT) is a new type of ophthalmic diagnosis method with non-contact, high resolution, and other characteristics, which has been used as an important reference for doctors to clinically diagnose ophthalmic diseases. As early detection and clinical diagnosis of retinopathy are crucial, it is necessary to change the time-consuming and laborious status quo of the manual classification of diseases. To this end, this study proposes a multi-classification recognition method for retinal OCT images based on an improved MobileNetV2 neural network. This method uses feature fusion technology to process images and designs an attention increase mechanism to improve the network model, greatly improving the classification accuracy of OCT images. Compared with the original algorithm, the classification effect has been significantly improved, and the classification accuracy, recall value, accuracy, and *F1* value of the proposed model reach 98.3%, 98.44%, 98.94% and 98.69%, respectively, which has exceeded the accuracy of manual classification. Such methods not only speed up the diagnostic process, reduce the burden on doctors, and improve the quality of diagnosis in actual diagnosis, but also provide a new direction for ophthalmic medical research.

**Key words:** retina; optical coherence tomography (OCT); attention mechanism; feature fusion; image classification

<sup>①</sup> 基金项目: 辽宁省自然科学基金 (2021-MS-272); 辽宁省教育厅项目 (LJKQZ2021088)

收稿时间: 2023-11-03; 修改时间: 2023-12-04; 采用时间: 2023-12-20; csa 在线出版时间: 2024-03-22

CNKI 网络首发时间: 2024-03-27

## 1 引言

近年来,患有视网膜疾病的人越来越多.根据《世界视力报告》,至少有22亿人患有视力障碍,其中一半以上可以预防<sup>[1]</sup>.随着患有眼部疾病的人数剧增,医生供不应求.由于专家的可用性较低而导致的时间压力加剧.已经证实,压力和时间压力会导致诊断错误.同时,多种眼部疾病分类在无形中给医生带来严重负担,导致很多病人无法得到恰当的治疗,从而对视力产生不可逆转的影响,严重情况下可能会导致失明.而定期检查眼睛可在初步诊断中发挥重要作用,可在初始阶段的治疗将阻止病情恶化或失明.因此,社会迫切需要人工智能技术来准确地分辨出各类眼部疾病.

目前,大量研究致力于计算机视觉,光学相干断层成像(optical coherence tomography, OCT)<sup>[2]</sup>自1991年发展至今,它不仅改善了各个国家的眼科医疗系统,同时也给社会带来了巨大的福音.随着OCT技术的逐渐成熟,其拥有着无限广阔的发展前景.OCT是一类三维断层且具有低损、非接触、高分辨、成像快<sup>[3-5]</sup>等优点的医学成像技术,可以探测生物组织中不同深度的反射或者散射信号.并且OCT使用低相干或白光干涉测量原理,并产生视网膜<sup>[6]</sup>的高分辨率横断面图像.OCT非常适用于诊断、进展评估和治疗反应评估.1991年,Huang等<sup>[7]</sup>在利用OCT系统对人眼视网膜和视盘进行成像研究,该研究具有一定的历史意义,为今后OCT技术的应用奠定了基础.由于人工智能、第5代(5G)电信网络和物联网等技术的集成,远程医疗在眼科领域的非接触式医疗保健服务方面取得了长足进步.它还有助于对孤雌生殖疾病有更多的了解,这将有助于创造新的治疗方法<sup>[8]</sup>.OCT图像特征如图1所示,医生可根据图像中的视网膜特征判断相关疾病.比如:OCT可用于视神经节细胞轴突损伤的定量分析,帮助评价许多神经退行性病变进展;神经纤维层在临床中可以反映视神经损伤的情况(如:青光眼,高度近视等),这层还可以反映一些中枢神经系统的疾病信息(如多发性硬化等神经退行病变).

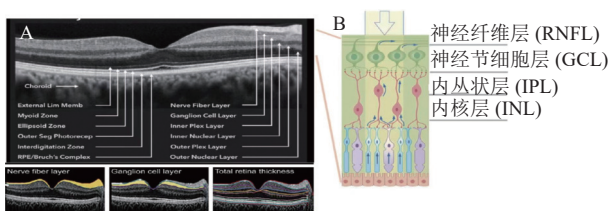


图1 OCT图像特征图

在实际临床应用中,利用OCT技术产生的视网膜疾病图像是分类眼科疾病的重要措施.医生可以通过观察OCT图像中视网膜层的特点来诊断眼部疾病,但由于OCT图像视网膜层边界对比度较低,人工分类费时费力.OCT已经成为视网膜疾病筛查的潜在成像工具,但分类视网膜多种疾病检测仍然是一个挑战,因此,对OCT图像多分类方法的研究具有重要意义.

近10年来,深度学习的发展越来越成熟,众多研究人员致力于将深度学习技术运用到现实生活,而深度学习在医学图像的分类中具有深远意义.卷积神经网络是一种通过连续的卷积训练在数据中找到相应模式的神经网络.网络在训练OCT图像过程中具有一定的局限性<sup>[9]</sup>,本文探讨的模型在一定程度上解决了该问题.简单的深度学习架构,如VGG和ResNet已通过迁移学习运用于OCT图像分类<sup>[10]</sup>.基于VGG-16的改进模型也被提出,它使用了基本预训练架构并添加了注意模块<sup>[11]</sup>,为分类研究提供了新方向,一定程度上有效地解决了当时对于OCT图像分类不准确的问题.2014年,牛津大学的visual geometry group提出VGG网络<sup>[12]</sup>,该网络的分类准确率获得显著提升,这项研究对今后的神经网络的分类工作奠定了基础.AI-Bander等<sup>[13]</sup>开发了一种基于正常眼与青光眼OCT图像二分类的卷积神经网络系统,分类准确性达到了88.2%.该网络虽然在一定程度上辅助医师诊断决策,但分类准确率不够高,容易在诊断中给医生带来干扰.此后,众多研究者对视网膜OCT图像分类做出了重大贡献.例如:Lee等<sup>[14]</sup>利用深度学习的方法有效解决了与年龄相关的黄斑变性二分类问题;Rasti等<sup>[15]</sup>应用多尺度卷积神经网络实现了对黄斑病变OCT图像的分类;Karri等<sup>[16]</sup>提出了一种基于迁移学习的糖尿病性黄斑水肿和干性年龄相关性黄斑变性光学相干断层扫描图像分类方法.他们的研究针对视网膜黄斑病变的分类研究具有重大突破,但二分类具有一定的局限性,在医生临床诊断中病人的眼部疾病是具有多样性的,无法在图像上准确的识别各种疾病.2018年,De Fauw等<sup>[17]</sup>提出了一种基于U-Net网络对视网膜OCT图像进行分割,进而用于视网膜病变的诊断.2020年,张添福等<sup>[18]</sup>提出了一种改进轻量级神经网络模型的方法,其主要技术特点是将深度可分离卷积代替为普通卷积层、全局平均池化代替全连接层,该网络对OCT图像分类准确率可高达

97%。但在 OCT 图像特征提取过程中,某些病变特征容易被忽略,造成信息丢失,使得网络准确的下降。因此,在网络中添加注意力机制以解决病变特征不明显的问题是分类研究的新方向。2017年 Hu 等<sup>[19]</sup>初次提出的通道注意力机制,但该技术在捕捉特征的方面仍有一定的局限性;2018年 Woo 等<sup>[20]</sup>提出的融合通道注意力以及空间注意力的 CBAM 机制,但 CBAM 机制需要更多的计算资源,计算复杂度更高;2020年 Wang 等<sup>[21]</sup>提出的改进通道注意力机制,在特征提取上较大的优势。经过近几年的科研工作者的努力,注意力机制能够有效地解决网络的准确率,但对多分类的研究仍有不足。

基于以上文献研究,在图像分类上依然存在特征不明显、学习率较低、OCT 图像存在不可避免的噪声、数字读取困难、分类类别有局限等问题。针对这些问题,本文提出了一种基于改进 MobileNetV2 神经网络对视网膜 OCT 图像多分类识别方法。其创新点 1: 在数据集训练之前融合新生血管、神经皮层缺失、视网膜神经纤维层变薄等多种疾病特征;创新点 2: 为增强模型对疾病特征的学习能力,在模型的输入输出位置上引入注意力机制。

如图 2 所示,图 2(a) 对应正常眼 OCT 图像;图 2(b) 对应黄斑裂孔 OCT 图像;图 2(c) 对应青光眼 OCT 图像;图 2(d) 对应视神经炎 OCT 图像,图 2(e) 对应糖尿病性视网膜病变 OCT 图像。本文方法主要对正常眼、黄斑裂孔、青光眼、视神经炎、糖尿病性视网膜病变进行分类。本文提出的方法不仅有效地解决了 OCT 图像特征不明显的问题,还引用注意力机制改善模型的性能,使分类结果更加可靠。MobileNetV2 作为一种轻量级的卷积神经网络,在 OCT 图像的多分类应用上更加灵活;此外本文此模型基础上增加特征融合以及注意力机制,使得数据集被充分利用,大幅度提升分类结果的准确率。

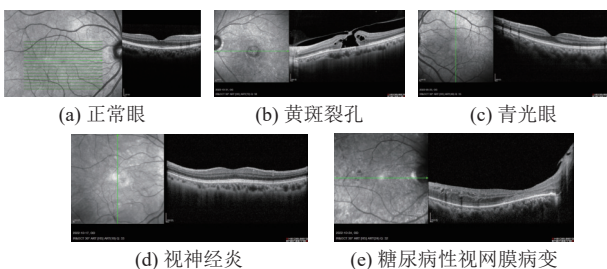


图 2 异常视网膜病变 OCT 图像

## 2 数据集与预处理

### 2.1 数据集

实验所采用的数据集来自 Retinal OCT 数据集和大连市第三人民医院提供的真实临床数据集。Retinal OCT 数据集共提供 1487 幅 OCT 视网膜病变图像,其中正常眼图像 366 幅、黄斑裂孔视网膜病变图像 256 幅、青光眼视网膜病变图像 287 幅、视神经炎视网膜病变图像 195 幅、糖尿病性视网膜病变图像 383 幅。大连市第三人民医院共提供 656 幅 OCT 视网膜病变图像,其中正常眼图像 98 幅、黄斑裂孔视网膜病变图像 90 幅、青光眼视网膜病变图像 150 幅、视神经炎视网膜病变图像 81 幅、糖尿病性视网膜病变图像 237 幅。数据集的数量信息如表 1 所示。

表 1 初始数据集

类型	Retinal OCT数据集	临床数据集
正常眼	366	98
黄斑裂孔	256	90
青光眼	287	150
视神经炎	195	81
糖尿病性视网膜病变	283	237
合计	1487	565

### 2.2 数据集预处理

由于 OCT 图像质量不稳定以及病变图像自身特点的问题,为了适应模型训练,在实验开始之前,需要对数据集进行预处理操作。首先,将数据集分为两大类,分别为训练集 DateTrain 和测试集 DateTest。其中 DateTrain 占 80%,用于模型训练;DateTest 占 20%,作为测试模型,用于评估最终的学习效果。

图像数据标准化:由于初始数据集的分辨率大小存在差异,因此,需要对 OCT 图像作统一化的批处理。具体操作为:在 Matlab 的环境下用 resize() 函数方法对数据集进行批处理操作,将 OCT 图像分辨率统一为 1264×596×3;利用相关函数对数据集进行顺时针 90°、180°、270° 旋转,如图 3 所示,以增加实验数据数量使实验结果更可靠。

图像优化:由于 OCT 图像在采集的过程中,由于镜头、光照等影响,容易导致图像存在较严重的噪声。因此,在训练数据之前需要对图像进行降噪处理。利用高斯函数对数据集做卷积,具体函数公式如式 (1)、式 (2) 所示:

$$p(a,b) = \exp\left(-\frac{a^2+b^2}{2\sigma^2}\right) \quad (1)$$



$$Q(a,b) = f(a,b) * p(a,b) \quad (2)$$

其中,  $P(a,b)$ 表示高斯函数,  $f(a,b)$ 表示原图像,  $Q(a,b)$ 表示去噪之后的图像,  $*$ 表示卷积操作.

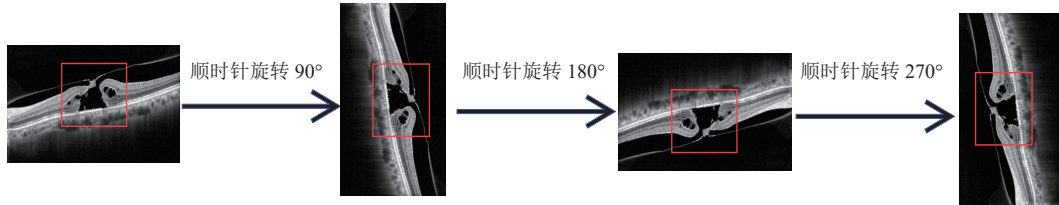


图3 数据集图像旋转的过程

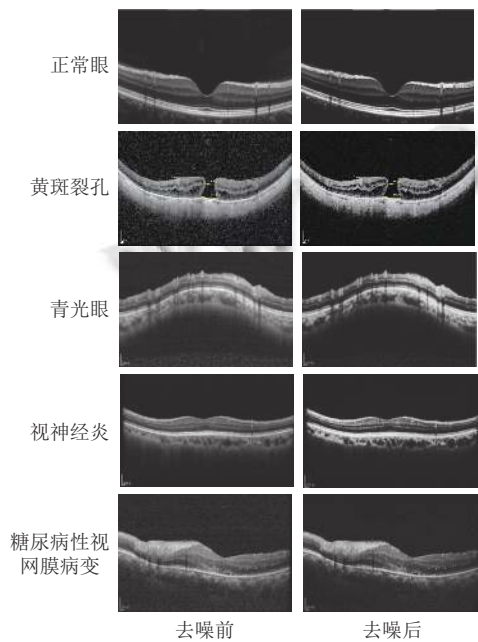


图4 预处理前后对比图

图4为5种疾病对应的去噪前后对比图, 降噪处理后的图像病变特征更加明显, 有利于网络训练, 能有效提升网络准确率.

将 MobileNetV2 块与左侧阴影突出显示的注意力机制块结合在一起.

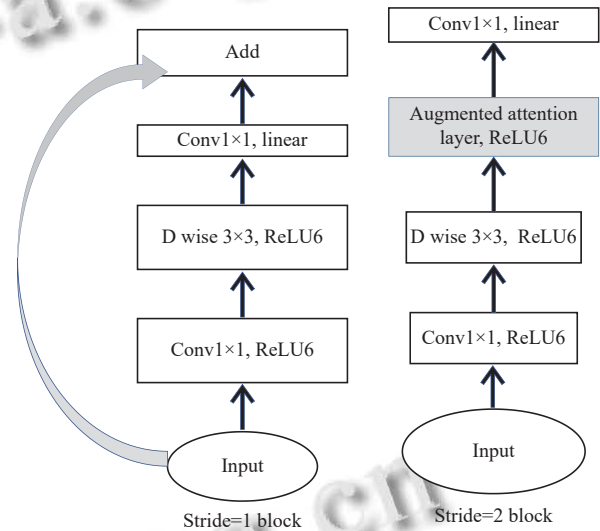


图5 MobileNetV2 网络模型

### 3 实验方法

#### 3.1 模型介绍

眼底成像和 OCT 是目前眼科广泛应用的两种诊断手段. 然而, OCT 眼底图像的分类研究在深度学习算法中更受欢迎. 随着深度学习相关网络模型不断成熟, CNN<sup>[22]</sup>、Alex-Net<sup>[23]</sup>等网络模型在图像的分类领域被陆续提出. 谷歌团队在 2018 年提出了 MobileNetV2 网络. MobileNetV2 作为一种轻量级的卷积神经网络模型, 它的优势在于: 该网络的精度高于 MobileNetV1<sup>[24]</sup> 网络, 且模型更小. MobileNetV2 结合了反向残差和线性瓶颈, 有效减少信息损失, 同时去掉了低维输出层后面的非线性激活层, 显著提高了模型的准确性和效率. 图5为 MobileNetV2 网络模型结构, 本文提出的架构

#### 3.2 模型设计及算法

图6展示了正常眼、黄斑裂孔、青光眼、视神经炎、糖尿病性视网膜病变4种疾病在OCT成像中的疾病特征. 由于视网膜病变特征繁多, 医生在诊断OCT图像眼部疾病时, 常依据是否有新生血管、黄斑中心凹陷程度、视网膜神经纤维层宽度等疾病特征. 不仅给医生的诊断过程造成了严重负担, 就连机器学习也难以准确学习复杂的眼底图像.

为了获得更复杂的眼底OCT图像特征, 在提出的网络中加入多特征融合 add, 从主干和分支中提取, 具体公式如下所示:

$$F^{add} = \sum_{i=1}^c (X_i + Y_i + Z_i + W_i) * K_i$$

$$= \sum_{i=1}^c X_i * K_i + \sum_{i=1}^c Z_i * K_i + \sum_{i=1}^c W_i * K_i \quad (3)$$

其中,  $F^{add}$  为多特征融合;  $K$  为卷积核;  $*$  为卷积;  $X_i$  为视网膜平伏程度、 $Y_i$  为新生血管、 $Z_i$  为黄斑中心凹陷程度、 $W_i$  为视网膜神经纤维层宽度,  $+$  指视网膜眼底图像的各特征信息的叠加和融合。

OCT 图像通常由矩阵的形式表示, 而矩阵由多个向量组成, 每个向量都代表着图像中的某些特征信息, 如边缘、纹理等。在多特征融合中可以通过卷积、激活函数和池化层的相关操作提取出 OCT 图像的病变特征, 在全连接层中, 所有的特征向量被展平成一个向量, 用来表示整个图像的特征, 最后采用逐元素相乘的方式实现多特征融合技术。

此外, 由于眼部 OCT 图像的病变区域非常小, 训练过程中无法捕捉相关特征, 容易影响实验结果。注意力机制在计算机视觉领域的快速应用和发展, 在不同的任务下取得了较理想的效果。为了提高模型在块内的特征学习能力, 我们还在第 2 层分组卷积层后引入了轻量级卷积块注意力机制。这种机制能在通道和空间

两个维度加权生成注意力图, 使网络能够关注到重要的通道维度特征以及空间维度特征, 本文模型中的门控注意力机制模块 (convolutional block attention module, CBAM)。如图 7 所示。其中,  $g$  表示网络训练中能够获取到的深层特征;  $x$  表示特征提取过程中容易获取到的浅层特征。深层特征是指深层网络提取的特征离输出较近, 一般是抽象的信息, 即语义信息等。浅层特征是指浅层网络提取的特征和输入比较近, 包含更多的像素点的信息, 如纹理、边缘、棱角信息。深层先进行上采样, 浅层特征层进行下采样, 二者进行特征融合, 再通过卷积和平均池化层进行特征提取。深层特征和浅层特征的提取在 CBAM 模块中都至关重要, CBAM 模块将深层特征和浅层特征融合并生成注意力图, 然后将该注意力图与浅层特征相乘, 用深层信息修剪冗余特征, 突出显著目标特征区域。标准卷积管道与生成的注意力图并行运行, 这两个管道最后合并为一个输出。

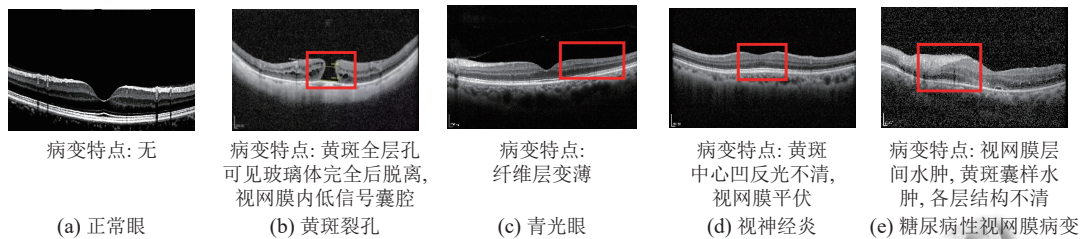


图 6 各类视网膜疾病特征

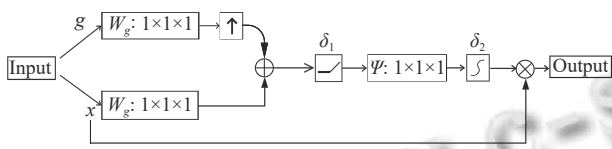


图 7 门控注意力机制模块

综上所述, 为了准确识别不同的病变特征, 本文将提出一种改进的深度学习模型 CBAM-MobileNetV2, 该模型是将多特征融合与 CBAM 模块共同融入 MobileNetV2 中, 完整的网络模型如图 8 所示。融合层将多种疾病特征从通道和空间两个维度相加融合, 采取矩阵广播扩展的方式实现空间维度特征和通道维度特征的对齐随后将特征图输入到 CBAM 模块。首先进行平均池化, 然后对特征图进行两次卷积。对于所提出的通道关注模块, 仅使用平均池化操作来压缩图像特征。由于采用单分支池化操作, 池化后输出特征图的复杂度较低。为了加强每个通道的特征学习, 将通道注意模块中

的全连接层替换为  $1 \times 1$  卷积核大小。这样, 卷积后得到的特征图被 Sigmoid 函数激活。将激活后的特征图与输入特征图相乘, 最后得到输出特征图。对于另一个空间注意模块, 其输入来自通道注意模块的输出特征图。首先, 依次对最大池化和平均池化进行处理。然后, 将处理后的特征图输入到卷积核为  $7 \times 7$  的卷积层中, 将被 Sigmoid 函数激活的特征图与输入特征图相乘, 输出最终的特征图。

### 3.3 算法流程

该网络设计过程如下。

(1) 数据集增强: 在原有数据集的基础上, 将数据集进行旋转、翻转、像素归一化等处理增加数据集数量, 使实验结果更具有说服力。

(2) 图像预处理: 利用高斯函数来消除噪声, 提高 OCT 图像的清晰度。

(3) 模型构建: 对原有 MobileNetV2 模型进行分析,

并结合提出的 CBAM 机制和多特征融合机制,对卷积层进行改进和优化,最终形成新的深度学习模型 CBAM-MobileNetV2.

(4) 网络训练: 将每种视网膜疾病的病变特征输入到模型中,实现多特征融合,系统学习 4 种视网膜疾病的特征信息.

(5) 网络测试: 将处理后的图像发送到网络,实现 5 种 OCT 图像的多疾病分类,验证所提模型的分精度.

实际上,实验具体流程如图 9 所示. 首先,利用数

据集增强处理和高斯函数对 OCT 图像进行初步处理,提高图像数量和清晰度;然后,将数据集统一划分,训练集与测试集以 8:2 的比例进行后续实验.在模型的每个卷积层学习增强图像的特征,并对学习到的特征进行多特征融合处理.同时,提出 CBAM 机制,对通道维度和空间维度上的注意力进行加权,使网络能够专注于对通道维度特征和空间位置信息特征的提取,从而提取出更准确的 OCT 图像病理特征.最后,通过 CBAM-MobileNetV2 模型的性能,获得了 OCT 图像中视网膜疾病的多重分类.

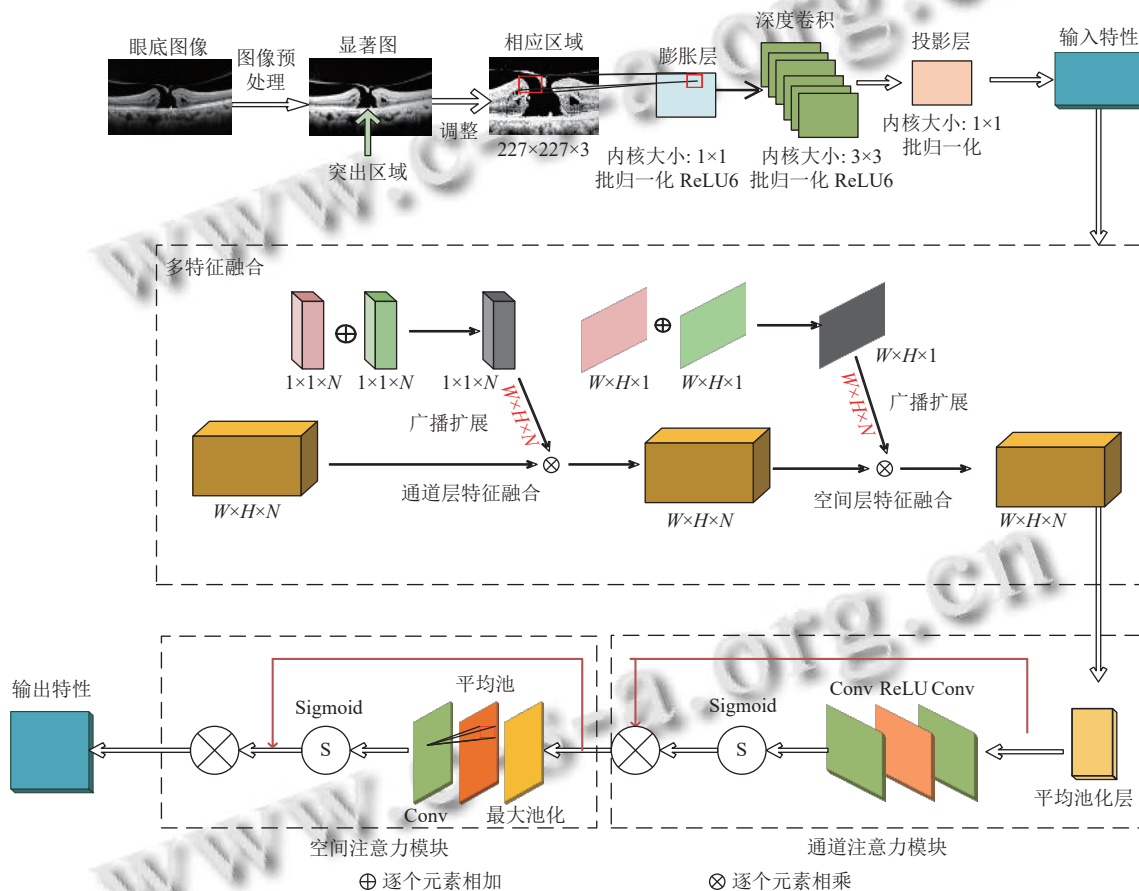


图 8 本文 MobileNetV2 算法

## 4 实验结果与分析

### 4.1 模型评估

针对多分类问题,模型评估指标分为两大类: example-based metrics, label-based metrics. 评价指标常被用来评估模型的好坏,混淆矩阵是人工智能领域中最常用的模型评价方法,常用的评价指标包括准确率、召回率等.

本文利用 label-based metrics 评估指标为标准,本

文以精确率 (Precision)、召回率 (Recall)、准确率 (Accuracy) 和 F1 值 (F1-score)<sup>[25]</sup> 为模型的评估指标,公式如式 (4)~式 (7) 所示. 精确率,指在阳性样本中正确分类的样本比例;召回率,测量所有阳性样本中正确分类的样本的比例;准确率,指样本正确分类的比例; F1-score,这是精密度和召回率之间的调和平均值. 这些初始指标被添加到多个职业和方程式的混淆矩阵中.



$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

$$F1-score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

其中,  $TP$  和  $TN$  分别表示正确分类的阳性和阴性样本数;  $FP$  和  $FN$  分别是错误分类的阴性和阳性样本数. 性能指标均为 $[0, 1]$ , 数值越大, 分类效果越好<sup>[26]</sup>.

### 4.2 实验结果

将经过预处理的 Retinal OCT 数据集和临床数据集应用到模型中, 经过大量训练得到的结果如图 10 所示, 其中第 6 行和第 6 列为最终结果, Retinal OCT 数据集的分类准确率为 98.0%, 临床数据集的分类准确率达到 98.3%, 损失率和精度曲线如图 11 所示. 相关评价指标实验结果如图 12 所示, Retinal OCT 数据集的精确率 ( $Precision$ )、召回率 ( $Recall$ )、 $F1-score$  值分别为 0.9876、0.9791、0.9816, 临床数据集的精确率 ( $Precision$ )、召回率 ( $Recall$ )、 $F1-score$  值分别为 0.9844、0.9894、0.9869. 实验结果表明, 基于改进的 MobileNetV2

网络模型具有较好的分类性能, 一定程度上能够帮助医生临床诊断.

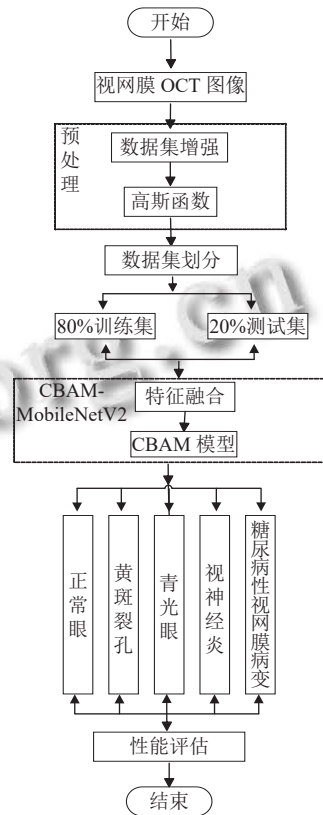


图 9 算法流程

Output class \ Target class	正常眼 OCT	糖尿病性视网膜病变 OCT	视神经炎 OCT	青光眼 OCT	黄斑裂孔 OCT	Sum	Accuracy
正常眼 OCT	12 6.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	12	100%
糖尿病性视网膜病变 OCT	0 0.0%	79 39.7%	0 0.0%	0 0.0%	0 0.0%	79	100%
视神经炎 OCT	0 0.0%	0 0.0%	27 13.6%	0 0.0%	0 0.0%	27	100%
青光眼 OCT	0 0.0%	0 0.0%	0 0.0%	50 25.1%	3 1.5%	53	94.3%
黄斑裂孔 OCT	0 0.0%	0 0.0%	0 0.0%	1 0.0%	27 13.6%	28	96.4%
Sum	12	79	27	53	28	199	98.0%
Accuracy	100%	100%	100%	98.0%	90.0%		98.0%
Loss	0.0%	0.0%	0.0%	2.0%	10.0%		2.0%

Output class \ Target class	正常眼 OCT	糖尿病性视网膜病变 OCT	视神经炎 OCT	青光眼 OCT	黄斑裂孔 OCT	Sum	Accuracy
正常眼 OCT	14 5.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	14	100%
糖尿病性视网膜病变 OCT	0 0.0%	95 39.7%	0 0.0%	0 0.0%	0 0.0%	95	100%
视神经炎 OCT	0 0.0%	0 0.0%	32 13.4%	0 0.0%	0 0.0%	32	100%
青光眼 OCT	0 0.0%	0 0.0%	0 0.0%	58 24.3%	3 1.5%	61	100%
黄斑裂孔 OCT	0 0.0%	0 0.0%	0 0.0%	4 1.7%	36 15.1%	40	90%
Sum	14	95	32	61	40	242	98.3%
Accuracy	100%	100%	100%	93.5%	100.0%		98.3%
Loss	0.0%	0.0%	0.0%	6.5%	10.0%		1.7%

图 10 实验结果

### 4.3 多模型比较

为了验证改进 MobileNetV2 模型对 OCT 图像多

分类的有效性, 我们将本文方法与常见轻量级网络 EfficientNetV2<sup>[27]</sup>、ShuffleNet<sup>[28]</sup>以及非轻量级网络

ResNet-50<sup>[29]</sup>、VGG-16<sup>[30]</sup>、黄斑变性<sup>[31]</sup>的眼底疾病分类性能进行对比实验. 记录其精确率、召回率以及  $F1$ -

$score$  的评估指标的值, 并依次对训练结果进行分析. 不同网络模型的实验对比结果如表 2 所示.

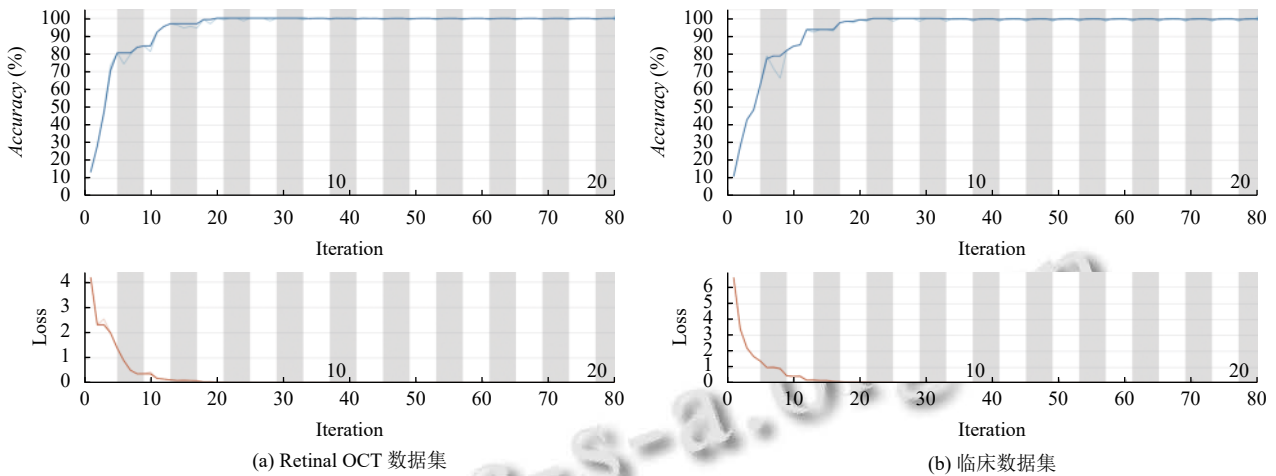


图 11 损失率和精度曲线

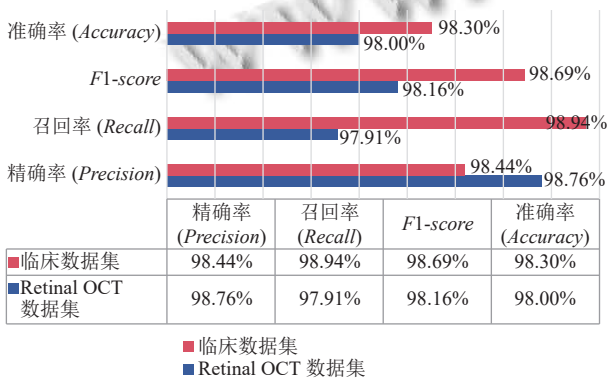


图 12 各评价指标实验结果

表 2 本文方法与其他方法的性能比较 (%)

网络模型	精确率 (Precision)	召回率 (Recall)	$F1$ 值 ( $F1$ -score)	准确率 (Accuracy)
EfficientNetV2	95.51	96.91	96.41	96.32
ShuffleNet	95.96	95.89	96.19	96.12
ResNet-50	97.08	97.56	96.11	97.51
VGG-16	95.51	94.91	94.41	94.8
黄斑变性	96.77	95.96	96.47	96.93
本文方法	<b>98.44</b>	<b>98.94</b>	<b>98.69</b>	<b>98.3</b>

结合表 2 的对比实验结果可知, 本文模型取得了 6 个模型中最高的分类准确率 98.3%, 相较于非轻量级模型 ResNet-50、VGG-16、黄斑变性高出 0.79%、3.5%、1.37%. 与轻量级模型相比, 本文模型的分类准确率分别比 EfficientNetV2、ShuffleNet 高 1.98%、2.18%. 本文提出的方法的召回率 (Recall)、准确率 (Accuracy) 和  $F1$  值 ( $F1$ -score) 均为最高.

由对比实验结果可知, 本文模型在临床数据集上

的性能不仅优于其他轻量级网络, 甚至优于非轻量级网络 ResNet-50、同类相关眼底疾病分类网络. ResNet-50 网络模型的关键就在于其结构中的残差单元, 因此它相较于其他网络模型分类准确率高, 但该网络模型堆叠了很多网络层, 使训练速度非常缓慢. 黄斑变性的双引导网络利用了分类任务和分割任务之间的强相关性, 但其分类具有局限性. 本文提出的方法在眼底疾病多分类领域具有较高的准确率, 相较于本文对比的网络模型更适用于眼底疾病多分类任务.

#### 4.4 实验分析

基于 MobileNetV2 网络模型上, 本文对其做出了以下 3 点改进: 降噪处理、多特征融合、注意力机制. 为验证改进 MobileNetV2 神经网络的必要性, 这里利用临床数据集进行消融实验, 实验中设置 8 组, 任意一种改进处理、任意两种改进处理、3 种改进处理以及无改进处理. 为了保证消融实验的可靠性, 每组实验的数据集均相同. 实验结果如表 3 所示.

表 3 改进模型处理的消融实验 (%)

处理方法	精确率	召回率	$F1$ 值	准确率
无处理	75.18	81.88	77.11	79.56
降噪处理	82.68	74.91	89.41	84.32
多特征融合	96.96	95.89	96.19	96.12
注意力机制	97.08	97.56	96.11	97.15
降噪处理+多特征融合	95.51	96.91	94.41	96.31
降噪处理+注意力机制	96.9	97.1	95.6	97.29
多特征融合+注意力机制	96.77	95.96	96.47	97.61
降噪处理+多特征融合+注意力机制	<b>98.44</b>	<b>98.94</b>	<b>98.69</b>	<b>98.3</b>



由表3可知,模型同时进行了降噪处理、多特征融合、注意力机制改进处理的分类准确率最高。分别比仅降噪处理、多特征融合、注意力机制和无改进处理的准确率分别高出13.98%、2.18%、1.15%和8.74%。同时,比提取任意两个特征的准确率高1.99%、1.01%和0.69%。实验表明,在实验前对数据集同时进行降噪处理以及多特征融合更有利于模型训练,而在网络模型训练中引入CBAM模型能有效地提高实验的分类准确率,因为CBAM模型具有较少的参数量和计算复杂度,并在所选的5种眼底疾病的多分类上取得了优异的分类结果。因此,在模型的基础上提出改进对实验分类中起着关键作用。

由于机器设备的不完美,OCT图像总是存在一些难以改善噪声,同时,MobileNetV2模型具有不稳定性。因此,该模型无法精确的识别并分类OCT图像。为了进一步提高模型分类准确率,本文将对分类错误的图像进行深度分析。发现只仅有1.7%的OCT图像出现分类错误的情况,除一些无法避免的特殊原因,若依靠模型的不断改进,大多数分类错误的OCT图像可以被纠正。

## 5 结论

本文针对视网膜OCT图像人工识别难度高、准确率不稳定等问题,提出了一种基于改进MobileNetV2神经网络对视网膜OCT图像多分类识别方法。该方法在原来MobileNetV2网络模型基础上增加特征融合以及注意力机制,有效地解决了模型不稳定性,使模型对图像分类准确率达到98.3%。同时,也验证了该方法具有较好的鲁棒性及泛化能力。

本文提出的方法,通过重构模型结构,引入多特征融合以及卷积注意力模块,设计了一种轻量级的小型卷积神经网络架构MobileNetV2,在提高图像特征的同时,大大降低了模型的计算成本和内存消耗。不仅帮助眼科医生减轻负担,还给世界各地相关疾病患者带来了更多的希望,同时在眼科医疗方面具有深远意义和良好的发展方向。今后,可以尝试利用本文方法,进一步提高该网络性能,提高分类准确率。

## 参考文献

- 1 World Health Organization (WHO). World report on vision. <https://www.iapb.org/advocate/eye-health-and-universal-health-coverage/world-report-on-vision/>. [2023-09-12].
- 2 张晓虹,陈卫伟.眼科光学断层成像仪的发展及应用.中国医疗设备,2012,27(6):118-120. [doi: 10.3969/j.issn.1674-1633.2012.06.043]
- 3 Drexler W, Fujimoto JG. State-of-the-art retinal optical coherence tomography. *Progress in Retinal and Eye Research*, 2008, 27(1): 45-88. [doi: 10.1016/j.preteyeres.2007.07.005]
- 4 Pizurica A, Jovanov L, Huysmans B, et al. Multiresolution denoising for optical coherence tomography: A review and evaluation. *Current Medical Imaging Reviews*, 2008, 4(4): 270-284. [doi: 10.2174/157340508786404044]
- 5 Beaufreire E, Boccara AC, Lebec M, et al. Full-field optical coherence microscopy. *Optics Letters*, 1998, 23(4): 244-246. [doi: 10.1364/OL.23.000244]
- 6 Ripandelli G, Coppé AM, Capaldo A, et al. Optical coherence tomography. *Seminars in Ophthalmology*, 1998, 13(4): 199-202. [doi: 10.3109/08820539809056053]
- 7 Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*, 1991, 254(5035): 1178-1181. [doi: 10.1126/science.1957169]
- 8 Fujimoto J, Swanson E. The development, commercialization, and impact of optical coherence tomography. *Investigative Ophthalmology & Visual Science*, 2016, 57(9): OCT1-OCT13. [doi: 10.1167/iovs.16-19963]
- 9 Kapoor R, Whigham BT, Al-Aswad LA. Artificial intelligence and optical coherence tomography imaging. *Asia-Pacific Journal of Ophthalmology*, 2019, 8(2): 187-194. [doi: 10.22608/APO.201904]
- 10 Hasan R, Langner H, Ritter M, et al. Investigating the robustness of pre-trained networks on OCT-dataset. *Proceedings of the 2019 Conference on Actual Problems of System and Software Engineering*. Moscow, 2019. 200-207.
- 11 Sitaula C, Hossain MB. Attention-based VGG-16 model for COVID-19 chest X-ray image classification. *Applied Intelligence*, 2021, 51(5): 2850-2863. [doi: 10.1007/s10489-020-02055-x]
- 12 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proceedings of the 3rd International Conference on Learning Representations*. San Diego: ICLR, 2015.
- 13 Al-Bander B, Al-Nuaimy W, Al-Tae MA, et al. Automated glaucoma diagnosis using deep learning approach. *Proceedings of the 14th International Multi-conference on Systems, Signals & Devices (SSD)*. Marrakech: IEEE, 2017. 207-210.

- 14 Lee CS, Baughman DM, Lee AY. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. *Ophthalmology Retina*, 2017, 1(4): 322–327. [doi: [10.1016/j.oret.2016.12.009](https://doi.org/10.1016/j.oret.2016.12.009)]
- 15 Rasti R, Rabbani H, Mehridehnavi A, *et al.* Macular OCT classification using a multi-scale convolutional neural network ensemble. *IEEE Transactions on Medical Imaging*, 2018, 37(4): 1024–1034. [doi: [10.1109/TMI.2017.2780115](https://doi.org/10.1109/TMI.2017.2780115)]
- 16 Karri SPK, Chakraborty D, Chatterjee J. Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. *Biomedical Optics Express*, 2017, 8(2): 579–592. [doi: [10.1364/BOE.8.000579](https://doi.org/10.1364/BOE.8.000579)]
- 17 De Fauw J, Ledsam JR, Romera-Paredes B, *et al.* Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 2018, 24(9): 1342–1350. [doi: [10.1038/s41591-018-0107-6](https://doi.org/10.1038/s41591-018-0107-6)]
- 18 张添福, 钟舜聪, 连超铭, 等. 基于深度学习特征融合的视网膜图像分类. *激光与光电子学进展*, 2020, 57(24): 258–265.
- 19 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 7132–7141.
- 20 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. In: Ferrari V, Hebert M, Sminchisescu C, *et al.*, eds. *Proceedings of the 2018 European Conference on Computer Vision*. Cham: Springer, 2018. 3–19.
- 21 Wang QL, Wu BG, Zhu PF, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020. 11531–11539.
- 22 董喜超, 高志军, 董春游. 基于改进的 AG-CNN 的视网膜 OCT 图像的黄斑病变识别方法. *智能计算机与应用*, 2021, 11(5): 163–169. [doi: [10.3969/j.issn.2095-2163.2021.05.032](https://doi.org/10.3969/j.issn.2095-2163.2021.05.032)]
- 23 Li SJ, Wang LZ, Li J, *et al.* Image classification algorithm based on improved AlexNet. *Journal of Physics: Conference Series*, 2021, 1813(1): 012051. [doi: [10.1088/1742-6596/1813/1/012051](https://doi.org/10.1088/1742-6596/1813/1/012051)]
- 24 毕鹏程, 罗健欣, 陈卫卫. 轻量化卷积神经网络技术研究. *计算机工程与应用*, 2019, 55(16): 25–35. [doi: [10.3778/j.issn.1002-8331.1903-0340](https://doi.org/10.3778/j.issn.1002-8331.1903-0340)]
- 25 Da Rocha DA, Ferreira FMF, Peixoto ZMA. Diabetic retinopathy classification using VGG16 neural network. *Research on Biomedical Engineering*, 2022, 38(2): 761–772. [doi: [10.1007/s42600-022-00200-8](https://doi.org/10.1007/s42600-022-00200-8)]
- 26 Li F, Liu Z, Chen H, *et al.* Automatic detection of diabetic retinopathy in retinal fundus photographs based on deep learning algorithm. *Translational Vision Science & Technology*, 2019, 8(6): 4. [doi: [10.1167/tvst.8.6.4](https://doi.org/10.1167/tvst.8.6.4)]
- 27 Wang XW, Zhang SL, Liang X, *et al.* Accurate and fast blur detection using a pyramid M-shaped deep neural network. *IEEE Access*, 2019, 7: 86611–86624. [doi: [10.1109/ACCESS.2019.2926747](https://doi.org/10.1109/ACCESS.2019.2926747)]
- 28 Zhao WD, Hou XQ, He Y, *et al.* Defocus blur detection via boosting diversity of deep ensemble networks. *IEEE Transactions on Image Processing*, 2021, 30: 5426–5438. [doi: [10.1109/TIP.2021.3084101](https://doi.org/10.1109/TIP.2021.3084101)]
- 29 Cun XD, Pun CM. Defocus blur detection via depth distillation. *Proceedings of the 16th European Conference on Computer Vision*. Glasgow: Springer, 2020. 747–763.
- 30 王盼红, 朱昌明. MIF-CNNIF: 一种基于 CNN 的交叉特征的多分类图像数据框架. *计算机科学*, 2022, 49(S2): 490–497.
- 31 Diao SY, Su JZ, Yang CQ, *et al.* Classification and segmentation of OCT images for age-related macular degeneration based on dual guidance networks. *Biomedical Signal Processing and Control*, 2023, 84: 104810. [doi: [10.1016/J.BSPC.2023.104810](https://doi.org/10.1016/J.BSPC.2023.104810)]

(校对责编: 孙君艳)