

DACSNNet:基于双注意力机制与分类监督的乳腺超声图像病变检测

李方, 王洁

引用本文

李方, 王洁. DACSNNet:基于双注意力机制与分类监督的乳腺超声图像病变检测[J]. 计算机科学, 2025, 52(9): 54-61.

LI Fang, WANG Jie. DACSNNet: Dual Attention Mechanism and Classification Supervision Network for Breast Lesion Detection in Ultrasound Images [J]. Computer Science, 2025, 52(9): 54-61.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

基于模仿学习的机器人演奏示教系统

Robot Performance Teaching Demonstration System Based on Imitation Learning
计算机科学, 2024, 51(11A): 240300063-5. <https://doi.org/10.11896/jsjcx.240300063>

面向风格的软件体系结构演化路径生成方法

Style-oriented Software Architecture Evolution Path Generation Method
计算机科学, 2024, 51(11A): 240100130-9. <https://doi.org/10.11896/jsjcx.240100130>

一种单阶段无监督可见光-红外跨模态行人重识别方法

Single Stage Unsupervised Visible-infrared Person Re-identification
计算机科学, 2024, 51(6A): 230600138-7. <https://doi.org/10.11896/jsjcx.230600138>

基于特征融合的边缘引导乳腺超声图像分割方法

Segmentation Method of Edge-guided Breast Ultrasound Images Based on Feature Fusion
计算机科学, 2023, 50(3): 199-207. <https://doi.org/10.11896/jsjcx.211200294>

基于可能回答集程序的多Agent信念协调

Belief Coordination for Multi-agent System Based on Possibilistic Answer Set Programming
计算机科学, 2020, 47(2): 201-205. <https://doi.org/10.11896/jsjcx.190100101>

DACSNet: 基于双注意力机制与分类监督的乳腺超声图像病变检测

李 方 王 洁

北京工业大学计算机学院 北京 100124

(leefwasx0921@163.com)

摘 要 超声成像是乳腺病变最常用的检测技术,基于深度学习的乳腺超声图像自动化病变检测引起了越来越多的研究人员关注。然而,大部分研究未能充分融合图像信息来增强特征,也未考虑到注意力模块的引入带来的模型复杂度增大和假阳率升高的问题。因此,对现有的 RetinaNet 模型进行改进,以 VMamba 为骨干网络,提出了基于双注意力机制与分类监督的病变检测网络(DACSNet)以提高乳腺超声图像中病变检测的准确性,并降低检测假阳率。具体来说,将医学领域的知识引入注意力模块,通过双注意力模块(DAM)来增强通道维度和空间维度的特征。DAM 仅涉及少量参数,且能有效提高模型的检测性能。此外,为了降低病变检测的假阳率,在模型中加入了分类监督模块(CSM)来融合病变分类信息,实现对疑似病变区域的二次关注。为了验证 DACSNet 的性能,在 3 组公开的乳腺超声图像数据集上进行了乳腺病变检测实验,结果证明了该方法的有效性。

关键词: 乳腺超声图像;病变检测;VMamba;双注意力模块;分类监督

中图分类号 TP391

DACSNet: Dual Attention Mechanism and Classification Supervision Network for Breast Lesion Detection in Ultrasound Images

LI Fang and WANG Jie

School of Computer Science, Beijing University of Technology, Beijing 100124, China

Abstract Ultrasound imaging is the most commonly used technology for breast lesion detection and automated lesion detection of breast ultrasound images has attracted increasing attention from researchers. However, most studies fail to fully integrate image information to enhance features, and they do not take into account the problems of the increased model complexity and neglect the issue of rising false positive rates. Therefore, this paper improves existing RetinaNet model by using VMamba as the backbone network and proposes a lesion detection network based on dual attention mechanism and classification supervision(DACSNet) to improve the accuracy of lesion detection in breast ultrasound images and reduce the false-positive rate. Specifically, this paper incorporates medical domain knowledge into the attention module, where the dual attention module(DAM) effectively enhances feature representation in both the channel and spatial dimensions. The DAM involves only a small number of parameters and effectively boosts the model's detection performance. Furthermore, to reduce the false-positive rate of lesion detection, a classification supervision module(CSM) is added the model to integrate lesion classification information, achieving secondary focus on suspected lesion areas. To verify the performance of DACSNet, experiments for breast lesion detection are conducted on three sets of publicly available breast ultrasound image datasets, and the experimental results demonstrate the effectiveness of this method.

Keywords Breast ultrasound image, Lesion detection, VMamba, Dual attention module, Classification supervision

1 引言

根据世界卫生组织国际癌症研究机构(GLOBOCAN 2022)公布的全球癌症流行统计,全球每年女性乳腺癌新发病例约 230 万例^[1],占全球所有癌症的 11.6%。乳腺癌是全世界女性死亡的第二大原因,早期筛查和诊断可以显著降低死亡率(40%或更多)。超声成像因具有多功能性、高安全性、低成本和高灵敏度特性而成为乳腺病变检测最常用的技术,但超声成像质量在很大程度上取决于机器设备、超声医师的操

作经验等多种因素。此外,由于散斑噪声和伪影的存在,以及其复杂性等原因,受过不同培训、具有不同临床经验的医生可能会对同一个超声成像做出不同的诊断结果^[2]。因此,实现对乳腺超声图像的自动化病变检测十分重要,它既能够为超声科医生提供客观建议,消除医生主观性误差,提高诊断的准确性和一致性,又能缩短医生对病情进行评估和判断的时间,提高病人从就诊到确诊的效率。

早期的研究主要应用传统数字图像处理技术和机器学习技术来实现乳腺病变检测。Drucker 等^[3]使用径向

到稿日期:2024-12-23 返修日期:2025-03-14

基金项目:国家自然科学基金(62476015)

This work was supported by the National Natural Science Foundation of China(62476015).

通信作者:王洁(wj@bjut.edu.cn)

梯度指数(RGI)滤波技术自动检测乳腺超声图像中的病变。Madabhushi等^[4]结合图像的强度、纹理特征、经验领域知识和方向梯度,使用基于形变模型的方法来寻找病变的边缘。这些传统方法大多是在表达能力和泛化能力有限的小数据集中得到验证的,而且相对于正常的乳腺组织而言,病变区域占比较小,在这种情况下,使用传统机器学习方法开发的检测方法检测精度较低。

随着深度学习的发展,许多研究人员成功地将深度学习应用于乳腺病变检测。Li 等^[5]开发的两阶段模型 BUSNet,集合了预处理、region proposal、边界回归和后处理,在 BUSI 数据集^[6]上达到了较高的检测精度。为了进一步提升模型的性能,研究者们开始将注意力引入到网络中,进一步提取图像中的信息。Laghmati 等^[7]提出了两种基于注意力的 U-Net (Attention3 U-Net 和 Attention4 U-Net),在原始 U-Net 架构上加入注意力机制,使模型更加关注某些局部信息。但以上模型未能充分增强网络特征的可辨别性。此外,在引入局部或全局信息的同时也引入了一些噪声,引起了假阳率的升高。

研究表明^[8],医学领域知识的引入可以有效地增强特征表示,提高模型的性能。Meng 等^[9]提出了双全局关注网络 DGANet,利用医学领域知识设计了双边空间注意模块和全局通道注意模块,有效增强了特征,获得了良好的性能。但其设计的注意力模块参数量较大,特别是双边空间注意模块需要计算图像的高斯距离矩阵,既耗时又冗余;同时,该方法也未考虑是否会导致假阳率升高的问题。

针对上述问题,文中提出了一个基于双注意力机制与分

类监督的网络(Dual Attention Mechanism and Classification Supervision Network for Breast Lesion Detection in Ultrasound Images, DACSNet),用于乳腺超声图像的病变检测。具体来说,根据医学领域知识设计了双注意力模块(Dual Attention Module, DAM)和分类监督模块(Classification Supervision Module, CSM)。DAM 将图像的通道信息与像素上下文信息进行了充分融合,增强了特征的辨别力,其参数量仅为 1.1×10^6 。CSM 将病变分类信息融合到提取的特征中,进一步加强疑似病变区域的特征,降低了模型检测的假阳率。

本文的主要贡献如下:

- 1)提出了基于 VMamba 的一阶段检测模型 DACSNet,用于乳腺超声图像的病变检测。
 - 2)提出了 DAM 聚合病变附近区域的上下文信息,并增强了重要特征的辨别力,提高了模型性能。DAM 仅涉及少量的参数。
 - 3)提出了 CSM 融合病变分类信息,对疑似病变的区域进行二次关注,降低了病变检测的假阳率。
- 本文第 2 章介绍了所提 DACSNet 方法;第 3 章给出了实验与结果分析;最后总结了研究并对未来的工作进行了展望。

2 DACSNet

DACSNet 的总体结构如图 1 所示。DACSNet 采用 RetinaNet^[10]作为整体框架,使用 VMamba-Tiny^[11]作为骨干网络。VMamba-Tiny 的不同阶段由不同数量的 VSS 模块组成。VSS 的详细结构如图 2 所示。

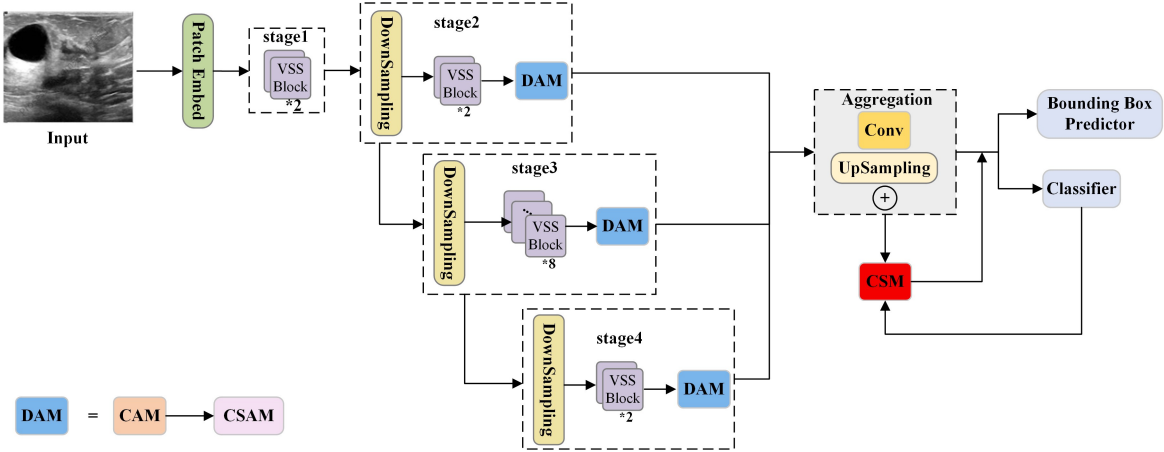


图 1 DACSNet 的结构

Fig. 1 Structure of DACSNet

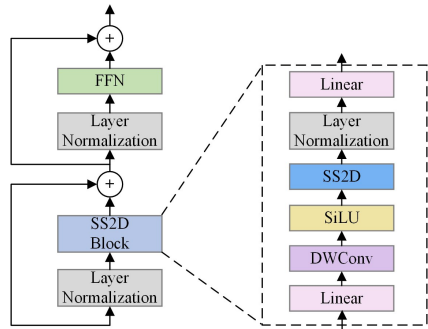


图 2 VSS 模块

Fig. 2 VSS block

首先,对图像进行位置编码,将其输入到由 4 个阶段组成的 VMamba 网络中进行特征提取。VMamba 各个阶段的 VSS 模块数量为[2,2,8,2]。然后,将 3 个 DAM 分别加入到 VMamba-Tiny 的 stage2, stage3, stage4 这 3 个阶段中,充分提取有效特征并抑制图像中斑点、伪影等噪声的干扰。3 个阶段提取的特征充分融合了局部信息和全局信息,将其输入到 RetinaNet 的 Neck 部分进行特征融合,得到 P2—P6 共 5 个特征。最后,将 P2—P6 特征输入到边界框回归器和分类器中进行初次预测,获得分类信息,并利用 CSM 将分类信息融合到 P2—P6 特征中进行二次预测,得到最终检测结果。下文将依次对 DAM 和 CSM 进行详细阐述。

2.1 双注意力模块

DAM由通道注意力模块(Channel Attention Module, CAM)和十字形空间注意力模块(Cross Spatial Attention Module, CSAM)构成。CAM充分探索了特征图中不同通道之间信息的关联度和重要度。CSAM充分利用了病变相邻区域的上下文信息,抑制了噪声的干扰。将特征先后输入到CAM和CSAM中以增强特征的辨别力。

2.1.1 通道注意力模块

根据BI-RADS,不同的医学特征是相互关联的,医生通常会根据多个特征对医学图像进行诊断。例如,一个既有不规则形状又有钙化的图像,其存在病变的概率显著高于仅具有其中一种特征的情况。此外,医生对不同的特征有不同的关注度。例如,超声科医生在对乳腺超声图像进行诊断时会更加关注病变的形态、方位以及边缘清晰度等。在经过深度学习网络提取得到的特征图中,每个通道都可以看作输入图像不同抽象层次特征的表示。在乳腺病变检测任务得到的特征图中,某些通道可能专门关注微小钙化点的分布,某些通道可能主要关注病变的边界清晰度或内部结构等,还有一些通道可能主要捕捉背景噪声或其他无关紧要的细节。因此,本文提出CAM来模拟通道特征之间的关联性,并增强重要通道特征维度的辨别力。CAM的详细结构如图3所示。

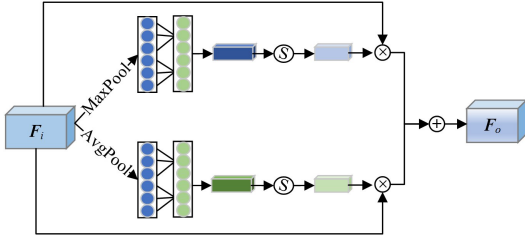


图3 CAM的结构

Fig. 3 Structure of CAM

为了高效计算通道注意力,受CBAM^[12]的启发,本文同时使用平均池化和最大池化来压缩输入特征 F_i 的空间维度,对特征图的空间信息进行聚合,得到平均池化特征 F_{avg}^c 和最大池化特征 F_{max}^c 。避免维度降低对于通道注意的学习十分重要^[13],而且适当的跨通道交互可以在降低模型复杂性的同时保持其性能。因此,首先将池化后的特征输入到1D卷积中,通过选择合适的卷积核 k 来决定局部跨通道交互的范围。为了不增加额外的参数,两个1D卷积之间共享参数,本研究选取 $k=3$ 进行实验。然后将平均池化特征 F_{avg}^c 和最大池化特征 F_{max}^c 分别通过sigmoid激活,得到两个通道注意力 A_{avg} , A_{max} ,如式(1)、式(2)所示。将原始输入特征 F_i 分别与得到的注意力 A_{avg} 和 A_{max} 相乘后求和,得到最终的通道加权特征 F_o ,如式(3)所示。

$$A_{max} = \sigma(\text{Conv}(\text{MaxPool}(F_i))) \quad (1)$$

$$A_{avg} = \sigma(\text{Conv}(\text{AvgPool}(F_i))) \quad (2)$$

$$F_o = F_i A_{max} + F_i A_{avg} \quad (3)$$

其中, σ 为sigmoid函数,Conv为卷积操作,MaxPool为最大池化操作,AvgPool为平均池化操作。

2.1.2 十字形空间注意力模块

通道注意力只能聚合不同特征之间的联系信息,无法聚合像素的上下文信息。而乳腺超声图像中病变的临近区域可

以反映回声模式及其与周围组织的对比度等信息。例如,低回声区表示可能存在液体积聚或实质成分,高回声点表示可能存在钙化点,钙化点在恶性病变中比较常见。不同的背景回声模式(如均匀或不均匀)可以反映不同的乳腺密度或潜在的基础疾病状态。因此,病变临近区域的信息对于准确检测和评估病变十分重要。然而,乳腺超声图像通常也存在背景噪声、伪影以及非目标结构的信息,这些因素可能会误导模型,导致检测结果错误。因此,我们提出了CSAM来聚合图像的空间信息。CSAM仅包括垂直方向与水平方向的计算,能够以更少的计算量和内存在空间上聚合像素的上下文信息,抑制噪声的干扰,增强模型对有效信息的辨别力。CSAM的详细结构如图4所示。

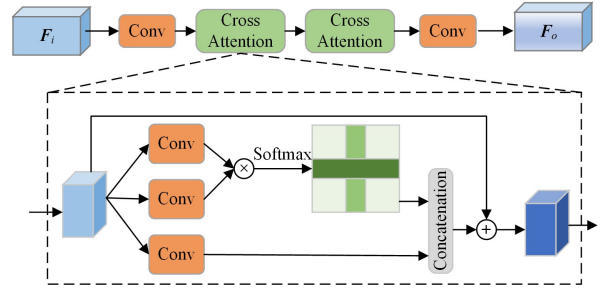


图4 CSAM的结构

Fig. 4 Structure of CSAM

为了减少计算量,首先应用卷积对输入特征 $F_i \in \mathbb{R}^{C \times W \times H}$ 在通道维度进行降维,获得特征映射 $F \in \mathbb{R}^{C' \times W \times H}$, $C' < C$ 。然后将 F 输入到十字注意力模块(Cross Attention, CA)中,生成一个新的特征图 F' ,该特征图将垂直和水平方向中每个像素的上下文信息聚合在一起。但特征图 F' 仅包含两个方向的信息。为了获得更丰富、更密集的上下文信息,将特征图 F' 再次输入到CA中,输出特征图 F'' 。最后应用卷积将特征图恢复到原来的维度,如式(4)所示。为了避免添加过多的额外参数,两个十字注意力模块共享参数。

$$F_o = \text{Conv}(\text{CA}(\text{CA}(\text{Conv}(F_i)))) \quad (4)$$

其中,CA代表十字注意力模块处理,Conv为卷积操作。

为了降低特征的维度,对特征 $F \in \mathbb{R}^{C' \times W \times H}$ 分别应用2个带有 1×1 卷积核的卷积层,得到特征图 Q 和 K , $\{Q, K\} \in \mathbb{R}^{C'' \times W \times H}$, $C'' < C'$, W 和 H 分别是特征图的宽度和高度, C'' 和 C' 为特征图的通道维数。对于 Q 中的每个位置 w ,可以得到一个向量 $Q_w \in \mathbb{R}^{C''}$,同时能从 K 中提取出与 w 在同一行或同一列的向量构成集合 K_w ($K_w \in \mathbb{R}^{(H+W-1) \times C''}$)。将 Q_w 和 K_w 进行点积计算,得到相关度矩阵 $D \in \mathbb{R}^{(H+W-1) \times (W \times H)}$,最后应用softmax得到注意力矩阵 A ,如式(5)所示:

$$A_u = \text{softmax}(Q_u^T K_u) \quad (5)$$

为了进行特征交互,对特征 $F \in \mathbb{R}^{C' \times W \times H}$ 再次应用带有 1×1 卷积核的卷积层,得到特征图 $V \in \mathbb{R}^{C'' \times W \times H}$ 。对于 V 中的每个位置 u ,可以得到一个集合 V_u ($V_u \in \mathbb{R}^{(H+W-1) \times C''}$),其中包含了与位置 u 在同一行或同一列的所有特征向量。通过注意力矩阵 A 对特征集合 V_u 进行加权求和,有选择地聚合上下文信息,最后与输入特征图 F 的 u 位置值相加得到增强特征向量 F_u' ,进而提高局部特征的辨别力,如式(6)所示:

$$F_u' = \sum_{i=0}^{H+W-1} A_{i,u} V_{i,u} + F_u \quad (6)$$

经过两个十字注意力模块的处理后,特征图 F' 中的每个像素都聚合了各个方向的信息,如图 5 所示。

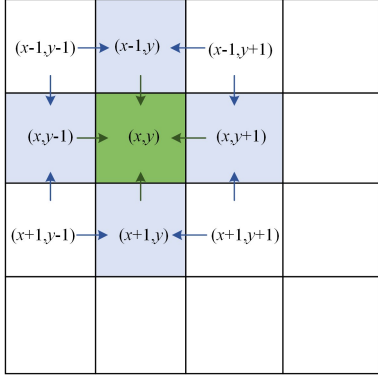


图 5 经过两个十字注意力模块处理的示例

Fig. 5 Example of processing through two cross attention modules

十字注意力模块对像素点的垂直方向和水平方向进行信息聚合,首先经过第一个十字注意力模块处理,像素点 $(x, y-1)$ 、 $(x-1, y)$ 、 $(x, y+1)$ 、 $(x+1, y)$ 包含了位置 $(x-1, y-1)$ 、 $(x-1, y+1)$ 、 $(x+1, y+1)$ 、 $(x+1, y-1)$ 的信息;然后经过第二个十字注意力模块的处理,像素点 (x, y) 聚合了位置 $(x, y-1)$ 、 $(x-1, y)$ 、 $(x, y+1)$ 、 $(x+1, y)$ 的信息,因此像素

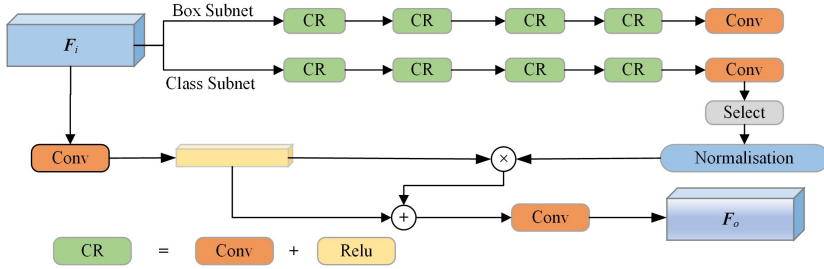


图 6 CSM 的结构

Fig. 6 Structure of CSM

本文利用分类预测分支的 anchor 来引导网络关注感兴趣的区域,进而抑制假阳性病变区域的干扰。首先,从分类特征图 F_i 中提取 anchor 信息,分别找到恶性和良性两种病变所对应的具有最大预测值的 anchor A_b 和 A_m ,对其进行拼接再归一化后得到特征 S ,如式(7)~式(9)所示。然后,利用卷积调整骨干网络提取的特征 F_i 的形状,得到特征 F_i' 。利用残差注意力融合^[14]来进行特征融合,将 F_i' 与特征 S 相乘之后再与 F_i' 相加得到融合特征。最后,利用卷积恢复融合特征的形状,得到新特征 F_o ,如式(10)所示。将输出特征 F_o 输入到分类预测分支和边界回归分支中进行二次预测。

$$S = \text{Norm}(\text{CAT}([A_b, A_m])) \quad (7)$$

$$A_b = \max(F_{i,b}) \quad (8)$$

$$A_m = \max(F_{i,m}) \quad (9)$$

$$F_o = \text{Conv}(\text{Conv}(F_i)S + \text{Conv}(F_i)) \quad (10)$$

其中, $A_b, A_m \in \mathbb{R}^{H \times W \times \text{Anchor}}$, CAT 表示连接操作, Conv 为卷积操作。

3 实验与分析

3.1 数据集

本研究使用了 3 个公共数据集,分别是 CVA-BUS^[15],

点 (x, y) 在经过两个十字注意力模块处理后聚合了其周围领域各个方向的信息。同时,在两次信息聚合的过程中,越靠近像素点 (x, y) 的区域在两次信息聚合过程中信息不断增加,较远区域的像素信息在注意力图中占比较少。因此,经过两个十字注意力模块处理后可以有效增强邻近区域的特征融合,并抑制图像远处的噪声干扰。

2.2 分类监督模块

基于深度学习的医学图像病变检测研究中,高灵敏度地检测出病变为首要任务,注意力模块的加入会提高模型的检测性能,但同时可能导致检测结果假阳率的增加,这对于实际诊断应用是一个严重的问题。在实际诊断中,医生在对疑似病变存有较大的不确定性时,往往会多次查看疑似病变区域及其周围组织的信息。此外,在获取医学影像数据的人工标注流程中往往会设置初检、二检,甚至三检。因此,本文加入了 CSM 将初次预测得到的分类信息与提取的特征进行融合,以进一步降低病变检测的假阳率。初次预测得到的分类信息提供了额外的监督信号,结合分类信息后,模型可以根据病变的具体类型调整对不同特征的关注度,并引导网络更加关注疑似病变区域,抑制无关的检测,从而降低病变检测的假阳率,使模型在提高病变检测率的同时进一步提高病变检测的可靠性。CSM 的详细结构如图 6 所示。

BUSI^[6], BUI^[16]。

1) CVA-BUS 数据集包含 188 个乳腺超声视频,其中 113 个是恶性视频,75 个是良性视频。这 118 个视频中有 25 272 张图像,每个视频的超声图像数量为 28 张到 413 张不等。

2) BUSI 数据集来自埃及开罗的 Baheya 妇女癌症早期发现和治疗医院,其包含的乳腺超声图像来自于 600 名年龄在 25 岁至 75 岁之间的女性患者。BUSI 包含 437 张良性图像,210 张恶性图像,133 张正常图像,共 730 张乳腺超声图像。

3) BUI 数据集由 Shahid Beheshti University of Medical Sciences 采集获得,图像来源于 18 岁以上的患者在接受治疗前或随后的影像学检查。BUI 包含 123 张恶性图像,109 张良性图像。

本研究对 CVA-BUS 数据集进行了图像挑选。首先经过检查删掉了 3 个重复的视频以及 2 个问题视频,剩余 183 个视频,其中 112 个为恶性,71 个为良性。然后以同一视频中图像的病变位置重叠率不超过 80% 为标准进行了图像过滤,筛选出了 3 557 张乳腺超声图像,其中良性病变 988 张,恶性病变 2 569 张。每个超声图像包含至少一个病变。对于 CVA-BUS 数据集,随机地从中选择 601 张图像(来源于 36 个

视频,约占数据集的 20%)作为测试集,剩余的图像作为训练集。对于 BUSI 和 BUI 数据集,由于数据集较小,因此将两个数据集共同使用。

3.2 实现细节

使用在 ImageNet 上预训练的 VMamba-Tiny 来初始化骨干网络,加入模块的参数采用随机初始化。使用 SGD 算法优化网络,权重衰减为 0.0001,动量为 0.9,批大小为 2,训练 25 epoch。初始学习率设为 2.5×10^{-4} ,使用 PyTorch 框架实现本文网络,并在单个 NVIDIA GeForce RTX 4060 Laptop GPU 上进行所有实验。

3.3 评价指标

为了定量评估不同方法对乳腺超声图像病变检测的效果,选择了 AP、AP50、AP75、检测精度(Precision)、召回率(Recall)、平均每张图像的假阳率(FPI,FP Per Image)、每秒预测的帧数(Frames Per Second,FPS)和 Params(参数量,输入图像大小为 512×512)作为评价指标。AP,AP50,AP75 遵循 COCO 评估器的定义;AP50,AP75 是在 IoU 阈值分别为 0.5,0.75 时计算的 AP。IoU,precision,recall,FPI 的定义如式(11)–式(14)所示。

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

(11)

表 1 在 CVA-BUS 数据集上不同方法的定量结果

Table 1 Quantitative results of different methods in CVA-BUS dataset

Method	AP ↑	AP50 ↑	AP75 ↑	Precision ↑ / %	Recall ↑ / %	FPI ↓ / %	Params ↓	FPS ↑
RetinaNet	19.6	40.7	17.1	81.72	82.80	18.63	3.7×10^7	9.69
ATSS	19.6	40.2	16.2	93.82	75.37	4.99	3.2×10^7	9.85
YOLOv5s-BiFPN-CBAM	21.1	39.4	20.2	88.73	74.80	11.25	1.8×10^7	10.31
Faster R-CNN	19.0	37.5	16.3	79.24	83.30	21.96	4.2×10^7	9.24
RTMDet_tiny	22.7	40.2	22.3	87.94	69.91	9.65	1.5×10^7	10.58
Deformable_DETR	19.5	39.6	17.5	90.20	80.66	8.81	4.3×10^7	8.97
DGNet	22.4	46.8	16.9	83.19	83.47	16.97	5.0×10^7	7.07
DACSNet(Ours)	25.5	48.1	23.4	89.32	82.97	9.98	4.5×10^7	7.81

除了 AP,AP50,AP75 外,还计算了不同方法的 precision,recall 和平均每张图像的假阳率(FPI)。ATSS 获得了最高的 precision(93.82%)和最低的 FPI(4.99%),但其 recall 仅为 75.37%,与 precision 相差过大。DGNet 获得了最高的 recall(83.47%),但其 FPI 高达 16.97%。本文所提的 DACSNet 在 precision(89.32%)和 recall(82.97%)两个指标上既接近最高值,又达到了较好的平衡,FPI(9.98%)也较低,说明 DACSNet 能实现更加精确的病变定位。实验表明,DACSNet 在乳腺超声图像中具有准确检测病变的能力。此外,DACSNet 的检测实时性达到了 7.81FPS,尽管其参数量与实时性不及其余检测方法,但相差不大。

除了定量分析之外,还选取了 5 个样本进行了检测结果可视化,如图 7 所示。从图 7(a)中可以看出,8 种检测方法都检测出了病变,但 DGNet 对病变分类错误。其中 DACSNet 预测的边界框置信度为 97.0,明显高于另外 7 种检测方法,证明了 DACSNet 具有最高的 AP 值。对于图 7(b)中的样本,RetinaNet,ATSS,DGNet 未能检测出乳腺超声图像中的病变;Faster R-CNN 检测出的边界框位置为假阳;Deformable_DETR 检测出两个病变,一个为假阳,一个类别错误;只有 YOLOv5s-BiFPN-CBAM,DACSNet 和 RTMDet_tiny 精

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

(12)

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

(13)

$$FPI = \frac{FP}{N}$$

(14)

其中,TP(真阳性)表示检测到正确标记的病变数量,FN(假阴性)表示未检测到的病变数量,FP(假阳性)表示检测到错误标记的病变数量,N 为图像数量。

3.4 与先进网络的比较实验

3.4.1 CVA-BUS 数据集上的比较

为了在同等模型大小的条件下与 DACSNet 进行比较,选取 RetinaNet^[11],ATSS^[17],YOLOv5s-BiFPN-CBAM^[18],Faster R-CNN^[19],RTMDet_tiny^[20],Deformable_DETR^[21],DGNet^[10]在 CVA-BUS 数据集上进行了检测实验,其中 RetinaNet 和 DGNet 选取了 ResNet50 作为骨干网络。实验的定量结果如表 1 所列。在 7 种比较方法中,RTMDet_tiny 的 AP(22.7)和 AP75(22.3)得分最高,DGNet 的 AP50(46.8)得分最高。本文所提的 DACSNet 在 AP,AP50,AP75 指标上都优于 RTMDet_tiny 和 DGNet,分别提高了 2.8,1.3,1.1。

确地检测出了图像中的病变,可见这 3 种方法能够精准检测出小目标病变,但 DACSNet 更优,其给出的边界框置信度(42.6)高于 YOLOv5s-BiFPN-CBAM(37.3)和 RTMDet_tiny(32.9)。在比较方法中,RTMDet_tiny 和 YOLOv5s-BiFPN-CBAM 的 AP 得分较高,与它们相比,DACSNet 的 AP 分别提高了 2.8 和 4.4,图 7(b)证实了此结果的真实性。对于图 7(c)中的样本,ATSS 未能检测到病变,RetinaNet,YOLOv5s-BiFPN-CBAM,Faster R-CNN,RTMDet_tiny,Deformable_DETR,DGNet 都出现了假阳性病变,只有 DACSNet 精确地检测出了图像中的病变且未出现假阳性病变。这表明,DACSNet 达到了较低的假阳率。对于图 7(d)中的样本,8 种方法均检测出了病变,但 ATSS,Deformable_DETR,RetinaNet,RTMDet_tiny 对病变分类错误。此外,RetinaNet,Faster R-CNN 与 DGNet 检测出了假阳性病变,证实了定量结果中 3 种方法具有较高的假阳率。YOLOv5s-BiFPN-CBAM 与 DACSNet 准确地检测出了病变,两种方法的 precision 和 FPI 十分接近。对于图 7(e)中的样本,ATSS 虽准确检测出了病变,但对病变的定位不够精确。ATSS 使用多种尺寸和宽高比的锚框来进行检测,所用的锚框类型比本文方法丰富,从而达到了较高的 precision,但同时也会带来定位的

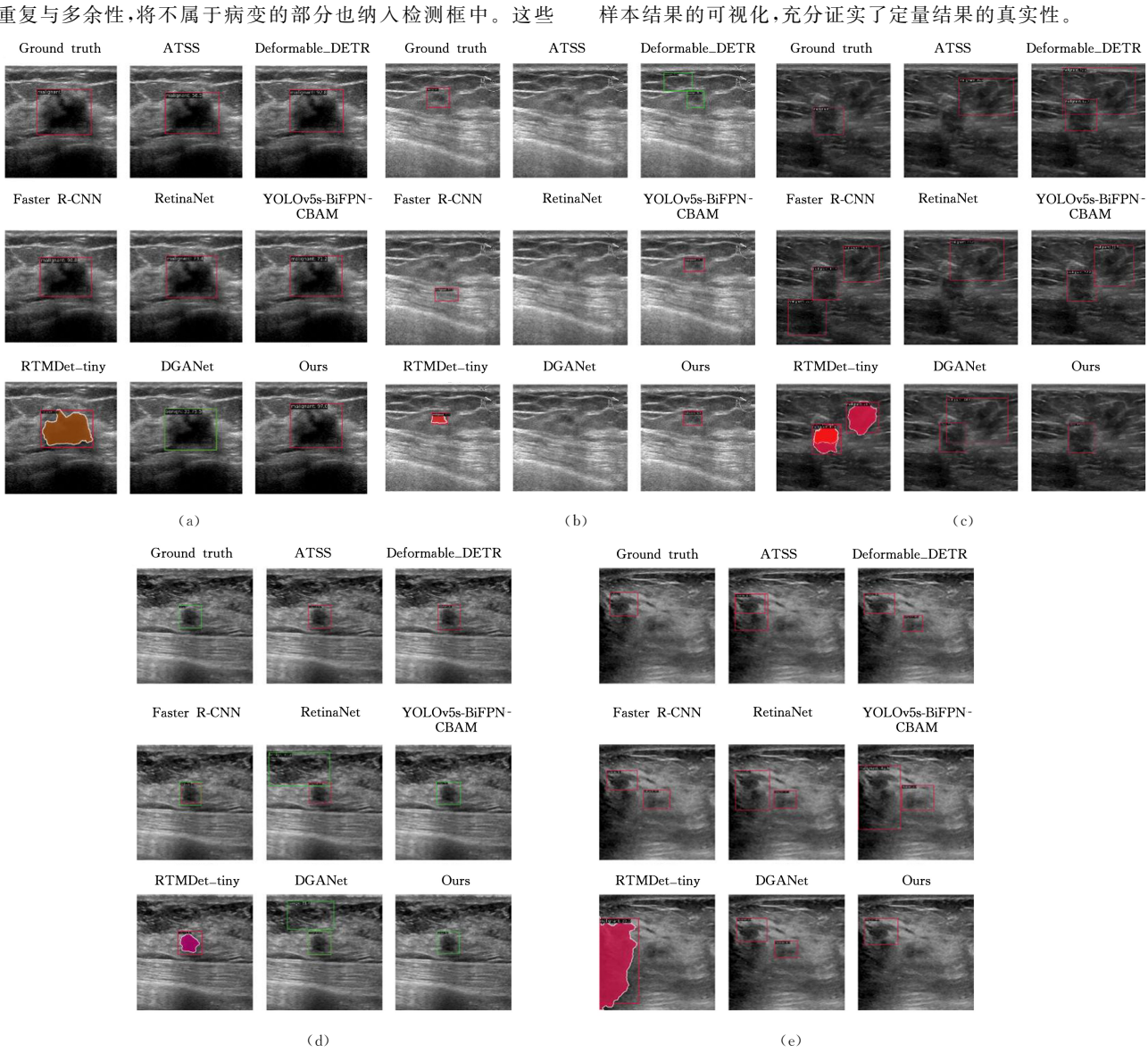


图 7 采用不同检测方法对 CVA-BUS 数据集中 5 个样本的可视化结果

Fig. 7 Visualization results of five samples in CVA-BUS dataset using different detection methods

3.4.2 BUSI+BUI 数据集上的比较

为了进一步评估 DACSNet 的泛化性能,在 BUSI 和 BUI 公共数据集上进行了乳腺病变检测实验。考虑到 BUSI 和 BUI 数据集的数据量较小,采用五重交叉验证实验来评估不同方法的性能。实验的定量结果如表 2 所列。在 8 种比较方法中,DACSNet 的 AP 和 AP50 得分最高,分别为 54.4 和 81.3,其 AP75 仅比 RetinaNet(61.0)低 0.2。Deformable_DETR 获得了最高的 precision 和最低的 FPI,分别为 92.97%

和 6.82%,但其 recall 仅为 87.51%,与 precision 略微失衡。DGANet 的 recall(91.43%)最高,但与 precision 相差较多,其 FPI 甚至高达 17.97%。DACSNet 的 precision 和 recall 分别为 92.17%和 91.09%,两个指标达到了很好的平衡且值都较高,同时 DACSNet 的 FPI 仅为 7.95%,证明其在保持高性能的同时假阳率也未过度升高。此结果进一步表明了 DACSNet 在乳腺超声图像病变检测中具有较高的检测精度和泛化性。

表 2 不同方法在 BUSI 和 BUI 数据集上的定量结果

Table 2 Quantitative results of different methods in BUSI and BUI dataset

Method	AP ↑	AP50 ↑	AP75 ↑	Precision ↑ / %	Recall ↑ / %	FPI ↓ / %
RetinaNet	53.3	78.4	61.0	85.91	89.86	15.12
ATSS	52.3	77.6	59.4	92.05	85.28	7.61
YOLOv5s-BiFPN-CBAM	50.1	76.5	53.8	90.52	83.65	9.37
Faster R-CNN	48.3	74.3	54.4	85.16	88.86	15.81
RTMDet_tiny	48.9	70.2	56.9	89.71	78.70	9.44
Deformable_DETR	48.6	75.1	56.7	92.97	87.51	6.82
DGANet	46.2	75.6	50.3	83.87	91.43	17.97
DACSNet(Ours)	54.4	81.3	60.8	92.17	91.09	7.95

3.5 消融研究

为了评估 CAM,CSAM,DAM 和 CSM 的有效性,在筛选后的 CVA-BUS 数据集上进行了消融研究实验,实验的基线是以 VMamba-Tiny 为骨干的 RetinaNet 网络。实验结果如表 3 所列。

为了分析 CAM 和 CSAM 模块的有效性,在基线模型的

基础上分别单独加入 CAM、单独加入 CSAM 以及同时加入 CAM 与 CSAM 进行实验。从表 3 中可看出,与基线相比,单独加入 CAM 和单独加入 CSAM 均使模型的 AP,AP50,AP75 指标得到了提高,也使 precision 和 recall 达到了更好的平衡,但 FPI 降低较少;同时加入 CAM 与 CSAM 使模型性能得到了进一步的提高,但 FPI 无明显变化。

表 3 消融实验的定量结果
Table 3 Quantitative results of ablation experiments

Method	AP ↑	AP50 ↑	AP75 ↑	Precision ↑ / %	Recall ↑ / %	FPI ↓ / %
Basic	23.6	47.3	18.9	81.85	90.24	20.13
Basic+CAM	25.1	48.8	22.4	87.54	89.42	17.81
Basic+CSAM	24.3	47.9	22.6	88.88	84.62	16.64
Basic+DAM	25.8	51.0	23.1	84.35	86.44	16.13
Basic+CSM	23.3	47.3	18.9	86.42	89.42	14.14
DACSNNet	25.5	48.1	23.4	89.32	82.97	9.98

为证实 CAM 与 CSAM 的兼容性,选取了两个可视化样本,如图 8 所示。对于图 8(a)中的样本,Basic+CAM 检测出了假阳性病变,而 Basic+CSAM 和 Basic+DAM 实现了精准检测,从而可以看出 CSAM 通过聚合检测区域的上下文信息,排除了左侧区域为病变的可能性,帮助模型实现了更精准的检测。对于图 8(b)中的样本,Basic+CAM 检测出了假阳性病变,Basic+CSAM 未检测出病变,Basic+DAM 精准检测出了病变。CAM 的加入使模型充分利用通道中的不同特征信息检测出真正的病变位置;CSAM 的加入使模型结合检测区域的上下文信息,排除了假阳性病变。因此,CAM 与 CSAM 使模型能够充分提取图像中的有效信息,达到更好的性能。

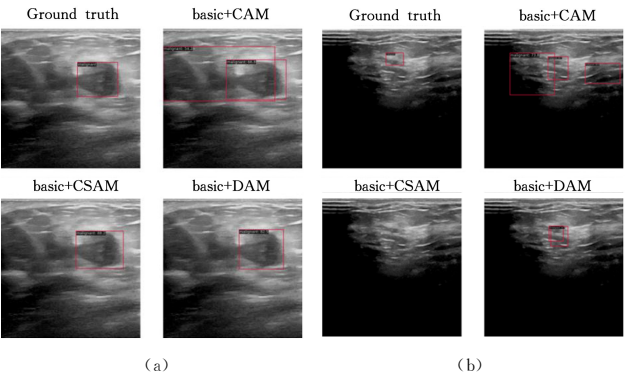


图 8 CAM+CSAM 消融实验的可视化结果

Fig. 8 Visualization results of CAM+CSAM ablation study

为了分析 DAM 与 CSM 的有效性,在基线模型的基础上分别单独加入 DAM、单独加入 CSM、同时加入 DAM 与 CSM 进行实验。从表 3 中可看出,与基线相比,DAM 的加入使 AP,AP50,AP75 分别提高了 2.2,3.7,4.2,同时也使 precision 和 recall 达到了更好的平衡,FPI 也降低了 4 个百分点,证明了 DAM 能够有效聚合不同特征的关联信息以及图像上下文信息,提高了模型的检测性能,减少了因误诊或漏诊引起的潜在危害。此外,CSM 的加入在保持基线性能的同时将 precision 提高了 4.57 个百分点,更重要的是 FPI 降低了近 6 个百分点,证明了初次预测分类信息的加入可以使模型更好地关注疑似病变区域,提高病变检测的可靠性。DAM 和

CSM 的整合不仅提高了基线的 AP,AP50,AP75 指标,还将 FPI 降至 9.98%。此结果表明了 DAM 和 CSM 的有效性。

为证实 DAM 和 CSM 的兼容性,选取了两个可视化样本进行分析,如图 9 所示。对于图 9(a)中的样本,Basic+DAM 检测了两个病变,其中一个为假阳;Basic+CSM 使模型利用初次预测的分类信息对疑似病变区域再评估,排除了假阳性病变,但其对病变分类给出了两种结果。而 DAM 与 CSM 的结合使 DACSNNet 没有出现假阳性检测且对病变的分类正确,并提高了其置信度分数(Basic+DAM 为 72.9,Basic+CSM 为 52.1,DACSNNet 为 78.6)。对于图 9(b)中的样本,Basic+DAM 检测出良性病变,Basic+CSM 经过二次评估后给出此病变也可能为恶性的结果,而两个模块的同时使用使 DACSNNet 整合相关信息,对病变的最终预测为恶性,且置信度分数高达 93.1。这充分证明了定量结果的真实性。

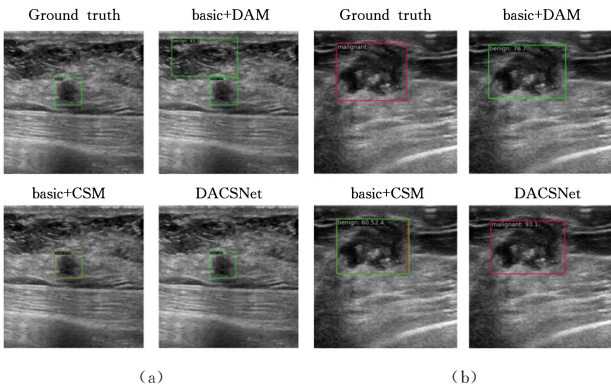


图 9 DAM+CSM 消融实验的可视化结果

Fig. 9 Visualization results of DAM+CSM ablation study

结束语 本文提出了一个基于双注意力机制与分类监督的网络 DACSNNet,以实现对乳腺超声图像的自动化病变检测。具体来说,将医学领域知识引入 DAM 和 CSM 模块的设计中。DAM 以较少的参数量和计算成本有效聚合了通道信息以及病变附近区域的上下文信息,抑制了超声图像中斑点等噪声的干扰,从通道维度和空间维度增强了特征的辨别力,使模型的检测能力增强。为了降低假阳率,利用 CSM 将初次预测的分类信息融入到提取的特征中,对疑似病变区域进行再关注,在提高检测精度的同时减少假阳率的升高。为了评

估 DACSNet 的性能,在 3 组公开数据集上进行了乳腺超声图像病变检测实验,实验结果证明了该方法具有在较低假阳率下提高乳腺超声图像病变检测性能的潜力。后续将把模型进一步应用于超声视频的检测,融合时间信息,进一步提高模型的检测性能。

参 考 文 献

[1] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA: a Cancer Journal for Clinicians, 2024, 74(3): 229-263.

[2] QIAN X, PEI J, ZHANG H, et al. Prospective assessment of breast cancer risk from multimodal multiview ultrasound images via clinically applicable deep learning[J]. Nature Biomedical Engineering, 2021, 5: 522-532.

[3] DRUKKER K, GIGER M L, HORSCH K, et al. Computerized lesion detection on breast ultrasound[J]. Medical Physics, 2002, 29(7): 1438-1446.

[4] MADABHUSHI A, METAXAS D N. Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions[J]. IEEE Transactions on Medical Imaging, 2003, 22(2): 155-169.

[5] LI Y J, GU H, WANG H Y, et al. BUS-net: A Deep Learning Model of Breast Tumor Lesion Detection for Ultrasound Images [J]. Frontiers in Oncology, 2022, 12: 12848271.

[6] AL-DHABYANI W, GOMAA M, KHALED H, et al. Dataset of breast ultrasound images[J]. Data in Brief, 2020, 28: 104863.

[7] LAGHMATI S, HICHAM K, CHERRADI B, et al. Segmentation of Breast Cancer on Ultrasound Images using Attention U-Net Model[J]. International Journal of Advanced Computer Science and Applications(IJACSA), 2023, 14(8): 770-778.

[8] XIE X, NIU J, LIU X, et al. A Survey on Incorporating Domain Knowledge into Deep Learning for Medical Image Analysis[J]. Medical Image Analysis, 2021(8): 101985.

[9] MENG H, LI X F, NIU J W, et al. DGANet: A Dual Global Attention Neural Network for Breast Lesion Detection in Ultrasound Images[J]. Ultrasound in Medicine & Biology, 2022, 49(1): 31-44.

[10] 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: 2999-3007.

[11] LIU Y, TIAN Y, ZHAO Y Z, et al. Vmamba: Visual state space model[J]. arXiv: 2401. 10166, 2024.

[12] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module[C]// Computer Vision—ECCV 2018. Cham: Springer, 2018.

[13] WANG Q L, WU B G, ZHU P F, et al. ECANet: Efficient Chan-

nel Attention for Deep Convolutional Neural Networks[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2020: 11531-11539.

[14] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2024-2036.

[15] LIN Z, LIN J H, ZHU L, et al. A New Dataset and a Baseline Model for Breast Lesion Detection in Ultrasound Videos[C]// Medical Computing and Computing Assisted Intervention. Cham: Springer, 2022: 614-623.

[16] ARDAKANI A A, MOHAMMADI A, MIRZA-AGHAZADEH-ATTARI M, et al. An open-access breast lesion ultrasound image database: Applicable in artificial intelligence studies[J]. Computers in Biology and Medicine, 2022, 152: 106438.

[17] ZHANG S F, CHI C, YAO Y Q, et al. Bridging the Gap Between Anchor-based and Anchor-free Detection via Adaptive Training Sample Se-lection[J]. arXiv. 1912. 02424, 2019.

[18] SUN X, WANG X Y, LIU J, et al. Classic YOLO Series Target Detection Algorithms and Their Application in Breast Cancer Detection [J]. Computer Systems and Applications, 2023, 32(12): 52-62.

[19] REN S, HE K, GIRSHICK R, et al. Faster RCNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.

[20] LYU C Q, ZHANG W W, HUANG H A, et al. RTM-Det: An Empirical Study of Designing Real-Time Object Detectors[J]. arXiv: 2212. 07784, 2022.

[21] ZHU X, SU W, LU L, et al. Deformable DETR: Deformable Transformers for End-to-End Object Detection[C]// International Conference on Learning Representations, 2021.



LI Fang, born in 1999, postgraduate. Her main research interests include medical image lesion detection and deep learning.



WANG Jie, born in 1972, Ph.D, associate professor. Her main research interests include logic programming, agent-oriented programming languages, and deep learning.

(责任编辑:柯颖)