

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

邢素霞¹, 申楠¹, 刘子骄¹, 鞠子涵¹, 何湘萍², 潘子妍¹, 王瑜¹

1. 北京工商大学人工智能学院, 北京 100048; 2. 北京海淀妇幼保健院乳腺科, 北京 100080

【摘要】乳腺癌的早期症状在乳腺钼靶图像中主要表现为微钙化点,微钙化区域的真假阳性检测对于乳腺癌早期筛查具有重要意义。首先,对DDSM乳腺数据集中的图像进行预处理,去除噪声及无关组织干扰;其次,基于空-频域差值图像技术实现了疑似微钙化点的分割,取得的敏感性为91.00%,但假阳性率也较高(34.00%),并根据疑似点的质心位置自动截取感兴趣区域;然后,通过超分辨率反馈网络算法进行微钙化区域超分辨率重建;最后,提取感兴趣区域的纹理特征,将Gentle AdaBoost算法和单层决策树算法相结合,构建强分类器GAB-DS对区域进行分类,将微钙化区域和正常组织分离开来,GAB-DS分类模型取得了96.25%的准确率、94.38%的敏感性以及98.13%的特异性。实验结果表明,该模型在微钙化区域检测上性能优越,可用于辅助临床乳腺癌检测及诊断,具有一定的临床应用价值。

【关键词】乳腺癌;微钙化区域;差值图像技术;超分辨率重建;GAB-DS

【中图分类号】R318;TP391.4

【文献标志码】A

【文章编号】1005-202X(2022)07-0840-10

Space-frequency domain difference with super-resolution feedback network for breast microcalcification detection

XING Suxia¹, SHEN Nan¹, LIU Zijiao¹, JU Zihan¹, HE Xiangping², PAN Ziyen¹, WANG Yu¹

1. School of Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China; 2. Department of Breast, Haidian District Maternal and Child Health Care Hospital, Beijing 100080, China

Abstract: The early symptoms of breast cancer are mainly characterized by microcalcifications in mammogram. The true and false positive detections of microcalcifications are of great significance for the early screening of breast cancer. After preprocessing the images from DDSM breast data for removing noise and irrelevant tissue interference, the suspected microcalcifications are segmented based on the space-frequency domain difference image technique, obtaining a sensitivity of 91.00% and a false positive rate of 34.00%. According to the centroid position of the suspected point, the region of interest is automatically intercepted. Then, super-resolution feedback network algorithm is used to realize the super-resolution reconstruction of the microcalcifications. Finally, the texture features of the region of interest are extracted, and the Gentle AdaBoost algorithm is combined with the single layer decision tree algorithm to construct a strong classifier GAB-DS for classifying the regions, separating the microcalcifications from the normal tissues. The accuracy, sensitivity and specificity achieved by GAB-DS model were 96.25%, 94.38% and 98.13%, respectively. The experimental results show that the model has superior performance in detecting microcalcifications, and can be used to assist in clinical detection and diagnosis of breast cancer, with high application value in clinic.

Keywords: breast cancer; microcalcification; difference image technique; super-resolution reconstruction; GAB-DS

前言

乳腺癌作为一种高发病率和高死亡率的恶性肿瘤,已是危害全球女性健康甚至生命的头号杀手。

世界卫生组织癌症专家于2021年2月2日表示,2020年全球1 000万人死于癌症,罹患癌症的人数达到1 930万^[1]。其中,乳腺癌病例激增226万例,占新增癌症病例的11.7%,首次超越肺癌,成为世界上最高发的癌症。相关统计数据预测,至2030年,全球范围内乳腺癌的发病数量将会达到1.1亿^[2],这一严峻形势引起了广泛关注。全球范围内,乳腺癌的有关研究尚处于探索阶段,发病原因尚未查明。定期进行乳腺癌检查以及发现后及时治疗,是降低发病率和死亡率、提升患者治愈率的最有效方法。乳腺癌早

【收稿日期】2022-01-20

【基金项目】国家自然科学基金(61671028);北京市自然科学基金-北京市教育委员会科技计划重点联合项目(KZ202110011015)

【作者简介】邢素霞,博士,副教授,主要研究方向:图像处理、嵌入式系统,E-mail: xingsuxia@163.com

期在乳腺钼靶图像中主要表现为微钙化区域,因此,能够检测出乳腺钼靶图像中的微钙化区域,对于乳腺癌筛查具有重要意义。微钙化点的大小一般为50~500 μm ,特别是恶性钙化点,极其细小,即便是经验丰富的医生仅凭肉眼也难以识别。在计算机检测中,也很容易把微钙化点误认为是噪声漏检,或是因检测精度较低而无法识别^[3]。在0.1 mm/像素的数字钼靶图像中,直径小于100 μm 的微钙化点无法显示,只有更大的微钙化点可能出现在图像中。此外,部分乳腺腺体组织与微钙化区域的亮度相近。上述这些特点使得乳腺微钙化区域的自动检测较为困难。因此,计算机辅助诊断(Computer Aided Diagnosis,

CAD)系统在乳腺钼靶图像微钙化区域的检测及分类上仍存在假阳性率(False Positive Rate, FPR)过高等问题,制约着CAD系统在临床上的应用与推广,如何科学有效地改进乳腺癌CAD技术成为解决问题的关键。

尽管近年来针对乳腺钼靶图像钙化区域的检测研究取得了一定进展,但仍然普遍存在着检测结果FPR较高、检测效率较低等不足。针对以上问题,本文提出了检测疑似微钙化点的空-频域差值方法,以及超分辨率重建技术,通过对超分放大后的感兴趣区域(Region of Interest, ROI)特征提取和分类,实现微钙化区域的计算机辅助检测,如图1所示。

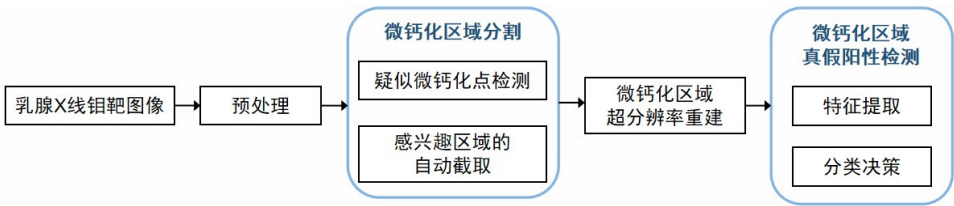


图1 乳腺癌CAD系统流程图
Figure 1 Flowchart of breast cancer CAD system

1 国内外研究现状

簇状钙化灶是乳腺癌的早期征兆,因此准确检测出乳腺钼靶图像中的微钙化点簇能极大程度地提高早期乳腺癌的确诊率,而分割出乳腺钼靶图像中微钙化点簇则是微钙化点簇检测的关键步骤。由于乳腺钙化区域检测存在很多难点,近几年来受到了越来越多学者的关注,国内外学者对乳腺微钙化点分割进行了深入研究,提出了各具特色的分割方法。梅明^[4]提出了将边缘点检测算法与水平集算法相结合的微钙化点分割模型,首先使用边缘点检测算法对微钙化点进行粗分割,然后运用水平集算法进一步细化分割,并且根据边缘点搜索的数据拟合出一个圆形,用作活动轮廓模型的初始轮廓。Ciecholewski^[5]首先从形态学上检测微钙化区域,确定发生微钙化的大致位置,然后对微钙化点进行分水岭分割,对 Digital Database for Screening Mammography (DDSM)数据集中200个512×512像素的ROI进行测试,得到了80.5%的相似性指数及75.7%的重叠率。Basile等^[6]提出一种分三阶段的聚类微钙化检测方法,首先进行图像预处理,然后基于Hough变换检测微钙化点的边缘,最后对微钙化点进行聚类识别,实现微钙化区域的检测,对364张乳腺钼靶图像进行检测,真阳性率为91.78%,FPR为2.87%。

2 实验方法

本文选用DDSM数据库的图像进行实验^[7]。DDSM数据库是由南佛罗里达大学于1999年提供的数字化乳腺图像库,该数据库是一个高分辨率的乳腺钼靶图像标准数据库,含2620个病例。本文实验的硬件环境为Intel(R) Core(TM) i7-9700K CPU @ 3.60 GHZ, 16 G内存;软件环境为Windows10操作系统, MATLAB R2019a和PyCharm 2019.1。

2.1 图像预处理

图像预处理是乳腺癌CAD系统的基础步骤,其目的在于提升微钙化点和背景组织间的对比度以及提高算法的运行速度。图像预处理模块包含灰度归一化、图像边缘裁剪、图像增强、去除标签噪声和乳腺区域自动分割5大步骤。经过图像预处理,去除了图像的标签噪声和多余结构干扰,利于后续微钙化区域的检测,提高了运算速度。整个预处理过程产生的图像如图2所示。

2.2 基于空-频域差值图像技术的ROI自动分割

从已有的研究成果来看,单一的检测方法并不能取得很好的结果,因此本文提出了空-频域组合的微钙化区域检测模型。文献[8]中,在进行图像差值处理后,采用区域生长算法进行疑似微钙化点的分割,而区域生长算法需要手动选择种子点,过程复杂。本文将区域生长算法替换为阈值分割,并增添

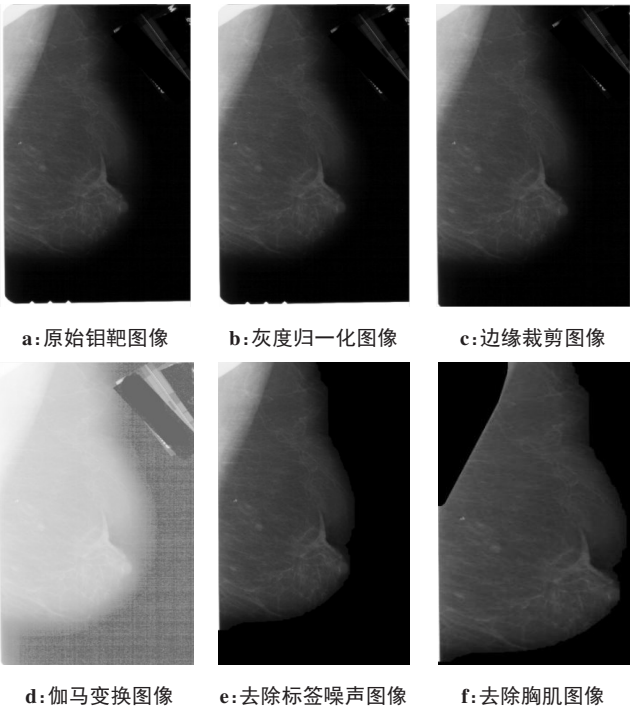


图2 预处理图像
Figure 2 Image preprocessing

了“找到疑似点的质心”和“画出最小外接矩形”两个步骤,不仅实现了疑似微钙化点的分割,还实现了钼靶图像上疑似微钙化点检测。

2.2.1 空-频域差值图像技术 从空间域角度来看,微钙化区域在乳腺钼靶图像中较周围正常组织亮;从频域角度看,微钙化点是淹没于极高频噪声和低频背景中的高频信号。因此,从空间域和频域的角度出发来对噪声进行消除,实现优势互补,获得更佳的效果。流程如图3所示。

(1)空域锐化。若去除胸肌后的钼靶图像为 $f(x,y)$,设 $\nabla^2 f$ 为拉普拉斯算子, $f(x,y)$ 的拉普拉斯变

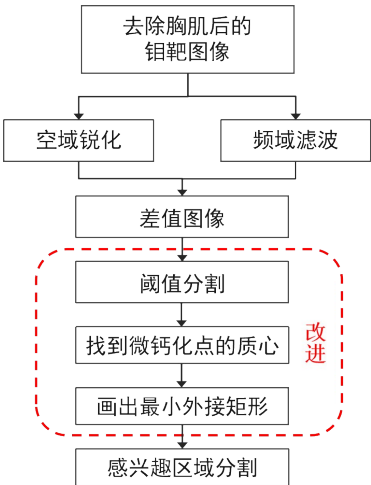


图3 疑似微钙化点的检测流程图
Figure 3 Flowchart of suspected microcalcification detection

换如下^[9]:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{1}$$

在数字图像处理中,需将式(1)转换为离散形式,由于式中含 x 和 y 两个变量,因此分别在 x 方向和 y 方向上进行二阶偏微分,得到:

$$\frac{\partial^2 f}{\partial x^2} = f(x+1,y) + f(x-1,y) - 2f(x,y) \tag{2}$$

$$\frac{\partial^2 f}{\partial y^2} = f(x,y+1) + f(x,y-1) - 2f(x,y) \tag{3}$$

式(1)中的二元图像 $f(x,y)$ 拉普拉斯变换可由两个分量相加得到:

$$\begin{aligned} \nabla^2 f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \\ &= [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)] - 4f(x,y) \end{aligned} \tag{4}$$

本文所使用的拉普拉斯算子为8邻域,它的增强效果更加显著,表示为:

$$\begin{aligned} \nabla^2 f &= [f(x-1,y-1) + f(x,y-1) + f(x+1,y-1) + f(x-1,y) \\ &\quad + f(x+1,y) + f(x-1,y+1) + f(x,y+1) + f(x+1,y+1)] - 8f(x,y) \end{aligned} \tag{5}$$

根据上式,可以得到本文所用的8邻域拉普拉斯算子的差分示意图(图4)和模板示意图(图5)。

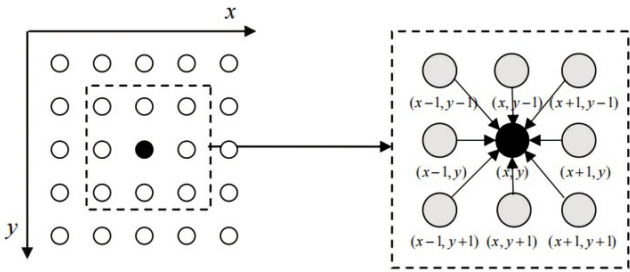


图4 拉普拉斯算子的差分示意图
Figure 4 Difference diagram of Laplacian

| | | |
|----|----|----|
| -1 | -1 | -1 |
| -1 | 8 | -1 |
| -1 | -1 | -1 |

图5 拉普拉斯算子的模板
Figure 5 Template of Laplacian

由差分示意图可以看出,拉普拉斯算子对于中心点(即黑色圆点)的8邻域分别进行了8次负的差分运算,将这8个值累加起来,相当于原图像的像素灰度值减去该点拉普拉斯算子卷积后的结果^[10],起

到了突出中心点灰度值的作用。

拉普拉斯是一种微分算子,它可以在增加图像灰度值突变的同时,减弱灰度缓慢变化的区域。将原始图像与拉普拉斯锐化处理后的图像叠加,这样不仅能够突出图像中的细节信息,同时还可以保留图像的背景信息:

$$g(x,y)=\begin{cases} f(x,y)+\nabla^2 f(x,y), & \text{掩膜中心系数为正} \\ f(x,y)-\nabla^2 f(x,y), & \text{掩膜中心系数为负} \end{cases} \quad (6)$$

(2)频率滤波。对图像 $f(x,y)$ 使用高斯低通滤波 5×5 模板进行平滑处理,消除高频信息,保留背景等低频信息,得到的低通图像为 $f_s(x,y)$ 。

高斯低通滤波即基于高斯函数的低通滤波^[11],它的变换函数如下:

$$H(u,v) = e^{-[D(u,v)/D_0]^2} \quad (7)$$

其中, $D(u,v)$ 表示从点 (u,v) 到傅里叶变换中心(即原点)的距离; D_0 是高斯低通滤波的截止频率,同时也是高斯函数的方差,表示高斯曲线的扩张程度^[12]。

(3)空-频域差值。将锐化后的图像 $f_e(x,y)$ 和滤波后的图像 $f_s(x,y)$ 进行差运算,产生一幅差值图像 $f_b(x,y)$:

$$f_b(x,y) = f_e(x,y) - f_s(x,y) \quad (8)$$

在图像相减过程中,产生的 $f_b(x,y)$ 图像中可能存在一些负值,由于像素值为负的点为低频背景,并不含微钙化点,所以本文对 $f_b(x,y)$ 进行取0操作^[13]:

$$f_b(x,y) = \max(0, f_b(x,y)) \quad (9)$$

可通过上述方法对乳腺图像进行差值处理得到只包含微钙化点和高频噪声的新图像 $f_b(x,y)$,然后通过最大类间方差法(Otsu)分割出微钙化点:

$$f_b(x,y) = \begin{cases} 1, & f_b(x,y) > T \\ 0, & f_b(x,y) \leq T \end{cases} \quad (10)$$

其中, T 为前景(即微钙化点)和背景的分割阈值。

通过空-频域差值、Otsu 阈值算法分割后的图像如图6所示。通过图6e发现,微钙化点得到了有效检测,为后续ROI的自动分割奠定了基础。

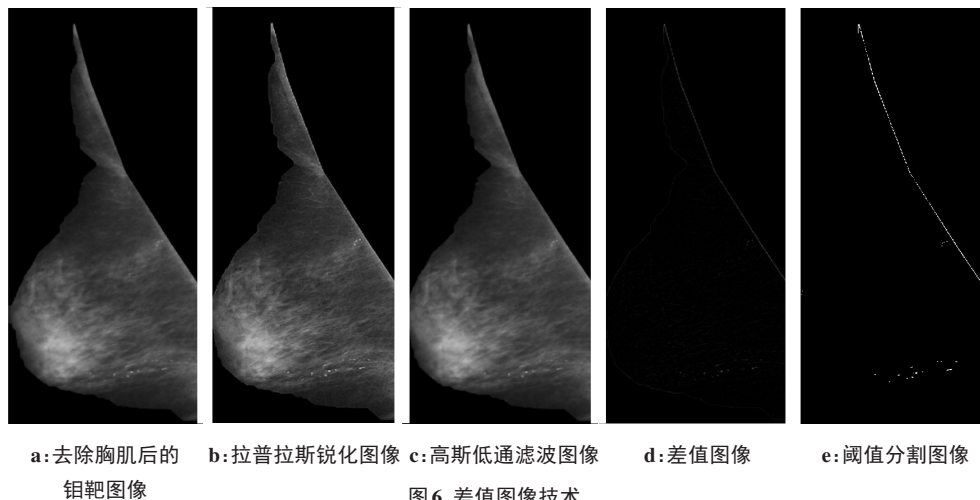


图6 差值图像技术

Figure 6 Difference image technique

2.2.2 ROI自动分割 由于乳腺部分边缘可能存在着微钙化点,所以在去除胸肌时不能将其完全去除,因此在阈值分割后的图像存在一组描述胸肌轮廓的离散点。通过膨胀操作,使这些离散点连接起来组成一条轮廓线(图7a),然后计算图像中的最大连通区域(即胸肌轮廓线),将其去除,此时图像中仅存在微钙化点。通过循环计算得到每个微钙化点的质心,并用“*”将其在二值图像上标注(图7b)。最后在钼靶图像上根据每个微钙化点的质心位置绘制出每个微钙化点所对应的最小外接矩形(图7c),实现疑似微钙化点的检测。

在医学上,单独存在的钙化点往往被认为是良性钙化。良性钙化点多比较分散,密度也较一致,还

具有边界清晰、形态规则的特点,而成簇分布的微钙化点是早期乳腺癌非常重要的诊断依据,钙化密集度越高,其恶性的可能性就越大。多数学者认为在乳腺钼靶图像中单位面积(cm^2)内存在3枚以上的微钙化点则定义为微钙化点簇,可诊断为乳腺癌^[14]。

本文根据疑似微钙化点的质心位置,并结合医学上对于微钙化点簇的定义,从上到下、从左到右自动截取 128×128 的疑似微钙化区域,每个微钙化区域含3枚以上的微钙化点,保存到指定文件夹中,如图8所示。

2.3 微钙化区域超分辨率重建

微钙化区域尺寸较小,且分辨率较低、特征不明显。由于存在上述特点,本文引入图像超分辨率

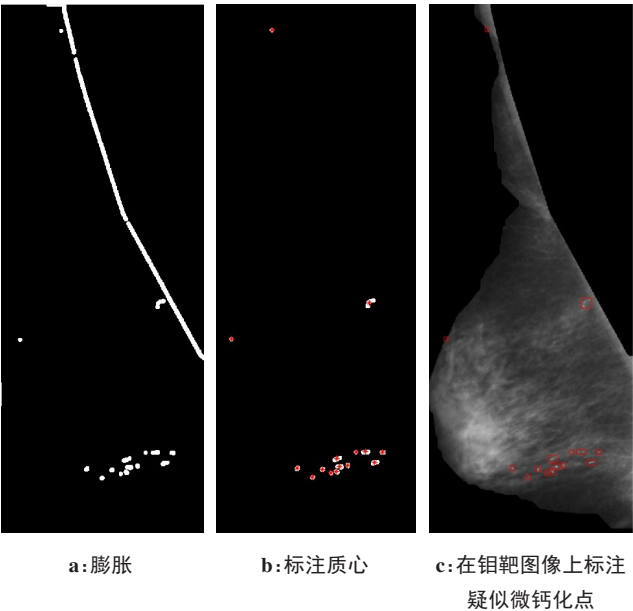


图 7 疑似微钙化点检测
Figure 7 Detection of suspected microcalcifications

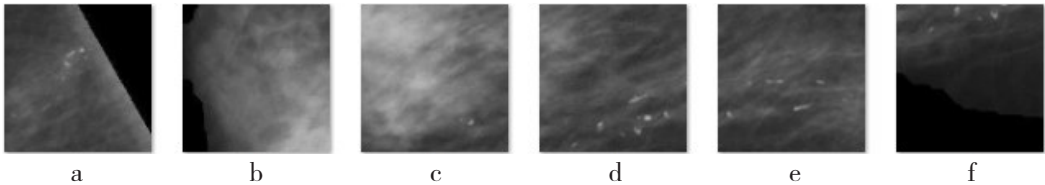


图 8 单幅乳腺图像所截取的疑似微钙化区域
Figure 8 Suspected microcalcifications in a single breast image
图 a~f 代表不同乳腺图像所截取的疑似微钙化区域

(Reconstruction Block, RB)3 部分组成。输入的 LR 图像先经过两层 3×3 的卷积层进行特征提取,然后与前一次迭代 FB 的输出进行级联与压缩后输入到本次迭代的 FB 中进行反馈学习;FB 输出一条分支作为下次迭代中

(Super Resolution, SR)重建技术对 ROI 图像进行放大。图像 SR 重建技术由 Tsia 在 20 世纪 80 年代首次提出,是指将低分辨率(Low Resolution, LR)图像进行重建,从而得到对应的高分辨率(High Resolution, HR)图像的过程^[15]。随着大量的科研人员投入到图像 SR 重建技术的研究中,SR 重建技术取得了重大发展,涌现出许多经典的 SR 重建算法。经过 SR 重建的图像细节更加清晰,视觉效果更好,有利于进一步的特征提取与分类。

本文使用超分辨率反馈网络(Super-Resolution Feedback Network, SRFBN)对图像进行 SR 算法重建^[16],结构如图 9 所示。SRFBN 有较强的早期重建能力,可以逐步生成最终的高分辨率图像,整个网络可以展开成 T 次迭代^[17],每次迭代都由低分辨率特征提取块(Low Resolution Feature Extractor Block, LRFB)、反 馈 块 (Feedback Block, FB) 和 重 建 块

FB 块的输入,另一条分支传入反卷积层进行上采样;最后经过一层 3×3 的卷积层得到重建的残差图像,与 LR 上采样图像相加后,得到这一次迭代的输出,经过 T 轮迭代后的输出,得到最终的高分辨率图像。

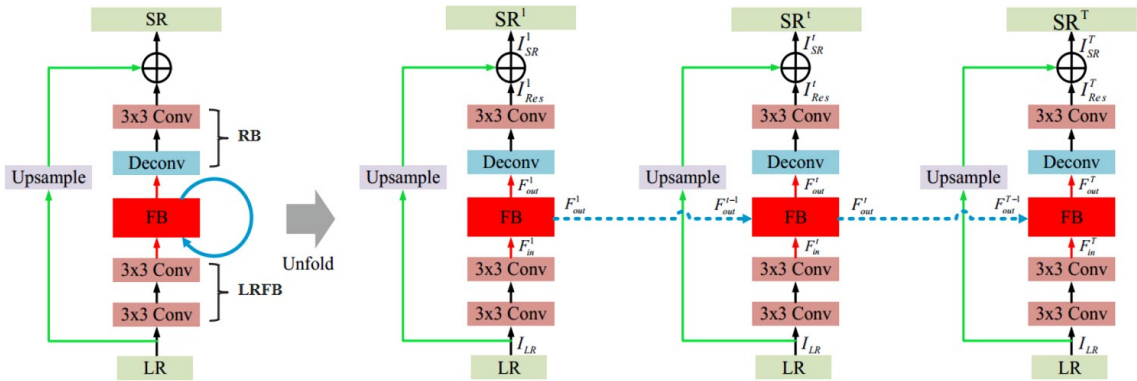


图 9 SRFBN 结构图
Figure 9 Structure diagram of SRFBN

2.4 微钙化区域真假阳性检测

2.4.1 特征提取 乳腺钼靶图像为灰度图像,钙化点的形态各异,如图 10 所示通过颜色特征和形状特征

难以区分。而乳腺组织区域中微钙化点区域相对正常区域在纹理方面差异明显^[12],因此可提取 ROI 的纹理特征,作为分类器的输入特征。

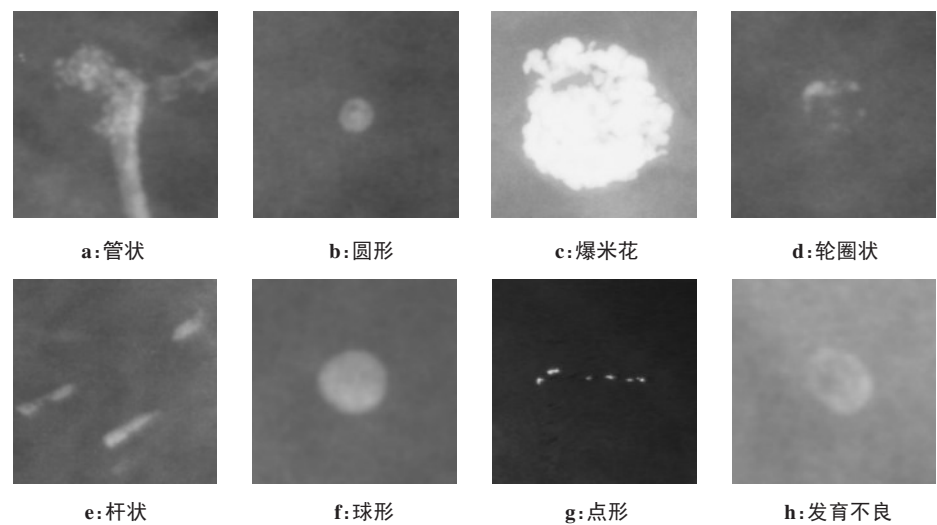


图10 钙化点类型
Figure 10 Types of calcifications

本文提取出ROI的Haralick纹理特征、灰度游程矩阵特征和方向梯度直方图(Histogram of Oriented Gradient, HOG)、局部二值模式(Local Binary Pattern, LBP)特征,共363维特征用于后续分类。

2.4.2 分类决策 完成疑似微钙化区域的特征提取后,如何区分正常乳腺组织和微钙化点区域是乳腺CAD系统的关键。多分类器的融合方法,可以放宽数据大小要求,平衡计算量,解决模型的不确定性。因此,采用单层决策树(Decision Stump, DS)作为弱分类器,通过Gentle AdaBoost算法集成强分类器GAB-DS。

3 实验结果与方法

3.1 基于空-频域差值的ROI分割实验

3.1.1 评价标准 为了验证和量化此检测方法对病灶检测的准确程度,采用敏感性(Sensitivity)、FPR对模型进行评价。

敏感性又称为真阳性率,即2.2节检测方法识别出的阳性区域数量与实际上所有样本中微钙化区域总数的比例:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \tag{11}$$

FPR是指正常组织被错判为微钙化区域的比例:

$$\text{FPR} = \frac{FP}{FP + TN} \times 100\% \tag{12}$$

设阳性代表微钙化区域,阴性表示非钙化区域。上式中,真阳性(True Positive, TP)表示被正确分割的微钙化区域样本个数;假阳性(False Positive, FP)表示被错误分割的微钙化区域样本个数;真阴性(True Negative, TN)表示被正确分割的正常组织样本个数;假阴性(False Negative, FN)表示被错误分割的正常组织样本个数。

3.1.2 结果与分析 为了验证本文提出的空-频域微钙化区域检测方法的可靠性,采用DDSM数据库中400张图像进行实验。在含有微钙化点的200张钼靶图像中,检测出含微钙化点的图像有182张;200张正常图像中,检测出含微钙化点的图像为68张。故微钙化点检测的敏感性为91.00%,FPR为34.00%。本文所采用的组合多种方法的微钙化区域检测模型取得了较高的检出率,但误检率也很高。可以通过后续的真假阳性分类实验,检测出真阳性的图像,降低FPR。

3.2 微钙化区域SR重建实验

3.2.1 主观评价标准 主观评价是以一定数量的人的视觉感为标准进行打分,比如视觉舒适度、结构层次和清晰度等方面,视觉感越好评分越高^[18]。主观评价是指通过人的肉眼从多个角度对不同算法重建出的图像进行主观评价,并对图像质量进行定性分析的过程。由于人的视觉对于图像的清晰度、边缘、图像亮度、对比度等信息比较敏感,所以当图像间的质量差异较大时,主观评价结果具有一定的参考价值。主观评价标准虽然能最直观地反映图像质量,但也存在一定的缺陷。首先,主观评价需要足够的观察者进行评价,并且耗时较长;此外,评价结果会受到观察者视觉习惯、生理状态以及外界环境等因素的影响,使得实验结果无法完全作为图像质量的评价标准,仅能起到参考价值。

3.2.2 客观评价标准 客观评价采用具体量化数据对重建结果进行评价,常用指标有峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似度(Structural Similarity, SSIM)两种。

PSNR评价指标是SR领域中广泛应用的指标之

—^[19], 将其原始图像和重建后图像进行逐个像素点的对比, 并计算两幅图像像素点之间的差值用来评价重建后图像的质量。PSNR的单位为dB, 若两幅图像之间的PSNR越大, 表示图像的质量越高, 失真越小。表达式如下:

$$\text{PSNR} = 10\log_{10}\left(\frac{\max_i I^2}{\text{MSE}}\right) \quad (13)$$

其中, $\max_i I$ 表示图像的最大像素值, 若每个像素都由8位二进制表示, $\max_i I$ 则为255; MSE为原始图像与重建图像之间的均方误差, 主要反映了重建图像与原始图像的偏离情况, MSE的值越小, 重建图像的质量越高, 其表达式如下:

$$\text{MSE} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (X_{\text{HR}}(i, j) - X_{\text{SR}}(i, j))^2 \quad (14)$$

其中, $m \times n$ 为原始图像与重建图像的大小, X_{HR} 表示原始的高分辨率图像, X_{SR} 为重建后的图像。

常用SSIM指标度量原图像和重建图像内部结构的相似度。SSIM指标综合了亮度、结构和对比度

3个方面整合计算, 其值越接近1, 则两幅图像的相似度越高, 重建效果越好。SSIM的表达式如下:

$$\text{SSIM}(i, j) = \frac{(2u_i u_j + C_1)(2\sigma_{ij} + C_2)}{(u_i^2 + u_j^2 + C_1)(\sigma_i^2 + \sigma_j^2 + C_2)} \quad (15)$$

其中, i 和 j 分别表示原图像与重建图像, u_i 和 u_j 分别为 x 和 y 的平均灰度值, σ_i 和 σ_j 为 i 和 j 的方差, σ_{ij} 为 i 和 j 的协方差, C_1 和 C_2 为常数。

3.2.3 结果与分析 为了验证所使用的SRFBN算法在微钙化区域图像分辨率提升上的优势, 本节主要采用SRFBN和Bicubic、基于卷积神经网络的SR重建开山之作SRCNN(Super Resolution Convolutional Network)算法进行了重建实验, 同时对实验结果在主观和客观两个方面进行了比较分析。

图11为原始图像与放大因子为4倍的Bicubic、SRCNN、SRFBN的高分辨率图像对比结果。从图中可能无法看出明显差别, 因此采用客观评价标准PSNR和SSIM分别对重建结果进行评价, 结果如表1和表2所示。

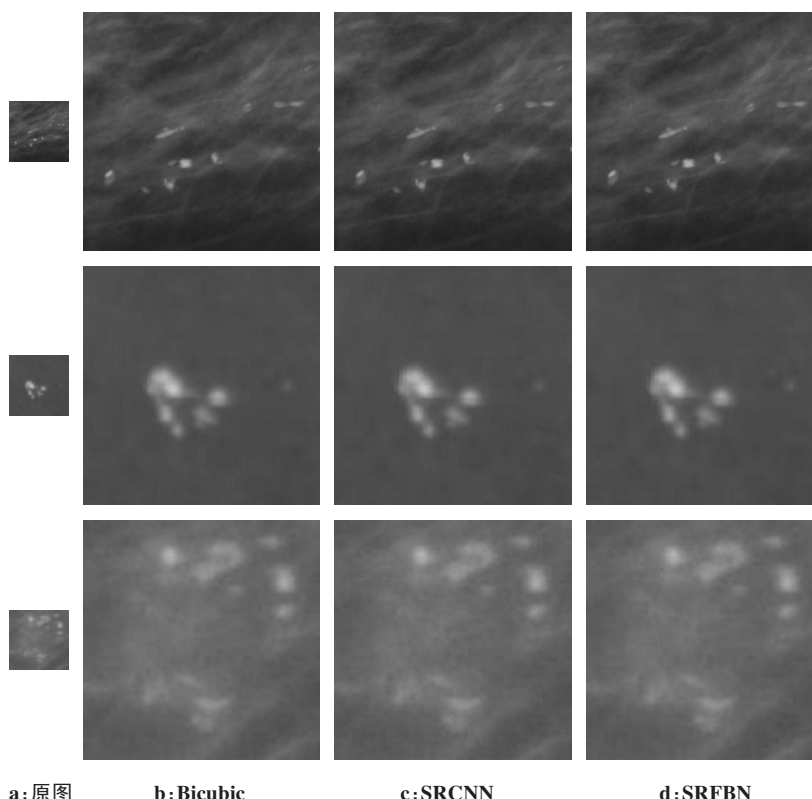


图11 放大因子为4倍的重建结果

Figure 11 Reconstruction results with magnification factor of 4

从表1和表2可以看出, 相比传统Bicubic方法, 基于学习的SR重建方法在PSNR和SSIM评价指标上有着显著的优势, 其中SRFBN算法在评价指标上全面优于SRCNN算法。综上所述, 选用SRFBN算法对实验数据集中的全部图像进行SR重建。

3.3 微钙化区域真假阳性检测实验

乳腺微钙化区域检测流程如图12所示。在GAB-DS模型中, 决策树的数量影响着分类器的性能。当值较大时, GAB-DS模型的分类误差小, 分类性能较好, 而模型的复杂性与值成正比, 值越大, 模

表 1 各超分辨率重建算法 PSNR 对比(dB)

Table 1 Comparison of PSNR among super-resolution reconstruction algorithms (dB)

| 图像 | Bicubic | SRCNN | SRFBN |
|--------|---------|-------|-------|
| 测试图像 1 | 32.44 | 36.72 | 37.83 |
| 测试图像 2 | 29.43 | 30.53 | 31.28 |
| 测试图像 3 | 26.46 | 28.89 | 30.56 |

表 2 各超分辨率重建算法 SSIM 对比

Table 2 Comparison of SSIM among super-resolution reconstruction algorithms

| 图像 | Bicubic | SRCNN | SRFBN |
|--------|---------|--------|--------|
| 测试图像 1 | 0.9144 | 0.9305 | 0.9486 |
| 测试图像 2 | 0.8260 | 0.8524 | 0.8638 |
| 测试图像 3 | 0.7045 | 0.7367 | 0.7521 |

型的复杂度越高,运行时间也越长。在 1 280 幅 ROI 图像中,随机选取 960 幅 ROI 图像为训练集,剩余的 320 幅为验证集。通过 MATLAB 中的集成学习工具

箱,搭建出 GAB-DS 模型。将决策树的数量设置为 500,绘制决策树的数量与分类错误率的关系曲线,如图 13 所示。

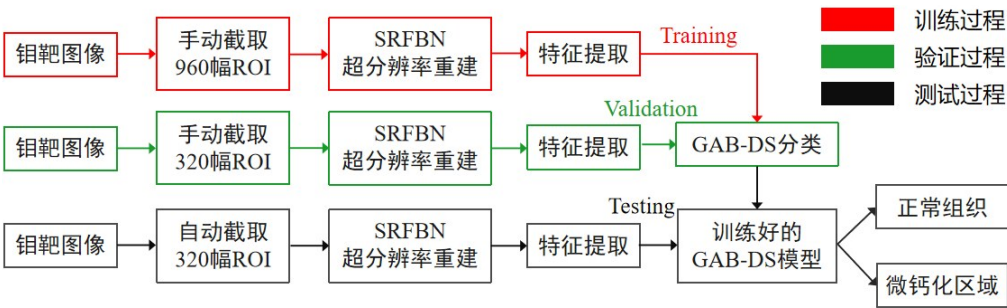


图 12 乳腺微钙化区域检测流程图

Figure 12 Flowchart of breast microcalcification detection

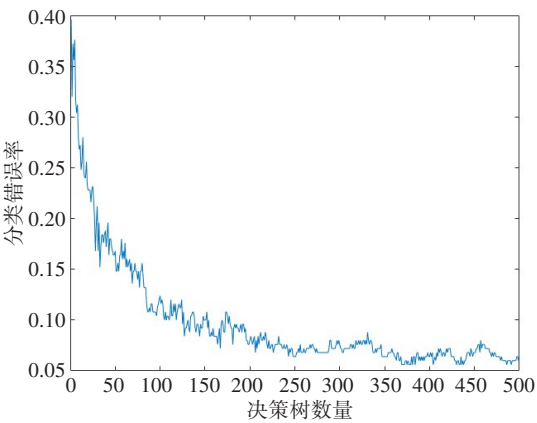


图 13 决策树数量与分类错误率的关系曲线

Figure 13 Relationship between the number of decision trees and the classification error rate

由图 13 可以看出,随着决策树数量 k 值的增大,分类错误率先快速下降,然后逐渐趋于平缓。图像中存在多个分类错误率最低点,综合分类准确率和运算复杂度考虑, k 应设置为 369(即第一个分类错误率最低点的横坐标)。

3.3.1 评价标准 为了验证和量化此分类模型的效果,记录实验结果中的 TP、FP、TN 和 FN,并计算出相应的准确率(Accuracy)、敏感性(Sensitivity)、特异性(Specificity)、阳性预测值(Positive Predictive Value,

PPV)、阴性预测值(Negative Predictive Value, NPV)及 F1 分数(F1-score)对训练的模型进行评价,其定义如下所示:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (16)$$

$$\text{Sensitivity} = \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (17)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}} \times 100\% \quad (18)$$

$$\text{PPV} = \text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (19)$$

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}} \times 100\% \quad (20)$$

$$\text{F1 - score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (21)$$

3.3.2 结果与分析 基于差值图像技术自动分割的疑似微钙化区域共 320 个,对照金标准图像打标签,用作 GAB-DS 模型的测试集来检验所使用的分类模型性能。根据上一节所述,将决策树的数量设置为 369,分别对原始数据集和 SR 重建后的数据集提取特征进行分类,实验结果如表 3 所示。数据集未进行 SR 重建时,准确率为 93.13%,已处于较高水平。数据集经 SR 重建后,分类准确率达到了 96.25%,敏感性为 94.38%,特异性为 98.13%,PPV 为 98.05%,NPV

为 94.58%,F1 分数为 0.961 8,GAB-DS 模型的各项指标均有显著提升。由超分前后的分类结果对比,可以看出图像 SR 重建有助于微钙化区域真假阳性检测。为了进一步验证 GAB-DS 算法的性能,将本文算法与其他算法进行实验比较,实验所用数据集均为 SR 重建后的图像数据,结果如表 4 所示。

表 3 数据集超分辨率重建前后分类对比
Table 3 Classification comparison for data sets before and after super-resolution reconstruction

| 数据集 | 准确率/% | 敏感性/% | 特异性/% | PPV/% | NPV/% | F1 分数 |
|---------|-------|-------|-------|-------|-------|---------|
| 超分辨率重建前 | 93.13 | 92.50 | 93.13 | 93.08 | 92.55 | 0.927 9 |
| 超分辨率重建后 | 96.25 | 94.38 | 98.13 | 98.05 | 94.58 | 0.961 8 |

PPV: 阳性预测值;NPV: 阴性预测值

表 4 不同算法的乳腺微钙化区域分类性能对照表
Table 4 Performances of different algorithms in breast microcalcification classification

| 算法 | 准确率/% | 敏感性/% | 特异性/% | PPV/% | NPV/% | F1 分数 |
|--------|-------|-------|-------|-------|-------|---------|
| SVM | 86.25 | 92.50 | 80.00 | 82.22 | 91.43 | 0.870 6 |
| KNN | 87.50 | 90.00 | 85.00 | 85.71 | 89.47 | 0.878 1 |
| AB-DS | 90.94 | 91.25 | 91.25 | 91.20 | 90.68 | 0.909 1 |
| GAB-DS | 96.25 | 94.38 | 98.13 | 98.05 | 94.58 | 0.961 8 |

为了使数据比较的结果更加直观,生成了基于不同方法的乳腺微钙化区域分类性能对照柱状图,如图 14 所示,本文所用模型在各项评价指标上均优

于其他模型。由此可见,所提出的检测模型在微钙化区域检测上具有较好的检测性能。

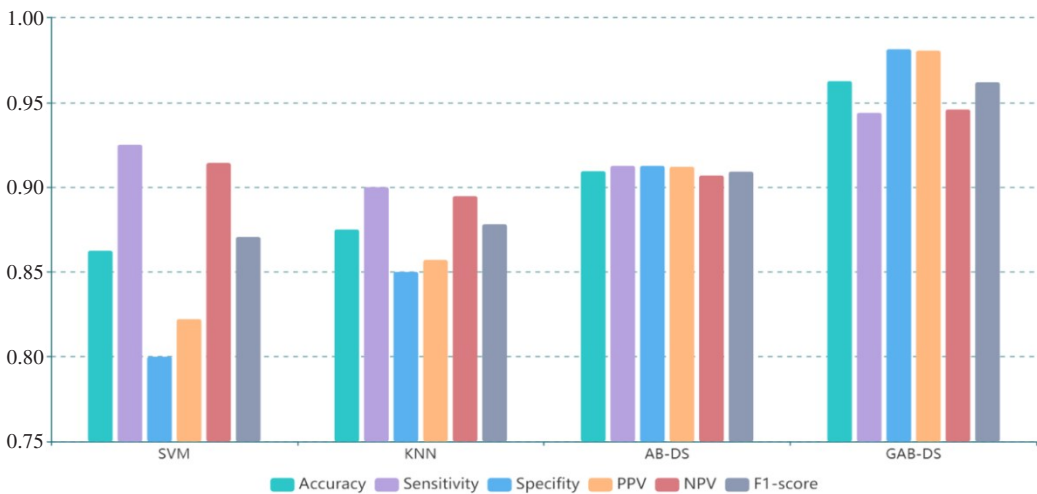


图 14 基于不同方法的乳腺微钙化区域分类性能对照柱状图
Figure 14 Histogram of performances of different methods in breast microcalcification classification

4 结束语

本文针对微钙化点检查精度不高的问题,提出了一种乳腺钼靶图像微钙化真假阳性检测方法。首先对乳腺钼靶图像进行了预处理,并采用差值图像技术和阈值分割算法实现了乳腺钼靶图像中疑似微钙化点的分割,根据疑似微钙化点的质心位置自动

截取 ROI,然后通过 SRFBN 算法进行微钙化区域 SR 重建,最后提取 ROI 的特征,将 Gentle AdaBoost 算法和单层决策树算法相结合,构建强分类器 GAB-DS 对区域进行分类,将微钙化区域和正常组织分离开来,分类正确率高达 96.25%。在辅助乳腺微钙化点检测中具有一定的临床应用价值。

【参考文献】

- [1] 曹毛毛,陈万青. GLOBOCAN 2020 全球癌症统计数据解读[J]. 中国医学前沿杂志(电子版), 2021, 13(3): 63-69.
Cao MM, Chen WQ. Interpretation on the global cancer statistics of GLOBOCAN 2020[J]. Chinese Journal of The Frontiers of Medical Science (Electronic Version), 2021, 13(3): 63-69.
- [2] 郭亚南. 乳腺钼靶X线病灶检测研究[D]. 兰州: 兰州大学, 2019.
Guo YN. Research on lesion detection for mammograms [D]. Lanzhou: Lanzhou University, 2019.
- [3] 钟明霞. 基于微钙化点检测的乳腺X线图像分割研究[D]. 杭州: 浙江工业大学, 2009.
Zhong MX. Research on breast X-ray image segmentation based on microcalcification detection [D]. Hangzhou: Zhejiang University of Technology, 2009.
- [4] 梅明. 全视野数字乳腺X线摄影中微钙化检测方法研究[D]. 武汉: 武汉科技大学, 2016.
Mei M. Research of microcalcification detection methods for breast cancer in full field digital mammography [D]. Wuhan: Wuhan University of Science and Technology, 2016.
- [5] Ciecholewski M. Microcalcification segmentation from mammograms: a morphological approach[J]. J Digital Imaging, 2016, 30(2): 172-184.
- [6] 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.
- [7] Heath M, Bowyer K, Kopans D, et al. Current status of the digital database for screening mammography[M]. Digital Mammography, 1998: 457-460.
- [8] 顾广娟. 乳腺X线影像微钙化簇检测技术研究[D]. 哈尔滨: 哈尔滨理工大学, 2009.
Gu GJ. Research of microcalcification cluster detection technology in mammograms [D]. Harbin: Harbin University of Science and Technology, 2009.
- [9] 冯丽,王娟. 基于Matlab的拉普拉斯图像增强算法与设计[J]. 电子技术与软件工程, 2020(16): 143-145.
Feng L, WANG J. Laplace image enhancement algorithm and design based on MATLAB [J]. Electronic Technology & Software Engineering, 2020(16): 143-145.
- [10] 李海. 基于FPGA的红外图像平滑与锐化算法研究及硬件设计[D]. 武汉: 华中科技大学, 2012.
Li H. Research and hardware implementation of FPGA based infrared image smoothing and sharpen algorithms [D]. Wuhan: Huazhong University of Science and Technology, 2012.
- [11] 王杰. 基于地物光谱矢量空间的遥感图像去噪方法研究[D]. 长春: 东北师范大学, 2014.
Wang J. The research for remote sensing image denoising method based on ground surface spectrum vector space [D]. Changchun: Northeast Normal University, 2014.
- [12] 王捷. 数字图像边缘检测方法的若干改进与应用研究[D]. 杭州: 浙江理工大学, 2010.
Wang J. Some improvement and application researches on edge detection methods of digital image [D]. Hangzhou: Zhejiang Sci-Tech University, 2010.
- [13] 张胜君. 基于乳腺X线图像的乳腺癌检测方法研究[D]. 北京: 北京交通大学, 2013.
Zhang SJ. Detection method of breast cancer based on mammogram [D]. Beijing: Beijing Jiaotong University, 2013.
- [14] 刘林. 基于OpenCV的乳腺X线医学图像处理与分析系统设计与实现[D]. 北京: 北京交通大学, 2012.
Liu L. Design and implementation of processing and analysis in mammograms based on OpenCV [D]. Beijing: Beijing Jiaotong University, 2012.
- [15] 吴磊. 基于深度学习的图像超分辨率重建及应用[D]. 合肥: 合肥工业大学, 2020.
Wu L. Research and application of image super-resolution reconstruction based on deep learning [D]. Hefei: Hefei University of Technology, 2020.
- [16] LI Z, Yang J, Liu Z, et al. Feedback network for image super resolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 3867-3876.
- [17] 陈文静,唐轶. 基于深度学习的图像超分辨率重建方法综述[J]. 云南民族大学学报(自然科学版), 2019, 28(6): 597-605.
CHEN WJ, TANG Y. A summary review of the methods for image super-resolution reconstruction based on deep learning[J]. Journal of Yunnan Minzu University (Natural Sciences Edition), 2019, 28(6): 597-605.
- [18] 迟嘉骅. 基于数字图像处理的皮肤病识别APP[J]. 价值工程, 2020(2): 246-247.
CHI JH. Skin disease recognition APP based on digital image processing[J]. Value Engineering, 2020(2): 246-247.
- [19] Xiao Y, Zhou L, Chen W. Wavefront control through multi-layer scattering media using single-pixel detector for high-PSNR optical transmission[J]. Optics and Lasers in Engineering, 2021, 139(12): 106453.

(编辑:薛泽玲)