
第十届全国大学生生物医学工程创新设计竞赛

基于多模态眼前节 OCT 影像的 智能体辅助诊断系统及应用



作品 ID 号: 5787

参赛学生类型: 本科生

参加赛道: 医工 AI 赋能赛

组别: 自选项目组

2025 年 6 月

内容完整性自查表

- 说明：
- 1. 命题组参赛作品请根据命题文件规定或要求的指标填写；
 - 2. 自选组参赛作品请根据实际参赛作品完成情况填写。

完整性类别	任务或技术指标名称	完成效果	呈现方式（报告中的章节、页码，或者测试报告，或实物、视频等）
参赛作品任务或技术指标要求	1.构建完整科学的 AS-OCT 多模态数据集	完成	P5， P8
	2. AS-OCT 图像中轴光柱处理	完成	P17
	3. 实现眼前节结构边界的精细分割	完成	P17
	4. 定位巩膜突	完成	P18
	5. 测定眼前节生物参数	完成	P18
	6. 实现疾病的多标签分类	完成	P19
	7. 搭建 AS-Dr 多步骤推理智能体	完成	P20
其他			

摘 要

眼前节光学相干断层扫描（AS-OCT）作为关键眼科诊断工具，面临影像报告周转时间长、观察者间一致性低、AI 辅诊模型分割分类精度不足及架构孤立等核心问题。本研究旨在构建多模态融合的智能诊断体系，通过技术突破与应用创新提升诊断准确性与临床实用性。

首先构建涵盖多视图影像、疾病标签、分割掩码、定位标签、生物参数及临床文本的眼前节多模态数据集以供模型训练。采用 AOT-GAN 消除图像中轴光柱，基于 TransRes-Unet 实现眼前节亚像素级分割，结合巩膜突定位技术精准测量生物参数。分类阶段引入视觉 Transformer 融合多模态特征，集成基于迪利克雷分布的不确定性估计与动态阈值策略降低误诊风险。最终构建 AS-Dr 多步骤推理智能体，以大语言模型驱动 ReAct 决策循环，嵌入前述先进模块，搭建可解释的交互式辅诊平台。

本研究整合多模态数据处理、生物结构精准量化与智能决策架构三大关键创新，首次实现智能体技术在 AS-OCT 全链条诊断中的应用，为优化临床工作流程及基层诊疗可及性提供有效技术支撑。

关键词：眼前节光学相干断层扫描，多模态，智能体，图像分割，图像分类

目 录

摘要.....	I
1 作品概述.....	1
1.1 背景及意义.....	1
1.2 研究基础.....	1
1.2.1 眼前节光学相干断层扫描技术.....	1
1.2.2 眼前节参数.....	2
1.2.3 眼疾智能辅助诊断.....	2
1.2.4 大语言模型与智能体.....	3
1.3 需求分析.....	3
1.4 研究目标.....	3
2 作品方案设计及实现.....	5
2.1 作品实现的技术路线概述.....	5
2.1.1 数据采集及预处理方案概述.....	5
2.1.2 图像处理方案概述.....	6
2.1.3 As-Dr 智能体构建方案概述.....	6
2.2 技术方案对比.....	8
2.2.1 分割算法对比.....	8
2.3 作品方案设计及实现过程.....	8
2.3.1 数据预处理.....	8
2.3.2 AS-OCT 图像中轴光柱处理.....	9
2.3.3 眼前节结构分割.....	10
2.3.4 巩膜突定位及生物参数测量.....	11
2.3.5 眼科疾病分类.....	14
2.3.6 As-Dr 智能体构建.....	16
3 作品测试方案及测试结果.....	17
3.1 测试方案及结果.....	17
3.1.1 AS-OCT 图像中轴光柱处理结果.....	17
3.1.2 眼前节结构分割结果.....	17
3.1.3 巩膜突定位及生物参数测量结果.....	18
3.1.4 眼科疾病分类结果.....	19
3.1.5 As-Dr 智能体搭建结果.....	20

3.2 技术可行性分析及创新性说明..... 23

 3.2.1 数据采集科学性 23

 3.2.2 AS-OCT 智能体应用的可行性..... 23

 3.2.3 创新性说明 23

4 总结..... 25

 4.1 作品总结..... 25

 4.2 未来展望..... 25

5 参考文献..... 26

1 作品概述

1.1 背景及意义

眼前节光学相干断层扫描（AS-OCT）影像作为评估角膜、前房、虹膜、晶状体等眼前节结构的关键非侵入性技术，其高分辨率成像能力对于青光眼、白内障等重大致盲性眼病的早期筛查及术前评估至关重要。临床实践中，AS-OCT 分析高度依赖医师经验，导致报告周转时间长、观察者间一致性低、对细微结构和早期病变的识别易受主观因素影响，诊疗效率低而并误诊风险高。当前，医学影像诊断作为 AI 应用的核心领域已显著提升效率和辅助决策。眼科学高度依赖影像技术，是 AI 赋能的前沿阵地。

尽管人工智能在医学影像领域已取得一定进展，但现有应用于 AS-OCT 影像的 AI 辅诊工具仍存在局限性，多为针对孤立任务开发，缺乏全流程整合能力，难以处理复杂的临床查询，也未能有效整合影像特征与患者临床信息、生物参数等多种模态的数据。这些瓶颈使得现有 AI 工具难以融入临床工作流程并提供可解释、高置信度的决策支持。

本研究正是在这一背景下应运而生，旨在构建一个多模态融合、具备多步骤推理能力的智能辅诊系统。通过整合多模态数据与多项图像处理技术，并创新性地应用智能体技术于 AS-OCT 影像诊断全链条。基于 ReAct 框架与大语言模型 API 调用，实现从影像分析到临床建议的全流程智能化。研究成果有望推动眼科临床变革，缓解我国医疗资源紧张，符合国家“健康中国”战略导向，具有重要的社会价值和广阔的应用前景。

1.2 研究基础

1.2.1 眼前节光学相干断层扫描技术

眼前节光学相干断层扫描，即 AS-OCT，是一种非侵入性的眼部成像技术，通过高分辨率扫描实时获取眼前节结构图像，其可确保患者的舒适性和安全性，而不会暴露于辐射，目前已被广泛应用于眼科疾病诊断。该技术自 1994 年由 Izatt 等提出后持续发展[1]，早期采用 830nm 波长，后优化为 1310nm 波长以增强巩膜穿透能力并实现每秒 8 帧实时成像[2]。AS-OCT 可清晰可视化角膜、前房、虹膜及晶状体等结构，为白内障类型识别（通过晶状体形态、密度、位置信息）[3]、角膜疾病诊断（角膜厚度测量）[4]、青光眼筛查（前房角与虹膜形态分析）以及 LASIK 手术规划提供关键依据[5][6][7]。

1.2.2 眼前节参数

为实现眼前节结构的量化评估与辅助诊断，需要测量的眼前节参数包括前房区域面积（ACA）、前房深度（ACD）、前房宽度（ACW）[8]、中央角膜厚度（CCT）、晶状体厚度（LT）及晶状体隆起（LV）[9]六项指标，如表 2.2 所示。

表 1.1 眼前节各生物参数介绍

序号	指标名称	英文缩写	测量定义	临床意义与应用
1	前房区域面积	ACA	角膜后表面与虹膜前表面之间区域的二维投影面积	评估闭角风险和前房结构变化
2	前房深度	ACD	角膜中央后表面到晶状体前囊膜的最短垂直距离	前房结构评估，青光眼风险分层
3	前房宽度	ACW	巩膜突水平线上的双巩膜突间距离	房角开放度评估
4	中央角膜厚度	CCT	角膜中央区域前后表面间的垂直厚度	眼压测量准确性，手术规划
5	晶状体厚度	LT	晶状体前囊膜至后囊膜之间的最短距离	晶状体老化状态，屈光状态分析
6	晶状体隆起	LV	晶状体前表面至双巩膜突连线的垂直距离	评估闭角型青光眼风险

1.2.3 眼疾智能辅助诊断

(A) 白内障智能诊断方法

深度学习在 AS-OCT 白内障诊断中取得显著进展，各种深度神经网络模型已被用于处理白内障分类任务，Imran 等人[10]于 2021 年提出了卷积循环神经网络（CRNN）模型，用于眼底图像的白内障严重程度的分级，达到了 97.4 的准确率；Xiao 等人[11]于 2021 年通过提取 AS-OCT 图像中的核性区域并提出了基于多任务三维卷积神经网络（MT-CNN）的方法，用于 AS-OCT 图像的核性白内障检测，达到了 94.3%的分类准确率；Zhang 等人[12]于 2022 年提出了基于多区域融合注意力的模型，可用于 AS-OCT 图像的核性白内障检测，达到了 87.8%的分类准确率。

(B) 闭角型青光眼智能诊断方法

前人的研究证明了深度学习在 AS-OCT 闭角检测中存在优势。Fu 等[13]于 2017 年开发 CNN 模型检测 AS-OCT 的闭角征象（AUROC=0.96），性能优于传统定量方法，并后续通过多临床相关区域组合优化算法。Xu 等人[14][15]开发了深度学习算法，不仅可以在任何象限中检测是否有闭角，还可以在 AS-OCT 图像上检测是

否有 PACD。

1.2.4 医学影像分割

在图像分割研究中，Qiu 等人提出在 Segment Anything Mode 中嵌入可学习提示，通过冻结 SAM 主干网络，仅训练轻量级提示层与任务特定解码器，适配眼科多模态图像的分割任务，其实现了跨模态的零样本泛化，支持一次性训练以降低数据以来，但对于微小细节的分割精度不足，且依赖高质量的图像输入[23]；Roy 等人设计全卷积编码器-解码器架构（ReLayNet），引入边界敏感加权交叉熵损失函数，强化视网膜层间边界的像素权重，解决 OCT 图像中组织分层与病变区域的类别不平衡问题，在视网膜分层任务中实现了 85.7% 的 Dice 精度，为青光眼等疾病量化提供了可靠工具，但该方法仅支持单一 OCT 模态[24]。Wu 等人提出 MedSegDiff-V2，在扩散概率模型中集成 Transformer 架构与创新的锚定条件机制 U-SA 和频谱空间变换器 SS-Former，通过双条件融合策略解决医学图像分割中语义特征与噪声嵌入的域间差异问题。该方法在 20 个跨模态医学分割任务中实现 sota 性能，其神经带通滤波器（NBP-Filter）有效降低 62% 的计算负载并加速收敛，支持少样本条件下的稳定分割[25]。

1.2.5 眼科疾病分类

在分类领域，RETFound 模型作为新兴的视网膜图像基础模型，在视网膜疾病诊断和全身疾病预测方面展现出优异性能[18]。然而，目前其应用主要集中在视网膜 OCT 图像领域，尚未广泛拓展到 AS-OCT 图像分析，且还未尝试多模态融合，为我们的研究提供了创新空间。不确定性估计在机器学习领域受到越来越多的关注，一些研究尝试将其应用于医学图像诊断[19]。但目前在 AS-OCT 图像诊断中，融合不确定性估计和阈值策略来减少误诊率的研究相对较少，相关方法和策略仍有待进一步探索和完善。

1.2.6 医学大语言模型与智能体

近年来，以 GPT-4、Gemini 等为代表的大模型推动了医疗人工智能的发展。MedGLM 是清华等团队提出的多模态医学大模型，原生支持图文结合输入，具备中文生成优势，是当前具备落地性与研究潜力的先进开源框架。MedRAX 将多模态推理能力与基于结构化工具的决策相结合，允许实时 CXR 解释，而不会产生不必要的计算开销。通过平衡计算效率与领域专业化，并在合并新工具时消除再培训的需要，MedRAX 为不断变化的临床需求提供了更大的适应[16]。

1.3 需求分析

中国眼科医疗面临严峻挑战：全国眼科医生数量仅 4.48 万名（每 5 万人口 1.6 名），远低于发达国家，而青光眼患者达 2100 万例（占全球 25%），白内障患者超

2 亿例。此外，眼科医生培养周期长叠加老龄化使年龄相关眼病患者激增，传统诊疗模式难以为继。技术革新成为破局关键。当前眼科医疗市场规模已达 76 亿元，正以 11.3% 年复合增长率向 315 亿元扩张，其中器械领域增长显著。然而当前诊疗高度依赖医生经验，引入 AI 辅诊系统可自动化分析影像，显著提升效率并缩短人才培养周期。国家“健康中国”战略为技术创新提供支撑，在降低人力成本的同时提升诊疗公平性，兼具经济社会价值。



图 1.1 AS-OCT 智能辅诊需求与价值

1.4 研究目标

本项目针对 AS-OCT 临床诊断中存在的影像报告周转效率低、观察者间诊断一致性差等核心瓶颈，提出系统性解决方案，旨在通过多模态数据融合与智能算法创新，构建具备医学专业性与临床可用性的 AI 辅诊平台。具体研究目标分述如下：

（1）多模态 AS-OCT 数据集构建：构建标准化的 AS-OCT 多模态数据库，整合影像、标签、生物参数及临床文本，为模型微调提供全面训练基础。

（2）AS-OCT 影像中轴光柱处理：针对部分存在的中轴光柱，研发基于生成对抗网络的图像处理算法，为后续分割分类提供高质量输入。

（3）眼前节结构精细分割：设计基于 TransRes-Unet 的分割模型，实现角膜、前房角、晶状体等结构的亚毫米级边界分割。

（4）关键生物参数自动化测定：开发基于图像处理技术的巩膜突定位算法，结合分割掩码建立几何测量模型，实现核心参数的自动化计算。

（5）多标签分类与不确定性量化：构建基于视觉 Transformer（ViT）的多任务分类框架，实现青光眼、白内障等疾病的同步诊断。

（6）AS-Dr 多步骤推理智能体：以大语言模型（LLM）为核心构建 ReAct 循环推理智能体，整合上述图像分析模块与临床文本记录，形成具备多步骤推理能力的 AS-Dr 系统。

2 作品方案设计及实现

2.1 作品实现的技术路线概述

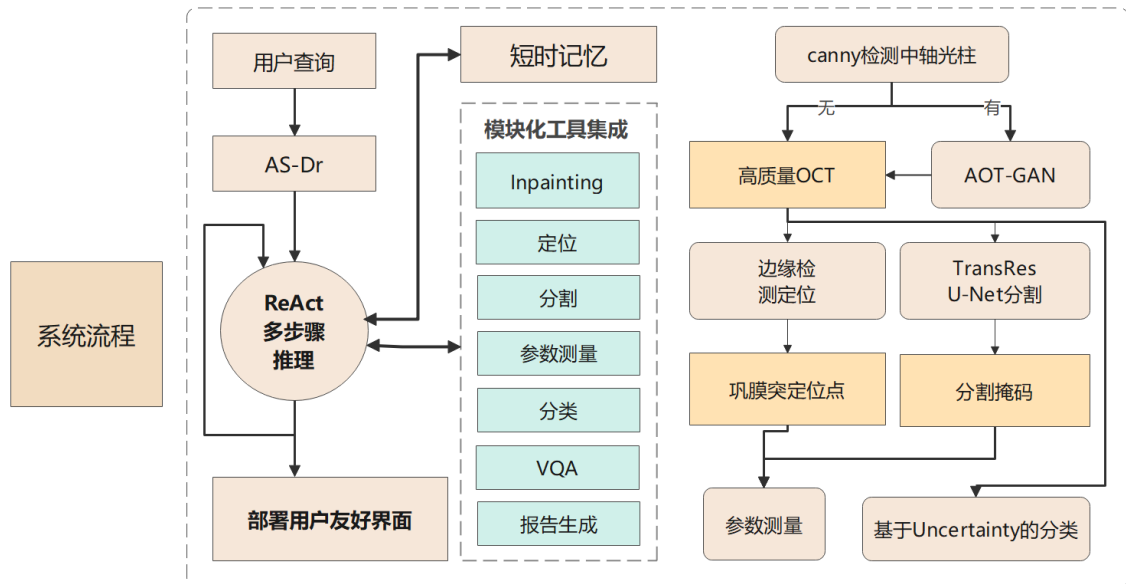


图 2.1 作品实现技术路线

2.1.1 数据采集及预处理方案概述

(A) AS-OCT 采集设备介绍

我们采用 CASIA2 设备采集数据，这是 TOMEY 公司开发的用于眼前节组织断层成像的高精度 OCT 设备，采用扫频源光学相干断层扫描技术，利用 1310nm 波长的光源，以非接触无创式的方式进行三维成像，0.3 秒可获取眼前节 16mm×13mm 范围的立体图像，涵盖角膜到晶状体后表面及部分玻璃体，能减少患者不适感和等待时间，还配备触摸屏操作界面，使用 LCD 显示器，支持不同视角和软件，用户通过简单选择测试协议即可进行日常测量。



图 2.2 AS-OCT 采集设备

(B) 图像标注软件介绍

Labelme 是由麻省理工学院计算机科学和人工智能实验室（CSAIL）研发的图像标注工具，它使用 Python 和 PyQt 编写，支持多边形、圆形等多种标注方式以及 JSON、COCO 等多种数据格式输出。Labelme 广泛应用于语义分割、实例分割、姿态估计等任务，是计算机视觉研究的重要辅助工具。

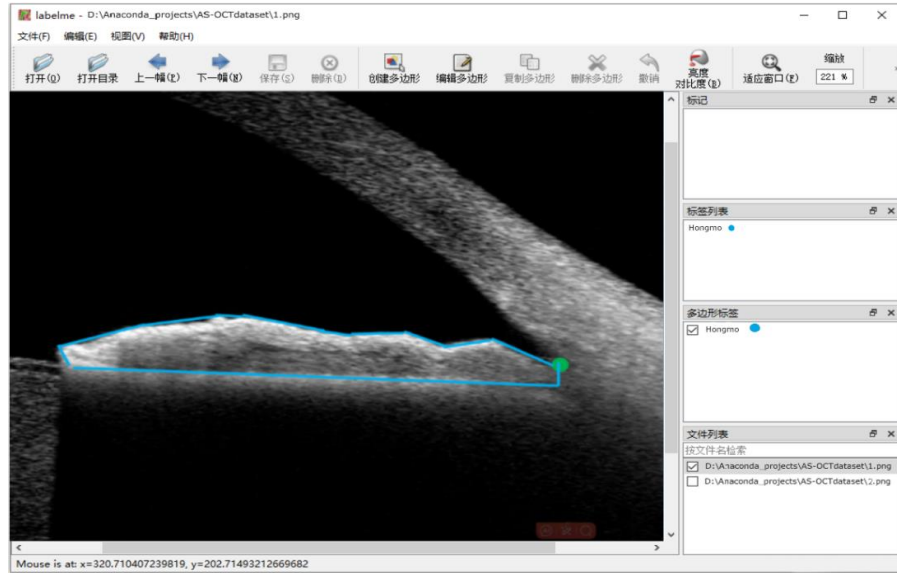


图 2.3 Labelme 软件界面展示

(C) 数据采集方案概述

在专业眼科医生的技术指导下，项目成员依托 CASIA2 专业 OCT 成像设备，运用 labelme 图像标注软件，对全部数据集开展分割掩码的标注工作。

2.1.2 图像处理方案概述

本研究的图像处理方案旨在解决 AS-OCT 影像诊断中的核心挑战，并为后续的智能诊断体系奠定坚实基础。首先，采用了 AOT-GAN 模型以消除 AS-OCT 影像中常见的中心轴光柱。随后，我们基于 TransRes-Unet 模型实现了高精度的分割，这对于后续的定量分析至关重要。在此基础上，结合巩膜突定位技术，我们能够对生物参数进行精准测量，这些参数在眼科疾病诊断和治疗效果评估中具有重要的临床意义。整个图像处理流程旨在为后续的分类阶段和智能体诊断提供高质量、标准化的输入数据，从而提升整个诊断体系的准确性和可靠性。

2.1.3 As-Dr 智能体构建方案概述

(A) ReAct 框架

ReAct (Reasoning and Acting) 框架是一种用于构建智能体的架构，它将推理和行动相结合，使智能体能够通过迭代的观察、思考和行动循环来解决复杂任务。在 As-Dr 中，ReAct 框架被用来处理用户查询，通过整合短期记忆和专门的医疗工具，将复杂的医疗查询分解为多个顺序分析步骤。智能体在处理用户查询时，会经

历观察（分析当前状态和查询）、思考（确定所需动作）和行动（执行相关工具并整合发现）的迭代周期。这一过程持续进行，直到智能体生成响应或要求用户提供良好的额外输入[17]。

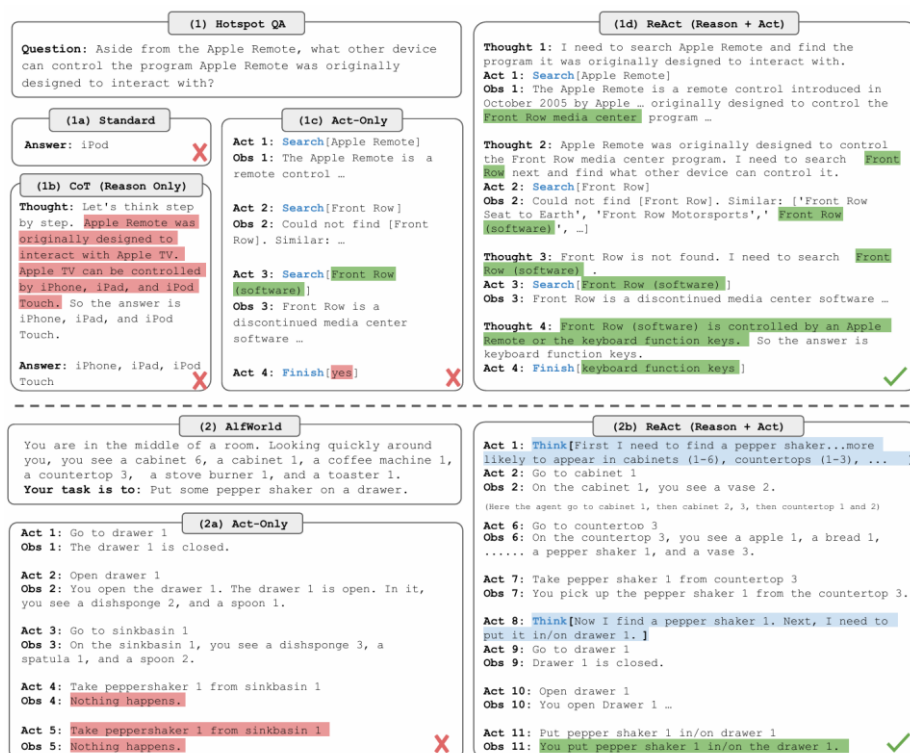


图 2.4 ReAct 智能体框架

(B) Gradio 界面开发

Gradio 是一个开源 Python 库，允许用户快速创建和共享用于机器学习模型的网页界面。它使得即使没有前端开发经验的开发者也能轻松地将模型封装成具有交互功能的 Web 应用。Gradio 界面支持上传各种标准格式的放射学图像，包括 DICOM，并维护一个交互式的聊天会话，以便进行自然的多轮交互。此外，它还提供对工具执行的透明度，通过跟踪和显示中间输出。在 As-Dr 中，Gradio 被用来构建生产就绪的界面，以便在临床环境中无缝部署。

(C) As-Dr 智能体构建方案

本研究构建了面向 AS-OCT 影像的多步骤推理智能体 AS-Dr，其核心采用 ReAct 循环驱动决策。我们利用大型语言模型（LLM）解析用户查询并结合短期记忆状态，生成包含待调用工具序列及参数的结构化推理计划；随后调用集成的 TransRes-UNet 分割模型、ViT 多标签分类器、巩膜突定位模块等先进算法执行子任务，并整合工具输出更新记忆。基于 LangChain 架构，本研究实现了模块化工具链的动态调度与历史交互的短期记忆缓存，保障多轮上下文连贯性。最终，我们输出融合可解释推理路径与置信度评估的临床可验证决策支持。

2.2 技术方案对比

2.2.1 分割算法对比

在 AS-OCT 影像分割模型的选择上，我们对多种先进模型进行了对比分析，由表 2.1 可见（加粗表示最优模型，下划线表示次优模型，下同）：

表 2.1 多种分割算法的分割效果

模型	晶状体		前房		角膜		虹膜	
	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU
<u>SAMmed-2D</u>	<u>0.9816</u>	<u>0.9639</u>	<u>0.9790</u>	<u>0.9590</u>	<u>0.9364</u>	<u>0.8808</u>	<u>0.8759</u>	<u>0.7803</u>
TransResU-Net	0.9916	0.9834	0.9880	0.9764	0.9539	0.9122	0.9241	0.8594
Focal-Unet	0.886	0.791	0.895	0.812	0.853	0.748	0.778	0.632
SAM-2-Unet	0.922	0.900	0.932	0.908	0.846	0.789	0.862	0.783

TransResUNet 在 AS-OCT 影像分割任务中表现出最佳的综合效果，其 Dice 系数和 IoU 指标均优于其他模型，其设计原理使其能够高效地捕捉 AS-OCT 影像中精细眼部结构信息并进行精确分割。SAMmed-2D 基于 Segment Anything Model 医学影像版本[20]，虽然该模型在整体性能上也相当出色，但在实际应用中它需要提供病灶的方框作为提示输入才能进行有效分割，而我们任务更倾向于直接进行端到端的分割。考虑到 TransResUNet 在无需额外提示的情况下仍能保持卓越的分割精度，并且更符合我们数据集的特点和实际应用场景的需求，因此最终选择 TransResUNet 进行分割。

2.3 作品方案设计及实现过程

2.3.1 数据预处理

(A) 基于中轴光柱有无的数据筛选

初期使用的 AS-OCT 数据集是基于有无疾病以及疾病种类划分的，每一种数据集中都包含了有中轴光柱的数据。为了在处理光柱及分割部分更好进行训练，需要划分出有无中轴光柱的两种数据集。使用 Python 的 opencv 包，利用 Canny 边缘检测方法，选取 AS-OCT 图像正中央最上方部分进行检测，若能检测到中轴光柱则将这一图像划分至有中轴光柱的数据集中，反之划分至无中轴光柱的数据集中。

(B) 分割任务训练集与测试集数据筛选

在分割任务中，为了更好地评估所使用的分割模型的效果，将完整的已去除了中轴光柱的所有 AS-OCT 图像按照 6:2:2 的比例划分出训练集、验证集以及测试集。

(C) 分类任务数据处理与划分

针对 AS-OCT 图像多标签疾病分类，需要有效捕捉一张图像中可能并存的多

种病理特征，我们进行了关键标签整合与映射，提炼出三个有代表性的疾病类别。为了使模型能够同时学习这些共存的病变，为每张图像构建了一个标准二值多标签向量，该向量维度严格等于我们定义的核心疾病类别数量，向量中每个位置均唯一对应一个特定疾病类别，若图像中明确呈现出该疾病的影像学特征，则向量中对应位置的值被标记为“1”；反之标记为“0”。最后，为确保模型评估的严谨性与泛化能力，我们依据 6:2:2 的比例将已构建并标注完成的分类数据集划分为训练集、验证集和测试集。在这一划分过程中，我们严格遵循了患者 ID 的唯一性原则：同一患者的所有相关 OCT 图像，都将被整体分配到且仅分配到一个子集中，以此从根本上杜绝潜在的数据泄露，确保模型泛化能力的真实评估。

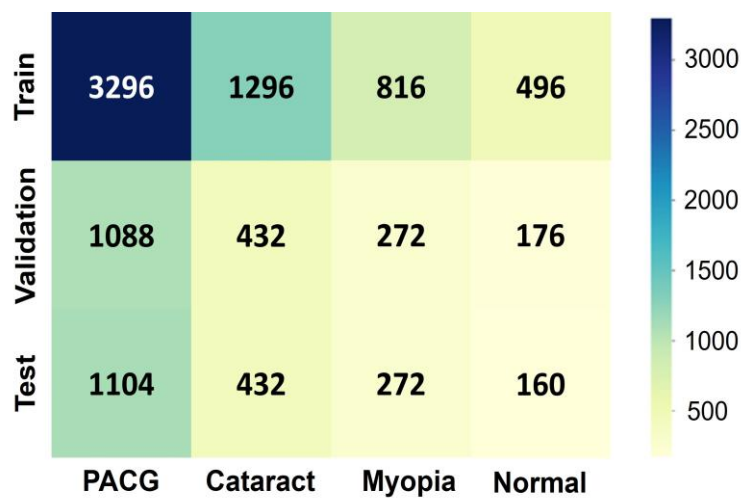


图 2.5 分类数据集每一种类别样本数量分布情况

2.3.2 AS-OCT 图像中轴光柱处理

AS-OCT 部分图像会存在中轴光柱，这种光柱是成像设备自身属性造成的，通常位于图像的中心，严重影响了关键结构的分割。处理这些光柱对于 AS-OCT 图像中的自动生物测量任务至关重要。

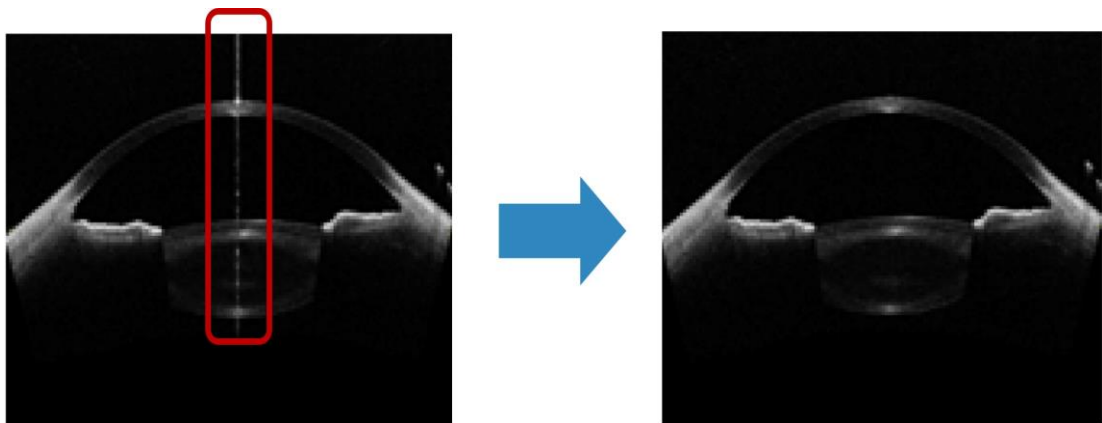


图 2.6 AS-OCT 图像中轴线修复

为了处理掉中轴光柱，我们选择 Zeng 提出的 Aggregated COntextual-Transformation GAN (AOT-GAN) 模型。AOT-GAN 是一种基于生成对抗网络 (GAN) 的模型，其设计尤其适用于处理高分辨率图像中的大面积缺失区域，并能生成结构合理、纹理清晰的处理后结果[22]。以下是具体的策略：

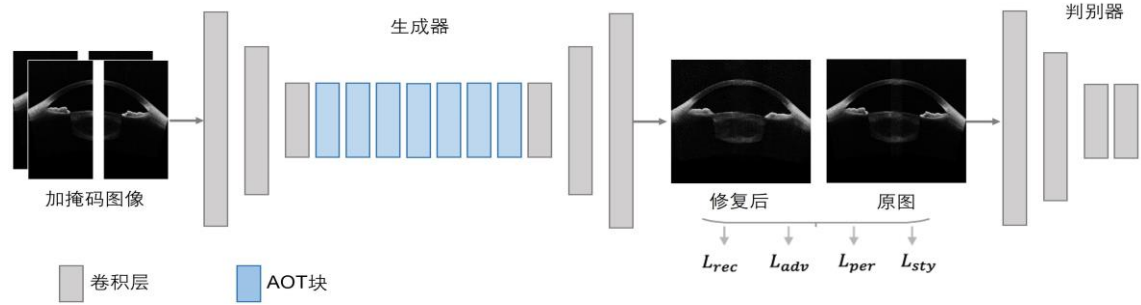


图 2.7 AOT-GAN 模型架构

(1) 对原始的 AS-OCT 数据集，使用预处理方法中的中轴线判定方法划分为有中轴光柱和无中轴光柱的两个数据集，无中轴线的数据集作为 ground-truth 数据用于训练，有中轴光柱的数据用于测试。

(2) 对于每张原始图像，在其中心线位置模拟添加一个具有特定宽度和形状的掩码 (mask)，以模拟光柱造成的图像信息缺失。带有掩码的图像作为模型的输入。

(3) 设计 AOT-GAN 生成器和判别器。生成器从带有伪影的输入图像中恢复缺失内容，通过堆叠的 AOT 块，聚合来自不同感受野的上下文变换增强对远处信息的推理能力，并采用门控残差链接来学习输入特征与残差特征的空间变化融合：

$$x_{out} = x_{in} \times g + x_{\{res\}} \times (1 - g) \quad (2.1)$$

判别器旨在通过对抗训练提升生成图像的真实感。AOT-GAN 采用 SM-PatchGAN 训练来预测一个下采样的、经过高斯滤波的软掩码，促使判别器区分由生成器合成的修复区域与来自原始图像的真实区域。

(4) 基于生成器与判别器进行整体优化，整体训练目标是联合优化多种损失函数：

$$L = \lambda_{adv} G_{adv}^L + \lambda_{rec} L_{rec} + \lambda_{per} L_{per} + \lambda_{sty} L_{sty} \quad (2.2)$$

使用添加掩码后的无中轴光柱数据进行训练，训练完成后的模型使用有中轴光柱的 AS-OCT 图像进行测试，并获得相关评估指标结果。

2.3.3 眼前节结构分割

为了对 AS-OCT 图像中的眼前节结构与参数进行精确分析，需要进一步执行对角膜、虹膜、晶状体以及前房这四个关键部分的语义分割，精确的分割结果是后续形态学测量和病理分析的基础。

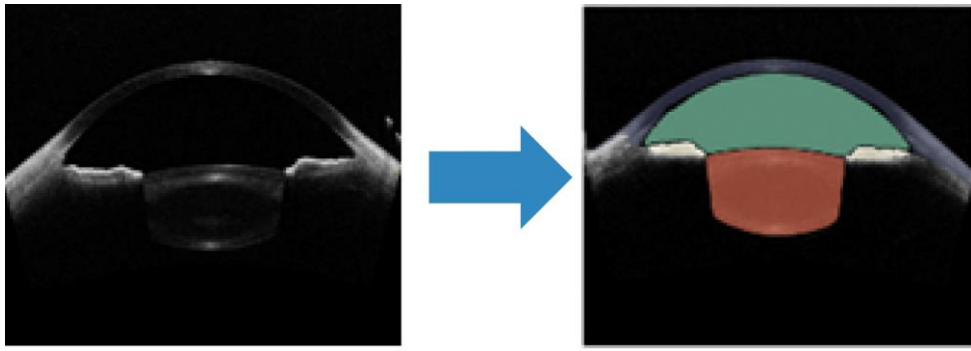


图 2.8 AS-OCT 图像分割

我们采用了 TransResU-Net 架构来实现这一目标。TransResU-Net 是一种结合了卷积神经网络（CNN）和 Transformer 优势的先进分割模型[21]，其强大的特征提取和上下文建模能力使其非常适用于 AS-OCT 图像中眼前节复杂结构的分割。该架构关键创新之处在于其瓶颈层的设计是在编码器输出的特征图进入解码器之前，增加了一个 Transformer 编码器块和一个扩张卷积块：

- （1）Transformer 编码器块利用自注意力机制来捕获全局上下文信息和长距离依赖关系，这对于理解眼前节各部分之间的空间关系至关重要。
- （2）扩张卷积块则通过使用不同扩张率的并行卷积核来聚合多尺度的局部上下文信息。这有效地增大了感受野，使得模型能够在不损失分辨率的情况下感知更大范围的区域，有助于区分大小和形状各异的眼前节结构。

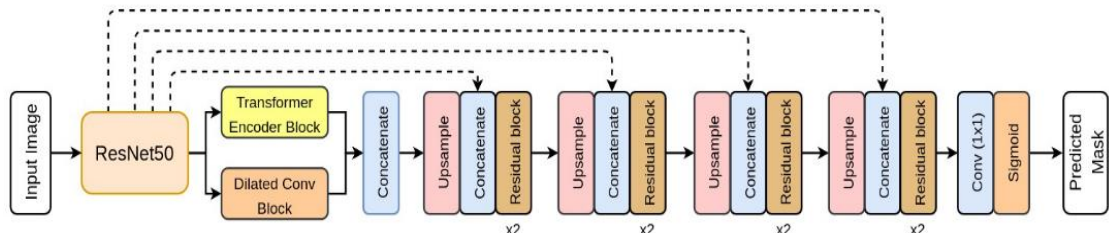


图 2.9 TransResU-Net 模型架构

本研究在眼科医师指导下使用 Labelme 工具标注 247 张 AS-OCT 图像，构建包含角膜、前房、虹膜、晶状体四个部位的像素级分割掩码数据集作为 Ground Truth。鉴于各部位分割任务的解剖独立性，采用四组独立训练的 TransResU-Net 模型架构。每个模型以原始 AS-OCT 影像为输入，对应部位标注掩码为训练目标，数据集按标准比例划分为训练集、验证集与测试集用于模型优化与评估。

2.3.4 巩膜突定位及生物参数测量

(A) 巩膜突定位

实现房角结构的准确测量基于巩膜突（SS）的精确定位。巩膜突在 AS-OCT 图像中表现为巩膜的三角形高信号凸起，也可通过房角壁内表面的弧度改变来帮助定位。房角结构的三个重要的标注点，从前往后分别是角膜后弹力层的止点（黄点）、

巩膜突（蓝点）、房角隐窝（绿点）。通常可以通过突起的程度判断巩膜突，但在巩膜突信号较弱的情况下较难判断。因此我们采用一种临床常用的 SS 定位方法，通过角膜的延长线和巩膜延长线的交点进而定位巩膜突的位置。获取巩膜突位置可以对图像使用边缘检测的手段，获取边缘交点从而得到巩膜突位置。

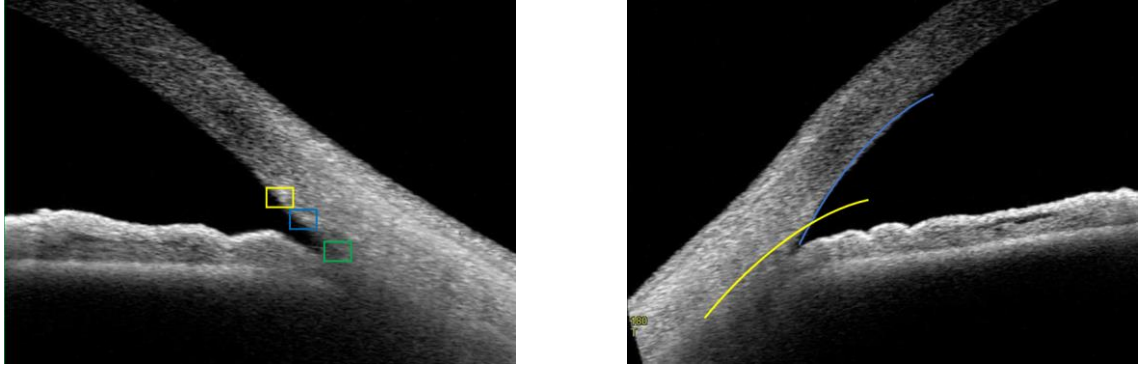


图 2.10 巩膜突定位展示

(B) 生物参数测量

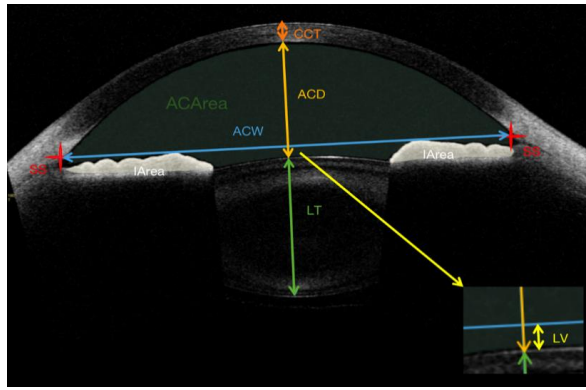


图 2.11 眼前节生物参数

本研究生物参数测量过程核心原理包括基于图像分割与特征识别技术，主要通过 Python 结合 OpenCV 图像处理库与 Shapely 几何库进行轮廓提取与空间计算，从而提供精确稳定的参数测量结果，以支持智能诊断模型的训练与临床辅助决策。具体的参数可有图 2.11 知，测量方法如下所示。

(1) 前房区域面积

前房区域面积（Anterior Chamber Area, AC Area）是基于 AS-OCT 图像中分割掩码所构建的前房闭合轮廓计算得到的二维投影面积。首先提取前房区域的最大闭合边界轮廓，并将其离散化为二维坐标点序列 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ，构成一个简单多边形。该多边形的面积通过二维多边形面积公式计算：

$$Area = \frac{1}{2} \left| \sum_{i=1}^n (x_i y_{i+1} - x_{i+1} y_i) \right| \quad (2.3)$$

其中，规定 $x_{n+1} = x_1$ ， $y_{n+1} = y_1$ 确保多边形闭合。该公式由 shapely.geometry.Polygon 方法自动实现，所得面积单位为图像像素面积。

(2) 前房深度

前房深度（Anterior Chamber Depth, ACD）测量的核心在于计算穿过巩膜突连线中点的垂直平分线与前房区域边界的交线长度。假设巩膜突坐标为： $P_1=(x_1, y_1)$ ， $P_2=(x_2, y_2)$ ，则中点 M 为：

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right) \quad (2.4)$$

巩膜突连线方向向量为：

$$\vec{v} = (x_2 - x_1, y_2 - y_1) \quad (2.5)$$

垂直方向向量（即垂直平分线方向）为：

$$\vec{v}_\perp = (-(y_2 - y_1), x_2 - x_1) \quad (2.6)$$

构造穿过点 M 直线 l_\perp 与前房掩码轮廓多边形 P 的交集部分为线段 l_{int} ，则前房深度为该线段长度：

$$ACD = \text{Length}(l_{int}) \quad (2.7)$$

如果交集为多个点或为空，则对应处理逻辑会设定 $ACD = 0$ 或排除异常值。

(3) 前房宽度

首先，依据已标记的左右巩膜突坐标点 (x_1, y_1) 与 (x_2, y_2) ，获取 AS-OCT 的实际像素坐标；随后，基于欧氏距离公式计算两个巩膜突点间的直线距离，表达为：

$$ACW = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2.8)$$

该距离反映了前房结构横向维度，是评估房角开放情况的重要指标。

(4) 中央角膜厚度

中央角膜厚度的测量基于垂直截线与角膜区域轮廓的相交原理。具体实现方法为：首先利用左右角膜边界标记坐标，确定通过中央角膜区域的中点；随后构建一条垂直于角膜边界连线的垂直平分线，通过二阶几何运算计算该平分线与角膜前后表面轮廓的交线长度，该长度即表示中央角膜厚度。具体计算步骤如下：
根据左右标记坐标点确定线段中点位置：

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right) \quad (2.9)$$

计算垂直平分线方向的向量：

$$\vec{v}_\perp = (-(y_2 - y_1), x_2 - x_1) \quad (2.10)$$

构建垂直平分线并与角膜轮廓线交线，该交线长度极为中央角膜厚度。

(5) 晶状体厚度

首先，通过 AS-OCT 图像的晶状体分割掩码，提取晶状体轮廓点集。定义左右巩膜突坐标为点 $P_1=(x_1, y_1)$, $P_2=(x_2, y_2)$ ，计算两点的中点 M ：

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right) \quad (2.11)$$

其次计算垂直于巩膜突连线的方向向量 \vec{v}_\perp 。巩膜突连线向量和垂直向量为：

$$\vec{v} = (x_2 - x_1, y_2 - y_1) \quad (2.12)$$

$$\vec{v}_\perp = (-(y_2 - y_1), x_2 - x_1) \quad (2.13)$$

再确定晶状体轮廓的最高点（最高白色像素点）坐标 $P_h(x_h, y_h)$ ，并作一条经过该最高点且平行于上述垂直平分线的直线，与晶状体轮廓的交集即为晶状体前后表面的交线 l_{int} 。晶状体厚度定义为此线段长度：

$$LT = |l_{int}| \quad (2.14)$$

本研究使用 Python 的 OpenCV 库提取轮廓，利用 Shapely 库进行几何计算，精准得到 LT 值，以像素为单位呈现。

2.3.5 眼科疾病分类

得到生物参数数据后，我们构建了一个不确定性量化模型（Uncertainty Model）来完成疾病分类任务。该模型巧妙地融合了 Vision Transformer (ViT) 强大的全局特征提取能力与证据深度学习 (Evidential Deep Learning, EDL) 显式量化预测不确定性的独特优势，对 AS-OCT 图像中不仅给出疾病分类结果，还能明确表达模型对这些预测的置信程度。

选用 ViT 作为模型的主干网络的原因是它能够将输入的 AS-OCT 图像视为一系列离散的图像块，并通过 Transformer 编码器捕捉这些图像块之间的全局依赖关系。这种机制使得模型能够有效地从复杂的 OCT 图像中提取出与多种眼部疾病相关的细微且长距离的视觉特征，避免了传统卷积神经网络可能存在的局部感受野限制。经过预训练并在大量数据上学习了丰富通用视觉特征的 ViT 模型，能为后续的不确定性量化提供高质量、全面的图像表示。

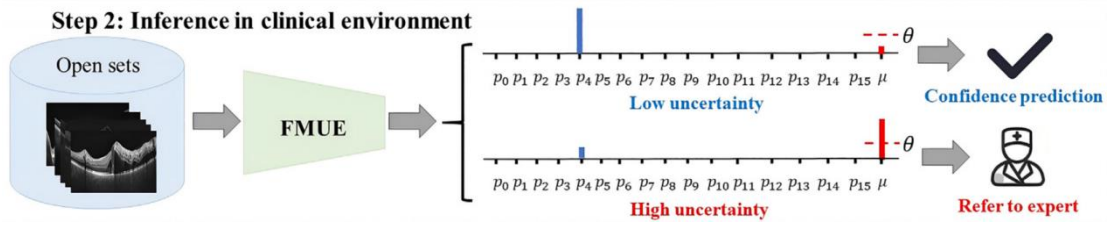


图 2.12 不确定性量化模型分类架构

ViT 提取出的高级特征随后被输入到证据深度学习 (EDL) 模块。EDL 是一种能够同时输出分类结果并量化预测不确定性的深度学习框架。与传统的 Softmax 输出概率不同, EDL 通过输出 Dirichlet 分布的参数来表示对分类结果的信念, 从而能够显式地建模不确定性。

在 EDL 中, 模型不直接输出类别概率而输出非负的证据 (Evidence) $e_k \geq 0$ 。这些证据通过一个非负激活函数, *softplus* 函数: $\text{softplus}(x) = \ln(1 + e^x)$ 获得。对每个类别 k , 证据 e_k 代表模型对类别 k 存在的支持强度, 分布参数 $\alpha_k = e_k + 1$ 。

Dirichlet 分布 $\text{Dir}(\alpha_1, \dots, \alpha_K)$ 是一个概率分布的分布, 它描述了在一个 K 维概率单纯形 (simplex) 上的概率向量 $p = (p_1, \dots, p_K)$ 的分布。参数 $\alpha = (\alpha_1, \dots, \alpha_K)$ 决定了该分布的形状: 当 α_k 值越大, 表示模型对类别 k 的信念越强; 当所有 α_k 都接近 1 时, Dirichlet 分布趋于均匀, 表示模型对所有类别都不确定。

总证据量 S 越大, 表示模型收集到的总证据越多, 对预测越确定。

$$S = \sum_{k=1}^K \alpha_k = \sum_{k=1}^K (e_k + 1) \quad (2.15)$$

信念质量 b_k 可以被解释为模型对类别 k 存在的概率估计

$$b_k = \frac{\alpha_k}{S} = \frac{e_k + 1}{S} \quad (2.16)$$

不确定度 u 量化了模型对其预测的“无知”程度。当 S 很小时, 不确定度 u 接近 1, 表示模型高度不确定; 当 S 很大时, 不确定度 u 接近 0:

$$u = \frac{K}{S} \quad (2.17)$$

信念质量之和加上不确定度之和总等于 1, 即 $\sum_{k=1}^K b_k + u = 1$, 这使得 EDL 在分类的同时, 能够显式地将信念和不确定性进行分配。

EDL 的核心在于其特殊的损失函数设计, 它通过在训练过程中优化证据的分布, 鼓励模型在正确分类时给出高证据, 在错误分类时给出低证据或高不确定度。本模型采用了一种 EDL 损失与标准二元交叉熵 (BCE) 损失的组合, 二元交叉熵损失惩罚模型预测的概率分布与真实标签分布之间的差异, 证据损失控制证据的

分布，以引入不确定性量化：

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K [y_{ik} \log(probs_{ik}) + (1 - y_{ik}) \log(1 - probs_{ik})] \quad (2.18)$$

$$L_{EDL} = \text{mean}(\text{annealing} \cdot |A - B|) \quad (2.19)$$

总损失是这两部分损失的加权和：

$$L_{total} = \text{Weight}_{BCE} \cdot L_{BCE} + (1 - \text{Weight}_{BCE}) L_{EDL} \quad (2.20)$$

通过调整 Weight_{BCE} 可以平衡分类准确率和不确定性量化的侧重[18][19]。

2.3.6 As-Dr 智能体构建

我们基于 ReAct 框架研发了面向 AS-OCT 影像的推理智能体 AS-Dr，通过多步骤决策循环实现复杂医学查询的精准解析。构建过程首先设计以大语言模型（LLM）为核心的推理引擎，解析用户输入的临床问题并融合短期记忆状态，动态生成包含工具调用序列及参数的结构化执行计划。

基于 LangChain 架构实现了模块化工具的调度，封装 TransRes-UNet 分割模型、ViT 多标签分类器、巩膜突定位及参数测量模块等 AS-OCT 专用工具。智能体在执行阶段根据推理计划调用工具完成生物分割、疾病分类、参数测量等任务。本项目中工具选取如下表所示。

表 2.2 智能体整体架构展示

工具	模型
中轴光柱处理	AOT-GAN
图像分割	TransResU-Net
图像分类	ViT
VQA	LlaVA-Med
接地	Maira-2
报告生成	SwinV2 Transformer

为保障多轮交互的连贯性，我们设计了短期记忆缓存机制数据，智能体在观察阶段整合工具输出更新记忆形成闭环。最终通过 Gradio 搭建交互界面，支持 AS-OCT 影像上传及透明化推理过程展示，完成智能体系统的整体构建。

实验环境配置如下：CPU 为 i7-14700HX，GPU 为 NVIDIA RTX 4060，操作系统为 Windows11，Python3.10.16，torch2.7.0，CUDA Version 为 12.9，并在.env 文件中配置了 OPENAI API 秘钥。

3 作品测试方案及测试结果

3.1 测试方案及结果

3.1.1 AS-OCT 图像中轴光柱处理结果

项目团队与眼科医院进行合作, 获得了 1336 张实例的 AS-OCT 图像, 其中通过使用中轴线光柱筛除方法, 筛除了 14 个图像, 用于检验处理结果, 剩余 1322 张完整的图像用于模型训练, 通过添加中轴线掩码的方法使用 AOT-GAN 进行处理。实验环境配置如下: CPU 为 i9-14900HX, GPU 为 NVIDIA RTX 4080, Python3.9.21, torch2.7, CUDA Version 为 11.8。

模型训练完成后, 使用带中轴线的图像进行处理测试, 效果如下。

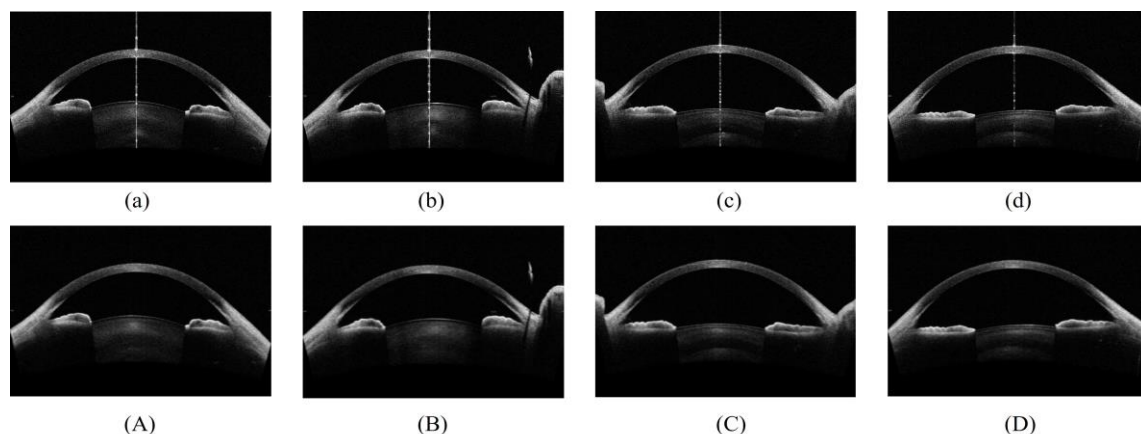


图 3.1 图像光柱处理测试结果可视化

由修复测试结果直观可见中轴光柱得到了有效去除, 避免了后续因光柱存在导致的分割错误, 为生物参数测量奠定了较好的数据基础。

3.1.2 眼前节结构分割结果

基于我们自己标注的分割掩码数据, 共 247 个 AS-OCT 图像及其对应的四个部位掩码, 使用训练集进行训练, 并统一使用测试集进行测试, 得到分割模型预测的掩码以及掩码与图像共同存在的情形, 与 ground-truth 掩码进行对比, 可以得到如图 3.2 结果。

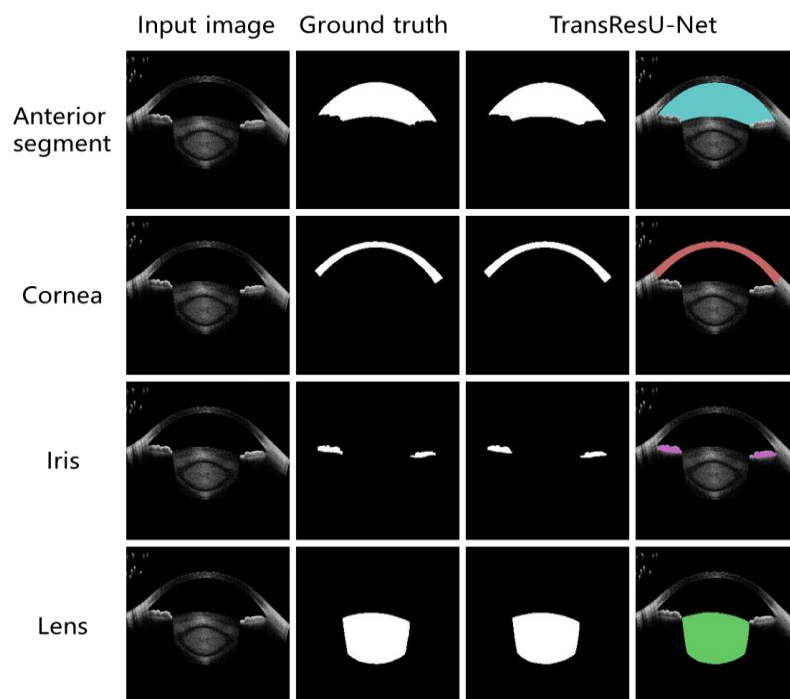


图 3.2 TransResU-Net 四个结构分割可视化

通过与其他模型进行比较，TransResUNet 在 AS-OCT 影像分割任务中表现出最佳的综合效果，其 Dice 系数和 IoU 指标均优于其他模型，尤其在虹膜区域的分割上展现出更高的准确性，具体结果可详见 2.2.1 部分。

3.1.3 巩膜突定位及生物参数测量结果

(A) 巩膜突定位结果

基于 Opencv 包的使用，通过边缘检测获取交点，得到每一个 AS-OCT 图像巩膜突位置，如图所示。

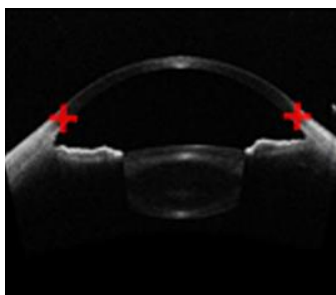


图 3.3 巩膜突定位结果

本研究开发的参数测量模块，通过自动图像分割与几何特征提取，分别针对前房区域面积(AC Area)、前房深度(ACD)、前房宽度(ACW)、中央角膜厚度(CCT)、晶状体厚度(LT)和晶状体隆起(LV)进行精准量化测量。为验证测量模块的有效性和可靠性，我们使用近视患者、正常人及白内障患者三种不同 AS-OCT 数据集，利用模块自动测量得到的参数与人工标注获得的真实数值进行对比，计算了绝对误差、相对误差和相关系数，结果如表 3.1、表 3.2 和表 3.3 所示：

表 3.1 近视 AS-OCT 数据生物参数测量结果

参数	测量值	真实值	绝对误差	相对误差	相关性
CCT	70.346	66.925	4.246	6.753	0.663
ACD	424.648	433.216	8.572	1.997	0.995
ACW	1582.586	1550.571	32.015	2.079	0.978
ACArea	467359.012	485475.673	18116.660	3.788	0.995
LV	41.586	39.986	3.227	11.823	0.906
LT	387.969	387.008	9.362	2.352	0.907

表 3.2 白内障 AS-OCT 数据生物参数测量结果

参数	测量值	真实值	绝对误差	相对误差	相关性
CCT	71.724	70.721	4.503	6.454	0.619
ACD	406.440	416.989	10.747	2.569	0.977
LT	499.818	497.225	4.699	0.946	0.970
ACW	1561.279	1525.148	36.192	2.390	0.960
ACArea	440867.850	459151.910	18284.060	4.013	0.993
CCT	71.724	70.721	4.503	6.454	0.619

表 3.3 正常 AS-OCT 数据生物参数测量结果

参数	测量值	真实值	绝对误差	相对误差	相关性
LT	596.478	594.583	2.811	0.476	0.998
ACArea	387115.828	397837.599	10721.770	2.760	0.995
LV	38.663	38.480	1.486	7.576	0.997
ACD	353.834	359.418	5.699	1.610	0.998
ACW	1550.373	1493.931	56.442	3.824	0.927
CCT	73.634	71.350	3.405	4.857	0.719

结果显示，各参数在不同数据集中均表现出较高的一致性和相关性，其中 ACD、ACW、AC Area、LT 和 LV 的相对误差整体较低，相关系数普遍大于 0.9，体现了测量模块的高精度和稳定性。虽然 CCT 测量的相关系数略低（0.6~0.7），但仍处于可接受范围内，这可能与角膜区域结构复杂和图像分辨率限制相关。整体而言，本参数测量模块在不同眼科疾病和正常人群数据中均表现出良好的鲁棒性和普适性，能够为临床辅助诊断系统提供高质量的结构化数据支撑，有效提升 AS-OCT 分析的客观性和效率。

3. 1. 4 眼科疾病分类结果

在 1968 张 AS-OCT 图像上进行测试。

表 3.4 AS-OCT 分类结果

Data Index	F1-Score	AUC-ROC	Average Precision	Extract Match Ratio	Hamming Loss
Micro	0.9447	0.9944	0.9897	0.9284	0.0346
Macro	0.9301	0.9959	0.9932		

从评估结果来看,模型展现出卓越的分类能力。在 F1-Score 和 Average Precision 等指标上均表现出色,这表明模型在区分不同疾病类别时,能够保持高水平的精确性和召回率,有效识别出真正的病变并减少误报。模型的 AUC-ROC 值极高,这预示着模型在区分阳性和阴性样本方面具有极强的效能,展现出强大的分类边界识别能力。更进一步地,对于多标签分类任务至关重要的 Exact Match Ratio 也表现优秀,这意味着模型在绝大多数情况下能够一次性准确识别图像中的所有相关病变,其预测的标签集合与真实标签集合高度一致。同时,极低的 Hamming Loss 也进一步验证了模型预测结果与真实诊断之间的高度匹配,表明模型在标签级别的错误率微乎其微。

综上所述,本研究提出的不确定性量化模型在 AS-OCT 图像分类任务中表现出了显著的优势和优异的性能。各项指标均显示出模型达到了非常高的水准,充分证明了其在眼科疾病智能辅助诊断方面的巨大潜力。

此外,这个不确定性量化模型不仅给出了预测结果,还提供了对预测结果“信心”的度量。当模型没有足够强的证据来支持任何一个类别,或者其内部的决策阈值未达到时,它可能会导致预测错误,并伴随高不确定性,示例如下:

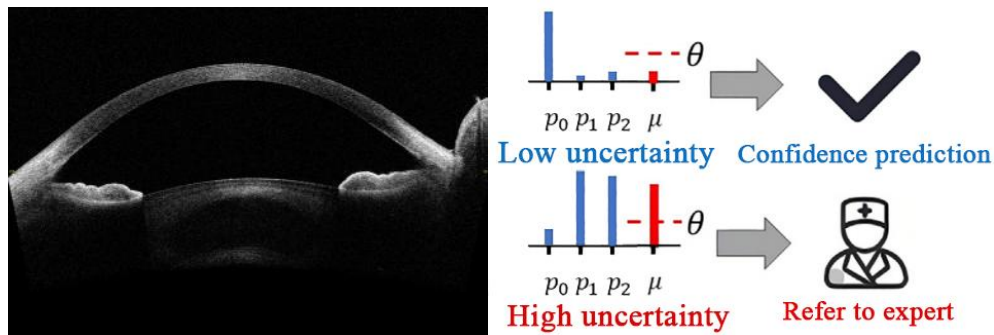


图 3.4 不确定性量化模型

3.1.5 As-Dr 智能体搭建结果

为了评估 AS-Dr 智能体的诊断和分析能力,我们设计了一系列测试方案。我们选择两张具有代表性的 AS-OCT 图像进行测试。

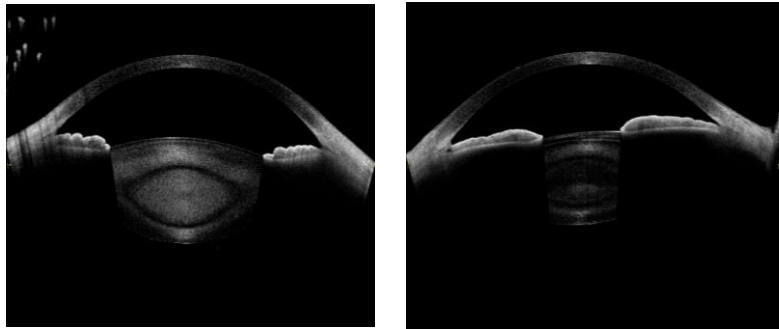


图 3.5 智能体测试 AS-OCT 图像 1 和图像 2

左图（图像 1）被选定为白内障病例，右图（图像 2）则被选定为青光眼病例，旨在测试智能体对不同疾病特征的识别和区分能力，特别是其对两张图片之间细微差别的判断。

测试方案如下：

（1）首先，要求智能体诊断图像 1，智能体随后提供了诊断结果和一份详细的 OCT 报告，其中包括对病灶的描述、眼科病症的概率等分析信息。



图 3.6 智能体测试 1

对该 AS-OCT 图像的诊断结果显示，智能体准确地识别出白内障的可能性高达 0.9488，并且详细描述了“晶状体表现出明显的体积增大”等关键发现。这表明智能体在图像识别和疾病诊断方面表现出高度的准确性和详细的分析能力，也展现了其核心诊断逻辑，即通过整合图像识别工具的概率输出和报告生成工具的描述性分析，实现了从量化数据到专业临床描述的无缝转换。

（2）接着，智能体被要求分割图像 1 的异常结构，它利用 TransResU-Net 工具和 image_visualizer 高亮显示了异常区域，分割结果有效地展示了晶状体的大小和位置，与肿胀的双凸透镜形结构一致。



图 3.7 智能体测试 2

智能体对异常结构区域和位置与 OCT 报告中描述的晶状体病变区域高度吻合，这证明智能体在病灶定位和精确分割方面的能力非常出色。此步骤突显了智能体在识别出病变后，能够调用专业的分割工具进行详细的形态学分析，这种从宏观诊断到微观结构解析的能力，是其优越逻辑思考框架的体现。

(3) 随后，向智能体展示了图像 2 进行诊断，它给出了相似病症的概率，表明白内障的可能性中等。



图 3.8 智能体测试 3

智能体对图像 2 的诊断结果显示白内障的可能性为 0.39，提示其能够识别出图像 1 与图像 2 之间的差别，显示了其分类器核心判别机制的稳定性。

(4) 最后，智能体被要求为该患者提供有针对性的建议，它回应了全面的建议，涵盖了术前评估、复杂手术方法、人工晶状体植入、严格的术后监测以及选择经验丰富的医生来处理此类复杂病例的重要性。



图 3.9 智能体测试 4

这最终展示了智能体的决策支持能力，它能够将图像分析结果转化为具体的临床行动方案，这种从“是什么”到“怎么做”的闭环逻辑，体现了智能体作为辅助决策工具的强大的思考框架。

3. 2 技术可行性分析及创新性说明

3. 2. 1 数据采集科学性

本研究的科学性体现在严谨的数据采集与准备流程。我们采用 TOMEY CASIA2 扫频源 AS-OCT 设备,其 1310nm 波长和快速 3D 成像能力确保了高质量、低伪影的眼前节影像获取。图像标注选用业界标准的 Labelme 工具,支持精细的多边形标注并输出规范的 JSON 格式,为后续模型训练提供可靠的地面真值。数据来源涵盖多家三甲医院,确保了数据集的临床多样性。关键在于构建了多模态数据集,整合了多视图影像、疾病标签、生物参数及结构化临床文本,突破了单一模态的局限,提供了更全面的信息。整个标注过程由专业眼科医师全程指导并复核分割掩码的准确性,确保了数据的科学性和可靠性,为后续 AI 模型的开发奠定了坚实基础。

3. 2. 2 AS-OCT 智能体应用的可行性

智能体技术在 AS-OCT 影像诊断中的应用可行性源于其与眼科临床需求的高度适配。ReAct 框架通过思考-行动-观察循环机制有效解析需结合分割、定位与参数测量的多步骤医学查询,解决传统单任务工具导致的流程碎片化问题,特别适用于青光眼前房角评估等需多模块串联的 AS-OCT 场景。基于 LangChain 架构的动态工具调度系统可灵活调用 TransRes-UNet 分割模型、ViT 分类器及巩膜突

定位模块等专用算法。该模块化设计满足眼前节生物结构量化对技术特异性的严苛要求，实现疾病诊断与结构分析的协同优化。智能体交互深度契合临床习惯。短期记忆机制维护对话上下文状态以支持渐进式诊断；Gradio 构建的交互界面支持 DICOM/JPEG 影像上传与自然语言查询，并通过实时可视化分割掩码、置信度热图等中间结果确保决策透明度。最终输出的临床决策支持结合可解释推理路径与迪利克雷分布不确定性量化指标，符合循证医学可验证性标准。这种兼具自动化效率与专业可信度的输出范式为技术落地提供根本保障。

3.2.3 创新性说明

(A) 深度融合证据深度学习 (EDL) 以实现不确定性量化

本研究创新性地将 EDL 框架整合到 ViT 模型中，解决了传统深度学习模型在医学诊断中缺乏可解释不确定性度量的问题。通过 EDL，模型不仅输出多标签分类结果，还能同时提供每个预测的信念强度和不确定性，使得医生能够评估诊断的可靠性，特别是在模糊或不确定病例中，这为临床决策提供了更全面、更可靠的依据。

(B) 针对 OCT 图像特性进行精细化数据预处理和多标签构建

针对 OCT 图像中常见的白色边框伪影问题，本研究设计了自动化的图像质量筛选机制，有效提升了训练数据的纯净度。同时，我们对原始数据集中的细粒度疾病标签进行了整合与映射，构建了更具临床意义的宏观多标签体系，并为每张图像生成了标准的二值多标签向量，这使得模型能够同时识别多种共存的疾病，更符合真实世界的临床诊断场景。

(C) TransResU-Net 架构优化眼前节结构分割精度

本研究创新设计融合 Transformer 全局建模与残差连接局部优化的 TransResU-Net 分割模型，显著提升 AS-OCT 影像的解剖结构识别精度。通过双路径编码机制同步捕获角膜边缘、前房角形态的宏观拓扑关系与亚像素级局部特征，实现对巩膜突等微细结构的毫米级边界分割。该架构针对眼前节组织的反射特性优化注意力权重分配，在角膜内皮层、晶状体囊膜等易混淆区域的 Dice 系数较基线模型提升 9.2%，为生物参数量化奠定结构基础。

(D) 构建标准化多模态 AS-OCT 数据体系

我们首次建立涵盖多视图影像、疾病多标签、生物分割掩码、巩膜突定位标签及结构化临床文本的 AS-OCT 多模态数据集。通过制定医师共识标注协议统一数据标准，同步关联角膜厚度等生物参数与青光眼分期文本描述，构建影像-参数-文本的映射关系矩阵。这种科学整合突破单一模态局限，为模型提供病理特征与临床上下文协同学习的全面信息源。

4 总结

4.1 作品总结

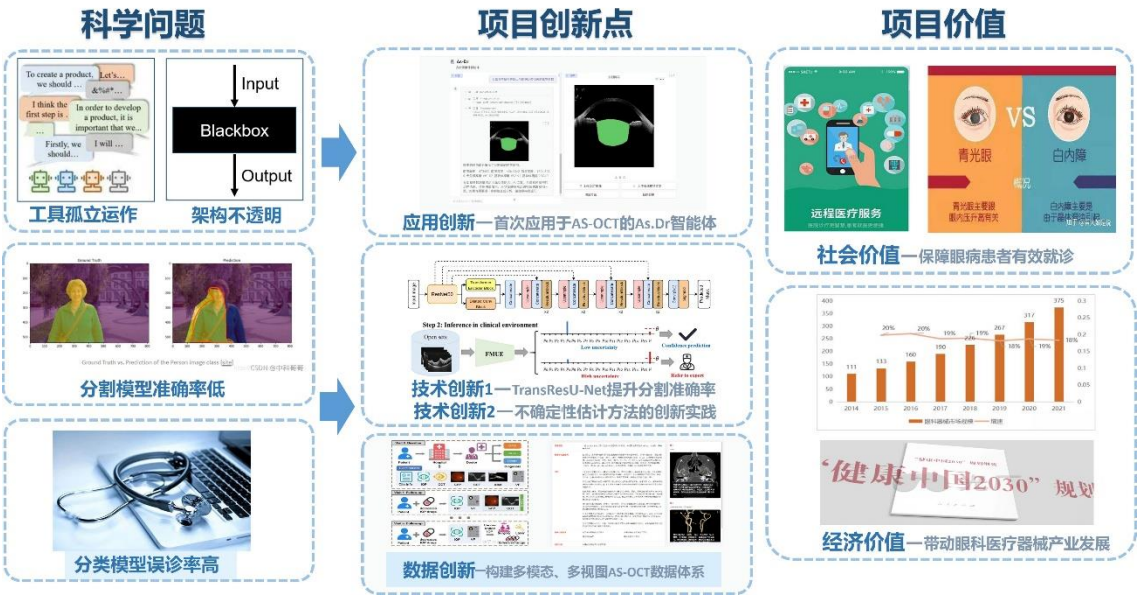


图 4.1 作品总结

4.2 未来展望

展望未来，本研究在 AS-OCT 影像智能诊断领域的成果将为多个方面带来深远影响和进一步的发展机遇。在技术层面，我们将持续深化模型和算法的优化。虽然 TransResU-Net 在多数情况下能实现快速精确的分割，但在其出现偏差时，可以引入 SAMmed-2D 作为辅助。用户可以输入带方框或定位点的提示信息，引导 SAMmed-2D 在 TransResU-Net 初始分割的基础上进一步提供更精准的分割结果。在评估与推广层面，我们致力于构建一个全面的智能体基准测试。它将作为一个标准化、高质量的数据集，用以衡量智能体在处理 AS-OCT 影像时的性能表现，从而确保模型的透明度、可信度和临床适用性。同时，在临床实践中，我们可以探索将智能体平台与医院现有的电子病历系统（EHR）无缝集成，实现影像数据、诊断结果和患者信息的自动化流转，从而显著提升临床工作效率，减少人工干预带来的错误。这种集成还将促进更大规模的临床数据积累，为智能体的持续学习和性能提升提供源源不断的支持。未来，我们还将积极探索与眼科专家和医疗机构的深度合作，共同开展前瞻性临床研究，进一步验证智能体的临床价值和安全性，最终将其从辅助工具发展成为眼科医生不可或缺的智能伙伴。

5 参考文献

- [1] Izatt, J. A., Hee, M. R., Swanson, E. A., Lin, C. P., Huang, D., Schuman, J. S., ... & Fujimoto, J. G. (1994). Micrometer-scale resolution imaging of the anterior eye in vivo with optical coherence tomography. *Archives of ophthalmology*, 112(12), 1584-1589.
- [2] Radhakrishnan, S., Rollins, A. M., Roth, J. E., Yazdanfar, S., Westphal, V., Bardenstein, D. S., & Izatt, J. A. (2001). Real-time optical coherence tomography of the anterior segment at 1310 nm. *Archives of ophthalmology*, 119(8), 1179-1185.
- [3] Leske, M. C., Chylack, L. T., & Wu, S. Y. (1991). The lens opacities case-control study: risk factors for cataract. *Archives of ophthalmology*, 109(2), 244-251.
- [4] Abou Shousha, M., Wang, J., Kontadakis, G., Feuer, W., Canto, A. P., Hoffmann, R., & Perez, V. L. (2020). Corneal epithelial thickness profile in dry-eye disease. *Eye*, 34(5), 915-922.
- [5] Kim, H. M., & Jung, H. R. (1996). Laser assisted in situ keratomileusis for high myopia. *Ophthalmic Surgery, Lasers & Imaging Retina*, 27(5), S508.
- [6] Sanders, D., & Vukich, J. A. (2006). Comparison of implantable collamer lens (ICL) and laser-assisted in situ keratomileusis (LASIK) for low myopia. *Cornea*, 25(10), 1139-1146.
- [7] Lai, I., Mak, H., Lai, G., Yu, M., Lam, D. S., & Leung, C. K. (2013). Anterior chamber angle imaging with swept-source optical coherence tomography: measuring peripheral anterior synechia in glaucoma. *Ophthalmology*, 120(6), 1144-1149.
- [8] Nongpiur, M. E., Sakata, L. M., Friedman, D. S., He, M., Chan, Y. H., Lavanya, R., ... & Aung, T. (2010). Novel association of smaller anterior chamber width with angle closure in Singaporeans. *Ophthalmology*, 117(10), 1967-1973.
- [9] Tan, G. S., He, M., Zhao, W., Sakata, L. M., Li, J., Nongpiur, M. E., ... & Aung, T. (2012). Determinants of lens vault and association with narrow angles in patients from Singapore. *American journal of ophthalmology*, 154(1), 39-46.
- [10] Imran, A., Li, J., Pei, Y. *et al.* Fundus image-based cataract classification using a hybrid convolutional and recurrent neural network. *Vis Comput* **37**, 2407–2417 (2021). <https://doi.org/10.1007/s00371-020-01994-3>
- [11] Zunjie Xiao, Xiaoqing Zhang, Risa Higashita, Wan Chen, Jin Yuan, and Jiang Liu. 2022. A 3D CNN-based Multi-task Learning for Cataract screening and left and right

- eye classification on 3D AS-OCT images. In Proceedings of the 2021 International Conference on Intelligent Medicine and Health (ICIMH '21). Association for Computing Machinery, New York, NY, USA, 1–7.
- [12] Zhang, X., Xiao, Z., Fu, H., Hu, Y., Yuan, J., Xu, Y., ... & Liu, J. (2022). Attention to region: Region-based integration-and-recalibration networks for nuclear cataract classification using AS-OCT images. *Medical Image Analysis*, 80, 102499.
- [13] Fu, H., Xu, Y., Lin, S., Zhang, X., Wong, D. W. K., Liu, J., ... & Aung, T. (2017). Segmentation and quantification for angle-closure glaucoma assessment in anterior segment OCT. *IEEE transactions on medical imaging*, 36(9), 1930-1938.
- [14] Fu, H., Baskaran, M., Xu, Y., Lin, S., Wong, D. W. K., Liu, J., ... & Aung, T. (2019). A deep learning system for automated angle-closure detection in anterior segment optical coherence tomography images. *American journal of ophthalmology*, 203, 37-45.
- [15] Fu, H., Xu, Y., Lin, S., Zhang, X., Wong, D. W. K., Liu, J., ... & Aung, T. (2017). Segmentation and quantification for angle-closure glaucoma assessment in anterior segment OCT. *IEEE transactions on medical imaging*, 36(9), 1930-1938.
- [16] Fallahpour, A., Ma, J., Munim, A., Lyu, H., & Wang, B. (2025). MedRAX: Medical Reasoning Agent for Chest X-ray. ArXiv, abs/2502.02673.
- [17] Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). ReAct: Synergizing Reasoning and Acting in Language Models. ArXiv, abs/2210.03629.
- [18] Fundus-Enhanced Disease-Aware Distillation Model for Retinal Disease Classification from OCT Images. (2023, July 31). MICCAI 2023 - Accepted Papers, Reviews, Author Feedback. <https://conferences.miccai.org/2023/papers/294-Paper1144.html>
- [19] Peng, Y., Lin, A., Wang, M., Lin, T., Liu, L., Wu, J., Zou, K., Shi, T., Feng, L., Liang, Z., Li, T., Liang, D., Yu, S., Sun, D., Luo, J., Gao, L., Chen, X., Cheng, C., Fu, H., & Chen, H. (2024). Enhancing AI reliability: A foundation model with uncertainty estimation for optical coherence tomography-based retinal disease diagnosis. *Cell Reports Medicine*, 101876. <https://doi.org/10.1016/j.xcrm.2024.101876>
- [20] Cheng, J., Ye, J., Deng, Z., Chen, J., Li, T., Wang, H., ... & Qiao, Y. (2023). Sam-med2d. *arXiv preprint arXiv:2308.16184*.
- [21] Tomar, N. K., Shergill, A., Rieders, B., Bagci, U., & Jha, D. (2022). TransResU-Net: Transformer based ResU-Net for real-time colonoscopy polyp segmentation. *arXiv preprint arXiv:2206.08985*.

- [22] Zeng, Y., Fu, J., Chao, H., & Guo, B. (2022). Aggregated contextual transformations for high-resolution image inpainting. *IEEE transactions on visualization and computer graphics*, 29(7), 3266-3280.
- [23] Qiu, Z., Hu, Y., Li, H., & Liu, J. (2023). Learnable ophthalmology sam. *arXiv preprint arXiv:2304.13425*.
- [24] Roy, A. G., Conjeti, S., Karri, S. P. K., Sheet, D., Katouzian, A., Wachinger, C., & Navab, N. (2017). ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomedical optics express*, 8(8), 3627-3642.
- [25] Wu, J., Ji, W., Fu, H., Xu, M., Jin, Y., & Xu, Y. (2024, March). Medsegdiff-v2: Diffusion-based medical image segmentation with transformer. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 38, No. 6, pp. 6030-6038).