

## 基于高效通道注意力的胸部 X 光片疾病分类算法

邵凌云<sup>1</sup>, 李强<sup>1</sup>, 关欣<sup>1\*</sup>, 丁学文<sup>2</sup><sup>1</sup>天津大学微电子学院, 天津 300072;<sup>2</sup>天津职业技术师范大学天津市现场总线控制技术工程中心, 天津 300222

**摘要** 深入研究不同肺部疾病的 X 射线光片, 有助于更清晰、准确地区分和预测各种疾病。基于此, 提出一种基于高效通道注意力机制的胸部 X 光片疾病分类算法。将高效通道注意力模块以密集连接的方式加入基础特征提取网络, 以增强特征通道中有效信息的传递, 同时抑制无效信息的传递; 使用非对称卷积块提高网络特征提取能力; 采用多标签损失函数解决多标签和数据不平衡的问题。将新型冠状病毒肺炎 X 光片添加到公开数据集 Chest X-ray 14 中构成数据集 Chest X-ray 15, 在该数据集上的实验结果表明, 所提基于高效通道注意力机制的胸部 X 光片疾病分类算法的平均 area under curve (AUC) 值达到 0.8245, 对气胸的 AUC 值达到 0.8829, 性能优于对比算法。

**关键词** 医用光学; 医学图像处理; 胸部 X 光片; 卷积神经网络; 高效通道注意力

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP220759

## Disease Classification Algorithm of Chest X-Ray Based on Efficient Channel Attention

Shao Lingyun<sup>1</sup>, Li Qiang<sup>1</sup>, Guan Xin<sup>1\*</sup>, Ding Xuwen<sup>2</sup><sup>1</sup>School of Microelectronics, Tianjin University, Tianjin 300072, China;<sup>2</sup>Tianjin Fieldbus Control Technology Engineering Center, Tianjin Vocational and Technical Normal University, Tianjin 300222, China

**Abstract** Extensive investigations of X-ray films of different lung diseases will help to precisely distinguish and predict various diseases. Herein, an algorithm for chest X-ray disease classification based on an efficient channel attention mechanism is proposed. The high-efficiency channel attention module is added to the basic feature extraction network in a densely connected manner to improve the transmission of effective information in the feature channel while inhibiting the transmission of invalid information. By using asymmetric convolution blocks to improve the ability of network feature extraction, the multilabel loss function is used to address multilabeling and data imbalance. The novel coronavirus pneumonia X-ray film is added to the public dataset, Chest X-ray 14, to form the dataset, Chest X-ray 15. The experimental results on this dataset show that the average area under curve (AUC) value of the proposed chest X-ray-film disease classification algorithm based on the efficient channel attention mechanism reaches 0.8245, and the AUC value for pneumothorax reaches 0.8829. Thus, the proposed algorithm is superior to comparison algorithms.

**Key words** medical optics; medical image processing; chest X-ray; convolutional neural network; efficient channel attention

## 1 引言

2019年12月,新冠肺炎疫情出现,引起全球的密切关注<sup>[1]</sup>。世界卫生组织将其命名为“2019年冠状病毒病”,简称COVID-19。截至2021年12月31日,美国约

翰霍普金斯大学实时统计数据显示,全球新型冠状病毒病毒病例高达2.87亿例,死亡病例超过543万<sup>[2]</sup>。逆转录聚合酶链反应(RT-PCR)常用于COVID-19的筛选<sup>[3]</sup>。但是RT-PCR灵敏度较低,仅为71%,需要重复检测,效率远低于胸部X光片检测。除新冠肺炎,心脏扩大、

收稿日期: 2022-02-17; 修回日期: 2022-04-07; 录用日期: 2022-05-25; 网络首发日期: 2022-06-05

基金项目: 国家自然科学基金(62071323, 61471263, 61872267)、天津市自然科学基金(16JCZDJC31100)、天津市科技计划项目(20YDTPJC01110)、天津大学自主创新基金(2021XZC-0024)

通信作者: \*guanxin@tju.edu.cn

肺气肿、肿块、肺结节和水肿等异常也是很多严重疾病的症状。及时诊断和治疗这些疾病能够有效降低死亡率。目前,诊断肺炎的主要方法还是胸部X光片检测,X光片检测也是医院常规体检项目<sup>[4]</sup>。但是随着肺部疾病患者逐年增多,仅依靠人工对X光片进行诊断工作量大,易漏诊,并且高强度工作很容易造成误诊。

随着深度学习的快速发展,卷积神经网络(CNN)已经在图像处理领域得到广泛应用。研究人员尝试让模型自己学习如何分配注意力,引发了注意力机制的发展<sup>[5]</sup>。Hu等<sup>[6]</sup>提出一种新的架构单元——压缩-激励(SE)模块,通过建模通道之间的相互依赖性,自适应地重新校准通道特征响应,在不同数据集上具有非常好的泛化能力。Li等<sup>[7]</sup>设计了一种名为选择性内核卷积(SKConv)的网络模型,该网络具有不同内核大小的多个分支,这些分支由Softmax注意力进行融合,神经元能够捕捉不同尺度的目标对象,且神经元接收域能够根据输入自适应调整。Woo等<sup>[8]</sup>提出一种新的思路,将注意力机制中的通道域和空间域结合设计卷积块注意力模块(CBAM)。卷积块注意力模块是一个轻量级通用模块,可以无缝集成到任何卷积神经网络架构中,适用性非常广。以上方法致力于开发更加复杂的注意力模块以获得更好的性能,但是没有考虑跨通道交互,这使得模型的复杂性进一步增大。许多研究人员将注意力机制和卷积神经网络融合,并引入计算机辅助诊断(CAD)系统。在胸部X光片分类领域:Wang等<sup>[9]</sup>在Chest X-ray 8数据集上使用预训练的ResNet50对胸部疾病进行分类并与其他3种网络模型进行比较,但是忽略了特征通道间的相关性导致分类准确率较低且数据集中只包含8种疾病,疾病种类较少;Ma等<sup>[10]</sup>使用注意力机制网络对肺部疾病进行识别与分类,该网络以ResNet101作为主干网络并加入SE模块来建模通道间的依赖关系,但检测结果的准确率较低;Yao等<sup>[11]</sup>提出一种结合ResNet

和DenseNet的多分辨率分类网络,该网络在分辨率高的病灶区定位图上表现突出,但是因缺乏对特征通道信息的关注而准确率不高;张智睿等<sup>[12]</sup>在DenseNet的基础网络上加入密集压缩激励块以增强特征通道的信息表达能力,在网络中以密集连接的方式引入压缩激励模块,形成一个特征通道注意力模块,增强通道中有效特征信息,同时抑制无效特征信息。但是没有考虑通道之间信息的交互以及疾病之间的相关性和依赖性,存在对部分疾病的分类准确率较低且对胸部疾病的整体分类准确率仍有待提高的问题。

针对上述研究中存在的问题,同时为进一步提高胸部疾病识别的准确率,本文提出一种基于高效通道注意力机制(ECA)的胸部X光片疾病分类算法,命名为dense efficient channel attention network (DECA-Net)。在DSENet<sup>[12]</sup>的基础上,引入高效通道注意力机制模块<sup>[13]</sup>,形成DECA-Net,主要作用是突出特征图中有效信息的特征通道,抑制无效信息的特征通道,与SE模块相比,ECA模块考虑跨通道交互,能够有效捕捉通道间的联系,在只增加少量参数的情况下,却能获得明显的性能增益;将网络中的 $3 \times 3$ 的方形卷积核替换为非对称卷积核<sup>[14]</sup>,以减小计算量;采用多标签损失函数(MSML)来解决多标签和数据不平衡的问题。

## 2 胸部疾病分类框架

### 2.1 DECA-Net 结构

所提DECA-Net整体框架如图1所示,基于DSENet进行改进,包含4个密集高效通道注意力模块:DECA block 1, DECA block 2, DECA block 3, DECA block 4,分别含有6、12、24、16个密集高效通道注意力块(DECA)。DECA结构如图2所示,其由全局平均池化、 $1 \times 1$ 卷积、Sigmoid激活函数和 $3 \times 3$ 非对称卷积块(ACB)组成,且每个DECA block之间由过渡层进行连接。

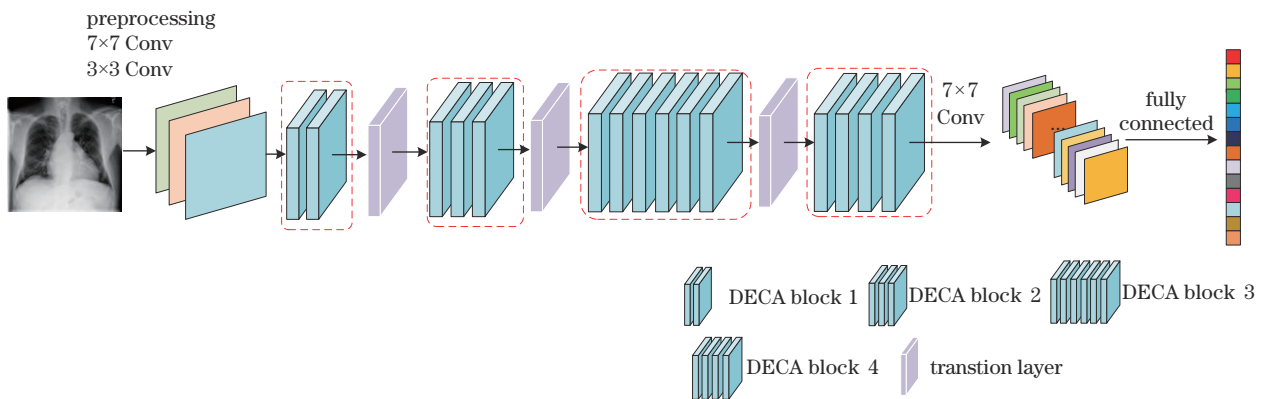


图1 胸部疾病分类网络框架图

Fig. 1 Framework diagram of chest disease classification network

首先,图像经过预处理、卷积与池化后输入DECA block 1。DECA block 1包含全局平均池化、 $1 \times 1$ 卷积

和 $3 \times 3$ 非对称卷积;全局平均池化的目的是通过池化操作把多维矩阵转换为特征向量替代全连接,极大地

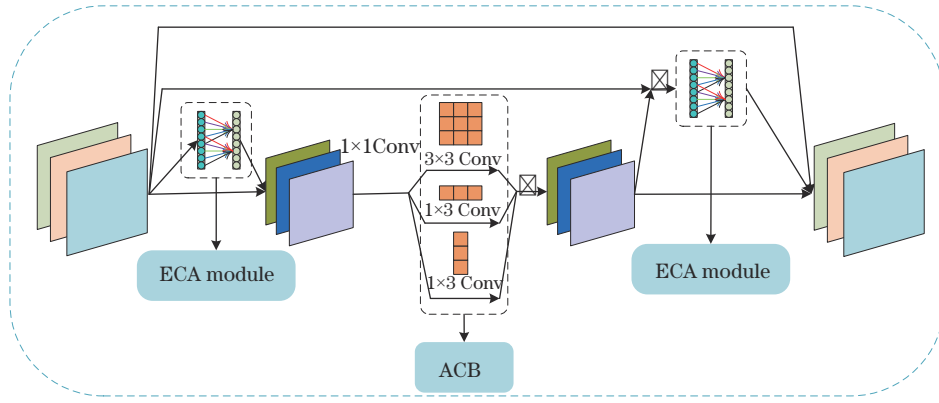


图 2 DECA 结构  
Fig. 2 Structure of DECA

减少了网络参数;  $1 \times 1$  卷积操作的目的是对输入特征图进行降维, 并融合各个通道的特征;  $3 \times 3$  非对称卷积的目的是在不引入额外的超参数的同时充分提取特征图的信息。DECA block 结合了密集连接和高效通道注意力机制的优势, 既可以充分提取特征信息并融合, 又可以在不增加网络参数的情况下提升网络模型的性能。

然后经过过渡层 (transition layer), 过渡层包含  $1 \times 1$  卷积层和  $2 \times 2$  平均池化层,  $1 \times 1$  卷积层将上一个 DECA block 的输出降维以避免信息冗余,  $2 \times 2$  平均池化层是下一个 DECA block 的预处理。过渡层输出的特征图进入 DECA block 2, 并进行多轮循环。特征图经过 DECA block 4 后进入全连接分类层。

最后, 全连接 (FC) 层对特征图进行  $7 \times 7$  全局平均池化操作, 建立 15 维特征向量的全连接, 从而输出图像的识别与分类结果, 实现一个完整的训练过程。网络的具体结构配置如表 1 所示。

### 2.2 DenseNet-121

深层网络可以提取更复杂的特征, 但是随着网络层数的增加, 网络的参数量过大会导致计算速度下降, 还会出现梯度消失等问题。ResNet<sup>[15]</sup> 模型的出现使得这些问题得到缓解, ResNet 改善了网络的非线性, 可以训练出更深的 CNN 模型, 提取更加复杂的图像特征。ResNet 模型的核心是通过建立层与层之间的“短路连接”构造残差结构, 这种结构能够加速网络收敛并有助于梯度的反向传播, 从而使训练出的网络能够解决梯度消失的问题。

DenseNet 是由 Huang 等<sup>[16]</sup> 于 2017 年提出的一种新的 CNN 结构, 其基本思路与 ResNet 一致, ResNet 通过残差学习有效缓解了深度网络的退化问题<sup>[13]</sup>。与 ResNet 的连接方式相比, DenseNet 将特征图大小匹配的所有层直接连接, 确保网络中各层之间最大信息流, 有效地缓解了梯度消失问题, 增强特征传播, 更适用于多标签分类问题。DenseNet 在参数和计算成本更少

表 1 DECA-Net 具体结构

Table 1 Specific structure of DECA-Net

Layer	Output size	DECA-Net-121
Convolution	$112 \times 112$	$7 \times 7$ Conv, stride 2
Pooling	$56 \times 56$	$3 \times 3$ max pool, stride 2
DECA block 1	$56 \times 56$	$\left\{ \begin{array}{l} \text{GAP} \\ 1 \times 1 \text{ Conv} \times k \\ \text{Sigmoid} \\ 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ ACB} \end{array} \right\} \times 6$
Transition layer 1	$56 \times 56$ $28 \times 28$	$1 \times 1$ Conv $2 \times 2$ average pool, stride 2
DECA block 2	$28 \times 28$	$\left\{ \begin{array}{l} \text{GAP} \\ 1 \times 1 \text{ Conv} \times k \\ \text{Sigmoid} \\ 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ ACB} \end{array} \right\} \times 12$
Transition layer 2	$28 \times 28$ $14 \times 14$	$1 \times 1$ Conv $2 \times 2$ average pool, stride 2
DECA block 3	$14 \times 14$	$\left\{ \begin{array}{l} \text{GAP} \\ 1 \times 1 \text{ Conv} \times k \\ \text{Sigmoid} \\ 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ ACB} \end{array} \right\} \times 24$
Transition layer 3	$14 \times 14$ $7 \times 7$	$1 \times 1$ Conv $2 \times 2$ average pool, stride 2
DECA block 4	$7 \times 7$	$\left\{ \begin{array}{l} \text{GAP} \\ 1 \times 1 \text{ Conv} \times k \\ \text{Sigmoid} \\ 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ ACB} \end{array} \right\} \times 16$
Classification layer	$1 \times 1$	$7 \times 7$ global average pool, stride 2 15 fully-connected

的情形下可实现比 ResNet 更优的性能。1 个 5 层的密集连接块如图 3 所示, 包括 5 个卷积层  $x_0, x_1, x_2, x_3, x_4$  和 4 个过渡层  $H_1, H_2, H_3, H_4$ 。

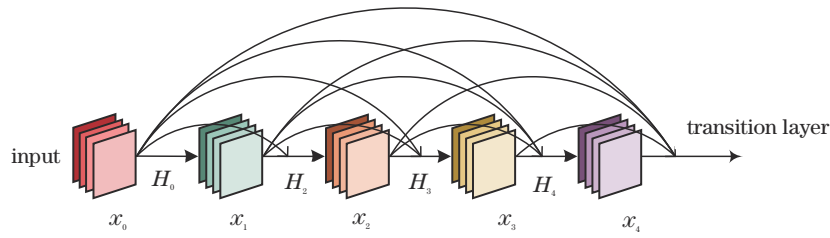


图 3 5 层密集连接块, 每一层都将上述所有特征图作为输入

Fig. 3 5-layer dense connected block, each layer taking all the preceding feature-maps as input

与 ResNet 相比, DenseNet 不需要学习冗余的特征图, 需要的参数更少。除此之外, DenseNet 的另一个优势是其改进的信息流动和整个网络的梯度更易于训练。每一层都有从损失函数和原始输入信号直接访问的梯度, 形成一个隐式的深度监督, 有助于训练更深层次的网络结构。此外, 密集连接具有正则化效应, 能够减少训练集规模较小任务的过拟合现象。

### 2.3 ECA 模块

相较于 SE 模块, ECA 模块避免降维, 使用一维卷积确定局部跨通道交互的覆盖范围。ECA 模块通过考虑每个通道及其  $k$  个相邻通道来捕获局部跨通道交互,  $k$  表示局部跨通道交互的覆盖范围, 既保证效率又保证效果。ECA 模块原理如图 4 所示, 对于聚合特征  $y \in \mathbf{R}^C$  不进行降维处理, 注意力通道权值表达式为

$$\omega = \sigma(\mathbf{W}y), \quad (1)$$

式中:  $\sigma$  代表 Sigmoid 激活函数;  $\mathbf{W}$  代表一个  $C \times C$  的参数矩阵;  $\omega$  为通道的权值。

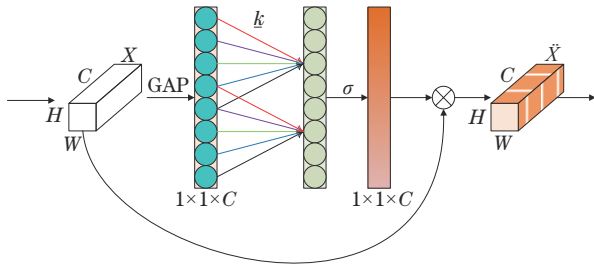


图 4 高效通道注意力模块

Fig. 4 Efficient channel attention module

其次, ECA 采用捕获跨信道交互的方法保证效率和有效性, 使用频率矩阵  $\mathbf{W}_k$  学习通道注意力,  $\mathbf{W}_k$  的表达式为

$$\begin{pmatrix} \omega^{1,1} & \dots & \omega^{1,k} & 0 & 0 & \dots & \dots & 0 \\ 0 & \omega^{2,2} & \dots & \omega^{2,k+1} & 0 & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & \omega^{C,C-k+1} & \dots & \omega^{C,C} \end{pmatrix}. \quad (2)$$

式(2)包括  $k \times C$  个参数, 避免不同通道之间完全独立, 目标是捕获局部的跨通道交互, 即只考虑每个通道与其  $k$  近邻之间的相互作用。  $y_i$  的权重可以表示为

$$\omega_i = \sigma \left( \sum_{j=1}^k \omega^j y_i^j \right), y_i^j \in \Omega_i^k, \quad (3)$$

式中:  $\Omega_i^k$  为  $y_i$  的  $k$  个相邻通道的集合。式(3)通过捕获局部跨通道交互建立局部约束机制, 从而避免所有通道的交互, 使模型的效率得到提高。每个通道的注意力模块涉及  $k \times C$  个参数, 为进一步降低模型复杂度并提升模型效率, 共享所有通道的权重, 即  $y_i$  的权重可以表示为

$$\omega_i = \sigma \left( \sum_{j=1}^k \omega^j y_i^j \right), y_i^j \in \Omega_i^k. \quad (4)$$

这种策略可以通过核大小为  $k$  的一维快速卷积来实现, 即

$$\omega = \sigma[\text{C1D}_k(y)], \quad (5)$$

式中: C1D 表示一维卷积。

交互的覆盖范围(即一维卷积的核大小  $k$ )与通道维度  $C$  成正比, 也就是说  $k$  和  $C$  之间可能存在一个映射  $\phi$ , 即

$$C = \phi(k), \quad (6)$$

式中:  $\phi$  通常是不确定的, 参数化指数函数是一个可行的选择。作为核函数的指数族函数(如高斯)被广泛用于处理未知映射问题。因此, 使用一个指数函数近似映射  $\phi$ , 即

$$C = \phi(k) = 2^{(\gamma \times k - b)}, \quad (7)$$

式中:  $\gamma \times k - b$  是最简单的线性函数。

然后, 给定通道维数  $C$  就可以自适应确定核大小  $k$ ,  $k$  的表达式为

$$k = \phi(C) = \left\lfloor \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}}, \quad (8)$$

式中: odd 表示最接近的奇数。在实验中分别设  $\gamma=2$ 、 $b=1$ 。通过  $\phi$  映射, 高维通道具有更大范围的相互作用, 低维通道通过非线性映射具有更小范围的相互作用。

### 2.4 非对称卷积

非对称卷积是指在普通方形卷积核中加入水平卷积和垂直卷积, 其原理如图 5 所示, 将每个  $3 \times 3$  卷积核替换为 1 个 ACB。ACB 由 3 个并行层组成, 分别是内核大小为  $3 \times 3$ 、 $1 \times 3$  和  $3 \times 1$  的卷积核。3 层分别进行批量归一化, 并将 3 层的输出求和得到

新的输出。ACB可以匹配各种网络结构以提升其性能并增强模型的鲁棒性,且没有引入需要调整的超

参数,因此在提升性能的同时不需要额外的推理计算时间。

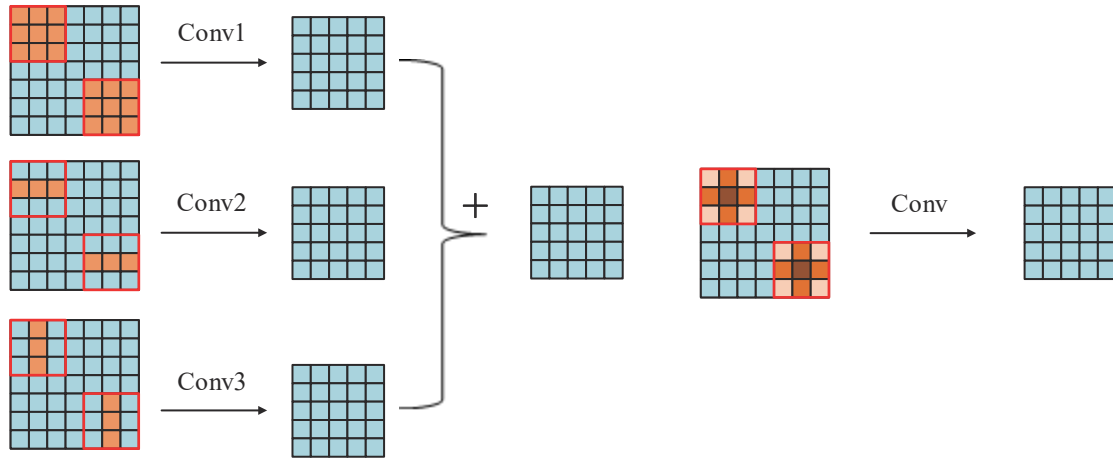


图5 非对称卷积原理图

Fig. 5 Schematic diagram of asymmetric convolution

所提 DECA-Net 结构中每一个卷积单元都包含 1 个卷积核大小为  $3 \times 3$  的卷积层。为提升网络的性能,增强网络的特征提取能力,将 DECA-Net 中所有的  $3 \times 3$  的方形卷积核都替换为 ACB。

## 2.5 多标签损失函数

Chest X-ray 14 数据集有以下特性:1)许多病理在视觉上是相似的;2)单次扫描中可能出现大量疾病,包括肺癌、肺结核和肺炎;3)该数据集属于不平衡数据集。对于上述问题,采用 MSML 来计算每种疾病的损失,即

$$E_{\text{MSML}} = \frac{1}{|Y_i|} \sum_{i \in Y_i} \frac{\exp x_i^i}{\exp x_i^i + \sum_{k \in \bar{Y}_i} \exp x_k^i}, \quad (9)$$

式中:  $Y_i$  表示当前样本的真实值集合;  $\bar{Y}_i$  表示当前样本预测值的集合;  $x_i^i$  表示真实值集合中样本的输入特征;  $x_k^i$  表示预测值集合中样本的输入特征。

MSML 可以解释为将多标签学习问题分解为多个独立的分类问题,同时能够学习每一类的单独分布。MSML 内部的 Sofamax 属性对于在指数大小的输出空间中促进学习过程至关重要。此属性可用于多标签空间学习,预测多个标签,还使该模型减轻了负类的过度拟合,使得难识别的疾病乘以一个较大的权重,易识别的疾病乘以一个较小的权重,以得到对疾病分类能力更强的网络模型。多标签损失函数为获得更广泛的医学数据的性能提供了新的方向。

## 3 实验与分析

### 3.1 数据集与预处理

实验使用由 NIH 临床中心<sup>[9]</sup>整理并公开的多标签数据集 Chest X-ray 14, 该数据集包含 30805 名患者的 112120 张正面视图的 X 射线图像及利用 NLP 从相关放射学报告挖掘的 14 类疾病的图像标签(每个图像可

以有多个标签)。数据集含有 14 类常见的肺部疾病,包括肺不张(atelectasis)、纤维化(fibrosis)、浸润(infiltration)、肺气肿(emphysema)、肺炎(pneumonia)、肺结节(nodule)等。由于新冠肺炎疫情的出现,将新冠肺炎也加入肺部疾病的分类问题中。因此,在 Chest X-ray 14 数据集的基础上增添 760 张新冠肺炎的 X 光射线图像<sup>[17]</sup>, 构建 Chest X-ray 15 数据集。图 6 展示了 Chest X-ray 15 中部分疾病的 X 射线图像。

Chest X-ray 15 数据集中的图像为灰度图,大小为  $1024 \text{ pixel} \times 1024 \text{ pixel}$ , 为适配网络模型同时减少计算量提高网络计算速率,将每张图像缩放至  $256 \text{ pixel} \times 256 \text{ pixel}$ , 并将其转换为 RGB 三通道格式,再随机选取中心点将图像裁剪为  $224 \text{ pixel} \times 224 \text{ pixel}$  大小并进行随机水平旋转实现数据增强。为改善引入数据对整个数据集造成的样本不平衡问题,对新引入的 760 张 X 光图像的对比度与饱和度分别随机增强  $85\% \sim 115\%$ , 将图像依 0.5 的概率水平翻转对数据进行扩增。最后将图像转换为向量格式并对其进行像素归一化处理。

### 3.2 参数设置与训练

为保证实验结果的准确性,将数据集分为训练集、验证集和测试集,比例为 7:2:1, 进行交叉验证以应对测试结果过于片面以及训练数据不足的问题。实验采用可以自适应调整学习率大小的 Adam 优化器, Adam 可根据损失函数的变化自适应地调整学习率的大小,当验证损失达到稳定时停止训练。在参数设置上,初始学习率设为 0.001, 动量参数设置为 0.9。为充分利用显存资源,加快训练速度,本次实验输入数据的 batch size 设为 192。

### 3.3 性能评价指标

采用接受者操作特性曲线(ROC)及其曲线下的面积(AUC)作为网络性能评价指标。ROC 是一个用

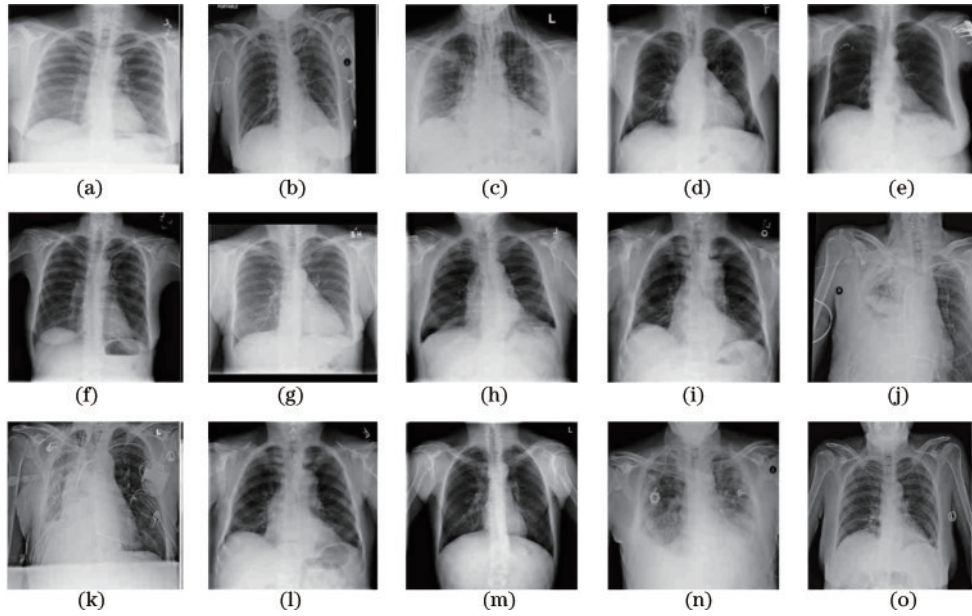


图 6 Chest X-ray 15 数据集中的 X 射线图像。(a) 未找到; (b) 肺炎; (c) 新冠肺炎; (d) 心脏扩大; (e) 疝气; (f) 浸润; (g) 肺结节; (h) 肺气肿; (i) 积液; (j) 胸膜增厚; (k) 气胸; (l) 肿块; (m) 纤维化; (n) 水肿; (o) 肺炎实变

Fig. 6 X-ray images in Chest X-ray 15 dataset. (a) No finding; (b) pneumonia; (c) COVID-19; (d) cardiomegaly; (e) hernia; (f) infiltration; (g) nodule; (h) emphysema; (i) effusion; (j) pleural thickening; (k) pneumothorax; (l) mass; (m) fibrosis; (n) edema; (o) consolidation

于度量分类中的非均衡性的工具,比 precision-recall ( $P-R$ ) 曲线有更好的稳定性,ROC 曲线上的每个点反映着对同一信号刺激的感受性,其横轴为负正类率 ( $R_{FPR}$ ),纵轴为真正类率 ( $R_{TPR}$ )。  $R_{FPR}$  表示所有的负样本中被预测为正例的占比,  $R_{TPR}$  表示所有的样本中被正确预测出来的占比,具体计算公式为

$$R_{TPR} = \frac{N_{FP}}{N_{FP} + N_{TN}}, \quad (10)$$

$$R_{TPR} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (11)$$

式中:  $N_{FP}$ 、 $N_{TP}$ 、 $N_{FN}$ 、 $N_{TN}$  分别表示样本中负正类、真正类、假负类、真负类的数量。

ROC 曲线反映分类器的分类能力,但 ROC 曲线会出现交叉的现象,有时仅通过观察 ROC 曲线并不能清晰地说明哪个分类器的效果更好。而 AUC 作为一个数值,能够量化 ROC 曲线的分类能力,其取值范围为  $[0.5, 1]$ ,表示分类时对单个样本预测为正类的概率大于预测为负类的概率的可能性。ROC 曲线越靠近左上角,AUC 值越大,相应疾病的分类准确度越高,说明网络对该类疾病的分类性能越好。取 15 种疾病 AUC 值的算术平均作为平均 AUC 值,平均 AUC 值是用来衡量网络整体分类性能的评价指标。

### 3.4 结果与分析

根据所提算法在 Chest X-ray 15 数据集上的分类结果绘制的 ROC 曲线如图 7 所示,分别展示了每种疾病的分类 ROC 曲线与 15 种疾病的算术平均值的 ROC 曲线,右下角给出了具体的 AUC 值。可见,曲线大部分位于左上方,且大部分疾病的 ROC 曲线都在算术平均 AUC 值

的曲线的左上方,表明该网络模型具有良好的分类性能。

为分析不同网络与算法对肺部疾病分类性能的影响,采用 ResNet50 和 DenseNet121 作为基准模型,与两个注意力机制模块分别进行组合:ResNet50 和 SE 模块组合 (RSE-Net), ResNet50 和 ECA 模块组合 (RECA-Net), DenseNet121 和 SE 模块组合 (DSENet121<sup>[10]</sup>), DenseNet121 和 ECA 模块组合 (DECA-Net) 即所提网络。在 Chest X-ray 15 数据集上进行 40 轮训练来测试各个组合模型与基准模型分类效果,结果如表 2 所示。

表 2 不同分类网络模型在 Chest X-ray 15 上的分类结果对比  
Table 2 Comparison of classification results of different classification network models on Chest X-ray 15

Network	Backbone	Average AUC
ResNet50	ResNet50	0.7468
ResNet50+SE	ResNet50	0.7642
ResNet50+ECA	ResNet50	0.7886
DenseNet121	DenseNet121	0.7952
DenseNet121+SE <sup>[10]</sup>	DenseNet121	0.8014
DECA-Net	DenseNet121	<b>0.8245</b>

由表 2 可知:相较于 ResNet50、ResNet50+SE、ResNet50+ECA、DenseNet121、DenseNet121+SE,所提 DECA-Net 在平均 AUC 值达到 0.8245,比单一模型 DenseNet121 提高 2.93 个百分点;比融合 SE 模块的 DenseNet121 的平均 AUC 值提高 2.31 个百分点,说明 ECA 模块性能优于 SE 模块;比融合 ECA 模块的 ResNet 的平均 AUC 高 3.59 个百分点,说明

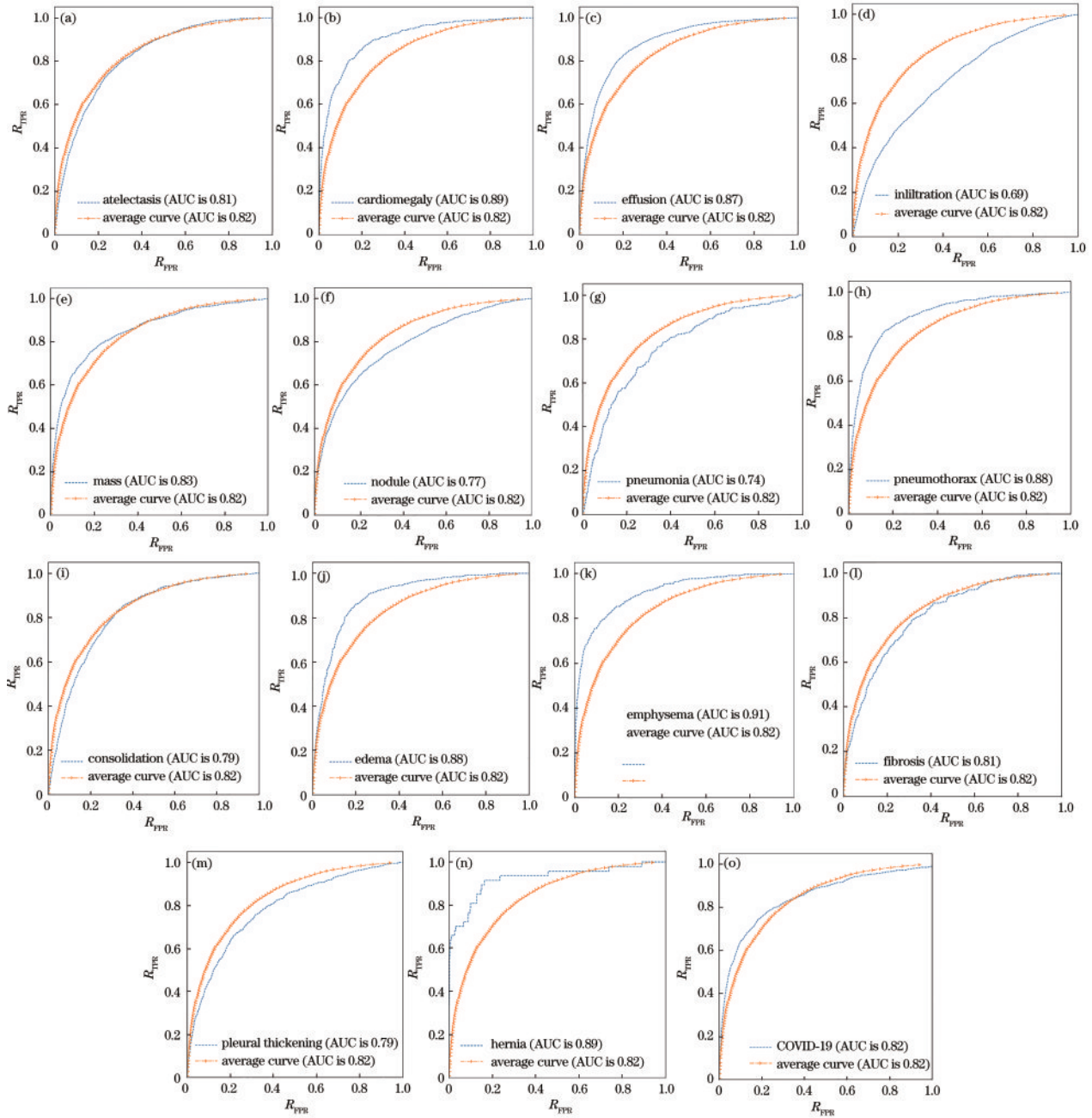


图 7 每种疾病和算术平均值的 ROC 曲线。(a)肺不张;(b)心脏扩大;(c)积液;(d)浸润;(e)肿块;(f)肺结节;(g)肺炎;(h)气胸;(i)肺实变;(j)水肿;(k)肺气肿;(l)纤维化;(m)胸膜增厚;(n)疝气;(o)新冠肺炎

Fig. 7 ROC curve and AUC value of proposed algorithm. (a) Atelectasis; (b) cardiomegaly; (c) effusion; (d) infiltration; (e) mass; (f) nodule; (g) pneumonia; (h) pneumothorax; (i) consolidation; (j) edema; (k) emphysema; (l) fibrosis; (m) pleural thickening; (n) hernia; (o) COVID-19

DenseNet在该数据集的分类性能优于ResNet。

表 3 为所提网络模型在 Chest X-ray 14 数据集上的分类 AUC 值与其他研究人员的研究结果比较,实验结果表明,所提算法对于 14 种胸部疾病分类的平均 AUC 值达到 0.8237, 优于 Wang 等<sup>[9]</sup>、Ma 等<sup>[10]</sup>、张智睿等<sup>[12]</sup>、Gündel 等<sup>[18]</sup>、Guan 等<sup>[19]</sup>、Teixeira 等<sup>[20]</sup>、Chen 等<sup>[21]</sup>的研究。就每种疾病的 AUC 值来看,所提算法对肺气肿的分类准确率最高,其 AUC 值达到 0.9027,但仍低于 Chen 等<sup>[21]</sup>的算法。但肺不张(atelectasis)和气

胸(pneumothorax)两种疾病的 AUC 值均高于以上几种算法。总体来看,所提分类算法对肺部疾病的分类问题上总体分类性能较好。

### 3.5 消融实验与分析

为验证模型中每个模块的有效性,在保持基本网络结构完整的前提下,在 Chest X-ray 15 数据集上进行了消融实验。实验方案如下:设置 4 组消融实验, Network\_1 代表基准模型, Network\_2 代表移除 ECA 模块后的模型, Network\_3 代表移除非对称卷积采用

表 3 不同胸部疾病分类算法在 Chest X-ray 14 数据集上 AUC 值的比较

Table 3 Comparison of AUC value of different chest disease classification algorithms on Chest X-ray 14 dataset

Disease	Algorithm of reference [9]	Algorithm of reference [10]	Algorithm of reference [13]	Algorithm of reference [18]	Algorithm of reference [19]	Algorithm of reference [20]	Algorithm of reference [21]	Proposed algorithm
Atelectasis	0.7003	0.7627	0.785	0.767	0.783	0.791	0.785	<b>0.8157</b>
Cardiomegaly	0.81	0.8835	0.8766	0.883	0.884	<b>0.898</b>	0.887	0.8657
Effusion	0.7585	0.8159	0.8628	0.828	0.832	<b>0.873</b>	0.831	0.8701
Infiltration	0.6614	0.6786	0.673	0.709	<b>0.708</b>	0.700	0.703	0.6948
Mass	0.6933	0.8012	0.804	0.821	<b>0.837</b>	0.832	0.833	0.8350
Nodule	0.6687	0.7293	0.7299	0.758	<b>0.800</b>	0.758	0.798	0.7683
Pneumonia	0.658	0.7097	0.7423	0.731	0.735	<b>0.767</b>	0.731	0.7548
Pneumothorax	0.7993	0.8377	0.8426	0.846	0.866	0.859	0.881	<b>0.8687</b>
Consolidation	0.7032	0.7443	0.7846	0.745	0.746	<b>0.800</b>	0.754	0.7952
Edema	0.8052	0.8414	0.8727	0.835	0.841	<b>0.889</b>	0.849	0.8647
Emphysema	0.833	0.8836	0.858	0.895	<b>0.937</b>	0.891	0.930	0.8942
Fibrosis	0.7859	0.8007	0.7754	0.818	0.82	0.789	<b>0.833</b>	0.8141
Pleural thickening	0.6835	0.7536	0.7563	0.761	<b>0.796</b>	0.771	0.782	0.7872
Hernia	0.8717	0.8763	0.8645	0.896	0.895	0.896	<b>0.921</b>	0.9027
Mean	0.7451	0.7941	0.802	0.807	0.82	0.822	0.823	<b>0.8237</b>

表 4 消融实验结果对比

Table 4 Comparison of ablation experiment results

Disease	Network_1	Network_2	Network_3	Network_4	DECA-Net
Atelectasis	0.7853	0.7838	0.7974	0.8073	<b>0.8101</b>
Cardiomegaly	0.8770	0.8702	0.8813	0.8877	<b>0.8919</b>
Effusion	0.8541	0.8498	0.8614	0.8689	<b>0.8701</b>
Infiltration	0.6696	0.6724	0.6845	<b>0.6946</b>	0.6892
Mass	0.8128	0.8090	<b>0.8306</b>	0.8302	0.8301
Nodule	0.739	0.7244	0.7595	0.7534	<b>0.7672</b>
Pneumonia	0.7282	0.7355	<b>0.7450</b>	0.7412	0.7389
Pneumothorax	0.8546	0.8623	0.8743	0.8806	<b>0.8829</b>
Consolidation	0.7818	0.7849	0.7908	0.7975	0.7931
Edema	0.8633	<b>0.8872</b>	0.8726	0.8805	0.8782
Emphysema	0.8742	0.8829	0.8924	0.9023	<b>0.9122</b>
Fibrosis	0.7722	0.7892	0.7859	0.796	<b>0.8054</b>
Pleural thickening	0.7607	0.7788	0.7627	0.7707	<b>0.7851</b>
Hernia	0.8571	<b>0.9011</b>	0.9004	0.8678	0.8893
COVID-19	0.8265	0.8157	0.8273	<b>0.8357</b>	0.8239
Mean	0.8038	0.8098	0.8177	0.8210	<b>0.8245</b>

普通  $3 \times 3$  的模型, Network\_4 代表移除多标签损失函数采用焦点损失函数的模型。4 组消融实验以及所提算法结果如表 4 所示。

由表 4 可知, 采用的 3 个模块对整体网络性能均有提升, 当 3 个模块组合使用对整体性能提升最为显著。移除 ECA 模块后, 模型的平均 AUC 值下降约 0.0147, 表明 ECA 模块的应用能够有效捕获跨通道之间的信息交互; 移除 ACB 模块后, 平均 AUC 值下降约 0.068, 表明 ACB 模块能够增强网络的特征提取能力; 移除多

标签损失函数而使用焦点损失函数代替后, 平均 AUC 值下降约 0.0035, 幅度较小, 但也能说明多标签损失函数有助于网络对疾病进行识别和分类。实验结果表明, 所提 DECA-Net 模型采用的 3 个模块均有助于提升模型性能。

## 4 结 论

针对目前对肺部疾病分类算法分类准确率不够高的问题, 提出一种基于高效通道注意力机制的胸部 X



光片疾病分类算法(DECA-Net)。该算法主要有以下 3 个方面的创新:1)在 DSENet121 的基础上引入高效通道注意力机制模块,增强特征通道信息的信息表达能力且考虑了跨通道之间的信息交互,使得网络对特征的提取能力进一步加强,使整体平均 AUC 提升 1.47 个百分点;2)在密集高效通道注意力块中采用非对称卷积,非对称卷积没有引入需要调整的超参数,在提升性能的同时不需要额外的推理计算时间,使整体平均 AUC 提升 0.68 个百分点;3)引入多标签损失函数——MSML 来解决多标签和数据不平衡的问题,使整体平均 AUC 值提升了 0.35 个百分点。实验结果表明,所提模型对胸部 X 光片疾病分类平均 AUC 达到 0.8245,说明该模型具有较高的分类准确度,并具有较强的实用性与可推广性。

## 参 考 文 献

- [1] Thompson R. Pandemic potential of 2019-nCoV[J]. *The Lancet: Infectious Diseases*, 2020, 20(3): 280.
- [2] 宋瑶, 刘俊. 改进 U-Net 的新冠肺炎图像分割方法[J]. *计算机工程与应用*, 2021, 57(19): 243-251.  
Song Y, Liu J. Improved U-net network for COVID-19 image segmentation[J]. *Computer Engineering and Applications*, 2021, 57(19): 243-251.
- [3] 李士雪, 单莹. 新型冠状病毒肺炎研究进展述评[J]. *山东大学学报(医学版)*, 2020, 58(3): 19-25.  
Li S X, Shan Y. Latest research advances on novel coronavirus pneumonia[J]. *Journal of Shandong University (Health Sciences)*, 2020, 58(3): 19-25.
- [4] Khobragade S, Tiwari A, Patil C Y, et al. Automatic detection of major lung diseases using Chest Radiographs and classification by feed-forward artificial neural network [C]//2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems, July 4-6, 2016, Delhi, India. New York: IEEE Press, 2016.
- [5] 陈文豪, 何敬, 刘刚. 引入注意力机制的卷积神经网络高光谱图像分类[J]. *激光与光电子学进展*, 2022, 59(18): 1811001.  
Chen W H, He J, Liu G. Convolutional neural networks for hyperspectral image classification with attention mechanism[J]. *Laser & Optoelectronics Progress*, 2022, 59(18): 1811001.
- [6] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [7] Li X, Wang W H, Hu X L, et al. Selective kernel networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 510-519.
- [8] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 3-19.
- [9] 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.
- [10] Ma Y B, Zhou Q H, Chen X S, et al. Multi-attention network for thoracic disease classification and localization [C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing, May 12-17, 2019, Brighton, UK. New York: IEEE Press, 2019: 1378-1382.
- [11] Yao L, Prosky J, Poblenz E, et al. Weakly supervised medical diagnosis and localization from multiple resolutions[EB/OL]. (2018-03-21)[2020-08-23]. <https://arxiv.org/abs/1803.07703>.
- [12] 张智睿, 李镛, 关欣. 密集挤压激励网络的多标签胸部 X 光片疾病分类[J]. *中国图象图形学报*, 2020, 25(10): 2238-2248.  
Zhang Z R, Li Q, Guan X. Multilabel chest X-ray disease classification based on a dense squeeze-and-excitation network[J]. *Journal of Image and Graphics*, 2020, 25(10): 2238-2248.
- [13] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [14] Ding X H, Guo Y C, Ding G G, et al. ACNet: strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 1911-1920.
- [15] 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.
- [16] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [17] Fares S, Gaurav D, Omar S. Kaggle-chest X-ray (Covid-19 & Pneumonia)[EB/OL]. [2022-02-05]. <https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia>.
- [18] Gündel S, Grbic S, Georgescu B, et al. Learning to recognize abnormalities in chest X-rays with location-aware dense networks[M]//Vera-Rodriguez R, Fierrez J, Morales A. *Progress in pattern recognition, image analysis, computer vision, and applications. Lecture notes in computer science*. Cham: Springer, 2019, 11401: 757-765.

- [19] Guan Q J, Huang Y P, Luo Y W, et al. Discriminative feature learning for thorax disease classification in chest X-ray images[J]. IEEE Transactions on Image Processing, 2021, 30: 2476-2487.
- [20] Teixeira V, Braz L, Pedrini H, et al. DualANet: dual lesion attention network for thoracic disease classification in chest X-rays[C]//2020 International Conference on Systems, Signals and Image Processing (IWSSIP), July 1-3, 2020, Niteroi, Brazil. New York: IEEE Press, 2020: 69-74.
- [21] Chen B Z, Zhang Z, Lin J Y, et al. Two-stream collaborative network for multi-label chest X-ray image classification with lung segmentation[J]. Pattern Recognition Letters, 2020, 135: 221-227.