

· 影像组学与人工智能专题 ·

基于胸部 X 线片的深度迁移学习模型早期诊断儿童肺炎支原体肺炎

孟名柱, 潘昌杰, 张浩, 陈晓霞

【摘要】 目的:探讨基于 VGG19 的深度迁移学习模型在儿童胸部 X 线片(胸片)早期诊断肺炎支原体肺炎(MPP)中的价值。**方法:**搜集常州市第二人民医院儿科就诊的患儿 3763 例,其临床及影像资料齐全。将患儿胸片分成细菌性肺炎组、MPP 组和病毒性肺炎组,并按照 9:1 随机分为训练集和测试集。另外选择 150 例患儿胸片作为验证集(细菌性肺炎、MPP 和病毒性肺炎各 50 例)。模型性能评价指标包括在训练集和测试集中的准确率(Ac)及在验证集中的精确度(Pr)、召回率(Rc)、F1 评分(F1)和 ROC 曲线下面积(AUC)。**结果:**VGG19 在训练集和测试集中的最高 Ac 分别为 0.99 和 0.95。细菌性肺炎组的 Pr、Rc、F1、AUC 分别为 0.87、0.82、0.85、0.92, MPP 组分别为 0.85、0.74、0.78、0.90, 病毒性肺炎组分别为 0.76、0.88、0.81、0.93。VGG19 对三组图像分类诊断的总体 AUC 为 0.92。**结论:**基于 VGG19 的深度迁移学习模型是儿童胸部 X 线片诊断 MPP 的可靠方法,能够帮助临床早期诊断 MPP。

【关键词】 肺炎支原体肺炎; 胸片; 儿童; 深度迁移学习; 人工智能

【中图分类号】 R725.6; R814.41 **【文献标识码】** A **【文章编号】** 1000-0313(2024)01-0037-05

DOI:10.13609/j.cnki.1000-0313.2024.01.007

开放科学(资源服务)标识码(OSID):



Early diagnosis of mycoplasma pneumonia in children based on chest X-ray images using deep transfer learning method MENG Ming-zhu, PAN Chang-jie, ZHANG Hao, et al. Department of Radiology, the Second Hospital of Changzhou Affiliated to Nanjing Medical University, Jingsu 213164, China

【Abstract】 Objective: This study was aimed to investigate the diagnostic value of a deep transfer learning method based on chest X-ray images for mycoplasma pneumonia (MPP) in children. **Methods:** A total of 3763 cases from the Department of Pediatrics of the Second People's Hospital of Changzhou were recruited in this study. Clinical data and radiographic data of each patient were completely recorded. The chest X-ray images were divided into bacterial, MPP, and viral groups. Each group was randomly divided into a train set and test set by a ratio of 9:1. Deep transfer learning based on the VGG19 model was established. Also, another 150 chest X-ray image samples with bacterial pneumonia, MPP and viral pneumonia were selected as a validation set; 50 cases each. Metrics for model performance evaluation include accuracy in the train set and test set, precision (Pr), recall rate (Rc), F1 score (F1) and area under the ROC curve (AUC) in the validation set. **Results:** Accuracies of VGG19 on the train set and the test set were 0.99 and 0.95, respectively. The values of Pr, Rc, F1 and AUC were bacterial group (0.87, 0.82, 0.85, 0.92), MPP group (0.85, 0.74, 0.78, 0.90) and viral group (0.76, 0.88, 0.81, 0.93), respectively. The VGG19 model achieved an overall AUC of 0.92 for three groups. **Conclusion:** The VGG19 model is a reliable method to classify chest X-ray images of MPP children, which is a promising AI-assisted diagnosis method for early diagnosis of MPP.

【Key words】 Mycoplasma pneumoniae pneumonia; Chest films; Child; Deep transfer learning; Artificial intelligence

肺炎支原体肺炎(mycoplasma pneumoniae pneu-

monia, MPP)是由于肺炎支原体(mycoplasma pneumoniae, MP)所引起的呼吸道感染及肺部的急性炎症性改变,以儿童和青年为主^[1]。一项前瞻性、多中心的儿童肺炎支原体流行病学监测研究显示,在新型冠状病毒肺炎流行期间,有 9.0% 的患儿检测到有 MP 感

作者单位:213164 江苏,南京医科大学附属常州第二人民医院医学影像科
作者简介:孟名柱(1984-),男,安徽凤阳人,硕士研究生,主治医师,主要从事人工智能辅助影像诊断工作。
通讯作者:陈晓霞, E-mail: czeycxs2022@163.com
基金项目:常州市科技局应用基础研究计划(CJ20220260)

染^[2]。MP 对呼吸道上皮细胞具有高亲和力,易引发机体毒性作用及严重的炎症反应,而且病程较长,可导致患儿免疫功能降低,对患儿生长发育造成一定影响。因此,MPP 的早期诊断及治疗非常重要。MPP 患者的临床表现缺乏特异性,目前确诊还需要实验室检查^[3]。MP 的检测方法主要分病原学检测和血清学检测两大类,不同的检测方法各有其特点,应依据不同需求和目的选择不同检测方法,其中聚合酶链反应(polymerase chain reaction,PCR)被认为是金标准,但该方法操作复杂,而且对实验人员及仪器的要求较高,对于一些基层医院及一般实验室难以开展。此外,PCR 比较耗时,不利于 MPP 的早期诊断。胸部 X 线评估对儿童 MPP 诊疗管理具有重要意义^[4]。但是由于不同病原体所致的肺炎在胸部 X 线片(简称胸片)中的表现具有相似性,即使是经验丰富的影像科医生,也难以通过胸片做出 MPP 的诊断^[3]。

因此,如果能够运用人工智能技术辅助 MPP 的诊断,在节约人力的同时还可以早期诊断 MPP。随着人工智能的不断发展,深度学习在医学图像分析领域的应用逐渐增多^[5]。其中,医学影像与人工智能的结合是最具发展前景的领域^[6]。Thakur 等^[7]应用深度学习方法在胸片和 CT 混合图像的数据集中对 COVID-19 患者进行分类诊断,结果显示准确率达98.28%。戴正行等^[8]的研究结果显示,相较于人工诊断,应用人工智能技术对肺结节的良、恶性进行诊断具有良好、可靠的诊断准确性。本文提出了一种基于 VGG19 的深度迁移学习(deep transfer learning,DTL)模型算法对 MPP 胸部 X 线片进行诊断和鉴别诊断,旨在实现 MPP 的自动检测,以期辅助影像科医师在胸片中对 MPP 进行早期诊断,这有助于 MPP 患儿及时有效的治疗,减少耐药性的发生。

材料与方法

1.病例资料

搜集 2017 年 1 月至 2021 年 12 月在常州市第二人民医院儿科就诊的患儿共 3763 例,临床资料及影像资料(胸部 X 线片)齐全。将患儿胸片分为细菌性肺炎组(简称细菌组,1273 例,其中男 673 例,女 600 例)、MPP 组(1246 例,其中男 962 例,女 284 例)和病毒性肺炎组(简称病毒组,1244 例,其中男 752 例,女 492 例)。本研究为回顾性研究,通过医院伦理审查委员会批准[编号(2020)KY234-01],免除患者知情同意。

本研究 MPP 组病例的纳入标准:即临床有肺炎的表现和胸部 X 线改变,且 MP 病原学检查阳性的患儿。MPP 诊断标准参照《儿童肺炎支原体肺炎诊治专

家共识(2015 年版)》^[9]。本研究细菌性肺炎和病毒性肺炎诊断标准参照《儿童社区获得性肺炎管理指南(2013 修订)(上)》^[10]。

2.影像学检查

胸片的数据采集分别在 6 台 X 线设备上,进行,后前位摄片(PA)占比 73.82%(表 1)。因胸部侧位 X 线片较少,不纳入本研究。

表 1 三组患儿数据采集 X 线设备分布

设备	PA	AP	管电压(kV)	管电流(mAS)
飞利浦 DR 4 台	1260	425	81	自动
AGFA DR 1 台	826	330	68	自动
西门子 DR 1 台	692	230	66	11.2
合计	2778	985		

注:PA,后前位摄片;AP,前后位摄片。

3.训练集、测试集和验证集

训练集和测试集图像由两部分构成。第一部分为网络公开数据集(<https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>):经过筛选,共获得此部数据图像 9882 幅^[11]。第二部分为自建数据集,共获得图像 3763 幅。运用程序指令将 3 组图像随机打乱,并按照 9:1 随机分为训练集(细菌组 6837 幅,MPP 组 1122 幅,病毒组 4327 幅)和测试集(细菌组 756 幅,MPP 组 124 幅,病毒组 480 幅)。另外选择 150 例患儿胸片作为验证集(细菌组、MPP 组和病毒组各 50 幅),验证集的图像没有在训练集中出现过。数据集总体分布见图 1。

4.计算机环境配置

采用 Windows 10 企业版 64 位操作系统,处理器为酷睿 i7 10700F,8 核心 16 线程,图形处理器(graphics processing unit, GPU)为 NVIDIA GeForce RTX2060,显存 6GB,程序运行时关闭其它所有应用程序。模型参数设置见表 2,DTL 模型结构设计见图 2。

表 2 模型参数设置

模型参数	参数值
旋转角度	60°
剪切范围	0.2
扩大范围	0.2
镜像翻转	True
垂直翻转	True
充填模式	Nearest
保存最好模型	True
早停	False
激活函数	Relu
分类函数	Softmax
丢失值	0.5
训练轮回	250

5.模型评估指标

模型评估指标中的准确率(accuracy, Ac)、损失值、精确度(precision, Pr)、召回率(recall rate, Rc)、F1

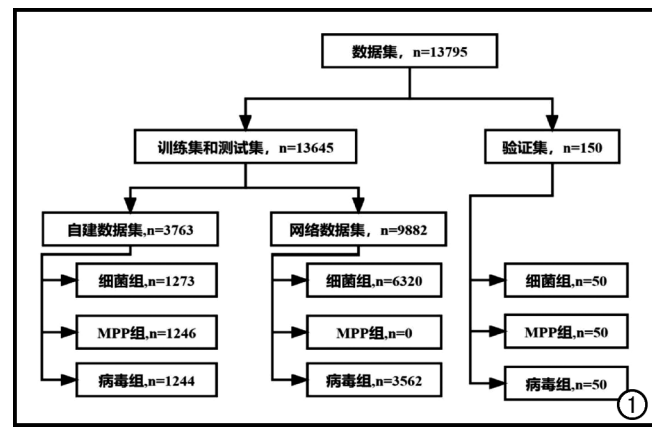


图 1 数据集分布图。

评分(F1 score, F1)和受试者工作特征(receiver operating characteristic, ROC)曲线的曲线下面积(area under curve, AUC)为模型运算得出。平均 Ac、平均 Rc 和平均 F1 为模型对验证集中三组数据验证结果的平均值,而不是模型在验证集中三组数据 Ac、Rc、F1 的平均数。计算公式分别如下:

$$Ac = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Pr = \frac{TP}{TP + FP} \quad (2)$$

$$Rc = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2 \times PPV \times Rc}{PPV + Rc} \quad (4)$$

在本研究中,TP 和 TN 分别表示模型正确诊断的三组图像数量,FP 和 FN 分别表示模型错误诊断的三组图像数量,AUC 为模型运算得出。F1 分数值为 0~1 之间,同时兼顾了分类模型的 Ac 和 Rc,可以认为是模型 Ac 和 Rc 的一种加权平均。

6. 统计学分析

采用 SPSS 23.0 软件进行统计学分析。患者年龄以均值±标准差表示,组间比较采用方差分析。以 $P < 0.05$ 差异有统计学意义。

结 果

1. 患儿年龄比较

三组患儿的年龄比较差异有统计学意义($F = 90.78, P < 0.01$),其中细菌组患儿年龄 $[(4.86 \pm 2.45)$ 岁]与病毒组 $[(4.94 \pm 2.52)$ 岁]比较差异无统计学意义($P = 0.384$),细菌组、病毒组患儿年龄与 MPP 组 $[(3.75 \pm 2.46)$ 岁]比较差异均有统计学意义(P 值均 < 0.01)。

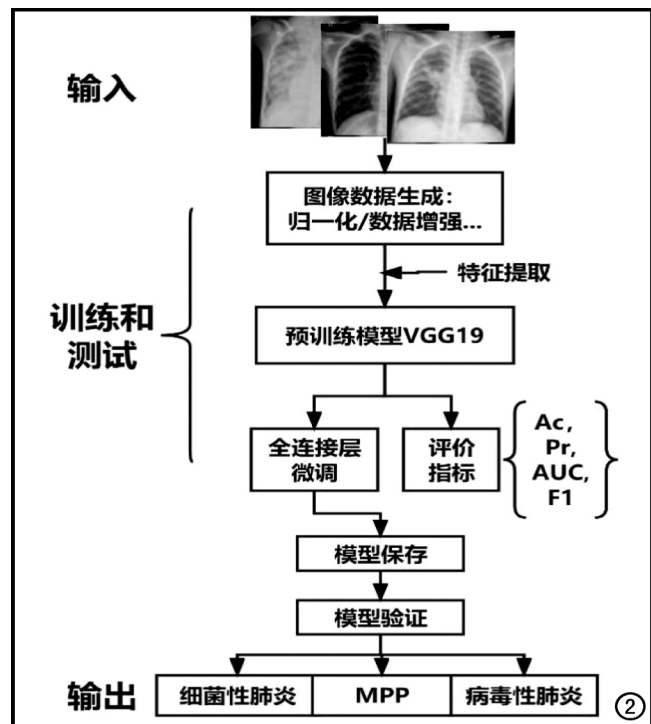


图 2 深度学习模型架构图。

2. 数据训练结果

学习曲线图(图 3)显示训练集准确率随着 epoch 的增加而逐渐增加,最高准确率为 0.99,且保持相对稳定。测试集准确率变化幅度较训练集大,最高准确率为 0.95。训练集的损失值随着 epoch 增加而逐渐减小,并趋于稳定。测试集的损失值在第 20 个 epoch 时最低,在此之后有逐渐增高的趋势。以上结果表明所建立的迁移学习模型在数据集中已经出现了过拟合现象。

3. 类激活热力图

类激活热力图是一种深度学习的可视化方法,它有助于了解一幅图像的哪一部分让卷积神经网络做出了最终的分类决策。这有助于对卷积神经网络的决策过程进行调试,特别是出现分类错误的情况下。类激活热力图是与特定输出类别相关的二维分数网格,对任何输入图像的每个位置都要进行计算,它表示每个

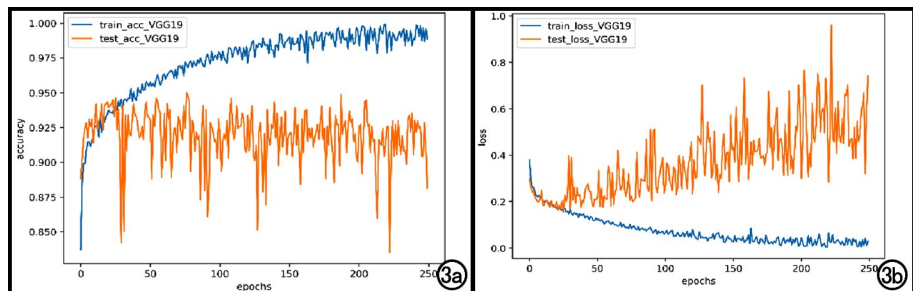


图 3 VGG19 在训练集和测试集中的学习曲线图。a)在训练集和测试集中的准确率;b)在训练集和测试集中的损失值。accuracy:准确率;epochs:训练的回合数,1 个 epoch 表示训练了 1 遍训练集中的所有样本;loss:损失值。

位置对该类别的重要程度。本研究结果显示卷积神经网络对细菌组胸片的激活部位与激活程度明显多于 MPP 组和病毒组(图 4)。

4.DTL 模型验证结果

混淆矩阵图(图 5)分析结果显示在验证集中,VGG19 对 MPP 组的诊断准确率为 0.74(37/50),低于细菌组的 0.82(41/50)和病毒组的 0.88(44/50)。

VGG19 模型在验证集中对细菌组、MPP 组和病毒组胸部 X 线片的平均 Pr、Rc 和 F1 值分别为 0.81、0.81 和 0.81。VGG19 在验证集中对细菌组的 Pr 值和 F1 值最高,分别为 0.87、0.85。VGG19 对病毒组的 Rc 值最高,为 0.88(表 3)。

VGG19 对三组图像分类诊断的总体 AUC 为 0.92,对病毒组分类诊断的 AUC 最高,为 0.93,对细菌组和 MPP 组分类诊断的 AUC 分别为 0.92 和 0.90(图 6)。

讨 论

不同病原所致的儿童肺炎胸片表现有所不同,细菌性肺炎以肺部实质性改变为主,MPP 胸片可

表 3 验证集验证结果分类报告

分类	Pr	Rc	F1	支持图像数 (幅)
细菌组	0.87	0.82	0.85	50
MPP 组	0.82	0.74	0.78	50
病毒组	0.76	0.88	0.81	50
平均/总数	0.81	0.81	0.81	150

见支气管周围斑片状影、网状影及局灶性实变影,但是由于影像征象的重叠以及人类视觉能力的局限性,通过胸片无法对细菌性肺炎与 MPP 进行鉴别诊断。本研究建立了基于 VGG19 的胸部病变 X 线三分类 DTL 模型,并在验证集中对此模型的鲁棒性进行了进一步验证。通过对性能指标的评估和比较,结果显示 VGG19 模型对三组病变均有较好的分类效果,总体分类诊断 AUC 达 0.92。

Jia 等^[12]建立的 DTL 模型对胸片进行分类诊断,在训练集中的准确率达 99.6%,这说明了 DL 在辅助医学影像诊断中的巨大潜力。本研究结果显示,基于 VGG19 模型在训练集和测试集中均得到了较高的准

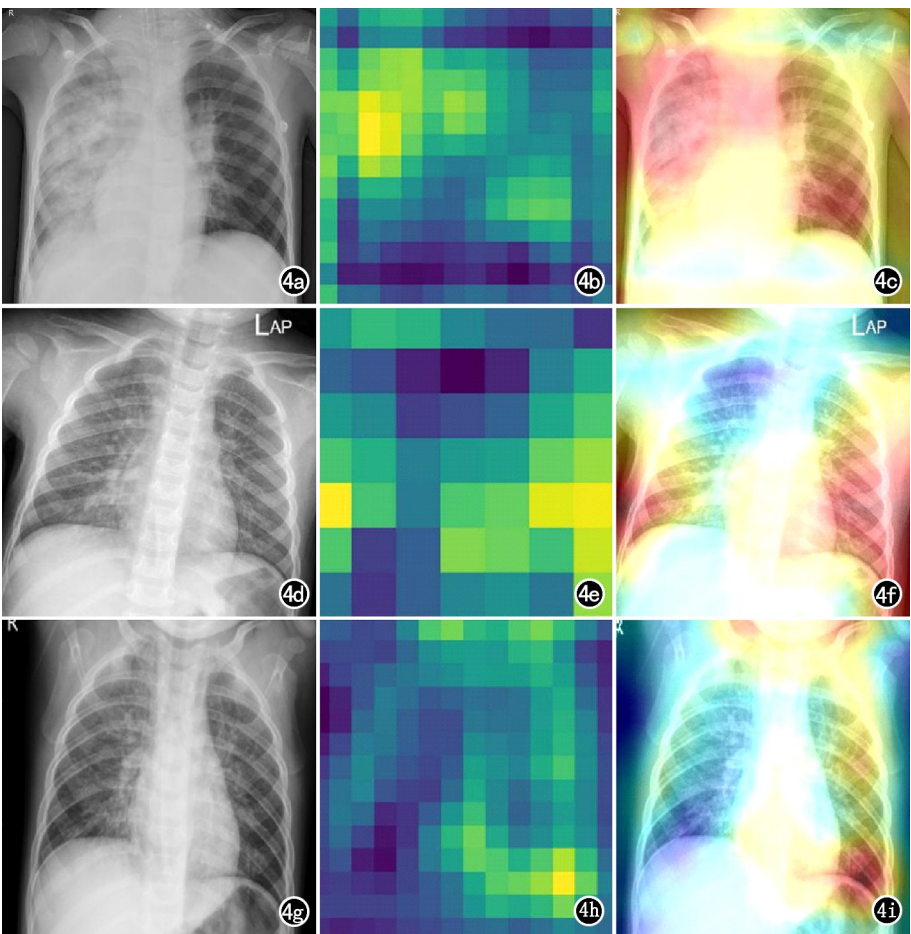


图 4 类激活热力图。a)细菌性肺炎的胸部 X 线图像;b)细菌性肺炎对应的类激活热力图;c)细菌性肺炎对应的类激活热力图融合图像;d)MPP 的胸部 X 线图像;e)MPP 对应的类激活热力图;f)MPP 对应的类激活热力图融合图像;g)病毒性肺炎的胸部 X 线图像;h)病毒性肺炎对应的类激活热力图 ;i)病毒性肺炎对应的类激活热力图融合图像。

确率,分别为 0.99 和 0.95,但在测试集中的损失值却有逐渐增高的趋势,这是模型过拟合现象。造成模型过拟合的主要原因有以下两点:首先为新任务的数据集较小,其次为所用的模型过于复杂。模型的过拟合是导致模型泛化能力差的重要原因。为了进一步验证本研究训练出的模型的泛化能力,我们又设置了验证集,而且此部分的数据图像没有在训练集和测试集中出现过。结果显示,在细菌组、MPP 组和病毒组的 AUC 分别为 0.92、0.90 和 0.93,而且总体 AUC 为 0.92。在临床实践中,放射科医生通过儿童胸部 X 片来对儿童肺炎类型进行分类诊断是比较困难的。这提示我们,DL 在辅助儿童胸部 X 片肺炎分类诊断方面的能力高于人工诊断。MPP 的早期正确诊断是 MPP 患儿得到及时有效治疗的前提,也是减少耐药性发生的重要基础。

此外,本研究还通过类激活热力图形式进一步展示了深度学习模型的过程可视化,通过类激活热力图

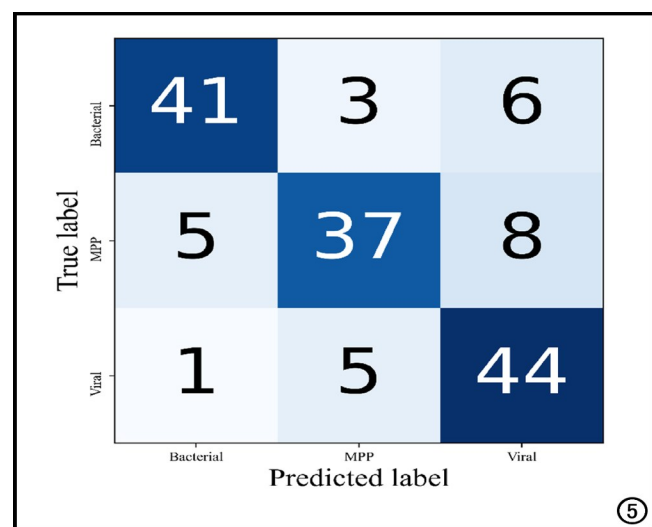


图5 VGG19 在验证集中的混淆矩阵图,图中的数字为模型正确分类的胸部 X 线片的数量。Bacteria:细菌组;MPP:MPP 组;Viral:病毒组

可直观地观察到卷积神经网络对不同图像的激活区域和方式各不相同,因此能够对细菌组、MPP 组和病毒组图像进行正确分类。这也为进行进一步的相关研究打下了基础。

本研究存在以下局限性:首先,本研究数据集 73.82% 为胸部后前位 X 线片,这导致网络对前后位胸片的学习不足,会影响模型的鲁棒性;其次,本研究数据为单中心,这也会影响模型的泛化能力。多中心来源的大样本数据集可以构建高泛化能力的深度学习模型^[13];再次,社区获得性肺炎中除细菌性肺炎及肺炎支原体肺炎以外,还有病毒性肺炎、真菌性肺炎及混合型感染肺炎^[14],本研究没有将后者纳入分析;最后,尚有其它深度学习模型可用于胸部 X 线片的分类诊断,这需要花费更多的资源和更多的精力。在此,笔者呼吁更多的专业技术人员参与这项研究。

综上所述,在通过胸部 X 线片对儿童细菌性肺炎、MPP 与病毒性肺炎的分类诊断中,基于 VGG19 的迁移学习模型是一种可靠的方法。该方法可以有效帮助放射科医生做出 MPP 诊断,为 MPP 的早期诊断奠定了基础,能够为临床医师制定合理的治疗策略提供可靠的循证医学证据,尤其是在实验室条件相对较差的医院将会起到更好的辅助诊断作用。

参考文献:

- [1] 中华医学会儿科学分会临床检验学组.儿童肺炎支原体呼吸道感染实验室诊断中国专家共识[J].中华检验医学杂志,2019,42(7):507-513.
- [2] Chen J,Zhang J,Lu Z,et al.Mycoplasma pneumoniae among Chinese Outpatient Children with Mild Respiratory Tract Infections during the Coronavirus Disease 2019 Pandemic[J].Microbiol Spectr,2022,10(1):e01550-01521.

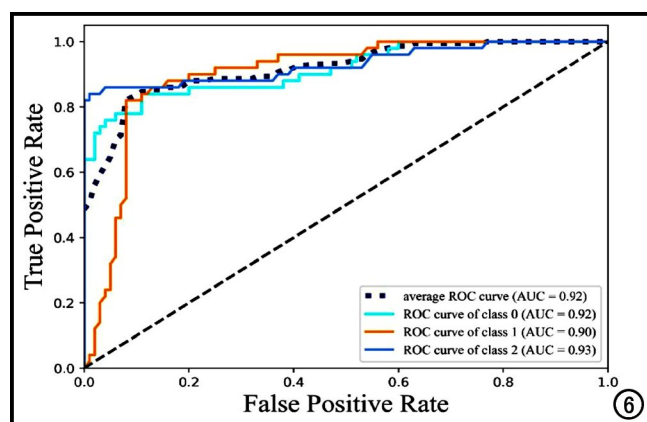


图6 VGG19 对三组图像分类诊断 ROC 曲线。average:模型对细菌组、MPP 组和病毒组 3 组图像分类的平均 ROC 曲线;class0:细菌组;class1:MPP 组;class2:病毒组。

- [3] 赵德育,邓欢.肺炎支原体肺炎实验室诊断的意义及进展[J].中华实用儿科临床杂志,2018,36(16):1205-1209.
- [4] Cho YJ,Han MS,Kim WS,et al.Correlation between chest radiographic findings and clinical features in hospitalized children with Mycoplasma pneumoniae pneumonia[J].PloS one,2019,14(8):e0219463.
- [5] Abumalloh RA,Nilashi M,Youssof Ismail M,et al.Medical image processing and COVID-19:A literature review and bibliometric analysis[J].J Infect Public Health,2022,15(1):75-93.
- [6] 朱文珍,胡琼洁.人工智能与医学影像融合发展:机遇与挑战[J].放射学实践,2019,34(9):938-941.
- [7] Thakur S,Kumar A.X-ray and CT-scan-based automated detection and classification of COVID-19 using convolutional neural networks (CNN)[J].Biomed Signal Process Control,2021,69(8):102920.
- [8] 戴正行,胡春洪,王希明,等.基于 DenseNet 网络深度学习法 CT 图像人工智能分析技术判断肺结节良恶性[J].放射学实践,2020,35(4):484-488.
- [9] 中华医学会儿科学分会呼吸学组.《中华实用儿科临床杂志》编辑委员会.儿童肺炎支原体肺炎诊治专家共识(2015 年版)[J].中华实用儿科临床杂志,2015,30(17):1304-1308.
- [10] 中华医学会儿科学分会呼吸学组.儿童社区获得性肺炎管理指南(2013 修订)(上)[J].中华儿科杂志,2013,51(10):745-752.
- [11] Loey M,El-Sappagh S,Mirjalili S.Bayesian-based optimized deep learning model to detect COVID-19 patients using chest X-ray image data[J].Comput Biol Med,2022,142(3):105213.
- [12] Jia G,Lam HK,Xu Y.Classification of COVID-19 chest X-ray and CT images using a type of dynamic CNN modification method[J].Comput Biol Med,2021,134(7):104425.
- [13] 安超,张晨,郑广平,等.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[J].放射学实践,2022,37(6):704-709.
- [14] 赵登峰,徐慧,陈恒,等.不同类型病原体感染社区获得性肺炎临床与影像特征及转归[J].中华医院感染学杂志,2021,31(18):2768-2772.

(收稿日期:2022-09-26 修回日期:2023-02-16)