

目前来看研究工作中主要存在以下问题需要深入研究：

### 1. 数据集需要进一步完善

目前现有的眼底图像数据集中存在一系列成像噪声，如黑色空间、低对比度、光线不足等问题，在后期研究工作中应该尝试使用更强大的图像处理技术来减少噪声，提高图像质量。数据集中的图片存在的黑色区域采取裁切预处理，虽然可以提高处理效率，但是也会造成一定程度信息丢失，特别是这些区域可能包含有效的解剖结构，在裁切操作中需要确保不会丢失特征提取关键信息。另一方面数据集中图片还存在 DR、PDR 类别不平衡现象，采用数据增强技术可能会引入偏见，需要进一步研究类别加权的策略，需要谨慎选择和调整数据增强方法，以确保增强后的图像仍然保持与原始图像相似的特征。

### 2. 模型组件的选择

前期模型构建中我们尝试了 InceptionV3、InceptionResNetV2 和 Xception 三种特征提取器的不同组合，后期研究工作需要交叉验证等方法评估不同特征提取器组合的效果，并针对本项目进行微调，更好地提取病变特征。通过实验进一步研究激活函数的选择，考虑其他非线性激活函数的效果，例如 Parametric ReLU (PReLU) 或 Exponential Linear Unit (ELU)，尝试在深度神经网络中使用不同激活函数的组合，例如在不同层使用不同的激活函数，以提高模型的表达能力。

### 3. 利用云算力进行模型训练、优化

前期模型训练、优化仅利用个人计算机进行小规模训练，耗费时间较长、效果较差。在后期研究过程中，利用云算力进行模型训练和优化可以提供更大的计算资源，加速实验和模型迭代的过程。利用云算力的分布式训练能力，将模型训练任务分配到多个计算节点上，以加速训练过程。分布式训练可以显著减少模型训练的时间，特别是对于深度学习模型。利用云端的自动调参工具，例如云服务商提供的超参数优化服务，自动搜索最佳的模型超参数，以提高模型性能。这可以节省大量的实验时间和计算资源。将实验所需的数据、模型参数和训练日志等存储在云端，方便管理和共享。云存储服务提供了高效的数据传输和存储方案，确保模型训练所需的数据能够迅速地被访问和利用。利用云端构建系统和持续集成服务，确保实验的可复现性和持续集成。这有助于快速检测模型改进的效果，并及时进行调整。虽然使用云服务器进行模型训练会增加成本开支，但使用云服

务商提供的成本管理工具，对模型训练和实验的成本进行监控和优化。及时调整资源配置，以在满足实验需求的同时降低成本。

下阶段工作计划旨在解决上述问题，探索更强大的图像处理技术，如去噪算法、对比度增强等，以降低成像噪声，提高眼底图像质量。尝试图像修复或填充算法，确保在裁切预处理中不丢失关键解剖结构信息。系统研究并实施类别加权的策略，解决 DR 和 PDR 类别不平衡问题，并避免数据增强引入的偏见。通过交叉验证等方法，全面评估不同特征提取器组合的性能，选择最适合眼底图像分类任务的模型组件。利用云计算平台进行大规模模型训练，以缩短实验周期，提高模型性能。使用分布式训练和自动调参服务优化实验流程。制定详细的实验计划，包括不同参数配置、数据处理方法和模型组件的组合，以系统评估系统性能。在云计算平台上执行实验计划，确保实验的高效执行和及时反馈。对实验结果进行详细分析，深入了解模型在不同条件下的表现，并挖掘可能的改进空间。