

文章编号: 2095-2163(2022)11-0197-07

中图分类号: TP391

文献标志码: A

# 基于双线性注意力金字塔网络的压疮等级识别

陈昱彤<sup>1</sup>, 邓 悟<sup>2</sup>, 何小海<sup>1</sup>

(1 四川大学 电子信息学院, 成都 610065; 2 四川大学 华西天府医院, 成都 610065)

**摘要:** 在实际临床应用场景下,压疮伤口等级分类较多且图像间差距小。针对使用图像识别技术对压疮图像分类难度大的问题,提出了基于双线性注意力金字塔的压疮等级识别网络 BAP-CNN。该网络以细粒度分类网络 APCNN 为基础(细粒度指类间差距小的情况符合压疮等级分类需求),设计了瓶颈注意力模块,并且采用双线性注意力池化的方法,从而提升模型的整体性能和识别准确率。实验结果表明,在细粒度视觉分类数据集和自建压疮伤口图像 SCU-PU 数据集上,改进后的网络 BAP-CNN 与基础网络 APCNN 以及经典细粒度网络 NTS、WSDAN 相比,模型的识别率均有所提升,证明了改进方法的有效性,以及在不同数据集下良好的泛化能力。

**关键词:** 压疮; 图像识别; 细粒度分类; 瓶颈结构; 双线性池化

## Pressure ulcer grade recognition based on bilinear attention pyramid network

CHEN Yutong<sup>1</sup>, DENG Wu<sup>2</sup>, HE Xiaohai<sup>1</sup>

(1 College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2 West China Tianfu Hospital, Sichuan University, Chengdu 610065, China)

**[Abstract]** In actual clinical application scenarios, there are many grades of pressure ulcer wounds and the gap between images is small. Aiming at the difficulty of using image recognition technology to classify pressure ulcer images, a pressure ulcer grade recognition network BAP-CNN based on bilinear attention pyramid is proposed. The network is based on the fine-grained classification network APCNN. The fine-grained means that the small gap between classes meets the requirements of pressure ulcer grade classification. The bottleneck spatial attention module is introduced, and the bilinear attention pooling method is used to improve the model. Therefore, overall performance and recognition accuracy of the model is improved. The experimental results show that on the fine-grained visual classification dataset and the self-built pressure ulcer wound image SCU-PU dataset, compared with the basic network APCNN and the classic fine-grained networks NTS and WSDAN, the recognition of the improved network BAP-CNN is better. The rate of improvement has been improved, which proves the effectiveness of the improved method and the good generalization ability under different datasets.

**[Key words]** pressure ulcer; image identification; fine-grained classification; bottleneck; bilinear pooling

## 0 引言

压疮,又称压力性溃疡(Pressure Ucler, PU),俗称褥疮<sup>[1]</sup>。压疮是临床上常见的一种慢性伤口,所有人都有可能患上压疮,且更高概率发生在重病、神经系统疾病、行动不便、姿势不良或畸形、皮肤受损或营养不良者身上<sup>[2]</sup>。因为压疮具有难以愈合的特性,故将压疮视作医疗质量的重要指标,并已成为公共卫生系统和经济支出上的一个不小负担<sup>[3]</sup>。但是目前 PU 的管理现状却远远不能令人满意,全球护理资源短缺也是不可忽视的事实。目前最普遍使用的 PU 风险评估工具是 Braden 量表,但是 Braden 量表在区分有压疮发生风险的重症监护患

者方面表现出预测有效性欠佳和准确性差的问题,可能无法充分反映重症监护患者的特征<sup>[4]</sup>。到目前为止,亟需一种能够帮助医护人员判别压疮分级的工具,以便及时采取治疗,预防轻度压疮向重度压疮的转变。

随着人工智能的不断发展,现已越来越多地被应用在日常生活中。而将人工智能应用在医疗领域的实践可行性也已获得了充分有效的验证。有研究指出,生物医学大数据(Biomedical Big Data)在学术领域和工业应用领域都是全球范围内到目前为止发展最快的领域之一<sup>[5-6]</sup>。本文探究关于 PU 智慧数字化的诊断方式,有助于减轻一线临床医生的负担,以及提高对病患的护理质量。

**基金项目:** 四川省科技计划项目(2020YFS0298)。

**作者简介:** 陈昱彤(1999-),女,硕士研究生,主要研究方向:计算机视觉、图像处理;邓 悟(1988-),男,硕士研究生,主要研究方向:图像处理、医疗信息化;何小海(1964-),男,博士,教授,博士生导师,主要研究方向:图像处理、模式识别、图像通信。

**通讯作者:** 何小海 Email: nic5602@scu.edu.cn

**收稿日期:** 2022-09-29

近些年,细粒度视觉分类(Fine-Grained Visual Categorization, FGVC)是计算机视觉领域颇具挑战的一个方向,细粒度是指类别更加细分、类间差异更加细小的分类任务<sup>[7]</sup>,符合压疮等级分类的需求。美国国家压疮咨询委员会(National Pressure Ulcer Advisory Panel, NPUAP)发布的《国际性压疮防治指南》中规定根据局部组织破坏的严重程度对压疮级别进行 I ~ IV 期<sup>[3]</sup>的划分,与医生共同完成压疮图像数据集 SCU-PU 的构建。数据集仅包含图像的类别标签,不包含额外的人工标注信息(Annotation),因此采用基于弱监督的细粒度图像分类方法<sup>[8]</sup>。

目前,主流的基于深度学习的弱监督细粒度图像分类方法大致可以分为 2 类<sup>[9]</sup>:基于特征编码的方法和基于局部区域定位的方法。对此可做阐述如下。

(1)基于特征编码的方法。是对特征的高阶信息进行编码,比如 Lin 等人<sup>[10]</sup>提出的 B-CNN 网络使用双线性池化来聚合 2 个独立卷积网络中的特征,计算 2 个 CNN 流输出特征通道的外积以捕获二阶信息。此外, Yu 等人<sup>[11]</sup>提出了跨层双线性池化的方法来捕获层级间的特征信息。

(2)基于局部区域定位的方法。是通过定位网络来明确地分离与特定对象局部区域相关联的细微

外观差异来促进细粒度分类。比如, Yang 等人<sup>[8]</sup>提出的 NTS 网络将金字塔特征应用于 FGVC 任务,在 CNN 上应用金字塔网络(Feature Pyramid Network, FPN)来学习区域定位器<sup>[12]</sup>,定位可能的信息区域有利于后续的分类任务。此外, Ding 等人<sup>[13]</sup>提出的 APCNN 中包含区域选取网络(Region Proposal Network, RPN)以弱监督的方式生成金字塔感兴趣区域<sup>[14]</sup>。

综合上述 2 种方式可以看出,高效获取特征和感兴趣区域是细粒度分类的关键,本文在 APCNN 网络的基础上优化了注意力特征提取模块,并且引入了双线性注意力池化机制,得到了改进的双线性金字塔网络 BAP-CNN 以此提高模型分类的准确率。实验结果表明,优化后的模型在自建的压疮图像数据集 SCU-PU 上取得了良好的识别率。

## 1 基础网络 APCNN

为了能够使用多尺度的特征<sup>[15]</sup>, APCNN 提出了一种新的注意力金字塔卷积神经网络,目的是更有效地集成高层语义和底层细节特征来进行细粒度分类,可以在不使用边框和其他零件标注的情况下联合学习多级信息和细化的特征。该网络与传统细粒度分类网络的不同在于将金字塔层级结构引入,图 1 为 APCNN 网络框图。

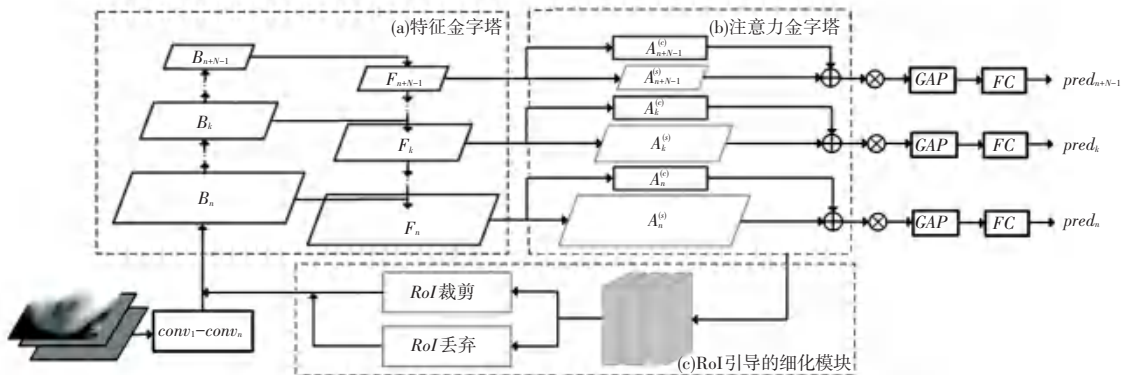


图 1 APCNN 网络示意图

Fig. 1 APCNN network diagram

首先,输入原始图像进入特征金字塔网络和注意力金字塔网络,遵循自上向下的特征路径和自下而上的注意路径,生成金字塔特征和金字塔注意。然后,RPN<sup>[14]</sup>以弱监督的方式生成金字塔感兴趣区域(Region of Interest, RoI),使用 RoI 引导的删除和缩放对图进行细化。将细化后的特征图送入细化阶段,已提取更多有利于分类的鉴别信息。每个金字塔级别都设置了单独的分类器,将特征图和注意力

图点乘后进行全局平均池化(Global Average Pooling, GAP)<sup>[16]</sup>,最后取平均得到分类结果。

### 1.1 注意力金字塔模块

一般的特征金字塔网络(Feature Pyramid Network, FPN)利用各个层次的特征定位不同尺度上的样本,以关注不同尺度上对象的细微差异<sup>[12]</sup>。APCNN 在一阶段通过在金字塔特征上引入额外的注意力机制,形成一个自下而上的注意力金字塔

(Attention Pyramid) 进一步增强 FPN 结构。注意力金字塔包括空间注意和通道注意, 生成过程如图 2 所示。

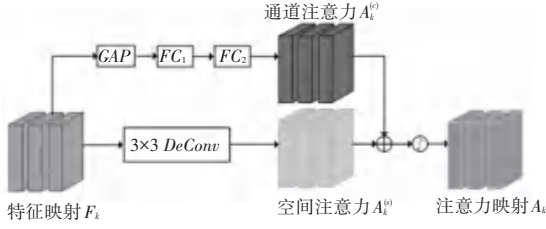


图 2 注意力金字塔模块

Fig. 2 Attention Pyramid Module

金字塔空间注意模块里每层将 FPN 中的特征映射  $F_k$  作为输入, 经过一个  $3 \times 3$  的反卷积层, 以压缩空间信息; 最后使用 *Sigmoid* 层来调整空间注意力输出的尺度, 按照公式(1)生成空间注意掩码  $A_k^{(s)}$ :

$$A_k^{(s)} = \sigma(v_c * F_k) \quad (1)$$

其中,  $\sigma$  表示 *Sigmoid* 函数; “ $*$ ”表示反卷积;  $v_c$  表示的是  $3 \times 3$  的卷积核。结果得到了基于多尺度的空间注意力金字塔  $\{A_n^{(s)}, A_{n+1}^{(s)}, \dots, A_{n+N-1}^{(s)}\}$ 。

金字塔通道注意模块设计思路基于 SE-Net<sup>[17]</sup>, 每层都将 FPN 中的特征映射  $F_k$  作为输入, 经过 *GAP* 层和 2 个全连接 (*FC*) 层; 最后使用 *Sigmoid* 层来调整通道注意力输出的尺度, 按照公式(2)生成通道注意掩码  $A_k^{(c)}$ :

$$A_k^{(c)} = \sigma(W_2 \cdot \text{ReLU}(W_1 \cdot \text{GAP}(F_k))) \quad (2)$$

其中,  $\sigma$  表示 *Sigmoid* 函数; *ReLU* 表示 *ReLU* 函数;  $W_1$  和  $W_2$  是 2 个 *FC* 层的权重矩阵; *GAP*( $\cdot$ ) 表示的全局平均池化函数, 具体如式(3)所示:

$$\text{GAP}(F_k) = \frac{1}{W \times H} \sum_{i=1}^H \sum_{j=1}^W F_k(i, j) \quad (3)$$

其中,  $W$  和  $H$  表示  $F_k$  的空间维度。

*GAP*<sup>[16]</sup> 的目的是对跨越空间维度, 对  $W \times H$  的特征映射进行压缩成一维的并且将感受野扩展到了全局, 能够提取更丰富的特征。2 个 *FC* 层可以更好地拟合通道间复杂的相关性, 自动学习分配特征权重, 提升对当前任务有用的特征通道, 抑制作用不大的特征通道。结果得到通道注意力金字塔  $\{A_n^{(c)}, A_{n+1}^{(c)}, \dots, A_{n+N-1}^{(c)}\}$ 。

从 2 个注意力分支中获取到了空间注意  $A_k^{(s)}$  和通道注意  $A_k^{(c)}$  后, 先对其中的主元素进行求和来实现高效梯度流, 求和后利用 *Sigmoid* 函数将每个元素归一化压缩在  $[0, 1]$  之间, 生成最终的注意力映射  $A_k$ 。

## 1.2 感兴趣区域引导的细化模块

细化模块是指 APCNN 在二阶段对金字塔底层特征  $B_n$  进行 *RoI* 引导的细化, 包括丢弃和裁剪, 从而进一步提高细化阶段的性能。

首先将先前模块得到的注意力映射  $A_k$  输入 RPN<sup>[14]</sup>。RPN 是一种候选框生成网络, 对于每个金字塔级别卷积感受野设计了比例和范围生成目标候选框, 并且对区域建议采用非极大值抑制 (Non-Maximum Suppression, NMS) 来减少区域冗余<sup>[18]</sup>, 这样一来就形成了感兴趣区域 (RoI) 金字塔  $R_{all} = \{R_n, R_{n+1}, \dots, R_{n+N-1}\}$ 。

*RoI* 引导的丢弃策略通过在特征图上随机删除连续区域来删除某些语义信息, 从而强制网络学习剩余区域的信息, 有效防止过拟合问题的产生<sup>[19]</sup>。首先从  $R_{all}$  中随机选取区域  $R_s$ , 然后从  $R_s$  中以相同概率选取信息区域  $r_s \in R_s$ , 并通过将 *RoI* 区域激活设置为 0, 得到丢弃掩码  $M$ , 推导得到的公式可写为:

$$M(i, j) = \begin{cases} 0 & (i, j) \in r_s \\ 1 & \text{otherwise} \end{cases} \quad (4)$$

通过式(5)在底层特征  $B_n$  上应用掩码  $M$ , 并进行归一化得到丢弃的特征映射  $D_n$ :

$$D_n = \sigma(B_n \times M) \quad (5)$$

不同于传统的随机丢弃, *RoI* 引导的丢弃可以直接擦除信息部分, 鼓励网络找到更多有区分性的区域, 从而提高准确率。

*RoI* 引导的裁剪策略通过合并所有 *RoI* 金字塔, 以弱监督的方式学习输入图像的最小边界矩形, 并得到分别表示合并边界矩形在  $x$  轴和  $y$  轴的最小和最大坐标, 表示为  $[t_{x1}, t_{x2}, t_{y1}, t_{y2}]$ , 根据式(6)从之前的丢弃的特征掩码  $D_n$  中提取该区域, 得到细化特征  $Z_n$ , 即:

$$Z_n = \varphi(D_n[t_{y1}:t_{y2}, t_{x1}:t_{x2}]) \quad (6)$$

其中,  $\varphi$  指的是双线性上采样函数。

需要指出的是, 细化操作只在训练过程中执行, 测试过程将跳过这一操作。将细化后的特征  $Z_n$  发送到细化阶段进行下一步预测, 再与原始阶段的预测进行平均, 得到最终预测结果。

## 2 改进网络 BAP-CNN

细粒度分类任务需要解决的 2 个难点在于:

- (1) 高层次的语义, 属于同一类别的物体图像呈现出不同的姿势和视角。
- (2) 低层次的特征, 子类别间的差异通常很细微。

针对以上问题,本文的改进网络 BAP-CNN 主要对基础网络 APCNN 做了 2 方面的改进:

(1)通过改进空间注意力提取的机制对注意力金字塔模块进行优化。

(2)通过改进池化机制对特征融合部分进行优化。

### 2.1 优化空间注意力模块

原始网络中的注意力金字塔模块由空间注意力和通道注意力两部分并联组成,2 个单独的注意力分支计算得出了最终的注意力图。其中,空间注意力模块采用了一个 3×3 的反卷积的方法,反卷积可以增加图像尺寸,从而实现压缩空间信息的目的。

但是反卷积有一个不可避免的缺点,输出的特征图容易出现棋盘效应<sup>[20]</sup>。棋盘效应是由于反卷积的“不均匀重叠”而导致的结果,使图像中某个部位的颜色比其他部位更深。尤其是当卷积核的大小不能被步长整除时,反卷积就会发生不均匀重叠的现象。并且这种棋盘状的现象会层层递进,影响此后的注意力图输出,最终影响整个网络的预测效果。

考虑到对空间注意力模块进行优化,本文参考 ResNet<sup>[21]</sup> 提出的 Bottleneck 瓶颈结构,改进后的瓶颈注意力模块(Bottleneck Attention Module, BAM) 结构如图 3 所示。在瓶颈注意力模块中,空间注意力分支生成过程如下。

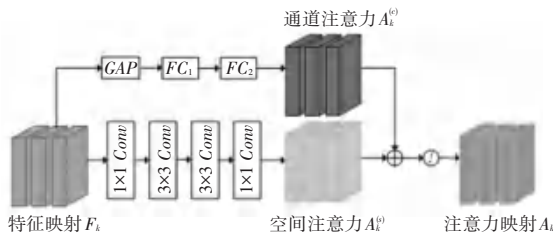


图 3 瓶颈注意力模块

Fig. 3 Bottleneck Attention Module

每层将 FPN 中的特征映射  $F_k$  作为输入,使用一个 1×1 的卷积将原特征映射降维以整合和压缩信道维度的特征图,并通过 2 个 3´3 的膨胀卷积<sup>[22]</sup> 来聚合具有更大感受野的上下文信息,继而使用 1´1 的卷积再次压缩空间注意力图;此后为了调整尺度比例,使用 Sigmoid 函数将空间注意掩码的每个元素归一化压缩在(0,1)之间,按照公式(7)生成空间注意掩码  $A_k^{(s)}$ :

$$A_k^{(s)} = \sigma(f_3^{1 \times 1}(f_2^{3 \times 3}(f_1^{3 \times 3}(f_0^{1 \times 1}(F_k)))))) \quad (7)$$

其中,  $\sigma$  指的是 Sigmoid 函数;  $f$  表示卷积运算,其上标为卷积核大小。

利用膨胀卷积来高效地扩大卷积的感受野,可

以得到比普通卷积更有效的空间映射,加强模型对图像每个像素和周围关系的理解。不仅如此, Bottleneck 瓶颈结构可以减少构造中的参数数量,从而节省计算的开销,提升了网络的速度。

### 2.2 双线性注意力池化

原始网络中在特征融合时使用了点乘之后全局平均池化(GAP)的方法<sup>[16]</sup>,其优势在于可以减少融合时特征的维度,极大地减少了网络的参数。但是缺点是 GAP 操作会造成收敛速度减慢,并且融合后的特征向量会受到注意力图强度的影响,容易丢失一些纹理信息。

针对细粒度图像分类的任务,利用双线性池化操作来融合 2 个独立卷积中的特征是一种经典的基准做法,被广泛地运用于提取图像高阶信息<sup>[10]</sup>。本文引入双线性注意力池化(Bilinear Attention Pooling, BAP)的特征融合方式,通过外积的重组方式,双线性池化能够将基础网络的一阶特征映射成更高阶的特征,显式地表达出更多的图像细节信息,从而提升分类网络的准确率。本文模型加入的双线性注意力池化的过程如图 4 所示。

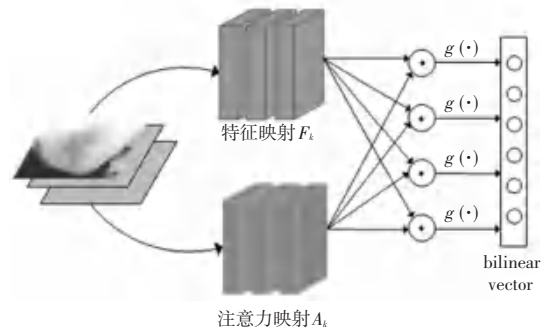


图 4 双线性注意力池化过程图

Fig. 4 Bilinear attention pooling process diagram

首先把网络前面阶段生成的特征图  $F_k$  和注意力图  $A_k$  进行逐元素相乘计算得到部分特征图,然后再对部分特征图进行池化和向量化操作后,得到双线性向量池  $P$ ,用于后续的分类和预测,计算过程可表示为:

$$P = \Gamma(A_k, F_k) = \begin{matrix} \otimes(A_1 \odot F_1) \\ \otimes(A_2 \odot F_2) \\ \dots \\ \otimes(A_n \odot F_n) \end{matrix} \begin{matrix} \circledast \\ \circledast \\ \circledast \\ \circledast \end{matrix} \begin{matrix} \text{vec} \\ \text{vec} \\ \dots \\ \text{vec} \end{matrix} \begin{matrix} \circledast \\ \circledast \\ \dots \\ \circledast \end{matrix} \begin{matrix} \circledast \\ \circledast \\ \dots \\ \circledast \end{matrix} \quad (8)$$

其中,  $g(\cdot)$  表示池化和向量化计算过程,可由如下公式计算求出:

$$g(\cdot) = \text{vec}(\text{SumPool}(\cdot)) \quad (9)$$

本文使用 BAP 对 2 种特征进行融合,再用于后续的分类任务,可以提升网络性能。并且双线性池

化的形式简单,有利于梯度的反向传播,从而实现端到端的训练。

### 3 实验结果与分析

#### 3.1 实验数据集

本文实验在 2 个常用的基准细粒度视觉分类 (FGVC) 数据集上进行训练和测试,包括 CUB-200-2001<sup>[23]</sup>和 FGVC-Aircraft<sup>[24]</sup>。所有数据集包含同一类别下的多组子类别。

为了验证改进后网络 BAP-CNN 的有效性,将改进后的模型在自建的压疮伤口图像数据集 SCU-PU 中进行训练和测试。根据美国国家压疮咨询委员会 (National Pressure Ulcer Advisory Panel, NPUAP) 发布的《国际性压疮防治指南》中规定 I ~ IV 期分期标准<sup>[3]</sup>,增广前的 SCU-PU 数据集包括压疮的 4 个等级,共 140 张图像。

由于 DeiT<sup>[25]</sup>表明数据增强可以实现高性能模型训练和测试,根据压疮图像的特征,为了保留更精细的图像特征信息,不干扰图像纹理特征,选取特定角度翻转和旋转的方式实现数据增强,以扩充压疮数据集协助网络训练。将 SCU-PU 数据集的每一张图像先进行水平、垂直、主对角线和副对角线四种翻转,再每 45°进行一次旋转,共旋转 7 次,最终扩充 40 倍的数据,经过以上数据增广后的 SCU-PU 可以更好地适用于网络训练。本文实验选用的所有数据集的分布情况见表 1。

表 1 实验数据集分布

Tab. 1 The distribution of the dataset used in the experiment

数据集	训练集	测试集	共计
CUB-200-2011	5 994	5 794	11 778
FGVC-Aircraft	6 667	3 333	10 000
SCU-PU	4 360	1 240	5 600

#### 3.2 实验平台与实验细节

本文实验使用的 GPU 为 Nvidia GeForce GTX1080Ti,操作系统为 Ubuntu 16.04,CUDA 版本为 10.0,Python 版本为 3.6,使用深度学习框架 Pytorch 1.3.0,TorchVision 0.4.1。

本文网络与基础模型对比,在 ResNet50<sup>[21]</sup>的基础上实现 BAP-CNN。具体来说,金字塔的层级结构选择 ResNet50 中剩余块 conv3、conv4 和 conv5,分别表示为 B3、B4 和 B5 在最后输出特征,其中细化模块在 B3 上进行。细化模块中的 drop-out 丢弃块率设置为 0.5。关于输入图像大小,对于细粒度视觉分类数据集,根据大多数对比文献中实验输入图像

大小为 448×448;对于自建压疮图像数据集 SCU-PU,在采集后,运用图像处理算法将输入图像大小调整为 224×224。

训练过程中 batchsize 设置为 16,epoch 为 100,初始学习率为 0.001,使用随机梯度下降 SGD 进行优化,设置动量为 0.9,衰减权重为 0.000 01,每隔 2 次迭代、调整一次学习率。

#### 3.3 评价指标与实验结果

为了评估 BAP-CNN 模型对细粒度分类任务和压疮伤口图像分类的性能,本文采用混淆矩阵来可视化计算准确率,并且将准确率作为本文实验的评价指标。准确率可以反映训练网络准确分类各等级的数量,准确率越大,表明正确分类的数量越多。准确率 Accuracy 的数学定义公式见如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

其中,TP、TN、FP、FN 分别表示真阳性、真阴性、假阳性和假阴性。

为了验证改进后的网络 BAP-CNN 的有效性,在 2 个基准细粒度视觉分类 (FGVC) 数据集 CUB-200-2001 和 FGVC-Aircraft 上进行训练和测试,并与当前主流代表性的细粒度分类网络 B-CNN<sup>[10]</sup>、NTS<sup>[8]</sup>和 WSDAN<sup>[26]</sup>进行比较,定量对比实验结果见表 2。

表 2 不同模型在 FGVC 数据集上对比

Tab. 2 Comparison of different models on the FGVC dataset

Models	Baseline	Accuracy /%	
		CUB-200-2001	FGVC-Aircraft
B-CNN <sup>[10]</sup>	VGG16	84.1	84.1
NTS <sup>[12]</sup>	ResNet50	87.5	91.4
WSDAN <sup>[27]</sup>	InceptionV3	89.4	93.0
APCNN <sup>[14]</sup>	ResNet50	88.4	93.4
BAP-CNN (The proposed)	ResNet50	90.3	94.5

表 2 列出了多个模型与本文中的 BAP-CNN 模型在 FGVC 数据集中的准确率 Accuracy 对比。其中,在 CUB-200-2001 数据集上,与经典的双线性卷积网络 B-CNN 模型相比,本文设计出的 BAP-CNN 模型的准确率从 84.1%增长到了 90.3%,提升了 6.2%,与经典的细粒度网络 NTS 和 WSDAN 相比,准确率分别提升了 2.8%、0.9%。而在 FGVC-Aircraft 数据集上,与经典网络 B-CNN、NTS 和 WSDAN 相比,准确率也均有提升,分别提升了 10.4%、3.1%和 1.5%。同时,在 2 个基准细粒度视觉

分类(FGVC)数据集上,本文的改进网络 BAP-CNN 与基础网络 APCNN 相比可知,识别准确率均能提升,分别提升了 1.9% 和 1.1%。说明本文 BAP-CNN 模型的方法具有较好的分类准确率,验证了改进网络的有效性。

本文 BAP-CNN 模型对基础网络主要进行了 2 个方面的改进,为了进一步验证本文改进方法对提升模型识别准确率性能的有效性,在基础网络 APCNN 上逐一进行模块的改进,通过消融实验来分别验证各个模块优化的有效性。实验中网络的 Baseline 均选用 ResNet50,实验结果见表 3。

表 3 消融实验在 FGVC 数据集上结果

Tab. 3 Results of ablation experiments on the FGVC dataset

Models	Accuracy/ %	
	CUB-200-2001	FGVC-Aircraft
APCNN <sup>[14]</sup>	88.4	93.4
APCNN+BAM	88.9	94.1
APCNN+BAP	89.4	93.8
BAP-CNN	90.3	94.5
(The proposed)		

从表 3 所列出的消融实验结果可知,在 2 个 FGVC 数据集上,对不同模块改进的 2 个网络在识别准确率上均有提升作用。与基础网络 APCNN 对比,改进了瓶颈注意力模块,在 FGVC-Aircraft 数据集上的识别准确率提升较大,提升了 0.9%;引入了双线性注意力池化的网络,在 CUB-200-2001 数据集上的识别准确率提升较大,提升了 1.0%。而且,根据表 3 的实验结果可知,同时改进了 2 个模块的本文网络 BAP-CNN 与基础网络对比在识别效果上有更明显的改进,分别提升了 1.9% 和 1.1%。

为了验证改进后网络 BAP-CNN 的有效性,将改进后的模型在自建的压疮伤口图像数据集 SCU-PU 中进行训练和测试,并与细粒度分类网络 NTS 和 WSDAN 在识别准确率上进行比较,定量对比实验结果见表 4。

表 4 不同模型在 SCU-PU 数据集上对比

Tab. 4 Comparison of different models on the SCU-PU dataset

Models	Baseline	Accuracy/ %
NTS <sup>[12]</sup>	ResNet50	67.2
WSDAN <sup>[27]</sup>	Inception V3	68.8
AP-CNN <sup>[14]</sup>	ResNet50	69.7
BAP-CNN	ResNet50	71.1
(The proposed)		

表 4 列出了多个模型与本文中的 BAP-CNN 模

型在自建的压疮伤口图像数据集 SCU-PU 中的准确率(Accuracy)对比。由实验数据可以看出,在压疮图像等级分类任务中相较于传统的细粒度分类网络,本文网络有更高的识别准确率,高于 NTS 网络 3.9%,高于 WSDAN 网络 2.3%。和基础网络 APCNN 对比,识别效果也提升了 1.4%,验证了本文方法有较好的分类准确性和良好的泛化能力。

为了进一步探究压疮图像各个等级分类任务的差异性,使用基础网络和本文网络 BAP-CNN 针对压疮图像的 I 期、II 期、III 期和 IV 期四种等级分别进行对比实验。实验中,网络的 Baseline 均选用 ResNet50,实验结果见表 5。

表 5 SCU-PU 数据集上压疮各等级识别率对比

Tab. 5 Comparison of the recognition rates of pressure ulcers at different levels on the SCU-PU dataset

Models	Accuracy/ %			
	I 期	II 期	III 期	IV 期
AP-CNN <sup>[14]</sup>	82.8	74.1	71.7	45.6
BAP-CNN(The proposed)	83.3	74.4	73.6	47.9

从表 5 所列出的对比实验结果可以看出,对于不同等级的压疮图像样本,识别的准确率不同。从 I 期到 IV 期压疮等级的样本,识别准确率依次降低。其中,I 期与 II 期、III 期的压疮等级识别率差异相对较小,而 I 期与 IV 期的压疮等级识别率相差较大,达到了 35% 左右。

压疮分级是根据局部组织破坏的严重程度进行划分的<sup>[3]</sup>,上述实验数据分析说明,不同程度的局部组织破坏会对压疮产生不同的影响。这也与实际的临床观察结果相对应,越高等级的压疮伤口、其局部组织破坏的程度越严重,可以提取到的特征差异性较小,所以导致了识别率不高的问题。

## 4 结束语

为了更好地识别不同等级的压疮图像,本文提出了一种基于双线性注意力金字塔网络的压疮等级识别方法。首先,输入原始图像进入金字塔网络,生成特征图和注意力图;然后,根据注意力图生成的 RoI 引导,对金字塔底层特征进行细化;最后,引入双线性注意力池化对特征图和注意力图进行特征融合得到双线性向量池,输入线性分类层得到识别结果。改进后的网络 BAP-CNN 在细粒度视觉分类数据集 CUB-200-2001 和 FGVC-Aircraft 上表现良好,相比基础网络和经典细粒度分类取得了更高的准确率,验证了网络改进的有效性。并且在自建压

疮图像数据集 SCU-PU 上测试得到了 71.1% 的识别准确率, 表明 BAP-CNN 在识别分类任务上具有较好的泛化能力。

但是在压疮图像数据集上却仍有不足, 采集到的不同等级的压疮图像数量存在差异, 数据集构建在未来还有亟待丰富扩充的空间。由于数据集存在不足对压疮等级识别准确率有一定影响, 所以本文网络在公开数据集上表现良好, 但是在自建数据集中识别准确率却较低, 针对如何提升不同等级的压疮的识别率, 在下一步工作中可以继续探索。

## 参考文献

- [1] ZUO Xiaolin, MENG Fanjie. A care bundle for pressure ulcer treatment in intensive care units [J]. *International Journal of Nursing Sciences*, 2015, 2(4): 340-347.
- [2] 褚万立, 郝岱峰. 美国国家压疮咨询委员会 2016 年压力性损伤的定义和分期解读 [J]. *中华损伤与修复杂志 (电子版)*, 2018, 13(01): 64-68.
- [3] KOTTNER J, CUDDIGAN J, CARVILLE K, et al. Prevention and treatment of pressure ulcers/injuries: The protocol for the second update of the international Clinical Practice Guideline 2019 [J]. *Journal of Tissue Viability*, 2019, 28(2): 51-58.
- [4] JIANG M, MA Y, GUO S, et al. Using machine learning technologies in pressure injury management: systematic review [J]. *JMIR medical informatics*, 2021, 9(3): e25704.
- [5] DUAN Y, EDWARDS J S, DWIVEDI Y K. Artificial intelligence for decision making in the era of Big Data - evolution, challenges and research agenda [J]. *International Journal of Information Management*, 2019, 48: 63-71.
- [6] YU K H, BEAM A L, KOHANE I S. Artificial intelligence in healthcare [J]. *Nature Biomedical Engineering*, 2018, 2(10): 719-731.
- [7] 邓旭冉, 闵少波, 徐静远, 等. 深度细粒度图像识别研究综述 [J]. *南京信息工程大学学报 (自然科学版)*, 2019, 11(06): 625-637.
- [8] YANG Ze, LUO Tiange, WANG Dong, et al. Learning to navigate for fine-grained classification [C]//*Proceedings of the European Conference on Computer Vision (ECCV)*. Munich, Germany: Amazon, 2018: 420-435. (12)
- [9] FU Jianlong, ZHENG Heliang, MEI Tao. Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii: IEEE, 2017: 4438-4446.
- [10] LIN T Y, ROYCHOWDHURY A, MAJI S. Bilinear CNN models for fine-grained visual recognition [C]//*Proceedings of the IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015: 1449-1457.
- [11] YU Chaojian, ZHAO Xinyi, ZHENG Qi, et al. Hierarchical bilinear pooling for fine-grained visual recognition [C]//*Proceedings of the European Conference on Computer Vision (ECCV)*. Munich, Germany: Amazon, 2018: 574-589.
- [12] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii: IEEE, 2017: 2117-2125.
- [13] DING Yifeng, MA Zhangyu, WEN Shaoguo, et al. AP-CNN: Weakly supervised attention pyramid convolutional neural network for fine-grained visual classification [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2826-2836.
- [14] REN Shaoqing, HE Kaiming, 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*, 2017, 39(6): 1137-1149.
- [15] ZHAO Qijie, SHENG Tao, WANG Yongtao, et al. M2det: A single-shot object detector based on multi-level feature pyramid network [C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Honolulu, Hawaii, USA: AAAI, 2019, 33(01): 9259-9266.
- [16] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization [C]//*Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. Las Vegas, Nevada: IEEE, 2016: 2921-2929.
- [17] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 7132-7141.
- [18] JIANG Borui, LUO Ruixuan, MAO Jiayuan, et al. Acquisition of localization confidence for accurate object detection [C]//*Proceedings of the European Conference on Computer Vision (ECCV)*. Munich, Germany: dblp, 2018: 784-799.
- [19] GHIASI G, LIN T Y, LE Q V. Dropblock: A regularization method for convolutional networks [J]. *arXiv preprint arXiv: 1810.12890*, 2018.
- [20] SUGAWARA Y, SHIOTA S, KIYA H. Checkerboard artifacts free convolutional neural networks [J]. *APSIPA Transactions on Signal and Information Processing*, 2019, 8.
- [21] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//*Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [22] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [J]. *arXiv preprint arXiv: 1511.07122*, 2015.
- [23] WAH C, BRANSON S, WELINDER P, et al. The caltech-ucsd birds - 200 - 2011 dataset [R]. USA: California Institute of Technology, 2011.
- [24] MAJI S, RAHTU E, KANNALA J, et al. Fine-grained visual classification of aircraft [J]. *arXiv preprint arXiv: 1306.5151*, 2013.
- [25] TOURVRON H, CORD M, DOUZE M, et al. Training data-efficient image transformers & distillation through attention [C]//*International Conference on Machine Learning*. PMLR, 2021: 10347-10357.
- [26] HU T, QI H, HUANG Q, et al. See better before looking closer: Weakly supervised data augmentation network for fine-grained visual classification [J]. *arXiv preprint arXiv: 1901.09891*, 2019.