

文章编号 1004-924X(2023)07-1074-11

## 基于密集双流聚焦网络的肺炎辅助诊断模型

周涛<sup>1,3</sup>, 叶鑫宇<sup>1,3\*</sup>, 陆惠玲<sup>2</sup>, 刘贇璨<sup>1,3</sup>, 常晓玉<sup>1,3</sup>

(1. 北方民族大学 计算机科学与工程学院, 宁夏 银川 750021;

2. 宁夏医科大学 理学院, 宁夏 银川 750003;

3. 北方民族大学 图像图形智能处理国家民委重点实验室, 宁夏 银川 750021)

**摘要:** X 光片对肺炎疾病的诊断具有重要作用,但其成像时易受噪声污染,导致肺炎疾病的影像学特征不明显和病灶特征提取不充分。针对上述问题,提出密集双流聚焦网络 DDSF-Net 的肺炎辅助诊断模型。首先设计残差多尺度块,利用多尺度策略提高网络对医学影像中不同尺寸肺炎病灶的适应性,采用残差连接提高网络参数的传递效率;然后设计双流密集块,采用全局信息流和局部信息流并行结构的密集单元,其中 Transformer 对全局上下文语义信息进行学习,卷积层进行局部特征提取,利用密集连接方式实现两种信息流的深浅层特征融合;最后,设计具有中心注意操作和邻近插值操作的聚焦块,利用裁剪医学影像尺寸来过滤背景噪声信息,利用插值对医学图像进行放大,增强病灶的细节特征。在肺炎 X 光片数据集中与典型模型进行对比,本文模型的准确率、精确率、召回率、F1、AUC 值和训练时间分别为 98.12%, 98.83%, 99.29%, 98.71%, 97.71% 和 15 729 s,准确率和 AUC 值较密集网络分别提升了 4.89% 和 4.69%。DDSF-Net 能够有效缓解肺炎影像学特征不明显和病灶特征提取不充分的问题,通过热力图和三份公共数据集进一步验证了本文模型的有效性和鲁棒性。

**关键词:** 医学图像处理;肺炎 X 光片;密集网络;残差多尺度块;全局和局部信息流;聚焦块

**中图分类号:** TP394.1; TH691.9 **文献标识码:** A **doi:** 10.37188/OPE.20233107.1074

## Pneumonia aided diagnosis model based on dense dual-stream focused network

ZHOU Tao<sup>1,3</sup>, YE Xinyu<sup>1,3\*</sup>, LU Huiling<sup>2</sup>, LIU Yuncan<sup>1,3</sup>, CHANG Xiaoyu<sup>1,3</sup>

(1. College of Computer Science and Engineering, North Minzu University, Yinchuan 750021, China;

2. College of Science, Ningxia Medical University, Yinchuan 750003, China;

3. Key Laboratory of Image and Graphics Intelligent Processing of State Ethnic Affairs Commission, North Minzu University, Yinchuan 750021, China)

\* Corresponding author, E-mail: 3303626778@qq.com

**Abstract:** X-ray images play an important role in the diagnosis of pneumonia disease, but they are susceptible to noise pollution during imaging, resulting in the imaging features of pneumonia being inconspicuous and an insufficient extraction of lesion features. A dense dual-stream focused network DDSF-Net is proposed in this paper for the development of an aided diagnosis model for pneumonia to address the above-

**收稿日期:** 2020-09-28; **修订日期:** 2020-10-28.

**基金项目:** 国家自然科学基金资助项目(No. 62062003); 宁夏自然科学基金资助项目(No. 2022AAC03149); 北方民族大学 2022 年研究生创新项目(No. YCX22198)

mentioned problems. The main steps of this method are as follows. First, a residual multi-scale block is designed, a multi-scale strategy is used to improve the adaptability of the network to different sizes of pneumonia lesions in medical images, and a residual connection is used to improve the efficiency of the network parameter transfer. Secondly, a dual-stream dense block is designed, a dense unit with a parallel structure for the global information stream and the local information stream is used, whereby the transformer learns global contextual semantic information. The convolutional layer performs local feature extraction, and a deep and shallow feature fusion of the two information streams is achieved using a dense connection. Finally, focus blocks with central attention operation and neighborhood interpolation operation are designed, background noise information is filtered by cropping the medical image size, and detailed features of lesions are enhanced by interpolating the medical images with magnification. In comparison with typical models used for a pneumonia X-ray dataset, the model introduced in this paper obtained better performance with a 98.12% accuracy, 98.83% precision, 99.29% recall, 98.71% F1, 97.71% AUC and 15729 s training time. Compared with DenseNet, ACC and AUC were improved by 4.89% and 4.69%, respectively. DDSF-Net effectively alleviates the problems of inconspicuous pneumonia imaging features and insufficient extraction of lesion features. The validity of this model and robustness of this paper are further verified by a heat map and three public datasets.

**Key words:** medical image processing; pneumonia X-ray images; dense network; residual multi-scale block; global and local information flow; focus block

## 1 引言

世界卫生组织指出,由于世界各地缺乏医疗资源和专业人员,超过1.5亿人感染肺炎<sup>[1]</sup>。肺部疾病主要通过X光片观察胸腔后前、前后和侧向投影内的异常病理模式来诊断<sup>[2]</sup>,X光片是肺炎检测中有效且极具成本效益的诊断工具,肺炎(尤其是新冠肺炎<sup>[3]</sup>)是一种致命的疾病,临床上通过观察气腔混浊、肺叶实变和间质混浊<sup>[3]</sup>进行鉴别,但X光片存在易受过厚软组织影响导致不透明度增加的问题,在人工诊断中不可避免导致误诊。而计算机辅助诊断有助于协助放射科医生识别X光片,其中可自动学习特征的深度学习密集网络(DenseNet)<sup>[4]</sup>模型是协助医生进行精准肺炎诊断的重要手段。Jain等<sup>[5]</sup>使用迁移学习的VGG, ResNet, Inceptionv3和DenseNet创建肺炎检测模型,在X光片识别肺炎中优势显著。Rajpurkar等<sup>[6]</sup>在大型公开X光片数据集ChestX-ray14上利用DenseNet实现了超过医师平均表现的优异性能。肺部各种感染具有相似的体征和症状,肺炎的放射学特征通常不是很典型,计算机辅助诊断模型在特定病因预测方面的作用有限<sup>[7]</sup>,容易出现诊断精度低且不稳定的情况。周涛等<sup>[4]</sup>使用预

训练的DenseNet提取深层特征,并结合非负稀疏协同表示分类器和残差相似度进行肺部疾病分类识别。Chen等<sup>[8]</sup>结合DenseNet和ResNet对早期特征进行重用和挖掘,捕获肺炎X光片中更具判别的特征。Li等<sup>[9]</sup>在X光片乳腺癌良恶性分类中将DenseNet第一个卷积层替换为多尺度结构,增加网络对不同规模特征的适应性,使模型更具泛化性和鲁棒性。Chen等<sup>[10]</sup>将DenseNet中密集块进行多个并行,提高精度的同时保持了训练稳定性。陈筱等<sup>[11]</sup>利用迁移学习的EfficientNet进行特征提取和疾病分类,并通过加权梯度类激活映射可视化解释模型诊断异常的原因。Liu等<sup>[12]</sup>利用过渡层和中心裁剪操作并行的DenseNet,生成丰富的多尺度肺结节特征,保存结核形态细节特征并过滤无关背景信息。Priya等<sup>[13]</sup>利用加权损失熵DenseNet缓解类别不平衡的问题,在X光片的胸部疾病检测中取得了较好性能。

基于密集网络的胸部X光片的计算机辅助诊断模型是近几年的研究热点。但是根据上述文献报道以及X光片自身的特点,该诊断模型还存在一些不足。具体来讲,从问题的角度来看,X光片在成像时易受噪声污染,导致肺炎影像学特征不明显和病灶特征提取不充分;从方法的角度

来看,密集网络存在梯度易消失和特征传递不足,以及难以捕获全局上下文语义信息的问题。本文从问题角度出发,充分考虑密集网络的特点,提出一种基于密集双流聚焦网络(Dense Dual-Stream Focused Network, DDSF-Net)的肺炎 X 光片计算机辅助诊断模型。针对深度神经网络的梯度易消失和特征传递不足的问题,设计残差多尺度块,替代 DenseNet 初始层,提高网络对医学影像中不同尺寸肺炎的适应性,其中多尺度可以聚集相关性强的特征以增强病灶信息,残差连接用于提高参数传递效率;针对 DenseNet 难以捕获全局上下文语义信息的问题,设计双流密集块,采用全局信息流和局部信息流并行的密集单元,具有长距离依赖属性的 Transformer 实现全局信息的提取和融合,卷积层实现局部信息的学习,采用密集连接方式通过学习前续密集单元的特征,让每个密集单元都可以直接利用损失和梯度,避免梯度消失,实现全局和局部信息流的深浅层特征融合;由于部分肺炎病灶区域小且具

有大量背景信息,病变特征难以捕获,如肺炎常见的多发小斑片影或间质改变等影像特征,设计具有中心注意操作和邻近插值操作的聚焦块,对 X 光片进行裁剪,插值放大影像尺寸,增强肺炎的影像学特征,保存病灶的形态细节并过滤无关背景信息。

## 2 原理

DenseNet121<sup>[4]</sup>网络共包含 4 个密集块,4 个密集块由不同数量的密集单元通过密集连接的方式组合而成,即任意两个密集单元之间进行直接连接。4 个密集块的密集单元数分别为 6, 12, 24, 16, 由批量归一化 BN、激活函数 ReLU、 $1 \times 1$  卷积, BN, ReLU,  $3 \times 3$  卷积级联而成。本文在 DenseNet121 基础上提出肺炎分类算法 DDSF-Net, 整体框架如图 1(a)所示,图 1(b)表示残差多尺度块;图 1(c)表示双流密集块的单层密集单元结构,由卷积路径和 Transformer 路径并行构成;

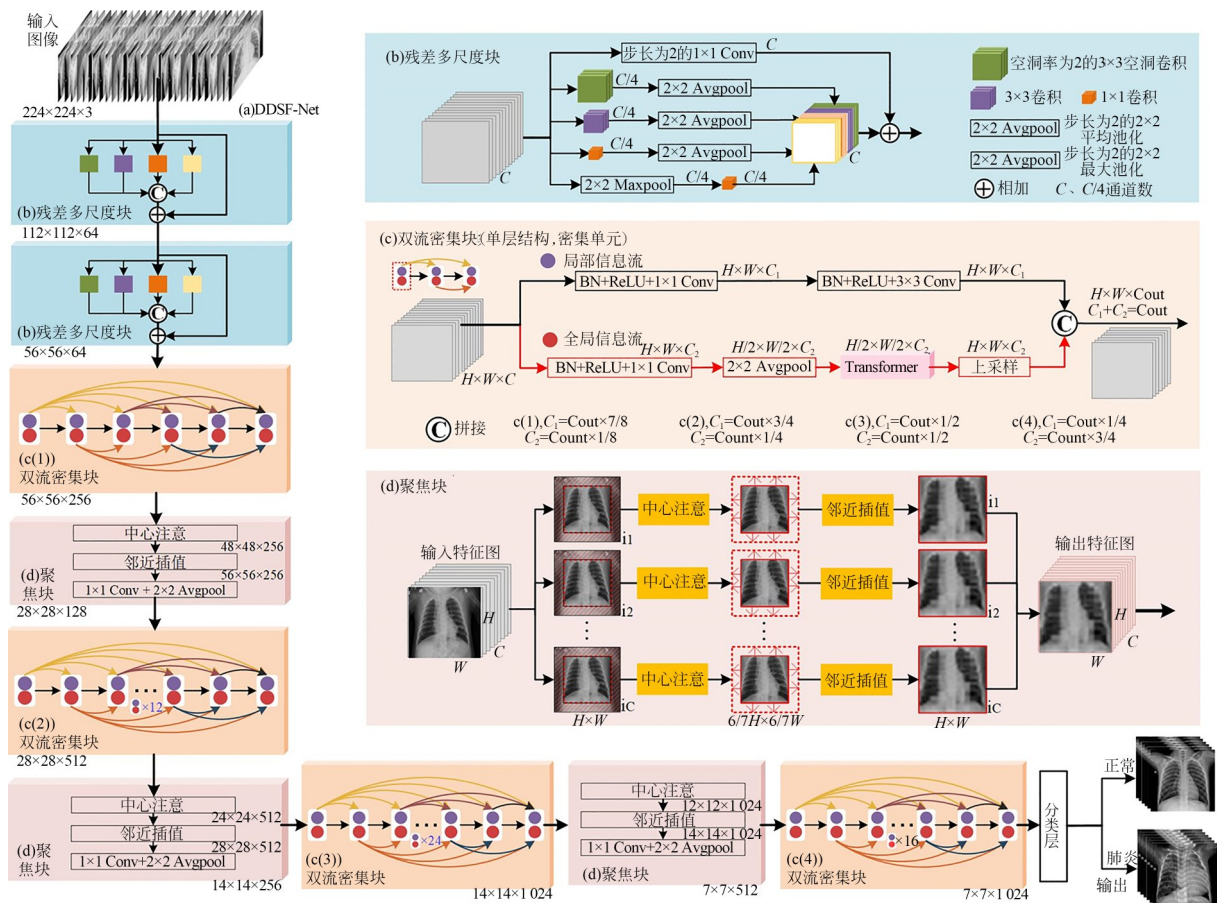


图1 DDSF-Net 整体框架  
Fig. 1 Overall framework of DDSF-Net

图 1(d)表示聚焦块,由中心注意操作和邻近插值操作串行构成。DDSF-Net 的具体操作流程如下:利用 2 个残差多尺度块对输入图像进行多尺度特征提取,然后利用串联的 4 个密集块和 3 个聚焦块对特征的局部和全局信息进行提取和增强,最后采用全连接层和 SoftMax 构成的分类层进行肺炎识别。

### 2.1 残差多尺度块

残差多尺度块是在多分支多尺度结构的基础上加入池化层和残差连接,如图 1(a)所示。2 个残差多尺度块替代 DenseNet 初始的  $7 \times 7$  卷积层和  $3 \times 3$  最大池化层。具体结构如图 1(b)所示。残差多尺度块采用多尺度四分支并行,同时使用 2 步长  $1 \times 1$  卷积实现残差连接,提高参数传递效率和缓解梯度弥散问题,四分支分别为扩张率为 2 的  $3 \times 3$  卷积、 $3 \times 3$  卷积、 $1 \times 1$  卷积和  $2 \times 2$  最大池化与  $1 \times 1$  卷积,每个分支的输出通道数为总输出通道数的八分之一,不同感受野的卷积有效地扩展网络宽度,提取多尺度特征,提高网络对不同规模特征的适应性;并采用步长为 2 的  $2 \times 2$  平均池化层,降低模型复杂度同时学习多尺度特征,防止网络过拟合。

残差多尺度块结合多尺度与残差连接的优势,利用不同感受野实现网络对医学影像中不同尺寸肺炎的适应性,还将相关性强的输出特征聚集,使得原本呈现稀疏矩阵分布的特征可以分解为多个密集分布的子集,减少冗余信息并提高计算性能;残差连接通过拟合残差项缓解梯度弥散和加速网络收敛,从而提高特征学习和模型表达的能力。

### 2.2 双流密集块

卷积神经网络通过卷积核对局部信息进行学习和特征提取,但卷积局部性限制了其学习全局上下文信息的能力。Naseer 等<sup>[14]</sup>的研究表明 Transformer 较 CNN 不依赖局部信息,灵活和动态感受野可以更好地提取有效特征,实现全局特征之间相互依赖关系的建模。利用 Transformer 计算权重来表达每个特征的不同重要性,但在网络浅层中引入,会导致模型参数量和计算量较大,以及丢失一些细节信息。本文设计双流密集块,将密集块中密集单元改进为双分支并行结构,分别是捕获局部信息流的卷积层和捕获全局信息流的 Transformer 层,结构如图 2 所示。两分

支的通道数按一定比例进行分配,充分结合卷积层和 Transformer 的优势,灵活高效地学习 X 光片中的肺炎特征,提高语义判别能力。同时,采用密集连接方式通过学习前续密集单元的特征,让每个密集单元都可以直接利用损失和梯度,避免梯度消失,实现全局和局部信息流的深浅层特征融合和隐性深度监督<sup>[15]</sup>,降低计算量的同时使网络学习全局和局部特征的相关性。

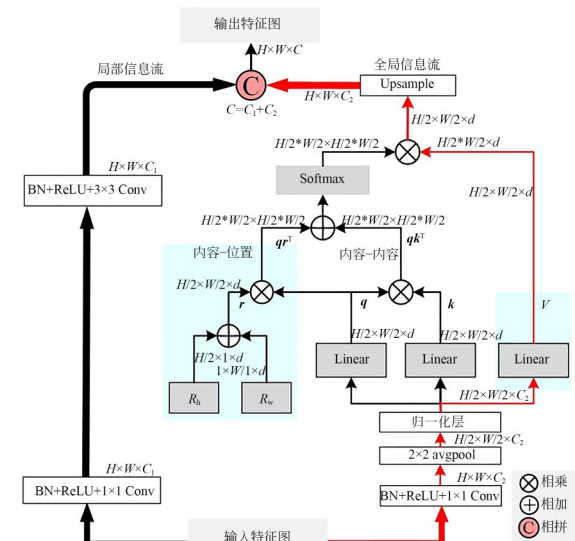


图 2 双流密集块(单层结构)

Fig. 2 Dual-stream dense block (single layer structure)

如图 2 所示,左分支卷积层学习局部信息流,右分支 Transformer 层学习全局信息流,卷积分支采用与密集网络相同的结构,由于 Transformer 直接对全部特征进行相关性计算,所需要学习的超参数更多,而浅层特征图中每个特征包含的信息量较少,过多使用 Transformer 会导致模型难以学到关键信息,还会导致过多的资源消耗。因此,两个分支在 4 个密集块中分别按 7 : 1, 3 : 1, 1 : 1 和 1 : 3 的比例,对全局和局部信息流的输出特征图通道数进行分配。

Transformer 分支首先利用下采样对输入进行空间尺寸压缩,其次通过线性(Linear)层将输入特征变化为 3 个矩阵查询( $q$ )、键值( $k$ )和值( $v$ ),然后通过缩放点积<sup>[14]</sup>和 Softmax 函数对  $q$  和键值  $k$  进行计算,生成注意力权重并应用于  $v$ ,最终输出特征图  $X$  的计算式为:

$$X = \text{Soft max} \left( \frac{qr^T + qk^T}{\sqrt{d_k}} \right) v, \quad (1)$$

其中:  $d_k$  表示键值  $k$  的维度,  $qk^T + qr^T$  表示注意力得分图。Transformer 将特征图在空间维度上展平, 输入长度为特征图的高乘宽, 图 2 中  $r$  为位置编码, 由高度编码  $R_h$  和宽度编码  $R_w$  相加所得, 青色框分别代表位置编码信息和值投影。本文使用的 Transformer 利用相对距离感知进行位置编码, 不仅考虑内容信息, 还考虑不同位置特征之间的相对距离, 从而能有效地关联特征的信息与位置, 实现全部特征之间相互依赖关系的建模, 使得 Transformer 也更适合 X 光片中的肺炎特征。

### 2.3 聚焦块

肺炎疾病常采用 X 光片进行诊断, 但影像所包含信息量大、背景复杂、且部分病灶区域小、病灶空间尺寸与形态上差异大等问题, 使得网络模型难以对整个影像直接识别。考虑到病灶所在的胸腔区域通常在图像中心, 卷积、池化和拼接操作都保持了图像信息的相对位置, 从而包含肺炎特征的胸腔区域也一直在特征图中心。因此, 设计具有中心注意操作和邻近插值操作的聚焦块, 对 X 光片中心区域进一步关注, 放大中心区域来生成丰富的多尺度特征; 聚焦块采用中心注意操作将影像四周像素进行舍弃, 这个操作无需额外超参数而不需要参数量, 且可以有效地过滤无关的背景信息; 还采用邻近插值操作将影像放大到原始尺寸, 用于保存 X 光片中病变的形态细节特征, 增强病灶的影像学特征。

由于卷积网络的位置不变性, 将聚焦块应用到网络中间特征图上也是有意义的。医学影像在输入网络前经过随机裁切, 通过初始的两层残差多尺度块, 影像边缘上的信息会被充分地学习, 为避免网络过多关注病灶出现概率较小的边缘区域, 同时需要为网络提供充足的信息量用于识别病灶, 聚焦块的使用频率不应过多或过少, 而每次密集块对特征充分学习后, 使用聚焦块既可以裁切边缘增强病灶的影像学特征, 还可以降低冗余度。本文将聚焦块应用于前 3 个密集块之后, 为直观显示聚焦块的作用, 如图 3 所示, 利用一组 X 光片模拟 3 个密集块特征图通过聚焦块时得到的中间特征图。

聚焦块输入是尺寸为  $h \times w \times c$  的密集块输出特征图, 其中  $h, w$  和  $c$  分别表示特征图的长、宽

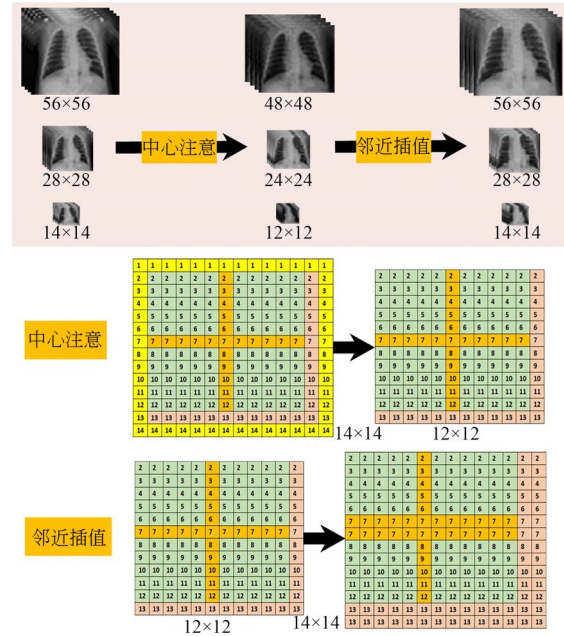


图 3 聚焦块的中间特征图

Fig. 3 Middle feature maps of focus block

和通道数, 采用中心注意操作舍弃密集块输出特征图的上、下、左、右四周各  $\frac{1}{14}h, \frac{1}{14}h, \frac{1}{14}w, \frac{1}{14}w$  像素, 此时中心注意操作的输出特征图尺寸为  $\frac{6}{7}h \times \frac{6}{7}w \times c$ , 然后采用邻近插值操作将特征图第 6 行或第 6 列的值插入到其下侧或右侧, 此时输出特征图尺寸为  $h \times w \times c$ , 网络可以继续执行其他模块, 从而实现胸腔和病灶区域的聚焦, 提升了网络的识别性能。另外, 由于肺炎病灶的空间尺寸与形态上差异大, 直径范围互异, 如细菌性肺炎阴影较实和感染区域较大、小叶性肺炎直径范围较小, 尽管通过不同感受野生成的多尺度特征对于分类任务十分重要, 但引入聚焦块, 利用特定操作直接对特征图进行中心区域放大, 生成更丰富的多尺度特征, 可以进一步提高网络的表达能力。

## 3 实验和讨论

### 3.1 数据集和实验参数

本文使用的数据集包含来自广州市妇幼保健中心整理并公开的 ChestXRy2017<sup>[16]</sup>, 包括 1 583 张正常图像和 4 273 张肺炎图像, 和来自 X

光片数据集<sup>[17]</sup>的2 313张正常图像和2 313张肺炎图像。按6:2:2比例分成训练集、验证集和测试集进行实验,使用水平或垂直随机翻转进行增强。本次实验环境为Windows Server 2019系统,256 GB内存,搭载两块3 GHz 36核处理器,并采用两块并行的TITAN V显卡加速图像处理,基于GPU的pytorch框架搭建网络,Adam优化器进行优化,采用0.01的初始学习率和每10周期0.9的衰减策略,采用0.2的随机丢弃率防止过拟合,设置权重衰减值为 $1 \times 10^{-4}$ ,训练周期为250,训练批处理大小为48。

### 3.2 评价指标

根据模型预测结果分类错误和正确的个数,得到真正类TP、假正类FP、假负类FN和真负类TN。准确率为全部类预测正确的比例,精确率(P)为正类且模型预测正确占所有正类的比例,召回率(R)为模型所预测出的正类占所有正类的比例,F1分数表示为:

$$F1 = \frac{2 \times P \times R}{P + R}. \quad (2)$$

ROC曲线是以敏感度即真正类率(TPR)为

纵轴、假正类率(FPR)为横轴进行绘制,TPR值等于召回率,FPR、特异度(TNR)可表示为:

$$FPR = \frac{FP}{FP + TN}, \quad (3)$$

$$TNR = \frac{TN}{FP + TN}. \quad (4)$$

将ROC曲线下面积定义为AUC(Area Under Curve),越靠近左上角AUC值越大,模型的排序能力和分类性能越好。评价指标越大表示模型越好。

### 3.4 消融实验与分析

为了评估模块的有效性,在DenseNet121基础上依次进行4组实验:实验一替换初始层为残差多尺度块;实验二添加残差多尺度块和双流密集块,以及添加残差多尺度块和通道数对半分配的双流密集块;实验三添加残差多尺度块和通道数按本文分配的双流密集块;实验四添加残差多尺度块、通道数按本文分配的双流密集块和聚焦块。消融实验结果如表1所示,不同模型的热力图<sup>[10]</sup>如图4所示。红色伪彩程度越深,表示网络对这个区域的关注度越高。

表1 消融实验结果对比

Tab. 1 Comparison of results of ablation experiments

模型	参数量 (M)	计算量 (G)	训练 时间/s	准确率/%	精确率/%	召回率/%	F1分数/%	AUC值/%
DenseNet121 <sup>[4]</sup>	60.88	5.732	16 670	93.44±2.1	96.04±1.8	92.69±1.6	94.33±1.7	93.60±2.0
实验一	58.25	5.563	16 591	94.63±1.5	96.46±1.2	94.60±1.1	95.41±1.2	94.64±1.3
实验二	81.19	4.382	15 934	96.35±1.9	97.07±1.7	96.74±1.6	96.90±1.7	96.27±2.0
实验二	76.31	4.068	15 672	96.89±1.8	97.41±1.5	97.30±1.4	97.36±1.5	96.79±1.8
实验三	76.31	4.149	15 729	98.01±1.1	98.53±0.8	98.09±0.9	98.31±0.8	97.99±1.0

实验一中,由于残差多尺度块的4条分支输出均为初始卷积层的四分之一,参数量和计算量小幅下降,准确率和AUC值提升1.27%和1.11%,多尺度特征能有效提升模型的识别能力,训练时间略微减少且评价指标波动幅度更小,残差连接可用于加速模型和稳定网络训练,从图4第三行可看出,网络对不同尺度的适应性得到提高,关注范围更大;相比于实验一,实验二的准确率和AUC值提高1.82%和1.72%,实验三的准确率和AUC值提高2.38%和2.27%,实验二和实验三同时学习肺炎的全局信息流和局

部信息流特征,并进行密集连接实现隐性深度监督,提升较为明显。采用本文2.2节方式分配通道数的性能更好,深层特征包含的信息量更多,能使Transformer更好地进行学习,也避免了难以学习到浅层特征的关键信息,模型参数量小幅增多,而Transformer并行计算效率更高,计算量分别下降21.23%和26.87%,以较高计算效率提高网络学习全局和局部信息的能力,从图4第四行可看出模型可以更好地学习病灶特征;实验三中,聚焦块操作没有引入超参数,DDSF-Net模型较实验二模型参数量不变、计算量小幅上升,

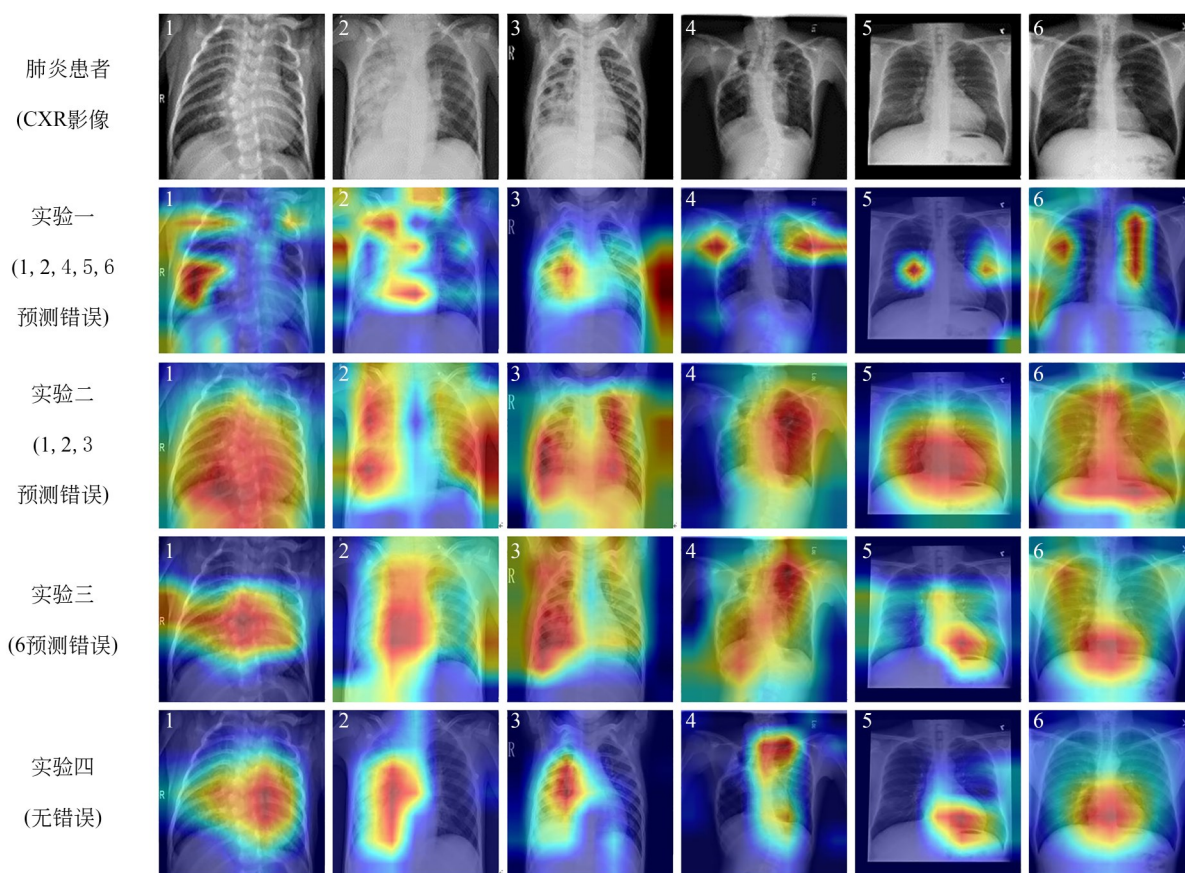


图4 各模型在肺炎X光片上的热力图

Fig. 4 Heat map of each models in pneumonia X-ray images

准确率和 AUC 值分别提升 1.15% 和 1.24%，对 X 光片进行边缘的逐步裁剪，可以增强病灶的影像学特征，从图 4 第五行可看出，网络对病灶区域关注更为精准，忽视边缘区域，能更好地学习病变特征。

### 3.5 对比实验与分析

将本文模型 DRLTransformer 与 11 个模型进行对比，在肺炎 X 光片数据集上的具体分类结果如表 2 所示。本文模型整体表现均优于其他方法，具有较好的检测精度和类别区分能力。

从表 2 可以得出，ResNet101 通过残差连接重构学习过程和重定向网络信息流，较 VGG19 性能和效率得到提升；密集网络 DenseNet121 与层数相近的残差神经网络 ResNet101 相比，在少量参数下能更精准地提取图像中具有区分度的特征，所用的时间代价也更少；SEResNet101 通过学习特征通道间相互依赖关系，以小幅时间代价获得性能提升；EfficientNetb0 和 EfficientNetb0 可从深度、宽度与分辨率进行缩放，以较高效率

获得不错的性能；基于架构搜索设计网络空间的 RegNetx032 以较少的参数量获得较好的性能；基于 Transformer 架构的 SwinTransformer 以 RegNetx032 近 6 倍的资源消耗获得了超过 CNN 架构的最佳性能；从卷积角度出发对其进行改进的 NATransformer，利用卷积的效率和归纳偏置使得资源消耗降低近一半，并获得更好的性能。本文模型 DDSF-Net 在代表模型整体大小的模型参数量、模型计算量上分别为 76.31M，4.149G，准确率达到 98.01%，AUC 值达到 97.99%，参数量略大于 DenseNet，而计算量小幅下降，准确率、精确率、召回率、F1 分数和 AUC 值较 DenseNet 分别提升 4.89%，2.59%，5.83%，3.98% 和 4.69%，与经典网络相比，本文模型的各项性能提升明显。

图 5 是 9 种模型在肺炎 X 光片数据集上的 ROC 曲线和 AUC，本文模型具有明显的优势和鲁棒性，能较好地学习和识别肺炎病灶特征。

表 2 肺炎 X 光片数据集中各模型的具体结果

Tab. 2 Specific results of each model in pneumonia X-ray dataset

模 型	参数量 (M)	计算量 (G)	训练 时间/s	准确率/%	精确率/%	召回率/%	F1分数/%	AUC 值/%
VGG19 <sup>[5]</sup>	1 096.09	39.294	26 613	89.07±2.4	94.81±2.2	86.16±2.9	90.28±2.2	89.69±2.8
ResNet101 <sup>[8]</sup>	324.28	15.664	21 291	91.32±1.5	93.46±1.5	91.68±1.7	92.56±1.3	91.24±1.9
DenseNet121 <sup>[4]</sup>	60.88	5.732	16 670	93.44±2.1	96.04±1.8	92.69±1.6	94.33±1.7	93.60±2.0
SeResNet101 <sup>[8]</sup>	360.47	15.692	22 413	93.83±1.3	95.12±1.2	94.38±1.5	94.75±1.3	93.72±1.4
EfficientNetb0 <sup>[18]</sup>	30.59	0.027	15 155	91.52±2.5	94.61±2.4	90.78±3.1	92.65±2.6	91.68±2.8
EfficientNetb4 <sup>[11]</sup>	133.91	0.068	19 299	93.97±1.9	96.18±1.7	93.48±1.5	94.81±1.7	94.08±1.9
RegNetx032 <sup>[18]</sup>	72.66	3.955	18 492	94.63±1.6	96.01±1.5	94.83±1.6	95.42±1.8	94.59±1.7
SwinTransformer <sup>[14]</sup>	411.75	18.836	24 327	95.29±2.1	95.75±2.0	96.29±2.2	96.02±2.1	95.08±2.3
NATransformer <sup>[19]</sup>	240.63	9.308	19 529	95.63±1.8	96.71±1.6	95.84±1.9	96.27±1.7	95.58±1.9
DDSF-Net	76.31	4.149	15 729	98.01±1.1	98.53±0.8	98.09±0.9	98.31±0.8	97.99±1.0

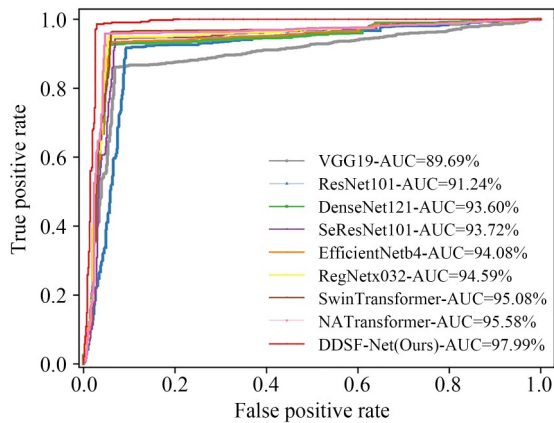


图 5 肺炎 X 光片数据集中各模型的 ROC 曲线

Fig. 5 ROC curves of each model in pneumonia X-ray dataset

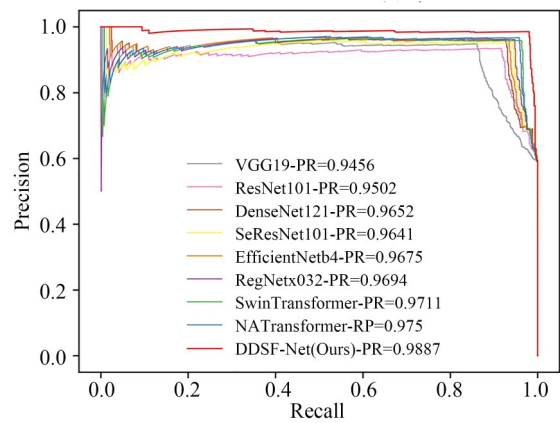


图 6 肺炎 X 光片数据集中各模型的 PR 曲线

Fig. 6 PR curves of each model in pneumonia X-ray dataset

图 6 为 9 种模型在肺炎 X 光片数据集上的 PR 曲线。准确率越高,召回率越低,向上方凸出、包围面积大的曲线代表模型效果好,可以看出,本文模型性能最优。

### 3.6 小目标病灶识别

本文将肺炎病灶区域所占像素面积小于  $32 \times 32$  的输入图像定义为包含小目标的图像,在 3.1 所述的数据集中随机挑选出 524 张包含肺炎病灶小目标的图像和 352 张正常图像,用于验证不同模型对体积偏小病灶的识别能力。在小目标肺炎数据集中,肺炎的影像学特征不明显和病灶特征提取不充分问题更为突出,可以很好地突出本文模型的性能和优势。5 个模型在小目标肺炎 X 光片数据集的对比结果如表 3 所示,各模

型的 ROC 曲线如图 7 所示。5 项指标和 ROC 曲线可以明显看出,本文模型可以更好地缓解肺炎数据集中肺炎影像学特征不明显和病灶特征提取不充分问题,因此能较好对包含小病灶目标的肺炎影像进行识别。

表 3 小目标肺炎 X 光片数据集中不同模型的指标

Tab. 3 Specifications of different models in small target pneumonia X-ray dataset

模 型	准确率	精确率	召回率	F1	AUC 值
ResNet101 <sup>[8]</sup>	86.24	90.00	88.73	89.36	85.16
DenseNet121 <sup>[4]</sup>	88.07	92.03	89.44	90.72	87.48
RegNetx032 <sup>[18]</sup>	90.83	94.20	91.55	92.86	90.51
NATransformer <sup>[19]</sup>	92.66	93.75	95.07	94.41	91.61
DDSF-Net	95.41	97.14	95.77	96.45	95.26

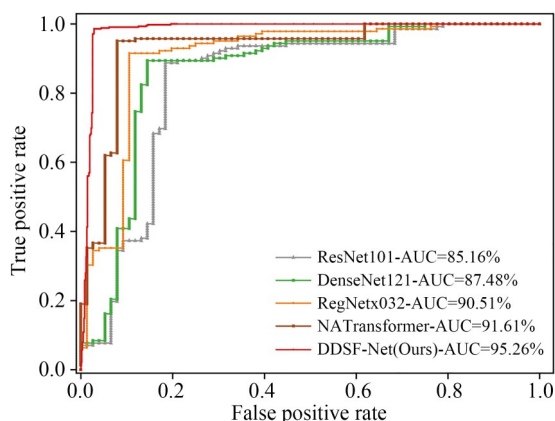


图7 小目标肺炎X光片数据集中各模型的ROC曲线  
Fig.7 ROC curves for each model in small target pneumonia X-ray dataset

### 3.7 不同公开数据集实验

为验证本文模型的鲁棒性和泛化能力,在3个公开的肺炎相关X光片数据集上进行对比实验,结果如表4~表6所示。第一个数据集来自文献[20],共234张正常图像和221张新冠肺炎图像。第二个数据集来自文献[21],共8552张正常、5674张肺炎图像和7598张新冠肺炎图像。第三个数据集来自文献[22],共510张正常和510张新冠肺炎图像,本文模型DDSF-Net均取得最好的性能。

表4 深度学习模型诊断新冠肺炎X光片

Tab.4 Deep learning model for diagnosing COVID-19 X-ray images (%)

模型	敏感度	特异度	准确率	精确率	F1
VGG19 <sup>[5]</sup>	82.96	93.96	92.33	—	—
Covid-caps <sup>[3]</sup>	90.00	95.80	95.70	—	—
DCNN <sup>[6]</sup>	97.91	91.87	—	93.36	—
GSEN <sup>[20]</sup>	93.40	98.08	95.60	92.73	95.50
DDSF-Net	98.36	98.08	98.23	98.08	98.08

### 参考文献:

[1] EZHILAN M, SURESH I, NESAKUMAR N. SARS-CoV, MERS-CoV and SARS-CoV-2: a diagnostic challenge[J]. *Measurement: Journal of the*

表5 深度学习模型诊断肺炎X光片

Tab.5 Deep learning model for diagnosing pneumonia X-ray images (%)

模型	敏感度	特异度	准确率	AUC
DenseNet121 <sup>[4]</sup>	91.00	87.00	88.13	90.00
EfficientNetb5 <sup>[11]</sup>	83.00	92.00	94.64	95.00
Covid-caps <sup>[3]</sup>	90.00	95.00	95.00	97.00
ViT-B32 <sup>[21]</sup>	96.00	96.00	96.00	99.10
DDSF-Net	98.55	98.97	98.71	99.53

表6 深度学习模型诊断肺炎X光片

Tab.6 Deep learning model for diagnosing pneumonia X-ray images (%)

模型	准确率	AUC
Dark COVID-Net <sup>[13]</sup>	0.870 2	—
AF-CAP <sup>[22]</sup>	0.991 6	0.988 0
DDSF-Net	0.996 3	0.991 7

## 4 结 论

本文提出一种面向X光片辅助诊断肺炎的密集双流聚焦网络模型,设计残差多尺度块提高网络的适应性,设计同时学习全局和局部特征的双流密集块,密集连接方式对两种信息流的深浅层特征进行融合,设计聚焦块过滤背景噪声信息,增强病灶的细节特征。本文模型DDSF-Net在肺炎X光片数据集中的准确率为98.01%,AUC为97.99%,在小目标肺炎数据集中的准确率为95.41%,AUC为95.26%,并且在3份公共肺炎相关数据集中获得了最优的精度,具有较好的鲁棒性和泛化能力。采用热力图可视化技术进一步验证本文模型,DDSF-Net可以有效缓解肺炎影像学特征不明显和病灶特征提取不充分的问题。在医学肺炎图像分类中,还应当注意模型轻量化程度,未来将继续设计轻量化的肺炎辅助诊断模型,以及研究更复杂的肺炎分类识别任务。

*International Measurement Confederation*, 2021, 168: 108335.

[2] Centers for Disease Control and Prevention [Z/OL]. <https://www.cdc.gov/pneumonia/prevention.html>, October 22, 2020.

- [3] TOĞAÇAR M, ERGEN B, CÖMERT Z. COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches[J]. *Computers in Biology and Medicine*, 2020, 121: 103805.
- [4] 周涛, 霍兵强, 陆惠玲, 等. 融合多尺度图像的密集神经网络肺部肿瘤识别算法[J]. *光学精密工程*, 2021, 29(7): 1695-1708.
- ZHOU T, HUO B Q, LU H L, *et al.* Lung tumor image recognition algorithm with densenet fusion multi-scale images[J]. *Opt. Precision Eng.*, 2021, 29(7): 1695-1708. (in Chinese)
- [5] JAIN R. Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning[J]. *Measurement*, 2020, 165: 108046.
- [6] RAJPURKAR P, IRVIN J, ZHU K, *et al.* CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning [EB/OL]. 2017: *arXiv*: 1711.05225. <https://arxiv.org/abs/1711.05225>.
- [7] ÇALLI E, SOGANCIOGLU E, VAN GINNEKEN B, *et al.* Deep learning for chest X-ray analysis: a survey [J]. *Medical Image Analysis*, 2021, 72: 102125.
- [8] CHEN B Z. DualCheXNet: dual asymmetric feature learning for thoracic disease classification in chest X-rays[J]. *Biomedical Signal Processing and Control*, 2019, 53: 101554.
- [9] LI H, ZHUANG S S, LI D A, *et al.* Benign and malignant classification of mammogram images based on deep learning[J]. *Biomedical Signal Processing and Control*, 2019, 51: 347-354.
- [10] CHEN B L, ZHAO T S, LIU J H, *et al.* Multipath feature recalibration DenseNet for image classification [J]. *International Journal of Machine Learning and Cybernetics*, 2021, 12(3): 651-660.
- [11] 陈筱, 朱向冰, 吴昌凡, 等. 基于迁移学习与特征融合的眼底图像分类[J]. *光学精密工程*, 2021, 29(2): 388-399.
- CHEN X, ZHU X B, WU CH F, *et al.* Research on fundus image classification based on transfer learning and feature fusion [J]. *Opt. Precision Eng.*, 2021, 29(2): 388-399. (in Chinese)
- [12] LIU Y J, HAO P Y, ZHANG P, *et al.* Dense convolutional binary-tree networks for lung nodule classification [J]. *IEEE Access*, 2018, 6: 49080-49088.
- [13] PRIYA K V. A federated approach for detecting the chest diseases using DenseNet for multi-label classification [J]. *Complex & Intelligent Systems*, 2022, 8(4): 3121-3129.
- [14] NASEER M, RANASINGHE K, KHAN S, *et al.* Intriguing properties of vision transformers [EB/OL]. 2021: *arXiv*: 2105.10497. <https://arxiv.org/abs/2105.10497>.
- [15] 景海钊, 史江林, 邱梦哲, 等. 基于密集残差块生成对抗网络的空间目标图像超分辨率重建[J]. *光学精密工程*, 2022, 30(17): 2155-2165.
- JING H ZH, SHI J L, QIU M ZH, *et al.* Super-resolution reconstruction method for space target images based on dense residual block-based GAN [J]. *Opt. Precision Eng.*, 2022, 30(17): 2155-2165. (in Chinese)
- [16] KERMANY D S, ZHANG K, GOLDBAUM M. Labeled optical coherence tomography (OCT) and chest X-ray images for classification [J]. *Mendeley data*, 2018
- [17] GIEŁCZYK A, MARCINIAK A, TARCZEWSKA M, *et al.* Pre-processing methods in chest X-ray image classification [J]. *PLoS One*, 2022, 17(4): e0265949.
- [18] RADOSAVOVIC I, KOSARAJU R P, GIRSHICK R, *et al.* Designing network design spaces [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 10425-10433.
- [19] HASSANI A, WALTON S, LI J, *et al.* Neighborhood Attention Transformer [EB/OL]. 2022: *arXiv*: 2204.07143. <https://arxiv.org/abs/2204.07143>.
- [20] ELARABY M E. A novel Gray-Scale spatial exploitation learning Net for COVID-19 by crawling Internet resources [J]. *Biomedical Signal Processing and Control*, 2022, 73: 103441.
- [21] CHETOUI M, AKHLOUFI M A. Explainable vision transformers and radiomics for COVID-19 detection in chest X-rays [J]. *Journal of Clinical Medicine*, 2022, 11(11): 3013.
- [22] BALASUBRAMANIAN K, ANANTHAMOORTHY N P, RAMYA K. An end-end deep learning framework for lung infection recognition

using attention-based features and cross average pooling [J]. *International Journal for Multiscale Computational Engineering*, 2022, 20(2): 67-82.

- [23] KHAN I A. CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest X-ray images [J]. *Computer Methods and Programs*

*in Biomedicine*, 2020, 196: 105581.

- [24] OUCHICHA C, AMMOR O, MEKNASSI M. CVDNet: a novel deep learning architecture for detection of coronavirus (Covid-19) from chest X-ray images [J]. *Chaos, Solitons, and Fractals*, 2020, 140: 110245.

#### 作者简介:



周 涛(1977—),男,宁夏同心人。博士,教授,2010年于西北工业大学获得博士学位,主要从事医学图像分析处理、深度学习、模式识别等方面的研究。E-mail: zhoutaonxmu@126.com

#### 通讯作者:



叶鑫宇(1999—),男,湖北天门人,硕士研究生,主要从事智能医学影像图像处理,计算机辅助诊断等方面的研究。E-mail: 3303626778@qq.com