



多模态眼底图像合成

李锡荣

中国人民大学 数据工程与知识工程教育部重点实验室
中国人民大学 信息学院 人工智能与媒体计算实验室

2022年6月18日

北京医师协会眼科专科医师分会AI眼科高峰论坛



眼科人工智能的数据困境

AI算法（deep learning）仍然是数据驱动（data-driven）

获取数据的主要途径

- 公开数据集（选项有限，且易过时）
- 自建数据集（耗时耗力，难以持续）

第三条道路？



未来的训练数据将是合成的？

Gartner®

Information Technology

Roles Experts Research & Tools Insights Events Client Success Stories



[Blog home](#) > [Blog post](#)



By 2024, 60% of the data used for the development of AI and analytics projects will be synthetically generated



By [Andrew White](#) | July 24, 2021 | 0 Comments

[Search all blog posts](#)

[Synthetic Data](#) [Synthetic Assets](#) [Digital Twin](#) [Data As An Asset](#)

[Data And Analytics](#) [Artificial Intelligence](#)

生成对抗网络 (GAN)

关键数据流: random noise \rightarrow GAN \rightarrow data

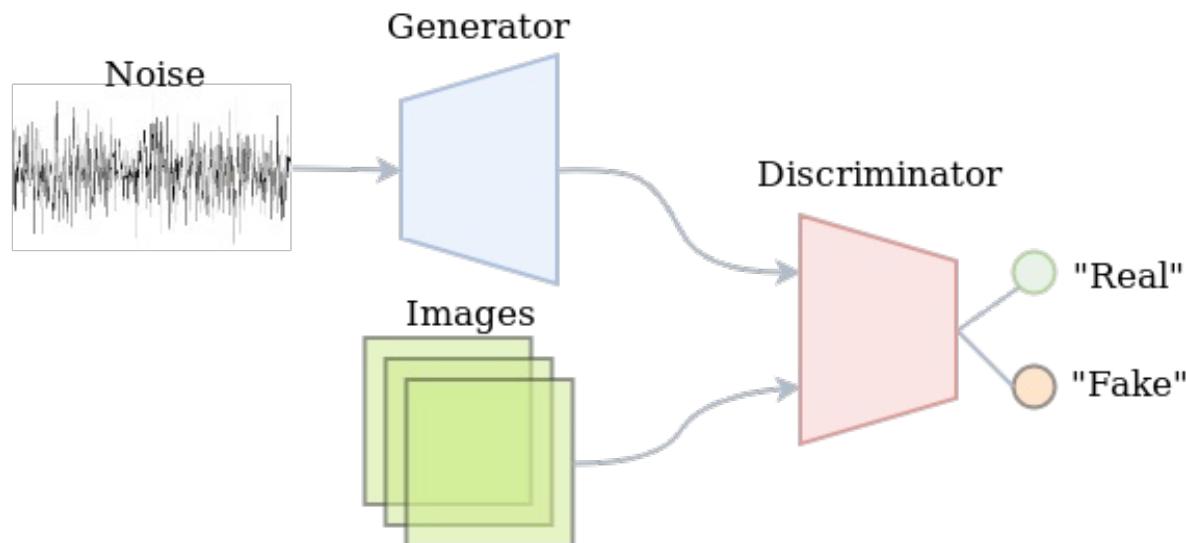


Image source: <https://medium.com/@jos.vandewolfshaar/semi-supervised-learning-with-gans-23255865d0a4>

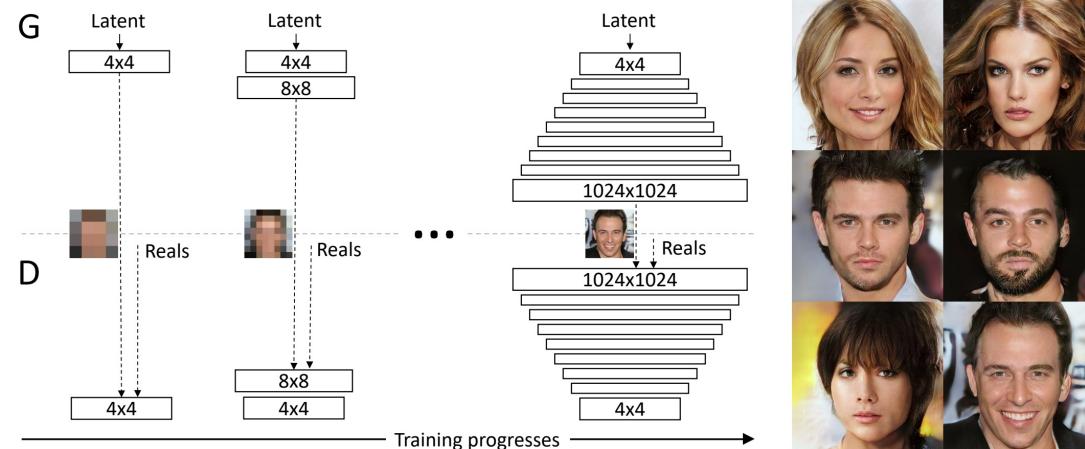
Goodfellow *et al.*, Generative Adversarial Nets, NIPS 2014



Assessment of Deep Generative Models for High-Resolution Synthetic Retinal Image Generation of Age-Related Macular Degeneration

Philippe M. Burlina, PhD; Neil Joshi, BS; Katia D. Pacheco, MD; T. Y. Alvin Liu, MD; Neil M. Bressler, MD

用 PGGAN 合成 normal / AMD 眼底彩照



挑战

随机噪声输入，
合成样本不可控

Karras *et al.*, Progressive growing of GANs for improved quality, stability, and variation, ICLR 2018



多模态眼底图像合成



IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. X, NO. X, XX 2022

1

Learning Two-Stream CNN for Multi-Modal Age-related Macular Degeneration Categorization

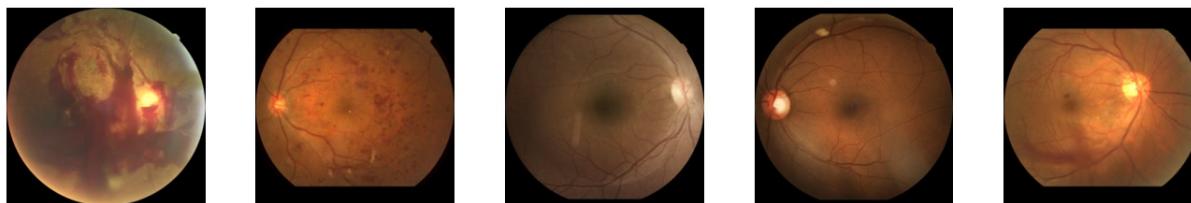
Weisen Wang, Xirong Li, Zhiyan Xu, Weihong Yu, Jianchun Zhao, Dayong Ding, Youxin Chen

研究问题

用什么替代随机噪声作为GAN的输入，使得合成样本更可控？

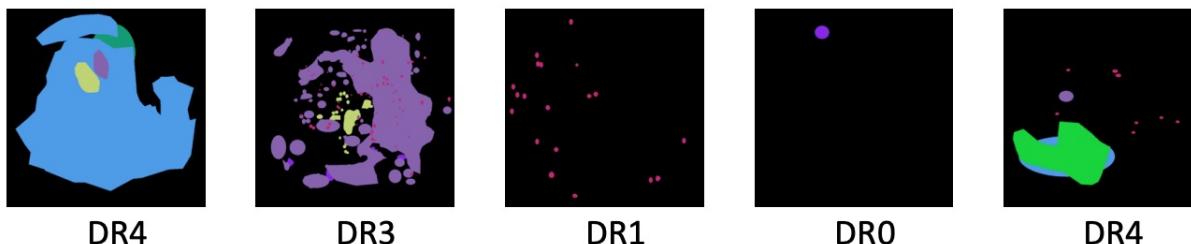
病灶图作为GAN的输入？

病灶分割图指示病变区域与类型



挑战

标注代价过大



Wei et al., Learn to Segment Retinal Lesions and Beyond, ICPR 2020

病灶图的近似方案

类激活图 Class Activation Map (CAM)

- 反映各区域对于预测结果的贡献

基本原理 $(a+b)*c$ 等价于 $(a*c + b*c)$

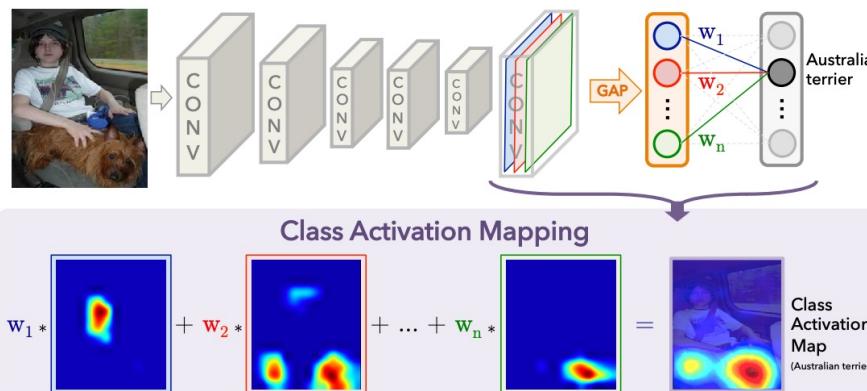
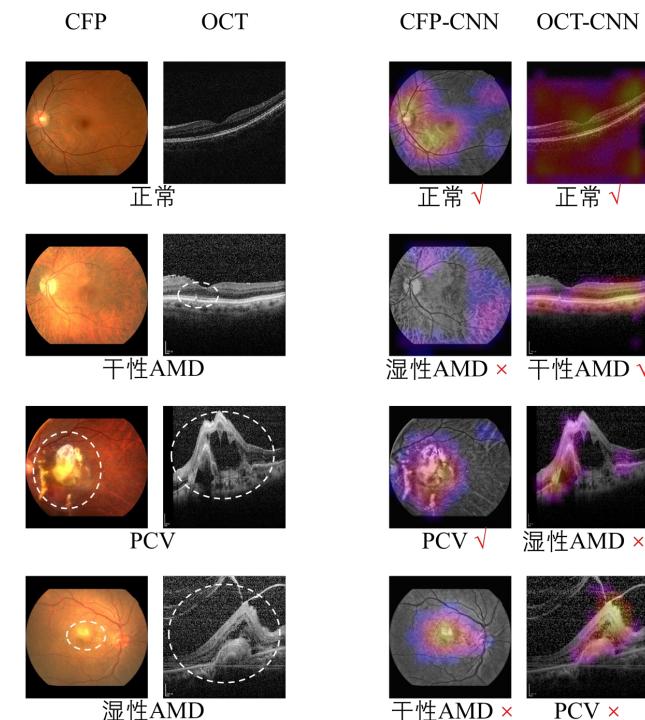


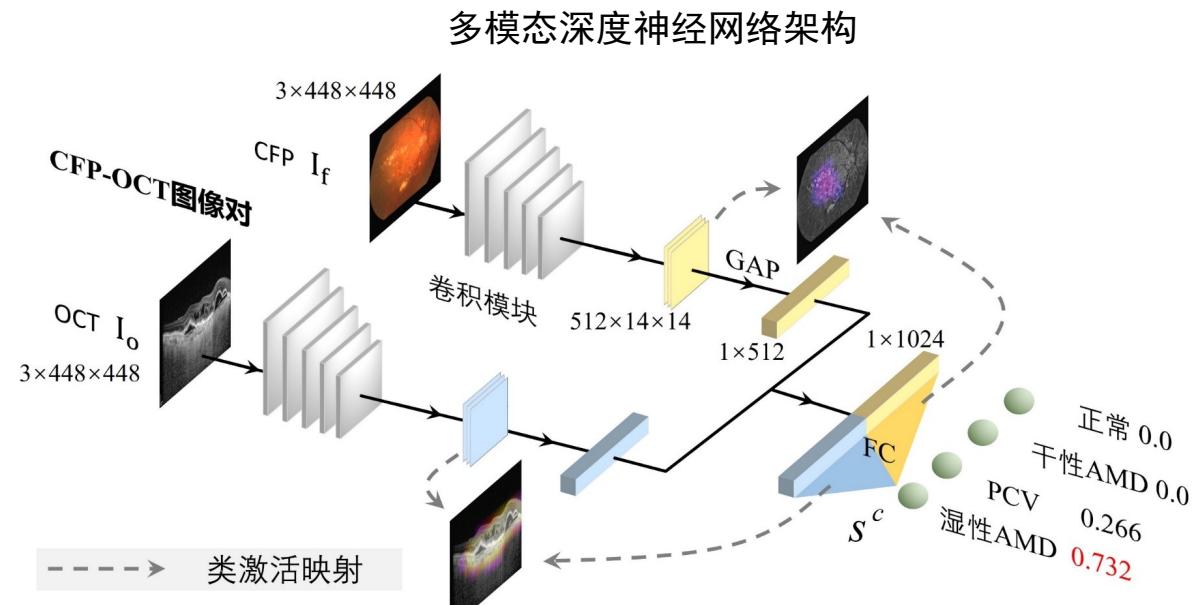
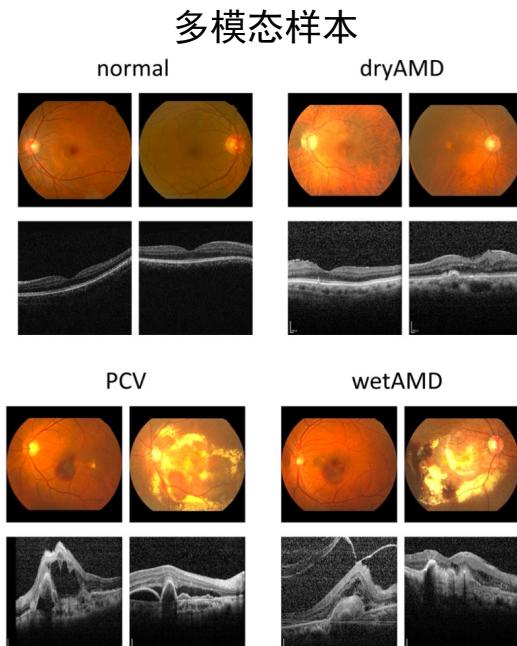
Figure 2. Class Activation Mapping: the predicted class score is mapped back to the previous convolutional layer to generate the class activation maps (CAMs). The CAM highlights the class-specific discriminative regions.



Zhou et al., Learning deep features for discriminative localization, CVPR 2016

我们的研究框架

任务设定：基于多模态眼底影像的AMD识别（四分类任务）





我们的研究框架

数据采集、标注及划分

- 北京协和眼科门诊
- 入组829位患者，1093只眼，1094张眼底彩照 (CFP)，1289张OCT图像
- 4个类别：Normal / dry AMD / wet AMD / PCV
- 参考FA / ICGA作为标注金标准
- 以眼为单位进行数据划分，避免模型过拟合

Class	Training set		Validation set		Test set	
	CFP	OCT	CFP	OCT	CFP	OCT
normal	155 (155)	156 (155)	20 (20)	20 (20)	20 (20)	20 (20)
dryAMD	67 (67)	33 (22)	20 (20)	35 (20)	20 (20)	38 (20)
PCV	259 (259)	289 (156)	20 (20)	44 (20)	20 (20)	47 (20)
wetAMD	453 (452)	531 (325)	20 (20)	38 (20)	20 (20)	38 (20)

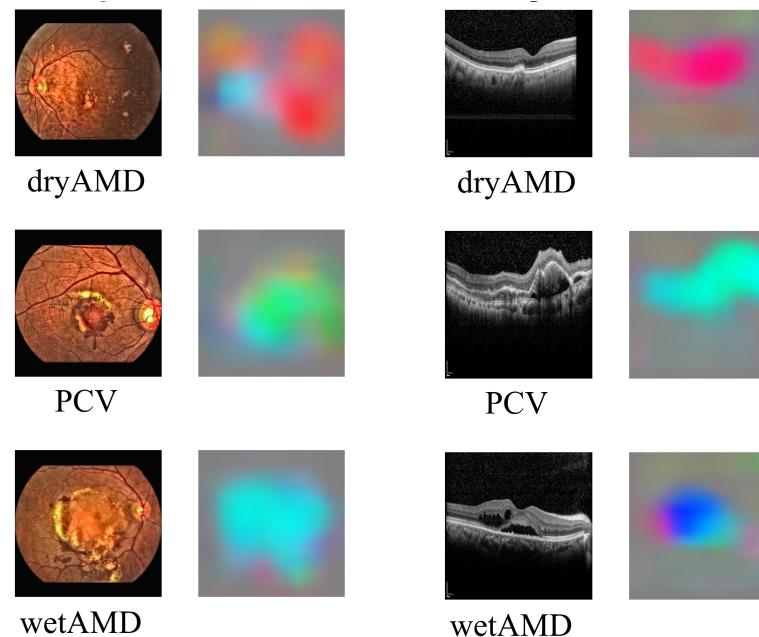
括号内为眼睛数量



单通道CAM → 三通道CAM

CAM^{dry} , CAM^{pcv} , CAM^{wet} 分别对应彩色图的Red, Green, Blue三通道

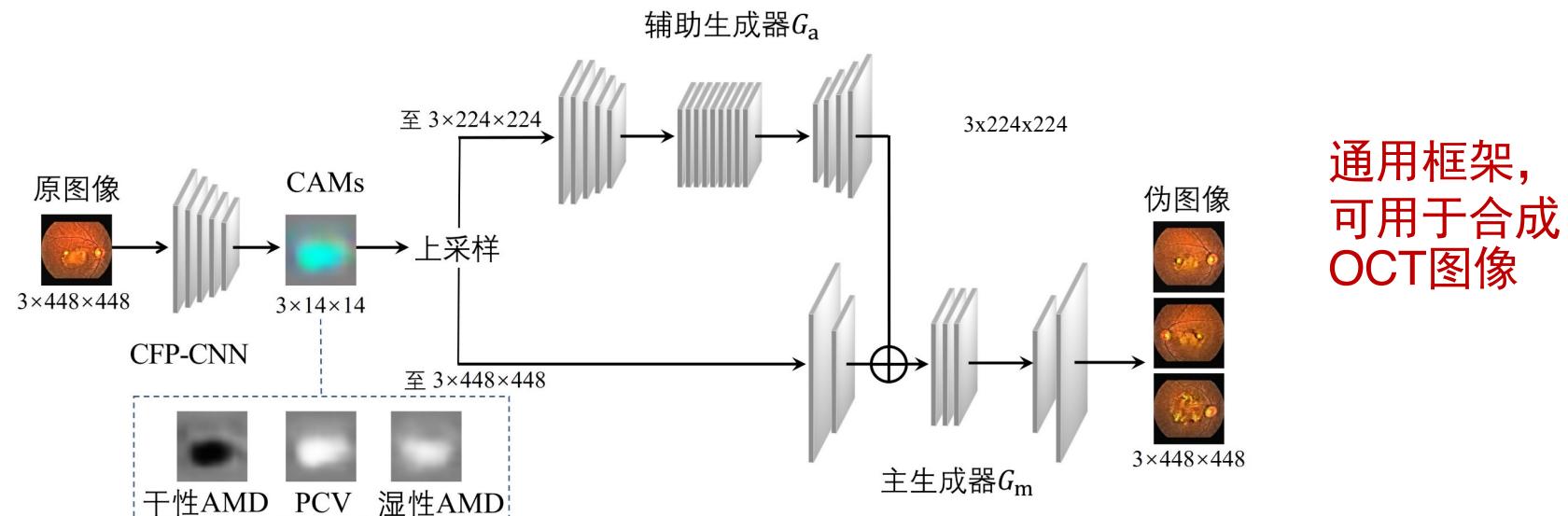
由灰度图变为彩色图，
更方便观察



CAM-conditioned Image Synthesis

三通道 CAM → 合成眼底图

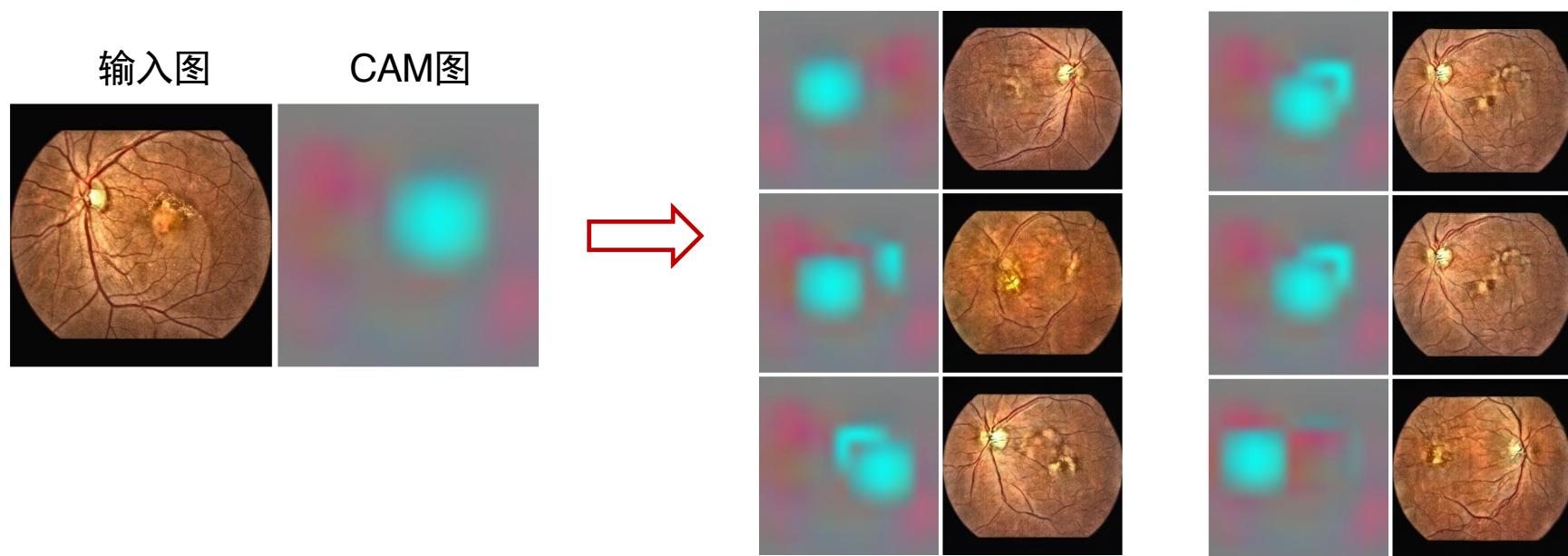
关键数据流：真实图像-> CNN -> CAM -> pix2pixHD -> 合成图



Wang et al., pix2pixHD, CVPR 2018

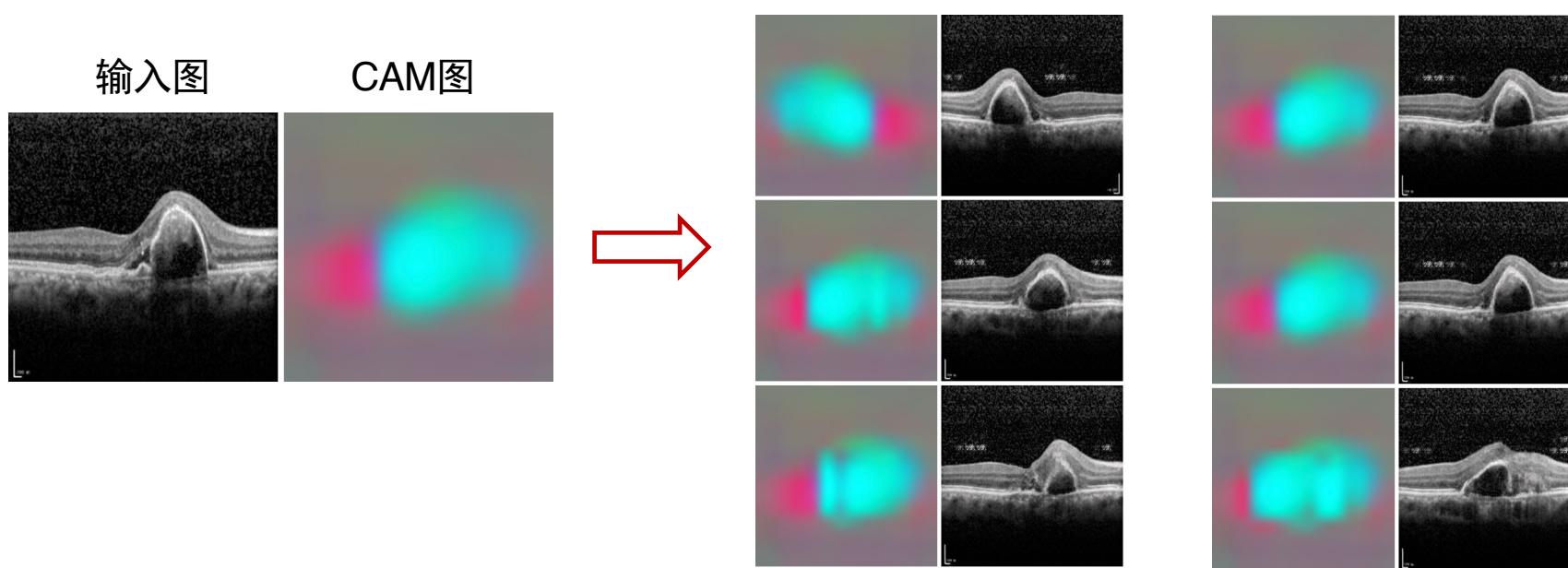
合成效果展示

通过操纵CAM图合成具有视觉多样性的眼底图像样本

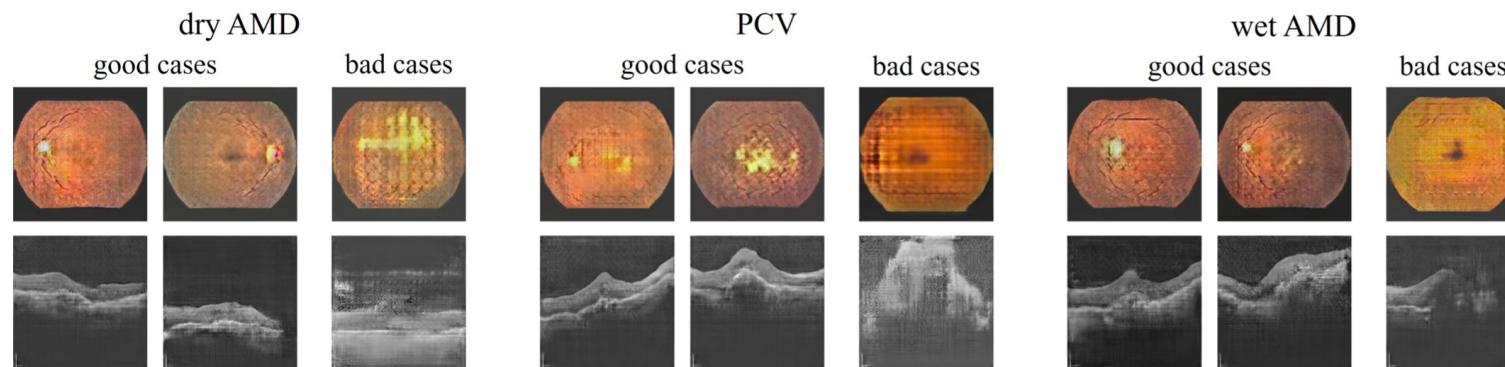


合成效果展示

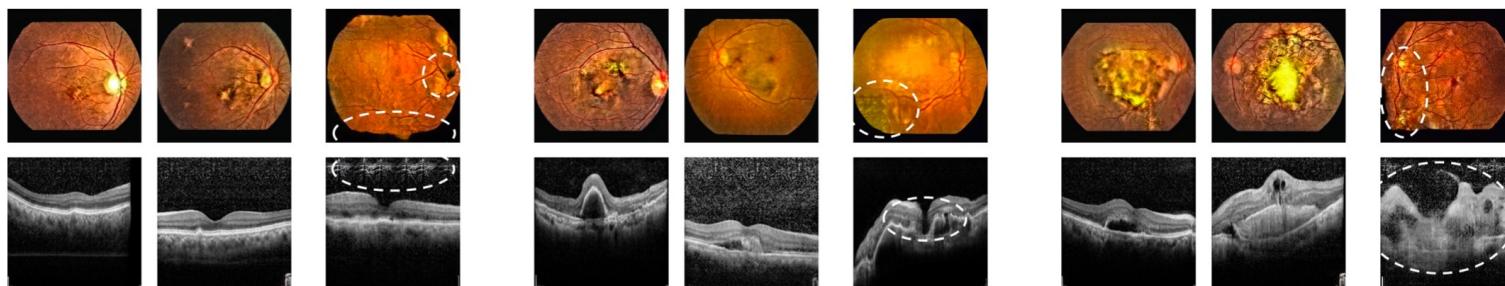
通过操纵CAM图合成具有视觉多样性的眼底图像样本



合成效果对比



(a) Label-conditioned GAN (**JAMA Ophthalmology 2019**)

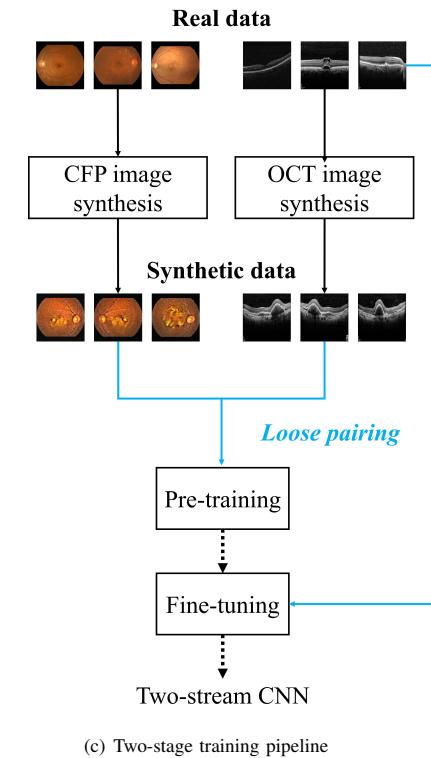
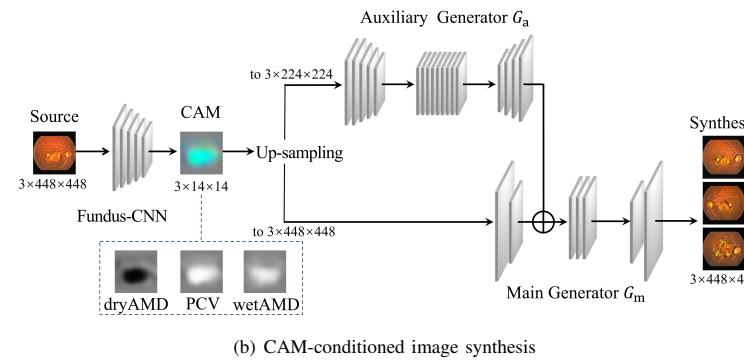
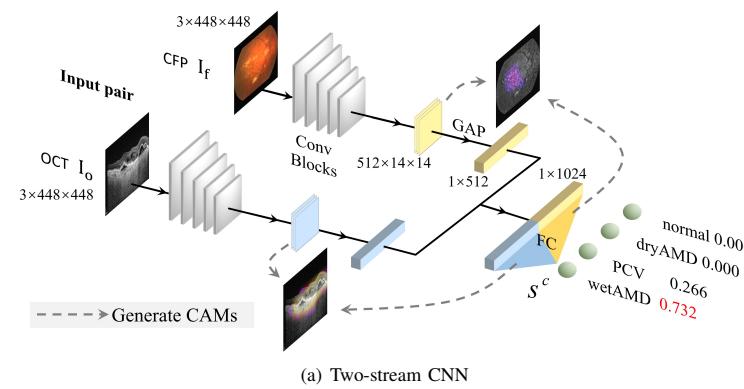


(b) CAM-conditioned GAN (**J-BHI 2022**)

整体技术方案

两阶段训练策略

- 用**合成**数据预训练
- 用**真实**数据微调





实验1：合成数据是否有助于改善模型？

基于多模态影像的AMD四分类性能

训练数据增强方案	F1 (Sen. & Spe.)	Accuracy
标准数据增强	0.872	0.804
Label-conditioned 合成	0.887 (1.7%)	0.825 (2.6%)
CAM-conditioned 合成	0.892 (2.3%)	0.846 (5.2%)
松弛配对 (MICCAI'19)	0.897 (2.9%)	0.837 (4.1%)
CAM-conditioned 合成 松弛配对 (MICCAI'19)	0.914 (4.8%)	0.863 (7.3%)



实验2：合成数据是否能替代真实数据？

基于单模态影像的AMD四分类性能

模态	数据类型	F1	Accuracy
OCT	真实数据	0.886	0.818
	合成数据	0.874	0.802
眼底彩照	真实数据	0.774	0.717
	合成数据	0.414	0.433



初步结论

合成数据可作为真实数据的有效补充，用于更大AI模型的预训练

合成数据不限量、可持续、可分享，促进数据流通共享

多模态数据合成技术有望在眼科AI中发挥关键作用



<https://github.com/li-xirong/mmc-amd>



致谢

合作伙伴

- 北京协和医院眼科 陈有信教授团队
- 北京致远慧图 (Vistel)



研究基金

- 北京市自然科学基金：面向常见眼底病识别的多模态可解释深度学习研究 (No. 4202033)
- 北京市自然科学基金-海淀原始创新联合基金：基于多模态影像的主要眼底黄斑疾病识别与自然语言诊断报告生成 (No. 19L2062)

✉ xirong@ruc.edu.cn