

DOI: 10.13382/j.jemi.B2407695

全局-局部特征融合的甲状腺细针穿刺活检全玻片 图像轻量化样本级分类*

高俊涛¹ 张菁^{1,2} 孙萌¹ 卓力^{1,2}

(1. 北京工业大学信息科学技术学院 北京 100124; 2. 计算智能与智能系统北京市重点实验室 北京 100124)

摘要: 细针穿刺活检全玻片图像 (FNAB-WSI) 的细胞学检查对甲状腺乳头状癌或良性结节性增生的诊断至关重要。由于样本级 FNAB-WSI 具有上亿像素的超高分辨率, 利用深度网络进行样本级别分类会消耗相当规模的计算资源。考虑到样本级 FNAB-WSI 兼具全局和细胞团局部细节特征, 提出了一种全局-局部特征融合的轻量化样本级分类方法。首先利用轻量化 GhostNet 网络提取全局特征, 通过设置卷积步长控制特征图谱尺寸, 并用特征切片与融合获取局部特征; 然后对全局和局部特征分别最大池化和降维, 进而融合为全局-局部特征; 最后全连接全局-局部特征, 并通过 softmax 分类器达成甲状腺样本级良恶性分类。在自建的 FNAB-WSI 样本级数据集上, 方法的各项性能指标上均超越了其他轻量化方法, 精度、召回率、准确率和 AUC 分别达到了最高的 89.9%、91.2%、91.7% 和 92.5%, 同时参数量方面具有可比性, 为 6.1×10^6 , 展现出了良好的平衡性能。方法不仅提高了样本级分类的准确性, 还通过减少参数量优化了模型的计算效率, 有望为甲状腺癌的临床诊断提供了一种有效的辅助工具。

关键词: 深度学习; 全玻片图像; 样本级分类; 轻量化; 全局-局部特征融合

中图分类号: TN911; TP391 文献标识码: A 国家标准学科分类代码: 510.40

Global-local feature fusion for lightweight sample-level classification of thyroid fine needle aspiration biopsy whole slide image

Gao Juntao¹ Zhang Jing^{1,2} Sun Meng¹ Zhuo Li^{1,2}

(1. School of Information Science and Technology, Beijing University of Technology, Beijing 100124, China; 2. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China)

Abstract: Cytologic examination of thyroid fine needle aspiration biopsy whole slide image (FNAB-WSI) is crucial for the diagnosis of papillary thyroid carcinoma or benign nodular hyperplasia. Due to the ultra-high resolution in sample-level FNAB-WSI, sample-level classification using deep networks consumes computational resources of considerable scale. Given that the sample-level FNAB-WSI has both global and cell cluster local detail features, a lightweight sample-level FNAB-WSI classification method with global-local feature fusion is proposed. Firstly, the global features are extracted using lightweight GhostNet, the feature map size is controlled by setting the convolutional stride, and the local features are obtained by feature slicing and fusion. Then, the global and local features are fused into global-local features after max-pooling and dimensionality reduction, respectively. Finally, the global-local features are fully connected to classify the benign-malignant FNAB-WSI by the softmax classifier. On the self-build FNAB-WSI sample-level dataset, our method surpasses other lightweight methods in all performance indicators, with 89.9% precision, 91.2% recall, 91.7% Acc, and 92.5% AUC, respectively, while the number of parameters is comparable to 6.1×10^6 , demonstrating a tradeoff result. The proposed method not only improves the accuracy of sample-level classification, but also optimizes the computational efficiency of the model by reducing the number of parameters, providing an effective auxiliary tool for clinical diagnosis of thyroid cancer.

Keywords: deep learning; whole slide image; sample-level classification; lightweight; global-local feature fusion

0 引言

甲状腺癌 (thyroid carcinoma) 是人体内分泌系统中最常见的肿瘤疾病之一。国际癌症研究机构的数据显示,全球新发甲状腺癌的数据量逐年递增^[1]。病理医生通常依据甲状腺细针穿刺活检 (fine needle aspiration biopsy, FNAB) 病理学全玻片图像 (whole slide image, WSI) 中细胞团的表现并结合样本总体特征,对甲状腺良性结节或恶性肿瘤进行综合判断^[2]。近年来,基于深度学习 (deep learning, DL) 的计算机视觉任务取得了空前的进展,也被广泛应用于 WSI 图像分析,在乳腺癌、前列腺癌、胃癌、神经母细胞瘤等^[3-6]领域发挥了良好的辅助作用。利用深度学习对甲状腺 FNAB-WSI 进行分类,有望为甲状腺癌的临床诊断提供有力支持,为患者的诊断提供精准信息。

病理医生在 FNAB-WSI 细胞学临床诊断时,通常从细胞团和样本级两级来完成 WSI 图像分类,其中样本级良恶性分类,是帮助医生进行甲状腺癌诊断最重要的任务之一。在开展本研究之前,分别提出了 FNAB-WSI 细胞团检测^[7]和多标签分类^[8]方法,取得了不错的结果,但是由于样本级甲状腺 FNAB-WSI 分辨率超高约有 60 000×60 000 pixels,且存在大量冗余信息,采用深度网络模型会消耗很高的计算资源,不利于临床疾病诊治,需要采用轻量化的分类主干网络,通过较低的计算量实现样本级分类。此外,样本级图像全局特征显著,同时图像局部细节中也包含细胞团的重要特征,往往良、恶样本中细胞团的特征区分更明显,因此,需要联合全局和局部特征提高样本级分类准确度。

近年来,深度网络轻量化主要采用轻量化模型压缩和轻量化网络结构设计两种思路。轻量化模型压缩指的是通过知识蒸馏、模型剪枝等压缩手段,降低模型的复杂度。例如, Li 等^[9]提出了一种残差引导的知识蒸馏网络,利用知识差距和残余,以教师和学生之间作为指导,训练一个更轻量化的学生网络;张雷等^[10]对批标准化层的缩放因子和卷积层卷积核参数同时添加约束,通过双准则通道剪枝减少误剪裁情况;Yeom 等^[11]提出了一种逐层相关性传播 (layer-wise relevance propagation, LRP) 的新型修剪框架,在模型压缩中加入了可解释性,减小模型大小的同时保证了准确率。但这些轻量化模型压缩策略大都依赖于一个高性能的成熟网络,对高分辨率的 WSI 图像不适用。

轻量化网络结构设计从模型设计时采用一些轻量化的思想,通过轻量卷积方式以减少卷积过程的计算量。例如, Yu 等^[12]提出了一种用于人体姿态估计的高效高分辨率网络 (lightweight high-resolution network,

LiteHRNet), 该网络引入了一个轻量级单元,即条件信道加权,以替换 ShuffleNet 中代价高昂的逐点 1×1 卷积,解决了大量使用 1×1 卷积导致的计算瓶颈问题,但是该方法更加适合于对位置敏感的计算机视觉任务,对分类效果提升不明显。程德强等^[13]引入坐标注意力,通过对特征进行降维处理来减少模型的参数量,使其重建方法具有高精度的同时也有可比的速度效果。Paoletti 等^[14]提出了一种基于轻量化 Ghost 模块的卷积神经网络 (convolutional neural network, CNN) 分类网络 GhostNet,在高光谱遥感图像分类任务上表现良好,其中 Ghost 模块基于一组内在特征图,应用一系列低成本线性变换生成更多特征图,实现了低参数量下的性能提升。近来有较多轻量化方法对该模块进行改进,如李忠科等^[15]结合 GhostConv 与 DWConv 改进了 YOLOv8 中的 C2f 模块;刘昶成等^[16]使用 GhostConv 改进了 SPPF 和 ASFF 模块并融入 YOLOv5 中,两者都有效减少了模型计算成本。总体而言, GhostNet 通过低成本的线性变换获得了更高的资源利用率,同时通过模块化设计提高了易用性,适于作为样本级分类的轻量化网络基础结构。

从整体来看,甲状腺样本级 FNAB-WSI 具有显著的全局特征,同时也包含许多小尺寸的细胞团目标,这些目标具有更精细的局部细节特征。全局特征提供了图像整体结构信息,而局部细节特征则对于精准识别和分类至关重要,因此处理过程中必须兼顾全局和局部特征。此外这些图像的超高分辨率,要求在处理这些图像时,需要考虑计算复杂度。文献[17]在处理此类图像时,用两种网络分别提取全局和局部特征,然而分别计算的复杂度高、费时费力。类似的,跨尺度的注意力交互可以联通不同尺度的信息,从而获得更准确的感兴趣区域结果,但是复杂跨尺度交互方式将带来额外的计算需求^[18]。为此,本文通过共享主干网络并结合后续不同步长的卷积支路来轻量地获取两种特征,并使用特征图拼接实现特征融合。

基于以上分析,本文提出了一种全局-局部特征融合的甲状腺 FNAB-WSI 轻量化样本级分类模型。设计思路如下:首先将样本级甲状腺 FNAB-WSI 送入轻量化的 GhostNet 网络提取全局特征,并通过高层卷积中的不同步长设置控制特征图的尺寸,挖掘局部特征;然后对全局和局部特征分别进行最大池化和降维,以提取有效信息进一步融合为全局-局部特征;最后,将获得的全局-局部特征图进行全连接,利用 softmax 分类器实现甲状腺 FNAB-WSI 样本级良恶性分类。

本文引入轻量化网络 GhostNet 主干网络,并通过不同的步长设置来有效获取全局和局部两种不同的特征,在保证高精度的同时减小了模型的参数量。通过融合全局和局部两种特征,使用简单的 softmax 分类器实现样本

的有效分类,并结合二元交叉熵损失和三元组损失进行模型优化。为了推动样本级 FNAB-WSI 的广泛研究,本文自建了一个甲状腺 FNAB-WSI 样本级图像数据集,可以供广大科研工作者使用。

1 方法

1.1 总体框架

提出的全局-局部特征融合的甲状腺 FNAB-WSI 轻量化样本级分类模型,总体架构如图 1 所示,包括如下 3 个部分: GhostNet 全局-局部特征提取,将甲状腺 FNAB-WSI 图像送入 GhostNet 网络,通过前 4 个层的共享卷积结构提取样本级图像的浅层特征作为图像特征,再使用步长相异的卷积层,分别处理图像特征获取全局、局部特征;全局-局部特征融合,对全局和局部特征分别进行最大池化和 1×1 卷积降维,经过特征融合获得全局-局部特征。

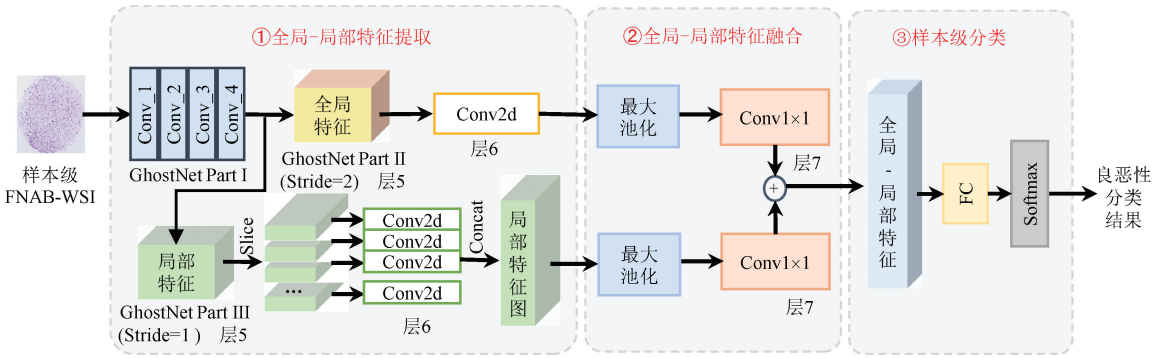


图 1 甲状腺 FNAB-WSI 图像样本级分类总体架构

Fig. 1 Overall framework for sample-level classification of thyroid FNAB-WSIs

1) 全局特征提取

全局特征提取中的 G-bneck 是最关键的结构,其核心组成模块是轻量化的 Ghost 模块。GhostNet 网络的层次结构和 G-bneck 的基本构成如图 2 所示,第 1 层初始层使用较大的卷积核(如 3×3),主要用于捕获基本的图像特征,该层的输出将被用作后续层的基础特征;第 2~4 中间层逐渐增加卷积核的数量,同时减少卷积核的尺寸(通常到 1×1),专注于提取更细致的特征,在这些层次的卷积结构中大部分采用步长为 2 的卷积,以快速减小特征图的尺寸,从而强制网络捕捉更广泛的图像内容;GhostNet 第 5 个卷积层的特征图作为全局特征。

Ghost 模块是由 G-bneck 核心组成的高效且轻量的架构。模块处理步骤如下。

步骤 1) 将上一阶段输入尺寸 $H_0 \times W_0$, 通道数 C_0 的输出特征图 $X \in R^{H_0 \times W_0 \times C_0}$ 作为输入,送入 Ghost 模块,假设 $f_{1 \times 1}$ 表示卷积滤波器,为卷积核为 1×1 的点卷积,则输

特征维度,减少计算复杂度,之后对降维后的特征进行融合;甲状腺 FNAB-WSI 样本级分类,将融合后的全局和局部特征图进行全连接后,利用 softmax 激活函数获得每一测试样本属于某一类别的概率值,实现样本级的良恶性分类。

1.2 甲状腺 FNAB-WSI 样本级特征提取

本节采用轻量级的 GhostNet 网络为主干提取甲状腺 FNAB-WSI 的样本级特征。该模型采用了类似与 ResNet 网络的残差连接结构,共包括 7 个层次模块的卷积,实现从浅到深的特征提取,每层卷积模块由轻量化 Ghost 模块构成的 G-bneck 结构。如图 1 所示,前 4 层卷积用于提取全局特征,第 5 层卷积设置步长为 2 提取全局特征,设置步长为 1 保证特征图尺寸不变,经过特征分块和融合提取局部特征,之后将全局和局部特征分别进行最大池化和 1×1 卷积降维,经过特征融合获得全局-局部特征。

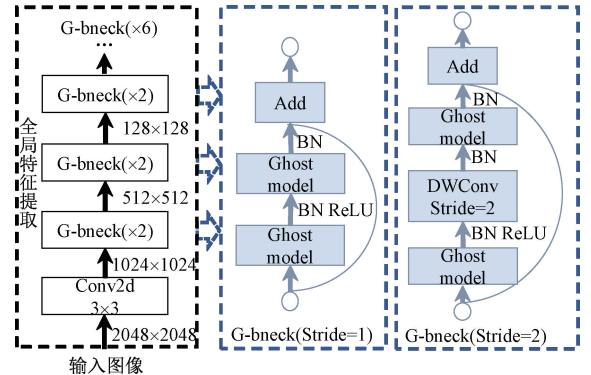


图 2 基于 G-bneck 的全局特征提取

Fig. 2 The global feature extraction by G-bneck

出 $Y \in R^{H_1 \times W_1 \times C_1}$ 为:

$$Y = X \times f_{1 \times 1} \quad (1)$$

这一步骤使用的卷积核数量远少于传统的卷积方

法,大大减少了计算过程的 FLOPs,输出的特征图 $\mathbf{Y}$ 将作为映射扩展的特征图基。

步骤 2) 对特征图 $\mathbf{Y} \in R^{H_1 \times W_1 \times C_1}$ 进行线性映射操作,通过一系列轻量的线性映射操作扩展这组基础特征图,生成更多的特征图,生成的特征图集合为 $\mathbf{X}'$ 。线性映射过程可以表示为 $\varphi(\cdot)$,为卷积或线性变换。将特征图 $\mathbf{Y}$ 按照通道划分,第 i 个内在特征可以表示为 $\mathbf{Y}_i'$;假设单个特征映射次数为 s , $\mathbf{X}_{ij}'$ 表示第 i 个内在特征的第 j 次映射获得的特征,映射过程可表示为:

$$\mathbf{X}_{ij}' = \varphi_{i,j}(\mathbf{Y}_i') \quad (2)$$

步骤 3) 将生成的特征图集合 $\mathbf{X}'$ 和上一步骤卷积的输出 $\mathbf{Y}$,通过 $\text{Concat}(\cdot)$ 函数沿通道维度连接,获得 Ghost 模块最终的输出特征图 $\mathbf{X}'' \in R^{H_0 \times W_0 \times C_2}$:

$$\mathbf{X}'' = \text{Concat}(\mathbf{X}', \mathbf{Y}) \quad (3)$$

相比于传统卷积,Ghost 模块在内存占用和计算量上均更小。假设 $k \times k$ 表示传统卷积和 Ghost 模块中的步骤一中的卷积核大小; $d \times d$ 表示线性映射过程中的平均核大小; r_s 为传统卷积模块和 Ghost 模块的计算量比; r_c 表示传统卷积和 Ghost 模块的参数量比,其计算过程可以表示为:

$$r_s = \frac{C_2 \cdot H \cdot W \cdot C_0 \cdot k \cdot k}{\frac{C_2}{s} \cdot H \cdot W \cdot C_0 \cdot k \cdot k + (s-1) \cdot \frac{C_2}{s} \cdot H \cdot W \cdot d \cdot d} = \frac{C_0 \cdot k \cdot k}{\frac{1}{s} \cdot C_0 \cdot k \cdot k + \frac{s-1}{s} \cdot d \cdot d} \approx \frac{s \cdot C_0}{s + C_0 - 1} \approx s \quad (4)$$

$$r_c = \frac{C_2 \cdot C_0 \cdot k \cdot k}{\frac{C_2}{s} \cdot C_0 \cdot k \cdot k + (s-1) \cdot \frac{C_2}{s} \cdot d \cdot d} \approx \frac{s \cdot C_0}{s + C_0 - 1} \approx s \quad (5)$$

由于计算量和参数量的比例关系,两比值的结果都为 s ,表征单个内在特征的映射次数,同时 $s < C_0$ 。可见,使用 Ghost 模块达到的参数量和计算量相比于传统卷积可以减少至原来的 $1/s$ 。

2) 局部特征提取

在采用上述的卷积结构提取甲状腺 FNAB-WSI 图像样本级全局图像特征之后,接下来在第 5 层通过设置不同的步长获取局部特征,其基本操作步骤如下。

步骤 1) 将上一小节中全局特征提取中第 5 层卷积结构获得的特征图表示为 $\mathbf{F}$,在局部特征提取阶段,设置步长(stride)为 1 的卷积 $\text{Conv}_{s=1}(\cdot)$ 保持特征图尺寸不变,以便捕获更细致的局部特征 $\mathbf{F}_l$:

$$\mathbf{F}_l = \text{Conv}_{s=1}(\mathbf{F}) \quad (6)$$

步骤 2) 将步长为 1 的局部不变特征图进行切片(Slice)处理,获取更具体的局部特征信息,切片后的第 i 个特征表示 $\mathbf{F}_{l,i}$ 。对切片后的特征图分别进行 2D 卷积 $\text{Conv}_{2D}(\cdot)$,每个切块的 2D 卷积共享相同的卷积和偏置项。将卷积之后的特征图进行 $\text{Concat}(\cdot)$ 融合为:

$$\mathbf{F}_l' = \text{Concat}(\mathbf{F}_{l,1}', \mathbf{F}_{l,2}', \dots, \mathbf{F}_{l,n}') = \text{Concat}\left(\text{Conv}_{2D}(\mathbf{F}_{l,1}), \text{Conv}_{2D}(\mathbf{F}_{l,2}), \dots, \text{Conv}_{2D}(\mathbf{F}_{l,n})\right) \quad (7)$$

3) 全局-局部特征融合

前两小节实现了全局特征和局部特征提取,接下来将全局和局部特征进行融合。为了提升计算效率,对全局特征 $\mathbf{F}_g$ 和局部特征 $\mathbf{F}_l'$ 分别进行最大池化 $\text{Maxpooling}(\cdot)$ 和 1×1 大小的卷积 $\text{Conv}_{1 \times 1}(\cdot)$ 降维,最后融合全局特征 $\mathbf{F}_g'$ 和局部特征 $\mathbf{F}_l''$,融合后的全局-局部特征 $\mathbf{F}'$ 被用来送入分类模块实现样本级分类。

$$\mathbf{F}_g', \mathbf{F}_l'' = \text{Conv}_{1 \times 1}(\text{Maxpooling}(\mathbf{F}_g, \mathbf{F}_l')) \quad (8)$$

$$\mathbf{F}' = \text{Concat}(\mathbf{F}_g', \mathbf{F}_l'') \quad (9)$$

1.3 甲状腺 FNAB-WSI 样本级分类

基于轻量化 GhostNet 网络获得了全局和局部特征,接下来的分类模块将融合后的全局-局部特征输入全连接层,然后通过分类器实现样本级分类。在分类器的选择上,选择 softmax 分类器实现良恶性样本级分类。通过计算当前测试样本属于良性还是恶性类别的概率,选择概率高的类别作为当前测试样本的类别。对于每个类别 i ,softmax 函数定义为:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (10)$$

式中: z_i 是模型输出层的第 i 个输入;分母为所有输出类别的指数值的总和。

对于样本分类,存在良性和恶性两种情况,假设分别对应 z_1 和 z_2 ,则 softmax 计算得到的类别输出概率 P_1 、 P_2 为:

$$\left\{ P_1 = \frac{e^{z_1}}{e^{z_1} + e^{z_2}}, P_2 = \frac{e^{z_2}}{e^{z_1} + e^{z_2}} \right\} \quad (11)$$

为了实现甲状腺 FNAB-WSI 图像样本级良恶性分类,本文采用二元交叉熵损失 $\Lambda_{cse}^{[19]}$ 和三元组损失 $\Lambda_{tri}^{[20]}$ 用于模型优化,最终损失函数可以表示为:

$$L = L_{cse} + L_{tri} \quad (12)$$

二元交叉熵损失 Λ_{cse} 用来衡量真样本标签和预测概率之间的差值,假设用 y 表示真实样本标签,正样本的 y 为 1,负样本的 y 为 0,用 p 表示预测概率,则二元交叉熵损失表示为:

$$L_{cse} = -[y \log(p) + (1-y) \log(1-p)] \quad (13)$$

三元组损失 Λ_{tri} 可以增大类间距离,减小类内距离,假设用 i 和 j 为同一类别的两个样本; k 为另一类的样

本; $D(\cdot)$ 表示欧氏距离计算; $F(\cdot)$ 表示特征提取等操作; m 表示阈值;则 $D(F(i), F(j))$ 表示同一类别特征的相似度; $D(F(i), F(k))$ 表示不同类别样本特征的相似度,则三元组损失表示为:

$$L_{tri} = [D(F(i), F(k)) - D(F(i), F(j)) + m]_+$$

(14)

式中: $[\cdot]_+$ 表示取值最小为0。

2 实验结果与分析

2.1 数据集

本文研究的甲状腺 FNAB-WSI 尚没有公开数据集,为此自建了一套甲状腺 FNAB-WSI 样本级图像数据集。该数据集收集并整理了广西科技大学第二附属医院病理科 2022 年 1 月~2022 年 12 月的甲状腺 FNAB 细胞学涂片(Thinprep2000, H&E 染色),并通过甲状腺病理科医生标注。数据集中的样本标签按医生诊断结果标注为良性或恶性。如表 1 所示,最初样本共 40 例,剔除低质量样本后,良性样本 19 例,恶性样本 21 例。为了较多保留实验的训练数据,后续实验中将数据集按照 8 : 2 比例划分为训练集和测试集,分别为 32 和 8 例。因数据集样本数量少,本文通过剪裁拼接、翻转镜像等方式扩充数据集,以提升模型的分类性能,同时避免模型过拟合。扩充后良性样本 224 例,恶性 228 例。

表 1 甲状腺 FNAB-WSI 数据集样本分布情况

Table 1 Distribution of various samples of thyroid FNAB-WSI dataset

	扩充前		扩充后	
样本类别	良性	恶性	良性	恶性
样本数量	19	21	224	228

2.2 评价指标

样本级分类的性能评价采用通用图像分类任务中的评价指标,分别为精确率(precision)、召回率(recall)、准确率(accuracy, Acc)、AUC(area under the curve),定义如式(15)~(17)所示。

$$Precision = \frac{TP}{TP + FP}$$

(15)

$$Recall = \frac{TP}{TP + FN}$$

(16)

$$Acc = \frac{TP + TN}{TP + TN + FN + FP}$$

(17)

式中: TP (true positive)为真阳性,表示正确检测到的阳性样本; TN (true negative)为真阴性,表示正确检测到的阴性样本; FP (false positive)为假阳性,即错误检测的阳

性样本; FN (false negative)为假阴性,即错误检测的阴性样本。鉴于 ROC 曲线显示了不同阈值下真正例率和假正例率的关系,AUC 值则表示 ROC 曲线下的面积,本文将用 AUC 综合评估模型性能。

此外,本文使用参数量(Params)、浮点运算数(FLOPs)和推理速度来评估模型的复杂度,前两者的计算可参考式(18)和(19),在实验中通过对应的 python 库引用计算统计。

$$Params_{conv} = (kernel \times kernel) \times C_{in} \times C_{out}$$

(18)

$$FLOPs_{conv} = (kernel \times kernel \times map \times map) \times C_{in} \times C_{out}$$

(19)

式中: $kernel$ 是卷积核的尺寸; C_{in} 和 C_{out} 分别是输入和输出特征的通道数; map 是输出特征图的尺寸。

2.3 实验设置

实验采用 Ubuntu20.04 (Linux) 操作系统,NVIDIA 2080ti GPU 和 Pytorch 深度学习框架。实验过程中涉及的超参数如下,优化器采用 SGD 优化算法,权重衰减设置为 0.000 1。初始学习率设置为 1×10^{-4} 。主干网络的初始权重来源于 ImageNet 数据集上的预训练权重,训练轮次(Epochs)设置为 200 次。共设计了 4 组实验,包括与其他主流的常规和轻量化分类方法的对比实验、模型参数量和复杂度对比、主观实验结果展示、消融研究,以证明提出的方法的有效性。

2.4 实验结果

1)与主流方法的比较

本实验将提出甲状腺 FNAB-WSI 图像轻量化样本级分类网络,和目前主流的常规分类网络、轻量化分类网络进行了对比,实验结果如表 2 所示。其中,常规分类网络选用了深度较浅的 ResNet50^[21]、ResNeXt50^[22]、DenseNet^[23]和 ConvNeXt^[24]作为对比方法,轻量化网络选用了 ShuffleNet^[25]、EfficientNet(B3)^[26]、MobileNetV2^[27]、MobileNetV3^[28]、Poolformer(S12)^[29]、FasterNet(T2)^[30]和 ParameterNet^[31]作为对比方法,其中括号中编号代表轻量化型号。从表 2 可以看出,常规卷积中 ConvNeXt 综合表现最佳, Precision、Recall、Acc 和 AUC 分别达到了 91.7%、92.3%、92.5%和 92.6%,每张图像测试时间为 1.869 s;在轻量化网络中,所提方法结果最高, Precision、Recall、Acc 和 AUC 分别达到了 89.9%、91.2%、91.7%和 92.5%,在检测速度上每张图像测试时间为 0.592 s。在上述 4 个精度指标上,所提方法相比 ConvNeXt, Precision、Recall、Acc 和 AUC 分别变化值为 -1.8%、-1.1%、-0.8%和-0.1%,在微弱的精度差异下,本文方法在测试时间上明显减少了 1.277 s,进一步地,相比速度最快的轻量化网络 MobileNetV3 时间逊色了 0.232 s,但是各精度指标都有相应提升(Precision、Recall、Acc 和

AUC 分别提升了 6.7%、10.9%、7.1% 和 4.8%), 这些综合结果证明了本文提出的基于全局和局部特征融合改进

GhostNet 模型在甲状腺 FNAB-WSI 样本级分类上具有更均衡的优势。

表 2 与其他常规和轻量化分类网络的对比

Table 2 Comparison with other conventional and lightweight classification methods						(%)
方法		Precision	Recall	Acc	AUC	测试时间/s
常规网络	ResNet50 ^[21]	85.4	88.3	88.9	89.9	1.679
	ResNeXt50 ^[22]	89.9	90.1	89.3	91.0	1.998
	DenseNet ^[23]	77.8	81.0	79.4	80.5	1.905
	ConvNeXt-T ^[24]	91.7	92.3	92.5	92.6	1.869
轻量化网络	ShuffleNetV2 ^[25]	76.3	77.1	76.9	78.4	0.451
	EfficientNet(B3) ^[26]	83.7	86.4	85.6	86.3	2.897
	MobileNetV2 ^[27]	76.7	76.6	78.1	79.2	0.499
	MobileNetV3 ^[28]	83.2	80.3	84.6	87.7	0.360
	Poolformer(S12) ^[29]	81.6	83.7	82.4	83.3	0.828
	FasterNet(T2) ^[30]	85.3	89.2	88.3	89.6	0.784
	ParameterNet ^[31]	85.9	90.7	88.6	91.3	0.846
	本文	89.9	91.2	91.7	92.5	0.592

2) 模型参数量和计算量对比

本实验对本文方法和其他轻量化分类网络的参数量和计算量进行了复杂度评估,如表 3 所示。可以看到,本文提出模型的参数量和浮点数为 6.1 MFLOPs 和 2.0 GFLOPs,在系列轻量化网络中表现中等。相比较最好的 ShuffleNetV2 模型在参数量和浮点数分别增加了 3.8 MFLOPs 和 1.79 GFLOPs,这主要是因为本模型除了提取图像的全局信息用于分类之外,还对局部特征信息进行了融合,这带来了一定的参数量和浮点数增益。但是总的来讲,本模型在准确率上具有显著优势,因此牺牲一定的计算效率是合理的。

表 3 模型参数量和计算量对比结果

Table 3 Number of parameters versus FLOPs		
模型	Params×10 ⁶	浮点数/GFLOPs
ShuffleNetV2 ^[25]	2.3	0.21
EfficientNet(B3) ^[26]	12	1.8
MobileNetV2 ^[27]	3.0	0.43
MobileNetV3 ^[28]	5.4	0.31
Poolformer(S12) ^[29]	11.9	1.82
FasterNet(T2) ^[30]	15.0	1.91
ParameterNet ^[31]	15.7	0.60
本文	6.1	2.0

3) 分类结果的可视化展示

为了直观展示提出模型的良恶性分类能力,本实验对测试集中的甲状腺 FNAB-WSI 图像样本进行了可视化,并展示了部分分类结果。图 3 从测试结果中选择了其中 8 组图像,图题为甲状腺 FNAB-WSI 的编号,左侧为

全玻片图像的整体样本图,右侧为样本局部细节放大的细胞团图,图像下方文字表示对应样本的良恶性真实标签和模型预测结果。

从实验结果可以看出,该模型在甲状腺 FNAB-WSI 图像的样本级分类数据集上具有良好的性能,模型大部分情况可以分别出甲状腺 FNAB-WSI 的类型。图 3(a)、(b)、(c)、(f) 表明了模型对于细胞团较多的甲状腺 FNAB-WSI 具有较好的分类能力;图 3(d)、(e)、(h) 则展示出模型对于细节信息较少的图像的不能完全正确分类,这是由于这几张图像从全局信息来看比较类似,局部细胞团的信息较少,因此导致了部分图像分类错误;另外对于杂质较多的图 3(g),模型能够正确分类,证明了本文模型对于杂质比较多的图像也能正确分类。综上所述,本文提出的全局-局部特征融合的甲状腺 FNAB-WSI 图像轻量化样本级分类模型总体上具有良好的效果。

4) 消融研究

实验通过消融研究评估方法中不同模块对甲状腺 FNAB-WSI 样本级良恶性分类性能的影响。实验比较了 ResNet、GhostNet、Ghost-ResNet 和本文方法的结果,实验结果如表 4 所示。其中, GhostNet、Ghost-ResNet 对比 ResNet 精度普遍降低,这是因为引入轻量化的 Ghost 模块来提升计算效率的方式带来了一定的准确率的损失。本文提出的方法通过融合局部特征, Precision、Recall、Acc 和 AUC 分别达到了 89.9%、91.2%、91.7% 和 92.5%,相比于 ResNet50 Precision、Recall、Acc 和 AUC 分别提升了 4.5%、2.9%、2.8% 和 2.6%,表明本文提出的全局和局部特征融合的轻量化分类模型在甲状腺 FNAB-WSI 数据集上带来了稳定的精度。

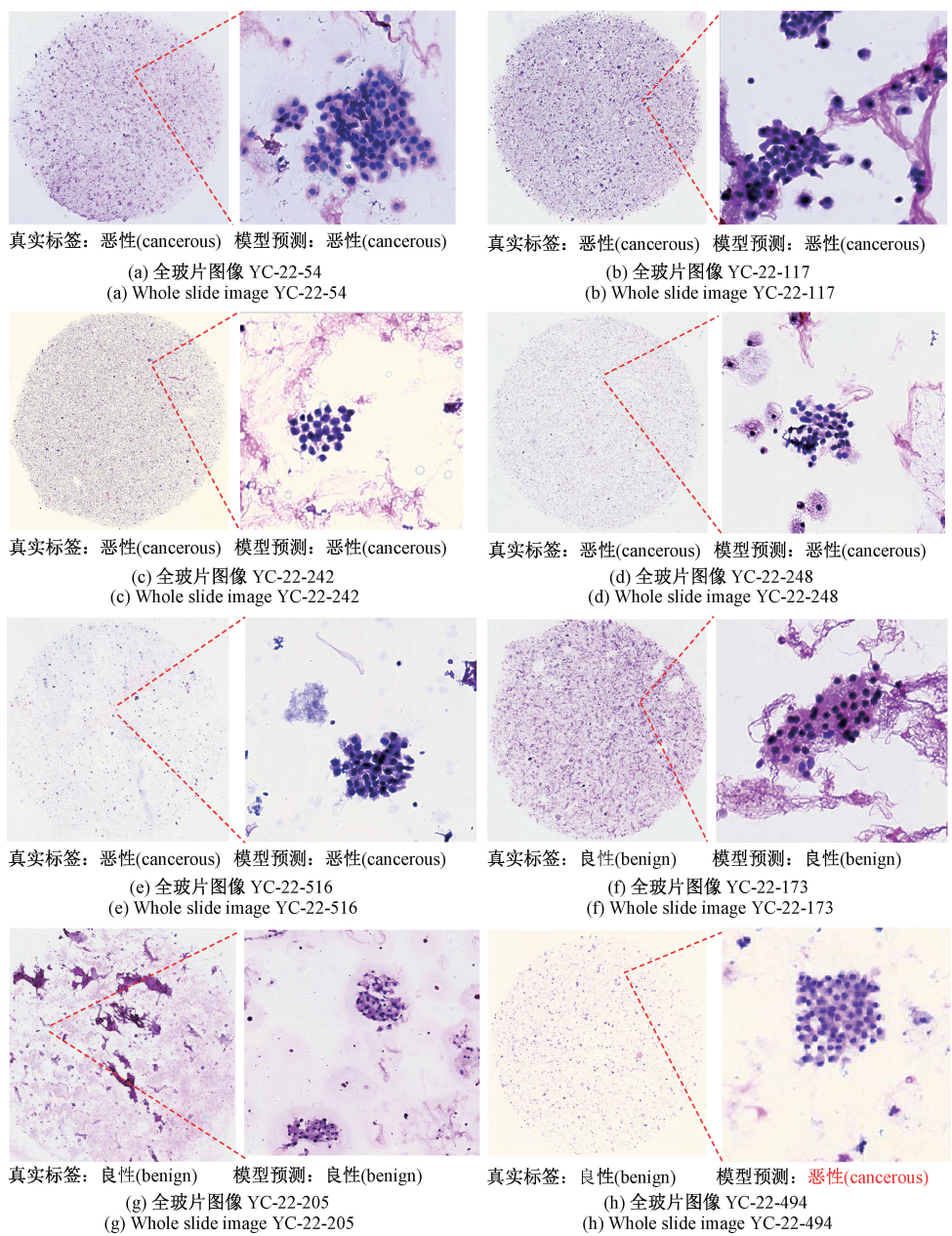


图 3 部分样本的可视化分类结果

Fig. 3 Visualization of classification results for selected samples

表 4 消融研究

Table 4 Ablation study (%)

模型	全局-局部	轻量化	Precision	Recall	Acc	AUC
	特征	Ghost				
ResNet	×	×	0.854	0.883	0.889	0.899
GhostNet	×	✓	0.826	0.796	0.831	0.832
Ghost + ResNet	×	✓	0.843	0.833	0.849	0.853
All	✓	✓	0.899	0.912	0.917	0.925

3 结 论

针对超高分辨率的样本级甲状腺 FNAB-WSI 图像分类模型会消耗大量的计算资源,和甲状腺 FNAB-WSI 具有显著的全局特征和细节信息的局部特征,提出了一种全局-局部特征融合的甲状腺 FNAB-WSI 轻量化样本级分类模型。首先,将样本级甲状腺 FNAB-WSI 送入轻量

化的 GhostNet 网络,网络前 4 个层用于提取全局特征,在第 5 层中设置卷积步长为 2 实现下采样控制全局特征图的尺寸,将分路径的卷积步长设为 1 控制特征图尺寸不变,并经过特征分块和卷积融合,提取局部特征;然后,将全局和局部特征分别进行最大池化和降维,融合为全局-局部特征;最后将融合后的特征经全连接送入 softmax 分类器,实现 FNAB-WSI 样本级分类。通过与其他主流的常规和轻量化分类网络对比、主观结果展示、模型参数量和复杂度对比、消融研究共 4 组实验,综合结果表明,所提方法各项指标上相比其他轻量化方法, Precision、Recall、Acc 和 AUC 分别达到了最高的 89.9%、91.2%、91.7%和 92.5%。研究结果表明本文的全局-局部特征融合的甲状腺 FNAB-WSI 轻量化样本级分类模型达到了良好的性能。本文通过数据增强扩充样本并完成分类,在样本量上仍有局限。在医学图像处理中,迁移学习和元学习可以减少对大量标注数据的需求,使模型快速适应新任务。未来研究将探索这些方法以提升样本级分类性能。

参考文献

- [1] CHHIKARA B S, PARANG K. Global cancer statistics 2022: The trends projection analysis [J]. Chemical Biology Letters, 2023, 10(1): 451-451.
- [2] RASTEGAR H, TORSHIZIAN A, YAGHOUBI M A, et al. Accuracy of fine-needle aspiration cytopathology to differentiate malignant and benign thyroid nodules with ≥ 4 cm diameter: A retrospective study [J]. Diagnostic Cytopathology, 2023, 51(4): 263-269.
- [3] ZHENG Y, LI J, SHI J, et al. Kernel attention Transformer for histopathology whole slide image analysis and assistant cancer diagnosis [J]. IEEE Transactions on Medical Imaging, 2023, 42(9): 2726-2739.
- [4] LI B, LI Y, ELICEIRI K W. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning [C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2021.
- [5] KHALILIBOROUjeni S, HE X, JIA W, et al. End-to-end metastasis detection of breast cancer from histopathology whole slide images [J]. Computerized Medical Imaging and Graphics, 2022, 102: 102-136.
- [6] RAKHEJA G, HANDA U, PUNIA R S, et al. Fine-needle aspiration cytology in soft tissue tumors: 5-year institutional experience [J]. Diagnostic Cytopathology, 2022, 50(10): 463-470.
- [7] SUN M, ZHANG J, ZHAO S, et al. Cell cluster detection of thyroid FNAB-WSI via deformable convolution with frequency channel attention [J]. Biomedical Signal Processing and Control, 2025, 100: 107023.
- [8] GAO J T, ZHANG J, SUN M, et al. Cellular spatial-semantic embedding for multi-label classification of cell clusters in thyroid fine needle aspiration biopsy whole slide images [J]. Pattern Recognition Letters, 2025, 188: 125-132.
- [9] LI X W, LI S Y, OMAR B, et al. ResKD: Residual-guided knowledge distillation [J]. IEEE Transactions on Image Processing, 2021, 30: 4735-4746.
- [10] 张雷, 童虎庆, 谢锦昌, 等. 基于 ARM 平台目标检测的轻量化方法 [J]. 电子测量技术, 2023, 46(12): 118-124.
ZHANG L, TONG H Q, XIE J CH, et al. Lightweight method of object detection based on ARM platform [J]. Electronic Measurement Technology, 2023, 46(12): 118-124.
- [11] YEOM S K, SEEGERER P, LAPUSCHKIN S, et al. Pruning by explaining: A novel criterion for deep neural network pruning [J]. Pattern Recognition, 2021, 115: 107899.
- [12] YU CH Q, XIAO B, GAO CH X, et al. Lite-hrnet: A lightweight high-resolution network [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2021.
- [13] 程德强, 陈杰, 寇旗旗, 等. 融合层次特征和注意力机制的轻量化矿井图像超分辨率重建方法 [J]. 仪器仪表学报, 2022, 43(8): 73-84.
CHENG D Q, CHEN J, KOU Q Q, et al. Lightweight super-resolution reconstruction method based on hierarchical features fusion and attention mechanism for mine image [J]. Chinese Journal of Scientific Instrument, 2022, 43(8): 73-84.
- [14] PAOLETTI M E, HAUT J M, PEREIRA N S, et al. GhostNet for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(12): 10378-10393.
- [15] 李忠科, 刘小芳. 基于轻量级 YOLOv8n 网络的 PCB 缺陷检测算法 [J]. 电子测量技术, 2024, 47(4): 120-126.
LI ZH K, LIU X F. PCB defect detection algorithm based on lightweight YOLOv8n network [J]. Electronic

- Measurement Technology, 2024, 47(4): 120-126.
- [16] 刘昶成, 邵文权, 李玲陶. 基于改进 YOLOv5s 的电力作业人员安全帽检测算法研究[J]. 国外电子测量技术, 2024, 43(2): 34-42.
- LIU CH CH, SHAO W Q, LI L T. Research on safety helmet detection algorithm for power operators based on improved YOLOv5s[J]. Foreign Electronic Measurement Technology, 2024, 43(2): 34-42.
- [17] ZHU Q, ZHONG Y, LIU Y, et al. A deep-local-global feature fusion framework for high spatial resolution imagery scene classification[J]. Remote Sensing, 2018, 10(4): 568.
- [18] 罗亦铭, 王建林, 田艳, 等. 基于注意力机制的甲状腺超声图像感兴趣区域定位方法[J]. 电子测量与仪器学报, 2023, 37(6): 39-47.
- LUO Y M, WANG J L, TIAN Y, et al. Location of regions of interest in thyroid ultrasound images based on attention mechanism [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(6): 39-47.
- [19] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.
- [20] FLORAIN S, DMITRY K, JAMES P. FaceNet: A unified embedding for face recognition and clustering [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 815-823.
- [21] HE K M, ZHANG X, REN S Q, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [22] XIE S, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1492-1500.
- [23] HUANG G, LIU ZH, VAN DER MAATEN L, et al. Densely connected convolutional networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [24] LIU ZH, MAO H Z, WU C Y, et al. A convnet for the 2020s [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2022.
- [25] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 11976-11986.
- [26] TAN M M, LE Q. Efficientnetv2: Smaller models and faster training [C]. Proceedings of the International Conference on Machine Learning, 2021: 10096-10106.
- [27] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetv2: Inverted residuals and linear bottlenecks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [28] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. Proceedings of the IEEE International Conference on Computer Vision, 2019: 1314-1324.
- [29] YU W H, SI CH Y, ZHOU P, et al. MetaFormer baselines for vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(2): 896-912.
- [30] CHEN J R, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [31] HAN K, WANG Y H, GUO J Y, et al. ParameterNet: Parameters are all you need for large-scale visual pretraining of mobile networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2024.

作者简介



高俊涛, 2023 年于北京工业大学获得学士学位, 现为北京工业大学硕士研究生, 主要研究方向为目标检测、深度学习。

E-mail: gaojt_2@emails.bjtu.edu.cn

Gao Juntao received his B. Sc. degree from Beijing University of Technology in 2023. Now he is a M. Sc. candidate at Beijing University of Technology. His main research interests include object detection and deep learning.



张菁(通信作者), 2008 年于北京工业大学获得博士学位, 现为北京工业大学教授、博士生导师, 主要研究方向为图像/视频处理、图像识别、计算机视觉和深度学习。

E-mail: zhj@bjtu.edu.cn

Zhang Jing (Corresponding author) received her Ph. D. degree from Beijing University of Technology in 2008. Now she is a professor and doctoral supervisor in Beijing University of

Technology. Her main research interests include image/video processing, image recognition, computer vision and deep learning.



孙萌, 2021 年于北京工业大学获得学士学位, 2024 年于北京工业大学获得硕士学位, 主要研究方向为图像识别、深度学习。
E-mail: sunmeng@emails.bjut.edu.cn

Sun Meng received her B. Sc. degree from Beijing University of Technology in 2021, and M. Sc. degree from Beijing University of Technology in 2024, respectively. Her main research interests include image recognition and deep learning.



卓力, 1992 年于电子科技大学获得学士学位, 1998 年于东南大学获得硕士学位, 2004 年于北京工业大学获得博士学位, 现为北京工业大学教授、博士生导师, 主要研究方向为图像/视频处理、图像识别和医学图像处理。
E-mail: zhuoli@bjut.edu.cn

Zhuo Li received her B. Sc. degree from University of Electronic Science and Technology of China in 1992, M. Sc. degree from Southeast University in 1998, and Ph. D. degree from Beijing University of Technology in 2004, respectively. Now she is a professor and doctoral supervisor in Beijing University of Technology. Her main research interests include image/video processing, image recognition, and medical image processing.