

• 胸部影像学 •

基于深度学习的胸部 X 线肺结核检测研究及多中心临床验证

安超, 张晨, 郑广平, 曹义, 杨根东, 印宏坤, 顾俊, 邹彤, 吴双, 王立非

【摘要】 目的:构建基于深度学习的胸部 X 线肺结核检测模型并通过多中心研究验证其效能及临床价值。**方法:**回顾性搜集 2600 例来自 3 个中心的胸部 X 线图像并随机分为训练集、验证集和测试集,构建基于 RetinaNet 架构的肺结核深度学习检测模型,并在 ChinaSet 和 MontgomerySet 胸部 X 线公开数据集以及来自深圳三院的外部临床测试集上对深度学习模型的鲁棒性进行外部测试。采用受试者工作特征曲线(receiver operator characteristic curve, ROC)评估模型效能。同时通过临床检测评估深度学习模型的重复性和再现性。**结果:**深度学习模型在内部测试集的 ROC 曲线下面积(AUC)为 0.967,在 ChinaSet、MontgomerySet 和深圳三院外部测试集的 AUC 分别为 0.95、0.93 和 0.976,具有较高的准确性和良好的鲁棒性。临床一致性评估证实了模型的重复性和再现性。**结论:**深度学习模型具备良好的效能,可以作为胸部 X 线影像结核病检测工具用于临床决策支持。

【关键词】 深度学习;放射摄影术,胸部;结核,肺;多中心研究

【中图分类号】 R-056; R816.4; R521; R181 **【文献标识码】** A

【文章编号】 1000-0313(2022)06-0704-06

DOI:10.13609/j.cnki.1000-0313.2022.06.006

开放科学(资源服务)标识码(OSID):



Development of a deep learning-based algorithm for pulmonary tuberculosis detection on chest radiograph: a multi-center study An Chao, ZHANG Chen, ZHENG Guang-pin, et al. Department of Radiology, the Third People's Hospital of Shenzhen, Guangdong 518100, China

【Abstract】 Objective: To develop a deep learning model for detecting tuberculosis on chest radiographs and evaluate its efficacy and clinical value on multi-center datasets. **Methods:** A total of 2600 chest radiographs from three hospitals were collected retrospectively, and a RetinaNet architecture-based convolutional neural network was proposed for the detection of tuberculosis. The chest radiographs were randomly divided into training, validation and internal testing datasets. Two public datasets (ChinaSet and MontgomerySet) and an independent dataset from The Third People's Hospital of Shenzhen were used for external testing. The repeatability and reproducibility of the deep learning model was confirmed through clinical validation. **Results:** The deep learning model had an AUC of 0.967 in the internal testing dataset, and showed good robustness with AUCs achieving 0.95, 0.93 and 0.976 in the ChinaSet, MontgomerySet and the independent testing dataset from The Third People's Hospital of Shenzhen. The clinical study demonstrated excellent repeatability and reproducibility of the deep learning model. **Conclusion:** The deep learning model showed good performance in detecting tuberculosis on chest radiographs, which could be used as a potential tool for assisting clinical decision.

【Key words】 Deep Learning; Radiography, thoracic; Tuberculosis, pulmonary; Multicenter study

我国是全球 30 个结核病流行严重的国家之一,每年新增肺结核病患者约 90 万例^[1],严重危害人民健康

作者单位: 518100 广东,深圳市第三人民医院放射科(安超、张晨、郑广平、曹义、杨根东、王立非); 518100 广东,广东省感染性疾病(结核病)临床医学研究中心(安超); 100025 北京,推想医疗科技股份有限公司(印宏坤、顾俊、邹彤、吴双)

作者简介: 安超(1987—),女,黑龙江人,博士,主治医师,主要从事人工智能辅助诊断。

通讯作者: 王立非, E-mail: wanglf007n@163.com

基金项目: 广东省科技计划项目(2020B1111170014)

和生命安全。肺结核病诊断的“金标准”是细菌学,即痰培养法,虽然灵敏度较高,但此法检验所需时间较长,且容易受其他杂杆菌影响结果^[2]。临床实践约有 2/3 的肺结核患者痰抗酸染色和痰结核分枝杆菌培养为阴性^[3],对于此类患者,传统的痰培养法无法准确筛查肺结核,往往会造成漏诊^[4]。胸部 X 线(chest X-rays, CXR)是发现结核病患者最简便、可推广、适宜基

层的筛查措施。X 线筛查具有较高的敏感度,但是特异度较差,高度依赖于医生的主观判断,具有一定的组内和组间观察者差异^[5],特别是对于基层医院具有一定的挑战性^[6]。近来,随着人工智能(AI)和深度学习技术的迅猛发展,深度学习算法被广泛应用于医学影像辅助检测和诊断。基于卷积神经网络的深度学习算法可以用于颅内血肿分割及出血类型识别^[7],在基于彩色眼底图像的糖尿病视网膜病变分级、皮肤癌与良性病变等疾病的诊断上,AI 已接近或达到临床专家水平^[8,9];通过大样本训练,AI 也被应用于眼底病变和小儿肺炎的诊断^[10];此外,深度学习技术还可用于心电图中心脏收缩功能障碍和心律失常的检测^[11]。目前也有一些利用深度学习检测胸部 X 线肺结核的报道^[12,13],但是这些研究都是基于公开数据集,缺乏外部独立测试集和临床数据集的验证,因此无法判断其鲁棒性和泛化性。另外,基于卷积神经网络的深度学习模型由于参数量巨大,之前报道的研究中使用的样本数量通常较小(约 1000 例左右或更少),容易出现过拟合的情况。目前尚未有研究报道胸部 X 线肺结核检测深度学习模型在多中心临床数据上的表现。

本研究旨在通过搜集多中心、大样本数据建立基于胸部 X 线的肺结核检出深度学习模型,并通过多中心的外部独立数据集来评估深度学习模型的效能、鲁棒性及临床应用价值。

材料与方法

1. 患者数据集资料

开发数据搜集:本研究回顾性搜集 2600 例来自国内 3 家医疗机构的胸部 X 线扫描影像作为开发数据集,所有数据均直接从 PACS 系统导出并经过脱敏处理。结核患者入组标准为:患者年龄>16 岁,且根据中华人民共和国卫生行业标准中肺结核诊断的规定(WS 288—2017)确诊为肺结核,具体依据:①DNA/RNA 或 GeneXpert 试剂盒检测证实肺结核阳性;②痰培养及抗酸杆菌阳性和结核分枝杆菌阳性;③组织活检提示结核病。排除标准:患者患有肿瘤性疾病,或由于支气管炎、肺炎及手术产生双肺上叶支气管增粗、肺纹理增浓导致难以与肺结核区别的。此外,本研究还搜集 407 例未患有肺结核的胸部 X 线扫描影像作为阴性对照。最终搜集数据中包含了 2193 例确诊肺结核患者数据,其中 1580 例为活动性病灶,613 例为非活动性病灶。

外部测试集数据搜集:本研究共搜集 2 组公开数据集^[14]和 1 组临床数据集作为外部测试集,用于深度学习模型性能的检测以及临床应用的评估,具体包括:①ChinaSet 公开数据集,大部分来自于深圳三院在

2012 年 9 月份搜集的病例,共包含 336 例结核病与 326 例非结核病患者数据,所有的胸部 X 线影像均采用飞利浦数字化 X 射线摄影系统(Philips DR digital diagnost system)采集,并存储为 3000×3000 分辨率的 PNG 图片格式;②MontgomerySet 公开数据集,共包含 58 例结核病和 80 例非结核患者的胸部 X 线影像,均来自于美国马里兰州蒙哥马利县,所有数据以 PNG 图片格式保存,分辨率为 4892×4020;ChinaSet 公开数据集和 MontgomerySet 公开数据集的下载地址为:<http://archive.nlm.nih.gov/repos/chestImages.php>;③深圳三院外部临床测试数据集,包含从 2012 年 1 月—2019 年 9 月之间搜集的 200 例结核病与 694 例非结核患者的 DICOM 格式胸部 X 线影像,均使用飞利浦数字 DR 系统进行捕获,分辨率约为 3000×3000。深圳三院外部临床测试数据集为独立搜集,与开发数据集来自不同医院,且与 ChinaSet 公开数据集来自于不同的患者。

数据标记:本项目目的是在胸部 X 线影像中实现肺结核病灶的检测,因此所有数据中的病灶均由放射科医生用矩形框进行标记。为了确保标记结果的准确性与一致性,在原有标记结果的基础上所有数据均由 1 名来自三甲医院的主治医生对照影像报告进行二次确认,对于有疑问的病例由另一名高年资主治医师进行最终审核确认,作为结核病灶位置判断的最终结果。

2. 深度学习模型的构建和验证

数据预处理:将开发数据集中的 2600 例数据随机划分为训练集(2122 例)、验证集(300 例)和内部测试集(178 例),分别进行深度学习模型的训练、参数调优和结果测试。深度学习一般要求样本数量充足,样本量越大训练出来的模型效果越好,泛化能力也越强。因此,我们采用了包含图像变化、旋转等数据增强技术对训练集中的图像进行了预处理,以提升模型的效能和鲁棒性。

模型构建:本研究通过构建基于类似于 RetinaNet 网络的深度学习模型实现肺结核的检测。使用 ResNeXt-50 作为骨架网络,相比于经典的 ResNet 残差网络,ResNeXt 在其基础上进一步提出了一种拓扑结构的转换模块,使得网络结构在宽度上变得更宽,拥有更强的特征提取能力。ResNeXt-50 的每一个模块中我们都采用了 Conv+BN+ReLU 的结构,其中 Conv 表示卷积层,用于提取图像特征;BN 表示批归一化(batch normalization),有助于网络的收敛和防止过拟合;ReLU 是一种激活函数,用于提升网络的非线性拟合能力。此外,为了保证不同尺寸肺结核征象的检出率,本研究在 ResNeXt 网络结构的基础上添加特征金字塔结构(feature pyramid networks,FPN)来有

效融合不同层次的图像特征;为保证一些样本量比较少的肺结核征象的检出率,采用难样本挖掘技术;为了保证收敛效果,避免过拟合,采用多任务的 loss 函数。本研究采用的深度学习模型是一种以 ResNeXt-FPN 为基础网络的一步(one-stage)检测模型(图 1)。模型的输入为 1 张 DR 胸片图像,不需要其他的临床信息。模型通过影像的特征分析,输出为该影像中包含肺结核征象的概率值。

为了提高模型训练效率,本研究还采用了迁移学习(transfer learning)的方法进行了预训练,预训模型参数的数据来源于 Image Net 的子集,其中训练集 1281167 张,验证集 50000 张,测试集 100000 张,总共 1000 个分类,在测试集上 top5 分类错误率达到 6.6%。

模型效能评估:肺结核诊断采用受试者工作曲线(receiver operating characteristic curve, ROC curve),在内部测试集和 2 个公开数据集上进行肺结核检测,以每张胸部 X 线片上是否检测出结核病灶作为 AI 诊断肺结核的结果,参比金标准进行 AI 诊断性能的分析。通过计算曲线下面积(area under curve, AUC)来评估模型的诊断效能,并基于约登指数所对应的阈值来确定模型的敏感性(sensitivity)、特异性(specificity)和准确性(accuracy)。

3. 深圳三院临床验证

样本量计算:据 WHO 2014 年发布的有关肺结核分诊(triage)筛查产品性能指标的最低要求,灵敏度需要达到 90%,特异度需要达到 70%。本研究引用 WHO 标准,采用目标值法对 AI 模型的有效性进行验证。主要评价指标:AI 模型在阳性肺结核影像病例上的灵敏度高于目标值 90%,优效界值为 0,实现优效;AI 模型在阴性肺结核影像病例上的特异度高于目标值 70%,优效界值为 0,实现优效。

对 AI 的有效性验证进行测试样本量估计,公式如下:

$$n = \frac{[Z_{1-\alpha/2} \sqrt{P_0(1-P_0)} + Z_{1-\beta} \sqrt{P_r(1-P_r)}]^2}{(P_r - P_0)^2}$$

其中 n 为最少样本量, $Z_{1-\alpha/2}$ 、 $Z_{1-\beta}$ 为标准正态分布的分位数,当双侧显著性 $\alpha=0.05$ 时, $Z_{1-\alpha/2}=1.96$, 当 $\beta=0.2$ 时, $Z_{1-\beta}=0.842$ 。

预期的灵敏度=95.5%,目标值=90%,计算得到灵敏度优效所需的最低样本量为 192.23。预期的特异度=75%,目标值=70%,计算得到特异度优效所需的最低样本量为 637.85。

临床一致性评估:为了检测 AI 模型在临床实际使用的情况,本研究对于临床一致性也进行了评估。首先对不同人员之间的结果一致性进行评估,分别由操作员 A、B 和 C 使用 AI 系统进行深圳三院外部临床测试数据集的肺结核检测,每次检测前均将正负样本随机打散,由观察人员记录 AI 诊断结果,整个过程中操作员 A、B、C 互不干扰,互不知情。将所有操作员的检测结果汇总并与已知金标准结果(临床确诊的结果)进行比对,计算各自的敏感性和特异性并评估 AI 模型的结果再现性。其次本研究还对结果的再现性进行了评估,由操作员 D 使用 AI 模型重复在深圳三院外部临床测试数据集上检测肺结核 3 次,每次检测前均将正负样本随机打散,每次检测间隔时间为 1 周。将操作员 D 3 次检测结果汇总并与已知金标准结果(临床确诊的结果)进行比对,评估 AI 模型的可重复性。

结 果

1. 内部验证结果

深度学习模型在内部测试集的 ROC 下曲线面积(AUC)为 0.967,对应最佳阈值下的敏感度为 95.32%,

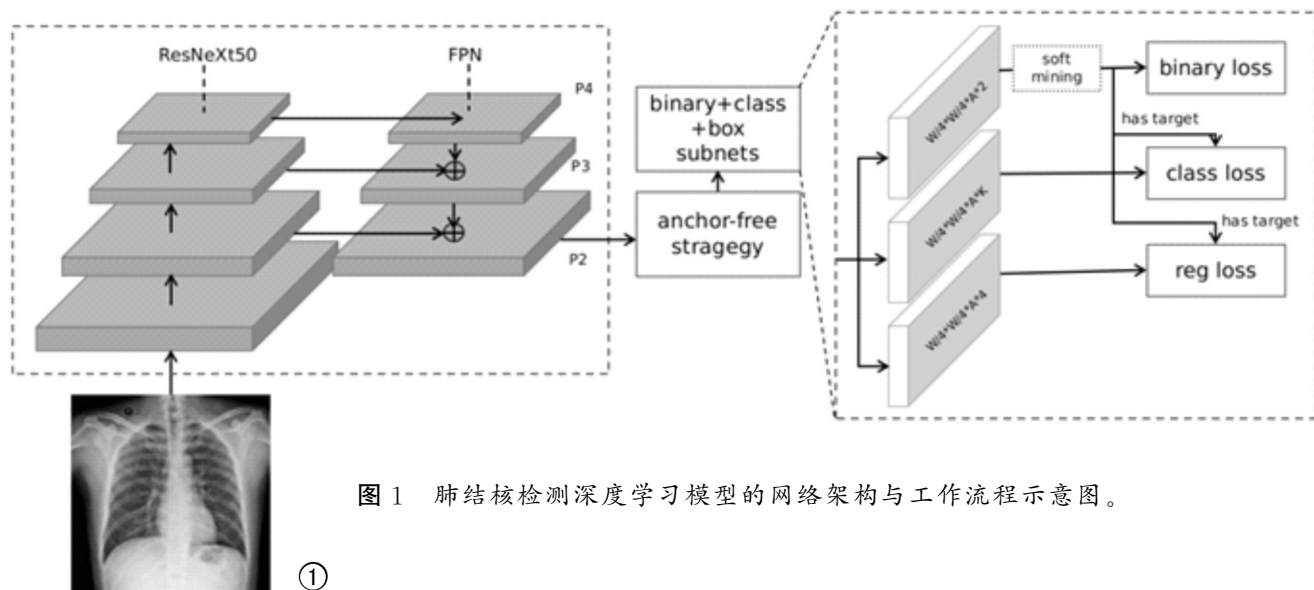


图 1 肺结核检测深度学习模型的网络架构与工作流程示意图。

特异度为 83.33%，AI 模型诊断肺结核的准确率为 94.94%。

2. 公开数据集验证结果

AI 模型在 ChinaSet 和 MontgomerySet 公开数据集上的测试结果 AUC 分别为 0.95 和 0.93(图 2)，与内部验证集的 AUC 基本一致，表现良好的鲁棒性。深度学习模型在 ChinaSet 和 MontgomerySet 公开数据集的外部测试详细结果如表 1。

表 1 深度学习模型在公开数据集上的具体验证结果

外部数据集	结核例数	非结核例数	AUC	准确率 %	敏感度 %	特异度 %
ChinaSet	336	326	0.95	91.39	88.39	94.48
MontgomerySet	58	80	0.93	86.23	75.86	93.75

3. 深圳三院临床数据集测试结果

深圳三院临床数据集包含 200 例肺结核样本和 694 例正常对照样本，满足有效性评价的测试样本量需求。数据集包含了来自 523 名男性和 371 名女性的胸部 X 线影像数据，其中有 596 例患者年龄在 16~35 岁，323 例患者年龄在 36~55 岁，65 例患者年龄在 56 岁以上，1 例患者年龄信息缺失。患者的性别和年龄分布符合临床实际情况。

ROC 分析结果表明 AI 模型在深圳三院临床数据集上也体现了优异的诊断效能，AI 模型的 AUC 为 0.976(图 3)，对应阈值为 0.17 时的敏感度为 97.50%，优于 WHO 规定的 90%；特异度为 77.52%，优于 WHO 规定的 70%，AI 模型诊断肺结核的准确率为 81.99%。

4. 临床一致性测试结果

基于设定好的阈值 0.17，AI 模型会将输出概率≤0.17 的图像判定为阴性，>0.17 的图像判定为阳性，则模型预测的各图像阴阳性与金标准一一对比，则可以得到以下指标。①TP：真阳，金标准为阳性且模型预测为阳性的图像数；②TN：真阴，金标准为阴性且模型预测为阴性的图像数；③FP：假阳，金标准为阴性而模型预测为阳性的图像数；④FN：假阴，金标准为阳性而

模型预测为阴性的图像数。

据 AI 预测结果和金标准的对比，操作员 A、B、C 分别的检测结果以及操作员 D3 次检测的结果汇总(表 2)。

表 2 深度学习模型临床一致性评估结果

时间	操作员	TP	TN	FP	FN	敏感度 %	特异度 %
2019.10.14	A	195	538	156	5	97.50	77.52
2019.10.14	B	195	538	156	5	97.50	77.52
2019.10.14	C	195	538	156	5	97.50	77.52
2019.10.15	D	195	538	156	5	97.50	77.52
2019.10.21	D	195	538	156	5	97.50	77.52
2019.10.28	D	195	538	156	5	97.50	77.52

测试结果表明在相同的环境条件下，不同的操作员或者同一操作员不同时间的检测结果均完全一致，说明 AI 模型对于胸部 X 线肺结节检测具有良好的重复性和再现性。

5. 案例分析

一位典型体检患者在本研究构建的人工智能辅助诊断系统下的结核病灶识别检出情况(图 4)，结果表明 AI 可准确的识别结核病灶并对其进行定位，有助于辅助临床医生进行肺结核诊断。

此外，我们也对 AI 模型的漏诊病例进行了分析，患者双肺纹理增强模糊，右肺及左上肺见结片状致密影，边界欠清，心脏大小形态正常范围，双侧膈面光滑，双侧肋膈角锐利(图 5a)，提示右肺及左上肺感染性病变，最终经病原学确诊为肺结核。AI 漏诊可能原因是由于双肺尖部局限性、对称性病变，且病变区与锁骨近端、第一前肋存在重叠，从而造成模型误判。患者右肺上野可见团片模糊影及纤维条索灶，右侧肺门上提，余双肺未见其他异常密度灶，主动脉迂曲增宽，余未见异常(图 5b)，右上肺改变提示陈旧结核可能，最终经过痰液病理检查确诊。AI 误判的可能原因是右肺尖部淡薄稍高密度影，病变范围较小导致结核征象不典型，同时受到肺尖部锁骨及第一前肋的影响造成模型漏诊。

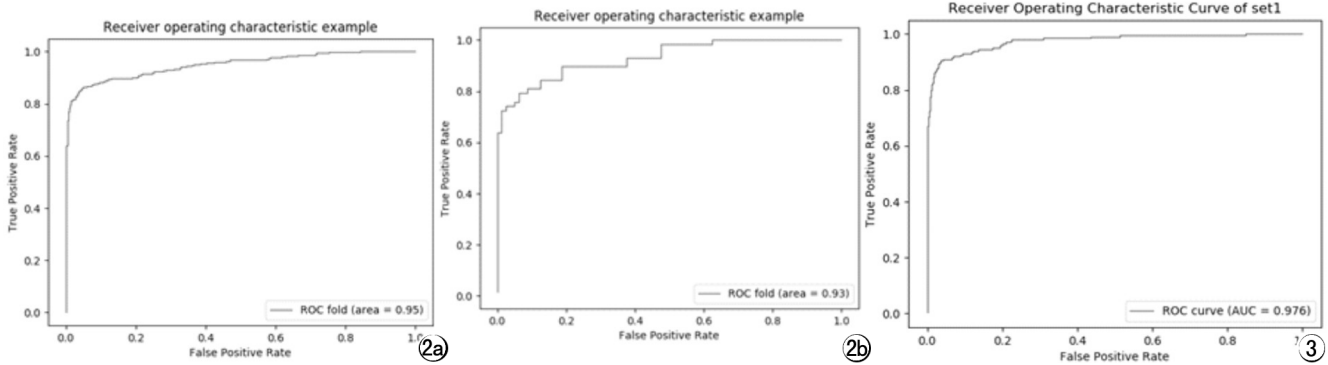


图 2 深度学习模型在公开数据集 ChinaSet(a)和 MontgomerySet(b)上的 ROC 曲线分析。 图 3 深度学习模型在深圳三院临床数据集上的 ROC 曲线分析。

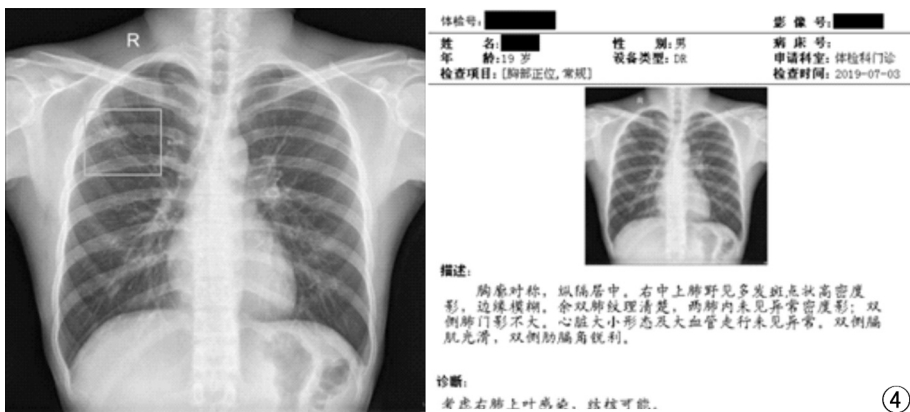


图 4 典型的在体检影像报告显示为结核疑似患者右肺上叶结核在被深度学习模型识别检出。

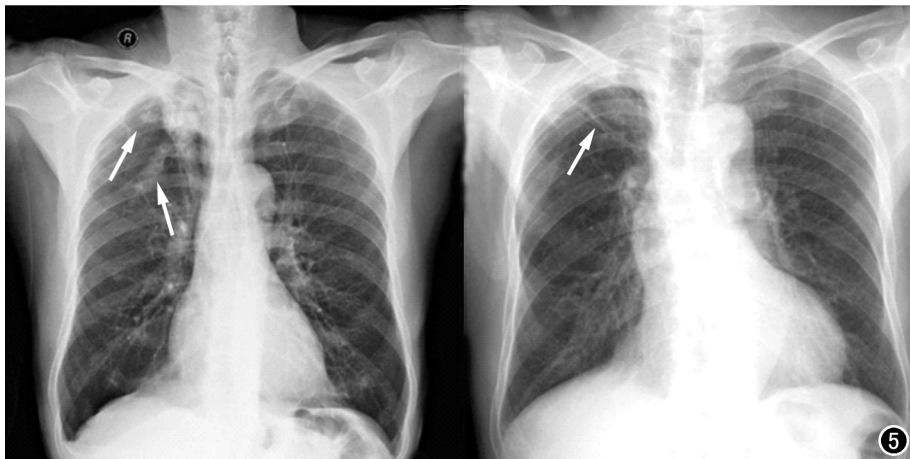


图 5 两例深度学习模型漏诊病例分析,肺结核病灶用白色箭头标出。

讨 论

本研究建立了基于胸部 X 线的肺结核检出深度学习模型,采用了基于 RetinaNet 网络的检测方法,检测精度高^[15]。内部测试数据中模型的 AUC 为 0.967,几乎达到了 100% 无漏诊。在公开数据集 ChinaSet 和 MontgomerySet 上,模型的 AUC 分别为 0.95 和 0.93,相比于内部测试集结果模型性能无明显下降,表明该模型不存在过拟合的现象,具有较高的鲁棒性。此外,我们还与商业软件 Qure.ai 在胸部 X 线肺结核诊断的结果进行了比较,不论是在 ChinaSet 公开数据集上 (AUC, 0.95 vs 0.87) 还是 MontgomerySet 数据集上 (AUC, 0.93 vs 0.91),我们建立的模型都有着更高的诊断准确率^[16]。在来自深圳三院外部验证数据集上模型的 AUC 仍然达到了 0.976,与模型在内部验证集以及 2 组公开数据集上的测试结果几乎完全一致,表明模型具有很好的泛化性。

相比于之前关于深度学习在胸部 X 线结核检测中的报道,本研究有以下优点:①本研究纳入了来自三家医院的 2600 例胸部 X 线数据用于模型训练,据我

们所知这是目前最大的肺结核胸部 X 线多中心数据库。多中心、大样本的数据为深度学习模型的构建提供了保障,实验结果也证明我们的 AI 模型有着良好的效能和优秀的鲁棒性。②本研究通过构建基于类似 RetinaNet 的 ResNeXt-FPN 网络架构进行肺结核病灶的检测,相比于之前报道中所采用的 AlexNet 和 GoogLeNet 等传统模型,本研究在胸部 X 线整图诊断的基础上进一步的实现了病灶定位,从而更有利于辅助临床医生进行诊断,同时源于 RetinaNet 的改良损失函数解决了目标检测模型中正负样本严重不平衡的问题,金字塔形的网络结构也充分利用了多尺度的特征,从而保证了检测精度。③本研究分别在公开数据集和来自临床的独立数据集上进行的模型性能的外部验证,充分验证了 AI 模型的鲁棒性和泛化性,为临床应用提供了数据支撑。④本研究还在深圳三院临床数据集中计算了最低样本量,确保了模型在敏感度和特异度上相

对于 WHO 标准的优效性验证,同时还进行了 AI 模型的临床一致性研究,充分证明了模型的临床应用可靠性。

AI 模型可以精确计算出每个病例的患病风险,通过人为设定一个阈值 (threshold) 来鉴别诊断正负样本,随着阈值从小到大逐渐升高,模型敏感性随之降低的同时特异性也随之升高,反之亦然。约登指数 (Youden Index) 是常用的评价筛查试验真实性的方法,其应用的前提是假设假阴性 (漏诊率) 和假阳性 (误诊率) 的危害性具有同等意义^[17]。针对不同疾病诊断的实际临床实验中,假阴性和假阳性的危害性往往并不相同,因此需要通过调节阈值大小来进行平衡。本研究基于 WHO 在肺结核诊断上的标准,在深圳三院临床测试数据集上通过人为设定阈值为 0.17,确保了模型在具有高敏感性 (97.50%) 的同时也有良好的特异性 (77.52%)。

本研究也存在局限性:首先,本研究为回顾性研究,未来还需要前瞻性的搜集更多病例来验证模型的临床应用效能;其次,本研究只分析了 AI 对于结核病灶的检出效能,还需要进行阅片实验来探索深度学习

模型与人类医生在结核检查的效能差异,以及评估医生在 AI 辅助下对结核检测准确率的提升;再次,临床验证结果表明虽然 AI 对于肺结核的检测准确率较高,但是对于病变特征不明显且有其他部位遮挡的结核病灶仍然存在漏诊的情况,需要针对性的进行模型优化以降低漏诊率;最后,本研究开发的 AI 模型只针对了肺结核病灶的检测,在训练过程中未纳入肺部其他病变,因此目前只能局限于肺结核的检测,未来将进一步拓展针对其他肺部病变的检测功能。

总之,本研究通过搜集多中心来源的大样本数据集,构建了具有高鲁棒性和泛化性的深度学习胸部 X 线影像肺结核检测模型,并采用了多个外部独立测试集进行性能验证,同时还在深圳三院临床数据集中对模型的优效性和可靠性进行了评估。结果表明基于深度学习的胸部 X 线肺结核病灶智能检测模型拥有较高的准确率和良好的可靠性,有助于更准确、有效地指导肺结核疾病的临床医疗决策。

参考文献:

- [1] Chakaya J, Khan M, Ntoumi F, et al. Global tuberculosis report 2020—reflections on the global TB burden, treatment and prevention efforts[J]. International J Infectious Diseases, 2021, 113(1): 7-12.
- [2] 李文文, 冯喜英, 关巍. 结核病诊断与治疗的研究进展[J]. 中华肺部疾病杂志(电子版), 2016, 9(2): 204-206.
- [3] 陈效友. 肺结核诊断的研究进展[J]. 中国临床医生, 2013, 41(3): 4.
- [4] 邱秀清, 朱秋霞, 陈玉明. 胸片普查对于肺结核阳性检出率的意义[J]. 现代养生, 2017, 32(14): 55-55.
- [5] Miller C, Lonnroth K, Sotgiu G, et al. The long and winding road of chest radiography for tuberculosis detection[J]. Eur Respir J, 2017, 49(5): 1700364.
- [6] 陆普选. 结核病对临床影像学诊断的挑战[J]. 放射学实践, 2011,

26(9):916-917.

- [7] 李娟, 汤翔宇, 朱文珍. 基于卷积神经网络的深度学习算法对颅内出血的类型识别及血肿分割一致性的研究[J]. 放射学实践, 2021, 36(1): 7-12.
- [8] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. Jama, 2016, 316(22): 2402-2410.
- [9] Andre E, Brett K, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115-118.
- [10] Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J]. Cell, 2018, 172(5): 1122-1131.e9.
- [11] Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network[J]. Nat Med, 2019, 25(1): 65-69.
- [12] Qin ZZ, Sander MS, Rai B, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems[J]. Sci Rep, 2019, 9(1): 15000.
- [13] Lakhani P, Sundaram B. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks[J]. Radiology, 2017, 284(2): 574-582.
- [14] Jaeger S, Candemir S, Antani S, et al. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases[J]. Quant Imaging Med Surg, 2014, 4(6): 475-477.
- [15] Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42(2): 318-327.
- [16] None. ECR 2018-Book of Abstracts[J]. Insights Into Imaging, 2018, 9(1): 473-473.
- [17] Li C, Chen J, Qin G. Partial Youden index and its inferences[J]. J Biopharm Stat, 2019, 29(2): 385-399.

(收稿日期: 2021-07-13 修回日期: 2021-12-15)