针对新型冠状病毒肺炎 X 射线图像识别的 DD-CovidNet 模型

王威, 胡亿洋, 王新*, 李骥, 李宇涛

(长沙理工大学计算机与通信工程学院 长沙 410114)
(wangxin@csust.edu.cn)

摘要: 受医疗资源紧张和医疗水平较低等因素的影响, 新型冠状病毒肺炎(coronavirus disease 2019, COVID-19)尚未得到有效控制. 利用深度学习方法在胸部 X 射线(chest X-ray, CXR)图像中检测感染者是一种安全有效的途径. 针对上述问题, 提出一种自动识别 COVID-19 的 CXR 图像的智能方法. 根据 CXR 图像的特点, 提出了对特征信息敏感的双路径多尺度特征融合(dual-path multi-scale fusion, DMFF)模块和密集空洞深度可分离卷积(dense dilated depthwise separable, D3S)模块, 分别提取浅层特征与深层特征. 在此基础上, 设计了高效的轻量级卷积神经网络——DD-CovidNet. DMFF 模块通过融合多尺度特征感知更丰富的浅层特征, D3S 模块通过强化特征传递与增大感受野提取更有效的类别区分特征. 在 2 个数据集上进行了实验验证, 结果表明, DD-CovidNet 模型对 COVID-19 识别的灵敏度为 96.08%, 精度与特异性均为 100.00%, 且具有较少的参数量与较快的分类速度. 与其他模型相比, DD-CovidNet 模型的检测速度更快, 检测结果更准确.

关键词: 新型冠状病毒肺炎; 胸部 X 射线图像; 卷积神经网络; 图像分类; DD-CovidNet
中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2021.18791

DD-CovidNet Model for X-Ray Images Recognition of Coronavirus Disease 2019

Wang Wei, Hu Yiyang, Wang Xin*, Li Ji, and Li Yutao

(School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114)

Abstract: Affected by the shortage of medical resources and low level of medical care, coronavirus disease 2019(COVID-19) has not yet been contained. It is a safe and effective way to detect infection in chest X-ray (CXR) images by deep learning. To solve the above problems, an intelligent method for automatic recognition of COVID-19 in CXR images is proposed. According to the characteristics of CXR images, a dual-path multi-scale feature fusion (DMFF) module and dense dilated depthwise separable (D3S) module are proposed to extract the shallow and deep features respectively. On this basis, an efficient and lightweight convolutional neural network—DD-CovidNet, is designed. DMFF module can sense more abundant spatial information by fusing multi-scale features. D3S module can extract more effective classification information by enhancing feature transfer and enlarging receptive field. The method is validated on two data sets. The experimental results show that the sensitivity of DD-CovidNet model for COVID-19 recognition is 96.08%, the precision and specificity are 100.00%, and it has less parameters and faster classification speed. Compared with other models, DD-CovidNet model has faster detection speed and more accurate detection results.

收稿日期: 2020-11-22; 修回日期: 2021-03-29. 基金项目: 国防科技创新特区项目(2019XXX00701); 湖南省自然科学基金(2019JJ80105); 长沙市科技计划(kq2004071); 湖南省研究生科研创新项目(CX20200882); 湖南省教育厅科学研究项目(20C1249). 王威(1974—), 男, 博士, 教授, 硕士生导师, CCF 会员, 主要研究方向为计算机视觉、模式识别; 胡亿洋(1997—), 女, 硕士研究生, 主要研究方向为计算机视觉、模式识别; 王新(1976—), 女, 硕士, 讲师, 论文通讯作者, 主要研究方向为计算机视觉、模式识别; 李骥(1981—), 男, 博士, 讲师, 硕士生导师, 主要研究方向为计算机视觉、模式识别; 李宇涛(1997—), 男, 硕士研究生, 主要研究方向为计算机视觉、模式识别.

Key words: coronavirus disease 2019 (COVID-19); chest X-ray (CXR) image; convolutional neural network (CNN); image classification; DD-CovidNet

由 2019 新型冠状病毒引起的 COVID-19 在极短的时间内遍布全球. 截至北京时间 2020 年 10 月 14 日 19 时, 全球累计确诊病例数达 38 432 125, 死亡病例数达 1 092 031, 治愈病例数为 28 609 848, 死亡率约为 3.00%.

COVID-19 检测方法包括逆转录聚合酶链反应(reverse transcription-polymerase chain reaction, RT-PCR)检测和肺部医学影像检测等. 基于 RT-PCR 的检测技术是 COVID-19 爆发初期最为广泛使用的诊断手段, 但其试剂盒有限且价格高, 诊断耗时长^[1]. RT-PCR 敏感性较低导致部分诊断结果呈假阴性, 使诊断存在一定的局限性. 肺部医学影像检测中, 临床广泛使用的医学成像主要有 X 射线(X-ray, XR)、计算机断层扫描(computed tomography, CT)以及核磁共振成像等. 目前针对新冠肺炎的辅助诊疗手段主要包括胸部 X 射线(chest X-ray, CXR)图像和 CT^[2]. 利用 CXR 图像筛查早期 COVID-19 患者引起了医学领域的极大关注. 与 CT 相比, CXR 图像易于发现病变的结构与组织, 其拍摄费用更低, 拍摄装置更为普及, 且辐射更低.

人工阅片是一项费时且易出错的劳动密集型工作. 人工智能技术可以通过准确描绘 XR 和 CT 图像中的感染情况提高工作效率, 便于后续量化. 因此, 计算机辅助诊断平台可以帮助医疗工作者做出临床决策, 进行疾病诊断、跟踪和预后^[3]. 辅助检测新冠肺炎的智能手段能够极大地减轻医护人员和资源供给的压力, 降低医疗人员的感染风险. 近年来, 深度学习方法在计算机视觉领域的应用取得了不错的成果, 诸多研究^[4-6]表明, 卷积神经网络(convolutional neural networks, CNN)具有优异的特征提取能力. 例如, Wang 等^[4]在 MobileNet 模型浅层引入空洞卷积, 增大局部感受野, 从而提高模型的分类准确率. 韩丽等^[5]提出一种基于结构感知深度学习的三维形状分类方法, 其具有较高的分类精度. 高翔等^[6]提出一种基于跨级注意力机制和多标签分类的高精度实时语义分割算法, 其在复杂场景分割中具有较好的鲁棒性. 深度学习技术在医疗保健, 尤其是医学影像方面的应用日益成熟. Roth 等^[7]把深度 CNN 应用于自动化计算机辅助检测, 提升了检测 CT 成像中的硬化性转移灶、淋巴结和结肠息肉的灵敏度. Yang 等^[8]提出一种快速变形的图像配准方法, 其使用深度编码

器-解码器网络作为预测模型, 取得了不错的预测精度.

近期, 很多研究者致力于如何利用深度学习方法从 CXR 图像中检测出 COVID-19. 由于开源 COVID-19 的 CXR 图像样本较少, 大多数论文中自动诊断 COVID-19 常用测试协议可能会产生偏差, 预测结果会更依赖源数据集的特征^[9]. Tartaglione 等^[10]应用深度学习技术识别 COVID-19 的 CXR 图像, 并提供对使用当前可用数据集获得的大量统计结果的批判性阅读. Apostolopoulos 等^[11]使用基于迁移学习方法的卷积神经网络自动检测 X 射线影像. Wang 等^[12]设计了针对从 CXR 图像中检测出 COVID-19 病例的深度卷积神经网络. Khan 等^[13]提出了基于 Xception 的深度卷积神经网络模型——CoroNet, 并使用基于 ImageNet 数据集的预训练方法识别 COVID-19 的 CXR 图像. Ozturk 等^[14]提出一种基于 DarkNet 层的 DarkCovidNet 模型, 自动检测 COVID-19 的 CXR 图像, 该模型具有较少的参数量.

1 CNN

近年来, CNN 模型被广泛用于计算机视觉领域, 如图像分类、目标检测和语义分割等, 其基本架构如图 1 所示. Krizhevsky 等^[15]设计的 AlexNet 模型获得了 2012 年 ImageNet 竞赛冠军, 应用局部响应归一化(local response normalization, LRN), Dropout 和 ReLU 等操作成功地解决了较深网络中梯度弥散问题. Simonyan 等^[16]通过重复交替使用小卷积核层和最大池化层, 构建了 16~19 层的牛津大学计算机视觉几何群(visual geometry group, VGG)网络模型, 其性能明显提升, 但计算开销十分庞大.

早期的 CNN 一般通过增加网络的深度或宽度提升性能, 但会导致过度拟合、网络退化和梯度消失等问题. Szegedy 等^[17]提出的 GoogLeNet 模型通过 Inception 模块提高参数利用率, 增强模型表征能力. He 等^[18]提出了易于优化的 ResNet 模型, 通过残差块中的跳跃连接解决网络退化问题. Huang 等^[19]提出基于密集连接的 DenseNet 模型, 通过特征重用减少了参数量与计算量, 同时能有效地缓解梯度消失的问题.

图 1 CNN 基本架构

随着移动设备的普及,越来越多的研究致力于神经网络模型的轻量化。Iandola 等^[20]提出基于 Fire 模块的 SqueezeNet 模型,模型尺寸较小,识别准确率达到 AlexNet 的水平,但并行计算能力较差。Howard 等^[21]针对 CNN 如何应用于移动设备的问题,提出了基于深度可分离卷积的流线型模型 MobileNet。Zhang 等^[22]提出的 ShuffleNet 模型,通过通道混洗与群卷积实现高效计算,具有不错的模型性能。

2 模型结构设计

2.1 DMFF 模块与 D3S 模块

CXR 图像具有类间特征相似度高且类内特征变异性低的特点,会造成模型偏差或过拟合^[23]。针对此问题,本文提出 2 种表征能力强且计算高效的模块——双路径多尺度特征融合(dual-path multi-scale fusion, DMFF)模块和密集空洞深度可分离卷积(dense dilated depthwise separable, D3S)模块。如图 2 和图 3 所示, h 为特征图的高度; w 为特征图

的宽度; c 为特征图的通道数; f 为卷积核的个数; k 为卷积核的尺寸; s 为步长。与传统卷积层中常用的激活函数 ReLU 相比,激活函数 h-swish 能降低延迟,适用于移动设备^[24]。

靠近输入的卷积层提取的特征包含纹理和颜色等信息,因此在网络浅层使用强表征能力的 DMFF 模块。DMFF 模块包含 2 条支路:一条支路使用空洞深度卷积层,即膨胀率为 2 的空洞卷积核替换普通卷积核的深度卷积层;另一条支路使用深度卷积层,使 2 条支路获得的感受野不同,从而实现多尺度特征融合,增强模型的表征能力。膨胀率为 2 的空洞卷积不会影响复杂度,因此 2 条支路的参数量与计算量相同。与 Inception 等模块相比,DMFF 模块使模型更适于并行计算。每条支路包含 2 个点卷积层:第 1 个点卷积层用于增加特征维度,丰富特征信息;第 2 个点卷积层关联深度卷积层或空洞深度卷积层的输出特征图的通道信息。本文将 DMFF 模块中的深度卷积层或空洞深度卷积层的步长设置为 2,实现下采样。与池化层相比,包含激活函数的卷积层能增强模型非线性。

图 2 DMFF 模块结构

图 3 D3S 模块结构

由于靠近输出的特征图语义信息丰富,因此在网络深层使用 D3S 模块. D3S 模块是基于密集连接的空洞深度可分离卷积层,将空洞深度卷积层的输入特征图与点卷积层输出特征图进行通道维度层面的拼接作为下一层的输入,强化特征传递.与普通卷积相比,空洞深度可分离卷积拥有更少的参数量与计算量以及更大的局部感受野,使模型有更低的复杂度与更高的性能.同时, D3S 模块通过密集连接与特征重用能有效地缓解梯度消失,增强特征并减少参数量.

2.2 DD-CovidNet 模型架构

基于 DMFF 模块和 D3S 模块,本文提出一种

新型轻量卷积神经网络模型——DD-CovidNet,并在此基础上提出一种自动检测新冠肺炎的方法,其流程如图 4 所示.

预处理后的 CXR 图像经过一层膨胀率为 2 的 3×3 空洞卷积,增加感受野;再利用 5 个 DMFF 模块对特征图的空间尺寸进行 5 次减半,利用 9 个 D3S 模块缓解梯度消失;经过 PyTorch 提供的自适应平均池化层,特征图的空间尺寸变为 1×1 ,并进行丢弃率为 0.2 的 Dropout 操作,提升泛化能力;经过一层点卷积层,增加特征维度;经过一层全连接层,弱化特征坐标信息对分类的影响;最后利用分类器输出类别信息.

图 4 检测方法流程图

2.3 模型复杂度与效率分析

引入膨胀率为 2 的空洞卷积不会改变参数量和计算量.同时,深度可分离卷积层比普通卷积层有更少的参数量与计算量,能降低模型复杂度.相比串行结构,并行结构能减小模型尺寸,降低模型时间复杂度.本文使用计算量和参数量衡量模型的复杂度.衡量模型高效性的直接指标是模型推理速度.模型计算量、内存访问成本、element-wise 操作等影响模型推理速度,其中内存访问成本对模型速度影响较大^[25].假设输入特征图的大小为 $h_1 \times w_1 \times c_1$,卷积核大小为 $c_1 \times k^2 \times c_0$,输出特征图的大小 $h_0 \times w_0 \times c_0$,标准卷积产生的参数量 P_{STD} 、计算量 F_{STD} 和内存访问成本 M_{STD} 分别为

$$P_{STD} = c_1 \times k^2 \times c_0 \quad (1)$$

$$F_{STD} = h_0 \times w_0 \times c_1 \times k^2 \times c_0 \quad (2)$$

$$M_{STD} = h_1 \times w_1 \times c_1 + h_0 \times w_0 \times c_0 + c_1 \times k^2 \times c_0 \quad (3)$$

深度卷积或膨胀率为 2 的空洞深度卷积产生的参数量 P_{DW} 、计算量 F_{DW} 和内存访问成本 M_{DW} 分别为

$$P_{DW} = c_1 \times k^2 \quad (4)$$

$$F_{DW} = h_0 \times w_0 \times c_1 \times k^2 \quad (5)$$

$$M_{DW} = h_1 \times w_1 \times c_1 + h_0 \times w_0 \times c_0 + c_1 \times k^2 \quad (6)$$

给定输入特征图 $h \times w \times c$ 与输出特征图

$\frac{h}{2} \times \frac{w}{2} \times (c+16)$, 3×3 标准卷积产生的参数量

P_{conv3} 、计算量 F_{conv3} 和内存访问成本 M_{conv3} 分别为

$$P_{conv3} = 9 \times c \times (c+16) \quad (7)$$

$$F_{conv3} = 9 \times \frac{h \times w}{4} \times c \times (c+16) \quad (8)$$

$$M_{conv3} = 9 \times c^2 + c \times \left(\frac{5}{4} \times h \times w + 144 \right) + 4 \times h \times w \quad (9)$$

DMFF 模块产生的参数量 P_{DMFF} 、计算量

F_{DMFF} 和内存访问成本 M_{DMFF} 分别为

$$P_{\text{DMFF}} = \frac{c^2}{2} + 17 \times c + 144 \quad (10)$$

$$F_{\text{DMFF}} = h \times w \times \left(\frac{5}{8} \times c^2 + \frac{57}{4} \times c + 68 \right) \quad (11)$$

$$M_{\text{DMFF}} = c^2 + \left(\frac{15}{4} \times h \times w + 33 \right) \times c + 4 \times h \times w + 272 \quad (12)$$

相对于 3×3 标准卷积, DMFF 模块的参数减少量 Δp_{DMFF} (即 P_{conv3} 与 P_{DMFF} 的差值)、计算减少量 ΔF_{DMFF} (即 F_{conv3} 与 F_{DMFF} 的差值) 和内存访问成本减少量 ΔM_{DMFF} (即 M_{conv3} 与 M_{DMFF} 的差值) 分别为

$$\Delta p_{\text{DMFF}} = \frac{17}{2} \times c^2 + 127 \times c - 144 \quad (13)$$

$$\Delta F_{\text{DMFF}} = h \times w \times \left(\frac{13}{8} \times c^2 + \frac{87}{4} \times c - 68 \right) \quad (14)$$

$$\Delta M_{\text{DMFF}} = 8 \times c^2 + c \times \left(11 - \frac{5}{2} \times h \times w \right) - 40 \times h \times w - 272 \quad (15)$$

给定输入特征图 $h \times w \times c$ 与输出特征图 $h \times w \times (c+16)$, 3×3 标准卷积产生的参数量 P_{conv3} 、计算量 F_{conv3} 和内存访问成本 M_{conv3} 分别为

$$P_{\text{conv3}} = 9 \times c \times (c+16) \quad (16)$$

$$F_{\text{conv3}} = 9 \times h \times w \times c \times (c+16) \quad (17)$$

$$M_{\text{conv3}} = 9 \times c^2 + (2 \times h \times w + 144) \times c + 16 \times h \times w \quad (18)$$

D3S 模块产生的参数量 P_{D3S} 、计算量 F_{D3S} 和内存访问成本 M_{D3S} 分别为

$$P_{\text{D3S}} = 9 \times c + 16 \times c \quad (19)$$

$$F_{\text{D3S}} = 9 \times h \times w \times c + 16 \times h \times w \times c \quad (20)$$

$$M_{\text{D3S}} = (3 \times h \times w + 25) \times c + 16 \times h \times w \quad (21)$$

相对于 3×3 标准卷积, D3S 模块的参数减少量 Δp_{D3S} (即 P_{conv3} 与 P_{D3S} 的差值)、计算减少量 ΔF_{D3S} (即 F_{conv3} 与 F_{D3S} 的差值) 和内存访问成本减少量 ΔM_{D3S} (即 M_{conv3} 与 M_{D3S} 的差值) 分别为

$$\Delta p_{\text{D3S}} = 9 \times c^2 + 119 \times c \quad (22)$$

$$\Delta F_{\text{D3S}} = 9 \times h \times w \times c^2 + 119 \times h \times w \times c \quad (23)$$

$$\Delta M_{\text{D3S}} = 9 \times c^2 + (119 - h \times w) \times c \quad (24)$$

在 DD-CovidNet 中, $c \geq 16$, 因此 Δp_{DMFF} , ΔF_{DMFF} , Δp_{D3S} 和 ΔF_{D3S} 都大于 0, 由此可得, DMFF 模块与 D3S 模块能降低模型复杂度. 一些

深度学习模型与 DD-CovidNet 的参数量与计算量的比较如表 1 所示. 由表 1 可知, 复杂度最低的 ShuffleNetV2^[25] 的参数量和计算量仍高于 DD-CovidNet, 分别是其的 160% 与 610%. 当 $c \geq 80$, $h \leq 14$, $w \leq 14$ 时, DMFF 模块的内存访问成本减少量 $\Delta M_{\text{DMFF}} > 0$. 采用 D3S 模块时, 输入图像经过下采样, 空间尺寸减小, 因此 $\Delta M_{\text{D3S}} > 0$. 可知, 合理应用 DMFF 模块和 D3S 模块能减少模型推理所需时间, 提升模型性能与效率.

表 1 不同模型参数量与计算量

模型	$10^{-6} \times$ 参数量	$10^{-6} \times$ 计算量
VGG19 ^[16]	18 736.81	137.04
GoogLeNet ^[17]	1 434.21	5.32
ResNet50 ^[18]	3 919.13	22.42
DenseNet121 ^[19]	2 731.91	6.62
SqueezeNet_1.0 ^[20]	702.71	0.73
MobileNet ^[21]	560.73	3.11
ShuffleNet ^[22]	142.02	0.91
MobileNetV2 ^[25]	311.13	2.13
ShuffleNetV2 ^[26]	144.72	1.22
DD-CovidNet	90.35	0.20

3 实验结果分析

3.1 数据集

实验使用的图像来自 2 个开源数据集. 普通肺炎与正常的 CXR 图像选自 Kaggle 数据集^[27], 其共有 5 863 幅图像, 本文从中选取 4 265 幅普通肺炎图像与 1 575 幅正常图像. 实验所用 COVID-19 的 CXR 图像选自 Cohen 等^[28]收集的肺炎数据集, 其包含 790 幅感染 COVID-19 或普通肺炎患者的 CXR 图像和 CT 图像. 本文从中选取 412 幅感染了 COVID-19 的 CXR 图像. 因此实验数据集共有 6 252 幅 CXR 图像, 实验中随机选取 5 526 幅图像作为训练集, 其中 1 341 幅正常图像, 3 875 幅普通肺炎图像, 310 幅 COVID-19 阳性图像. 剩余 726 幅图像为测试集, 其中 234 幅正常图像, 390 幅普通肺炎图像, 102 幅 COVID-19 阳性图像. 数据集部分示例如图 5 所示. 可以看出, CXR 图像具有类内变异性低且类间相似度高的特点.

参考大多数文献中评估医学图像分类深度学习模型的标准, 实验使用准确率 A 、精度 P 、灵敏度 S_1 、特异性 S_2 和 F_1 分数作为评估模型的指标, 公式为

图 5 实验数据集图像示例

$$A = \frac{T_p + T_N}{T_p + T_N + F_p + F_N} \quad (25)$$

$$P = \frac{T_p}{T_p + F_p} \quad (26)$$

$$S_1 = \frac{T_p}{T_p + F_N} \quad (27)$$

$$S_2 = \frac{T_N}{T_N + F_p} \quad (28)$$

$$F_1 = \frac{T_p}{2 \times T_p + F_p + F_N} \quad (29)$$

其中, T_p 表示该类被预测正确的样本数; F_N 表示该类被误判的样本数; T_N 表示其他类被预测正确的样本数; F_p 表示其他类被误判的样本数。

3.2 预处理以及参数设置

本文把实验图像处理成固定分辨率 224×224 像素, 再对训练集进行数据增强: 以 50% 的概率对图像进行角度为 60° 的水平翻转与 45° 的垂直翻转, 对图像的亮度值、对比度值和饱和度值进行 60%~140% 的均匀变换处理。通过数据增强, 实际用于训练的图像数是训练集的 5 倍。

实验模型卷积层的权重参数使用 Kaiming 正态分布初始化。归一化层中的权值参数初始为 1, 偏置参数初始为 0。全连接层中的权值参数用均值为 0、偏差为 0.01 的正态分布初始化。实验在同一配置环境下进行, 如表 2 所示。

实验模型的初始学习率均为 0.01, 进程数量为 8。每组实验训练 100 个 epoch, 损失函数为交叉熵损失函数。在不同周期分别使用 Adam 优化器与 SGD 优化器。前 20 个 epoch 使用 betas 介于 0.9 和 0.999 之间的 Adam 优化器使模型快速收敛。后 80 个 epoch 使用动量为 0.9、权重衰减为 0.000 5 的 SGD 优化器寻找模型最优解。在整个训练过程中, 使用等间隔调整学习率的方法, 即每 10 个 epoch,

表 2 实验环境平台

属性	配置信息
操作系统	Ubuntu 16.04.5 LTS
CPU	Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30 GHz
GPU	GeForce GTX TITAN X
CUDNN	CUDNN 6.0.21
CUDA	CUDA 8.0.61
框架	Pytorch
开发环境	PyCharm
语言	Python

学习率衰减至原来的 0.9。训练集与测试集的批大小分别为 32 和 16。

3.3 模型有效性验证与结果分析

实验首先验证以下 2 个结论: (1) 与普通卷积相比, DMFF 模块和 D3S 模块具有更加优秀的特征提取能力, 有利于提升模型性能。(2) DD-CovidNet 靠近输入端使用 DMFF 模块, 靠近输出端使用 D3S 模块, 有利于增加模型分类精度。使用 5 种网络模型进行对比实验: 使用普通卷积替换 DD-CovidNet 中的 DMFF 模块和 D3S 模块得到 F-ConvNet, 替换 DD-CovidNet 中的 D3S 模块得到 DMFF-ConvNet, 替换 DD-CovidNet 中的 DMFF 模块得到 Conv-D3SNet, 全部使用 DMFF 模块得到 F-DMFFNet, 全部使用 D3S 模块得到 F-D3SNet。

图 6 所示为消融实验的准确率曲线。表 3 所示为每隔 20 个 epoch 记录模型在测试集上的准确率。由图 6 和表 3 可知, 与 F-ConvNet 相比, F-DMFFNet 的分类准确率提高了近 2%。这说明 DMFF 模块特征提取能力优于普通卷积, 利于增强模型分类能力。与 F-DMFFNet 相比, DD-CovidNet 的识别准确率增加了近 3%, 说明靠近输出端使用 D3S 模块能进一步提升模型性能。与 F-ConvNet 相比, F-D3S 的分类准确率提高了近 2%, 说明 D3S 模块的表征能力优于普通卷积, 有利于提升模型识别能力。与 F-D3S 相比, DD-CovidNet 的识别准确率增加了

图 6 不同模块消融实验模型准确率曲线图

表 3 不同模型分类准确率对比 %

模型	epoch				
	20	40	60	80	100
F-ConvNet	68.66	78.82	83.15	87.08	88.32
DMFF-ConvNet	70.33	79.45	84.43	90.16	90.14
Conv-D3SNet	72.75	79.91	85.14	90.21	90.18
F-DMFFNet	85.87	86.23	88.82	90.72	91.30
F-D3SNet	84.92	86.14	87.97	90.35	90.27
DD-CovidNet	86.87	89.13	92.86	94.57	94.52

注. 粗体表示最佳结果.

4%, 说明靠近输入端使用 DMFF 模块能进一步提高模型分类精度.

为了证明 DMFF 模块与 D3S 模块对特征信息更敏感, 将 DD-CovidNet 中第 1 个 DMFF 模块、D3S 模块和 F-ConvNet 中对应位置的普通卷积生成的特征图进行可视化, 如图 7 和图 8 所示. 图 7 显示与普通卷积相比, DMFF 模块提取的纹理等细节特征更为丰富. 图 8 显示, D3S 模块提取的含类别区分信息的抽象特征更为饱满.

为了模型轻量化, 本文在网络浅层进行下采样, 但该操作会引起特征信息的丢失, 造成模型分类精度的损失. 因此, 实现下采样操作的模块是否具有优异的特征提取能力尤为关键. 深层特征十分抽象但包含重要的类别识别信息, 因此, 需要强化特征提取与传递的能力. 为了验证 DMFF 模块的并行连接与 D3S 模块的密集连接的合理性, 在实验对比中, 保留 DMFF 模块支路 1 得到 DMFF1-D-CovidNet 模型, 保留支路 2 得到 DMFF2-D-CovidNet 模型, 进行 DMFF 模块的消融实验. 同时, 去掉 D3S 模块的密集连接得到 noDense-Net 模型, 进行密集连接的消融实验. 实验结果如图 9 所示, 从中可以看出, DMFF1-D-CovidNet 模型、DMFF2-D-CovidNet 模型与 noDense-Net 模型的分类型准确率均低于 DD-CovidNet 模型. 可知, DMFF 模块的并行

a. 第 1 个 DMFF 模块

b. 对应位置普通卷积

图 7 DMFF 模块与对应位置普通卷积提取的特征图

a. 第 1 个 D3S 模块

b. 对应位置普通卷积

图 8 D3S 模块与对应位置普通卷积提取的特征图

图 9 不同支路消融实验模型准确率曲线图

连接与 D3S 模块的密集连接可以提升模型识别准确率.

为了进一步体现 DD-CovidNet 模型的有效性, 本文使用 VGG19^[16], GoogLeNet^[17], ResNet50^[18]等 CNN 模型, 以及 SqueezeNet_1.0^[20], MobileNet^[21], ShuffleNet^[22]等轻量级 CNN 模型进行对比实验, 结果如表 4 所示.

表 4 不同模型各项实验指标结果对比

模型	准确率/%	精度/%	灵敏度/%	特异性/%	F_1 值/%	速度/(幅·s ⁻¹)
VGG19 ^[16]	93.11	96.09	92.93	96.47	93.02	909
GoogLeNet ^[17]	92.56	95.29	91.56	95.78	92.06	1 149
ResNet50 ^[18]	93.53	96.01	93.15	96.53	93.34	980
DenseNet121 ^[19]	93.11	95.98	92.75	96.38	92.92	1 040
SqueezeNet_1.0 ^[20]	67.91	45.83	50.51	64.16	57.93	1 388
MobileNet ^[21]	88.53	90.14	87.25	91.84	87.89	1 538
ShuffleNet ^[22]	87.02	90.08	86.17	92.31	86.59	1 724
MobileNetV2 ^[25]	89.26	91.89	88.51	93.16	88.89	2 000
ShuffleNetV2 ^[26]	92.01	91.92	91.74	96.29	91.87	2 083
DD-CovidNet	94.57	96.52	93.68	96.16	94.90	2 128

注. 粗体表示最佳结果.

由表 4 可知, DD-CovidNet 的分类准确率、精度、灵敏度和 F_1 值均高于其他模型, 分别为 94.57%, 96.52%, 93.68% 与 94.90%。DD-CovidNet 的模型推理速度最快, 每秒能识别 2 128 幅图像。在对比的传统 CNN 模型中, ResNet50^[18] 分类准确率最高, 为 93.53%, 但仍比 DD-CovidNet 低了 1.04%, 且其参数量与计算量分别为 DD-CovidNet 的 4338% 和 11210%。在对比的轻量级 CNN 模型中, ShuffleNetV2^[26] 的分类准确率最高, 为 92.01%, 但比 DD-CovidNet 低了 2.56%, 其参数量和计算量分别是 DD-CovidNet 的 160% 和 610%。不同于 GoogLeNet^[17] 中的 Inception 结构与 ShuffleNet^[22] 的 ShuffleNet 单元, DD-CovidNet 浅层网络每条支路的复杂度一样, 并使用空洞深度卷积增加其中一条支路的局部感受野, 实现多尺度特征融合, 增强模型表征能力, 提升模型性能。不同于 DenseNet121^[19] 在整个网络使用密集连接, DD-CovidNet 只在网络深层使用密集深度可分离连接, 提升特征图利用率。

表 5 所示为 DD-CovidNet 在各类上的部分性能指标的具体表现, 图 10 所示为 DD-CovidNet 在测试集上的混淆矩阵。由实验结果可知, 在扩充的数据集上, DD-CovidNet 识别的平均精度、灵敏度和特异性均高于 90%, 尤其是 DD-CovidNet 识别 COVID-19 的精度、灵敏度和特异性分别为 100.00%, 96.08% 和 100.00%, 远高于 COVID-19 的 CXR 图像基线灵敏度 69%^[29]。因此, 对特征信息

表 5 DD-CovidNet 模型的精度、灵敏度和特异性 %

类别	精度	灵敏度	特异性
COVID-19	100.00	96.08	100.00
正常	98.04	85.47	99.19
普通肺炎	91.51	99.49	89.29
平均值	96.52	93.68	96.16

真实标签	COVID-19	98	2	2
	正常	0	200	34
	普通肺炎	0	2	388
		COVID-19	正常	普通肺炎
		预测标签		

图 10 DD-CovidNet 的混淆矩阵

敏感的 DD-CovidNet 能有效地识别 COVID-19 的 CXR 图像。

此外, 一些 CNN 模型改进的自动检测 COVID-19 的 CXR 图像的深度学习方^[11-14]与 DD-CovidNet 模型的实验结果对比, 如表 6 所示。

表 6 其他识别 COVID-19 的 CXR 图像方法与 DD-CovidNet 方法结果对比

方法	样本		模型	准确率/%	10 ⁻⁶ × 参数量
	数量	类别			
Apostolopoulos 等 ^[11]	224	COVID-19	Xception	92.85	33.00
	700	普通肺炎			
	504	正常			
Wang 等 ^[12]	358	COVID-19	Covid-Net	93.30	11.75
	5 538	普通肺炎			
	8 066	正常			
Khan 等 ^[13]	284	COVID-19	CoroNet	94.59	33.00
	657	普通肺炎			
	310	正常			
Ozturk 等 ^[14]	125	COVID-19	DarkCovidNet	87.02	1.10
	500	普通肺炎			
	500	正常			
本文	412	COVID-19	DD-CovidNet	94.57	0.20
	4 265	普通肺炎			
	1 575	正常			

由表 6 可知, 在 4 个对比模型中, DarkCovidNet^[14] 参数量最少, 但仍为 DD-CovidNet 的 550%, 且其分类准确率比 DD-CovidNet 低了 7.55%。CoroNet^[13] 总体分类准确率最高, 比 DD-CovidNet 高了 0.02%, 但其参数量是 DD-CovidNet 的 165 倍。此外, 它对 COVID-19 的精度与灵敏度分别为 97% 和 89%, 比 DD-CovidNet 的分别低了 3.00% 与 7.08%。因此, 相对而言, DD-CovidNet 模型能更快速、更准确地识别 COVID-19 的 CXR 图像。

4 结 语

针对 CXR 图像的特点, 本文提出 2 个对特征信息敏感的模块 DMFF 和 D3S。基于这 2 个模块, 提出一种具有表征能力强且复杂度低的 CNN 模型——DD-CovidNet。实验结果表明: DD-CovidNet 的整体准确率、精度、灵敏度、特异性和 F_1 值分别为 94.57%, 96.52%, 93.68%, 96.16% 和 94.90%; 识别 COVID-19 的精度、灵敏度和特异性分别为 100.00%, 96.08% 和 100.00%。通过对实验结果进行分析与比较发现, DD-CovidNet 模型能更快速准确地识别 COVID-19 的 CXR 图像。DD-CovidNet 虽有不错的实验结果, 但仍需进一步临床研究与测试。目前开源的 COVID-19 的 CXR 图像稀缺, 实

验的 COVID-19 阳性 CXR 图像数量较少, 可能导致模型识别存在一定偏差, 还需要根据实际应用进行优化. 后续将努力收集 COVID-19 的 CXR 图像, 并研究新的图像处理技术以克服数据有限所带来的困难.

参考文献(References):

- [1] Xie X Z, Zhong Z, Zhao W, *et al.* Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing[J]. *Radiology*, 2020, 296(2): 200343
- [2] Ng M Y, Lee E Y P, Yang J, *et al.* Imaging profile of the COVID-19 infection: radiologic findings and literature review[J]. *Radiology Cardiothoracic Imaging*, 2020, 2(1): Article No.e200034
- [3] Shi F, Wang J, Shi J, *et al.* Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19[J]. *IEEE Reviews in Biomedical Engineering*, 2021, 14: 4-15
- [4] Wang W, Hu Y Y, Zou T, *et al.* A new image classification approach via improved MobileNet models with local receptive field expansion in shallow layers[J]. *Computational Intelligence and Neuroscience*, 2020, 2020: Article No.8817849
- [5] Han Li, Piao Jingyu, Lan Pengyan, *et al.* 3D shape classification method based on shape-aware deep learning[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2021, 33(1): 29-38(in Chinese)
(韩丽, 朴京钰, 兰鹏燕, 等. 结构感知深度学习的三维形状分类方法[J]. *计算机辅助设计与图形学学报*, 2021, 33(1): 29-38)
- [6] Gao Xiang, Li Chungeng, An Jubai. Real-time image semantic segmentation based on attention mechanism and multi-label classification[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2021, 33(1): 59-67(in Chinese)
(高翔, 李春庚, 安居白. 基于注意力和多标签分类的图像实时语义分割[J]. *计算机辅助设计与图形学学报*, 2021, 33(1): 59-67)
- [7] Roth H R, Lu L, Liu J M, *et al.* Improving computer-aided detection using convolutional neural networks and random view aggregation[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1170-1181
- [8] Yang X, Kwitt R, Styner M, *et al.* Quicksilver: fast predictive image registration - a deep learning approach[J]. *Neuroimage*, 2017, 158(1): 378-396
- [9] Maguolo G, Nanni L. A critic evaluation of methods for COVID-19 automatic detection from X-ray images[OL]. [2020-11-22]. <https://arxiv.org/abs/2004.12823v4>
- [10] Tartaglione E, Barbano C A, Berzovini C, *et al.* Unveiling COVID-19 from Chest X-ray with deep learning: a hurdles race with small data[OL]. [2020-11-22]. <https://arxiv.org/abs/2004.05405>
- [11] Apostolopoulos I D, Mpesiana T A. COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks[J]. *Physical and Engineering Sciences in Medicine*, 2020, 43: 635-640
- [12] Wang L, Lin Z Q, Wong A. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images[J]. *Scientific Reports*, 2020, 10: Article No.19549
- [13] Khan A I, Shah J L, Bhat M M. CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest X-ray images[J]. *Computer Methods and Programs in Biomedicine*, 2020, 196: 105581
- [14] Ozturk T, Talo M, Yildirim E A, *et al.* Automated detection of COVID-19 cases using deep neural networks with X-ray images[J]. *Computers in Biology and Medicine*, 2020, 121: 103792
- [15] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C] // *Proceedings of Advances in Neural Information Processing Systems*. Cambridge: MIT Press, 2012: 1097-1105
- [16] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[OL]. [2020-11-22]. <https://arxiv.org/abs/1409.1556>
- [17] Szegedy C, Liu W, Jia Y Q, *et al.* Going deeper with convolutions[C] // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2015: 1-9
- [18] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image Recognition[C] // *Proceedings of the IEEE International Conference on Computer Vision*. Los Alamitos: IEEE Computer Society Press, 2016: 770-778
- [19] Huang G, Liu Z, van der Maaten L, *et al.* Densely connected convolutional networks[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2017: 2261-2269
- [20] Iandola F N, Moskewicz M W, Ashraf K, *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[OL]. [2020-11-22]. <https://arxiv.org/abs/1602.07360>
- [21] Howard A G, Zhu M L, Chen B, *et al.* MobileNets: efficient convolutional neural networks for mobile vision applications[OL]. [2020-11-22]. <https://arxiv.org/abs/1704.04861>
- [22] Zhang X Y, Zhou X Y, Lin M X, *et al.* ShuffleNet: an extremely efficient convolutional neural network for mobile devices[OL]. [2020-11-22]. <https://arxiv.org/abs/1707.01083>
- [23] Chollet F. Xception: deep learning with depthwise separable convolutions[OL]. [2020-11-22]. <https://arxiv.org/abs/1610.02357>
- [24] Rajaraman S, Siegelman J, Alderson P O, *et al.* Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays[J]. *IEEE Access*, 2020, 8: 115041-115050
- [25] Sandler M, Howard A, Zhu M L, *et al.* Inverted residuals and linear bottlenecks: mobile networks for classification, detection and segmentation[OL]. [2020-11-22]. <https://arxiv.org/abs/1801.04381v2>
- [26] Ma N N, Zhang X Y, Zheng H T, *et al.* ShuffleNet V2: practical guidelines for efficient CNN architecture design[OL]. [2020-11-22]. <https://arxiv.org/abs/1807.11164>
- [27] Mooney P. Chest X-ray images (Pneumonia)[OL]. [2020-11-22]. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>
- [28] Cohen J P, Morrison P, Dao L. COVID-19 image data collection[OL]. [2020-11-22]. <https://arxiv.org/abs/2003.11597>
- [29] Wong H, Lam H, Fong A, *et al.* Frequency and distribution of chest radiographic findings in COVID-19 positive patients[J]. *Radiology*, 2020, 296(2): E72-E78