

基于 Attention U-Net 的乳腺 X 线图像微钙化检测模型的临床应用

孙晓琪^{1,2}, 蔡思清², 任艳楠²

1. 福建医科大学附属泉州第一医院影像科, 福建 泉州 362000; 2. 福建医科大学附属第二医院放射科, 福建 泉州 362000

【摘要】目的:通过开发基于 Attention U-Net 的乳腺 X 线图像微钙化检测模型, 实现微钙化的高效检出, 并探究不同性质钙化、不同乳腺密度对该深度学习模型微钙化检测性能的影响。**方法:**回顾性分析接受乳腺常规 X 线检查的 347 例患者的 694 幅图像。通过低年资医师独立阅片, 高年资医师审核的方式, 建立微钙化检出的参考标准。进行神经网络训练, 建立深度学习模型。以钙化面积和数量分别计算, 并采用精确率、召回率、F1 分数、交并比等指标评估微钙化检测性能, 分析不同性质钙化(良性 vs 恶性)、不同乳腺密度(a+b 类 vs c+d 类)对深度学习模型微钙化检测性能的影响。**结果:**深度学习模型对微钙化检测的精确率为 $85.12\% \pm 18.39\%$ (以钙化面积计算) 和 $76.72\% \pm 19.85\%$ (以钙化数量计算); 召回率为 $78.18\% \pm 19.25\%$ (以钙化面积计算) 和 $85.12\% \pm 18.39\%$ (以钙化数量计算); 交并比为 $68.29\% \pm 21.39\%$ (以钙化面积计算) 和 $67.13\% \pm 23.84\%$ (以钙化数量计算); F1 分数为 $78.96\% \pm 17.70\%$ (以钙化面积计算) 和 $77.65\% \pm 9.37\%$ (以钙化数量计算)。深度学习模型在不同钙化性质(良性 vs 恶性)中的精确率、召回率、交并比、F1 分数之间差异均无统计学意义 ($P > 0.05$), 在不同乳腺密度(a+b 类 vs c+d 类)中对微钙化检测的精确率、召回率、交并比、F1 分数之间差异无统计学意义 ($P > 0.05$)。**结论:**基于 Attention U-Net 的乳腺 X 线图像微钙化检测模型能够对乳腺微钙化进行有效的检测、有助于乳腺微钙化的定量研究, 同时该模型稳定性强, 钙化性质及乳腺密度对该模型的检测性能无影响。

【关键词】乳腺 X 线图像; 微钙化; 人工智能; 乳腺密度

【中图分类号】R318; R816.4

【文献标志码】A

【文章编号】1005-202X(2024)06-0716-08

Clinical application of mammogram microcalcification detection model based on Attention U-Net

SUN Xiaoqi^{1,2}, CAI Siqing², REN Yannan²

1. Department of Imaging, Quanzhou First Hospital Affiliated to Fujian Medical University, Quanzhou 362000, China; 2. Department of Radiology, the Second Affiliated Hospital of Fujian Medical University, Quanzhou 362000, China

Abstract: Objective To develop a mammogram microcalcification detection model (DL model) based on Attention U-Net for realizing the efficient detection of microcalcifications, and to investigate the effects of breast density and microcalcification type on the microcalcification detection performance of the DL model. **Methods** A retrospective analysis was performed on 694 images from 347 patients undergoing mammography. Through the independent image diagnosis by junior physicians and review by senior physicians, the reference standard for microcalcification detection was established. Neural network training was performed to establish a DL model. The performance of the model for microcalcification detection was evaluated using precision rate, recall rate, intersection over union (IoU) and F1-score which were calculated based on calcification area or quantity; and the effects of microcalcification type (benign vs malignant) and breast density (a+b vs c+d) on the model performance were also analyzed. **Results** For detecting microcalcifications by the DL model, the precision rate, recall rate, IoU and F1-score were $85.12\% \pm 18.39\%$, $78.18\% \pm 19.25\%$, $68.29\% \pm 21.39\%$ and $78.96\% \pm 17.70\%$ when the calculation was based on calcification area, and those were $76.72\% \pm 19.85\%$, $85.12\% \pm 18.39\%$, $67.13\% \pm 23.84\%$ and $77.65\% \pm 9.37\%$ when the calculation was based on calcification quantity. The differences in precision rate, recall rate, IoU, F1-score of DL model in different microcalcification types (benign vs malignant) and breast densities (a+b vs c+d) were insignificant. **Conclusion** The developed mammogram microcalcification detection model based on Attention U-Net can effectively detect breast microcalcifications and is conducive to the quantitative research on breast microcalcifications. Meanwhile, the model exhibits high stability, and the breast density and microcalcification type have trivial effects on the microcalcification detection performance of the model.

Keywords: mammogram; microcalcification; artificial intelligence; breast density

【收稿日期】2024-01-08

【基金项目】福建省自然科学基金(2021J01257); 吴阶平医学基金(3206750.2021-06-35)

【作者简介】孙晓琪, 硕士, 医师, 研究方向: 乳腺疾病的影像诊断, E-mail: 812157370@qq.com

【通信作者】蔡思清, 副教授, 主任医师, 研究方向: 乳腺疾病、骨质疏松等, E-mail: 1920455696@qq.com

前言

乳腺癌作为女性恶性肿瘤中发病率最高的疾病,严重威胁着我国女性的身心健康^[1]。早期乳腺癌的5年生存率接近100%,而进展期乳腺癌的5年生存率只有约28%^[2]。由于乳腺X线摄影广泛应用于乳腺癌筛查中,大量早期乳腺癌得以被及时发现、治疗,显著降低了癌症相关致死率^[3]。乳腺微钙化灶是早期乳腺癌的一个常见征象,临床上约有30%~50%无法触诊到的乳腺癌可仅表现为钙化灶^[4]。相比于超声、磁共振检查,乳腺X线摄影对微钙化的检出敏感性 & 诊断特异性更高^[5]。随着女性两癌(乳腺癌、宫颈癌)筛查的兴起,乳腺X线影像数据大量增加。然而培养一个合格的乳腺X线影像诊断医师周期较长,现有诊断医师水平不一,漏诊、误诊较为普遍。同时乳腺X线影像判别耗时耗力,乳腺X线影像诊断医师长期连续阅片容易导致视觉疲劳、注意力集中度下降,影响诊断的准确率。

人工智能技术的出现给医学影像诊断带来了新的发展机遇。近年来基于深度学习的图像处理技术取得了飞跃式发展,并成功应用于医学影像领域,辅助医生进行疾病诊断,大大提高了医生的诊断效率。许多研究在乳腺X线图像上微钙化的检测方面做出了积极的探索。Cai等^[6]提出基于卷积神经网络及手工标注的乳腺钙化区域检测算法,该算法的准确率为88.59%,敏感性为88.43%,特异性为86.89%。邢素霞等^[7]提出基于空-频域差值与超分辨率反馈网络的乳腺微钙化检测模型,该模型准确率为96.25%,敏感性为94.38%,特异性为98.13%。申楠等^[8]提出基于Adaboost决策树算法的乳腺微钙化真假阳性检测模型,该模型在DDSM数据库上取得91.75%的准确率,91.75%的敏感性,91.79%的特异性。

2018年,Oktay等^[9]提出一种新的机器学习模型-Attention U-Net,并首次应用于CT图像上胰腺的分割。目前这种机器学习模型很少被报道运用于乳腺X线图像微钙化的识别检测上。为进一步提高乳腺X线图像上微钙化的检测准确率,降低检测假阳性率,本研究拟开发基于Attention U-Net的乳腺X线图像微钙化检测模型,并基于模型检测到的微钙化面积、数量来评价模型的性能,比较该模型在不同性质钙化、不同乳腺密度中对微钙化的检测性能。

1 资料与方法

1.1 研究对象

回顾性分析2018年1月~2020年1月10 219例于福建医科大学附属第二医院放射科接受乳腺X线检

查的女性患者资料。由于本研究属于回顾性研究,经福建医科大学附属第二医院伦理委员会审查免除患者知情同意(伦理审批号:[2021]福医附二伦理审字(45)号)。

纳入标准:(1)单侧乳腺病灶,且病灶内有多个微钙化灶(直径<1 mm);(2)乳腺X线检查前未经过任何诊疗,包括手术、穿刺及放化疗等;(3)图像质量、拍摄条件及拍摄投照体位符合中华医学会制定的《乳腺影像检查方法专家共识》^[10];(4)所有病灶均取得病理结果(穿刺或手术)。最终有347例患者资料符合此标准,被纳入研究中。患者年龄21~75岁,平均年龄(47.1±10.5)岁。

所有乳腺X线检查均采用美国Hologic公司Selenia Dimensions数字乳腺机,行常规双侧头尾位(Cranio-Caudal, CC)和内外斜位(Medio-Lateral Oblique, MLO)投照,必要时加拍补充投照位。乳腺机采用全自动曝光,根据乳腺厚度和密度全自动选择所有成像参数。最终得到CC位图像347幅,MLO位图像347幅,共计694幅图像。

1.2 乳腺X线图像上微钙化灶标注

低年资医师(资质不足5年)利用3D Slicer软件,首先选取病灶区域,然后使用大津(OTSU)算法,通过调节阈值获得精准的标注结果。低年资医生标注完成之后,由高年资医师(10年及以上乳腺X线诊断经验)进行审核,对不合格的标注图像进行重新标注,直到最后通过审核。图1和图2展示了乳腺微钙化灶的标注分割及面积数量的量化。钙化面积以像素表示,每个像素大小为70 μm×70 μm。钙化数量由分割后得到的连通域数量计算得出。在标注过程中,放射科医师可通过调整窗宽窗位进行观察,并且标注时间不受限制。694幅乳腺X线图像按照594:100随机划分为训练集和测试集,进行微钙化检测模型训练。为了验证模型在不同乳腺密度中微钙化的检出效果,本研究按照ACR 2013年第5版BI-RADS标准对乳腺密度进行分类,其中训练集中a+b类图像57幅,c+d类图像537幅;测试集中a+b类图像19幅,c+d类图像81幅。

1.3 基于 Attention U-Net 的乳腺X线图像微钙化检测模型

本研究使用的核心算法为Attention U-Net^[9](图3),是一种基于Attention Gate (AG)的改良U-Net模型^[11],主要应用于医学影像自动分割。U-Net是基于全卷积网络(Fully Convolutional Network, FCN)^[12]的一种改进模型,其网络结构呈U型,包含下采样和上采样两部分结构,这两部分结构也被称为收缩路径和扩张路径。收缩路径为经典的VGG16网络的结

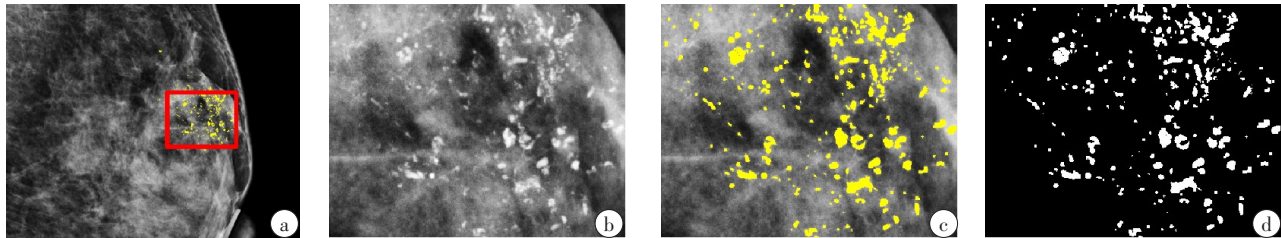


图1 乳腺X线图像微钙化灶标注及分割效果图

Figure 1 Labeling and segmentation of breast microcalcifications in mammogram

a: 47岁女性,左侧内外斜位图像,病理证实为浸润性癌;b:微钙化局部放大图;c:微钙化融合显示图(黄色代表微钙化组织);d:微钙化分割二值图

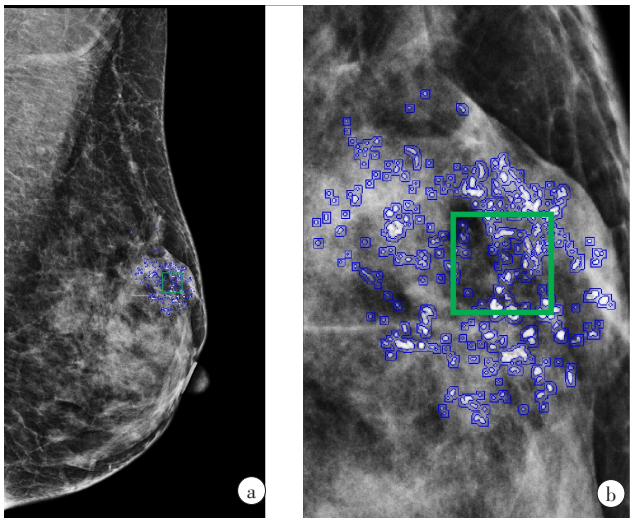


图2 乳腺X线图像微钙化灶量化示意图(a)及其放大图(b)

Figure 2 Quantization of breast microcalcifications (a) and the partial zoom-in of breast microcalcifications (b) in mammogram

钙化数量:257个,单位面积(绿色框)钙化数量:47个。钙化总面积:2680个像素。钙化面积<10个像素的钙化灶数量:131个;10个像素<钙化面积<20个像素的钙化灶数量:75个;钙化面积>20个像素的钙化灶数量:51个。单位像素大小:70 μm×70 μm,钙化区域的灰度值:2984,钙化周边灰度值(蓝色框):1945

构^[13],使用3×3的卷积核,卷积步长为1的卷积层,每两次卷积后,接一个2×2的最大池化层,来缩小特征图的尺寸。在扩张路径中,上采样时,基于前一个网络层通过反卷积来实现,并将上采样后的结果同对应的下采样阶段的特征图进行融合。相较于FCN, U-Net得到的网络特征更加丰富,同时它仅需要一次训练,便可以获得精细的分割结果,无需分阶段训练。Attention Gate在推理时会自动关注有显著特征的区域,在分割时先确定待分割的器官或组织的感兴趣区(Region of Interest, ROI),再进行小区域的分割。Attention Gate可通过抑制无关区域中特征的激活来提高模型敏感度和准确性,以进行密集的标志预测,并且可以在维持高准确率的同时免除使用外部器官定位模型的必要性。乳腺X线图像微钙化检测模型算法流程分为训练阶段和测试阶段,见图4。

训练集数据采用imgaug库(<https://github.com/aleju/imgaug>)进行增强,设置的是随机增强方式,每张训练图像增强为4张,从而将训练数据集提升到594×4幅。训练阶段使用增广后的训练集数据进行

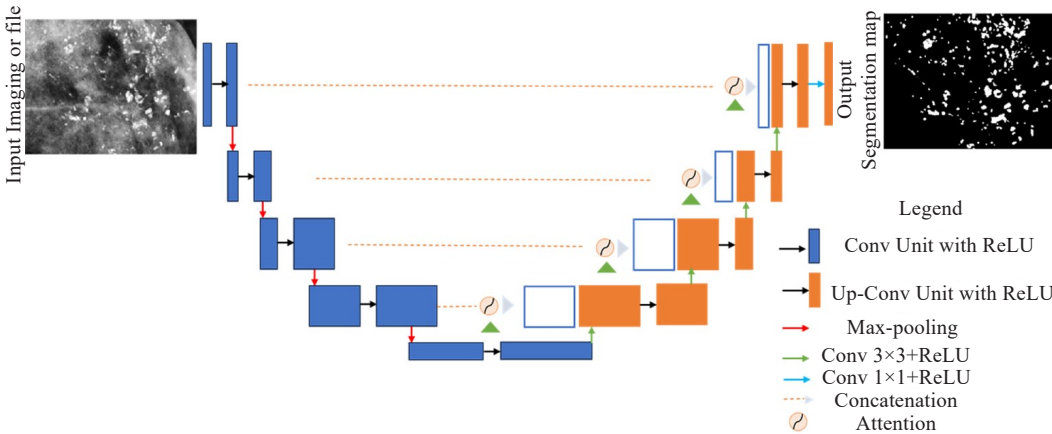


图3 乳腺X线图像微钙化检测模型架构图

Figure 3 Architecture diagram of the mammogram microcalcification detection model

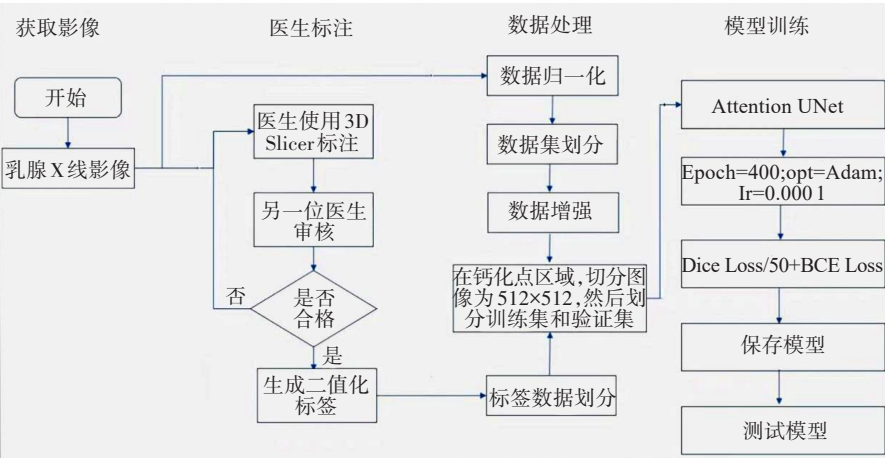


图4 乳腺X线图像微钙化检测模型实验流程图

Figure 4 Experimental flowchart of the mammogram microcalcification detection model

训练,训练参数:Epoch=400,学习率=0.000 1,优化器为 Adam,损失函数 Dice Loss 和 BCE Loss 见式(1)、式(2):

$$\text{Dice Loss}=1-(2\cdot|A\cap B|)/(|A|+|B|)$$
 (1)

$$\text{BCE Loss}=-[y\log(p)+(1-y)\log(1-p)]$$
 (2)

Dice Loss 是一种用于图像分割任务的损失函数,它通过计算模型预测结果与真实标签的相似度来衡量损失^[14]。 A 表示预测结果的二值化结果, B 表示真实标签的二值化结果, $|A|$ 和 $|B|$ 分别表示它们的总面积, $|A\cap B|$ 表示它们的重叠面积。Dice系数为0~1,1表示完全重叠,0表示无重叠。BCE Loss 是交叉熵损失函数的一种特例,它可以衡量模型预测结果与真实标签之间的差异,只应用在二分类任务中^[15]。 y 表示真实标签, p 表示模型预测的概率值。本研究中,Dice Loss 和 BCE Loss 混合使用,Total loss=Dice Loss/50+BCE Loss。

训练过程中以模型检测的钙化面积、人工标注的钙化面积差异计算精确率(Precision)、召回率(Recall)、交并比(IoU)、F1 分数(F1-score),4个指标变化趋势见图5。

本研究中,乳腺X线图像微钙化检测模型使用的是五折交叉验证技术,训练集在训练过程中被分为5份,进行5次训练得到5个模型,训练完后在测试集上进行测试,对于5个模型采用求和平均的方式输出测试集的结果,将检测出来的结果与人工标注的结果进行比较用来评估微钙化检测的性能。

1.4 评价指标

本研究使用精确率、召回率(敏感性)、IoU、F1 分数等指标对模型微钙化检测性能进行评价。以上4个指标采用钙化面积及钙化数量分别计算。见式

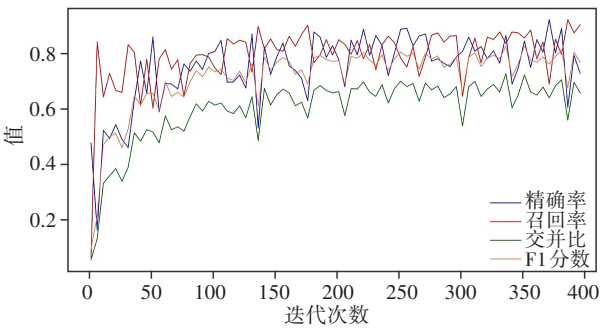


图5 精确率、召回率、交并比、F1分数在训练时(以钙化面积计算)的变化趋势图

Figure 5 Trend plots of precision rate, recall rate, IoU, and F1-score during training (calculation based on calcification area)

(3)~式(6):

$$\text{Precision} = \frac{TP}{TP + FP}$$
 (3)

$$\text{Recall} = \frac{TP}{TP + FN}$$
 (4)

$$\text{IoU} = \frac{TP}{TP + FN + FP}$$
 (5)

$$\text{F1 - score} = \frac{2TP}{2TP + FN + FP}$$
 (6)

其中,真阳性(True Positive, TP)表示模型正确检测的钙化灶面积(或数量);假阳性(False Positive, FP)表示模型错误检测的钙化灶面积(或数量);假阴性(False Negative, FN)表示模型未检测出的钙化灶面积(或数量)。精确率代表模型正确检测的钙化灶面积(或数量)占模型检测出的所有钙化灶面积(或数量)的比例。召回率代表模型正确检测的钙化灶面积(或数量)占人工标记的钙化灶面积(或数量)的比例。IoU指模型正确检测的钙化灶面积(或数量)占

人工标记的钙化灶面积(或数量)、模型错误检测的钙化灶面积(或数量)总和的比例。F1分数指模型的精确率和召回率的调和平均数,最大为1,最小为0,数值越大说明模型质量越高。

1.5 统计学方法

采用SPSS 26.0软件进行统计学分析。连续变量用均数±标准差表示,用独立样本*t*检验方法;分类变量用例(%)表示,用卡方检验(χ^2)方法。 $P<0.05$ 为差异有统计学意义。

2 结果

2.1 研究对象的一般临床特征

347例(694幅图像)女性患者的一般临床特征见表1。在训练集良性病例组和恶性病例组中,月经状态、乳腺密度、BI-RADS分类差异均有统计学意义($P<0.001$)。在测试集良性病例组和恶性病例组中,月经状态、BI-RADS分类差异有统计学意义($P<0.05$)。

表1 研究对象的一般临床特征[例(%)]
Table 1 Clinical characteristics of the subjects [cases (%)]

分类	训练集			测试集		
	良性(<i>n</i> =226)	恶性(<i>n</i> =368)	<i>P</i> 值	良性(<i>n</i> =44)	恶性(<i>n</i> =56)	<i>P</i> 值
月经状态			<0.001			0.022
绝经前	151(66.8)	158(42.9)		29(65.9)	24(42.9)	
绝经后	75(33.2)	210(57.1)		15(34.1)	32(57.1)	
乳腺密度			<0.001			0.084
a+b类	5(2.2)	52(14.1)		5(11.4)	14(25.0)	
c+d类	221(97.8)	316(85.9)		39(88.6)	42(75.0)	
BI-RADS分类			<0.001			<0.001
2	3(1.3)	8(2.2)		1(2.3)	0(0.0)	
3	8(3.5)	0(0)		2(4.5)	0(0.0)	
4	215(95.1)	221(60.1)		41(93.2)	37(66.1)	
5	0(0.0)	139(37.8)		0(0.0)	19(33.9)	

2.2 深度学习模型微钙化检测性能

在测试集中,分别以钙化面积和钙化数量计算得到的精确率、召回率,两组差异有统计学意义

($P<0.05$);而交并比、F1分数两组差异没有统计学意义($P>0.05$),见表2。

表2 不同方式计算的评价深度学习模型微钙化检测性能的指标比较(%)
Table 2 Comparison of different metrics to evaluate the performance of deep learning model in microcalcification detection (%)

组别	精确率	召回率	交并比	F1分数
以钙化面积计算	85.12±18.39	78.18±19.25	68.29±21.39	78.96±17.70
以钙化数量计算	76.72±19.85	85.12±18.39	67.13±23.84	77.65±9.37
<i>P</i> 值	0.002	0.014	0.717	0.619

以钙化面积计算,在不同乳腺密度中(a+b类 vs c+d类),精确率、召回率、交并比、F1分数差异无统计学意义($P>0.05$)。在不同钙化性质中(良性 vs 恶性),精确率、召回率、交并比、F1分数差异无统计学意义($P>0.05$),具体见表3。

以钙化数量计算,在不同乳腺密度中(a+b类 vs c+d类),精确率、召回率、交并比、F1分数差异无统计学意义($P>0.05$)。在不同钙化性质中(良性 vs 恶性),精确率、召回率、交并比、F1分数差异无统计学意义($P>0.05$),具体见表3。

表3 测试集中深度学习模型微钙化检测性能比较(%)

Table 3 Comparison of the performance of deep learning model in detecting different breast densities and microcalcification types (%)

指标	乳腺密度			钙化性质		
	a+b类	c+d类	P值	良性	恶性	P值
以钙化面积计算						
精确率	87.86±13.46	84.47±19.38	0.473	85.20±20.34	85.05±16.56	0.968
召回率	84.19±20.61	76.77±18.77	0.131	79.11±18.36	77.44±20.05	0.669
交并比	76.49±21.61	66.37±21.01	0.063	68.52±21.85	68.12±21.22	0.925
F1分数	84.72±17.11	77.61±17.67	0.115	79.01±18.52	78.92±17.20	0.981
以钙化数量计算						
精确率	79.88±19.11	75.99±20.06	0.444	78.94±20.44	74.99±19.38	0.326
召回率	87.62±16.30	84.86±22.95	0.622	84.67±22.66	85.95±21.28	0.772
交并比	72.17±21.17	65.95±24.39	0.309	68.76±25.32	65.85±22.75	0.547
F1分数	82.05±15.41	76.62±20.13	0.274	78.57±20.36	76.93±18.17	0.677

2.3 不同乳腺X线图像微钙化检测模型的效能对比

为了验证本文中基于 Attention U-Net的乳腺X线图像微钙化检测模型的效能,将该模型结果与近年来其他学者的研究结果进行对比。由于文献[6-8, 16-17]报道中乳腺微钙化检测模型的性能评估主要是基于模型检测到的微钙化区域样本个数,因此本文另外随机选取在福建医科大学附属泉州第一医院影像科行乳腺X线检查患者的100幅图像,包括微钙化图像46幅、无钙化(正常)图像54幅,采取该方法计算相应评价指标并进行比较,其中本文模型的精确率为97.78%、召回率为95.65%,与近年来其他学者的研究结果进行对比^[6-8, 16-17],本文模型的精确率与召回率总体上具有优势(表4)。

表4 不同乳腺X线图像微钙化检测模型的效能对比(%)

Table 4 Comparison among different mammogram microcalcification detection models (%)

模型	精确率	召回率
本文模型	97.78	95.65
Cai等 ^[6]	89.32	88.43
邢素霞等 ^[7]	98.05	94.36
申楠等 ^[8]	92.35	91.75
Sannasi Chakravarthy等 ^[16]	86.78	90.08
Rehman等 ^[17]	85.50	97.00

3 讨论

本研究开发了基于 Attention U-Net的乳腺X线

图像微钙化检测模型,并基于模型检测到的微钙化面积、数量来评价模型的性能,笔者发现该模型对微钙化检测的精确率、召回率、交并比、F1分数均较理想,表明该模型能够对乳腺微钙化进行有效检测。同时本研究还评估了该模型在不同性质钙化中、不同乳腺密度中对微钙化的检测性能,其精确率、召回率、交并比、F1分数之间的差异无统计学意义,表明微钙化的性质、不同的乳腺密度并不影响该模型对乳腺微钙化的检测性能。

乳腺微钙化的检测及性质判断不管对诊断医师还是对机器学习而言都是一项极具挑战性的任务。本研究中乳腺X线影像像素大小是2560×3328,但是钙化灶区域面积占比在大多数图像中均低于1%,导致图像分割任务极度不均衡。同时本研究中训练集数据量偏少,只有594幅图像。为解决这些问题,本研究对训练集数据进行增强。数据增强能有效的扩增训练数据集,通过数据增广获得大量的训练数据,从而缓解数据集较小而网络泛化性不足的问题。对于乳腺微钙化图像而言,数据增强并不会改变其特征,反而会使得算法对乳腺微钙化的位置信息不再特别敏感,更多的关注到乳腺微钙化病变区域同周边组织的差异,有利于微钙化的识别。在数据训练过程中,精确率、召回率、交并比、F1分数的变化波动幅度相对平稳,得益于本研究使用Dice Loss和BCE Loss混合函数进行训练,防止了因梯度变化剧烈而导致训练困难。BCE Loss+Dice Loss混合函数结合2种不同的损失函数,综合考虑了像素级别的分类准确性(通过BCE Loss表示)及分割结果的相似性(通过Dice Loss表示)。BCE Loss可以在训练早期加快

模型的收敛速度,而Dice Loss可以促使模型生成更平滑和连续的分割结果^[14-15]。本研究中基于Attention U-Net的乳腺X线图像微钙化检测模型精确率为97.78%(基于微钙化区域样本个数)、召回率为95.65%(基于微钙化区域样本个数),与近年来其他学者的研究结果进行对比^[6-8, 16-17],本文模型的精确率与召回率总体上具有优势。既往文献报道中乳腺微钙化检测模型的性能评估主要是基于模型检测到的微钙化区域样本个数^[6-8, 16-17],而本文主要采用模型检测到的微钙化的面积、数量来评价模型的性能,计算得出模型的精确率为85.12%(基于微钙化的面积)和76.72%(基于微钙化的数量),召回率为78.18%(基于微钙化的面积)和85.12%(基于微钙化的数量),均较理想,有助于后续乳腺微钙化的定量临床应用研究。

微钙化是乳腺病变中最常见的表现形式,在乳腺癌的早期发现中起着重要作用,临床上有30%~50%触诊阴性却最终诊断为乳腺癌的病例可仅表现为微钙化灶^[4]。根据成份的不同,乳腺微钙化可分为由草酸钙构成的I型微钙化以及由羟基磷灰石复合物构成的II型微钙化。I型微钙化仅见于乳腺良性病变,而II型微钙化既可以见于良性病变也可以见于恶性病变^[18]。虽然乳腺X线图像无法准确区分I型微钙化和II型微钙化,但是BI-RADS分类却可以从钙化的密度、形态、分布等方面预测钙化的良恶性。拟恶性钙化一般呈簇状、线样分布,形态上多表现为细点状、针尖状或无定形,在乳腺X线图像上容易漏诊。Mordang等^[19]报道了恶性钙化在乳腺X线图像上的检出敏感性只有45.5%。计算机辅助检测系统可提高无定形钙化等拟恶性钙化的检出敏感性。王小琦等^[20]研究中,计算机辅助检测系统对可疑钙化的检出敏感性为99.64%,显著高于年轻医生。Basile等^[21]提出一种新的全自动乳腺X线图像钙化检测模型,该模型对BI-RADS 3类钙化的检出敏感性最低(81.15%),而对BI-RADS 2类钙化的检出敏感性最高(98.26%),BI-RADS 4类次之(94.92%)。周娟等^[22]使用北京深睿博联科技有限责任公司开发的Dr.Wise乳腺X线影像辅助诊断系统,对解放军总医院第五医学中心的数据进行评估,发现人工智能辅助检测系统对BI-RADS ≥ 3 类的钙化检出敏感性均达到100%,而医师对于无定形钙化漏检率较高。上述研究采用BI-RADS分类对钙化性质进行预测分类,表明计算机辅助检测系统对乳腺X线图像上钙化的检出率不受BI-RADS分类影响。然而上述研究缺乏病理结果对照,而本研究乳腺微钙化均有病理结果。本研究开发的基于Attention U-Net的乳腺X线图像

微钙化检测模型对乳腺微钙化的检测性能与钙化性质无关,兼顾精确率与敏感性,F1分数较高,大于75%。

乳腺因包含的纤维腺体组织与脂肪组织比例不同,X线影像下显影密度存在差异。乳腺密度通常指放射线密度,一般以X线图像上放射致密组织所占的比重来量化^[23]。乳腺密度与年龄、激素使用情况、体重、遗传等因素相关。在美国大约50%的乳腺X线检查者为致密型乳腺^[24]。乳腺密度是乳腺癌的独立风险因子,致密型乳腺的患癌风险是非致密型乳腺的4~6倍^[25]。同时,乳腺密度可影响乳腺癌的检出率,较小的乳腺肿块、腺体结构扭曲及浅淡钙化灶等病变容易被高密度背景所掩盖,导致乳腺癌风险被低估^[26]。既往研究发现在致密型乳腺中X线检查对乳腺癌检出敏感性仅为62.2%~68.1%^[27]。近几年来,乳腺人工智能检测系统被用于乳腺X线图像中帮助医生发现乳腺病变,如肿块、微钙化、结构扭曲等。既往文献在乳腺密度对人工智能检测系统的影响方面进行深入的研究。周娟等^[22]发现乳腺密度对Dr.Wise乳腺X线影像辅助诊断系统的钙化检出率无影响。宋张骏等^[28]使用西安百利信息科技有限公司开发的MammoWorks™乳腺X线图像智能诊断系统,发现MammoWorks™ 3.5.2.6版及更新后的3.5.4.43版对不同乳腺密度的图像假阳性标记及假阴性标志的差异具有统计学意义。本研究开发的基于Attention U-Net的乳腺X线图像微钙化检测模型对乳腺微钙化的检测性能与乳腺密度无关,不管在致密型还是非致密型乳腺中微钙化检测的敏感性均高于80%。人工智能辅助检测系统对致密型乳腺中微钙化的检测具有优势,可提高乳腺微钙化病灶检出率。

本研究有以下几个局限性。第一,这是一项单中心回顾性研究,笔者纳入了有限数量的微钙化病例($n=347$),而且a+b类病人较少,可能存在选择偏倚。第二,大部分乳腺钙化呈现为点状,没有清晰的边界,而且早期微钙化很难被资历较浅的医生观察到,容易导致误标注和漏标注。本文采用低年资医师标注,高年资医师审核的方式,尽量避免了误标注和漏标注的出现。第三,采用手动方式标注乳腺微钙化,存在耗时、耗力以及主观性的问题。未来笔者将采用半监督学习的标注方式对多中心大样本乳腺数据进行更深入的分析研究,完善算法模型,并将该模型运用于临床实践当中,帮助医生发现隐蔽的微钙化灶,通过微钙化面积及数量的量化,对比其治疗前后、随访过程中的细微变化。

4 结 论

本研究建立了基于 Attention U-Net的乳腺X线图像微钙化检测模型,该模型能够对乳腺微钙化进行有效的检测,有助于乳腺微钙化的定量研究,同时该模型稳定性强,乳腺密度及微钙化性质对该模型的检测性能无影响。

【参考文献】

- [1] Zheng RS, Zhang SW, Zeng HM, et al. Cancer incidence and mortality in China, 2016[J]. J Natl Cancer Cent, 2022, 2(1): 1-9.
- [2] Miller KD, Nogueira L, Devasia T, et al. Cancer treatment and survivorship statistics, 2022 [J]. CA Cancer J Clin, 2022, 72(5): 409-436.
- [3] Ren WH, Chen MY, Qiao YL, et al. Global guidelines for breast cancer screening: a systematic review[J]. Breast, 2022, 64: 85-99.
- [4] Karale VA, Ebenezer JP, Chakraborty J, et al. A screening CAD tool for the detection of microcalcification clusters in mammograms[J]. J Digit Imaging, 2019, 32(5): 728-745.
- [5] Chan YS, Hung WK, Yuen LW, et al. Comparison of characteristics of breast cancer detected through different imaging modalities in a large cohort of Hong Kong Chinese women: implication of imaging choice on upcoming local screening program[J]. Breast J, 2022, 2022: 3882936.
- [6] Cai HM, Huang QJ, Rong WT, et al. Breast microcalcification diagnosis using deep convolutional neural network from digital mammograms [J]. Comput Math Methods Med, 2019, 2019: 2717454.
- [7] 邢素霞, 申楠, 刘子骄, 等. 基于空-频域差值与超分辨率反馈网络的乳腺微钙化区域检测技术[J]. 中国医学物理学杂志, 2022, 39(7): 840-849.
Xing SX, Shen N, Liu ZJ, et al. Space-frequency domain difference with super-resolution feedback network for breast microcalcification detection [J]. Chinese Journal of Medical Physics, 2022, 39(7): 840-849.
- [8] 申楠, 邢素霞, 何湘萍, 等. 基于Adaboost-决策树算法的乳腺微钙化区域真假阳性检测[J]. 中国医学物理学杂志, 2021, 38(8): 940-945.
Shen N, Xing SX, He XP, et al. True-and false-positive detections of breast microcalcifications based on Adaboost-decision tree algorithm [J]. Chinese Journal of Medical Physics, 2021, 38(8): 940-945.
- [9] Oktay O, Schlemper J, Folgoc LL, et al. Attention U-net: learning where to look for the pancreas [EB/OL]. (2018-05-20). <https://arxiv.org/abs/1804.03999>.
- [10] 中华医学会影像技术分会, 中华医学会放射学分会. 乳腺影像检查技术专家共识[J]. 中华放射学杂志, 2016, 50(8): 561-565.
Imaging Technology Branch of Chinese Medical Association, Radiology Branch of Chinese Medical Association. Expert consensus on breast imaging [J]. Chinese Journal of Radiology, 2016, 50(8): 561-565.
- [11] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [12] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(4): 640-651.
- [13] Liu SY, Deng WH. Very deep convolutional neural network based image classification using small training sample size[C]//2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). Piscataway, NJ, USA: IEEE, 2015: 730-734.
- [14] Milletari F, Navab N, Ahmadi SA. V-net: fully convolutional neural networks for volumetric medical image segmentation [EB/OL]. (2016-06-15). <https://arxiv.org/abs/1606.04797>.
- [15] Creswell A, Arulkumaran K, Bharath AA. On denoising autoencoders trained to minimise binary cross-entropy [EB/OL]. (2017-10-09). <https://arxiv.org/abs/1708.08487>.
- [16] Sannasi Chakravarthy SR, Rajaguru H. Detection and classification of microcalcification from digital mammograms with firefly algorithm, extreme learning machine and non-linear regression models: a comparison[J]. Int J Imaging Syst Technol, 2020, 30(1): 126-146.
- [17] Rehman KU, Li JQ, Pei Y, et al. Computer vision-based microcalcification detection in digital mammograms using fully connected depthwise separable convolutional neural network [J]. Sensors (Basel), 2021, 21(14): 4854.
- [18] Tot T, Gere M, Hofmeyer S, et al. The clinical value of detecting microcalcifications on a mammogram [J]. Semin Cancer Biol, 2021, 72: 165-174.
- [19] Mordang JJ, Gubern-Mérida A, Bria A, et al. The importance of early detection of calcifications associated with breast cancer in screening[J]. Breast Cancer Res Treat, 2018, 167(2): 451-458.
- [20] 王小琦, 刘鹏, 陈贻, 等. 基于深度学习的乳腺X线摄影钙化检测系统对乳腺可疑钙化的检出效能[J]. 中国医学影像技术, 2019, 35(12): 1784-1788.
Wang XQ, Liu P, Chen Z, et al. Efficiency of mammography detection system based on deep learning for breast suspicious calcifications[J]. Chinese Journal of Medical Imaging Technology, 2019, 35(12): 1784-1788.
- [21] Basile TM, Fanizzi A, Losurdo L, et al. Microcalcification detection in full-field digital mammograms: a fully automated computer-aided system[J]. Phys Med, 2019, 64: 1-9.
- [22] 周娟, 王婷婷, 李明, 等. 基于深度学习的乳腺X线摄影钙化检出系统评估[J]. 中华放射学杂志, 2019, 53(11): 968-973.
Zhou J, Wang TT, Li M, et al. Evaluation of mammography calcification detection system based on deep learning [J]. Chinese Journal of Radiology, 2019, 53(11): 968-973.
- [23] Nazari SS, Mukherjee P. An overview of mammographic density and its association with breast cancer[J]. Breast Cancer, 2018, 25(3): 259-267.
- [24] Wengert GJ, Helbich TH, Kapetas P, et al. Density and tailored breast cancer screening: practice and prediction - an overview[J]. Acta Radiol Open, 2018, 7(9): 2058460118791212.
- [25] Archer M, Dasari P, Evdokiou A, et al. Biological mechanisms and therapeutic opportunities in mammographic density and breast cancer risk[J]. Cancers (Basel), 2021, 13(21): 5391.
- [26] Bell RJ. Mammographic density and breast cancer screening [J]. Climacteric, 2020, 23(5): 460-465.
- [27] Vourtsis A, Berg WA. Breast density implications and supplemental screening[J]. Eur Radiol, 2019, 29(4): 1762-1777.
- [28] 宋张骏, 王虎霞, 赵静, 等. 基于深度学习技术的乳腺智能检测系统在触诊阴性乳腺肿瘤诊断中的应用[J]. 中国临床研究, 2021, 34(3): 323-328.
Song ZJ, Wang HX, Zhao J, et al. Application of breast intelligent detection system based on deep learning technology in palpation negative breast tumor diagnosis [J]. Chinese Journal of Clinical Research, 2021, 34(3): 323-328.

(编辑: 薛泽玲)