



浙江省大学生人工智能竞赛



# 智绘医疗

## ——双原型学习驱动的眼底增强分布诊断与报告生成系统

*Dual Prototype Learning-Driven Fundus Enhanced Distribution Diagnosis and Report Generation System*

啊对对队

赛道：场景应用挑战赛

专题：医疗与安全场景专题赛道

# 市场痛点：眼类病未能早筛给人们带来严重的经济负担与健康隐患



以眼眶病筛查时期对疗效、风险、花费的影响为例

眼眶病早期



0~6个月

眼眶病中期



6~9个月

眼眶病晚期



9~12个月后

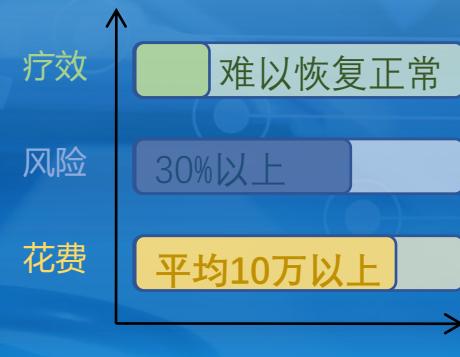
药物治疗为主



药物+放疗



药物+放疗+手术



人口老龄化加剧

眼科疾病在全球范围内影响着**数亿**人生活

我国是全球眼病患者最多的国家

- 中国糖尿病患者数量位居全球第一，达**1.164亿**。
- 约**每3位**糖尿病患者中就有**1位**糖尿病视网膜病变。

眼眶病早期筛查可减少**95%**的医疗费用，降低**30%**的致盲致死率

# 技术痛点1：传统诊断模式不适用于早期筛查

## 专科人力不足

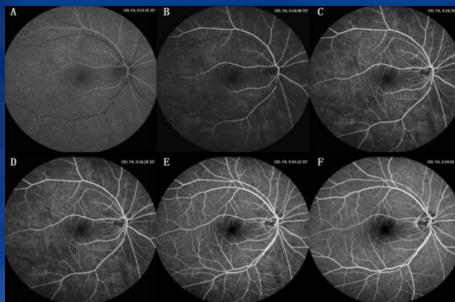


眼科医生  
小于5万人

眼科医生培养周期长、门槛高  
工作负荷强、职业风险高

不可靠！

## 基层技术不够

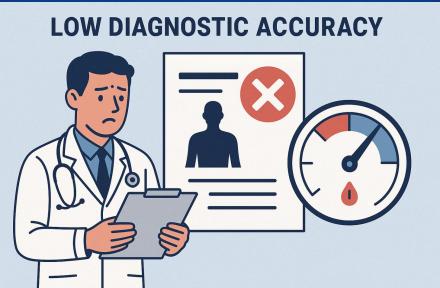


基层影像  
读片困难

眼底镜检测误诊率22%  
基层医生诊断准确率仅63%  
诊断一致性仅中等

延病情！

## 诊断准度欠佳



诊断主观性强  
诊断准确率低

对经验、专注力要求高  
疲劳状态下误判风险增加

需要借助人工智能进行早期筛查

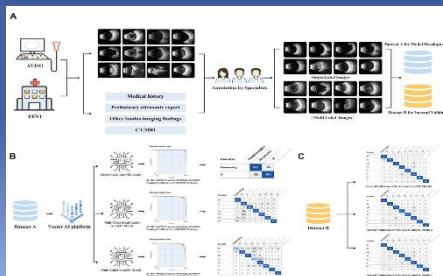


传统诊断模式尚未实现  
有效且准确的早期筛查

# 技术痛点2：机器/深度学习方法难以实际临床应用

## 眼病AI筛查壁垒重重

### 机器学习方法



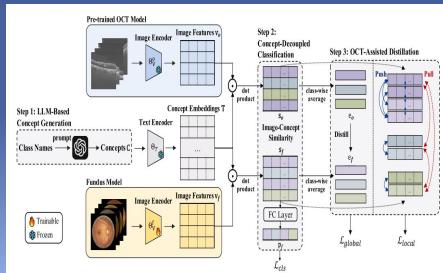
随机森林/支持向量机等

特征选择困难

泛化能力差

处理非线性复杂关系弱

### 深度学习方法



单模态/多模态  
/跨模态融合等

不具备可解释性

过度依赖图像

忽略临床信息

推理流程  
无法可视化

医学先验知识  
融合缺乏

黑箱模型

Input → Black Box → Output

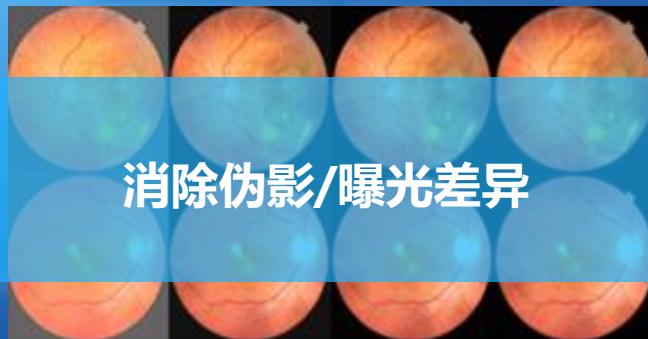
难应用！

实际投产与临床  
应用困难

极端病例  
泛化能力弱

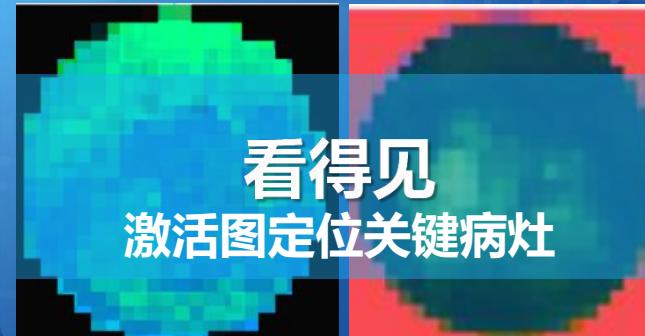
## 双原型学习驱动的诊断系统：可解释・可追溯・临床可核验

### 1、影像净化



消除伪影/曝光差异

### 2、病灶激活图



看得见  
激活图定位关键病灶

### 数据导入

### 影像净化

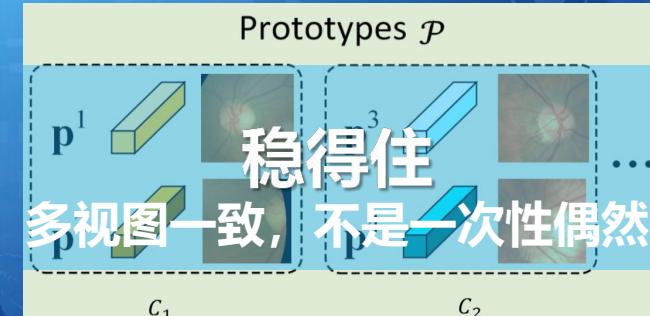
### 激活图生成

### 3、原型对照



对得上  
原型对照，可追溯训练样本

### 4、一致性



Prototypes  $\mathcal{P}$   
 $p^1$        $p^2$        $p^3$   
稳得住 多视图一致，不是一次性偶然  
 $c_1$                $c_2$

### 原型对照

### 一致性检验

### 诊断报告生成

#### 一键预测

预测结果	AMD/糖尿病/有髓神经纤维
糖尿病	0.816
白内障	0.042
高血压	0.293
其它疾病/异常	0.919
青光眼	0.127
AMD	0.874
近视	0.178

#### 报告

- 老年性黄斑变性 (AMD)  
诊断：黄斑区退行性病变，可能出现视力下降、视物变形和视野中心暗点。  
治疗建议：  
中期AMD：建议补充微量营养素（如维生素C、维生素E、锌等），并定期随访。  
生活方式管理：戒烟、控制血压、定期使用Amsler方格表监测视力变化。
- 糖尿病视网膜病变 (DR)  
诊断：视网膜微血管损害，非增生型 (NPDR)，并发糖尿病性黄斑水肿 (DME)。  
治疗建议：

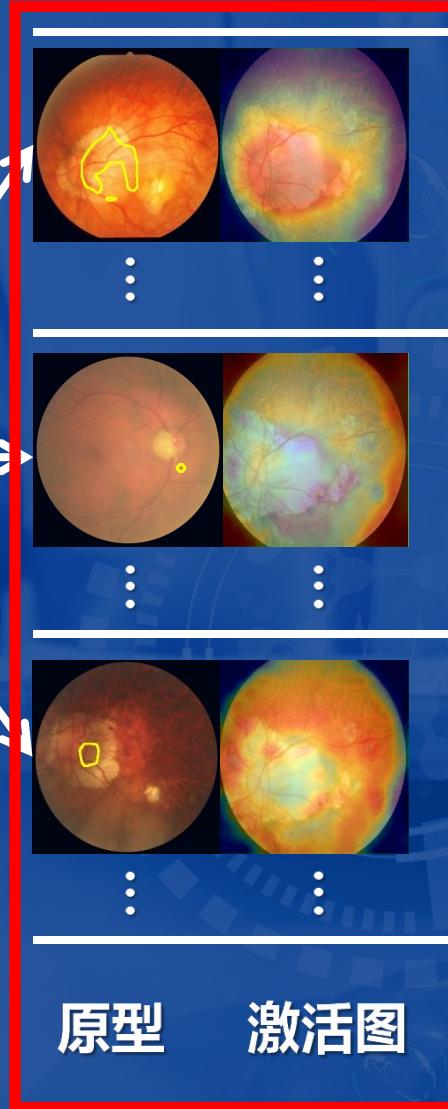
可追溯证据链：每次诊断都附“可视证据 + 原型旁证 + 一致性指标”，支持医生逐条核验

# 系统推理流程演示：可追溯证据链：每次诊断都附“可视证据 + 原型旁证 + 一致性指标”

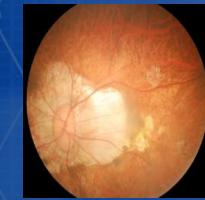


测试图片  
(经预处理)

对比



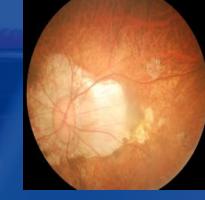
近视	0.960	$\times$	1.169	Risk $\geq 0.5$ 判阳性
青光眼	0.754	$\times$	1.023	0.87 近视的输出概率/风险
高血压性视网膜病变	0.698	$\times$	1.123	0.42 青光眼的输出概率/风险 0.31 高血压性视网膜病变的输出概率/风险



视图1



视图2



视图3

计算同一图像在不同视图输入下模型预测结果的稳定程度（得分方差的反向归一化）

一致性指标

# 项目历程



## 打破医研壁垒，追求临床实效

### 攻克实践调研关



前往浙医二院眼科中心等地  
调研、洽谈合作事宜、研究  
通过伦理审查，取得影像数  
据使用许可

### 攻克算法创新关



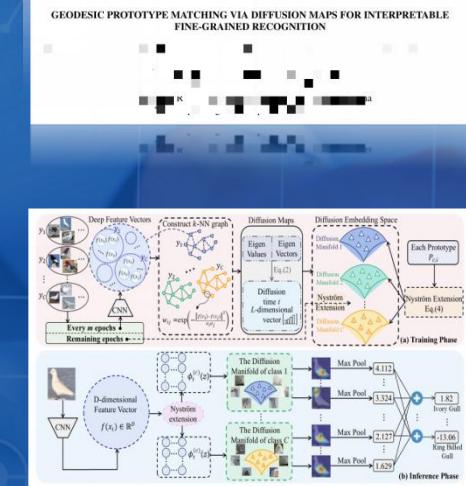
完成原型模型的搭建与训  
练、前后端v1.0版本均  
完成部署

### 攻克临床试用关



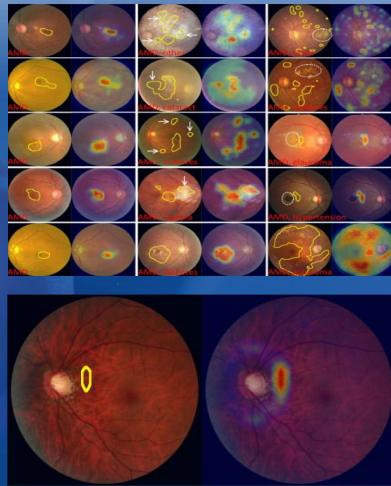
产品在浙医二院眼科中心  
等地进行试用，获得推荐  
信两封、产品试用报告两  
份、合作意向书两份

### 攻克技术迭代关



提出 GeoProto框架，以  
类流形取代欧式空间。在  
流体上学习与匹配原型。  
并发表CCF-B论文1篇

### 原型数据扩充关



在原有的七类疾病原型库  
上继续扩充，增加系统可  
诊断疾病类别

2025年1月-2月

2025年2月-3月

2025年3-4月

2025年4月-8月

2025年8月-至今

## 核心模块1：基于FundusDiff的数据预处理模块

### 扩散概率模型

- 条件扩散
- 去噪概率扩散

### 基于图像块的任意尺寸增强

- 条件引导
- 卷积压缩
- 传递潜在表示

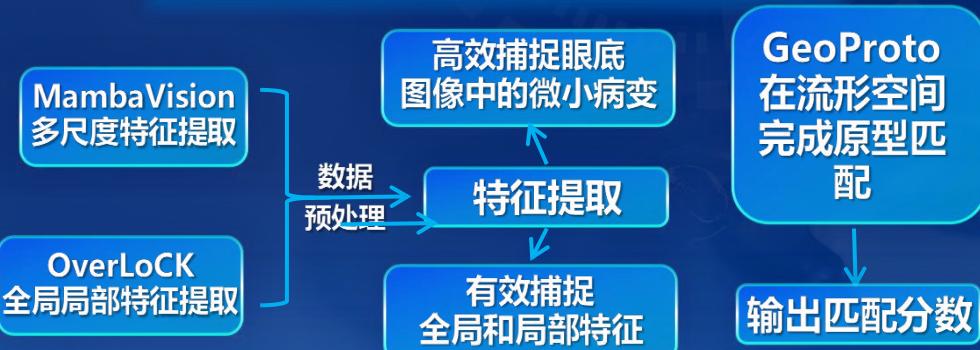
定向生成

### 损失函数与亮度控制

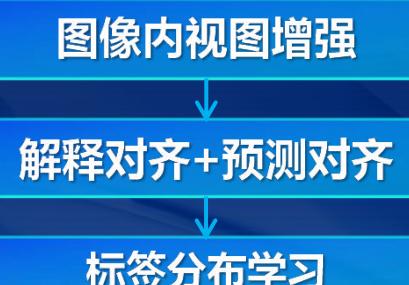
- 自适应亮度调节
- 曝光控制损失
- 光照平滑损失

## 核心模块2：双原型学习（DPL）驱动的诊断模块

### 双Backbone与GeoProto



### 多标签增强框架



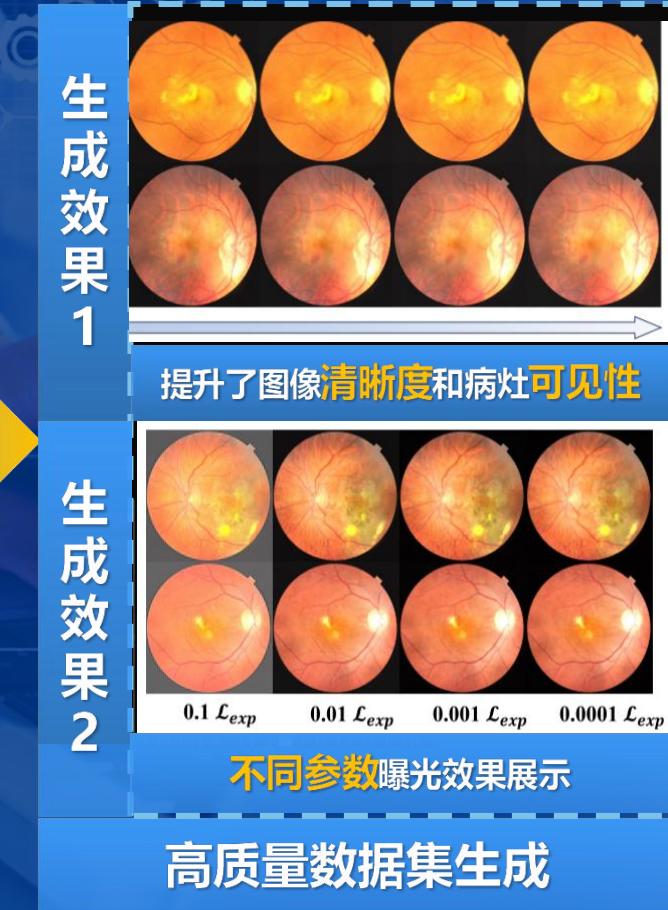
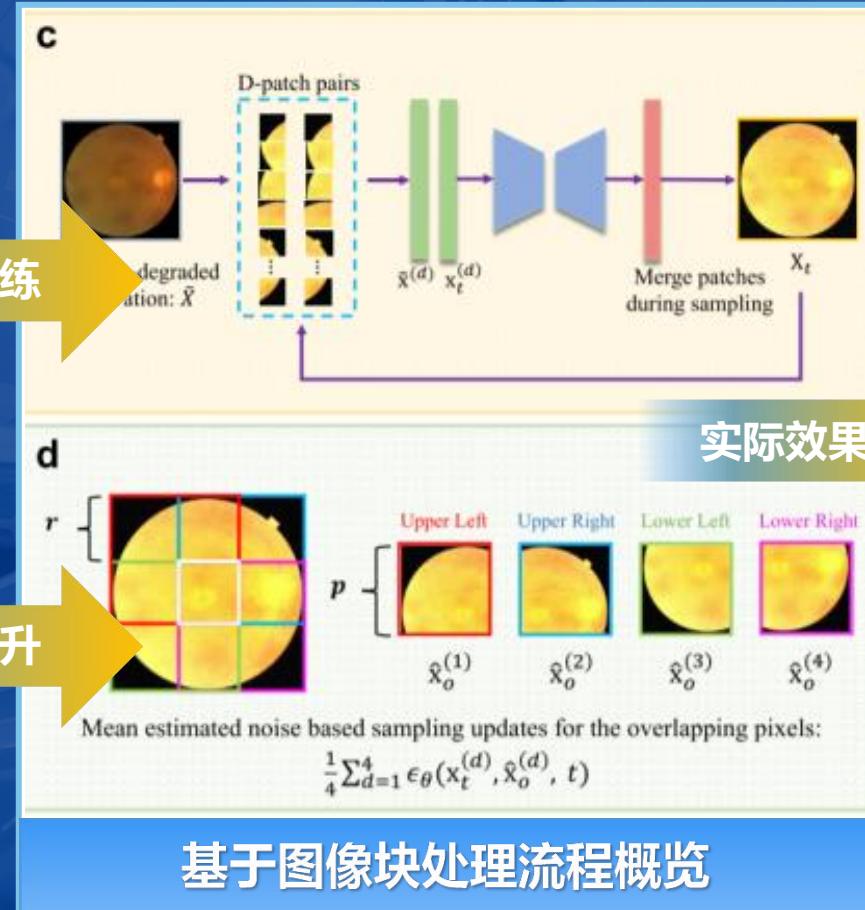
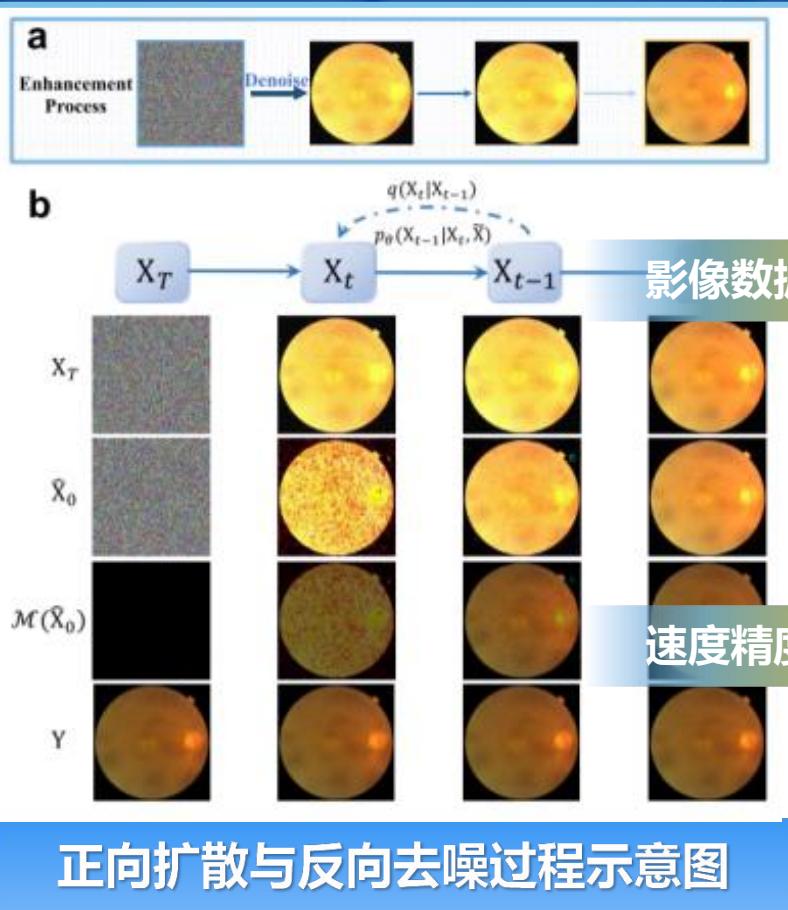
### 综合诊断系统平台



# 核心模型1——基于FundusDiff的数据预处理

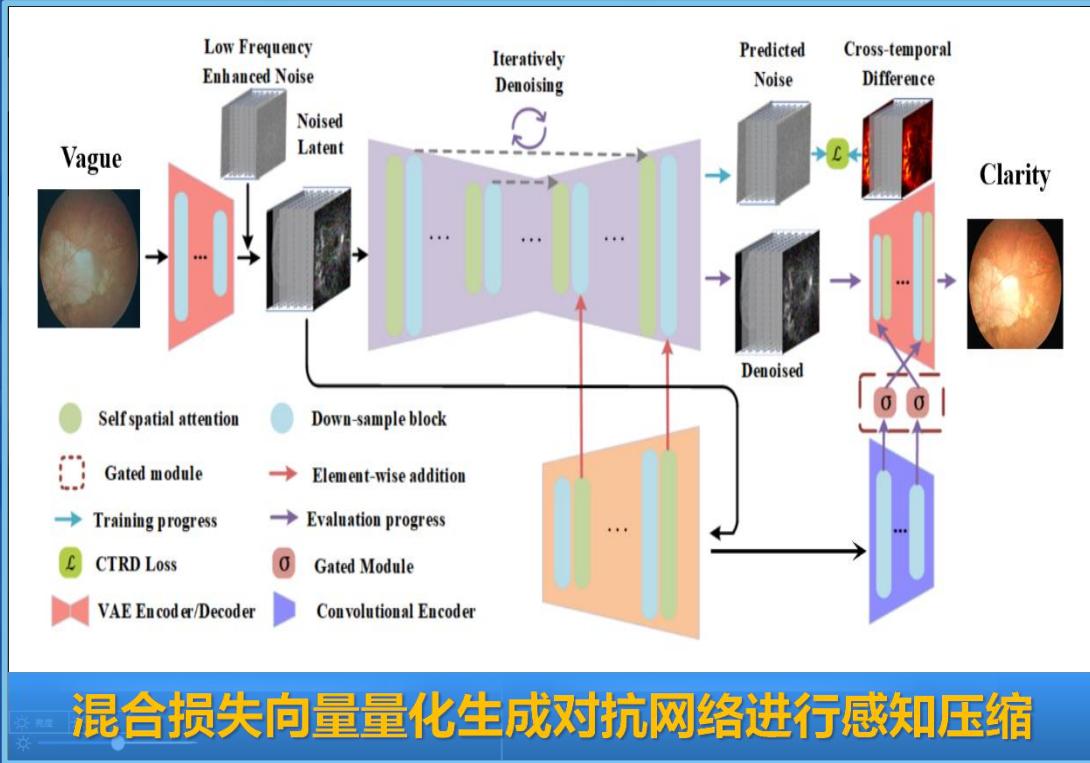
## 基于扩散模型的眼底图像质量增强架构

利用扩散模型强大的图像生成能力赋能眼底图像生成，引导图像逐步从噪声中恢复为高质量状态



## 基于扩散模型的眼底图像质量增强架构

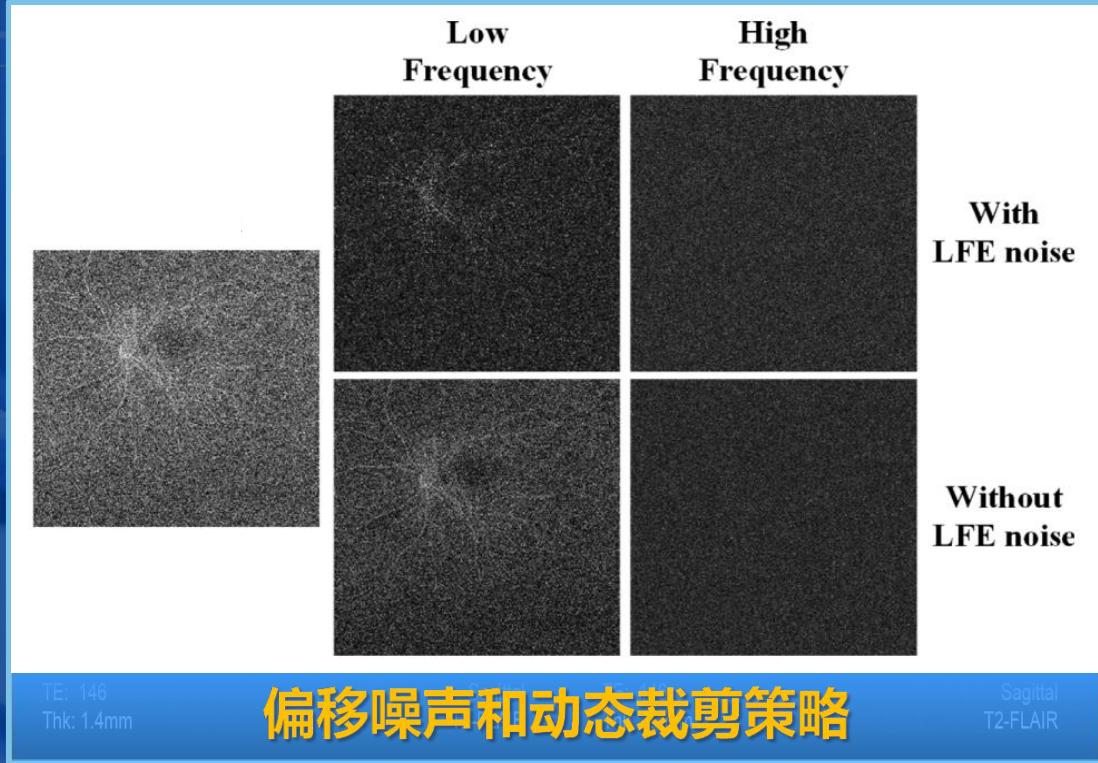
### 感知压缩实现细节



解决问题：解放了模型的计算需求，…

达到效果：使眼底医学影像生成更清晰、更快速

### 偏移噪声可视化细节



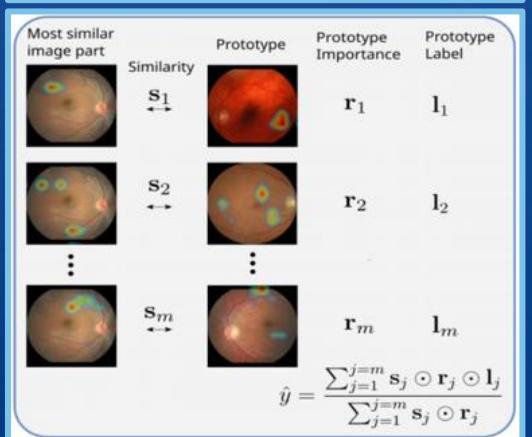
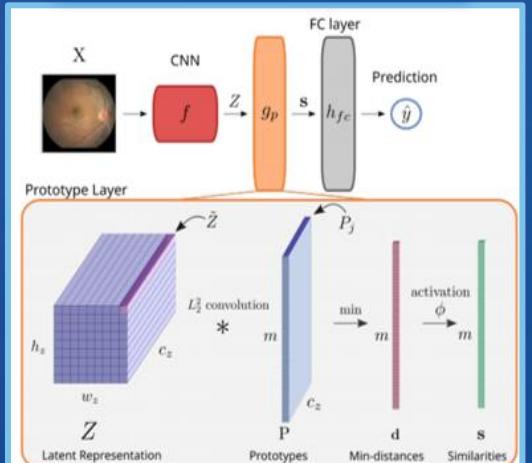
解决问题：解决传统扩散模型常见的色差问题

达到效果：还原眼底医学影像的原始色彩分布，提高生成结果的真实性

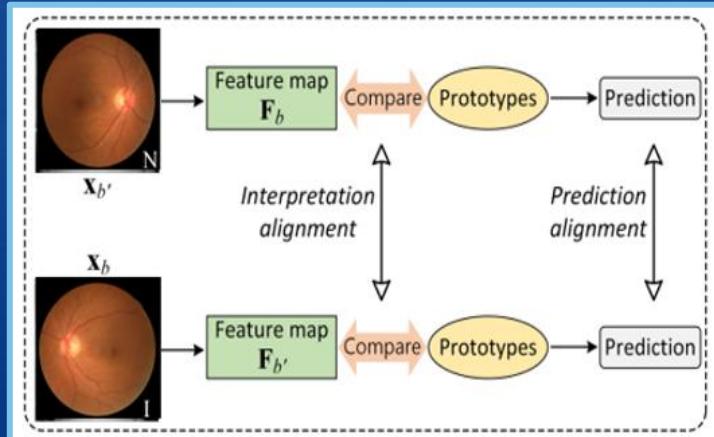
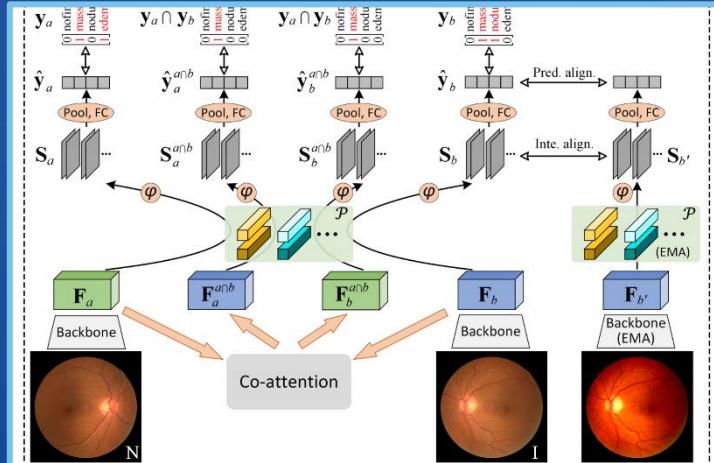
# 核心模型2——基于DPL的诊断框架

## 双原型学习 (DPL) 驱动的诊断框架

### 传统原型学习



### 项目提出的双原型学习



### 双原型学习与传统原型学习比较

双原型学习  
(本项目) VS 传统  
原型学习

标签区分 解耦区分 无法区分

病灶学习 全面清晰 无法识别  
复杂病灶

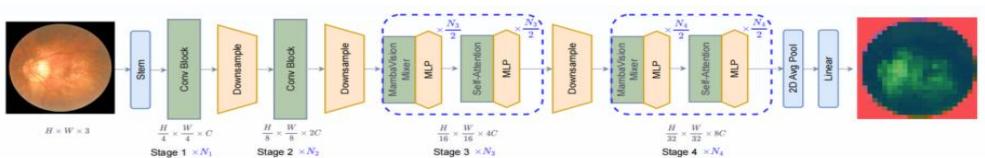
原型质量 可代表病灶 难以完全  
代表病灶

项目提出的双原型学习框架  
解决了传统原型学习  
难以区分多标签病灶的技术痛点

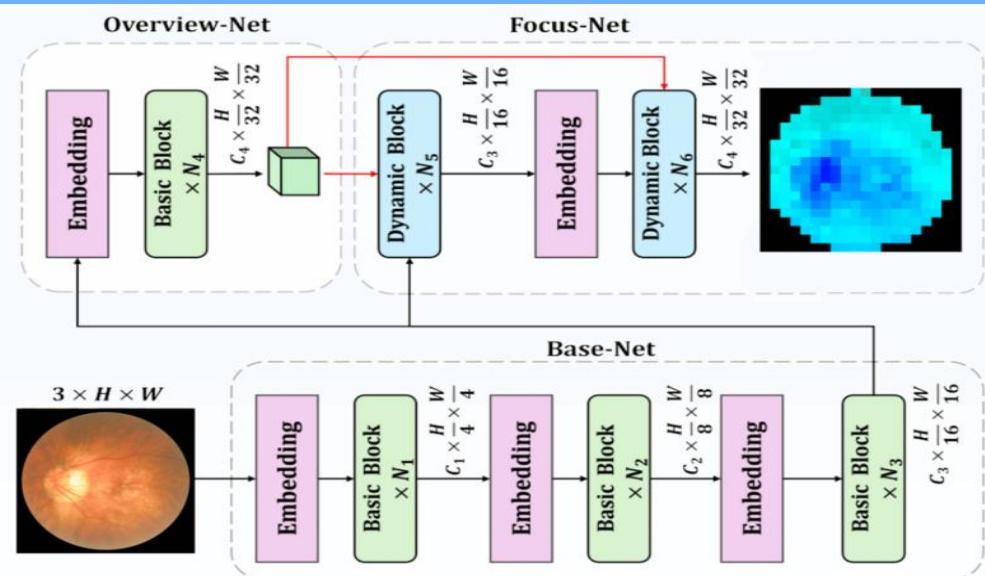
## 2.1 特征提取：双Backbone架构捕捉特征



### 双Backbone实现性能效率双提升

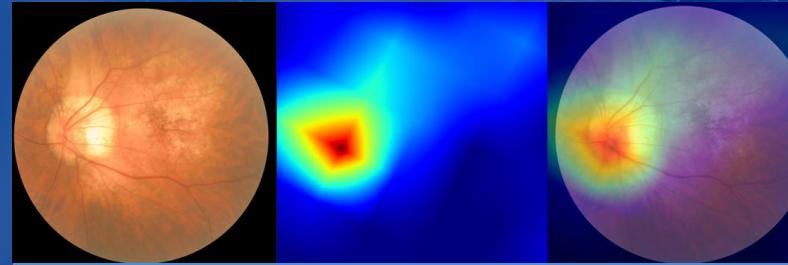


MambaVision高效捕捉眼底图像中的微小病变

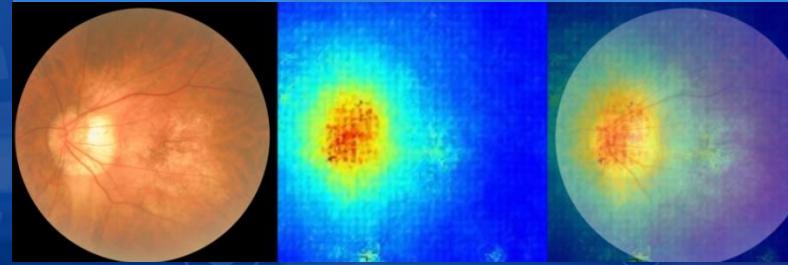


OverLoCK能够有效捕捉全局和局部特征

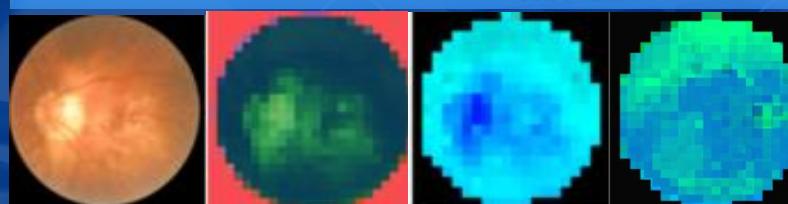
将 MambaVision 的全局上下文建模能力和 OverLoCK 的局部细节捕捉能力相结合，显著增强了特征的表达能力，并有效捕获了多尺度信息



MambaVision自注意力特征



OverLoCK全局特征捕捉能力



原图与特征融合图可视化

这种融合方案提升了模型在复杂场景中的鲁棒性和适应能力，在性能和效率上实现显著提升

创新成效

语义表达  
能力增强

全局与局部均覆盖

有效防止  
信息缺失

低层高层信息全融合

高效适应  
下游任务

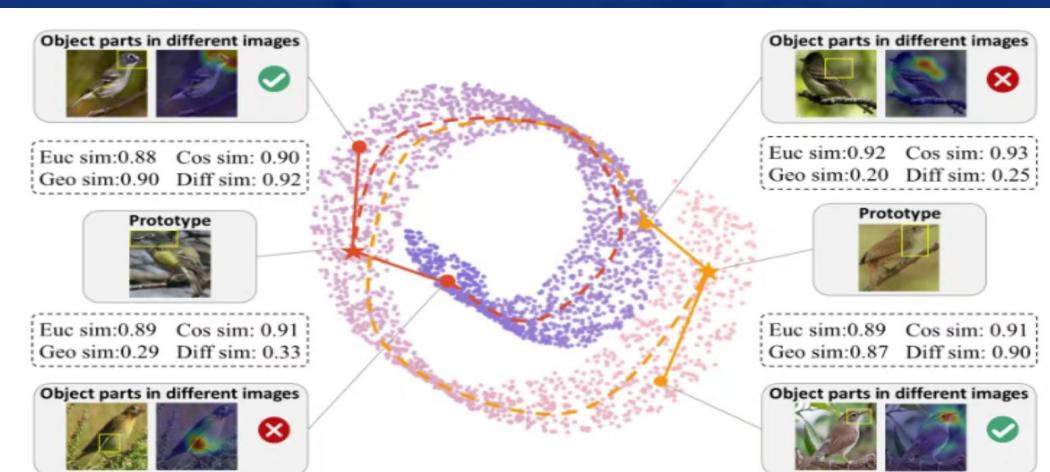
性能与效率双提升

## 2.2 原型匹配：提出GeoProto用流形几何提升可解释AI精度



### 告别欧氏距离“抄近道”！让原型匹配变得更符合语义

#### 传统模型VS我们的模型



传统模型：使用简单的欧氏距离比较图片特征原型  
我们的模型：不走直线，沿着数据本身所在的弯曲流形表面走，真正反映两个特征点的语义相似度。

特征空间往往是弯曲的，  
走直线“抄近道”反而会出错

#### GeoProto模型效果

Table 1: Classification accuracy (%) on two benchmark datasets. Best results are highlighted as first, second and third.

Method	CUB-200-2011 [6]						Stanford Cars [7]					
	V16	V19	R34	R50	D121	D161	V16	V19	R34	R50	D121	D161
Baseline	70.1	70.8	75.4	78.1	77.3	79.2	80.5	81.2	83.1	85.0	84.4	85.3
ProtoPNet [1]	69.4	71.8	71.5	80.3	73.1	74.9	79.6	80.4	80.7	84.5	81.2	82.1
TesNet [8]	74.8	76.7	75.4	85.4	78.3	78.7	81.8	82.5	81.8	87.1	83.4	83.1
ProtoPool [9]	74.3	74.7	75.4	82.6	77.5	79.7	80.3	81.1	81.1	84.6	82.4	83.2
ProtoKNN [10]	76.2	76.9	77.1	80.1	78.9	80.6	83.5	84.1	83.9	85.3	85.3	85.7
ProtoConcepts [11]	69.8	71.5	72.9	76.8	73.7	74.4	80.6	81.5	82.1	84.1	82.4	83.0
ST-ProtoPNet [2]	75.3	76.9	76.8	83.9	77.6	79.3	81.7	82.8	82.5	85.8	82.9	83.7
SDFA-SA [12]	75.7	77.1	76.8	85.6	80.3	79.9	81.6	82.3	81.8	86.6	84.2	83.4
MGProto [13]	78.8	79.0	80.1	86.2	80.3	82.1	83.4	84.4	87.2	84.2	85.7	
CBC [14]	78.3	78.6	80.3	85.5	79.6	82.3	83.6	83.7	84.9	86.9	83.9	86.0
ProtoArgNet [15]	77.9	78.7	79.2	85.1	79.1	81.5	83.0	82.9	84.2	85.0	84.0	84.3
GeoProto (Ours)	80.5	80.1	82.1	87.8	81.2	84.3	85.1	85.1	86.4	88.9	85.7	87.4

在CUB-200-2011（鸟类）和Stanford Cars（汽车）两个经典细粒度识别数据集上的准确率。

准确率全面超越了其他所有基于原型的可解释方法。在CUB数据集上使用ResNet-50时，GeoProto达到了 87.8% 的准确率。

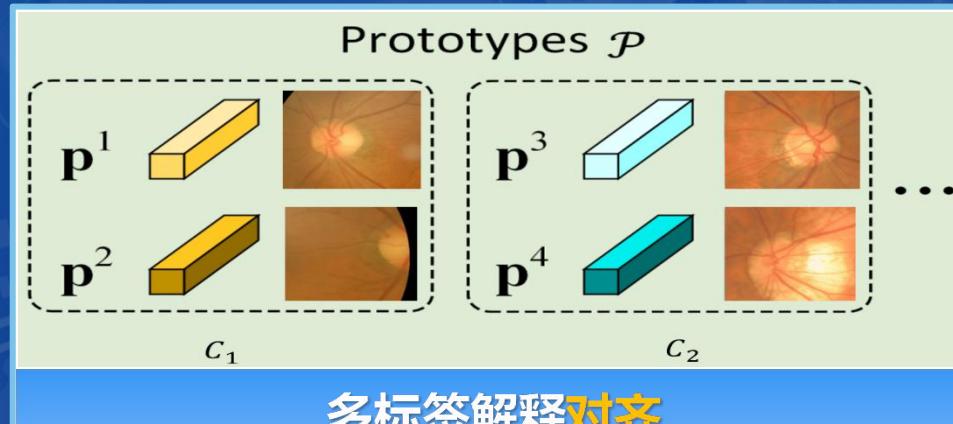
## 2.3 多标签训练稳健化：视图增强 + 预测对齐 + 标签关系建模



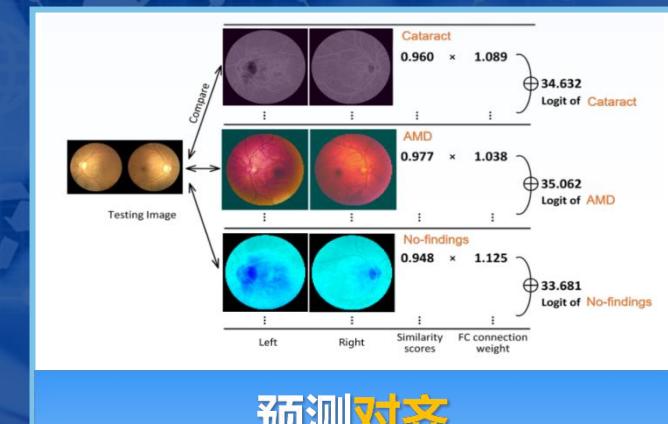
4个训练模块：提升数据多样性、保证同图预测稳定、建模标签共现关系”



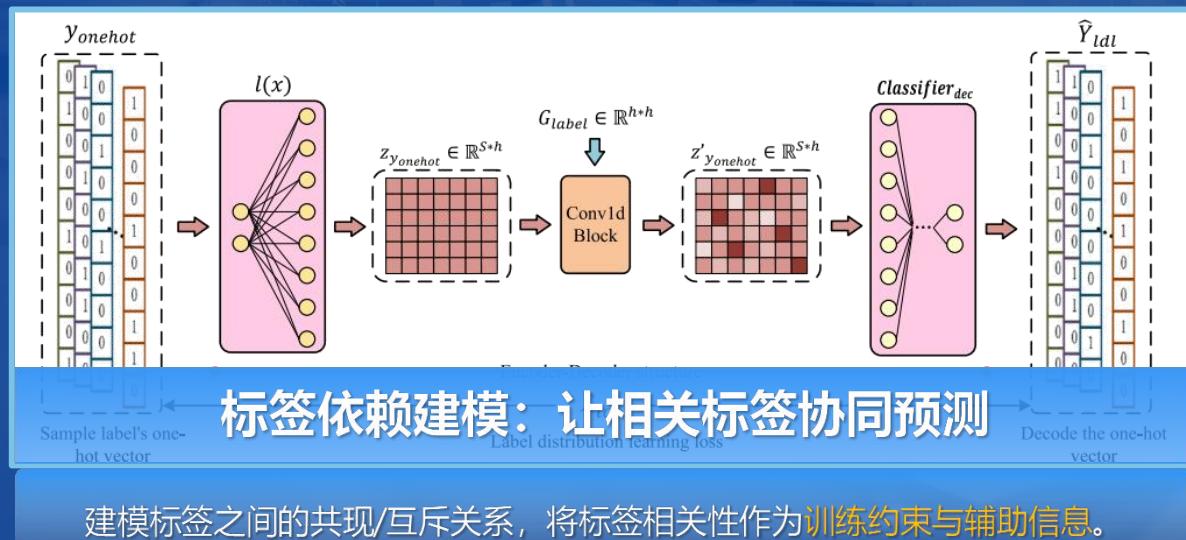
生成两个视图增强训练集多样性



解释对齐的目标是确保模型在不同视图下生成的相似性图，通过强制不同视图的相似性图对齐，可以增强模型解释的稳健性



计算视图之间的预测结果的对齐损失，确保模型在不同视图下生成一致的预测结果



训练数据  
多样性

视图增强

创新成效  
视图相似度  
一致性  
双对齐策略

增强样本  
标签关系  
标签分布学习

## 2.4 医疗大模型报告生成



### 本地化部署保证生成报告数据安全

#### 引入Chuanhu本地化部署架构

个人信息、病史、  
家族史、诊断结果

多种格式文件



患者信息数据库

原型库、激活图  
眼底图像token

合作  
医院

信息接入

本地化部署

个性化诊断报告

精准定位病变区域

可解释性的诊断

诊疗意见

#### 诊断报告生成

一键预测

预测结果

AMD/糖尿病/有髓神经纤维

糖尿病 0.816

青光眼 0.127

白内障 0.042

AMD 0.874

高血压 0.293

近视 0.178

其它疾病/异常 0.919

数据



定制化  
报告生  
成

报告

1. 老年性黄斑变性 (AMD)

诊断：黄斑区退行性病变，可能出现视力下降、视物变形和视野中心暗点。

治疗建议：

中期AMD：建议补充微量营养素（如维生素C、维生素E、锌等），并定期随访。

生活方式管理：戒烟、控制血压、定期使用Amsler方格表监测视力变化。

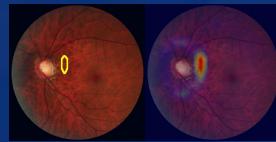
2. 糖尿病视网膜病变 (DR)

诊断：视网膜微血管损害，非增生型 (NPDR)，并发糖尿病性黄斑水肿 (DME)。

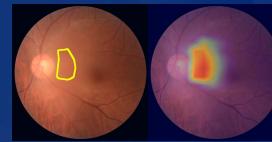
治疗建议：

## 原型库展示+AMD原型分析

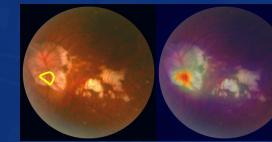
### 原型库展示



青光眼



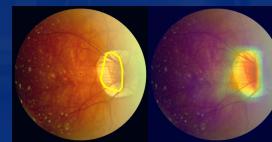
白内障



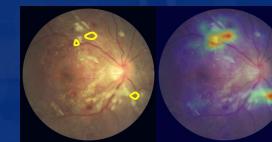
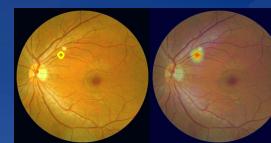
近视



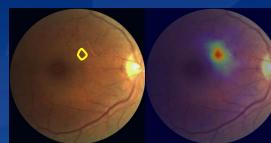
老年性黄斑变性



正常

其他疾病  
(视神经萎缩)

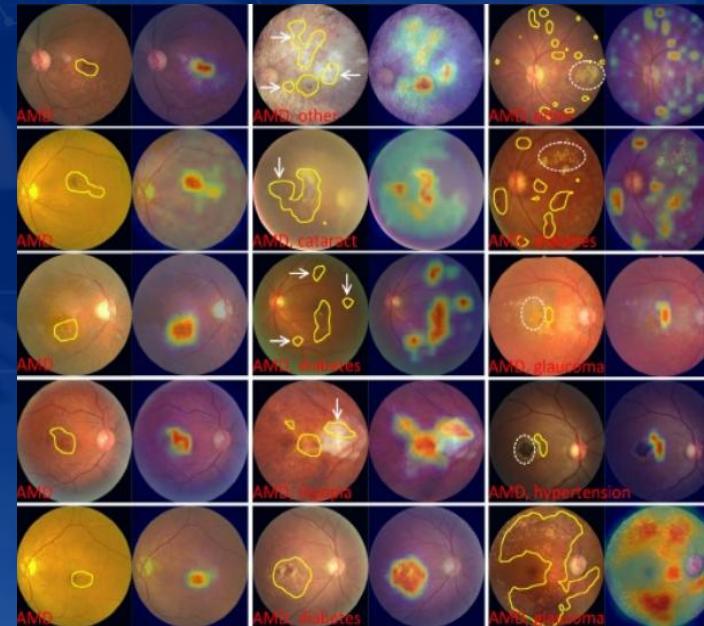
糖尿病视网膜病变



高血压性视网膜病变

涵盖7类疾病+正常类别的**大型原型库**

### 不同方法的典型AMD病变原型



其他对比方法生成的原型则容易与其他疾病混淆，  
在某些情况甚至完全忽略了实际的病变

# 开发工具



## 前端技术栈

前端



Vue.js 的灵活性与响应式;  
TypeScript 的类型系统和开发效率;  
Element Plus 的高质量组件库支持;

## 后端技术栈

后端



图像处理能力强  
(OpenCV, PIL);  
数据分析与建模能力优秀 (Scikit-learn, SciPy);

模块化设计  
易于组合

开源  
社区活跃

高扩展性  
可定制性

适合数据  
驱动型项目

## 深度学习建模的技术栈

模型



强大的灵活性  
(PyTorch);  
丰富的视觉支持工  
具 (torchvision);  
与后端图像处理库  
配合良好;

这套技术栈结合了前端的高效交互、  
后端的强大图像处理能力，以及模  
型层的深度学习支持，能够快速构  
建智能化、可视化的AI应用系统。

# 合作机构与系统展示



同合作机构建立紧密联系，不断优化系统

## 医疗机构



## 科研机构



## 多模态眼底影像数据库



智绘医疗 双目眼底智能诊断系统

左眼眼底图像

右眼眼底图像

报告

1. 老年性黄斑变性 (AMD)  
诊断：黄斑区退行性病变，可能出现视力下降、视物变形或视野中心暗点。  
治疗建议：  
中症AMD：建议补充微量元素营养素（如维生素C、维生素E、锌等），并定期随访。  
生活习惯建议：严格控制血糖、血压和血脂，以延缓疾病进展。  
PDT治疗：全视网膜光凝 (PRP) 或联合抗VEGF药物治疗。

2. 建议定期眼底检查 (DR)  
诊断：视网膜血管病变，非增生型 (NPDR)。并发糖尿病性视网膜病变 (DME)。  
治疗建议：  
控制基础疾病：严格控制血糖、血压和血脂，以延缓疾病进展。  
PDT治疗：全视网膜光凝 (PRP) 或联合抗VEGF药物治疗。

智绘医疗 双目眼底智能诊断系统-病历页

姓名	年龄	性别	身份证号	电话
张三	53	男	xxxxxxxxxxxxxx	xxx xxxx xxxx

日期	眼底	记录	诊断
2024-11-19		主述双眼酸胀、干涩。门诊可见豹纹状眼底，脉络膜毛细血管层萎缩，大血管裸露，弥漫性视网膜萎缩斑块	系统诊断：近视 (0.67) 医生诊断：中度病理性近视改变
2025-01-27		主诉酸胀好转，前来复查。门诊可见脉络膜淡斑，视网膜萎缩轻微	系统诊断：近视 (0.75) 医生诊断：中度病理性近视改变
2025-03-18		主诉双眼干涩。门诊可见动脉血管壁增宽，呈银丝状改变，动静脉比例轻微异常，交叉处静脉轻微压痕	系统诊断：近视 (0.76)，高血压 (0.82) 医生诊断：高血压病变轻度 (1级)

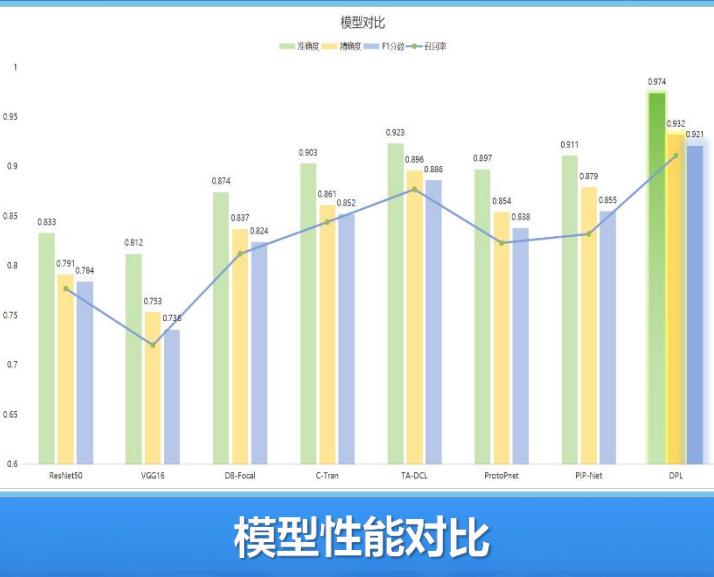
# 数据测试



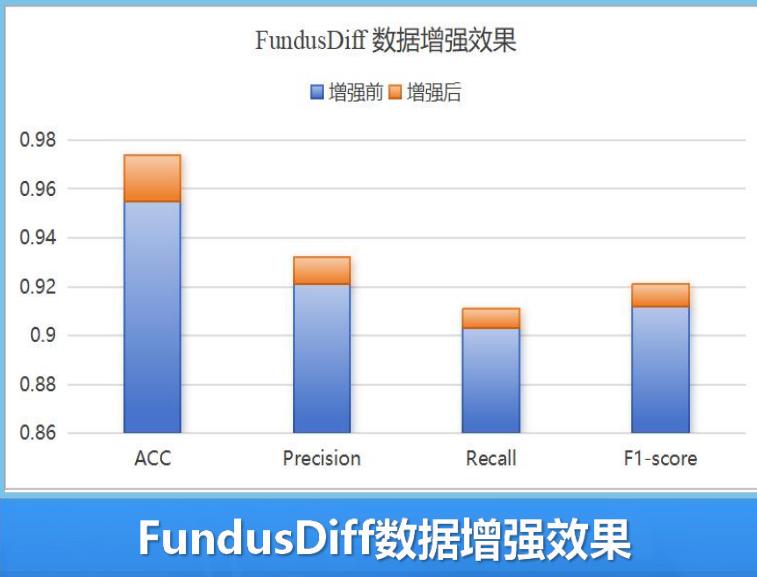
提出的DPL在所有指标均处于最优

方法	ResNet50	VGG16	DB-Focal	C-Tran	TA-DCL	ProtoPnet	PIP-Net	DPL (我们的模型)
可解释性	无	无	无	无	无	有	有	有
准确度	0.833	0.812	0.874	0.903	0.923	0.897	0.911	0.974
精确度	0.791	0.753	0.837	0.861	0.896	0.854	0.879	0.932
召回率	0.777	0.720	0.812	0.844	0.877	0.823	0.832	0.911
F1分数	0.784	0.736	0.824	0.852	0.886	0.838	0.855	0.921
来源	CVPR	ICLR	ECCV	CVPR	Media	NIPS	CVPR	自行研发

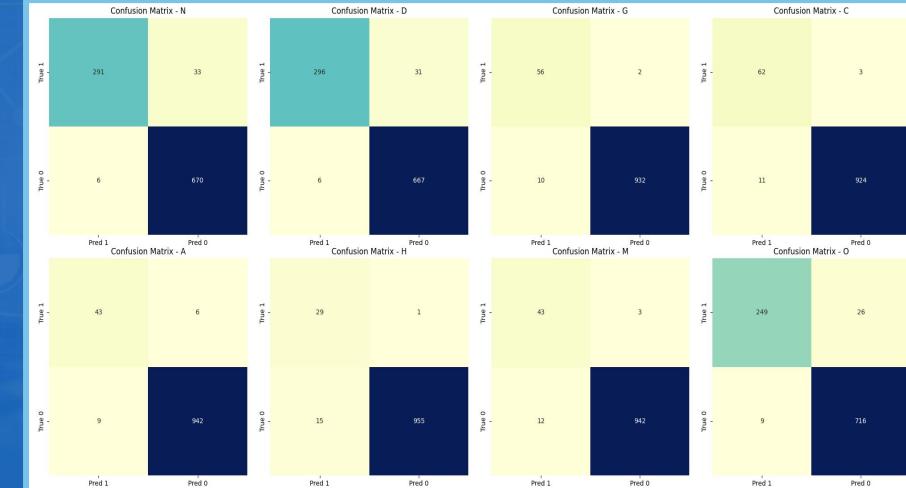
# 模型性能测试



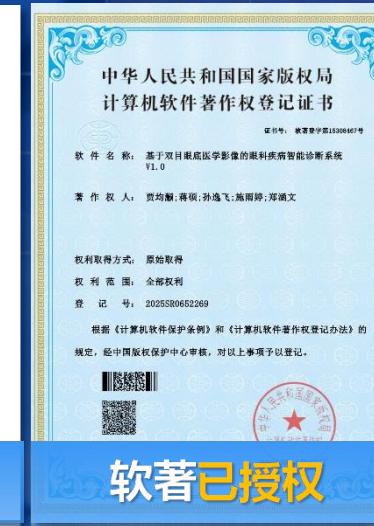
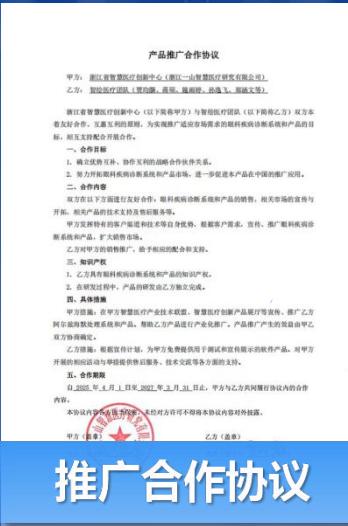
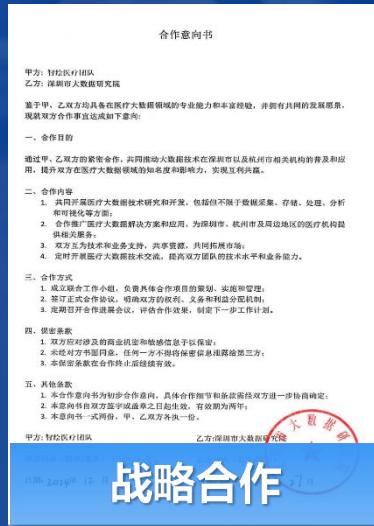
## 模型性能对比



## FundusDiff数据增强效果



## 各标签类别混淆矩阵



智绘医疗

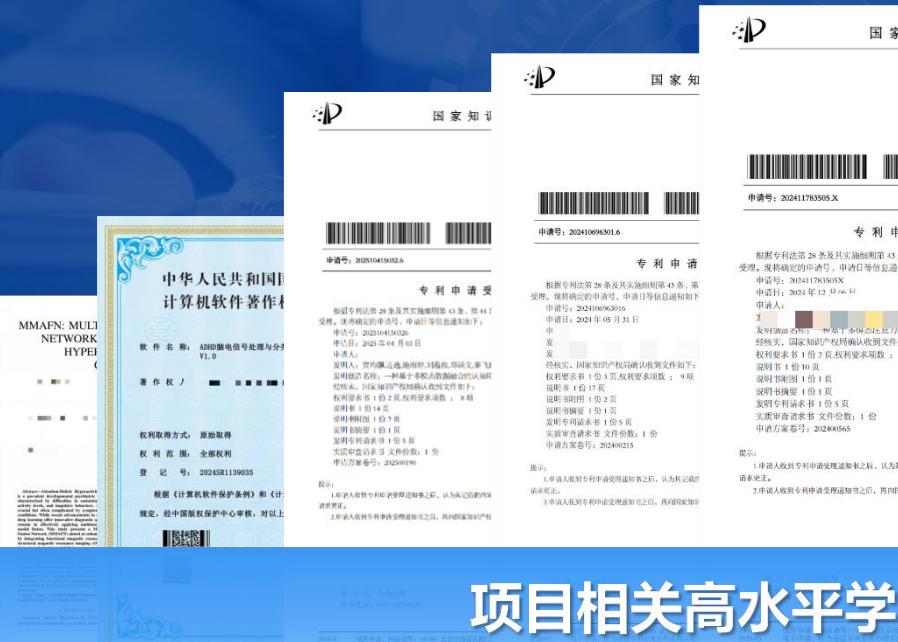
# 技术壁垒



## 5项国家发明专利，4项软著，2篇学术论文，构筑坚实技术壁垒

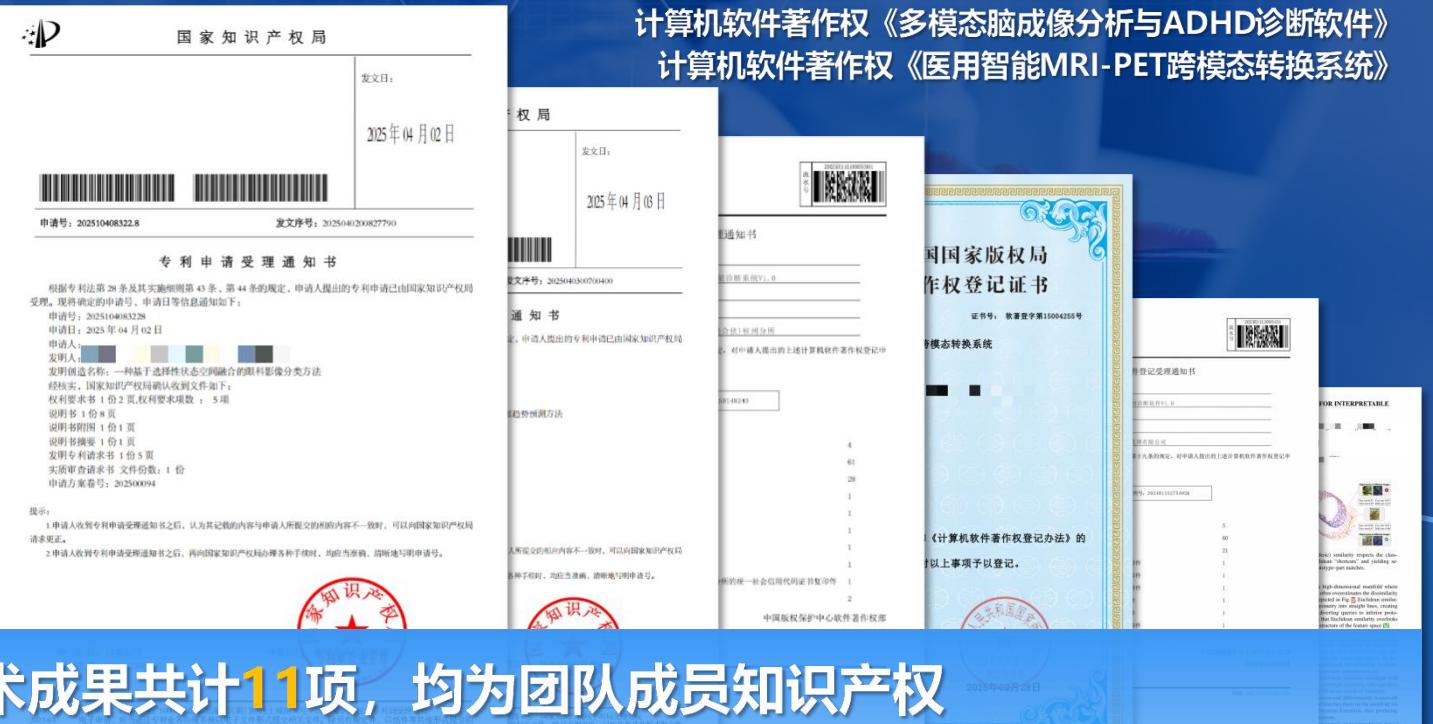
### 国家发明专利5项（其中团队成员第一发明人4项）

- 一种基于选择性状态空间融合的眼科影像分类方法（受理）
- 一种基于多模态注意力融合网络的 ADHD 分类诊断方法（受理）
- 一种基于多模态数据融合的认知障碍进展趋势预测方法（受理）
- 一种基于VMD-WmRMR-DF的注意力缺陷多动障碍分类方法（公开）
- 一种基于 GL-LCM 模型的高分辨率胸部 X 光图像骨抑制方法（受理）



### 其他知识产权/学术成果6项（均为智慧医疗领域）

- 计算机视觉领域国际顶级会议 DLCV Oral MMAFN: Multi-Modal Attention Fusion ...
- 信号处理领域国际顶级会议 ICASSP Geodesic Prototype Matching via Diffusion ...
- 计算机软件著作权《基于双目眼底医学影像的眼科疾病智能诊断系统》
- 计算机软件著作权《ADHD脑电信号处理与分类诊断系统》
- 计算机软件著作权《多模态脑成像分析与ADHD诊断软件》
- 计算机软件著作权《医用智能MRI-PET跨模态转换系统》



项目相关高水平学术成果共计11项，均为团队成员知识产权

# 拓展价值

## 探讨模型其他业务价值

实际问题



知识学习



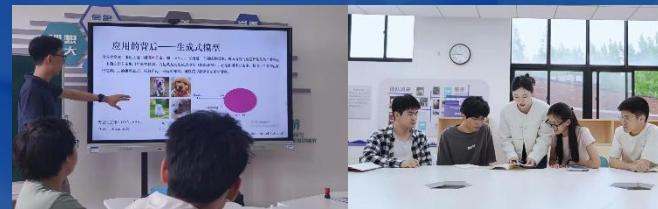
理论研究



实践创新



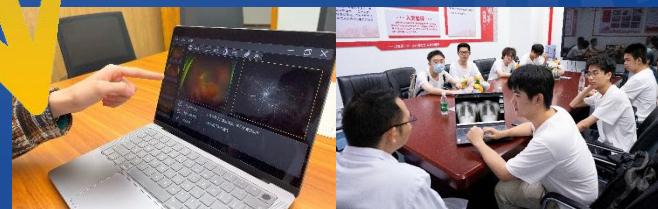
项目负责人带领团队实地调研



项目负责人组织开展科研交流

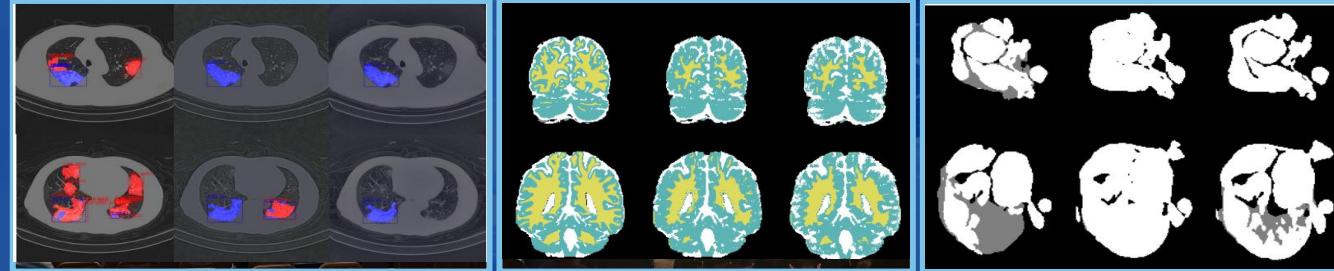


团队深入研究生物医学成像

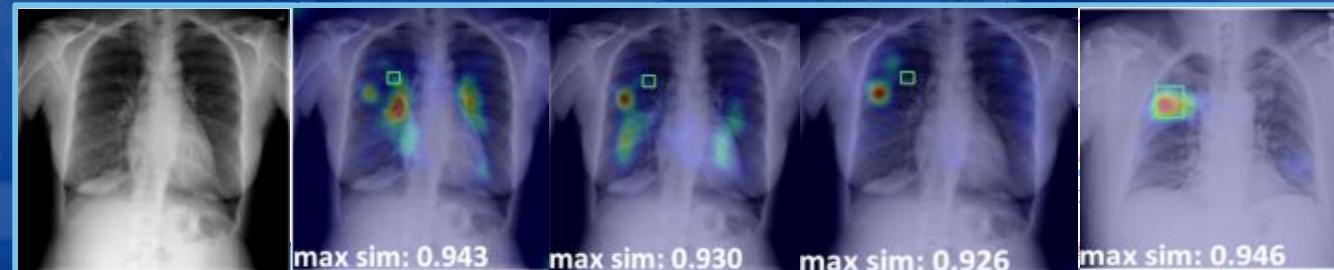


团队跟进临床评估与产品试用

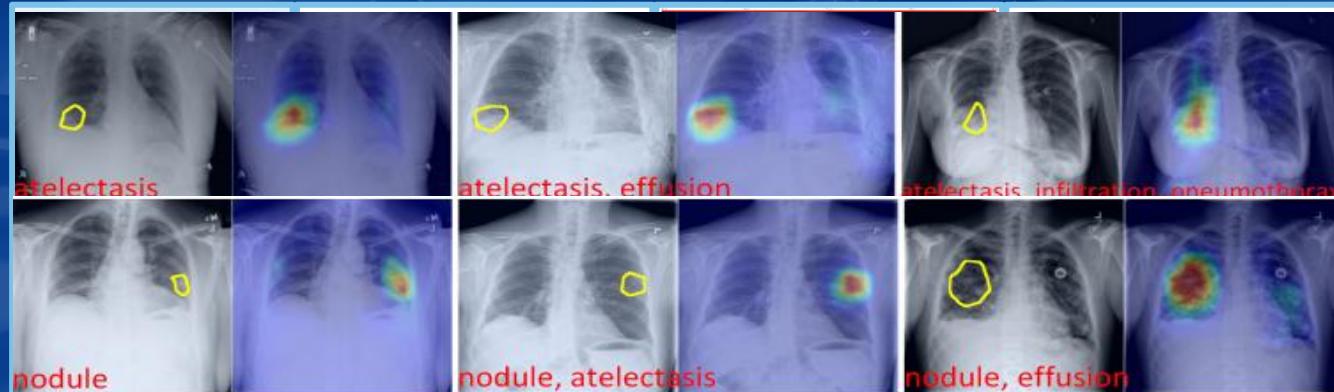
## FundusDiff在其他医学影像生成上的应用



## 相似度对齐在胸部X光成像上的应用



## DPL在胸部X光成像上的应用



# 专家顾问



深圳市大数据研究院  
Shenzhen Research Institute of Big Data

## 王昌林

深圳市大数据研究院副研究员

研究领域：智能视觉、医疗影像分析、  
计算机辅助诊疗、计算机视觉  
发表国际顶级期刊和会议30余篇  
担任多个国际顶级会议(CCF A类)审稿人



## 张少冲

深圳市眼科医院院长

清华大学深圳国际研究生院校外导师  
发表国际顶级期刊和会议论文150余篇  
教育部提名国家科学技术进步奖一等奖  
主编《眼科解剖与临床》等多部专著



浙江大学医学院附属第二医院  
The Second Affiliated Hospital Zhejiang University School of Medicine

## 陈芝清

浙医二院眼科中心副主任

研究领域：玻璃体视网膜疾病  
中华医学会眼科学分会眼病理学组副组长  
在国内外学术期刊发表论文30多篇  
中国医师协会眼科学分会眼病理学组委员

# 进程安排

## 深化医研协同 | 凝聚专家智策 | 创新多维突破

2家合作医院

3位专家顾问

5个项目亮点

 编写文档

梳理项目研发全流程文档，  
提交**专利与软著申请**，协同  
推进成果转化进程

 需求分析

通过**实地走访**多家医院眼科中心，系统梳理传统诊疗模式与深度学习应用间的协同瓶颈

选用**双原型学习**框架，  
采用标签分布学习等策略解决类别不均衡问题

 技术选型

 算法优化

设计FundusDiff**图像质量优化算法**，来解决原始数据集部分质量较低、部分标签较少等问题

借助前端、后端、深度学习技术栈开发了眼科疾病诊断系统

 功能开发

 测试与验证

成功获得由合作医院出具的**2份产品试用报告**，并取得深圳市眼科医院**医学伦理委员会审查批件**

分别与深圳市大数据研究院、浙江省智慧医疗创新中心签订**合作协议**，成功达成战略合作意向并建立**产品推广合作关系**

 反馈闭环

# 社会价值



## 医患价值



### 智慧医疗新产业

- 医疗资源配置情况大大优化
- 基层医疗压力得到有效缓解
- 推动“分级诊疗”有效落地

链接医患，智绘医疗

健康中国，惠及民生



### 患者就诊新体验

- 潜在眼底病灶更易发现更早干预
- 显著提升眼部疾病识别精度
- 眼科专家经验的依赖程度降低

## 产业价值



### 直接价值

- 减少误诊、漏诊的经济损失
- 减少30%以上的重复检查费用
- 催生智能眼科设备新赛道

5年内 直接 带动就业

100+

5年内 间接 带动就业

1000+



### 间接价值

- 推动健康产业的结构性升级
- 带动上游光学元件等产业链发展
- 为健康中国战略提供核心驱动力



浙江省大学生人工智能竞赛



智绘医疗



# 智绘医疗

人工智能新技术赋能眼科疾病诊断新场景，

我们一直在路上！