

基于细粒度的肺炎识别方法^{*}

杨杰之, 唐万梅, 皮家甜, 汪建良

(重庆师范大学 计算机与信息科学学院, 重庆 401331)

摘要:【目的】针对肺炎识别案例中存在的数据集数量分配不均、数量少、类别间差异小等问题,提出一种基于卷积神经网络判别模块的肺炎识别方法。【方法】首先,将网络骨干设定为预训练的 121 层 DenseNet 网络,并冻结相关参数,以迁移学习的方式来解决数据量少的问题,再将网络的中间层定义一组额外的卷积滤波器,通过学习这组滤波器,可以捕获类别之间的高信号区域,而不需要额外的边框注释,以此解决肺炎识别中的类间样本差异小的问题。最后定义网络的损失函数为 FocalLoss 以解决数据集数量分配不均的问题。【结果】用提出的新方法在 Chest-X-Ray Image 数据集上进行实验。实验结果显示:该方法的准确率达到 95%,比传统的迁移学习准确率提升了 10%,比未进行迁移学习的轻量级卷积神经网络的准确率则普遍提升了 20%。【结论】该方法能够针对肺部 X 光片做出是否患有肺炎的判断,且定位出感染区域。

关键词:计算机视觉;卷积神经网络;肺炎识别;细粒度识别;迁移学习;FocalLoss

中图分类号:TP391.41

文献标志码:A

文章编号:1672-6693(2021)03-0100-07

自新型冠状病毒肺炎疫情发生以来,截至 2021 年 5 月 22 日,中国国内确诊病例 107 807 例,全球确诊病例 167 011 807 例,且疫情在全球仍在不断蔓延^[1]。目前,新冠肺炎的检测方式主要是核酸检测,肺部 X 光图像主要作为肺炎诊疗的辅助依据。由于新冠肺炎阳性与阴性的肺部 X 光图像差别不是特别明显,所以每张 X 光图像均需要放射科医生花费一定的时间和精力进行图像比对,在存在大量待诊断的新冠肺炎疑似病例的情况下,就亟需一种关于肺炎 X 光图像的自动识别方法。

随着深度学习技术的不断发展,基于卷积神经网络的肺炎图像识别得到了广大研究者的关注。Rajpurkar 等人^[2]提出 ChexNet 用于肺炎识别,即对迁移学习的 121 层的 DenseNet 加以改进,以 sigmoid 激活函数作为网络的分类器,他们的方法在 Chest-X-Ray14 数据集上达到了 86% 的识别准确率,这一网络诊疗的推断已经超越了许多放射科医生。这种方法还可以通过 Zhou 等人^[3]提出的可视化方法进行呈现,进而精确定位出各类肺炎出现的区域,但该模型过于简单并不能以细粒度的方式进行很好的判别,识别率仍有提升空间。Vianna 等人^[4]使用 Pytorch 深度学习框架,分别使用迁移的 AlexNet^[5],SqueezeNet^[6],ResNet^[7]以及 InceptionV3^[8]神经网络在 Chest-X-Ray 数据集上进行训练,都取得了不错的训练效果,但是由于未采取数据增强的方式,数据的不平衡性导致识别阳性的准确率与识别阴性的准确率不够平衡,且前者高于后者。

由于肺炎 X 光图像差异较小,本文考虑利用细粒度识别进行分析。细粒度识别通常利用局部区域信息捕捉细微的差异,早期细粒度识别方式不仅需要类别的标签而且需要研究者们进行区域的注释,例如 Zhang 等人^[9]通过将注释后的边框信息以及标签信息进行端到端的训练,这使得整个细粒度识别的过程类似于目标检测,虽然很大程度上提高了模型准确率,但是标注数据的时间成本过高,训练所耗时间也过长。为了降低人工标注数据的成本,Yang 等人^[10]通过利用 Faster-RCNN^[11]中的区域建议网络机制,让网络自动地去找高信号区域,并对每个高信号区域进行监督学习,再在网络的终端采用融合的策略进行端到端的训练。这一方法在 CUB-200-2011 鸟类数据集中的识别准确率达到 86.7%,但是该方法的网络训练参数过多,模型预测时间过长,添加可视化方法更是需要大量时间计算,并不利于辅助诊断。Wang 等人^[12]提出 VGG-DL 的方式进行端到端的训练,

^{*} 收稿日期:2020-10-09 修回日期:2021-05-22 网络出版时间:2021-05-21 11:06

资助项目:国家自然科学基金(No. 11971083);重庆市教育委员会科技项目(No. KJQN201800521);重庆市基础研究与前沿探索项目(No. cstc2018jcyjAX0470);重庆师范大学 2019 年研究生科研创新项目(No. YKC19014)

第一作者简介:杨杰之,男,研究方向为计算机视觉、深度学习,E-mail:Yjz357406659@126.com;通信作者:唐万梅,教授,博士,E-mail:1093895431@qq.com

网络出版地址:https://kns.cnki.net/kcms/detail/50.1165.N.20210521.1055.002.html

使用判别机制从而使得神经网络不使用区域建议或者标注的区域信息也能够精准的定位图像的高信号区域,但是对于肺炎数据不平均的问题,识别率仍有提升的空间。

考虑到肺炎数据集相对较少且存在不平均以及类间差别小等问题,本文将从细粒度识别的方向进行研究,具体过程如下:

- 1) 以 DenseNet121^[13] 神经网络为基础,进行迁移;
- 2) 在该神经网络上构建判别滤波器^[11]进行区域判别;
- 3) 使用 FocalLoss^[14] 作为损失函数,解决数据不平均的问题。

1 肺炎识别网络设计

1.1 肺炎识别网络框架

由于可用于训练的数据少,而 DenseNet 能够加大对前层特征的使用,从而有效地避免梯度爆炸,所以本文将在 Imagenet 上训练的 DenseNet121^[13] 进行迁移,并通过文献[12]的方式对网络整体架构进行改进。由于细粒度识别的任务不仅需要依赖于全局的特征,而且还需要图像局部区域的特征,所以肺炎识别的模型需要一种能够兼顾全局与局部的架构。

针对以上分析,本文将网络分类层进行微调后得到肺炎识别网络框架(图1),该框架主要由3个部分组成:1) 全局判别模块(G-Module);2) 判别模块(P-Module);3) 跨通道池化模块(C-Module)。最终的目标函数为将上述3个部分的损失加权相加所得。

1.2 G-Module

针对肺部 X 光图像的肺炎识别任务中,网络模型需要对整张图进行理解所以需要一个全局的判别模块,而 G-Module 中最常用的是全局平均池化(GAP)以及全连接层,GAP 被许多高效的卷积神经网络采用,经大量实验证明 GAP 对于非细粒度的任务能够防止过拟合。但是 Wu 等人^[14] 发现,对于细粒度分类的任务,在用于训练细

粒度的数据量不够庞大的数据集上,没有 GAP 的卷积神经网络架构比有 GAP 的架构有着更高的准确率。这是由于 GAP 无法对网络进行权重区域区分,利用全连接层则能解决这一问题。但是全连接层的参数量又过于庞大,而对于不够庞大的数据集容易产生过拟合的问题。为了使网络在降低参数量的前提下不损失精度,因此本文的 G-Module 由 7th~12th 的 DenseNetLayer 和 Streaming Module 组成。7th~12th 的 DenseNetLayer 提取浅层的语义信息,深层的语义信息由 Streaming Module^[15] 进行提取。在 Straming Module 的设计上,首先选择以 0.5 为节点隐藏率以防止过拟合,采用尺寸为 7×7、步长为 1 的深度卷积核;接下来对卷积之后的特征图进行 Batch Normlization 的操作;然后进行 ReLu6 非线性激活;最后使用个数为 2(因为肺炎识别是一个二分类的问题)、步长为 1 的 1×1 卷积。这样既达到了替代全连接层的效果,又减少了网络的参数量,加快了训练速度,也起到了对全局区域判别的作用。

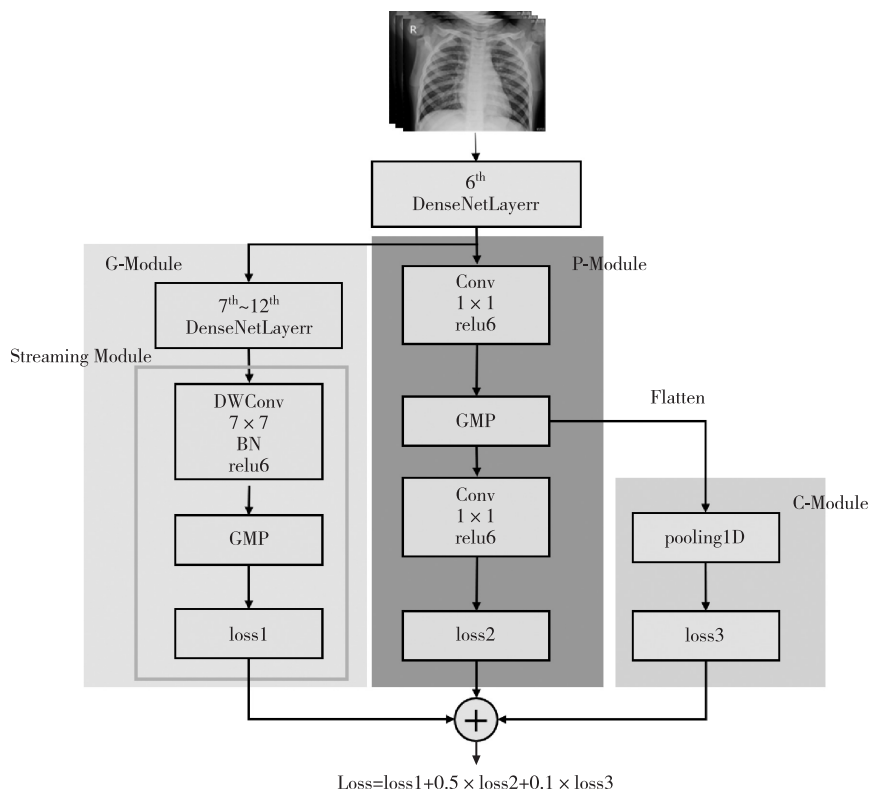


图1 肺炎识别网络框架

Fig. 1 Pneumonia recognition network framework

1.3 P-Module

网络中负责甄别“块”的核心组件为 1×1 的卷积以及全局最大池化(GMP)。在 P-Module 中,将 1×1 的卷积核作为一个“块”的检测器,原始图像通过 6 个 DenseNetLayer 进行特征提取,得到 $C \times H \times W$ 的特征图。其中 C 为特征图的通道数, W 为特征图的宽, H 为特征图的高。在该特征图上的每个 $C \times 1 \times 1$ 的张量被感受野机制所影响,因此都能在原图中找到对应的“块”。

判别区域在细粒度的识别任务中处于非常局部的范围,这就意味着模型需要找到一些稍小的判别区域。由于涉及到“块”的精准定位,特征图传递至网络后层,对于传递至后层的特征图上的单个 $C \times 1 \times 1$ 的张量,基于该张量的感受野机制,会使得它在原图中定位到较大的区域,而且相邻的步长也较大,不利于“块”的精准定位。针对 224×224 大小的原图,本文选取经过第 6 个 DenseNetLayer 后的特征图,特征图的大小为 56×56 ,在特征图上每个点映射回原图的感受野为 67,步长为 4,对于肺炎识别这一细粒度识别任务,感受野和步长已经足够小和稠密。感受野的计算公式为:

$$F(i) = (F(i+1) - 1) \times S_{\text{Stride}} + S_{\text{Ksize}}$$

其中: $F(i)$ 为第 i 层的感受野, S_{Stride} 为第 i 层步长, S_{Ksize} 为第 i 层卷积核或采样核的尺寸,网络的输入尺寸为 224×224 ,本文网络经过第 6 个 DenseNetLayer 后的感受野为 67×67 ,步长为 4,可以甄别图像细微差别,满足细粒度识别的要求。

特征图通过学习 20 个(因为目标类是两类,每类通过 10 个 1×1 滤波器提取特征) 1×1 的卷积后得到新的特征图,在这个新的特征图上通过全局最大池化找到最大值的位置,新的特征图上最大值所在位置,通过感受野的机制映射回原图就可以找到一个带有肺部病灶区域的“块”。

1.4 C-Module

由于在判别模块中的 1×1 滤波器不能总是触发,本文通过引入跨通道池化层和 softmax 分类层来直接对 1×1 的过滤器进行监督。

跨通道池化层的监督工作如下,由于肺炎识别为一个 2 分类的任务,每个类有 10 个检测器;那么需要的 1×1 个过滤器的数量为 20,提取后的特征图通过 GMP 获取最大响应。C-Module 将 P-Module 中的 GMP 层的特征图进行 Flatten 操作,即:将 GMP 计算后得出的特征图以一维张量进行展开,展开后此时结果为一个 20 维的向量,将该向量分成两组,每组 10 个值取平均值作为其中一个类的响应,从而得到一个 2 维向量。由于平均池化倾向于影响所有池化的过滤器,而最大池化只影响响应最大的过滤器,所以最终使用平均池化使目标类别具备平衡的响应。

1.5 FocalLoss 损失函数

肺炎数据集中阳性样本与阴性样本的比例为 3 : 1,对于深度学习模型来讲,阳性样本就是足以正确分类的简单样本。简单样本经过交叉熵损失函数计算后损失值会很小,但是又因为简单样本相对而言数量较大,会使得简单样本对损失给出巨大的贡献从而影响网络的收敛性。为此 Lin 等人^[16]提出了 FocalLoss 以解决数据的平衡性问题。

对于二分类问题目标类别 $y \in \{+1, -1\}$,目标类别的概率 $p \in [0, 1]$, p_t 定义为: $p_t = \begin{cases} p, & \text{当 } y=1 \text{ 时} \\ 1-p, & \text{否则} \end{cases}$;

FocalLoss 损失定义为: $L_{\text{Floss}} = a_t (1 - p_t)^\gamma \log_{10}(p_t)$,其中: $a_t = \begin{cases} a, & \text{当 } y=1 \text{ 时} \\ 1-a, & \text{否则} \end{cases}$ 为平衡因子,且为 $[0, 1]$ 之间的常

数,以解决数据集分布不均的问题, γ 为常数,本文中令 $\gamma=2$ 。不难看出当 p_t 接近 1 时,此时样本属于容易分类的简单样本,则损失函数被降低权重,使得神经网络将精力投放至复杂样本上。

2 肺炎识别网络设计

本实验所采用的硬件配置为:CPU 型号为 Intel Core i7-7700HQ,主频为 2.80 GHz;GPU 型号为 NVIDIA GTX1050,显存 4 GB。软件环境为:Windows 10 系统,采用 GPU 进行运算,编程语言为 Python 3.6,pytorch 1.0.0,CUDA 8.0,Cudnn 7.1.4。

2.1 数据集及预处理

本文实验中所采用的数据集(Chest-X-Ray Image)来源于 kaggle 网站(<https://www.kaggle.com/>)

paultimothymooney/chest-xray-pneumonia)。该数据集有阴性肺炎图像(Normal)以及阳性肺炎图像(Pneumonia)共两类,其中分为训练集(5 232 张)、验证集(16 张)以及测试集(624 张),训练集中阳性(3 875 张),阴性(1 341 张)。

为了对数据进行增强,本文将图像缩放裁剪为 224×224 大小的尺寸,然后水平翻转,并利用随机缩放比例将原始图像剪切生成新图像,预处理后的数据示意图如图 2 所示。

2.2 实验过程及参数设置

图像经过裁剪、数据增强等预处理后放入神经网络进行端到端的学习,神经网络共输出 3 个结果:1) G-Module;2) P-Module;3) C-Module。3 个模块每个都拥有各自的 FocalLoss 损失和不同的权重,最终的损失为 3 个模块的损失加权之和,模型训练过程中 3 个模块的损失权重分别分配为 $1:0.5:0.1$ 。测试输出预测结果的权重设置为 $1:0.5:0.1$ 。超参数如 Batchsize、初始学习率(Learning rate, LR)、权重衰减系数 Weight_decay 还有 FocalLoss 损失函数的两个超参数 α 和 γ 的具体设置如表 1 所示。

3 实验结果分析

3.1 各网络对比实验

本文以准确率、召回率以及模型参数作为评估指标。准确率(P_{Acc})与召回率(P_{Recall})计算公式分别为: $P_{Acc} = \frac{V_{TP} + V_{TN}}{V_{TP} + V_{TN} + V_{FP} + V_{FN}}$, $P_{Recall} = \frac{V_{TP}}{V_{TP} + V_{FN}}$ 。式中: V_{TP} 表示为将正类预测为正类的数量, V_{TN} 表示为将负类预测为负类的数量, V_{FP} 表示为将负类预测为正类的数量, V_{FN} 表示为将正类预测为负类的数量。由于阴性样本与阳性样本比例失调,为了突出模型的效果,本文以阴性样本作为正例,阳性样本作为负例,计算准确率及召回率。

由于数据量少,采用大型神经网络容易产生过拟合的影响,所以实验分别使用轻量级网络 MobileNetV3^[17]、GhostNet^[18] 和通过迁移学习后的 VGG16^[19]、ResNet18^[20]、DenseNet121^[13]、VGG-DL^[12]、ChexNet^[2] 以及细粒度识别方法 Bilinear-CNN^[21] 与本文方法进行对比,模型以相同的超参数、损失函数进行训练,将迁移学习的网络前层冻结,分类层进行微调,实验结果如表 2 所示。

表 2 中,数据部分的前两行为轻量级网络 MobileNetV3^[17] 与 GhostNet^[18] 的实验结果,虽然轻量级网络能够有效地避免过拟合,但是由于数据集少,所以性能未达到最佳。本文方法相对于未迁移学习的 MobileNetV3 准确率提升了 18%,召回率提升了 21%;相对于轻量级网络 GhostNet 准确率提升了 21%,召回率提升了 25%;相对于细粒度识别网络 Bilinear-CNN 准确率提升了 10%,召回率提升了 13%;相对于迁移学习后的 VGG16^[19] 准确率提升了 13%,召回率提升了 16%;相对于迁移学习后的 ResNet18^[20] 准确率提升了 9%,召回率提升了 10%;相对于迁移学习后的 VGG16-DL^[12] 网络准确率提升了 9%,召回率提升了 11%。迁移学习效果最好的为 121 层的 DenseNet^[13],相对于迁移学习后的 DenseNet121 准确率提升了 8%,召回率提升了 8%;相对于迁移学习后的 ChexNet 准确率提升了 5%,召回率提升了 4%。同时本文对模型的参数量以及单张图像的预测时间进行了对比,轻量级网络 GhostNet 以及 MobileNetV3,参数量最少,单张图像预测时间最短,但是准确率难以达到应用标准,在召回率超过 80% 的模型中,本文所提方法参数量最少,且单张图像的预测时间最短,与表现最佳的迁移学习的 DenseNet121 相比,参数量少 6M,单张图像的预测时间少 0.007 s。

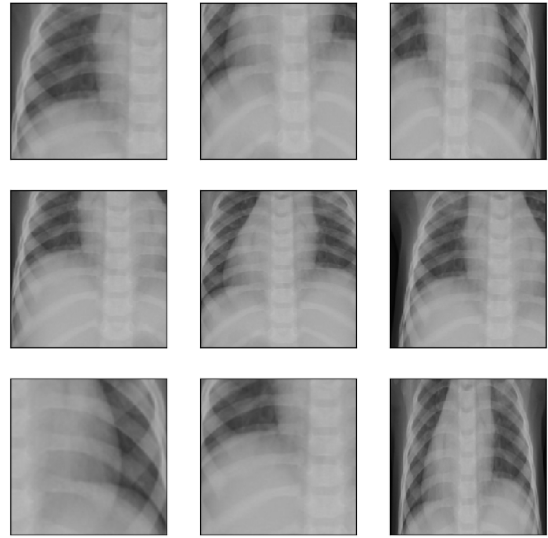


图 2 数据预处理

Fig. 2 Data preprocessing

表 1 超参数设置表

Tab. 1 Hyper parameters setting table

超参数	值
Batchsize	16
LR	$1e-3$
Weight_decay	$1e-4$
α	0.75
γ	2

表 2 各模型效果对比

Tab. 2 Comparison of the effects of each model

方法	准确率/%	召回率/%	参数量/MB	预测时间/s
MobileNetV3	77.3	72.3	17	0.032
GhostNet	74.6	68.2	15	0.032
VGG16_FT	82.3	77.3	59	0.079
Bilinear-CNN	85.6	80.2	58	0.079
ResNet18_FT	86.7	82.6	44	0.053
DenseNet121_FT	87.4	84.5	33	0.046
VGG16_DFL	86.3	82.3	57	0.059
ChexNet	90.1	88.7	33	0.047
本文方法	95.6	93.4	27	0.039

3.2 模块实验对比

为验证模型的鲁棒性,实验以准确率为指标分别选取仅训练 G-Module 和仅训练 P-Module 以及训练 G-Module 与 P-Module 和 C-Module 加权后输出的准确率进行对比,实验结果如表 3 所示。

在 C-Module 中,表 3 中数据部分的 3~4 行在 C-Module 上做了平均池化(AVGpool)以及最大池化(MAXpool)的对比,结果显示平均池化的效果比最大池化的效果提升了 1.3%。实验结果表明使用 G-P-AVGpool 的准确率比仅使用 G-Module 的准确率提升了 1.1%,比仅使用 P-Module 准确率提升了 7%,G-Module 与迁移学习后的 DenseNet 相比准确率提升了 7.1%。P-Module 与迁移后的 DenseNet 相比准确率也提升了 1.2%。

3.3 损失函数对比

表 4 中为验证 FocalLoss 损失函数对于本文模型的有效性,本文选取二分类交叉熵(BCELoss),多分类交叉熵(CELoss)与 FocalLoss 进行对比,FocalLoss 相比 BCELoss 提升了 1.2%,相对于 CELoss 提升了 2%。

表 3 模块效果对比

Tab. 3 Comparison of the effects of each module

模块	准确率/%
G-Module	94.5
P-Module	88.6
G-P-MAXpool	94.3
G-P-AVGpool	95.6

表 4 损失函数对比

Tab. 4 Comparison of the effects of each loss function

模块	准确率/%
BCELoss	94.4
CELoss	93.6
FocalLoss	95.6

3.4 G-Module 可视化方法

为使网络具备可解释性,本文在 G-Module 上取出经过深度可分离卷积层的特征图,通过 Grad-Cam 方法^[22]生成热图进行可视化处理,找到最具疾病指示性的图像区域,如封三彩图 3 所示。封三彩图 3b 中红色区域为肺炎患者感染区域。热图的计算公式为:

$$V_{\text{heatmap}} = \sum_k V_{\text{mean}} \left(\frac{\partial y^c}{\partial A^k} \right) \times A^k.$$

肺炎识别为二分类问题,由于数据集的样本数量较少,网络在训练的过程中存在着只对固定的区域进行判别的情况,所以生成的热图只能代表这部分感染区域。

3.5 P-Module 可视化方法

为了使网络对感染区域的定位更加精确,本文在 P-Module 上,取出经过第 6 个 DenseNetLayer 后经过 1×1 卷积后的特征图,并将它以热图的形式进行可视化。封三彩图 4 中红色区域为峰值区域,再利用特征图上最大值的索引通过感受野机制反向映射回原图,找到响应最大的高信号区域,这些区域就是可能存在感染的区域(图 5)。

4 结语

本文提出了一种针对细粒度肺炎识别的方法,以3种模块的协调来解决细粒度识别的问题。通过 FocalLoss 损失函数解决数据集分布不均的问题,通过反向映射的方法使神经网络对于肺炎识别任务进行可视化呈现,但是由于实验硬件设备的限制,模型并未达到最佳效果,考虑到后续的研究工作,可使用更好的硬件设备对部分超参数进行调整,也可以利用更前沿的医学图像数据特征提取的方法对图像进行预处理,使得图像的特征更为明显,对于数据集量较少的情况,也可以采用更好的数据增强方式进行增强。

在可视化上,由于阳性肺炎的种类可继续细分,可以利用本文方法对图像进行进一步地多分类任务训练,通过本文的可视化方法,可以找到更为细致的肺炎感染区域,为临床医学提供辅助诊断的依据。

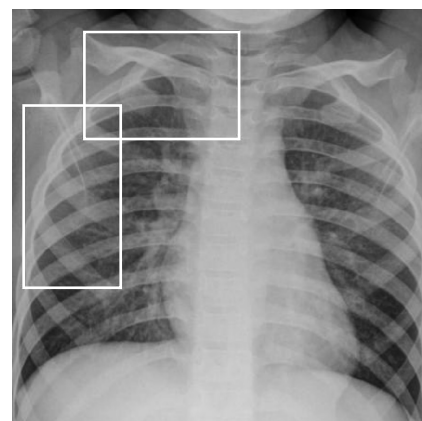


图5 P-Module 可视化

Fig. 5 P-Module visualization

参考文献:

- [1] WHO. Coronavirus (COVID-19) dashboard[EB/OL]. (2021-05-22)[2021-05-22]. <https://covid19.who.int/table?tablechartType=heat>.
- [2] RAJPURKAR P, IRVIN J, ZHU K, et al. Chexnet: radiologist-level pneumonia detection on chest X-rays with deep learning[EB/OL]. (2017-12-25)[2021-05-22]. <https://arxiv.org/abs/1711.05225>.
- [3] ZHOU B L, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA; IEEE, 2016: 2921-2929.
- [4] VIANNA V P. Study and development of a computer-aided diagnosis system for classification of chest X-ray images using convolutional neural networks pre-trained for imagenet and data augmentation[EB/OL]. (2018-06-03)[2021-05-22]. <https://arxiv.org/pdf/1806.00839.pdf>.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//2012 Advances in Neural Information Processing Systems, Lake Tahoe, USA; IEEE, 2012: 1097-1105.
- [6] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: alexnet-level accuracy with 50x fewer parameters and <0.5 MB model size[EB/OL]. (2016-11-04)[2021-05-22]. <https://arxiv.org/pdf/1602.07360.pdf>.
- [7] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA; IEEE, 2016: 770-778.
- [8] 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, Las Vegas, USA; IEEE, 2016: 2818-2826.
- [9] ZHANG N, DONAHUE J, GIRSHICK R, et al. Part-based R-CNNs for fine-grained category detection[C]//2014 European Conference on Computer Vision(ECCV). Zurich; Springer, 2014: 834-849.
- [10] YANG Z, LUO T, WANG D, et al. Learning to navigate for fine-grained classification[C]//2018 Springer European Conference on Computer Vision (ECCV). Munich, Germany; Springer, 2018: 420-435.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
- [12] WANG Y, MORARIU V, DAVIS L S. Learning a discriminative filter bank within a CNN for fine-grained recognition[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA; IEEE, 2018: 4148-4157.
- [13] HUANG G, LIU Z, Van Der MAATEN L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA; IEEE, 2017: 4700-4708.
- [14] WU B C, WAN A, YUE X Y, et al. Shift: a zero flop, zero parameter alternative to spatial convolutions[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA; IEEE, 2018: 9127-9135.
- [15] ZHANG P, ZOU F, WU Z, et al. FeatherNets: convolutional neural networks as light as feather for face anti-spoofing[C]//Proceedings of 2019 IEEE Computer Vision and Pattern Recognition, Piscataway, USA; IEEE, 2019: 1574-1583.
- [16] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, Venice, Italy; IEEE, 2017: 2980-2988.

- [17] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [C]//2019 IEEE International Conference on Computer Vision, Seoul;IEEE,2019;1314-1324.
- [18] HAN K, WANG Y, TIAN Q, et al. Ghostnet; more features from cheap operations[EB/OL]. (2020-03-13) [2021-05-22]. <https://arxiv.org/abs/1911.11907>.
- [19] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10) [2021-05-22]. <https://arxiv.org/abs/1409.1556>.
- [20] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA;IEEE,2016;770-778.
- [21] LIN T, ROYCHOWDH A, MAJI S. Bilinear CNN models for fine-grained visual recognition[C]//2015 IEEE International Conference on Computer Vision, Santiago, Chile;IEEE,2015;1449-1457.
- [22] SELVARAJU R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]//2017 IEEE International Conference on Computer Vision, Venice, Italy;IEEE,2017;618-626.

Fine-Grained Pneumonia Recognition Method

YANG Jiezhi, TANG Wanmei, PI Jiatian, WANG Jianliang

(School of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: [Purposes] Aiming at the problems of uneven distribution of data sets, small numbers, and small differences between categories in pneumonia identification cases. [Methods] A pneumonia identification method is proposed based on the convolutional neural network discriminant module. First, the backbone of the network is a pre-trained 121-layer DenseNet network, and the relevant parameters are frozen to solve the problem of low data volume by means of transfer learning, and then a set of additional convolution filters are defined in the middle layer of the network. The group filter can capture the high signal area between the categories without the need for additional border annotations, so as to solve the problem of small sample differences between the categories in pneumonia recognition; finally define the loss function of the network as FocalLoss to solve the number of data sets of the problem of uneven distribution. [Findings] The proposed method is tested on the Chest-X-Ray Image dataset. The experimental results show that the accuracy of the method has reached 95%, which is 10% higher than the traditional transfer learning, and is lighter than the non-transfer learning. Convolutional neural networks have generally increased by 20%. [Conclusions] This method can determine whether there is pneumonia based on lung X-rays and locate the infected area.

Keywords: computer vision; convolutional neural network; pneumonia recognition; fine-grained recognition; transfer learning; FocalLoss

(责任编辑 许 甲)

(接正文96~97页)

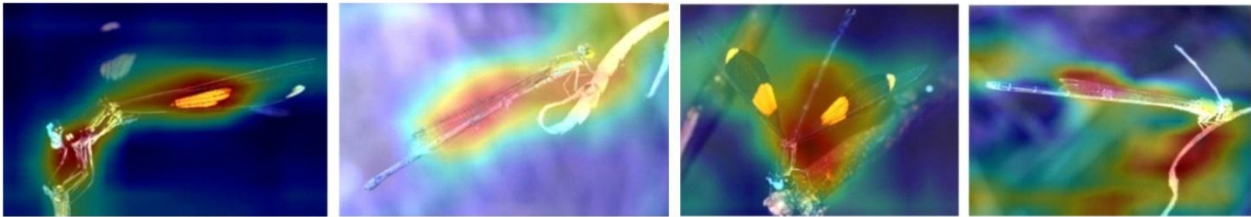


图5 蜻蜓数据集集中的高信息区域
Fig. 5 High information area in the dragonfly dataset

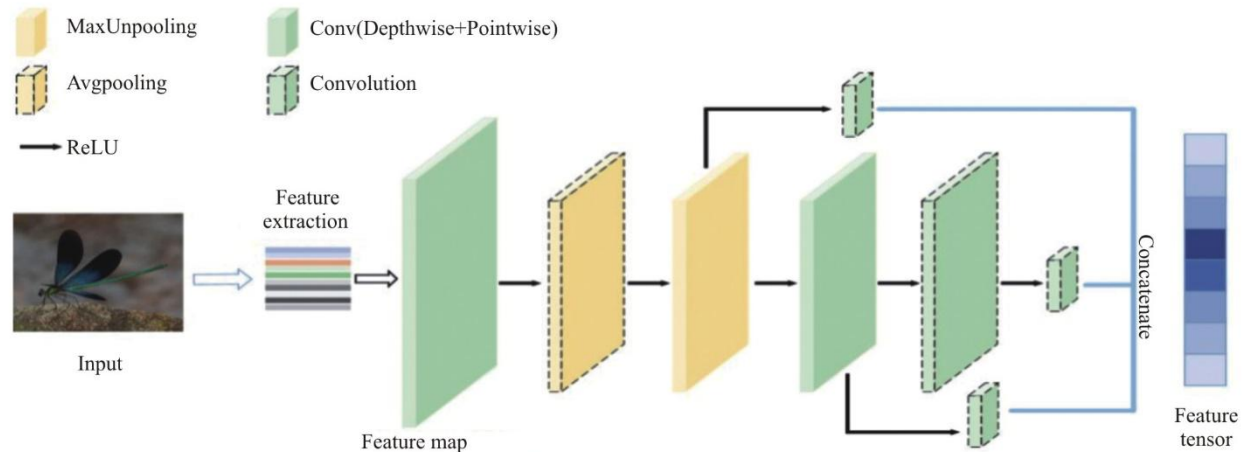


图6 QT-RPN结构图
Fig. 6 QT-RPN structure diagram

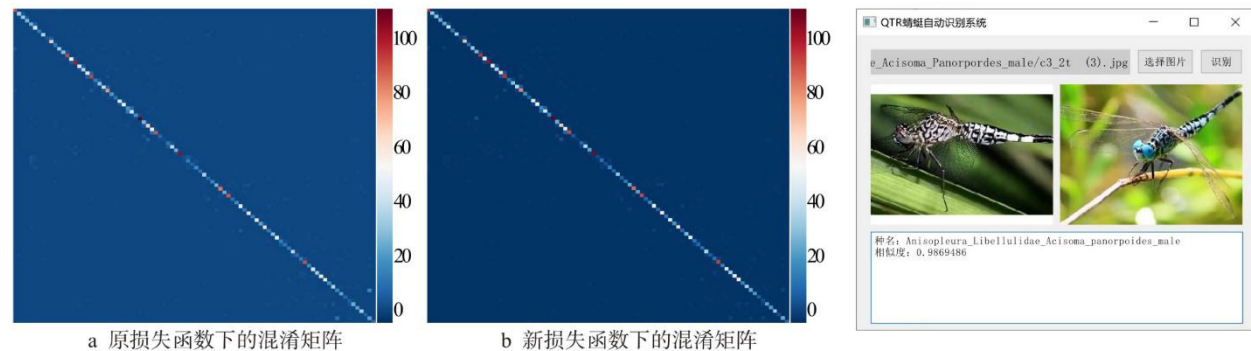


图7 不同损失函数下的测试结果混淆矩阵
Fig. 7 Confusion matrix of test results under different losses

图8 蜻蜓目识别GUI
Fig. 8 Odonata recognition GUI

(接正文104页)

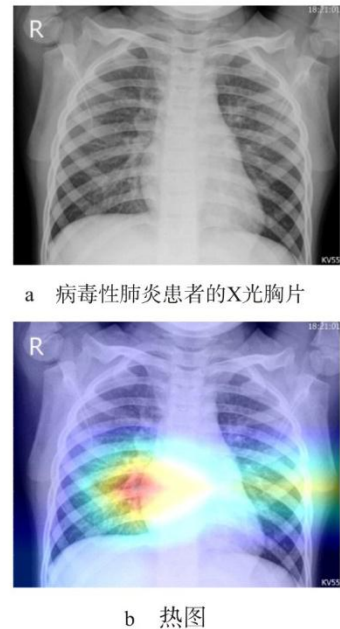


图3 G-Module可视化
Fig. 3 G-Module visualization

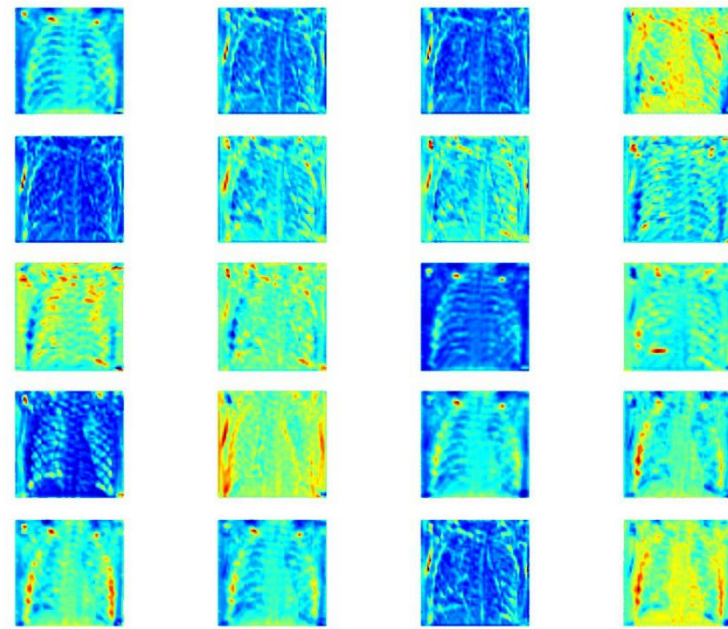


图4 第6个DenseNetLayer后经过1×1卷积后的特征图
Fig. 4 Feature map after 1×1 convolution after the 6th DenseNetLayer