

DOI:10.16136/j.joel.2023.10.0428

注意力引导梯形金字塔融合网络识别新冠肺炎X射线影像

葛斌*, 彭曦晨, 孙倩倩, 袁政

(安徽理工大学 计算机科学与工程学院, 安徽 淮南 232000)

摘要: 新型冠状病毒肺炎(corona virus disease 2019, COVID-19)严重影响人类社会和经济的发展, 威胁人类的健康。如何更准确、快速地排查感染病毒的患者, 使用卷积神经网络(convolutional neural network, CNN)的方法识别 COVID-19 胸部 X 射线影像, 完成计算机自动辅助诊断。但是, 由于识别精度不高, 难以准确判断是否感染了 COVID-19。为了提高网络模型对 COVID-19 胸部 X 射线影像的识别性能, 首先提出注意力引导梯形金字塔融合网络(attention steered trapezoid pyramid fusion network, ASTP-Net), 该网络可以附加在不同的 CNN 上, 有效地利用模型中深层与浅层网络的特点; 其次提出注意力引导块(attention steered block, AS Block), 通过通道和空间注意力, 强调通道和空间中的有效语义信息, 弱化无效的干扰信息, 高效地聚合加权信息。最终实验结果表明: 将 ASTPNet 附加在 VGG16/19、ResNet34/50 和 ResNeXt 上, 识别精度有了显著提升; 应用于自建的 COVID-19 数据集, 与其他 CNN 方法相比, ASTP-ResNet34 性能更优, 二分类和三分类准确率分别达到了 98.40% 和 97.10%, 能够精准判断是否感染了 COVID-19。

关键词: 新型冠状病毒肺炎(COVID-19); 卷积神经网络(CNN); 胸部 X 射线影像; 注意力; 融合网络
中图分类号: TP391 **文献标识码:** A **文章编号:** 1005-0086(2023)10-1111-10

Attention steered trapezoid pyramid fusion network for COVID-19 X-ray image recognition

GE Bin*, PENG Xichen, SUN Qianqian, YUAN Zheng

(School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan, Anhui 232000, China)

Abstract: The corona virus disease 2019 (COVID-19) is severely affects the development of society and economy, and threatens human health. In order to solve the problem that how to identify patients infected with the virus more accurately and quickly, convolutional neutral network (CNN) methods are used to identify COVID-19 chest X-ray images. However, due to the low recognition accuracy of CNN, it is difficult to accurately determine whether a patient is infected with COVID-19. In order to improve the recognition performance of the network for COVID-19 chest X-ray images, firstly, the attention steered trapezoid pyramid fusion network (ASTPNet) is proposed. The ASTPNet can be attached to different CNNs. The characteristics of deep and shallow networks in the model are effectively utilized. Secondly, the attention steered block (AS Block) is proposed to aggregate the weighted information more efficiently to emphasize effective semantic information in channels and spaces, and weaken ineffective interference information through channel and spatial attention. The results show that the accuracy is significantly improved after attaching the ASTPNet to VGG16/19, ResNet34/50 and ResNeXt. When applied to the self-built COVID-19 dataset, and compared with other CNN methods, ASTP-ResNet34 has the better

* E-mail: pengxc2021@163.com

收稿日期: 2022-06-08 修订日期: 2022-08-10

基金项目: 国家重点研发计划“智能机器人”重点专项(2020YFB1314103)、国家自然科学基金(3210071479)、安徽省自然科学基金(2108085QF258)和安徽省高等学校自然科学研究项目(KJ2020A0299)资助项目

performance. The accuracy reaches 98.40% (two classes) and 97.10% (three classes). It can accurately determine whether the infection of COVID-19.

Key words: corona virus disease 2019(COVID-19); convolutional neural network (CNN); chest X-ray image; attention; fusion network

0 引 言

新型冠状病毒肺炎(corona virus disease 2019, COVID-19)蔓延全球,世界卫生组织向世界发布了全球卫生警告,截止 2022 年 6 月,累计确诊人数为 5.2 亿,死亡 630 万,死亡率为 1.15%,其病毒的传播速度快、细胞扩散广、活性强、威胁大^[1]。感染病毒最常见的生理反应有咳嗽、发烧、呕吐和呼吸困难等症状,但通过症状描述的方式难以准确判断是否感染了新冠病毒。临床检测 COVID-19 公认的方法有试剂盒检测法和胸部 X 射线扫描等,然而 COVID-19 的潜伏期长达 14 d,试剂盒检测法需要达到一定病毒浓度才有效,导致假阴性率高达 15%左右,并且检测出结果需要 4—6 h,严重的疫情地区对检测试剂盒的需求量极大,难以及时提供大规模的试剂盒^[2]。但是胸部 X 射线检测方法能通过分析胸片准确、实时地检测出疑似患者是否感染了 COVID-19,从而立刻隔离感染患者,节省了大量的宝贵时间^[3],因此这种方法是评估疑似患者的重要途径。

由于胸部 X 射线影像大多是由放射科医生视觉诊断,普通的肺炎和 COVID-19 也有相似的特点,导致大量的图片需要长时间分析,无法保证诊断的实时准确性。近年来,随着深度学习的发展,目前许多研究人员和学者使用卷积神经网络(convolutional neural network, CNN)识别 COVID-19 胸部 X 射线影像,主要研究构建模型完成计算机自动辅助检测,以及如何提高网络的识别性能。MURALIDHARAN 等^[4]认为基于胸部 X 射线影像的人工检测 COVID-19 是一个既费时又费力的繁琐过程,需要采用计算机辅助诊断系统对 COVID-19 疾病进行自动检测,以骨干网络 DenseNet 作为识别模型,对胸部 X 射线影像数据集进行学习、训练;BAGHDADI 等^[5]以 CNN 模型为基础,提出了麻雀搜索的仿生算法优化模型的训练阶段,模仿动物觅食和避开危险的行为,帮助相关工作人员诊断 COVID-19 影像,其特点是响应速度更快并且性能高效;GOUR 等^[6]提出一种堆叠的 CNN 模型,从 VGG 和 Xception 得到预训练权重,加载到自建的堆叠 CNN 上,实现自动识别 COVID-19 影像,缓解数据集类别不均衡的问题,但是堆叠不同的模型导致模型参数过于复杂,迭代周

期过长。OZTURK 等^[7]使用目标检测主流模型 YOLO 的骨干网络 DarkNet 构建出 DarkCovid-Net,共有 17 层卷积层,同时为了平衡数据集加入了普通肺炎的数据样本形成三分类数据集,最后利用可视化热力图方法,发现高响应区域可能是病状区域,能够实现基于热力图评价治疗效果的方法;BARGSHADY 等^[8]提出了 Inception 结合 CycleGan 的网络模型,利用 GAN 和 SSA-CycleGan 获得新数据集以增强原始的训练数据集,方法的局限性在于数据集样本不足,如果有更多的数据样本, CycleGan 生成的样本将会更丰富。OWAIS 等^[9]提出了一种集成设计方法,并行连接两种轻量级网络 MobileNet 和 ShuffleNet,再级联输入 FCNet 中,但是在模型构建中,仅考虑到两种轻量级网络输出的连接,并没有构建模块内部连接。

为了进一步提高模型网络的识别精度,最大限度地利用浅层和深层网络特点,合理运用空间和通道中的重要因素,解决数据集类别不平衡问题,本文做出了如下贡献:

1) 提出了一种注意力引导的梯形金字塔融合网络(attention steered trapezoid pyramid fusion network, ASTPNet),它可以附加在不同的 CNN 模型上,在基础网络上设置多层分支,提取深层和浅层网络的不同特征图,利用深层网络的特点加强浅层网络,复用浅层网络的特征图,提升模型的识别精度。

2) 通过通道和空间注意力方法引导特征融合的多层分支,并且聚合注意力权重,将模型的关注点聚焦到有效的特征,抑制干扰分辨目标的特征。

3) 自建类别均衡的 COVID-19 胸部 X 射线影像数据集,包含 1383 张影像,使用 ASTPNet 结合 VGG16/19、ResNet34/50、ResNeXt 进行训练,实现自动对胸部 X 射线影像进行识别诊断。

1 研究方法

1.1 图像预处理

首先对所有图像进行预处理操作,由于 CNN 中批量输入的图像尺寸相同,然而数据集中所包含的图像分辨率并不统一,有 1024×1024 、 685×615 、 972×741 等,医学放射性图像的分辨率都比较高,现有的 GPU 设备显存不够,处理时间太长。

因此通过改变大小(Resize)方法将所有图像的尺寸转变为 256×256 的特征图,这种方法只改变了图像长宽比,保留了原始图像的特征,再将图片随机缩放裁剪变为 224×224 的大小,以满足特征图输入尺寸的要求,减少输入特征数量和 GPU 运算时间,以默认概率 0.5 对图像进行水平翻转,实现数据增强。实验中使用 Pytorch 的深度学习框架,处理的是以张量表示的图像数据,对每个通道进

行归一化处理,缓解数据之间的差异,使其满足标准正态分布。

1.2 ASTPNet 框架

为了提升网络对 COVID-19 的识别精度,提出 ASTPNet 如图 1 所示。ASTPNet 是基于特征融合和并联空间、通道注意力的网络,该网络由基础模块(base module)、特征图模块(feature module)和注意力引导的梯形金字塔融合模块(ASTP

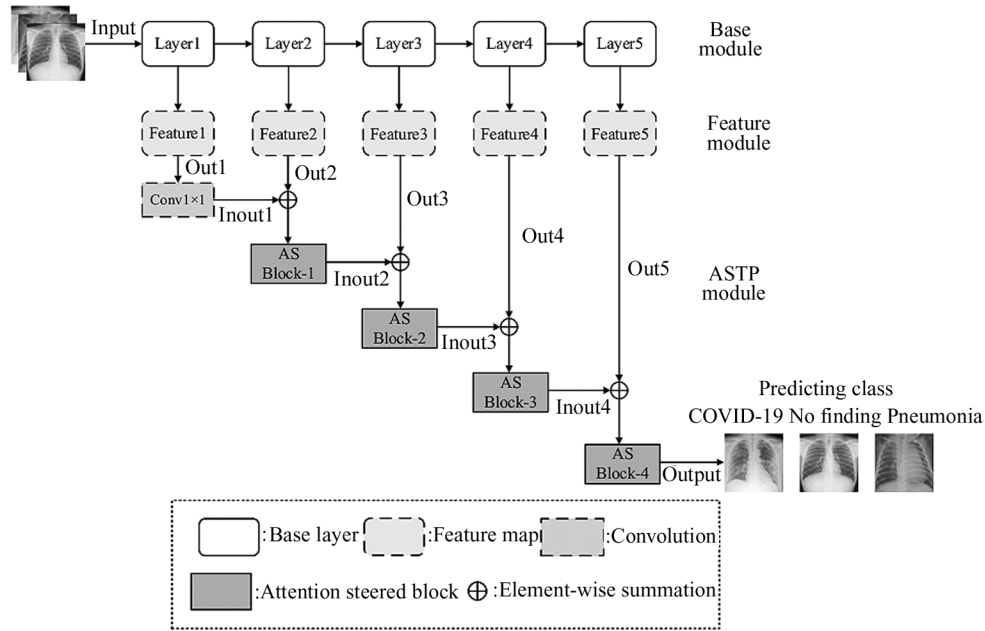


图 1 ASTPNet 网络结构

Fig. 1 Network structure of the ASTPNet

module)所构成。

1.2.1 梯形金字塔融合模块

将预处理后的特征图输入基础网络,将基础网络分成 5 个 Layer 层,在每个 Layer 层上设置 1 个 Layer 分支,从每个 Layer 分支中提取特征图。第 1 层网络的特征图是浅层网络,特征图的分辨率高但是可学习到的语义信息少,使用 1×1 步距为 2 的卷积核改变通道以及特征图的大小,便于与第 2 层的输出进行元素级级联操作。通过 4 次的级联操作,进入分类器,如式(1)所示:

$$F_r = AS(AS(AS(AS(Conv_{1 \times 1}(F_1) \oplus F_2) \oplus F_3) \oplus F_4) \oplus F_5), \quad (1)$$

式中, F_1 为提取出的第 1 张特征图, AS 是注意力引导模块, F_r 是输出的特征图, $\oplus$ 是元素级级联操作。

分类器中没有使用传统的随机概率失活结合全连接层的方法,而是使用全局平均池化(global average pooling, GAP)完成分类器任务,它本质上是一

种正则化方法,能够不固定地输入图像的尺寸并且抑制过拟合现象,还能够大大降低分类器的参数量。

1.2.2 注意力引导模块

由于深层特征图包含的语义信息量更大,使用高层语义加强低层语义输出^[10],为了进一步提取特征,提出了注意力引导块(attention steered block, AS Block),以图 2 的 AS Block-1 为例。

经过 1×1 卷积得到的 Inout1 与 Out2 进行元素级相加,将特征图分别经过通道注意力(channel attention)和空间注意力(spatial attention)。通道注意力的主要目的在于加强通道特征之间的有效信息,排斥对识别无效的语义信息,消除通道中干扰因素的影响^[11]。使用 GAP 计算每张特征图所有像素点的平均值,得到包含每一张特征图的权重值,赋予模型分辨不同特征的能力,如式(2)所示:

$$y = \text{Average}(\sum_{i=1}^n F_i), \quad (2)$$

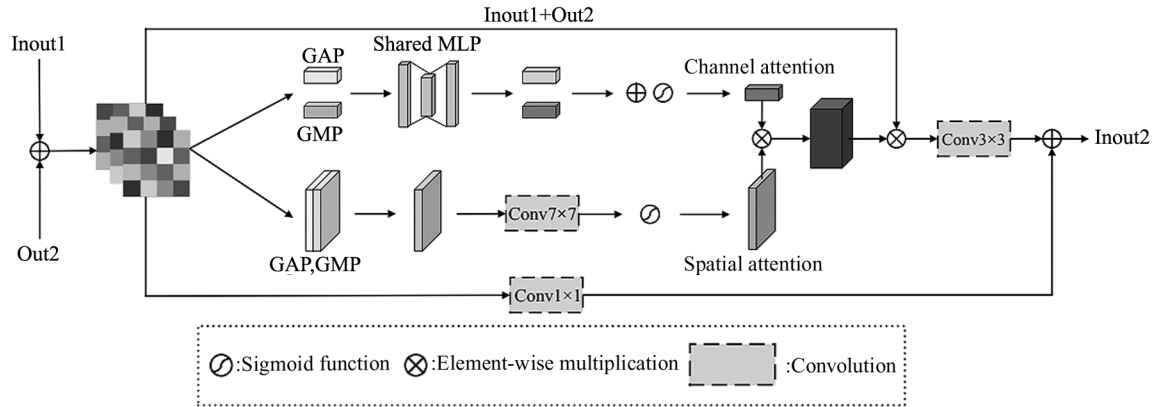


图2 AS Block-1 模块

Fig. 2 The AS Block-1 module

式中, F_i 为第 i 张特征图, n 为特征图的总数, y 为均值输出。

使用全局最大池化方法(global max pooling, GMP), 获取每一张特征图中最重要的区域, 即每一张特征图权重的最大值, 如式(3)所示:

$$x = \text{Max} \left(\sum_{i=1}^n F_i \right). \quad (3)$$

通道注意力分为压缩操作和激励操作, GAP 和 GMP 为压缩操作, 获得平均全局权重信息和最大的权重信息, 式(4)为压缩操作:

$$\mathbf{Z}_c = F_s(\mathbf{u}_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \mathbf{u}_c, \quad (4)$$

式中, F_s 为压缩操作, $\mathbf{u}_c$ 是通道索引为 c 的特征图张量, $H \times W$ 为特征图的长和宽, $\mathbf{Z}$ 为压缩后的特征向量。

Shared MLP 层由两个 1×1 的卷积核和一个 ReLU 激活函数组成。特征向量 $\mathbf{Z}$ 通过激励操作, 并联相加两种权重向量, 再使用 Sigmoid 激活函数将权重向量非线性归一化至 $[0, 1]$ 中, 得到既包含特征的平均值也没有丢失最重要区域的权重向量 $\mathbf{S}$, 如式(5)~(7)所示:

$$\sigma(x) = \frac{1}{e^{-x} + 1}, \quad (5)$$

$$\delta(x) = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases}, \quad (6)$$

$$\mathbf{S} = F_e(\mathbf{Z}, \mathbf{A}, \mathbf{M}) = \sigma(\mathbf{A}_2 \delta(\mathbf{A}_1 \mathbf{Z}) \oplus \mathbf{M}_2 \delta(\mathbf{M}_1 \mathbf{Z})), \quad (7)$$

式中, F_e 为激励操作, $\mathbf{S}$ 是输出权重向量, $\mathbf{Z}$ 为输入的特征向量, $\mathbf{A}_1$ 、 $\mathbf{A}_2$ 、 $\mathbf{M}_1$ 和 $\mathbf{M}_2$ 是经过 GAP 和 GMP 并用卷积降维和升维后的权重值, δ 是 ReLU 激活函数, σ 是 Sigmoid 激活函数, $\oplus$ 是并联元素级相加。

对于空间注意力, 同样使用两种全局提取特征

的方法, 提取后的权重向量经过 7×7 的卷积核和 Sigmoid 激活函数得到代表空间关系的非线性权重特征值, 如式(8)所示, 主要是用于寻找像素点之间的空间关系, 通过空间权重关系判断哪一部分是识别相关的重要区域, 并强调这些重要的区域, 抑制不关键的空间区域。

$$F_c = \sigma(\text{Conv}_{7 \times 7}(\text{concat}(\mathbf{W}_A; \mathbf{W}_M))), \quad (8)$$

式中, F_c 为提取空间注意力权重操作, concat 是维度为 1 的连接, $\mathbf{W}_A$ 和 $\mathbf{W}_M$ 是空间注意力中经过 GAP 和 GMP 得到的权重特征图。

最后将空间和通道注意力的特征权重值元素级相乘得到一个尺寸为 $C \times H \times W$ 的注意力特征块, 注意力特征块与原特征图加权得到包含注意力特征权重的特征图, 重新对原特征图分配通道和空间权重, 完成特征图重标定。使用步距为 2、尺寸为 3×3 的卷积核, 调整输出特征图尺寸与下一层特征图级联, 其次聚合注意力引导的语义信息确定特征提取方向, 最大限度强化注意力作用^[12]。在捷径分支中使用步距为 2、尺寸为 1×1 的卷积核对特征图进行降维, 减少了计算量和参数量, 并防止了梯度消失和梯度爆炸, AS Block 处理过程的伪代码如下所示:

输入: 分支 $\text{Inout}(i)$ 和 $\text{Out}(i+1)$;

输出: $\text{Inout}(i+1)$;

1) 利用元素级相加方法级联 $\text{Inout}(i)$ 和 $\text{Out}(i+1)$, 得到 n 张新的特征图;

2) 将 n 张特征图通过式(2)和式(3)处理, 获得平均和最大的全局权重信息;

3) 将两种全局权重信息输入 Shared MLP 中, 通过两个 1×1 的卷积和式(4), 完成压缩操作 F_s ;

4) 采用式(7)激励操作 F_e , 得到通道注意力输出;

5) 利用式(8)的方法,连接步骤2)输出的两种全局权重信息,再通过 7×7 的卷积核和 σ 激活函数处理,完成空间注意力操作 F_c ;

6) 将步骤4)和步骤5)的输出进行元素级相乘,得到注意力特征块;

7) 将注意力特征块加权步骤1)输出的原始特征图,完成特征图重标定操作;

8) 使用 3×3 的卷积核聚合处理后,与捷径分支元素级相加,得到最终的输出 $Inout(i+1)$ 。

2 实验过程与分析

2.1 数据集

数据集来源为 COHEN 的团队建立的一个公共开放的 COVID-19 胸部 X 射线影像数据集(<https://github.com/ieee8023/COVID-chestxray-dataset>),从公共资源和医院收集到 125 张感染 COVID-19 的影像,如图 3(a)所示,以及 WANG 等^[13]建立的胸部 X 射线数据库,获取了 500 张未发现症状的影像,COVID-19-625 二分类数据集共计 625 张,与 OZ-TURK 所使用的数据集相同,如图 3(b)所示。但是考虑到两类的数据量不平衡,部分影像质量不高,添加了 Kaggle 数据集中 258 张感染 COVID-19 和 500 张如图 3(c)所示的普通肺炎的影像,自建 COVID-19-1383 三分类数据集共包含感染 COVID-19 影像 383 张,未发现症状影像 500 张,普通肺炎影像 500 张,共 1383 张胸部 X 射线影像。按照 4:1 划分训练集和测试集,分别对两种 COVID-19 数据集进行训练和测试。此外,选取 CIFAR10/100 公共数据集验证模型的鲁棒性和普适性,分别为 6 万张图片的 10 分类和 100 分类。

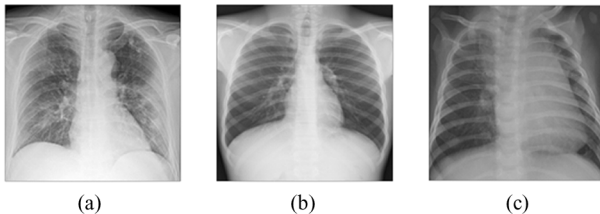


图3 胸部X射线影像:(a)感染COVID-19;
(b)未发现;(c)普通肺炎

Fig.3 Chest X-ray image:(a) Infection with COVID-19;
(b) No finding; (c) Common pneumonia

2.2 实验环境与评价指标

本文的实验环境为 Ubuntu 18.04 操作系统, NVIDIA GeForce RTX3090 显卡, Intel Xeon Plati-

num 8260C CPU, CUDA 版本为 11.1, cuDNN 为 8.0.5, Pytorch1.8.1 深度学习框架。

选取 Top-1、Top-5 准确率(Accuracy)、召回率(Recall)、查准率(Precision)、特异度(Specificity)和 F1 得分(F1-score)作为评价模型性能的指标,其中准确率作为评价模型性能的最重要指标,它们的定义为:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}, \quad (9)$$

$$Precision = \frac{TP}{TP + FP}, \quad (10)$$

$$Recall = \frac{TP}{TP + FN}, \quad (11)$$

$$Specificity = \frac{TN}{TN + FP}, \quad (12)$$

$$F1-score = \frac{2 \times Precision \times Recall}{Precision + Recall}. \quad (13)$$

以正样本为未发现病状、负样本为 COVID-19 的数据集为例,解释 TP、TN、FP、FN。TP(true positive)是实际为未发现病状,被分类为未发现病状的数量;TN(true negative)是实际为未发现病状,被分类为感染 COVID-19 的数量;FP(false positive)是实际为感染 COVID-19,被分类为未发现病状的数量;FN(false negative)是实际为感染 COVID-19,被分类为感染 COVID-19 的数量。

2.3 实验结果与分析

2.3.1 公共数据集实验

在 CIFAR10/100 公共数据集上验证模型性能的鲁棒性和普适性,选择 5 种主流的 CNN 模型,探究 ASTPNet 在不同的模型上的有效性。分别对 VGG16/19、ResNet34/50 和 ResNeXt 添加 ASTPNet 前后的网络进行训练,实验参数设置为:迭代周期 Epoch 为 300,优化器选择随机梯度下降,使用交叉熵损失函数,学习率采用动态学习率,前 100 次学习率为 0.1,之后每 50 次迭代减小为之前的 1/10。在 CIFAR100 数据集中,由于 CIFAR100 数据集共有 6 万张图片,分为 100 个类别,识别类别多、难度大,所以选择 Top-1 和 Top-5 准确率作为评价指标。Top-5 准确率是在测试图片的 100 个类别中,选取前 5 个最大的类别概率,如果标签类别在这 5 个中,那么判定为正确,同理,Top-1 准确率就是最大概率的类别是否与标签相同。表 1 显示了加入 ASTPNet 的模型与原模型的实验对比结果,相较于原网络模型,识别准确率都有 1%—2% 的提升,这表明所提出的 ASTPNet 在不同的模型上都有很好的表现,其中

ASTPNet-ResNeXt 的模型性能最好,在 CIFAR10 和 100 的 Top-1 和 Top-5 准确率为 95.56 % 和 80.44 %、95.32 %。

表 1 CIFAR10/100 数据集改进模型对比结果

Tab. 1 Comparison results of improvement models in CIFAR10/100 dataset

Model	CIFAR10		CIFAR100	
	Top-1/%	Top-1/%	Top-5/%	Top-5/%
VGG16	92.64	74.26	91.75	
ASTPNet-VGG16	93.72	74.57	92.32	
VGG19	92.31	72.14	89.47	
ASTPNet-VGG19	93.38	74.07	91.62	
ResNet34	94.21	77.79	93.36	
ASTPNet-ResNet34	95.13	78.05	94.19	
ResNet50	94.09	77.62	94.30	
ASTPNet-ResNet50	94.29	79.44	95.15	
ResNeXt	94.82	78.84	94.85	
ASTPNet-ResNeXt	95.56	80.44	95.32	

2.3.2 新冠肺炎数据集实验

上述实验验证了 ASTPNet 的有效性和普适性,使用 ASTPNet-VGG16/19、ASTPNet-ResNet34/50、ASTPNet-ResNeXt 网络模型,对 COVID-19-625 二分类数据集和自建三分类数据集 COVID-19-1383 进行训练及测试,其中 COVID-19-625 二分类数据集包含感染 COVID-19 和未发现病状 2 个类别,COV-

ID-19-1383 三分类数据集包含感染 COVID-19、未发现病状和普通肺炎 3 个类别。实验迭代周期设置为 200 次,选择 Adam 优化器,学习率设置为 0.000 1,使用交叉熵损失函数。

图 4 展示了 ASTPNet-ResNet34 在 COVID-19-1383 三分类数据集的训练和测试迭代曲线。网络模型训练是一个从获取权重参数,再迭代更新,直至达到稳定状态的过程。在训练前网络模型的各层中不存在权重参数,训练初期首先学习一个新批次的数据样本得到初始的权重参数,然后多次迭代其他批次的数据样本,直到学习过所有的数据样本,迭代学习数据集的所有样本是一个周期,此时权重参数变化幅度较大,数据权重不稳定,所以在训练初期测试曲线表现为震荡的状态。然而训练了多个周期后,权重参数趋于稳定状态,对样本的识别率逐渐提升,损失值逐渐减小,所以在训练后期曲线起伏波动减小,训练损失曲线处于滑梯式稳定下降状态。直到 160 次迭代周期之后,损失值不再减小,准确率不再增长,损失值和准确率在收敛值附近浮动,权重参数的数据表现已经达到稳定状态。网络模型在自建数据集的整个训练过程表现正常,体现出自建数据集和网络模型的合理性。

图 5(a)显示了结合 ASTPNet 的网络模型在 COVID-19-625 二分类数据集的 Accuracy、Recall、

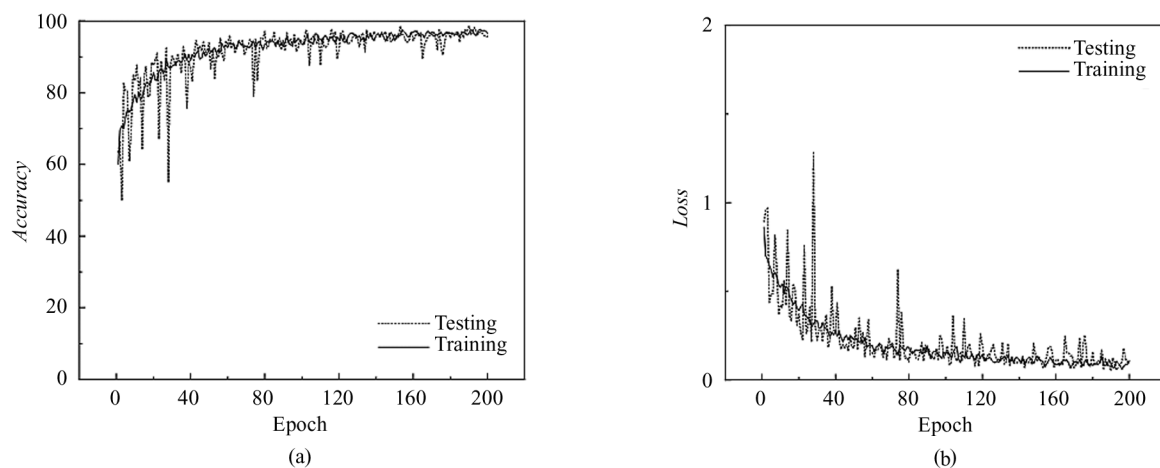


图 4 ASTPNet-ResNet34 训练和测试迭代曲线:(a) 准确率;(b) 损失

Fig. 4 ASTPNet-ResNet34 training and testing iterative curve:(a) Accuracy;(b) Loss

Precision、Specificity 和 F1-score 结果。ASTPNet-ResNet 34 和 ASTP-ResNeXt 准确率达到 98.40 %,已经能够准确地识别感染 COVID-19 的胸部 X 射线影像。图 5(b)显示了结合 ASTPNet 的网络模型在 COVID-19-1383 三分类数据集上的评价指

标结果,ASTPNet-ResNet34 准确率为 97.10 %。ASTPNet-VGG16/19 是卷积层传统堆叠的模型,另外 3 种都是残差网络。由综合结果可见残差网络在胸部 X 射线影像上的识别效果明显优于卷积层传统堆叠的模型。虽然在二分类数据集上 ASTPNet-

ResNeXt 表现与 ASTPNet-ResNet34 大致相同,并且 ASTPNet-ResNeXt 网络层数更深,但在 CI-FAR10/100 上的表现优于 ASTPNet-ResNet34,然而在三分类数据集上两者准确率相差 1.45%。此

外,ASTPNet-ResNet50 与 ASTPNet-ResNet34 相比,在二分类和三分类上准确率相差 0.80% 和 3.26%,两者层数相差 16 层。实验结果表明:1) 残差网络更适合识别胸部 X 射线影像;2) 层数过深的

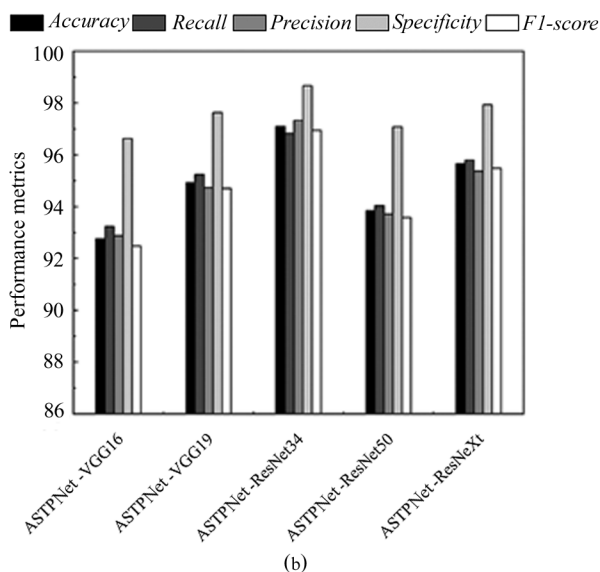
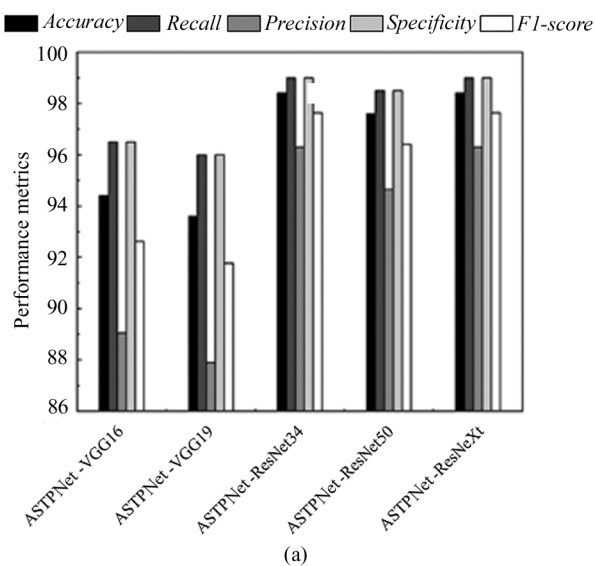


图 5 COVID-19 数据集实验结果:(a) COVID-19-625 数据集实验结果;(b) COVID-19-1383 数据集实验结果

Fig. 5 The experimental results in COVID-19:(a) The experimental results in COVID-19-625;

(b) The experimental results in COVID-19-1383

残差网络并不适用于识别胸部 X 射线影像。

2.3.3 混淆矩阵

由于评价指标数据难以全面体现模型性能,所以使用混淆矩阵能够更直观地分析 TP 、 TN 、 FP 、 FN 和每种类别的准确率,更全面地评估模型的性能。选择效果较好的 ASTPNet-ResNet34 和 ASTPNet-ResNeXt 遍历测试集构建混淆矩阵,如图 6 所

示,行代表真实值,列代表模型的预测值。ASTPNet-ResNet34 对 COVID-19 和普通肺炎类别的识别错误率为 0%,ASTPNet-ResNeXt 对 COVID-19 类别的识别准确率为 97.37%,对普通肺炎类别的识别准确率为 98%,两种模型对未发现病状类别识别准确率为 92%,说明模型对不同的肺炎类型可以准确识别,已经能够实现对疑似患者的排查。

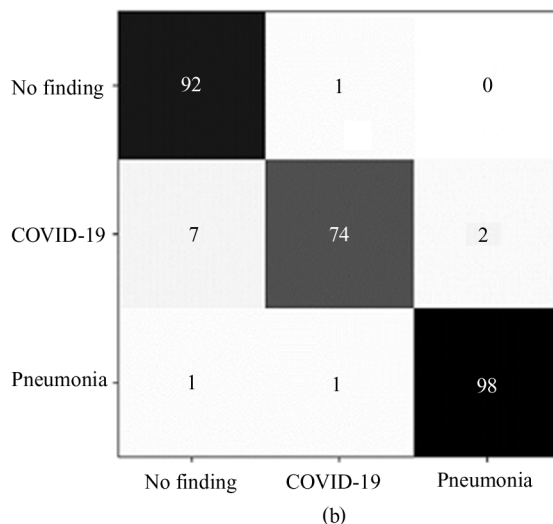
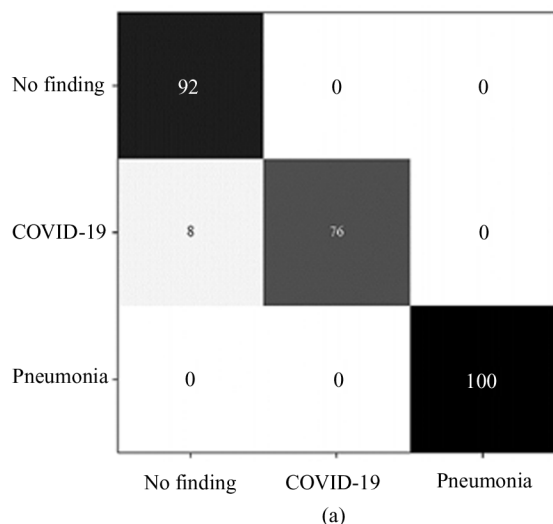


图 6 混淆矩阵:(a) ASTPNet-ResNet34 混淆矩阵;(b) ASTPNet-ResNeXt 混淆矩阵

Fig. 6 Confusion matrix:(a) Confusion matrix of the ASTPNet-ResNet34;(b) Confusion matrix of the ASTPNet-ResNeXt

2.3.4 可视化分析

最后利用 Grad-Cam 方法^[14]绘制可视化图像对比 ASTPNet-ResNet34 和原网络模型。Grad-Cam 能够体现出 CNN 的可解释性,探究网络对感染 COVID-19 胸部 X 射线影像的关注区域,网络根据关注区域识别目标,选择图7进行可视化识别和分析。

对图8进行识别,网络的识别结果为正确,将网



图 7 感染 COVID-19 胸部 X 射线影像

Fig. 7 Chest X-ray image of infected with COVID-19

络结构导入 Grad-Cam 中,分别输出如图 8 的(a) Feature2、(b) Feature3、(c) Feature4、(d) Feature5,与 Feature module 结构对应的(e) AS Block-1、(f) AS Block-2、(g) AS Block-3、(h) AS Block-4。浅层网络输出的(a)、(b)、(e)、(f)关注的高响应区域集中在胸部轮廓和肺部纹理上,表现出浅层网络还不能够分辨 COVID-19 的病变位置。由于 COVID-19 的感染,在影像上会出现双肺的磨玻璃样阴影和胸腔积液等特征,深层网络输出的(c)、(d)、(g)、(h)高响应区域集中在肺部的一些病变区域,(d)、(h)更为明显,通过对这些病变区域识别,能够准确判断是否感染 COVID-19。

ASTP module 中的(e) (f)与 Feature module 中的(a)、(b)相比,(e)、(f)已经能够细致地标出纹理特征,而(a)、(b)并不具备这样的能力,(e)、(f)效果明显优于(a)、(b)。(c)、(d)与(g)、(h)相比,(g)已经能够模糊地辨别出一些病变位置,而(c)只能分辨纹理特征,(h)中的高响应区域多于(d),病变区域的标注更加集中。通过 Grad-Cam 方法输出可视化图片,验证出 ASTPNet 对原网络模型的提升效果显著,高响应区域能够初步标注病变区域。

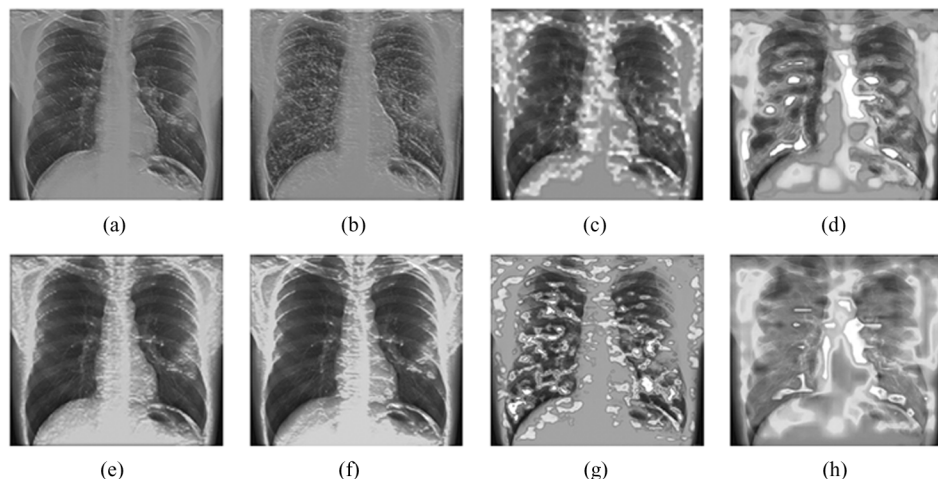


图 8 Grad-Cam 可视化图像:(a) Feature2 输出;(b) Feature3 输出;
(c) Feature4 输出;(d) Feature5 输出;(e) AS Block-1 输出;
(f) AS Block-2 输出;(g) AS Block-3 输出;(h) AS Block-4 输出

Fig. 8 Grad-Cam visual images:(a) Output of the Feature2;
(b) Output of the Feature3;(c) Output of the Feature4;
(d) Output of the Feature5;(e) Output of the AS Block-1;
(f) Output of the AS Block-2;(g) Output of the AS Block-3;(h) Output of the AS Block-4

2.3.5 与其他方法的比较

将识别效果最好的 ASTPNet-ResNet34 与其他

使用 CNN 识别胸部 X 射线影像的方法进行对比,如表 2 所示,与文献中的其他方法相比,本文所提出的

表 2 与其他方法识别准确率的比较
Tab. 2 Comparison of recognition accuracy with other methods

Study	Model used	Number of cases	Accuracy/%
GAYATHRI et al. ^[15]	FFNN	504COVID-19(+)542No Finding(-)	95.78
ISMAEL et al. ^[16]	ResNet50+SVM	180COVID-19(+)200No Finding(-)	95.79
PANDIT et al. ^[17]	VGG-16	224COVID-19(+)504No Finding(-)	96.00
		224COVID-19(+)504No Finding(-)700Pneumonia	92.53
LI et al. ^[18]	EfficientNet	231COVID-19(+)1583No Finding(-)	97.23
		231COVID-19(+)1583No Finding(-)4007Pneumonia	93.49
OZTURK et al. ^[7]	DarkCovidNet	125COVID-19(+)500No Finding(-)	98.08
		125COVID-19(+)500 No Finding(-)500Pneumonia	87.02
Proposed	ASTPNet-ResNet34	125COVID-19(+)500No Finding(-)	98.40
		383COVID-19(+)500No Finding(-)500Pneumonia	97.10

方法识别精度更高。

3 结 论

本文提出了将 ASTPNet 作为附加模块添加在主流的 CNN 模型上。采用通道和空间注意力引导特征融合的方式,强调对识别目标有效的特征,抑制不利于识别的特征,并且采用元素级级联的方法融合浅层与深层网络提取出的特征图,依靠深层网络语义性强的特点加强浅层网络,复用浅层网络,最大限度地利用每一层输出的特征图语义信息,提高了网络模型的识别精度。此外,将所提出的网络应用在胸部 X 射线影像上识别 COVID-19、普通肺炎和未发现病状 3 个类别,实验证明了 ASTPNet 的普适性和先进性,能够实现计算机自动辅助诊断。在之后的工作中,将继续对 COVID-19 胸部 X 射线的放射性影像进行研究,为医学临床做出一定的贡献,进一步优化网络的结构,提升网络模型的诊断精度。

参考文献:

[1] ZU Z Y, JIANG M D, XU P P, et al. Coronavirus disease 2019 (COVID-19): a perspective from China[J]. Radiology, 2020, 296(2): E15-E25.

[2] BAHGAT W M, BALAHA H M, ABDULAZEEM Y, et al. An optimized transfer learning-based approach for automatic diagnosis of COVID-19 from chest X-ray images[J]. PeerJ Computer Science, 2021, 7: e555.

[3] BERNHEIM A, MEI X, HUANG M, et al. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection[J]. Radiology, 2020, 295(3): 200463.

[4] MURALIDHARAN N, GUPTA S, PRUSTY M R, et al. De-

tection of COVID19 from X-ray images using multiscale deep convolutional neural network[J]. Applied Soft Computing, 2022, 119: 108610.

[5] BAGHDADI N A, MALKI A, ABDELALIE M S F, et al. An automated diagnosis and classification of COVID-19 from chest CT images using a transfer learning-based convolutional neural network[J]. Computers in Biology and Medicine, 2022, 144: 105383.

[6] GOUR M, JAIN S. Automated COVID-19 detection from X-ray and CT images with stacked ensemble convolutional neural network[J]. Biocybernetics and Biomedical Engineering, 2022, 42(1): 27-41.

[7] 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.

[8] BARGSHADY G, ZHOU X, BARUA P D, et al. Application of CycleGAN and transfer learning techniques for automated detection of COVID-19 using X-ray images[J]. Pattern Recognition Letters, 2022, 153: 67-74.

[9] OWAIS M, YOON H S, MAHMOOD T, et al. Light-weighted ensemble network with multilevel activation visualization for robust diagnosis of COVID19 pneumonia from large-scale chest radiographic database[J]. Applied Soft Computing, 2021, 108: 107490.

[10] AHMAD Z, JINDAL R, MUKUNTHA N S, et al. Multi-modality helps in crisis management: An attention-based deep learning approach of leveraging text for image classification[J]. Expert Systems with Applications, 2022, 195: 116626.

[11] ZHOU W, CHEN P. A deep attention mechanism method

- for maritime salient ship detection in complex sea background[J]. Optoelectronics Letters, 2021, 17(7): 438-443.
- [12] ZHANG Z, XUE W, FU G. Unsupervised image-to-image translation by semantics consistency and self-attention[J]. Optoelectronics Letters, 2022, 18(3): 175-180.
- [13] WANG X, PENG Y, LU L, et al. ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, USA. New York: IEEE, 2017: 2097-2106.
- [14] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 618-626.
- [15] GAYATHRI J L, ABRAHAM B, SUJARANI M S, et al. A computer-aided diagnosis system for the classification of COVID-19 and non-COVID-19 pneumonia on chest X-ray images by integrating CNN with sparse autoencoder and feed forward neural network[J]. Computers in Biology and Medicine, 2022, 141: 105134.
- [16] ISMAEL A M, SENGÜR A. Deep learning approaches for COVID-19 detection based on chest X-ray images[J]. Expert Systems with Applications, 2021, 164: 114054.
- [17] PANDIT M K, BANDAY S A, NAAZ R, et al. Automatic detection of COVID-19 from chest radiographs using deep learning[J]. Radiography, 2021, 27(2): 483-489.
- [18] LI J, ZHAO G, TAO Y, et al. Multi-task contrastive learning for automatic CT and X-ray diagnosis of COVID-19[J]. Pattern Recognition, 2021, 114: 107848.

作者简介:

葛 斌 (1974—), 男, 博士, 教授, 主要研究方向为物联网技术.