



Research
Coronavirus Disease 2019—Article

利用深度学习系统筛查新冠病毒肺炎

徐小微^{a, #}, 蒋贤高^{b, #}, 马春莲^{c, #}, 杜鹏^{d, #}, 李旭坤^d, 吕双志^e, 俞亮^a, 倪勤^a, 陈燕飞^a, 苏俊威^a, 郎观晶^a, 李永涛^a, 赵宏^a, 刘俊^a, 徐凯进^a, 阮凌翔^e, 盛吉芳^a, 裘云庆^a, 吴炜^{a, *}, 梁廷波^{f, *}, 李兰娟^{a, *}

^a State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China

^b Department of Infectious Disease, Wenzhou Central Hospital, Wenzhou 325000, China

^c Department of Infectious Disease, The First People's Hospital of Wenling, Wenling 317500, China

^d Artificial Intelligence Lab, Hangzhou AiSmartVision Co., Ltd., Hangzhou 310012, China

^e Department of Radiology, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China

^f Department of Hepatobiliary and Pancreatic Surgery & Key Lab of Pancreatic Diseases Research of Zhejiang Province & The Innovation Centre for the Study of Pancreatic Diseases of Zhejiang Province & Clinical Medical Research Center of Hepatobiliary and Pancreatic Diseases in Zhejiang Province & Precision Innovation Center of the Diagnosis and Treatment of Hepatobiliary and Pancreatic Diseases of Zhejiang University, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China

ARTICLE INFO

Article history:

Received 23 March 2020

Revised 21 April 2020

Accepted 28 April 2020

Available online 27 June 2020

关键词

COVID-19

位置敏感机制深度学习网络

计算机断层扫描

摘要

实时逆转录聚合酶链反应 (RT-PCR) 检测早期新冠病毒肺炎 (COVID-19) 患者的痰液或鼻咽拭子中的病毒RNA阳性率较低。同时, COVID-19的计算机断层扫描 (CT) 影像学的临床表现有其自身的特点, 不同于甲型流感病毒性肺炎 (IAVP) 等其他类型的病毒性肺炎。本研究旨在应用深度学习技术, 建立COVID-19、IAVP及健康人群肺部CT的早期筛查模型。本研究共采集618份CT样本, 其中219份样本来自110例COVID-19患者 (平均年龄50岁, 其中男性63例, 占57.3%), 224份样本来自224例IAVP患者 (平均年龄61岁, 其中男性156例, 占69.6%), 175份样本来自健康人群 (平均年龄39岁, 其中男性97例, 占55.4%)。所有CT样本均来自浙江省三家COVID-19定点收治医院。我们首先利用胸部CT图像集的三维 (3D) 深度学习模型分割出候选感染区域, 然后利用位置敏感机制深度学习网络将这些分离的图像归类为COVID-19、IAVP以及与感染无关 (ITI) 的图像, 并且输出相应置信度得分。最后, 用Noisy-OR贝叶斯函数计算每份CT病例的感染类型及总置信度。测试数据集的实验结果表明, 从整体CT病例来看, 本研究利用深度学习系统建立的COVID-19患者的早期筛查模型的总体准确率为86.7%。该模型有望成为一线临床医生诊断COVID-19的一种有效的辅助方法。

© 2020 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. 引言

2019年年底, 新冠病毒肺炎 (COVID-19) 首先在中国湖北省武汉市被证实[1–4]。2020年1月24日,

Huang等[5]总结了41例COVID-19患者的临床特征, 提示常见的首发症状为发热、咳嗽、肌痛或疲劳。41例患者均有肺炎表现, 而且肺部计算机断层扫描 (CT) 提示异常。这些患者出现的并发症包括急性呼吸窘迫综合

* Corresponding author.

E-mail address: 1198042@zju.edu.cn (W. Wu), liangtingbo@zju.edu.cn (T. Liang), ljl@zju.edu.cn (L. Li).

These authors contributed equally to this work.

征、急性心脏损伤和继发感染等。其中13名(31.7%)患者被送入重症监护室(ICU),6名(14.6%)患者死亡。之后,香港大学的Kok研究团队[6]首次发现了COVID-19在人与人之间传播的证据。

中国《新型冠状病毒肺炎诊疗方案(试行第七版)》[7]建议,通过核酸检测或特异性抗体检测对患者进行COVID-19的病原学确认。之后几个月,核酸检测的可及性有了很大提高,但仍存在操作要求高、耗时长、阳性率相对较低等缺点[8,9]。研究表明,新冠病毒感染者的口腔拭子、肛门拭子和血液核酸检测阳性率分别为53.3%、26.7%和40.0% [10]。另外,因为新冠病毒特异性抗体的出现存在一个窗口期,所以抗体检测对于早期筛查也存在漏诊可能。Long等[11]报道,新冠病毒特异性免疫球蛋白G(IgG)和免疫球蛋白M(IgM)的血清转化中位数时间均为发病后的13 d左右。

在疫情初期,放射学专家就注意到COVID-19病例的肺部CT影像学有独特表现,不同于甲型流感病毒性肺炎(IAVP)等其他类型的病毒性肺炎的CT影像学表现,如图1所示。因此,临床医生建议利用肺部CT影像学作为COVID-19的早期诊断标准之一[12],以弥补核酸诊断的不足,从而遏制COVID-19的传播。

随着计算机技术的飞速发展,数字图像处理技术在医学领域得到了广泛应用,包括器官分割和图像增强修复,为后续的医学诊断提供了支持[13,14]。卷积神经网络(CNN)等具有较强非线性建模能力的深度学习技术在医学图像处理中也有着广泛的应用[15-18]。国内外研究者在肺结节的诊断[19]、良性和恶性肿瘤的分型[20,21]、肺结核的分析和疾病预测[22-24]等方面开展了大量研究。

本研究利用多个CNN模型对CT图像数据集进行分类,计算了COVID-19的感染概率,并利用深度学习系统建立了COVID-19患者早期筛查模型。这项工作有助于COVID-19患者的早期筛查。

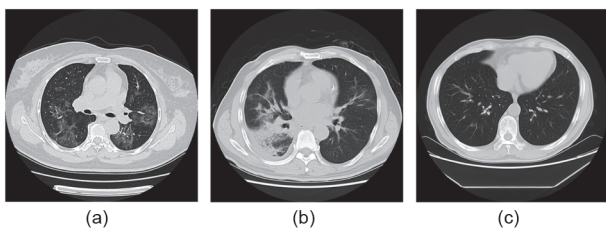


图1. 典型病例的肺部横断面CT图像。(a) COVID-19;(b) IAVP;(c)无肺炎表现。(a)和(b)均为症状出现后10 d内的表现。

2. 样本和方法

2.1. 实验数据集

本研究共收集肺部CT标本618份,其中219份来自2020年1月19日至2月14日浙江大学医学院附属第一医院、温州市中心医院、温岭市第一人民医院收治的110例COVID-19患者(平均年龄50岁,其中男性63例,占57.3%)。这三家医院都是浙江省指定的COVID-19患者收治医院。所有COVID-19患者均经痰液或鼻咽拭子的实时逆转录聚合酶链反应(RT-PCR)检测确诊,排除了部分肺部CT影像学无表现的病例。此外,同一病例的CT数据集收集时间至少有两天的间隔,以确保样本的多样性。在冬春季节,IAVP与COVID-19疑似病例有重要的鉴别诊断意义,故本研究采用224例IAVP(包括H1N1、H3N2、H5N1、H7N9等)患者(平均年龄61岁,其中男性156例,占69.6%)和175例健康人群(平均年龄39岁,其中男性97例,占55.4%)的CT样本,共计399例CT样本作为对照病例。这些对照病例均来自浙江大学医学院附属第一医院。IAVP的诊断经痰液及鼻咽拭子中甲型流感病毒RNA的RT-PCR检测证实。198例(90.4%)COVID-19病例和196例(87.5%)IAVP病例处于疾病早期或进展期($P > 0.05$),其余9.6%的COVID-19病例和12.5%的IAVP病例处于重症或危重症期。

本研究经浙江大学医学院附属第一医院伦理委员会批准,并按照相关指南和规定进行。所有患者和(或)其法定监护人在研究前均签署了知情同意书。

618份CT样本中,528份CT样本(85.4%)用于训练和验证集,其中包括COVID-19患者189例、IAVP患者194例、健康人群145例。其余90例CT样本(14.6%)作为测试集,包括COVID-19患者30例、IAVP患者30例、健康人群30例。此外,测试集中不包含已经训练过的患者的CT样本。

2.2. 实验流程

图2显示了本研究中COVID-19诊断报告生成的整个过程。首先,对CT图像进行预处理,提取有效的肺部区域。其次,利用三维CNN模型对多个候选图像进行分割,收集中心图像和与之相邻的两个图像供进一步分析。再次,利用图像分类模型将所有的候选图像分为三组:COVID-19组、IAVP组和与感染无关(ITI)组。利用来自同一个三维区域的三张图像投票选出该候选区域的类型和整体置信度得分。最后,使用Noisy-OR贝

叶斯函数计算CT样本的总体感染类型和概率[25]。

3. 理论和计算

3.1. 数据集预处理和候选区域分割

本研究在数据预处理和候选区域分割阶段使用了之前肺结核研究[23]所建立的方法和模型。该网络模型采用VNet[26]主干，以感知残差(inception-residual, IR)网络结构[27]作为特征提取部分，然后用区域建议网络(reginal proposal network, RPN)分割候选区域。肺结核感染病灶具有粟粒型、浸润型、干酪样型、结核球型、空洞型等多种结构和类型。前期研究证实该模型对于肺结核的多型性病灶具有较好的分割效果。尽管该模型是为肺结核病灶分割而训练的，但经专业放射科医生验证，它足以分离病毒性肺炎的候选感染区域。

本文使用的三维分割过程有别于传统的VNet或三维u-net像素级分割操作[28]。由于健康区域与感染灶之间的边界通常是模糊的，因此很难标记肺炎病变区的像素级掩模。此外，病变区的多变三维结构也加剧了这项工作的难度。因此，本文使用的分割操作更类似于目标检测算法中的操作，即采用RPN结构，用三维边界框代替像素级分割掩模来捕捉感兴趣区域。

此外，肺结核研究采用了VNet-IR-RPN模型进行分割和分类。因为本研究只进行分割操作，因此只保留了与分割相关的边界框回归部分。

3.2. 图像数据处理和数据扩容

大量与本研究无关的非感染区也可能被三维分割模型分割出来，包括血管、肺纤维化结构、钙化点和识

别错误的健康区域等。因此，被分割的候选区域除了COVID-19组和IAVP组外，还增加了一个ITI组。

本研究从618份CT样本(219份COVID-19样本、224份IAVP样本和175份健康样本)的三维分割模型中共生成了3957个候选立方体。只有靠近立方体中间的区域包含这个感染病灶的最大信息。因此，只收集中心图像及与其相邻的两个图像来代表该三维区域，供进一步分类。接下来，由两位专业放射科医生将所有分割出的图像块手动分为两类，即肺炎图像块和ITI图像块。根据临床诊断结果，前一类的图像被自动识别为COVID-19或IAVP。

从上述步骤共获得11 871个图像块，其中包括2634张COVID-19图像、2661张IAVP图像和6576张ITI图像。根据先前的数据集分配，训练和验证集有528个CT样本，这些样本对应10 161张(85.6%)图像，其中包括2301张COVID-19图像、2244张IAVP图像和5616张ITI的图像。剩下的1710张(14.4%)图像被保留用于测试数据集。

COVID-19和IAVP病例的抽样采纳率被扩大三倍，以平衡ITI的样本数量，从而减少不同图像类型分布不均对现有数据集的影响。同时，为了增加训练样本数，防止数据过度拟合，对样本使用了随机剪切、左右翻转、上下翻转、镜像操作等通用的数据扩展机制。

3.3. 分类的深度学习模型

3.3.1. 位置敏感的深度学习分类网络

Kanne [29]和Chung等[30]的研究显示了COVID-19病例肺部CT的三个显著特征，即磨玻璃样病灶、多分布于胸膜附近的病灶、双肺多发感染灶，如图3所示。

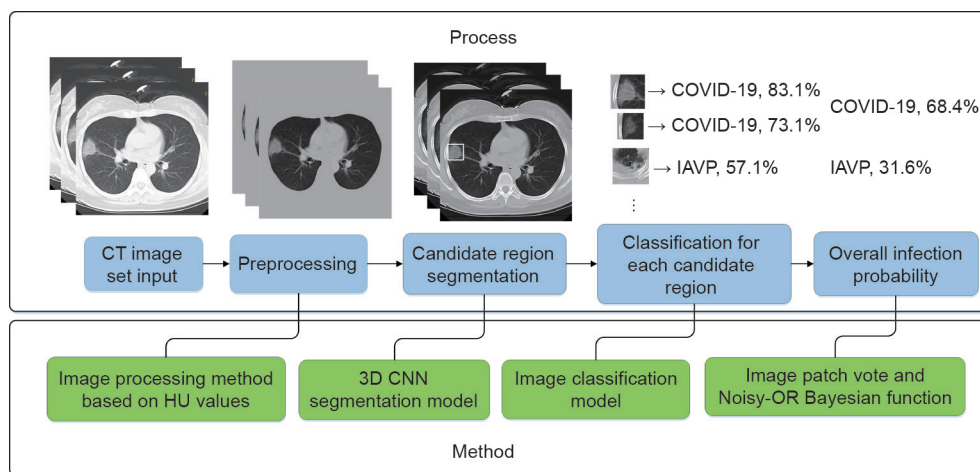


图2. COVID-19诊断流程图。HU: 亨氏单位。

根据这些发现，我们对模型进行了优化，并设计了图像分类模型来区分不同感染的外观和结构。此外，利用病灶与边缘的相对距离作为附加权重对模型进行训练，从而得到位置敏感的深度学习模型。靠近胸膜的感染病灶更容易被认为是COVID-19类型。

每个病灶与边缘的相对距离可以按照以下步骤进行计算。

步骤1: 测量从掩模到病灶中心的最小距离[双向箭头, 如图3 (c) 所示];

步骤2: 获得肺部图像最小外切矩形的对角线[图3 (d)];

步骤3: 利用从步骤1获得的距离除以从步骤2获得的对角线计算病灶与边缘的相对距离。

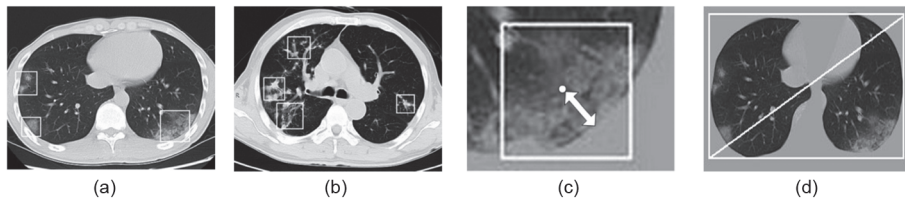


图3. (a) COVID-19图像，有三个磨玻璃感染病灶；(b) IAVP图像，有四个感染病灶；(c) 从掩模到该病灶中心的最小距离（双向箭头）；(d) 该肺部图像最小外切矩形的对角线。

3.3.2. 深度学习的网络结构

本研究评估了两个CNN分类模型，如图4所示。一种是相对传统的基于残差网络（ResNet）的模型[31]，另一种是在第一种网络结构的基础上，通过在全连接层添加位置注意机制来提高整体准确率的模型。该位置注意机制被添加到第一个全连接层，以最大限度地提高该因素对整个网络的影响。

采用经典的ResNet-18网络结构进行图像特征提取。使用池化操作进行数据降维，以防止过拟合，并提高泛化能力。

卷积层的输出被展平为256维特征向量，然后使用全连接网络将其转换为16维特征向量。对于位置敏感机制的深度学习网络，首先将病灶与边缘的相对距离规

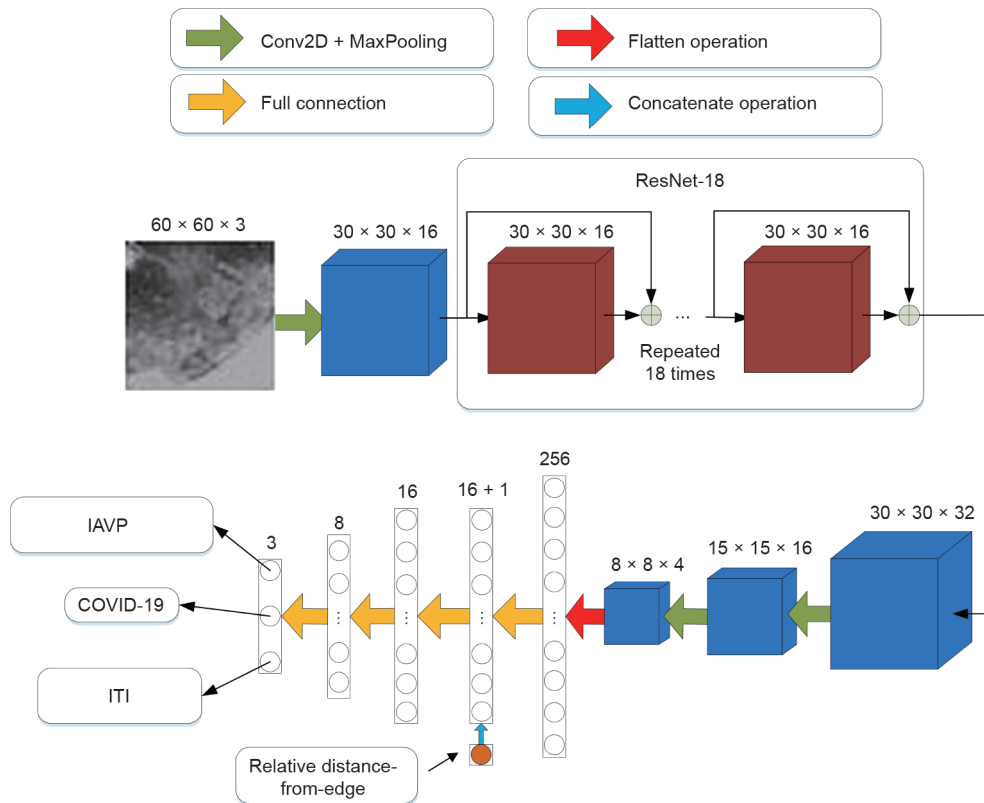


图4. 基于ResNet-18的传统分类模型的网络结构（无位置敏感机制）。在ResNet-18的主干网络上，通过在全连接层连接位置敏感机制，建立面向位置敏感的深度学习模型，从而提高整体准确率。Conv2D: 二维卷积。

范化为相同的数量级, 然后连接到这个全连接网络结构中。接下来, 按照三个全连接层输出最终的分类结果和置信度得分。

3.4. 诊断报告

3.4.1. 投票选出候选区域类型

受机器学习技术中Bagging预测算法[32]的启发, 一个候选区域由三个图像表示, 即中心图像及两个邻域图像。这三张图像用以下策略进行投票, 以确定整块候选区域的类型:

(1) 如果至少有两个图像被分为同一类型, 则选择在该类型中具有最大置信度得分的图像代表这块候选区域;

(2) 否则, 选择置信度得分最大的图像代表这块区域。

如果该候选区域被确认为ITI类型, 则此区域将在下一步中被忽略。

3.4.2. 利用 Noisy-OR 贝叶斯函数推导总体感染率

COVID-19的一个显著特征是在一个CT病例中有多个独立的感染灶[29,30]。如果一个患者的CT图像有两个COVID-19区域(两个区域都有50%的概率), 那么总体感染概率远远大于50%是合理的。因此, 使用Noisy-OR贝叶斯函数的概率公式可计算一种感染类型的总感染置信度得分(C), 具体如下所示:

$$C = 1 - \prod_i (1 - C_i) \quad (1)$$

式中, C_i 表示第*i*个区域的置信度。

根据两种类型的置信度得分($C_{\text{COVID-19}}$ 和 C_{IAVP})进行推断, 然后根据占据优势的置信度得分将该CT样本分为相应的组。

此外, 使用以下策略导出整个CT样本的置信度, 以便临床医生合理地参考:

(1) 如果 $C_{\text{COVID-19}}$ 和 C_{IAVP} 均为0, 则该CT样本属于未感染病例;

(2) 如果其中一个的置信度值等于0, 则直接导出另一个的置信度值作为该CT样本的置信度;

(3) 否则, 使用softmax函数生成两个置信度得分。

$$S_j = \frac{e^{C_j}}{\sum_k e^{C_k}} \quad (2)$$

式中, $j, k \in (\text{COVID-19, IAVP})$ 。二者的 S_j 均作为每种感染类型的置信度得分导出。softmax操作将 S_j 总和标准化为100%, 并且没有改变感染类型的判断结果。一旦系统发现CT影像存在COVID-19可疑区域, 则建议临床医生进行人工复查。

4. 结果

4.1. 评估计算平台

测试服务器采用Intel i7-8700k中央处理器(CPU)和NVIDIA GPU GeForce GTX 1080ti。处理时间在很大程度上取决于一组CT图像的层数。从数据预处理到报告输出, 处理一组70层的CT图像平均用时不到30 s。

4.2. 训练过程

本研究采用交叉熵(分类模型中最经典的损失函数之一)作为损失函数。当训练迭代数达到1000次以上时, 损失值并没有明显的减少或增加, 说明模型收敛到一个相对较优的状态, 而且没有明显的过度拟合。两个分类模型的损失值和准确率的训练曲线如图5所示。与原始的ResNet相比, 采用位置敏感机制的网络在训练数据集上取得了更好的性能。

4.3. 模型在测试数据集的表现

4.3.1. 性能测量方法

混淆矩阵是用来描述一个分类模型在一组真实值

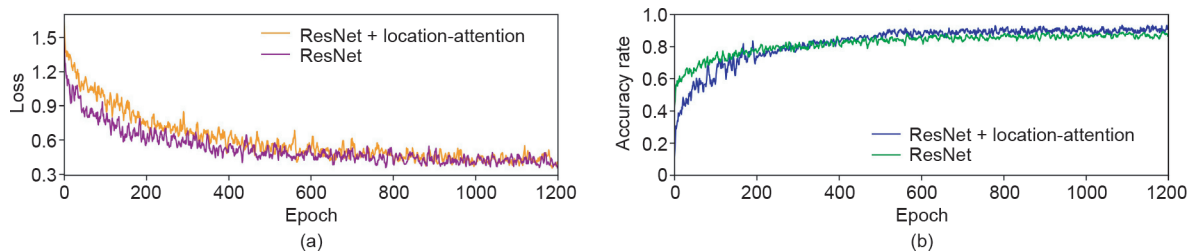


图5. 两种分类模型的损失值(a)和准确率(b)的训练曲线。

已知的测试数据上的性能的表。它将算法性能可视化。方法的准确率 (accuracy, A) 决定了预测值的准确性。正确率 (precision, P) 表示预测结果中正样本的准确率。召回率 (recall, R) 表示发现了多少正确的结果。 f_1 -score使用正确率和召回率的调和来计算平衡的结果。下列方程式说明如何计算这些值, 其中TP、TN、FP和FN分别为真阳性 (true positive)、真阴性 (true negative)、假阳性 (false positive) 和假阴性 (false negative)。

$$A = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$R = \frac{TP}{TP + FN} \quad (5)$$

$$f_1\text{-score} = \frac{2 \times P \times R}{P + R} \quad (6)$$

4.3.2. 图像预处理和分割

从每组样本中随机抽取90例CT样本 (COVID-19组30例、IAPV组30例、健康组30例) 作为测试集。测试集的选择遵循以下规则: 该病例的任何CT样本在前一阶段没有接受过训练, 以避免模型学习类似的CT。此外, 对图像预处理和分割的阈值进行优化, 使之更适合当前的研究。在图像预处理阶段, 将用于二值化再采样图像的亨氏单位 (HU) 值的阈值提高到-200, 以最大限度地滤除有效肺组织。对于分割模型VNet-IR-RPN, 其阈值被配置成最低以最大化分割候选区域, 尽管许多正常区域也可能被错误分割出来。后续, 我们会利用分

类模型进行过滤以去除健康区域。即使这样, 我们仍注意到COVID-19组的一个CT病例没有被分割出任何COVID-19或IAPV的候选区域, 因此被错误地归类为健康组, 如图6所示。这个病例的感染病灶太稀薄, 几乎无法被人眼察觉, 也无法被分割模型捕获。

4.3.3. 单个图像块的分类

从90个测试集CT样本中共获得1710张图像块, 其中包括357张COVID-19图像块、390张IAPV图像块和963张ITI图像块 (真值)。为了确定哪种方法最优, 使用混淆矩阵来评估每种方法的性能。共评估了两种网络结构, 即没有位置敏感机制的网络结构和有位置敏感机制的网络结构, 如表1和表2所示。

两种模型的平均 f_1 -score和总准确率分别为0.750/0.764和78.5%/79.4%。利用位置敏感机制可以提高COVID-19组和IAPV组的诊断准确率, 其中COVID-19组的诊断准确率提高了5.0% (260/273), 而IAPV组的诊断准确率提高了1.4% (276/280)。上述证据表明, 第二种具有位置敏感机制的模型取得了较好的效果。因此, 这一模型被用于本研究的其余部分。

此外, 由于本研究使用ITI组来去除干扰因素, 因

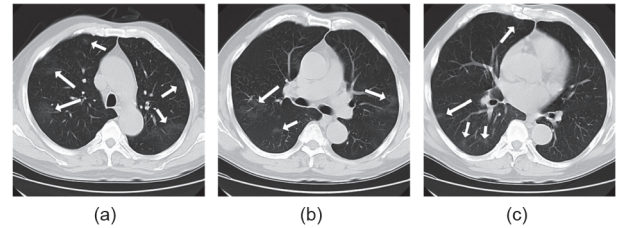


图6. 所有CT图像 (a~c) 都来自一个单独的CT病例。箭头所示区域为感染病灶。

表1 COVID-19、IAPV和ITI的两个分类模型的混淆矩阵

Group	COVID-19 (M_1/M_2)	IAPV (M_1/M_2)	ITI (M_1/M_2)
COVID-19 (M_1/M_2)	260/273	47/32	50/52
IAPV (M_1/M_2)	55/46	276/280	59/64
ITI (M_1/M_2)	75/77	81/82	807/804

M_1 : the ResNet model; M_2 : the ResNet model with the location-attention mechanism.

表2 COVID-19、IAPV和ITI的两个分类模型的召回率、正确率、 f_1 -score和准确率^a

Group	R	P	f_1 -score
COVID-19 (M_1/M_2)	0.728/0.765	0.667/0.689	0.696/0.725
IAPV (M_1/M_2)	0.708/0.718	0.683/0.711	0.695/0.714
ITI (M_1/M_2)	0.838/0.835	0.881/0.874	0.859/0.854

^aThe overall accuracy rates of M_1 and M_2 for the three groups are 78.5% and 79.4%; the accuracy rates of M_1 and M_2 for COVID-19 and IAPV groups are 71.8% and 74.0%.

此，其在最后一步将被忽略，且不被贝叶斯函数计算在内。为了与下一步保持一致，我们进一步比较了前两组的平均 f_1 -score和准确率，分别为0.720和74.0%。

4.3.4. 针对候选区域的投票

每个投票图像块代表整个候选区域。共识别出570个候选三维区域，其中包括119个COVID-19候选三维区域、130个IAVP候选三维区域和321个ITI的候选三维区域（真值）。投票结果的混淆矩阵及相应的召回率、正确率和 f_1 -score见表3和表4。

三个组的平均 f_1 -score和总准确率分别为0.856和89.3%，比上一步分别提高了12.0%和12.5%。前两组的平均 f_1 -score为0.806分、准确率为78.3%，分别提高了11.9%和5.8%。

4.3.5. CT样本整体分类结果

用Noisy-OR贝叶斯函数对优势感染类型进行识别。最终报告中输出三种结果：COVID-19感染、IAVP感染和健康人群。实验结果总结在表5和表6中。

仅计算COVID-19组和IAVP组的平均 f_1 -score和准确率（分别为0.843和85.0%），并将其与先前的结果进行比较，结果分别提高了4.6%和8.6%。

通过上述步骤，观察到平均 f_1 -score和准确率的持

表3 投票后COVID-19、IAVP和ITI组的混淆矩阵

Group	COVID-19	IAVP	ITI
COVID-19	97	15	7
IAVP	18	98	14
ITI	5	2	314

表4 投票后COVID-19、IAVP和ITI组的召回率、正确率、 f_1 -score和准确率^a

Group	R	P	f_1 -score
COVID-19	0.815	0.808	0.811
IAVP	0.754	0.852	0.800
ITI	0.978	0.937	0.957

^aThe overall accuracy rate for the three groups was 89.3%; and the average accuracy rate for COVID-19 and IAVP groups was 78.3%.

表5 COVID-19、IAVP和健康人群的Noisy-OR贝叶斯函数输出的混淆矩阵

Group	COVID-19	IAVP	Healthy
COVID-19	26	3	1
IAVP	4	25	1
Healthy	2	1	27

续改善。COVID-19和IAVP的分类准确率由74.0%（单幅图像块）提高到78.3%（图像候选区域），再提高到85.0%（根据主要感染类型分类的全部CT病例）。从三个分组（COVID-19组、IAVP组和健康人群组）的角度计算，总体分类准确率为86.7%。此外，模型还自动导出一系列带有突出显示的感染病灶的图像（图7）。

5. 讨论

COVID-19已经造成严重的公共卫生和安全问题，因此成为全球关注的焦点[33–35]。在COVID-19早期，一些患者已经有阳性的肺部影像学表现，但他们没有咳嗽或痰液和鼻咽拭子中的核酸检测结果为假阴性。这些患者没有被诊断为疑似或确诊病例。因为没有得到及时隔离或治疗，因此，这些患者成为潜在的感染源。而CT检查成本低、操作方便，通常医生会对每一例早期有发热、呼吸症状的患者进行常规CT检查。利用CT图像对疑似COVID-19患者进行筛查，可以提高COVID-19的早期检出率，缓解实验室核酸检测的压力。

根据先前的研究[29,30,36]，COVID-19患者的肺部

表6 COVID-19、IAVP和健康人群的Noisy-OR贝叶斯函数导出的召回率、正确率、 f_1 -score和准确率^a

Group	R	P	f_1 -score
COVID-19	0.867	0.813	0.839
IAVP	0.833	0.862	0.847
Healthy	0.900	0.931	0.915

^aThe overall accuracy rate for the three groups is 86.7% and the average accuracy rate for COVID-19 and IAVP groups was 85.0%.

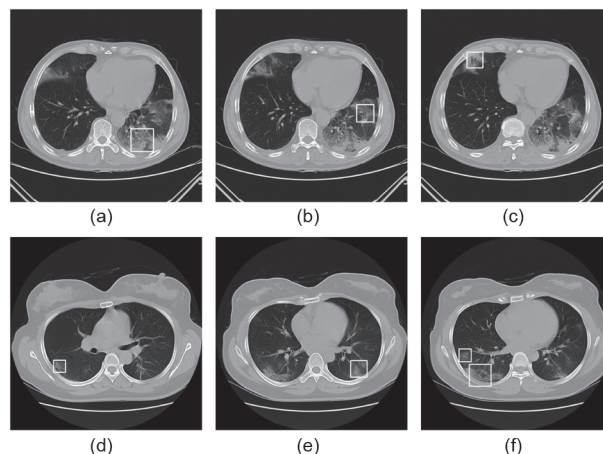


图7. 在原始图像上用边界框标注两个CT病例的病灶区。(a)、(b)和(c)图像来自同一例IAVP患者。其余三幅图像来自同一例COVID-19患者。肺部分割区域的原始图为一个三维立方体结构，为方便读取仅在中心位置的图片中标识出边界框。

CT有几种特征性的表现,如以胸膜为主要分布部位的局灶性磨玻璃影、伴有周围磨玻璃状“晕征”的实变影,以及不同大小的多发性实变影和网格状高密度影。有经验的放射科医师可以根据这些特征对COVID-19的可能性作出判断,但这种判断容易受到一些主观因素和个人熟练程度的影响。相比之下,基于深度学习系统的筛查模型通过对图像信息的数字化和标准化处理,显示出更具体、更可靠的结果,从而帮助医生做出更准确的临床决策。

目前临床上已有许多人工智能辅助诊断模型,而且部分模型已得到广泛应用,如肺结节诊断系统。这个系统是2019年6月由Ardila等[20]提出的一种深度学习算法,该算法使用患者当前和既往的肺部CT图像来预测患肺癌的风险。该模型对6716例患者进行测试,准确率接近94.4%,对另外1139例独立临床验证集病例进行测试的准确率也接近94.4%。在此后的临床使用中,人工智能辅助诊断模型除了具有较高的精度外,还显示出比人类诊断更快、更高效的特点。

本研究利用深度学习技术设计了一个分类网络,以区分COVID-19与IAPV。在网络结构方面,采用经典的ResNet进行特征提取,并将具有位置敏感机制和不具有位置敏感机制的网络模型进行比较。实验表明,具有位置敏感机制的网络模型能较好地地区分COVID-19病例。此外,我们还研究了多种增强方法,如图像块投票和Noisy-OR贝叶斯函数,以确定主要感染类型。所有这些都使平均 f_1 -score和准确率得到不断提高。

当然,这项研究有一些局限性。首先,COVID-19的临床表现与IAPV、社区获得性肺炎、机化性肺炎和嗜酸性粒细胞性肺炎等的临床表现相似。然而,因本研究样本数量和种类有限,所以只比较了COVID-19和IAPV的CT临床表现。今后应继续增加训练样本和测试样本的数量和种类,开展更多的多中心临床研究,以提高模型的准确性和应对复杂的临床情况。其次,COVID-19的临床诊断还需结合患者的接触史、旅行史、首发症状和实验室检查等资料来共同完成。此外,我们还应努力改进分割和分类模型,同时设计更好的排他模型进行训练,提高模型的分割和分类精度,并利用较大的数据集验证该算法的泛化性能。

6. 结论

在这个多中心的案例研究中,我们提出了一种利

用深度学习技术自动筛选COVID-19患者CT图像的新方法。具有位置敏感机制的网络模型可将COVID-19、IAPV和健康病例进行分类,总准确率为86.7%,有望成为一线临床医生有力的辅助诊断方法。

致谢

本研究获国家科技重大专项基金(2018ZX10101-001)支持。

Author contributions

Wei Wu, Tingbo Liang, Lanjuan Li and Xiaowei Xu initiated the project and provided clinical expertise and guidance on the study design. Xukun Li and Peng Du designed the network architecture and data/modeling infrastructure, training, testing setup, and statistical analysis. Xiaowei Xu, Wei Wu, Xukun Li, and Peng Du wrote the manuscript. Xiangao Jiang, Chunlian Ma, ShuangzhiLv, Liang Yu, Yanfei Chen, Junwei Su, Guanjing Lang, Yongtao Li, Hong Zhao, Kaijin Xu, and LingxiangRuan collected the datasets and interpreted the data. Xiaowei Xu, XianGao Jiang, Chun-Lian Ma, and Peng Du contributed equally to this article.

Compliance with ethics guidelines

Xiaowei Xu, Xiangao Jiang, Chunlian Ma, Peng Du, Xukun Li, ShuangzhiLv, Liang Yu, Qin Ni, Yanfei Chen, Junwei Su, Guanjing Lang, Yongtao Li, Hong Zhao, Jun Liu, Kaijin Xu, LingxiangRuan, Jifang Sheng, YunqingQiu, Wei Wu, Tingbo Liang, and Lanjuan Li declare that they have no conflict of interest or financial conflicts to disclose.

References

- [1] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020;382(8):727–33.
- [2] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020;382(13):1199–207.
- [3] Cohen J, Normile D. New SARS-like virus in China triggers alarm. *Science* 2020;367(6475):234–5.
- [4] Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 2020;25(3):23–30.
- [5] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients

- infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395(10223):497–506.
- [6] Chan JFW, Yuan S, Kok KH, To KKW, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020;395(10223):514–23.
- [7] National Health Commission of the People's Republic of China; National Administration of Traditional Chinese Medicine. Diagnosis and treatment protocol for novel coronavirus pneumonia (trial version 7) [Internet]. Beijing: National Health Commission of the People's Republic of China; [cited 2020 Mar 8]. Available from: http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294av_7dfe4cef80dc7f5912eb1989/files/ce3e6945832a438eae415350a8ce964.pdf. Chinese.
- [8] Loeffelholz MJ, Tang YW. Laboratory diagnosis of emerging human coronavirus infections—the state of the art. *Emerg Microbes Infect* 2020;9(1):747–56.
- [9] Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA*. Epub 2020 Mar 11.
- [10] Zhang W, Du R, Li B, Zheng X, Yang X, Hu B, et al. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. *Emerg Microbes Infect* 2020;9(1):386–9.
- [11] Long Q, Deng H, Chen J, Hu J, Liu B, Liao P, et al. Antibody responses to SARS-CoV-2 in COVID-19 patients: the perspective application of serological tests in clinical practice [preprint]. medRxiv. Posted 2020 Mar 20.
- [12] National Health Commission of the People's Republic of China; National Administration of Traditional Chinese Medicine. Diagnosis and treatment protocol for novel coronavirus pneumonia (trial version 5) [Internet]. Beijing: National Health Commission of the People's Republic of China; [cited Feb 5, 2020]. Available from: <http://www.nhc.gov.cn/yzygj/s7653p/202002/d4b895337e19445f8d728fcaf1e3e13a/files/ab6bec7f93e64e7f998d802991203cd6.pdf>. Chinese.
- [13] Liu X, Guo S, Yang B, Ma S, Zhang H, Li J, et al. Automatic organ segmentation for CT scans based on super-pixel and convolutional neural networks. *J Digit Imaging* 2018;31(5):748–60.
- [14] Gharbi M, Chen J, Barron JT, Hasinoff SW, Durand F. Deep bilateral learning for real-time image enhancement. *ACM Trans Graph* 2017;36(4):118.
- [15] Hesamian MH, Jia W, He X, Kennedy P. Deep learning techniques for medical image segmentation: achievements and challenges. *J Digit Imaging* 2019;32(4):582–96.
- [16] Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Zhou J, et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur Radiol* 2019;29(11):6163–71.
- [17] Nardelli P, Jimenez-Carretero D, Bermejo-Pelaez D, Washko GR, Rahaghi FN, Ledesma-Carbayo MJ, et al. Pulmonary artery-vein classification in CT images using deep learning. *IEEE Trans Med Imaging* 2018;37(11):2428–40.
- [18] Zhu W, Huang Y, Zeng L, Chen X, Liu Y, Qian Z, et al. AnatomyNet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy. *Med Phys* 2019;46(2):576–89.
- [19] Huang P, Park S, Yan R, Lee J, Chu LC, Lin CT, et al. Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study. *Radiology* 2018;286(1):286–95.
- [20] Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019;25(6):954–61.
- [21] Esteva A, Kuprel B, Novoa R A, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542(7639):115–8.
- [22] Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* 2017;284(2):574–82.
- [23] Wu W, Li X, Du P, Lang G, Xu M, Xu K, et al. A deep learning system that generates quantitative CT reports for diagnosing pulmonary tuberculosis. 2019. arXiv:1910.02285v1.
- [24] Li L, Huang H, Jin X. AE-CNN classification of pulmonary tuberculosis based on CT images. In: Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME); 2018 Oct 19–21; Zhejiang, China. New York: IEEE; 2018.
- [25] Oniško A, Druzdzel M, Wasyluk H. Learning Bayesian network parameters from small data sets: application of Noisy-OR gates. *Int J Approx Reason* 2001;27(2):165–82.
- [26] Milletari F, Navab N, Ahmadi SA. V-Net: fully convolutional neural networks for volumetric medical image segmentation. 2016. arXiv:1606.04797v1.
- [27] Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, inception-resnet and the impact of residual connections on learning. 2016. arXiv:1602.07261.
- [28] Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D u-net: learning dense volumetric segmentation from sparse annotation. 2016. arXiv:1606.06650.
- [29] Kanne JP. Chest CT findings in 2019 novel coronavirus (2019-nCoV) infections from Wuhan, China: key points for the radiologist. *Radiology* 2020;295(1):16–7.
- [30] Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology* 2020;295(1):202–7.
- [31] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2015. arXiv:1512.03385.
- [32] Breiman L. Bagging predictors. *Mach Learn* 1996;24(2):123–40.
- [33] Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet* 2020;395(10223):470–3.
- [34] Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First case of 2019 novel coronavirus in the United States. *N Engl J Med* 2020;382(10):929–36.
- [35] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020;395(10223):507–13.
- [36] Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect Dis* 2020;20(4):425–34.