

医学算法与模型前沿简报

今天重点关注：四中心冠脉 CTA 报告生成评价、中国团队参与的临床结构化指标、糖尿病纵向轨迹的可解释概率模型、眼底多视图识别脑卒中，以及跨国胸片队列的联邦 LoRA 适配。

正式发布。论文 1、4 属于最近24-48小时的新稿；论文 2 扩展至近10天，论文 3 为近30天内高度相关的“眼底+脑卒中”背景研究。下载 本期 PDF，或查看 上一期简报。

1. 冠心病 | 四中心 CCTA 报告生成基准与临床结构化指标

A Multicenter Benchmark and Clinically Structured Metric for Coronary CTA Report Generation

arXiv 2609.00909 | 2026-09-01 | 北京协和医院、天津胸科医院、北京世纪坛医院、新疆医科大学第一附属医院等

模型/指标结构

- 输入为患者级 3D CCTA 序列，基准统一评估 7 个开源 3D 医学视觉语言模型生成的报告。
- CSMCCTA 先用临床规则与关键词抽取，把报告转成患者级、血管级及“血管-节段-异常-严重程度”四元组。
- 根据参考报告与预测报告共同支持的最细解剖层级进行比较，避免报告粒度不一致造成不公平惩罚。
- 各临床变量以 F1 计分，再用 70 例专家评分拟合非负权重；另留 30 例作不重叠验证。

数据与优势

四院共 818 个患者-报告对、3,021 个 CTA

序列。指标不只衡量文字相似度，而是显式检查冠脉位置、斑块类型、狭窄程度及报告层级。它与专家评分的相关系数为 0.97，高于 13 个文本或医学报告指标；同义表述替换时分数不下降，逐步删减临床信息时分数随之降低。

结果与临床边界

CCTA 专用 C2RG 在四院均排名最高，CSMCCTA 分别为 0.456、0.520、0.355 和 0.479，但距离理想表现仍远。部分通用模型生成的无关报告比例最高达

98.7%。这是报告生成与评价研究，并非对冠心病未来事件的预测，也未验证自动报告能否改善诊断或治疗结局。

短板

专家权重来自单中心 100 例内部样本，虽然 30 例与拟合样本不重叠，但不是独立外部专家验证。参考报告本身存在完整性和颗粒度差异；四院数据未公开。模型排名只适用于当前 7 个模型和本基准，不能外推为通用优劣结论。

中文摘要

这项中国多中心研究建立了冠脉 CTA 自动报告基准，并提出按患者、血管和冠脉节段分层的临床结构化指标。指标把自由文本转为可核对的临床变量，再在双方共同支持的解剖层级上比较，因而比 BLEU、ROUGE 或单纯语义相似度更能识别位置和严重程度错误。四院数据证明 CCTA 专用模型明显优于通用 3D 模型，但总体性能仍不足以支持临床自动化。研究的突出价值是把“报告写得像不像”推进到“临床事实是否在正确位置且严重程度正确”。

English summary

This multicenter study introduces a benchmark for coronary CTA report generation and a clinically structured metric that converts free text into patient-, vessel-, and segment-level variables. Scoring is performed at the finest anatomical level supported by both reports, with expert-calibrated non-negative weights. The metric correlated strongly with held-out expert ratings and was robust to equivalent wording while responding monotonically to omitted clinical information. A CCTA-specific model ranked best at all four hospitals, whereas several generalist 3D models often produced irrelevant reports. The study evaluates report fidelity, not diagnosis, prognosis, or downstream patient benefit.

资源：论文页 · PDF直链 · 官方代码/指标仓库 · 训练权重：本研究未发布统一权重 · 四中心数据：临床数据，未提供公开下载

2. 慢病 | T2DM 多器官并发症的结构化 FHMM 轨迹模型

A Structural FHMM for Interpretable Disease Trajectories in T2DM

arXiv 2608.24328 | 2026-08-25 | EPFL/ETH Zürich 与药物流行病学团队

模型结构

- 输入为每6个月聚合一次的纵向 EHR：10类慢性并发症、HbA1c/血脂/血压/BMI/GFR 等实验室指标及用药信息。
- Factorial HMM
用多条相互独立、同步演化的离散隐状态链共同解释观察值；状态按健康恶化方向排序，并限制每步最大跳转和末状态吸收。
- 发射参数施加单调性与组 Lasso 约束；初始状态结合既往4年用药、年龄、吸烟和饮酒信息。
- 使用变分 EM 优化 ELBO，将部分推断复杂度降至 $O(K^2MT)$ ；隐状态期望再经 K-means 聚合为典型轨迹。

数据与优势

IQVIA IMRD-UK/THIN 队列纳入 88,201

名首次接受非胰岛素降糖药的成人，索引日为首次处方日期，要求至少1年既往记录；中位随访约4年。患者按 80%/10%/10% 分到训练、验证和测试集，并按并发症分布分层。相比单一 HMM，多链结构能把视网膜病变、高血压、心肾、梗死等不同进展轴拆开解释。

发现与终点性质

模型识别出骨质疏松、心血管、梗死、代谢综合征、高血压、视网膜病变、血管及心肾等轨迹链；例如心肾链共同关联心梗、卒中、心衰、CKD、血脂异常与肥胖。这是对既往和随访 EHR 的无监督/生成式轨迹刻画，并不是预先指定时间窗的个体风险预测器，也没有给出可用于临床阈值决策的 AUC 或校准结果。

短板

数据来自单一英国数据库且需商业许可；因关键检验缺失排除了 100,325 人，可能形成明显选择偏倚。慢病一旦出现即被视为持续存在，用药只主要影响初始状态；隐链独立、单调恶化和吸收状态都是强结构假设。作者明确提醒，药物与状态的关系是相关性，不是因果效应。

中文摘要

研究用结构化因子隐马尔可夫模型分析2型糖尿病患者的长期多器官演变。多条隐状态链同步产生并发症和实验室指标，并通过单调性、有限跳转与吸收状态增强可解释性。模型在8.82万名新启用非胰岛素降糖药的患者中发现了视网膜、高血压、心血管、心肾等不同轨迹。它适合发现群体层面的疾病演化模式，但不能等同于经过外部校准的未来事件预测模型；单数据库、严重缺失筛选和较强结构假设限制了推广性。

English summary

The study models longitudinal type 2 diabetes records with several parallel hidden Markov chains whose joint state explains chronic comorbidities and laboratory measurements. Ordered states, restricted transitions, monotone emissions, and variational EM make the representation more interpretable and computationally tractable. In 88,201 newly treated adults, the model recovered distinct cardiovascular, renal, retinal, hypertensive, and metabolic trajectories. These trajectories are descriptive latent patterns rather than calibrated forecasts with a fixed prediction horizon. Generalizability is limited by a single licensed UK database, extensive exclusion for missing laboratory values, and strong assumptions about irreversible disease progression.

资源：论文页 · PDF直链 · 官方代码：未发现公开仓库 · 权重：未公开 · IQVIA IMRD 官方申请页（商业/申请访问，非开放下载）

3. 眼底 + 脑卒中 | 双眼四视图 Braided Vision Transformer

Braided Vision Transformer for Stroke Detection in Multi-view Retinal Fundus Imaging

arXiv 2608.14722 | 2026-08-12 | VTT Technical Research Centre of Finland

模型结构

- 每位受试者输入双眼各两张 224×224 眼底图：黄斑中心视图与视神经乳头中心视图，共四路。
- 四个图像嵌入通过两层“编织式”交叉注意力连接，使双眼与不同视野间的信息逐级交换。
- 四路输出展平拼接，经 ReLU 全连接层后用 Softmax 输出二分类（卒中/TIA 对健康）或三分类（卒中、TIA、健康）。
- 约 7.97M 参数；小幅旋转和平移增强。论文报告总体以分类损失训练，但未给出可复现代码。

数据与优势

Stroke-Data 来自芬兰 Oulu、Kuopio 大学医院，含 220 人：73 例卒中、26 例 TIA、121 名健康对照，共 802 张眼底图。采用患者级 4:1

训练/测试分离并做分层5折交叉验证。多视图设计比单图更接近临床同时观察双眼和不同眼底区域的习惯，参数量也低于 DenseNet-121 版多视图基线。

结果与终点性质

二分类 AUC 0.752、敏感度 0.647、特异度 0.727、准确率 0.691，优于普通 ViT 的 AUC 0.729；单视图基线仅约 0.49，多视图约 0.71。三分类中 TIA 的 AUC 仅 0.618、敏感度 0.154。研究识别的是已发生卒中/TIA 与健康对照的横断面差异，不是未来卒中风险预测。

短板

样本量小、单一地区、没有独立外部验证或校准；12K 图像是增强后的训练量，不能当作独立样本。缺失视图有时以对侧眼水平翻转替代，可能引入不真实的双眼一致性。若病例采集时间接近急性事件，模型可能学习事件后改变或采集偏差，尚不能用于人群筛查。

中文摘要

BViT 将双眼的黄斑和视盘中心眼底图作为四路输入，通过编织式交叉注意力建模眼间与视野间关系。220人的内部患者级交叉验证显示，它在卒中/TIA 与健康对照的二分类中优于普通 ViT 和既有多视图

CNN，证明多视图融合可能提取更多脑血管相关视网膜特征。不过 TIA 敏感度很低，样本与地区单一，也没有未来事件时间窗，因此它目前属于“眼底识别既有脑血管事件”的概念验证，而非卒中风险预测工具。

English summary

BViT processes four 224×224 fundus views-macula- and optic-disc-centered images from both eyes-and links them through braided cross-attention before classification. Patient-level splitting and stratified five-fold evaluation on 220 participants showed better binary stroke/TIA discrimination than vanilla ViT and prior multi-view CNN baselines. The improvement supports cross-eye, cross-field fusion, but the evidence remains preliminary: the cohort is small and local, augmented images are not independent samples, missing views may be synthesized from the contralateral eye, and TIA sensitivity is poor. This is cross-sectional event detection, not prospective stroke prediction.

资源：论文页 · PDF直链 · 官方代码：未发现公开仓库 · 权重：未公开 · Stroke-Data：作者采集，未提供公开下载或申请页

4. 其他代表性模型 | 跨国胸片队列的联邦 LoRA + BiomedCLIP

Federated LoRA Adaptation of BiomedCLIP Across Four International Chest X-Ray Cohorts

arXiv 2609.02101 | 2026-09-02 | 美国、越南、西班牙四个公开队列

模型结构

- 主干为 BiomedCLIP：冻结 PubMedBERT 文本编码器，图像编码器为用 1,500 万图文对预训练的 ViT-B/16。
- 在每个 Transformer 块融合后的 QKV 投影加入 LoRA， $r=8$ 、 $\alpha=16$ 、 $\text{dropout}=0.1$ ，仅约 0.25% 参数可训练。
- 本地用图像-文本对称对比损失训练；提示文本最多列出3个阳性征象。
- 客户端只交换 LoRA 更新。服务器先还原 BA 乘积、等权平均，再用截断 SVD 投回秩8，避免直接平均 A、B 因子造成失真。

数据、成本与优势

使用 NIH ChestX-ray14、CheXpert、VinDr-CXR、PadChest，训练/验证/测试图像总量分别约 308K/37K/54K。图像统一为 224×224。每客户端每轮上传约 1.13 MB，比交换完整模型小99%以上；5轮联邦训练，每轮本地1个 epoch，Adam 学习率 2×10^{-4} 、batch 32。

结果与验证边界

五个共有征象的平均测试 AUC 从冻结主干的 0.687 升至 FedAvg 的 0.802，接近集中训练参考 0.812。相较各队列单独训练，CheXpert 与 VinDr 分别提高 0.038、0.016，NIH 持平，PadChest 下降 0.028。NIH、CheXpert、PadChest 为患者级划分；VinDr 因无患者ID只能图像级划分。

短板

只运行一个随机种子，置信区间反映测试抽样而非训练不确定性；VinDr 可能存在同患者泄漏。评价只用 AUC，没有校准或临床阈值；联邦学习减少原始数据汇集，但本身不等于隐私保护。论文代码仓库当前仅有说明文件并标注“后续更新”，适配后的 LoRA 权重尚未发布。

中文摘要

研究把 BiomedCLIP 的图像编码器用 LoRA 分别适配到四个国家队列，再通过乘积空间平均与 SVD 压缩聚合客户端更新。训练仅交换约1.13 MB 的适配器，显著降低通信成本；四队列共有五病种的平均 AUC 达0.802，接近集中式训练。其关键经验是不能直接平均 LoRA 的两个低秩因子，而应先重建完整更新后再投影。结果具备工程参考价值，但单种子、缺少校准、VinDr 图像级划分和未完成的代码仓库意味着复现证据仍不完整。

English summary

This work adapts BiomedCLIP across four international chest X-ray cohorts using rank-8 LoRA modules on fused QKV projections. Clients exchange only compact adapters, and the server averages reconstructed update products before using truncated SVD to recover low-rank factors. The approach increased mean shared-label AUC from 0.687 to 0.802 while using roughly 1.13 MB per client per round. It approached a centralized reference without pooling raw images. However, results come from one seed, calibration was not assessed, one cohort used image-level splitting, and federated training alone does not provide formal privacy guarantees. The official repository currently contains a placeholder rather than runnable code.

资源：论文页 · PDF直链 · 官方代码仓库（当前仅README，代码待更新） · BiomedCLIP官方权重 · 本研究LoRA权重：未公开 · NIH数据 · CheXpert · VinDr-CXR · PadChest

今日复现建议

优先级	项目与理由	预计门槛	资源完整度
首选	CSMCCTA：中国多中心、临床层级定义清楚，可先用少量匿名化报告测试结构化指标，不必训练大模型。	中等；主要是中文/英文临床实体抽取、冠脉术语映射与专家复核。	论文与官方指标仓库可用；原始四院影像不公开。
结构最值得借鉴	FedLoRA-BiomedCLIP：LoRA乘积空间聚合和通信成本控制很实用。	中高；需要4个公开数据集、多客户端模拟和至少一张较大显存GPU。	基础权重与数据齐；论文代码仓库尚未更新，适配权重缺失。
眼底方向	BViT：四视图编织式融合直接契合“眼底+脑血管事件”。	中高；网络不大，但同类患者级多视图卒中数据难获得。	论文齐；代码、权重和数据均未公开。
解释性研究	Structural FHMM：适合做糖尿病-心肾-眼底多病共演化分析。	高；需长周期EHR清洗、缺失处理和概率模型实现。	论文齐；数据需申请/许可，代码和权重未公开。

证据说明：本简报优先检索最近24小时，数量或病种覆盖不足时扩展至7天；高度契合“眼底+全身疾病”的背景研究扩展至30天并明确标注。所有摘要均依据论文自行改写，资源状态只按作者、论文页、官方仓库、官方模型页及数据平台核验。内容仅用于科研情报与复现评估，不构成临床诊断或治疗建议。生成时间：2026-09-03 17:00（Asia/Shanghai）。