

基于医生知识引导的眼底多疾病诊断模型研究

侯伊林 11912636

黄晨露 11910935

2022 年 6 月 8 日

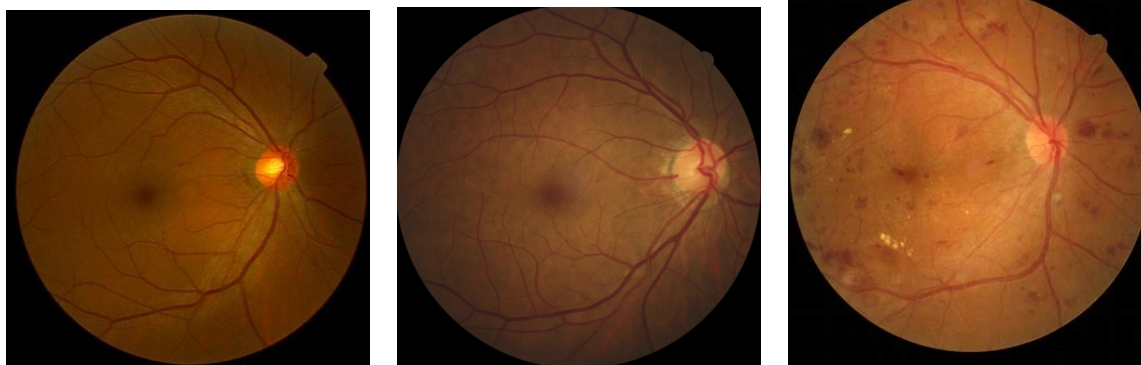
1 研究背景

1.1 糖尿病性视网膜病变

1.1.1 定义

糖尿病性视网膜病变 (DR) 是慢性进行性糖尿病导致的视网膜微血管渗漏和阻塞从而引起一系列的眼底病变, 如微动脉瘤、硬性渗出、棉絮斑、新生血管、玻璃体增殖、黄斑水肿甚至视网膜脱离。

在临床上, 通常根据是否出现新生血管, 将 DR 分为两大类: 增殖型糖尿病性视网膜病变和非增殖型糖尿病性视网膜病变。其中, 非增殖型糖尿病性视网膜病变依据患病程度又分为一级、二级、和三级。这里我们明确本研究提到的早期 DR 为非增殖性视网膜病变的一级, 其表现为眼底出现少数微动脉瘤 (MA)。在眼底彩照中, 微动脉瘤表现为暗红色的孤立小圆点, 它们的直径通常小于 125 μ m。



(a) 正常

(b) 早期

(c) 中期

图 1: 彩色眼底图像

1.2 临床诊断需求

早期 DR 的诊断对医生专业水平要求较高, 需要医生经过专业培训, 经验积累才能做出较为准确的诊断。导致在大规模筛查中医疗资源紧张。最后, 由于早期 DR 仅表现微动脉瘤, 临床上漏诊率较高。因此, 利用计算机图像处理技术对眼底图像进行检测, 进行大规模筛查眼底病是一种有价值的方案。

1.3 青光眼

青光眼 (glaucoma) 是一组以视乳头萎缩及凹陷、视野缺损及视力下降为共同特征的疾病，病理性眼压增高、视神经供血不足是其发病的原发危险因素，视神经对压力损害的耐受性也与青光眼的发生和发展有关。在房水循环途径中任何一环发生阻碍，均可导致眼压升高而引起的病理改变，但也有部分患者呈现正常眼压青光眼。青光眼是导致人类失明的三大致盲眼病之一，它的疾病进程极为缓慢，在一段很长的时间后慢慢使得病人失去视力，因此，青光眼的早期诊断十分重要。

青光眼的眼底表现为视盘中带红色神经组织的丧失，和视盘凹陷的增加，即中央白色的部位增大，以及视盘边缘有出血现象。

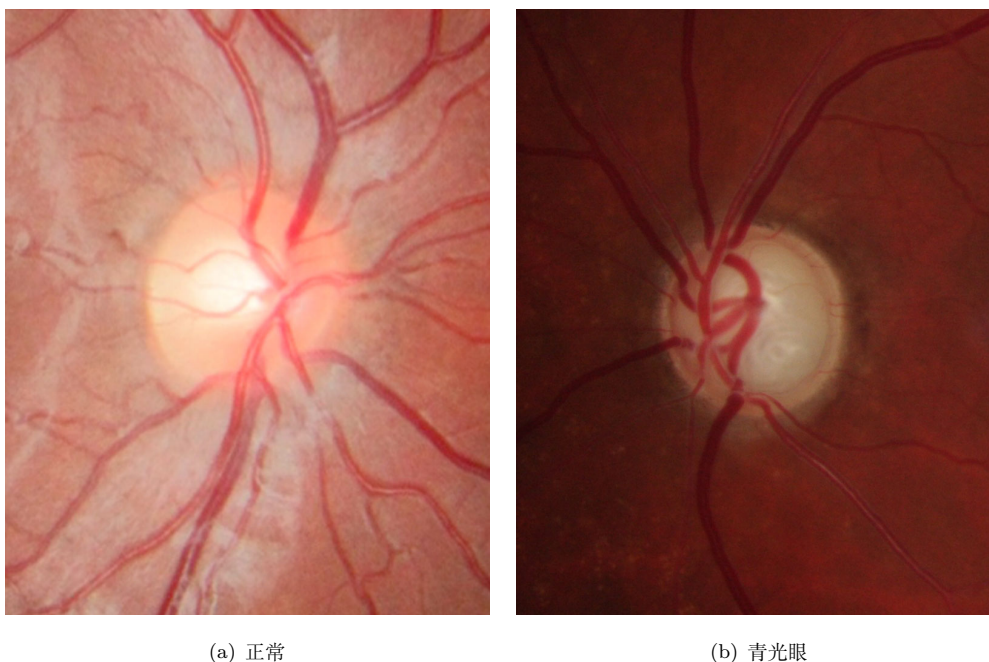


图 2: 正常人与青光眼患者视盘对比

1.4 老年黄斑变性

老年黄斑变性 (AMD) 是一种影响眼睛后部视网膜 (黄斑) 中央区域的疾病。黄斑变性通常是高龄退化的自然结果，随着年龄增加，视网膜组织退化，变薄，引起黄斑功能下降。老年病变常常开始只影响一只眼睛，且早期症状不明显。随着病情的发展，会表现出中心视力模糊，有时候会出现黑斑或黑点，错误判断距离和高度，识别颜色困难等症状。AMD 是世界上第三大最常见的致盲原因，也是人口老龄化的高收入国家的主要致盲原因。全球约有 5% 的失明是由 AMD 引起的。据估计，2020 年全球有 1.96 亿人患有 AMD，2040 年将增加到 2.88 亿人。

AMD 的眼底表现为早期可见黄斑区色素紊乱，中心凹反光不清，有散在的玻璃疣。发病晚期，黄斑部可有金属样反光，视网膜色素上皮萎缩呈地图状，可见囊样变性。

1.5 研究目的

本研究通过眼动追踪技术获取医生诊断时的注意力，将医生注意力作为临床先验知识，设计一种知识注入型的神经网络模型，在眼底图像上实现对不同眼底疾病的精确诊断。

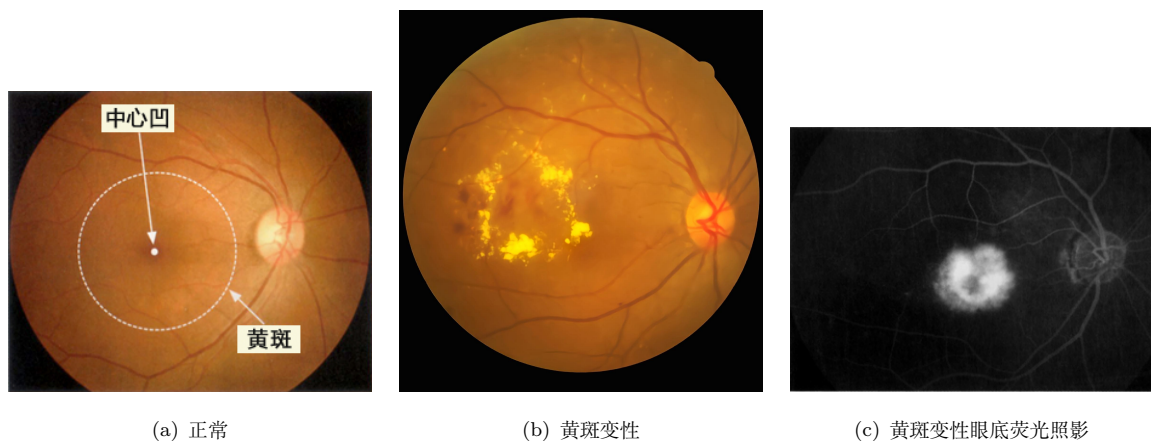


图 3: 老年黄斑变性眼底图像

2 前期工作

2.1 医生注意力数据集制作

本项目组在组内医生的辅助下，挑选了正常和早期 DR 的眼底彩照。我们将使用 Tobii 屏幕式眼动仪，记录医生诊断时的眼动注意力，以热点图的形式叠加在原图像并导出。

2.2 眼动仪实验流程

在本研究中，每一张热点图都由专业医生进行诊断，组内成员全程在旁辅助，使用眼动仪并对机器进行操作，进行对数据的导入导出。

1. 设置眼动仪参数
2. 将挑选后的原始眼底彩照按设置参数导入眼动仪。
3. 人机相对位置标准化，受试者视线垂直于眼动仪显示屏，面部与屏幕相距 55cm。
4. 眼动仪校准，受试者配合视线紧盯屏幕上的运动圆点。
5. 开始实验。实验中保证受试者不被干扰，眼动仪半径一米以内无人员，防止眼动仪测量出现偏差。
6. 试验结束后，对可视化图像进行参数设定（如下），导出可视化图像（热点图）。

热点半径	30px
背景	透明
透明度	100%

2.3 视觉误差处理

在实验过程中，我们发现生成的热点图存在视觉误差，即医生的注视区域和真实病灶区域存在微小偏差。热点图的生成原理是利用高斯曲线对注视点周围区域进行拟合，我们可对该高斯分布关联区域的大小进行设定。将 $r=30\text{px}$ 转变为 60px ，使得热点区域可完全覆盖病灶区域。

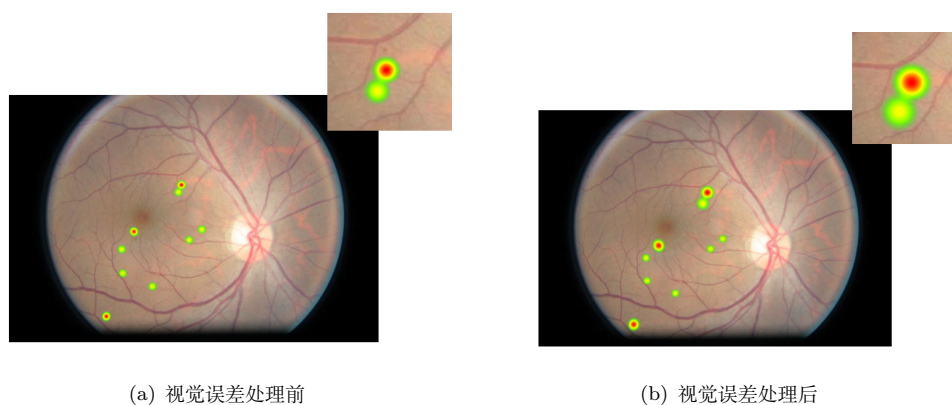


图 4: 视觉误差处理前后对比

2.4 数据标准化处理

2.4.1 收集 DR 数据

在原有的组内采集的数据集的基础上，用公共数据集（kaggle 数据集）进行了补充，现有 206 张早期 DR 眼底图像，506 张正常眼底图像。

2.4.2 去除低质量眼底图像

主要去除过曝、过暗和被污染的眼底图像。

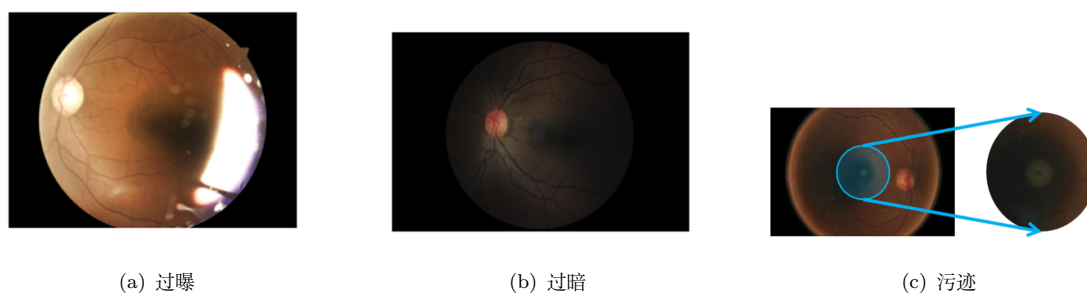


图 5: 彩色眼底图像

2.4.3 统一尺寸

以眼球半径为基准，将所有图像调整至同一尺寸。对于每一张眼底图像，因为图像基本满足上下对称，所以我们取图像正中间一行，记录有颜色的像素个数，除以 2 即可得到半径。根据这个半径对图像进行缩放。

2.4.4 去除黑边

与 2.1.3 的方法类似，找出每一张眼底图像黑边的范围，裁剪去除即可。

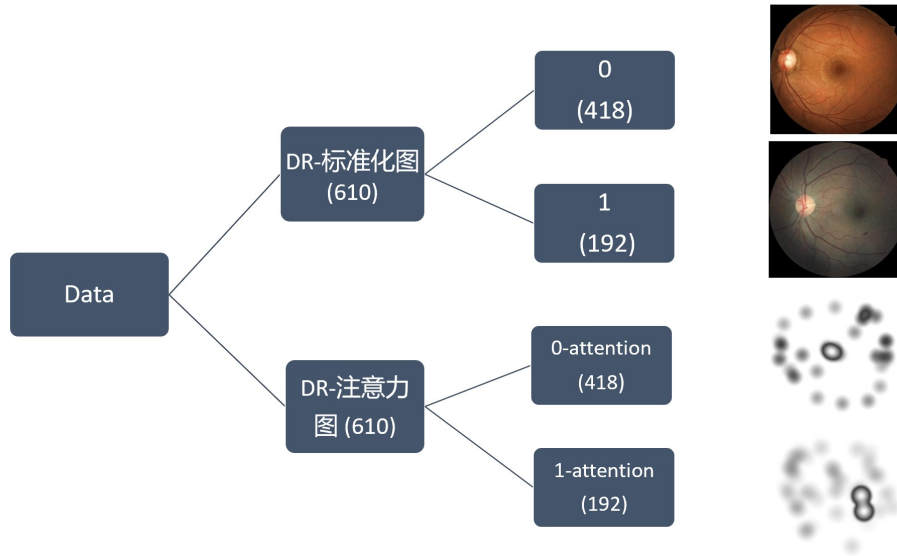


图 6: 完整数据集架构

3 使用 U-Net 框架建模

3.1 U-Net 架构

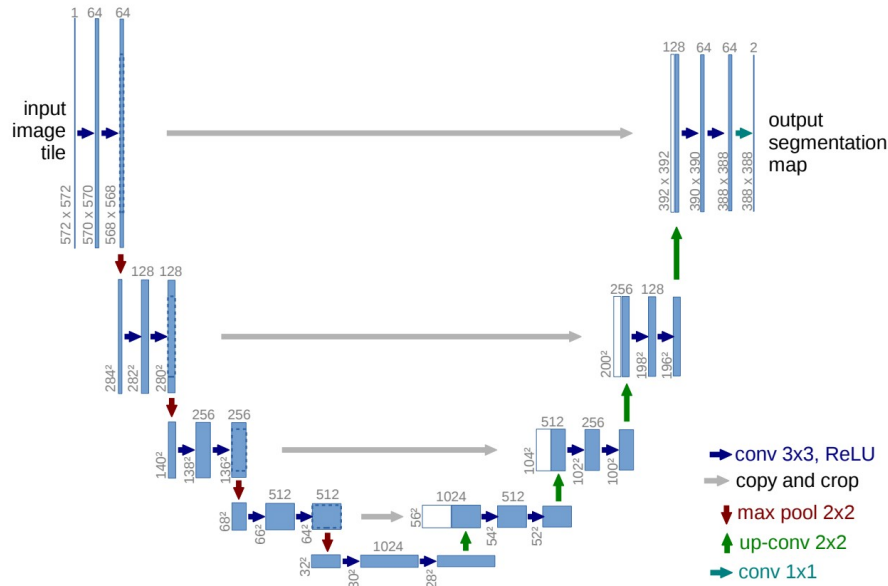


图 7: U-Net 网络结构

U-Net 是一个 U 型对称的网络，常用来解决图像分割问题。U-Net 借鉴了 FCN 网络，其网络结构包括两个对称部分：前面一部分网络与普通卷积网络相同，使用了 3x3 的卷积和池化下采样，能够抓住图像中的上下文信息（也即像素间的关系）；后面部分网络则是与前面基本对称，使用的是 3x3 卷积和上采样，以达到输出图像分割的目的。此外，网络中还用到了特征融合，将前面部分下采样网络的特征与后面上采样部分的特征进行了融合以获得更准确的上下文信息，达到更好的分割

效果。

医疗图像的数量相对于自然图像要少，获得有临床意义的标签的难度也更大，所以用于医疗图像的网络模型机构不宜太复杂，否则易造成过拟合，影响模型效能。U-Net 的模型参数可以通过调成模型层数、每层的通道数来缩小，与其他模型比做到轻量级，十分适用于医疗图像的训练。

3.2 结合眼动图的 U-Net 模型

在本实验中，我们提出了 U-Net 架构框架中的分类问题，并在瓶颈层增加了一个多类分类块。U-Net 的编码臂和瓶颈臂可以是任何标准的预训练分类器，不需要全连接层。这两个组合将作为分类器的特征编码器。该网络的 CNN 解码器部分运行反卷积层来预测静态眼球注视热图。该方法的优点是可以联合训练输出注视静态注视热图，并进行多分类预测。对于测试集的眼底彩照图像，输入一张眼底彩照原图到该网络，可以输出预测的诊断结果，并生成与疾病相关的最重要位置的的概率热图。

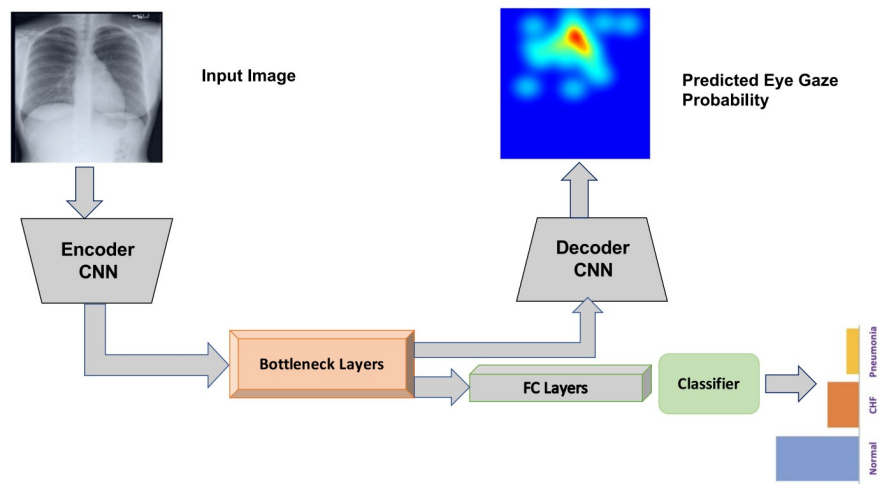


图 8: 结合眼动图的 U-Net 结构模型

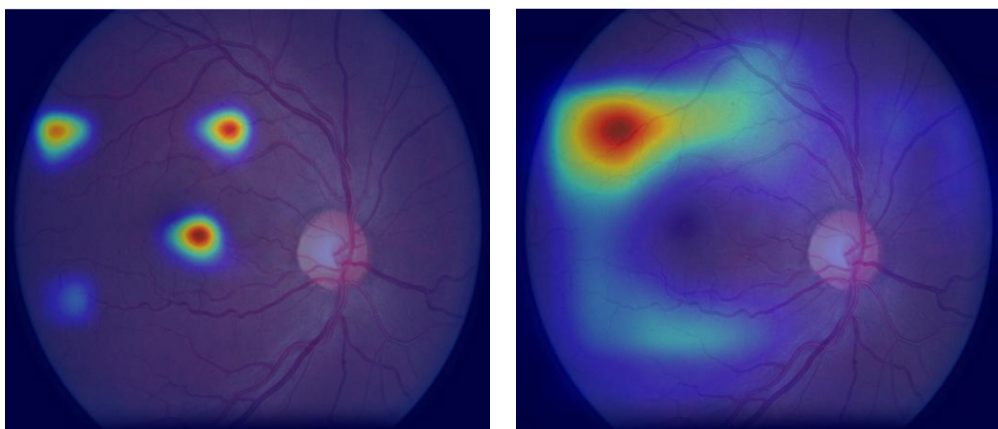
我们使用一个预先训练的 EfcientNet-b0 作为编码器和瓶颈层。分类头为自适应平均池，然后是增肥层、dropout 层和线性输出层。该解码器 CNN 由三个卷积和上采样层组成。损失函数是由分类损失函数和分割损失函数加权得到，两者都采用了二元交叉熵损耗函数。

$$Loss = w_1 * L_{classification} + w_2 * L_{segmentation}$$

3.3 实验结果

我们用训练好的模型导出关于训练集的 CAM 图与医生注意力图进行对比。图 9 中左图是医生注意力热图，图中有明显红色区域的三个部分都有病灶（微血管瘤）的分布，表明这几个区域在医生诊断该图片时受到了更多的光顾。右图是用训练好的模型导出关于训练集的 CAM 图，如图所示，在左上角一块区域热力图颜色明显更深，证明在训练网络过程中该区域的特征对训练网络的影响最大。这两幅图对比可以得出，两个热力图的共同点是都关注到了左上方的部分区域，说明我们的网络结构确实学习到了部分医生的注意力，从而帮助分类器分得正确的结果。

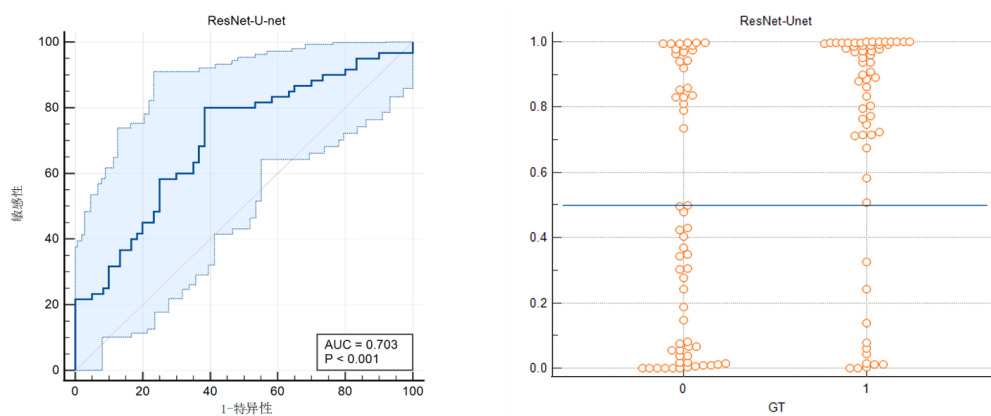
经过调参后，模型的准确率最高达到了 0.703. 以 0.4982 为阈值，大于阈值时判定为 1，小于阈值判定为 0。分类的敏感性最高可以达到 0.80，特异性最高可以达到 0.617。



(a) 医生注意力

(b) 注意力预测图

图 9: 正常人与青光眼患者视盘对比



(a) ROC 曲线

(b) 敏感性和特异性分析

图 10: 模型分类准确率、敏感性、特异性

4 评价标准

本实验的卷积神经网络为二分类模型。我们定义诊断为 DR 为正样本，诊断为正常为负样本。我们利用以下几个指标对模型性能进行评判。

$$\begin{aligned} accuracy &= \frac{tp+tn}{tp+fp+fn+tn} \\ precision &= \frac{tp}{tp+fp} \\ recall &= \frac{tp}{tp+fn} \\ specificity &= \frac{fp}{fp+tn} \\ f-score &= (1 + \beta^2) \frac{precision*recall}{\beta^2*precision+recall} \end{aligned}$$

5 研究计划

5.1 增加补充实验

增加一组实验：用普通人注意力图注入模型。

5.1.1 实验背景

普通人的注意力图的特点：相比专业医生，普通人只能了解到眼底图像和疾病的基本知识，不具备诊断疾病的能力，所以观察眼底图像获得的注意力图会有较大噪声。

5.1.2 结果预测

实验分组：(a) 不加入注意力图；(b) 加入医生注意力；(c) 加入普通人注意力。

预想的实验结果：(b) 要优于 (a) 和 (c)。且由于 (c) 有较大干扰，(a) 可能优于 (c) 或者和 (c) 准确度相似。

5.1.3 实验进度

已采集 DR 疾病的普通人注意力图共 1200 张，DR0 和 DR1 各 600 张。采集注意力图所用的眼底图像与采集医生注意力图一样，其他采集流程也保持一致。

5.1.4 实验计划

(1) 补充更多医生和普通人的注意力图，注入模型进行实验。

(2) 改进原有的医生注意力数据集：因为医生的注意力也会存在噪声干扰。可以利用普通人观察眼底图的特点，减弱这些干扰。

(3) 增加模型的实际作用：可以通过区分注意力图像，来判断一个人是否已经具备了区分 DR0 和 DR1 的能力，即是否可以被称为“医生”。

5.2 研究进度和未来规划

阶段	研究进度
第一次答辩	文献、方法调研; 完成 DR 注意力数据集整理; 选择模型并设计实验思路;
第二次答辩	将整理好的数据集用于当前模型进行分类; 整理收集青光眼、黄斑变性等眼底数据集;
第三次答辩	继续对数据集做注意力标注; 对当前模型进行改进和调试; 由普通人进行注意力实验（组内成员），对比医生诊断图片和普通人 诊断图片的注意力;

参考文献

- [1] 刘卫, 张勇进. 糖尿病性视网膜病变 [J]. 国外医学. 眼科学分册, 2005(05):74-79.
- [2] Thomas Walter, Pascale Massin, Ali Erginay, Richard Ordonez, Clotilde Jeulin, Jean-Claude Klein, Automatic detection of microaneurysms in color fundus images, Medical Image Analysis, Volume 11, Issue 6, 2007, Pages 555-566, ISSN 1361-8415.
- [3] Maximilian Ilse and Jakub M. Tomczak and Max Welling, Attention-based Deep Multiple Instance Learning, 2018, arXiv:1802.04712.
- [4] Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, 234–241 (Springer, 2015).
- [5] Karargyris, A., Kashyap, S., Lourentzou, I. et al. Creation and validation of a chest X-ray dataset with eye-tracking and report dictation for AI development. Sci Data 8, 92 (2021).