

## 基于改进 Inception-ResNet-v2 的儿童肺炎识别与分类

杨俊豪, 马志庆\*, 魏国辉, 赵爽

山东中医药大学智能与信息工程学院, 山东 济南 250355

**摘要** 针对儿童肺炎图像难以准确诊断的问题, 提出一种基于改进 Inception-ResNet-v2 的分类识别方法, 以提高对不同儿童肺炎图像的识别准确率。以 Inception-ResNet-v2 为基础网络, 引入多尺度通道注意力模块, 促进网络识别和检测极端尺度变化下的目标。加大网络 stem 层的卷积核大小, 增大模型初期的有效感受野。为避免模型出现过拟合, 减少了激活函数的使用, 并使用 SiLU 激活函数来代替 ReLU 激活函数。针对数据集 Chest X-ray 中数据量较少的问题, 对输入图像进行一定角度的旋转和随机水平的翻转以增强原始数据。实验结果表明, 所提方法在儿童肺炎数据的二分类中准确率达到 97.9%, 三分类中准确率达到 85.8%, 证明该方法能够有效提高儿童肺炎的识别精度。

**关键词** 图像处理; 儿童肺炎; 多尺度通道注意力模块; 感受野; 激活函数

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP221774

## Recognition and Classification of Childhood Pneumonia Based on Improved Inception-ResNet-v2

Yang Junhao, Ma Zhiqing\*, Wei Guohui, Zhao Shuang

College of Intelligence and Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan 250355, Shandong, China

**Abstract** To address the issue of difficulty in accurately diagnosing children's pneumonia images, a classification and recognition method based on improved Inception-ResNet-v2 is proposed to improve the recognition accuracy of various types of children's pneumonia images. A multiscale channel attention module based on Inception-ResNet-v2 is introduced to promote network recognition and detection of targets under extreme scale changes. The size of the network stem layer's convolution kernel and the effective receptive field are increased at the start of the model. To avoid overfitting the model, the activation function is reduced in size, and the SiLU activation function is used instead of the ReLU activation function. To address the issue of less amount of data in the Chest X-ray dataset, the input image is rotated at a specific angle and flipped horizontally at random to improve the original data. The experimental results show that the proposed method has a 97.9% accuracy in the second classification of children's pneumonia data and an 85.8% accuracy in the third classification, demonstrating that the method can effectively improve the recognition accuracy of children's pneumonia.

**Key words** image processing; pneumonia in children; multiscale channel; attention module receptive field; activation function

## 1 引言

在不发达和发展中国家, 由于城市的快速发展, 空气污染正在以惊人的速度出现, 常见的肺部疾病如肺炎、肺结节、肺癌、肺气肿等的发病率是巨大的。其中, 肺炎是一种影响肺部的急性呼吸道感染, 是一种致命

的疾病。而肺结节为大多数肺癌的早期症状<sup>[1]</sup>。肺气肿是慢性、阻塞性、经常进行性和破坏性肺部疾病的一种表现<sup>[2]</sup>。此外, 胸腔积液、胸部肿块和气胸等也是某些疾病的症状。由于儿童抵抗力较低, 在感染肺炎的患者中儿童占很大一部分比重。据估计每年约有 70 万儿童死于肺炎<sup>[3]</sup>。而在患病的早期, 儿童如果能

收稿日期: 2022-06-06; 修回日期: 2022-06-09; 录用日期: 2022-08-29; 网络首发日期: 2022-09-09

基金项目: 国家自然科学基金(61702087)、山东省研究生教育质量提升计划(SDYAL20050)、山东省医药卫生科技发展计划(202109040649)、党建+新工科+工程教育认证背景下生物医学工程专业研究生创新创业人才培养模式研究(SDYJG1943)、山东省教育厅“混合式”教学模式教学管理研究(ZYZ2019009)

通信作者: \*mzhq126@163.com

得到有效的治疗将有机会痊愈。胸部 X 光片是诊断肺炎的常见临床方法<sup>[4]</sup>。然而,从胸部 X 光图像诊断肺炎是一项艰巨的任务,即使是专业放射科医生也面临着挑战。儿童肺炎在 X 光中的表现往往不清楚,其中可能会混淆有其他疾病。因此有必要借助计算机辅助诊断来帮助放射科医生从 X 光图像中诊断肺炎。

近年来,随着深度学习的发展,卷积神经网络(CNN)在图像处理方面取得了巨大的成功<sup>[5]</sup>。Saraiva 等<sup>[6]</sup>使用多层感知器和神经网络对肺炎进行分类和检测,并采用 k-折交叉验证,取得了不错的结果。在 CNN 中,为了进一步提高模型的精度,通常做法是加深网络层数,增大数据集。Rajpurkar 等<sup>[7]</sup>提出一种全新的深层次的卷积神经网络 CheXNet,并将其在目前最大的 X 光片存储库 Chest X-ray 14<sup>[8]</sup>数据集上训练、验证,在与多名放射科医生对比中发现,CheXNet 在检测过程中的表现优于放射科医生。随着越来越多的机器学习应用场景出现,由于监督学习需要大量的数据标注,迁移学习受到了越来越多的关注。Ayan 等<sup>[9]</sup>使用 VGG 和 Xception 神经网络模型通过迁移学习和微调在儿童肺炎数据集上进行二分类时分别取得 87% 和 82% 的准确率。Chouhan 等<sup>[10]</sup>采用迁移学习的方法将 AlexNet、DenseNet121、Inception-v3、GoogleNet、ResNet18 模型在 ImageNet 数据集上预训练提取特征,并将特征传到各个模型的分层器,在 Chest X-ray 数据集上的准确率达到 96.39%。

然而以上方法均属于肺炎的二分类方法,虽然取得的结果较好,但并没有深入具体的肺炎分类任务中去,在临床诊断中很难取得针对性效果。肺炎主要有三种类型,由于肺炎图像特点,不同类型肺炎差异较小,因此利用深度学习的方法对其准确分类具有一定挑战。

目前,肺炎三分类检测对临床有积极的辅助作用。Barhoom 等<sup>[11]</sup>利用基于 VGG 架构的自定义模型实现了对肺炎的三分类任务,并取得了较好的结果,但由于模型参数量较大,训练过程较为缓慢。Acharya 等<sup>[12]</sup>利用 Deep Simese CNN 架构解决肺炎的多分类问题,并在 kaggle 数据上进行了验证。Rahman 等<sup>[13]</sup>对 4 种不同预训练的卷积神经网络用迁移学习的方法在 kaggle 数据集上进行肺炎三分类,取得一定的进步。

针对上述问题,本文提出一种改进的 Inception-ResNet-v2<sup>[14]</sup>分类方法对儿童肺炎进行准确分类。该方法拥有更好的泛化性和鲁棒性,能更加准确识别儿童肺炎的类型。

## 2 所提方法

在网络中引入多尺度通道注意力模块(MS-CAM)<sup>[15]</sup>,将初始特征融合并与另一个注意力模块交替集成,能够使网络动态地、自适应地以上下文规模感知的方式融合接收到的特征,提高模型的精度<sup>[16]</sup>,从而

关注不同类型儿童肺炎图像的差异性,并抑制其他冗余信息带来的影响<sup>[17]</sup>。提高网络前期卷积核大小,增大网络的有效感受野。通过减少激活函数的使用来防止网络过拟合。通过对儿童肺炎图像数据集进行数据预处理<sup>[18]</sup>和数据增强等操作来增强数据量,以帮助模型对图像进行特征提取和分类。

### 2.1 Inception-ResNet-v2

Inception 网络最早是由 Google 在 2014 年提出 GoogleNet<sup>[19]</sup>后开始的。最开始的 Inception-v1 引用不同尺寸的卷积核来捕捉不同尺度的特征,去除网络最后一层全连接层,并将其改为全局平均池化层,大大减少参数量、加快训练速度的同时也达到了减轻过拟合的效果。Inception-v2 提出了著名的 batch normalization(BN)<sup>[20]</sup>层,BN 层加快了模型的训练速度和收敛速度。Inception-v3<sup>[21]</sup>将二维卷积拆分为两个较小的一维卷积,减少了参数量,起到了扩展非线性变换的作用。Inception-v4 比起之前的网络使用更加复杂的 stem 模块,stem 模块引用并行结构、不对称卷积结构,可以在信息损失足够小的情况下降低计算量。Inception-ResNet 将 Inception 和 ResNet 结合,使用 Inception 结构来拟合残差结构,有 Inception-ResNet-v1 和 Inception-ResNet-v2 两个版本。Inception-ResNet-v1 在 Inception-v3 上加入 ResNet,Inception-ResNet-v2 在 Inception-v4 上加入 ResNet 和相应的残差网络,网络收敛速度更快。Inception-ResNet-v2 由 stem、Inception-ResNet-A、Reduction-A、Inception-ResNet-B、Reduction-B、Inception-ResNet-C 等模块堆叠而来。Inception-ResNet-v2 既可以加速训练防止梯度弥散,又能获得同一网络层稀疏或非稀疏特征,其结构如图 1 所示。

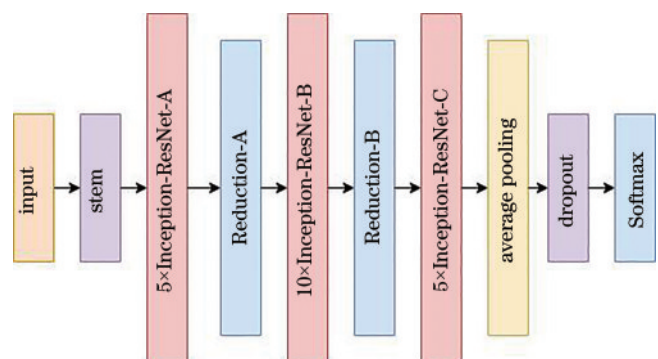


图 1 Inception-ResNet-v2 模型

Fig. 1 Inception-ResNet-v2 model

### 2.2 注意力机制

注意力机制是用来自动学习和计算输入数据对输出数据的贡献大小的,其本质就是定位感兴趣的信息,抑制无用信息。由于儿童肺炎数据集的噪声比较多,儿童病毒性肺炎、细菌性肺炎差异不明显,这就需要网络本身能够提取出更加有用的信息,从而做出更加准

确的判断。所提方法在 Inception-ResNet 层和 Reduction 层之间加入迭代注意特征融合模块 (IAFF), 在 Inception-ResNet 中加入 MS\_CAM, 如图 2 所示。

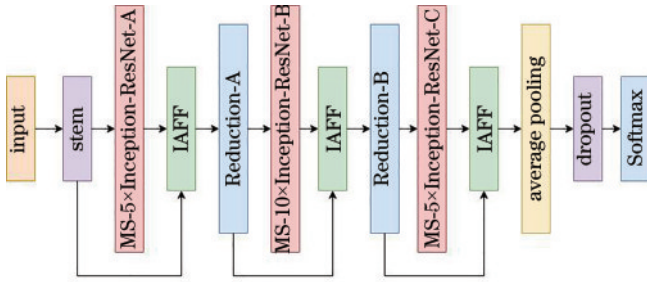


图 2 改进的 Inception-ResNet-V2 模型  
Fig. 2 Improved Inception-ResNet-v2 model

近年来, 大多网络中引入的注意力机制模块都是基于全通道的注意力机制, 对同一层中的多个组进行动态加权平均。尽管这种注意力机制为特征融合提供了非线性, 但它们的初始整合过于单一, 缺乏整合不同尺度注意特征的能力。所使用的 MS\_CAM 不仅能够有效解决不同尺度注意特征不一致的问题, 而且同时强调分布在全局的大型物体和突出分布在局部的小型物体, 促进网络识别和检测极端尺度变化下的目标, 其结构如图 3 所示。

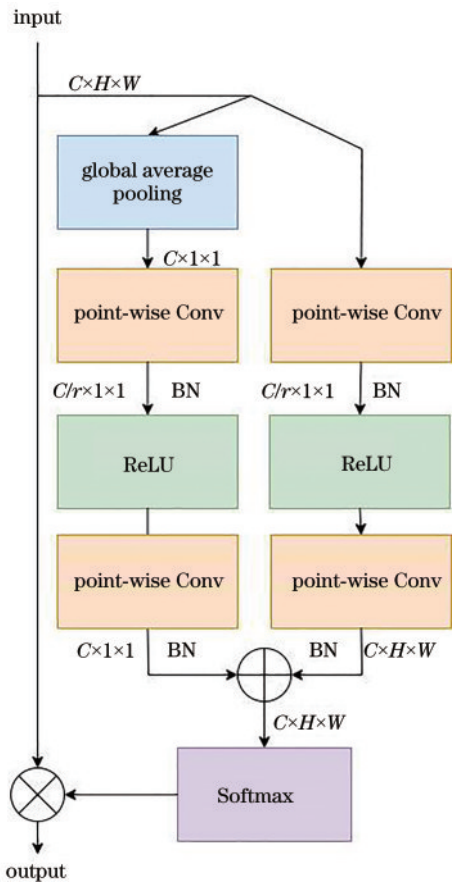


图 3 MS\_CAM 结构  
Fig. 3 MS\_CAM structure

如今, 许多特征融合多采用简单的求和或者串联的方式, 仍无法做到跨层的特征融合, 然而这只是提供了一个固定的特征图的线性聚集, 完全不知道这种组合是否适合特定的对象, 使得模型表现受限。所提方法使用 IAFF 模块为各种特征融合场景使用统一的方法进行融合, 并且 IAFF 做到了跨层的特征融合, 提高了网络在不同特征尺度对融合特征的质量以及网络的泛化性, IAFF 结构如图 4 所示。

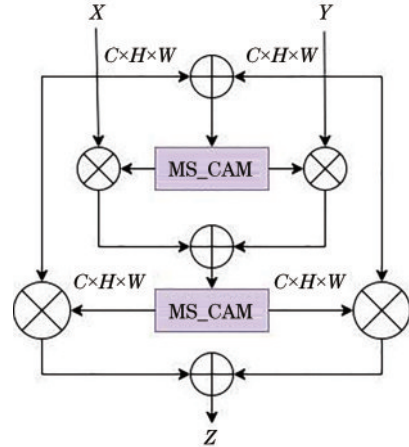


图 4 IAFF 结构  
Fig. 4 IAFF structure

### 2.3 改进的 stem 层

stem 层的前 3 层采用的是卷积核大小为  $3 \times 3$  的卷积, 在第 1 个数据拼接过后左分支采用的是  $1 \times 1$  卷积后接  $3 \times 3$  卷积, 右分支第 1 层和第 4 层使用的是  $1 \times 1$  卷积和  $3 \times 3$  卷积, 这样虽然在一定程度上降低了参数量, 但其实有效感受野并不大。由于数据集中图像分辨率较高, 图像像素值方差较小, 因此选用更大的  $7 \times 7$  卷积。通过多次实验, 把 stem 层的前 3 层卷积的卷积核大小由  $3 \times 3$  改为  $7 \times 7$ , 第 1 个数据拼接过后左分支改为 2 个卷积核大小为  $7 \times 7$  的卷积, 右分支第 1 层和第 4 层同样替换为  $7 \times 7$  的卷积, 如图 5 所示。以上修改大大增强了模型初期的有效感受野, 使得获取的全局特征更好。

### 2.4 激活函数的修改

原 Inception-ResNet-v2 中每一次卷积以及 Inception-ResNet 层最后都会引入 ReLU 激活函数<sup>[22]</sup>, 虽然激活函数能为模型带来非线性, 能够更好地把输入图片映射到输出, 但一定程度上过多的激活函数会使得模型更加复杂, 从而导致过拟合。所提方法在原 Inception-ResNet-v2 的基础上去掉每一次卷积过后的激活函数, 降低了模型的复杂程度, 加快了模型收敛速度。在原 Inception-ResNet 层的最后一层, 采用 SiLU 激活函数来代替原激活函数。SiLU 具备无上界、有下界、平滑、非单调的特性, 特别是在深层的网络中有着比 ReLU 激活函数更好的效果。

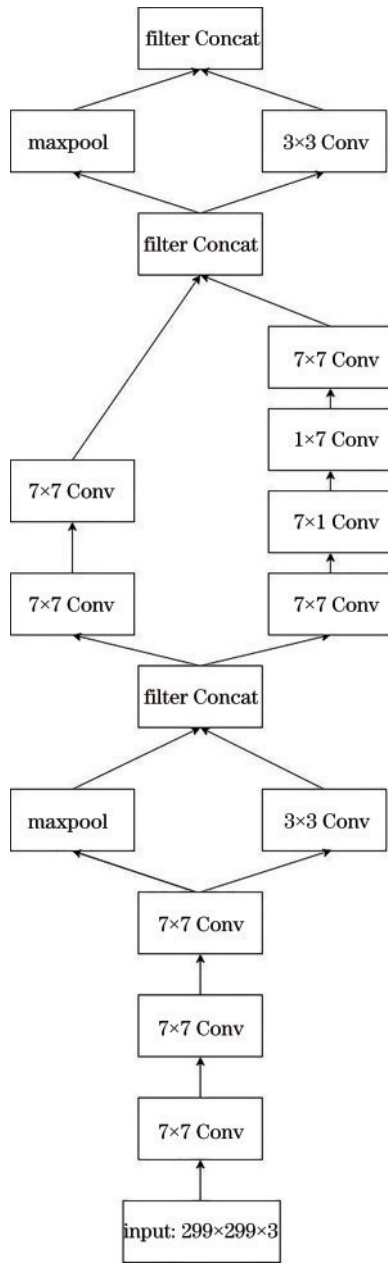


图 5 改进的 stem 模块  
Fig. 5 Improved stem module

### 3 实验及结果分析

#### 3.1 实验平台

本实验是在使用 Python 语言的基础上利用 PyTorch 深度学习框架进行的,实验基于 GeForce GTX1080Ti 的 Ubuntu 系统。

#### 3.2 数据集

本实验所用数据集是 Chest X-ray<sup>[23]</sup>,该公共数据集是基于广州妇女儿童医学中心 1~5 岁儿科患者的 X 射线扫描数据库制作的。该数据集一共包含 5856 张标记为正常和肺炎两个类别的胸部 X 光片,图像格式为 JPEG。根据数据集中标注肺炎类型再将其分为病毒性肺炎和细菌性肺炎,具体分布情况如表 1 所示。实验中采用 8:1:1 的比例来划分数据集,其中,80% 的肺炎图像作为训练集,10% 的肺炎图像作为测试集,10% 的肺炎图像作为验证集。图片样例如图 6 所示。

表 1 儿童肺炎分级数量

Table 1 Number of grades of pneumonia in children

Category	Type of pneumonia	Quantity
0	Bacterial pneumonia	2780
1	Normal	1583
2	Viral pneumonia	1493

#### 3.3 模型训练策略

为了适应 Inception-ResNet-v2 的输入,将图像大小调整为  $299 \times 299$ ,在训练集中采用翻转、随机角度旋转的方法进行数据增强<sup>[24]</sup>。

##### 3.3.1 优化器

优化器在神经网络中是用来更新和计算影响模型训练和模型输出的网络参数的,因此选择一个合适的优化器在模型构建中十分重要。由于所用数据集存在较大噪声,所以采用自适应矩估计(Adam)<sup>[25]</sup>优化器,Adam 能够在加快训练速度的同时,很好地避免局部最优解。在 Adam 优化器中梯度为  $g$ ,第 1 矩阵估计和第 2 矩阵估计为  $m$  和  $v$ ,参数变量为  $\theta$ ,在设置了初始学

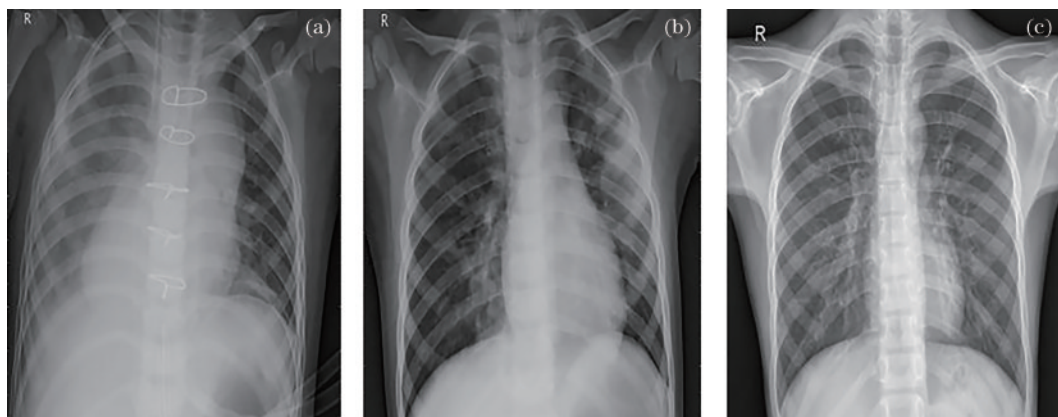


图 6 数据集样例。(a)细菌性肺炎;(b)病毒性肺炎;(c)正常肺  
Fig. 6 Sample dataset. (a) Bacterial pneumonia; (b) viral pneumonia; (c) normal

习率  $\alpha$  后对参数进行优化:

$$t \leftarrow t + 1, \quad (1)$$

$$g_t \leftarrow \nabla_{\theta} f_i(\theta_{t-1}), \quad (2)$$

$$m_t \leftarrow \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad (3)$$

$$v_t \leftarrow \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \quad (4)$$

$$\hat{m}_t \leftarrow m_t / (1 - \beta_1^t), \quad (5)$$

$$\hat{v}_t \leftarrow v_t / (1 - \beta_2^t), \quad (6)$$

$$\theta_t \leftarrow \theta_{t-1} - \alpha \times \hat{m}_t / (\sqrt{\hat{v}_t + \epsilon}), \quad (7)$$

式中:  $\beta_1$  和  $\beta_2$  为矩估计的自然衰减率;  $f$  为参数的随机目标函数值;  $t$  为时间步长。具体设置为:  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 10^{-8}$ , 初始学习率  $\alpha = 0.0001$ 。

### 3.3.2 学习率的调整

学习率是深度学习中重要的超参数。如果学习率设置过小, 会使得损失下降速度太慢, 增加模型的复杂度; 如果学习率设置过大, 梯度可能无法达到最小值, 甚至可能无法收敛。合理设置学习率能够使得模型在合适的时间收敛到局部最小值。所提方法采用指数减缓的方法来使学习率按训练轮数的增长指数插值递减, 其表达式为

$$\alpha = \alpha_0 \times \lambda^{N_{\text{epoch}}}, \quad (8)$$

式中:  $\alpha_0$  表示初始学习率;  $\lambda$  为学习率调整倍数的底数, 取 0.95;  $N_{\text{epoch}}$  为训练所迭代的次数。

### 3.3.3 损失函数

儿童肺炎是一个多分类问题, 本研究选用交叉熵损失函数。交叉熵损失函数是分类中常用的损失函数, 交叉熵是用来度量两个概率分布的差异性的, 用来衡量模型学习到的分布和真实分布的差异, 其具体公式为

$$L = -\frac{1}{N} \sum_{i=0}^{N-1} \sum_{k=0}^{K-1} y_{i,k} \ln p_{i,k}, \quad (9)$$

式中:  $y_{i,k}$  表示第  $i$  个样本的真实标签为  $k$ ;  $K$  和  $N$  分别表示标签值和样本的总数;  $p_{i,k}$  表示为第  $i$  个样本表示为  $k$  标签值的概率。

### 3.4 评价指标

为了验证所提改进 Inception-ResNet-v2 算法的效果, 使用准确率 ( $R_{\text{accuracy}}$ )、精确率 ( $R_{\text{precision}}$ )、召回率 ( $R_{\text{recall}}$ )、特异度 ( $R_{\text{specificity}}$ ) 及混淆矩阵来对模型性能进行进一步的测评, 计算公式为

$$R_{\text{accuracy}} = \frac{N_{\text{TP}} + N_{\text{TN}}}{N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}} + N_{\text{TN}}}, \quad (10)$$

$$R_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (11)$$

$$R_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (12)$$

$$R_{\text{specificity}} = \frac{N_{\text{TN}}}{N_{\text{TN}} + N_{\text{FP}}}, \quad (13)$$

式中:  $N_{\text{TP}}$  表示真阳性的数量;  $N_{\text{FP}}$  表示假阳性的数量;  $N_{\text{FN}}$  表示假阴性的数量;  $N_{\text{TN}}$  表示真阴性的数量。

混淆矩阵是对分类问题的预测结果的总结, 每一行之和表示该类别的真实样本数量, 每一列之和表示被预测为该类别的样本数量。

### 3.5 实验结果

采用混淆矩阵对儿童肺炎分类结果进行评估, 根据式(10)~(13)分别计算儿童肺炎三分类和二分类的准确率、精确率、召回率、特异度, 得到的三分类混淆矩阵如表 2 所示, 其中, 标签 0 表示细菌型肺炎, 标签 1 表示正常, 标签 2 表示病毒性肺炎, 每一列代表预测的类别, 每一列的总数表示预测为该类别的数据数目, 每一行代表数据的真实归属类别, 每一行的数据总数表示该类别的数据实例数目。可以看出, 大部分数据都集中在对角线上, 但结果仍未达到理想状态。课题组怀疑原因在于所用数据集中标签 2 病毒性肺炎的数据量远小于标签 0 细菌性肺炎的数据量, 加之两种类型的图片差异较小, 数据集本身噪声比较多, 导致三分类混淆矩阵没有达到预期值。而对于二分类, 准确率为 97.9%, 精确率为 97.6%, 召回率为 97.3%, 特异度为 97.2%。

表 2 融合迭代注意特征融合模块和改进 Inception-Resnet-v2 的分类混淆矩阵

Category	0	1	2
0	260	5	43
1	3	151	2
2	28	2	92

### 3.6 结果比较

为了验证模型的有效性, 在数据集 Chest X-ray 上与现有的研究成果进行二分类对比。选用 Teacher module<sup>[26]</sup>、GIV3<sup>[27]</sup>和 Chouhan<sup>[10]</sup>的方法与所提方法进行对比。其中, Teacher module 在 Inception-v3 模型上进行预训练, 并通过知识蒸馏的方法将训练好的有效信息提取到 AlexNet 模型中, GIV3 使用 ImageNet 数据集训练好的 GoogleNet Inception-v3 网络模型进行特征提取, 增加了特征融合层, 并使用随机森林分类器进行分类预测。

实验结果如表 3 所示。在与其他研究成果对比过程中, 由于所用评价指标不同, 因此选用相同评价指标准确率来对比实验结果。从表 3 可以看出, 所提方法

表 3 二分类实验对比

Table 3 Comparison of two classification experiments

Method	$R_{\text{accuracy}} / \%$
Teacher module	94.87
Method of reference [10]	96.39
GIV3	96.77
Proposed method	97.9

在准确率上优于其他方法,证明了其有效性。

而对于三分类,改进的 Inception-ResNet-v2 的分类混淆矩阵结果如表 2 所示,由于表 3 中的对比算法仅用作二分类任务,为进一步验证网络的效果,将其与

VGG16<sup>[28]</sup>、ResNet50<sup>[29]</sup>、Inception-v3、Inception-ResNet-v1、Inception-ResNet-v2 进行对比,其准确率,精确率,召回率,特异度如表 4 所示。可以发现,在各方面,所提方法均优于其他方法。

表 4 实验结果对比

Table 4 Comparison of experimental results

Method	Category	$R_{\text{accuracy}} / \%$	$R_{\text{precision}} / \%$	$R_{\text{recall}} / \%$	$R_{\text{specificity}} / \%$
VGG16	bacteria		82.1	76.4	85.8
	normal	79.5	94.2	93.8	97.7
	virus		61.0	68.0	83.8
ResNet50	bacteria		79.8	80.2	81.9
	normal	81.3	93.3	94.6	97.3
	virus		66.3	63.4	88.7
Inception-v3	bacteria		80.7	86.0	82.7
	normal	82.2	95.1	95.7	98.1
	virus		66.6	62.8	91.0
Inception-ResNet-v1	bacteria		76.7	86.0	76.2
	normal	78.8	94.3	93.6	98.2
	virus		67.7	54.2	89.8
Inception-ResNet-v2	bacteria		77.4	86.0	78.4
	normal	79.1	91.9	74.2	97.0
	virus		67.2	52.8	90.4
Proposed method	bacteria		84.4	89.3	83.7
	normal	85.8	96.8	95.6	98.8
	virus		75.4	67.2%	93.3%

采用消融实验来验证多尺度通道注意力模块、改进 stem 层、替换激活函数等 3 种优化策略的效果,并进行了 5 组对比实验,实验结果如表 5 所示。

从表 5 可以看出,实验 1 是不包括 3 种优化策略的 Inception-ResNet-v2 原模型,实验 2 是只加上修改的激活函数的模型,实验 3 是只加上改进的 stem 层的模型,实验 4 是加上多尺度通道注意力模块的模型,

实验 5 是包含 3 种优化策略的模型。其中,实验 3 改进的 stem 层对于模型的提升最明显,其准确率、精确率、召回率、特异度分别提高 4.3 个百分点、4.8 个百分点、4.9 个百分点、2.1 个百分点。实验 2 和实验 4 相对于原模型准确率也分别提高了 3.2 个百分点和 3.9 个百分点,精确率、召回率、特异度也有明显的提升。

表 5 消融实验结果对比

Table 5 Comparison of ablation experiment results

Number	Method	$R_{\text{accuracy}} / \%$	$R_{\text{precision}} / \%$	$R_{\text{recall}} / \%$	$R_{\text{specificity}} / \%$
Experiment 1	Inception-ResNet-v2	79.1	78.8	78.4	88.6
Experiment 2	Inception-ResNet-v2+SiLU	82.3	81.3	82	90.4
Experiment 3	Inception-ResNet-v2+new stem	83.4	83.6	83.3	90.7
Experiment 4	Inception-ResNet-v2+MS_CAM	83.0	82.2	82.2	90.8
Experiment 5	Proposed method	85.8	85.5	84.0	95.2

选用大型 COVID-19 肺部 CT 数据集来验证所提方法的鲁棒性,总共收集了来自 466 名患者的 7953 张 COVID-19 图像、来自 604 名患者的 6893 张正常图像以及来自 60 名患者的 2618 张 CAP 图像。在实验中只关注被诊断为 COVID-19 的患者,因此删除了 CAP 标签以获得二分类的结果。选用了 Hartono 等<sup>[30]</sup>和

Aytac 等<sup>[31]</sup>的方法与所提方法进行对比,实验结果如表 6 所示。

综合以上实验结果可知,所提优化策略能够有效增强模型对儿童肺炎图像的识别能力,提高分类精度,同时其他肺部数据集中也有较好的表现,证明了该方法的有效性。

表 6 COVID-19数据集实验结果对比

Table 6 Comparison of experimental results of the COVID-19 dataset

Method	$R_{accuracy} / \%$
Method of reference [31]	93.0
Method of reference [30]	97.0
Proposed method	98.9

## 4 结 论

提出一种改进的 Inception-ResNet-v2 分类算法用于儿童肺炎图像的检测。通过融合多尺度通道注意力模块有效提高模型对不同尺度特征的提取能力。增大 stem 层的卷积核大小,提升模型的有效感受野。减少激活函数的使用并且采用数据增强的方法来防止模型过拟合。实验结果表明,相比于其他方法,所提方法在儿童肺炎二分类和三分类任务中各方面指标均有所提高,具有更好的泛化性。但在实验过程中并未形成完整的分类系统,在后续的工作中会尝试把实验的所有过程开发成完整的辅助系统。

### 参 考 文 献

- [1] Tandon Y K, Bartholmai B J, Koo C W. Putting artificial intelligence (AI) on the spot: machine learning evaluation of pulmonary nodules[J]. *Journal of Thoracic Disease*, 2020, 12(11): 6954-6965.
- [2] Taraseviciene-Stewart L, Voelkel N F. Molecular pathogenesis of emphysema[J]. *The Journal of Clinical Investigation*, 2008, 118(2): 394-402.
- [3] Hashmi M F, Katiyar S, Keskar A G, et al. Efficient pneumonia detection in chest X-ray images using deep transfer learning[J]. *Diagnostics*, 2020, 10(6): 417.
- [4] World Health Organization. Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children[R]. Geneva: World Health Organization, 2001.
- [5] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]//NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems-Volume 1, December 3-6, 2012, Lake Tahoe, Nevada, United States. New York: ACM Press, 2012: 1097-1105.
- [6] Saraiva A A, Santos D B S, Costa N J C, et al. Models of learning to classify X-ray images for the detection of pneumonia using neural networks[C]//Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies, February 22-24, 2019, Prague, Czech Republic. Setúbal: SciTePress, 2019.
- [7] Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning[EB/OL]. (2017-11-14) [2022-02-04]. <https://arxiv.org/abs/1711.05225>.
- [8] Wang X S, Peng Y F, 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]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 3462-3471.
- [9] Ayan E, Ünver H M. Diagnosis of pneumonia from chest X-ray images using deep learning[C]//2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), April 24-26, 2019, Istanbul, Turkey. New York: IEEE Press, 2019.
- [10] Chouhan V, Singh S K, Khamparia A, et al. A novel transfer learning based approach for pneumonia detection in chest X-ray images[J]. *Applied Sciences*, 2020, 10(2): 559.
- [11] Barhoom A M A, Abu-Naser S S. Diagnosis of pneumonia using deep learning[J]. *International Journal of Academic Engineering Research (IJAER)*, 2022, 6(2): 48-68.
- [12] Acharya A K, Satapathy R. A deep learning-based approach towards the automatic diagnosis of pneumonia from chest radio-graphs[J]. *Biomedical and Pharmacology Journal*, 2020, 13(1): 449-455.
- [13] Rahman T, Chowdhury M E H, Khandakar A, et al. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray[J]. *Applied Sciences*, 2020, 10(9): 3233.
- [14] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[C]//AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA. New York: ACM Press, 2017: 4278-4284.
- [15] Dai Y M, Gieseke F, Oehmcke S, et al. Attentional feature fusion[C]//2021 IEEE Winter Conference on Applications of Computer Vision, January 3-8, 2021, Waikoloa, HI, USA. New York: IEEE Press, 2021: 3559-3568.
- [16] 陈思思, 陈明惠, 马文飞. 基于多通道的光学相干层析成像视网膜图像自动分类研究[J]. *中国激光*, 2021, 48(23): 2307001.  
Chen S S, Chen M H, Ma W F. Research on automatic classification of optical coherence tomography retina image based on multi-channel[J]. *Chinese Journal of Lasers*, 2021, 48(23): 2307001.
- [17] 张祥东, 王腾军, 朱劭俊, 等. 基于扩张卷积注意力神经网络的高光谱图像分类[J]. *光学学报*, 2021, 41(3): 0310001.  
Zhang X D, Wang T J, Zhu S J, et al. Hyperspectral image classification based on dilated convolutional attention neural network[J]. *Acta Optica Sinica*, 2021, 41(3): 0310001.
- [18] Gu Y, Lu X Q, Zhang B H, et al. Automatic lung nodule detection using multi-scale dot nodule-enhancement filter and weighted support vector machines

- in chest computed tomography[J]. *PLoS One*, 2019, 14 (1): e0210551.
- [19] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA. New York: IEEE Press, 2015.
- [20] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning-Volume 37, July 6-11, 2015, Lille, France. New York: ACM Press, 2015: 448-456.
- [21] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2818-2826.
- [22] Xu B, Wang N Y, Chen T Q, et al. Empirical evaluation of rectified activations in the convolutional network[EB/OL]. (2015-05-05) [2022-05-04]. <https://arxiv.org/abs/1505.00853>.
- [23] Kermany D S, Goldbaum M, Cai W J, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J]. *Cell*, 2018, 172(5): 1122-1131.
- [24] 王朝晖, 康欢, 陈多芳, 等. 轻量化深度网络辅助于无透镜计算显微图像的细胞分类[J]. *中国激光*, 2022, 49 (5): 0507204.  
Wang Z H, Kang H, Chen D F, et al. Lightweight deep learning network assisted cell classification using lensless computational microscopic imaging data[J]. *Chinese Journal of Lasers*, 2022, 49(5): 0507204.
- [25] Newey W K. Adaptive estimation of regression models via moment restrictions[J]. *Journal of Econometrics*, 1988, 38(3): 301-339.
- [26] 邓棋, 雷印杰, 田锋. 用于肺炎图像分类的优化卷积神经网络方法[J]. *计算机应用*, 2020, 40(1): 71-76.  
Deng Q, Lei Y J, Tian F. Optimized convolutional neural network method for classification of pneumonia images[J]. *Journal of Computer Applications*, 2020, 40 (1): 71-76.
- [27] 何新宇, 张晓龙. 基于深度神经网络的肺炎图像识别模型[J]. *计算机应用*, 2019, 39(6): 1680-1684.  
He X Y, Zhang X L. Pneumonia image recognition model based on the deep neural network[J]. *Journal of Computer Applications*, 2019, 39(6): 1680-1684.
- [28] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2022-05-04]. <https://arxiv.org/abs/1409.1556>.
- [29] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [30] Hartono A P, Luhur C R, Indriyani C A, et al. Evaluating deep learning for CT scan COVID-19 automatic detection[C]//2021 International Conference on ICT for Smart Society (ICISS), August 2-4, 2021, Bandung, Indonesia. New York: IEEE Press, 2021.
- [31] Aytac U C, Güneş A, Ajlouni N. A novel adaptive momentum method for medical image classification using convolutional neural network[J]. *BMC Medical Imaging*, 2022, 22(1): 34.