

双信道注意力机制神经网络的肺炎图像分类研究

曾德洋, 杨鑫荣

上海理工大学, 光电信息与计算机工程学院, 上海

收稿日期: 2024年1月4日; 录用日期: 2024年1月20日; 发布日期: 2024年2月29日

摘要

为提高诊断肺炎患者的时效性和精确性, 本文提出了三种基于双信道注意力机制的肺部X光图像检测模型。采用了21,158张肺部X光图像, 并对原始图像进行图像增强、几何变化等处理。将预处理后的X光图像分别输入三种模型并检测, 最后采用识别的准确度、精确度、灵敏度、F1-Score等指标评估了各个模型的性能。实验结果表明, 基于双通道注意力机制的DenseNet201模型效果最优, 其准确度、精确度、灵敏度、F1-Score分别为95.112%、96.2%、95.4%、95.9%, 其中准确度已超过近年来他人4分类肺炎模型, 基于双通道注意力机制的DenseNet201模型有利于提高肺炎图像的分类检测效率, 可以应用于快速医学筛查。

关键词

医学X光图像, CBAM, ResNet, DenseNet, ResNext

Research on Pneumonia Image Classification Based on Dual Channel Attention Mechanism Neural Network

Deyang Zeng, Xinrong Yang

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Jan. 4th, 2024; accepted: Jan. 20th, 2024; published: Feb. 29th, 2024

Abstract

In order to improve the timeliness and accuracy of the diagnosis of pneumonia patients, this paper

文章引用: 曾德洋, 杨鑫荣. 双信道注意力机制神经网络的肺炎图像分类研究[J]. 理论数学, 2024, 14(2): 733-745.

DOI: 10.12677/pm.2024.142072

proposes three lung X-ray image detection models based on dual channel attention. This paper uses 21,158 lung X-ray images, then these original images are processed with image enhancement and geometric changes. Next, the preprocessed X-ray images are input into three different models for detection. Finally, the performance of each model is evaluated by using evaluation factors such as accuracy, precision, sensitivity, and F1-Score. The results indicate that the DenseNet201 model based on dual channel attention has gained greater performance increasing up on average of 95.112%, 96.2%, 95.4%, and 95.9% for accuracy, precision, sensitivity, and F1-Score, respectively. In terms of accuracy, it has surpassed that of others' 4 classification pneumonia models in recent years. The DenseNet201 model based on dual channel attention mechanism is beneficial to improve the classification and detection efficiency of pneumonia images, this model can be applied to medical rapid screening.

Keywords

Medical X-Ray Images, CBAM, ResNet, DenseNet, ResNext

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

传统肺炎检测采用逆转录聚合酶链式反应(RT-PCR)方法来诊断患者患病与否,然而该方法存在一定缺陷,其主要表现为诊断耗时过长且具有较高假阳性率。与正常肺部 X 光图像相比,肺炎 X 光图像具有较多阴影区域,然而其他肺部问题,如肺水肿等也会增加肺部 X 光图像阴影区域,这无疑增加了通过肺部 X 光图像判断肺炎患病与否的难度。肺部 X 光图像具有丰富病理信息,经验丰富的医师可快速通过肉眼观测肺部 X 光图像并给出合理判断,大多数缺乏丰富经验的医生往往不能给出合理判断,但,即便是具有丰富经验的医生也会由于疲劳度等不可抗因素给出错误诊断结果[1]。综上所述,我们迫切需要一种更加准确的肺炎 X 光图像检测手段。

当下人工智能已经展现出卓越的性能,其中基于卷积神经网络的医学图像分类技术也愈加成熟。使用卷积神经网络对肺部 X 光图像进行预处理、训练、特征提取、依据特征信息进行分类,最终精确预测肺炎类别,在提高肺炎检测准确率的同时,减轻了医生的医疗压力。因此,将基于卷积神经网络的深度学习技术应用于肺炎检测是一种非常有效的诊断解决方案[2]。

2. 相关工作

大量研究表明,利用深度学习对肺部 X 光图像进行检测比人工肉眼检测具更高的准确度以及更少的检测时间。周天奕[3]等,提出了一种模糊逻辑引导的多粒度深度神经网络,可充分提取病理图像特征信息,该模型在组织病理图像分类上具有较高的准确率。J. Zhang [4]等提出了置信度感知异常检测(CAAD)模型,该模型在区分病毒性肺炎和非病毒性肺炎的任务中取得了和人工检测相近的结果,该模型的准确度达到 83.16%,灵敏度达到 71.7%,虽然在分类结果上与人工不相上下,但对区分病毒性肺炎和非病毒性肺炎的速度远超人工区分,且比人工区分更具稳定性。J. Wang [5]提出了一种人工智能肺炎检测系统 AIRBiS, AIRBiS 基于高性能人工神经网络和交互式用户界面,可实现有效的操作和监控,其评估结果表明,其所提出的系统对肺炎的检测准确率达到 94.4%。E. J. Palomo [6]等等提出了一种用于胸部 X 射线图像中肺炎检测的计算机辅助诊断系统,该系统基于卷积神经网络(CNNs),该系统的二分类肺炎分类准

确率达到 98.59%。S. Pappula [7]等提出了一种自动肺炎诊断方法 CAD (计算机辅助诊断), 可显著减少收集 CXR 图像数据的时间和成本。裴颂文[8]等人提出了一种用于图像降噪的生成对抗网络模型(TriGAN), 对 CT 图像进行降噪并增强图像特征, 与 DnCNN 以及 GAN 等网络相比, 图像峰值信噪比平均提高 4.5%, 结构相似性平均提高 1.5%。王科平[9]等人, 提出基于伪全局 Swin Transformer 的遥感图像识别算法, 有效提升模型对所有目标的感知能力。V. Dediu [10]提出了一种轻量级的筛查工具, 可以协助医生高效的对肺炎患者进行分类。J. Joby [11]等提出了 PneumoGAN 的模型, PneumoGAN 在公开二分类肺炎数据集上的准确率、召回率和 F1-score 分别为 87.71%、91.4%和 89.52%。H. Agrawal [12]等利用迁移学习和图像增强的概念, 设计了一种新的基于图像处理的深度学习方法来检测肺炎。D. Abin [13]等利用图像去雾技术来提高原始肺炎图像的质量, 在数据集层面来优化肺炎分类准确性。S. Jamil [14]等提出了一种基于 AlexNet 和 SVM 的肺炎检测算法, 与传统机器学习算法相比, 该算法表现出了更好的性能, 并且简单易用。Yadavendra [15]等将 Mask R-CNN 应用于肺部 X 图像不透明区域快速检测并取得了优异效果。Budi Nugroho [16]等提出了使用 RMSprop 优化算法来提升神经网络性能, 在三分类肺炎数据集上的分类精确率达到 88%, 准确率到达 87.99%, 敏感度达到 86%, F1-score 达到 87%。谢星星[17]等提出了一种动态特征融合网络, 降低了输入图像中目标尺度对特征融合的影响。许金亚[18]等人针对医学图像小目标病灶区域检测, 提出了一种具有注意力机制的特征金字塔模型, 该模型较 3DCE 模型平均检测精度提高了 5.2%, 较 MSE 模型平均精度提高了 4.4%。江泽涛[19]等提出了一种结合特征增强和多尺度感受野的低照度目标检测算法, 在低照度目标检测精度方面取得了较好的表现。

上述方法实验中涉及的肺部 X 光图像广泛应用于肺炎分类检测任务中, 但由于肺部 X 光图像数据集存在较大类别不平衡性, 导致分类结果与现实情况具有一定的偏差。

本文提出并对比分析了 3 种肺炎图像分类模型, 首先对实验数据进行预处理, 减小由于数据集的不平衡性所带来的影响; 随后将融合双信道注意力模块的网络模型在 ImageNet 数据集上进行迁移学习, 使得模型更具健壮性; 最后将处理后的数据集放入模型中进行训练、分类, 最终得到最优分类模型。

3. 基于双信道注意力机制肺炎图像分类方法

3.1. 双信道注意力机制模块

双信道注意力机制模块(CBAM)由通道注意力模块(CAM)及空间注意力模块(SAM)组成, CAM 用于提取特征图通道特征, SAM 用于提取特征图空间特征。图像进入神经网络得到对应特征图, 首先经过两个并行的 MaxPool 和 AvgPool 对特征图进行初步处理, 使得特征图大小由 $C \times H \times W$ 变为 $C \times 1 \times 1$, 然后再经 ShareMLP 模块处理, 先将通道数压缩为原来的 $1/n$, 再扩张到原通道数, 再经 ReLU 激活函数得到两个激活后的结果。将两个输出结果逐元素相加, 随后通过 sigmoid 激活函数得到 CAM 模块的输出结果, 最后将输出结果与原始特征图相乘, 使特征图维度恢复为 $C \times H \times W$ 。SAM 将 CAM 的输出结果通过 MaxPool 和 AvgPool 得到两个维度为 $1 \times H \times W$ 的特征图, 并对这两个特征图进行 Concat 操作, 通过 7×7 卷积变为 1 通道的特征图, 随后经 sigmoid 处理得到 SAM 的特征图, 最后将输出结果再与原始特征图相乘恢复特征图初始维度。

通道注意力公式如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (1)$$

空间注意力公式如下:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (2)$$

其中: F 代表输入特征图, AvgPool 表示平均池化, MaxPool 表示最大池化, MLP 为多层感知机, σ 表示 sigmoid 激活函数, f^{7*7} 表示一个 $7*7$ 的卷积, $M_c(F)$ 表示经通道注意力模块后得到的特征图, $M_s(F)$ 表示经空间注意力机制后得到的特征图。

3.2. 残差模型

ResNet 的核心思想是使用残差模块(Residual Block)构建深度神经网络。传统神经网络每一层输入都会经过一系列线性变换和非线性激活函数后输出给下一层。由于信号在逐层传递的过程中不可避免地会产生信息损失, 因此随着网络深度的增加, 在模型训练过程会出现梯度消失和梯度爆炸等问题, 导致网络深度受到限制, ResNet 通过引入残差模块来解决这个问题。残差模块由两个卷积层和一个跨层连接组成, 跨层连接将输入信号直接传递到输出端, 使得输入信号可以直接流到下一层, 从而解决了信号传递过程中的信息损失问题。残差模块使得网络深度可以更加容易地加深, 同时也提高了网络的精确度和稳定性。残差模块公式如下。

$$y = \sigma(F(x, W) + x) \quad (3)$$

其中, y 代表残差块输出值, σ 代表激活函数, F 代表残差函数, x 代表输入特征, W 代表权重。

3.3. DenseNet 模型

DenseNet 的核心思想是在各层之间增加密集连接。DenseNet 各层都接收来自前面所有层的输出。这种密集连接使网络更加有效地利用参数, 在避免梯度消失和梯度爆炸的同时, 提高了网络的信息流动性和特征重用性。

DenseNet 主要特点包括: 1) 密集连接: 每个层都与前面所有层连接, 使信息更加充分地流动。2) 特征重用: 由于每个层都能够接收前面所有层的输出, 可以更好地利用共享的特征。3) 参数共享: 由于每个层都可以接收前面所有层的输出, 可以大大降低网络中参数的数量。4) 降低梯度消失和梯度爆炸: 由于每个层都能够接收前面所有层的输出, 可以避免在网络深度加深时出现的梯度消失和梯度爆炸的问题。5) 减少过拟合: DenseNet 中引入了 Dropout 和 Batch Normalization 等技术, 可有效减少模型过拟合。DenseNet 模型公式如下。

$$X_i = H_i([X_1, X_2, X_3, \dots, X_{i-1}]) \quad (4)$$

其中 X_1, X_2, X_3, X_{i-1} 代表特征图, H_i 代表非线性变换, $[]$ 表示将所有特征图按在通道维度上进行拼接。

3.4. ResNext 模型

ResNext 模型采用一种全新的卷积操作, 称为分组卷积(Group Convolution), 该卷积操作将输入通道分组, 将各分组的通道分别进行卷积计算, 最后将各分组的结果进行拼接, 得到最终的输出。与传统卷积操作相比, 分组卷积可减少运算量, 提高计算效率, 与此同时 ResNext 模型融合了 ResNet 的残差连接(Residual Connection)思想, 即将输入与输出进行加和操作, 将前一层的输出直接传递到后一层, 从而避免了由于网络深度增加带来的梯度消失问题, 使得网络更加容易训练。ResNext 模型公式如下所示。

$$y = x + \sum_{i=1}^C \tau_i(x) \quad (5)$$

其中 x 代表输入特征, $\tau_i(x)$ 代表分支结构, C 代表卷积组大小。

3.5. 基于 CBAM 的神经网络模型结构

本文提出并构建了三个基于 CBAM 的神经网络模型, 分别如图 1、图 2 和图 3 所示。其中图 1 为

ResNet50 + CBAM 模型结构, 图 2 为 DenseNet201 + CBAM 模型结构, 图 3 为 ResNext50 + CBAM 模型结构。三个模型接收到的输入图像维度均为 $1*224*224$ 。除固有结构差异外, 三种模型之间的核心差异在于双通道注意力模块(CBAM)的插入位点不同。

ResNet50 + CBAM 模型由以下几部分组成, 首先是一个维度为 $1*224*224$ 的图像输入层, 其后是一个 $7*7$ 的卷积层, 卷积后进入 4 个不同的瓶颈模块组, 每个瓶颈模块组由 3 个卷积层和一个 CBAM 模块拼接而成, 其中第一个瓶颈模块组额外多一个最大池化层, 各层的具体参数如下图所示, 其中 Conv 代表卷积操作, c 代表通道数, s 代表卷积步长, p 代表 padding 数。ResNet50 + CBAM 模型结构以及具体参数, 如图 1 所示。

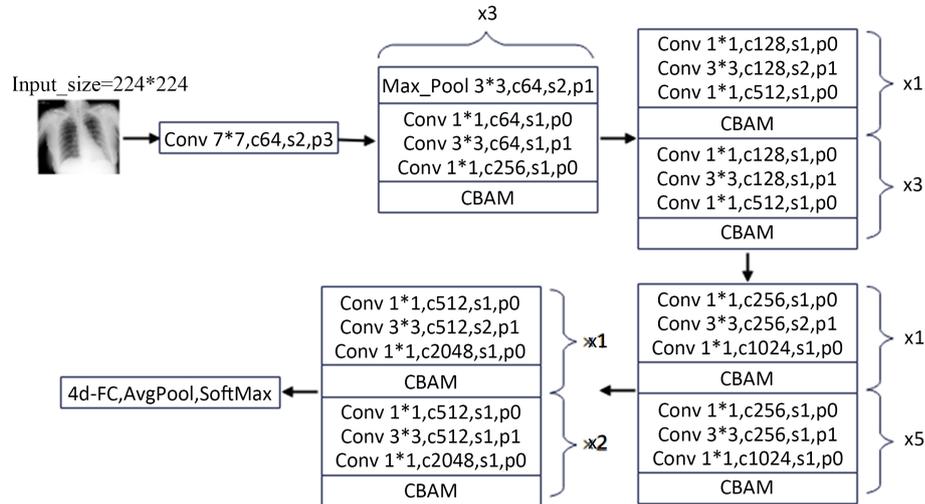


Figure 1. ResNet50 + CBAM model structure
图 1. ResNet50 + CBAM 模型结构

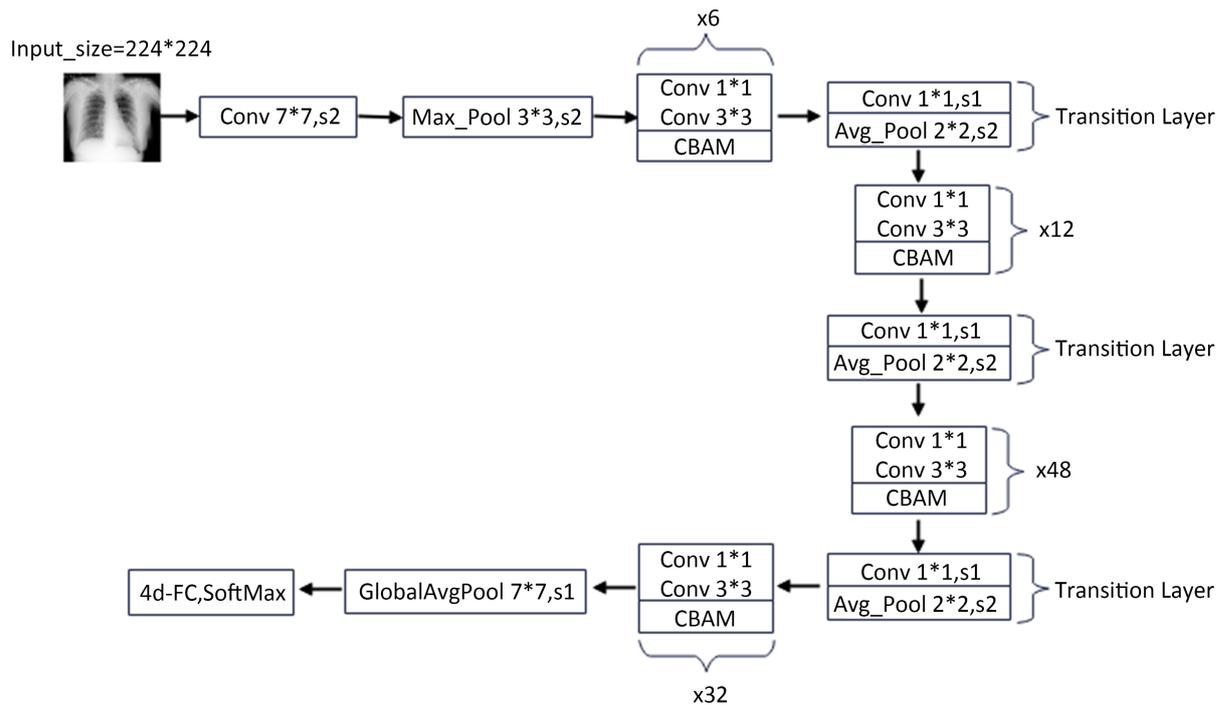


Figure 2. DenseNet201 + CBAM model structure
图 2. DenseNet201 + CBAM 模型结构

DenseNet201 + CBAM 模型整体结构由以下几部分组成, 首先是一个维度为 224*224 的输入图像输入层, 其后是一个 7*7 的卷积层, 卷积后进入 4 个不同的 DenseBlock 层, 每个 DenseBlock 层由若干个体积结构构成, 每个小体积结构又由一个 1*1 卷积、一个 3*3 卷积, 以及一个 CBAM 模块构成, 相邻 DenseBlock 结构间存在一个 Transition 层, Transition 层由一个 1*1 卷积以及一个 2*2 的平均池化层构成, 最后分类输出层由 7*7 平均池化和 4 个完全连接的分类层组成。DenseNet201 + CBAM 模型结构如图 2 所示。

ResNext50 + CBAM 模型由以下几部分组成, 首先是一个大小为 224*224 的输入图像, 其后是由 64 个 7*7 的卷积核构成的卷积层以及一个 3*3 的最大池化层, 池化后进入 4 个组卷积层, 分别重复 3、4、6、3 次, 每个组卷积层都由 2 个 1*1 的卷积以及一个 Group 数为 32 的分组卷积构成, ResNext50 + CBAM 模型结构以及具体参数如图 3 所示。

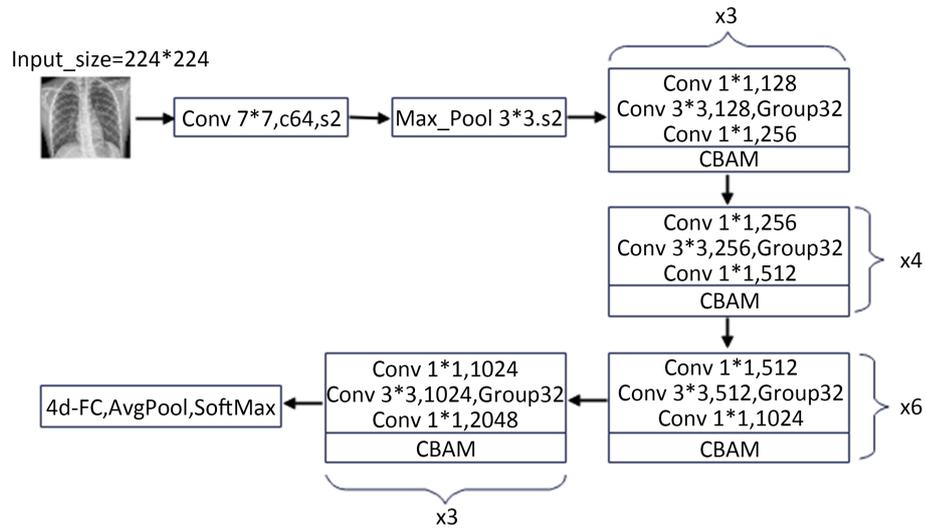


Figure 3. ResNext50 + CBAM model structure
图 3. ResNext50 + CBAM 模型结构

在三个模型的训练过程中, 使用交叉熵作为损失函数来计算每次反向传播的损失值。优化器选择 SGD (随机梯度下降), 学习率为 0.001, 动量为 0.5。选择 SGD 的优点是, SGD 引起的波动有利于优化方向从当前局部极小点跳到另一个更好的局部极小点。因此, 对于非凸函数, 它会收敛到一个好的局部极值点, 甚至全局极值点。此外, 优化器一次只更新一个样本的模型参数, 训练速度快, 有利于模型的快速收敛。交叉熵损失函数及 SGD 优化器公式如下。

$$Loss = -\sum y^i * \log p^i \tag{6}$$

其中 Loss 表示交叉熵损失值, y 表示真实标签, p 表示模型预测概率值。

$$\omega_{t+1} = \omega_t - \rho \nabla_{\omega} L(\omega_t; x_i; y_i) \tag{7}$$

其中 ω_t 表示某一时刻模型参数, x 代表样本, y 代表样本标签, $\nabla_{\omega} L(\omega_t; x_i; y_i)$ 代表在 (x_i, y_i) 点处模型梯度, ρ 代表学习率。

4. 实验

4.1. 实验流程

实验总体流程, 如图 4 所示, 包括三个重要阶段: 1) 通过迁移学习初始化模型参数; 2) 合理的图像

预处理; 3) 选择最优的骨干网络和 CBAM 的最佳组合。

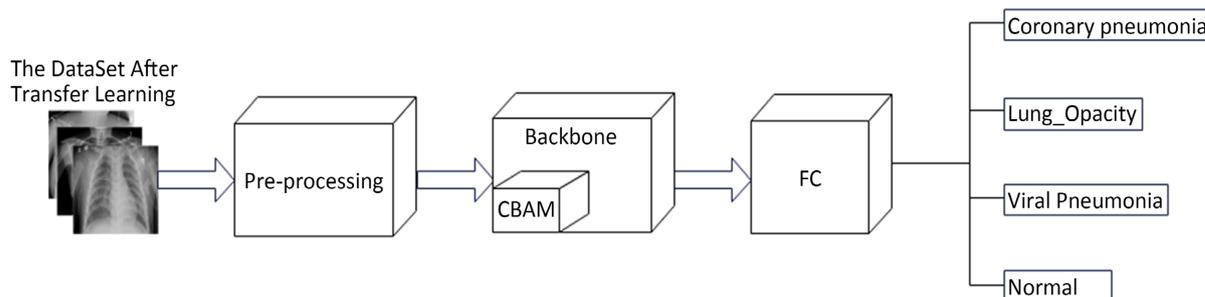


Figure 4. Experimental flowchart

图 4. 实验流程

本文使用 ImageNet 公开数据集对模型进行预训练, 使得模型具有更加合理的初始化参数。图像识别算法的准确性与图像质量强相关, 从 Kaggle 中获取的原始肺炎数据集并不能很好的直接放入到深度学习模型中进行训练。为提高模型分类准确性, 本实验在数据集的图像中对每个像素乘以 $1/255$ 进行标准化处理, 使得数据集在 0 到 1 的范围内进行归一化, 从而使数据集在像素强度方面保持一致。此外, 我们所获取的数据集中, 不同类别的数据量有较大差异。因此, 本文进行了相关的数据增强, 使得不同类别数据量保持平衡。数据增强的方法有很多, 包括 1) 基于几何变换的扩增技术, 如 Flipping, Cropping, Rotation, Translation, Noise injection 等。2) 颜色空间变换, 通过在特征通道上进行操作, 改变图像颜色分布, 进而增强模型对色彩变化的鲁棒性。3) 卷积核滤波器, 通过卷积操作对图像进行一系列滤波处理, 以增强模型泛化能力, 4) Mixing images, 通过将多个图像融合生成新图像, 从而使得模型更加关注图像间的相关性。5) Random erasing, 通过随机擦除图像某区域, 引入局部信息损失, 使得模型在不完整的图像信息中进行训练, 有利于提高模型鲁棒性和泛化能力。本文使用 rotation 和 flipping 增强, 针对于本实验, rotation 操作使得模型能够学习图像不同角度下的特征, 避免模型对于训练数据中特定角度的过拟合, flipping 操作有助于减小模型对图像中某些方向上的特定特征依赖, 使得模型更具泛化能力。几何变换后的肺炎图像, 如图 5 所示。

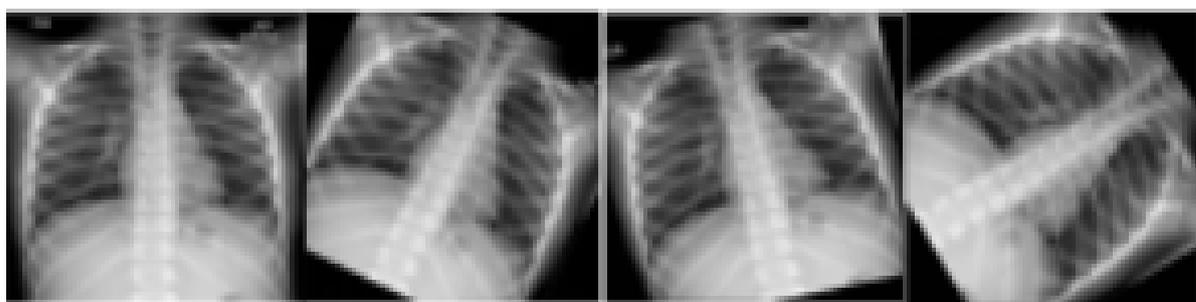


Figure 5. Image of pneumonia after geometric changes

图 5. 几何变换后的肺炎图像

4.2. 数据集及评估指标

4.2.1. 数据集

本实验数据集取自 Kaggle 公开肺炎数据集, 共 21,157 张肺部 X 光图片, 分为 4 类, 分别是: Coronary pneumonia (3610 张), Lung Opacity (6010 张), Viral Pneumonia (1345 张), Normal (10192 张), 典型图像

数据如图 6 所示。

四种类型的肺炎图像在数据集四个子分类中的具体分布情况如表 1 所示, 为确保实验的科学有效性以及模型的健壮性, 本实验预先将不同子分类按 8:1:1 的比例随机分配到训练集、测试集及验证集中。

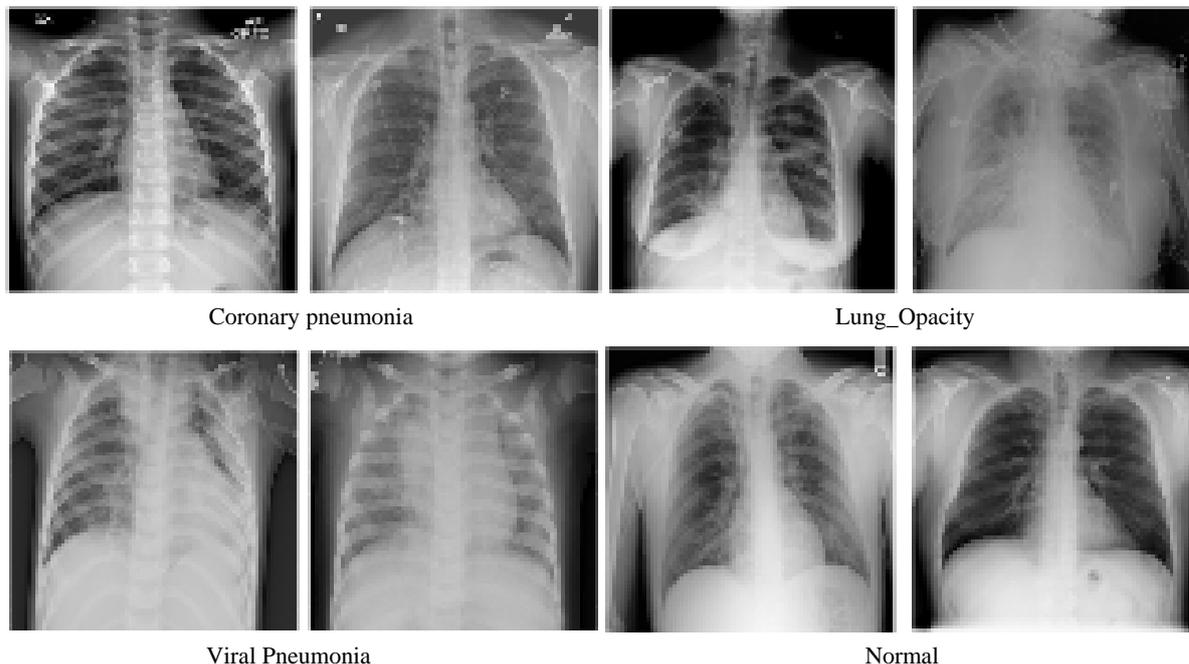


Figure 6. Pneumonia dataset

图 6. 肺炎数据集

Table 1. Dataset partitioning

表 1. 数据集划分情况

数据集	Coronary pneumonia	Lung Opacity	Viral Pneumonia	Normal	Total
训练集	2888	4808	1072	8152	16920
测试集	362	602	135	1020	2119
验证集	362	602	135	1020	2119

4.2.2. 评估指标

深度学习模型性能除了和模型本身结构、参数有关外, 还与模型评估指标密切相关, 针对不同模型及应用场景, 选择合理的评估指标对模型性能的合理判断极为重要。

为了多方位多角度的评估本实验模型真实性能, 本文引入了 4 中不同评估指标, 分别为: 1) Accuracy; 2) Precision; 3) Sensitivity; 4) F1-Score。

1) 模型 Accuracy 值代表模型分类准确度, 表示在所有样本分类中正确分类的比例。

$$Accuracy = (TP + TN) / (TP + FP + TN + FN)$$

2) 模型 Precision 值代表模型分类的精确度, 表示在所有被预测为正类别的样本中, 真实属于正类别的比例。

$$Precision = TP / (TP + FP)$$

3) 模型 Sensitivity 值代表模型分类的特异性, 表示在所有实际正类别样本中成功预测为正类别的比例。

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

4) 模型 F1-Score 值兼顾了模型准确率和灵敏性, 是准确率和灵敏性的调和平均值。

$$\text{F1-score} = 2 \cdot \text{Precision} \cdot \text{Recall} / (\text{Precision} + \text{Recall})$$

上式中, TP 为模型正确预测为正类别的样本数, TN 为模型正确预测为负类别的样本数, FP 为模型错误预测为正类别的样本数, FN 为模型错误预测为负类别的样本数。

4.3. 基于 CBAM 的改进 ResNet50 模型的肺炎 CT 图像识别

原始肺部 X 光图像经 rotation, flipping 增强后, 按 batch_size = 16 为一组, 输入 ResNet50 + CBAM 模型进行训练, 按图 1 所示结构, 在 ResNet50 基模型不同位点融入 CBAM 特征提取模块, 强化模型特征提取力度, 图像在模型内部经过一系列池化、卷积、CBAM 处理等操作得到最终的输出结果。

经过改进的 ResNet50 模型在训练集上的 loss 变化曲线以及在验证集上的 Accuracy 变化曲线如图 7 所示, 在 400 个 epoch 后 loss 以及 Accuracy 变化趋于稳定, 经 500 个 epoch 后 ResNet50 + CBAM 模型在验证集上的四分类肺炎分类最大准确率达到 89.504%。

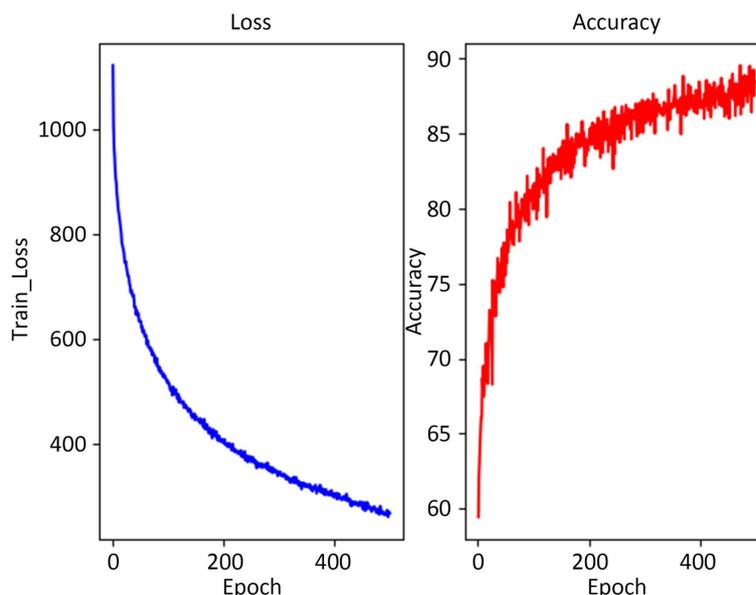


Figure 7. Changes in Loss and Accuracy values of ResNet50 + CBAM Model
图 7. ResNet50 + CBAM 模型 Loss 值与 Accuracy 值变化情况

4.4. 基于 CBAM 的改进 DenseNet201 模型的肺炎 CT 图像识别

图像增强后的原始肺部 X 光图像按 batch_size = 16 为一组, 输入 DenseNet201 + CBAM 模型进行训练, 按图 2 所示结构, 图像经一系列 DenseBlock 层、Transition 层进行特征提取, 在每个 DenseBlock 中, 每经一个由 1*1 以及 3*3 卷积构成的小体结构后, 额外进行一次 CBAM 特征提取, 使得模型提取到更加深层的图像信息, 图像在模型内部图像经过一系列池化、卷积、CBAM 处理等操作得到最终的输出结果。

经过改进的 DenseNet201 模型在训练集上的 loss 变化曲线以及在验证集上的 Accuracy 变化曲线如图 8 所示, 在大约 350 个 epoch 后 loss 以及 Accuracy 变化趋于稳定, 经 500 个 epoch 后 DenseNet201 + CBAM

模型在验证集上的四分类肺炎分类最大准确率达到 95.112%。

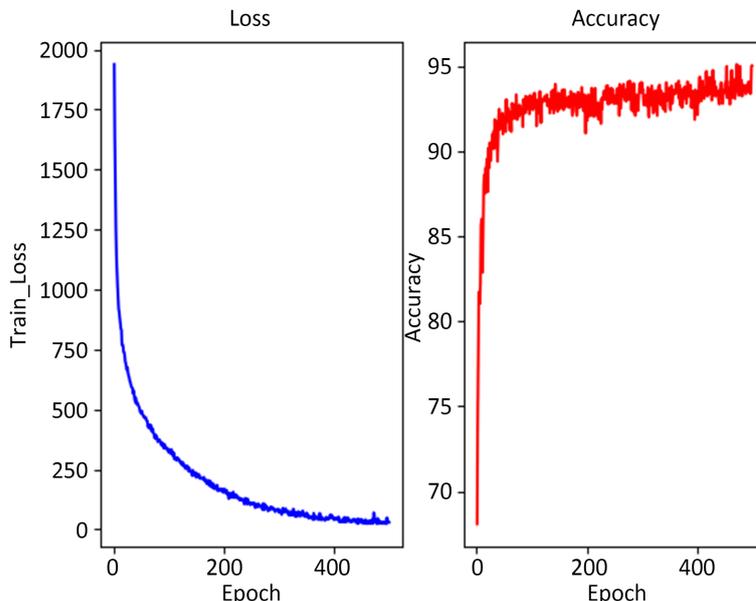


Figure 8. Changes in Loss and Accuracy values of DenseNet201 + CBAM Model
图 8. DenseNet201 + CBAM 模型 Loss 值与 Accuracy 值变化情况

4.5. 基于 CBAM 的改进 ResNext50 模型的肺炎 CT 图像识别

图像增强后的原始肺部 X 光图像按 $batch_size = 16$ 为一组, 输入 ResNext50 + CBAM 模型进行训练, 按图 3 所示结构, 分别在每个组卷积层中的分组卷积块后加入 CBAM 特征提取模块, 将图像经过组卷积后的特征在空间以及通道维度上进行二次特征提取, 使得图像特征被充分利用, 图像在模型内部图像经过一系列池化、卷积、CBAM 处理等操作得到最终的输出结果。

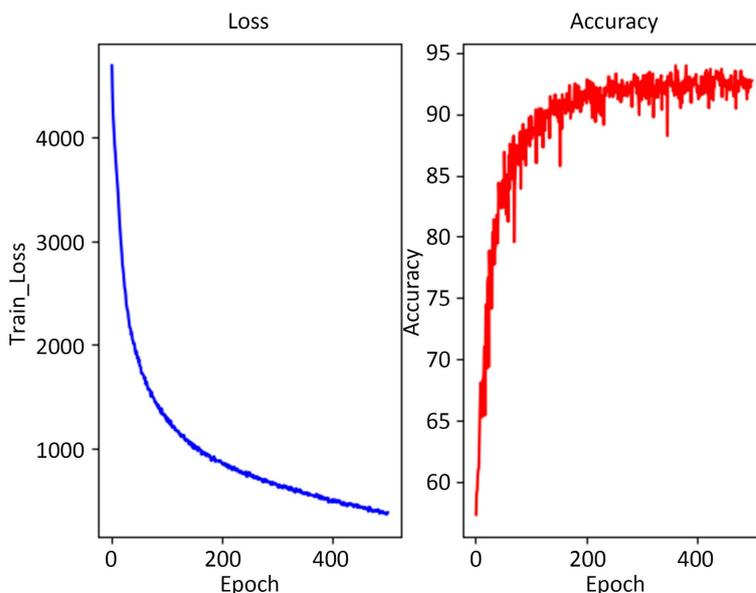


Figure 9. Changes in Loss and Accuracy values of ResNext50 + CBAM Model
图 9. ResNext50 + CBAM 模型 Loss 值与 Accuracy 值变化情况

经过改进的 ResNext50 模型在训练集上的 loss 变化曲线以及在验证集上的 Accuracy 变化曲线如图 9 所示, 在大约 400 个 epoch 后 loss 以及 Accuracy 变化趋于稳定, 经 500 个 epoch 后 ResNext50 + CBAM 模型在验证集上的四分类肺炎分类最大准确率达到 93.948%。

4.6. 结果与分析

本文经实验得到的三种四分类肺炎模型的性能指标如图 10 所示。从中可以看出, DenseNet201 + CBAM 在准确度、精确度、灵敏度以及 F1-Score 均显著优于 ResNet50 + CBAM 模型以及 ResNext50 + CBAM 模型, 最大值分别为 95.112%、96.2%、95.4% 和 95.9%, 这体现了 DenseNet201+CBAM 模型的潜在卓越性, 同时也凸显了 CBAM 模块对模型性能改进的积极作用。深入分析图 7、图 8 和图 9, 可以更为具体地观察到 DenseNet201 + CBAM 模型相较于 ResNet50 + CBAM 模型和 ResNext50 + CBAM 模型的性能优势。首先, 在 loss 值的收敛速度方面, 随着训练批次的逐渐增加, DenseNet201 + CBAM 模型呈现出更为迅速的模型损失收敛趋势。这表明 DenseNet201 + CBAM 模型更为迅速地适应了训练数据, 为其更高效的学习能力提供了有力支持。其次, 模型在准确度提升速率方面也展现出明显差异, 进一步印证了 DenseNet201 + CBAM 模型在特征提取和分类性能方面的卓越表现。

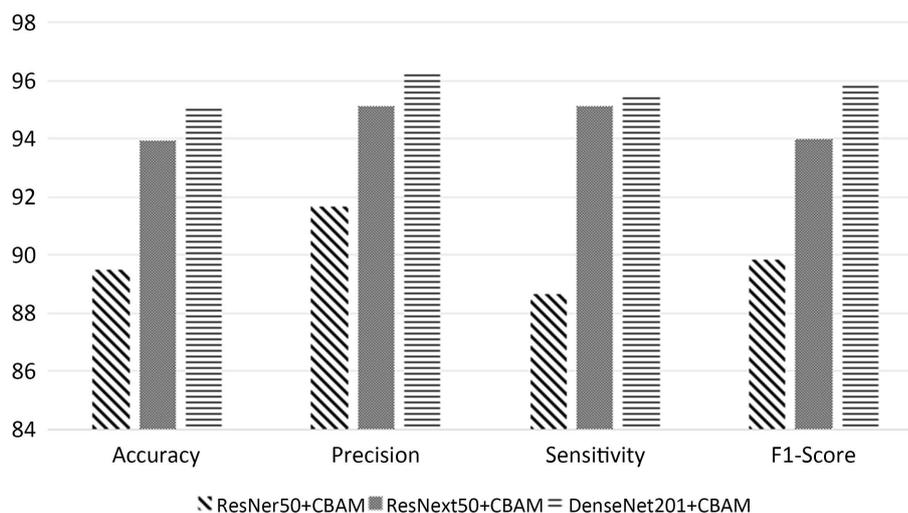


Figure 10. Comparison results of the three models proposed in this article on Accuracy, Precision, Sensitivity, and F1 Score

图 10. 3 种模型在 Accuracy、Precision、Sensitivity、F1-Score 上的对比结果

肺炎分类是目前医学图像分类的热门话题, 但绝大多数肺炎分类都是二元分类, 多分类模型较少。本文将通过实验获得的最优模型 DenseNet201 + CBAM 与近年来他人肺炎多分类模型进行了比较, 结果如表 2 所示。

Table 2. Comparison of accuracy of four classification pneumonia models in recent years

表 2. 近年来四分类肺炎模型分类准确度比较

Reference	Year	# of classes	Accuracy
P. Nagapuri [20]	2022	4	91%
A. Jha [21]	2022	4	94.65%
S. S. Selvi [22]	2022	4	95.01%
Proposed	2023	4	95.112%

由表 2 可知, 本文提出的基于 CBAM 改进的 DenseNet201 模型比近年来提出的四分类肺炎分类模型具有更好的准确性。

5. 结束语

本文将特征提取模块 CBAM 与不同网络模型进行融合, 探究并对比分析了 CBAM + DenseNet 模型在四分类肺炎数据集上的性能, 本文提出的基于 CBAM 的 DenseNet201 四分类肺炎模型在分类准确度上达到 95.112%, 超过近些年其他肺炎四分类模型, 实验结果证明该模型适用于多分类肺炎分类检测, 临床诊断时可为医生提供可靠帮助。基于 CBAM 的 DenseNet201 四分类肺炎分类模型不仅在医学图像分类领域具有实质性作用, 同时也为深度学习方法在其他复杂任务中的应用提供了有益的经验 and 指导。本文的研究为未来医学影像分析和疾病诊断领域的深入研究提供了新的思路和方法。

未来工作将结合更多临床数据进一步提高 DenseNet201 + CBAM 的预测能力, 以实现快速区分不同肺炎患者, 并在肺炎患者得到快速救治的同时, 更大程度地减小医疗压力。

参考文献

- [1] Meenakshi, P., Bhavana, K. and Nair, A.K. (2022) Pneumonia Detection Using X-Ray Image Analysis with Image Processing Techniques. 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, 22-24 June 2022, 1657-1662. <https://doi.org/10.1109/ICCES54183.2022.9835798>
- [2] Ko, H., Ha, H., Cho, H., Seo, K. and Lee, J. (2019) Pneumonia Detection with Weighted Voting Ensemble of CNN Models. 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, 25-28 May 2019, 306-310. <https://doi.org/10.1109/ICAIBD.2019.8837042>
- [3] 周天奕, 丁卫平, 黄嘉爽, 鞠恒荣, 姜舒, 王海鹏. 模糊逻辑引导的多粒度深度神经网络[J]. 模式识别与人工智能, 2023, 36(9): 778-792.
- [4] Zhang, J., et al. (2021) Viral Pneumonia Screening on Chest X-Rays Using Confidence-Aware Anomaly Detection. *IEEE Transactions on Medical Imaging*, **40**, 879-890. <https://doi.org/10.1109/TMI.2020.3040950>
- [5] Wang, J., Nakamura, M. and Abdallah, A.B. (2022) Efficient AI-Enabled Pneumonia Detection in Chest X-Ray Images. 2022 IEEE 4th Global Conference on Life Sciences and Technologies (LifeTech), Osaka, 7-9 March 2022, 470-474. <https://doi.org/10.1109/LifeTech53646.2022.9754850>
- [6] Palomo, E.J., Zafra-Santisteban, M.A. and Luque-Baena, R.M. (2022) Pneumonia Detection in Chest X-Ray Images Using Convolutional Neural Networks. 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRINE), Rome, 26-28 October 2022, 16-21. <https://doi.org/10.1109/MetroXRINE54828.2022.9967590>
- [7] Pappula, S., Nadendla, T., Lomadugu, N.B. and Revanth Nalla, S. (2023) Detection and Classification of Pneumonia Using Deep Learning by the Dense Net-121 Model. 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, 17-18 March 2023, 1671-1675. <https://doi.org/10.1109/ICACCS57279.2023.10113110>
- [8] 裴颂文, 樊静, 沈天马, 顾春华. 面向低剂量 CT 图像的多生成器对抗网络降噪模型的研究[J]. 小型微型计算机系统, 2020, 41(12): 2582-2587.
- [9] 王科平, 左鑫浩, 杨艺, 费树岷. 基于伪全局 Swin Transformer 的遥感图像识别算法[J]. 模式识别与人工智能, 2023, 36(9): 818-831.
- [10] Dediu, V. and Udrea, A. (2023) Pneumonia Detection from X-Ray Images Using Convolutional Neural Networks. 2023 24th International Conference on Control Systems and Computer Science (CSCS), Bucharest, 24-26 May 2023, 61-66. <https://doi.org/10.1109/CSCS59211.2023.00019>
- [11] Joby, J., Raju, R.J., Job, R., Thomas, S.M. and Arun, K.S. (2021) PneumoGAN: A GAN based Model for Pneumonia Detection. 2021 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), Ernakulam, 2-4 September 2021, 102-106. <https://doi.org/10.1109/ACCESS51619.2021.9563341>
- [12] Agrawal, H. (2021) Pneumonia Detection Using Image Processing and Deep Learning. 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, 25-27 March 2021, 67-73. <https://doi.org/10.1109/ICAIS50930.2021.9395895>
- [13] Abin, D., Thepade, S.D. and Dhore, S. (2021) An Empirical Study of Dehazing Techniques for Chest X-Ray in Early

- Detection of Pneumonia. 2021 *2nd International Conference for Emerging Technology (INCET)*, Belagavi, 21-23 May 2021, 1-5. <https://doi.org/10.1109/INCET51464.2021.9456201>
- [14] Jamil, S., Abbas, M.S., Fawad, Zia, M.F. and Rahman, M.U. (2021) A Deep Convolutional Neural Network Based Framework for Pneumonia Detection. 2021 *International Conference on Digital Futures and Transformative Technologies (ICoDT2)*, Islamabad, 20-21 May 2021, 1-5. <https://doi.org/10.1109/ICoDT252288.2021.9441539>
- [15] Yadavendra and Chand, S. (2021) Pneumonia Lung Opacity Detection and Segmentation in Chest X-Rays by Using Transfer Learning of the Mask R-CNN. 2021 *7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, 19-20 March 2021, 1035-1040. <https://doi.org/10.1109/ICACCS51430.2021.9441864>
- [16] Nugroho, B. and Yuniarti, A. (2022) Performance of Root-Mean-Square Propagation and Adaptive Gradient Optimization Algorithms on Covid-19 Pneumonia Classification. 2022 *IEEE 8th Information Technology International Seminar (ITIS)*, Surabaya, 19-21 October 2022, 333-338. <https://doi.org/10.1109/ITIS57155.2022.10010119>
- [17] 谢星星, 程焱, 姚艳清, 姚西文, 韩军伟. 动态特征融合的遥感图像目标检测[J]. 计算机学报, 2022, 45(4): 735-747
- [18] 许金亚, 裴颂文. 面向 CT 图像小目标病灶区域的智能化检测算法研究[J]. 小型微型计算机系统, 2023, 44(8): 1792-1797.
- [19] 江泽涛, 翟丰硕, 钱艺, 肖芸, 张少钦. 结合特征增强和多尺度感受野的低照度目标检测[J]. 计算机研究与发展, 2023, 60(4): 903-915.
- [20] Nagapuri, P., et al. (2022) Deep Learning Hybrid for X-Ray Lung Disease Detection. 2022 *Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE)*, Bangalore, 16-17 December 2022, 1-7. <https://doi.org/10.1109/ICATIECE56365.2022.10047575>
- [21] Jha, A., John, E. and Banerjee, T. (2022) Transfer Learning for COVID-19 and Pneumonia Detection Using Chest X-Rays. 2022 *IEEE 65th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Fukuoka, 7-10 August 2022, 1-4. <https://doi.org/10.1109/MWSCAS54063.2022.9859403>
- [22] Selvi, S.S., Agarwal, N., Barkur, P., Mishra, Y. and Kumar, A. (2022) A Novel Deep Learning Algorithm for Covid Detection and Classification. 2022 *2nd International Conference on Intelligent Technologies (CONIT)*, Hubli, 24-26 June 2022, 1-5. <https://doi.org/10.1109/CONIT55038.2022.9847880>