用于胸部X光图像中 COVID-19识别的改进型深度学习模型设计：一种知识蒸馏方法

AmirReza BabaAhmadi*

德黑兰大学工程学院机械工程学院，伊朗德黑兰（电子邮箱：babaahmadi.amir@ut.ac.ir）

Sahar Khalafi

伊斯法罕理工大学电气与计算机工程系，伊朗伊斯法罕

Masoud ShariatPanahi

伊朗德黑兰，德黑兰大学工程学院机械工程学院

Moosa Ayati

伊朗德黑兰，德黑兰大学工程学院机械工程学院

摘要

背景：COVID-19 对社会产生了重大影响，因此需要准确识别并提供适当的医疗治疗。为解决这一问题，人们探索了各种手动和自动特征提取技术。然而，深度学习模型的计算需求限制了其普及，尤其是在资源有限的机构和社会中。

目的：本研究旨在实现两个主要目标：降低模型在嵌入式设备、移动设备和传统计算机上运行时的计算成本，并提升模型相较于现有方法的性能。研究旨在确保模型在医疗识别任务中具有出色的性能和准确率。

方法：采用 VGG19 和 ResNet50V2 两种神经网络，以增强从数据集中提取特征的能力。这些网络提取语义特征，并在全连接分类器层中合并其特征向量，从而实现对正常病例和新冠肺炎病例的满意分类效果。为减轻计算负担，采用了以在移动设备和嵌入式设备上高效运行而著称的网络 MobileNetV2。通过知识蒸馏，将知识

从教师网络（ResNet50V2 和 VGG19）迁移至学生网络（MobileNetV2），从而提升其识别 COVID-19 的性能。

结果：采用预训练网络和5折交叉验证来评估所提出的方法。该模型在检测感染病例和正常病例方面的准确率达到98.8%，F1分数达到99.1%。

意义：结果表明，所提出的方法性能优越。学生模型通过交叉验证获得了可接受的准确率和F1分数，适用于普通计算机、嵌入式系统以及临床专家的手机。该方法为COVID-19识别提供了一种经济高效的解决方案，有助于提高检测的可及性并实现准确诊断。与以往研究相比，我们的方法在准确率、F1分数及其他相关指标方面表现更优，至少提高了1%。

关键词：COVID-19，深度学习，医学图像分析，知识蒸馏，胸部X光图像，教师-学生模型

1. 背景与目标

近年来，COVID-19一直严重威胁人类的生命与健康，并对全球医疗卫生系统造成了不利影响。不幸的是，COVID-19具有高度传染性和传播性。因此，开发能够快速诊断COVID-19并避免其不良影响的预测系统势在必行。众多科学家开展了多项研究，旨在开发达到人类水平的预测和诊断系统，帮助社区共同抗击这一疾病。

这些方法通常能够有效地利用计算机断层扫描和胸部X光图像检测 COVID-19。然而，其致命缺陷在于，自动特征提取技术，尤其是深度学习模型，需要大量计算资源才能执行特定任务。许多医院和机构无力购置能够运行这些算法的昂贵硬件系统。

此外，缺乏医学专家可能会危及许多发展中国家民众的健康。在一些发展中国家，能够根据医学图像识别疾病的人类专家十分稀缺，这是不容忽视的现实。因此，本研究尝试提供一种替代方案来解决这一问题。

识别 COVID-19 是迈向治疗的第一步。本文的主要目标是尝试解决这一步骤。COVID-19 通常伴有咳嗽、感冒以及

呼吸急促等症状。根据世界卫生组织的说法，呼吸系统问题是 COVID-19 的致命症状，可通过医院和医疗诊所获取的 CT扫描或胸部X光片进行识别。

近年来，深度学习及其相关技术已广泛应用于计算机视觉任务，例如医学图像分析。相关示例包括肿瘤检测[1], 糖尿病视网膜病变[2],等。

鉴于医学图像分析是计算机视觉和深度学习领域最令人振奋的应用之一，自动特征提取算法对于以可与人类专家相媲美的水平进行疾病诊断至关重要。由于初步诊断具有关键作用，研究人员已将这一环节的准确率确定为一项重大挑战。

1.2 相关工作

以下部分将考察相关工作和研究，以便为我们的研究建立坚实的框架。在[3], 有人提出了一种基于堆叠卷积神经网络并结合多个预训练网络的模型；然而，该研究存在较高的计算成本。在[4], 有人使用了一种采用改进卷积核的基于残差层的神经网络，用于预测正常胸部X 射线图像和肺炎胸部X 射线图像中是否存在 COVID-19。该研究解决了梯度消失和退化问题；然而，其改进的卷积核并未显著降低计算成本，因此不适合用于边缘设备。

在 [5], 中，提出了一种基于集成学习分类器的模型，以解决不平衡数据集的问题。然而，使用集成模型存在若干缺点，例如可解释性降低、部署复杂度增加以及计算需求更高。

另一篇论文 [6] 提出了将 LSTM（用于检测）与卷积神经网络（用于识别）相结合的方法。该方法的准确率指标达到 99.4%。然而，LSTM 主要用于捕捉序列数据中的时间依赖关系，可能无法充分利用医学图像处理中至关重要的空间关系。此外，该方法缺乏可解释性。

在 [7], 中介绍了一种采用长短期记忆网络和生成对抗网络架构的模型。该方法在二分类任务中无需额外的网络进行特征提取。然而，生成对抗网络可能会出现模式崩塌，即生成图像的变化有限，或无法充分体现目标分布的多样性。在医学图像处理中，充分捕捉解剖结构或

病理变化的全部范围至关重要，而模式崩塌可能导致图像不真实或存在偏差，从而限制生成对抗网络在某些应用中的实用性。评估和验证生成对抗网络生成的医学图像也颇具挑战，因为计算机视觉中使用的传统评估指标可能无法充分反映生成图像的质量或临床相关性。

论文 [8]，使用基于卷积神经网络（CNN）的深度神经网络，并在不同层中添加额外组件，以研究医学图像中 COVID-19 的存在情况。遗憾的是，这项工作未关注计算成本，其唯一目标是提高准确率和评估指标。

在 [9]，中，研究人员结合 CNN 和支持向量机（SVM）分析 COVID-19 胸部X光图像。然而，使用 SVM 进行分类可能耗时较长，且无法处理海量数据。使用 SVM 的另一项缺点是，训练过程将不再是端到端的。

在 [10]，中，研究人员提出了一种新框架，利用 DarkNet53 处理胸部X光图像。DarkNet53 最初被设计为 YOLO 网络的定制骨干网络，用于目标检测，因此可能不适用于分类任务。

论文 [11] 开展了一项比较研究，以确定性能最佳的预训练深度学习架构。研究人员考察了 15 个模型，以找出最优架构；根据这项研究，VGG19 被认为是准确率最高的预训练模型。

另一项比较研究 [12] 使用 DenseNet201、InceptionV3、ResNet50 及其他几种新型预训练架构，将 COVID-19 与正常病例和肺炎区分开来。然而，这两篇论文都仅关注比较预训练网络的性能，未考虑这些网络在医院环境中的实际适用性。

在论文 [13], 中，介绍了一种基于生成对抗网络和卷积神经网络的新模型。该研究使用多个分类器执行分类任务。此外，论文还利用一个独特的数据集开发了 COVID-19 分割模型。然而，我们认为，使用多个分类器对分类任务并无显著效果，还会带来较高的计算需求，因此这种方法不适用于实际应用。

在 [14], 中，研究者对预训练架构进行了比较，并加以修改以减少可训练参数、加快算法训练。这是一项有价值的贡献；然而，还必须解决遗忘（灾难性遗忘）问题，而本文并未对此进行讨论。

在[15], 中，研究考察了多种预训练架构，包括 AlexNet、VGG16 和 ResNet系列。此外，本文重点评估了迁移学习方法中使用的冻结层及其他配置。

[16] 开展了一项研究，旨在确定最适合COVID-19识别的候选模型。研究结果表明，Inception网络在COVID-19检测方面取得了出色的成果。

[17] 开展了一项比较研究，使用了DenseNet121、ResNet50、VGG16和VGG19。这些研究有助于为COVID-19识别寻找合适的骨干网络。然而，它们没有讨论计算成本和部署问题，也没有为改进架构的技术层面作出贡献。

在[18], 中，研究执行了一项二分类任务，并对预训练的SqueezeNet、DenseNet121、ResNet50和ResNet18网络的最后一层进行了微调。然而，这项工作的不足之处在于，仅靠微调无法解决训练这些网络所需计算资源过高以及其部署配置方面的挑战。

论文 [19] 提出了一种基于投票的流程，用于评估 VGG16、InceptionV3 和 ResNet50 的预测结果。最后一层会同时比较所有上述网络的性能，并选出最佳结果。虽然该模型表现良好，但将其部署到实际应用中较为复杂，且需要进行大量修改以降低计算开销。

文献 [20], 评估了预处理对基于CNN的网络结果的影响。然而，仅进行预处理只能提升性能，无法解决这些网络带来的计算负担。

文献 [21], 介绍了为胸部X光（CXR）和胸部X光医学图像开发的一种定制网络 COVID-Net。然而，COVID-Net 结构复杂，其部署过程需要投入大量精力来降低计算负担。

另一项研究 [22]使用最先进的算法和 CNN架构比较了 COVID-19、细菌性肺炎和病毒性肺炎的分类结果。尽管这项研究颇有意义，但它仍然

采用仅经过少量修改的预训练网络，前述问题依然没有得到解决。

在另一项研究中，[23], 研究人员使用主成分分析（PCA）提高特征提取的效率。然而，我们认为 PCA 存在线性限制且对离群值敏感，因此不太适合当前任务。

在[24], 该研究中，研究人员采用了增加卷积层并使用经过修改的卷积核的 YOLO 网络，从胸部X光图像中检测 COVID-19。如前所述，采用修改后卷积核的 YOLO 非常适合目标检测任务，但并未针对图像分类进行专门优化，这可能会给分类任务带来问题。

在[25], 该研究中，研究人员考察了 Efficient-Net 系列在正常病例中检测 COVID-19 和肺炎的表现。该网络非常适合实际应用。然而，这项工作并未着重解决现实世界实际应用中的缓解问题。训练主要在学术环境中进行。

在[26], 该研究中，研究人员介绍了一种名为 CoroNet 的网络，该网络基于 Xception 网络，可进行多分类，包括病毒性肺炎、细菌性肺炎和 COVID-19 类别。该架构轻量，但其准确率和其他指标表现不够理想，因此不适合作为现实世界应用的候选方案。

还有一些相关工作采用模糊增强深度学习方法，利用胸部X光图像进行 COVID-19识别。作者声称，他们的算法能够提升在不平衡数据集和便携式胸部X光图像上的性能 [27]。

另一项研究提出了一种新颖的神经进化算法，用于利用胸部X光图像进行 COVID-19识别。研究人员使用 K 近邻（KNN）算法，将感染病例与正常病例区分开来 [28]。然而，我们认为 KNN 无法处理海量数据，因此并非现代深度学习算法的理想分类器。

我们审阅的所有论文都没有同时解决缓解问题和降低计算开销的问题。因此，我们致力于开发一种能够兼顾这两个方面的方法，本文余下部分将对此进行详细阐述。

为克服这些挑战，我们探索了知识蒸馏技术（也称为教师-学生模型）的应用。我们的主要目标有两个：首先，在胸部X光图像中实现对 COVID-19 的高准确率检测；其次，大幅降低训练所需的计算

能力。降低计算需求后，资源匮乏的国家或城市中可获得的普通计算机也能用于训练。此外，我们还希望将训练好的模型应用于低功耗芯片移动设备，以便高效地从医学图像中识别疾病。

为实现这些目标，我们提出了一种新颖的架构，利用双网络提取更丰富的语义特征。我们将这些知识迁移至轻量级网络，从而便于将其高效集成到软件工程和生产流水线中。为构建二分类数据集，我们合并了两个来自 Kaggle 的公开数据集，详情见第 4 节。

我们的研究取得了极具前景的结果和评估指标。我们实现了高达 98.8% 的准确率和 99.1% 的 F1分数，超越了此前发表的研究成果。因此，我们的方法学展现出巨大潜力，能够降低计算资源需求，同时提高专为嵌入式设备设计的神经网络的准确率。

主要贡献：

1. 计算效率：通过采用知识蒸馏，我们的方法显著降低了训练过程中的计算需求，使资源匮乏地区也能使用普通计算机。此外，该方法还缓解了深度神经网络中出现的遗忘问题。2. 移动设备兼容性：我们的模型旨在高效部署于低功耗芯片移动设备，从而实现医学图像疾病识别的低成本便携应用。3. 提高准确率：我们的方法学取得了 98.8% 的准确率和 99.1% 的 F1分数，表现出优越性能，超越此前发表的研究成果。4. 嵌入式设备的应用潜力：我们的方法学有望提高准确率 o

f

专为嵌入式设备设计的神经网络，拓展其在资源受限环境中的应用范围。

本文结构如下：第二节概述了以往针对该任务的研究成果。第三节介绍了知识蒸馏方法，并解释了将其纳入算法的原因。第四节讨论了用于开展本研究的数据集。第五节介绍了训练阶段和评估指标。最后，第六节呈现了

所提出模型的评估结果，并将其与相关工作进行比较；第七节总结了本文的研究发现。

2. 方法

如今，深度神经网络彻底改变了计算机视觉领域。其应用已在众多领域得到广泛研究，包括自动驾驶汽车 [29], 医学图像分析 [30], [31], 和农业 [32], 等。卷积神经网络已成为计算机视觉、自然语言处理、语音处理和视频分类任务中进行自动特征提取的最有效工具。

本文展示了基于卷积神经网络的神经网络如何改进涉及 COVID-19 和正常病例的二分类任务中的语义特征提取。DenseNet [33], ResNet [34], VGG[35], Xception [36], 和 Mobile-NetV2 [37] 是一些最强大的预训练网络。VGG19 是 VGG16 的扩展版本，包含十六个卷积层和三个全连接层。此外，它还使用五个最大池化层和一个最终的 SoftMax层。

ResNet50V2 是 ResNet50 的增强版本，性能优于 ResNet101 等先前版本。它在不同模块之间引入了若干新的连接，使其在 ImageNet竞赛中取得了很高的准确率。MobileNetV2 是一种采用深度卷积层的卷积网络，通过减少可训练参数，在降低计算成本的同时提高准确率。由于采用了倒置残差块，它适合嵌入手机、平板电脑及其他常规嵌入式设备中。

我们训练并评估了大量成对神经网络，以选出用于构建教师网络的最佳架构。此外，我们还针对学生架构训练了多个预训练神经网络。最终，我们选择了 VGG19 和 ResNet50V2，因为与其他架构相比，它们表现出色。选择学生架构的理由也同样如此。

由于 ResNet50V2 和 VGG19 生成的输出层（特征向量）大小相同，因此将它们的特征向量合并，以便为指定任务生成更丰富的语义特征集。在该架构中添加了一个卷积核大小为 1、包含 1024 个滤波器的卷积神经网络层。随后，将网络的输出展平并输入全连接层。值得注意的是，最后一个卷积神经网络层未使用激活函数。该

分类层由全连接层中的 64 个神经元组成，丢弃率为 0.5。随后又添加了一个全连接层作为最终层，其神经元数量等于该问题中的类别数。

整个架构被称为教师网络，而在更广泛的教师-学生模型或知识蒸馏框架中，MobileNetV2 充当学生网络。教师模型的结构如图1所示。ResNet50V2 和 VGG19 的张量特征尺寸均为 10*10*2048，拼接特征的尺寸为 10*10*4096。

在我们的研究中，我们采用 5折交叉验证来选择最优超参数。交叉验证是一种稳健的方法，广泛用于超参数调优和模型评估。采用 5 折交叉验证时，我们将数据集划分为五个子集或折。我们反复使用其中四折训练模型，同时用剩余一折进行验证。此过程重复五次，每次使用不同的一折作为验证集。通过系统地调整超参数，并评估模型在五折上的表现，我们确定了整体性能最佳的超参数配置。使用 5折交叉验证使我们能够有效评估和微调模型的超参数，从而确保研究结果的可靠性和泛化能力。

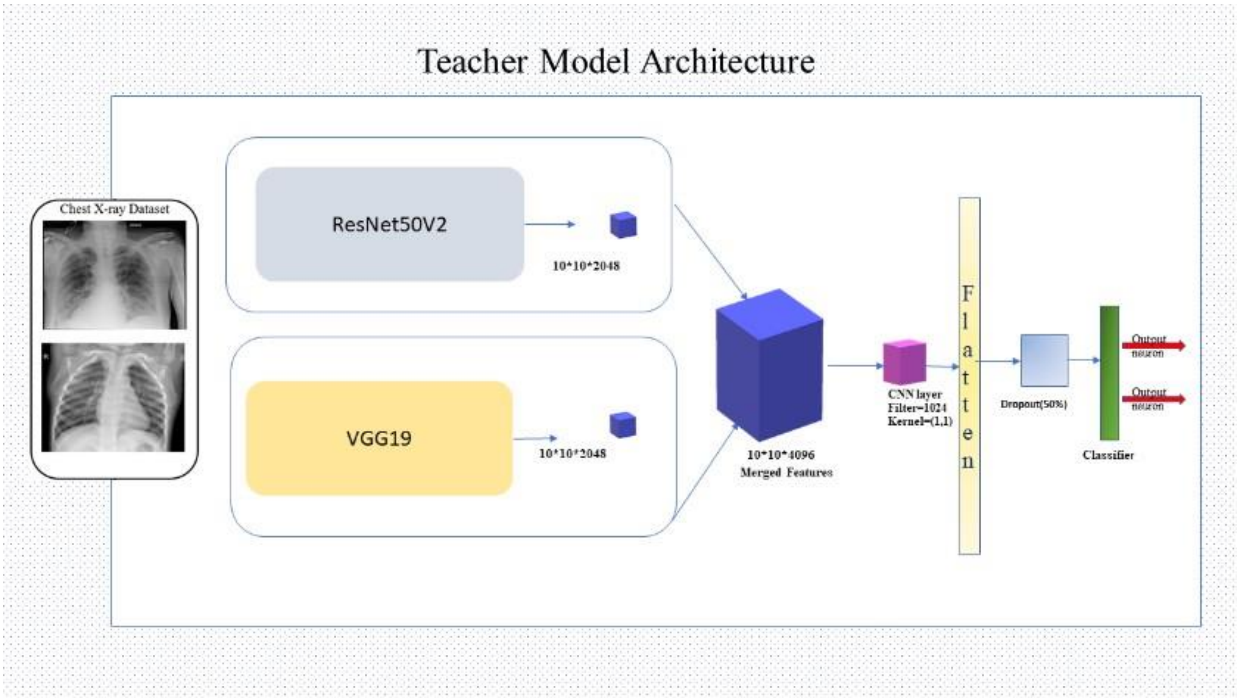


图1. 教师模型架构

3. 知识蒸馏

Hinton 等人于 [38] 开创了知识蒸馏。知识蒸馏（KD）是一种训练较小网络以模仿较大网络行为的过程。设计复杂网络作为教师，旨在学习更复杂的特征并取得更好的结果。然而，在标准计算机或嵌入式设备上运行我们的网络时，我们常常会遇到内存大小和计算成本方面的限制，从而引发各种问题。因此，我们需要找到解决这些问题的方案。

为了将知识从教师传递给学生，需要进行加权平均（均值）。初始目标函数是采用软目标的交叉熵，它通过较小网络中的Softmax函数在较高温度下计算得出。软目标由较大的架构或网络生成。第二个目标函数是采用有效标签的交叉熵，它使用学生模型在温度为零时的Softmax输出。

教师网络和学生网络开始并行接收训练数据。教师模型在其输出中采用带温度的Softmax函数。相比之下，学生模型会生成两个不同的输出。第一个输出是带温度的Softmax，第二个输出则是标准Softmax。学生模型的目的是生成经过软化的概率，使其与教师模型的输出相对应。知识蒸馏的损失使用以下公式计算：

$$L_{KD} = \alpha KL(p.q)T^2 + (1-\alpha)L(W_s.x) \qquad (1)$$

这里，p 和 q分别表示学生网络和教师网络在特定温度（T)下生成的概率，而 KL表示Kullback-Leibler散度，用于衡量两个概率分布之间的差异程度。学生模型在温度 $T=1$ 下的交叉熵为（$LW_s.x$）。根据 [38], 和 均为超参数，其中 的值越大，学生模型的学习体验越好。

α T α

在蒸馏阶段，只应在学生网络中执行反向传播，因为教师网络已经对其参数进行了微调。在整个蒸馏过程中，知识会从教师模型传递到学生模型。值得注意的是，学生模型的训练速度可以比教师模型更快。有关蒸馏过程的更多详情，请参阅 [39]。知识蒸馏过程如图 2 所示。

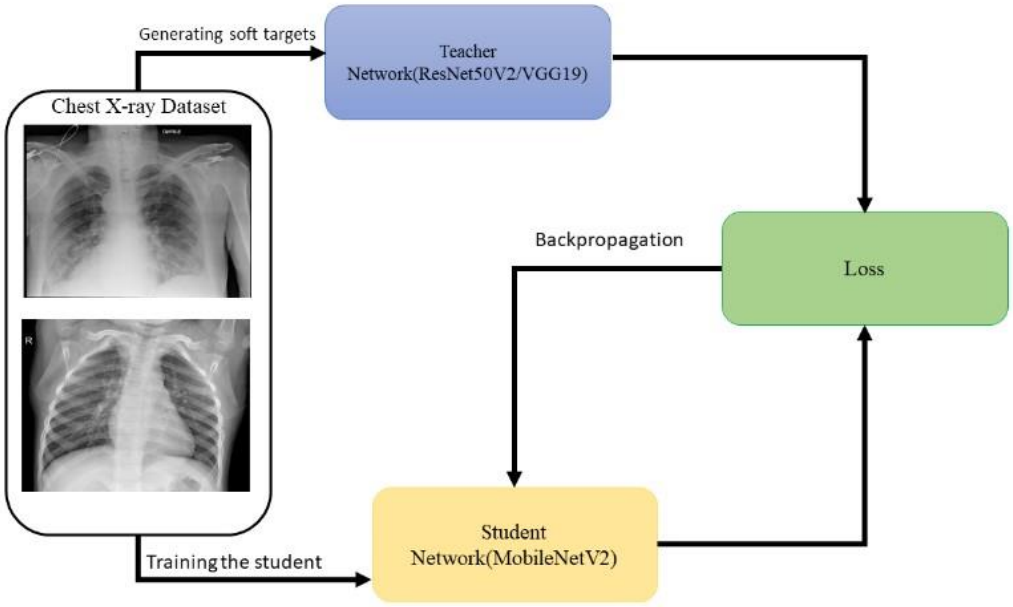


图 2. 用于 COVID-19 检测的知识蒸馏 [38]

4. 数据集

使用了两个公开数据集来训练所提出的深度学习模型并构建所需的数据集。首先，使用了一个位于 (https://github.com/ieee8023/covid-chestxray-dataset) ↗) 的公开数据集作为 COVID-19 阳性样本。随后，利用位于 (https://www.kaggle.com/c/rsna-pneumonia-detection-challenge) ↗) 的数据集收集阴性样本（正常病例）。

合并两个数据集后，共得到 118 例新冠肺炎病例和 8,851 例正常病例。由此可见，形成的数据集存在类别不平衡，阳性病例（COVID-19）的数量远少于正常胸部X光医学图像的数量。为解决这一问题，我们采用了采样技术。主要方法是为二分类任务从每个类别中选取相同数量的样本。我们采用过采样方法增加 COVID-19（阳性病例）样本的数量，确保正类和负类的样本数量相同。

过采样后，阳性病例数量增加至 8,851，而阴性（正常）样本数量保持不变。需要注意的是，本研究未纳入任何肺炎图像。肺炎可分为多种类别，如严重急性呼吸综合征、链球菌、急性呼吸窘迫综合征和肺孢子菌。将所有这些类别视为单一类别并不切实际。因此，

肺炎类型分类器的开发被推迟至未来的调查中。图3展示了从数据集中获得的部分 COVID-19 患者图像和正常图像。



图 3(a)：健康人

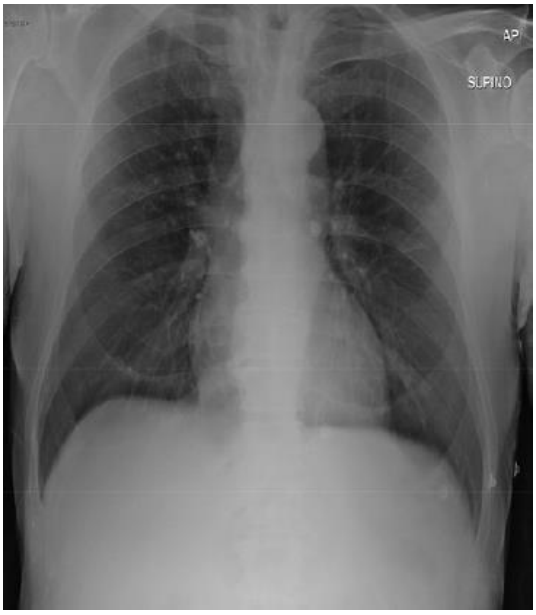


图 3(b)：COVID-19 患者

4.1 数据增强

在进行数据增强之前，先对图像进行归一化，以减少梯度消失和梯度爆炸可能引发的问题。随后，将图像调整为 224x224 的尺寸。为了

提高模型对医学图像中各种变化的适应能力，采用了数据增强技术。假设这些变化不会改变标签（真实标签）的定义。在为该数据集考虑的各种数据增强方法中，采用了 0 至 20 度范围内的随机旋转。这不仅有助于防止过拟合，还能加快学习曲线的收敛。

5. 训练阶段

约 80% 的数据集用于训练阶段，其余 20% 留作测试阶段。为确保准确评估模型性能，采用了K折交叉验证技术。所使用的损失函数为二元交叉熵，优化器为 Adam，学习率为 1e-5。交叉验证共使用 5 折（k=5）。批量大小设为 32，模型总共训练了 120 轮。

5.1 评估指标

为了评估模型在测试数据集上的性能，需要定义几项具体标准。真阳性指在阳性病例和正常病例中正确识别出 COVID-19 阳性病例。真阴性指准确识别正常病例。假阳性是指将 COVID-19 病例误判为正常病例。假阴性是指将正常病例误判为 COVID-19 病例。

精确率定义为真阳性数与真阳性数和假阳性数之和的比值。

$$Precision = \frac{TP}{TP + FP} \qquad (1)$$

敏感度是真阳性数量与真阳性和假阴性数量之和的比值。

$$Sensitivity(recall) = \frac{TP}{TP + FN} \qquad (2)$$

特异度是真阴性数量与假阳性和真阴性数量之和的比值。

$$Specificity = \frac{TN}{TN + FP} \qquad (3)$$

最后，我们采用了另外三项指标来评估所提出的模型在本研究所用数据集上的表现。这些指标包括 F1分数、平衡准确率，以及传统的准确率指标。

$$\text{精确率 召回率 F分数} = \frac{2 * \text{精确率} * \text{召回率}}{\text{精确率} + \text{召回率}} \qquad (4)$$

$$Balanced\ Accuracy = \frac{specificity + sensitivity}{2} \qquad (5)$$

$$Accuracy = \frac{TP + TN}{Positive + Negative} \qquad (6)$$

6. 结果与讨论

本节介绍将我们的方法应用于上述数据集所得到的结果。评估中使用的测试图像在分析前未曾被算法见过。因此，这些结果表明，所提出的算法能够有效处理此前未见过的医学图像。

表1展示了教师模型的训练结果。显然，教师模型在检测胸部X光图像方面表现良好，准确率和F1分数均令人满意。

接下来，观察表2中的 MobileNetV2（学生模型）结果。学生模型在分类任务中的表现令人满意。然而，尽管结果良好，我们仍决定通过知识蒸馏方法提升其分类能力。

表3展示了学生网络完成知识蒸馏过程后的结果。因此，学生网络的表现与文献中现有方法相当，在某些情况下，其准确率和F1分数甚至优于以往发表的研究。

表1：教师网络（ResNet50V2/VGG19）的评估结果

折叠编号	**Acc**	精确率	特异度	召回率（敏感度）	F1分数	平衡-**Acc**
1	0.978	0.978	0.978	0.978	0.978	0.978
2	0.989	0.978	1.000	1.000	0.988	1.000
3	1.000	1.000	1.000	1.000	1.000	1.000
4	0.989	0.978	1.000	1.000	0.988	1.000
5	1.000	1.000	1.000	1.000	1.000	1.000
平均值	**0.992**	**0.987**	**0.996**	**0.996**	**0.991**	**0.996**

表2：知识-蒸馏前学生网络（MobileNetV2）的评估结果

折叠编号	**Acc**	精确率	特异度	召回率（敏感度）	F1分数	平衡-**Acc**
1	0.987	0.978	1.000	1.000	0.992	1.000
2	0.974	0.976	0.976	0.976	0.972	0.976
3	0.974	0.976	0.976	0.976	0.972	0.976
4	0.987	0.978	1.000	1.000	0.992	1.000
5	0.987	0.978	1.000	1.000	0.992	1.000
平均值	**0.980**	**0.977**	**0.990**	**0.990**	**0.984**	**0.990**

表3：知识蒸馏后学生网络（MobileNetV2）的评估结果

折叠编号	**Acc**	精确率	特异度	召回率（敏感度）	F1分数	平衡-**Acc**
1	0.974	0.976	0.976	0.976	0.972	0.976
2	1.000	1.000	1.000	1.000	1.000	1.000

3	0.978	0.978	1.000	1.000	0.992	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000
5	0.989	0.978	1.000	1.000	0.988	1.000
平均值	**0.988**	**0.987**	**0.996**	**0.996**	**0.991**	**0.996**
%改进 0.8 %		1.0 %	0.6 %	0.6 %	0.7 %	0.6 %

表4：各架构的总参数量

教师网络（ResNet50V2/VGG19）	49,222,390
学生网络（MobileNetV2）	2,334,966
参数量减少（%）	-95.3 %

值得强调的是，在评估指标表现方面，学生网络优于其先前版本（未进行知识蒸馏）。这一改进可归因于知识蒸馏的作用，它在一定程度上提升了MobileNetV2的性能。此外，知识蒸馏还有助于网络缓解神经网络常见的遗忘问题。

学生模型的表现表明，它适用于嵌入式系统中的医学识别任务；由于采用了深度卷积层和知识蒸馏，该模型所需的计算量极少。

表4展示了各架构中的参数数量。值得注意的是，知识蒸馏不仅提升了学生模型的性能，还在保持性能的同时将参数数量减少了约95.3%。因此，学生模型成为COVID-19识别任务的可行候选方案。

6.1 与相关文章的比较

与相关工作进行全面比较存在一些需要指出的挑战。首先，现有大量用于 COVID-19检测的胸部X光数据集

，相关研究使用单个数据集或来自多个来源的组合数据集。数据集使用方式的多样性使直接比较颇具挑战。其次，有些研究使用其他研究人员无法获取的私有数据集，进一步阻碍了直接比较。第三，分类任务各不相同，例如二分类或多分类，这增加了比较过程的复杂性。此外，所采用的模态也带来了另一层难度，无论是 CT扫描图像还是胸部X光图像，因为不同算法是专门针对COVID-19识别中的不同模态设计的。

需要注意的另一重要一点是，尽管以往的相关工作采用了知识蒸馏，但与我们的研究相比，它们在所使用的数据集、架构和模态方面存在显著差异。因此，不宜将我们的研究与这些工作直接进行比较。

此外，我们的研究着重于在最大化评估指标的同时，尽可能降低计算成本和参数数量。我们优先开发高效模型，而非构建参数过多的复杂模型，以便与最先进的方法进行比较。尽管面临这些固有挑战，我们仍努力采用相同的数据集和方法学进行训练与评估，并通过交叉验证将本研究与相关研究进行比较。这些比较结果汇总于表5。

表5。与相关工作对比

方法	类别	准确率(%)	F1分数(%)	敏感度(%)	特异度(%)
我们的工作-学生网络	二分类	98.80	99.10	99.60	99.60
我们的工作-教师网络	二分类	99.20	99.10	99.60	99.60
Duran-Lopez等人 al. [40]	二分类	94.4	93.14	92.53	96.33
Khan等人 [41]	二分类	99.0	98.5	99.3	98.6
Sahlol等人 [42]	二分类	98.7	98.21	未报告	未报告
Al-Waisy等人 [43]	二分类	未报告	98.68	98.72	98.63

AlMohimeed 等人。[44]	多分类	98.48	98.48	98.48	未报告
Chakraborty 等人。[45]	多分类	96.43	93.0	93.68	99.0
Loey 等人 [46]	多分类	96.0	未报告	未报告	未报告
Ieracitano 等人 [27]	二分类	80.9	未报告	82.5	76.6
Duong 等人 [47]	多分类	91.93	93.8	95.2	未报告
Deb 等人 [48]	多分类	85.71	90	90	90
Soundrapandiyan 等人 [49]	多分类	96.1	未报告	未报告	未报告
Lasker等人 [50]	多分类	97.28	未报告	未报告	未报告
Agrawal等人 [51]	二分类	93.6	96.65	93.6	100

根据表5所示的比较结果，我们设计的教师模型在所有评估指标上均优于以往研究，包括准确率、特异度、F1分数和敏感度。对于蒸馏后的学生网络，我们的教师网络优于所有已报告的结果，并展现出最先进的性能。

考虑到计算成本，如果不存在降低此类成本的限制，我们的教师模型可以被视为适合实施和部署的可行候选方案。然而，对于计算能力较低的设备，可以放心使用我们的学生网络，因为它在性能和计算效率之间实现了良好的平衡。

7. 结论

本文提出了一种用于识别 COVID-19 医学胸部X射线图像的新方法。通过采用过采样和数据增强技术，解决了数据集不平衡这一挑战。利用五折交叉验证方法对所提出模型的性能进行了评估，以确保其稳健性和准确率。

应用知识蒸馏后，该算法取得了令人瞩目的成果，准确率达到 98.8%，F1分数达到 99.1%。这些高性能指标证明了所提出的方法在 COVID-19识别任务中的有效性。

然而，为了进一步提升该算法，建议纳入来自不同国家的更多样化数据集。纳入此类数据集将有助于提高该算法的泛化能力和准确率。

针对未来在医学图像数据集方面的研究，建议探索多个方向，包括研究增量学习技术、运用自监督深度学习方法、探讨视觉Transformer架构，以及研究对抗攻击下的知识蒸馏。这些方向有望推动医学图像分析领域的发展，并进一步改进COVID-19识别算法。

声明

利益冲突：

通讯作者代表所有作者声明不存在利益冲突。

伦理批准：

本研究论文不适用此项内容。

资金支持：

本文完成过程中未获得任何外部资金或财务支持。研究由团队独立开展，仅依靠团队自身的努力和资源完成。尽管没有资金支持并不会降低这项工作的价值，作者仍感谢在研究过程中提供帮助或支持的所有人士。本文遵循伦理规范，并恰当引用了资料来源。如有需要，作者愿意提供进一步说明。

作者贡献：

AmirReza BabaAhmadi：概念化、方法学、调查、代码编写——初稿撰写。

Sahar Khalafi：数据预处理、正式分析与评估结果、可视化、初稿撰写。

Masoud ShariatPanahi：监督、项目管理、撰写——审阅与编辑 g.

Mousa Ayati：监督、项目管理、撰写——审阅与编辑。

所有作者均已阅读并批准最终稿件。

数据和材料的可用性：

本文研究中使用的所有数据集均可在 Kaggle 上免费获取。

参考文献

[1] K. Paeng、S. Hwang、S. Park 和 M. Kim，“乳腺组织病理学中肿瘤增殖评分预测的统一框架”，载于《计算机科学讲义》（包括《人工智能讲义》和《生物信息学讲义》子系列），2017 年，第 10553 卷 LNCS，第 231–239 页，数字对象标识符：10.1007/978-3-319-67558-9_27。[2] J. de la Torre、A. Valls 和 D. Puig，“用于糖尿病视网膜病变疾病分级的可解释深度学习分类器”，Neurocomputing，第 396 卷，第 465–476 页，2020 年 7 月，数字对象标识符：10.1016/J.NEUCOM.2018.07.102。[3] A. Gupta、Anjum、S. Gupta 和 R. Katarya，“InstaCovNet-19：一种用于通过胸部X光检测 COVID-19患者的深度学习分类模型”，《应用软计算》，第 99 卷，2021 年 2 月，数字对象标识符：10.1016/J.ASOC.2020.106859。[4] C. Ouchicha、O. Ammor 和 M. Meknassi，“CVDNet：一种用于从胸部X光图像中检测冠状病毒（Covid-19）的新型深度学习架构”，Chaos, Solitons and Fractals，第 140 卷，2020年11月，数字对象标识符：10.1016/J.CHAOS.2020.110245。[5] P. Saha、M. S. Sadi 和 M. M. Islam，“EMCNet：利用 X 射线图像自动诊断 COVID-19，usin

g

卷积神经网络与机器学习分类器集成，”Informatics Med. Unlocked，第 22 卷，2021 年 1 月，数字对象标识符：10.1016/J.IMU.2020.100505。[6] M. Z. Islam、M. M. Islam 和 A. Asraf，“结合深度 CNN-LSTM 网络，利用 X 射线图像检测新型冠状病毒（ COVID-19 ） ”，Informatics Med. Unlocked，第 20 卷，2020 年 1 月，数字对象标识符：10.1016/J.IMU.2020.100412。[7] S. Sheykhivand 等，“开发一种高效深度神经网络，利用胸部X光图像自动检测 COVID-19”，Alexandria Eng. J.，第 60 卷，第 3 期，第 2885–2903 页，2021 年 6 月，数字对象标识符：10.1016/J.AEJ.2021.01.011。[8] E. Hussain、M. Hasan、M. A. Rahman、I. Lee、T. Tamanna 和 M. Z. Parvez，“CoroDet：一种基于深度学习的分类

方法，利用胸部X光图像检测 COVID-19”，Chaos, Solitons and Fractals，第 142 卷，2021 年 1 月，数字对象标识符：10.1016/J.CHAOS.2020.110495。[9] Y. KUTLU 和 Y. CAMGÖZLÜ，“使用深度卷积神经网络从 X 射线图像中检测冠状病毒病（COVID-19）”，Nat. Eng. Sci.，第 6 卷，第 1 期，第 60–74 页，2021 年 1 月，数字对象标识符：10.28978/NESCIENCES.868087。[10] R. C. Joshi 等，“一种基于深度学习的 COVID-19 自动诊断框架，使用胸部X光图像”，Biocybern. Biomed. Eng.，第 41 卷，第 1 期，第 239–254 页，2021 年 1 月，数字对象标识符：10.1016/J.BBE.2021.01.002。[11] M. M. Rahaman 等，“使用深度学习从胸部X光图像中识别 COVID-19 样本：迁移学习方法比较”，J. Xray. Sci. Technol.，第 28 卷，第 5 期，第 821–839 页，2020 年，数字对象标识符：10.3233/XST-200715。[12] K. El Asnaoui 和 Y. Chawki，“利用 X 射线图像和深度学习自动检测冠状病毒病”，J. Biomol. Struct. Dyn.，第 1–12 页，2020 年，数字对象标识符：10.1080/07391102.2020.1767212。[13] M. Loey、F. Smarandache 和 N. E. M. Khalifa，“在缺乏胸部 COVID-19 X 光数据集的情况下：一种基于生成对抗网络和深度迁移学习的新型检测模型”，Symmetry (Basel).，第 12 卷，第 4 期，2020 年 4 月，数字对象标识符：10.3390/SYM12040651。[14] R. K. Singh、R. Pandey 和 R. N. Babu，“COVIDScreen：一种可解释的深度学习框架，利用胸部X光片对 COVID-19 进行鉴别诊断”，Neural Comput. Appl.，2021 年，数字对象标识符：10.1007/S00521- 020-05636-6。[15] S. Sarv Ahrabi、M. Scarpiniti、E. Baccarelli 和 A. Momenzadeh，“用于 COVID-19 疾病检测的深度学习架构在准确率与复杂度方面的比较”，Computation，第 9 卷，第 1 期，第 1–20 页，2021 年 1 月，数字对象标识符：10.3390/COMPUTATION9010003。[16] I. Lorencin 等，“使用 X 射线图像和卷积神经网络自动评估 COVID-19患者的肺部状况”，J. Pers. Med.，第 11 卷，第 1 期，第 1–31 页，2021 年 1 月，数字对象标识符：10.3390/JPM11010028。[17] R. Jain、M. Gupta、S. Taneja 和 D. J. Hemanth，“基于深度学习的胸部X光图像 COVID-19 检测与分析”，Appl. Intell.，第 51 卷，第 3 期，第 1690–1700 页，2021 年 3 月，数字对象标识符：10.1007/S10489-020- 01902-1。[18] S. Minaee、R. Kafieh、M. Sonka、S. Yazdani 和 G. Jamalipour Soufi，“Deep-COVID：利用深度迁移学习从胸部X光图像预测 COVID-19”，Med. Image Anal.，第 65 卷，2020 年 10 月，数字对象标识符：10.1016/J.MEDIA.2020.101794。[19] S. C. R 和 R. K. Dubey，“基于深度学习的混合模型：使用胸部

X 射线”，2020年8月，数字对象标识符：10.36227/TECHRXIV.12839204.V1.

[20] J. D. Arias-Londono、J. A. Gomez-Garcia、L. Moro-Velazquez 和 J. I. Godino-Llorente，“将人工智能应用于胸部X光图像以自动检测 COVID-19：一种深思熟虑的评估方法”，IEEE Access，2020年，数字对象标识符：10.1109/ACCESS.2020.3044858。[21] L. Wang、Z. Q. Lin 和 A. Wong，“COVID-Net：一种用于从胸部X光图像中检测新冠肺炎病例的定制深度卷积神经网络设计”，Scientific Reports，第10卷，第1期，2020年12月，数字对象标识符：10.1038/S41598-020-76550-Z。[22] I. D. Apostolopoulos 和 T. A. Mpesiana，“COVID-19：利用迁移学习与卷积神经网络从 X 射线图像中自动检测”，Physics and Engineering Sciences in Medicine，第43卷，第2期，第635–640页，2020年6月，数字对象标识符：10.1007/S13246-020-00865-4。[23] T. Garg、M. Garg、O. P. Mahela 和 A. R. Garg，“采用迁移学习的卷积神经网络进行 COVID-19 识别：不同方法的比较研究”，人工智能，第1卷，第4期，第586–606页，2020年12月，数字对象标识符：10.3390/AI1040034。[24] T. Ozturk、M. Talo、E. A. Yildirim、U. B. Baloglu、O. Yildirim 和 U. Rajendra Acharya，“使用基于 X 射线图像的深度神经网络自动检测 COVID-19 病例”，Computers in Biology and Medicine，第121卷，2020年6月，数字对象标识符：10.1016/J.COMPBIOMED.2020.103792。[25] E. Luz 等，“探索一种有效且高效的深度学习模型，用于 X 射线图像中的 COVID-19 模式检测”，Research on Biomedical Engineering，2021年4月，数字对象标识符：10.1007/S42600-021-00151-6。[26] A. I. Khan、J. L. Shah 和 M. M. Bhat，“CoroNet：一种用于从胸部X光图像中检测和诊断 COVID-19 的深度神经网络”，Computer Methods and Programs in Biomedicine，第196卷，第105581页，202

0,

数字对象标识符：10.1016/J.CMPB.2020.105581.[27] C. Ieracitano 等，“一种基于模糊增强深度学习的方法，用于从便携式胸部X光图像中早期检测新冠肺炎”，Neurocomputing，第 481 卷，第 202–215 页，2022年4月，数字对象标识符：10.1016/J.NEUCOM.2022.01.055。[28] S. M. J. Jalali、M. Ahmadian、S. Ahmadian、R. Hedjam、A. Khosravi 和 S. Nahavandi，“基于 X 射线图像的

d

采用进化深度学习方法进行COVID-19检测，”《Expert Syst. Appl.》，第201卷，第116942页，2022年9月，数字对象标识符：10.1016/J.ESWA.2022.116942。[29] C. Häne 等人，“利用多摄像头系统实现自动驾驶汽车的三维视觉感知：标定、建图、定位与障碍物检测”，《Image Vis. Comput.》，第68卷，第14–27页，2017年12月，数字对象标识符：10.1016/J.IMAVIS.2017.07.003。[30] X. Qiu、Z. Liu、M. Zhuang、D. Cheng、C. Zhu 和 X. Zhang，“基于 CNN1 和 CNN2 的磁共振

膝关节半月板损伤磁共振成像诊断及其与计算机断层扫描的比较分析”，《Computer Methods and Programs in Biomedicine》，第106297页，2021年7月，数字对象标识符：10.1016/J.CMPB.2021.106297。[31] C. Liao、C. Wang、J. Bai、L. Lan 和 X. Wu，“用于连续无线胶囊内窥镜图像帧中感兴趣区域配准的深度学习”，《Computer Methods and Programs in Biomedicine》，第208卷，第106189页，2021年9月，数字对象标识符：10.1016/J.CMPB.2021.106189。[32] C. Ren、D. K. Kim 和 D. Jeong，“农业领域深度学习综述：技术及其应用”，《信息处理系统杂志》，第16卷，第5期，第1015–1033页，2020年，数字对象标识符：10.3745/JIPS.04.0187。[33] G. Huang、Z. Liu、L. van der Maaten 和 K. Q. Weinberger，“密集连接卷积网络”，《2017年计算机视觉与模式识别会议论文集——第30届IEEE计算机视觉与模式识别会议》，第2017-January卷，第2261–2269页，2016年8月，访问日期：2021年7月26日。[在线]。可访问：https://arxiv.org/abs/1608.06993v5。[34] K. He、X. Zhang、S. Ren 和 J. Sun，“用于图像识别的深度残差学习”，载于《IEEE计算机学会计算机视觉与模式识别会议论文集》，2016年12月，第2016-December卷，第770–778页，数字对象标识符：10.1109/CVPR.2016.90。[35] K. Simonyan 和 A. Zisserman，“用于大规模图像识别的超深卷积网络”，2015年9月，访问日期：2021年6月17日。[在线]。可访问：http://www.robots.ox.ac.uk/。[36] F. Chollet，“Xception：采用深度可分离卷积的深度学习”，载于《2017年计算机视觉与模式识别会议论文集——第30届IEEE计算机视觉与模式识别会议》，2017年11月，第2017-January卷，第1800–1807页，数字对象标识符：10.1109/CVPR.2017.195。[37] M. Sandler、A. Howard、M. Zhu、A. Zhmoginov 和 L. C. Chen，“MobileNetV2：倒置残差与线性瓶颈”，载于《IEEE计算机学会计算机视觉与模式识别会议论文集》，2018年12月，第4510–4520页，数字对象标识符：10.1109/CVPR.2018.00474。[38] G. Hinton、O. Vinyals 和 J. Dean，“神经网络中的知识蒸馏”，2015年3月，访问日期：2021年7月26日。[在线]。可访问：https://arxiv.org/abs/1503.02531v1。[39] J. Gou、B. Yu、S. J. Maybank 和 D. Tao，“知识蒸馏：综述”，《国际计算机视觉杂志》，第129卷，第6期，第1789–1819页，2020年6月，数字对象标识符：10.1007/s11263-021-01453-z。[40] L. Duran-Lopez、J. P. Dominguez-Morales、J. Corral-Jaime、S. Vicente-Diaz 和 A. Linares-Barranco，“COVID-XNet：用于诊断和定位胸部X光图像中COVID-19的定制深度学习系统”，《应用科学》2020年，第10卷，页面5683，第10卷，第16期，第5683页，2020年8月，数字对象标识符：10.3390/APP10165683。[41] A. I. Khan、J. L. Shah 和 M. M. Bhat，“CoroNet：一种用于检测和诊断的深度神经网络

基于胸部X光图像识别 COVID-19，《Computer Methods and Programs in Biomedicine》，第 196 卷，2020年11月，数字对象标识符：10.1016/J.CMPB.2020.105581。[42] A. T. Sahlol、D. Yousri、A. A. Ewees、M. A. A. Al-qaness、R. Damasevicius 和 M. A. Elaziz，《使用深度特征和分数阶海洋捕食者算法进行 COVID-19 图像分类》，《科学报告》2020 101，第 10 卷，第 1 期，第 1–15 页，2020年9月，数字对象标识符：10.1038/s41598-020-71294-2。[43] A. S. Al-Waisy 等人，《COVID-CheXNet：用于识别胸部X光图像中 COVID-19 病毒的混合深度学习框架》，《软计算》，第 27 卷，第 5 期，第 2657–2672 页，2023年3月，数字对象标识符：10.1007/S00500-020-05424-3/TABLES/8。[44] A. AlMohimeed、H. Saleh、N. El-Rashidy、R. M. A. Saad、S. El-Sappagh 和 S. Mostafa，《基于堆叠集成深度学习，结合胸部X光图像和疾病症状诊断 COVID-19》，《诊断学》2023，第 13 卷，第 1968 页，第 13 卷，第 11 期，第 1968 页，2023年6月，数字对象标识符：10.3390/DIAGNOSTICS13111968。

[45] S. Chakraborty、B. Murali 和 A. K. Mitra，“一种利用胸部X光图像检测 COVID-19 的高效深度学习模型”，《国际环境研究与公共卫生杂志》2022 年，第 19 卷，第 2013 页，第 19 卷，第 4 期，第 2013 页，2022 年 2 月，数字对象标识符：10.3390/IJERPH19042013。

[46] M. Loey、S. El-Sappagh 和 S. Mirjalili，“基于贝叶斯的优化深度学习模型，利用胸部X光图像数据检测 COVID- 19 患者”，Computers in Biology and Medicine，第 142 卷，第 105213 页，2022 年 3 月，数字对象标识符：10.1016/J.COMPBIOMED.2022.105213。[47] L. T. Duong、P. T. Nguyen、L. Iovino 和 M. Flammini，“使用深度神经网络和迁移学习从胸部X 射线和肺部计算机断层扫描图像中自动检测 Covid-19”，《应用软计算》，第 132 卷，第 109851 页，2023 年 1 月，数字对象标识符：10.1016/J.ASOC.2022.109851。[48] S. D. Deb、R. K. Jha、R. Kumar、P. S. Tripathi、Y. Talera 和 M. Kumar，“CoVSeverity-Net：一种从胸部X 射线图像估计 COVID-19 严重程度的高效深度学习模型”，Research on Biomedical Engineering，第 39 卷，第 1 期，第 85–98 页，2023 年 3 月，数字对象标识符：10.1007/S42600-022-00254-8/FIGURES/13。[49] R. Soundrapandiyan、H. Naidu、M. Karuppiah、M. Maheswari 和 R. C. Poonia，“基于人工智能的小波与堆叠深度学习架构，用于从胸部X光图像中检测冠状病毒（COVID-19）”，《计算机与电气工程》，第 108 卷，第 108711 页，2023 年 5 月，数字对象标识符：10.1016/J.COMPELECENG.2023.108711。[50] A. Lasker、M. Ghosh、S. M. Obaidullah、C. Chakraborty 和 K. Roy，“LWSNet——一种用于从 X光图像中区分 Covid-19 和肺炎的新型深度学习架构”，《多媒体工具与应用》，第 82 卷，第 14 期，第 21801–21823 页，2023 年 6 月，数字对象标识符：10.1007/S11042-022-14247-3/TABLES/7。[51] S. Agrawal、V. Honnakasturi、M. Nara 和 N. Patil，“利用深度学习模型和迁移学习从 X 射线图像中检测 COVID-19”，《SN计算机科学》，第 4 卷，第 4 期，第 1–12 页，2023 年 7 月，

数字对象标识符：10.1007/S42979-022-01655-3/FIGURES/6.