

医学算法与模型前沿简报

本期重点看“结构是否值得复现”：多模态对比学习如何利用无标签 PET、稀疏门控 MoE 如何压缩成临床指标、视觉场 masked autoencoder 如何学习神经解剖，以及通用开源 VLM 在单层 MRI 上为何可能胜过医学专用模型。

本期说明：近24小时及近7天内没有检索到足够覆盖各病种、且方法和资源信息完整的新论文，因此扩展到近30天（论文发布日期 2026-08-12 至 2026-08-26）。下载本期 PDF · 查看上一期

1. 冠心病 | PET影像 + 临床表格

A Two-Stage Multimodal Contrastive Framework for PET-Based Prediction of Obstructive Coronary Artery Disease

medRxiv · 2026-08-24 · REFINE PET 多中心注册研究

模型结构

- 输入：15通道心肌灌注 PET 极坐标图，加定量 PET/CT 指标及年龄、性别、BMI。
- 第一阶段：图像编码器与表格编码器分别产生患者向量；用同一患者的图像-表格作为正对，以对称 InfoNCE/CLIP 式损失和可学习温度完成无标签预训练。
- 第二阶段：拼接两路向量，经全连接头输出阻塞性 CAD logit；编码器使用较低学习率微调，分类采用 BCE-with-logits。
- 解释：图像侧 integrated gradients，表格侧 kernel SHAP。

优势与适用场景

把大量没有造影标签的临床 PET 先用于跨模态表征学习，再用少量造影标签微调。预训练 12,225 例、微调 968 例，并在六个独立中心 1,865 例外部验证；外部 AUC 0.85，优于从头训练的 0.83。适合“无标签影像多、金标准标签昂贵”的医学影像场景。

终点与验证边界

终点是 PET 后 180 天内侵入性冠脉造影定义的阻塞性 CAD（左主干 $\geq 50\%$ ，其他主要心外膜冠脉 $\geq 70\%$ ），属于诊断性检测，不是未来事件预测。开发与外部队列无患者、站点或时间重叠；但仅纳入资料完整且接受造影的患者，存在选择偏倚。

结构短板

依赖专业 PET MPI、15通道派生图和定量参数，迁移到普通 CT、SPECT 或基层场景并不直接。表格与图像在患者层面拼接，仍可能受中心设备和重建流程影响；参数量、训练代码和权重未公开，复现门槛高。

中文摘要式概述

研究提出两阶段 PET 多模态框架，用成对极坐标图与临床/定量特征做无标签对比预训练，再以造影标签微调阻塞性冠心病分类器。模型在六个外部中心取得 0.85 AUC，并在与临床灌注评分相同特异度时提高敏感度和阴性预测值。优势是利用常规积累却缺少金标准标签的数据；局限是造影选择人群、完整病例分析以及对专用 PET 工作流的依赖。结果支持辅助检测，不证明长期预后改善。

English summary

This study aligns 15-channel PET polar maps with quantitative imaging and basic clinical variables through symmetric contrastive pretraining, then fine-tunes the encoders with an angiography-labeled classification head. The model reached an external AUC of 0.85 across six independent sites and improved sensitivity and negative predictive value at a clinically matched specificity. Its main strength is label efficiency: routinely acquired PET studies without angiography can still shape the representation. Important limits include selection into angiography, complete-case inclusion, dependence on specialized PET processing, and the absence of

released training code or weights. The endpoint is contemporaneous obstructive CAD, not future cardiovascular events.

资源： 论文页 · PDF直链 · 代码：未公开 · 权重：未公开 · 数据：REFINE PET 临床注册数据，未提供公共下载

2. 慢病/心代谢 | 中国队列 + 外部验证

A Deep Learning-Derived Insulin Resistance Index for Cardiovascular Risk Prediction

medRxiv · 2026-08-12 · REACTION (福建) / CHARLS / NHANES

模型结构

- 输入：9项常规变量；学习型稀疏 mask 自动调节特征权重。
- 并行 MLP 专家组成 Mixture-of-Experts，产生14个候选 IR 表征；L1 正则促稀疏，KL 散度用于知识蒸馏。
- 最终选择只依赖 FPG、体重、腰围、高血压史和 HbA1c 的 DNN-IR。
- 下游再用 LR、RF、XGBoost、LightGBM、AdaBoost比较该指标与19个传统 IR 指标。

优势与适用场景

MoE 可拟合不同年龄、性别和体型下的代谢异质性，稀疏 mask

又把最终工具压缩为5项常规变量，临床录入成本低。REACTION 内部验证 AUC 0.84；CHARLS 7年随访中，DNN-IR 每升高1 SD，调整后新发 CVD 风险增加23%。

数据与时间窗

REACTION 6,298人用于横断面动脉僵硬开发/内部验证（训练5,000、测试1,298）；CHARLS 7,047人用于7年新发 CVD 外部验证；NHANES

13,889人连接死亡登记，验证心血管及全因死亡。三个数据源承担的任务不同，不能把横断面 baPWV 训练直接解释成事件预测训练。

结构短板

候选指数先围绕 $\text{baPWV} \geq 1800 \text{ cm/s}$ 学习，baPWV 是动脉僵硬替代终点而非冠脉造影。最终 DNN-IR 的外部分析主要依赖回归关联，仍可能有残余混杂；代码、权重和完整公式实现未发布，降低独立复现性。

中文摘要式概述

中国团队以福建 REACTION 队列开发稀疏门控

MoE，把九项候选临床特征压缩成由空腹血糖、体重、腰围、高血压史和 HbA1c 构成的

DNN-IR。该指标在内部数据上优于传统胰岛素抵抗指标，并在 CHARLS 与美国 NHANES 中关联新发心血管病和死亡。亮点是特征少、跨国验证且关注心代谢异质性；但训练标签是横断面动脉僵硬，外部结局分析不等同于端到端生存模型，也不能据此推断干预 DNN-IR 会改变结局。

English summary

The investigators used a sparse masking layer and parallel MLP experts to derive candidate insulin-resistance representations from routine variables. They selected a five-variable index using fasting glucose, weight, waist circumference, hypertension history, and HbA1c. Development relied on arterial stiffness in the Fujian REACTION cohort, while CHARLS and NHANES supplied independent longitudinal checks for incident cardiovascular disease and mortality. The compact input set and cross-country evaluation are practical strengths. However, the learning target was a cross-sectional surrogate, downstream outcome analyses were association models rather than a unified survival network, and residual confounding remains possible. No reproducible implementation or trained weights were identified.

资源： 论文页 · PDF直链 · 代码：未公开 · 权重：未公开 · CHARLS申请页 · NHANES官方页 · REACTION福建数据：联系作者申请

3. 眼科 + 神经系统 | 视野/OCT关联

CHIASM: A Self-Supervised Visual Field Encoder for Neuro-Ophthalmology

medRxiv · 2026-08-25 · 视野功能 + OCT结构交叉验证

模型结构

- 每只眼52个 Humphrey 视野检测位置作为 token; 4层 Transformer、维度128、4头注意力、pre-norm、dropout 0.1。
- 随机遮挡30%-50%位置，以2层解码器和 Huber loss 重建；编码器约537,088参数，最终 mean pooling 得到128维向量。
- 用视网膜坐标初始化位置编码；从33个“垂直中线”潜变量训练逻辑回归，区分神经性与非神经性视野缺损。

优势与适用场景

模型小、输入结构清楚，并把自监督潜变量与视觉通路解剖对应。它不是简单做青光眼诊断，而是审计青光眼训练集里可能混入的脑卒中、出血或肿瘤型视野损害，属于“眼科 + 神经系统”很实用的关联任务。

数据与验证

UWHVF

共28,943个视野、3,871名患者，23,223个无标签视野用于预训练，患者分组交叉验证避免同人泄漏；外部应用于 Harvard-GF 3,300名患者的视野+OCT。分类器平衡准确率0.78、AUC 0.85；在1,748名青光眼标签患者中标出20个疑似神经型病例，并由专家、规则及RNFL结构证据交叉支持。

结构短板

监督探针的专家标签规模有限，只有单随机种子且作者声明检查点并非位级可复现。外部队列用于无微调筛查，但“前20名”阈值是保守展示，不估计真实污染率；尚无前瞻性临床工作流程验证。

中文摘要式概述

CHIASM 用 masked autoencoder 从无标签 Humphrey 视野中学习128维表示，再以具有垂直中线解剖意义的潜变量训练简单分类器，识别更符合神经系统损害而非青光眼的视野模式。其训练、监督探针和外部应用数据相互独立，并在 Harvard-GF 中用 OCT 的 RNFL 保留情况验证结构-功能不一致。模型参数少、解释路径清楚，适合数据集质控；但目前只证明可筛出可疑模式，不能替代神经影像或临床诊断。

English summary

CHIASM represents each monocular Humphrey field as 52 spatial tokens and learns a 128-dimensional embedding by reconstructing randomly masked locations. A linear probe restricted to anatomically interpretable vertical-meridian dimensions distinguishes neurological from non-neurological loss. Patient-grouped development and zero-shot application to Harvard-GF reduce leakage, while OCT retinal nerve fiber layer measurements provide an independent structural check. The compact transformer and public source datasets make the concept attractive for auditing glaucoma corpora. Still, the supervised neurological labels are limited, the operating threshold was chosen to review only the highest-ranked cases, and the work does not establish prevalence or clinical diagnostic performance.

资源：论文页 · PDF直链 · 代码：论文补充材料提供方法细节，未确认独立模型仓库 · 权重：未确认公开 · UWHVF数据 · Harvard-GF数据/基线代码

4. 其他代表性模型 | 脑转移瘤 MRI + 开源VLM

Benchmarking Open-Source Vision-Language Models for Brain Metastasis Assessment on Single-Slice CE-MRI

medRxiv · 2026-08-26 · 韩国 AI Hub 安全区离线评测

模型结构与输入

研究不是训练新模型，而是统一评测6个4B-8B开源 VLM。输入为单张轴位增强 T1 MRI PNG，加固定结构化提示，输出病灶有无、数量、侧别、部位、强化、坏死、水肿和占位效应。表现最佳的 MiniCPM-V 4.5 以 Qwen3-8B 为语言骨干、SigLIP2-400M 为视觉编码器，总规模约8B，并用视觉 resampler 压缩图像 token。

优势与适用场景

所有模型部署在单张 V100 32GB 的离线安全区，证明医院内网可落地。MiniCPM-V 4.5 的敏感度78.3%、特异度85.0%，平衡性最好；通用模型未必逊于医学专用模型。适合本地原型与模型筛选，而不是直接作为临床产品。

数据与验证

从 AI Hub 1,000例脑转移瘤数据中连续选择60名患者；每人一张病灶阳性切片，再匹配同患者不同解剖层面的阴性切片，共120张。两名神经放射科医师建立参考标准。相同患者同时贡献正负图，适合配对比较，但样本小且只测单层图像。

结构短板

单切片 PNG 丢失3D上下文、序列信息和DICOM强度语义；没有训练/微调，也没有外部中心验证。医学专用模型未胜出并不代表通用 VLM 已具临床可靠性，病灶计数和细粒度征象表现仍有限。

中文摘要式概述

该研究在断网医院安全区，用统一提示比较三种通用和三种医学专用开源视觉语言模型对脑转移瘤增强 MRI 单切片的理解能力。MiniCPM-V 4.5 获得最佳平衡准确性，说明小型通用 VLM 的视觉编码与多模态训练可能比“医学专用”标签更关键。模型及权重可下载，复现推理相对容易；然而数据只有60名患者、120张配对切片，缺少三维与多序列上下文，也没有独立外部验证，结果应视为离线基准而非诊断证据。

English summary

Six open-source vision-language models were tested offline on paired positive and negative single-slice contrast-enhanced T1 MRI images. MiniCPM-V 4.5 achieved the most balanced detection profile, whereas medical-purpose models did not consistently outperform general-purpose alternatives. The study is valuable because it uses deterministic decoding, a shared prompt, an isolated hospital-style environment, and openly downloadable model weights. Its evidence remains narrow: only 60 patients were included, images were converted from DICOM to PNG, volumetric and multi-sequence context was removed, and no external site was evaluated. The findings support model selection experiments, not autonomous brain-metastasis diagnosis.

资源： 论文页 · PDF直链 · MiniCPM官方代码 · MiniCPM-V 4.5权重 · AI Hub数据申请入口 ·
研究脚本： 论文补充材料，未见独立仓库

今日复现建议

项目	建议	门槛	资源完整度
最值得复现	CHIASM：结构小、患者级划分明确、两个公开数据源，且 “眼科+神经系统” 关联价值高。	中等；需整理视野空间位置和PD/TD规则，单卡即可。	PDF ☐ 数据 ☐ 基线代码/数据仓库 ☐ 专用权重 ☐
最快跑通	MiniCPM-V 4.5 脑MRI评测：可直接下载权重并复用论文提示。	中高；BF16约9B，建议 ≥24-32GB显存或量化。	PDF ☐ 模型代码 ☐ 权重 ☐ 数据需申请 ☐
最有结构创新	PET-CAD 两阶段对比学习：非常适合无标签多、金标准少的场景。	高；专用PET数据与预处理难获得。	PDF ☐ 代码 ☐ 权重 ☐ 数据 ☐
最贴近基层慢病	DNN-IR：只需5项常规变量，但当前实现材料不足。	中等；数据可申请，需自行重建MoE。	PDF ☐ 公共队列 ☐ 代码 ☐ 权重 ☐

证据说明：本简报为研究情报与复现线索，不构成临床诊断或治疗建议。所有摘要均依据论文信息重新撰写；资源链接仅列官方论文、作者/机构仓库、官方模型页和数据平台。生成时间：2026-09-02 17:00 (Asia/Shanghai)。