

DOI:10.11918/j.issn.0367-6234.201904005

基于超声和钼靶的两阶段乳腺癌诊断系统

管红娇,张英涛,唐降龙

(哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001)

摘要: 全球范围内乳腺癌发病率持续上升,乳腺癌的异质性导致恶性肿瘤在超声声像上呈现不同程度交叉和重叠,只利用一种图像信息不能得到满意的分类效果. 因此本文提出基于超声和钼靶的两阶段乳腺癌诊断系统:第一阶段用拒绝分类方法对乳腺超声图像进行识别,对置信度高的肿瘤赋予其类别标签,对不确定的肿瘤不予分类;第二阶段对这些没有分类的肿瘤使用钼靶图像进行辨别. 该系统利用多模态图像信息,筛查难以辨识的超声图像,辅以钼靶图像信息,对乳腺肿瘤进行诊断. 本研究采用的超声和钼靶图像数据由哈尔滨医科大学肿瘤医院和哈尔滨医科大学第一附属医院提供,分别对拒绝分类方法和两阶段诊断系统进行实验验证和方法对比. 与只使用乳腺超声特征的诊断系统相比,本文提出的两阶段诊断系统性能更优:正确率为 92.59%,AUC 为 0.933 3,G-mean 为 93.09%、敏感性为 86.67%,特异性为 100%,阳性预测值为 100%,阴性预测值为 85.71%,马修斯相关系数为 0.861 9. 实验结果表明,增加钼靶图像信息可提高只使用超声图像的单一模态诊断系统的性能.

关键词: 计算机辅助诊断;乳腺癌;超声;钼靶;拒绝分类;支持向量机

中图分类号: TP391

文献标志码: A

文章编号: 0367-6234(2019)11-0008-08

Two-stage breast cancer diagnosis system based on ultrasound and mammogram images

GUAN Hongjiao, ZHANG Yingtao, TANG Xianglong

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

Abstract: The incidence of breast cancer continues to rise worldwide. Due to the heterogeneity of breast cancer, there is an overlap between benign and malignant tumor images that only using a type of images cannot obtain satisfactory classification results. This paper proposes a two-stage breast cancer diagnosis system based on ultrasound and mammogram images. In the first phase, an abstaining classification method is used to classify breast ultrasound (BUS) images, in which some BUS tumors are classified with high confidence, and the uncertain tumors are not classified. These unclassified tumors are then classified using mammogram images in the second stage. Supplemented by mammogram information, the system can diagnose breast cancer by utilizing multimodal image information to screen for unrecognizable ultrasound images. The ultrasound and mammogram images used in this study were provided by the Cancer Hospital of Harbin Medical University and the First Affiliated Hospital of Harbin Medical University. The abstaining method and the two-stage diagnosis system were validated in experiments. Compared with diagnostic systems using only BUS features, the proposed two-stage diagnostic system provided better performance with the accuracy of 92.59%, AUC of 0.933 3, G-mean of 93.09%, sensitivity of 86.67%, specificity of 100%, positive predictive value of 100%, negative predictive value of 85.71%, and Matthew's correlation coefficient of 0.861 9. Experimental results demonstrate that adding mammogram information can increase the performance of the diagnosis system that uses BUS images only.

Keywords: computer-aided diagnosis; breast cancer; ultrasound; mammogram; abstaining classification; support vector machine

乳腺癌是全球女性癌症发病和死亡的主要原因^[1],全球范围内乳腺癌发病率仍在增加,特别是在亚洲国家^[2-3]. 早筛查早治疗是降低乳腺癌死亡率的重要预防手段^[4-5]. 乳腺超声和钼靶是检测乳

腺癌的两个重要方式. 研究表明,钼靶筛查可降低死亡率,且已成为许多发达国家筛查乳腺癌的首选方法^[6]. 然而,在年轻或致密乳房的乳腺癌患者中,钼靶诊断敏感性有所降低^[3,7]. 另外,乳腺钼靶检查使乳房暴露于 X 射线辐射,不建议孕妇进行常规检查. 相比之下,超声检查无放射性、操作灵活且费用较低,因此,许多发展中国家使用超声影像筛查乳腺癌. 但是,超声图像存在斑点噪声,对比度较低,由于

收稿日期: 2017-04-01

作者简介: 管红娇(1990—),女,博士研究生;

唐降龙(1960—),男,教授,博士生导师

通信作者: 唐降龙, tangxl@hit.edu.cn

乳腺癌的异质性以及超声声像复杂多样, 良性和恶性肿瘤的声像特征存在交叉重叠^[8-9], 可使用图像处理 and 机器学习方法解决超声诊断的这些问题。

目前检测和诊断乳腺癌的方法主要集中于单模态图像, 例如超声^[10-14]、钼靶^[15-18]、磁共振成像^[19]等。临床试验表明, 结合超声和钼靶可提高早期乳腺癌的检出率^[3]。考虑上述超声和钼靶的优缺点, 若由超声图像可得到可靠的诊断结果, 则无需进行钼靶检查; 由超声图像无法判断的肿瘤, 进一步使用钼靶进行诊断。鉴于此, 本文提出两阶段乳腺癌诊断系统, 对超声和钼靶图像进行级联分类, 辨别乳腺肿瘤良恶性。实验结果表明, 该系统性能优于只使用超声的诊断系统。

1 图像采集

本研究分两步验证提出的两阶段诊断系统, 首先验证第一阶段的拒绝分类方法(下称“拒绝方法验证”), 其次验证两阶段诊断系统的级联分类方法(下称“级联方法验证”)。本研究获取每位参与者知情同意, 患者隐私得到完全保护。

拒绝方法验证实验图像: 从哈尔滨医科大学第一附属医院收集了 219 名乳腺肿瘤患者的超声图像, 包含 92 张恶性和 127 张良性图像, 成像设备为具有 5 ~ 14 MHz 线性探针的 VIVID 7 (GE, Horten, Norway) 超声仪。219 例乳腺肿瘤已通过手术和病理检查确定其良性或恶性。

级联方法验证实验图像: 从哈尔滨医科大学肿瘤医院收集了 52 名患者的超声和钼靶图像, 成像设备分别为具有 5 ~ 14 MHz 线性探针的 VIVID 7 (GE, Horten, Norway) 超声仪和 Senographe Essential (GE, America) 钼靶仪。所有乳腺肿瘤通过手术和病理检查确定其良恶性。

验证级联分类方法需要超声图像和钼靶图像相互对应, 都对应于同一乳腺肿瘤, 因此需筛选合适的图像数据。由于肿瘤位置较深或乳房组织致密等原因, 钼靶图像不能清晰呈现病灶区域, 52 名患者中有 30 名因此情况被排除, 保留的 22 名患者为 7 名良性和 15 名恶性。由于样本量非常小且不平衡, 为了增加数量少的良性肿瘤图像的数量, 对于 7 名良性患者各挑选两张图像, 而对于 15 名恶性患者, 各挑选一张图像。具体挑选过程: 对于超声图像, 7 名良性患者中只有 5 名患者能各挑出较清晰的两张图像, 剩余 2 名良性患者和 15 名恶性患者各挑选一张清晰的超声图像, 因此超声图像集共有 12 张良性和 15 张恶性图像; 为与超声图像集对应, 对于钼靶图像, 为上述 5 名良性患者挑选头尾位 (craniocaudal,

CC) 和侧斜位 (mediolateral oblique, MLO) 两个视角的钼靶图像, 剩余 2 名良性患者和 15 名恶性患者都选择侧斜位的钼靶图像, 因此钼靶图像集共有 12 张良性和 15 张恶性图像。收集大量相同患者的超声和钼靶图像是一项艰难的工作, 涉及筛查患者和图像, 仅做超声或钼靶一项检查的患者的图像不能采用, 而两项检查都做的患者, 若图像没有清晰呈现病变区域或成像质量不高不采用。

综上, 验证拒绝分类方法使用 219 例超声图像, 验证级联分类方法使用包含 27 例超声图像和 27 例钼靶图像的联合数据集。

2 两阶段诊断系统

图 1 为两阶段诊断系统的训练和测试流程图。本文提出的诊断系统分为 3 个步骤: 图像分割, 特征提取和选择, 基于超声和钼靶图像的两阶段级联分类。第一阶段根据给定条件用超声图像识别部分置信度高的肿瘤, 第二阶段根据钼靶特征对未识别的肿瘤进行分类。

2.1 图像分割

本文使用基于细胞自动机 (cellular automata, CA) 的方法^[20]对超声肿瘤区域进行分割。该分割方法对斑点噪声和模糊肿瘤边界具有鲁棒性。CA 描述了能量传输和状态转换的动态机制, 当相邻细胞间能量不再扩散且每个细胞状态不再改变, CA 便达到稳定状态。CA 中定义了 3 个概念: 包含所有细胞状态的细胞状态集, 细胞邻域系统, 以及表示邻域系统中细胞状态转换规则的状态转移函数。

在基于 CA 的图像分割中, 细胞和细胞空间分别对应图像像素和包含肿块的矩形区域 (region of interest, RoI)。本研究中 RoI 由经验丰富的放射科医师进行标记。像素 p 的状态定义为一个三元组 (l_p, s_p, e_p) , 其中 l_p 是像素的标签, 定义为

$$l_p = \begin{cases} 0, & p \text{ 是未标记像素;} \\ 1, & p \text{ 是前景像素;} \\ 2, & p \text{ 是背景像素。} \end{cases}$$

本研究中, 背景种子集初始化为 RoI 边界上的像素, 前景种子集由肿块内部某个圆上的像素组成。除了前景集和背景集中的像素之外, RoI 区域其余像素是未标记像素。 s_p ($\in [0, 1]$) 表示像素 p 属于某类别 (前景或背景) 的强度, 初始化为

$$s_p = \begin{cases} 0, & p \text{ 是未标记像素;} \\ 1, & p \text{ 是种子像素。} \end{cases}$$

e_p 为像素 p 的能量, 定义为该像素的灰度值。基于 CA 分割肿瘤就是根据状态转移函数把每个未被标记的像素分配到前景或背景集中, 最后前景集中

的像素构成肿瘤区域.

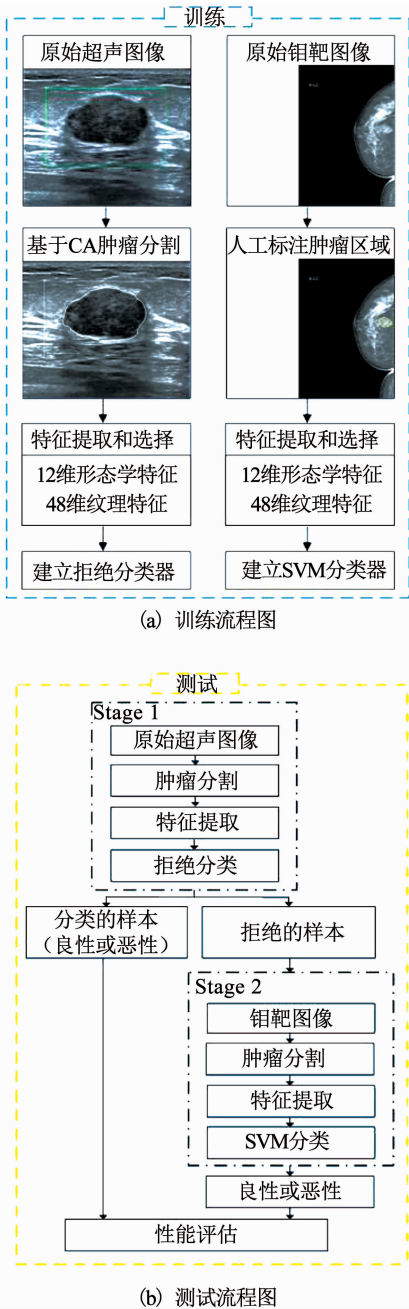


图 1 两阶段诊断系统的训练和测试流程图

Fig. 1 Flow diagram of training and testing processes of the proposed two-stage diagnosis system

从图像局部信息和全局信息两方面考虑设计状态转换函数. 对于肿块的模糊边界, 近邻像素的灰度值相似, 但这些像素邻域的局部纹理可能不同. 基于这个假设, 根据不同邻域之间相关纹理和自相关纹理对局部邻域信息进行建模. 对全局图像信息建模时考虑像素的灰度值和空间位置. 由于种子像素的标记已知, 因此使用种子像素来区分空间位置. 该分割方法采用两个邻域系统: Moore 邻域系统用于获取相邻像素的精确相似性, 使用 von Neumann 邻域系统避免过滤掉孤立像素的近邻像素. 详细的 CA

分割方法请参考文献[20].

肿瘤区域分割在诊断系统中起着重要作用, 影响提取特征的质量和分类性能. 乳腺钼靶图像肿瘤边界由哈尔滨医科大学肿瘤医院具有丰富临床经验的放射科医师手动标注.

2.2 特征提取和选择

对每个超声或钼靶图像提取描述肿瘤形状 ($s_1 - s_7$)、边缘 ($b_1 - b_4$) 以及方向 (θ) 的 12 维形态学特征^[14,21-22] (见表 1) 和基于灰度共生矩阵 (gray-level co-occurrence matrix, GLCM) 的 48 维纹理特征^[23]. 纹理特征采用 4 个描述子, 对比度 (contrast)、相关性 (correlation)、能量 (energy) 和同质性 (homogeneity), 每个描述子提取 4 个角度 ($0^\circ, 45^\circ, 90^\circ, 135^\circ$) 和 3 个距离 (1, 2, 3) 的特征, 因此共有 $48(4 \times 4 \times 3)$ 维纹理特征.

太多特征易导致模型复杂, 出现过拟合, 增加计算量, 本研究采用 infinite latent feature selection (ILFS) 算法^[24] 进行特征选择. ILFS 是一种过滤式 (filter) 特征选择方法, 包括特征排序 (ranking) 和选择 (selection) 两个步骤. 详细算法请参考文献[24]. 通常根据交叉验证程序选择前 m' 个特征 (假设有 m 个原始特征, $m' < m$). 本研究通过交叉验证选择性能最优的且数量最少的特征建立模型并分类.

表 1 形态学特征

Tab. 1 Morphological features

形态学特征	描述
s_1	似圆度
s_2	最小外接矩形的长宽比
s_3	径向距离的长宽比
s_4	平均归一化径向长度
s_5	归一化径向长度标准差
s_6	平均归一化径向长度嫡
s_7	面积比
b_1	边界粗糙度
b_2	小叶数目
b_3	椭圆归一化周长
b_4	椭圆归一化轮廓
θ	方向角度

2.3 拒绝分类

2.3.1 相关工作

拒绝分类方法根据一定条件筛选不确定样本, 暂时不赋予其类别以减少错误分类. 拒绝分类方法已应用于安全相关领域, 例如生物医学图像分类、故障检测^[25-26]等.

在传统二类分类问题中, 分类规则为

$$P_i(m) = \begin{cases} +, s(m) > t; \\ -, s(m) \leq t. \end{cases}$$

式中: t 为决策阈值, $s(m)$ 为样本 m 属于正类的置信得分. 在拒绝分类中, 分类规则为

$$P_{t_1, t_2}(m) = \begin{cases} +, s(m) > t_2; \\ -, s(m) < t_1; \\ R, t_1 \leq s(m) \leq t_2. \end{cases} \quad (1)$$

式中: t_1 和 t_2 为拒绝阈值. 根据得分分类器求得置信得分 s , 本研究使用支持向量机 (support vector machine, SVM) 作为得分分类器.

利用拒绝规则式 (1) 判断样本的类别, 需要建立拒绝分类模型确定拒绝阈值 t_1 和 t_2 . 主流的拒绝分类模型有以下两种. 一是最小化期望代价拒绝模型^[27-28]

$$\min_{t_1, t_2} R_{\text{cost}}(t_1, t_2).$$

其中期望代价定义为

$$R_{\text{cost}}(t_1, t_2) = p + (C_{\text{fn}} \cdot r_{\text{fn}}(t_1) + C_{\text{tp}} \cdot r_{\text{tp}}(t_1) + C_{\text{fp}} \cdot r_{\text{fp}}(t_1, t_2)) + p - (C_{\text{in}} \cdot r_{\text{in}}(t_1) + C_{\text{fp}} \cdot r_{\text{fp}}(t_2) + C_{\text{m}} \cdot r_{\text{m}}(t_1, t_2)).$$

式中: $p +$ 和 $p -$ 分别为正类和负类的先验概率, C_{in} 、 C_{tp} 、 C_{fp} 和 C_{fn} 分别为真阴性、真阳性、假阳性和假阴性的错分代价, r_{in} 、 r_{tp} 、 r_{fp} 和 r_{fn} 分别为真阴性率、真阳性率、假阳性率和假阴性率, C_{tp} 和 C_{m} 分别为正类和负类的拒绝代价, r_{tp} 和 r_{m} 分别为被拒绝分类的正类和负类样本的比例. 该拒绝模型通过求解最小化代价问题, 确定拒绝阈值, 因此需要提前设置完整代价矩阵 (C_{in} 、 C_{tp} 、 C_{fp} 、 C_{fn} 、 C_{tp} 和 C_{m}), 但实际应用中难以获得或估计这些代价.

第二种是有界拒绝 (bounded abstention, BA) 模型^[29]:

$$\min_{t_1, t_2} \frac{C_{\text{fn}} \cdot N_{\text{fn}}(t_1) + C_{\text{fp}} \cdot N_{\text{fp}}(t_2)}{N_{\text{in}}(t_1) + N_{\text{fp}}(t_2) + N_{\text{tp}}(t_2) + N_{\text{fn}}(t_1)},$$

$$\text{s. t. } r_{\text{rej}}(t_1, t_2) \leq k_{\text{max}}.$$

式中: N_{fn} (N_{fp}) 为被错误分类的正类 (负类) 样本数量, N_{in} (N_{tp}) 为正确分类的负类 (正类) 样本数量, r_{rej} 为总体拒绝率, 即所有未分类的样本占总样本的比例, k_{max} 为 BA 模型超参数. BA 模型通过求解有约束的最小化问题确定拒绝阈值. 该模型同样需要设置错分代价 C_{in} 和 C_{fp} , 若 C_{in} 和 C_{fp} 设为相同值, 则 BA 模型变为在总体拒绝率约束下最小化错误率, 不适用于不平衡数据.

2.3.2 拒绝分类模型

本文拒绝分类使用前期工作提出的 BA2Cs 模型^[30], 该模型分别约束正类和负类拒绝率, 同时最大化 ROC 曲线下面积 AUC:

$$\max_{t_1, t_2} R_{\text{auc}}(t_1, t_2),$$

$$\text{s. t. } \begin{cases} r_{\text{tp}}(t_1, t_2) \leq p_{\text{max}} \\ r_{\text{m}}(t_1, t_2) \leq n_{\text{max}} \\ t_1 < t_2 \end{cases}$$

式中: p_{max} 和 n_{max} 为 BA2Cs 模型的超参数, $R_{\text{auc}}(t_1, t_2)$ 定义为^[31]

$$R_{\text{auc}}(t_1, t_2) = \frac{1 + r_{\text{tp}}(t_1, t_2) - r_{\text{fp}}(t_1, t_2)}{2}.$$

对比以前拒绝分类模型, BA2Cs 模型有以下优点: 1) 不需要设置代价, 通过分别约束正负类的拒绝率控制每类性能. 虽然这些模型都需要设置参数, 但是相比无界的代价值, 拒绝参数 p_{max} 和 n_{max} 是 $[0, 1]$ 内有界值, 参数设置有针对性, 且不同拒绝参数下性能可由性能-拒绝曲线直观展示. 2) BA2Cs 模型优化 AUC, 而不是错误率, AUC 对不平衡类分布和不相等错分代价不敏感^[32-33]; 分别控制两类的拒绝率, 而不是只约束总体拒绝率, 处理不平衡数据时更鲁棒. BA2Cs 模型的详细介绍及其有效性验证请查看文献[30].

使用穷举搜索法求解 BA2Cs 优化问题. 假设已使用得分分类器求得 n 个训练样本的置信得分, n 个得分在 $[s_{\text{min}}, s_{\text{max}}]$ 区间内. 穷举搜索即把 $[s_{\text{min}}, s_{\text{max}}]$ 区间内的值分配给 t_1 和 t_2 , 计算相应的 $R_{\text{auc}}(t_1, t_2)$, $r_{\text{tp}}(t_1, t_2)$ 和 $r_{\text{m}}(t_1, t_2)$. 最后, 获得最大 AUC 且同时满足两个约束条件的参数组合即为经验最优的拒绝阈值. 由于 $[s_{\text{min}}, s_{\text{max}}]$ 区间内有无限个实数, 实验中, 以 $(s_{\text{max}} - s_{\text{min}}) / k$ 为步长划分原区间, 并用 k 个小区间的端点值作为 t_1 和 t_2 的候选值, k 取值 200.

3 实验

3.1 实验流程和评价指标

由于本研究中超声和钼靶图像数量较少, 因此实验采用留一法 (leave-one-out) 交叉验证程序, 恶性样本为正例, 良性样本为负例.

拒绝方法验证实验流程: 数据集包含 219 个样例, 其中 92 个正例和 127 个负例样本. 每次选择 218 个样本建立拒绝分类器, 对剩余一个样本进行测试, 测试结果可能为良性、恶性或无标签, 整个过程执行 219 次, 计算性能指标.

级联方法验证实验流程: 第一阶段用拒绝分类方法对 27 个 (12 个正例和 15 个负例) 超声样本构成的数据集进行分类: 每次选择一个超声样本作为测试样本, 用剩余的 26 个超声样本作为训练集学习拒绝分类器, 即用训练样本确定两个拒绝阈值 t_1 和 t_2 ; 在测试阶段, 首先计算测试样本的置信得分, 然

后根据拒绝规则式(1)给该测试样本分配标签或暂时拒绝分类. 如果该测试样本被拒绝, 则执行第二阶段分类: 把未分配标签的肿瘤对应的钼靶样本作为测试样本, 剩余的钼靶样本作为训练集学习一个 SVM 分类器, 对测试样本进行分类. 整个过程执行 27 次, 直到每个超声样本作为测试样本, 性能指标综合两阶段的分类结果计算.

评估分类系统的性能指标有正确率 (M_{acc})、AUC (M_{auc})、G-mean (M_g)、敏感性 (M_{sen})、特异性 (M_{spe})、阳性预测值 (M_{ppv})、阴性预测值 (M_{npv}) 和马修斯相关系数 (M_{mcc}), 分别定义为

$$M_{acc} = \frac{N_{tp} + N_{tn}}{N_{tp} + N_{fn} + N_{tn} + N_{fp}}$$

$$M_{auc} = \frac{M_{sen} + M_{spe}}{2}$$

$$M_g = \sqrt{M_{sen} \times M_{spe}}$$

$$M_{sen} = \frac{N_{tp}}{N_{tp} + N_{fn}}$$

$$M_{spe} = \frac{N_{tn}}{N_{tn} + N_{fp}}$$

$$M_{ppv} = \frac{N_{tp}}{N_{tp} + N_{fp}}$$

$$M_{npv} = \frac{N_{tn}}{N_{tn} + N_{fn}}$$

$$M_{mcc} = \frac{N_{tp} \times N_{tn} - N_{fp} \times N_{fn}}{\sqrt{(N_{tp} + N_{fp}) \times (N_{tp} + N_{fn}) \times (N_{tn} + N_{fp}) \times (N_{tn} + N_{fn})}}$$

3.2 实验参数设置

拒绝方法验证实验参数设置: 拒绝分类方法采用线性核 SVM 作为得分分类器. 拒绝超参数 p_{max} 和 n_{max} 设为相同值, 为从 0.01 到 0.2, 以 0.01 为步长, 计算每组参数下的性能指标, 并画出性能-拒绝曲线. 没有选择大于 0.2 的参数是因为更大的拒绝率在实际中没有意义.

级联方法验证实验参数设置: 由于 SVM 适合小样本学习, 因此该实验使用 SVM, 即第一阶段拒绝分类方法的得分分类器为 SVM, 第二阶段对钼靶特征进行分类使用 SVM, 实验参数是 MATLAB R2017a 默认参数, 但是分类时选择用线性核还是径向基核. 表 2 和表 3 为使用线性核 SVM 和径向基核 SVM 分别对超声和钼靶特征及其选择后特征进行分类的结果. 可看出, 在超声集上使用线性核 SVM 分类性能更优, 在钼靶集上使用径向基核 SVM 分类性能更优, 因此得分分类器使用线性核 SVM, 第二阶段使用径向基核 SVM 进行分类.

严格来说, 拒绝超参数 p_{max} 和 n_{max} 为被拒绝样本的比例, 但是由于本实验样本数非常小, 因此 p_{max} 和 n_{max} 设为拒绝样本个数, 分别设置为 [1, 5] 区间内的整数, 这样共有 25 个参数组合. 对于每个参数组合, 用拒绝分类器对超声样本进行分类, 对于被拒绝的肿瘤, 用钼靶特征进行分类, 最后统计两阶段综合分类结果.

表 2 线性核 SVM 和径向基核 SVM 在超声数据集上的分类结果

Tab.2 Results of using SVM with linear and radius basis function kernels on ultrasound datasets

分类方法	特征选择	$M_{acc}/\%$	M_{auc}	$M_g/\%$	$M_{sen}/\%$	$M_{spe}/\%$	$M_{ppv}/\%$	$M_{npv}/\%$	M_{mcc}
线性核 SVM	无	81.48	0.816 7	81.65	80.00	83.33	85.71	76.92	0.629 8
	有	88.89	0.891 7	89.13	86.67	91.67	92.86	84.62	0.779 0
径向基核 SVM	无	74.07	0.725 0	71.10	86.67	58.33	72.22	77.78	0.474 3
	有	81.48	0.791 7	76.38	100.00	58.33	75.00	100.00	0.661 4

注: 表中每个指标最优值加粗显示

表 3 线性核 SVM 和径向基核 SVM 在钼靶数据集上的分类结果

Tab.3 Results of using SVM with linear and radius basis function kernels on mammogram datasets

分类方法	特征选择	$M_{acc}/\%$	M_{auc}	$M_g/\%$	$M_{sen}/\%$	$M_{spe}/\%$	$M_{ppv}/\%$	$M_{npv}/\%$	M_{mcc}
线性核 SVM	无	96.30	0.958 3	95.74	100.00	91.67	93.75	100.00	0.927 0
	有	96.30	0.958 3	95.74	100.00	91.67	93.75	100.00	0.927 0
径向基核 SVM	无	100.00	1.000 0	100.00	100.00	100.00	100.00	100.00	1.000 0
	有	100.00	1.000 0	100.00	100.00	100.00	100.00	100.00	1.000 0

注: 表中每个指标最优值加粗显示

表 4 拒绝方法验证实验的对比结果

Tab. 4 Results of compared methods in the experiment of validating the rejection method

分类方法	特征选择	$M_{acc}/\%$	M_{auc}	$M_g/\%$	$M_{sen}/\%$	$M_{spe}/\%$	$M_{ppv}/\%$	$M_{npv}/\%$	M_{mcc}
文献[11]	无	65.75	0.631 3	60.97	46.74	79.53	62.32	67.33	0.279 1
线性核 SVM	无	87.21	0.868 8	86.85	84.78	88.98	84.78	88.98	0.737 6
	有	88.58	0.880 6	88.00	84.78	91.34	87.64	89.23	0.765 0
径向基核 SVM	无	86.30	0.851 9	84.91	78.26	92.13	87.80	85.40	0.717 8
	有	88.13	0.873 7	87.24	82.61	92.13	88.37	87.97	0.755 3
BA2C _S ($p_{max} = n_{max} = 0.04$)	无	88.94	0.895 1	89.43	93.18	85.83	82.83	94.50	0.781 6
BA2C _S ($p_{max} = n_{max} = 0.07$)	无	90.82	0.908 0	90.80	90.70	90.91	87.64	93.22	0.812 3
BA2C _S ($p_{max} = n_{max} = 0.03$)	有	88.37	0.890 9	88.99	93.26	84.92	81.37	94.69	0.771 1
BA2C _S ($p_{max} = n_{max} = 0.07$)	有	91.18	0.914 3	91.42	92.86	90.00	86.67	94.74	0.821 3

注:表中每个指标最优值加粗显示

3.3 实验结果

3.3.1 拒绝方法验证实验结果

图 2 为不同拒绝率参数下拒绝分类方法的性能指标曲线:(a)为原始特征上的性能曲线(b)为特征选择后的性能曲线.从图 2 可看到,整体来说,随着拒绝率的上升,性能指标值逐渐增大,这验证了拒绝方法的有效性.另外,当拒绝参数大于 0.07,性能变化趋于平缓.

选取两个方法进行实验对比,一是文献[11]中的乳腺超声诊断方法,二是使用线性核 SVM 和径向基核 SVM 对包含 219 个样本的超声集进行分类.对比实验选择文献[11]中性能较好的 HOG 特征,使用随机森林分类,即利用 bagging 集成多个决策树的分类结果,在本实验中,分别测试了 1~40 个决策树的随机森林,当决策树个数为 23 时,性能最优.进行方法对比时,对于原始特征集展示参数 p_{max} 和 n_{max} 为 0.04 和 0.07 时的结果,对于特征选择后的特征集展示参数为 0.03 和 0.07 时的结果,方法结果总结如表 4 所示.从表中看出,拒绝分类方法的性能优于其他对比方法,且拒绝参数为 0.07 时在特征选择后的数据集上建模的拒绝方法表现最优.另外,只比较线性核 SVM 和拒绝方法,发现后者性能优于前者,这表明通过拒绝部分不确定样本,可提高分类性能,进一步验证了拒绝方法的有效性.

3.3.2 级联方法验证实验结果

本实验对比方法为文献[11]采用超声图像分类方法和使用线性核 SVM 对包含 27 个样本的超声集进行分类的方法.文献[11]中用随机森林进行分类,通过多次测试,选择性能最优的 10 个决策树组成的随机森林作为分类器.

由于级联方法中 p_{max} 和 n_{max} 分别设置 5 个不同值,因此有 25 个参数组合的结果,选择分类结果最好的那组进行对比,实验对比结果见表 5.相比文献[11]和使用线性核 SVM 超声诊断的方法,提出的基于超声和钼靶的级联分类方法获得更好的分类性能.尤其是,级联分类方法结果优于超声诊断方法,这表明第二阶段通过增加钼靶信息提高了整体系统的性能.

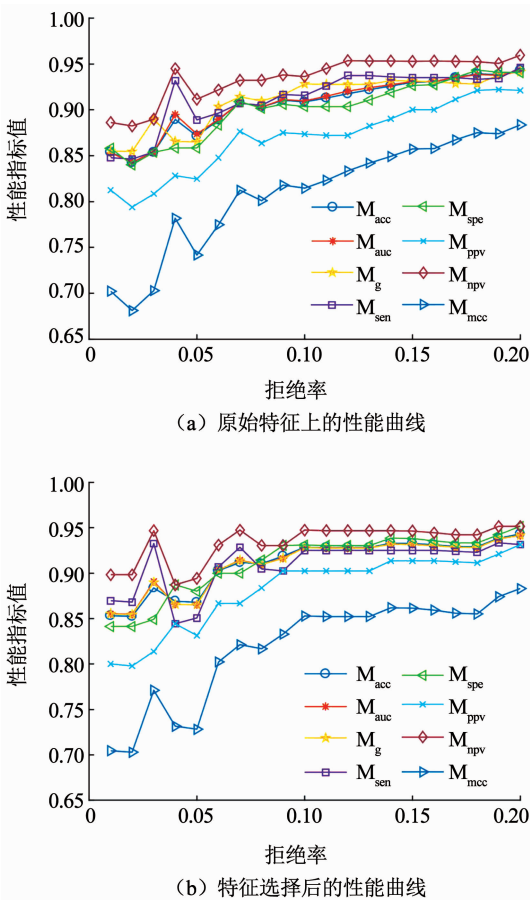


图 2 不同拒绝率参数下拒绝分类方法的性能指标

Fig. 2 Results of the proposed rejection method under different abstaining parameters

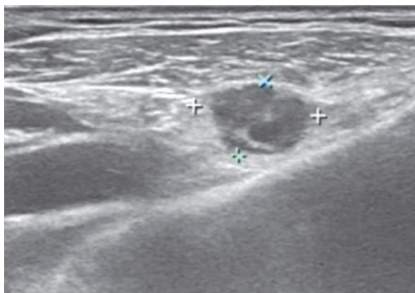
表 5 级联方法验证实验的对比结果

Tab. 5 Results of compared methods in the experiment of validating the cascade method

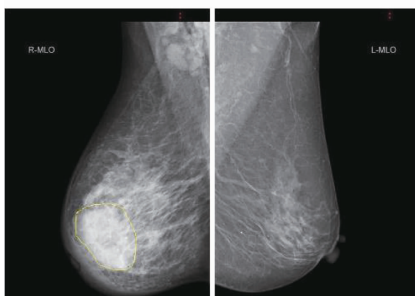
分类方法	特征选择	$M_{acc}/\%$	M_{auc}	$M_g/\%$	$M_{sen}/\%$	$M_{spe}/\%$	$M_{ppv}/\%$	$M_{npv}/\%$	M_{mcc}
文献[11]	无	74.07	0.725	71.10	86.67	58.33	72.22	77.78	0.474 3
本文超声方法	无	81.48	0.816 7	81.65	80.00	83.33	85.71	76.92	0.629 8
	有	88.89	0.891 7	89.13	86.67	91.67	92.86	84.62	0.779 0
本文两阶段方法	无	92.59	0.933 3	93.09	86.67	100.00	100.00	85.71	0.861 9
	有	92.59	0.933 3	93.09	86.67	100.00	100.00	85.71	0.861 9

注:表中每个指标最优值加粗显示

图 3 提供了一个被拒绝的恶性肿瘤的(a)超声和(b)钼靶图像.用 SVM 对该恶性肿瘤的超声样本进行分类,该样本被错分为良性,但是用拒绝分类器对该样本分类时,该样本未被赋予标签,当第二阶段中对这个肿瘤的钼靶样本进行分类时得到了正确的诊断结果.



(a) 超声图像



(b) 钼靶图像

图 3 一个被拒绝分类的恶性肿瘤的超声和钼靶图像

Fig. 3 Breast ultrasound and mammogram images of a rejected malignant tumor

4 结 论

早期筛查和及时治疗可以降低乳腺癌的死亡率.本文提出了结合乳腺超声和钼靶的计算机辅助诊断系统,该系统使用两阶段级联分类方法辨别肿瘤良恶性,其中拒绝分类器可以筛选不确定肿瘤样本,结合多模态信息重新对这些样本进行诊断.相比只利用超声图像的诊断方法,该系统的两阶段级联方法分类性能显著提高,这表明利用钼靶图像辅助超声诊断有利于提高乳腺肿瘤诊断性能.

参 考 文 献

- [1] FERLAY J, COLOMBET M, SOERJOMATARAM I, et al. Estimating the global cancer incidence and mortality in 2018; GLOBOCAN sources and methods[J]. International Journal of Cancer, 2019, 144(8): 1941–1953
- [2] FAN Lei, STRASSER-WEIPPL K, Lunjie, et al. Breast cancer in China[J]. The Lancet Oncology, 2014, 15(7): e279–e289
- [3] OHUCHI N, SUZUKI A, SOBUE T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): A randomised controlled trial[J]. The Lancet, 2016, 387(10016): 341–348
- [4] CUBASCH H, DICKENS C, JOFFE M, et al. Breast cancer survival in Soweto, Johannesburg, South Africa; A receptor-defined cohort of women diagnosed from 2009 to 11[J]. Cancer Epidemiology, 2018, 52: 120–127
- [5] OHSUMI S, SAKAMOTO G, TAKASHIMA S, et al. Long-term results of breast-conserving treatment for early-stage breast cancer in Japanese women from multicenter investigation[J]. Japanese Journal of Clinical Oncology, 2003, 33(2): 61–67
- [6] FEIG S. Cost-effectiveness of mammography, MRI, and ultrasonography for breast cancer screening[J]. Radiologic Clinics, 2010, 48(5): 879–891
- [7] LEE H, CHEN Y P P. Image based computer aided diagnosis system for cancer detection[J]. Expert Systems with Applications, 2015, 42(12): 5356–5365
- [8] CHEN S C, CHEUNG Y C, SU C H, et al. Analysis of sonographic features for the differentiation of benign and malignant breast tumors of different sizes[J]. Ultrasound in Obstetrics and Gynecology, 2004, 23(2): 188–193
- [9] LI Lujing, ZHOU Xinchuan, ZHAO Xinbao, et al. B-mode ultrasound combined with color Doppler and strain elastography in the diagnosis of non-mass breast lesions: A prospective study[J]. Ultrasound in Medicine & Biology, 2017, 43(11): 2582–2590
- [10] SINGH B K, VERMA K, PANIGRAHI L, et al. Integrating radiologist feedback with computer aided diagnostic systems for breast cancer risk prediction in ultrasonic images; An experimental investigation in machine learning paradigm[J]. Expert Systems with Applications, 2017, 90: 209–223
- [11] ABDEL-NASSER M, MELENDEZ J, MORENO A, et al. Breast tumor classification in ultrasound images using texture analysis and super-resolution methods[J]. Engineering Applications of Artificial Intelligence, 2017, 59: 84–92
- [12] ZHOU Shichong, SHI Jun, ZHU Jie, et al. Shearlet-based texture

- feature extraction for classification of breast tumor in ultrasound image[J]. *Biomedical Signal Processing and Control*, 2013, 8(6): 688–696
- [13] CAI Lingyun, WANG Xin, WANG Yuanyuan, et al. Robust phase-based texture descriptor for classification of breast ultrasound images[J]. *Biomedical Engineering Online*, 2015, 14(1): 26
- [14] SHAN Juan, ALAM S K, GARRA B, et al. Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods[J]. *Ultrasound in Medicine & Biology*, 2016, 42(4): 980–988
- [15] TAI Shenchuan, CHEN Z S, TSAI W T. An automatic mass detection system in mammograms based on complex texture features[J]. *IEEE Journal of Biomedical and Health Informatics*, 2014, 18(2): 618–627
- [16] SHASTRI A A, TAMRAKAR D, AHUJA K. Density-wise two stage mammogram classification using texture exploiting descriptors[J]. *Expert Systems with Applications*, 2018, 99: 71–82
- [17] BEURA S, MAJHI B, DASH R. Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer [J]. *Neurocomputing*, 2015, 154: 1–14
- [18] CHAKRABORTY J, MIDYA A, RABIDAS R. Computer-aided detection and diagnosis of mammographic masses using multi-resolution analysis of oriented tissue patterns[J]. *Expert Automated Breast Cancer Detection and Classification Systems with Applications*, 2018, 99: 168–179
- [19] BANAIE M, SOLTANIAN-ZADEH H, SALIGHEH-RAD H R, et al. Spatiotemporal features of DCE-MRI for breast cancer diagnosis[J]. *Computer Methods and Programs in Biomedicine*, 2018, 155: 153–164
- [20] LIU Yan, CHENG H D, HUANG Jianhua, et al. An effective approach of lesion segmentation within the breast ultrasound image based on the cellular automata principle[J]. *Journal of Digital Imaging*, 2012, 25(5): 580–590
- [21] CHENG H D, SHAN Juan, JU Wen, et al. Automated breast cancer detection and classification using ultrasound images: A survey[J]. *Pattern Recognition*, 2010, 43(1): 299–317
- [22] JOO S, YANG Y S, MOON W K, et al. Computer-aided diagnosis of solid breast nodules; Use of an artificial neural network based on multiple sonographic features[J]. *IEEE Transactions on Medical Imaging*, 2004, 23(10): 1292–1300
- [23] 刘研. 基于双模态乳腺超声图像的良恶性分类及其关键技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2013
- LIU Yan. Research on key technologies of benign and malignant classification based on dual-mode breast ultrasound images[D]. Harbin, Harbin Institute of Technology, 2013
- [24] ROFFO G, MELZI S, CASTELLANI U, et al. Infinite latent feature selection: A probabilistic latent graph-based ranking approach[C]//*Proceedings of the IEEE International Conference on Computer Vision*. Washington, DC, USA: IEEE Press, 2017: 1398–1406
- [25] LIN Dongyun, SUN Lei, TOH K A, et al. Biomedical image classification based on a cascade of an svm with a reject option and subspace analysis[J]. *Computers in Biology and Medicine*, 2018, 96: 128–140
- [16] WANG Zhanwei, WANG Zhiwei, HE Suwei, et al. Fault detection and diagnosis of chillers using Bayesian network merged distance rejection and multi-source non-sensor information[J]. *Applied Energy*, 2017, 188: 200–214
- [27] TORTORELLA F. An optimal reject rule for binary classifiers[C]//*Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*. Berlin, Heidelberg: Springer, 2000: 611–620
- [28] TORTORELLA F. Reducing the classification cost of support vector classifiers through an ROC-based reject rule[J]. *Pattern Analysis and Applications*, 2004, 7(2): 128–143
- [29] PIETRASZEK T. On the use of ROC analysis for the optimization of abstaining classifiers[J]. *Machine Learning*, 2007, 68(2): 137–169
- [30] GUAN Hongjiao, ZHANG Yingtao, CHENG H D, et al. BA2Cs: Bounded abstaining with two constraints of reject rates in binary classification[J]. *Neurocomputing*, 2019, 357: 125–134
- [31] LÓPEZ V, FERNÁNDEZ A, GARCÍA S, et al. An insight into classification with imbalanced data; Empirical results and current trends on using data intrinsic characteristics[J]. *Information Sciences*, 2013, 250: 113–141
- [32] FAWCETT T. ROC graphs: Notes and practical considerations for researchers[J]. *Machine Learning*, 2004, 31(1): 1–38
- [33] PRATI R C, BATISTA G E, MONARD M C. A survey on graphical methods for classification predictive performance evaluation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2011, 23(11): 1601–1618

(编辑 苗秀芝)