

深度学习联合超声在乳腺BI-RADS分类中的价值应用

任 静^{1,2*}, 韩保凤², 路子妍^{1,2}, 卞福勤², 马 芳^{1,2#}

¹安徽医科大学研究生学院, 安徽 合肥

²合肥市第二人民医院超声医学科, 安徽 合肥

收稿日期: 2025年8月23日; 录用日期: 2025年9月17日; 发布日期: 2025年9月26日

摘 要

目的: 评估人工智能深度学习(DL)软件优化乳腺超声BI-RADS分类的价值及联合诊断效能。方法: 以病理结果作为金标准, 回顾性分析237例乳腺肿块(良性153例, 恶性84例)。由低年资医师按ACR BI-RADS评估肿块特征, 对比DL系统诊断性能, 对不一致病例分级调整后实施联合诊断。结果: 超声医师诊断结果的ROC曲线下面积为0.710, AI软件为0.838; 二者诊断不一致的病例进行联合诊断后, AUC提升至0.957。在形态学特征(形状、方向)的判断上, 医师与AI的一致性较高; 在钙化判断上一致性中等; 而在内部回声、边缘、后方回声的判断上一致性较差。结论: 与单一超声医师诊断相比, AI深度学习诊断系统联合超声医师优化BI-RADS分类后对乳腺癌的良恶性诊断效果更好, 联合诊断能够更好地为临床决策提供更高的价值, 降低不必要的活检率, 避免超声诊断存在的主观性缺陷, 人工智能辅助医师在诊断的过程中实现了量化及标准化, 具有一定的临床推广意义。

关键词

乳腺, 超声, 乳腺影像报告和数据系统(BI-RADS), AI深度学习(DL)

The Evaluating the Utility of Deep Learning Combined with Ultrasound in Breast BI-RADS Classification

Jing Ren^{1,2*}, Baofeng Han², Ziyang Lu^{1,2}, Fuqin Bian², Fang Ma^{1,2#}

¹Graduate School, Anhui Medical University, Hefei Anhui

²Department of Ultrasound Medicine, Hefei Second People's Hospital, Hefei Anhui

Received: Aug. 23rd, 2025; accepted: Sep. 17th, 2025; published: Sep. 26th, 2025

*第一作者。

#通讯作者。

文章引用: 任静, 韩保凤, 路子妍, 卞福勤, 马芳. 深度学习联合超声在乳腺 BI-RADS 分类中的价值应用[J]. 临床医学进展, 2025, 15(9): 1843-1850. DOI: 10.12677/acm.2025.1592691

Abstract

Objective: To evaluate the value and combined diagnostic efficacy of artificial intelligence (AI) deep learning (DL) software in optimizing BI-RADS classification for breast ultrasound. **Methods:** Using pathological results as the gold standard, we retrospectively analyzed 237 breast masses (153 benign, 84 malignant). Junior physicians assessed mass features according to ACR BI-RADS. DL system performance was compared, and discordant cases underwent adjusted grading for combined diagnosis. **Results:** The AUC for ultrasound physicians was 0.710, compared to 0.838 for the AI software. Combined diagnosis of discordant cases significantly increased the AUC to 0.957. Physician-AI agreement was high for morphological features (shape, orientation) and moderate for calcification assessment, but poor for internal echogenicity, margins, and posterior acoustic features. **Conclusion:** Compared to single-physician diagnosis, the AI DL system combined with physician assessment significantly improves diagnostic accuracy for breast lesion malignancy. This combined approach enhances clinical decision-making, reduces unnecessary biopsies, mitigates subjective diagnostic limitations, and introduces quantification and standardization, demonstrating significant clinical value for broader adoption.

Keywords

Mammary Gland, Ultrasonic, Breast Imaging Reporting and Data System (BI-RADS), AI Deep Learning (DL)

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

2022 年全球癌症统计数据表明,乳腺癌位居女性新发癌种首位,成为导致女性死亡主因(占比约 15%) [1]。超声(Ultrasound, US)以其简单、方便、无辐射、价格低廉等优点,被广泛推荐为乳腺早期筛查的首选检查方式[2]。然而,传统的超声检查方式,医师往往会根据自己的经验来进行判断,缺乏统一、规范的标准,最后就会导致结果大相径庭。美国放射学会(American College of Radiology, ACR)制定的超声乳腺影像报告和数据系统(Breast ultrasound breast imaging reporting and data system, US BI-RADS)为乳腺超声的诊断报告提供了统一的规范。然而超声检查存在“同图异病”、“同病异图”等特点,导致超声医师操作过程中存在一定的主观因素,其中 BI-RADS 3 类与 4 类乳腺良恶性跨度较大(3%~94%为 BI-RADS 4 类恶性),从而会导致假阳性率较高,所以如何精确分级是目前研究的难点[3] [4]。

常规的超声作为乳腺的早期筛查方法在大规模的社区人群中的应用受到明显的制约,为了突破这一限制,北京协和医院通过技术创新(如光声成像、自动乳腺容积成像 ABVS)和全国性筛查体系建立,推动了中国乳腺超声诊断从 A 超时代迈向多模态 AI 时代[5]。近年来,人工智能(Artificial intelligence, AI)在各个领域的持续爆发,在超声医疗领域也得到了大量的应用,人工智能辅助超声可以帮助医师快速准确地筛查出优先需要诊断的病例,使筛查的灵敏度和特异度得到了很大的提升,从而有效降低了漏诊率,提高诊断的准确性,降低活检率,特别是对年资较低的医师具有一定的参考价值[6]。

本研究将比较超声医师和人工智能深度学习(Deep Learning, DL)模型联合 BI-RADS 分类对良恶性的诊断价值,探讨深度学习模型对乳腺癌的诊断价值以及是如何辅助医师提高诊断水平,从而实现更为准

确的 BI-RADS 分级,降低活检率。

2. 资料与方法

2.1. 一般资料

回顾性纳入 2020 年 9 月至 2024 年 12 月于安徽医科大学附属合肥市第二人民医院接受乳腺超声检查且有手术病理或穿刺病理的患者。最终纳入 230 例女性患者(共 237 个结节)。年龄范围 16~88 岁(均值 $\pm$ 标准差: 46.1 ± 14.4 岁),所有病灶均经组织病理学确诊(良性 153 例,恶性 84 例)。

纳入标准:① 超声评估为 BI-RADS 3 类或 4 类(4a/4b/4c)病灶;② 原始超声图像完整且未添加测量标识(避免 AI 分析干扰);③ 具有明确的手术或穿刺活检病理结果。

排除标准:① AI 检测系统无法自动分割病灶;② 正在接受新辅助化疗或放疗者;③ 病灶最大径超过整个探头扫查范围;④ 图像显示不标准(病灶显示不清/伪影遮挡)。

2.2. 常规超声检查

超声检查收集的所有乳腺图像数据均来自安徽省合肥市第二人民医院超声医学科,超声仪器型号包括 SUMSUMG、ACUSON Sequoia Silver、迈瑞 Resona R9s 等,探头使用频率为 5~14 MHz 的宽频线阵探头。

检查流程:患者取仰卧位,充分暴露双侧乳腺及腋窝区域。由两位具有乳腺浅表超声经验的医师(5 年以上经验)在未知病理结果的情况下,采用多切面放射状扫查方式,动态扫查病灶,观察并记录肿块的大小、形态、内部回声、方向、边缘特征、后方回声特征及彩色血流频谱情况等,当两名医师诊断结果不一致时,两名医师进行商讨决定,并以 DICOM 格式存取相关声像图资料。上述操作超声医师均按照安徽省超声质量控制要求和《超声医学》第 6 版乳腺超声操作规范要点及诊断标准。

2.3. 乳腺超声 AI 实时辅助检测系统

进行上述超声诊断的医师接着使用乳腺超声 AI 实时辅助诊断系统(杏脉·瑞声超影超声实时辅助检测系统,上海杏脉信息科技有限公司)进行检测,每个病灶至少选取两张超声图片以获得 AI 的最佳诊断性能。AI 服务器直接连接存有乳腺资料影像的计算机 HDMI 接口,AI 系统以 64 帧/秒的速度实时地勾勒出可疑病灶(见图 1),通过深度(DL)分析,提取超声诊断医师肉眼无法识别的良恶性病变特征,结合深度学习(DL)分析和 ACR BI-RADS 指南的全属性自动分析,生成智能化 US BI-RADS 诊断结果,内容包括肿块的大小、形状、方位、边缘、内部回声、后方回声等,智能化结果见图 1。

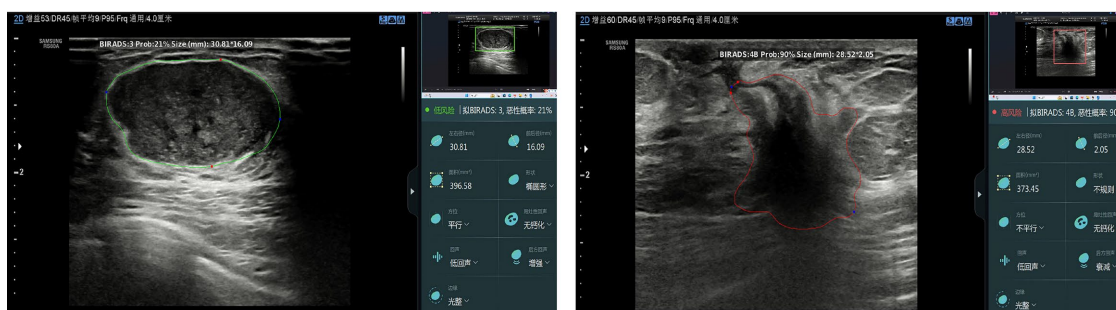


Figure 1. AI-based automatic identification and outlining of benign and malignant breast nodules

图 1. AI 自动识别、勾画乳腺良恶性结节的结果

2.4. 应用 AI 深度学习后升降级的标准

对比医师诊断结果和 AI 深度学习系统诊断结果,要优化调整两者诊断不一致的 BI-RADS 分类结果,

将 4A 类及其以上归为恶性范围, 3 类归为良性范围。当医师组诊断为 4A, AI 组诊断为 3 类时, 将 4A 降为 3 类, 其余诊断参照超声医师结果不发生变化。进行联合调级后综合判断, 再分别与医师组和 AI 组诊断系统诊断结果进行比较。

3. 统计学方法

所有数据录入 Excel 表建库, 分类数据用四格表的形式进行整理。统计分析使用了 SPSS 27.0 统计学软件进行数据分析。计量资料采用 $\bar{x} \pm s$ 表示, 计数资料采用频数及构成比(%)表示, 组间比较用 χ^2 检验。组织学病理结果作为金标准, 将乳腺病变良恶性进行二分类, 计算两种方法诊断效能: 灵敏度(sensitivity, Sen)、特异度(specificity, Spe)、假阳性率(false positive rate, FPR)、假阴性率(false negative rate, FNR)、阳性似然比(positive likelihood ratio, LR^+)、阴性似然比(negative likelihood ratio, LR^-)。参照病理良恶性, 绘制计算受试者工作特征(receiver operating characteristic, ROC)曲线, 并对曲线下面积(Area under the curve, AUC)进行计算。不同诊断方法的 ROC 曲线下面积比较采用 Z 检验。不同诊断方式使用 Kappa 值对同一结节结果一致性检验, 简称 K 值, K 值的取值范围为-1~1, 一般 K 值 ≥ 0.75 为一致性极好, K 值在 0.4~0.75 为中、高度一致, K 值 ≤ 0.4 为一致性差。使用 Z 检验及相应的 P 值用于评估 Kappa 系数(K 值)的统计学显著性。以 $P < 0.05$ 认为差异有统计学意义。

公式: 活检率计算公式 = (超声医师/AI/两者联合)超声检查后进行活检的病例数/进行超声检查的总病例数 $\times 100\%$

4. 结果

4.1. 术后病理结果

有明确手术或穿刺病理结果的结节有 237 个。所有乳腺病变的病理学类型见表 1, 乳腺良性病灶病理类型中最多的是乳腺纤维腺瘤(75/153, 49.02%), 乳腺恶性病灶病理类型最多的是浸润性癌(65/84, 77.38%)。

Table 1. Distribution of pathological types of 237 cases of breast lesions

表 1. 237 例乳腺病变病理类型分布

良性病灶病理类型	频数/例	百分比/%	恶性病灶病理类型	频数/例	百分比/%
乳腺纤维腺瘤	75	49.02%	浸润性癌	65	77.38%
乳腺腺病	59	38.56%	导管原位癌	7	8.33%
乳腺良性增生性疾病	8	5.22%	梭形细胞肿瘤	2	2.38%
分叶状肿瘤	4	2.61%	浸润性小叶癌	2	2.38%
浆细胞性乳腺炎	2	1.31%	黏液癌	2	2.38%
肉芽肿性炎	2	1.31%	混合型浸润癌	1	1.19%
泌乳性腺瘤	1	0.6%	导管内乳头状癌	3	3.57%
管状腺瘤	2	1.31%	浸润性髓样癌	2	2.38%
总计	153	100%	总计	84	100%

4.2. 医师组与 AI 组的诊断效能

组织病理学作为最终诊断的“金标准”, 医师组与 AI 组分别得出的诊断效能对照结果(见表 2)。AI 的阳性似然比($LR^+ = 3.347$)是医师($LR^+ = 1.755$)的 1.9 倍, 反映了 AI 阳性诊断对乳腺恶性肿瘤的预测能力

更高。AI 的阴性似然比($LR^- = 0.050$)低于医师的阴性似然比($LR^- = 0.054$)，表明了 AI 排除乳腺肿块疾病的能力更好。

Table 2. Diagnostic performance comparison of two methods for selected breast lesions

表 2. 两种不同诊断方法对入选乳腺病灶的诊断效能对比

	Sen	Spe	FPR	FNR	LR^+	LR^-
医师诊断	0.976	0.444	0.556	0.024	1.755	0.054
AI 诊断	0.964	0.712	0.288	0.036	3.347	0.050

4.3. 超声医师独立诊断结果、AI 独立诊断结果，两者联合诊断结果以及各自 ROC 曲线

本研究将 4A 及其以上病变定为活检指征。超声医师单独诊断恶性 167 例(活检率 70.46%，167/237)；AI 单独诊断恶性 125 例(活检率 52.74%，125/237)；联合诊断结果：恶性 119 例(活检率 50.21%，119/237)。应用 AI 深度学习软件后，联合诊断的活检率较超声医师单独诊断的活检率降低(70.46%到 50.21%)。医师诊断、AI 诊断的 AUC 分别为 0.710 (95% CI: 0.646~0.774)、0.838 (95% CI: 0.787~0.890)。ROC 曲线见图 2，ROC 曲线分析证实 AI 在维持高敏感度(>0.96)的同时提升 AUC 值(0.892 vs. 0.811)。超声联合 AI 诊断的 AUC 为 0.840 (95% CI: 0.787~0.892)。

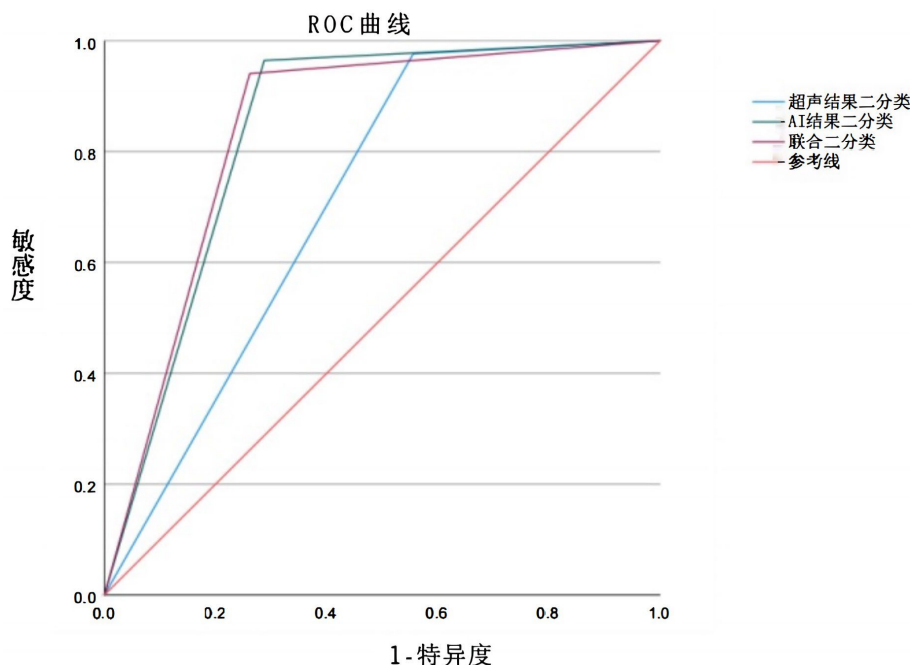


Figure 2. ROC Curves for radiologist versus AI breast cancer diagnosis using BI-RADS categories

图 2. 医师、AI 以及两者联合二分类诊断乳腺癌的 ROC 曲线

4.4. 医师和 AI 诊断对病灶组学特征分析

采用 ACRBI-RADS 标准进行 BI-RADS 分类，对乳腺肿块特征分析，见表 3，经统计学分析后得出医师和 AI 诊断在形状方面 K 值为 0.621，在方向特征中 K 值为 0.446，两者均在 0.4~0.75 之间，为中、高度一致，具有统计学意义($P < 0.05$)，医师和 AI 诊断在内部回声、后方回声及边缘组学特征诊断一致性差，无统计学意义($P > 0.05$)；对于钙化的判断虽然显著，但是 Kappa 值仅为 0.393。

Table 3. Radiologist and AI analysis of lesion radiomics features
表 3. 医师和 AI 对病灶组学特征分析

特征	医师/例	AI/例	K 值	Z 值	P 值
形状			0.621	10.127	0.000
圆/椭圆形	91	131			
不规则形	146	106			
方向			0.446	7.048	0.000
平行	212	200			
不平行	25	37			
局灶回声			0.393	6.072	0.000
无钙化	180	189			
有钙化	56	47			
内部回声			0.111	2.307	0.021
低回声	230	205			
非低回声	7	32			
后方回声			0.134	2.822	0.005
无改变	210	142			
有改变	26	94			
边缘			0.149	3.949	0.000
光整	119	214			
不光整	116	21			

5. 讨论

目前乳腺癌的影像学筛查手段包括超声、乳腺 X 线摄影(钼靶)、磁共振成像(MRI)等,其中乳腺 X 线摄影(钼靶)对钙化结节的乳腺病灶检测效果佳,是欧美国家的主要筛查方式,但其在致密型乳腺组织中的灵敏度、特异度不高,因此不适用于腺体致密的亚洲女性[7];磁共振成像虽然具有优异的软组织的分辨能力,但因操作复杂、费用昂贵,通常不作为初筛的首选。相比之下,超声以其便捷、价廉、无创等优势,已成为目前筛查的首选,广泛应用于乳腺结节的良恶性鉴别。

近年来,人工智能在乳腺超声(特别是基于 AI-US BI-RADS 标准)的病灶检测和生存预测的发展迅猛。相关文献从 2020 年的 12 篇激增至 2021 年的 546 篇。且早期研究多采用传统机器学习(Machine Learning, ML)算法(如支持向量机 SVM)用于乳腺 X 线图像分类。到 2015 年后,DL (尤其是 CNN)成为主流技术,逐渐成为提升乳腺癌检测和生存预测的准确性的主流技术[8]。传统的人工智能需要依赖手动勾画病灶区域(ROC)。然而,由于受低分辨率及肿块边界模糊等因素的影响,手动分割操作困难,其结果仅仅局限于肿块的二分类,难以进行细致的 BI-RADS 分类,应用存在明显的局限性。此外,超声的可重复性很容易受到操作者的经验、患者个体差异以及仪器设备性能的影响。但是本研究采用的人工智能技术是基于深度学习框架和先进的数据处理原理,能够识别最小 2 mm 的肿块。通过对图像本身中内部的特征进行的深度分析(形状、边缘、内部回声、后方回声等),从而直接进行完全化的自动流程,并且生成相应的 BI-RADS 分级,这种方式有望改善超声图像对操作者、患者状态和仪器设备的依赖性,从而提升检查的标准化和可重复性。

本研究没有进行分层比较有经验医师与无经验医师的诊断,此次研究核心在于 AI 联合医师能否创造出一定的价值,本研究结果显示, AI 诊断的特异度(0.712)显著高于医师(0.444),该结果与 Wei Q 等[9]通过前瞻性多中心的研究得出的结果一致,证实了 AI 在乳腺结节良恶性鉴别方面具有优势。值得注意的是, AI 组的假阳性率(0.288)相比超声诊断医师(0.556)降低,提示其可以降低良性病灶的误判率,提高阳性预测值。此外, AI 的阳性似然比($LR^+ = 3.347$)明显高于医师水平(1.755),提示 AI 阳性结果对恶性病变的预测效能更强,这一点在乳腺 BI-RADS 4 类病灶的鉴别中尤为重要,这与 Lu X 等[10]人的研究结果一致。本研究中,联合诊断的活检率较医师单独诊断活检率减低(从 70.46%下降至 50.21%),这与 Browne JL 等[11]得出的可以“避免医师 BI-RADS 4 级中的大量活检”结论一致。Shen 等[12]人基于大量的数据亦证实,在人工智能的辅助下假阳性率降低 37.3%,将活检率减少 27.8%,同时保持同等的灵敏度。联合诊断的 AUC 为 0.840 (95% CI: 0.787~0.892),能够提高乳腺病变良恶性判断的特异度和准确度,这与 Lyu 等[13]联合使用得出的降低 BI-RADS4 类结节的过度活检率较一致。以上均说明了联合诊断方法比单一的超声诊断或 AI 诊断效果更好。

此外,本研究对医师和 AI 对乳腺病灶分析统计发现(见表 3),形态学特征(形状、方向)的 Kappa 值最高(0.621/0.446),且 $P < 0.001$,说明 AI 对肉眼可辨的形态判断高度显著;对于钙化的判断虽然显著,但是 Kappa 值仅为 0.393,提示该深度学习软件在临床使用中仍需谨慎。在内部回声、后方回声及边缘的一致性较差,分析一致性较差的原因可能是内部回声的“低回声”判定依赖于周围组织的灰度对比,但乳腺组织背景回声存在个体差异(如致密型乳腺),导致了 AI 的误判;后方回声受声衰减(如钙化后方声影)及增强效应(如囊肿后方增强)等的影响, AI 难以区分真实病理改变与伪影;而边缘的评估往往在于医师结合病史多切面动态的扫查,而此次研究使用的 AI 系统仅进行了单帧图像,遗漏了三维边界信息,也有可能是 AI 未将病史信息与病灶信息相融合,这与高思琦等[14]研究结果较一致。

此次研究的 AI 所有误诊病例中, 1 例较为典型的误诊病例: AI 系统将一例最大直径为 2.8 cm 的乳腺肿块诊断为良性肿块,而超声诊断为恶性。分析原因可能为医师发现该肿块彩色频谱内可见较为丰富的血流信号,中央部可见粗大血管,探及高阻动脉频谱,因此考虑为恶性肿块;而 AI 系统未能整合内部血流征象,仅从肿块的二维病灶特征,给予了良性肿块的诊断。以上表明 AI 系统目前尚只能作为一种辅助医师诊断方法,提升超声医师的诊断信心和准确性,优化乳腺肿块诊断。

本研究的不足: (1) 样本量局限: 本研究为单中心研究,且纳入的样本量相对有限,未来可以进行扩大样本量,进行多中心前瞻性的研究。(2) 未进行多模态融合: 本研究仅进行了单一模态超声,未能结合其他影像(X 射线钼靶、MRI)或分子影像学等。(3) 存在选择偏倚: 研究对象仅包括活检证实的病变。因此,不包括 BI-RADS 2 类病变。(4) 没有进行有经验医师和无经验医师的分层比较,以后研究将进一步完善。

综上所述,这款超声实时辅助系统能够通过辅助医师更好地提高对乳腺癌的诊断水平,降低活检率,通过联合更好有效地诊断乳腺癌良恶性,具有一定的价值。

声 明

该病例报道已获得病人的知情同意。

基金项目

蚌埠医科大学科研课题计划(2023byzd251)。

参考文献

- [1] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., *et al.* (2024) Global Cancer Statistics

- 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, **74**, 229-263. <https://doi.org/10.3322/caac.21834>
- [2] Santos, R., Ribeiro, A.R. and Marques, D. (2023) Ultrasound as a Method for Early Diagnosis of Breast Pathology. *Journal of Personalized Medicine*, **13**, Article No. 1156. <https://doi.org/10.3390/jpm13071156>
- [3] Lee, J. (2017) Practical and Illustrated Summary of Updated BI-RADS for Ultrasonography. *Ultrasonography*, **36**, 71-81. <https://doi.org/10.14366/usg.16034>
- [4] Uehiro, N., Horii, R., Iwase, T., Tanabe, M., Sakai, T., Morizono, H., *et al.* (2013) Validation Study of the UICC TNM Classification of Malignant Tumors, Seventh Edition, in Breast Cancer. *Breast Cancer*, **21**, 748-753. <https://doi.org/10.1007/s12282-013-0453-7>
- [5] Yang, M., Jin, S.Q., Zhu, Q.L., *et al.* (2021) Development and Experience of Ultrasonic Diagnosis Technology for Breast Cancer. *Scientia Sinica Vitae*, **51**, 1092-1100.
- [6] 姜珏, 何鑫, 曹茹, 等. 医准智能软件在 4 类乳腺结节良恶性鉴别诊断中的应用价值[J]. 中国超声医学杂志, 2023, 39(7): 750-754.
- [7] Syed, A.H. and Khan, T. (2022) Evolution of Research Trends in Artificial Intelligence for Breast Cancer Diagnosis and Prognosis over the Past Two Decades: A Bibliometric Analysis. *Frontiers in Oncology*, **12**, Article ID: 854927. <https://doi.org/10.3389/fonc.2022.854927>
- [8] 贾卫, 吴凯平, 陈伟贤. 基于 CiteSpace 软件的我国乳腺癌筛查研究现状及热点可视化分析[J]. 广西医学, 2023, 45(16): 1998-2002.
- [9] Wei, Q., Yan, Y.J., Wu, G.G., *et al.* (2022) The Diagnostic Performance of Ultrasound Computer-Aided Diagnosis System for Distinguishing Breast Masses: A Prospective Multicenter Study. *European Radiology*, **32**, 4046-4055. <https://doi.org/10.1007/s00330-021-08452-1>
- [10] Lu, X., Lu, Y., Zhao, W., Qi, Y., Zhang, H., Sun, W., *et al.* (2025) Ultrasound-Based Deep Learning Radiomics for Multi-Stage Assisted Diagnosis in Reducing Unnecessary Biopsies of BI-RADS 4A Lesions. *Quantitative Imaging in Medicine and Surgery*, **15**, 2512-2528. <https://doi.org/10.21037/qims-24-580>
- [11] Browne, J.L., Pascual, M.Á., Perez, J., Salazar, S., Valero, B., Rodriguez, I., *et al.* (2023) AI: Can It Make a Difference to the Predictive Value of Ultrasound Breast Biopsy? *Diagnostics (Basel, Switzerland)*, **13**, Article No. 811. <https://doi.org/10.3390/diagnostics13040811>
- [12] Shen, Y., Shamout, F.E., Oliver, J.R., Witowski, J., Kannan, K., Park, J., *et al.* (2021) Artificial Intelligence System Reduces False-Positive Findings in the Interpretation of Breast Ultrasound Exams. *Nature Communications*, **12**, Article No. 5645. <https://doi.org/10.1038/s41467-021-26023-2>
- [13] Lyu, S., Zhang, Y., Zhang, M., Zhang, B., Gao, L., Bai, L., *et al.* (2022) Diagnostic Value of Artificial Intelligence Automatic Detection Systems for Breast BI-RADS 4 Nodules. *World Journal of Clinical Cases*, **10**, 518-527. <https://doi.org/10.12998/wjcc.v10.i2.518>
- [14] 高思琦, 牛司华, 黄剑华, 等. 超声人工智能在乳腺良恶性疾病鉴别诊断中的应用[J]. 中国超声医学杂志, 2021, 37(7): 752-755.