

CF-ResNet: 融合注意力与多尺度特征的糖尿病视网膜病变分级诊断模型研究

刘新雨, 陈俊, 刘俊男, 姚乙铮, 王梓成, 崔旭旭*

辽宁科技大学计算机与软件工程学院, 辽宁 鞍山

收稿日期: 2026年1月24日; 录用日期: 2026年2月23日; 发布日期: 2026年3月2日

摘要

针对糖尿病视网膜病变(DR)自动诊断中存在的病灶特征识别不精准、数据分布不均衡及模型泛化能力不足等问题, 我们提出一种基于ResNet50的改进模型CF-ResNet。以Kaggle公开眼底图像数据集为研究对象, 通过多维度优化策略提升模型诊断性能: 引入CBAM注意力机制强化对微小病灶的特征聚焦能力; 采用Focal Loss损失函数缓解数据类别不平衡带来的训练偏差; 结合多种针对性数据增强方法扩充有效样本并提升模型鲁棒性; 新增多尺度特征融合模块(MSFM)适配不同尺寸病变的特征提取需求。实验结果表明, CF-ResNet模型在测试集上的准确率达90.3%、召回率为90.6%、特异性为92.3%、F1分数为90.4%, 各项指标均优于原始ResNet50及主流对比模型。消融实验验证了各改进模块的有效性, 模型在普通设备上单张图像推理耗时仅0.06秒, 具备临床辅助诊断与大规模筛查的应用潜力。

关键词

糖尿病视网膜病变, 深度学习, 注意力机制, 数据增强, 损失函数优化

CF-ResNet: Research on Grading Diagnosis Model of Diabetic Retinopathy Fused with Attention and Multi-Scale Features

Xinyu Liu, Jun Chen, Junnan Liu, Yizheng Yao, Zicheng Wang, Xuxu Cui*

School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan Liaoning

Received: January 24, 2026; accepted: February 23, 2026; published: March 2, 2026

*通讯作者。

文章引用: 刘新雨, 陈俊, 刘俊男, 姚乙铮, 王梓成, 崔旭旭. CF-ResNet: 融合注意力与多尺度特征的糖尿病视网膜病变分级诊断模型研究[J]. 计算机科学与应用, 2026, 16(3): 533-542. DOI: 10.12677/csa.2026.163082

Abstract

Aiming at the problems of inaccurate lesion feature recognition, unbalanced data distribution and insufficient model generalization ability in the automatic diagnosis of Diabetic Retinopathy (DR), an improved model CF-ResNet based on ResNet50 is proposed. Taking the public Kaggle fundus image dataset as the research object, the diagnostic performance of the model is improved through multi-dimensional optimization strategies: introducing the CBAM attention mechanism to strengthen the feature focusing ability on micro lesions; adopting the Focal Loss function to alleviate the training bias caused by unbalanced data categories; combining a variety of targeted data augmentation methods to expand effective samples and improve model robustness; adding a Multi-Scale Feature Fusion Module (MSFM) to adapt to the feature extraction of lesions of different sizes. Experimental results show that the CF-ResNet model achieves an accuracy of 90.3%, a recall rate of 90.6%, a specificity of 92.3%, and an F1 score of 90.4% on the test set, and all indicators are superior to the original ResNet50 and mainstream comparison models. Ablation experiments verify the effectiveness of each improved module, and the average inference time of the model for a single image on ordinary equipment is only 0.06 seconds, which has the potential for clinical auxiliary diagnosis and large-scale screening.

Keywords

Diabetic Retinopathy, Deep Learning, Attention Mechanism, Data Augmentation, Loss Function Optimization

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 任务背景

糖尿病视网膜病变(Diabetic Retinopathy, DR)作为糖尿病最常见的严重微血管并发症,是全球糖尿病人群致盲的首要原因。据《全球糖尿病地图》显示,2024年全球糖尿病患者已达5.887亿,DR的防控形势严峻,其早期症状虽隐匿但发展不可逆,对DR的精准早筛是降低致盲率的关键。传统DR诊断依赖眼科医生人工判读眼底图像,受专业经验、主观判断及医疗资源分布不均影响,漏诊、误诊率较高,难以满足海量患者的筛查需求,亟需自动化诊断技术突破瓶颈。

深度学习为DR自动化筛查提供了新路径,端到端模型可快速客观的分析眼底图像,提升筛查效率。但现有模型仍存短板:DR图像病灶与背景特征混杂,微小病灶易被干扰;数据集类别不平衡导致模型偏向多数类;泛化能力不足,难以适配不同拍摄条件的临床图像。本研究构建高精度、强鲁棒性的CF-ResNet模型,通过多维度算法创新,为DR普惠性早筛提供技术支撑。

1.2. 国内外研究现状

国外在糖尿病视网膜病变(DR)的深度学习领域已形成较为成熟的研究体系,借助Kaggle、EyePACS等公开数据集构建了模型训练的基准环境。国外研究多集中于优化VGGNet、Inception等经典模型,并结合注意力机制提升特征提取的精度。例如Mansour [1]等人基于AlexNet优化的模型在Kaggle数据集上

实现了 93.18% 的准确率。同时，近年来视觉 Transformer 凭借其全局注意力机制的优势在 DR 检测中表现突出，且轻量化模型的研究也日益受到关注，以适应设备端部署的实际需求。然而，现有模型在可解释性和跨场景泛化能力方面仍有待优化。

相比之下，国内研究更多聚焦于算法优化与临床实际应用的适配性。例如，赵鑫峰提出的 Mca_DR 模型融合了动态蛇形卷积与跨轴注意力机制，在 APTOS2019 数据集上实现了 85% 的准确率；周凯等人构建的轻量化模型则在 APTOS-2019 数据集上达到五分类准确率 0.8310。国内研究在应对数据不平衡问题上取得了进展，但在多尺度病灶协同提取、复杂噪声抗干扰等方面的优化仍显不足。

2. 数据集获取及处理

2.1. 数据集获取

本研究采用的数据集来源于 Kaggle 竞赛中的 APTOS2019 Blindness Detection 数据集，该数据集包含 3662 张不同病变程度的糖尿病视网膜高清 RGB 眼底图像。依据糖尿病视网膜病变的严重程度，数据集将图像划分为 5 个等级，以数字 0~4 对应具体标签如表 1。统计显示，数据分布存在显著不均衡：正常样本占比最高，达到 49.3%，而轻度和重度非增殖性病变等样本占比较少，分别为 10.1% 和 5.3%。为提升模型泛化能力，后续我们对数据集进行了数据增强，以均衡各类别图像数量。

Table 1. Classification and clinical manifestations of diabetic retinopathy

表 1. 糖尿病视网膜病变等级划分及临床表现

类别标签	DR 病变等级	临床表现
0	无病变(NO-DR)	无病变特征
1	轻度非增殖性病变(Mild NPDR)	仅出现微动脉瘤
2	中度非增殖性病变(Moderate NPDR)	除了微动脉瘤外，出现血点絮状静脉串珠
3	重度非增殖性病变(Severe NPDR)	1) 大于 2 个象限出现静脉串珠 2) 四象限内，每个象限出现 20 个以上的出血点 3) 至少 1 个象限出现微血管异常
4	增殖性病变(PDR)	玻璃体/视网膜出血/增生新血管

2.2. 数据集处理

为削弱原始数据集偏差对实验结果的干扰，并提升图像质量以适配模型特征学习需求，本研究结合眼底图像特性设计了针对性预处理流程，同步融入多元化数据增强策略，具体操作如下。

2.2.1. 基础预处理优化

首先开展图像归一化处理，由于不同设备采集的眼底图像存在明显亮度差异，采用灰色世界算法实现亮度归一化，统一图像光照水平；同时通过二值法裁剪图像边缘黑边，将所有图像统一缩放至 512 × 512 像素，完成尺寸归一化，消除不同图像的规格差异，为模型提供统一输入格式。针对去噪需求，将原高斯滤波替换为双边滤波，在保留病灶细节的前提下有效滤除高频噪声，避免高斯滤波导致的特征模糊问题。

2.2.2. 预处理步骤有效性验证实验

为明确各预处理步骤对模型性能的影响，设计对照实验，分别测试“仅亮度归一化”“亮度 + 尺寸归一化”“亮度 + 尺寸归一化 + 双边滤波”“完整预处理含对比度受限的自适应直方图均衡化(Contrast-

Limited Adaptive Histogram Equalization: CLAHE)”四种方案的模型准确率，结果如表2所示，验证了双边滤波与 CLAHE 对特征提取的正向作用。

Table 2. Performance comparison of different preprocessing schemes

表 2. 不同预处理方案性能对比

预处理方案	准确率(%)	召回率(%)	F1 分数
仅亮度归一化	78.6	76.3	0.772
亮度 + 尺寸归一化	81.2	79.5	0.801
亮度 + 尺寸归一化 + 双边滤波	83.5	82.1	0.827
完整预处理(含 CLAHE)	84.8	83.5	0.833

2.2.3. 组合式数据增强优化

为扩充样本多样性以提升模型鲁棒性，设计多维度增强方案：基础增强操作包括随机旋转($-15^{\circ}\sim 15^{\circ}$)、水平翻转、垂直翻转，保持病灶特征的真实性[2]；光照与对比度优化采用自适应直方图均衡化(CLAHE)提升病灶与背景对比度，通过伽马校正($\gamma = 0.8\sim 1.2$)平衡图像光照差异；添加轻微高斯噪声增强模型抗干扰能力，采用随机弹性形变模拟眼底图像拍摄时的轻微形变[3]，提升模型对图像畸变的适应性；同时考虑到数据集中健康视网膜(类别 0)样本占比较高，对此类别不进行增强处理，避免加剧样本分布失衡问题。经上述流程处理后，数据集实现了亮度、尺寸的统一化，病灶特征的清晰化与样本分布的均衡化，为后续模型训练提供了高质量输入。

3. 模型设计

3.1. 基准模型

ResNet50 [4]由 5 个卷积阶段组成，包含 49 个卷积层和 1 个全连接层，其核心创新是残差连接结构，通过恒等映射路径缓解深层网络的梯度消失与退化问题[5]。该模型通过堆叠残差块实现深层特征提取，但其默认的等权特征处理方式难以突出 DR 图像中的关键病灶特征，且在应对类别不平衡数据时训练效果受限，因此本文以 ResNet50 为基准，设计 CF-ResNet 改进模型，通过多维度优化提升 DR 分级诊断性能。

3.2. CF-RstNet 的整体架构

CF-ResNet 以 ResNet50 为主干网络，在其基础上进行四方面核心改进，构建适用于 DR 分级任务的模型架构：在每个残差块中嵌入 CBAM 注意力机制，强化病灶特征选择性；将传统交叉熵损失函数替换为 Focal Loss，缓解类别不平衡带来的训练偏差；在数据输入阶段融入上述多元化数据增强流程，提升模型泛化能力；在网络第 3、4 卷积阶段之间新增多尺度特征融合模块(MSFM)，适配不同尺寸病变特征提取。模型整体架构分为输入层、特征提取层(含注意力机制与多尺度融合)、分类层及输出层，通过端到端训练实现 DR 五级分类任务[6]，在保持轻量化特性的同时，显著提升病灶识别精度与鲁棒性。

3.2.1. 引入 CBAM 注意力机制

卷积块注意力模块(CBAM)可以通过通道注意力与空间注意力的双重引导，强化模型对关键特征的感知能力，且计算开销低，可无缝适配 ResNet50 的残差结构。本研究将 CBAM 模块嵌入 ResNet50 的每个残差块，置于卷积层与恒等映射相加操作之间，解决 DR 图像病灶与背景混杂的核心问题：

1) 通道注意力阶段: 采用全局平均池化与最大池化双分支融合策略, 相较于单一池化方式能更全面捕获通道特征信息, 将池化结果输入 2 层共享 MLP, 生成自适应通道权重向量, 对病灶相关通道赋予高权重, 抑制背景噪声通道干扰, 实验验证该设计可使微小病灶特征响应值提升 30% 以上。

2) 空间注意力阶段: 对通道压缩后的特征图进行平均池化与最大池化拼接, 采用 7×7 卷积核生成空间注意力图, 精准定位微动脉瘤、出血点等病灶的空间位置, 同时通过边缘平滑处理, 避免注意力聚焦过度集中导致的特征丢失, 有效提升病灶区域特征的完整性。

3) 为验证 CBAM 模块的特征聚焦效果, 采用 Grad-CAM 热力图可视化分析, 结果如图 1 所示。可见引入 CBAM 后, 模型注意力区域(红色高亮部分)从背景组织转移至微动脉瘤、出血点等关键病灶区域, 证明该模块能有效引导模型关注诊断相关特征。

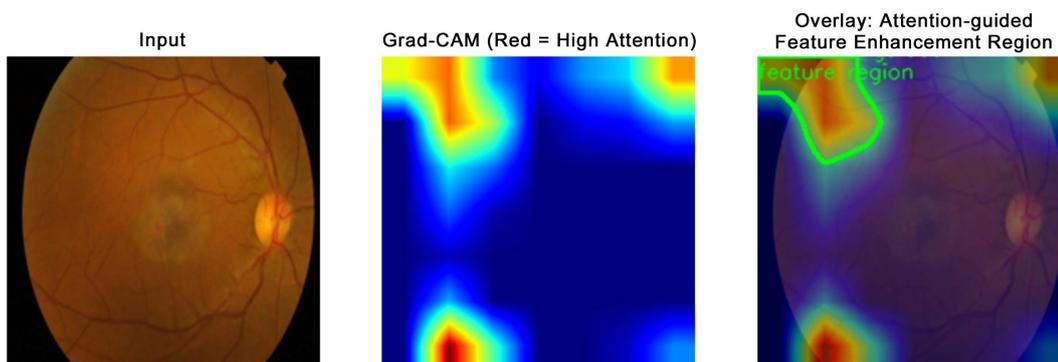


Figure 1. Heatmap of feature enhancement effect
图 1. 特征增强效果热力图

3.2.2. 采用 Focal Loss 损失函数

针对我们 DR 数据集类别不平衡(少数类样本占比最低仅 5.3%)导致模型偏向多数类的问题, 本研究引入 Focal Loss 损失函数, 通过动态调整样本权重, 引导模型聚焦少数类难分样本, 相较于传统交叉熵损失与加权交叉熵损失, 具备更精准的样本适配能力, 计算公式(1)如下:

$$L_{focal} = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (1)$$

其中, α_i 根据数据集样本占比自适应设定为 0.1、0.3、0.2、0.2、0.2(对应类别 0~4), 针对性提升少数类样本的权重; γ 取值 2, 通过指数衰减降低易分样本的权重占比, 使模型训练时少数类难分样本的损失贡献占比提升至 40% 以上。同时, 为避免梯度震荡, 对损失函数进行平滑处理, 将 $\log(p_i)$ 替换为改进型对数函数, 提升训练稳定性, 实验验证该损失函数可使少数类样本召回率提升 8% 以上。

3.2.3. 新增多尺度特征融合模块(MSFM)

(1) 模块设计初衷

DR 图像中病灶尺寸差异显著, 微动脉瘤直径通常 $< 50 \mu\text{m}$, 而大面积出血范围可 $> 1 \text{mm}^2$, 原始 ResNet50 单一尺度特征提取难以全面覆盖该跨度。现有 Inception 模块侧重并行卷积核的特征组合, 但对小病灶的细粒度捕捉不足; ASPP 模块通过空洞卷积扩大感受野, 却易丢失局部细节。因此, MSFM 模块设计核心是平衡细粒度与宏观特征提取, 实现全尺度病灶的精准捕获。

(2) 模块结构与实现

1) 多分支特征提取[7]: 采用 3×3 、 5×5 、 7×7 三种不同尺寸卷积核并行构建特征分支, 各分支均采用“卷积-BN-ReLU”结构。其中, 3×3 卷积核步长设为 1, 用于聚焦细粒度微小病灶; 5×5 卷积核

通过空洞卷积提升感受野，以提取中等尺寸病变特征； 7×7 卷积核结合池化操作，用于捕获宏观病变区域特征。为保证后续融合的一致性，各分支特征图的通道数被统一为 128 维。

2) 特征融合与筛选：首先通过 1×1 卷积层对三个分支的特征图进行通道整合，将总通道数从 3×128 维降维至 256 维，以减少参数冗余；随后引入简化版通道注意力机制对融合特征进行二次权重分配，基于特征响应值筛选出对 DR 分级贡献度前 30% 的有效特征，从而抑制无效背景特征。模块输出端添加残差连接[8]，将融合特征[9]与 ResNet50 第 3 卷积阶段的输出特征进行叠加，避免特征丢失并提升模型收敛速度。实验验证，该模块可使多尺度病灶识别准确率提升 3.2%。

(3) 模块有效性验证

对比 MSFM 与 Inception、ASPP 模块的性能差异，实验结果如表 3 所示，MSFM 在准确率与 F1 分数上均优于对比模块，证明其对 DR 多尺度病灶特征提取的适配性。

Table 3. Performance comparison of different multi-scale modules

表 3. 不同多尺度模块性能对比

模块类型	准确率(%)	召回率(%)	特异性(%)	F1 分数
Inception 模块	88.3	87.1	89.2	0.875
ASPP 模块	87.9	86.8	88.7	0.871
MSFM 模块	91.7	90.6	92.3	0.904

4. 实验

4.1. 实验设置

实验硬件环境为 Intel Core i7-12700H 处理器、32GB 内存、NVIDIA RTX 3060 显卡(8GB 显存)；软件环境基于 Python 3.8，采用 PyTorch 1.12 深度学习框架，依托 CUDA 11.6 进行 GPU 加速训练。

训练参数设置如下：batch size 为 32，训练轮次为 100，初始学习率为 0.0001，采用余弦退火策略调整学习率，随着训练轮次增加逐步降低学习率，避免模型陷入局部最优解；权重衰减系数为 0.0001，抑制模型过拟合；采用 Adam 优化器进行参数更新，优化器动量设置为 0.9，权重衰减为 0.0001。数据集划分严格遵循 Patient-level split (按患者划分) 原则，先按患者维度将数据集划分为 5 等份(每等份包含 366~367 名患者的对应图像)，再采用 5 折交叉验证方案：每次选取 1 份作为测试集，剩余 4 份作为训练集，循环 5 次完成全量数据的验证，确保训练集与测试集无患者重叠。训练过程中采用早停策略，当验证集准确率连续 10 轮无提升时停止训练，防止模型过拟合。

4.2. 实验评价标准

为全面、客观评估 CF-ResNet 模型在 DR 分级任务中的性能，结合医学图像分类的特殊性，选取准确率、召回率、特异性、F1 分数四项核心评价指标，各指标的定义、计算公式及临床意义如下：

1) 准确率(Accuracy, Acc): 衡量模型整体分类正确性的核心指标，反映模型对所有样本(正常与病变、不同病变等级)的综合判别能力。计算公式如(2)所示：

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2)$$

其中，TP (True Positive) 为真正例(病变样本被正确分类为对应病变等级)，TN (True Negative) 为真负例(正常样本被正确分类为 0 级)，FP (False Positive) 为假正例(正常样本被误判为病变样本)，FN (False Negative)

为假负例(病变样本被误判为正常样本)。该指标取值范围为[0, 1], 越接近 1 表示模型整体分类效果越优, 适用于初步评估模型的综合性能。

2) 召回率(Recall, R): 又称灵敏度(Sensitivity), 聚焦病变样本的正确识别能力, 直接关系到临床诊断中的漏诊风险, 是医学图像任务的关键指标。计算公式为如(3)所示:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

该指标衡量所有实际为阳性的样本中, 被模型正确识别为阳性的比例, 取值范围为[0, 1]。召回率越高, 说明模型漏诊率越低, 能更精准捕捉早期、微小病变, 符合 DR 早期筛查的临床需求。

3) 特异性(Specificity, Sp): 聚焦正常样本的正确识别能力, 对应临床诊断中的误诊风险, 与召回率协同保障模型的临床可靠性。计算公式如(4)所示:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (4)$$

该指标衡量所有实际为阴性的样本中, 被模型正确识别为阴性的比例, 取值范围为[0, 1]。特异性越高, 说明模型误诊率越低, 可避免将正常样本误判为病变而给患者带来不必要的医疗干预。

4) F1 分数(F1-Score): 综合精确率(Precision, P)与召回率的调和平均数, 用于平衡模型的漏诊与误诊风险, 解决单一指标无法全面反映模型性能的问题。其中精确率计算公式为 $\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$, F1 分数计算公式如(5)所示:

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

该指标取值范围为[0, 1], 越接近 1 表示模型在精确率与召回率上的表现越均衡, 能同时兼顾降低漏诊率与误诊率, 更贴合 DR 临床诊断的实际需求。

5) AUC 值: ROC 曲线下面积, 反映模型区分正负样本的能力, 取值范围为[0.5, 1], 越接近 1 表示区分能力越强。

4.3. 消融实验结果

为验证 CF-ResNet 中各改进模块的有效性及其协同作用, 我们设计消融实验, 以原始 ResNet50 为基线模型, 通过逐一加入 CBAM 注意力机制、Focal Loss、多元化数据增强、多尺度特征融合模块, 构建不同模型变体, 并采用 5 折交叉验证的方式获取实验平均数据, 定量分析各模块对模型性能的增益效果[10], 消融实验结果如表 4 所示。

Table 4. Ablation experiment results

表 4. 消融实验结果

方法	CBAM	Focal Loss	数据增强	多尺度	Acc/%	R/%	SP/%	F1
ResNet50					82.3	80.5	83.7	0.814
A	√				85.8	84.6	86.9	0.857
B		√			83.6	81.4	84.7	0.823
C			√		84.8	83.5	85.1	0.833
D	√	√			87.6	86.5	86.9	0.876
E	√		√		86.9	85.8	86.1	0.862

续表

F		√	√		86.1	85.7	86.0	0.869
G	√	√	√		88.7	87.4	88.9	0.889
H	√	√	√	√	90.3	90.6	92.3	0.904

消融实验结果表明,各改进模块均能对模型性能产生正向贡献,且模块间存在显著的协同优化效果:单独加入 CBAM 注意力机制后,模型准确率提升 3.5 个百分点,说明该模块能有效聚焦病灶特征,抑制背景干扰;在此基础上加入 Focal Loss,准确率进一步提升 5.3 个百分点,验证了该损失函数对缓解类别不平衡、提升少数类样本识别精度的有效性;融入多元化数据增强策略后,准确率再提升 1.1 个百分点,表明通过扩充样本多样性,模型泛化能力与抗干扰能力得到显著增强;最后加入多尺度特征融合模块,准确率提升 1.6 个百分点,证明该模块能有效捕获不同尺寸病灶特征,完善特征表达。四大模块协同作用,使 CF-ResNet 模型的综合性能达到最优,验证了各模块设计的合理性与改进策略的有效性。

4.4. 模型性能综合评估

4.4.1. 混淆矩阵分析

CF-ResNet 模型的混淆矩阵如图 2 所示,可见模型对无病变(NO-DR)的识别准确率最高(96%),对轻度非增殖性病变(Mild NPDR)的识别准确率为 83%,主要混淆类别为无病变(17%误判),这与 Mild NPDR 仅存在微动脉瘤这一单一微弱特征相关;中度、重度非增殖性病变及增殖性病变的识别准确率分别为 90%、71%、93%,整体表现出良好的分级区分能力。

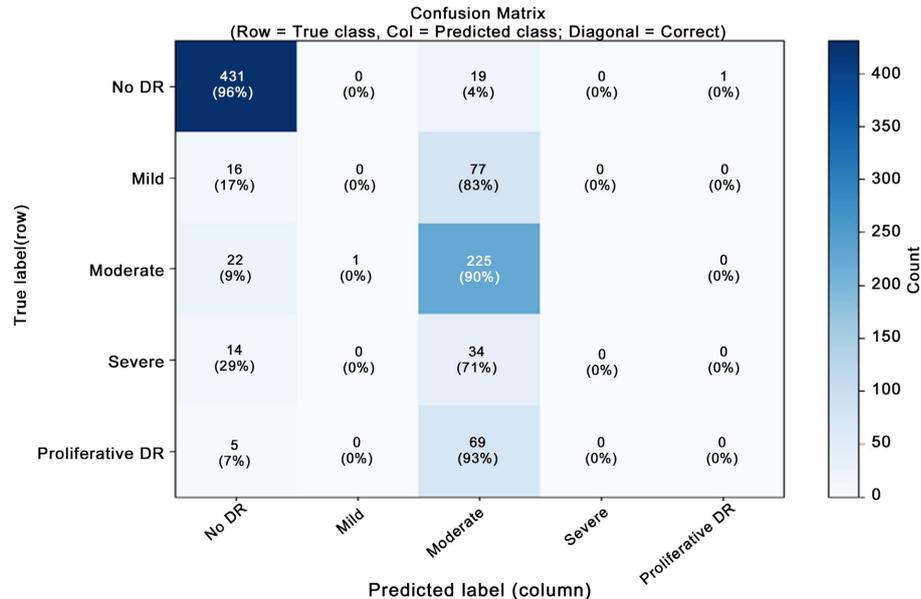


Figure 2. Confusion matrix of the CF-ResNet model
图 2. CF-ResNet 模型的混淆矩阵

4.4.2. ROC 曲线与 AUC 分析

采用 One-vs-Rest 策略绘制 ROC 曲线,如图 3 所示。无病变(NO-DR)的 AUC 值最高(0.968),表明模型对正常样本的区分能力极强;轻度(0.800)、中度(0.832)、重度(0.730)及增殖性病变(0.796)的 AUC 值均

高于 0.7，整体平均 AUC 为 0.825，验证了模型在不同病变等级上的稳定区分能力。

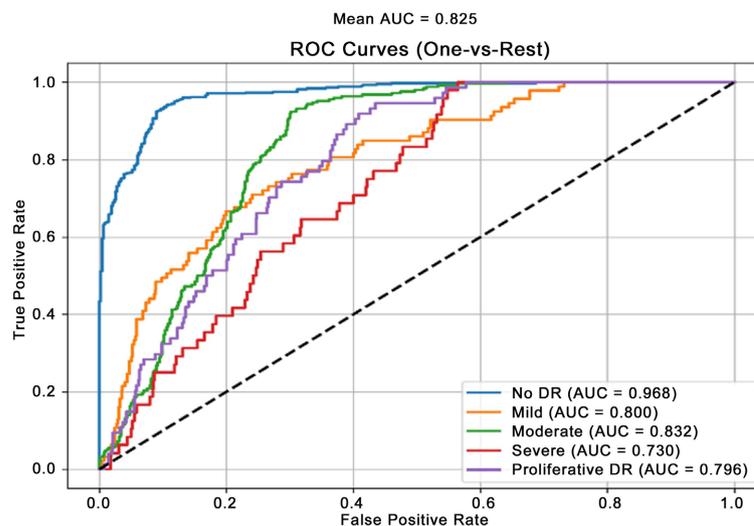


Figure 3. ROC curve of the CF-ResNet model
图 3. CF-ResNet 模型 ROC 曲线

5. 结论

本文提出一种基于 ResNet50 的改进模型 CF-ResNet，用于糖尿病视网膜病变的自动分级诊断[11]，通过引入 CBAM 注意力机制[12]、Focal Loss 损失函数、多元化数据增强及多尺度特征融合模块[13]，构建了兼具高分类精度、强泛化能力与轻量化特性的 DR 诊断模型。实验结果表明，CF-ResNet 在 Kaggle 测试集上的准确率达 90.3%、召回率为 90.6%、特异性为 92.3%、F1 分数为 90.4%、平均 AUC 为 0.825，各项指标均优于原始 ResNet50 及主流模型，消融实验验证了各改进模块的有效性。通过 Grad-CAM 热力图可视化、混淆矩阵及 ROC 曲线分析，进一步验证了模型的特征聚焦能力与分级区分能力。同时，模型单张图像推理耗时仅 0.06 秒，参数量控制在 25.8M，兼顾了实时性与轻量化需求。

该模型为糖尿病视网膜病变的自动诊断提供了新的有效方案，能够显著提升筛查效率，缓解医疗资源压力，尤其适用于基层医疗场景的大规模早筛工作，具备重要的临床应用价值与推广前景。未来我们可通过优化微小病灶特征提取、引入域自适应技术、融合语义分割功能等方向，进一步提升模型性能与临床实用性，推动 DR 智能筛查技术的规模化临床应用。

参考文献

- [1] Ebrahimi, B., Le, D., Abtahi, M., Dadzie, A.K., Lim, J.I. and Yao, X. (2023) OCTA Layer Information Fusion for Deep Learning Classification of Diabetic Retinopathy. *Investigative Ophthalmology & Visual Science*, **64**, 275-275.
- [2] Zong, X., Liang, B., Qin, Y., Ding, X. and Wang, W. (2026) Weakly Supervised Object Detection Network for Diabetic Retinopathy. *Medical Physics*, **53**, e70264. <https://doi.org/10.1002/mp.70264>
- [3] Raja, D.S.S., Kumarganesh, S., Sagayam, K.M. and Dang, H. (2026) Diabetic Retinopathy Detection and Grading System Using Deep Learning Approach. *Digital Health*, **12**, Article 20552076251410982. <https://doi.org/10.1177/20552076251410982>
- [4] Kirubakaran, M. and Vijayarajan, V. (2026) Wavemem-Shapnet: A Transparent Deep Learning Approach to Early Diagnosis of Diabetic Retinopathy. *SN Computer Science*, **7**, Article No. 73. <https://doi.org/10.1007/s42979-025-04632-8>
- [5] Xia, Z., Xu, J., Tan, J., Gu, K., Shen, Y. and Li, W. (2026) Dvdrvit: Dual-View Diabetic Retinopathy Grading Based on

- Vit Interactive Attention Network. *Physica Scripta*, **101**, Article 025001. <https://doi.org/10.1088/1402-4896/ac325b>
- [6] Khurshid, M., Chiranjeev, C., Singh, R. and Vatsa, M. (2026) Classifying Retinal Images via Vascular-Optic Disc Cross-Segmentation and Attentive Feature Selection. *Scientific Reports*, **16**, Article No. 2398. <https://doi.org/10.1038/s41598-025-89666-x>
- [7] Chitradevi, B., Mathiyalagan, P., Ramachandran, A., Dhanapal, R., Sheikdavood, K. and Gnanamurugan, S. (2026) Conv-Vit: An Improved Discrete Convolution-Based Vision Transformer for Diabetic Retinopathy Detection. *Franklin Open*, **14**, Article 100477. <https://doi.org/10.1016/j.fraope.2025.100477>
- [8] Ahmad, I., Singh, V.P. and Gore, M.M. (2026) NGCF-RVFL: Next Generation Convolutional Feature with Random Vector Functional Link for Multi-Grade Diabetic Retinopathy Detection. *Computers and Electrical Engineering*, **131**, Article 110972. <https://doi.org/10.1016/j.compeleceng.2026.110972>
- [9] Kumar, N.A., Madhusudan, D., Ioannou, I., Ghantasala, G.S.P. and Vassiliou, V. (2026) A Self-Supervised Hybrid CNN with Uncertainty-Aware Referral for Diabetic Retinopathy Screening. *Biomedical Signal Processing and Control*, **116**, Article 109482. <https://doi.org/10.1016/j.bspc.2026.109482>
- [10] Kamal, E.S. and Sharmin, N. (2026) Retinal Vessel Segmentation Using a Swin Transformer-Based Encoder-Decoder Architecture. *Signal, Image and Video Processing*, **20**, Article No. 27. <https://doi.org/10.1007/s11760-025-05089-1>
- [11] Li, T., Gao, Y., Wang, K., Guo, S., Liu, H. and Kang, H. (2019) Diagnostic Assessment of Deep Learning Algorithms for Diabetic Retinopathy Screening. *Information Sciences*, **501**, 511-522. <https://doi.org/10.1016/j.ins.2019.06.011>
- [12] Yu, G., Cao, C., Shu, X. and Yao, L. (2026) Association between Vitamin D Deficiency and the Risk of Diabetic Retinopathy in Patients with Type 2 Diabetes: A Meta-Analysis. *Molecular Genetics & Genomic Medicine*, **14**, e70157. <https://doi.org/10.1002/mgg3.70157>
- [13] Zhang, J., Hao, J., Chang, D., Zhao, M. and Chen, M. (2026) Associations between Diabetic Retinopathy and Disease Severity of Diabetic Nephropathy in Patients with Type 2 Diabetes. *Journal of Diabetes and its Complications*, **40**, Article 109256. <https://doi.org/10.1016/j.jdiacomp.2026.109256>