融合局部与全局特征的轻量化乳腺肿瘤分类模型

宋青青， 陈海龙， 王馨怡， 张瀚匀， 李 骜
（哈尔滨理工大学 计算机科学与技术学院，哈尔滨 150080）

摘 要：乳腺组织中的复杂空间结构与纹理细节特征是肿瘤分类的关键因素。针对当前网络对特征提取不充分、特征间依赖关系匮乏及分类不平衡与速度慢等问题，提出了一种轻量化局部与全局特征融合网络。首先，设计了轻量化多维动态卷积模块（lightweight multi-dimensional dynamic convolution，LMDC）提取乳腺图像中的关键局部特征，结合了不同输入图像的空间、通道等维度特征的独特性，又减少了模型参数量。其次，设计了轻量化圆形循环膨胀卷积模块（circular recurrent dilated-convolution，CIR-C）增加网络感受野，加强了乳腺特征的空间联系。最后，在Focal loss损失函数中引入类别权重调整因子，解决乳腺肿瘤类别不平衡问题。实验结果表明：本文模型在BreakHis与BACH数据集上分别取得了98.1%、99.0%的识别准确率，模型参数量和计算量仅为9.04 MB和1.89 GB，极大提升了乳腺肿瘤分类性能。

关键词：乳腺肿瘤；轻量化模型；动态卷积；特征融合；自注意力机制
DOI：10.15938/j.jhust.2026.04.005
中图分类号：TP391.5 **文献标志码**：A **文章编号**：1007-2683(2026)04-0041-15

A Lightweight Breast Tumor Classification Model Combining Local and Global Features

SONG Qingqing， CHEN Hailong， WANG Xinyi， ZHANG Hanyun， LI Ao
（School of Computer Science and Technology，Harbin University of Science and Technology，Harbin 150080，China）

Abstract：The complex spatial structure and texture details in breast tissue are key factors in tumor classification. In this paper，a lightweight local and global feature fusion network is proposed to solve the problems of inadequate feature extraction，lack of dependency between features，unbalanced classification and slow classification. First，a lightweight multi-dimensional dynamic convolution（LMDC）module is designed to extract key local features from breast images. This module captures the uniqueness of spatial and channel dimensions across different input images while reducing model parameters. Secondly，a lightweight circular recurrent dilated convolution module（CIR-C）is designed to increase the receptive field of the network and strengthen the spatial connections of breast features. Finally，the classification weight adjustment factor was introduced into the Focal loss function to solve the classification imbalance problem of breast tumors. Experiments demonstrate that the model proposed in this paper achieves recognition accuracies of 98.1% and 99.0% on the BreakHis and BACH datasets respectively. The model's parameter size is merely 9.04 MB and its computational cost is only 1.89 GB，which significantly enhances the classification performance of breast tumors.

Keywords：breast tumour；lightweight model；dynamic convolution；feature fusion；self-attention mechanism

收稿日期：2024-10-21
基金项目：国家自然科学基金面上项目(62071157).
作者简介：宋青青(2001—)，女，硕士研究生；
王馨怡(2000—)，女，硕士研究生.
通信作者：陈海龙(1975—)，男，博士，教授，E-mail：hrbustchl@163.com.

0 引 言

乳腺癌是女性群体中最高发的癌症之一,近年,乳腺癌的发病率一直呈上升趋势[1]。早期的诊断和干预治疗是降低乳腺癌死亡率的关键,常见的传统诊断方法包括乳腺X射线、乳腺超声、核磁共振、活检及乳腺组织病理图像检查[2],且组织病理学检查是临床诊断乳腺癌的最优标准[3]。然而,医生通过显微镜来观察乳腺病理组织图像,不仅耗时耗力,且不同病理学家的诊断结果也存在差异。因此,设计一种高效而准确的乳腺肿瘤分类方法具有重要的临床应用价值。

随着科技的快速发展,深度学习技术在乳腺肿瘤检测中得到了广泛应用。卷积神经网络(convolutional neural network,CNN)通过局部感受野自动提取图像中的细节特征(如边缘、纹理),可用于乳腺肿瘤检测。TOMA 等[4]使用 ResNet34、ResNeXt、SENet、DenseNet、NASNet 与 Wide ResNet34 等网络对乳腺肿瘤特征进行学习,并取得了良好成果。许学斌等[5]提出了融合乳腺组织空间特征信息和通道特征信息的分类模型 BCSCNet,通过 DenseNet 网络建立不同层次特征的联系,并进行空间特征融合。同时,嵌入压缩激励网络(squeeze-and-excitation network,SENet)学习不同通道特征的重要性,在 Breakhis 数据集取得了 95.72% 的多分类准确率。但由于 CNN 存在局部相关性与权值共享性,导致其提取的特征缺乏全局特征信息交流。

在 CNN 中,较深的网络层只能获取初始层的一些低级信息,导致 CNN 网络失去了特征的全局上下文联系。虽然通过增加卷积核的数量可以提高表示能力,但会大大增加计算负担。为解决 CNN 网络的缺陷,Vision Transformer(ViT)模型[6]利用自注意力机制捕捉全局特征信息联系。文[7]使用预训练的 ViT 模型有效地处理组织病理图像中的空间信息,并用监督对比损失来增强样本中显著特征提取的判别能力。CHEN 等[8]通过局部 ViT 块分别提取乳腺 X 线照片的两侧特征,将这些特征合并输入全局 ViT 块,以实现乳腺良恶性肿瘤的分类。由此可见,ViT 模型可以有效地提取图像中的全局特征,但同类肿瘤间乳腺特征存在巨大差异,如病理图像中的细胞形态、组织结构和纹理特征,只依靠局部或全局特征无法实现更准确的乳腺肿瘤分类。

CNN 网络与 ViT 网络存在特征提取互补性,将其有效结合可极大提升乳腺肿瘤分类水平。ZHUANG 等[9]构造了 ResNet101 和 Swin Transformer 的并行结构,并结合 CBAM 机制捕捉特征空间与通道依赖关系,充分利用了 Transformer 的全局建模能力与 ResNet101 的局部特征提取能力,在 Breakhis 数据集上实现了 96.6% 的分类准确率。ZOU 等[10]设计了 DCET-Net 双流网络,采用 2 个独立 CNN 从乳腺组织图像不同部分提取局部特征,再通过 Transformer 捕捉复杂的空间关系。李广丽等[11]提出了在线学习并融合特征的 MVT-OFML,以 ResNet-50 提取局部特征,用多个 ViT 编码器多视角编码乳腺特征序列,结合注意力机制排序乳腺病理特征的重要性,融合相邻视角输出强化全局上下文特征。

此外,模型面临平衡轻量级要求与精确分类乳腺肿瘤的重大挑战。MobileNet[12]、ShuffleNet[13]等轻量化模型应运而生。方祯祺等[14]采用分组卷积与深度卷积代替 EfficientNet 网络中的普通卷积,有效降低了模型的参数量和计算复杂度。此外,在分组卷积后进行通道混洗,并对空间和通道特征进行加权,增强信息融合与网络鲁棒性。文[15]通过将 Vanilla 卷积分解为深度非对称卷积和点卷积缓解网络分割负担,参数量下降 29.54M。文[16]重构 ShuffleNet 架构并引入 PAConv 并行卷积层,结合深度可分离卷积,网络参数量降低了 78.63%。为解决 Vision Transformer 计算复杂度高的问题,SwinTransformer[17]、Mobile Vit[18]将特征图划分成了多个不相交的区域,在窗口之间先进行局部注意力运算,随后再通过跨区域的信息交流来增强全局建模能力,在保证计算效率的同时,减轻了网络的运算负担。

综上所述,深度学习技术在乳腺肿瘤诊断中发挥着重要作用,但乳腺肿瘤的分类任务还存在以下不足:①传统的静态卷积神经网络对不同图像使用相同卷积核,忽略了乳腺组织图像的特异性和多样性。②多网络特征融合虽能弥补单一网络弊端,但是部分 CNN 与 Transformer 参数量大、计算复杂,二者结合会加剧网络负担。③乳腺肿瘤病理图像数据集类别样本不均,肿瘤特征差异大,易引发训练时的类别分类不平衡。

结合以上研究的优缺点,本文提出基于局部和全局特征融合的 LMDC-MVCD 轻量化网络,具体如下:①设计轻量级多维度卷积 LMDC 模块提取局部特征,动态生成依赖于原始图像的多维度细节信息(包括空间、通道和卷积核等维度)增强图像重要特征表达,并进行轻量化处理,提升网络运算效率。②设计全局特征提取模块(MoblieViT-CIR-C,MVCD)加强乳腺组织图像不同区域的特征信息交

流,同时提出圆形循环膨胀卷积 CIR-C 扩大网络感受野,更好地捕捉全局信息。③针对乳腺肿瘤类别分类不平衡问题,为 Focal Loss 引入动态权重调整因子,合理分配网络对难/易分类样本的关注度。本文在公开数据集 BreakHis 与 BACH 上取得了良好的准确率,为今后组织病理图像分析提供了重要参考。

1　网络结构设计

本文提出了局部特征与全局特征融合的乳腺肿瘤分类模型,该模型由 LMDC 局部特征提取模块、MVDC 空间特征提取模块 、FFM 特征提取模块及分类器组成。

1.1　LMDC 模块

乳腺肿瘤图像具有丰富的纹理及轮廓等细节特征,对其进行局部特征提取是进行乳腺肿瘤多分类的关键步骤。本文提出轻量级通道与空间多维度动态卷积模块 LMDC,该模块遵循 Mobliev2 模型[12]一般架构,将传统卷积模型中的残差结构替换为多维度动态的深度卷积(depthwise conv,DW)与点卷积(point-wise conv,PW)组合的倒残差结构,如图 1 所示,增强网络局部特征提取的能力,并减少了模型参数。

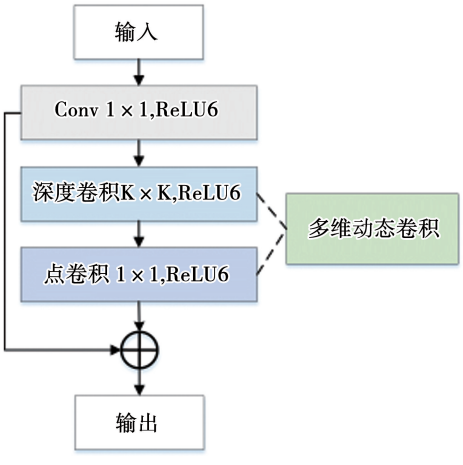


图 1　LCDC 结构图

Fig. 1　LCDC structure diagram

1.1.1　深度可分离卷积

标准卷积输入特征图大小为[C_{in},H_{in},W_{in}],卷积核大小为[K,K],输出特征图尺寸为[C_{out},H_{out},W_{out}],不考虑偏置,其参数量如下:

$$C_{in}K^2C_{out} \tag{1}$$

式中:C_{in}为输入通道维数;C_{out}为输出通道维数;K 为卷积核大小;H 为特征图的宽度;W 为特征图的高度。

图 2 为深度可分离卷积示意图,其将标准卷积分成[C_{in},K,K]大小的通道卷积操作与[C_{in},1,1] × C_{out}的点卷积操作。前者用来调整特征图的尺寸与输入各通道内的特征提取,后者用来调整输出特征图的通道数,其运算参数量如下:

$$C_{in}K^2 + C_{in}C_{out} \tag{2}$$

后者与前者参数量之比如式(3)所示,可见后者极大减少了模型的参数量。

$$\frac{1}{C_{out}} + \frac{1}{K^2} \tag{3}$$

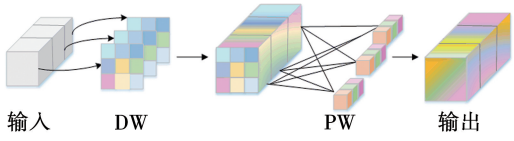


图 2　深度可分离卷积示意图

Fig. 2　Depth separable convolution schematic

1.1.2　轻量化多维动态卷积模块

传统卷积网络通过一组静态卷积核提取局部特征,且初始卷积核与输入特征无关,提取特征时会降低不同输入样本特征的差异性。目前大多数研究在特征图上采用注意力机制,增强输入特征图的细节表示,增加了网络模型的负担,而动态卷积的权重源于输入特征图[19],具有输入依赖性,且参数量较前者大大降低。本文基于此设计了轻量级多维动态卷积 LMDC 模块,如图 3 所示。LMDC 模块先后在卷积核空间 $K \times K$、输入通道 C_{in}、输出通道 C_{out}3 个维度更新原始卷积核权重。其中,空间维度应用于 DW 卷积,通道维度作用于 PW 卷积,3 个维度的权重分别为

$$A_{K\times K} = \{a_{k,1}, a_{k,2}, a_{k,3}, \cdots, a_{k,k\times k}\} \tag{4}$$

$$A_{C_{in}} = \{a_{c,1}, a_{c,2}, a_{c,3}, \cdots, a_{c,C_{in}}\} \tag{5}$$

$$A_{C_{out}} = \{a_{f,1}, a_{f,2}, a_{f,3}, \cdots, a_{f,C_{out}}\} \tag{6}$$

式中:$a_{k,i}$为卷积核空间权重;$a_{c,i}$为输入通道维度权重;$a_{f,i}$为输出通道维度权重。

具体来说,对于卷积核空间维度,动态调整全局特征权重可以提升特征图中纹理信息的表达能力,从而帮助模型更好地捕捉关键细节。因此,本文采用空间注意力机制(SAM)[20]作为卷积核空间维度的权重 A_1,$A_{K\times K}$为 DW 卷积核空间权重,如式(7) ~

(8)所示。

$$A_1 = \mathrm{Conv2d}(\mathrm{Concat}(\mathrm{Avgpool}(x), \mathrm{Maxpool}(x))) \tag{7}$$

$$A_{K\times K} = \mathrm{sigmoid}(\mathrm{Softpool}(A_1)) \tag{8}$$

式中:Conv2d 为二维卷积;AvgPool 和 Maxpool 分别表示对输入特征[C,H,W]在通道维度进行空间平均池化操作和最大池化操作,得到 2 个[1,H,W]尺寸大小的空间特征权重图,将其进行 Concat 拼接,之后利用卷积核(体积为 1×2×3×3)进行二维卷积运算,得到[1,H,W]尺寸的混合特征空间特征图。此外,为了匹配动态卷积核大小尺寸,本文采用了 SoftPool(如式(8)所示)[21],可以在保持池化层功能的同时,尽可能减少池化过程中带来的信息损失,更好地保留信息特征并且放大更强的特征激活,其表达式如下:

$$\mathrm{SoftPool}(\alpha_R) = \sum_{i\in R}\frac{e^{\alpha_i}\alpha_i}{\sum_{j\in R}e^{\alpha_j}} \tag{9}$$

式中:R 为池化区域大小;α_i、α_j 为池化区域内的每个特征元素。

在卷积核输入通道和输出通道维度,利用通道注意力机制对卷积核输入通道进行动态加权,自适应调整卷积核输入通道的重要性。传统的通道注意力机制,如 SE 通道注意力机制[22],使用全局平均池化(global average pooling,GAP)均值获得通道的权重,无法代表不同特征通道的重要性。因此,本文采用频率通道注意力机制(FCA)[23]作为通道维度权重。FCA 机制将 GAP 推广到二维离散余弦变换(discrete cosine transform,DCT),对特征图引入多个频率分量,将其组合在一起作为通道权重来表征特征之间的相关性,从而保留了特征的细节多样性,公式如下:

$$A_2 = \mathrm{FC}(\mathrm{Concat}(\mathrm{Freq}(x))) \tag{10}$$

$$A_3 = A_2 \tag{11}$$

$$A_{C_{\mathrm{in}}} = \mathrm{Sigmoid}(A_2) \tag{12}$$

$$A_{C_{\mathrm{out}}} = \mathrm{Sigmoid}(\mathrm{Conv1d}(A_3)) \tag{13}$$

式中:FC 为全连接层;Conv1d 为一维卷积;A_2 为输入通道权重;A_3 为输出通道权重。Freq 为对输入特征进行二维 DCT 频率分量变换,并进行拼接操作,得到多光谱特征向量通道符。将这些向量输入全连接层进行学习加权获取注意力权重,并对其进行 Sigmoid 处理,得到卷积核通道权重。最后,在卷积核输出通道维度使用一维卷积调整通道数。

通过以上操作,LMDC 模块能够捕捉到细粒度的乳腺局部特征,为后续的特征融合提供丰富的细节信息。

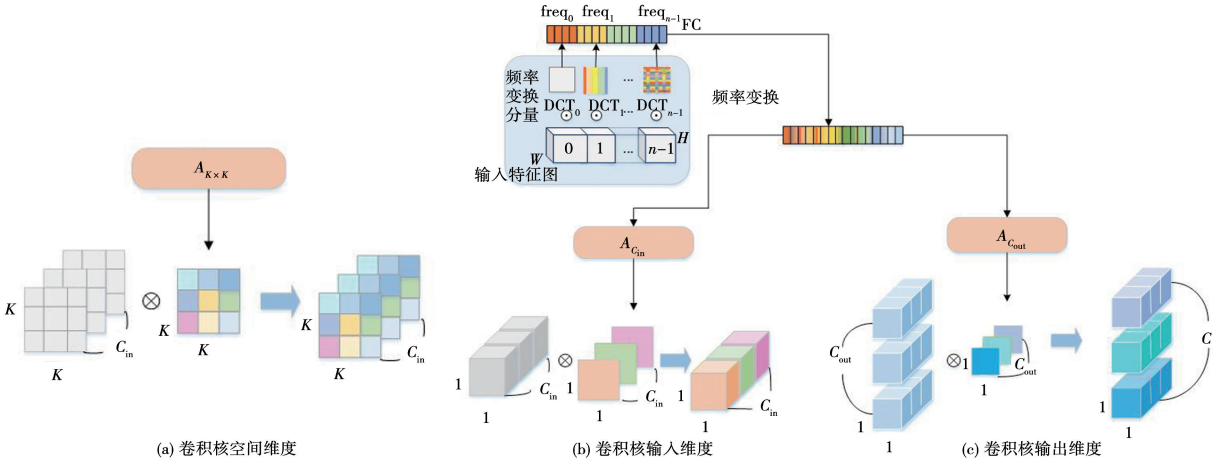


图 3 多维动态卷积计算过程

Fig. 3 Multi-dimensional dynamic convolution calculation process

1.2 MVCD 模块

局部特征能够揭示病变区域的宏观模式和组织结构变化,但不同类型的乳腺病变区域与组织图像之间的全局特征联系存在差异,因此全局特征对肿瘤分类也十分重要。相比于 CNN,ViT 利用自注意力机制实现不同区域信息的长期依赖联系。为了减少网络计算复杂度并建立乳腺组织全局上下文联系,本文设计了轻量化全局特征提取分支 MVCD 模块,如图 4 所示。其中,LightVit 模块计算不同区域的注意力联系,Patch Merging 操作加强乳腺组织的空间特征交流,CIR-Convotion 增加网络感受野。

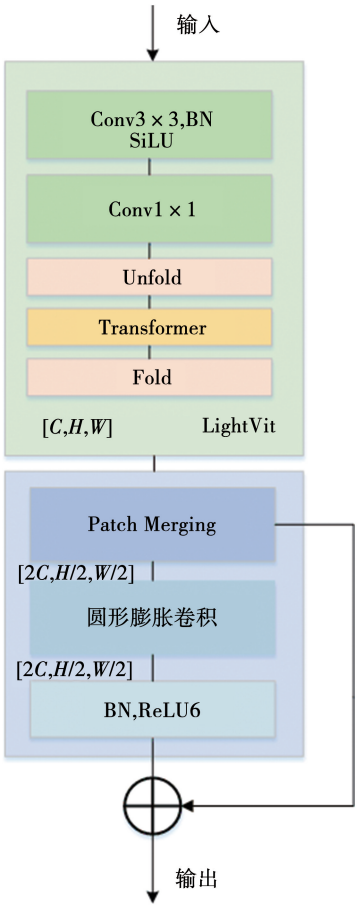


图 4　MVCD 结构图

Fig. 4　MVCD structure diagram

1.2.1　LightViT 模块

Transformer 中的计算开销主要来源于自注意力层,原始的自注意力计算将每个令牌(Token)与所有的 Token 进行自注意力计算,加大了网络计算负担。本模块中,遵循 MobileVit 架构,首先利用 3 × 3 卷积与 1 × 1 卷积进行简单特征提取,之后按照每层图像块(Patch)尺寸对特征图[C,H,W]进行划分,如图 5 所示。对于每个 Patch 中的 Token,只与位置相同的 Token 进行自注意力计算,以更少的参数学习更好的特征表达。

1.2.2　圆形膨胀卷积操作 CIR-C

$H \times W$ 为特征图大小,P 为每层 Patch 尺寸,P_{num} 为特征图 Patches 数目。在 LighVit 中自注意力计算时,1 个Patch 内的特征只与其余 Patch 内的 $P_{num}-1$ 个空间位置建立依赖关系,Patch 尺寸越大,网络所构建的全局特征联系就越稀疏,且处于相同 Patch 内的特征无法建立空间联系,导致空间信息匮乏,故本文设计了一个圆形膨胀卷积操作解决上述问题。

本文在 CIR-Convotion 模块前先将特征图[C,H,W]进行 Patch Merging 下采样[17]处理,特征图的相邻像素按照$\frac{H}{2} \times \frac{W}{2}$尺寸划分为 4 倍 Patch,每个 Patch 中的相同空间位置在通道维度进行 Concat 拼接,得到[$4C$,$\frac{H}{2}$,$\frac{W}{2}$]尺寸的特征图。

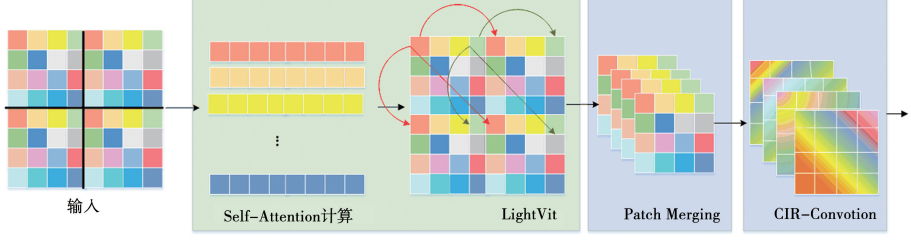


图 5　MVDC 模块特征提取原理图

Fig. 5　Feature extraction schematic diagram of MVDC module

特征图每经过一次卷积,便可提取原特征图所有 Patch 相关信息。其次,为了保持原特征的空间结构,CIR-Convotion 模块在[H,W]尺寸的输入特征图的 W 列后添加原始 $n(K-1)$列,在 H 行添加 $n(K-1)$行,CIR-Convotion 选择深度可分离膨胀卷积核[24]加强不同区域的特征交流并增加模型感受野。$K \times K$ 为膨胀卷积核空间尺寸,n 为膨胀因子,CIR-Convotion 模块如图 6 所示,该层卷积感受野如式(14)所示。

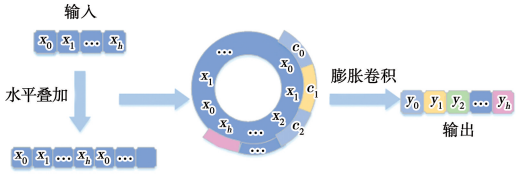


图 6　CIR-Convotion 原理图

Fig. 6　CIR-Convotion schematic

$$S=[n(K-1)+1][n(K-1)+1] \tag{14}$$

i 和 j 为在 $H\times W$ 特征图中空间位置的特征，其在行方向的感受野 S_R 如式(15)所示，在列方向的感受野 S_C 如式(16)所示。

$$S_R=[i,(i+n)\bmod H,\cdots,(i+(k-1)n)\bmod H] \tag{15}$$

$$S_C=[j,(j+n)\bmod W,\cdots,(j+(k-1)n)\bmod W] \tag{16}$$

式中：i,j 均从 0 开始；mod 为取模运算。

通过以上设计，MVCD 模块能够捕捉到乳腺组织的宏观模式和结构信息，为乳腺肿瘤分类提供了至关重要的全局特征支持。

1.3 FFM 融合模块

LMDC 模块提取丰富的乳腺组织纹理细节特征，MVCD 模块聚焦空间结构，提取图像中的宏观结构空间特征。但多尺度特征简单组合方式易丢失关键信息、削弱特征表达。FFM 模块通过交互融合局部特征与全局特征，从而强化其互补性和协同作用，如图 7 所示。将局部特征与全局特征拼接得到 y_1，如式(17)所示。

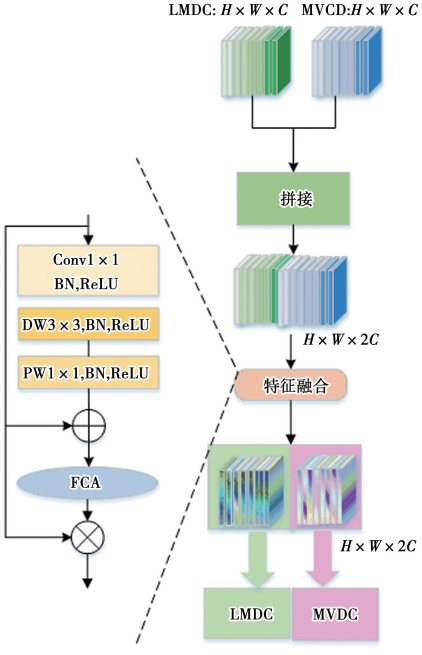


图 7 FFM 结构图

Fig. 7 FFM structure diagram

而后，使用 2 个 1×1 卷积、1 个 3×3 卷积及 BN 层、ReLU 激活函数组成的 Conv 模块来构建双分支的特征融合 y_2，实现特征深度融合。之后形成如式(18)所示的残差连接以缓解模型训练中的梯度消失和梯度爆炸问题。再将 FCA 注意力机制作用于 y_2 生成融合特征注意力权重，对融合特征点乘得到加权特征 y_3，从而增强局部特征与全局特征之间的相关性。最后，沿着通道维度划分特征，将其返回到两分支网络继续训练。特征融合如式(19)所示：

$$y_1=\mathrm{Concat}(x_1,x_2) \tag{17}$$

$$y_2=\mathrm{Conv}(y_1)+y_1 \tag{18}$$

$$y_3=y_2\mathrm{Sigmoid}(\mathrm{FCA}(y_2)) \tag{19}$$

FFM 模块通过多尺度特征自适应加权促进局部与全局特征交互，有效融合 LMDC 与 MVCD 模块的特征，提高乳腺肿瘤分类性能。

1.4 LCMC-MVCD 网络架构

针对乳腺组织图像纹理信息丰富与整体结构复杂的特点，本文设计了 LCMC-MVCD 乳腺肿瘤分类模型，如图 8 所示。预处理后的乳腺组织图像首先经过卷积下采样 Stem 层，Stem 层由 3×3 的卷积、标准化(batch normalization，BN)和 ReLU6 激活函数组成，用来提取初步特征并减少图像的空间尺寸，以便后续的深层网络更高效地学习复杂的乳腺特征。之后通过 LMDC 模块与 MVCD 模块进行多尺度特征提取，FFM 模块进行特征融合与交互，最后利用全局平均池化与全连接层组成的分类器输出乳腺肿瘤类别。具体而言，特征提取分为 4 个并行阶段，每个阶段由 LMDC、MVCD 模块共同组成，层数为[3,4,6,3]，并且随着网络层数加深，特征图通道增加一倍，尺寸减半。之后通过 FFM 模块进行细节与空间特征互补融合，使得模型能够更好地捕捉特征之间的相关性，提高整体分类性能。

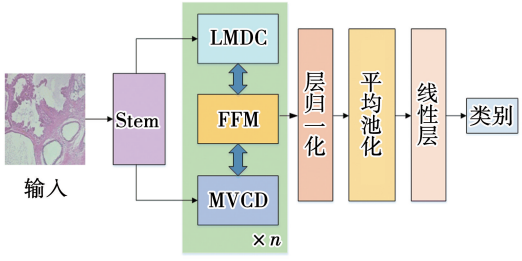


图 8 整体网络结构图

Fig. 8 Overall network structure diagram

1.5 乳腺肿瘤类别不平衡的动态损失函数

经过多次实验得知，不同种类乳腺肿瘤存在较大差异，其中叶状肿瘤与导管癌分类效果较差，腺病与乳头状癌等分类效果较好。因此，需要在每次训练中动态提升难分类肿瘤权重，降低易分类肿瘤权重。本文选择了 Focal loss 函数[25]进行优化，如式(20)所示。Focal loss 函数利用动态缩放的交叉熵损失来解决类别不平衡的问题，通过在反向传播过程中对难分类样本进行加权，对易分类样本进行降

权来提高模型的整体准确率。

乳腺肿瘤分类不平衡直观地表现在每个类别的分类准确率上,其中较难分类肿瘤的准确率低于平均分类准确率。由于类别分类准确率应与损失函数类别权重呈负相关,与平均准确率相差越大,网络对某类乳腺肿瘤关注度越高。基于此,受 Sigmoid 激活函数启发,本文设 $\alpha_{\mathrm{L}i}$ 为每类肿瘤损失函数权重,当某类乳腺肿瘤在训练过程中分类准确率下降且低于平均准确率时,便可增加其损失占比。此外,还设置了 d_i 加强网络对样本的关注差异,与平均准确率相差越大,权重越大。

$$F_L = \sum_{i=0}^{N-1} -\alpha_{\mathrm{L}i}(1-p_i)^{\gamma}\log p_i \tag{20}$$

$$\alpha_{\mathrm{L}i} = \frac{\dfrac{1}{1+\mathrm{e}^{d_i}}}{\sum_{i=0}^{N_i-1}\dfrac{1}{1+\mathrm{e}^{d_i}}} \tag{21}$$

$$d_i = \varepsilon(a_i - M) \tag{22}$$

式中:N 为总类别数;N_i 为第 i 类样本总数;$\alpha_{\mathrm{L}i}$ 为第 i 类别权重值;a_i 为上一 epoch 训练中 i 类分类准确率;M、ε 分别为所有类别准确率的均值与标准差倒数;p_i 为第 i 类模型预测概率;γ 为损失函数的调制因子,这里取 2;$(1-p_i)^{\gamma}$ 可调节样本对损失的贡献。

2　实验与结果分析

2.1　数据集

本文在 BreakHis 与 BACH 数据集上进行实验,BreakHis 数据库是针对乳腺组织病理学研究的开放性数据库,由 7909 张乳腺组织活检切片生成的图像样本组成,该数据集为三通道 RGB 图片,图片格式为 PNG,统一大小为 700 ×460 像素。根据显微镜下肿瘤细胞的形态将图像分为良性肿瘤和恶性肿瘤,良性由腺病(adenosis,AD)、纤维腺瘤(fibroadenoma,FA)、叶状肿瘤(phyllodes tumor,PT)、管状腺瘤(tubular adenoma,TA)组成,恶性肿瘤由导管癌(ductal carcinoma,DC)、小叶癌(lobular carcinoma,LC)、黏液癌(mucinous carcinoma,MC)、乳头状癌(papillary carcinoma,PC)组成,各类由 40、100、200、400 放大倍数的乳腺组织图像组成。表 1 为 BreakHis 数据集 8 类肿瘤组织图像统计。BACH 数据集包括了正常、良性、原位癌与浸润性癌 4 类,每个样本的分辨率为 2048 ×1536,该数据集包含 400 张均匀分布的乳腺病理组织图像,每个类别均为 100 张原始样本。

表 1　BreakHis 数据集乳腺组织图像数量统计表

Tab. 1　BreakHis dataset breast tissue image count table

类型	40 ×	100 ×	200 ×	400 ×	总计
AD	114	113	111	106	444
FA	253	260	264	237	1 014
PT	109	121	108	115	453
TA	149	150	140	130	569
DC	864	903	896	788	3451
LC	156	170	163	137	626
MC	205	222	196	169	792
PC	145	142	135	138	560
总计	1 995	2 081	2 013	1 820	7 909

2.2　数据处理

为了减弱噪声对实验分类精度的影响及模型过拟合问题,本文实验中通过图像标准化、随机裁剪、随机旋转角度的方式对 BreakHis 数据集进行预处理和数据增强[26],有助于提高模型的训练稳定性和收敛速度。

1)图像标准化:利用图像的方差和标准差将数据集中的输入值映射到[0,1]范围。

2)随机图像旋转:将图像从 -45°随机旋转到 +45°。

3)随机裁剪:将图像随机裁剪到指定大小,在本文实验中,图像尺寸统一调整为 224 ×224。

考虑到深度学习需要大量数据来提高网络分类性能,为防止过拟合,本文采用 Augmentor Python 库的水平和垂直翻转、随机旋转、随机对比度、颜色调整等为 BACH 数据集生成了 500 个样本,以此扩大数据集样本数量。

2.3　实验环境

实验环境配置如表 2 所示。

表 2　实验配置

Tab. 2　Experimental configuration

实验工具	具体配置
操作系统	64 位 win10 系统
CPU	Intel Core i5 - 13500H
GPU	NVIDIAGeForce RTX4090
编程语言	Python
深度学习框架	Pytorch

2.4　评价准则

准确性(α_{ACC})是乳腺肿瘤分类的主要评估指标[27]。为了进一步分析模型的性能,本实验还引入了 $F_{1\text{-Score}}$、平衡准确性(α_{BA})、马修相关系数

(α_{MCC})[28]、混淆矩阵、参数量和算力(每秒浮点运算次数,GFLOPs)等多种评价指标,计算公式如下:

$$R = \frac{TP}{TP + FN} \tag{23}$$

$$S = \frac{TN}{TN + FP} \tag{24}$$

$$P = \frac{TP}{TP + FP} \tag{25}$$

$$F_{1-Score} = \frac{2RP}{R + P} \tag{26}$$

$$\alpha_{BA} = \frac{R + S}{2} \tag{27}$$

$$\alpha_{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{28}$$

式中:TP 为模型将某一类别肿瘤正确识别的样本数;FP 为模型将某一类别样本错误分类为其他类别的样本数;FN 为模型错误地将某一类别样本分类为其他类别的样本数;TN 为模型正确地将某一类别以外的样本分类为其他类别的样本数。

2.5 实验参数设置

本文将经过预处理的数据集以 8:2 的比例随机拆分为训练集和测试集。训练过程中采用 AdamW 优化器,批量大小为 32,初始学习率为 10^{-4},训练周期总数为 150,采用余弦退火算法学习率调度器来控制学习速率衰减以达到更好的训练效果。

2.6 消融实验

为验证 LMDC 模块、MVCD 模块、多尺度融合模块的有效性,本文在 BreakHis 与 BACH 数据集分别进行了消融实验。首先,对两模块及其基线模型进行单独实验,进而将其进行组合测试。最后,为了验证动态损失函数的有效性,将 Baseline3 中的 Focal loss 函数的每类肿瘤权重设为相同进行验证。消融实验性能指标如表 3 ~4 所示。

表 3 5 种模型在 BreakHis 数据集的评价指标对比

Tab. 3 Comparison of evaluation metrics of five models on the BreakHis dataset

模型	α_{ACC}/%	$F_{1\text{-Score}}$/%	α_{BA}/%	α_{MCC}/%	参数量/M	算力/($10^9 s^{-1}$)
Baseline1(MoblieV2)	96.4	96.8	97.4	96.4	2.23	0.33
Baseline2(MoblieVit)	96.2	95.8	98.0	95.4	0.95	0.27
LMDC	97.0	97.1	97.8	96.7	2.80	0.49
MVCD	97.1	97.4	98.0	96.9	5.85	1.02
Baseline3(Focal loss)	97.7	97.8	98.4	97.4	9.04	1.8
本文模型	98.1	98.3	98.9	97.9	9.04	1.89

表 4 5 种模型在 BACH 数据集的评价指标对比

Tab. 4 Comparison of evaluation metrics of five models on the BACH dataset

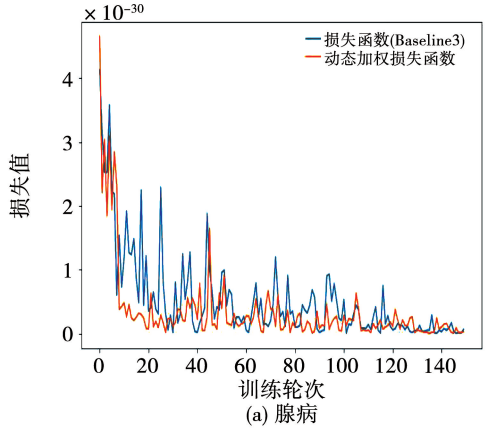
模型	α_{ACC}/%	$F_{1\text{-Score}}$/%	α_{BA}/%	α_{MCC}/%	参数量/M	算力/($10^9 s^{-1}$)
Baseline1(MoblieV2)	95.2	95.4	95.2	93.6	2.23	0.33
Baseline2(MoblieVit)	96.5	96.4	96.3	95.3	0.95	0.27
LMDC	96.7	96.6	96.7	95.6	2.80	0.49
MVCD	97.3	97.4	97.4	96.2	5.85	1.02
Baseline3(Focal loss)	98.1	97.9	98.0	97.5	9.04	1.8
本文模型	99.0	99.1	99.2	98.6	9.04	1.89

如表 3 ~4 所示,相较于 Baseline1 模型,LMDC 模块在引入多维动态卷积核后,在增加少数参数量的情况下,BreakHis 数据集分类准确率提升 0.6%,$F_{1\text{-Score}}$、α_{BA}、α_{MCC}分别提升 0.3%、0.4%、0.3%;BACH 数据集分类准确率提升 1.5%,$F_{1\text{-Score}}$、α_{BA}、α_{MCC} 分别提升 1.2%、1.5%、2.0%,表明 LMDC 模块可使网络更有效地捕捉乳腺组织中的关键特征。BreakHis 数据集实验中,MVCD 模块较基线网络 Baseline2 的分类准确率提升 0.9%,$F_{1\text{-Score}}$、α_{MCC} 分别增加了 1.6%、1.5%;而在 BACH 数据集中准确率提升 0.8%,这表明 MVCD 模块提取了乳腺组织中更加广泛的全局结构特征。相较于 Baseline3,BreaKHis 数据集中本文模型的 $F_{1\text{-Score}}$、α_{BA}、α_{MCC} 均提升 0.5%,BACH 数据集提升 1.2%、1.2%、1.1%,有效缓解了乳腺肿瘤分类不平衡问题。本文模型可以有效融合乳腺组织两分支的特征,在 BreakHis 数据集上达到了 98.1% 的分类

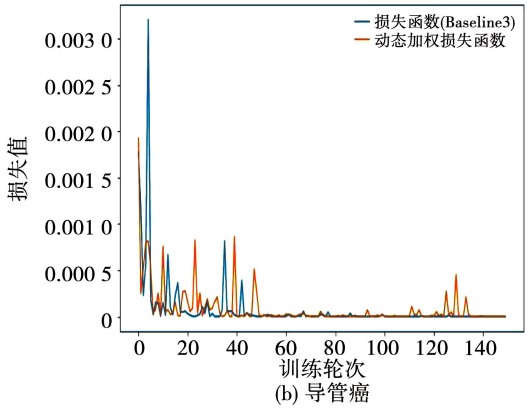
准确率,其他指标较 LMDC、MVCD 模块均有所提升;在 BACH 数据集上达到了 99.0% 的分类准确率,其他指标较 LMDC、MVCD 模块也均有所提升。

2.7　动态损失函数性能研究

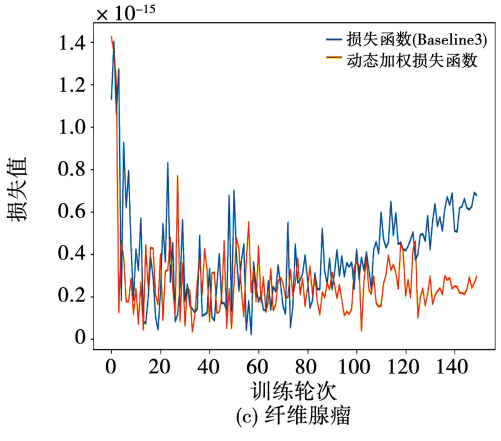
为了更直观地显示动态损失函数的性能,图 9 列举了 BreaKHis 数据集中 Baseline3(Focal loss)与本文模型(动态加权损失函数)在训练中 8 类乳腺肿瘤的损失值对比。其中蓝色线为 Baseline3 模型,橙色线为本文设计模型。



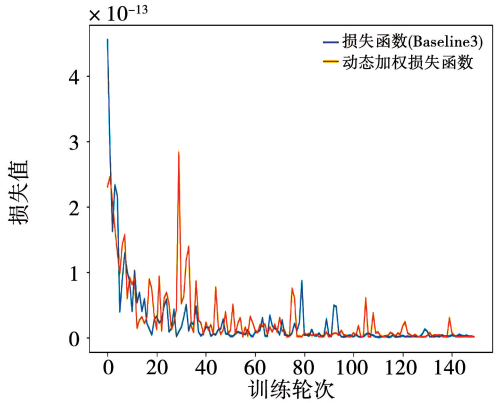
(a) 腺病



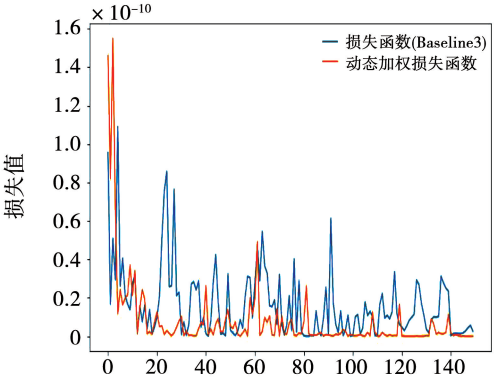
(b) 导管癌



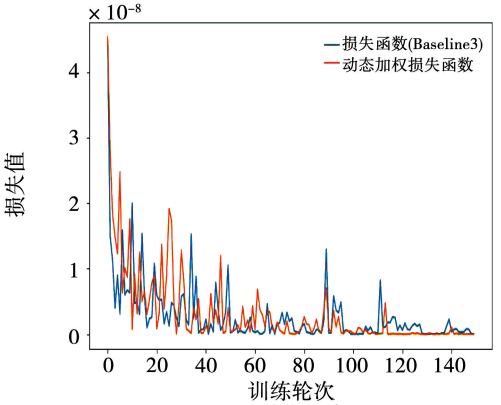
(c) 纤维腺瘤



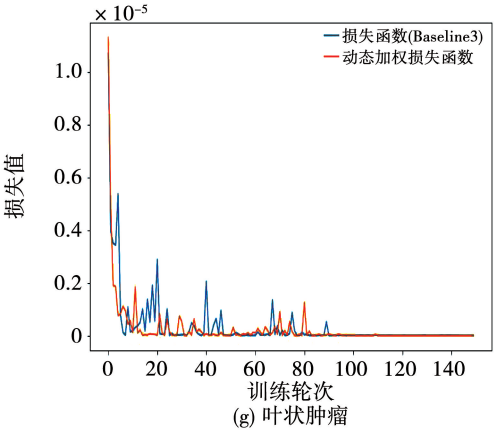
(d) 小叶癌



(e) 黏液癌



(f) 乳头状癌



(g) 叶状肿瘤

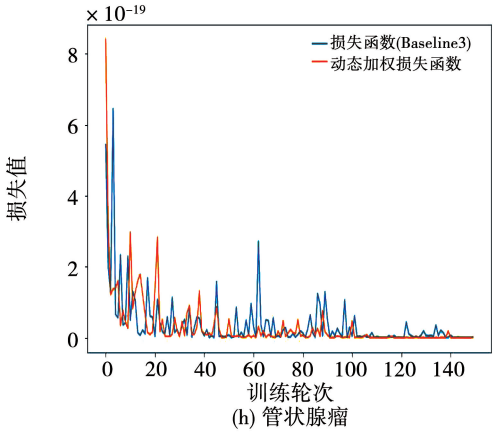


图 9 损失函数值对比

Fig. 9 Comparison of loss function values

表 5 为 BreakHis 数据集中 8 类乳腺肿瘤准确率,表 6 为 BACH 数据集中 4 类乳腺肿瘤准确率。如表 5 所示,腺病(AD)、黏液癌(MC)等肿瘤的分类准确率较高,且由图 9(a)、(e)可知动态加权函数的损失值大多低于 Baseline3 模型的损失值,表明模型将其判定为较易分类样本,故权重降低,损失值减小。由于导管癌(DC)与小叶癌(LC)的分类准确率低于平均分类准确率,对应图 9(b)、(d)中 Baseline3 模型的损失值大多高于本文模型的损失值,且由于小叶癌相较于导管癌的准确率更低,Baseline3 模型的损失值波动较大,表明模型将其判定为较难分类样本,故增加权重,加强了模型对其的关注能力。对于 BACH 数据集,良性与原位癌的分类准确率低于四类肿瘤平均准确率,差值为 3.1%、0.6%,通过动态加权损失函数的调整,良性与原位癌的准确率已上升至 97.55%、99.2%,与平均分类准确率差距缩小至 1.5%、0.2%。

表 5 BreakHis 数据集中 8 类乳腺肿瘤准确率(%)

Tab. 5 Accuracy of eight types of breast tumors in the BreakHis dataset

模型	AD	DC	FA	LC	MC	PC	PT	TA	平均准确率/%
Baseline3(Focal loss)	100	93.7	97.2	91.5	100	97.8	100	99.2	97.4
本文模型	100	95.4	98	94.3	100	99	98.2	100	98.1

表 6 BACH 数据集中 4 类乳腺肿瘤准确率(%)

Tab. 6 Accuracy of four types of breast tumors in the BACH dataset

模型	良性	原位癌	浸润性癌	正常	平均准确率/%
Baseline3(Focal loss)	95.0	97.5	100.0	100.0	98.1
本文模型	97.5	99.2	99.2	100.0	99.0

2.8 可视化

为了深入研究 LMDC 与 MVCD 模块对网络特征表示学习效果的提升作用,本文采用 Grad-Cam 技术[29]可视化乳腺组织图像不同区域特征对网络分类性能的贡献。本文分别展示了消融实验中的 5 种模型在 BreakHis 与 BACH 数据集上的可视化效果,如表 7~8 所示。

表 7~8 中红色区域代表网络重点关注区域,体现了不同模型提取特征注意力水平。由表可见,乳腺分类关注组织中的细胞轮廓、纹理等特征,LCDC 模块可以提取大多数乳腺组织中的重要特征,忽略图像中不相关的信息;MVCD 模型的感受野更为丰富,提取特征范围更广。相比之下,本文模型既关注了乳腺组织图像中的细节特征,又建立了全局信息交流,提高了网络识别精度。

2.9 对比实验

为进一步验证模型分类的有效性,在 BreakHis 与 BACH 数据集上将本文模型与近年热门网络进行对比,还选取了 ResNet34[30]局部特征分支、Swin-Transformer 空间特征分支及二者融合(Res-SwT)的模型进行对比,结果如表 9~10 所示。

由表可见,本文模型分类性能均优于其他网络。首先,相较空间与局部单独特征分支,融合后的 Res-SwT 模型在 BreakHis 数据集和 BACH 数集的准确率、$F_{1\text{-Score}}$、α_{BA}、α_{MCC}值均有所提高,验证了融合特征对于乳腺肿瘤分类的有效性,但参数量与运算量上升了一倍,大幅度增加了网络负担。其次,在 BreakHis 数据集上,本文模型比 Res-SwT 准确率提升 0.6%,但参数量减少了大约 5/6 倍,并且对于主流轻量化模型 MobileV3、ShuffleNet 等,本文模型通过对卷积网络与 Transformer 架构进行轻量化设计减少了网络复杂度,而且还大幅度提高了乳腺肿瘤的识别能力,达到了分类能力与网络性能的平衡。

表 7　BreakHis 数据集的 5 种模型热力图可视化对比

Tab. 7　Heatmap visualization comparison of five models in the BreakHis dataset

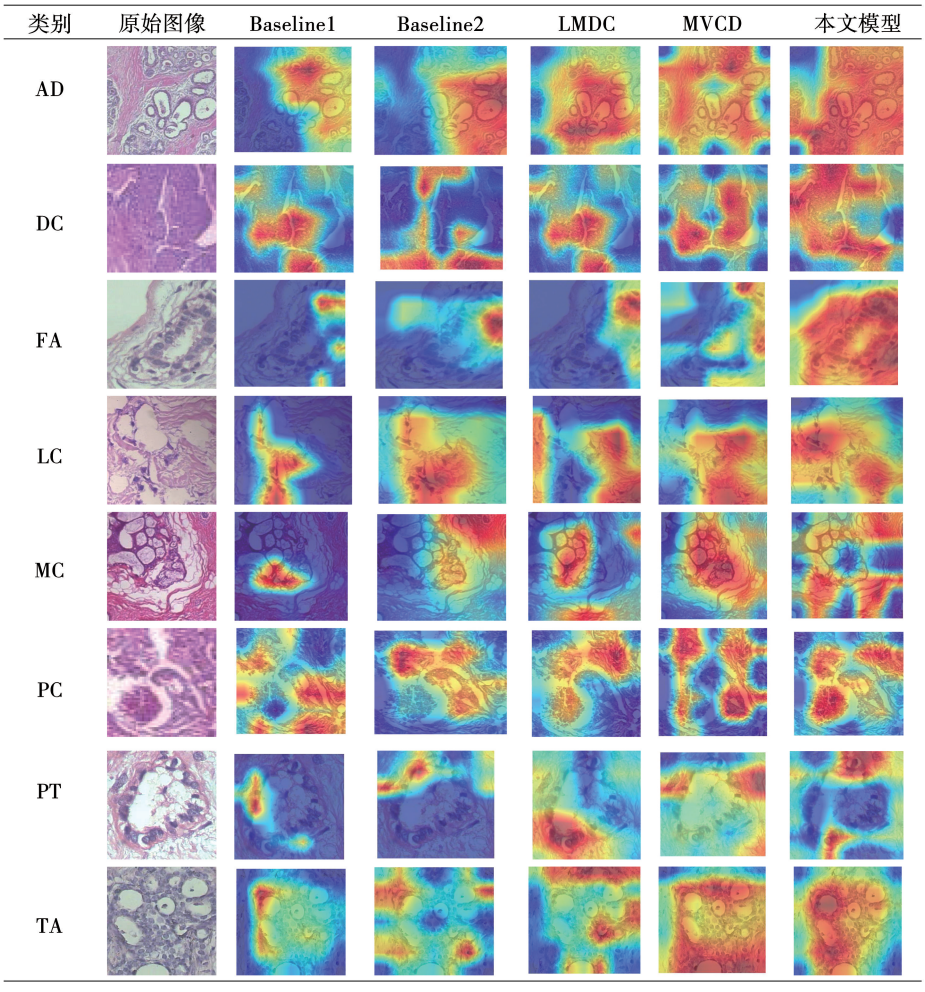


表 8　BACH 数据集的 5 种模型热力图可视化对比

Tab. 8　Heatmap visualization comparison of five models in the BACH dataset

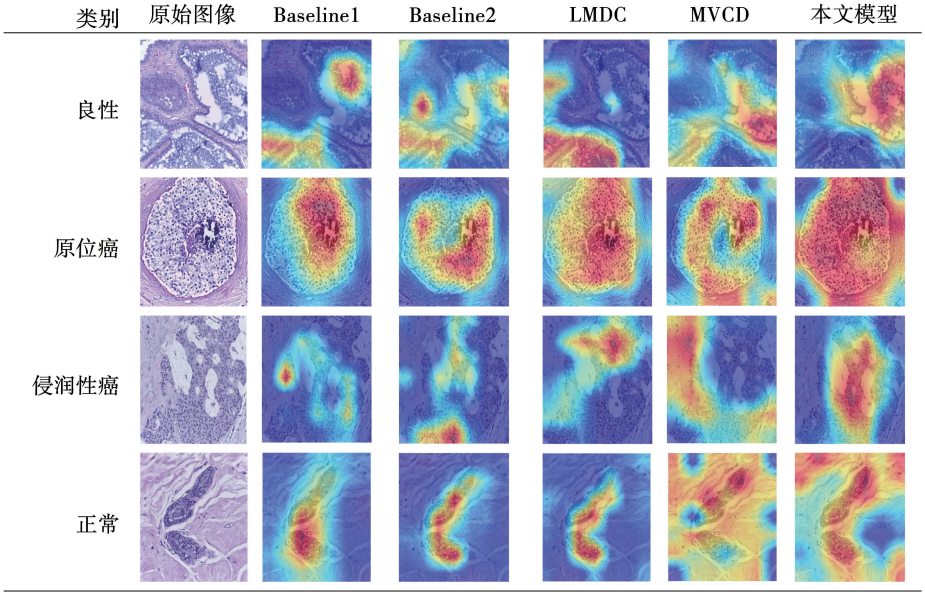


表9 Breakhis数据集评价指标对比

Tab. 9 Comparison of evaluation metrics of the Breakhis dataset

模型	α_{ACC}/%	$F_{1\text{-Score}}$/%	α_{BA}/%	α_{MCC}/%	参数量/M	算力/($10^9 s^{-1}$)
ResNet34[30]	91.4	90.9	91.3	90.1	23.37	4.03
SwinTransformer-Tiny	96.6	97.3	97.4	97.1	27.50	4.37
Res-SwT	97.5	97.7	98.0	97.2	54.31	8.23
MobileV3[31]	96.7	97.0	97.1	96.5	4.21	0.23
ShuffleNet	91.4	90.1	90.0	87.8	0.35	0.044
EfficientNet[32]	96.4	96.9	96.8	96.5	4.02	0.41
ConvNext[33]	96.5	96.3	97.5	95.7	27.80	4.45
ParC-Net[34]	92.2	92.0	93.2	90.1	21.69	3.98
EdgeViTs[35]	91.6	91.0	90.1	89.8	12.73	1.90
CvT[36]	92.1	92.5	92.3	91.2	19.55	4.06
ResNeXt Student(2023)[37]	96.4	95.0	95.3	94.4	13.46	—
DAFLNet(2023)[38]	95.5	95.2	95.9	94.5	—	—
Vit variant(2024)[39]	94.8	94.8	93.5	92.9	36.38	—
DRDA-Net(2024)[40]	97.5	98.1	—	—	—	—
GLNET(2024)[41]	91.5	91.3	87.4	—	58.7	—
本文模型	98.1	98.3	98.9	97.9	9.04	1.89

表10 BACH数据集评价指标对比

Tab. 10 Comparison of evaluation metrics of the BACH dataset

模型	α_{ACC}/%	$F_{1\text{-Score}}$/%	α_{BA}/%	α_{MCC}/%	参数量/M	算力/($10^9 s^{-1}$)
ResNet34[30]	94.8	94.7	94.6	93.1	23.37	4.03
SwinTransformer-Tiny	97.5	97.8	97.5	96.7	27.50	4.37
Res-SwT	98.1	98.2	98.1	97.5	54.31	8.23
MobileV3	96.7	96.7	96.8	95.6	4.21	0.23
ShuffleNet	88.8	88.7	88.7	85.2	0.35	0.044
EfficientNet	97.1	97.2	96.9	96.2	4.02	0.41
ConvNext	80.5	80.3	80.5	74.1	27.80	4.45
ParC-Net	98.5	98.4	98.7	98.0	21.69	3.98
EdgeViTs	97.0	97.1	97.3	96.2	12.73	1.90
CvT	98.2	98.3	98.8	97.7	19.55	4.06
GARL-Net(2023)[42]	97.5	97.6	—	95.1	—	—
SAVT(2024)[43]	90.0	89.9	90.0	86.6	—	—
SPCZP-CNN(2024)[44]	93.8	93.7	—	—	—	—
GLNET(2024)	90.3	87.7	86.5	—	58.7	—
本文模型	99.0	99.1	99.2	98.6	9.04	1.89

之后,相较于局部与全局特征串行融合的分类模型 EdgeViTs[35]、CvT[36]等,本文采用的特征融合方式能够自适应地选择不同尺度上的重要特征进行加权融合,并通过通道注意力机制进行特征间的交互,避免重要的细节或上下文信息丢失。最后,与近年研究相比,本文模型的 $F_{1\text{-Score}}$、α_{BA}、α_{MCC}指标良好,BreakHis 数据集上分别为 98.3%、98.9%、97.9%,BACH 数据集上分别为 99.1%、99.2%、98.6%,相较于其他现有模型有明显优势。在 2 个数据集中,本文模型与近年研究相比,在 $F_{1\text{-Score}}$、α_{BA}、α_{MCC}均有所提高,证明了本文模型可以缓解类别不平衡现象,从而进一步提升乳腺肿瘤分类性能。

3　结　语

本文针对乳腺组织图像中复杂的特征信息,提出了局部与全局特征融合的轻量化网络,实现乳腺肿瘤类别分类,提升计算机辅助诊断的临床应用价值。其中,LMDC 模块使用轻量化动态卷积模块综合了通道、空间维度中关键特征的重要性,有效地凸显了乳腺组织纹理细节,抑制了噪声干扰,减少了模型参数量;MVCD 模块使用自注意力机制建立不同区域的信息联系,并通过圆形循环卷积增大空间结构特征交流范围,从而增强了乳腺组织全局特征的依赖关系;特征融合模块 FFM 综合了细节纹理与空间上下文信息,有效克服了单一尺度特征分类中的局限性;动态加权损失函数依据不同肿瘤类别分类性能自适应调节各类别权重,提高对部分较难分类肿瘤的识别能力,有效缓解了乳腺肿瘤多类别分类不平衡问题。本文模型在 BreakHis 数据集达到了 98.1%的分类准确率,在 BACH 数据集达到了 99.0%的分类准确率,参数量仅为 9.04 M。未来将通过对肿瘤定位推进乳腺癌的治疗方案优化。

参 考 文 献:

[1] AZAMJAH N, SOLTAN Z Y, ZAYERI F. Global Trend of Breast Cancer Mortality Rate: A 25-year Study[J]. Asian Pacific Journal of Cancer Prevention, 2019, 20(7): 2015.

[2] 齐宏宇, 吕朝阳, 安广宇, 等. 不同年龄乳腺癌患者早期诊断的方法学比较[J]. 中国医学前沿杂志(电子版), 2019, 11(2): 63.
QI Hongyu, LV Zhaoyang, AN Guangyu, et al. Comparison of Methods for Early Diagnosis of Breast Cancer among Different Age Groups[J]. Chinese Journal of the Frontiers of Medical Science (Electronic Version), 2019, 11(2): 63.

[3] 姜金杰, 韩海林, 付焕杰, 等. 低剂量 DBT 和 FFDM 在早期乳腺癌筛查中的价值研究[J]. 国际放射医学核医学杂志, 2024, 48(5):303.
JIANG Jinjie, HAN Hailin, FU Huanjie, et al. Value Research of Low-Dose Digital Breast Tomosynthesis and Full-Field Digital Mammography in Screening Early Breast Cancer [J]. International Journal of Radiation Medicine and Nuclear Medicine, 2024, 48(5): 303.

[4] TOMA T A, BISWAS S, MIAH M S, et al. Breast Cancer Detection Based on Simplified Deep Learning Technique with Histopathological Image Using BreaKHis Database[J]. Radio Science, 2023, 58(11): 1.

[5] 许学斌, 张佳达, 刘伟, 等. 融合空间和通道特征的高精度乳腺癌分类方法[J]. 计算机应用, 2021, 41(10): 3025.
XU Xuebin, ZHANG Jiada, LIU Wei, et al. High-precision Classification Method for Breast Cancer Fusing Spatial Features and Channel Features[J]. Journal of Computer Applications, 2021, 41(10): 3025.

[6] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An Image Is Worth 16 × 16 Words: Transformers for Image Recognition at Scale[J]. arXiv preprint arXiv: 2010.11929, 2020.

[7] SHIRI M, SUN J. Supervised Contrastive Vision Transformer for Breast Histopathological Image Classification [J]. arXiv preprint arXiv: 2404.11052, 2024.

[8] CHEN X, ZHANG K, ABDOLI N, et al. Transformers Improve Breast Cancer Diagnosis from Unregistered Multi-view Mammograms [J]. Diagnostics, 2022, 12(7): 1549.

[9] ZHUANG J, WU X, MENG D, et al. A Swin Transformer and Residualnetwork Combined Model for Breast Cancer Disease Multi-classification Using Histopathological Images[J]. Instrumentation, 2024, 11(1): 112.

[10] ZOU Y, CHEN S, SUN Q, et al. DCET-Net: Dual-stream Convolution Expanded Transformer for Breast Cancer Histopathological Image Classification[C]//2021 IEEE International Conference on bioinformatics and biomedicine (BIBM), 2021: 1235.

[11] 李广丽, 叶艺源, 吴光庭, 等. 联合多视角 Transformer 编码与在线融合互学习的乳腺癌病理图像分类模型[J]. 电子学报, 2024, 52(7): 2369.
LI Guangli, YE Yiyuan, WU Guangting, et al. Breast Cancer Pathological Image Classification Model via

Combining Multi-view Transformer Coding and Online Fusion Mutual Learning[J]. Acta Electronica Sinica, 2024, 52(7): 2369.

[12] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 4510.

[13] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 6848.

[14] 方祯祺, 李雪, 莫红. 基于改进 EfficientNet 的乳腺肿瘤诊断[J]. 智能科学与技术学报, 2023, 5(4): 505. FANG Zhenqi, LI Xue, MO Hong. Diagnostic of Breast Tumors Based on Improved EfficientNet [J]. Chinese Journal of Intelligent Science and Technology, 2023, 5(4): 505.

[15] LEI T, SUN R, DU X, et al. SGU-Net: Shape-guided Ultralight Network for Abdominal Image Segmentation [J]. IEEE Journal of Biomedical and Health Informatics, 2023, 27(3): 1431.

[16] MA X Y, SUN L, GAO J P, et al. LMCNet: A Lightweight and Efficient Model for Multi-classification of Breast Cancer Images [J]. Signal, Image and Video Processing, 2025, 19(1): 175.

[17] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 9992.

[18] MEHTA S , RASTEGARI M . MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer [J]. arXiv preprint arXiv:2110.02178, 2021.

[19] LI C, ZHOU A J, YAO A B. Omni-dimensional Dynamic cConvolution [J]. arXiv preprint arXiv: 2209.07947, 2022.

[20] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [C]//Proceedings of the European Conference on Computer Vision (ECCV),2018: 3.

[21] STERGIOU A, POPPE R, KALLIATAKIS G. Refining Activation Downsampling with Softpool[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10357.

[22] HU J, SHEN L, SUN G. Squeeze-and-excitation Networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132.

[23] QIN Z, ZHANG P, WU F, et al. FCANet: Frequency Channel Attention Networks [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 763.

[24] KHAN S D, BASALAMAH S. Deep Learning Framework for Enhanced PCB Surface Defect Detection Leveraging Multi-scale Feature Aggregation and Contextual Information [J]. Multimedia Tools and Applications, 2026, 85(2): 167.

[25] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection [C]//2017 IEEE International Conference on Computer Vision (ICCV), 2017: 2999.

[26] PEREZ L, WANG J. The Effectiveness of Data Augmentation in Image Classification Using Deep Learning[J]. arXiv preprint arXiv:1712.04621, 2017.

[27] ALTHNIAN A, ALSAEED D, ALBAITY H, et al. Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical Domain[J]. Applied Sciences, 2021, 11(2): 796.

[28] JURMAN G, RICCADONNA S, FURLANELLO C. A Comparison of MCC and CEN Error Measures in Multi-class Prediction[J]. PLOS One, 2012, 7(8): e41882.

[29] SELVARAJU R 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 (ICCV), 2017: 618.

[30] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770.

[31] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 1314.

[32] TAN M, LE Q. Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks [C]//International Conference on Machine Learning, 2019: 6105.

[33] LIU Z , MAO H , WU C Y ,et al. A ConvNet for the 2020s[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11976.

[34] ZHANG H, HU W, WANG X. ParC-Net: Position Aware Circular Convolution with Merits from ConvNets and Transformer[C]//European Conference on Computer Vision, 2022: 613.

[35] PAN J, BULAT A, TAN F, et al. EdgeViTs: Competing Light-weight CNNS on Mobile Devices with Vision Transformers[C]//European Conference on Computer , 2022: 294.

[36] WU H, XIAO B, CODELLA N, et al. CVT: Introducing Convolutions to Vision Transformers [C]//2021 IEEE/

CVF International Conference on Computer Vision (ICCV), 2021: 22.

[37] SEPAHVAND M, ABDALI M F. Joint Learning Method with Teacher-Student Knowledge Distillation for On-device Breast Cancer Image Classification [J]. Computers in Biology and Medicine, 2023, 155: 106476.

[38] 赵晓平, 王荣发, 孙中波, 等. 改进 DenseNet 的乳腺癌病理图像八分类研究[J]. 计算机工程与应用, 2023, 59(5): 213.
ZHAO Xiaoping, WANG Rongfa, SUN Zhongbo, et al. Research on Eight Classifications of Breast Cancer Pathological Images Based on Improved DenseNet [J]. Computer Engineering and Applications, 2023, 59 (5): 213.

[39] ABIMOULOUD M L, BENSID K, ELLEUCH M, et al. Vision Transformer Based Convolutional Neural Network for Breast Cancer Histopathological Images Classification [J]. Multimedia Tools and Applications, 2024, 83 (39): 86833.

[40] AHMADI M, KARIMI N, SAMAVI S. A Lightweight Deep Learning Pipeline with DRDA-Net and MobileNet for Breast Cancer Classification [J]. arXiv preprint arXiv:2403.11135, 2024.

[41] KHAN S U R, ZHAO M, ASIF S, et al. GLNET: Global-Local CNN's-based Informed Model for Detection of Breast Cancer Categories from Histopathological Slides [J]. The Journal of Supercomputing, 2024, 80 (6): 7316.

[42] PATEL V, CHAURASIA V, MAHADEVA R, et al. GARL-Net: Graph Based Adaptive Regularized Learning Deep Network for Breast Cancer Classification[J]. IEEE Access, 2023, 11: 9095.

[43] BARONI G L, RASOTTO L, ROITERO K, et al. Optimizing Vision Transformers for Histopathology: Pretraining and Normalization in Breast Cancer Classification [J]. Journal of Imaging, 2024, 10 (5): 108.

[44] KUMAR A, SINGH C, SACHAN M K. A Moment-based Pooling Approach in Convolutional Neural Networks for Breast Cancer Histopathology Image Classification [J]. Neural Computing and Applications, 2025, 37 (2): 1127.

(编辑:温泽宇)