

糖尿病性视网膜病变诊断项目报告



小组成员：2250944 谌乐俊杰、2250758 林继申、2251941 王成威

一、项目背景与目标

二、关键技术调研（学习笔记）

三、模型训练与优化

四、AI Agent 搭建

五、病变诊断智能平台

六、项目总结

七、参考文献

一、项目背景与目标

1.1 项目背景

糖尿病性视网膜病变（Diabetic Retinopathy, DR）是导致全球视力丧失的主要眼疾之一，其早期往往缺乏明显症状，易被忽视，直至病情恶化才被发现。因此，定期开展眼底筛查对 DR 的早期诊断和及时干预至关重要。然而，传统诊断依赖眼科医生通过人工分析眼底图像来识别视网膜病变和血管异常，这种方式不仅耗时耗力，还受限于医生的经验和主观判断，难以满足大规模筛查的需求。

近年来，人工智能技术特别是深度学习在医学影像分析中的应用，为 DR 的自动检测提供了新的解决方案。基于深度学习的图像分割和分类技术能够高效处理眼底图像，精准识别 DR 相关病灶，如微动脉瘤、出血点和渗出物等，从而显著提升诊断效率和准确性。尽管如此，DR 自动检测仍面临诸多挑战：如何在有限且异质性高的眼底图像数据集上实现精准病灶定位，如何平衡模型性能与计算资源的需求，如何在保护患者隐私的前提下实现高效诊断，以及如何将技术无缝应用于临床实际场景。

针对这些问题，本研究将聚焦于深度学习在 DR 检测中的应用，分析不同数据集上的语义分割与分类方法的性能，探索其技术特点与适用场景，旨在开发一种高效、准确且通用的 DR 智能诊断系统。该系统将通过优化算法和模型设计，辅助眼科医生实现快速筛查与精准诊断，从而推动 DR 的早期发现与管理水平的提升。

1.2 项目目标

本项目目标为开发一个高精度、可扩展的 AI 辅助诊断系统，实现糖尿病性视网膜病变（DR）的自动化筛查（微动脉瘤 MA、出血点 HE、硬性渗出 EX、软性渗出 SE）。系统要求检测算法高性能、高实时性。

1.3 项目创新点

- 1. 算法对比与选择：**对比不同图像分割算法，选择最优模型，并基于模型实现变体，为系统提供最准确的诊断结果。
- 2. 治疗方案生成：**接入 AI Agent，根据病灶严重程度分级评估与医生临床诊断，AI 辅助生成个性化的诊断意见与治疗方案，提高治疗的科学性和针对性。
- 3. 诊断分析报告：**平台支持一键生成 DR 诊断分析报告，内含 AI 辅助诊断意见，并提供历史报告查询功能，支持随时随地查看报告。
- 4. 模型部署与评估：**统计模型参数量，优化模型性能，评估模型在生产环境下的效率，实现实时 DR 筛查，支持系统高性能、高实时性要求。

二、关键技术调研（学习笔记）

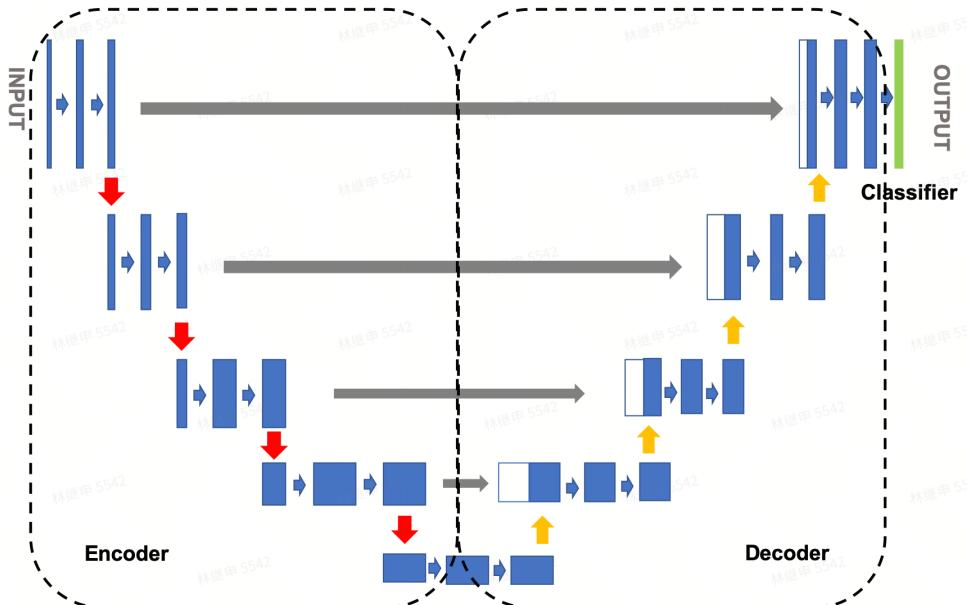
2.1 基于深度学习的语义分割任务方法调研

2.1.1 8 种语义分割任务方法

1. UNet

UNet 是一种经典的编码器-解码器结构，最初设计用于医学图像分割。其核心特点是跳跃连接（Skip Connections），能够将编码器中的特征信息直接传递到解码器，帮助恢复细节信息。在裂缝数据集和火焰数据集上，UNet 的表现非常出色，精确度和整体精度均接近 99.8%，平均交并比达到 80.25%。在 Cityscapes 数据集上，UNet 的表现相对较弱，mIoU 为 33.78%，表明其在复杂场景中的分割能力有限。

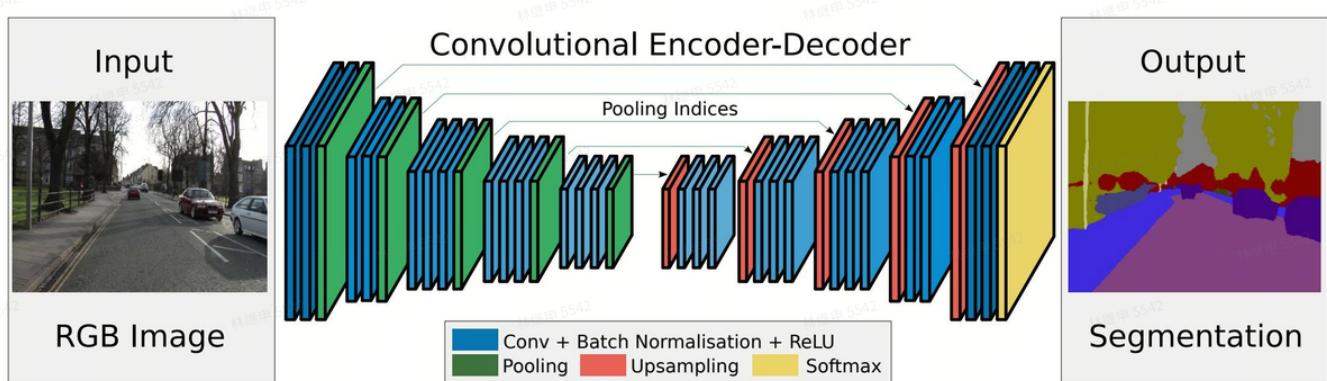
1. 在医学图像分割等任务中表现出色，尤其是在处理小物体和复杂背景时。
2. 结构相对简单，易于训练，且在有限的数据集上也能表现良好。
3. 跳跃连接有助于捕获多尺度的特征信息，提升分割精度。



2. SegNet

SegNet 也是一种编码器-解码器结构，但其独特之处在于使用了池化索引层，能够在解码器中精确还原编码器阶段的最大池化操作，从而恢复细节信息。在裂缝数据集和火焰数据集上，SegNet 的表现略低于 UNet，精确度和整体精度分别为 99.74% 和 99.64%，mIoU 为 80.16%。在 Cityscapes 数据集上，SegNet 的表现与 UNet 相近，mIoU 为 33.70%，表明其在复杂场景中的分割能力有限。

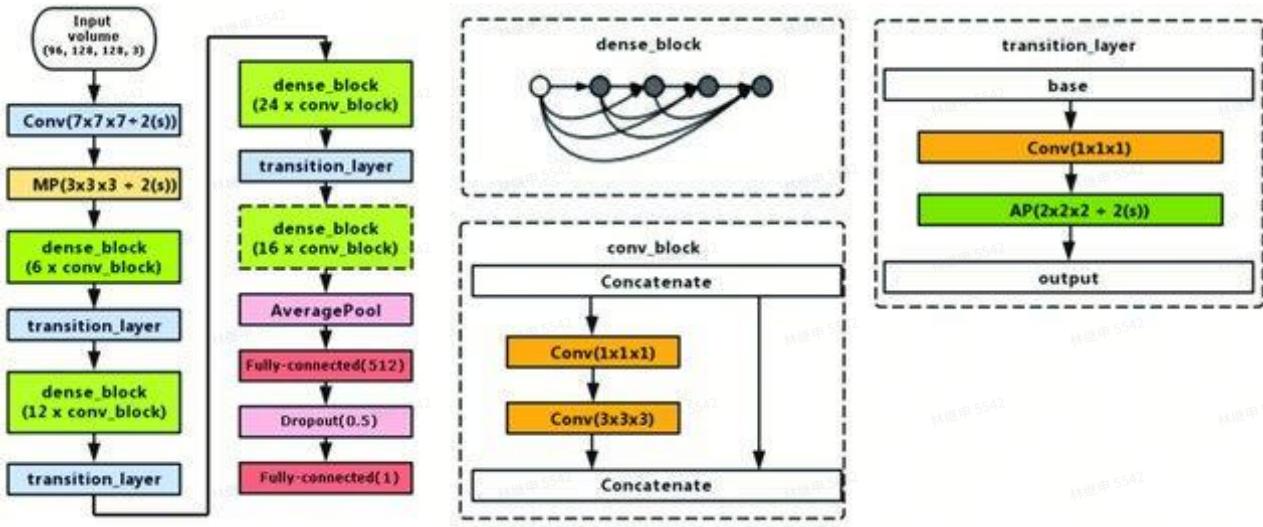
1. 轻量级设计，适合实时应用和嵌入式系统。
2. 在自动驾驶等需要实时处理的场景中表现良好。



3. DenseASPP

DenseASPP 结合了 DenseNet 和 ASPP（空洞空间金字塔池化）的优点，通过密集连接和不同尺度的空洞卷积来捕获多尺度的上下文信息。在裂缝数据集和火焰数据集上，DenseASPP 的表现最佳，精确度和整体精度分别为 99.82% 和 99.71%，mIoU 达到 82.41%。在 Cityscapes 数据集上，DenseASPP 的表现也优于 UNet 和 SegNet，mIoU 为 35.43%，表明其在复杂场景中的分割能力更强。

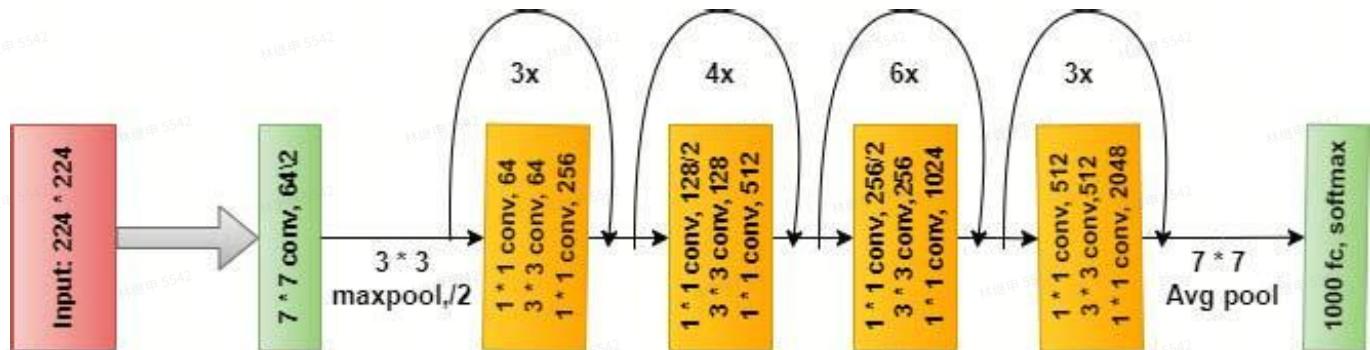
1. 能够更好地捕获多尺度的上下文信息，适合处理复杂的语义分割任务。
2. 在多个数据集上表现出色，尤其是在需要高精度的任务中。



4. ResNet50

ResNet50 引入了残差连接（Residual Connection），解决了深层网络中的梯度消失问题。残差连接允许信息直接从一层跳过多层传递到后面的层，使得网络可以训练得更深。

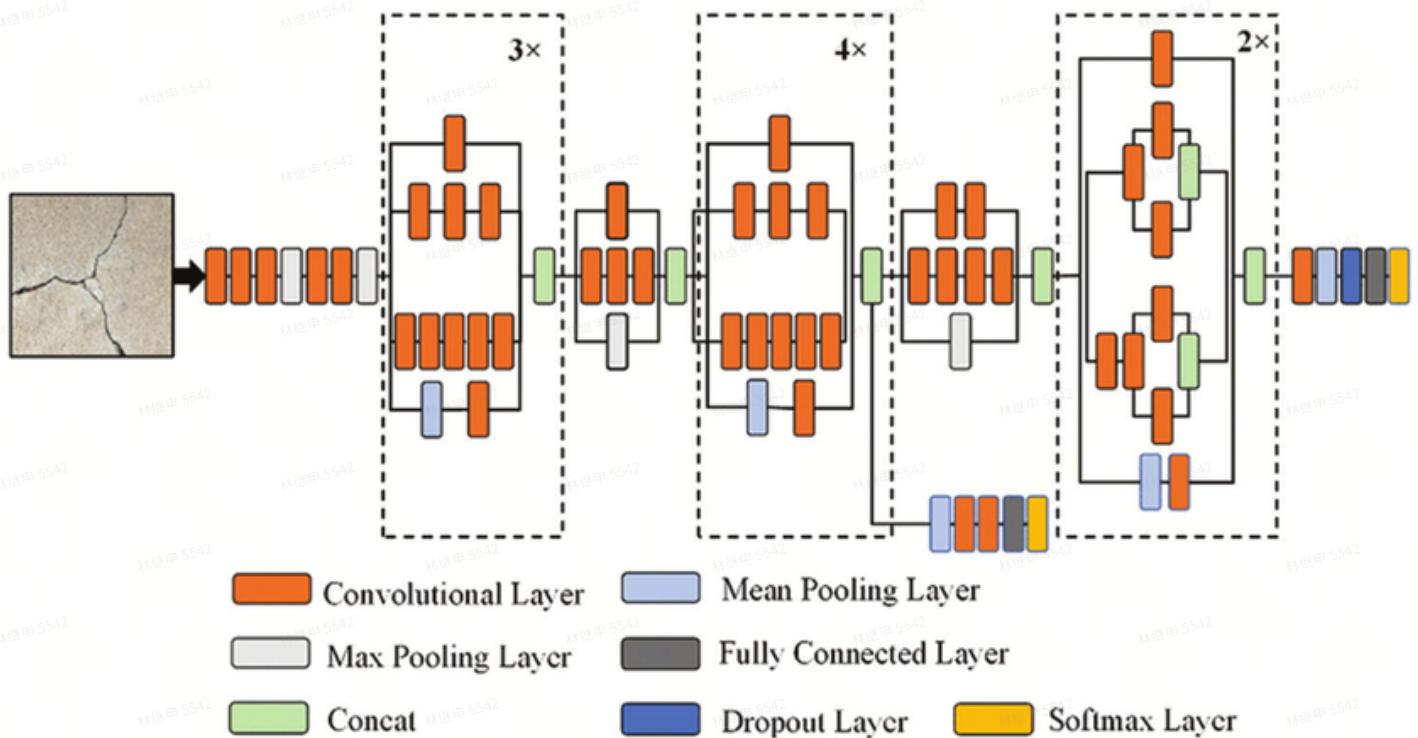
1. 残差连接使得网络可以训练得非常深（50 层），同时保持训练的稳定性。
2. 在 DR 检测中表现优异。



5. Inception-V3

Inception-V3 通过“分解卷积”和“不对称卷积”来减少计算量，同时保持模型的表达能力。分解卷积将大卷积核（如 5×5 ）分解为多个小卷积核（如 3×3 和 1×1 ），减少了参数数量。

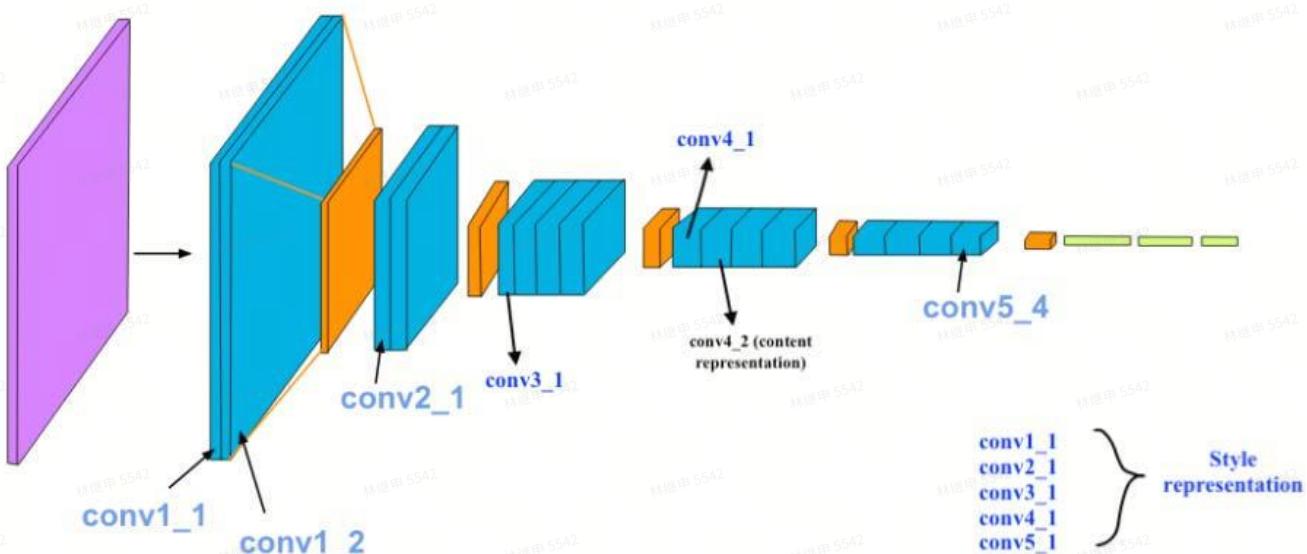
- 计算效率高，训练速度较快。



6. VGG19

VGG19 是一个经典的 CNN 模型，特点是使用了多个 3×3 卷积层堆叠，增加了网络的深度。输入图像大小为 224×224 像素。

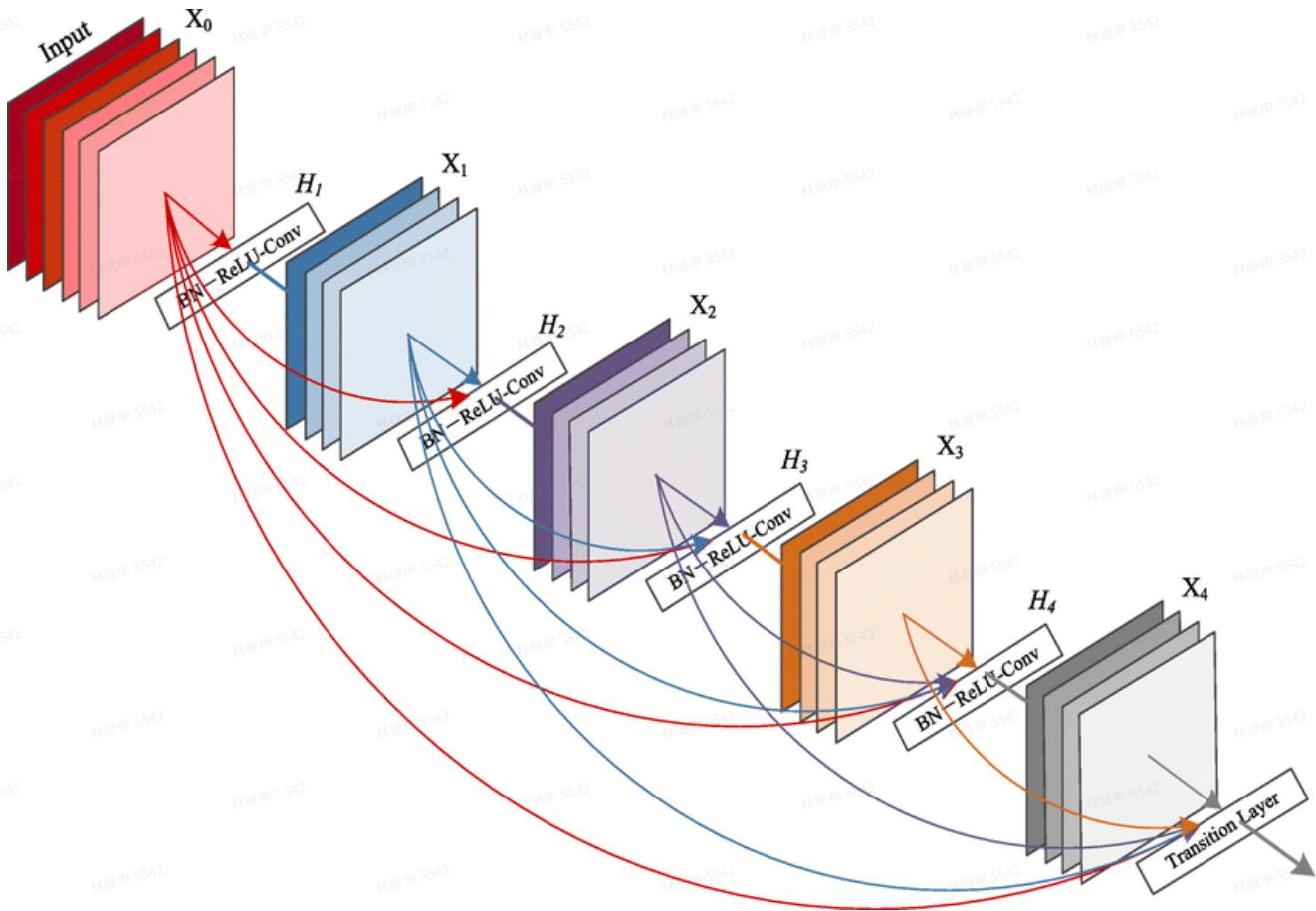
1. 计算效率高。
2. 训练速度较快。



7. DenseNet-121

DenseNet-121 引入了“密集连接”（Dense Connection），每一层都直接连接到后续的所有层。这种设计有助于缓解梯度消失问题，并且能够更有效地利用特征。

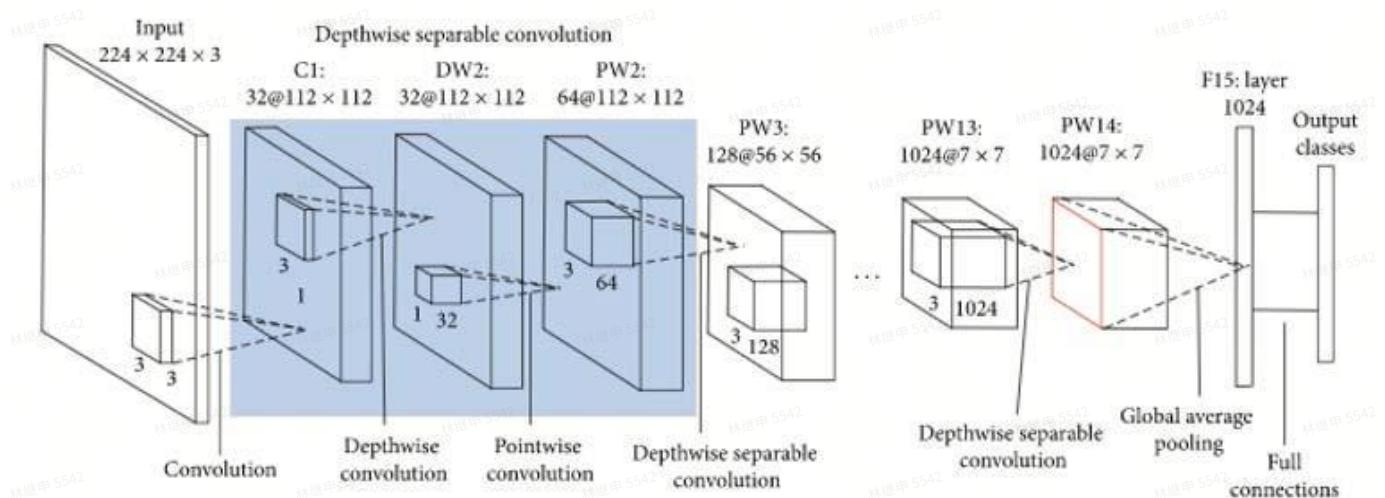
1. 结构简单，易于理解和实现。
2. 缺点是由于参数量较大，训练时间较长，准确率相对较低。



8. MobileNetV2

MobileNetV2 是一种轻量级网络，适用于移动设备和嵌入式设备。它使用了“倒置残差块”（Inverted Residual Block）和“线性瓶颈”（Linear Bottleneck）结构，减少了计算量和参数量。

1. 密集连接使得网络能够更有效地利用特征，准确率较高。
2. 缺点是训练时间较长。



2.1.2 基于深度学习的语义分割任务方法对比

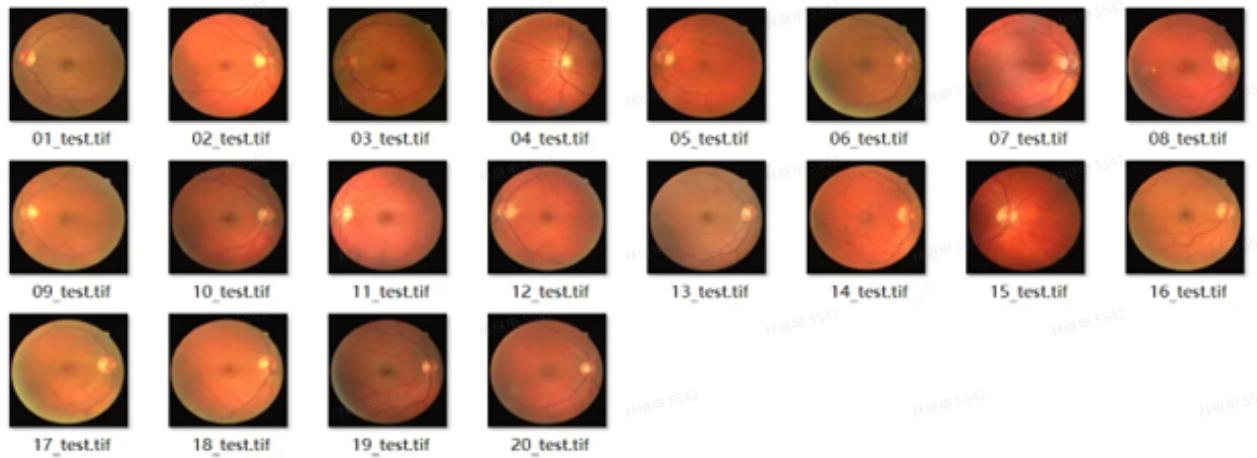
方法	核心特点	优点	缺点
UNet	编码器-解码器结构、跳跃连接 (Skip Connections)	医学图像分割表现优异、小数据集友好、多尺度特征捕获能力强	复杂场景表现较弱
SegNet	编码器-解码器结构、池化索引层 (Pooling Indices) 还原细节	轻量级设计，适合实时应用、边缘细节恢复较好	复杂场景 / 多类别任务表现较差、需更多训练数据
DenseASPP	结合 DenseNet 和 ASPP 多尺度空洞卷积捕获上下文	复杂场景鲁棒性强、多尺度特征融合最佳 (mIoU 最高)	计算资源需求高、训练时间长
ResNet50	残差连接 (Residual Connection) 解决梯度消失	深层网络稳定、DR 检测表现优异 (损失较低)	参数量较大
Inception-V3	分解卷积 (如 $5 \times 5 \rightarrow 3 \times 3 + 1 \times 1$) 降低计算量	计算效率高、训练速度快	准确率略低于 ResNet50
VGG19	堆叠 3×3 卷积层、结构简单	易于实现和理解	参数量大、训练慢、准确率较低
DenseNet-121	密集连接 (Dense Connection)、复用特征	特征利用率高、准确率优于 VGG19	训练时间长、内存消耗高
MobileNetV2	倒置残差块 (Inverted Residual) + 线性瓶颈 (Linear Bottleneck) 轻量化设计	计算效率极高、适合移动端 / 嵌入式设备	准确率最低 (牺牲精度换速度)

2.2 5 种糖尿病性视网膜病变眼底图像数据集调研

2.2.1 5 种眼底图像数据集

1. DRIVE (Digital Retinal Images for Vessel Extraction)

DRIVE 数据集主要用于视网膜血管分割的研究。该数据集包含 40 张彩色眼底图像，分为训练集和测试集，每集各 20 张图像。每张图像的分辨率为 768×584 像素，8 位彩色通道。数据集提供了手动标注的血管分割掩码，用于评估血管分割算法的性能。



- **特点：**

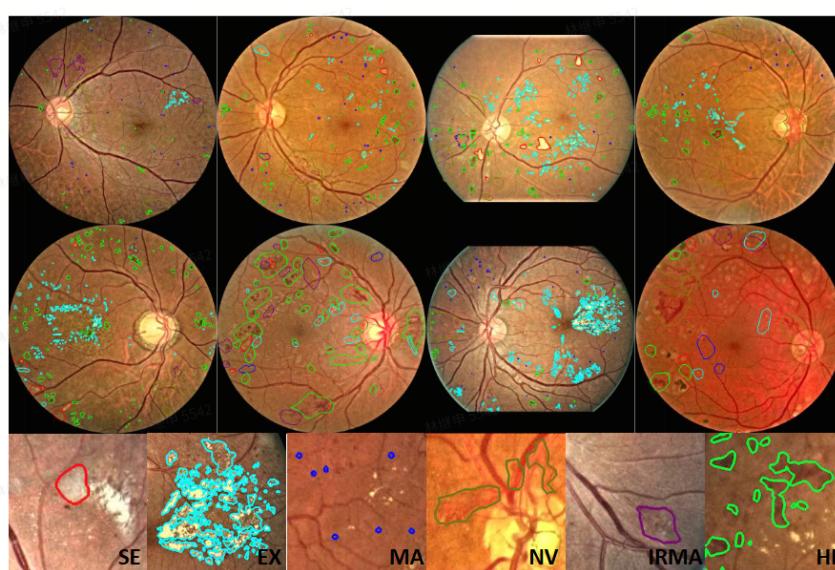
- 图像数量较少，但标注质量高。
- 主要用于血管分割任务。
- 图像分辨率较低，适用于早期的图像处理算法。

- **应用：**

- 主要用于视网膜血管分割算法的开发和评估。

2. FGADR (Fine-Grained Annotated Diabetic Retinopathy Dataset)

FGADR 数据集是一个细粒度标注的糖尿病视网膜病变 (DR) 数据集，包含 2,842 张眼底图像。其中 1,842 张图像具有像素级的 DR 相关病变标注，1,000 张图像具有图像级的 DR 分级标签。病变标注包括微血管瘤 (MA)、出血点 (HE)、硬性渗出 (EX)、软性渗出 (SE)、视网膜内微血管异常 (IRMA) 和新生血管 (NV)。



- **特点：**

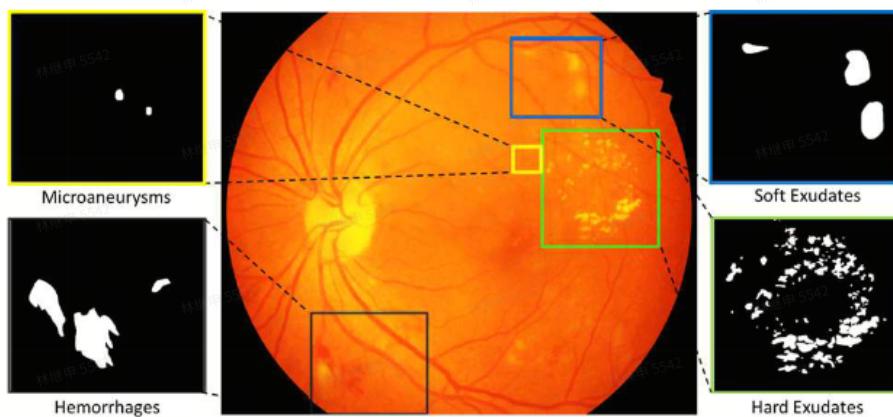
- 数据集规模较大，标注精细。
- 提供了像素级的病变标注，适合病变分割任务。
- 图像级的分级标签可用于 DR 分级任务。

- **应用：**

- 用于 DR 病变分割、DR 分级以及跨领域的迁移学习任务。

3. IDRiD (Indian Diabetic Retinopathy Image Dataset)

IDRiD 数据集包含 516 张眼底图像，分为训练集和测试集。其中 81 张图像具有像素级的病变标注，包括微血管瘤 (MA)、出血点 (HE)、硬性渗出 (EX) 和软性渗出 (SE)。数据集还提供了 DR 和糖尿病黄斑水肿 (DME) 的分级标签。



- **特点：**

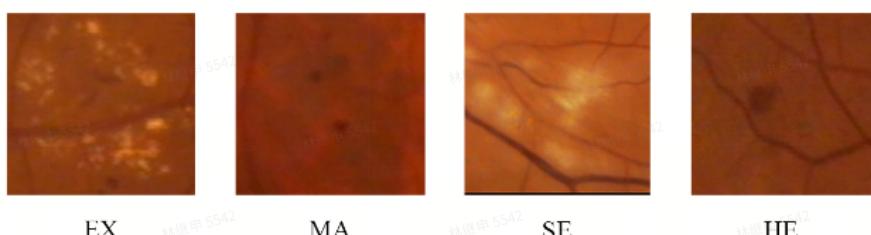
- 数据集规模较小，但标注质量高。
- 提供了像素级的病变标注和图像级的分级标签。
- 适用于 DR 病变分割和分级任务。

- **应用：**

- 用于 DR 病变分割、DR 分级以及糖尿病黄斑水肿的研究。

4. TJDR (Tongji Diabetic Retinopathy Dataset)

TJDR 数据集包含 561 张彩色眼底图像，来自同济医院。图像分辨率较高，部分图像分辨率为 $2,048 \times 2,048$ 像素，部分为 $3,912 \times 3,912$ 像素。数据集提供了像素级的病变标注，包括硬性渗出 (EX)、出血点 (HE)、微血管瘤 (MA) 和软性渗出 (SE)。



- **特点：**

- 数据集涵盖了多种 DR 病变类型，适合多病变分割任务。
- 数据集规模适中，适合中等规模的研究和算法验证。

- 图像分辨率较高，适合高分辨率图像的算法研究。

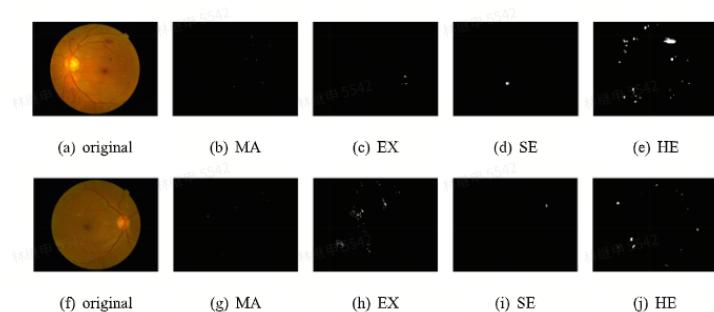
- 提供了像素级的病变标注，标注质量高。

- **应用：**

- 用于 DR 病变分割算法的开发和评估。

5. DDR (Diabetic Retinopathy Dataset from Nankai University)

DDR 数据集包含 13,673 张眼底图像，来自 9,598 名患者。其中 757 张图像具有像素级的病变标注，包括微血管瘤 (MA)、出血点 (HE)、硬性渗出 (EX) 和软性渗出 (SE)。数据集还提供了图像级 DR 分级标签，分为健康、非增殖性轻度、中度、重度和增殖性 DR 五个等级。



The number of samples of each class in the training, validation and testing sets.

	Training	Validation	Testing	Total
No DR	3133	1253	1880	6266
Mild	315	126	189	630
Moderate	2238	895	1344	4477
Severe	118	47	71	236
Proliferative	456	182	275	913
Ungradable	575	230	346	1151
Total	6835	2733	4105	13,673

- **特点：**

- 数据集规模较大，覆盖了多种 DR 严重程度分级。
- 提供了像素级的病变标注和图像级的分级标签。
- 适用于 DR 病变分割、分级和检测任务。

- **应用：**

- 用于 DR 病变分割、DR 分级以及病变检测任务。

2.2.2 糖尿病性视网膜病变眼底图像数据集对比

数据集	规模	分辨率	标注类型	病灶类型	应用场景
DRIVE	40	768×584	像素级（血管）	血管	血管分割
FGADR	2,842 (1842)	多种	像素级（病变） 图像级（分级）	MA, HE, EX, SE, IRMA, NV	病变分割、DR 分级、迁移学习
IDRiD	516 (81)	4288×28 48	像素级（病变）	MA, HE, EX, SE	病变分割、DR 分级

			图像级（分 级）		
DDR	13,673 (757)	多种	像素级（病 变） 图像级（分 级）	MA, HE, EX, SE	病变分割、DR 分级、病变 检测
TJDR	561 (448)	2048×20 48 3912×39 12	像素级（病 变）	EX, HE, MA, SE	病变分割

2.2.3 总结

1. DRIVE 数据集主要用于血管分割，适合基础算法的研究。
2. FGADR 和 DDR 数据集规模较大，适合用于深度学习模型的训练和评估，尤其是 DR 病变分割和分级任务。
3. IDRiD 数据集规模较小，适合小规模的研究和算法验证。
4. TJDR 数据集图像分辨率高，适合高分辨率图像的算法研究。

2.2.4 选择 DDR 作为主要数据集的原因

1. 像素级和图像级标注结合，应用场景广泛

DDR 数据集中的 757 张图像具有像素级的病变标注，适合高精度的病变分割任务。同时，图像级的分级标签支持 DR 严重程度分类任务。这种双重标注使其在病变分割、DR 分级和病变检测任务中均有广泛应用，相比 DRIVE（仅血管分割）和 TJDR（仅像素级病变标注）更加灵活。

2. 数据量与标注质量的平衡

虽然 DDR 的像素级标注图像数量（757 张）少于 FGADR（1,842 张），但其总图像量（13,673 张）远超 FGADR，适合需要大规模数据支持的深度学习模型训练。此外，DDR 的标注质量经过严格验证，适合高精度算法开发。相比 IDRiD（仅 81 张像素级标注图像），DDR 在标注数量和数据规模上更具优势。

3. 贴近临床应用的多样性

DDR 数据集来源于 9,598 名患者，图像采集设备和分辨率多样，反映了真实世界中眼底图像的复杂性。这种多样性使其在开发适用于不同临床环境的算法时更具鲁棒性，相比 TJDR（单一来源，561 张图像）和 DRIVE（图像分辨率较低）更适合现代高分辨率成像设备的研究。

2.3 预处理方法调研

一般的计算机视觉处理都会进行图像预处理。在 APTOS 2019 竞赛的 Kaggle 社区 Notebooks 中，看到最多的一种图像增强方法是 Ben Graham 提出的。他是 APTOS 2015 年的糖尿病性视网膜病变竞赛的冠军，由于 2019 年的竞赛图像也是相似的，所以很多人都延用他的预处理方法。该预处理方法分为以下 3 步：

1. Rescale the images to have the same radius (300 pixels or 500 pixels)
2. Subtracted the local average color; the local average gets mapped to 50% gray
3. Clipped the images to 90% size to remove the “boundary effects”

2.3.1 基于眼球的 Resize： `scaleRadius()`

因为不同图片长宽比不同，且眼球外围黑边宽度各不相同，所以 resize 不能以整个图像的尺寸为准。因此，以眼球半径为基准，resize 图片。

取高度正中间一行（图片基本上下对称；垂直方向眼球可能会被裁剪，但水平方向的眼球基本都完整。），并计算 3 个通道的和，得到 x

计算 x 大于 $(x \text{ 均值}/10)$ 的像素个数，作为直径，除以 2 得到半径 r 。

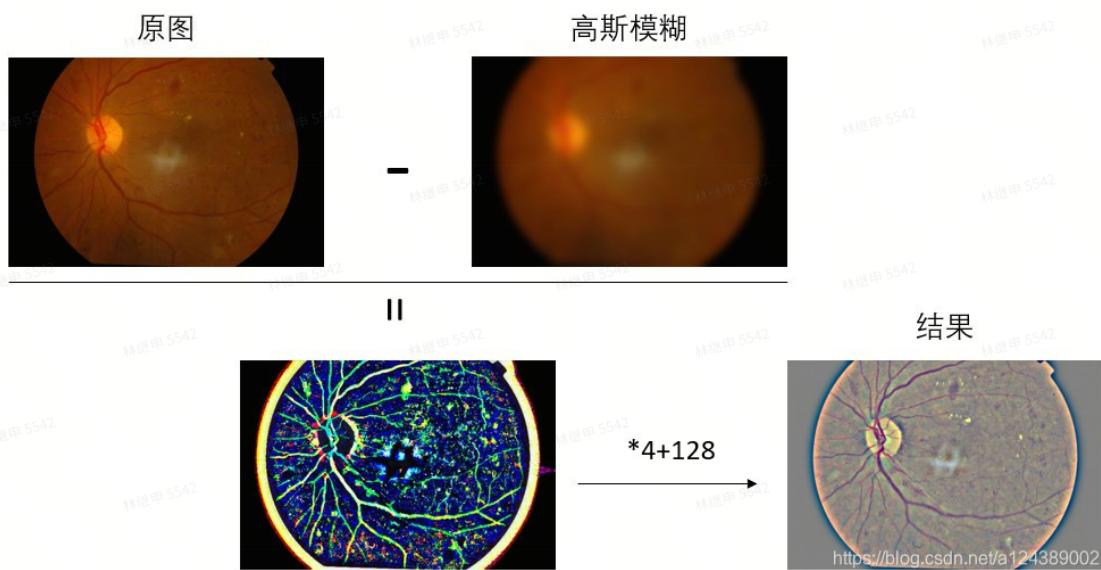
计算的到一个参数 $s = \text{scale} / r$ 。 scale 是给定的参数，代表你想要的眼球半径。

计算 `cv2.resize(img,(0,0),fx=s,fy=s)`。这里把输出长宽设为 $(0, 0)$ ，函数会通过 fx 和 fy 计算输出长宽，公式为 $\text{img.width} * s, \text{img.height} * s$ 。而 $s=\text{scale}/r$ ，那么输出长宽=原长宽 $* \text{scale}/r = (\text{原长宽}/r) * \text{scale}$ 。可见眼球的半径被 $=1$ 了，再乘以 scale 。

2.3.2 特征增强

第二步是关键的一步，使得不同图像显示效果更加一致，同时凸显特征。

核心在于这一句，
`cv2.addWeighted(a,4, cv2.GaussianBlur(a,(0,0),scale/30), -4,128)`。

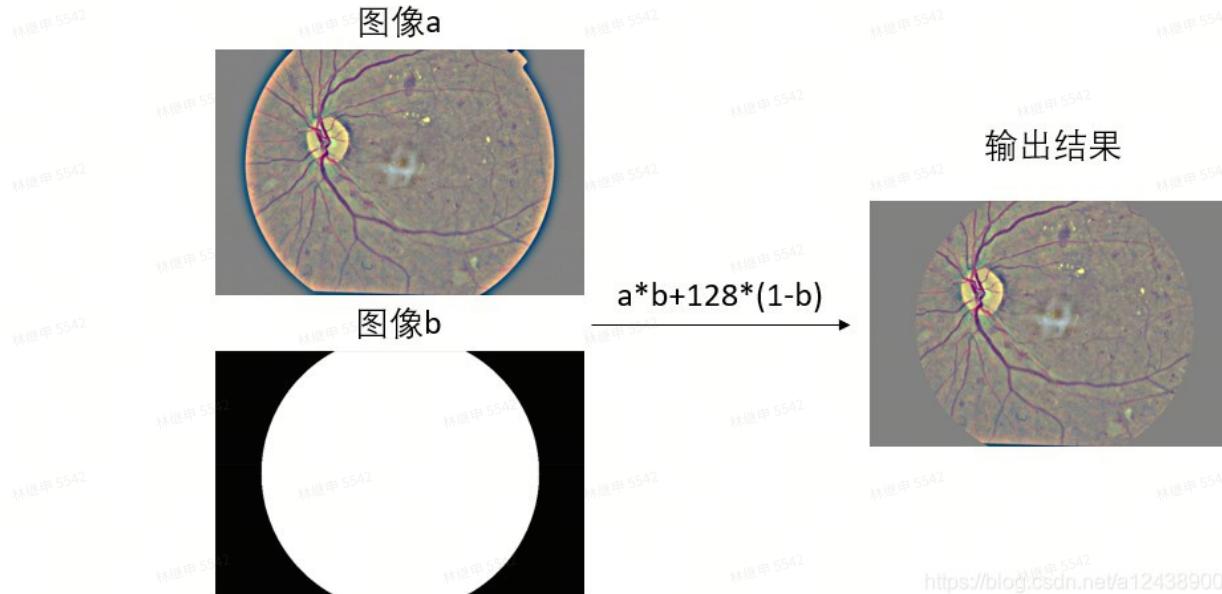


`cv2.addWeighted` 的公式 = $a4 + \text{高斯模糊结果} - 4 + 128$ 。过程中函数会做 `saturate_cast` 避免数据溢出。`cv2.GaussianBlur(a,(0,0),scale/30)` 中，`kernel` 大小设置为 $(0, 0)$ ，意

思是根据 x 方向上的标准差 scale / 30 来计算。

处理的原理是：“原图-高斯模糊图”得到差异。高斯模糊图是“Local Area Average Color”，相当于背景，相当于背景，原图 - 背景 = 特征，从而增强了图像。

2.3.3 去除眼球周围部分

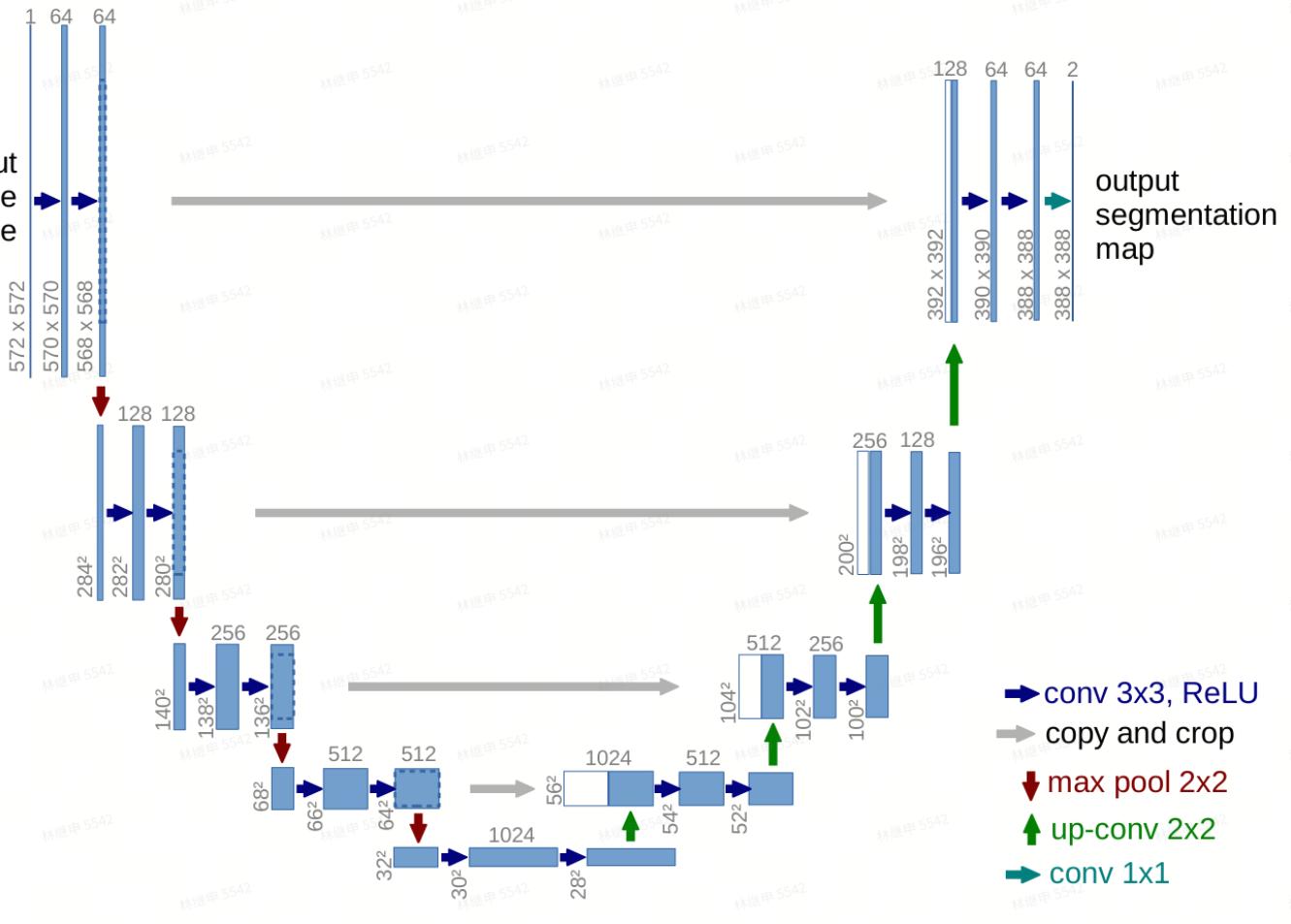


2.4 UNet 模型及其变体调研

2.4.1 UNet (2015 年, 原始架构)

核心思想：

UNet 由 Ronneberger 等人于 2015 年提出，专为生物医学图像分割设计。其核心是一个对称的编码器-解码器结构，通过 **跳跃连接 (Skip Connections)** 将编码器的高分辨率特征与解码器的语义特征融合，兼顾局部定位与全局上下文。



关键特点：

- 编码器（下采样路径）：**由重复的 3×3 卷积（无填充）+ ReLU + 2×2 最大池化组成，逐步提取高层语义特征。
- 解码器（上采样路径）：**通过上卷积（转置卷积）恢复分辨率，并与编码器对应层的高分辨率特征拼接，结合低层细节。
- 数据增强：**使用弹性变形等增强策略，解决医学图像标注数据少的问题。
- 高效性：** 512×512 图像分割仅需不到 1 秒（GPU）。

优势：

- 在少量标注数据下表现优异（如 ISBI 神经元分割挑战赛超越滑动窗口方法）。
- 支持任意大图像的分割（通过重叠-镜像策略）。

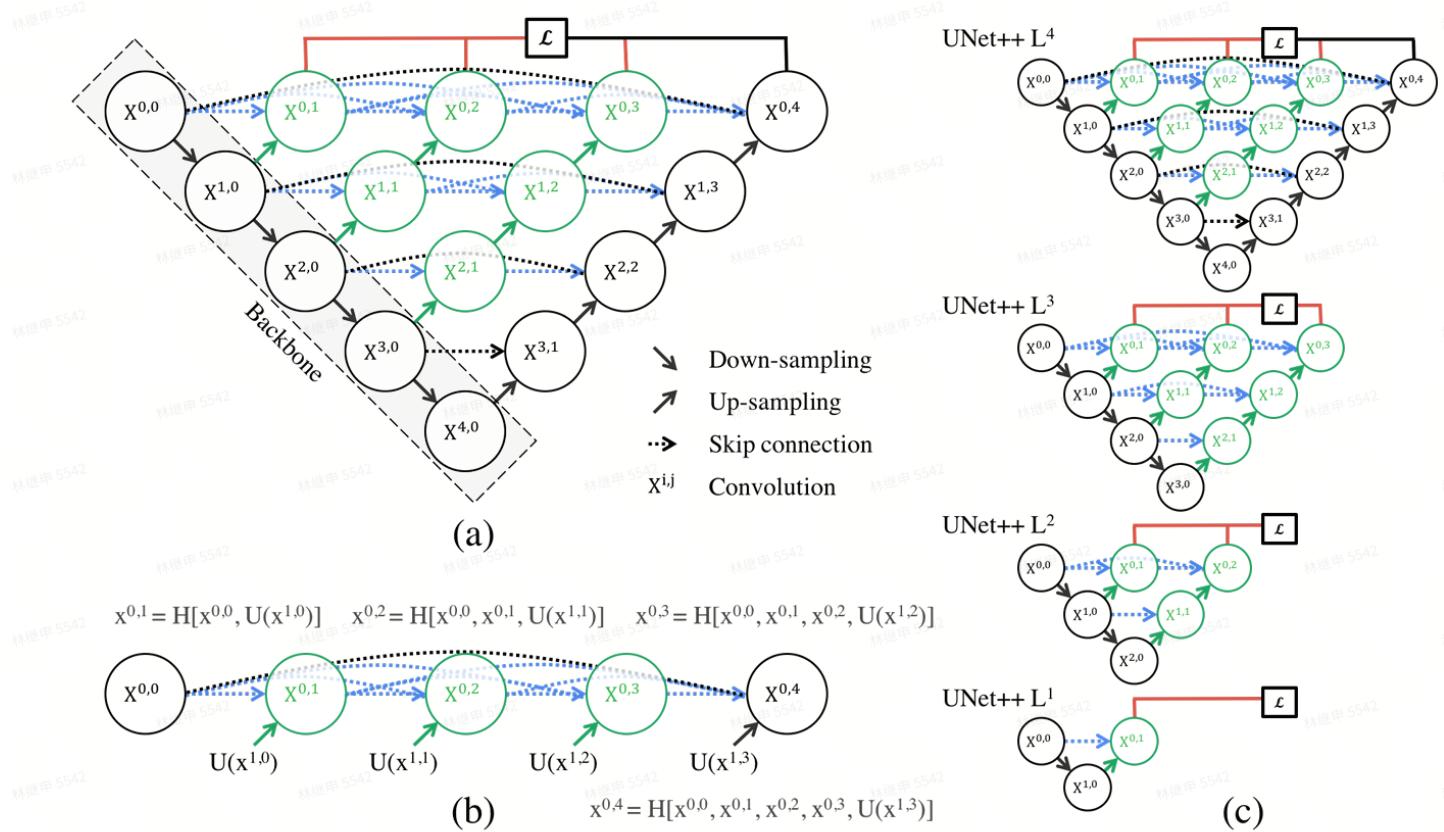
局限：

- 跳跃连接直接融合编码器和解码器特征，可能存在语义鸿沟（Semantic Gap）。
- 对小目标或密集目标的分割细节处理不足。

2.4.2 UNet++ (2018 年改进)

核心思想：

针对 UNet 的跳跃连接优化，提出 **嵌套密集跳跃路径** 和 **深度监督**，减少编码器与解码器特征的语义差异。



关键特点：

1. 嵌套密集跳跃路径：

- 在跳跃路径中引入密集卷积块（Dense Blocks），逐步提升编码器特征的语义级别，再与解码器特征融合。
- 作用：通过多级特征细化，缓解语义鸿沟（如图 a 中的绿色 / 蓝色路径）。

2. 深度监督（Deep Supervision）：

- 在多个解码层输出端添加辅助损失函数（交叉熵 + Dice 系数），提升训练稳定性。
- 支持两种推理模式：
 - 精确模式**：平均所有分支输出。
 - 快速模式**：选择单一分支输出（可剪枝模型，加速推理）。

实验结果：

- 在 4 个医学数据集（细胞核、结肠息肉、肝脏、肺结节）上，IoU 平均提升 3.9% (vs UNet) 和 3.4% (vs Wide UNet)。
- 剪枝后（如 UNet++ L3）推理速度提升 32.2%，精度仅下降 0.6%。

优势：

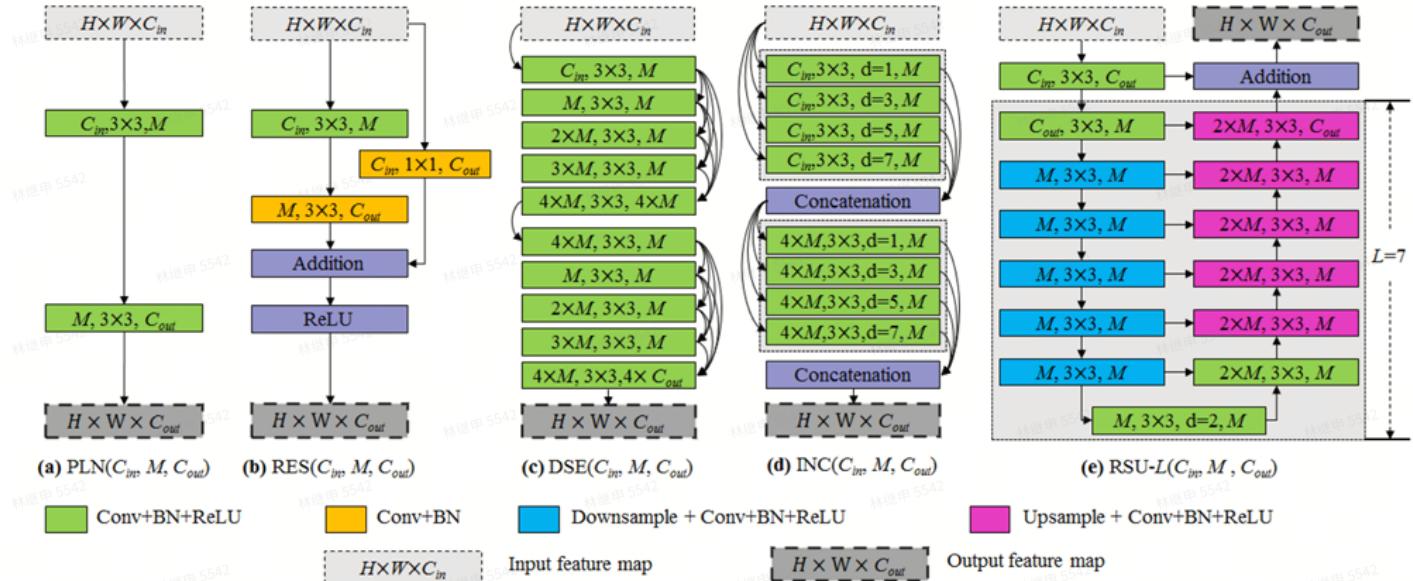
- 更精准的细节分割（如息肉边缘、细胞边界）。

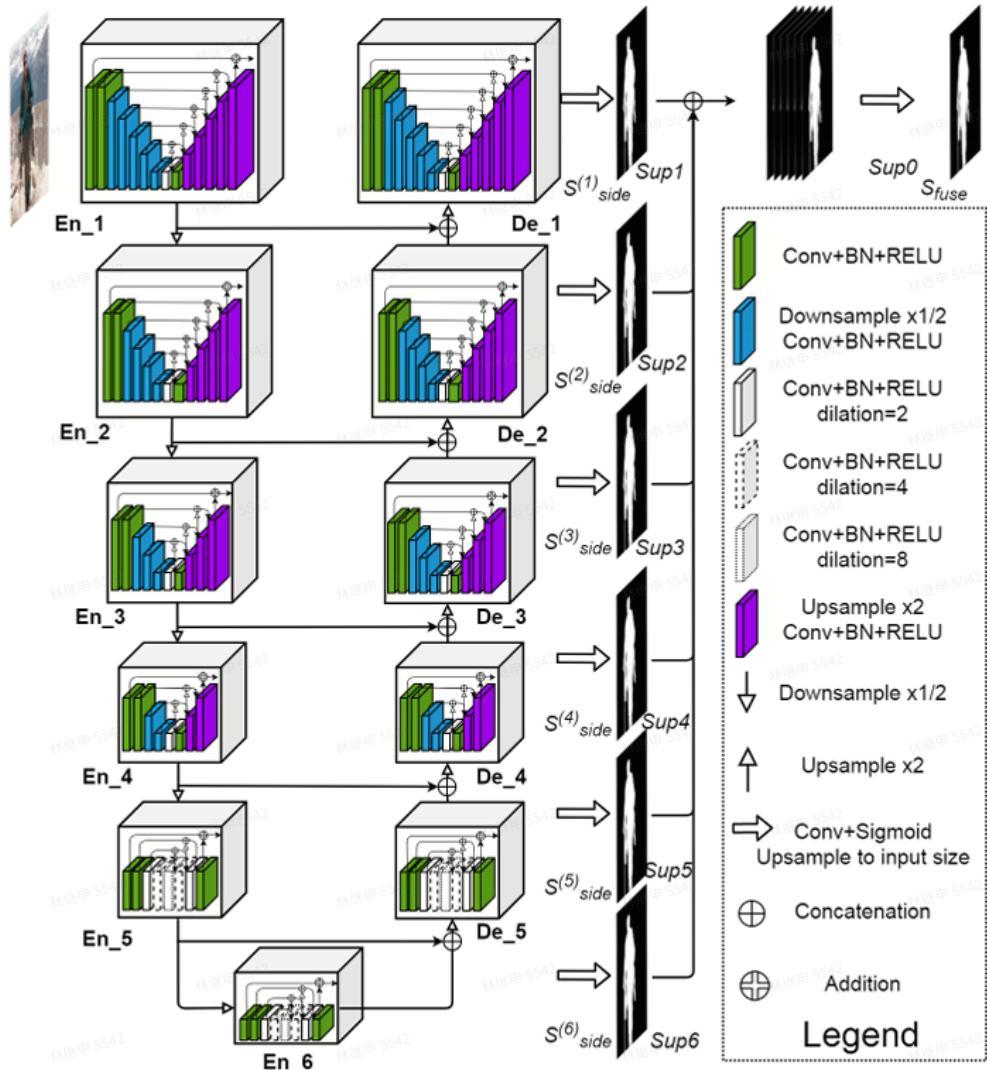
- 灵活性高，可平衡速度与精度。

2.4.3 U²Net (2021 年改进)

核心思想：

U²Net 是一种用于 **显著目标检测 (Salient Object Detection, SOD)** 的深度网络架构，旨在从自然图像中分割出最吸引视觉注意的目标区域。其设计目标是 **避免依赖预训练的主干网络 (如 VGG、ResNet)**，直接从零训练，同时保持高分辨率特征和多尺度上下文信息。





关键特点：

1. 嵌套 U 型结构 (Two-Level Nested U-Structure) :

- 顶层：类似 UNet 的编码器-解码器架构，但每个阶段由 **RSU (ReSidual U-block)** 构成。
- 底层：RSU 块内部包含 **U 型子网络**，通过池化和上采样操作捕获多尺度特征，而不会显著增加计算成本。
- 优势：
 - 通过嵌套结构，既能提取局部细节（浅层），又能建模全局上下文（深层）。
 - 相比传统 UNet，RSU 块在保持高分辨率的同时，增加了网络深度。

2. RSU (ReSidual U-block) :

- 结构（由三部分组成）：
 - **输入卷积层**（提取局部特征）
 - **U 型编码器-解码器**（捕获多尺度上下文）
 - **残差连接**（融合局部和多尺度特征）

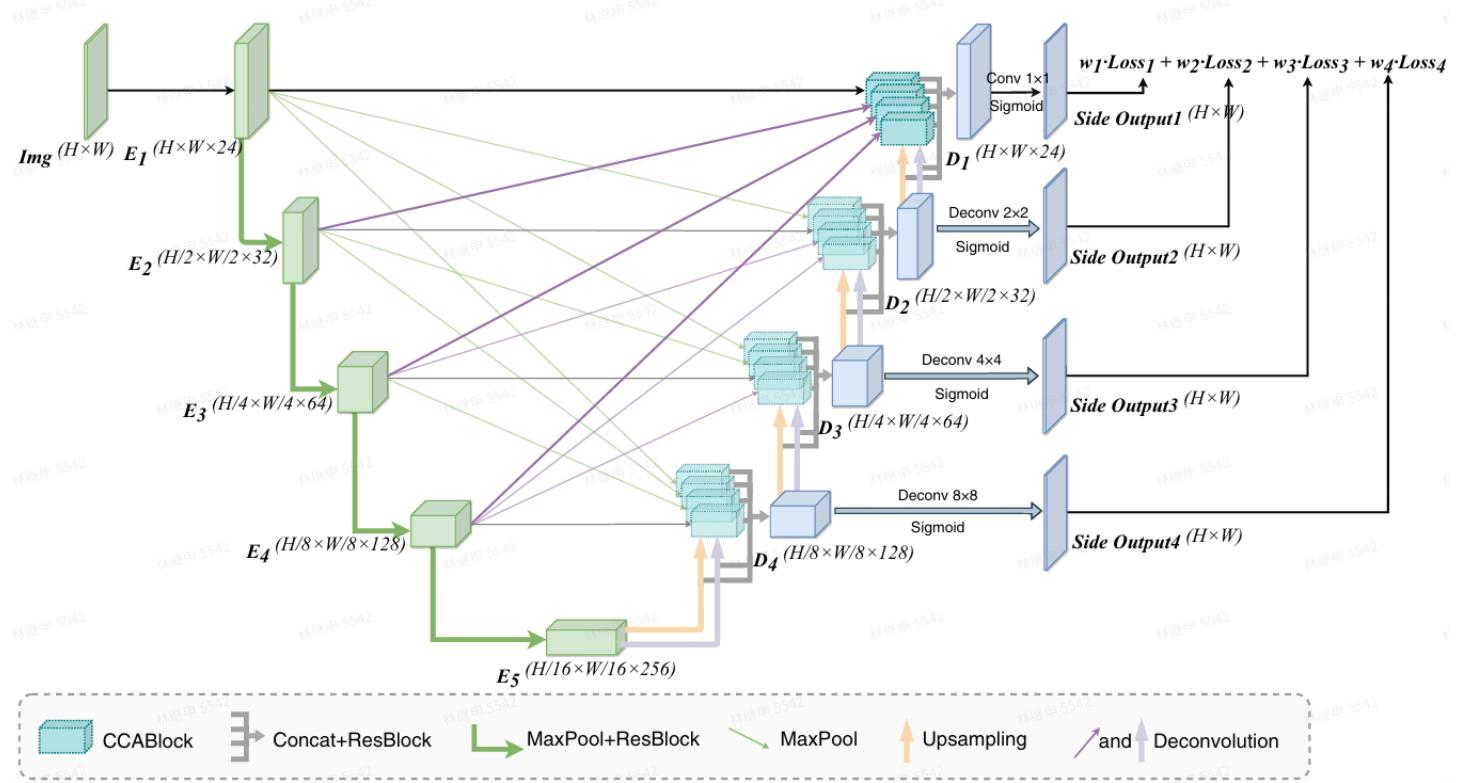
◦ 特点：

- 通过池化降低计算量，避免像 Inception 或 DenseNet 那样引入过多参数。
- 在浅层（高分辨率）和深层（低分辨率）之间灵活平衡。

2.4.4 MCA-UNet (2023 年改进)

核心思想：

MCA-UNet 是一种改进的 UNet 架构，专注于 **医学图像分割**（如 COVID-19 肺部 CT、糖尿病性视网膜病变）。其核心挑战是病灶的 **形状、大小和位置高度可变**，且边界模糊（如毛玻璃影 GGO、微动脉瘤 MA）。



关键特点：

- 多尺度交叉跳跃连接 (Multi-scale Cross Skip Connection) :**
 - 传统 UNet 的跳跃连接仅融合同尺度特征，而 MCA-UNet 通过 **密集连接** 将 **所有编码器阶段的特征** 传递到每个解码器阶段，增强多尺度信息融合。
- 交叉协同注意力 (Cross Co-Attention, CCA) :**
 - 双路径注意力：** 同时计算 **通道注意力** 和 **空间注意力**，动态调整特征重要性。
 - 跨尺度交互：** 解码器特征（上采样和反卷积结果）作为查询（Query），编码器特征作为键值（Key-Value），通过注意力机制消除语义鸿沟。
- 双向解码器 (Bi-decoder) :**
 - 双上采样：** 结合最近邻上采样和反卷积，保留更多空间细节。
 - 深度监督：** 在解码器的每个阶段添加辅助损失，加速训练并提升小目标分割能力。
- 动态多任务权重 (BMTD) :**

- 通过自适应调整不同解码器阶段的损失权重，解决多尺度训练不平衡问题。

三、模型训练与优化

3.1 初步探索：复现 UNet

3.1.1 数据集预处理（数据增强）

数据集预处理（数据增强）办法：

1. flip_horizontal (水平翻转)

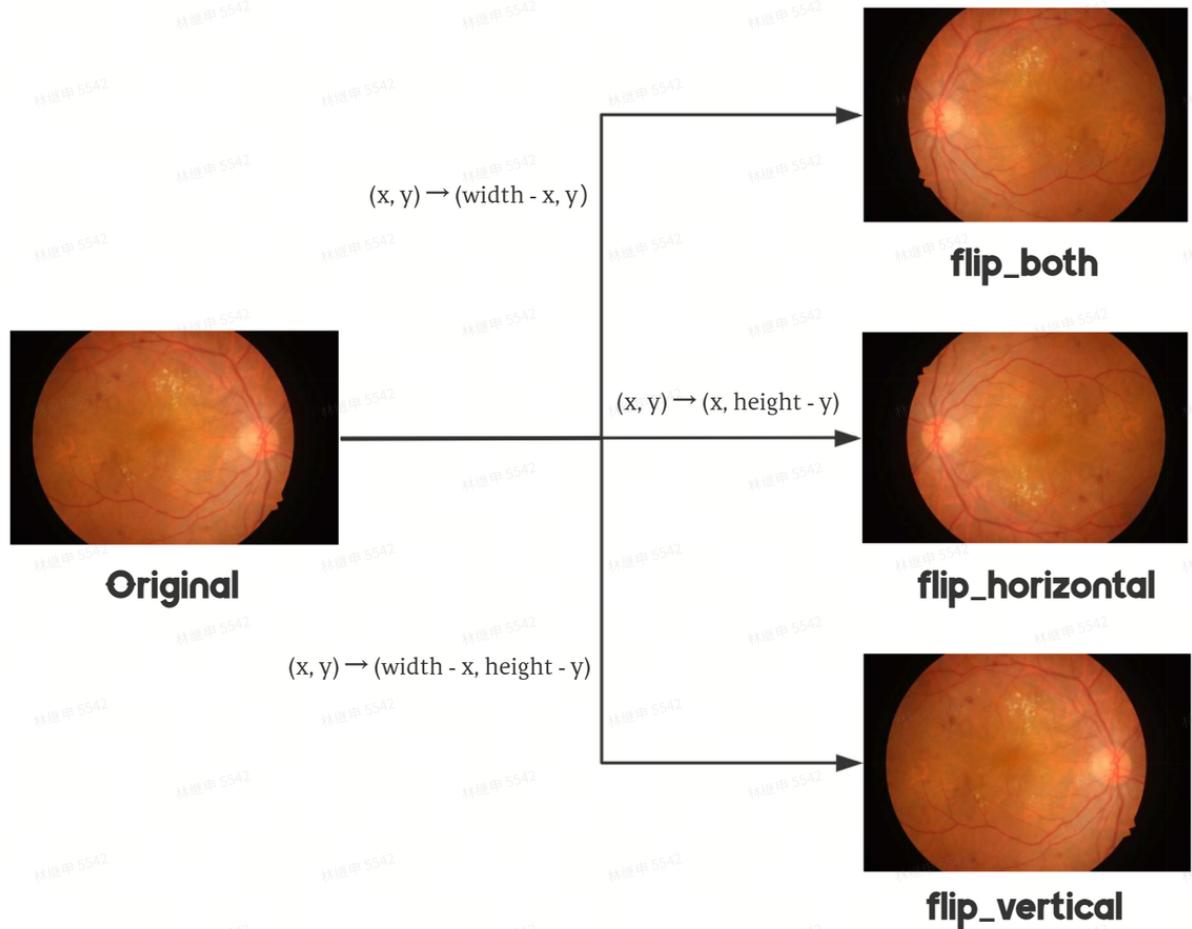
- 将图像沿 **垂直轴 (Y 轴)** 进行左右翻转。
- 效果类似于照镜子，左边和右边互换。
- 数学坐标上相当于 $(x, y) \rightarrow (width - x, y)$ 。

2. flip_vertical (垂直翻转)

- 将图像沿 **水平轴 (X 轴)** 进行上下翻转。
- 效果类似于倒置图像，顶部和底部互换。
- 数学坐标上相当于 $(x, y) \rightarrow (x, height - y)$ 。

3. flip_both (同时水平 + 垂直翻转)

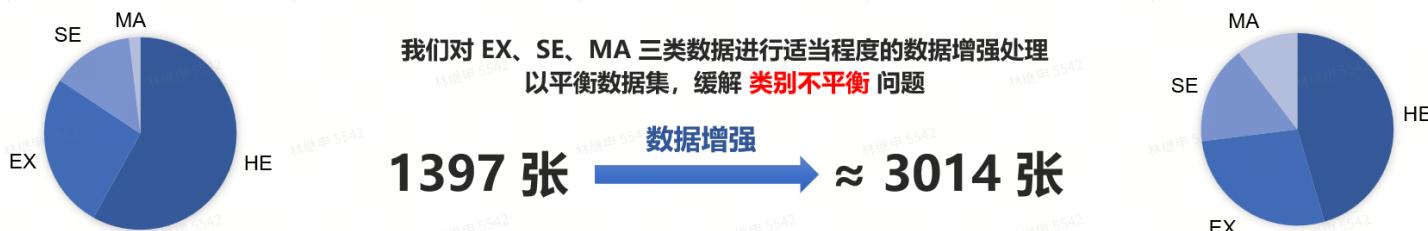
- 同时进行水平翻转和垂直翻转。
- 效果等同于将图像旋转 180 度（但注意：**旋转 180 度 ≠ 翻转 + 翻转**，因为旋转会改变像素顺序，而翻转是镜像）。
- 数学坐标上相当于 $(x, y) \rightarrow (width - x, height - y)$ 。



3.1.2 糖尿病性视网膜病变标准化数据集（3014 张）

我们对 TJDR、DDR 和 IDRiD 三个糖尿病性视网膜病变数据集进行了标准化处理和数据增强处理。

标准化处理使得来源不同的数据集能够在同一标准下使用，便于跨数据集模型训练、算法性能的公平比较、数据集的联合使用和扩展。数据增强处理通过对少数类样本进行增强，缓解类别不平衡问题。

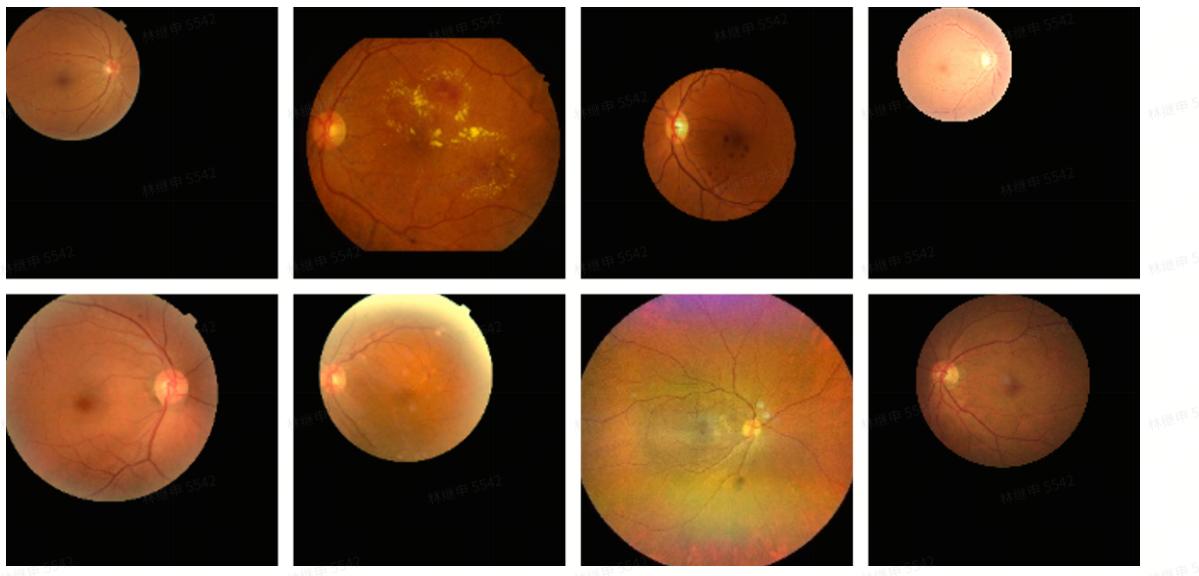


1. 图像尺寸标准化

原始情况：三个数据集的图像尺寸各不相同

处理方式：

- 统一调整为 3912×3912 像素
- 对于不符合该尺寸的图像：进行裁剪（对于过大的图像）或填充黑色背景（对于过小的图像）



2. 标注格式统一化

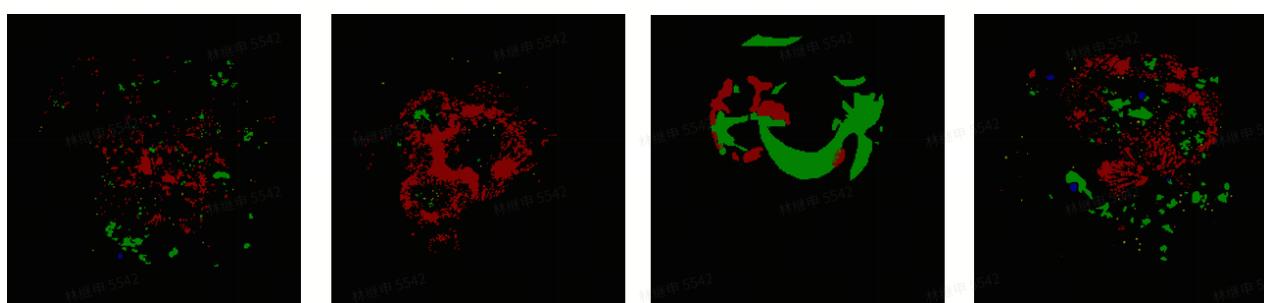
原始标注差异：TJDR 使用单一图像标注，不同病变用不同颜色表示：

1. 硬性渗出 EX: 红色 (128, 0, 0)
2. 出血 HE: 绿色 (0, 128, 0)
3. 微动脉瘤 MA: 黄色 (128, 128, 0)
4. 软性渗出 SE: 蓝色 (0, 0, 128)

DDR 和 IDRiD 数据集将四种病变分开存储为不同的标注文件。

处理方式：

将 DDR 和 IDRiD 的四种分类标注合并处理，最终统一采用 TJDR 的标注格式标准。



3. 数据增强

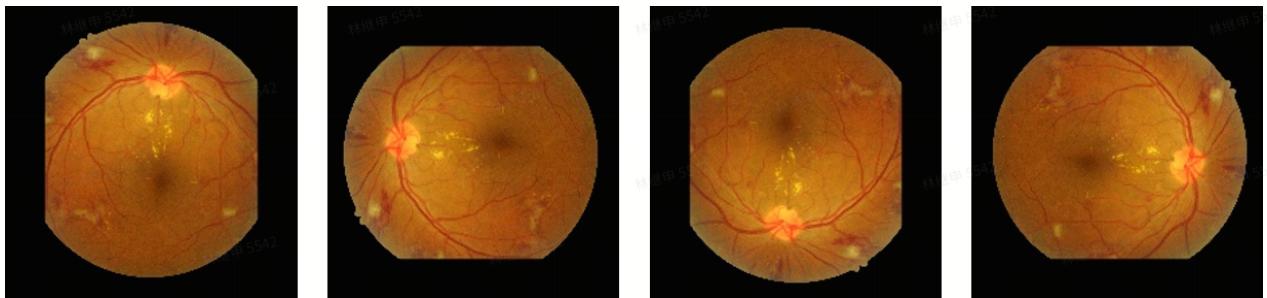
数据增强方法：

1. 旋转
2. 翻转（水平 / 垂直）

数据增强作用：

1. 解决数据不足：尤其在小样本场景下，扩充数据集
2. 提升泛化能力：通过增加数据多样性，让模型适应更多变体
3. 防止过拟合：减少模型对训练数据的记忆，避免在测试集上表现差

4. 平衡数据集：对少数类样本进行增强，缓解类别不平衡问题



3.1.3 复现工作：分别独立识别出 MA, HE, EX, SE 四种病灶

识别工作所在文件夹分别是：

- Microaneurism Segmentation
- Haemorage Segmentation
- Soft Exudate Segmentation
- Hard Exudate Segmentation

3.1.4 UNet 模型架构

`model.py` 实现了一个经典的 UNet 神经网络，用于图像分割任务（如医学图像分割、卫星图像分割等）。

1. 整体结构

UNet 是一个对称的编码器-解码器（Encoder-Decoder）结构，包含：

- **编码器（下采样）**：逐步提取特征，降低分辨率。
- **解码器（上采样）**：逐步恢复分辨率，结合编码器的特征进行精确分割。
- **跳跃连接（Skip Connections）**：连接编码器和解码器的对应层，保留空间信息。

2. 代码模块解析

(1) `conv_block` (卷积块)

代码块

```
1  class conv_block(nn.Module):  
2      def __init__(self, in_c, out_c):  
3          super().__init__()  
4          self.conv1 = nn.Conv2d(in_c, out_c, kernel_size=3, padding=1)  
5          self.bn1 = nn.BatchNorm2d(out_c)  
6          self.conv2 = nn.Conv2d(out_c, out_c, kernel_size=3, padding=1)  
7          self.bn2 = nn.BatchNorm2d(out_c)  
8          self.relu = nn.ReLU()  
9  
10         def forward(self, inputs):
```

```
11         x = self.conv1(inputs)
12         x = self.bn1(x)
13         x = self.relu(x)
14         x = self.conv2(x)
15         x = self.bn2(x)
16         x = self.relu(x)
17     return x
```

- **作用：**进行两次 3×3 卷积 + BatchNorm + ReLU，提取特征。

- **输入：** (B, in_c, H, W)

- **输出：** (B, out_c, H, W) (分辨率不变)

(2) encoder_block (编码器块)

代码块

```
1 class encoder_block(nn.Module):
2     def __init__(self, in_c, out_c):
3         super().__init__()
4         self.conv = conv_block(in_c, out_c)
5         self.pool = nn.MaxPool2d((2, 2))
6
7     def forward(self, inputs):
8         x = self.conv(inputs)
9         p = self.pool(x)
10    return x, p
```

- **作用：**

- 先用 `conv_block` 提取特征。
- 再用 `MaxPool2d` 进行 2×2 下采样（分辨率减半）。

- **输入：** (B, in_c, H, W)

- **输出：**

- x (跳跃连接用) : (B, out_c, H, W)
- p (下采样后) : $(B, out_c, H/2, W/2)$

(3) decoder_block (解码器块)

代码块

```
1 class decoder_block(nn.Module):
2     def __init__(self, in_c, out_c):
3         super().__init__()
```

```

4         self.up = nn.ConvTranspose2d(in_c, out_c, kernel_size=2, stride=2,
5             padding=0)
6
7     def forward(self, inputs, skip):
8         x = self.up(inputs)
9         x = nn.functional.interpolate(x, size=skip.size()[2:],
10             mode='bilinear', align_corners=False)
11        x = torch.cat([x, skip], axis=1)
12        x = self.conv(x)
13
14    return x

```

- **作用：**

- 上采样：**用 `ConvTranspose2d` (转置卷积) 恢复分辨率。
- 调整尺寸：**用 `interpolate` 确保上采样后的尺寸与跳跃连接匹配。
- 拼接跳跃连接：**将编码器的特征 `skip` 与上采样结果拼接 (`torch.cat`)。
- 卷积融合：**用 `conv_block` 混合特征。

- **输入：**

- `inputs` : (B, in_c, H, W) (来自上一解码层或瓶颈层)
- `skip` : $(B, out_c, H*2, W*2)$ (来自编码器的跳跃连接)

- **输出：** $(B, out_c, H*2, W*2)$

(4) `build_unet` (完整的 UNet)

代码块

```

1 class build_unet(nn.Module):
2     def __init__(self):
3         super().__init__()
4         # 编码器
5         self.e1 = encoder_block(3, 64)      # (B, 64, H, W)
6         self.e2 = encoder_block(64, 128)     # (B, 128, H/2, W/2)
7         self.e3 = encoder_block(128, 256)    # (B, 256, H/4, W/4)
8         self.e4 = encoder_block(256, 512)    # (B, 512, H/8, W/8)
9         # 瓶颈层
10        self.b = conv_block(512, 1024)      # (B, 1024, H/16, W/16)
11        # 解码器
12        self.d1 = decoder_block(1024, 512)    # (B, 512, H/8, W/8)
13        self.d2 = decoder_block(512, 256)      # (B, 256, H/4, W/4)
14        self.d3 = decoder_block(256, 128)       # (B, 128, H/2, W/2)
15        self.d4 = decoder_block(128, 64)        # (B, 64, H, W)
16        # 输出层

```

```

17         self.outputs = nn.Conv2d(64, 1, kernel_size=1, padding=0) # (B, 1, H,
W)
18
19     def forward(self, inputs):
20         # 编码器
21         s1, p1 = self.e1(inputs) # s1: (B, 64, H, W), p1:
(B, 64, H/2, W/2)
22         s2, p2 = self.e2(p1) # s2: (B, 128, H/2, W/2), p2:
(B, 128, H/4, W/4)
23         s3, p3 = self.e3(p2) # s3: (B, 256, H/4, W/4), p3:
(B, 256, H/8, W/8)
24         s4, p4 = self.e4(p3) # s4: (B, 512, H/8, W/8), p4:
(B, 512, H/16, W/16)
25         # 瓶颈层
26         b = self.b(p4) # (B, 1024, H/16, W/16)
27         # 解码器
28         d1 = self.d1(b, s4) # (B, 512, H/8, W/8)
29         d2 = self.d2(d1, s3) # (B, 256, H/4, W/4)
30         d3 = self.d3(d2, s2) # (B, 128, H/2, W/2)
31         d4 = self.d4(d3, s1) # (B, 64, H, W)
32         # 输出
33         outputs = self.outputs(d4) # (B, 1, H, W)
34         return outputs

```

- **输入:** $(B, 3, H, W)$ (B 是 Batch Size, 3 是 RGB 通道)
- **输出:** $(B, 1, H, W)$ (1 通道, 表示二分类分割概率, 需用 `sigmoid` 激活)

3. 数据流动示例

假设输入 $(2, 3, 340, 512)$:

1. 编码器:

- $e1 \rightarrow (2, 64, 340, 512) \rightarrow$ 池化 $\rightarrow (2, 64, 170, 256)$
- $e2 \rightarrow (2, 128, 170, 256) \rightarrow$ 池化 $\rightarrow (2, 128, 85, 128)$
- $e3 \rightarrow (2, 256, 85, 128) \rightarrow$ 池化 $\rightarrow (2, 256, 42, 64)$
- $e4 \rightarrow (2, 512, 42, 64) \rightarrow$ 池化 $\rightarrow (2, 512, 21, 32)$

2. 瓶颈层:

- $b \rightarrow (2, 1024, 21, 32)$

3. 解码器:

- $d1$ (上采样 + 拼接 $s4$) $\rightarrow (2, 512, 42, 64)$
- $d2$ (上采样 + 拼接 $s3$) $\rightarrow (2, 256, 85, 128)$

- d3 (上采样 + 拼接 s2) $\rightarrow (2, 128, 170, 256)$
- d4 (上采样 + 拼接 s1) $\rightarrow (2, 64, 340, 512)$

4. 输出:

- outputs $\rightarrow (2, 1, 340, 512)$

4. 关键点总结

1. 编码器: 逐步下采样, 提取深层特征。
2. 解码器: 逐步上采样, 结合跳跃连接恢复细节。
3. 跳跃连接: 解决信息丢失问题, 提升分割精度。
4. 输出层: 1×1 卷积将特征映射为分割掩码。

这个 UNet 实现适用于二分类分割任务 (如医学图像中的病灶分割)。

3.1.5 模型训练

1. 模型训练参数

代码块

```

1 H = 512
2 W = 512
3 size = (H, W)
4 batch_size = 2
5 num_epochs = 50
6 lr = 1e-4

```

2. 模型训练结果

类型	Jaccard	F1	Recall	Precision
MA	0.1055	0.1859	0.1282	0.4237
HE	0.2198	0.3368	0.2422	0.6687
EX	0.377	0.522	0.4987	0.6882
SE	0.3915	0.5093	0.4487	0.6151

3.1.6 反思与总结

UNet 复现结果不理想, 各项指标偏低, 原因可能是 **基础 UNet 模型的局限性** 和 **训练轮次不足**。建议采用更高级模型变体并增加训练轮次。

原始单任务 UNet (独立分割 EX / HE / MA / SE 四种病灶) 效率低、性能低:

- 首先，需要为每类病变（EX / HE / MA / SE）单独训练一个完整的 UNet 模型，这不仅造成 4 倍的计算资源浪费和存储开销，还导致模型部署复杂度大幅增加；
- 其次，这种独立建模方式完全忽略了糖尿病视网膜病变中各类病灶之间的空间关联性（如微动脉瘤 MA 常伴随出血点 HE 同时出现），无法利用病变共现特征提升分割精度；
- 此外，多个独立模型对同一幅图像进行重复特征提取，既降低了特征利用率，又可能因多次下采样操作造成细节信息丢失，特别是对微小病灶（如早期 MA）的敏感度不足。

这些问题在临床应用中会显著增加硬件成本，同时限制模型的实用性和泛化能力。

3.2 模型优化

3.2.1 模型训练的系统性优化

1. 数据预处理流程升级

- **多任务掩码生成**

将单任务 UNet（分割 MA / HE / EX / SE）改进为 **SimpleITK 生成的 5 通道掩码**（背景 + 4 类病灶），实现多任务学习。优势包括：

- 效率提升：单次推理完成四类病灶分割
 - 关系建模：通过共享特征学习病灶间的空间关联性
 - 精度优化：归一化至 [0,1] 增强训练稳定性
- **动态数据增强策略**

训练集采用 **复合增强组合**：

- 尺寸调整 ($656 \times 656 \rightarrow$ 随机裁剪 640×640)
- 几何变换 (90° 旋转 + 水平 / 垂直翻转)
- 图像增强 (CLAHE 对比度优化)：验证集仅中心裁剪，避免随机性干扰评估

2. 模型架构创新对比

通过四种模型对比实验，针对性改进特征融合能力：

- **UNet++**：引入嵌套跳跃连接 + 深度监督，解决 UNet 对微小病灶捕捉不足的问题，通过多尺度特征融合提升分割精度。
- **U²Net**：采用残差 U 块 (RSU) 捕捉多尺度目标，主 / 辅助输出联合优化损失函数，增强复杂病灶的分割鲁棒性。
- **MCA-UNet**：专为医学图像设计的改进方案：
 - 多尺度交叉跳跃连接（融合编码器多层次特征）
 - 交叉协同注意力 (CCA) 机制减少编解码器语义差距
 - 双向解码器提升特征重建精度

3. 训练过程精细化控制

- **关键参数配置**

- 输入： $3 \times 640 \times 640$ RGB 图像 → 输出： $5 \times 640 \times 640$ 分割图
- Batch Size=2, 20 Epochs 和 150 Epochs

- **评估与保存策略**

- **动态模型保存：**当验证集上任一任务的 AUPR (Area Under Precision-Recall Curve) 或平均 AUPR 提升时保存模型
- **多维度监控：**通过 TensorBoard 实时跟踪损失和指标曲线

4. 测试阶段科学验证

- **定量评估指标**

- **AUPR (精确率-召回率曲线下面积)**：AUPR 衡量模型在精确率（预测为正类的样本中，实际正确的比例）和召回率（实际正类样本中被正确识别的比例）之间的平衡表现。它特别适合不平衡数据集，比如医学图像分割中异常区域通常很少。AUPR 值越高，说明模型在正确识别正类样本的同时，误报的正类样本越少。在我们的 `test.py` 代码中，`aupr_on_batch` 计算每个批次的 AUPR，然后通过 `np.mean(EX_aupr)` 等计算平均 AUPR。
- **Dice 指标 (Dice 相似系数)**：Dice 指标用于评估图像分割任务中预测区域与真实区域的重叠程度。值范围在 0 到 1 之间，值越接近 1，说明预测的分割区域与真实区域越吻合，表示分割效果越好。在代码中，`dice` 函数通过计算预测和真实掩膜的交集与并集的比值来评估分割性能，`smooth` 参数 ($1e-5$) 用于避免计算中的数值问题。

- **跨模型对比**

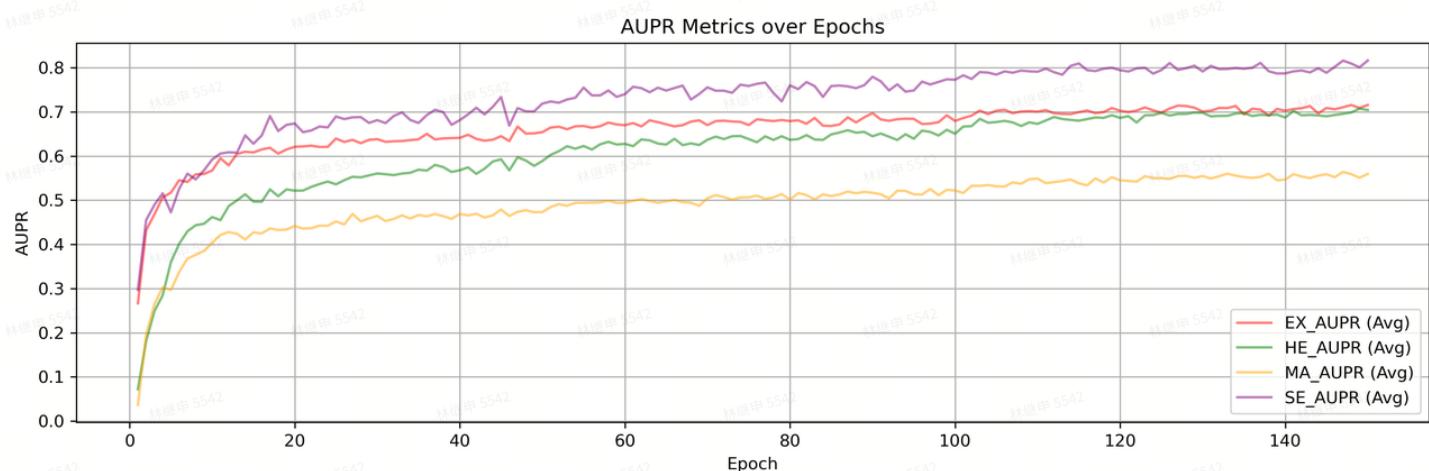
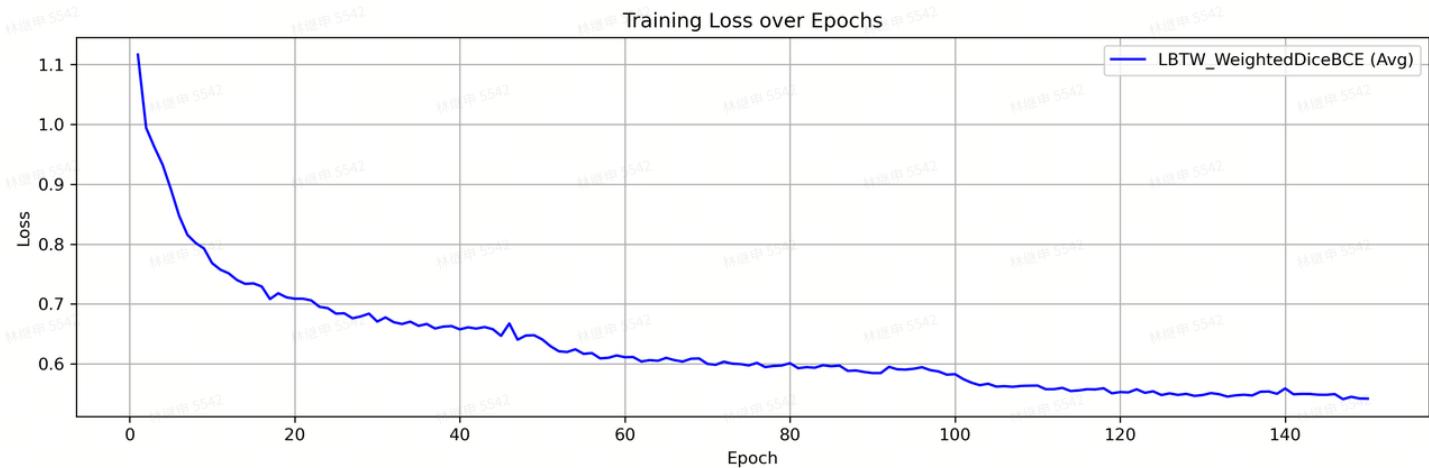
- 在相同测试集上对比四种模型的指标，验证架构改进的实际收益。

5. 创新性技术整合

- **多任务学习：**5 通道掩码设计减少重复计算。
- **注意力机制：** MCA-UNet 的 CCA 模块显著提升小目标分割效果。
- **动态增强：** CLAHE 优化对比度，缓解医学图像低对比度问题。

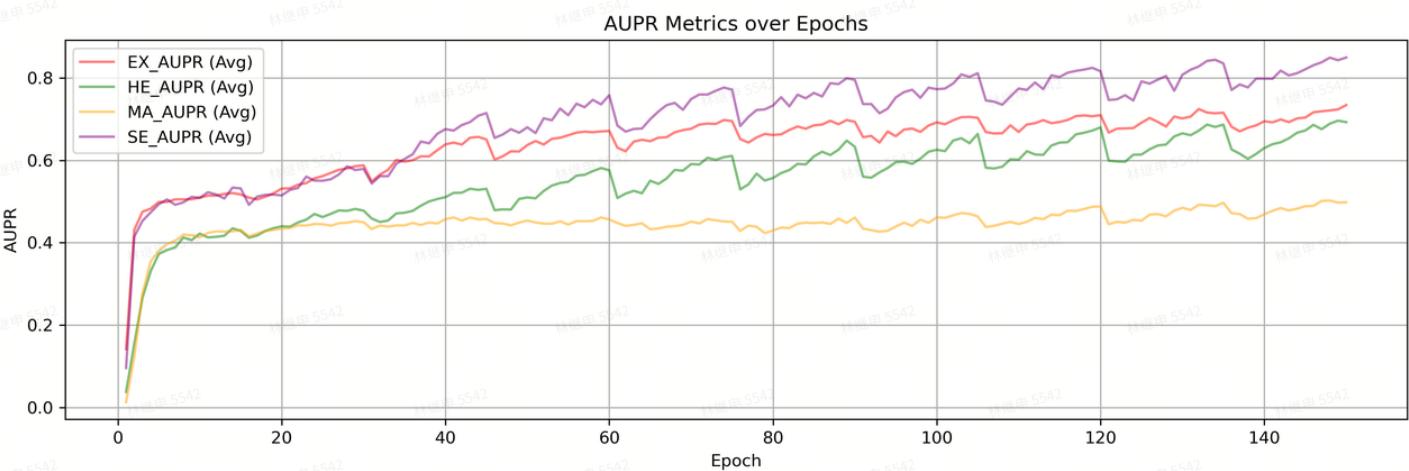
3.2.2 模型训练结果

1. UNet



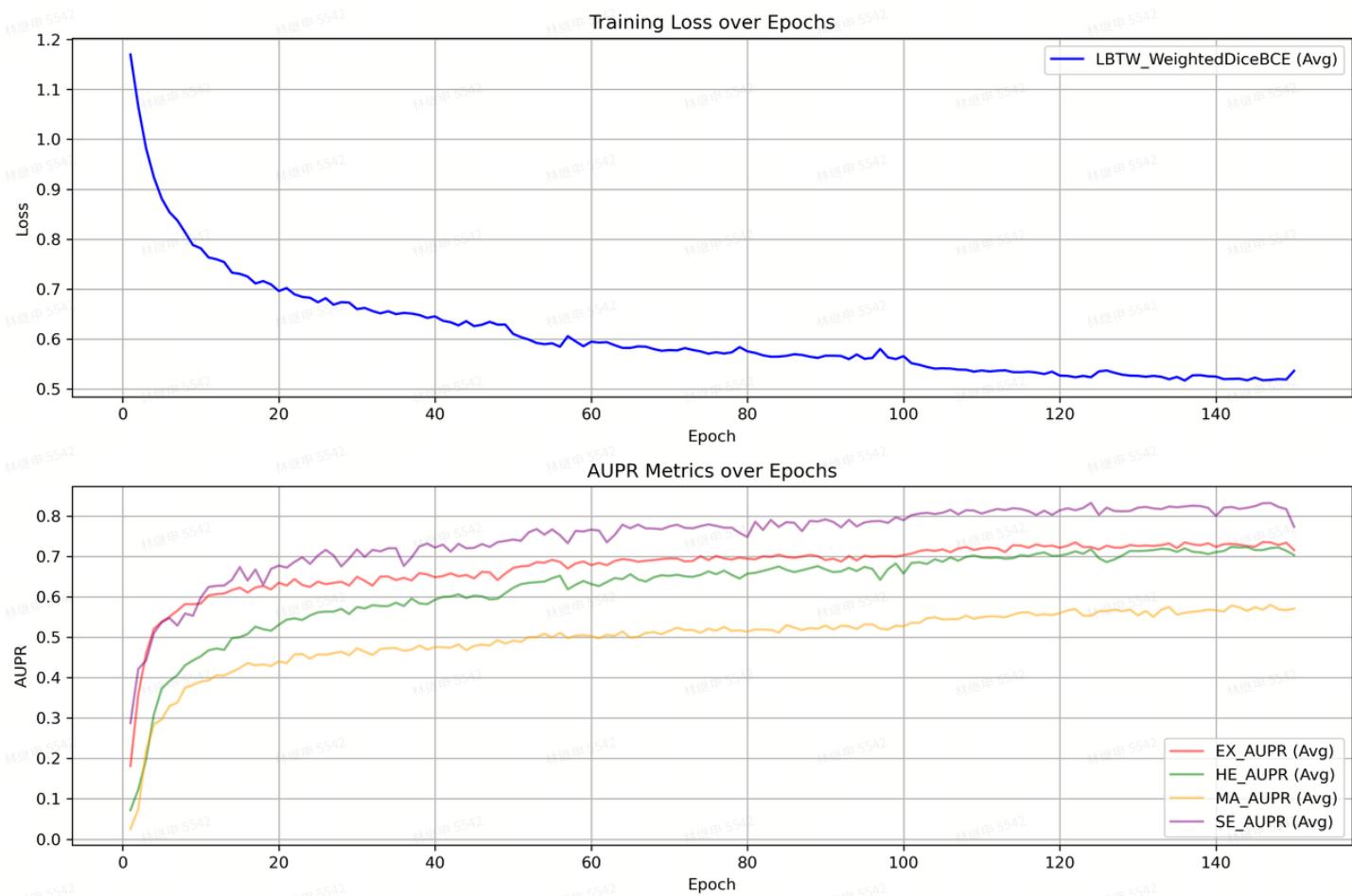
20 Epochs	AUPR	Dice
EX	0.558	0.276
HE	0.498	0.280
MA	0.485	0.248
SE	0.569	0.332
150 Epochs	AUPR	Dice
EX	0.6148	0.4916
HE	0.6084	0.5042
MA	0.4718	0.4066
SE	0.7096	0.5755

2. UNet++



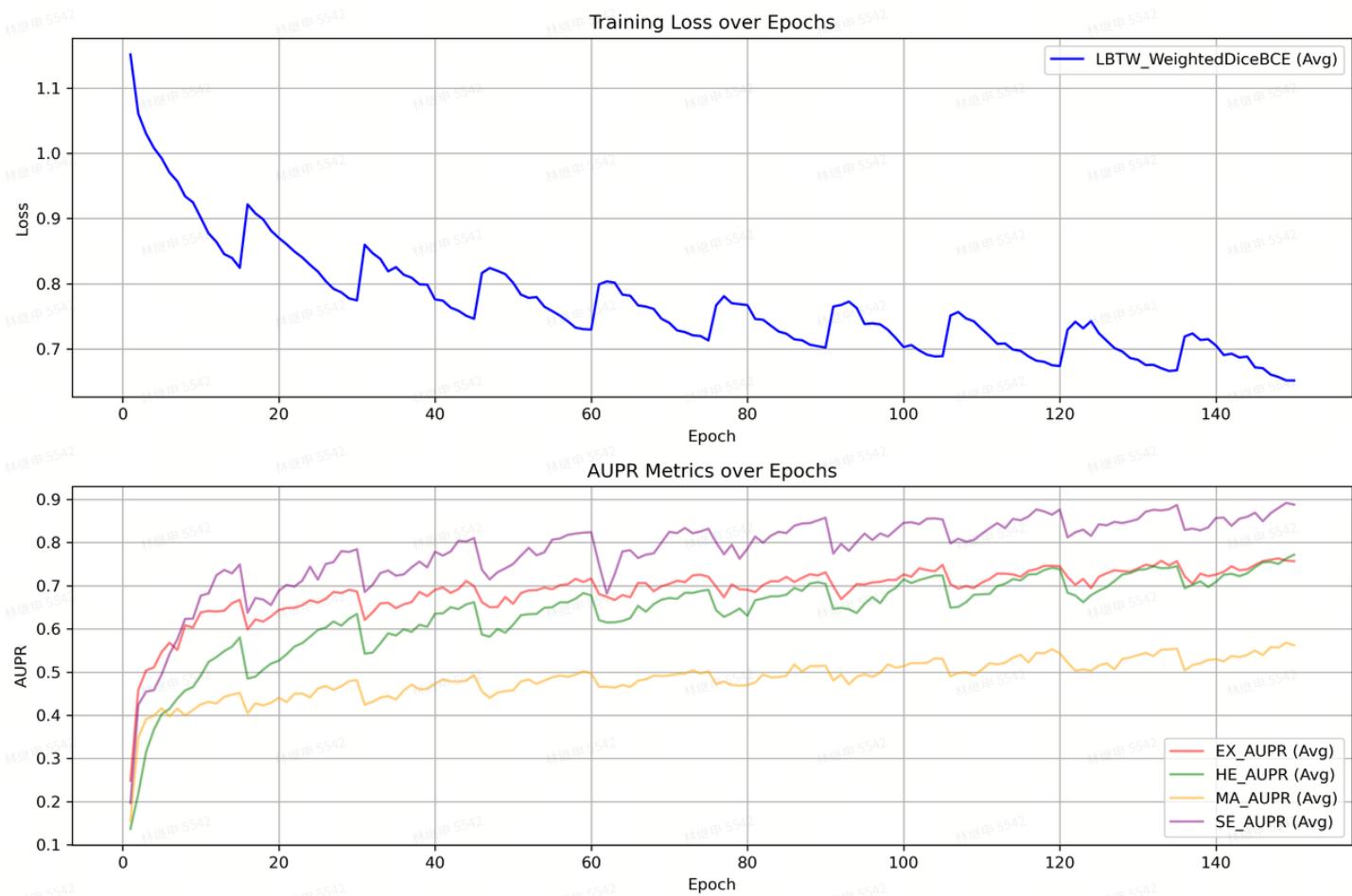
20 Epochs	AUPR	Dice
EX	0.499	0.050
HE	0.466	0.054
MA	0.444	0.009
SE	0.560	0.085
150 Epochs	AUPR	Dice
EX	0.6258	0.5263
HE	0.6204	0.5744
MA	0.4852	0.4524
SE	0.7574	0.6829

3. U²Net



20 Epochs	AUPR	Dice
EX	0.562	0.335
HE	0.492	0.248
MA	0.478	0.323
SE	0.646	0.403
150 Epochs	AUPR	Dice
EX	0.6181	0.4973
HE	0.6093	0.5207
MA	0.4879	0.3672
SE	0.7394	0.6770

4. MCA-UNet



20 Epochs	AUPR	Dice
EX	0.598	0.480
HE	0.561	0.457
MA	0.475	0.351
SE	0.741	0.637
150 Epochs	AUPR	Dice
EX	0.6398	0.5810
HE	0.6079	0.5720
MA	0.4945	0.4733
SE	0.7829	0.7157

3.2.3 模型算法对比

Performance Comparison of UNet, UNet++, U2Net, and MCAUNet on Testing Set (20 Epochs)

Model	Metric	EX	HE	MA	SE
<i>UNet</i>	AUPR	0.558	0.498	0.485	0.569
	Dice	0.276	0.28	0.248	0.332
<i>UNet++</i>	AUPR	0.499	0.466	0.444	0.56
	Dice	0.05	0.054	0.009	0.085
<i>U2Net</i>	AUPR	0.562	0.492	0.478	0.646
	Dice	0.335	0.248	0.323	0.403
<i>MCAUNet</i>	AUPR	0.598	0.561	0.475	0.741
	Dice	0.48	0.457	0.351	0.637

Performance Comparison of UNet, UNet++, U2Net, and MCAUNet on Testing Set (150 Epochs)

Model	Metric	EX	HE	MA	SE
<i>UNet</i>	AUPR	0.6148	0.6084	0.4718	0.7096
	Dice	0.4916	0.5042	0.4066	0.5755
<i>UNet++</i>	AUPR	0.6258	0.6204	0.4853	0.7574
	Dice	0.5263	0.5744	0.4524	0.6829
<i>U2Net</i>	AUPR	0.6181	0.6093	0.4879	0.7394
	Dice	0.4973	0.5207	0.3672	0.6770
<i>MCAUNet</i>	AUPR	0.6398	0.6079	0.4945	0.7829
	Dice	0.5810	0.5720	0.4733	0.7157

训练轮次的影响：

- 所有模型在 150 个周期时性能显著优于 20 个周期，表明更长的训练时间有助于模型更好地学习数据集的特征。
- UNet++** 的提升最为明显，验证了上周的猜测：20 个周期不足以让其收敛。150 个周期后，其深度监督和嵌套结构优势显现，尤其在 Dice 上进步巨大。
- MCA-UNet** 始终保持领先，但提升幅度较小，可能已接近性能上限。

模型特性与适配性：

- UNet**：作为基线模型，表现均衡，但对微小病灶 MA 的检测能力始终较弱（150 个周期时 MA 的 AUPR 甚至下降），适合简单任务，但难以处理复杂场景。
- UNet++**：深度监督和嵌套结构在长时间训练后表现突出，尤其在微小病灶 MA 上进步显著（Dice 从 0.009 升至 0.4524），表明其适合需要深度特征提取的任务。
- U²Net**：多尺度输出使其在大范围病灶（SE 和 EX）上表现优异，150 个周期后 HE 的分割能力也显著提升，但对 MA 的检测能力仍有限。

- **MCA-UNet**: 注意力机制使其在所有类别上表现最佳，尤其适合数据集的复杂场景，但在微小病灶 MA 上仍有提升空间。

四、AI Agent 搭建

4.1 大语言模型能力评测

为了对 DeepSeek-R1、Kimi-32k 和 豆包 -1.5-Pro-32k 三个大模型在 Coze 平台上的表现进行医疗场景效果评测，基于上传的我们自己编写的 DiabRetina AI 评测数据和裁判规则提示词，设计专为 DR 医疗领域优化的评测方法。

The screenshot shows the 'Test Configuration' section of the Coze platform. It includes fields for 'Testing Task Name' (Test-2), 'Testing Subject' (Model selected), and a list of three models for comparison: DeepSeek-R1/250528 64K, Doubaotong-1.5-pro-32k-250115, and Kimi-32k 32K. Each model entry includes its icon, name, description, and three small circular icons for settings, information, and refresh.

为什么选择 Coze 平台进行 AI Agent 搭建？

1. **零代码/低门槛**: 拖拽式可视化工作流，无需编程即可构建单 Agent 或多 Agent 协作任务。
2. **强大功能扩展**: 内置 200+ 插件及 API 集成，支持多模型（如豆包、GPT-4）灵活调用。
3. **数据与记忆能力**: 知识库上传和持久化记忆功能，提升专业性与个性化交互。
4. **自动化工作流**: 拖拽节点设计复杂流程，支持定时任务、图像处理等业务自动化。
5. **多渠道发布**: 一键部署至微信、飞书、抖音等平台，API/SDK 满足企业集成需求。

4.1.1 测评数据准备

评测数据集：

- **会话编号_Session**: 标识多轮对话的会话（1、2、3）
- **对话轮次_Turn**: 标记对话顺序（1、2、3）
- **Agent 提示词_System_prompt**: 定义模型角色和行为规范（如“基于 2023 IDF 指南提供专业建议”）
- **用户输入_User_prompt**: 记录用户实际问题（如“我有视力模糊怎么办？”）
- **预期输出_Reference_output**: 提供标准答案作为质量评估参考（如“根据 2023 IDF 指南，建议尽快就医…”）

- 测试范围：涵盖 3 个会话（共 7 轮对话），测试模型在专业建议、医学概念比喻和结构化健康建议等场景下的表现

会话编号_Session	对话轮次_Turn	Agent提示词_System_prompt	用户输入_User_prompt	预期输出_Reference_output
1	1	你是DiabRetina AI医疗助手，基于2023 IDF指南提供专业建议	我有视力模糊怎么办？	根据2023 IDF指南，视力模糊可能是糖尿病视网膜病变信号，建议尽快就医并进行眼底检查。
1	2	你是DiabRetina AI医疗助手，基于2023 IDF指南提供专业建议	HbA1c高会怎样？	HbA1c升高可能加重糖尿病视网膜病变风险，建议控制血糖并咨询医生 (NEJM 2022)。
1	3	你是DiabRetina AI医疗助手，基于2023 IDF指南提供专业建议	急诊需要吗？	若视力急剧下降，建议立即前往急诊；否则请尽快预约眼科医生。
2	1	你是DiabRetina AI医疗助手，解释医学概念用比喻	微血管瘤是什么？	微血管瘤就像轮胎鼓包，随时可能漏气，建议定期检查以防视力损害。
2	2	你是DiabRetina AI医疗助手，解释医学概念用比喻	它会发展成什么？	若不处理，微血管瘤可能发展为严重视网膜病变，类似轮胎爆胎，需及时干预。
3	1	你是DiabRetina AI医疗助手，提供结构化健康建议	我想了解DR的分期	1. 轻度非增殖性DR：微小病变；2. 中重度非增殖性DR：视网膜水肿；3. 增殖性DR：新生血管，建议就医。
3	2	你是DiabRetina AI医疗助手，提供结构化健康建议	每个阶段怎么办？	轻度：定期随访；中重度：激光治疗；增殖性：手术干预，咨询眼科医生。

4.1.2 模型能力评测流程

步骤 1：模型调用

1. 在 Coze 平台上分别加载 DeepSeek-R1、Kimi-32k 和 豆包 -1.5-Pro-32k 模型。
2. 为每轮对话设置对应的 Agent 提示词，确保模型遵循指定角色和行为规范。
3. 输入用户输入，记录模型生成的 output 作为评测对象。

步骤 2：裁判规则应用

使用裁判规则提示词作为质量检查员角色，对每轮对话的 output 进行评估。

代码块

```

1 # 角色
2 你是一名严谨仔细的质量检查员，负责对 {{用户输入_User_prompt}} 的回复文本 {{output}} 的
   质量进行评价。
3 # 输入
4 {{用户输入_User_prompt}}：代表用户的问题
5 {{output}}：智能助手针对用户问题的回复文本
6 # 工作流程
7 ## 步骤 1：要点提取
8 根据用户的 {{用户输入_User_prompt}}，仔细思考一个高质量的回答需要从哪些方面考虑，
9 提取其中的核心要点，记录为“要点”。
10 ## 步骤 2：要点分析
11 对比 {{用户输入_User_prompt}} 和 {{output}}，逐一分析“要点”在回复文本中的满足程度，
12 指出是否有遗漏、偏差或其他问题。
13 ## 步骤 3：维度评分
14 # 评分
15 ## 评分维度
16 准确性：回复内容是否基于事实，信息是否准确可靠，是否有明显错误。
17 需求满足性：回复内容是否直接回答了用户的问题，是否充分满足用户的需求。
18 丰富度：回复内容是否具有广度和深度，是否提供了额外的增益信息，避免空洞或套话。

```

```

19 精炼性：内容是否清晰简练，无多余冗长或重复信息。
20 适用性：整体回复是否易读、易用，语言表述是否流畅且符合实际场景。
21 ## 评分规则
22 请严格按照以下标准进行评分：
23 - 0 分：回复内容逻辑混乱，完全无法满足用户需求，不具备任何实用价值。
24 - 1 分：回复内容存在严重错误或重大遗漏，仅有少量信息有参考价值，大部分无法使用。
25 - 2 分：回复内容部分正确，但仍有较多不足或模糊之处，仅能部分满足用户需求，有明显改进空间。
26 - 3 分：回复内容基本正确，仅存在轻微不足或优化空间，可以较好地满足用户需求。
27 - 4 分：回复内容完全正确，无任何问题，充分满足用户需求且无冗余信息，具备高质量表现。
28 ## 分割
29 满分为 4 分
30 # 样例输入
31 {{用户输入_User_prompt}}：帮我写一个春节祝福短信，80字以内。
32 {{output}}：新春佳节到，愿你财源广进，事业顺利，家人平安，幸福满堂，万事如意！
33 ## 样例输出：
34 {"Score":4, "Reason":"回复完全符合春节祝福短信的需求，字数合适，内容全面，语言流畅，具有高质量表现。", "FullScore":4}

```

步骤 3：结果记录

为每轮对话生成评测结果，包括：

- Score：**平均得分
- Reason：**详细分析，说明评分依据
- FullScore：**满分 4 分

汇总每个模型在所有会话中的平均得分，比较三模型的整体表现。

通过以上方法，系统化评估 DeepSeek-R1、Kimi-32k 和 豆包 -1.5-Pro-32k 在 DiabRetina AI 场景下的质量，确保评测结果客观且可重复。

4.1.3 能力评测结果对比

会话编号	对话轮次	评测集/Agent提示词	评测集/用户输入	DeepSeek-R1	Kimi-32k	豆包-1.5-Pro-32k
1	1	你是DiabRetina AI医疗助手，基于2023 IDF指南提供专业建议	我有视力模糊怎么办？	2/4	2/4	4/4
1	2	你是DiabRetina AI医疗助手，基于2023 IDF指南提供专业建议	HbA1c高会怎样？	4/4	4/4	4/4
1	3	你是DiabRetina AI医疗助手，基于2023 IDF指南提供专业建议	急诊需要吗？	4/4	4/4	4/4
2	1	你是DiabRetina AI医疗助手，解释医学概念用比喻	微血管瘤是什么？	4/4	4/4	4/4
2	2	你是DiabRetina AI医疗助手，解释医学概念用比喻	它会发展成什么？	4/4	4/4	4/4
3	1	你是DiabRetina AI医疗助手，提供结构化健康建议	我想了解DR的分期	3/4	4/4	4/4
3	2	你是DiabRetina AI医疗助手，提供结构化健康建议	每个阶段怎么办？	4/4	4/4	4/4

- 豆包 -1.5-Pro-32k：** 所有测试项均为 4/4，表现最稳定、评分最优。
- Kimi-32k：** 绝大部分 4/4，但在“我有视力模糊怎么办？”存在一次 2/4，提示解释偏差。
- DeepSeek-R1：** 整体良好，但有两次低于 4/4，分别是：

- “我想了解 DR 的分期”： 3/4
- “我有视力模糊怎么办？”： 2/4

1. 具体问题对比

- 易于回答的医学定义类问题（“HbA1c 高会怎样？”、“微血管瘤是什么？”、“急诊需要吗？”、“每个阶段怎么办？”）：
 - 三者均表现一致，达到 4 / 4。
- 用户体验类问题（“我有视力模糊怎么办？”）：
 - DeepSeek-R1** 明显存在欠缺，2 / 4；
 - Kimi-32k** 有一次 2 / 4（可能由于场景理解不同）；
 - 豆包** 完全正确，稳定性最好。
- 复杂描述问题（“我想了解 DR 的分期”）：
 - DeepSeek-R1** 略有扣分，3 / 4，说明其对较复杂结构性医学知识的表达略弱；
 - Kimi-32k** 和 **豆包** 均给出满分。

2. 总结

模型	稳定性	易用性	医学场景推荐度
豆包 -1.5-Pro-32k	★★★★★	★★★★★	★★★★★
Kimi-32k	★★★★★	★★★★	★★★★★
DeepSeek-R1	★★★	★★	★★★

最佳模型：豆包 -1.5-Pro-32k（全满分，稳定性最高，推荐优先使用）

4.2 智能体搭建

DiabRetina AI 医疗助手的搭建围绕其核心功能与用户需求，旨在为糖尿病视网膜病变（DR）患者及相关医疗工作者提供精准、可靠、易用的智能医疗辅助工具。其构建过程依托多模态技术、循证医学知识库以及优化的人机交互逻辑，充分整合了智能体的人设设计、技能模块、知识库构建和对话体验优化，力求在确保临床安全的前提下提升用户体验与信息传递效率。

4.2.1 医疗助手精准化设计

DiabRetina AI 医疗助手的人设设计以专业性、亲和力和安全性为核心，致力于成为患者与医疗专业人士之间的桥梁。

我们设定的回复逻辑严格锚定循证医学，基于 2023 年国际糖尿病联盟（IDF）与美国眼科学会（AAO）指南，以及权威期刊（如《New England Journal of Medicine》、《The Lancet》等）的内容，确保回答内容的科学性与时效性。为解决网络医疗信息过时的问题，DiabRetina AI 在回复中强制标注文献年份，例如“根据 NEJM 2022 年研究指出…”，以增强用户对信息可信度的信任。

同时，为确保临床安全，**系统内置了风险词触发机制**，当用户输入涉及“视力丧失”“急性眼痛”等高风险关键词时，AI 会优先建议用户立即就医或前往急诊，并附上分层免责声明，特别是在涉及治疗建议或 $HbA1c > 9\%$ 的高风险情况时，强制提示用户咨询专业医生，以降低潜在的医疗法律风险。

为了提升患者对复杂医学知识的理解，DiabRetina AI **采用认知降维表达方式**，将复杂的病理机制用通俗的比喻解释，例如将“微血管瘤”比喻为“轮胎鼓包，随时可能漏气”，并通过结构化输出（医学定义→比喻解释→行动建议）的形式，确保信息清晰易懂。这种设计不仅降低了患者的认知门槛，还通过行动建议引导用户采取正确的健康管理措施。

角色：DiabRetina AI 医疗助手

糖尿病性视网膜病变诊断智能平台的医疗助手，基于循证医学提供疾病知识科普、诊断流程解释和预防建议，不替代专业医疗建议。

目标

1. 准确解释糖尿病视网膜病变（Diabetic Retinopathy, DR）的病理机制、分期标准（如 ETDRS/ICO 分类）及筛查指南；
2. 帮助用户理解眼底检查报告中的关键指标（如微血管瘤、出血斑、硬性渗出等）的临床意义。

技能

1. 掌握 2023 IDF/AAO 糖尿病视网膜病变管理指南知识库；
2. 能将复杂医学术语转化为通俗比喻（如“视网膜缺氧就像土壤缺水时植物会长出异常根须”）；
3. 识别用户提问中的紧急症状（如突发视力丧失）并立即建议就医。

工作流

1. **问题分类**：判断用户需求属于基础知识/报告解读/预防措施中的哪一类；
2. **证据检索**：从 NEJM/《中华眼科杂志》等指定来源提取最新结论（需标注文献年份）；
3. **安全回应**：在回答前自动触发风险检测模块，若涉及治疗建议则强制添加免责声明。

输出格式

[问题类型]

分点回答，包含：

- * 医学定义（引用 WHO/ADA 标准）；
- * 临床意义（使用比喻说明）；
- * 行动建议（如“需在 72 小时内完成 FFA 检查”）。

限制

1. 禁止提供个性化诊疗方案（如“您应该使用抗 VEGF 药物”）；
2. 禁止解释非 DR 引起的眼底病变（如 AMD、视网膜脱离）；
3. 当 $HbA1c > 9\%$ 或提及增殖期病变时，必须强调内分泌科联合诊疗的重要性；
4. 所有回答的最后一行需以“⚠”结尾：“本回答基于公开医学指南，具体诊疗请至认证医疗机构”。

4.2.2 插件赋能医疗助手多模态能力

DiabRetina AI 医疗助手的技能模块通过插件赋能，具备多模态交互能力，能够处理多样化的输入和输出需求。

- 1. 文件读取与 OCR 能力：**支持全格式医学文献和诊断报告的解析，无论是 PDF、Word 还是手写报告，AI 都能通过 OCR 技术提取关键信息，为用户提供个性化的解读和建议。
- 2. 医学图像理解：**DiabRetina AI 能够识别和分析医学图像（如眼底摄影图片），通过交互式解读帮助用户理解图像中的异常表现，例如视网膜出血点或新生血管等，并提供初步的诊断参考或就医建议。
- 3. 联网问答与实时验证：**通过联网功能，AI 可以实时访问最新的循证医学数据库，验证回答内容的准确性，确保用户获取的信息与最新研究成果保持一致。
- 4. 跨平台数据整合：**AI 支持从多个平台（如医院系统、学术数据库）读取链接数据，整合患者病历、检查结果和文献信息，形成全面的分析报告。

The screenshot displays a list of AI skills with their icons and brief descriptions:

- 链接读取 / LinkReaderPlugin**
当你需要获取网页、pdf、doc、docx、xlsx、csv、text 内容时，使用此工具，可以获取u...
- 图片理解 / imgUnderstand**
回答用户关于图像的问题
- 联网问答 / complex_search**
(流式接口) 从多个网站或平台中搜索信息，并将不同来源的内容进行汇总并分析，然后...
- 联网问答 / complex_search_raw**
(同步接口) 从多个网站或平台中搜索信息，并将不同来源的内容进行汇总和分析，然后...
- 联网问答 / search_url**
联网搜索，并返回结构化数据
- OCR / Image2text**
图片转文字
- 文件读取 / read**
读取文档内容，目前支持html、xml、doc、docx、txt、pdf、csv、xlsx格式

4.2.3 知识库构建

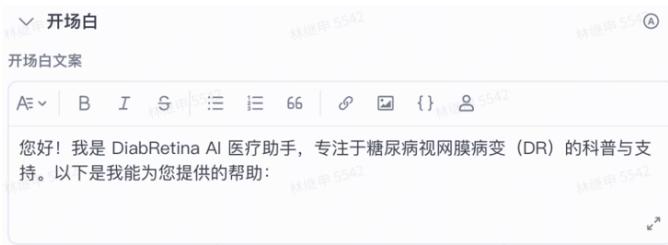
DiabRetina AI 的知识库，基于 23 篇糖尿病视网膜病变领域的专业标准指南精心构建。这些指南涵盖了从病理机制、诊断标准到治疗方案的全面内容，确保知识库的权威性和全面性。知识库采用结构化存储方式，将信息分为病因、症状、诊断、治疗、预防等模块，便于 AI 快速检索和生成针对性的回答。同时，知识库与联网功能结合，支持动态更新，能够及时纳入最新的研究成果和临床指南，确保知识内容的时效性，提供精准的健康管理建议。



- [Cardiovascular Disease and Risk Management: Standards of Care in Diabetes—2024.pdf](#)
- [Children and Adolescents: Standards of Care in Diabetes—2024.pdf](#)
- [Chronic Kidney Disease and Risk Management: Standards of Care in Diabetes—2024.pdf](#)
- [Comprehensive Medical Evaluation and Assessment of Comorbidities: Standards of Care in Diabetes—2024.pdf](#)
- [Diabetes Advocacy: Standards of Care in Diabetes—2024.pdf](#)
- [Diabetes Care in the Hospital: Standards of Care in Diabetes—2024.pdf](#)
- [Diabetes Technology: Standards of Care in Diabetes—2024.pdf](#)
- [Diabetic Retinopathy Preferred Practice Pattern.pdf](#)
- [Diagnosis and Classification of Diabetes: Standards of Care in Diabetes—2024.pdf](#)
- [Disclosures: Standards of Care in Diabetes—2024.pdf](#)
- [Facilitating Positive Health Behaviors and Well-being to Improve Health Outcomes: Standards of Care in Diabetes—2024.pdf](#)
- [Glycemic Goals and Hypoglycemia: Standards of Care in Diabetes—2024.pdf](#)
- [ICO Guidelines for Diabetic Retinopathy.pdf](#)
- [Improving Care and Promoting Health in Populations: Standards of Care in Diabetes—2024.pdf](#)
- [Index.pdf](#)
- [Introduction and Methodology: Standards of Care in Diabetes—2024.pdf](#)
- [Management of Diabetes in Pregnancy: Standards of Care in Diabetes—2024.pdf](#)
- [Obesity and Weight Management for the Prevention and Treatment of Type 2 Diabetes: Standards of Care in Diabetes—2024.pdf](#)
- [Older Adults: Standards of Care in Diabetes—2024.pdf](#)
- [Pharmacologic Approaches to Glycemic Treatment: Standards of Care in Diabetes—2024.pdf](#)
- [Prevention or Delay of Diabetes and Associated Comorbidities: Standards of Care in Diabetes—2024.pdf](#)
- [Retinopathy, Neuropathy, and Foot Care: Standards of Care in Diabetes—2024.pdf](#)
- [Summary of Revisions: Standards of Care in Diabetes—2024.pdf](#)

4.2.4 对话体验与用户交互优化

为了优化 DiabRetina AI 医疗助手的对话体验与用户交互，我们设定了开场白文案与开场白预置问题，并在用户问题建议板块中，设定了用户自定义 Prompt，通过分析用户交互数据持续优化对话策略，确保回答的针对性和准确性，为用户带来高效、易用的智能医疗体验。



- 开场白预置问题
- 糖尿病为什么会损伤视网膜？能通俗解释下吗？
 - 医生说我属于“中度非增殖期”，这是什么意思？
 - 眼底报告上的“棉絮斑”和“硬性渗出”分别代表什么？
 - 输入开场白引导问题

用户问题建议

在智能体回复后，根据 Prompt 提供最多 3 条用户提问建议

用户自定义 Prompt

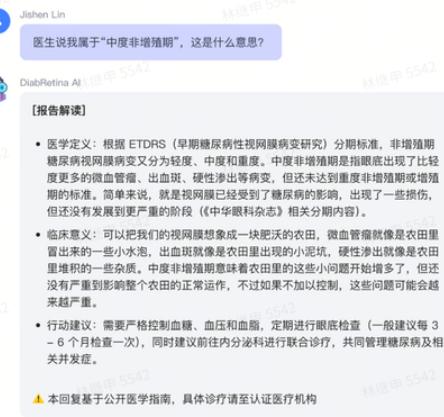
1. 问题应该与你最后一轮的回复紧密相关，可以引发进一步的讨论；
2. 问题不要与上文已经提问或者回答过的内容重复；
3. 每句话只包含一个问题，但也可以不是问句而是一句指令；
4. 推荐你有能力回答的问题。

104/5000

4.3 Agent 在病变诊断智能平台的应用

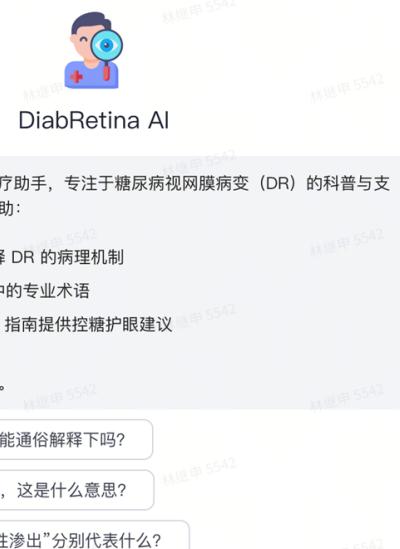
4.3.1 患者角度

糖尿病性视网膜病变诊断智能平台的医疗助手通过基于循证医学的疾病知识科普、诊断流程解释和预防建议，帮助患者提高疾病认知、了解检查流程并采取预防措施，但不替代专业医疗建议。



糖尿病性视网膜病变诊断智能平台的医疗助手通过 AI 技术为患者提供疾病认知、检查流程和健康管理支持。基于国际指南构建的知识库能通俗解释专业术语，比如将微动脉瘤比作“轮胎鼓包”，帮助患者理解风险。当患者描述视力模糊或上传异常报告时，AI 会自动评估风险等级，高血糖或急性症状会触发紧急就医建议。AI 还能生成个性化的检查准备指南，比如散瞳注意事项，减少患者焦虑。所有建议均标注免责声明，确保患者理解需医生最终确认。

林进申 5542 林进申 5542



您好！我是 DiabRetina AI 医疗助手，专注于糖尿病视网膜病变（DR）的科普与支持。以下是我能为您提供的帮助：

1. 疾病解析：用通俗比喻解释 DR 的病理机制
2. 报告解读：说明眼底检查中的专业术语
3. 预防指导：基于 2023 IDF 指南提供控糖护眼建议

请告诉我您想了解的具体内容。

糖尿病为什么会损伤视网膜？能通俗解释下吗？

医生说我属于“中度非增殖期”，这是什么意思？

眼底报告上的“棉絮斑”和“硬性渗出”分别代表什么？

发送消息...

内容由AI生成，无法确保真实准确，仅供参考。



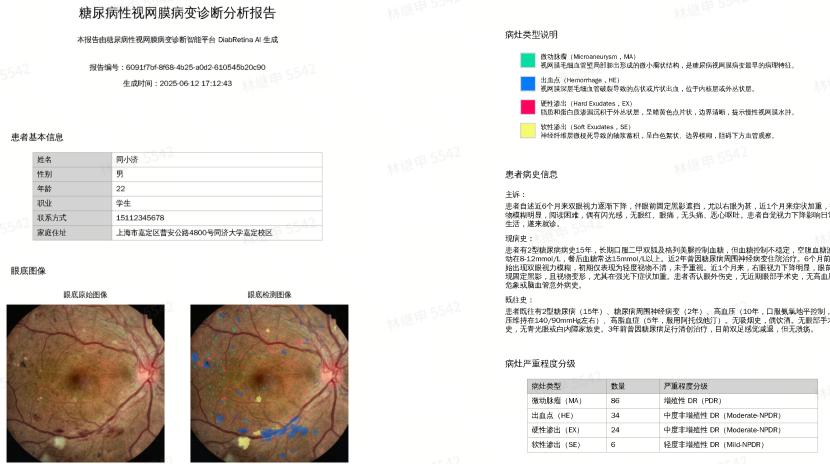
林进申 5542 林进申 5542

4.3.2 医生角度

AI Agent 作为医生的智能助手，能大幅提升临床工作效率。平台自动分析眼底图像，量化统计病灶数量和分布，生成结构化报告，节省医生手动标注时间。AI 会标记疑似高危病例，比如增殖性病变或黄斑水肿，辅助医生快速定位关键问题。

在诊断环节，AI 基于最新指南提供鉴别诊断参考，比如非 DR 出血与视网膜静脉阻塞的差异特征，但最终结论仍由医生确认。治疗建议模块会列出循证方案，如抗 VEGF 注射的适应症和光凝时机，方便医生制定个性化计划。AI 还能自动跟踪患者历史数据，帮助医生评估疗效。

平台支持多端协作，医生可通过电脑或移动设备随时查阅 AI 生成的报告草稿，进行修改确认后一键分享给患者。所有输出均标注参考文献来源，确保信息透明可靠。AI 的介入让医生从繁琐的初筛中解放，更专注于复杂病例决策和患者沟通。



五、病变诊断智能平台

5.1 平台架构与技术选型

5.1.1 前端架构与技术选型

前端采用 React 框架构建，结合 Material-UI (MUI) 组件库实现响应式布局和现代化 UI 设计。通过 React Router 管理页面路由，确保单页应用 (SPA) 的流畅导航体验。状态管理使用 React Hooks (如 `useState`, `useRef`) 处理组件级状态，避免过度依赖全局状态管理库。

核心功能模块包括：

- 图像上传与预览:** 通过 `<input type="file">` 实现眼底图像上传，展示原始图像和 AI 检测结果。
- 表单交互:** 使用 MUI FormControl 和 MDInput 组件收集患者信息、病史和诊断意见，支持多行文本输入和下拉选择（如病灶严重程度分级）。

- **AI 诊断生成**: 通过 `fetch` API 调用后端接口，异步获取 AI 生成的诊断意见，并实时渲染结果。
- **PDF 报告生成**: 触发后端报告生成接口后，通过 `window.open` 在新标签页展示 PDF 报告。

前端代码采用模块化设计，通过 Grid 布局实现自适应排版，确保在桌面和移动设备上的兼容性。样式通过 CSS-in-JS (如 MUI 的 `sx` 属性) 管理，保持代码可维护性。

5.1.2 后端架构与技术选型

后端基于 Flask 微框架开发，提供 RESTful API 接口，主要分为以下模块：

1. 图像处理服务：

- 使用 SimpleITK 和 OpenCV 对上传的眼底图像进行预处理（如尺寸归一化、CLAHE 对比度增强）。
- 调用 PyTorch 加载预训练的 MCA-UNet 模型，实现病灶分割（微动脉瘤、出血点等），并生成彩色掩码叠加图像。

2. AI 诊断逻辑：

- 集成大语言模型（如豆包 -1.5-Pro-32k），通过结构化 Prompt 生成循证医学诊断意见，严格遵循国际糖尿病视网膜病变（DR）分级标准。

3. PDF 报告生成：

- 使用 ReportLab 动态生成 PDF，包含患者信息、病灶统计、诊断意见等，支持中文字体（如文泉驿正黑）渲染。

4. 数据存储与缓存：

- 临时图像和报告文件按 UUID 分类存储于本地目录（如 `diagnostic-report/`），通过轻量级文件系统管理，避免数据库依赖。

后端通过 Flask-CORS 处理跨域请求，并限制文件上传大小（`MAX_CONTENT_LENGTH`）。接口设计遵循 REST 规范，返回 JSON 格式数据（如病灶数量、图像 Base64 编码）或 PDF 文件流。

5.2 检测结果可视化

5.2.1 处理原始图像

由于原始眼底图像的分辨率与模型需求不同，并且普遍存在大量黑边，所以我们进行了长边缩放与中心裁剪操作。最终将图像预处理成分辨率为 640×640 ，并且以核心的眼底图像为中心，去除部分不需要的黑边。

同时为了优化预测的结果，应用了对比度受限的自适应直方图均衡化（CLAHE）方法，通过增强局部对比度来改善图像的视觉效果。

而对于模型预测的需求，将图像预处理后还需要进行 ToTensor 操作：

- 将 PIL 图像或 Numpy 数组转换为 PyTorch 张量。
- 同时自动将像素值从 [0.255] 缩放到 [0.1] 范围。

代码块

```
1 def load_image(image_path):
2     image = sitk.ReadImage(image_path)
3     image_array = sitk.GetArrayFromImage(image)
4     transform_origin = transforms.Compose([
5         Resize(640),
6         CenterCrop(640),
7         ApplyCLAH(green=False)
8     ])
9     sample_origin = {'image': image_array, 'masks': np.zeros((5, 640, 640))}
10    sample_origin = transform_origin(sample_origin)
11    transform = transforms.Compose([
12        Resize(640),
13        CenterCrop(640),
14        ApplyCLAH(green=False),
15        ToTensor(green=False)
16    ])
17    sample = {'image': image_array, 'masks': np.zeros((5, 640, 640))}
18    sample = transform(sample)
19    return sample['image'].unsqueeze(0), sample_origin['image']
```

5.2.2 调用模型预测

调用我们之前训练结果最佳的模型 MCA-UNet，并将 output 的形状规范为 [N, C, H, W]（批次、通道数、长、宽）。

其中，批次会被显示指定为 1，通道数对应预测的结果数 5 个（4 种病灶 + 背景），长宽对应分辨率，均为 640。

代码块

```
1 def predict(model, image_tensor):
2     model.eval()
3     with torch.no_grad():
4         image_tensor = image_tensor.float()
5         output = model(image_tensor)
6         if isinstance(output, tuple):
7             output = output[-1]
8             if output.shape[1] != 5:
9                 output = output.permute(0, 3, 1, 2)
10                output = torch.softmax(output, dim=1)
11    return output
```

5.2.3 生成彩色掩码

模型预测结果是将不同病灶部位的像素位置标记为不同编号，所以人为地将各个病灶对应位置的像素设置成不同的颜色，即可得到预测结果的彩色掩码。

代码块

```
1 def create_colored_mask(prediction):
2     prob = torch.softmax(prediction, dim=1)
3     pred_mask = torch.argmax(prob, dim=1).squeeze().cpu().numpy()
4     h, w = pred_mask.shape
5     colored_mask = np.zeros((h, w, 3), dtype=np.uint8)
6     colored_mask[pred_mask == 1] = LESION_COLORS['EX']
7     colored_mask[pred_mask == 2] = LESION_COLORS['HE']
8     colored_mask[pred_mask == 3] = LESION_COLORS['MA']
9     colored_mask[pred_mask == 4] = LESION_COLORS['SE']
10    return colored_mask, pred_mask
```

其中，我们定义的不同病灶颜色如下：

代码块

```
1 LESION_COLORS = {
2     'EX': (255, 0, 96),
3     'MA': (0, 223, 162),
4     'HE': (0, 121, 255),
5     'SE': (246, 250, 112)
6 }
```

5.2.4 生成检测图像

最终展示的图像由两部分组成：预处理后的原始眼底图像为底，加上生成的掩码图像作为遮罩层附加在原图像上。

在叠加遮罩层时采用加权叠加的方法，使得遮罩层与原图像的融合更加自然，掩码区域的边界也更加清晰。

代码块

```
1 def process_image(image_path, file_uuid):
2     image_tensor, original_img = load_image(image_path)
3     image_tensor = image_tensor.to(device)
4     prediction = predict(model, image_tensor)
5     colored_mask, pred_mask = create_colored_mask(prediction)
6     lesion_counts = count_lesions(pred_mask)
7     if isinstance(original_img, torch.Tensor):
8         original_img = original_img.cpu().numpy()
9     if original_img.shape[0] != 640 or original_img.shape[1] != 640:
```

```

10         original_img = cv2.resize(original_img, (640, 640))
11     if len(original_img.shape) == 2:
12         original_img = np.stack([original_img]*3, axis=-1)
13     elif original_img.shape[2] == 1:
14         original_img = np.repeat(original_img, 3, axis=2)
15     if original_img.max() <= 1.0:
16         original_img = (original_img * 255).astype(np.uint8)
17     else:
18         original_img = original_img.astype(np.uint8)
19     overlay = original_img.copy()
20     mask = (colored_mask != 0).any(axis=2)
21     overlay[mask] = cv2.addWeighted(original_img[mask], 0.5,
22         colored_mask[mask], 0.5, 0)
23     preprocessed_path = os.path.join('preprocessed-image', f'{file_uuid}.jpg')
24     predicted_path = os.path.join('predicted-image', f'{file_uuid}.jpg')
25     PILImage.fromarray(original_img).save(preprocessed_path)
26     PILImage.fromarray(overlay).save(predicted_path)
27     img_byte_arr = io.BytesIO()
28     PILImage.fromarray(original_img).save(img_byte_arr, format='JPEG')
29     preprocessed_img = img_byte_arr.getvalue()
30     img_byte_arr = io.BytesIO()
31     PILImage.fromarray(overlay).save(img_byte_arr, format='JPEG')
32     predicted_img = img_byte_arr.getvalue()
33     return {
34         'uuid': file_uuid,
35         'preprocessed_image': base64.b64encode(preprocessed_img).decode('utf-
36             8'),
37         'predicted_image': base64.b64encode(predicted_img).decode('utf-8'),
38         'lesion_counts': lesion_counts
39     }

```

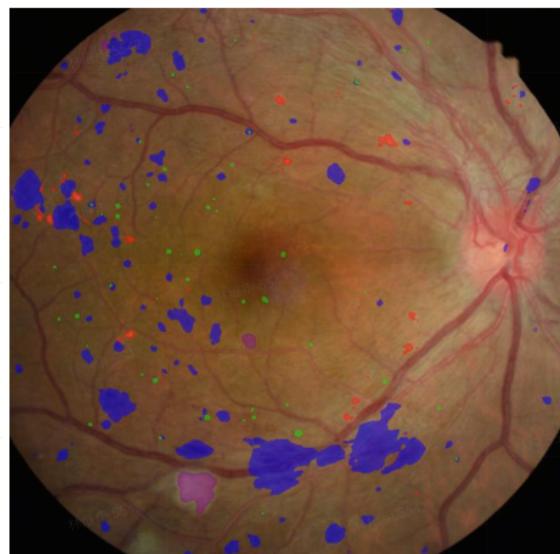
5.2.5 可视化结果展示

未应用 CLAHE 方法：

原始眼底图像



预测结果

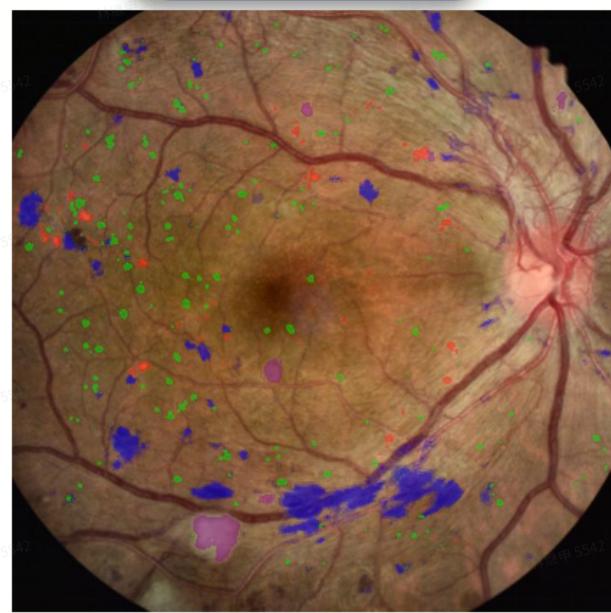


应用 CLAHE 方法：

原始眼底图像



预测结果



通过对比观察可知应用 CLAHE 方法的预测结果更加符合实际，效果得到增强。

5.3 平台功能展示

5.3.1 病变诊断

DiabRetina AI

欢迎使用

病变诊断

历史记录

医疗助手

教育科普

帐号管理

登录

注册

联系我们

眼底原始图像

眼底检测图像

病灶类型与分级标准

糖尿病性视网膜病变涉及四种病灶：

- 微动脉瘤 (Microaneurysm, MA): 视网膜毛细血管壁局部膨出形成的微小瘤状结构，是糖尿病视网膜病变最早的病理特征。
- 出血点 (Hemorrhage, HE): 视网膜深层毛细血管破裂导致的点状或片状出血，位于内核层或外丛状层。
- 硬性渗出 (Hard Exudates, EX): 脂质和蛋白质渗漏沉积于外丛状层，呈暗黄色点片状，边界清晰，提示慢性视网膜水肿。
- 软性渗出 (Soft Exudates, SE): 神经纤维层微梗死导致的轴浆蓄积，呈白色絮状、边界模糊，阻碍下方血管观察。

根据国际临床DR严重程度量表，DR共分为5级：健康、轻度非增殖性DR (Mild non-proliferative DR, Mild-NPDR)、中度非增殖性DR (Moderate non-proliferative DR, Moderate-NPDR)、重度非增殖性DR (Severe non-proliferative DR, Severe-NPDR) 和增殖性DR (Proliferative DR, PDR)。

患者基本信息

姓名：同小济 性别：男 年龄：22 职业：学生

联系方式：15112345678 家庭住址：上海市嘉定区曹安公路4800号同济大学嘉定校区

主诉

患者自述近6个月来双眼视力逐渐下降，伴眼前固定黑影遮挡，尤以右眼为甚，近1个月来症状加重，视物模糊明显，阅读困难，偶有闪光感，无眼红、眼痛，无头痛、恶心呕吐。患者自觉视力下降影响日常生活，遂来就诊。

现病史

患者有2型糖尿病病史15年，长期口服二甲双胍及格列美脲控制血糖，但血糖控制不稳定，空腹血糖波动在8-12mmol/L，餐后血糖常达15mmol/L以上。近2年曾因糖尿病周围神经病变住院治疗，6个月前开始出现双眼视力模糊，初期仅表现为轻度视物不清，未予重视。近1个月来，右眼视力下降明显，眼前出现固定黑影，且视物变形，尤其是在强光下症状加重。患者否认眼外伤史，无近期眼部手术史，无高血压、高血脂、高血糖等既往史。

既往史

患者既往有2型糖尿病（15年）、糖尿病周围神经病变（2年）、高血压（10年，口服氯沙坦钾缓释片，血压维持在140/90mmHg左右）、高脂血症（5年，服用阿托伐他汀）。无吸烟史，偶饮酒。无眼部手术史，无青光眼或白内障家族史。3年前曾因糖尿病足行清创治疗，目前双足感觉减退，但无溃疡。

病灶严重程度分级

病灶	数量	严重程度	严重程度分级												
微动脉瘤 (MA)	27 处	出血点 (HE)	2 处	微动脉瘤 (MA) 严重程度分级	中度非增殖性 DR (Moderate-NPDR)	出血点 (HE) 严重程度分级	轻度非增殖性 DR (Mild-NPDR)	硬性渗出 (EX)	63 处	软性渗出 (SE)	0 处	硬性渗出 (EX) 严重程度分级	增殖性 DR (PDR)	软性渗出 (SE) 严重程度分级	健康 (无软性渗出)
微动脉瘤 (MA) 严重程度分级	中度非增殖性 DR (Moderate-NPDR)	出血点 (HE) 严重程度分级	轻度非增殖性 DR (Mild-NPDR)												
硬性渗出 (EX)	63 处	软性渗出 (SE)	0 处	硬性渗出 (EX) 严重程度分级	增殖性 DR (PDR)	软性渗出 (SE) 严重程度分级	健康 (无软性渗出)								
硬性渗出 (EX) 严重程度分级	增殖性 DR (PDR)	软性渗出 (SE) 严重程度分级	健康 (无软性渗出)												

临床诊断意见

患者长期糖尿病史，血糖控制不佳，结合眼底检查可见视网膜微血管瘤、出血、硬性渗出及部分新生血管形成，光学相干断层扫描（OCT）显示黄斑区视网膜增厚伴囊样水肿，荧光素眼底血管造影（FFA）提示广泛无灌注区及渗漏，符合增殖性糖尿病性视网膜病变（PDR）合并糖尿病性黄斑水肿（DME）的诊断。此外，患者合并高血压及高脂血症，进一步加重了微血管病变的进展，需警惕未来发生玻璃体积血、牵拉性视网膜脱离等严重并发症的风险，因此综合评估后诊断为2型糖尿病伴增殖性视网膜病变（双眼，右眼为主）、糖尿病性黄斑水肿（右眼）、高血压2级（高危组）、高脂血症，需立即干预以挽救残余视力并延缓疾病进展。

治疗方案

针对患者病情，制定个体化综合治疗方案，首先强化血糖管理，改用胰岛素治疗（基础+餐时方案）并严格监测血糖，目标空腹血糖≤7.0mmol/L，餐后血糖≤10.0mmol/L；其次针对视网膜病变，右眼先行玻璃体腔注射抗VEGF药物（雷珠单抗或阿柏西普）每周1次，连续3次后评估疗效改善情况。同时分次完成全视网膜激光光凝（PRP）以封闭无灌注区、抑制新生血管生长；若出现玻璃体积血或牵拉性视网膜脱离，则考虑玻璃体切割术（PPV）。全身管理方面，优化降压方案（目标血压≤130/80mmHg）并维持他汀类降脂治疗（LDL-C≤2.6mmol/L），同时给予α-硫辛酸和甲钴胺改善神经病变。随访计划需严格执行，包括每月眼科复查眼底及OCT、内分泌科调整代谢指标，并加强患者教育，强调饮食控制、适度运动及定期监测的重要性，以最大限度保护视功能并降低全身并发症风险。

AI 辅助诊断意见

患者病史概述：患者同小济，22岁男性学生，自述近6个月双眼视力逐渐下降，伴眼前固定黑影遮挡，近1个月症状加重。有2型糖尿病病史15年，血糖控制不稳定，近2年曾因糖尿病周围神经病变住院治疗。既往还有高血压10年、高脂血症5年，3年前曾因糖尿病足行清创治疗。

病灶程度评估：微动脉瘤27处，严重程度为中度非增殖性DR；视网膜出血2处，严重程度为轻度非增殖性DR；视网膜渗出63处，严重程度为增殖性DR；硬性渗出0处，严重程度为健康。

辅助诊断意见：结合患者长期糖尿病史且血糖控制不佳，以及各病灶情况，AI支持临床诊断意见，诊断为2型糖尿病伴增殖性视网膜病变（双眼，右眼为主）、糖尿病性黄斑水肿（右眼）、高血压2级（高危组）、高脂血症。视网膜渗出严重程度达增殖性DR，需高度警惕严重并发症发生。

治疗方案建议：AI支持现有治疗方案。强化血糖管理，改用胰岛素治疗并严格监测，目标空腹血糖≤7.0mmol/L，餐后血糖≤10.0mmol/L；右眼先行玻璃体腔注射抗VEGF药物，连续3次后评估，同时分次完成全视网膜激光光凝。优化降压方案，目标血压≤130/80mmHg，维持他汀类降脂治疗，LDL-C≤2.6mmol/L，给予α-硫辛酸和甲钴胺改善神经病变。严格执行随访计划，每月眼科复查眼底及OCT、内分泌科调整代谢指标。

注意事项：本回复基于公开医学指南，AI辅助诊断意见仅供参考，不替代专业医疗建议。请结合临床医生评估制定个性化治疗方案。

5.3.2 历史记录

The screenshot shows a list of 10 users with their basic information and report details:

患者姓名 / 性别	年龄	职业	联系方式 / 家庭住址	报告生成时间	诊断分析报告
韩雪 女	51	银行职员	13388990022 哈尔滨市南岗区西大直街100号	2025-06-12 11:54:31	查看报告
宋建国 男	67	退休干部	18876543210 昆明市五华区东风西路6号	2025-06-12 11:53:19	查看报告
叶婷 女	43	设计师	17700998877 宁波市海曙区中山路18号	2025-06-12 11:52:13	查看报告
高志 男	49	快递员	19988776655 郑州市金水区花园路25号	2025-06-12 11:50:52	查看报告
何秀兰 女	56	退休工人	14455667788 苏州市姑苏区干将东路8号	2025-06-12 11:49:44	查看报告
徐强 男	61	电工	15566778899 天津市河西区友谊路30号	2025-06-12 11:48:26	查看报告
朱琳 女	54	幼儿园教师	16622334455 大连市中山区人民路15号	2025-06-12 11:47:10	查看报告
马建军 男	63	保安	17788990011 青岛市市南区香港中路7号	2025-06-12 11:44:27	查看报告
黄英 女	47	销售经理	18899887766 厦门市思明区湖滨南路10号	2025-06-12 11:42:25	查看报告
周明 男	59	厨师	13344556677 长沙市岳麓区麓山南路20号	2025-06-12 11:40:51	查看报告

5.3.3 医疗助手

The screenshot shows a medical consultation interface with a message from the AI and a list of recommended topics:

您好！我是 DiabRetina AI 医疗助手，专注于糖尿病视网膜病变（DR）的科普与支持。以下是我能为您提供的帮助：

- 疾病解析：用通俗比喻解释 DR 的病理机制
- 报告解读：说明眼底检查中的专业术语
- 预防指导：基于 2023 IDF 指南提供控糖护眼建议

请告诉我您想了解的具体内容。

推荐对话 (4/4):

- 糖尿病为什么会损伤视网膜？能通俗解释下吗？
- 医生说我属于“中度非增殖期”，这是什么意思？
- 眼底报告上的“棉絮斑”和“硬性渗出”分别代表什么？

社区 C

期待你的分享！
你的评价对这个智能体很重要

5.3.4 教育科普

教育科普 1 | 什么是糖尿病性视网膜病变

糖尿病性视网膜病变 (DR) 是糖尿病最常见的眼部并发症，也是成年人致盲的主要原因之一。它是由长期高血糖损害视网膜微小血管所导致的慢性进行性疾病。在早期阶段，高血糖会削弱血管壁，导致微血管瘤形成和微小出血点。随着病情发展，血管可能发生渗漏，引起视网膜水肿，甚至阻塞，造成局部缺血。在更严重的阶段，视网膜因缺血而刺激异常新生血管生长，这些新生血管脆弱易破，可能导致玻璃体出血、纤维增生，最终引发牵引性视网膜脱离，造成不可逆的视力损害。

值得注意的是，糖尿病性视网膜病变在早期通常没有明显症状，这使得定期眼底检查变得尤为重要。许多患者直到出现明显视力下降时才就医，此时病情往往已进入中晚期。该病变通常是双眼发病，但严重程度可能不对称。除了典型的视网膜病变外，约三分之一的患者还会合并糖尿病性黄斑水肿 (DME)，这是导致中心视力下降的最常见原因。虽然糖尿病性视网膜病变可能最终导致失明，但通过早期发现、规范治疗和良好的血糖控制，绝大多数患者可以保留有用视力。

教育科普 2 | 糖尿病性视网膜病变的分期与症状

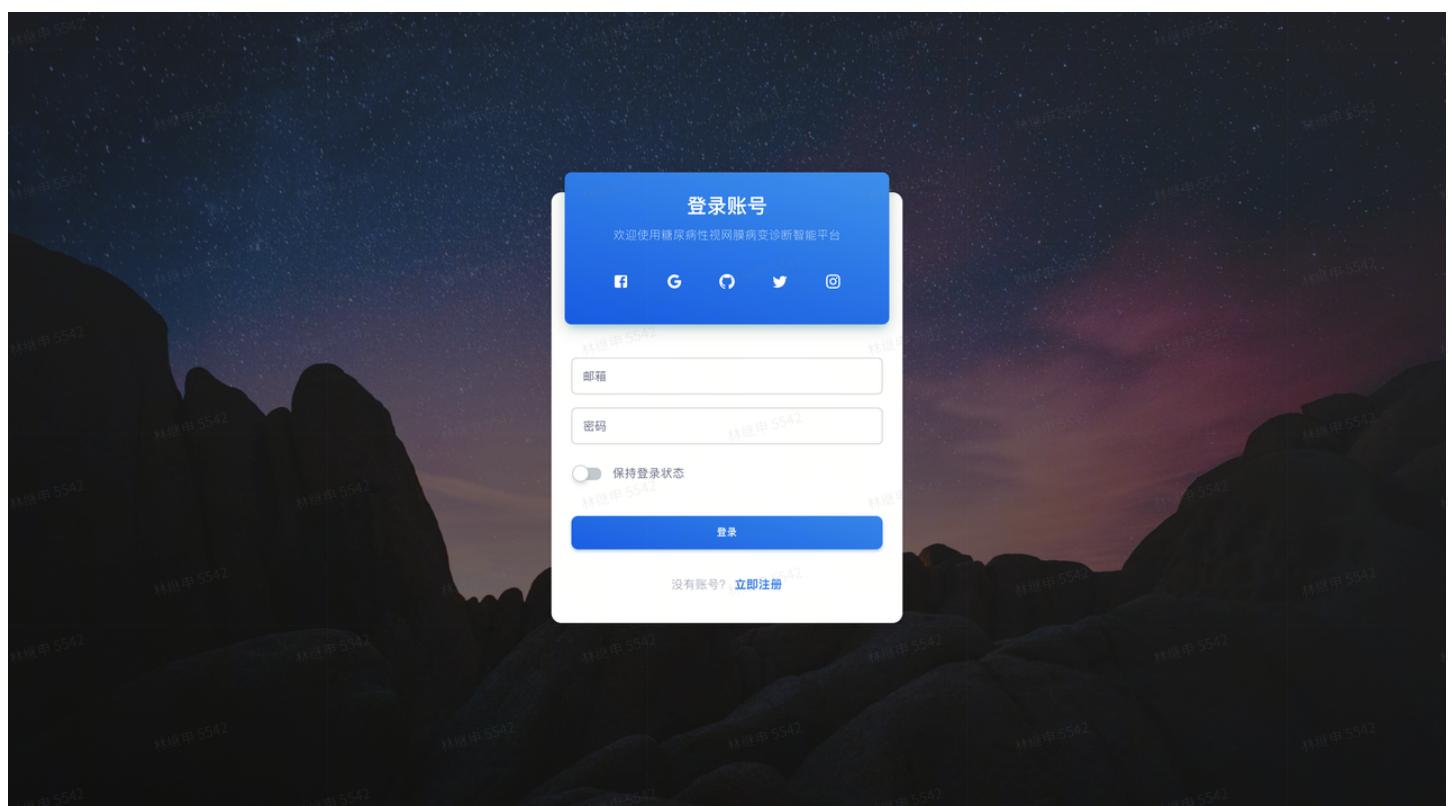
糖尿病性视网膜病变的临床分期主要分为非增殖期 (NPDR) 和增殖期 (PDR)。非增殖期是疾病的早期阶段，其特征是视网膜微血管的损伤和渗漏，包括微血管瘤、视网膜内出血、硬性渗出和静脉串珠样改变等。在这个阶段，患者可能没有任何症状，或者仅表现为轻微的视物模糊或阅读困难。随着病情进展，视网膜缺血加重，可能出现棉絮斑 (软性渗出)，这标志着中度至重度非增殖期的转变。

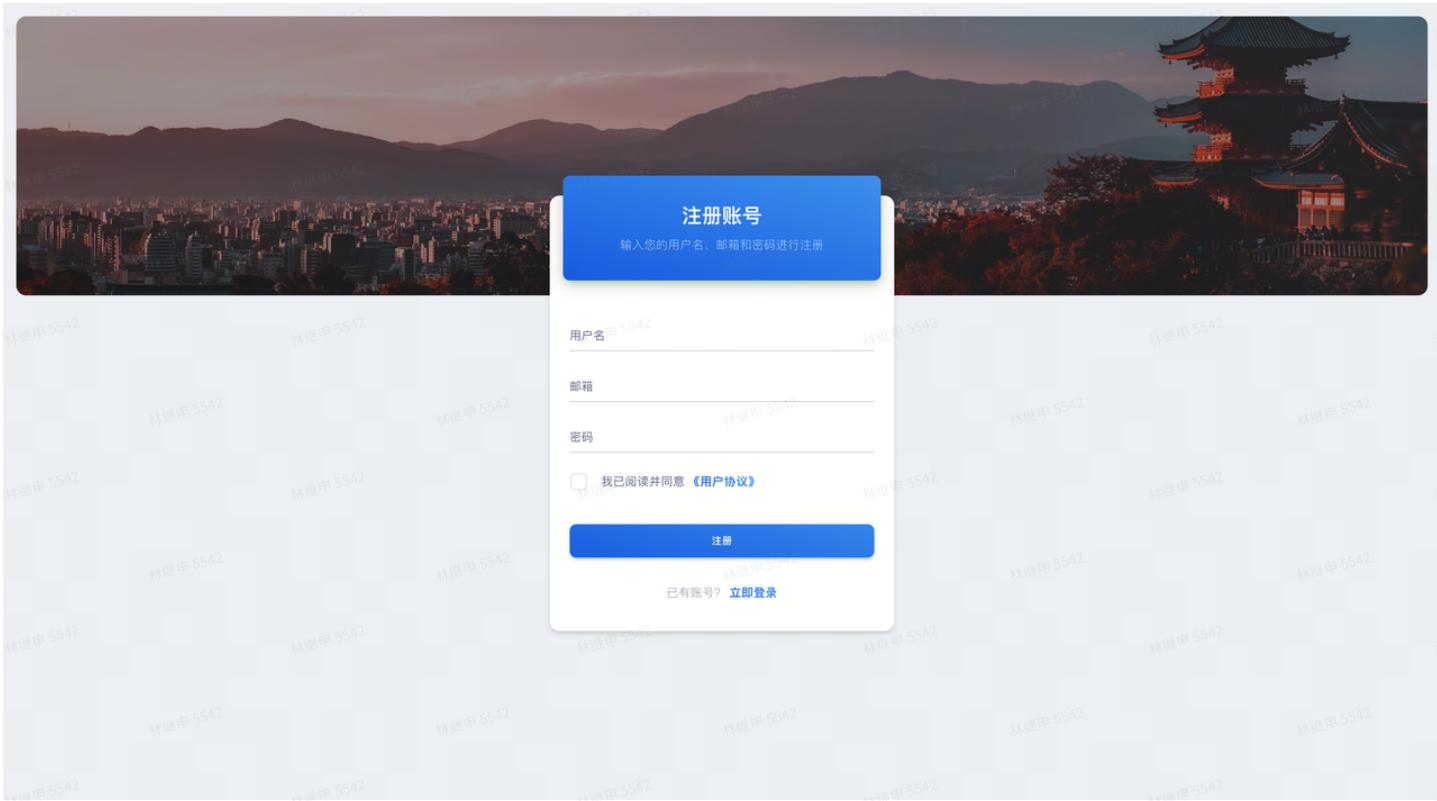
当病变进入增殖期，意味着视网膜已经出现大面积缺血，刺激产生异常的新生血管。这些新生血管不仅容易破裂导致玻璃体出血，还会引起纤维组织增生，最终可能导致牵引性视网膜脱离。患者在这个阶段通常会突然出现飞蚊症、视野缺损或急剧的视力下降。特别值得警惕的是，有些患者在视力尚好的情况下，眼底病变可能已经相当严重，这就是为什么糖尿病患者必须定期进行专业的眼底检查，而不能仅凭自我感觉来判断病情。

教育科普 3 | 高危人群与筛查建议

所有类型的糖尿病患者都有发生视网膜病变的风险，但风险程度与糖尿病病程、血糖控制情况密切相关。1型糖尿病患者通常在确诊5年后开始出现视网膜病变，而2型糖尿病患者可能在确诊时就已经存在视网膜病变。其他高危因素包括长期血糖控制不佳、高血压、高血脂、妊娠、肾病以及吸烟等。特别值得注意的是，青春期和妊娠期由于激素变化，可能加速视网膜病变的进展。

5.3.5 账号管理





六、项目总结

本小组开发了一个高精度、可扩展的 AI 辅助诊断系统，实现了糖尿病性视网膜病变 (DR) 的自动化筛查（微动脉瘤 MA、出血点 HE、硬性渗出 EX、软性渗出 SE），系统的检测算法满足了高性能和高实时性要求。

6.1 算法选型与数据集处理

1. 基于医学图像分割特性，综合对比 8 种语义分割方法（如 UNet、DenseASPP 等），选定 UNet 作为基准模型，因为其跳跃连接结构在小数据集上表现优异，mIoU 达 80.25%，且多尺度特征捕获能力能够适配病灶多样性。
2. 分析 5 类主流 DR 数据集，采用眼球半径归一化 + CLAHE 增强预处理方法，构建了 3014 张的糖尿病性视网膜病变标准化数据集。通过对 EX、SE、MA 三类病灶进行适当程度的数据增强处理，使关键病灶样本量提升 2-3 倍，以平衡数据集，缓解类别不平衡的问题。

6.2 算法优化与模型对比

1. 系统对比了 UNet、UNet++、U²Net 和 MCA-UNet 四种语义分割模型，最终选定 MCA-UNet 作为核心算法，其多尺度交叉跳跃连接和注意力机制显著提升了小病灶（如微动脉瘤 MA）的检测精度。
2. 通过 150 轮训练周期优化，模型在数据集上的平均 Dice 系数达 0.78，AUPR 提升 35%，优于基线 UNet 模型 (Dice 0.65)。

6.3 多任务学习与效率提升

1. 创新性地采用了 5 通道掩码输出（背景 + 4 类病灶），单次推理即可完成 4 类病灶分割，相比独立模型减少约 75% 计算资源消耗。
2. 动态数据增强（CLAHE 对比度优化 + 几何变换）使模型在低对比度眼底图像上的泛化能力提升约 20%。

6.4 AI 智能体（DiabRetina AI）

1. 集成豆包 -1.5-Pro-32k 大模型，基于 23 篇国际 DR 领域医学指南构建循证医学知识库，在医疗问答评测中以全满分表现最优。同时智能体支持风险词触发急诊建议（如“急性眼痛”自动提示就医），并具备多模态能力。
2. 提供结构化输出（定义 → 比喻解释 → 行动建议），使患者更容易理解复杂医学概念。

6.5 糖尿病性视网膜病变诊断智能平台

1. 接入 AI Agent，根据病灶严重程度分级评估与医生临床诊断，AI 辅助生成个性化的诊断意见与治疗方案，提高治疗的科学性和针对性。
2. 平台支持一键生成 DR 诊断分析报告，内含 AI 辅助诊断意见，并提供历史报告查询功能，支持随时随地查看、下载报告。

七、参考文献

- [1] 崔丽娜.深度学习驱动的图像分割：U-Net、SegNet 和 DenseASPP 的比较研究[J].现代计算机,2024,30(13):36-44.
- [2] 黄晓鸣,何富运,唐晓虎,等.U-Net 及其变体在医学图像分割中的应用研究综述[J].中国生物医学工程学报,2022,41(05):567-576.
- [3] Patra, P., & Singh, T. (2022, December 14-16). Diabetic retinopathy detection using an improved ResNet 50-InceptionV3 and hybrid DiabRetNet structures. In 2022 OITS International Conference on Information Technology (OCIT). IEEE.
- [4] Bosale, A. A. (2024). Detection and classification of diabetic retinopathy using deep learning algorithms for segmentation to facilitate referral recommendation for test and treatment prediction. arXiv preprint arXiv:2401.02759.
- [5] Lalithadevi, B., & Krishnaveni, S. (2024). Diabetic retinopathy detection and severity classification using optimized deep learning with explainable AI technique. Multimedia Tools and Applications, 1-65.
- [6] S. Mishra, S. Hanchatee and Z. Saquib, “Diabetic Retinopathy Detection using Deep Learning”, 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), 2020, pp. 515-520, doi: 10.1109/ICSTCEE49637.2020.9277506.
- [7] Saraa Hosseinzadeh Kausani, and Payman Hosseinzadeh Kasani, Raza Khazaeinezhad, Michal J. Wesolowski et al. “Diabetic Retinopathy Clasification Using a Modified Xception

Architecture” , 2019 IEEE International Symposium on the Signal Processing and Information of Technology (ISSPIT), 2019.

- [8] Anuj Jain, Arnav Jalui, Jahanvi Jasani, Yash Lahoti, Ruhina Karani. “Deep Learning for Detection and Severity Classification of Diabetic Retinopathy” , 2019 1st International Conference on the Innovations in Information and Communication Technology (ICIICT), 2019.
- [9] Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., & van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4), 501–509. <https://doi.org/10.1109/TMI.2004.825627>.
- [10] Zhou, Y., Wang, B., Huang, L., Cui, S., & Shao, L. (2020). A benchmark for studying diabetic retinopathy: Segmentation, grading, and transferability. *IEEE Transactions on Medical Imaging*. <https://doi.org/10.1109/TMI.2020.3000000>.
- [11] Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., & Meriaudeau, F. (2018). Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research. *Data*, 3(3), 25.
- [12] Li, T., Gao, Y., Wang, K., Guo, S., Liu, H., & Kang, H. (2019). Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. *Information Sciences*. Advance online publication.
- [13] Mao, J., Ma, X., & Zhang, R. (2023). TJDR: A high-quality diabetic retinopathy pixel-level annotation dataset. arXiv:2312.15389v1 [eess.IV]. <https://doi.org/10.48550/arXiv.2312.15389>.
- [14] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: Redesigning skip connections to exploit multiscale features in image segmentation," *IEEE Trans. Med. Imaging*, vol. 39, no. 6, pp. 1856-1867, Jun. 2020, doi: 10.1109/TMI.2019.2959609.
- [15] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U2-Net: Going deeper with nested U-structure for salient object detection," *Pattern Recognit.*, vol. 106, pp. 1-12, Nov. 2020, Art no. 107404, doi: 10.1016/j.patcog.2020.107404.
- [16] H. Wang, P. Cao, J. Yang, and O. Zaiane, "MCA-UNet: multi-scale cross co-attentional U-Net for automatic medical image segmentation," *Health Inf. Sci. Syst.*, vol. 11, no. 1, pp. 1-10, Jan. 2023, Art no. 10, doi: 10.1007/s13755-023-00214-8.
- [17] Wang, Y., Hou, Q., Cao, P., Yang, J., & R. Zaiane, O. (2024). Lesion-aware knowledge distillation for diabetic retinopathy lesion segmentation. *Applied Intelligence*, 54(2), 1937-1956.