

# 基于改进YOLOv5的视网膜黄斑病变 分类检测算法

王楠楠, 吴其洲, 王召巴\*, 金永

(中北大学信息与通信工程学院, 山西太原 030051)

**摘要:** 视网膜黄斑病变是失明的重要原因之一, 人工筛查容易出现误诊, 而基于深度学习的自动诊断有助于早期的检测和治疗。提出一种基于改进YOLOv5的黄斑病变分类检测算法, 针对黄斑病变图像细微特征融合不充分的问题, 将YOLOv5颈部的PANet特征融合模块替换为加权双向特征金字塔网络, 实现高效的多尺度特征融合, 以获得更好的黄斑病变细节特征; 针对小目标病变检测能力差的问题, 在模型中引入SK注意力机制, 通过自适应地调整感受野, 增强对黄斑病变区域特征的捕获。对比实验证明所提算法可将小目标检测准确率由原来的91.9%提升至94.2%, 全类平均准确率由原来的93.4%提升到96.6%, 且在相同条件下, 该算法的表现优于其他目标检测网络模型。

**关键词:** 目标检测; 视网膜黄斑病变; 加权双向特征金字塔网络; 注意力机制

**中图分类号:** TP391

**文献标识码:** A

**doi:** 10.62756/csjs.1671-7449.2025029

**引用格式:** 王楠楠, 吴其洲, 王召巴, 等. 基于改进YOLOv5的视网膜黄斑病变分类检测算法[J]. 测试技术学报, 2025, 39(2): 130-137.

WANG Nannan, WU Qizhou, WANG Zhaoba, et al. Classification and detection algorithm of macular disease based on improved YOLOv5[J]. Journal of Test and Measurement Technology, 2025, 39(2): 130-137.

## Classification and Detection Algorithm of Macular Disease Based on Improved YOLOv5

WANG Nannan, WU Qizhou, WANG Zhaoba\*, JIN Yong

(School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

**Abstract:** Macular disease is one of the important causes of blindness, manual screening is prone to misdiagnosis, and automatic diagnosis based on deep learning helps early detection and treatment. A classification and detection algorithm for macular lesions based on improved YOLOv5 was proposed. To solve the problem of insufficient fusion of fine features in macular lesions images, a weighted bidirectional feature pyramid network replaced the PANet feature fusion module of YOLOv5 neck to achieve efficient multi-scale feature fusion to obtain better-detailed features of macular lesions. To solve the problem of poor detection ability of small target lesions, the SK attention mechanism was introduced into the model to enhance the capture of regional features of macular lesions by adjusting the receptive field adaptively. Comparative experiments show that the proposed algorithm can improve the detection accuracy of small targets

**收稿日期:** 2024-09-27

**基金项目:** 山西省自然科学基金资助项目(202103021224202); 山西省归国留学人员科研基金资助项目(20210038)

**作者简介:** 王楠楠(2001—), 女, 硕士生, 主要从事计算机视觉和图像处理研究。E-mail: wangnn1081@163.com。

\* **通信作者:** 王召巴(1967—), 男, 教授, 博士, 主要从事信息探测处理、无损检测、人工智能研究。E-mail: wangzb@nuc.edu.cn。

from 91.9% to 94.2%, and the average accuracy of the whole class from 93.4% to 96.6%. Moreover, under the same conditions, the algorithm performs better than other target detection network models.

**Key words:** object detection; macular degeneration of the retina; weighted bidirectional feature pyramid network; attention mechanism

## 0 引言

人类对外部世界的认知超过 80% 来自于视觉,而这些信息中有 90% 以上来源于黄斑<sup>[1]</sup>。黄斑是视网膜光学中心,也是视网膜视力最敏锐的区域,黄斑的健康确保了视觉中央区域敏锐、细节化的视功能。然而,随着人们生活水平的提高和寿命不断延长,黄斑疾病的发病率也不断提高,黄斑疾病可能造成严重的视力损伤,该病会对视力产生不可逆转的影响,病情严重的患者将面临视力丧失的风险。因此,及时进行视网膜黄斑病变的早期筛查和诊断尤为重要<sup>[2-3]</sup>。

解决这一问题的有效方案是建立一种高效且准确的视网膜黄斑病变图像识别模型协助医生对视网膜黄斑病变患者进行诊断。从 2012 年至今,深度学习在图像识别领域取得了显著的进展,是众多专家学者的研究热点。Lee 等<sup>[4]</sup>使用 VGG16 网络对年龄相关性黄斑变性(AMD)与正常视网膜的光学相干断层扫描(OCT)图像进行识别,取得了 0.928 的曲线下面积(AUC)分类精度。2020 年, Bhadra 等<sup>[5]</sup>提出了一种多层卷积神经网络,针对玻璃膜、脉络膜新生血管、糖尿病性黄斑水肿和正常图像进行四分类识别,其 4 个类别的分类准确率分别达到了 89%、92%、98% 和 96%,平均准确率为 96.5%。尽管该方法在总体分类效果上表现良好,但对玻璃膜疣这一类疾病的分类准确率较低。2023 年,杨文意等<sup>[6]</sup>提出了一种改进的卷积神经网络模型用于黄斑病变分类,在卷积神经网络中引入了多尺度特征融合模块和注意力机制,该方法在四分类任务中达到了 96.8% 的平均准确率。

Karri 等<sup>[7]</sup>提出了一种基于迁移学习的视网膜病变识别算法,选用 GoogLeNet 作为特征提取网络,该算法在糖尿病性黄斑水肿、干性年龄相关性黄斑变性和正常图像分类的准确率分别为 86%、89% 和 99%。连超铭等<sup>[8]</sup>基于迁移学习提出了一种改进的深度学习网络用于自动分类视网膜图像,该网络对视网膜病变图像的分类准确率

达到了 97.3%。

以上研究在视网膜 OCT 病变分类任务中效果显著,但仍面临以下两个问题:1) 视网膜 OCT 图像中存在大量冗余,现有模型无法充分捕捉视网膜图像中的细微特征,从而导致分类效果不理想;2) 玻璃膜疣病变位置小且形态模糊,这增加了该类疾病分类的难度,目前其识别准确率相对较低。

针对上述问题,本文提出一种基于改进 YOLOv5 的视网膜黄斑病变检测方法。主要工作有以下 3 个方面:

1) 在专业医生指导下,对病变图像进行标注,构建视网膜黄斑病变数据集,该数据集包含脉络膜新生血管、糖尿病性黄斑水肿、玻璃膜疣 3 类 OCT 图像各 2 000 张,总计 6 000 张,用于多目标的视网膜黄斑病变检测;

2) 提出了一种基于改进 YOLOv5 的视网膜黄斑病变多目标检测模型,为了解决现有模型在捕捉黄斑病变图像细微特征方面的不足,在特征提取网络中,采用加权特征金字塔网络替代原有的 Concat 模块,通过调整权重融合不同尺度的特征信息,增强网络对黄斑病变区域的定位能力,从而获得更丰富的黄斑病变图像特征;

3) 针对玻璃膜疣小目标检测精度较差的问题,在该模型的特征融合网络中引入了 SK 注意力机制模块。旨在不增加网络复杂度的前提下,扩大感受野并按不同比例获取上下文信息,有效增强对黄斑病变区域的特征捕获,从而提升小目标病变检测的效果。

## 1 模型方法

目前主流的目标检测算法主要分为两个类型:一是 RCNN 系列算法,实现复杂,计算量大,推理速度慢,但检测精度高;二是 YOLO<sup>[9]</sup>和 SSD<sup>[10]</sup>系列算法,训练速度快,但检测精度较低。因此,为了同时兼顾视网膜黄斑病变检测的实时性和准确性,以及针对玻璃膜疣这一小目标病变识别准确率较低的问题,本文以 YOLOv5s 模型为基础,对 YOLOv5s

模型的部分模块进行精简与改进。

### 1.1 改进YOLOv5黄斑病变检测模型

YOLO单阶段目标检测算法具有适应性强、速度快的特点<sup>[11]</sup>。YOLOv5根据不同的网络深度分为YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x 4种模型。随着网络深度的增加,模型参数量逐渐增大,不利于后续模型在网络设备上的部署和推理。因此,本文选择了轻量化的YOLOv5s模型作为基线网络。

目前提出的相关分类方法中,玻璃膜疣类别的分类准确率较视网膜OCT图像中的其他3个类

别有待提高。为改善视网膜黄斑病变检测算法模型的性能,本文对YOLOv5s原始网络做出如下改进:将YOLOv5s颈部的滤镜聚合网络模块替换为加权双向特征金字塔网络,通过强化特征的双向流动来提升特征融合的效率 and 有效性,简单快速的进行多尺度特征融合,同时在特征融合网络的结合加入SK注意力模块,提升多尺度特征的目标检测精度,最后,采用SIoU损失函数代替原模型中的CIoU损失函数,在增强模型对小目标和尺寸变化大类型病变检测能力的同时可获得更准确的边界框回归,进一步加快收敛速度。改进后的YOLOv5s模型的网络架构如图1所示。

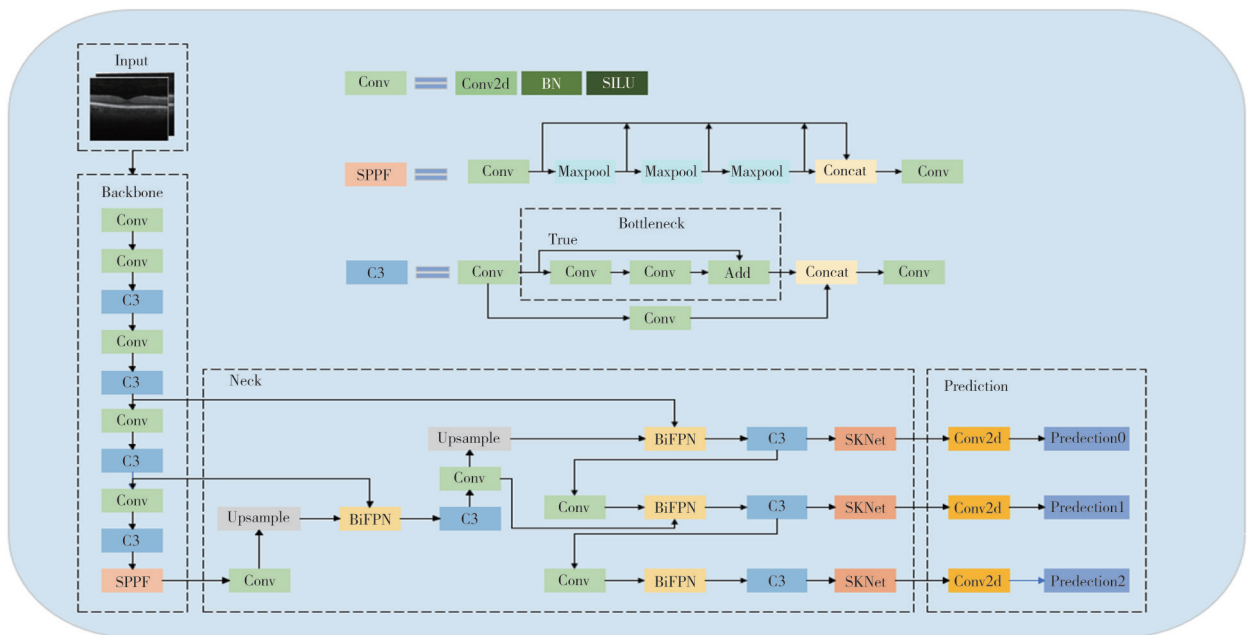


图1 改进YOLOv5s网络结构

Fig.1 Improved YOLOv5s network structure

### 1.2 引入加权双向特征金字塔网络

在原始YOLOv5s模型中,图像特征通过PANet<sup>[12]</sup>进行聚合和输出。然而,当模型处理不同分辨率的视网膜黄斑病变OCT图像时,由于病变区域在图像中的位置和大小各异,这种方法面临位置信息偏差和难以突出小目标特征的问题。此外,模型未能充分利用不同尺度的目标特征,从而降低了特征融合的有效性,严重影响了OCT图像的分类效果。

BiFPN<sup>[13]</sup>是一种在PANet基础上进行改进的双向特征融合网络,其核心思想主要有两个部分:高效的双跨尺度连接和加权特征图融合<sup>[14]</sup>,如图2(c)所示。为了减少计算量,BiFPN删除了

第3层和第7层中只有一个输入的特征融合节点,因为该节点与前一个节点所传递的信息是相同的;同时,BiFPN增加了跨尺度连接方式(图2(c)中横向曲线3个连接),从原始输入到输出节点添加一条额外的边,通过自上而下与自下而上两条特征融合路径,将深层特征与浅层特征跨尺度融合,以便在不增加太多成本的情况下融合更多的功能。这种模块化设计使得BiFPN能够方便地融入各种神经网络架构,特别是在处理复杂背景和多尺度物体时,展现出优异的性能,从而增强模型的整体表现和泛化能力。本文BiFPN加权特征融合计算方法为

$$O = \sum_i \frac{w_i I_i}{\varepsilon + \sum_j w_j}, \quad (1)$$

式中:  $O$  为多路径融合后的特征信息;  $I_i$  为输入特征信息;  $w_i$  为路径  $i$  对应的学习权重;  $\epsilon$  为常数。

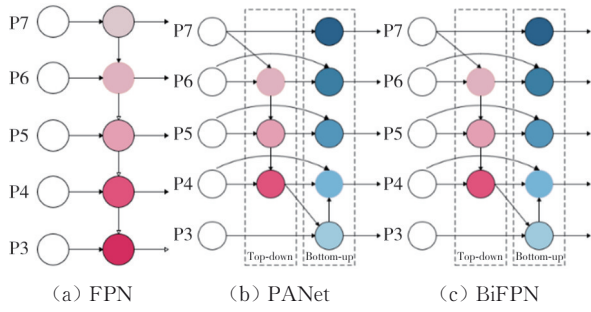


图 2 不同特征金字塔结构对比

Fig. 2 Contrast of pyramid structure with different features

### 1.3 引入SK注意力机制

在医学图像分类任务中,经典的深度学习网络通常通过增加网络的深度来扩展感受野,旨在捕获

更丰富的上下文信息。然而,这种方法在处理分辨率较小的特征图时往往会出现信息损失,特别是在复杂的医学图像中,细节信息的丢失会影响最终的分类效果。

为解决该问题,本文提出在特征融合网络中加入SK注意力机制模块<sup>[15]</sup>,其核心思想是利用通道注意力机制来动态调整不同特征通道的权重,使网络能够更有效地整合来自不同层次的特征,通过对特征图进行加权处理,自动识别出有助于分类的关键特征,同时抑制噪声和冗余信息的影响。这种方法不仅能够提升特征图的表达能力,还能在保持高效性的同时减少信息损失,提升网络对多目标视网膜病变检测的效果。SKNet是一个轻量级的嵌入式模块,由Split, Select和Split 3个操作来实现<sup>[16]</sup>,其结构如图3所示。

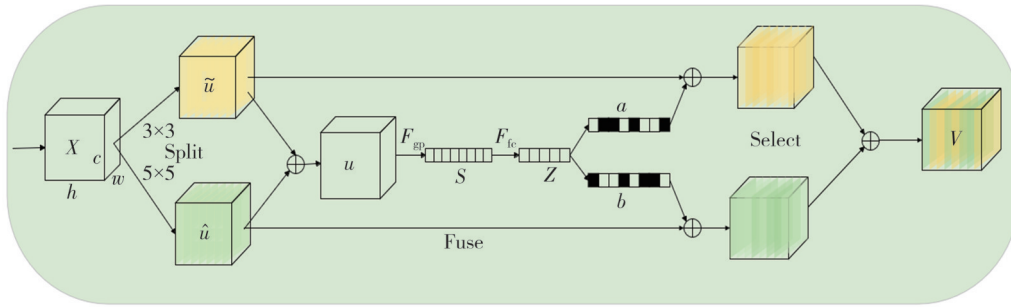


图 3 SKNet 结构

Fig. 3 SKNet structure

通过选择性地应用不同尺度的卷积核,SKNet 能够有效地融合多尺度的特征信息。这有助于网络捕捉不同层次的视觉特征,提高了特征的代表能力,虽然引入了多尺度卷积核,但由于选择模块的存在,SKNet 只会选择一部分卷积核进行计算,从而减少了计算成本,保持了网络的高效性。

Split 操作:对于任意给定的特征映射  $X \in \mathbb{R}^{H \times W \times C}$ ,SK注意力机制对特征图  $X$  进行卷积核为3和5的分支分离卷积转换得到特征图  $\tilde{U}$  和  $\hat{U}$ :  $\tilde{I}: X \rightarrow \tilde{U} \in \mathbb{R}^{H \times W \times C}$  和  $\hat{I}: X \rightarrow \hat{U} \in \mathbb{R}^{H \times W \times C}$ 。其中  $\tilde{I}$  和  $\hat{I}$  均由高效的分组卷积,批量标准化和Relu函数组成。

Fuse 操作由两个全连接层、一个全局平均池化及Relu激活函数组成,对融合多个感受野信息  $U$  进行全局平均池化操作后通过一个完全连接层降低维度以提高效率,编码方式如式(2)和式(3)所示。

$$S = F_{gc}(U_c) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j), \quad (2)$$

$$Z = F_{fc}(s) = \delta(B(W_s)), \quad (3)$$

式中:  $H$  和  $W$  分别为与当前注意力模块相对应的输入特征图的高度和宽度;  $S$  为融合多个感受野信息  $U$  的第  $c$  通道输出; 紧凑特征  $Z \in \mathbb{R}^{d \times 1}$ ;  $\delta$  为Relu激活函数;  $B$  为批量标准化。

Select 操作:为动态选择不同空间尺度的信息,SKNet 通过紧凑特征  $Z$  来指导跨通道软注意力机制。具体地,对每个分离通道的特征图应用softmax函数得到每个通道的权重,如式(4)和式(5)所示。

$$a_c = \frac{e^{A_c}}{e^{A_c} + e^{B_c}}, \quad (4)$$

$$b_c = \frac{e^{B_c}}{e^{A_c} + e^{B_c}}, \quad (5)$$

式中:  $A, B \in \mathbb{R}^{c \times d}$ ;  $a, b$  为  $\tilde{U}$  和  $\hat{U}$  通过跨通道注意力机制计算得到的每个通道的权重矩阵,  $A_c \in \mathbb{R}^{1 \times d}$  是  $A$  的第  $c$  行,  $a_c$  为  $a$  的第  $c$  个元素,  $B_c, b_c$  同理。

在两分支的情况下,最终得到的特征映射  $V$  是通过各个卷积核的注意力权重获得的,如式(6)所示。

$$V_c = a_c \times \tilde{U}_c + b_c \times \hat{U}_c, \quad a_c + b_c = 1, \quad (6)$$

式中:  $V=[V_1, V_2, \dots, V_c]$ ,  $V_c \in \mathbb{R}^{H \times W}$ 。

为了更好地捕获复杂黄斑病变图像的多尺度特征, 本文在此基础上进行扩展, 采用4种不同尺度的卷积核(即  $\text{kernels}=[1, 3, 5, 7]$ ), 使网络能够更灵活地选择合适的卷积核尺度, 以实现多尺度病变特征的动态融合。

综上所述, 本文通过选择性地应用不同尺度的卷积核, 使网络能够自适应地调整感受野的大小, 从而兼顾多个尺度的病变信息。在不增加网络复杂性的前提下, 有效提升了网络对玻璃膜疣这一小目标病变类别的特征提取能力。

#### 1.4 损失函数

YOLOv5算法的损失函数由分类损失、目标置信度损失、目标框与预测框位置损失(也称边界框损失)等3个部分组成。其中, YOLOv5原始模型的边界框损失函数为 CIoU loss。CIoU 是在 IoU 的基础上提出的一种改进, 其通过引入预测框和真实框的最小外接矩形来获取预测框、真实框在闭包区域中的比重。但 CIoU 在实际训练过程中存在以下两个问题: 1) CIoU 损失函数在处理小目标时不够敏感; 2) CIoU 损失函数对于长宽比例不平衡的目标检测框不够稳定, 长宽比比例差异较大会导致不稳定的梯度更新, 影响模型的收敛性和性能。因此, 为解决上述问题, 本文采用一种新的损失函数 SIOU。SIOU 损失函数是一种考虑到边界框匹配方向的目标检测损失函数, 由角度、距离、形状和 IoU 4 个部分构成。

角度成本

$$\Delta = 1 - 2 \sin^2 \left( \arcsin x - \frac{\pi}{4} \right), \quad (7)$$

距离成本

$$\Delta = \sum_{k=x,y} (1 - e^{-\gamma \rho_k}), \quad (8)$$

形状成本

$$\Omega = \sum_{k=w,h} (1 - e^{-w_k})^\theta, \quad (9)$$

IoU 成本

$$IoU = \frac{|A \cap B|}{|A \cup B|}, \quad (10)$$

式中:  $\Delta$  为角度损失, 并对距离误差进行重新定义;  $\rho_{k=x} = \left( \frac{b_{cx}^{gt} - b_{cx}}{c_w} \right)$ ,  $\rho_{k=y} = \left( \frac{b_{cy}^{gt} - b_{cy}}{c_h} \right)$ ,  $\gamma = 2 - \Delta$ ;  $\Delta$  为距离损失;  $\Omega$  为对 CIoU 中欧氏距离计

算方法的改进;  $w_{k=w} = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})}$ ,  $w_{k=h} = \frac{|h - h^{gt}|}{\max(h, h^{gt})}$ ,  $\theta$  为控制形状误差关注程度;  $A$  为预测框;  $B$  为真实框。

结合 IoU 和上述公式最终得出目标框损失函数表达式为

$$L_{\text{box}} = 1 - IoU + \frac{\Delta + \Omega}{2}. \quad (11)$$

## 2 实验结果与分析

为验证本文提出方法的有效性与先进性, 通过消融实验及对比试验进行分析。消融实验通过逐步引入两个关键创新点并与基线网络进行比较, 分析各个组件对整体性能的影响。对比试验则选取了当前主流的目标检测算法作为参考, 测试其在相同条件下的性能表现。通过这两种实验方法系统地评估了本文方法在精度和效率上的优势。

### 2.1 视网膜黄斑病变数据集构建

由于医疗图像数据的获取难度较大, 且标注过程需要专业医生的指导, 目前尚无公开标注的视网膜黄斑病变数据集。因此, 本文自建视网膜黄斑病变数据集, 步骤如下: 1) 数据收集: 在 Kaggle 公开的视网膜 OCT 图像分类数据集中手动筛选, 去掉重复或者病变特征不完整的图像, 针对脉络膜新生血管(CNV)、糖尿病性黄斑水肿(DME)和玻璃膜疣(Drusen) 3 种病变, 如图 4 所示, 分别选取了 2 000 张图像, 共计 6 000 张图像作为训练集, 另外选取 750 张图像作为测试集, 数据集之间没有任何交叉; 2) 图像标注: 在专业医生的指导下, 根据表 1 视网膜黄斑病变的临床表现特征, 使用精灵标注助手软件对 OCT 图像中病变的位置和类型进行标注, 生成了包含病变坐标、类别、图像大小等信息的 XML 文件, 并转化为 YOLO.txt 格式, 生成数据的真实标签。

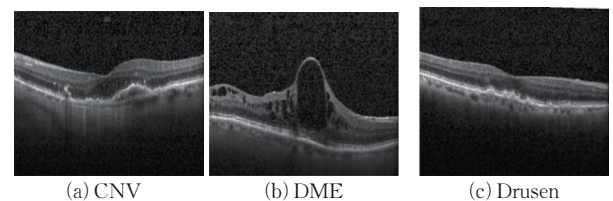


图 4 视网膜黄斑病变示例图

Fig. 4 Example diagram of macular degeneration

表 1 视网膜黄斑病变临床表现

病变类型	临床表现
Drusen	色素上皮层上的连续驼峰状凸起
CNV	较 Drusen 更大的拱形凸起, 通常呈单峰状
DME	在 Henle 纤维层中形成较大的圆形孔洞

### 2.2 实验环境

实验的深度学习框架为 Python 3.9.16、Pytorch 1.12.1 和 CUDA 11.1, 采用的 CPU 为 Interl(R) Xeon(R) CPU E5-2690 V3@2.60 GHz, GPU 为 NVIDIA GeForce RTX 2080Ti, 模型训练的超参数设置如表 2 所示。

表 2 实验超参数设置

超参数名称	超参数设置
Batch Size	16
优化器	SGD
动量系数	0.937
初始学习率	0.01
权重衰减系数	5e-4
迭代次数	300

### 2.3 评价指标

在测试集上对训练网络进行性能测试, 模型采用精确率、召回率和准确率等指标进行评估, 评价指标计算如下:

精准率

$$Precision = \frac{TP}{TP + FP},$$

召回率

$$Recall = \frac{TP}{TP + FN},$$

准确率

$$Accuracy = \frac{TP}{TP + FN + FP + TN},$$

式中:  $TP$  为目标病变类别正确分类的数量;  $TN$  为其他病变类别正确分类的数量;  $FP$  为目标病变类别分类错误的数量;  $FN$  为误检为其他病变类别的数量。

### 2.4 实验分析

#### 2.4.1 模型消融实验

为了验证每个组件在模型中的贡献, 本文开展了消融实验, 比较了改进后的 YOLOv5s 模型、仅加入 BiFPN 模块、仅加入 SKNet 模块和基线网络的性能表现, 实验结果如表 3 所示。

由表 3 可以看出, 加入 BiFPN 模块后, 网络的平均准确率相比基线网络提升了 2.2 个百分点。这表明 BiFPN 模块的引入能够更好地捕捉目标物体的关键信息, 同时减少背景噪声的干扰, 提升病变检测的准确率。加入 SK 注意力机制后, Drusen 病变的识别效果明显改善, 准确率提升了 2.7 个百分点。实验结果表明, SK 注意力机制的引入能够根据输入特征的差异动态调整特征权重, 确保在特征提取过程中小目标病变的细节信息得以保留, 并有效抑制无关特征, 从而显著提升了 Drusen 的识别效果。改进后模型的平均准确率较基线网络提升了 3.2 个百分点, CNV、DME、Drusen 3 个类别分别提升了 2.8、3.7、2.3 个百分点, 改进后模型的检测速度为 8.1 帧/ms, 相比基线网络慢 1.9 帧/ms, 说明在黄斑病变分类检测任务中, 改进后模型在准确率提升的同时仍能保持较优的实时性。

表 3 消融实验结果对比

方法	类别	Accuracy/%	Precision/%	Recall/%	平均准确率/%	检测速度/(帧·ms <sup>-1</sup> )
YOLOv5s	CNV	95.9	91.9	93.8	93.4	6.2
	DME	92.4	85.9	84.5		
	Drusen	91.9	83.9	72.7		
YOLOv5s+BiFPN	CNV	97.9	92.8	96.9	95.6	7.7
	DME	95.7	89.9	92.9		
	Drusen	93.4	84.6	89.0		
YOLOv5s+SKNet	CNV	97.9	94.5	94.6	96.0	6.9
	DME	95.6	90.7	89.7		
	Drusen	94.6	87.3	86.0		
YOLOv5s+BiFPN+SKNet	CNV	98.7	94.5	97.9	96.6	8.1
	DME	96.1	90.8	92.9		
	Drusen	94.2	91.6	87.5		

#### 2.4.2 对比实验

为进一步分析本文改进模型在视网膜黄斑病变检测上的性能, 将本文所提的改进 YOLOv5 病

变检测模型与 YOLOv4、YOLOv5s、YOLOv7、YOLOv8s、SSD 等主流目标检测模型进行对比, 结果如表 4 所示。

表4 不同算法的检测结果对比  
Tab. 4 Comparison of detection results of different algorithms

算法	Accuracy/%			平均准确率/%	参数量/M	FLOPS/G
	CNV	DME	Drusen			
YOLOv4	87.4	92.7	90.6	90.2	52.5	119.83
YOLOv5s	95.9	92.4	91.9	93.4	7.02	15.8
YOLOv7	97.3	92.0	88.0	92.4	36.49	103.5
YOLOv8s	96.9	94.6	91.4	94.3	11.2	28.6
SSD	91.9	92.6	90.5	91.7	20.0	66.9
本文算法	98.7	96.1	94.2	96.6	38.17	40.6

由表4可以看出:1)无论是平均准确率或是Drusen小目标的分类准确率, YOLOv5s作为主干网络比SSD作为主干网络在视网膜OCT病变检测中更有优势;2)YOLO系列算法随着版本的不断改进,病变检测精度也得到了提升,其中YOLOv7模型的检测精度与基线网络YOLOv5s相比较差, YOLOv8s和YOLOv5s相比有较高的检测精度,但是在Drusen类别准确率上比YOLOv5s低0.5个百分点;3)改进YOLOv5s模型的数量和FLOPS与原模型相比增加了31.15M和24.8G,结合表3数据,虽然模型复杂度有所增加但对其检测速度的影响较小,改进后的模型在准确率提升的同时仍能保持较优的实时性;4)采用YOLOv5s作为基线网络在视网膜OCT病变检测中具有检测精度高、检测速度快特点,本文算法在3种病变的检测效果和整体检测精度都较为理想。

## 2.5 可视化结果

为了直观地验证本文算法的检测效果,将3种病变的检测结果与原YOLOv5s模型的检测结果做对比,如图5所示。

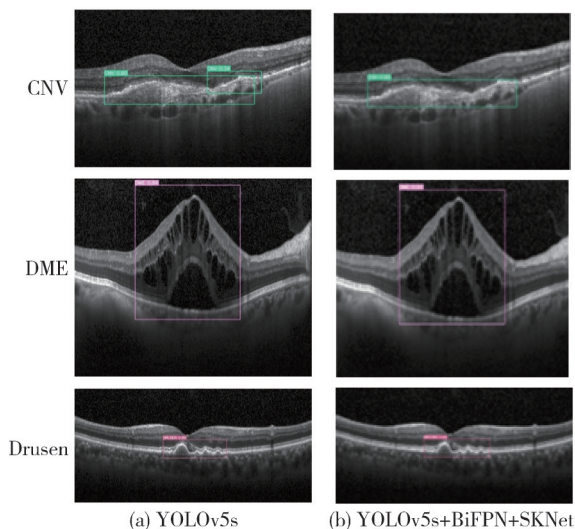


图5 改进前后模型检测结果对比  
Fig 5 Comparison of model detection results before and after improvement

从图5可以看出, YOLOv5s模型的检测精度较低,同时在CNV类别出现了误检现象,改进后的模型显著提高了3种黄斑病变类别的检测精度,降低了误检的概率,有效地验证了本文算法相对原始模型的性能提升。

## 3 结论

本文提出一种基于改进YOLOv5的视网膜黄斑病变分类检测算法。将加权双向金字塔网络融合模块应用于颈部,并结合引入SK注意力的推理层,在保证检测速度的基础上提升检测精度。使用Siou损失函数以提高定位的精度和鲁棒性。实验结果表明,该算法提升了基于视网膜OCT黄斑病变图像的检测效果,将整体检测精度提升到96.6%, Drusen的小目标检测精度达到94.2%。

该模型与传统的识别算法相比,可以在较快的训练速度下保持更高的识别率,为眼科医生提供了一种便捷、高效的诊断辅助工具。

## 参考文献:

- [1] 陈彬彬, 楼丽霞, 叶娟. 中国眼病疾病负担现状及三十年变化趋势[J]. 浙江大学学报(医学版), 2021, 50(4): 420-428.  
CHEN Binbin, LOU Lixia, YE Juan. Eye diseases burden in China in the past 30 years[J]. Journal of Zhejiang University (Medical Sciences), 2021, 50(4): 420-428. (in Chinese)
- [2] 韩璐. 基于深度学习的视网膜病变分类与检测任务研究[D]. 哈尔滨: 哈尔滨工程大学, 2022.
- [3] 周鹏. 基于深度学习的眼底疾病诊断研究[D]. 南京: 南京航空航天大学, 2020.
- [4] LEE S C, BAUGHMAN M D, LEE Y A. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images[J]. Ophthalmology Retina, 2017, 1(4): 322-327.
- [5] BHADRA R, KAR S. Retinal disease classification from optical coherence tomographical scans using multi-layered convolution neural network[C]//2020 IEEE

- Applied Signal Processing Conference (ASPCON), IEEE, 2020; 212-216.
- [6] 杨文意, 陈雯, 周兰, 等. 一种用于黄斑病变分类的改进卷积神经网络模型[J]. 西华师范大学学报(自然科学版), 2023, 44(3): 318-325.
- YANG Wenyi, CHEN Wen, ZHOU Lan, et al. An improved convolutional neural network model for macular diseases classification [J]. Journal of China West Normal University (Natural Science), 2023, 44(3): 318-325. (in Chinese)
- [7] KARRI S P K, CHAKRABORTY D, CHATTERJEE J. Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration[J]. Biomedical Optics Express, 2017, 8(2): 579-592.
- [8] 连超铭, 钟舜聪, 张添福, 等. 光学相干断层扫描视网膜图像的迁移学习分类[J]. 激光与光电子学进展, 2021, 58(1): 262-268.
- LIAN Chaoming, ZHONG Shuncong, ZHANG Tianfu, et al. Transfer learning-based classification of optical coherence tomography retinal images[J]. Laser & Optoelectronics Progress, 2021, 58(1): 262-268. (in Chinese)
- [9] KAROLINA P. Vehicle detection and masking in UAV images using YOLO to improve photogrammetric products[J]. Reports on Geodesy and Geoinformatics, 2022, 114(1): 15-23.
- [10] SHEPING Z, DINGRONG S, SHUHUAN W, et al. DF-SSD: an improved SSD object detection algorithm based on DenseNet and feature fusion[J]. IEEE Access, 2020, 8: 24344-24357.
- [11] ZHU X, LÜ S, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2778-2788.
- [12] WANG K, LIEW J H, ZOU Y, et al. Panet: Few-shot image semantic segmentation with prototype alignment [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9197-9206.
- [13] 徐明升, 祝俊辉, 干家欣, 等. 基于 YOLOv5 的钢材表面缺陷检测算法[J]. 无线电工程, 2024, 54(2): 351-359.
- XU Mingsheng, ZHU Junhui, GAN Jiabin, et al. Steel surface defect detection algorithm based on YOLOv5[J]. Radio Engineering, 2024, 54(2): 351-359. (in Chinese)
- [14] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [15] LI X, WANG W, HU X, et al. Selective kernel networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 510-519.
- [16] 肖维颖, 王健, 李文顺. 松树株数识别的 YOLOv5 轻量化算法研究[J]. 森林工程, 2023, 39(4): 126-133.
- XIAO Weiyang, WANG Jian, LI Wenshun. Research on YOLOv5 lightweight algorithm for pine tree strain identification [J]. Forest Engineering, 2023, 39(4): 126-133. (in Chinese)