

案例分析-基于超声图像的疾病智能诊断

王瑞轩

<http://www.isee-ai.cn/~wangruixuan/>

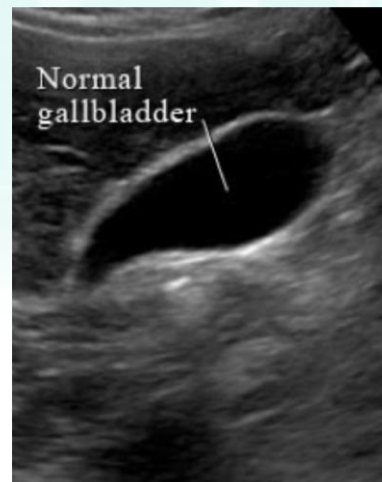
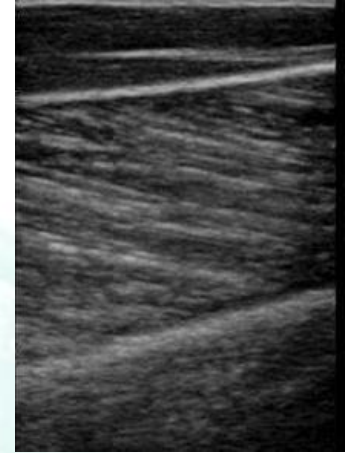
SUN YAT-SEN University



机器智能与先进计算
教育部重点实验室

Ultrasound Imaging

- ❑ 可以对身体各部位软组织成像





基于超声图像的胆道闭锁智能诊断

中山大学计算机学院与附属第一医院超声科合作完成

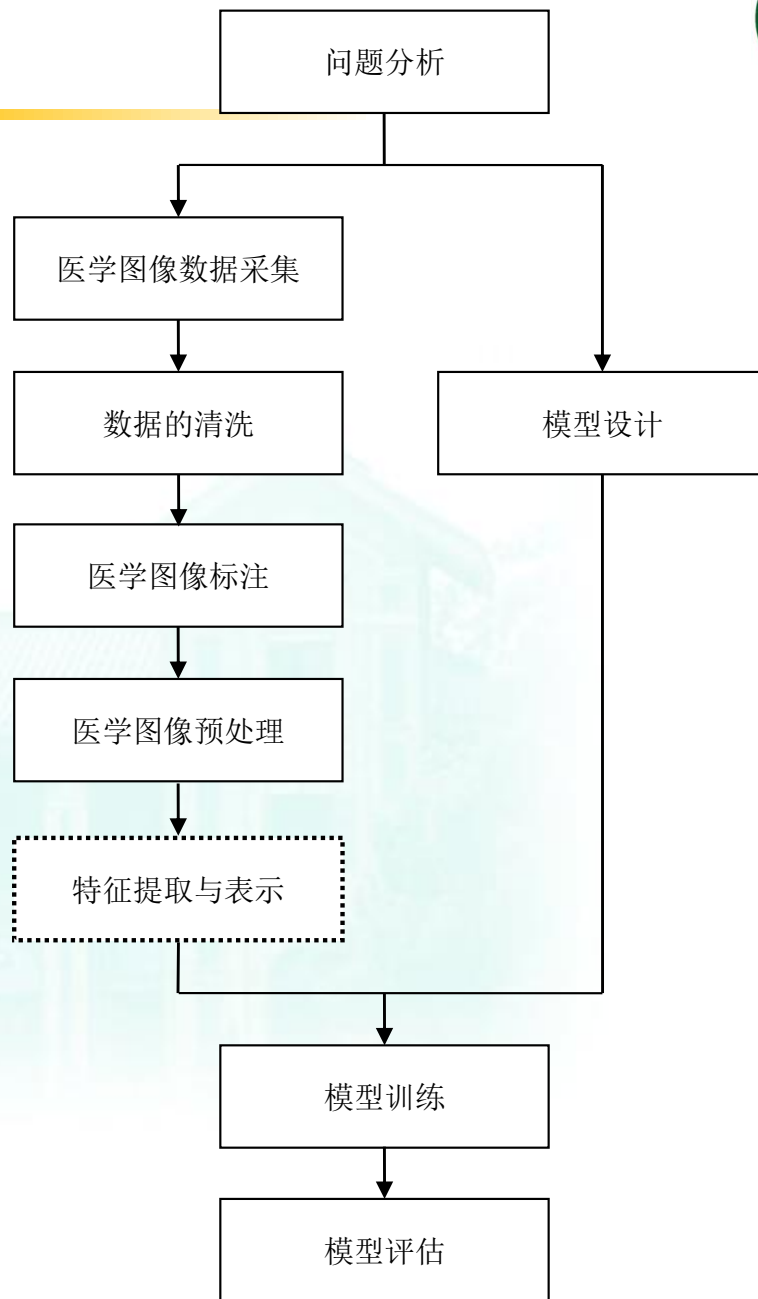
目标：利用人工智能技术实现基于超声图像的疾病诊断

项目背景：罕见病，有经验医生很少，基层医院无法诊断

意义：希望提升医生的诊断性能，尤其是帮助基层医院，
将AI技术真正服务于社会

医学图像分析基本流程

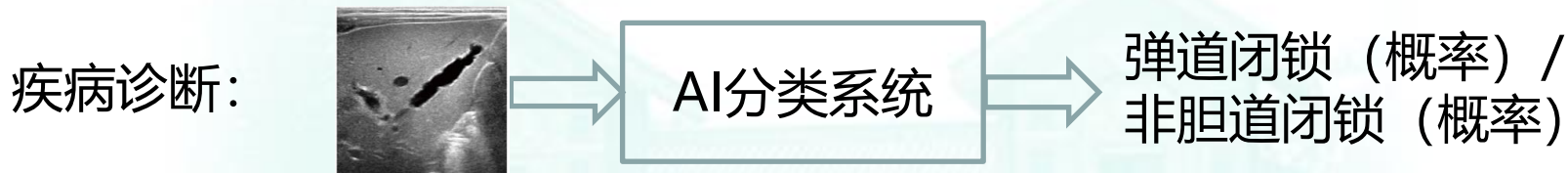
- ❑ 项目初期准备
- ❑ AI模型训练
- ❑ AI模型评估



项目初期：项目分析

❑ 目标：梳理医生需求，明确要解决的AI问题

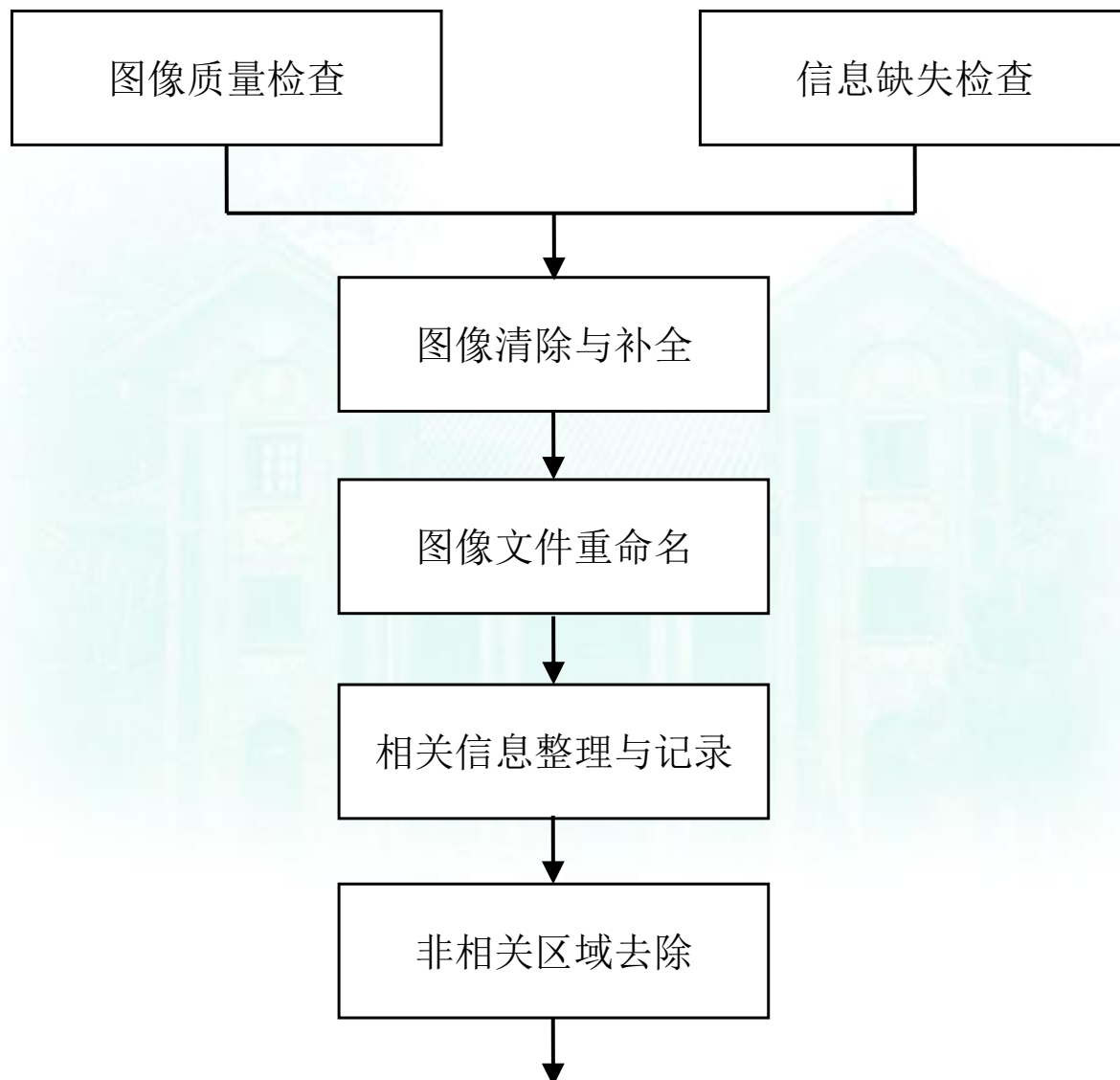
- 关键：交流讨论，站在对方角度思考
- 明确研究问题：AI系统输入与输出各是什么



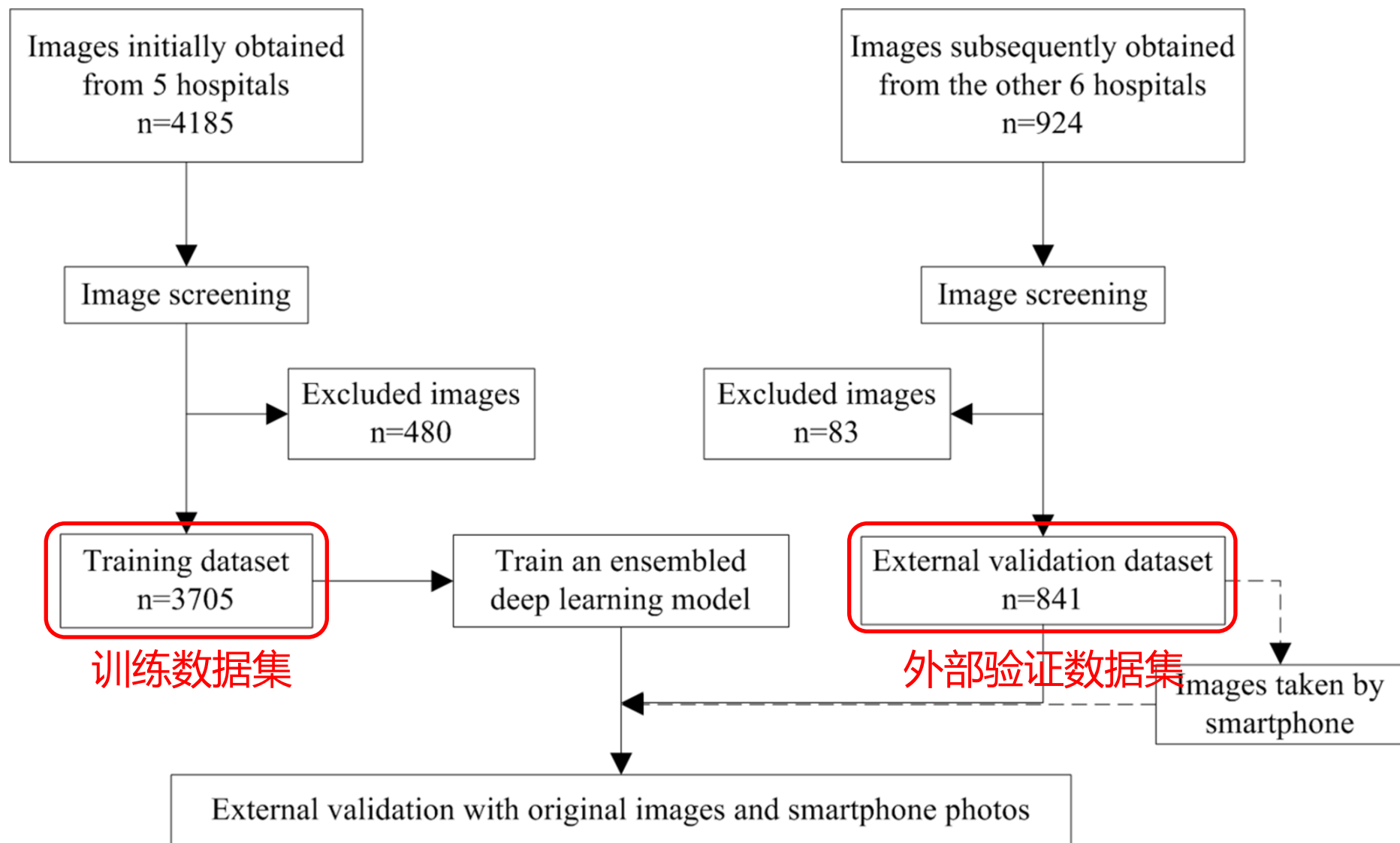
❑ 可能遇到的问题

- 医生实际需求不止一个 ——> 明确后选择一个
- 医生以为AI什么都能做 ——> AI很笨，还不如一个婴儿
- 医生以为不需要很多数据 ——> 需要足够多数据“教”AI模型
- 医生以为不需要标注信息 ——> 无标注，不AI

项目初期：数据收集与清洗



项目初期：数据收集与清洗



项目初期：数据收集与清洗

收集的原始图像（示例）

胆道闭锁



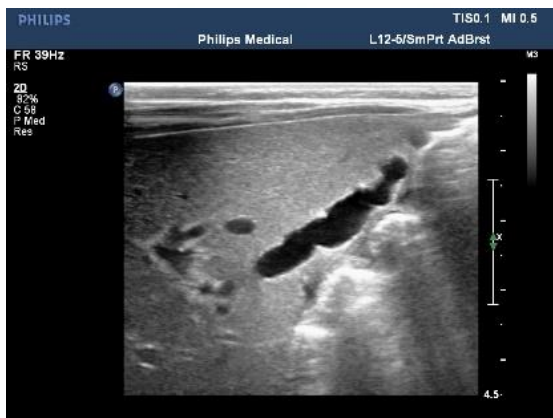
非胆道闭锁



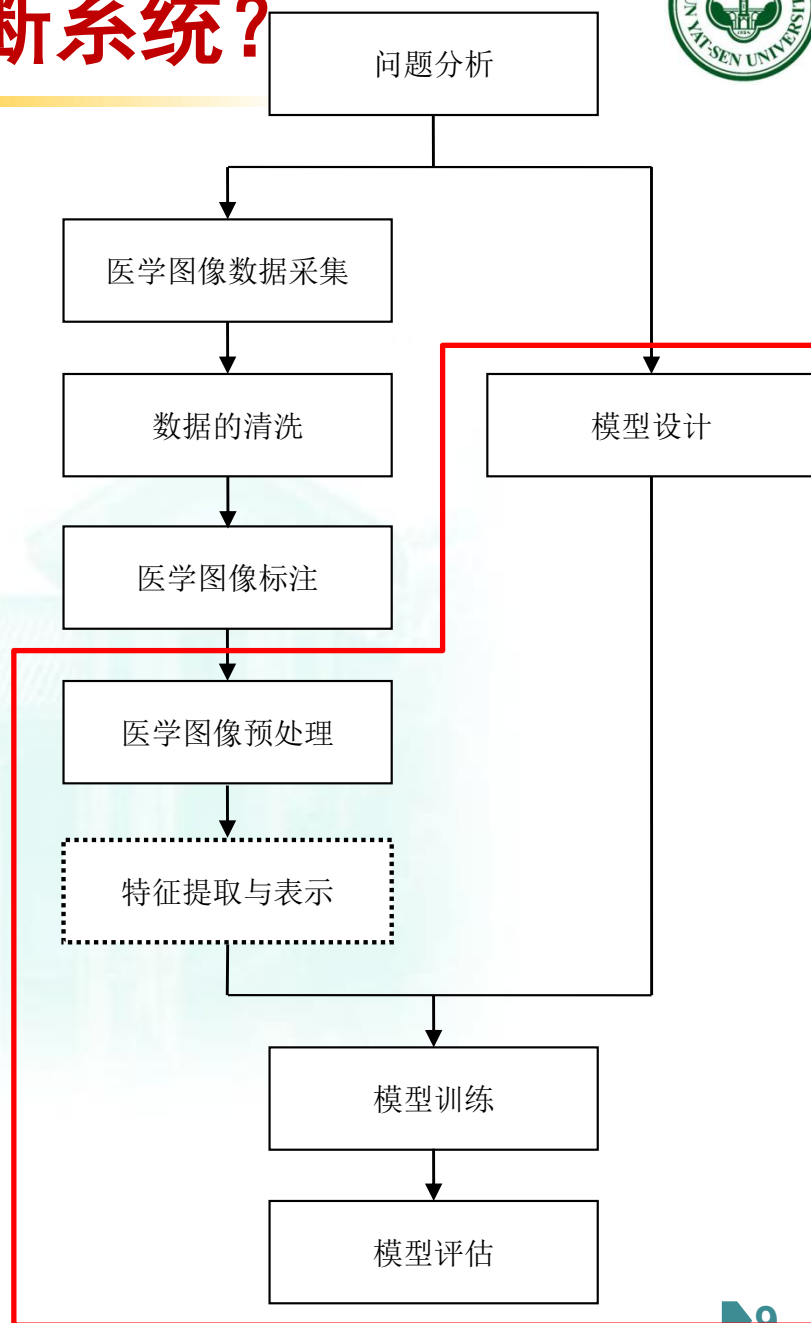
注意“类别不均衡”：胆道闭锁图像明显少于非胆道闭锁图像的数量

如何获得胆道闭锁智能诊断系统？

分组讨论 (10分钟)



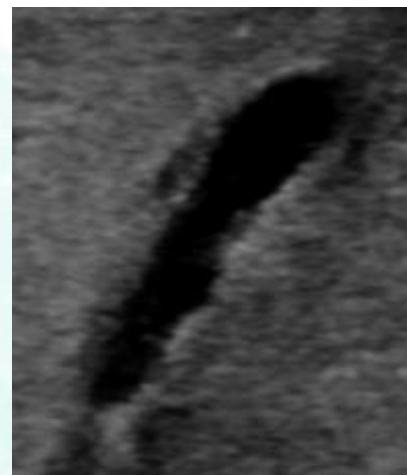
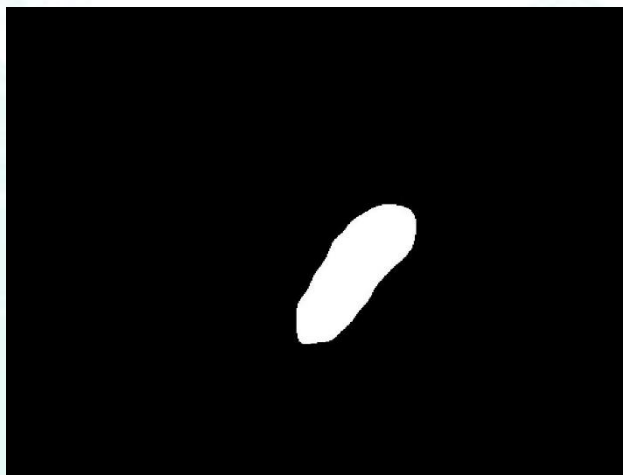
提醒：医生观察胆道闭锁区域及边缘进行诊断



AI模型训练：训练集

□ 图像预处理

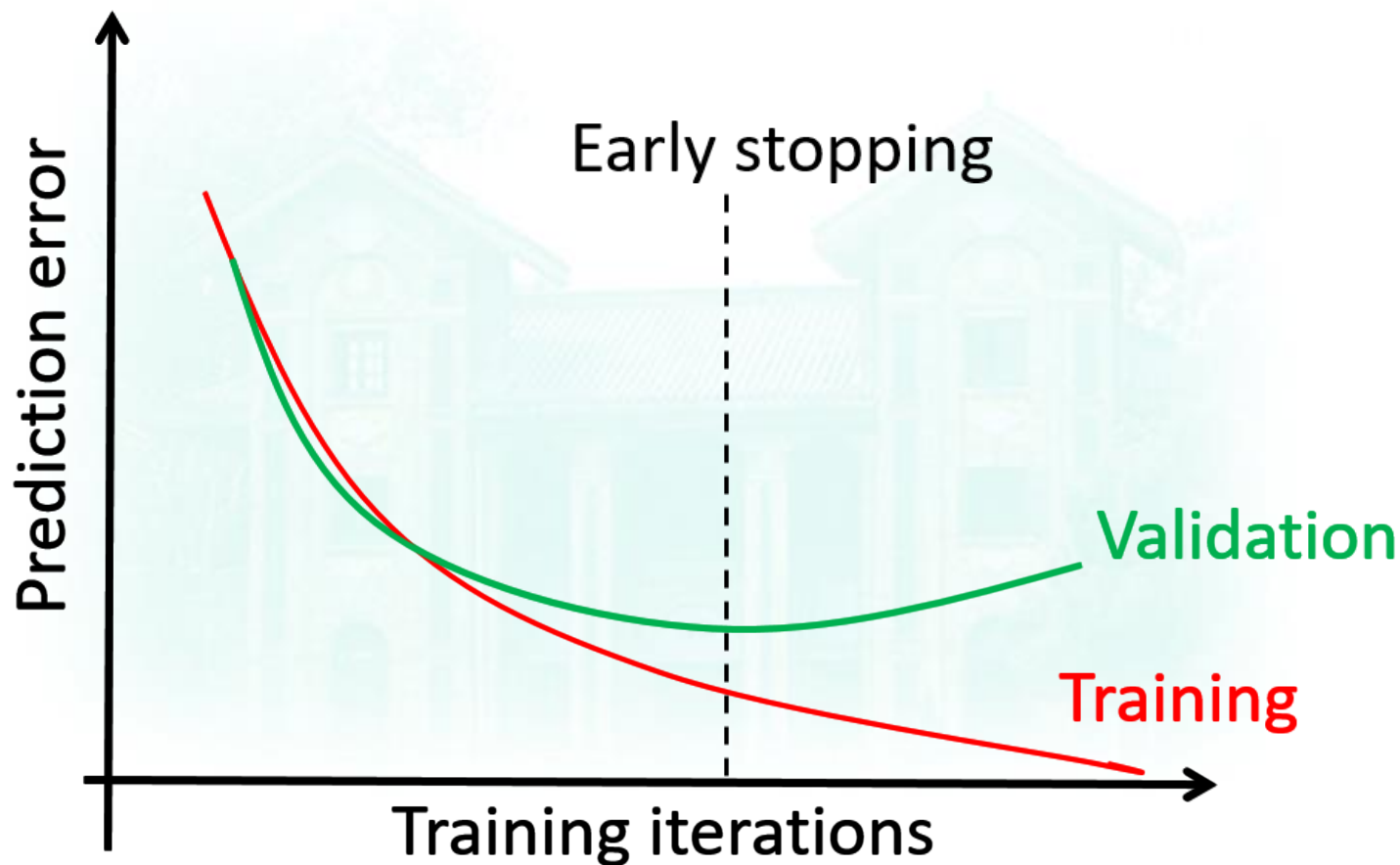
- 去除图像中不必要或干扰的区域
- 对像素值进行归一化（如：均值为0，方差为1）



- ## □ 原始训练数据集可划分为：训练集（80%）+验证集（20%）
- 训练集：用于训练模型
 - （内部）验证集：用于内部评估模型，决定何时停止训练模型

AI 模型训练：提升模型泛化性能

- 利用内部验证集：（内部验证集来自哪里？）



AI 模型训练：提升模型泛化性能

□ 数据增广 (Data augmentation)

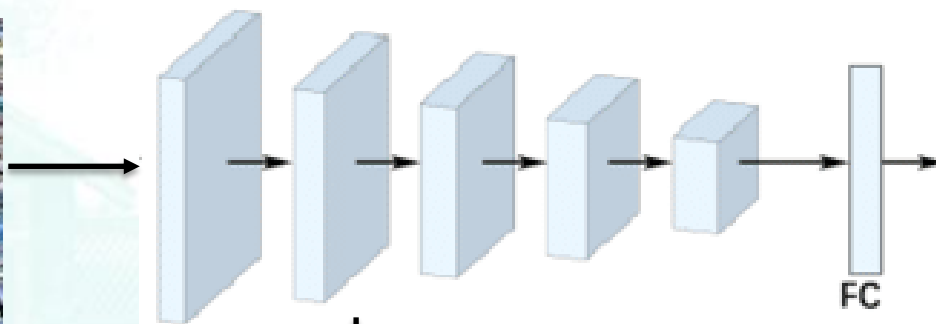


旋转，缩放，平移，左右翻转，变形，颜色与亮度改变等

AI 模型训练：提升模型泛化性能

□ 迁移学习 (Transfer Learning)

Step 1: 利用大规模自然图像预训练模型



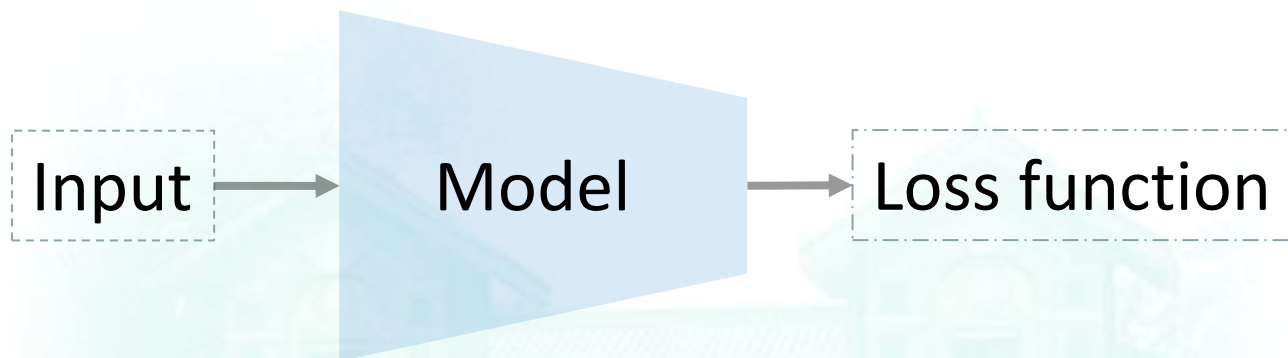
Step 2: 迁移模型大部分参数



Step 3: 利用超声图像数据训练微调模型参数

AI 模型训练：提升模型泛化性能

❑ 不平衡分类：重平衡策略



Oversampling;
Augmentation

Transfer learning;
Ensemble model

Class weighting
(Cost sensitive learning);
Instance weighting
(e.g., focal loss)

AI 模型训练：提升模型泛化性能

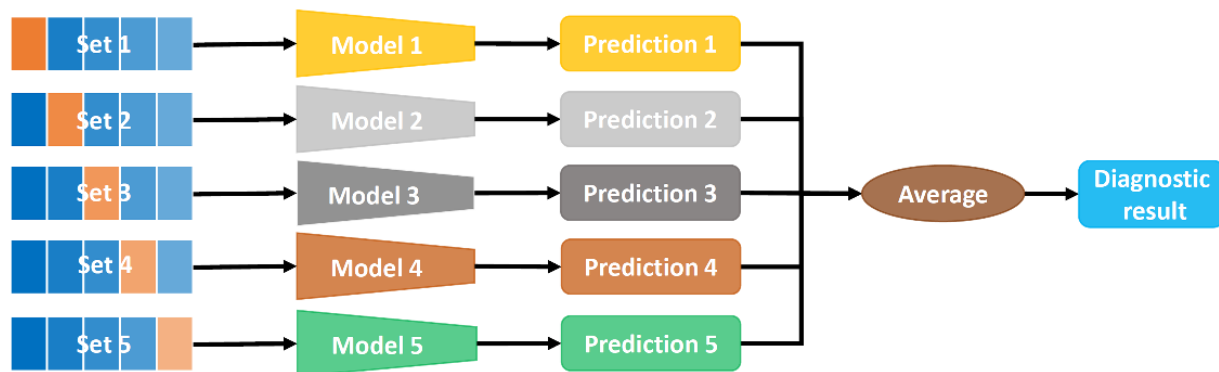
❑ 集成多个AI模型

- 每个模型：80%训练集来训练，另20%决定何时停止训练
- 预测时利用多个模型预测结果的平均值

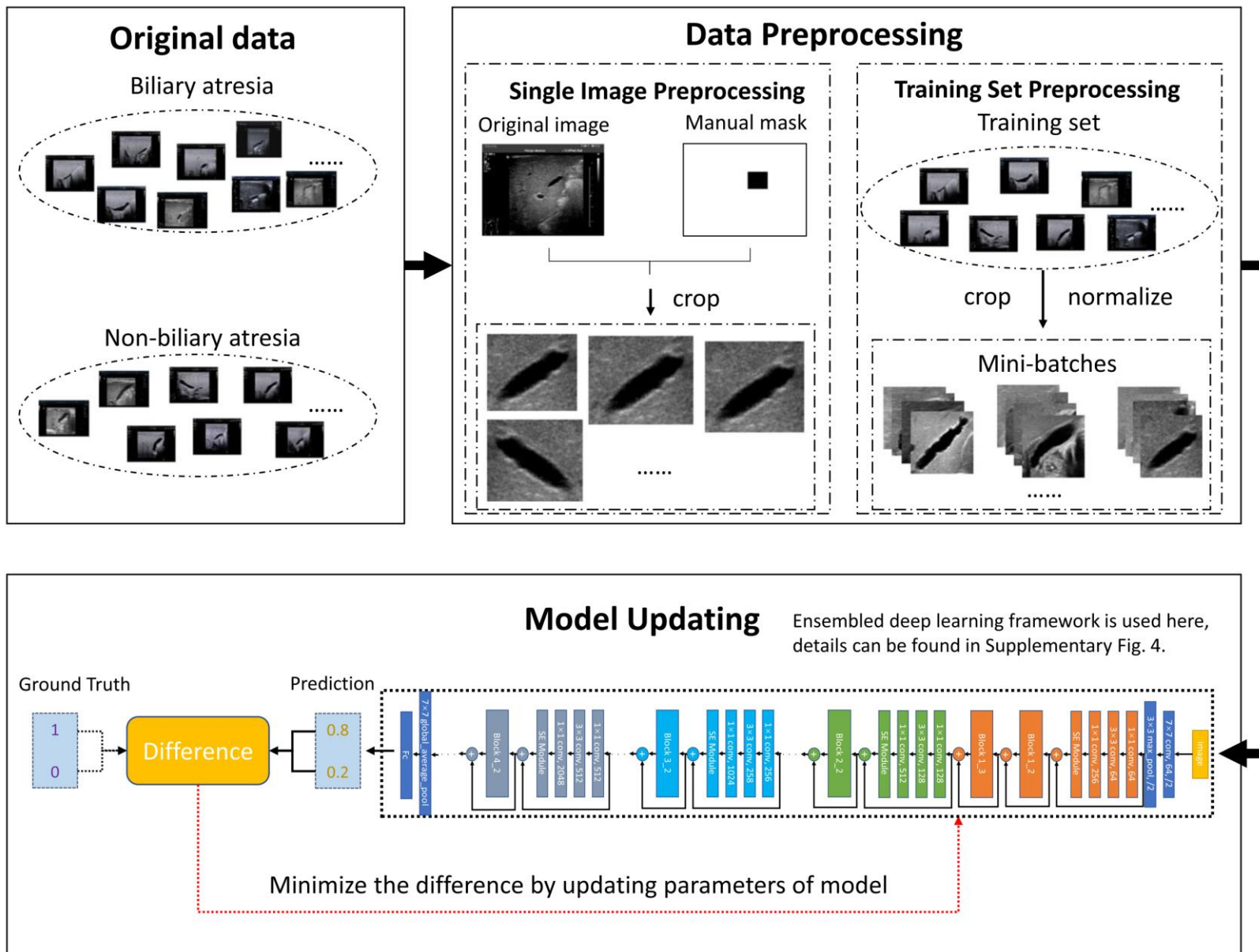
5-Fold Cross Validation

Set 1	Validation	Fold 2	Fold 3	Fold 4	Fold 5
Set 2	Fold 1	Validation	Fold 3	Fold 4	Fold 5
Set 3	Fold 1	Fold 2	Validation	Fold 4	Fold 5
Set 4	Fold 1	Fold 2	Fold 3	Validation	Fold 5
Set 5	Fold 1	Fold 2	Fold 3	Fold 4	Validation

Ensembled deep learning framework

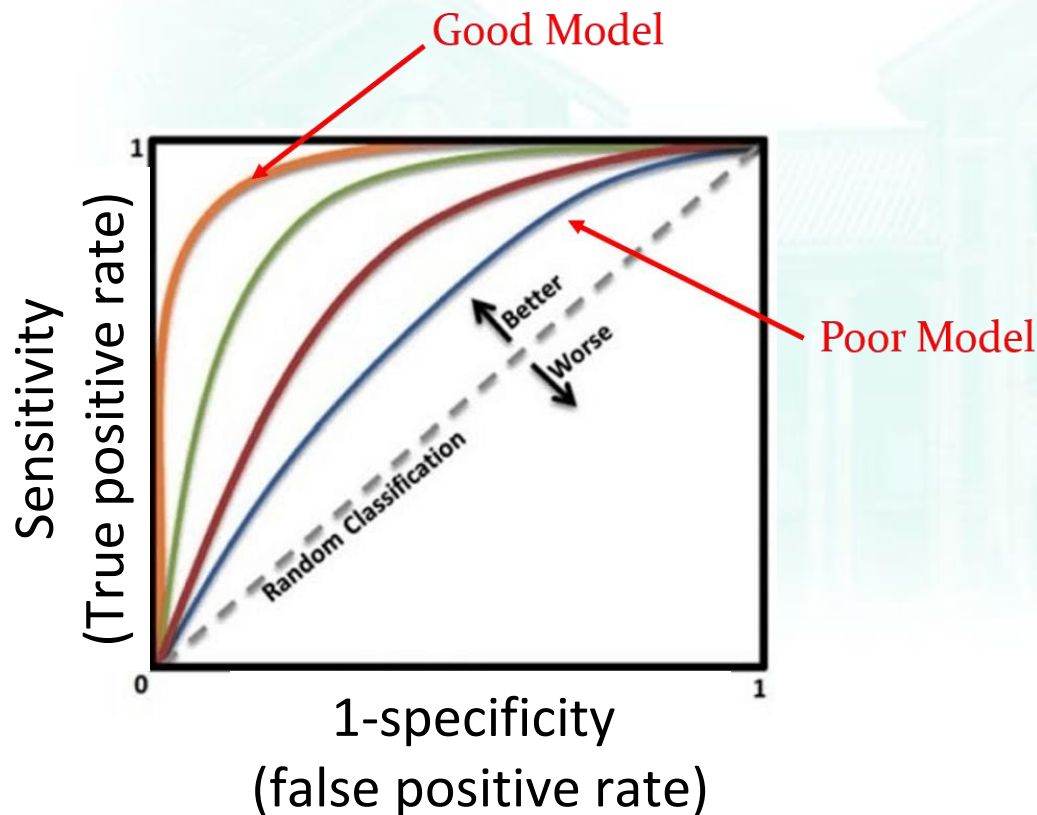


AI 模型训练：单个模型训练流程



AI模型评估：评估指标

- 定量评估指标(注：不同任务的评估指标可能不同)
 - 敏感性：所有真正阳性病例被检测为阳性的百分比
 - 特异性：所有真正阴性病例被检测为阴性的百分比
 - ROC曲线及曲线下面积 (AUC)



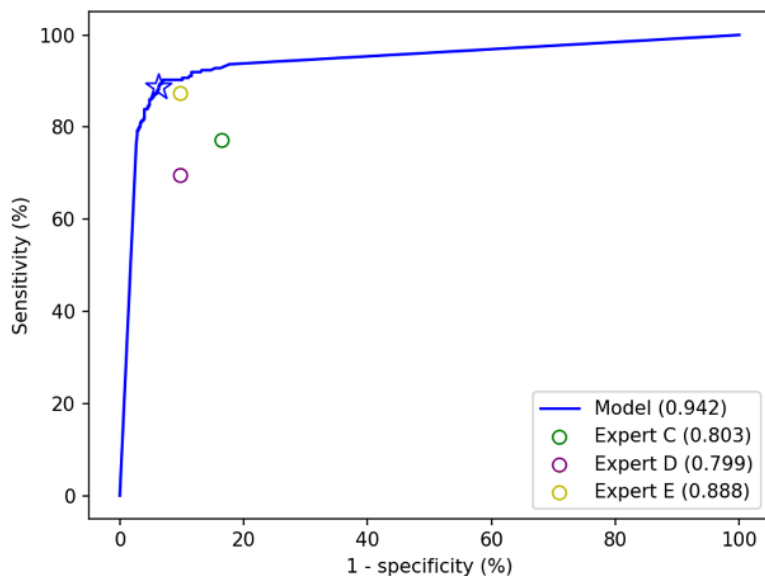
问题：如何根据模型预测结果与真实患病结果得到ROC曲线？

Hint: 模型预测输出是概率值

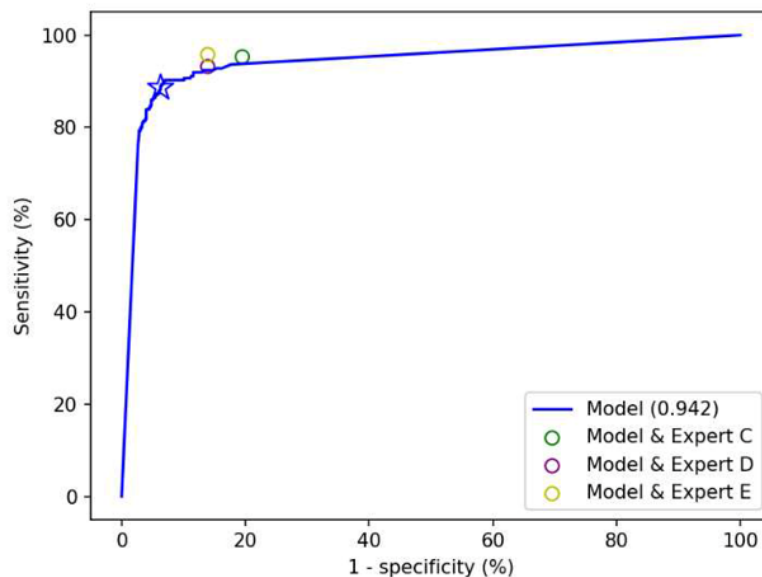
AI模型评估：与专家对比

- ❑ 模型 pk 专家
- ❑ 模型+专家 pk 专家/模型

AI模型好于专家诊断性能

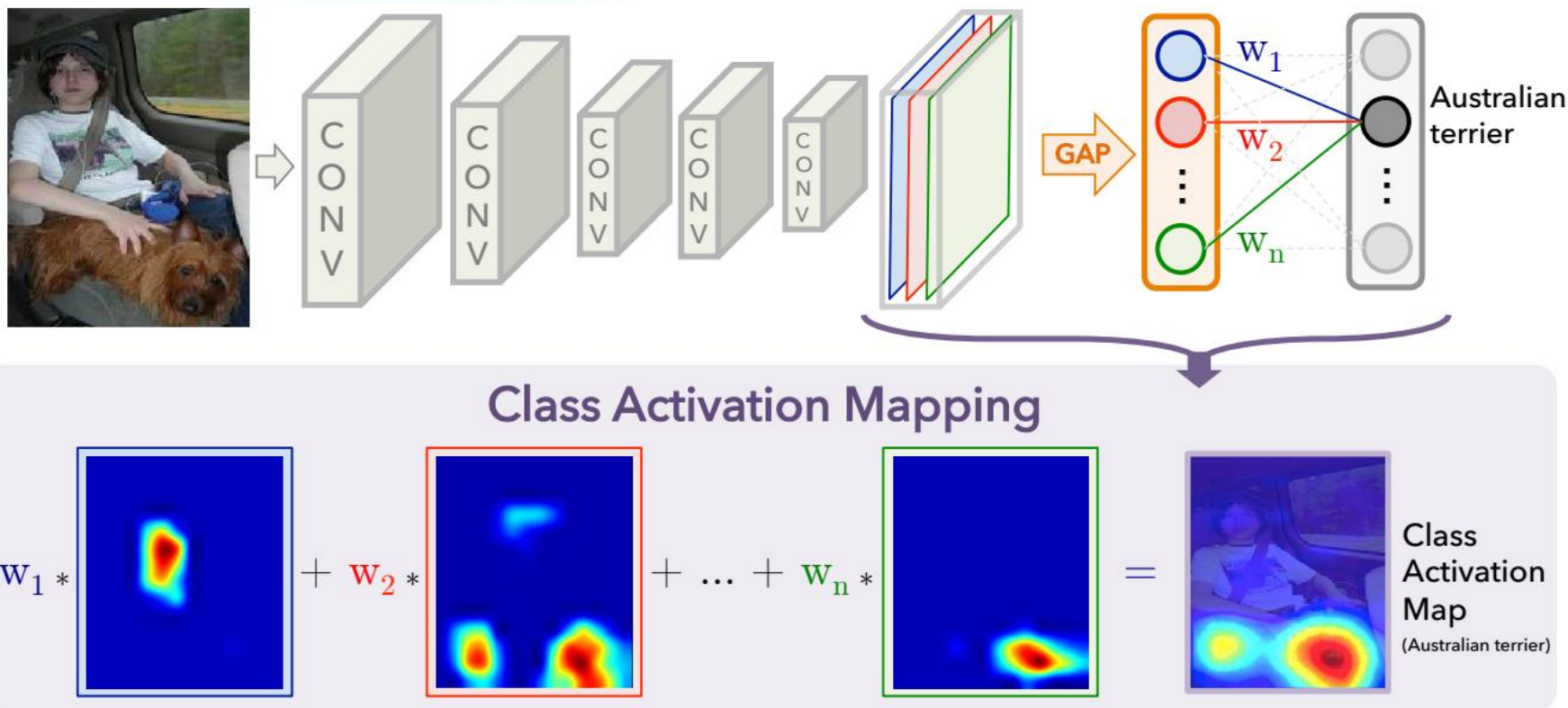


AI模型+专家好于模型
(在敏感性方面)



AI模型评估：解释模型的每次预测

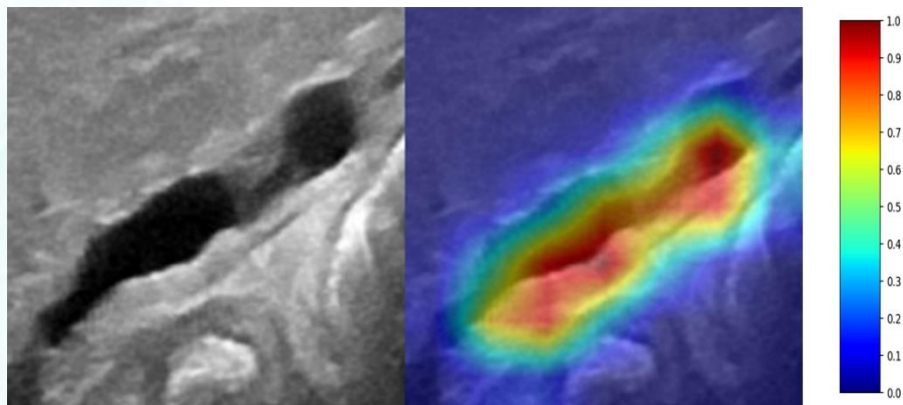
- 类激活图（CAM）：最后卷积层输出特征图的特定线性组合可解释模型预测时关注的图像区域



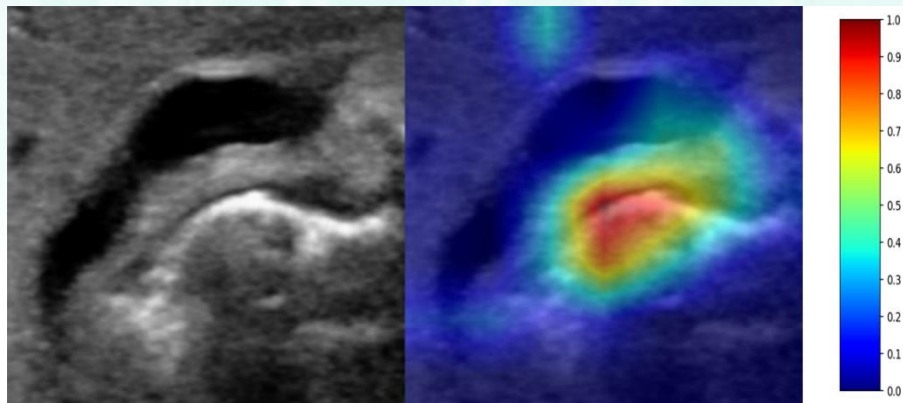
AI模型评估：解释模型的每次预测

- AI模型预测的初步解释
 - AI模型与医生关注区域：99%预测时一致

诊断正确，
关注区域一致

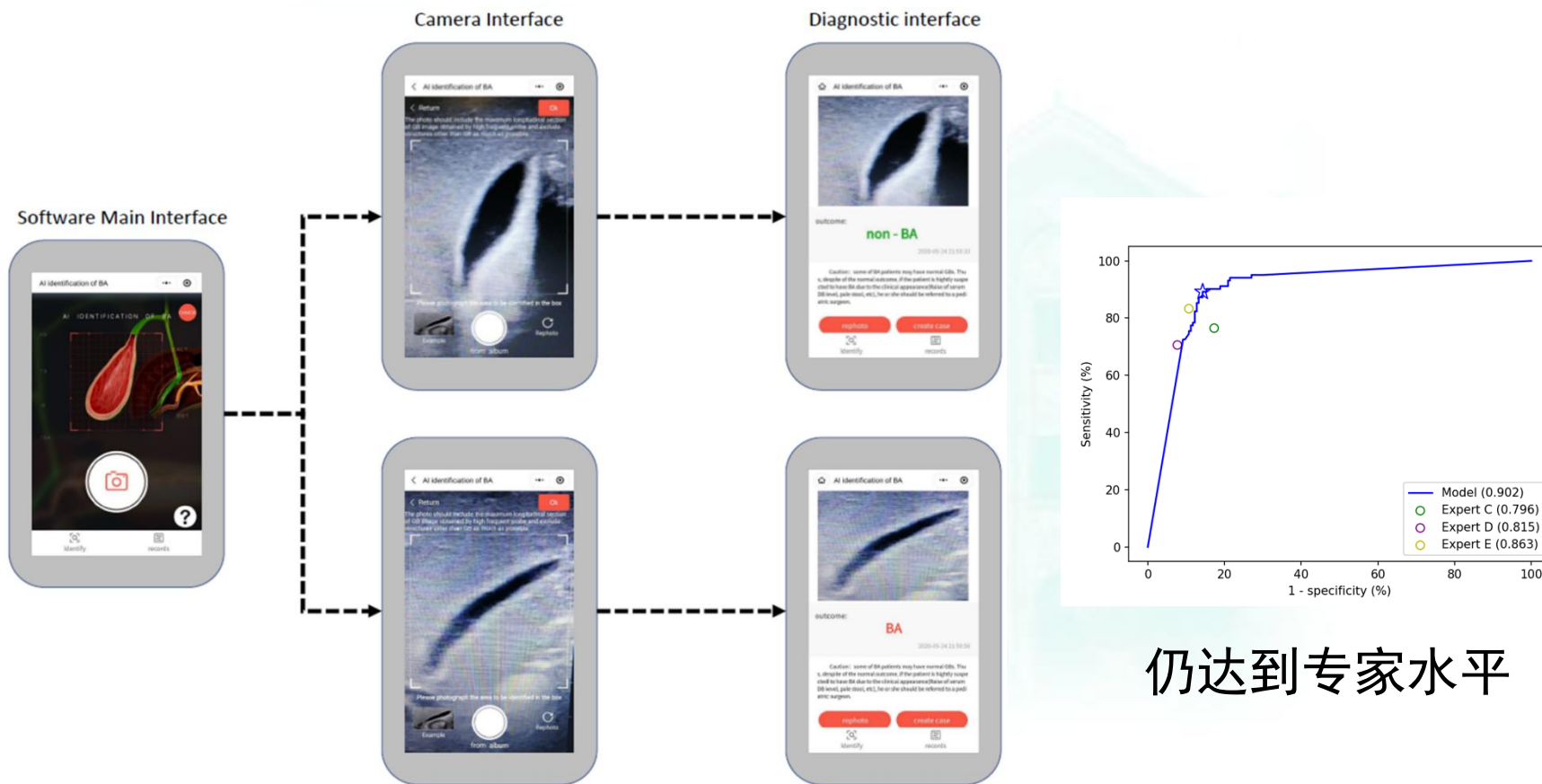


诊断正确，
关注区域不一致



应用及挑战

- 手机APP：手机拍照上传云端，实时显示诊断结果



仍达到专家水平

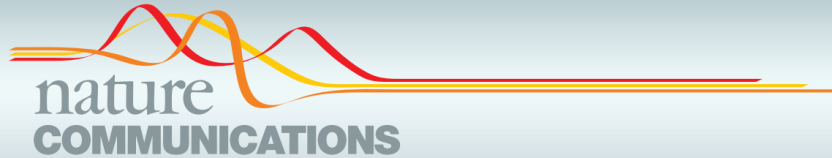
应用及挑战

- ❑ 手机APP面临的挑战：拍摄图像质量较差



- ❑ 其它改进：结合更多数据（如临床指标）提升诊断性能

该项工作发表在《自然-通讯》子刊
















ARTICLE



<https://doi.org/10.1038/s41467-021-21466-z>

OPEN

Ensembled deep learning model outperforms human experts in diagnosing biliary atresia from sonographic gallbladder images

Wenying Zhou ^{1,12}, Yang Yang ^{2,12}, Cheng Yu ^{3,12}, Juxian Liu ^{4,12}, Xingxing Duan ⁵, Zongjie Weng⁶, Dan Chen ⁷, Qianhong Liang⁸, Qin Fang ⁹, Jiaojiao Zhou ⁴, Hao Ju ¹⁰, Zhenhua Luo ¹¹, Weihao Guo¹, Xiaoyan Ma⁷, Xiaoyan Xie ^{1,13}✉, Ruixuan Wang ^{2,13}✉ & Luyao Zhou ^{1,13}✉

Source code: <https://github.com/youngyzzZ/Sonographic-Gallbladder-Images-for-BA-Diagnosis>